Conserving approximations in direct perturbation theory: new semianalytical impurity solvers and their application to general lattice problems

2008 J. Phys.: Condens. Matter 20 365217

(http://iopscience.iop.org/0953-8984/20/36/365217)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **20** (2008) 365217 (24pp) doi:10.1088/0953-8984/20/36/365217

Conserving approximations in direct perturbation theory: new semianalytical impurity solvers and their application to general lattice problems

Norbert Grewe[1], **Sebastian Schmitt**[1], **Torben Jabben**[1] and **Frithjof B Anders**[2]

[1] Institut für Festkörperphysik, Technische Universität Darmstadt, Hochschulstraße 6, D-64289 Darmstadt, Germany

[2] Institut für Theoretische Physik, Universität Bremen, PO Box 330 440, D-28334 Bremen, Germany

Received 18 April 2008, in final form 21 July 2008
Published 19 August 2008
Online at stacks.iop.org/JPhysCM/20/365217

Abstract

For the treatment of interacting electrons in crystal lattices, approximations based on the picture of effective sites, coupled in a self-consistent fashion, have proven very useful. Particularly in the presence of strong local correlations, a local approach to the problem, combining a powerful method for the short-ranged interactions with the lattice propagation part of the dynamics, determines the quality of results to a large extent. For a considerable time the noncrossing approximation (NCA) in direct perturbation theory, an approach originally developed by Keiter for the Anderson impurity model, was a standard for the description of the local dynamics of interacting electrons. In the last couple of years exact methods like the numerical renormalization group (NRG), as pioneered by Wilson, have surpassed this approximation as regarding the description of the low-energy regime. We present an improved approximation level of direct perturbation theory for finite Coulomb repulsion U, the crossing approximation 1 (CA1), and discuss its connections with other generalizations of NCA. CA1 incorporates all processes up to fourth order in the hybridization strength V in a self-consistent skeleton expansion, retaining the full energy dependence of the vertex functions. We reconstruct the local approach to the lattice problem from the point of view of cumulant perturbation theory in a very general way and discuss the proper use of impurity solvers for this purpose. Their reliability can be tested in applications to, for example, the Hubbard model and the Anderson-lattice model. We point out shortcomings of existing impurity solvers and improvements gained with CA1 in this context.

This paper is dedicated to the memory of Hellmut Keiter.

1. Introduction

In a key paper [1], Keiter and Kimball in 1971 described a new perturbational method for treating the problem of an impurity with strong local Coulomb matrix elements, embedded in a metallic host. Their guideline was to preserve the local correlations from the outset, contrary to Hartree–Fock theory or to other decoupling schemes, and to keep the interpretation of individual contributions as physical processes. The particular difficulties to be surmounted arose from the fact that choosing hybridization or intersite transfer of single particles as the perturbation leaves an interacting local shell as the unperturbed part of the Hamiltonian. In such a case the well-known machinery of Feynman diagrammatics cannot be used, including Wick's theorem and linked cluster expansions. The solution found used time-ordered pieces of Feynman processes, visualized as Goldstone diagrams, and organizing them in the form of Brillouin–Wigner perturbation theory with real energy variables. In this early formulation of the theory the need to regularize vanishing energy denominators

prevented extensive studies to infinite perturbational orders, which are necessary in the presence of infrared divergencies, encountered, for example, in the Kondo problem. These can be thought of as to arise from degeneracies in a classical part of the Hamiltonian; they are then lifted by quantum fluctuations. Early applications of the technique can be found in [2], where leading logarithmically divergent terms are summed to all orders to generate finite results, and in [3, 4], where also generalizations of the formalism to more general local shell structures and to the genuine lattice problem were discussed.

The technique of Keiter and Kimball originally was designed in connection with the Kondo problem. A revival occurred around 1980, when metallic compounds exhibiting the intermediate valence phenomenon stayed in the focus of experimentalists [5], and somewhat later, when the existence of very heavy quasiparticles (heavy fermions) was revealed in such compounds containing ions with active 4f or 5f shells [6]. In particular the discovery of heavy fermion superconductors [7] spurred the investigation of lattice models with strong local correlations. A breakthrough in theory occurred in 1983, when it was learnt how to handle direct perturbation theory completely in the complex plane, thus circumventing the regularization problem. The first independent extensive studies [8, 9] of impurity problems were based on the leading skeleton diagram contributions to the dynamics of ionic shell states, requiring the solution of self-consistently coupled singular integral equations. The contributing diagrams are completely characterized as not containing crossing band electron lines. This approximation (NCA) was shown to adequately describe dynamical properties like excitation spectra in addition to thermodynamic quantities, the latter of which being known from Wilson's implementation of the numerical renormalization group [10]. NCA allows us to study [8] the temperature-dependent formation of the Abrikosov–Suhl resonance (ASR) [11] with an assessment of the systematic shortcomings via a comparison with the resonant level model, the limit of vanishing spin degeneracy. NCA can also be characterized as a conserving approximation in the sense of Baym and Kadanoff [12] and can consistently be applied to calculate various properties of the strongly correlated impurity problem [9]. Also in 1983, a reformulation of perturbation theory was presented [13], which allowed for the use of Feynman diagrammatics via the introduction of auxiliary particles (slave bosons). Correlated local states are reintroduced in this approach by a constraint on the larger Hilbert space and a corresponding projection onto the physical sector after resummations. The slave-boson method stands in one-to-one correspondence to the formulation via direct perturbation theory and thus contributes to the same line of development.

Merits and shortcomings of the NCA for the Anderson impurity with infinite local Coulomb repulsion, meanwhile, are well known, for example as a result of the early numerical studies [8, 14] or of the exact analysis of the case with a flat electron–hole symmetric conduction band at zero temperature [15]. NCA, for example, captures the exponential part of the Kondo temperature T_K, the dynamically generated energy scale below which a local Fermi liquid is formed due to spin compensation, but not the prefactor. It does not furnish the correct values of the threshold exponents, connected with the time development of ionic states as known from the x-ray absorption problem [16, 17]. Its accuracy increases with increasing (orbital) degeneracy of the ionic level, and NCA may even become a fully acceptable approximation for a multi-channel situation [18].

The need for improved treatments of the impurity problem with correlated electrons, however, turned out to be even more important, when it became a building block in theories of lattice systems [19–21], using the concept of effective sites. It turned out that, for example, the coherence forming in the low temperature regime of lattices can only be correctly retrieved by a proper incorporation of the effective local Fermi liquid. The necessary increase in the quality of results for lattice models was largely driven by the use of numerically exact methods as impurity solvers, i.e. quantum Monte Carlo (QMC) and numerical renormalization group (NRG) [22]. NRG in particular, in its extension of the original static version to dynamical quantities [23], has been developed as a useful and convenient tool for a description of the low-energy region [24]. When combined with the (cluster-) dynamical mean field theory (DMFT) it opens the perspective for a proper description of ground states and correlations in the lattice.

Still there remain problems: (1) nonlocal matrix elements of the Coulomb interaction can be strong, too, and may enforce a nonlocal approach with links between sites from the outset, such as, for example, in pyrochlore lattices, where for certain fillings classical ground states characterized by the tetrahedron rule build the arena for quantum fluctuations [25]. (2) The use of clusters as larger local building blocks in a lattice theory greatly enlarges the local space to be diagonalized in the beginning [26, 27] and enforces a restricted choice of states and/or the use of simpler impurity solvers like, for example, a simplified finite U-version of NCA (SNCA), especially if a self-consistent determination of one-particle states and matrix elements is aimed at as part of the solution of a lattice problem [28]. (3) The numerical methods mentioned do not work equally well in different energy regimes.

NRG, for example, often leads to a poor description at high excitation energies. Depending on tunable parameters in lattice models corresponding to temperature, pressure, chemical composition, doping, etc, a wealth of phases and corresponding transitions is found; even a transmutation of underlying pictures or concepts may occur, as, for example, from magnetism of (nearly) stable local magnetic moments coupled via short-ranged exchange interactions to itinerant forms of magnetism to be described by an effective Stoner theory for bands of itinerant quasiparticles [21, 29]. In these cases a good description of a large regime of excitation energies is desired, which requires impurity solvers, which work equally well in a broad range of energies and allows for a controlled approach to known limiting cases. Facing these difficulties for a more complete understanding and description of lattice systems, further developments of methods with an analytical background seem necessary.

In this paper we will present a version of direct perturbation theory for an (effective) impurity, which combines

elements of some existing improvements of NCA and goes beyond them in some respects. An essential ingredient of this new stage of approximation are processes with crossing band electron lines, as depicted along a (imaginary) time axis, and hence the name 'crossing approximation 1' (CA1). In the following section 2 we will give a very short account of direct perturbation theory, describe CA1 and comment on its relation to NCA and its extensions as accounted for in the literature so far. Section 3 contains some calculated spectra for the Anderson impurity model and a comparison with former approximations and, most important, to results obtained with the NRG. As will become apparent, CA1 turns out to be a rather good impurity solver in the whole range of energies and can be applied to calculate excitation spectra, ionic propagators and susceptibilities. A conceptual bridge between impurity and lattice theories is outlined in section 4, which allows us to recover the results of lattice theories like XNCA [20] and DMFT [22, 30] in a general fashion.

As an example for the usefulness of CA1 in this context a calculation for the Anderson lattice is presented. The concluding section 5 contains remarks about an application of direct perturbation theory to susceptibilities and magnetic phases of lattice models and about improvements regarding nonlocal correlations as well as a new impurity solver, i.e. a CA2 project.

2. Description of CA1

2.1. Introductory remarks on direct perturbation theory

A typical set-up for the application of direct perturbation theory uses a Hamiltonian $H = H_{0\ell} + H_{0c} + V$ with the following parts: $H_{0\ell} \equiv H_{0\ell}(\{f_{m\sigma}, f^\dagger_{m\sigma}\})$ contains the dynamics of interacting electrons in local one-particle states with quantum numbers m and σ (pseudospin) and is expressed via corresponding annihilation (creation) operators $f^{(\dagger)}_{m\sigma}$. $H_{0c} \equiv H_{0c}(\{c_{\underline{k}\sigma}, c^\dagger_{\underline{k}\sigma}\})$ describes a reservoir of noninteracting electrons in Bloch states (a band index is suppressed here), and $V = V(\{f_{m\sigma}, f^\dagger_{m\sigma}, c_{\underline{k}\sigma}, c^\dagger_{\underline{k}\sigma}\})$ is a hybridization or transfer between local and band states, which is likewise expressed via elementary one-particle processes. $H_{0\ell}$ acts on a local Fock space of finite dimension, typically one or a few valence shells or orbitals, and can in principle be diagonalized. A basis of eigenstates $|n_0, M\rangle$ ('ionic states') is denominated by a local particle number n_0 and a set of many-body quantum numbers M specifying angular momenta or crystal field levels. With the operators $X_{n_0'M',n_0M} \equiv |n_0'M'\rangle\langle n_0M|$ and corresponding n_0-particle energies E_{n_0M} the local Hamiltonian is

$$H_{0\ell} = \sum_{n_0,M} E_{n_0M} X_{n_0M,n_0M}, \tag{1}$$

where only projectors onto the eigenstates appear. The terms 'local' or 'ionic' do not necessarily imply one single atom. The formalism equally well applies to local subsystems of molecular type or to local clusters. A transcription of V to local many-body states involves via

$$f^{(\dagger)}_{m\sigma} = \sum_{n_0} \sum_{M,M'} \alpha^{(*)}_{m\sigma}(n_0 - 1M', n_0M) X^{(\dagger)}_{n_0-1M',n_0M} \tag{2}$$

the set of ionic transfer operators $X_{n_0'M',n_0M}$ with $n_0' = n_0 \pm 1$. Using V as the perturbation, processes of direct perturbation theory are constructed from elementary absorption or emission events of band electrons from a local shell state at fixed (imaginary) times with amplitudes given by the coefficients $\alpha_{m\sigma}(n_0'M', n_0M)$ in (2) [3]. Insofar it can be expected that, for example, the partition function can be cast into a form where the contribution of each particular ionic state becomes apparent:

$$\begin{aligned} Z = \mathrm{Tr}\, \mathrm{e}^{-\beta H} &= \oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi \mathrm{i}}\, \mathrm{e}^{-\beta z}\, \mathrm{Tr}_\ell \mathrm{Tr}_c (z - H)^{-1} \\ &= Z_{0c} \sum_{n_0,M} \oint_{\mathcal{C}} \frac{\mathrm{d}z}{2\pi \mathrm{i}}\, \mathrm{e}^{-\beta z} P_{n_0M}(z) \\ &= Z_{0c} \sum_{n_0,M} \int \mathrm{d}\omega\, \mathrm{e}^{-\beta\omega} \varrho_{n_0M}(\omega), \end{aligned} \tag{3}$$

where the contour $\mathcal{C}$ encircles all singularities of the integrand in a counterclockwise fashion. $Z_{0c} = \mathrm{Tr}_c\, \mathrm{e}^{-\beta H_{0c}}$ is the partition function for the band part alone and $\varrho_{nM_0}(\omega) = -\frac{1}{\pi}\mathrm{Im}\, P_{n_0M}(\omega + \mathrm{i}\delta)$ is the spectral intensity of the ionic state $|n_0M\rangle$, which evolves with the propagator $P_{n_0M}(z)$. A straightforward concept of irreducibility with respect to intermediate ionic states allows for the introduction of irreducible ionic self-energies, with analytical properties as usual:

$$P_{n_0M}(z) = \big(z - E_{n_0M} - \Sigma_{n_0M}(z)\big)^{-1}, \tag{4}$$

and a corresponding perturbation expansion. The processes contributing to these self-energies $\Sigma_{n_0M}(z)$ will be constructed from skeleton diagrams in the following, so that the ionic propagators $P_{n_0M}(z)$ are to be determined self-consistently from a set of coupled integral equations.

Representations for general Green functions, which in correspondence with the partition function (3) are expressed as convolutions of ionic propagators, can also be derived along the lines sketched above. We consider in particular the local one-particle Green function:

$$\begin{aligned} F_{m\sigma}(\tau) &= -\langle T(f_{m\sigma}(\tau) f^\dagger_{m\sigma})\rangle \\ &= -\sum_{n_0,\widetilde{n}_0} \sum_{M_1,M_1'} \sum_{M_2,M_2'} \alpha_{m\sigma}(n_0 - 1M_1', n_0M_1) \\ &\quad \times \alpha^*_{m\sigma}(\widetilde{n}_0 - 1M_2', \widetilde{n}_0M_2) \\ &\quad \times \langle T(X_{n_0-1M_1',n_0M_1}(\tau) X_{\widetilde{n}_0M_2,\widetilde{n}_0-1M_2'})\rangle, \end{aligned} \tag{5}$$

the Fourier coefficients of which at Matsubara frequencies $\omega_n = \frac{(2n+1)\pi}{\beta}$ ($n \in \mathbb{Z}$) give, after analytical continuation $\mathrm{i}\omega_n \to \omega + \mathrm{i}\delta$ to the upper border of the real frequency axis, the local one-particle excitation spectrum $\varrho_{m\sigma}(\omega) = -\frac{1}{\pi}\mathrm{Im} F_{m\sigma}(\omega + \mathrm{i}\delta)$. The complete set-up of nonstandard direct perturbation theory is well documented, including the diagrammatic rules for the processes to be discussed in the following [3, 8, 31].

2.2. General features of crossing and noncrossing approximations for the SIAM

The first comprehensive studies of direct perturbation theory for the Anderson impurity model (SIAM):

$$\hat{H} = \sum_{\sigma} \left(\epsilon_\ell \, \hat{f}^\dagger_\sigma \hat{f}_\sigma + \frac{U}{2} \hat{n}^f_\sigma \hat{n}^f_{\bar{\sigma}} \right) + \sum_{\underline{k},\sigma} \epsilon_{\underline{k}} \hat{c}^\dagger_{\underline{k}\sigma} \hat{c}_{\underline{k}\sigma} + \frac{1}{\sqrt{N}} \sum_{\underline{k},\sigma} \left(V_{\underline{k}} \hat{c}^\dagger_{\underline{k}\sigma} \hat{f}_\sigma + \text{h.c.} \right) \tag{6}$$

concentrated on the limit $U = \infty$ of infinite local Coulomb repulsion and were termed NCA [8, 9, 32]. They were based on the leading skeletons of order V^2 to the ionic self-energies $\Sigma_0(z)$ and $\Sigma_{1\sigma}(z)$ and furnished a qualitatively correct picture, e.g. for the temperature-dependent formation of the Abrikosov–Suhl resonance (ASR), the most prominent many-body signature of the Kondo effect [33] in the local one-particle spectrum of the model. The particular aspect of degeneracy ν of the local level, i.e. $\nu = 2$ for the two possible z components of spin in the original SIAM but higher ν as in Ce compounds with $\nu = 6$ becoming possible through orbital degeneracy, drew much attention: in the limit $\nu \to \infty$, using a proper scaling $V \to \frac{V}{\sqrt{\nu}}$, the NCA results become increasingly valid [2, 13, 34], and for $\nu \to 1$, on the other hand, a trivially solvable resonant level model emerges. Whereas the limit of large ν gave reason for classification schemes of diagrams in orders of $\frac{1}{\nu}$, the limit $\nu = 1$ aroused hopes of reconstructing an exact solution of a simple model by direct perturbation theory, thus completely clarifying the systematics of diagrams for all cases of ν.

This hope was not fulfilled up to now, although Keiter presented an exact solution of SIAM for the zero-bandwidth limit unravelling the full diagrammatics for this simpler case [35]. It became clear then that progress with the direct perturbation approach had to be worked out stepwise by including more important classes of skeleton diagrams into the calculations.

