

$$G(\omega) = \int d\mathbf{k} \left(\omega + \mu - H(\mathbf{k}) - \Sigma_c(\omega) \right)^{-1}$$

$$G_b^{-1}(\omega) = \Sigma_c(\omega) + G^{-1}(\omega)$$

$$G_b^{-1}(\omega) \approx \omega + \mu - H_c - \Gamma [\omega - E]^{-1} \Gamma^+$$

$$H_{\text{full}} = H_{\text{loc}} + \sum_{\sigma \in \uparrow\downarrow} E_{\sigma \sigma \sigma} a_{\sigma \sigma}^+ a_{\sigma \sigma} + \sum_{\ell \in \uparrow\downarrow} \Gamma_{\ell \ell} a_{\ell \ell}^+ c_{\ell \ell} + \text{H.C.}$$

$$\Sigma_c(\omega) = G_b^{-1}(\omega) - G_c^{-1}(\omega)$$

DMFT at 25: Infinite Dimensions

Eva Pavarini, Erik Koch, Dieter Vollhardt and Alexander Lichtenstein (Eds.)

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation

German Research School for
Simulation Sciences GmbH

**Lecture Notes of the Autumn School on
Correlated Electrons 2014**

Eva Pavarini, Erik Koch, Dieter Vollhardt and Alexander Lichtenstein (Eds.)

DMFT at 25: Infinite Dimensions

Autumn school organized by the
DFG Research Unit 1346
Dynamical Mean-Field Approach
with Predictive Power for
Strongly Correlated Materials

at Forschungszentrum Jülich
15 – 19 September 2014

Schriften des Forschungszentrums Jülich
Reihe Modeling and Simulation

Band / Volume 4

ISSN 2192-8525

ISBN 978-3-89336-953-9

Bibliographic information published by the Deutsche Nationalbibliothek.
The Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the
Internet at <http://dnb.d-nb.de>.

Publisher: Forschungszentrum Jülich GmbH
 Institute for Advanced Simulation

Cover Design: Grafische Medien, Forschungszentrum Jülich GmbH

Printer: Druckerei Schloemer, Düren

Copyright: Forschungszentrum Jülich 2014

Distributor: Forschungszentrum Jülich
 Zentralbibliothek, Verlag
 D-52425 Jülich
 Phone +49 (0)2461 61-5368 · Fax +49 (0)2461 61-6103
 e-mail: zb-publikation@fz-juelich.de
 Internet: <http://www.fz-juelich.de>

Schriften des Forschungszentrums Jülich
Reihe Modeling and Simulation, Band / Volume 4

ISSN 2192-8525
ISBN 978-3-89336-953-9

The complete volume ist freely available on the Internet on the Jülicher Open Access Server (JUWEL) at
<http://www.fz-juelich.de/zb/juwel>

Neither this book nor any part of it may be reproduced or transmitted in any form or by any
means, electronic or mechanical, including photocopying, microfilming, and recording, or by any
information storage and retrieval system, without permission in writing from the publisher.

Contents

Preface

1. From Gutzwiller Wave Functions to Dynamical Mean-Field Theory
Dieter Vollhardt
2. Electronic Structure of Correlated Materials:
Slave-Boson Methods and Dynamical Mean-Field Theory
Gabriel Kotliar
3. Dynamical Mean-Field Theory:
Materials from an Atomic Viewpoint beyond the Landau Paradigm
Antoine Georges
4. Development of the LDA+DMFT Approach
Alexander Lichtenstein
5. Projectors, Hubbard U , Charge Self-Consistency, and Double-Counting
Tim Wehling
6. Linear Response Functions
Eva Pavarini
7. Continuous-time QMC Solvers for Electronic Systems in Fermionic and Bosonic Baths
Fakher Assaad
8. Quantum Cluster Methods
Erik Koch
9. Making Use of Self-Energy Functionals: The Variational Cluster Approximation
Michael Potthoff
10. Dynamical Vertex Approximation
Karsten Held
11. Functional Renormalization Group Approach to Interacting Fermi Systems:
DMFT as a Booster Rocket
Walter Metzner
12. Correlated Electron Dynamics and Nonequilibrium Dynamical Mean-Field Theory
Marcus Kollar
13. Theoretical Description of ARPES: The One-Step Model
Ján Minár
14. Introduction to Photoemission Spectroscopy
Michael Sing
15. Challenges from Experiment:
Correlation Effects and Electronic Dimer Formation in Ti_2O_3
Hao Tjeng

Index

Preface

Dynamical mean-field theory (DMFT) has opened new perspectives for dealing with strong electronic correlations and the associated emergent phenomena. This successful method has exploited the experience previously gained with single-impurity models, e.g., the Anderson model, transferring it to many-body lattice problems. The basis for this breakthrough was the realization, 25 years ago, that diagrammatic perturbation theory greatly simplifies in the limit of infinite dimensions, so that the self-energy becomes local. Nowadays DMFT, combined with *ab-initio* density-functional techniques, is the state-of-the art approach for strongly correlated materials. The lectures collected in this volume range from recounting the development of the dynamical mean-field theory to applications of the LDA+DMFT approach to real materials and modern developments. Among the latter, topics covered are modern impurity solvers, the calculation of two-particle Green functions, and method extensions beyond the single-site approximation. Lectures on photoemission spectroscopy provide the necessary contact to experiments. The goal of the school is to introduce advanced graduate students and up to the modern approaches to the realistic modeling of strongly-correlated systems.

A school of this size and scope requires support and help from many sources. The DFG Research Unit FOR 1346 provided the framework for the school and a large part of the financial support. The Institute for Advanced Simulation and the German Research School for Simulation Sciences at the Forschungszentrum Jülich provided additional funding and were vital for the organization of the school and the production of this book. The Institute for Complex Adaptive Matter (ICAM) offered travel grants for selected international participants.

The nature of a school makes it desirable to have the lecture-notes available when the lectures are given. This way students get the chance to work through the lectures thoroughly while their memory is still fresh. We are therefore extremely grateful to the lecturers that, despite tight deadlines, provided their manuscripts in time for the production of this book. We are confident that the lecture notes collected here will not only serve the participants of the school but will also be useful for other students entering the exciting field of strongly correlated materials.

We are grateful to Mrs. H. Lexis of the Verlag des Forschungszentrum Jülich and to Mrs. D. Mans of the Graphische Betriebe for providing their expert support in producing the present volume on a tight schedule. We heartily thank our students and postdocs who helped in proofreading the manuscripts, often on quite short notice: Michael Baumgärtel, Khaldoon Ghanem, Esmaeel Sarvestani, Amin Kiani Sheikhabadi, Hermann Ulm, Guoren Zhang, and, in particular, our native speaker Hunter Sims.

Finally, our special thanks go to Dipl.-Ing. R. Hözlle for his invaluable advice on the innumerable questions concerning the organization of such an endeavour, and to Mrs. L. Snyders and Mrs. E. George for expertly handling all practical issues.

Eva Pavarini, Erik Koch, Dieter Vollhardt, and Alexander Lichtenstein

August 2014

1 From Gutzwiller Wave Functions to Dynamical Mean-Field Theory

Dieter Vollhardt

Center for Electronic Correlations and Magnetism
University of Augsburg

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 1.1 | Modeling of correlated lattice electrons | 2 |
| 2 | Approximation schemes for correlated electrons | 4 |
| 2.1 | Mean-field theories | 4 |
| 2.2 | Variational wave functions | 5 |
| 3 | Gutzwiller wave functions | 5 |
| 3.1 | Gutzwiller approximation | 6 |
| 3.2 | Connection to Fermi liquid theory | 8 |
| 4 | From the Gutzwiller approximation to infinite dimensions | 9 |
| 4.1 | Calculation of expectation values with the Gutzwiller wave function | 9 |
| 4.2 | Lattice fermions in infinite spatial dimensions | 13 |
| 5 | Dynamical mean-field theory for correlated lattice fermions | 18 |
| 5.1 | Derivation of the self-consistent DMFT equations | 18 |
| 5.2 | The LDA+DMFT approach to correlated materials | 21 |
| 6 | Summary and outlook | 22 |

1 Introduction

Correlations between the degrees of freedom of d and f electrons lead to a wealth of fascinating phenomena, which include Mott metal-insulator transitions [1–3], the Kondo effect [4], heavy fermion behavior [5], band-ferromagnetism [6], high-temperature superconductivity [7], colossal magnetoresistance [8], and other Fermi liquid instabilities [9]. In particular, the very sensitive dependence of the properties of correlated materials on external parameters such as temperature, pressure, magnetic field, or doping make them interesting not only for fundamental research but also for future technological applications, e.g., the construction of sensors and switches, and the development of electronic devices with novel functionalities [10].