In the following we discuss the systematics of these approximations for the SIAM by concentrating on the vertices, which allows for writing down several quantities in a compact and rigorous form, i.e. the ionic self-energies

$$\begin{aligned} \Sigma_0(z) &= \sum_\sigma \int \mathrm{d}x \; D_\sigma(x) f(x) \Lambda_{0,1\sigma}(z,x) P_{1\sigma}(z+x), \\ \Sigma_{1\sigma}(z) &= \int \mathrm{d}x f(x) \big[D_\sigma(-x) \Lambda_{0,1\sigma}(z+x,-x) P_0(z+x) \\ &\quad + D_{-\sigma}(x) \Lambda_{2,1\sigma}(z+x,-x) P_2(z+x) \big], \\ \Sigma_2(z) &= \sum_\sigma \int \mathrm{d}x D_{-\sigma}(-x) f(x) \Lambda_{2,1\sigma}(z,x) P_{1\sigma}(z+x), \end{aligned} \tag{7}$$

and the local one-particle Green function (5), which in the special case of the SIAM contains only two contributions:

$$F_\sigma(z) = \frac{1}{Z_\ell} \oint_{\mathcal{C}} \frac{\mathrm{d}z'}{2\pi \mathrm{i}} \mathrm{e}^{-\beta z'} \big[\Lambda_{0,1\sigma}(z',z) P_0(z') P_{1\sigma}(z+z') + \Lambda_{2,1-\sigma}(z',-z) P_2(z') P_{1-\sigma}(z-z') \big]. \tag{8}$$

Here we have introduced the hybridization intensity $D_\sigma(\epsilon) = \frac{V(\epsilon)^2}{N} \sum_{\underline{k}} \delta(\epsilon - \epsilon_{\underline{k}\sigma}) = V(\epsilon)^2 \varrho_{0c}(\epsilon)$ and the (perturbed) local partition function $Z_\ell = \frac{Z}{Z_{0c}}$. Expressions for higher Green functions take an analogous form; we only add here a formula for the dynamical magnetic susceptibility (leaving out prefactors $(\frac{1}{2} g \mu_\beta)^2$):

$$\chi_{\mathrm{mag}}(z) = -\frac{1}{Z_\ell} \oint_{\mathcal{C}} \frac{\mathrm{d}z'}{2\pi \mathrm{i}} \mathrm{e}^{-\beta z'} \times \sum_\sigma \Lambda_{\sigma,\sigma}(z',z) P_{1\sigma}(z') P_{1\sigma}(z+z'), \tag{9}$$

which involves a separate kind of vertex $\Lambda_{\sigma,\sigma}$. Equations (7)–(9) are graphically represented in figure 1; observe identities like $\Lambda_{0,1\sigma}(z-z',z') = \Lambda_{1\sigma,0}(z,z')$ for setting up the equations from there.

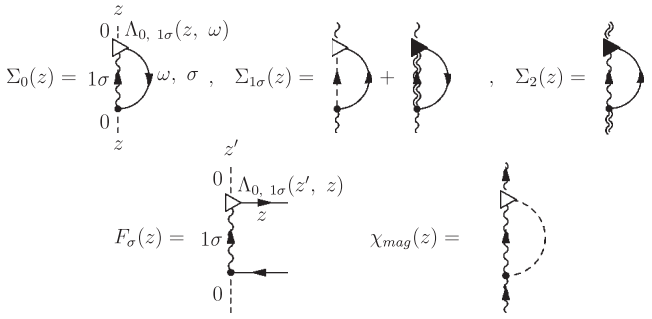

Figure 1. Diagrammatic representation of ionic self-energies, local one-particle Green function and magnetic susceptibility for a single-impurity Anderson model (SIAM) in direct perturbation theory. Physical processes are arranged vertically along an imaginary time axis (broken line) which bears an energy variable z after Laplace transformation. The presence of an electron in the local shell is indicated via a wiggly line on this time axis. Excitations of band electrons (straight lines) take place at hybridization vertices (dots on the time axis). Due to time-rotational invariance all vertex corrections in these diagrams can be collected at one of the vertices, which is drawn as a triangle.

In the SNCA, which can be viewed as the simplest nontrivial approximation for all values of U, the vertex functions are all taken without any vertex corrections:

$$\text{SNCA}: \quad \Lambda_{0,1\sigma} = \Lambda_{2,1\sigma} = \Lambda_{\sigma,\sigma} \equiv 1 \,. \tag{10}$$

The original NCA constitutes the $U \to \infty$ limit hereof and is obtained by ignoring the doubly occupied state, i.e. by setting $P_2 \equiv 0$.

The NCA, and as a consequence the SNCA for general values of U, can only furnish qualitative insight into the dynamics of the SIAM, since it is plagued by shortcomings. These are revealed in the following ways: (1) comparison with the resonant level limit $\nu \to 1$ ends in an insufficient fit to the virtual scattering resonance in the excitation spectrum, in particular when the intermediate valence regime is approached [8]. (2) Form and position of the ASR are not in accord with the Friedel sum rule, most important for $\nu = 2$, and correspondingly the local self-energy $\widetilde{\Sigma}_\sigma(z) = \Sigma_\sigma(z) + \mathrm{i}\Delta_\mathrm{A} = z - \epsilon_\ell + \mathrm{i}\Delta_\mathrm{A} - F_\sigma(z)^{-1}$ does not comply with local Fermi-liquid properties [36] ($\Delta_\mathrm{A} = \pi V(0)^2 \varrho_{0c}(0)$ is the Anderson width and $\epsilon_\ell = E_{1\sigma} - E_0$ the local one-particle level). (3) Threshold exponents, as taken from the ionic propagators (see below) with values $\alpha_0 = \frac{1}{3}$ and $\alpha_{1\sigma} = \frac{2}{3}$ in the NCA, do not agree with the values known from the x-ray absorption problem [17]. (4) An exact analytical solution of the

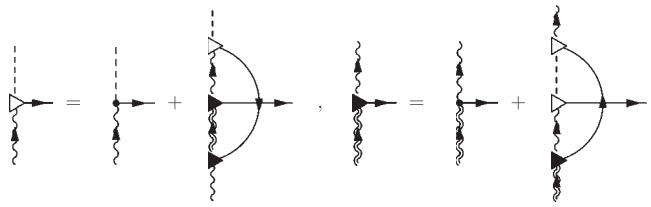

Figure 2. Vertex structure of the 'full NCA' (FNCA). The hierarchy of vertex corrections is generated by two coupled integral equations. The bare vertices (first terms) are subsequently crossed by one more band excitation, which ends below and above in a full vertex, respectively. The diagrams are taken as skeletons, i.e. the local lines are dressed with the full ionic propagators.

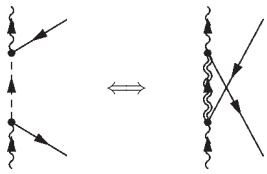

Figure 3. Sequences of two elementary excitation processes, which together constitute the lowest order exchange coupling vertex remaining after a Schrieffer–Wolff transformation of SIAM to the s–d-model.

NCA version of equations (7) for zero temperature and a flat conduction band density of states symmetric around the Fermi energy reveals spurious features near the ASR [15], namely a sharp spike showing up at the Fermi level below a ('pathology') temperature T_{p}, being still lower than T_{K} in the Kondo regime.

Point (3) deserves some further comments, because it hints at the particular singular structure of the ionic propagators $P_{n_0 M}(z)$, which causes difficulties in the numerical solution of the system (7) of integral equations and also in subsequent procedures like (8) and (9) involving convolutions of several of the $P_{n_0 M}$. As explained, for example in [17], these propagators develop a common threshold at an energy $\omega = E_{\mathrm{g}} < \epsilon_\ell$ for zero temperature, due to a slow algebraic decay $P_{n_0 M}(t) \sim \mathrm{e}^{-\frac{\mathrm{i}}{\hbar} E_{\mathrm{g}} t}/t^{\alpha_M}$ in the time domain, i.e.

$$\begin{aligned} &\mathrm{Im}\, P_{n_0 M}(\omega - \mathrm{i}\delta) \sim 1/(\omega - E_{\mathrm{g}})^{1-\alpha_M}, \\ &\alpha_0 = \frac{n_\ell^2}{\nu}, \qquad \alpha_{1\sigma} = 1 - 2\frac{n_\ell}{\nu} + \frac{n_\ell^2}{\nu}, \end{aligned} \tag{11}$$

$0 < n_\ell \leqslant 1$ being the occupation of the local level in the Kondo regime $0 < \Delta_{\mathrm{A}} < -\epsilon_\ell < U$. At $T = 0$, E_{g} is the lower endpoint of a branch cut in the functions $P_{n_0 M}(z)$ along the real axis $z = \omega > E_{\mathrm{g}}$; it is this particular divergent behaviour—for $n_\ell \lesssim 1$ and $\nu = 2$ one has $1 - \alpha_0 \gtrsim \frac{1}{2}$ and $0 < 1 - \alpha_{1\sigma} \lesssim \frac{1}{2}$—which needs care and makes numerical calculations to higher orders much more time-consuming than NCA or SNCA, due to multiple convolutions of these singular structures.

2.3. Generalizations of SNCA: PNCA, ENCA and FNCA

The first useful generalization of NCA to the SIAM with general values of the Coulomb repulsion U was proposed and investigated in 1989 [37]. It was called 'full NCA' (FNCA) and is visualized diagrammatically in figure 2. One recognizes a particular subsystem of integral equations, which serves to generate a class of vertex corrections (again as skeletons) extending to infinite order. This particular choice was motivated by an attempt to include as many exchange counterparts as possible to those processes, which already contribute to the ionic propagators in SNCA, see the appendix in [37]. As Keiter repeatedly has pointed out [38], the balance between processes which transform into each other by a reversal of partial time orderings, as shown in figure 3, is necessary to obey the Pauli principle and to comply with universality in the Kondo limit, where in accord with the Schrieffer–Wolff transformation from SIAM to the s–d-exchange model [39] the characteristic energy scale $k_{\mathrm{B}} T_{\mathrm{K}}$ is expressed via an effective exchange coupling constant $I = \frac{V^2}{\epsilon_\ell} - \frac{V^2}{\epsilon_\ell + U}$; figure 3 just visualizes the two contributions to this I [11, 40].

The system of five integral equations according to figures 1 and 2 was solved for finite U in [37], and the results for the ionic propagators and the corresponding excitation spectra were compared to some simpler calculation schemes. Whereas pronounced discrepancies from SNCA showed up, e.g. regarding the important energy scales, the so-called 'enhanced NCA' (ENCA) already captured important improvements.

In ENCA all vertices on the right-hand side of the two equations in figure 2 are taken as bare ones. Then, only the leading contributions to the infinite series of vertex corrections contained in FNCA are included; among the latter are running n-particle cascades between the initial and final state during the excitation by the external electron (iterate the vertex in the middle of the last diagram) as well as long-time memory effects between initial and final states through chains of internal excitations, before or after the external excitation occurs (iterate the respective vertices on top and at the bottom of the diagram). Since the ENCA has proven as a good compromise between accuracy and the calculational effort to be invested in an impurity solver for lattice problems (see also section 4), we cite the explicit expressions for the vertex corrections, which have to be solved together with the system (7) of self-energy equations:

$$\begin{aligned} \Delta\Lambda^{(\mathrm{ENCA})}_{0,1\sigma}(z, z') &= \int \mathrm{d}\epsilon\, D_{-\sigma}(\epsilon) f(\epsilon) \\ &\quad \times\ P_{1-\sigma}(z + \epsilon) P_2(z + z' + \epsilon), \\ \Delta\Lambda^{(\mathrm{ENCA})}_{2,1-\sigma}(z, z') &= \int \mathrm{d}\epsilon\, D_{-\sigma}(\epsilon)(1 - f(\epsilon)) \\ &\quad \times\ P_{1\sigma}(z - \epsilon) P_0(z - z' - \epsilon). \end{aligned} \tag{12}$$

Calculations of the local one-particle spectrum in [37] were then based on the ENCA and led to an improved many-body scale and a better understanding of the many-body dynamics of SIAM, in particular at finite values of U.

Whereas ENCA takes into account the vertex corrections up to order $\mathcal{O}(V^2)$, and FNCA in addition certain classes up to infinite order, both do not include the fully crossing diagram of order $\mathcal{O}(V^4)$ shown in figure 4(a).

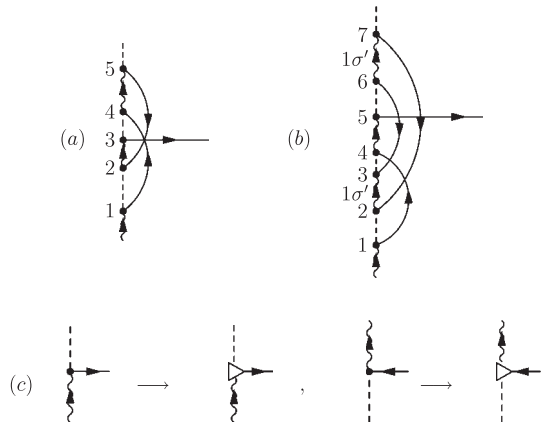

Figure 4. Part (a) shows a fully crossing vertex correction of order $\mathcal{O}(V^4)$. The two processes shown in parts (a) and (b) are of the same order regarding an expansion in the degeneracy ν of the singly occupied local state. With the self-consistent replacements shown in part (c) and with fully dressed local lines they constitute the 'post-NCA' (PNCA), a theory for $U = \infty$, in which the doubly occupied local state is projected out.

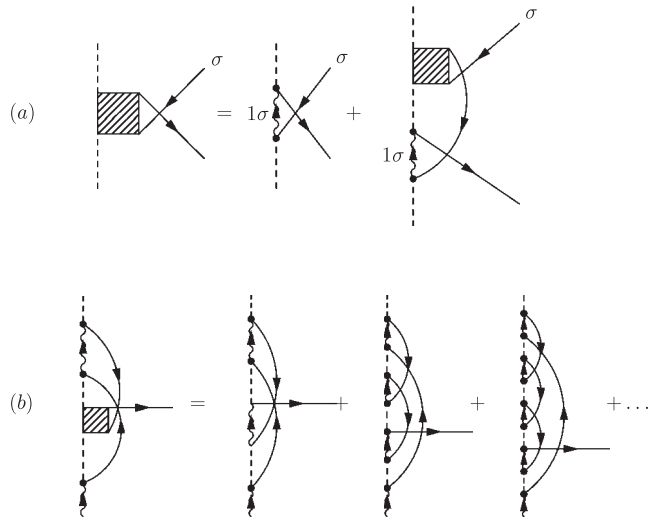

Figure 5. Vertex corrections summed in the 'conserving T-matrix approximation' (CTMA) for the SIAM at $U = \infty$ as viewed from direct perturbation theory. Part (a) T-matrix which, when substituted into the diagram of figure 4(a), generates the sequence of vertex corrections shown in part (b).

This vertex correction is the lowest nonvanishing one in the $U = \infty$-theory and was frequently used to discriminate 'crossing' and 'noncrossing' approximations.

In order to investigate the role of such fully crossing diagrams the SIAM at infinite U was investigated in 1994 with help of a 'post-NCA'(PNCA) [41]. This approximation scheme was set up along the lines of a $\frac{1}{\nu}$ expansion and collected all vertex corrections up to $\mathcal{O}(\frac{1}{\nu^2})$, i.e. all contributions to the ionic self-energies up to this order. Therefore, also the vertex correction shown in figure 4(b) was taken into account, which has two more powers of V compared with figure 4(a), but due to $\sum_{\sigma'=1}^{\nu} (\frac{V}{\sqrt{\nu}})^6 = \frac{V^6}{\nu^2}$ is of the same order $\frac{1}{\nu^2}$ thanks to the closed ring with spin summation over σ' between vertices 2, 3, 6 and 7. Actually, and in close analogy to the FNCA, the bare vertices in figures 4(a) and (b) were all replaced by full ones, as indicated in figure 4(c), and the coupled system of vertex corrections (now including all orders) was solved, again self-consistently together with the system (7) of ionic self-energies. Convergence could be reached on not too large timescales by the use of parallel computing. Progress over the original NCA turned out to be essential: apart from a corrected many-body scale $k_B T_K$, the local Fermi-liquid properties improved considerably, the position of the ASR near the Fermi level agreed much better with the one implied by Friedel's sum rule, and also the threshold exponents α_0 and α_1 were shifted towards the values of (11), although agreement with these values or with a variant according to [42] was not conclusive.

Due to the considerable numerical effort, regarding the multiple overlapping integrations over functions with rich structure, an extension of PNCA to finite values of U seemed not possible in 1994, since many more diagrams involving the doubly occupied state would have to be added.

2.4. Other approximation schemes in the literature

Before the new approximations CA1 and a CA2 project (in section 5) will be explained, we shortly comment on two approximation schemes, which have been proposed and investigated over the last ten years. In the so-called 'symmetrized finite-U NCA' (SUNCA) special emphasis is laid on the chains of scattering events [43] mentioned above in connection with the FNCA [37]. This scheme is conserving (Φ-derivable in the sense of Kadanoff and Baym) like all other approximations mentioned in this section; moreover it can be characterized as involving just a subclass of the FNCA-diagrams. Although the relevant papers are written with help of the slave-boson formalism, the formulation is fully equivalent to direct perturbation theory as pointed out above. The evaluation of the local one-particle spectrum is based on a full infinite subclass of vertex corrections and thus goes beyond the ENCA calculations. These vertex corrections with long scattering chains are easily visualized with the help of the FNCA diagrams of figure 2: iterate the vertex equations with respect to the upmost vertex only. The results underline the progress reached with ENCA and FNCA [37].

Whereas SUNCA is applicable to the finite-U case and can be placed into a scheme of repeated vertex corrections with single line crossings (a more general version being FNCA), the 'conserving T-matrix approximation' (CTMA) [44] again is restricted to infinite U and stresses the importance of chains of scattering events for band electrons off the local shell over the whole duration of the external excitation process. These are argued to contain those significant contributions, which are known to lead to the correct singular threshold behaviour of x-ray absorption spectra as predicted by Mahan [45] and calculated by Nozières *et al* [16].

Correspondingly, essentially exact threshold exponents are expected from the CTMA. This approximation can be characterized with reference to the fully crossing diagram of figure 4(a): the middle part between vertices 2 and 3 becomes the lowest contribution to a T matrix, which is fully determined by the implicit equation shown in figure 5(a). It generates the

sequence of vertex corrections with scattering chains shown in figure 5(b). Observe that only the first of these is contained in PNCA.

In spite of a superficial resemblance, already the second contribution is different from figure 4(b), which is more easily recognized by counting the number of independent spin summations. Indeed, CTMA-results [46] point to considerably improved values of the threshold exponents; nevertheless, the description of the local Fermi-liquid formation, similar to PNCA, is still not fully satisfactory. Both of these approximations involve time-consuming numerical calculations; up to now a generalization to the even more demanding case of finite U values has not been reported.

Other approximation schemes involving additional simplifying assumptions for the ionic propagators and vertex functions, be it either in a non-conserving [47] or conserving fashion [48], will not be considered here. Although they may be useful with respect to computational effort, they have only been justified for the case of large orbital degeneracy.

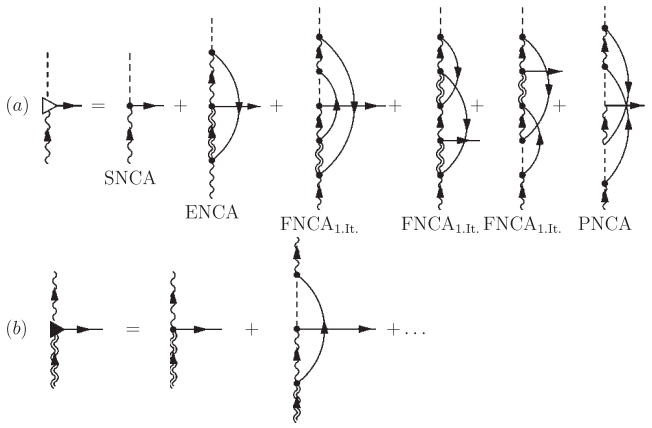

Figure 6. CA1 collects all vertex corrections for general (finite and infinite) values of U up to order $\mathcal{O}(V^4)$; in part (a) these are shown explicitly for one of the two vertices. The analogous construction for the other vertex is indicated in part (b) by dots. Again local lines are dressed, i.e. the diagrams are used as skeletons; the vertex points on the right-hand side, however, are bare ones.