The importance of interactions between electrons in a solid was realized already at the outset of modern solid state physics. Namely, the report by de Boer and Verwey [11] on the surprising properties of materials with partially filled $3d$ -bands such as NiO prompted Mott and Peierls [12] to postulate that theoretical explanations of these properties must include the electrostatic interaction between the electrons. Explicit calculations soon confirmed this conjecture. At the same time it turned out that theoretical studies of interacting many-fermion systems are highly demanding. Here the development of the dynamical mean-field theory (DMFT) marks a methodological breakthrough. Indeed, by replacing the d -dimensional lattice of a correlated-electron solid by a single quantum impurity, which is self-consistently embedded in a bath provided by the other electrons, the DMFT opened the way for comprehensive theoretical investigations of correlation phenomena in electronic lattice models and materials. The starting point for the development of this powerful new many-body approach was the discovery, 25 years ago, that diagrammatic perturbation theory for interacting lattice fermions is much simpler in infinite spatial dimensions than in finite dimensions and, in particular, that the self-energy is then purely local. The current Autumn School on Correlated Electrons *DMFT at 25: Infinite Dimensions* commemorates this anniversary. Starting with a discussion of the properties of the Gutzwiller variational wave function and the mean-field-type Gutzwiller approximation, which provides a simplistic but robust, non-perturbative theoretical approach to correlated lattice fermions, I will describe the steps which eventually led to the formulation of the DMFT.

1.1 Modeling of correlated lattice electrons

The simplest model for interacting electrons in a solid is the one-band Hubbard model, which was introduced independently by Gutzwiller, Hubbard and Kanamori [13–15]. In this model the interaction between the electrons is assumed to be strongly screened, i.e., purely local. The Hamiltonian \hat{H} is the sum of two terms, the kinetic energy \hat{H}_{kin} and the interaction energy \hat{H}_I (here and in the following operators are denoted by a hat):

$$\hat{H} = \sum_{\mathbf{R}_i, \mathbf{R}_j} \sum_{\sigma} t_{ij} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma} + U \sum_{\mathbf{R}_i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (1)$$

where t_{ij} is the hopping amplitude, U is the local Hubbard interaction, $\hat{c}_{i\sigma}^{\dagger} (\hat{c}_{i\sigma})$ is the creation (annihilation) operator of an electron with spin σ in a Wannier orbital localized at lattice site \mathbf{R}_i ,

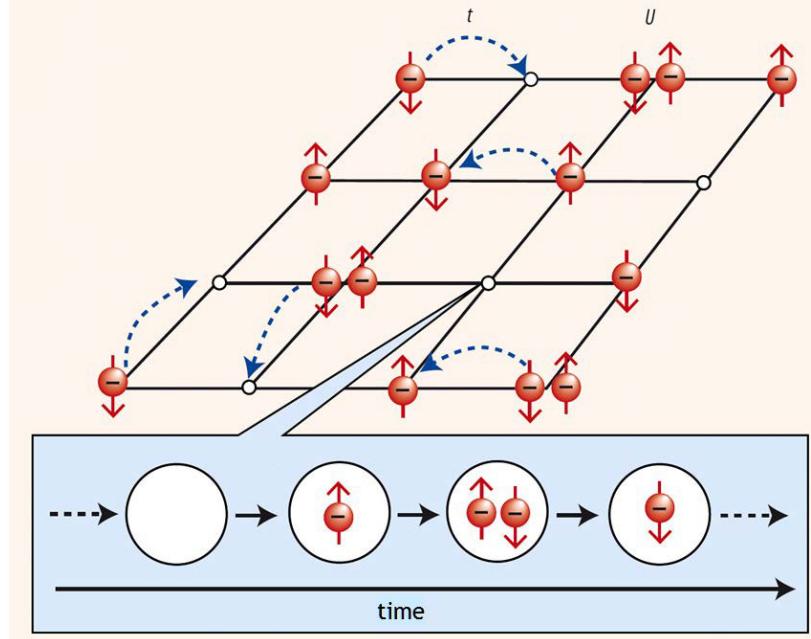


Fig. 1: Schematic illustration of interacting electrons in a solid described by the Hubbard model. The ions enter only as a rigid lattice, here represented by a square lattice. The electrons, which have a mass, a negative charge, and a spin (\uparrow or \downarrow), are quantum particles that move from one lattice site to the next with a hopping amplitude t . The quantum dynamics thus leads to fluctuations in the occupation of lattice sites as indicated by the time sequence. A lattice site can either be unoccupied, singly occupied (\uparrow or \downarrow), or doubly occupied. When two electrons meet on a lattice site, which is only possible if they have opposite spin because of the Pauli exclusion principle, they encounter an interaction U .

and $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$. The Hubbard interaction can also be written as $\hat{H}_I = U \hat{D}$ where $\hat{D} = \sum_{\mathbf{R}_i} \hat{D}_i$ is the number operator of doubly occupied sites of the system, with $\hat{D}_i = \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$ as the local operator for double occupation. The Fourier transform of the kinetic energy

$$\hat{H}_{\text{kin}} = \sum_{\mathbf{k},\sigma} \varepsilon_{\mathbf{k}} \hat{n}_{\mathbf{k}\sigma} \quad (2)$$

is defined by the dispersion $\varepsilon_{\mathbf{k}}$ and the momentum distribution operator $\hat{n}_{\mathbf{k}\sigma}$. A schematic picture of the Hubbard model is shown in Fig. 1. For strong repulsion U double occupations are energetically unfavorable and are therefore suppressed. In this situation the local correlation function $\langle \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \rangle$ must not be factorized, since otherwise correlation phenomena are eliminated from the beginning. Therefore Hartree-Fock-type mean-field theories, which do factorize the interaction, cannot explain the physics of strongly correlated electrons.

The Hubbard model looks deceptively simple. However, the competition between the kinetic energy and the interaction leads to a complicated many-body problem, which is impossible to solve analytically except in dimension $d = 1$ [16]. This model provides the basis for most of the theoretical research on correlated electrons during the last few decades.

2 Approximation schemes for correlated electrons

Theoretical investigations of quantum-mechanical many-body systems are faced with severe technical problems, particularly in those dimensions which are most interesting to us, namely $d = 2, 3$. This is due to the complicated quantum dynamics and, in the case of fermions, the non-trivial algebra introduced by the Pauli exclusion principle.

In view of the fundamental limitations of exact analytical approaches one might hope that, at least, modern supercomputers can provide detailed numerical insights into the thermodynamic and spectral properties of correlated fermionic systems. However, since the number of quantum mechanical states increases exponentially with the number of lattice sites L , numerical solutions of the Hubbard model and related models are limited to relatively small systems. This shows very clearly that there is still a great need for analytically tractable approximation methods [17], in particular for non-perturbative approximation schemes which are applicable for all input parameters.

2.1 Mean-field theories

In the theory of classical and quantum many-body systems an overall description of the properties of a model is often obtained within a *mean-field theory*. Although the term is frequently used it does not have a precise meaning, since there exist numerous ways to derive mean-field theories. One construction scheme is based on a factorization of the interaction, as in the case of the Weiss mean-field theory for the Ising model, or the Hartree-Fock theory for electronic models. The decoupling implies a neglect of fluctuations (or rather of the correlation of fluctuations; for details see Ref. [18]) and thereby reduces the original many-body problem to a solvable problem where a single spin or particle interacts with a mean field provided by the other particles. Another, in general unrelated, construction scheme makes use of the simplifications that occur when some parameter is assumed to be large (in fact, infinite), e.g., the magnitude of the spin S , the spin degeneracy N , the number Z of nearest neighbors of a lattice site (the coordination number), or the spatial dimension d .¹ Investigations in this limit, supplemented, if possible, by an expansion in the inverse of the large parameter,² often provide valuable insight into the fundamental properties of a system even when this parameter is not large. One of the best-known approximations obtained in this way is the Weiss mean-field theory for the Ising model [19]. This is a prototypical “single-site mean-field theory,” which becomes exact not only in the limit $Z \rightarrow \infty$ or $d \rightarrow \infty$, but also for an infinite-range interaction. It contains no unphysical singularities and is applicable for all values of the input parameters, i.e., coupling parameters, magnetic field, and temperature.

¹For regular lattices both a dimension d and a coordination number Z can be defined. However, there exist other lattices, such as the Bethe lattice, which cannot be associated with a physical dimension d although a coordination number Z is well-defined.

²In three dimensions one has $Z = 6$ for a simple cubic lattice, $Z = 8$ for a bcc lattice, and $Z = 12$ for an fcc-lattice. The parameter $1/Z$ is therefore quite small already in $d = 3$.