2.5. CA1 approximation

CA1 is designed to describe SIAM in the full range of values for the local Coulomb repulsion U with good accuracy and likewise for dynamical properties at general excitation energies ω. Being a straightforward collection of all vertex corrections up to order $\mathcal{O}(V^4)$ (as skeletons) it is conserving and contains the leading contributions from all of the approximations sketched above. More precisely, it can be defined by the set of vertex corrections shown in figure 6(a), plus the corresponding ones for the other vertex, indicated by points in figure 6(b). In successive order the diagrams may be characterized as follows: the first terms on the rhs are the bare vertices and define SNCA. ENCA additionally contains the following vertex correction with a single crossing electron or hole line, respectively. The next three vertex corrections can be viewed as originating from the ENCA diagram by dressing each of the vertices with a single crossing line successively, i.e. they represent the first iteration in the FNCA scheme. The last diagram is the fully crossing contribution not contained in the FNCA; it is the leading vertex correction in both approximations for infinite U, PNCA and CTMA.

Whereas CA1 is explored in the following together with the other approximations mentioned, a CA2-project will be designed to add more vertex iterations like those included in FNCA and longer scattering chains like those of the CTMA, which can be incorporated via a T-matrix formalism.

3. Results from CA1 and comparison with other impurity solvers

3.1. General remarks on the quality of impurity solvers

The following criteria have frequently been applied to judge the quality of impurity solvers in connection with the Anderson impurity model: (1) ionic propagators have to obey the correct threshold behaviour, in accord with the relevant work on orthogonality catastrophe and excitonic correlations in the case with spin degeneracy [17, 45, 49, 50]. (2) The infrared divergences of the perturbation series produce a characteristic low-energy scale, usually referred to as the Kondo temperature T_K, which should faithfully be reproduced by the approximation. (3) The many-body resonance (Abrikosov–Suhl resonance, ASR) forming at temperatures of the order of the Kondo scale T_K and lower is pinned at a position near the Fermi level, which is determined by Friedel's sum rule. (4) The shape of the ASR has to comply with a form of the self-energy, which guarantees local Fermi-liquid properties.

In the following, the four semianalytical impurity solvers introduced hitherto for the SIAM with general, in particular, finite values of U, i.e. SNCA, ENCA, FNCA and SUNCA will be compared to the new CA1. Special emphasis will be laid on the above four criteria. As a reference, NRG calculations for the SIAM are also presented, which in the low-energy regime should provide a reliable bias. They have been produced with help of the very effective numerical procedures presented in [51]. All other calculations have been performed with a software package written for the solution of a number of impurity and lattice models, in which various impurity solvers can be combined with different methods for the lattice aspects. Since it is based on adaptive strategies for an all-purpose use, no particular provisions have been taken to optimize numerical strategies for the SIAM in the deep Kondo limit. Nevertheless, the program package seems to work very reliably, although numerical convergence problems and approximation errors become visible for certain extreme choices of model parameters. In particular, the convolution of several nearly singular factors in an integrand like that of (8) at very low temperatures needs a thorough analytical preparation and consumes much numerical effort, and likewise the iteration of vertex parts depending on two energy variables for SUNCA, FNCA and CA1.

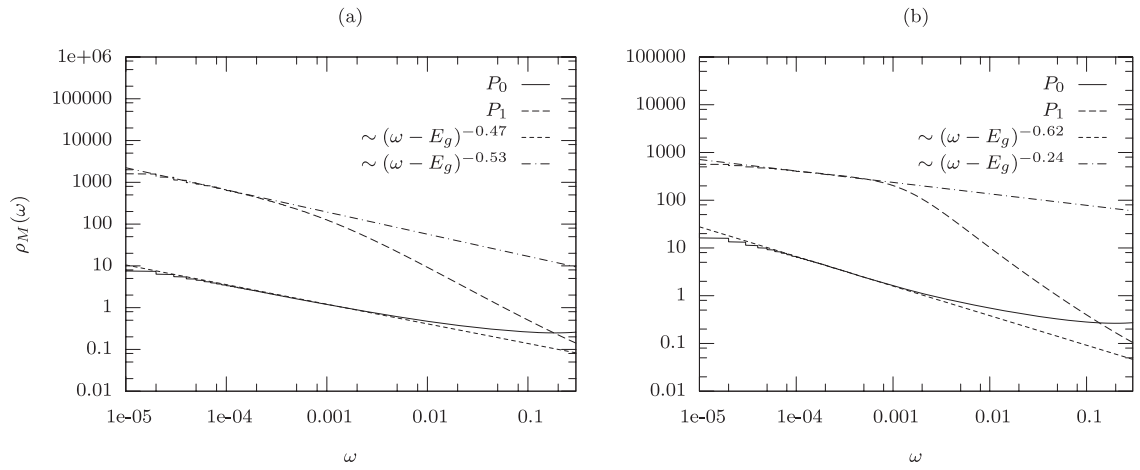

Figure 7. Double-logarithmic plot of ionic spectra, centred at the threshold, for a SIAM in SNCA (a) in the symmetric case with $\epsilon_\ell = -1.0,\ U = 2.0,\ \beta = \infty$ and (b) in an asymmetric case with $\epsilon_\ell = -1.0,\ U = 3.0,\ \beta = \infty$, both for an Anderson width $\Delta_A \equiv \pi V^2 \varrho_{c\sigma}^{(0)}(0) = 0.3$ and a 3d–sc band density of states $\varrho_{c\sigma}^{(0)}(\omega)$, centred at $\omega = \mu = 0$; the bandwidth is 6 units. The threshold exponents can be read off as one plus the slope of the asymptotic tangents drawn in the figures.

3.2. Ionic threshold behaviour

The shortcomings of SNCA have already been mentioned in the foregoing section. It proves worthwhile, however, to check how the known values of threshold exponents for this approximation are recovered in the calculations; this gives valuable hints as to how results for the other approximations should be interpreted and, generally, how reliably the algorithms work. It is interesting to note here that the information given about the NCA-threshold exponents (i.e. the case $U = \infty$) in connection with (11) is not complete when regarding SNCA at finite U. The treatment of [15] based on an ansatz for ionic self-energies is easily generalized:

$$\Sigma_M(\omega + \mathrm{i}\delta) \approx E_\mathrm{g} - E_M - \frac{\mathrm{i}}{A_m}(\omega - E_\mathrm{g})^{1-\alpha_M}$$
$$\rightarrow\ P_{n_0 M}(\omega + \mathrm{i}\delta) \approx -\mathrm{i}A_M(\omega - E_\mathrm{g})^{\alpha_M - 1}. \qquad (13)$$

It furnishes the same asymptotic region near $\omega = E_\mathrm{g}$ as in the case $U = \infty$, characterized by $\alpha_0 = \frac{1}{3}$, $\alpha_{1\sigma} = \frac{2}{3}$ for (spin-) degeneracy $\nu = 2$, for the whole unsymmetric regime ($\epsilon_\ell \equiv E_{1\sigma} - E_0 < -\Delta$ and) $2\epsilon_\ell + U > 0$. Asymptotically, the propagator P_2 does not contribute here. For U approaching the value $-2\epsilon_\ell$ from above, however, P_2 becomes equal to P_0, which leads to

$$\alpha_1 = \frac{\nu}{2+\nu}, \qquad \alpha_0 = \alpha_2 = \frac{2}{2+\nu} \qquad (2\epsilon_\ell + U = 0), \qquad (14)$$

ν being the degeneracy of ionic state $|1\sigma\rangle$. For the model with spin degeneracy only this means $\alpha_0 = \alpha_1 = \alpha_2 = \frac{1}{2}$. These values coincide with the presumably exact ones taken from (11). Naturally this does not imply that SNCA becomes correct for the symmetric SIAM. As we will see below, for example, the shape of the ASR still reveals serious shortcomings. The transition from $U = \infty$ to $U = -2\epsilon_\ell$ happens in a gradual way: the former asymptotic regime around $\omega = E_\mathrm{g}$ shrinks to zero and a different regime takes over, which was originally situated at higher values of $|\omega - E_\mathrm{g}|$ and developed the exponents of (14).

In figures 7 and 8 we show the results of a rather precise SNCA calculation of a symmetric SIAM in the deep Kondo regime, part (a), and an asymmetric one, part (b). The impurity states locally hybridize with a tight-binding simple-cubic conduction band in three dimensions of width 6, the Fermi level and band centre lie at energy $\omega = 0$ and van Hove singularities at $\omega = \pm 1$. Shown in figure 7 are the spectra of the relevant ionic propagators P_0 and $P_{1\sigma}$ for $T = 0$ on a doubly logarithmic scale with origin at the corresponding threshold energies E_g. Numerical resolution is somewhat below 10^{-4}, acting as an effective temperature cutoff. The asymptotic regime is entered only one order of magnitude higher, at about $\omega - E_\mathrm{g} \approx 10^{-3}$. From the slope of the tangents drawn one reads off the exponents $\alpha_0 \approx 0.53$, $\alpha_1 \approx 0.47$ in the symmetric case, and $\alpha_0 = 0.38$, $\alpha_1 = 0.76$ in the asymmetric one. Even if by a proper extrapolation, using several low values of the temperature, these numbers can be brought closer to the exactly known ones, i.e. (0.5, 0.5) and (0.33, 0.67), an uncertainty of about 10% remains. This is enough, however, to identify and discriminate the two different situations. With lower numerical accuracy or higher temperatures $T > 10^{-4} \approx T_\mathrm{K}/200$ the asymptotic regime will be hard to attain.

Figure 8(a) demonstrates that, in spite of accurate threshold exponents, the shape of the ASR at $\omega = 0$ in the one-particle-spectrum for low T comes out as a quite unphysical spike. Part (b) of this figure shows how the ASR is deformed, when U is raised; the position of the ASR is clearly temperature-dependent, and its flank at $\omega = 0$ develops a pathological steepness. This reminds us of the pathological structure found at infinite U with a flat conduction band density of states [15]. Since for our calculations a three-dimensional simple-cubic tight-binding band structure was used, van Hove singularities (i.e. kinks at $\omega = \pm 1$) leave their traces in the spectrum; their visibility also constitutes a test for the numerical procedures used.

A good qualitative insight into the relation between the five semianalytical impurity solvers under consideration can be obtained from figure 9, which shows the spectrum of

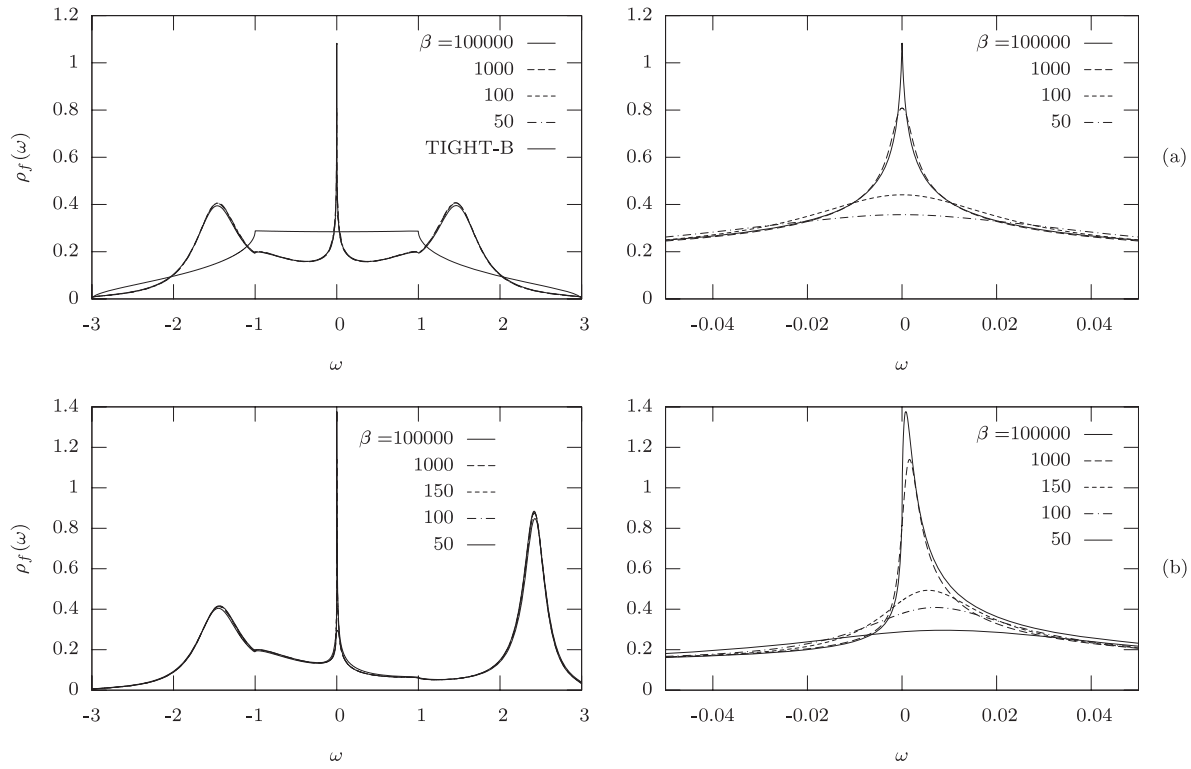

Figure 8. One-particle excitation spectrum of SIAM in SNCA, parameter values as in figure 7, β values as specified.

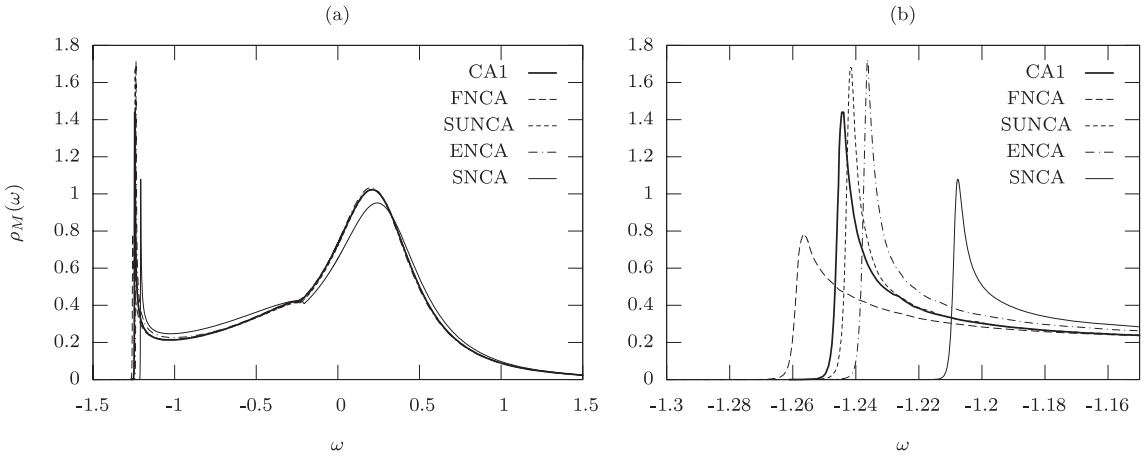

Figure 9. Spectral density of the empty ionic state $M = 0$ for a SIAM, calculated with the five semianalytical approximations discussed in the text, parameter values as in figure 7, $\beta = 1000$. The shifted thresholds allow for fits of different quality to the Kondo temperature T_K.

$P_0 = P_2$ for the symmetric SIAM discussed before at temperature $T = 10^{-3}$. Whereas part (a) gives an overall view with the threshold to the left and a broad one-particle resonance to the right, corresponding to a distribution of contributing frequencies around $\omega = -E_{1\sigma} - \Delta E_{1\sigma}$ (E_0 is set to zero in all calculations), part (b) with a much finer energy resolution points to the discrepancies between the different approximations visible in the low-energy regime. One recognizes threshold peaks at different values of E_g, in increasing order for FNCA, CA1, SUNCA, ENCA and SNCA. Differences in E_g directly reflect the ability of the approximations to reproduce the Kondo scale, which in this regime can be expressed as [52]

$$[k_B]T_K = a\sqrt{I}\exp\left[-\frac{\pi}{I}\right], \qquad I = -\frac{2U\Delta_A}{\epsilon_\ell(\epsilon_\ell + U)}. \quad (15)$$

Choosing $a = \frac{U}{2\pi}$ for $U \leqslant$ bandwidth W [37] we obtain $T/T_K = 0.04$, i.e. the spectra represent the temperature range well below T_K, even with slightly different choices of the coefficient a. If one accepts, for the moment, that the FNCA with lowest E_g furnishes the closest approximation to the real T_K, as, for example, is implied by the NRG calculation (see below), then the following conclusion can be drawn: ENCA, SUNCA and CA1 all improve considerably

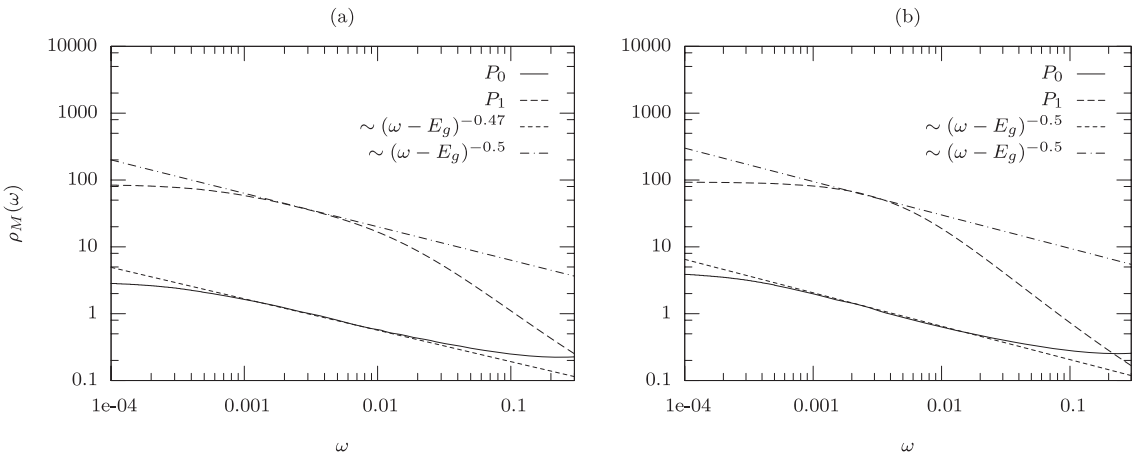

Figure 10. Double-logarithmic plot of ionic spectra, centred at the threshold, for a SIAM in ENCA, parameter values as in figure 7.

on the SNCA. The leading vertex correction already included in the ENCA contributes the essential part to this effect, whereas the additional terms further taken into account in SUNCA and CA1, respectively, have a relatively smaller impact. This agrees with the original investigation of ENCA and FNCA [37], where it was shown, that the ENCA already captures the right exponential behaviour of T_K for the SIAM at finite U, whereas the inclusion of further vertex corrections then only improves on the prefactor in this scale.