2.2 Variational wave functions

Another useful approximation scheme for interacting quantum many-body systems makes use of variational wave functions. They allow for approximate but explicit and physically intuitive investigations of correlations among quantum particles and are particularly valuable in situations where standard perturbation theory fails. Correlation problems where variational wave functions have been employed include such diverse examples as the quantum liquids Helium-3 and Helium-4 [20], rotons in superfluid ^4He [21], nuclear physics [22], and the fractional quantum Hall effect [23]. Variational wave functions received renewed attention in the study of heavy fermions [24, 25] and high- T_c superconductivity [26].

The general strategy is to construct an explicit wave function of the form

$$|\Psi_{\text{var}}\rangle = \hat{C}|\Psi_0\rangle \quad (3)$$

where $|\Psi_0\rangle$ is a tractable one-particle starting wave function on which a correlation operator $\hat{C}(\lambda_1, \dots, \lambda_n)$ acts. The latter depends on variational parameters λ_i and describes the microscopic interaction between the particles in an approximate way. This wave function is then used to calculate the expectation value of an operator $\hat{\mathcal{O}}$ as

$$\langle \hat{\mathcal{O}} \rangle_{\text{var}} = \frac{\langle \Psi_{\text{var}} | \hat{\mathcal{O}} | \Psi_{\text{var}} \rangle}{\langle \Psi_{\text{var}} | \Psi_{\text{var}} \rangle}. \quad (4)$$

In particular, by calculating and minimizing the ground state energy $E_{\text{var}} = \langle \hat{H} \rangle_{\text{var}}$, where \hat{H} is the Hamiltonian, the variational parameters contained in \hat{C} (and perhaps also in $|\Psi_0\rangle$) can be determined. These parameters are used to suppress those configurations in $|\Psi_0\rangle$ which for given interaction strength are energetically unfavorable. The variational principle guarantees that E_{var} provides a rigorous upper bound for the exact ground state energy.

3 Gutzwiller wave functions

For the Hubbard model, (1), the simplest variational wave function of the form (3) is the so-called Gutzwiller wave function

$$|\Psi_G\rangle = g^{\hat{D}} |\text{FG}\rangle \quad (5a)$$

$$= \prod_{R_i} [1 - (1 - g)\hat{D}_i] |\text{FG}\rangle, \quad (5b)$$

where $g^{\hat{D}}$, with $0 \leq g \leq 1$, is the correlation operator and $|\text{FG}\rangle$ is the ground state of the non-interacting Fermi gas. Hence the correlation operator globally reduces the amplitude of those spin configurations in $|\text{FG}\rangle$ with too many doubly occupied sites. The limit $g = 1$ corresponds to the non-interacting case, while $g \rightarrow 0$ describes the limit $U \rightarrow \infty$. Indeed, for $g \rightarrow 0$ one finds

$$g^{\hat{D}} \Big|_{g=0} = \prod_{R_i} [1 - \hat{D}_i] \equiv \hat{P}_G. \quad (6)$$

The projection operator \hat{P}_G eliminates all configurations with doubly occupied sites (*Gutzwiller projection*). The ground state energy in terms of the Gutzwiller wave function is then given by

$$E_G = \langle \hat{H} \rangle_G \equiv \frac{\langle \Psi_G | \hat{H} | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle}. \quad (7)$$

By replacing $|FG\rangle$ with a more general starting wave function one can also describe states with broken symmetry; examples are the antiferromagnetic Hartree-Fock wave function (spin density wave)

$$|SDW\rangle = \prod_{\mathbf{k},\sigma} [u_{\mathbf{k}} \hat{a}_{\mathbf{k}\sigma}^\dagger + \sigma v_{\mathbf{k}} \hat{a}_{\mathbf{k}+\mathbf{Q},\sigma}^\dagger] |0\rangle, \quad (8a)$$

where \mathbf{Q} is half a reciprocal lattice vector and $|0\rangle$ is the vacuum, and the BCS wave function [27]

$$|BCS\rangle = \prod_{\mathbf{k}} [u_{\mathbf{k}} + v_{\mathbf{k}} \hat{a}_{\mathbf{k}\uparrow}^\dagger \hat{a}_{-\mathbf{k}\downarrow}^\dagger] |0\rangle, \quad (8b)$$

which after projection leads to a resonating valence bond state (RVB) [26].