As will be shown below, the good estimate of T_K furnished by the FNCA does not imply that FNCA behaves well in all other respects, e.g. concerning the four points mentioned in the beginning. What is apparent, however, is the pronounced and qualitatively similar threshold behaviour visible in all of the five approximation schemes applied to the ionic spectra.

Since ENCA, SUNCA and FNCA all reduce to the NCA in the limit $0 > \epsilon_\ell$ fixed, $U \to \infty$, it is to be expected that, in the asymmetric case, $2\epsilon_\ell + U > 0$ an ultimate asymptotic regime very near to $\omega = E_g$ exists, where the ionic spectra are ruled by the NCA-threshold exponents. This must not necessarily be true for the CA1 with its fully crossing vertex correction, which does not vanish in this limit. It can nevertheless be anticipated that, as a precursor, a regime with 'better' threshold exponents at somewhat higher values of $|\omega - E_g|$ occurs also for ENCA, SUNCA and FNCA. In figure 10(a) a corresponding evaluation of ENCA is shown for a temperature which again is far below T_K and also below the numerical resolution of about 10^{-3} or somewhat less. The exponents for the symmetric case in figure 10(a), as read off in the range $10^{-3} \leqslant \omega \leqslant 10^{-2}$, are clearly near the value 0.5, whereas in the asymmetric case of figure 10(b) only $\alpha_0 \approx 0.5$ is really conclusive; a value $\alpha_1 \approx 0.5$ can be justified only if a tangent is drawn in the reduced range between $\omega \approx 10^{-3}$ and $\omega \approx 0.5 \times 10^{-2}$ before the steeper decrease sets in. Comparing figures 7(b) and 10(b) this seems to be a reasonable procedure. It must be remembered here that the better approximation cannot be evaluated with the same numerical accuracy, at least not using the program package in its present form. Figures 11(a)–(c) show the threshold behaviour for the asymmetric model (again with $\epsilon_1 = -1.0$, $U = 3.0$ and $T \ll T_K$) obtained with SUNCA, FNCA and CA1. Similar conclusions as for the ENCA can be drawn here: in all cases the threshold exponents in the accessible asymptotic regime come out near the value 0.5. At least for the CA1 this gives reason to hope for an essential improvement of the true asymptotics of ionic spectra over the SNCA.

3.3. One-particle excitation spectra

For a comparison of the one-particle excitation spectra obtained with the five approximation schemes a calculation with the NRG is also taken into account. This should give an impression of the exact result, at least in a low-energy regime near $\omega = \mu = 0$.

In figure 12 we present the results for the symmetric SIAM and in figure 13 for the asymmetric one, at $\beta = 1/k_B T = 100$ and $\beta = 150$, respectively, implying T being roughly half T_K, other parameters as before. First looking at parts (a) of these figures, the following fact seems remarkable: peaks in the NRG spectra are considerably broader compared with the other cases and the features due to the van Hove singularities in the band DOS are smeared out much more. At least part of this, in particular at higher excitation energies, should be due to numerical procedures used in the NRG calculation: it is based on a discrete set of energy eigenvalues, which is considerably spaced near the band edges and becomes logarithmically denser for smaller excitation energy ω; interpolation then has a smoothening effect. Whereas near the original resonances at $\omega \leqslant \epsilon_\ell$ and $\omega \geqslant \epsilon_\ell + U$ the five semianalytical approximations are certainly closer to the truth than the NRG curves, the situation is not completely clear in the energy region around the ASR, although here the NRG is most trustworthy.

The peak value close to $\omega = 0$ at $T = 0$ is given via Friedel's sum rule to be $\varrho_{f\sigma}(\omega) = 1/\pi\Delta_A \approx 1.06$; this is faithfully reproduced by the NRG, whereas the width of the ASR might already be somewhat exaggerated by the NRG. In effect, however, we take the NRG-ASR as our measure of quality for the other approximations with respect to the low-energy regime. Regarding the full range of excitation energies, on the other hand, any of the other approximations (except SNCA) might be more appropriate, depending on the purpose of the calculation. With the halfwidth of the ASR taken as

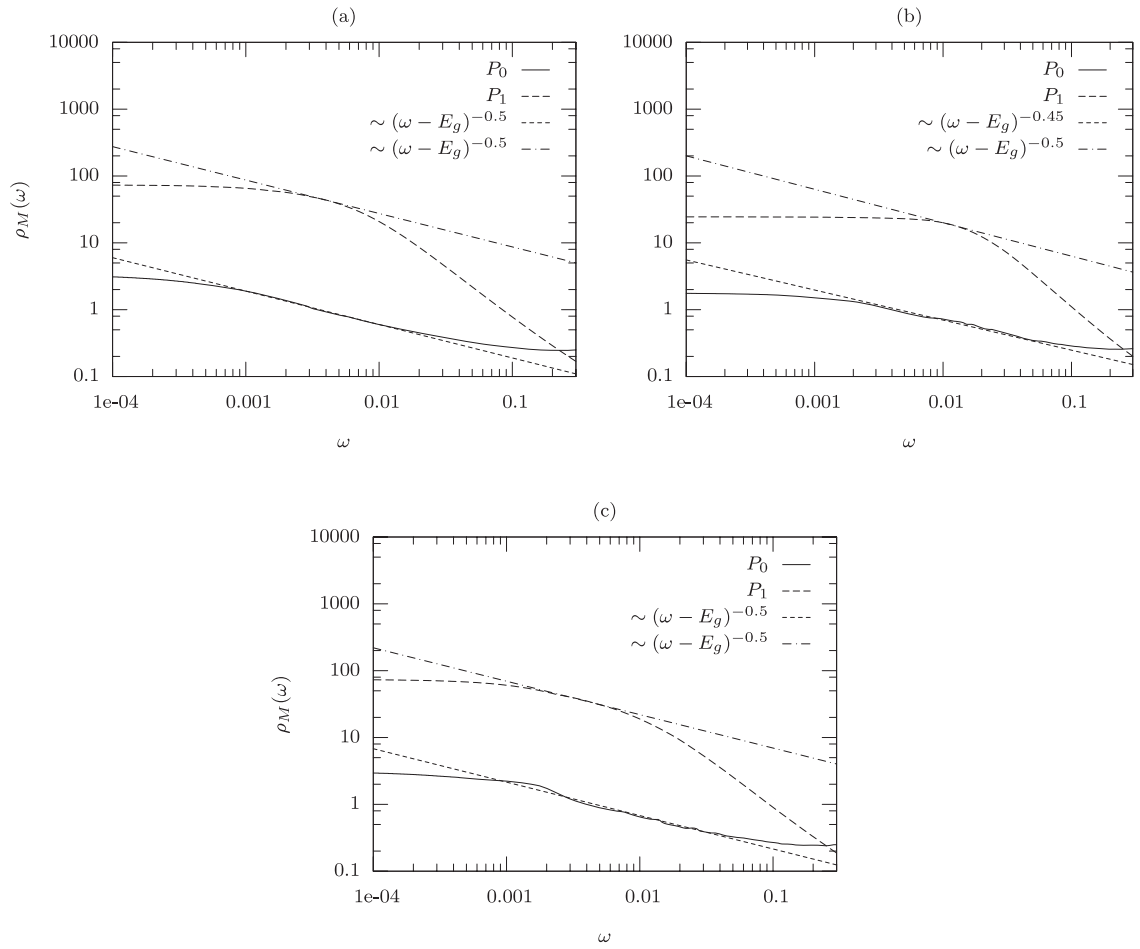

Figure 11. Double-logarithmic plot of ionic spectra, centred at the threshold, for a SIAM in SUNCA (part (a)), FNCA (part (b)) and CA1 (part (c)), all for the asymmetric case with parameter values as in figure 7.

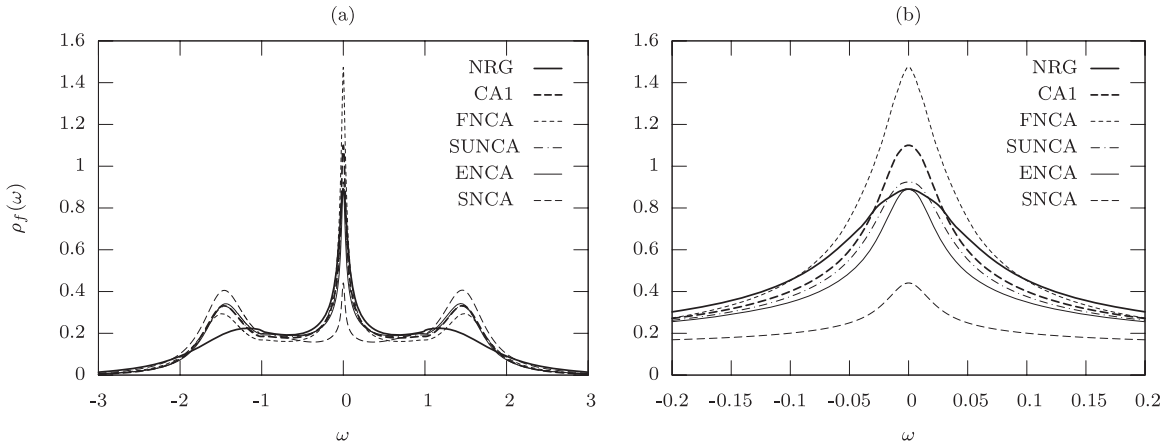

Figure 12. One-particle excitation spectrum of a symmetric SIAM at $\beta = 100$, other parameters as in figure 7(a), in the five semianalytical approximations discussed in the text and, additionally, calculated with the numerical renormalization group (NRG). Part (a) reveals shortcomings of the NRG-method at large excitation energies, whereas in the low-energy region of part (b) the NRG-curve can be used as a reference for the other approximations.

measure for the many-body scale T_K, certainly the FNCA, with its vertex corrections systematically iterated through all orders, compares most favourably with the NRG. On the other hand, the FNCA spectrum clearly exaggerates the height and thus the total weight of the ASR: the limit of $1/\pi\Delta_A$ for the peak height becomes violated even more for lower temperatures. Insofar, CA1 seems to represent a good compromise and even ENCA does not work too bad.

The lesson to be learnt from these results is that improved semianalytical impurity solvers of this type should incorporate

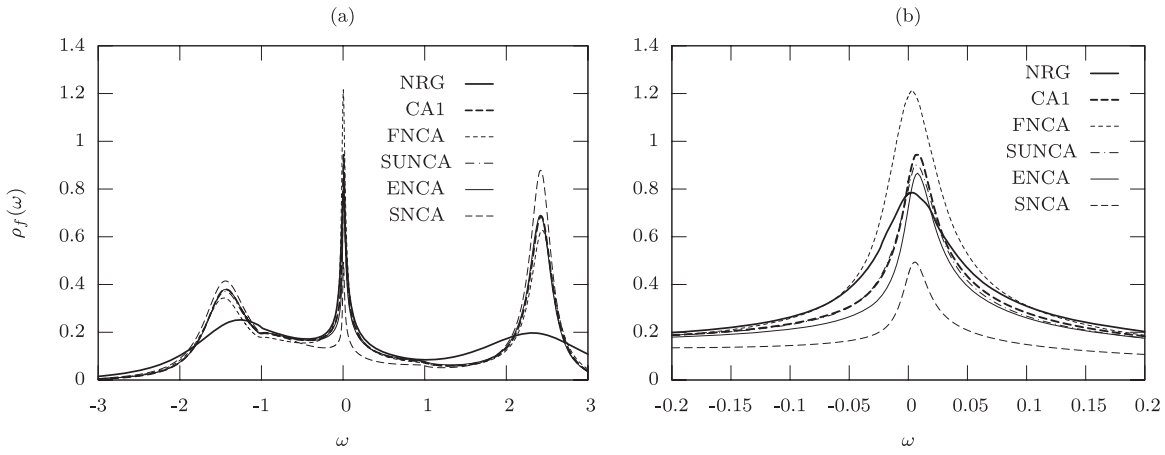

Figure 13. Analogue of figure 12 for the asymmetric case with $\beta = 150$, other parameters as in figure 7(b).

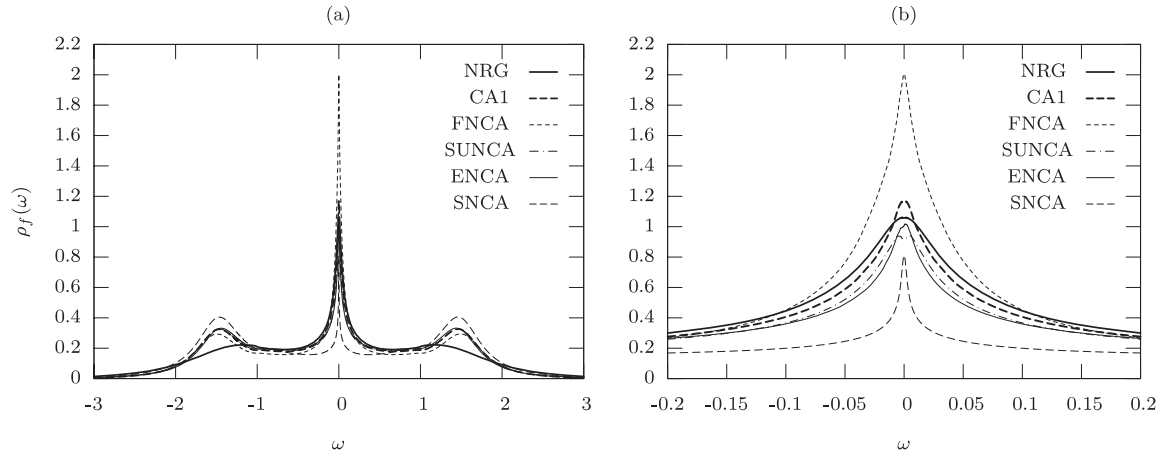

Figure 14. Like figure 12, but with $\beta = 1000$, i.e. at $T \ll T_K$.

skeleton diagrams of two types in a well-balanced way: classes of iterated vertex corrections have to be accompanied by chains of iterated particle-scattering events, being related to each other as exchange partners [37, 44, 46]. In light of the discussion in section 2, CA1 serves as a further step in this direction. Regarding the regime at high excitation energies $\omega \approx \epsilon_\ell + U$ the comparison of figures 12(a) and13(a) reveals a trend, which for even larger values of U becomes more and more pronounced and which apparently is not well captured by the NRG, at least with its present numerical performance: the resonance due to double occupancy of the local shell becomes sharper with increasing U, in particular when $\epsilon_\ell + U$ reaches the order of the upper band edge $\omega \approx \frac{1}{2}W = 3$. Beyond this value, the peak keeps its weight but rapidly loses its width and finally vanishes as a single spectral line out of the accessible region. This is faithfully reproduced by any of the five semianalytical impurity solvers under consideration.

Figure 14 gives an impression about qualities and failures of the five approximations as applied to the full calculation of one-particle spectra at very low temperatures: $\epsilon_\ell = -1.0$, $U = 2.0$ and $\beta = 1000$, i.e. $T/T_K = 0.04$ have been chosen here. Whereas the ionic spectra in all five cases come out rather reliably with the procedures used in our program package, the subsequent folding of ionic propagators and defect propagators (see, e.g., [14]) can produce spurious results near $\omega = 0$. With very sharp thresholds in all quantities at low T slight displacements of the maxima (as a consequence of numerical procedures and rounding errors) can have a large effect on the integrals containing several of these quantities. Although figure 14, too, supports the conclusions drawn before, the SUNCA curve and to a somewhat lesser degree the FNCA curve, show a spurious double-peak structure near $\omega = 0$, supposedly due to such threshold shifts. In addition, the FNCA curve should not be taken too seriously very close to $\omega = 0$, although its shape is in accord with the numerically more precise SNCA calculation in figure 8(a). FNCA overestimates the ASR peak height strongly, whereas CA1, in spite of a too high peak value, rather favourably compares with the NRG curve.

3.4. Fermi-liquid properties

As a final test for our new CA1 the local Fermi-liquid properties are inspected in figure 15. For this purpose the imaginary part of the self-energy

$$-\mathrm{Im}\,\Sigma_{f\sigma}(\omega + \mathrm{i}\delta) = \frac{\pi \varrho_{f\sigma}(\omega)}{(\mathrm{Re}G_{f\sigma}(\omega))^2 + (\pi \varrho_{f\sigma}(\omega))^2} \quad (16)$$

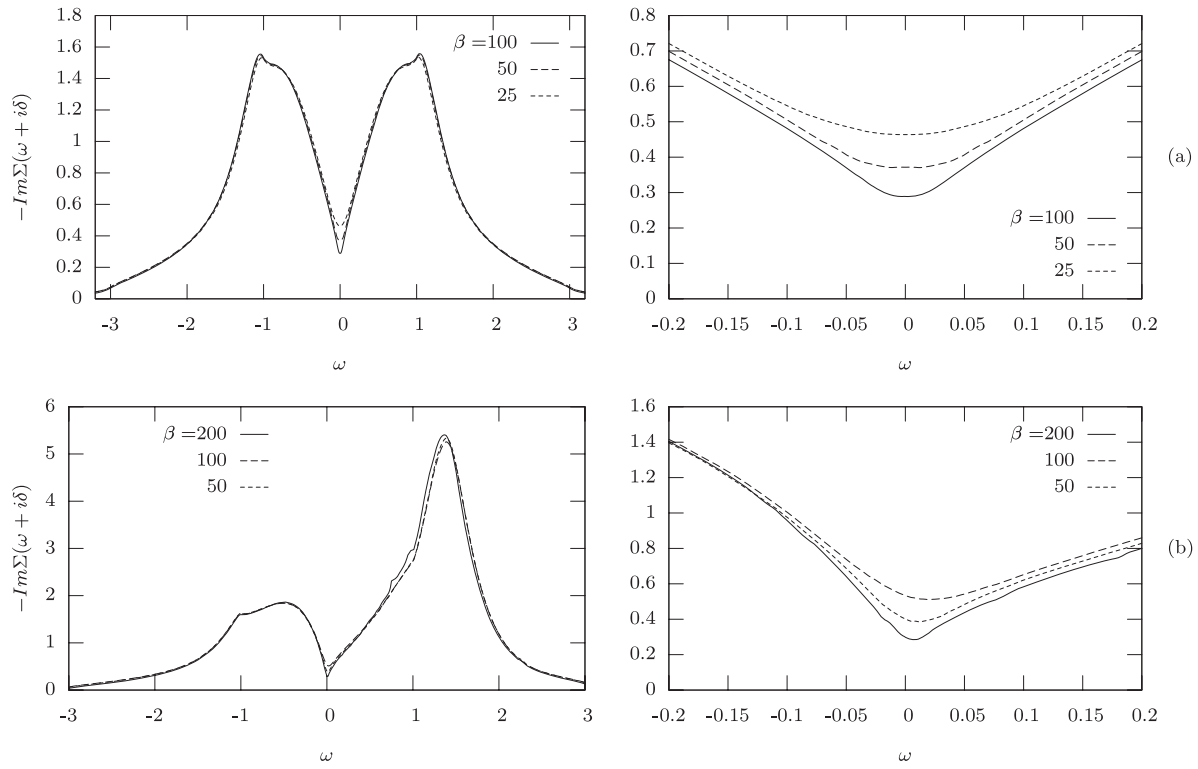

Figure 15. Imaginary part of the self-energy of local electrons (absolute value) for a SIAM, calculated within CA1, as a test for local Fermi-liquid properties. Temperatures as shown, other parameters as in figure 7.

is shown for a few temperatures near and well below T_K. The formation of a minimum at $\omega = 0$ obviously takes place. In the asymmetric case of figure 15(b) a displacement of this minimum away from the Fermi level with growing temperature is recognized, similar but somewhat weaker than has been reported before for (S)NCA and ENCA [37], as well as for the PNCA, the latter being a $U = \infty$ theory with crossing contributions to vertex corrections in high orders [41]. The CA1 result for the value of $-\mathrm{Im}\Sigma_{f\sigma}(\omega + \mathrm{i}\delta)$ at its minimum falls short of the exact limiting value $(\pi\varrho_{f\sigma}(\omega))^{-1} = \Delta_A = 0.3$ for the lowest temperatures. This is in accord with the too high ASR value seen, for example, in figure 14. In a quantitative sense, however, this result improves on ENCA and, much more, on (S)NCA. Obviously, the minimum can well be fitted by a parabola as long as the temperature does not become so low that numerical deficiencies near $\omega = 0$ become predominant. Also its position and height, as well as the quadratic coefficient, may be compared to exact results for Fermi liquids [11, 58], e.g. as a guideline for corrective measures when using these impurity solvers in lattice calculations. These conclusions are similar to those for the PNCA published before [41]. Furthermore, it is a remarkable fact how dramatic the scattering rate raises for increasing excitation energies. In the range of the ionic resonances $\omega = \epsilon_\ell$ and $\omega = \epsilon_\ell + U$ it becomes high enough to completely prevent locally a band picture even for the c-electrons. This will become even more evident in the lattice calculations of the next section, where the effect occurs on each lattice site and thus affects the whole Bloch states.

4. Impurity solvers and lattice theories

4.1. Cumulant perturbation theory of strongly correlated lattice models

The importance of impurity solvers for approximate solutions of lattice problems came to light in theories like ATA [19], LNCA [21] and XNCA [20], which all used a picture of effective sites and relied on NCA as the best available implementation in those days. These three forms of effective site theories aimed at a solution of the Anderson-lattice model in the context of the heavy fermion problem, and thus were based on a particular local shell structure with well localized f states and extended c states, with only the former being subject to a local Coulomb repulsion. In most cases, transfer also was restricted to the c-states only, which together makes possible a reduction from the matrix formulation, envisaged in section 2 and shortly outlined below, to a scalar formalism.

The three theories differed in the way in which the dynamics on general lattice sites influenced the representative effective site considered: in ATA only the coherent build up of quasiparticle bands from scattering by independent Anderson impurities on the lattice sites was taken into account, thus ignoring exhaustion problems and important renormalization effects. LNCA used self-consistently modified local excitations at the effective site; unlike XNCA, however, it introduced weight factors for nonlocal processes in order to approximately factorize the partition function into contributions from an unperturbed band and independent effective sites. XNCA finally established the form of self-consistency between effective site and surrounding medium

which becomes exact in the limit of infinite spatial dimension and which nowadays is regarded as the characteristic defining feature of the DMFT.

In a local approach a lattice Hamiltonian $H = \sum_\nu H_{0\nu} + \sum_{\nu\neq\nu'} V_{\nu,\nu'}$ is built up from local Hamiltonians $H_{0\nu}$ on lattice sites R_ν, each of the type $H_{0\ell}$ considered in section 2, and nonlocal parts $V_{\nu,\nu'}$, which contain one-particle terms like transfer or hybridization and possibly two-particle terms, i.e. nonlocal interactions between electrons on two different sites ν and ν'. $V_{\nu,\nu'}$ is usually expressed via the elementary creation and annihilation operators $f^{(\dagger)}_{\nu m\sigma}$ which define the local Fock spaces, whereas the diagonalization of $H_{0\nu}$ involves the 'ionic states' as described in section 2. In principle there is complete freedom in defining what should be such a local subsystem, a single ionic shell, or an ion or a molecular complex, or even a cluster of ions or complexes. The need to diagonalize them, even when restricted to e.g. the low-energy regime, may however set limitations to the size of what can be regarded as local. At the outcome $H_{0\nu}$ will be a finite matrix in the space of local many-body eigenstates $|\nu\, n_0\, M\rangle$. A one-particle transfer term contained in $V_{\nu,\nu'}$ will, for example, induce changes $|\nu\, n_0\, M_1\rangle \rightarrow |\nu\, n_0 - 1\, M_2\rangle$, $|\nu'\, n_0'\, M_1'\rangle \rightarrow |\nu'\, n_0'+1\, M_2'\rangle$.

A convenient way of keeping the formalism simple is to work generally with the original one-particle quantum numbers (m,σ) and to built up local matrix Green functions of the type $\underline{\underline{G}}_\nu(z) = (G_{m\sigma,m'\sigma'}(z))$ for one-particle propagation, and corresponding higher ones; these matrices involve information about the composition of the local many-body eigenstates, e.g. via the coefficients of fractional occupancy. Thus an overall $(N \times N)$-matrix formalism for lattice processes is established with N being the number of local one-particle states taken into account and local excitations of single electrons are translated via (2) into the dynamics of many-body eigenstates. Nonlocal interactions can be viewed as (simultaneous) two-particle transfers and be handled in an analogous fashion.

The local approach treats the local dynamics exactly, as apparent in the many-body eigenstates. In the presence of strong and dominating local interaction matrix elements this seems a natural starting point for a lattice theory. However, it turns to a disadvantage when a perturbation expansion in terms of the nonlocal parts $V_{\nu,\nu'}$ of the Hamiltonian is to be set up. An attempt to factorize, for example, a general contribution to the partition function with the help of Wick's theorem stops at an intermediate level: in the case of, for example, a pure one-particle nearest-neighbour transfer mechanism, time-ordered expectation values of products of creation and annihilation operators belonging to sites which knot together different propagation paths of single particles, remain as unfactorized parts of internally connected contributions B_λ. These also contain matrix elements t of $V_{\nu,\nu'}$, symmetry factors r_λ, a sign $(-1)^{\chi_\lambda}$, site summations and time integrations, as well as an indicator function F assuming values 0 or 1, which realizes site exclusions between all the B_λ occurring via $n_\lambda = 1$. Schematically:

$$\frac{Z}{Z_0} = \sum_{\{n_\lambda=0,1\}} \left(\prod_{\lambda=1}^{\infty} B_\lambda^{n_\lambda}\right) F(\{n_\lambda\}),$$

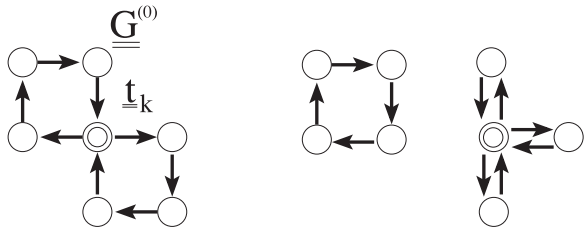

Figure 16. This contribution to the partition function is a product of three disconnected pieces, each of them containing single-particle loops. In two of the pieces loops are glued together at sites (knots), which are marked with a double circle. These nodes give rise to cumulant vertices for local two-and three-particle interactions, respectively.

$$\begin{aligned} B_\lambda = \frac{(-1)^{\chi_\lambda}}{r_\lambda}\Big(& \prod_{\text{sites } \mu \text{ involved as knots}} \int \mathrm{d}\tau_\mu^{(1)} \int \mathrm{d}\tau_\mu^{(1)'} \cdot \ldots \\ & \cdot \langle T\big(f_{m_\mu^{(1)}\sigma_\mu^{(1)}}(\tau_\mu^{(1)})\, f^\dagger_{m_\mu^{(1)'}\sigma_\mu^{(1)'}}(\tau_\mu^{(1)'}) \cdot \ldots\big)\rangle_0 \Big) \\ & \cdot \prod_{\text{chains } \mathcal{C} \text{ of transfers between knots}} \; \prod_{\text{intermediate sites } \nu \text{ along } \mathcal{C}} \\ & \times \int \mathrm{d}\tau_\nu^{(1)} \int \mathrm{d}\tau_\nu^{(1)'} \cdot \ldots \\ & \cdot t \cdot \langle T\big(f_{m_\nu^{(1)}\sigma_\nu^{(1)}}(\tau_\nu^{(1)}) \cdot f^+_{m_\nu^{(1)'}\sigma_\nu^{(1)'}}(\tau_\nu^{(1)'})\big)\rangle_0 \cdot \ldots \cdot t. \end{aligned} \tag{17}$$

In general, local correlations between all of the excitations caused at the nodes by intersite transfers remain. The emerging picture is that of the disconnected sets of loops over the lattice, each set being internally glued together at certain sites with four or more intersite transfer legs, see figure 16 for a simple example. Moreover, while keeping the topological structure of such a graph, the position of sites involved cannot be summed freely over the lattice, thus preventing a convenient momentum space formulation. As a consequence, a linked cluster theorem is also not available for Green functions of the lattice problem, since a partition function, incorporating all diagrams not linked to external sources, cannot be factored out: the partition function can also not be represented as an exponential of a sum of single connected graphs, thus preventing a straightforward extensivity property of the thermodynamic potential.

A way out of this dilemma uses a representation of higher-order local time-ordered expectation values at nodes as a sum over products of successively smaller ones, which in total cancel out except for the original highest term; when appropriately grouped together they define a cumulant expansion of local n-particle Green functions, containing a set M_n of n local destruction operators and a set $\widetilde{M}_n$ of creation operators:

$$\begin{aligned} G_n(M_n;\widetilde{M}_n) = & \sum_{\text{all partitions } P \text{ of } M_n \text{ and } \widetilde{P} \text{ of } \widetilde{M}_n \text{ into subsets } \left(N_{n_q}^{(q)},\widetilde{N}_{n_q}^{(q)}\right)} \\ & \times (-1)^{\chi_P+\chi_{\widetilde{P}}} \prod_q G^c_{n_q}(N_{n_q}^{(q)},\widetilde{N}_{n_q}^{(q)}), \\ G_2^c\big(f_{m_\mu^{(1)}\sigma_\mu^{(1)}}(\tau_\mu^{(1)}),\; & f_{m_\mu^{(2)}\sigma_\mu^{(2)}}(\tau_\mu^{(2)}); \\ & f^\dagger_{m_\mu^{(2)'}\sigma_\mu^{(2)'}}(\tau_\mu^{(2)'}),\; f^\dagger_{m_\mu^{(1)'}\sigma_\mu^{(1)'}}(\tau_\mu^{(1)'})\big) \end{aligned}$$

$$= G_2(1,2;2',1') - [G_1(1;1')G_1(2;2') - G_1(1;2')G_1(2;1')], \ldots . \quad (18)$$

The right-hand side of this expression can be viewed as containing two contributions, each playing a different role in the expansion: the maximally decomposed terms, being products of only one-particle Green functions at this site, together just furnish the result which would be obtained if Wick's theorem were valid for the local dynamics. The rest of the terms represent all possible local decompositions of the knot in the graph into products of independent irreducible knots of equal or lower order with together the same number of intersite legs as the original expectation value; among them are possibly local one-particle Green functions but at least one of higher order, i.e. with more than two intersite legs.

An evident approximation can be obtained as follows. Keeping only the maximally decomposed terms, all loops of the diagram, which were glued together originally, would be disconnected, the partition function would become an exponential of the sum of all different loops, and a free summation of all site positions over the lattice would leave only contributions differing by the number of links. Likewise, for example, the lattice-one-particle Green function would contain, after dividing out the partition function, a sum over paths differing only by the number of links, which can be summed as a geometric series (in matrix space).

The results for a one-particle Green function and thermodynamic potential F can be given explicitly as

$$\underline{\underline{G}}^{\mathrm{FT}}_{\underline{k}}(z) = [\underline{\underline{G}}^{(0)}(z)^{-1} - \underline{\underline{t}}_{\underline{k}}]^{-1},$$

$$F^{\mathrm{FT}} - F^{(0)} = \frac{1}{\beta}\sum_{\mathrm{i}\omega_n}\sum_{\underline{k}}(\mathrm{Tr}[\ln \underline{\underline{G}}^{(\mathrm{FT})}_{\underline{k}}(\mathrm{i}\omega_n)] - \mathrm{Tr}[\ln \underline{\underline{G}}^{(0)}(\mathrm{i}\omega_n)]), \quad (19)$$

with $\underline{\underline{G}}^{(0)}(z)$ the Green function from the known solution of the isolated local subsystem. The above results may be viewed as a very general form of the well-known Hubbard-I approximation [53], to which they reduce when the local subsystem is a simple s-shell and $V_{\nu,\nu'}$ contains only nearest-neighbour one-particle transfers i.e. in the case of the Hubbard model.

The rest of the terms in the cumulant-decomposition furnishes all possible combinations of one-particle scattering events of this site, serving as nodes with two intersite legs in simple transfer chains, and of higher-order nodes with more intersite legs; together they realize a restricted way of glueing together the original connected set at this site. The contributions resulting from all of these local cumulant decompositions in the set can be re-interpreted in terms of a perturbation expansion with respect to an infinite set of local n-particle cumulant interactions, with n between 2 and infinity.

The 'noninteracting' starting point of this expansion is the aforementioned generalized Hubbard-I-theory free of such interactions ('free theory'). A perturbation expansion with the corresponding 'free' propagators for the particles and the set of all local cumulant vertices as interaction terms along the conventional lines a la Feynman thus faithfully produces all contributions to partition function and lattice Green functions and re-introduces the applicability of the linked cluster theorem and all benefits connected with it into this new form of the theory.

The price paid for the conceptual progress described above for the lattice problem, i.e. the applicability of conventional methods in perturbation theory for the lattice aspects, lies in the large number of vertices appearing and in their dynamical nature, i.e. their multiple time dependences. It should also be clear that the perturbation series obtained in this way are based, although looking conventional in a superficial way, on quite unconventional definitions of connectedness and irreducibility: whereas the diagrams remaining i.e. for a one-particle Green function are linked to the external sources and consist of one connected piece (we exclude anomalous terms here), being glued together via local cumulant vertices, they would in most cases fall apart into several unconnected pieces, only one of them bearing the two external links, when the cumulants are made explicit in terms of the original expectation values.

The pieces without links remaining in this case as factors are partially due to the partition function in the denominator of the original expression and must consistently be included for a proper renormalization of, for example, excitation spectra. It is therefore not a trivial problem to define consistent approximations in cumulant perturbation theory.

A straightforward evaluation of cumulant vertices even in low orders requires some effort but can be managed e.g. with methods of direct perturbation theory [54]. The somewhat lengthy expression of the local two-particle cumulant has been published before [55] and served as basis for calculations of one-particle properties of the Hubbard model along the lines of Hartree-like expansions [56], also in a self-consistent fashion [54]. It should be clear, and can in fact be proven [54], that these approximations are not able to capture the interesting many-body aspects in the low temperature regime for large values of the local Coulomb repulsion U, whereas leading effects of band splitting, which are present in principle already in the Hubbard-I approximation, and of band deformation can be captured. What is needed for a proper approximation of strong correlation effects are summations of infinitely many processes with cumulant vertices up to infinite order to cope with the problem of long-time decay of correlations and of the infrared problems connected with them.

The local starting point of the cumulant expansion and its formulation in real space open the possibility of a local infinite order resummation, which would be more difficult to recognize after unrestricted Fourier-transformation. In the latter $\underline{k}$-space version of cumulant perturbation theory, due to the unrestricted site summations, processes taking place on the same site in a diagram cannot be identified anymore and may appear completely uncorrelated.

Moreover, also individual parts of the partition function in the original denominator connected to the same site and necessary for a proper local normalization are hidden in the cumulant and cannot readily be identified. A proper conserving approximation of infinite order should unite such pieces of local processes in a consistent way. This furnishes the guideline for a 'locally complete approximation': collect all those diagrammatic contributions to the one-particle-Green

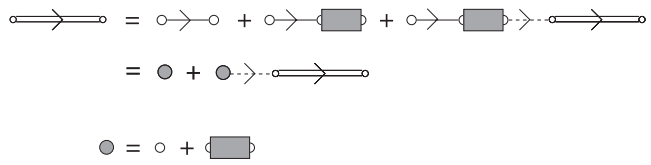

Figure 17. Different forms of the Dyson equation for the case of one-particle transfer only, in cumulant perturbation theory. Care has to be taken in specifying nonlocal connections, which is done here via the broken transfer lines being made explicit. Single unbroken lines relate to the propagators in free theory, the hatched rectangular box is one-particle irreducible in cumulant (and not in regular) perturbation theory.

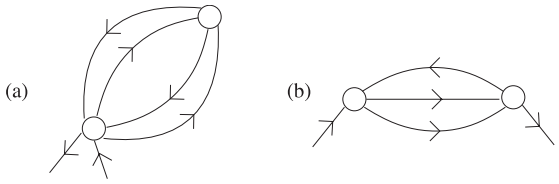

Figure 18. Contributions neglected in the self-energy matrix in the 'locally complete' (LC) approximation, one-particle transfer only. Part (a) shows a process, in which two loops, based at the same site, are correlated at a different site by a two-particle cumulant vertex. Part (b) shows a process in which the external particle is extracted at a site different from where it was injected.

function of cumulant perturbation theory in real space, where both external legs belong to the same site, regardless whether they belong to the same or to different vertices situated at this site.