3.1 Gutzwiller approximation

In addition to introducing the wave function (5a) Gutzwiller constructed a non-perturbative approximation scheme that allowed him to obtain an explicit expression for the ground state energy of the Hubbard model [13, 28].³ We will see in Sec. 4.1.1 that this Gutzwiller approximation yields the exact result for expectation values calculated with Gutzwiller wave functions in the limit of infinite spatial dimensions ($d = \infty$). The idea behind the Gutzwiller approximation is easily understood [29, 30] and will be illustrated below by calculating the norm $\langle \Psi_G | \Psi_G \rangle$. Working in configuration space the ground state of the Fermi gas can be written as

$$|FG\rangle = \sum_D \sum_{\{i_D\}} A_{i_D} |\Psi_{i_D}\rangle, \quad (9)$$

where $|\Psi_{i_D}\rangle$ is a spin configuration with D doubly occupied sites and A_{i_D} the corresponding probability amplitude. The sum extends over the whole set $\{i_D\}$ of different configurations with the same D , and over all D . For a system with L lattice sites and N_σ electrons of spin σ (σ -electrons) the number N_D of different configurations in $\{i_D\}$ is given by the combinatorial expression

$$N_D = \frac{L!}{L_\uparrow! L_\downarrow! D! E!}, \quad (10)$$

where $L_\sigma = N_\sigma - D$ and $E = L - N_\uparrow - N_\downarrow + D$ are the numbers of singly occupied and empty sites, respectively. Since $|\Psi_{i_D}\rangle$ is an eigenstate of \hat{D} , the norm of $|\Psi_G\rangle$ reads

$$\langle \Psi_G | \Psi_G \rangle = \sum_D g^{2D} \sum_{\{i_D\}} |A_{i_D}|^2. \quad (11)$$

³By studying lattice electrons with a local Coulomb repulsion Gutzwiller wanted to understand the origin of ferromagnetism in metals.

The Gutzwiller approximation effectively amounts to neglecting spatial correlations between the spins of the electrons. The probability $|A_{i_D}|^2$ is then the same for all configurations of electrons on the lattice, i.e., is given by the classical combinatorial result for uncorrelated particles

$$|A_{i_D}|^2 = P_\uparrow P_\downarrow. \quad (12)$$

Here $P_\sigma = 1/\binom{L}{N_\sigma} \simeq n_\sigma^{N_\sigma} (1 - n_\sigma)^{L - N_\sigma}$, with $n_\sigma = N_\sigma/L$, is the probability for an arbitrary configuration of σ -electrons. In this case (11) reduces to

$$\langle \Psi_G | \Psi_G \rangle = P_\uparrow P_\downarrow \sum_D g^{2D} N_D. \quad (13)$$

In the thermodynamic limit the sum in (13) is dominated by its largest term corresponding to a value $D = \bar{D}$, where $\bar{D} = L\bar{d}$ is determined by

$$g^2 = \frac{\bar{d}(1 - n_\uparrow - n_\downarrow + \bar{d})}{(n_\downarrow - \bar{d})(n_\uparrow - \bar{d})}. \quad (14)$$

Equation (14) has the form of the law of mass action where, however, the correlation parameter g^2 rather than the Boltzmann factor regulates the dynamical equilibrium between the concentrations of singly occupied sites on one side of this “chemical reaction” and that of doubly occupied sites and holes on the other.⁴ Eq. (14) uniquely relates \bar{d} and g , such that g may be replaced by the quantity \bar{d} . The calculation of the expectation values of the kinetic and the interaction energy of the Hubbard model proceeds similarly [30]. The Gutzwiller approximation for quantum mechanical expectation values, which is based on the counting of classical spin configurations, belongs to the class of *quasiclassical approximations*.

3.1.1 Brinkman-Rice transition

The ground state energy per lattice site of the Hubbard model as a function of the variational parameter $\bar{d}(g)$ is then found as

$$E_G[\bar{d}(g)]/L = \sum_\sigma q_\sigma(\bar{d}, n_\uparrow, n_\downarrow) \varepsilon_{0,\sigma} + U\bar{d}, \quad (15)$$

which is to be minimized with respect to \bar{d} . Here $\varepsilon_{0,\sigma}$ is the energy of non-interacting σ -electrons and $q_\sigma \leq 1$ may be viewed as a reduction factor of the kinetic energy (or the band width) due to correlations. In particular, for $n_\uparrow = n_\downarrow$ one has $q_\sigma \equiv q = 2(1 - \delta - 2\bar{d})(\sqrt{\bar{d} + \delta} + \sqrt{\bar{d}})^2/(1 - \delta^2)$, where $\delta = 1 - n$, with $n = n_\uparrow + n_\downarrow$ as the particle density, and $\varepsilon_\uparrow = \varepsilon_\downarrow$. So one finds that within the Gutzwiller approximation the correlations only lead to a multiplicative renormalization of the non-interacting kinetic energy. Brinkman and Rice [32, 33] showed that in the special case $n_\sigma = 1/2$ (half-filled band) the minimization of (15) yields

$$q = 1 - \bar{U}^2, \quad (16a)$$

$$\bar{d} = (1 - \bar{U})/4, \quad (16b)$$

$$E/L = -|\varepsilon_0|(1 - \bar{U})^2, \quad (16c)$$

⁴It is interesting to note that (14), with g^2 replaced by the Boltzmann factor $e^{-\beta U}$, is the exact result for the Hubbard model with infinite-range hopping [31].

where $\bar{U} = U/(8|\varepsilon_0|)$ and $\varepsilon_0 = \varepsilon_{0\uparrow} + \varepsilon_{0\downarrow}$. Eq. (16c) implies that the ground state energy E , which equals $-L|\varepsilon_0|$ at $U = 0$, increases with U and vanishes at a finite critical value $U_c = 8|\varepsilon_0|$, since the density of doubly occupied sites \bar{d} (and hence the reduction factor q) vanishes at this point. The fact that $E_{\text{kin}} \rightarrow 0$ and $E_{\text{I}} \rightarrow 0$ for $U \rightarrow U_c$ means that the particles become localized, which implies that a charge current can no longer flow. So the Gutzwiller approximation actually describes a Mott-Hubbard metal-insulator transition at a finite interaction strength (*Brinkman-Rice-transition*). It occurs only for $n_\sigma = 1/2$. A transition to a localized, paramagnetic state with $E = 0$ at a finite value of U clearly does not describe the behavior of the electrons completely. It is well known that for $U \gg t$ localized spins couple antiferromagnetically, which leads to a lowering of the energy $E = 0$ by an amount $E_{AF} \propto -t^2/U$. This effect is not included in the Gutzwiller approximation, since spatial correlations were explicitly neglected. On the other hand, the magnetic coupling is an additional effect, which can be derived within second-order perturbation theory from the localized state. Therefore, as long as one is not too close to $U = U_c$ the overall results of the Gutzwiller approximation are not invalidated by the magnetic coupling and give important insight into the correlation-induced approach to the localized state.

3.2 Connection to Fermi liquid theory

Since the results of the Gutzwiller approximation describe correlated, paramagnetic fermions with a renormalized kinetic energy one can make contact with Landau's Fermi liquid theory [32, 30]. In particular, it turns out that the reduction factor q in (15) describes the discontinuity of the momentum distribution n_k at the Fermi level and may thus be identified with the inverse effective mass ratio $(m^*/m)^{-1}$ of the quasiparticles. Since $m^*/m = q^{-1} < \infty$ for $U < U_c$, the system is a Fermi liquid, i.e., a metal. At $U = U_c$ the effective mass diverges and the system becomes an insulator.

One can use (15) to calculate the spin susceptibility $\chi_s = \chi_s^0(m^*/m)/(1 + F_0^a)$ and compressibility $\kappa = \kappa^0(m^*/m)/(1 + F_0^s)$ within the Gutzwiller approximation, where χ_s^0 and κ^0 are the results for the non-interacting Fermi gas [32, 30]. For $n_\sigma = 1/2$, and assuming Galilei invariance, one finds [30] $m^*/m \equiv 1 + \frac{1}{3}F_1^s = 1/(1 - \bar{U}^2)$. The corresponding Fermi liquid parameters are given by

$$F_0^a = p \left(\frac{1}{(1 + \bar{U})^2} - 1 \right), \quad (17a)$$

$$F_0^s = p \left(\frac{1}{(1 - \bar{U})^2} - 1 \right), \quad (17b)$$

$$F_1^s = \frac{3\bar{U}^2}{1 - \bar{U}^2}, \quad (17c)$$

where $p = 2|\varepsilon_0|N(0)$, with $N(0)$ as the density of states at the Fermi energy. For typical symmetric densities of states one finds $p \simeq 1$. Hence, for $U \rightarrow U_c$ the Landau parameter F_0^a levels off and saturates at $\simeq -3/4$, while F_0^s increases much faster than linearly and eventually

diverges. In particular, for $U \rightarrow U_c$ the Wilson ratio remains constant:

$$\frac{\chi_s/\chi_s^0}{m^*/m} = \frac{1}{1 + F_0^a} \rightarrow \text{const.} \quad (18)$$

So the strong increase of χ_s as a function of U for $U \rightarrow U_c$ is mainly due to the rapid increase of the effective mass ratio m^*/m and not due to an incipient ferromagnetic instability [34], which would demand $F_0^a \rightarrow -1$.