The formulation of a Dyson equation for the one-particle-Green function in cumulant perturbation theory is straightforward, but needs some care in distinguishing local and nonlocal parts of the propagation process: certain intersite transfers have to be made explicit. In a straightforward way one-particle-irreducible pieces can be identified, which either are linked to external sources or to transfers. It is useful to amputate local factors $\underline{\underline{G}}^{(0)}(z)$ (the Green functions of isolated local subsystems) at their two ends, thus defining irreducible cumulant self-energies $\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)$. Since the two ends can be situated at two different sites, they will generally be $\underline{k}$-dependent after Fourier-transformation. Two equivalent forms of the Dyson-equation are visualized in figure 17; their algebraic form is:

$$\begin{aligned}\underline{\underline{G}}_{\underline{k}}(z) &= \underline{\underline{G}}^{(\mathrm{FT})}_{\underline{k}}(z) + \underline{\underline{G}}^{(\mathrm{FT})}_{\underline{k}}(z)\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)\underline{\underline{G}}^{(0)}(z) \\ &\quad + \underline{\underline{G}}^{(\mathrm{FT})}_{\underline{k}}(z)\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)\underline{\underline{G}}^{(0)}(z)\underline{\underline{t}}_{\underline{k}}\underline{\underline{G}}_{\underline{k}}(z) \\ &= \underline{\underline{\tilde{G}}}_{\underline{k}}(z) + \underline{\underline{\tilde{G}}}_{\underline{k}}(z)\underline{\underline{t}}_{\underline{k}}\underline{\underline{G}}_{\underline{k}}(z) \\ &\Rightarrow \underline{\underline{G}}_{\underline{k}}(z) = \left[\underline{\underline{\tilde{G}}}_{\underline{k}}(z)^{-1} - \underline{\underline{t}}_{\underline{k}}\right]^{-1}, \end{aligned} \tag{20}$$

with

$$\underline{\underline{G}}^{(\mathrm{FT})}_{\underline{k}}(z) = \left[\underline{\underline{G}}^{(0)}(z)^{-1} - \underline{\underline{t}}_{\underline{k}}\right]^{-1} \equiv \left[z - \underline{\underline{\Sigma}}^{(0)}(z) - \underline{\underline{t}}_{\underline{k}}\right]^{-1} \tag{21}$$

$$\underline{\underline{\tilde{G}}}_{\underline{k}}(z) = \underline{\underline{G}}^{(0)}(z) + \underline{\underline{G}}^{(0)}(z)\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)\underline{\underline{G}}^{(0)}(z), \tag{22}$$

where $\underline{\underline{\Sigma}}^{(0)}(z)$ is known from the solution of the local subsystem. The irreducible cumulant self-energy $\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)$ defined above must not be confused with the standard self-energy $\underline{\underline{\Sigma}}^{(\mathrm{st})}_{\underline{k}}(z)$ defined via

$$\underline{\underline{G}}_{\underline{k}}(z) = \left[z - \underline{\underline{\Sigma}}^{(\mathrm{st})}_{\underline{k}}(z) - \underline{\underline{t}}_{\underline{k}}\right]^{-1}. \tag{23}$$

The connection between them, which can be expressed as

$$\underline{\underline{\Sigma}}^{(\mathrm{st})}_{\underline{k}}(z) = \underline{\underline{\Sigma}}^{(0)}(z) + \underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)\left[1 + \underline{\underline{G}}^{(0)}(z)\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)\right]^{-1}, \tag{24}$$

sheds some light on the possible momentum dependence of the self-energy. We expect that this may contribute to the question of how to incorporate nonlocal correlations into lattice theories (see [27, 60] and references therein).

In a local approximation, the two external links of $\underline{\underline{\Sigma}}^{(\mathrm{amp})}_{\underline{k}}(z)$ are restricted to be situated at the same site, which eliminates the $\underline{k}$ dependence from $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$. This type of approximation is very much in the tradition of the early effective site theories mentioned above. It brings formal advantages but involves shortcomings concerning the neglect of certain nonlocal correlations.

Still it does not lead to an easily tractable calculational scheme, which is due to classes of remaining nonlocal correlations in $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$ as indicated in figure 18(a); part (b) of this figure, on the other hand, shows correlations not included in the local form of $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$. Eliminating diagrams like the one shown in figure 18(a) leads to a restricted form $\underline{\underline{\Sigma}}^{(\mathrm{lc})}(z)$ of $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$, which is at least 'locally complete' and may be characterized as follows: In a diagram contributing to $\underline{\underline{\Sigma}}^{(\mathrm{lc})}(z)$ cumulant vertices of order $n \geqslant 2$ (four or more external links) are forbidden to appear, since they would correlate two or more one-particle loops based at the same site. Therefore one-particle loops generally will contain insertions with two external links (and no more), which are of the same class as those contributing to $\underline{\underline{\Sigma}}^{(\mathrm{lc})}(z)$; in this way a complete hierarchical structure of independent loops remains.

Amazingly, the locally complete approximation thus defined can quite generally be brought into a form suited for a straightforward solution, in principle without further approximations or restrictions. It turns out to be equivalent to general formulations of the XNCA or the DMFT methods [20, 54]. The key to a simpler formulation of the locally complete approximation of cumulant perturbation theory lies in the reduction of all nonlocal topological elements in a diagram to independent one-particle loops as described above. This makes it possible to trace back the $\underline{k}$-summed form of (20), i.e.

$$\underline{\underline{G}}(z) \equiv \frac{1}{N}\sum_{\underline{k}} \underline{\underline{G}}_{\underline{k}}(z) = \frac{1}{N}\sum_{\underline{k}} \left[\underline{\underline{\tilde{G}}}(z)^{-1} - \underline{\underline{t}}_{\underline{k}}\right]^{-1}, \tag{25}$$

to the irreducible part $\underline{\underline{\tilde{T}}}(z)$ of the (unrestricted) loop propagator

$$\underline{\underline{T}}(z) = \frac{1}{N}\sum_{\underline{k}} \underline{\underline{T}}_{\underline{k}}(z), \qquad \underline{\underline{T}}_{\underline{k}}(z) = \underline{\underline{t}}_{\underline{k}}\underline{\underline{G}}_{\underline{k}}(z)\underline{\underline{t}}_{\underline{k}}. \tag{26}$$

It turns out that $\underline{\underline{G}}(z)$, as the local one-particle-Green function (i.e. both external sources are at the same site), can be constructed as a functional $\underline{\underline{G}}[z; \underline{\underline{\varrho}}_{\widetilde{T}}(\omega)]$ of the irreducible loop spectrum $\underline{\underline{\varrho}}_{\widetilde{T}}(\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,\underline{\underline{\widetilde{T}}}(\omega + \mathrm{i}\delta)$ and $\underline{\underline{\widetilde{G}}}(\underline{\underline{G}}, \underline{\underline{\widetilde{T}}})$ has a simple expression via $\underline{\underline{G}}$ and $\underline{\underline{\widetilde{T}}}$. Therefore (26) reduces to an implicit equation for $\underline{\underline{\widetilde{T}}}(z)$ and consequently also furnishes solutions for $\underline{\underline{\widetilde{G}}}(z)$, $\underline{\underline{G}}_{\underline{k}}(z)$ and $\underline{\underline{G}}(z)$.

The original loop propagator $\underline{\underline{T}}(z)$ embodies the possibility that the loop connects to its basic site at several intermediate instances due to the unrestricted site summations contributing to $\underline{\underline{G}}_{\underline{k}}(z)$ in cumulant perturbation theory. The whole loop may be viewed as a repetition of irreducible pieces, i.e. loops without such intermediate connections to its basic site. The relation between both loop propagators takes the form of a Dyson equation:

$$\underline{\underline{T}}(z) = \underline{\underline{\widetilde{T}}}(z) + \underline{\underline{T}}(z)\underline{\underline{\widetilde{G}}}(z)\underline{\underline{\widetilde{T}}}(z) \Rightarrow \underline{\underline{\widetilde{G}}}(z) = \underline{\underline{\widetilde{T}}}(z)^{-1} - \underline{\underline{T}}(z)^{-1}, \tag{27}$$

which was first recognized in [57]. Equation (27) generalizes earlier attempts for implementing a 'site-exclusion principle' for the propagation of quasiparticles under the influence of strong local correlations [3].

It is easy to visualize the effect of loops on the dynamics of the basic site of $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$, i.e. the one bearing the two external links, when, for example, the picture suggested by direct perturbation theory is used. Apart from the fact that a projection of $\underline{\underline{\varrho}}_{\widetilde{T}}(\omega)$ onto a local 1-P-state is used in all formulae instead of the spectrum $\varrho^{(0)}_{c\sigma}(\omega)$ of band electrons, the way to calculate $\underline{\underline{G}}(z)$ is the same as in the impurity problem considered in section 2.

Insofar, an effective impurity is constructed, and the spectrum of the irreducible loop propagator may be viewed as a matrix of real, frequency-dependent effective external fields, constituting a 'bath' or a 'dynamical mean field'. Conclusions about $\underline{\underline{\widetilde{G}}}(z)$ are to be traced back to the definition $\underline{\underline{\widetilde{G}}}(z) = \underline{\underline{G}}^{(0)}(z) + \Delta\underline{\underline{\widetilde{G}}}(z)$, $\Delta\underline{\underline{\widetilde{G}}}(z) = \underline{\underline{G}}^{(0)}(z)\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)\underline{\underline{G}}^{(0)}(z)$ in (20), reproduced here for the locally complete approximation.

At first glance it may seem as if the diagrams contributing to $\underline{\underline{\widetilde{G}}}(z)$ in cumulant perturbation theory via $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$ would just reproduce the original contributions of the effective impurity problem. However, this cannot be true, since the introduction of cumulants also serves the purpose of removing restrictions from site summations, giving $\underline{\underline{G}}_{\underline{k}}(z)$, the simple form used in (25). $\underline{\underline{\widetilde{G}}}(z)$ thus contains compensation terms for loop contributions produced by $\underline{\underline{t}}_{\underline{k}}$ in the denominator, i.e. in the corresponding geometric series with local parts $\underline{\underline{\widetilde{G}}}(z)$ and links $\underline{\underline{t}}_{\underline{k}}$. One possible way of uncovering the relation between $\underline{\underline{G}}(z)$ and $\underline{\underline{\widetilde{G}}}(z)$ consists in formulating the difference between both quantities just as the contribution for loops to be compensated, i.e.[3]

$$\underline{\underline{\widetilde{G}}}(z) - \underline{\underline{G}}(z) = -\underline{\underline{\widetilde{G}}}(z)\underline{\underline{T}}(z)\underline{\underline{\widetilde{G}}}(z). \tag{28}$$

[3] An alternative statement of this relation is

$$\underline{\underline{\widetilde{G}}}(z)\underline{\underline{T}}(z) = \underline{\underline{G}}(z)\underline{\underline{\widetilde{T}}}(z)\,,$$

which is realized by analysing the expansions of $\underline{\underline{G}}(z)$ and $\underline{\underline{T}}(z)$ in terms of irreducible loops $\underline{\underline{\widetilde{T}}}(z)$:

When (27) is inserted into (28) to eliminate $\underline{\underline{T}}(z)$ in favour of $\underline{\underline{\widetilde{T}}}(z)$, one obtains

$$\underline{\underline{\widetilde{G}}}(z) = \left[\underline{\underline{G}}(z)^{-1} + \underline{\underline{\widetilde{T}}}(z)\right]^{-1} \equiv \underline{\underline{\widetilde{G}}}(\underline{\underline{G}}, \underline{\underline{\widetilde{T}}})(z), \tag{29}$$

thus completing the reduction of (25) to an implicit equation for $\underline{\underline{\widetilde{T}}}$ as envisaged above.

It should finally be remarked that the contributions in cumulant perturbation theory to the quantities considered here, which may be classified as connected pieces not linked to the external sources in the original picture, are absorbed in the proper normalization of spectra and one-particle Green functions; they originate from a division by the partition function as explained above. Using a consistent locally complete summation of these contributions, they become absorbed in the partition function of the effective site problem. In direct perturbation theory, for example, this is taken into account by properly normalized defect propagators.

4.2. Periodic Anderson model: part I

The formal development outlined above leads to a result, which constitutes a matrix generalization of DMFT. The original concern of LNCA and XNCA was the physics of the Anderson-lattice model:

$$\hat{H} = \sum_{\sigma,\nu}\left(\epsilon_\ell\, \hat{f}^\dagger_{\nu\sigma}\hat{f}_{\nu\sigma} + \frac{U}{2}\,\hat{n}^f_{\nu\sigma}\hat{n}^f_{\nu\bar{\sigma}}\right) + \sum_{\underline{k},\sigma} t_{\underline{k}}\,\hat{c}^\dagger_{\underline{k}\sigma}\hat{c}_{\underline{k}\sigma} + \sum_{\underline{k},\sigma}\left(V_{\underline{k}}\,\hat{c}^\dagger_{\underline{k}\sigma}\hat{f}_{\underline{k}\sigma} + \mathrm{h.c.}\right) \tag{30}$$

and it should shortly be explained how a scalar form of the equations is achieved for this case. The local subsystem here involves for the simplest case a basis of four one-particle states on a given lattice site ν, the two f states with spin up and down subject to the Coulomb repulsion U, and two c states which do not interact with each other or with the f states. All matrices considered above are four by four, with $\underline{\underline{t}}$ being spin-diagonal and transferring only c and f electrons to nearest-neighbour c states with respective matrix elements t and V.

Although unphysical in most cases, a purely local hybridization V is often used for simplicity; it can be treated in close analogy to the nearest-neighbour case for the reason explained in the following. Since the c electrons remain noninteracting, Wick's theorem can be used for them. Consequently, no cumulant vertices of order $n \geqslant 2$ exist with links to c Green functions. $\underline{\underline{G}}^{(0)}(z)$ is block-diagonal with a diagonal c block and a f block, and likewise is $\underline{\underline{G}}^{(0)}(z)^{-1}$, which is used in the amputation of vertices.

The block structure mentioned is 2×2 with respect to spin; without magnetic splitting and with spin-preserving transfer and hybridization one may fix a spin direction σ and treat these blocks as scalars. As a consequence, $\underline{\underline{\Sigma}}^{(\mathrm{amp})}(z)$ and also $\Delta\underline{\underline{\widetilde{G}}}(z)$ have nonzero matrix elements only in the diagonal, i.e. only $\widetilde{G}_{ff\sigma}$ and $\widetilde{G}_{cc\sigma}(z) = \widetilde{G}^0_{cc\sigma}(z)$ enter the calculation according to the definition in (20). Nondiagonal elements come into play only via $\underline{\underline{\widetilde{T}}}(z)$, since propagation along a loop can mix c and f

states. Equation (29) now gives after matrix inversions:

$$\begin{aligned}\widetilde{T}_{ff\sigma}(z) &= \widetilde{G}_{ff\sigma}(z)^{-1} - \frac{G_{cc\sigma}(z)}{N_\sigma(z)},\\ \widetilde{T}_{cc\sigma}(z) &= \widetilde{G}_{cc\sigma}(z)^{-1} - \frac{G_{ff\sigma}(z)}{N_\sigma(z)},\\ \widetilde{T}_{cf\sigma}(z) &= \widetilde{T}_{fc\sigma}(z) = \frac{G_{cf\sigma}(z)}{N_\sigma(z)},\\ N_\sigma(z) &= G_{ff\sigma}(z)G_{cc\sigma}(z) - G_{cf\sigma}(z)^2.\end{aligned} \tag{31}$$

These matrix elements would all have to be used if the four states were locally correlated. However, in the simple form of the Anderson model the local c states do not directly influence the f-state dynamics. Therefore one can reduce the local problem to one of f states by combining site-irreducible loops in a way that only a restricted form of irreducibility with respect to the f states on the basic site is realized, i.e. the c level on this site is treated like a different site.

The combined loop propagator connects local f states and obeys a generalized Dyson equation of the form $\widetilde{T}^{(\mathrm{red})}_{ff\sigma}(z) = \widetilde{T}_{ff} + \widetilde{T}_{fc}\widetilde{G}_{cc}\widetilde{T}_{cf} + \widetilde{T}_{fc}\widetilde{G}_{cc}\widetilde{T}_{cc}\widetilde{G}_{cc}\widetilde{T}_{cf} + \cdots$ and hence may be summed up to

$$\widetilde{T}^{(\mathrm{red})}_{ff\sigma}(z) = \widetilde{T}_{ff\sigma}(z) + \frac{\widetilde{T}_{cf\sigma}(z)^2\widetilde{G}_{cc\sigma}(z)}{1 - \widetilde{G}_{cc\sigma}(z)\widetilde{T}_{cc\sigma}(z)}\,. \tag{32}$$

If now the quantities $\widetilde{T}$ on the rhs are replaced via (31), one obtains the scalar equivalent to (29):

$$\widetilde{T}^{(\mathrm{red})}_{ff\sigma}(z) = \widetilde{G}_{ff\sigma}(z)^{-1} - G_{ff\sigma}(z)^{-1}\,. \tag{33}$$

As is clear from this construction, the spectrum of $\widetilde{T}^{(\mathrm{red})}$ is to be used in the effective site problem, i.e. $G_{ff\sigma}(z) = G_{ff\sigma}[z; \varrho_{\widetilde{T}^{(\mathrm{red})}_{ff\sigma}}(\omega)]$.

Also the inverse of the last term in (33), i.e. $G_{ff\sigma}(z)$, has to be calculated via (25) using a matrix inversion, which involves the nondiagonal transfer matrix $\underline{\underline{t}}_{\underline{k}} = t_{\underline{k}}\,\underline{\underline{1}} + V_{\underline{k}}\,\underline{\underline{\sigma}}_x$. After explicitly formulating this step the reduction from a matrix to a scalar form of the theory for the Anderson-lattice model is completed.

Whereas there might exist easier ways of setting up the XNCA/DMFT-self-consistency cycle for this model [20, 62], the above argument generally demonstrates the connection between the universal matrix formulation and possible reduced schemes. If, for example, electrons in c states would interact locally (but again not with those on f states), the matrix problem would reduce to two scalar problems of type (28), which would be coupled only via the lattice summation, i.e. via the $\underline{k}$ sums in (25). This also points to a possible treatment of more general models with inequivalent sites.

Finally, since formally the local c state acts like a different site on its f states, a local hybridization can be treated in the same way as outlined above: the first and last transfer step, represented in (26) by factors $\underline{\underline{t}}_{\underline{k}}$, now carries a $\underline{k}$-independent matrix element V. This only enters the calculation in a modified $\underline{\underline{t}}_{\underline{k}} = t_{\underline{k}}\,\underline{\underline{1}} + V\underline{\underline{\sigma}}_x$ to be used in (25).

The Anderson-lattice model furnishes a good testing ground for the quality of impurity solvers because of the particular impact of coherence in the half-filled case. In the symmetric situation of the simple version without orbital degeneracy considered above, with two electrons per site, the Luttinger theorem predicts a Fermi surface filling the whole first Brillouin zone, which should lead to the formation of an excitation gap due to Bragg scattering in the quasiparticle DOS for temperature T approaching zero [59, 61]. This signature of onsetting coherence is hard to recover in approximations, since it requires a pronounced structure in the self-energy.

On the one hand, the increasing lifetime of quasiparticles with $T \to 0$ near the Fermi level requires the (near) cancellation of the term Δ_{A} (see section 5) in the local self-energy by the buildup of scattering during propagation along loops through the lattice (see above), and on the other hand, formation of a gap should go along with a narrow peak in $\mathrm{Im}\,\Sigma$ signalling strongly increased resonant scattering. This implies a fine balance between different contributions, which is easily destroyed by inconsistent approximations.

4.3. Hubbard model

In order to elucidate the effect of increasing lifetime as $T \to 0$ one may consider the Hubbard model:

$$\hat{H} = \sum_{\sigma,\underline{k}} t_{\underline{k}}\,\hat{c}^\dagger_{\underline{k}\sigma}\hat{c}_{\underline{k}\sigma} + \frac{U}{2}\sum_{\sigma,\nu}\hat{n}^c_{\nu\sigma}\hat{n}^c_{\nu\bar{\sigma}} \tag{34}$$

where at half-filling and zero temperature scattering should be absent near $\omega = 0$ in the Fermi-liquid phase, but which should not develop a coherence gap: since according to Luttinger's theorem the Fermi surface now lies well inside the first Brillouin zone no Bragg scattering should be effective there.

Figure 19(a) shows the corresponding local DOS, obtained within different approximations, two of them within the locally complete scheme (DMFT) using SNCA and ENCA as impurity solvers, respectively.

A reduction process from the matrix-formalism analogue to the Anderson-lattice model is not necessary for the Hubbard model, since the latter is of scalar type from the outset.

The other approximations are Hartree–Fock, effectively meaning $U = 0$ in the half-filled case, and Hubbard-I ('free theory'). The last two cases show no temperature dependence, whereas $\beta = 100$ and $\beta = 10$ are chosen for SNCA and ENCA in order to produce comparable heights of the many-body resonance, which should reach the Hartree–Fock value at $T = 0$.

In a $\underline{k}$ resolution this DOS produces the quasiparticle band structure, as shown for the ENCA calculation in figure 19(b) along the [111] direction of the simple-cubic Brillouin zone. At high excitation energies the two split bands, which in Hubbard-I approximation contain sharp resonances with reduced spectral weight are so much washed out that a concept of band electrons can hardly be justified here; this is in accord with the conclusion in the last section about local scattering near an impurity. The narrow band of pronounced quasiparticle resonances around $\omega = \mu = 0$ shows no splitting and gap

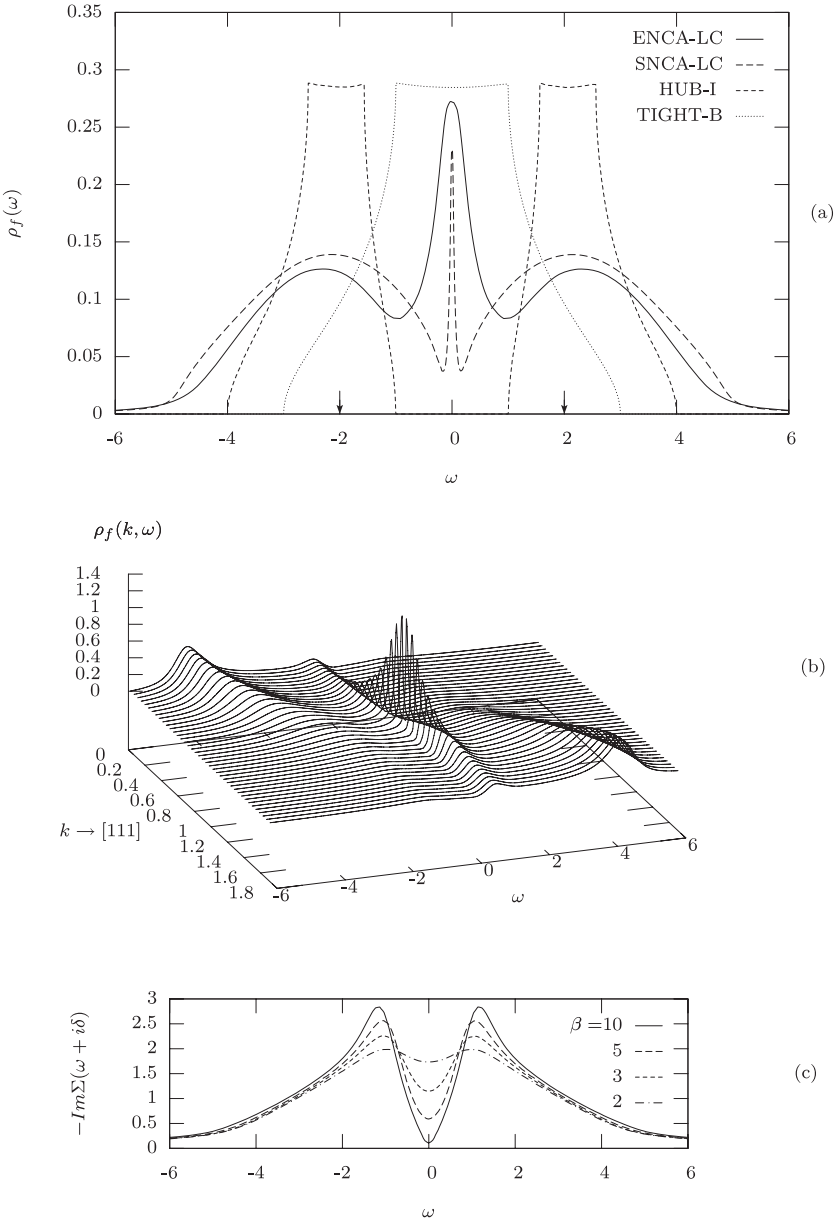

Figure 19. One-particle excitation spectrum of the Hubbard model, calculated within XNCA/DMFT, using ENCA as an impurity solver, for the half-filled case with nearest-neighbour hopping in a 3d-sc lattice. Parameter values $\epsilon = -2$, $U = 4$ and β not too large favour a metallic phase with a Fermi liquid. In part (a) the local DOS is compared for different approximation schemes, also including Hartree–Fock (resulting in a tight-binding band of width 6) and the Hubbard-I approximation. Beside ENCA with $\beta = 10$ also SNCA with $\beta = 100$ is used as impurity solver. The arrows indicate the positions of the ionic levels ϵ and $\epsilon + U$. Part (b) shows the $\underline{k}$-resolved excitation spectrum along the [111] direction, from which a quasiparticle band structure may be derived within XNCA/DMFT-ENCA. The imaginary part of the self-energy (absolute value) in part (c) visualizes the formation of the Fermi liquid with decreasing temperature (k_B)$T = \beta^{-1}$.

formation. The corresponding decrease of scattering is shown in figure 19(c): in lowering the temperature from $\beta = 2$ to 10 the imaginary part of the self-energy forms a steep and nearly quadratic minimum as a sign of Fermi-liquid formation.

Quite generally it can be stated that the ENCA impurity solver for the Hubbard model produces good results and accomplishes an acceptable tradeoff between accuracy and numerical effort. For not too low temperatures it thus represents a reliable and usable alternative to impurity solvers like NRG, QMC, MPT, etc, known from the literature (see, for example, [22, 63]).

4.4. Periodic Anderson model: part II

The Anderson-lattice model behaves differently than the Hubbard model, as shown in figures 20 and 21. Parts 20(a) and (b) contain the local one-particle excitation spectra for c and f electrons, respectively, each calculated with three approximations of increasing complexity. The tight-binding approach for the band (c) states treats all interactions on the mean field level and furnishes the connected curve known from the Hubbard model with its edge-like van Hove singularities at $\omega = \pm 1$ and square-root band edges at $\omega = \pm 3$.

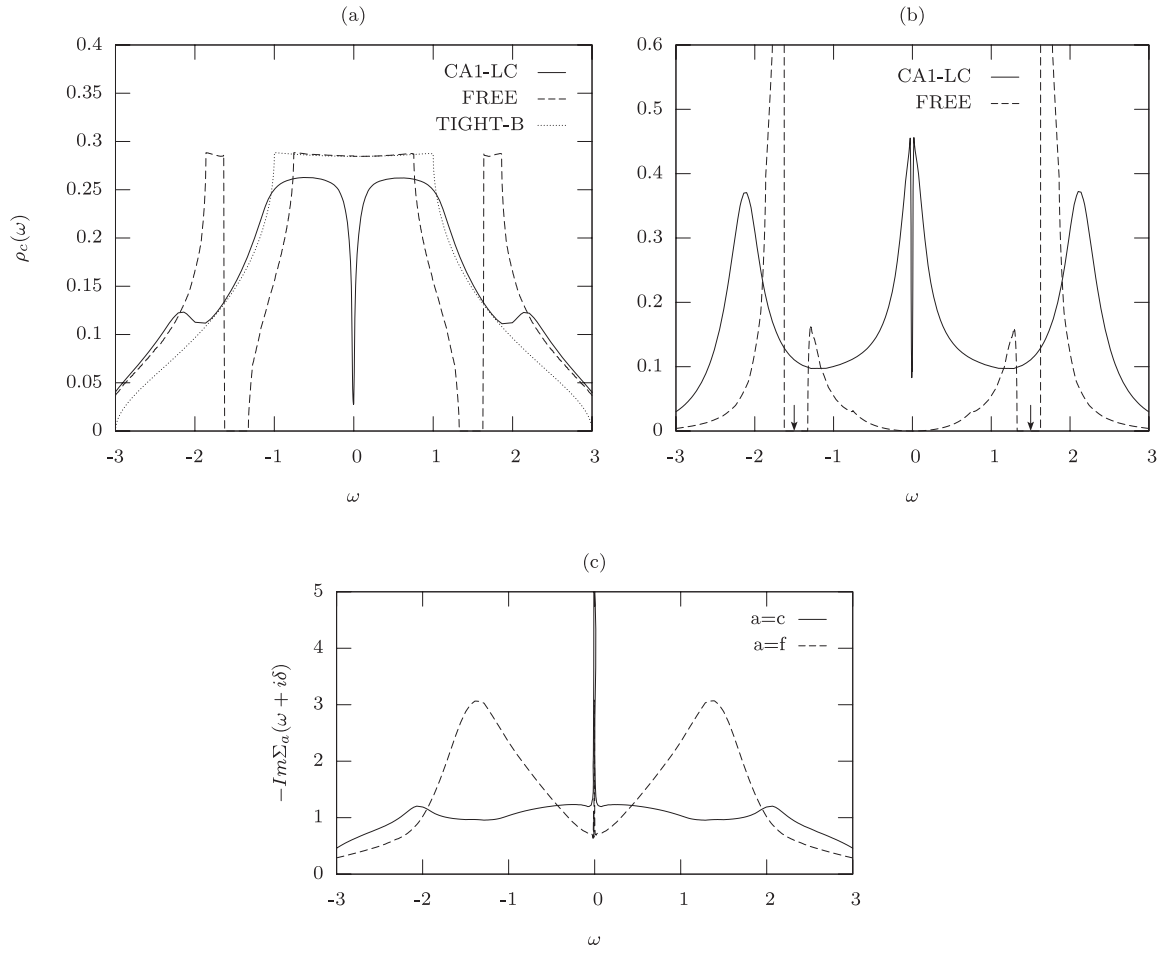

Figure 20. One-particle excitation spectra of band (c-) electrons and of local (f-) electrons for an Anderson-lattice model, calculated within XNCA, using CA1 as impurity solver, for the half-filled case (two electrons per site), with a tight-binding c-band of width 6 in a 3d-sc lattice and local hybridization. Parameters are $\epsilon_\ell = -1.5$, $U = 3$, $\beta = 15$ and $\Delta_A \equiv \pi V^2 \varrho_{c\sigma}^{(0)}(0) = 0.3$. Parts (a) and (b) show the local density of c and f electrons, respectively, both in comparison with a Hartree–Fock result and a calculation within 'free theory'. Part (c) contains the imaginary part of the 'local self-energies' (absolute value) for c and f electrons. The spikes seen in the middle of the gap region can cause numerical problems.

Hybridization with the correlated local f states at one-particle energies $\omega = \epsilon_\ell = -1$ and $\omega = \epsilon_\ell + U = +1$ produces the gaps visible in the result of the free theory, which furnishes three disconnected bands with variable spectral weights. The gaps are somewhat displaced by level repulsion, which is an effect of hybridization, too. Interactions, much better taken into account in the locally complete approximation using a CA1-impurity solver, wash out the two gaps and produce a repulsion of c weight away from the Fermi level $\omega = 0$ as a consequence of the formation of the many-body resonance with predominant f character. This resonance is clearly seen in figure 20(b), where also the two local one-particle levels of the isolated f shell as arrows and the spectrum of the free theory are shown. In the latter, two gaps are recognized as counterparts of those in figure 20(a); the f states acquire dispersion through the mixture with the band and share, in corresponding regions, the effect of level repulsion and gap formation. Interestingly, the CA1-impurity solver at the low temperature considered, i.e. $\beta = 15$, is able to describe the formation of the gap in the narrow region of quasiparticle states near $\omega = 0$, which was to be expected as a consequence of coherence in the Anderson lattice. This effect is connected with a strong increase of scattering at $\omega = 0$, see figure 20(c) where imaginary parts of the 'local' self-energies:

$$\widetilde{\Sigma}_{a\sigma}(z) = z - \epsilon_a - \Big(\frac{1}{N}\sum_{\underline{k}} G_{aa\underline{k}\sigma}(z)\Big)^{-1}, \qquad (z = \hbar\omega + \mathrm{i}\delta, \quad a = c, f) \tag{35}$$

are shown. For a proper incorporation of the narrow structure at $\omega = 0$ a fine numerical resolution is needed and even more so, since the XNCA/DMFT self-consistency cycle contains critical subtractions of local and lattice contributions, see, e.g., (27) and the remarks at the end of section 4.2. Judging the quality of an impurity solver for the description of the lattice coherence one has to assert numerical accuracy in this region of small energies.

Whereas the original gaps become smeared also in the f spectrum, the coherence gap should become perfect in the limit $T \to 0$, as is indeed observed with the NRG as an impurity solver [62]. This leads to the narrow gapped quasiparticle band structure, to be seen in the $\underline{k}$-resolved spectra of figures 21(a)

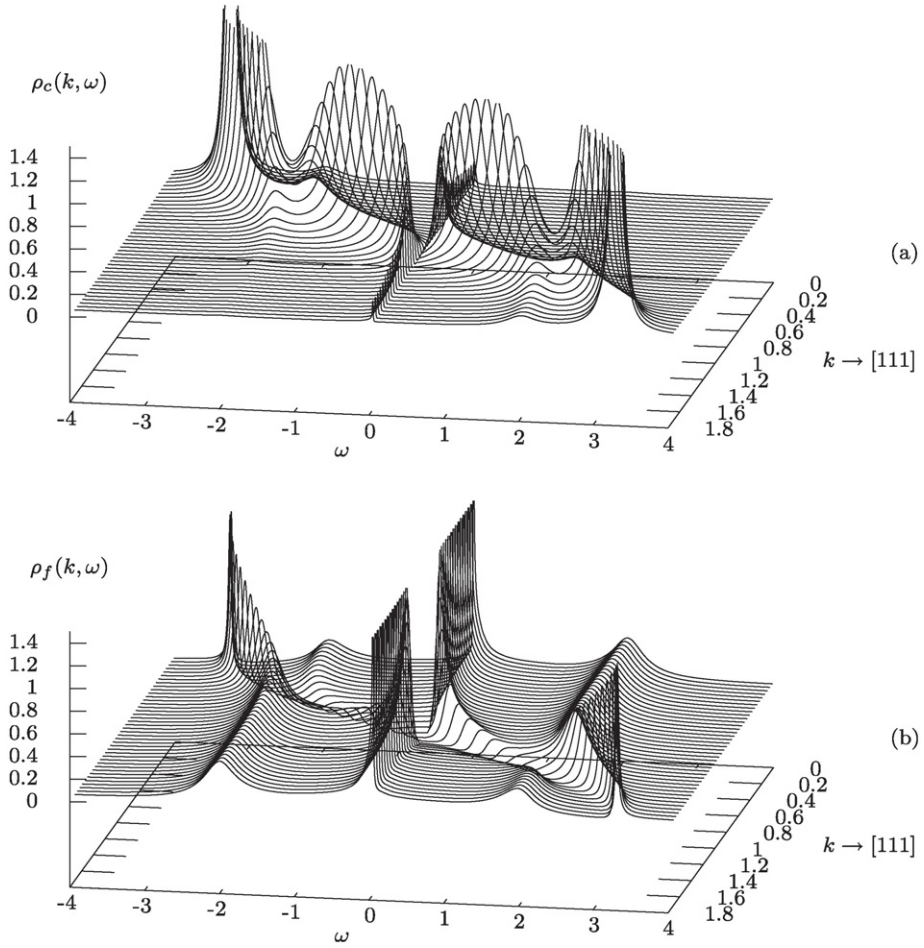

Figure 21. The $\underline{k}$-resolved excitation spectra of c and f electrons, respectively, for the same parameters as in figure 20 are drawn along the [111] direction. They demonstrate, like figures 20(a) and (b), the formation of hybridization pseudogaps, smeared by lifetime effects, in the high-energy region and of a narrow and complete coherence gap at the Fermi energy.

and (b), and to the broad structures, smeared out by the interactions, at higher excitation energies.

Realistic band structures, in addition to the consequences of local level hierarchies, attain their structure, such as gaps and crossings, from the $\underline{k}$ dependence of matrix elements like the one-particle hybridization. In the presence of strong local correlations, however, the quasiparticle bands forming inside the region of conduction states develop their own structure, possibly also with gaps, crossings, etc. The Luttinger theorem, as explained above, links some of the features seen in a description via independent band particles with the truly interacting case. Insofar, one can be assured, that Bragg scattering also opens gaps in quasiparticle bands in correspondence with, e.g., the effect of hybridization in conventional band structure theory. A particular nice example for this correspondence is furnished by the ionic Hubbard model. A calculation with the NRG as local impurity solver [61] demonstrates the development of a gap, when the Brillouin zone is halved by increasing the ionicity of sublattices. So far, only the NRG was able to calculate such effects properly. In extensive calculations for realistic systems the semianalytical impurity solvers constitute an additional useful tool, since they work reliably in both the high-and low-energy region.

The $\underline{k}$ dependence of the hybridization involves a second interesting aspect. In the versions of the Anderson model (30) studied here numerically, a local, $\underline{k}$-independent hybridization V was used. Although it was shown in section 4.4 that hybridizations, e.g. to nearest neighbours, can be formally treated in the same manner, important differences arise: excitation spectra calculated with SNCA or ENCA as an impurity solver [65] already show a substantial suppression of the many-body correlations derived from the Kondo effect in the symmetric case with hybridization to nearest-neighbour sites. The reason seems to be that $V_{\underline{k}}$ becomes zero at the band centre, e.g. near the Fermi level, so that band states at low energies cannot effectively screen the local moment. This conclusion is in accord with an investigation of an Anderson impurity model with a pseudogap density of band states [66], which vanishes like $(\omega - \mu)^r$ at the chemical potential with a characteristic exponent r. Suppression of screening was found for $r \geqslant 1/2$. Since a linearly vanishing $V_{\underline{k}}$ implies a quadratically vanishing effective hybridization strength $\pi V_{\underline{k}}^2 \rho_c(t_{\underline{k}})$ these findings may be relevant here, too, although the effective medium in the lattice case is determined from the self-consistency cycle. This again points to the highly nontrivial circumstances of gap formation in the quasiparticle

band structure of correlated lattice systems and underlines the importance of a thorough investigation of the reliability of impurity solvers in the low-energy regime.

Finally it has to be mentioned that the Fermi-liquid phase might not be thermodynamically stable for all parameter values used in our calculations for the Hubbard- and the Anderson-lattice model. Susceptibilities can point to instabilities towards other possible ground states [67], which is among a variety of methods being applied to the investigation of phase diagrams.

5. Conclusion and outlook

The foregoing sections have demonstrated the considerable progress, which has been made in the development of impurity solvers via direct perturbation theory and their application to impurity and lattice problems with strongly correlated electrons. Our presentation has emphasized a unified view on several approximations of this kind, which have been proposed in the past, and on a new one, the CA1, discussed here for the first time.

All of these approximations can be characterized in a systematic fashion as skeleton expansions in terms of time-ordered local perturbational processes along the lines laid out by [31]. As such, they furnish coupled implicit integral equations for propagators, which in general have to be solved numerically; this gives rise to the notion 'semianalytic'. ENCA, SUNCA, FNCA and CA1 include different classes of vertex corrections; the first three of these approximations and SCNA reduce for $U \to \infty$ to the old NCA. CA1, on the other hand, contains fully crossing vertex corrections of fourth order in the hybridization.

It turns out that for the quality of the approximation it is important to include ladders for repeated particle scattering and higher-order vertex corrections in a well-balanced way. This is apparently accomplished best by the CA1, which however does not iterate special subclasses to infinite order like SUNCA and, more generally, FNCA. Comparison with NRG calculations in the spirit of Wilson's approach reveals that even CA1 has deficiencies at low temperatures and excitation energies. At higher energies, however, the situation is reversed: in an overall view taking into account the complete spectral region, the semianalytical impurity solvers, with the possible exception of SNCA, perform quite well, and even the ENCA, as the least complicated of them, may be used for qualitative investigations.

The unified view developed here also concerns the construction of approximations for lattice problems and the use of the impurity solvers therein. It was shown that appropriate choices of local building blocks, each containing a set of internally correlated one-particle states, and a selection of paths for propagation through the lattice can be consistently combined in a matrix formulation for a calculation of partition function and Green functions.

It was explained how in certain simple situations, such as encountered in the Hubbard model or the Anderson-lattice model with noninteracting band states, the formalism reduces to a scalar one and how a locally complete selection of local processes then leads to the well-known XNCA and DMFT approximations.

Our starting point was a cumulant expansion for all local n-particle vertices, which renders the application of the linked cluster theorem and of unrestricted site summations possible, thus enabling a convenient $\underline{k}$-space representation of quantities. Generalized dynamical fields have been introduced and traced back to matrix propagators along closed loops. The neglect of all cumulant vertices of order $n \geqslant 2$ leaves as a natural 'free theory' the matrix generalization of the Hubbard-I approximation, which can readily be evaluated in explicit form. Restricted selection of cumulant vertices to finite order allows us to define, for example, Hartree-type approximations [54, 56], which are not yet well investigated but supposedly are of restricted usefulness for the regime of low temperatures and excitation energies.

The locally complete approximations, on the other hand, combine advantages of real space as well as of $\underline{k}$-space formulations with a better treatment of infrared divergences up to infinite order. Nowadays this can be implemented by a variety of local impurity solvers, among which we have concentrated here on the class based on direct perturbation theory. As applications of the formalism we have presented local and $\underline{k}$-resolved one-particle-excitation spectra for Hubbard- and Anderson-lattice model and have discussed characteristic similarities and differences. It has proven useful to connect this discussion with the foregoing treatment of the SIAM as the prototypical effective impurity. In particular, the formation of a Fermi liquid could be illuminated in this way, emphasizing a local point of view.

We will conclude with a short perspective on possible future developments on the basis of our local approach and with some critical remarks about its shortcomings. Improvements of semianalytical impurity solvers could be based on CA1 and proceed along directions laid down in the simpler case of the $U = \infty$ version of SIAM [41, 44].

Two different approaches could be combined to develop such a CA2 theory. The fully crossing fourth order vertex corrections could be reinforced by certain diagrams of sixth order like in figure 4(b), and possibly iterated further, which had proven beneficial for spectral properties at large U in the frame of the PNCA [41]. The CTMA [44, 46], on the other hand as the second of these theories for $U = \infty$, stresses the role of long ladders of crossing particle lines.

A generalization to finite U can be accomplished by solving the system of four coupled T-matrix equations shown in figure 22. The resulting T matrices then are to be inserted into the fully crossing vertex corrections contained in CA1; they replace parts shown as lowest-order contributions envisaged in figure 3. The first T matrix of figure 22, for example, replaces the diagram in figure 4(a), leading to the sequence of figure 5(b). Both measures together should again be well balanced in the sense discussed above. Although we expect further essential quantitative improvements by such additions to CA1, the resulting CA2-impurity solver still cannot be expected to be perfect.

It lies at the heart of the infrared problems in SIAM or its generalizations that no treatment based on a restricted selection

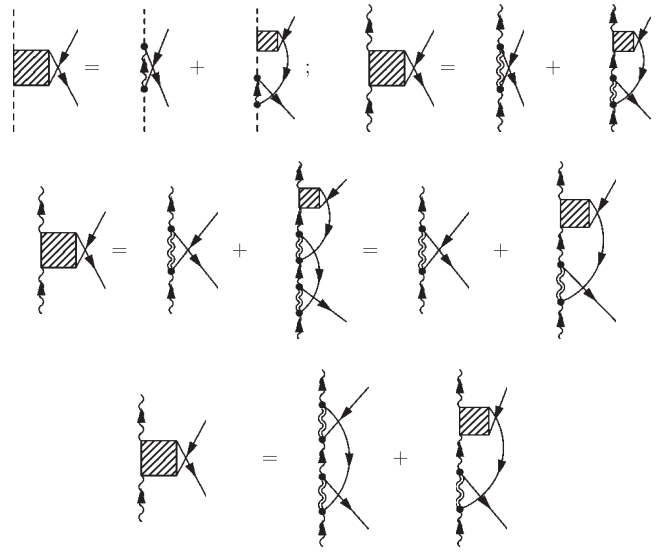

Figure 22. Higher-order corrections to CA1 can be implemented via the four T-matrix equations shown graphically. The first of these was considered in figure 5. The other three equations would generalize $U = \infty$-theories like PNCA and CTMA to finite values of U (CA2-project).

of perturbational processes to infinite order can adequately describe the complete energy range down to $\omega = 0$. From a practical point of view, however, for many purposes this will not really be necessary, for example when other types of correlations in concentrated systems intervene, e.g. producing magnetic states. One problem will be left anyway, even in this case: the numerical expense connected with approximations like CA1, and even more so with a hypothetical CA2, is considerable. Certainly there will be a need for improved algorithms or for the use of parallel computing. At the end, these higher semianalytical impurity solvers may turn out impractical compared with e.g. Quantum Monte Carlo or NRG-methods. Their ability, on the other hand, to describe the regime at higher temperatures and excitation energies very well is already shared by the ENCA, which does not need so much numerical effort.

Impurity solvers in general serve as a key ingredient in the local approach to lattice systems. Current research aims at the inclusion of more realistic local building blocks containing several orbitals or clusters of sites and at a better treatment of nonlocal correlations. With an optimal selection of a localized basis of Wannier states screened direct interactions are hopefully short-ranged and can either be treated within a small cluster as a local building block or via an extension of the perturbation in the Hamiltonian to double-transfer between neighbours, as indicated in section 2.

This, as well as a multi-orbital situation, can in principle be handled by the matrix-propagator formalism presented in this paper; it has been demonstrated in section 4 how matrix-loop propagators systematically serve to define appropriate dynamical fields acting on effective sites. Long ranged correlations, however, need an extension of the locally complete approximation for their proper incorporation.

The cumulant approach outlined above can serve as a basis for such generalizations. The nonlocal processes additionally to be taken into account in, for example, self-energies or susceptibilities connect local cumulants with four external legs or more on different sites via more than one loop. The resulting correlations between propagating particles or holes and between the states of local blocks on different sites may generate different types of long-range order, e.g. phases of itinerant or local-moment magnetism.

Technically, the difficulty in setting up such appropriate nonlocal theories seems to lie in a consistent choice of diagram classes as skeletons without overcounting, so that basic requirements are fulfilled, such as conservation laws and analytic properties like positivity of excitation spectra. This requires, apart from a strict use of cumulant subtractions, the consequential implementation of the concept of irreducibility [21]. Hopefully, work along these lines will soon lead to improved forms of, for example, band structure theories for correlated electrons and will shed more light on the mechanisms behind the formation of exotic ground states in transition metal compounds.

Acknowledgments

One of the authors (NG) expresses his gratitude to the Max Planck Institut für Physik Komplexer Systeme in Dresden and to its director Professor P Fulde for their hospitality and the opportunity for discussions and extensive numerical calculations contributing to this work.

This research was supported in parts (FBA) by the DFG project AN 275/5-1. FBA also acknowledges supercomputer support by the NIC, Forschungszentrum Jülich under project no. HHB000.

References

[1] Keiter H and Kimball J C 1971 *Int. J. Magn.* **1** 233
Keiter H and Kimball J C 1971 *J. Appl. Phys.* **42** 1460
[2] Bringer A and Lustfeld H 1977 *Z. Phys.* B **28** 213
Lustfeld H and Bringer A 1978 *Solid State Commun.* **28** 119
[3] Grewe N and Keiter H 1981 *Phys. Rev.* B **24** 4420
[4] Grewe N 1982 *Valence Instabilities* ed P Wachter and H Boppart (Amsterdam: North-Holland) p 21
[5] Early review articles are Güntherodt G 1976 *Configurations of 4f Electrons in Rare Earth Compounds 'Festkörperprobleme XVI/Advances in Solid State Physics'* (Braunschweig: Vieweg & Sohn) p 95
Wohlleben D 1976 *J. Phys. Coll.* **C4** 231
[6] Grewe N and Steglich F 1991 *Heavy Fermions, in Handbook on the Physics and Chemistry of Rare Earths* vol 14, ed K A Gschneidner and L Eyring (Amsterdam: Elsevier)
[7] Steglich F, Aarts J, Bredl C D, Lieke W, Meschede D, Franz W and Schäfer H 1979 *Phys. Rev.* **43** 1892
[8] Grewe N 1983 *Z. Phys.* B **52** 193
Grewe N 1983 *Z. Phys.* B **53** 271
[9] Kuramoto Y 1983 *Z. Phys.* B **53** 37
Kuramoto Y 1984 *Z. Phys.* B **54** 293
[10] Wilson K G 1975 *Rev. Mod. Phys.* **47** 773
Krishnamurthy H R, Wilkins J W and Wilson K G 1980 *Phys. Rev.* B **21** 1003
[11] For an early review see Grüner G and Zawadowski A 1974 *Rep. Prog. Phys.* **37** 1497

[12] Baym G and Kadanoff L P 1961 *Phys. Rev.* **124** 287
Baym G 1962 *Phys. Rev.* **127** 1391
[13] Coleman P 1983 *Phys. Rev.* B **28** 5255
Coleman P 1984 *Phys. Rev.* B **29** 3035
[14] Kuramoto Y and Kojima H 1984 *Z. Phys.* B **57** 95
[15] Kuramoto Y and Müller-Hartmann E 1985 *J. Magn. Magn. Mater.* **52** 122
[16] Nozières P and De Dominicis C T 1969 *Phys. Rev.* **178** 1097
Nozières P and De Dominicis C T 1969 *Phys. Rev.* **178** 1084
Nozières P and De Dominicis C T 1969 *Phys. Rev.* **178** 1097
[17] Menge B and Müller-Hartmann E 1988 *Z. Phys.* B **73** 225
[18] For a review see Cox D L and Zawadowski A 1998 *Adv. Phys.* **47** 599
[19] Grewe N 1984 *Solid State Commun.* **50** 19
[20] Kuramoto Y 1985 *Theory of Heavy Fermions and Valence Fluctuations* ed T Kasuya and T Saso (Berlin: Springer) p 152
Kim C I, Kuramoto Y and Kasuya T 1990 *J. Phys. Soc. Japan* **59** 2414
[21] Grewe N 1987 *Z. Phys.* B **67** 323
Grewe N, Pruschke T and Keiter H 1988 *Z. Phys.* B **71** 75
[22] Georges A and Kotliar G 1996 *Rev. Mod. Phys.* **68** 13
[23] Brito J J S and Frota H O 1990 *Phys. Rev.* B **42** 6378
Costi T A and Hewson A C 1990 *Physica* B **163** 179
Costi T A and Hewson A C 1992 *Phil. Mag.* B **65** 1165
[24] Bulla R, Costi T and Pruschke T 2008 *Rev. Mod. Phys.* **80** 395
[25] Pollmann F, Runge E and Fulde P 2006 *Phys. Rev.* B **73** 125121
[26] Schumann R 2002 *Ann. Phys., Lpz.* **1** 49
[27] Maier T, Jarrel M, Pruschke T and Hettler M H 2005 *Rev. Mod. Phys.* **77** 1027
[28] Held K, Nekrasov I A, Blümer N, Anisimiov V I and Vollhardt D 2001 *Int. J. Mod. Phys.* B **15** 2611
Kotliar G and Vollhardt D 2004 *Phys. Today* **57** 53
[29] Grewe N and Pruschke T 1985 *Z. Phys.* B **60** 311
[30] Metzner W and Vollhardt D 1989 *Phys. Rev. Lett.* **62** 324
[31] Keiter H and Morandi G 1984 *Phys. Rep.* **109** 227
[32] Bickers N E, Cox D L and Wilkins J W 1987 *Phys. Rev.* B **36** 2036
Bickers N E 1987 *Rev. Mod. Phys.* **59** 845
[33] Kondo J 1964 *Prog. Theor. Phys.* **32** 37
[34] Ramakrishnan T V 1981 *Valence Fluctuations in Solids* ed L M Falicov and W Hanke (Amsterdam: North-Holland) p 13
Ramakrishnan T V and Sur K 1982 *Phys. Rev.* B **26** 1798
[35] Keiter H 1982 *Z. Phys.* B **49** 209
[36] Nozières P 1974 *J. Low Temp. Phys.* **17** 31
[37] Pruschke T and Grewe N 1989 *Z. Phys.* B **74** 439
[38] Keiter H 1985 *Z. Phys.* B **60** 337
Keiter H and Qin Q 1990 *Z. Phys.* B **79** 397
[39] Schrieffer J R and Wolff P A 1966 *Phys. Rev.* **149** 491
[40] Mühlschlegel B 1968 *Z. Phys.* **208** 94
Coqblin B and Schrieffer J R 1969 *Phys. Rev.* **185** 847
[41] Anders F B and Grewe N 1994 *Europhys. Lett.* **26** 551
Anders F B 1995 *J. Phys.: Condens. Matter* **7** 2801
[42] Grunenberg J and Keiter H 1991 *Physica* B **171** 39
[43] Haule K, Kirchner S, Kroha J and Wölfle P 2001 *Phys. Rev.* B **64** 155111
[44] Kroha J, Wölfle P and Costi T A 1997 *Phys. Rev. Lett.* **79** 261
[45] Mahan G D 1967 *Phys. Rev.* **153** 882
Mahan G D 1967 *Phys. Rev.* **163** 612
[46] Kroha J and Wölfle P 2005 *J. Phys. Soc. Japan* **74** 16
[47] Sakai O, Shimizu1 Y and Kaneta Y 2005 *J. Phys. Soc. Japan* **74** 2517
Sakai O, Motizuki M and Kasuya T 1988 *Core-Level Spectroscopy in Condensed Systems Theory* ed J Kanamori (Berlin: Springer) p 45
Kang K and Min B I 1996 *Phys. Rev.* B **54** 1645
[48] Otsuki J and Kuramoto Y 2006 *J. Phys. Soc. Japan* **75** 064707
[49] Anderson P W 1967 *Phys. Rev. Lett.* **18** 1049
Anderson P W 1967 *Phys. Rev.* **164** 352
[50] Schotte K D and Schotte U 1969 *Phys. Rev.* **182** 479
[51] Peters R, Pruschke T and Anders F B 2006 *Phys. Rev.* B **74** 245114
[52] Hewson A C 1993 *The Kondo Problem to Heavy Fermions* (Cambridge: Cambridge University Press) p 63
[53] Hubbard J 1963 *Proc. R. Soc.* A **276** 238
[54] Grewe N 1998 Lokale Theorie *Lecture Notes* (TU Darmstadt); available on request
[55] Grewe N 2005 *Ann. Phys., Lpz.* **14** 611
Vladimir M I and Moskalenko V A 1990 *Theor. Math. Phys.* **82** 301
[56] Sherman A 2006 *Phys. Rev.* B **73** 155105
Sherman A 2006 *Phys. Rev.* B **74** 035104
Vakaru S I, Vladimir M I and Moskalenko V A 1990 *Theor. Math. Phys.* **85** 1185
[57] Craco L and Gusmão M A 1995 *Phys. Rev.* B **52** 17135
Craco L and Gusmão M A 1996 *Phys. Rev.* B **54** 1629
Consiglio R and Gusmão M A 1997 *Phys. Rev.* B **55** 6825
[58] Friedel J 1952 *Phil. Mag.* **43** 153
Yamada K 1974 *Prog. Theor. Phys.* **53** 970
Yamada K 1975 *Prog. Theor. Phys.* **54** 316
[59] Martin R M and Allen J W 1979 *J. Appl. Phys.* **50** 7561
Martin R M 1982 *Phys. Rev. Lett.* **48** 362
[60] Rubtsov A N, Katsnelson M I and Lichtenstein A I 2008 *Phys. Rev.* B **77** 033101
[61] Jabben T, Grewe N and Anders F B 2005 *Eur. Phys. J.* B **44** 47
[62] Jarrell M 1995 *Phys. Rev.* B **51** 7429
Pruschke T, Bulla R and Jarrell M 2000 *Phys. Rev.* B **61** 12799
Grenzebach C, Anders F B, Czycholl G and Pruschke T 2006 *Phys. Rev.* B **74** 195119
[63] Jarrell M 1992 *Phys. Rev. Lett.* **69** 168
Pruschke T, Jarrell M and Freericks J K 1995 *Adv. Phys.* **44** 187
[64] Schmitt S and Grewe N 2005 *Physica* B **359–361** 777
[65] Grewe N and Schmitt S 2006 unpublished
[66] Bulla R, Pruschke T and Hewson A C 1997 *J. Phys.: Condens. Matter* **9** 10463
Koller W, Hewson A C and Meyer D 2005 *Phys. Rev.* B **72** 045117
[67] Schmitt S and Grewe N 2008 in preparation