As first pointed out by Anderson and Brinkman [35] and discussed in detail in Ref. [30], the behavior expressed by (17c) and (18) is indeed observed in the prototypical Fermi liquid Helium-3 (${}^3\text{He}$). Normal-liquid ${}^3\text{He}$ is an isotropic, strongly correlated fermionic system. The effective mass m^* and the spin susceptibility χ_s of the quasiparticles are strongly enhanced, while the compressibility κ is strongly reduced. Normal-liquid ${}^3\text{He}$ has therefore been called an *almost-localized* Fermi liquid.

4 From the Gutzwiller approximation to infinite dimensions

My 1984 Review of Modern Physics article [30] explained Gutzwiller's variational approach to the Hubbard model, the Gutzwiller approximation, and the Brinkman-Rice transition and thereby drew attention to the usefulness of this non-perturbative investigation scheme for correlated fermions. Nevertheless there remained questions about the nature of the Gutzwiller approximation, whose results are simple and mean-field-like. In fact, the latter feature is one of the reasons why the results of the Gutzwiller approximation, which originally had been derived for lattice fermions, are applicable even to liquid ${}^3\text{He}$ [30, 36]. The question was, therefore, whether the Gutzwiller approximation could also be derived by other, more conventional methods of quantum many-body theory in some limit. During 1983-84 I discussed this question with several colleagues, in particular with Andrei Ruckenstein at Bell Laboratories, Murray Hill, in 1983. At that time, Andrei tried to understand whether it was possible to generalize the Brinkman-Rice transition to correlated electronic systems in the presence of disorder [37]. This eventually led him and Gabi Kotliar to formulate a functional integral representation of the Hubbard and Anderson models in terms of auxiliary bosons, whose simplest saddle-point approximation (*slave-boson mean-field theory*) reproduces exactly the results of the Gutzwiller approximation [38]. Thus they had shown that the results of the Gutzwiller approximation could also be obtained without the use of the Gutzwiller variational wave function. We will return to this mean-field theory in Section 4.1.1.

4.1 Calculation of expectation values with the Gutzwiller wave function

As mentioned earlier, mean-field theories can be constructed in different ways. In particular, the Gutzwiller approximation, which was originally based on the quasi-classical counting of electronic configurations on a real-space lattice [13, 28], had been re-derived as a saddle-point approximation for electrons expressed in terms of auxiliary bosons [38]. At the same time, the

question of whether the Gutzwiller approximation could also be derived in a controlled way by calculating expectation values with the Gutzwiller wave function using conventional many-body perturbation theory was still open. In 1986, I suggested to Walter Metzner, then a diploma student of physics at the Technical University of Munich, to calculate the ground-state energy of the one-dimensional Hubbard model with the Gutzwiller wave function by means of many-body perturbation theory. It turned out that expectation values of the momentum distribution and the double occupation can be expressed as power series in the small parameter $g^2 - 1$, where g is the correlation parameter in the Gutzwiller wave function (5a).⁵ The coefficients of the expansions are determined by diagrams which are identical in form to those of a conventional Φ^4 theory. However, lines in a diagram do not correspond to one-particle Green functions of the non-interacting system, $G_{ij,\sigma}^0(t)$, but to one-particle density matrices, $g_{ij,\sigma}^0 = \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} \rangle_0$. Walter showed that it was possible to determine these coefficients to all orders in $d = 1$. This facilitated the exact analytic calculation of the momentum distribution and the double occupation, and thereby of the ground state energy of the Hubbard chain, in terms of the Gutzwiller wave function [40, 41].

In particular, for $n = 1$ and $U \gg t$ the ground state energy obtained with the Gutzwiller wave function in $d = 1$ has the form [40, 41]

$$E_G = - \left(\frac{4}{\pi} \right)^2 \frac{t^2}{U} \frac{1}{\ln \overline{U}}, \quad (19)$$

where $\overline{U} = U/(8|\varepsilon_0|)$, with $\varepsilon_0 < 0$ as the energy of the non-interacting particles. Hence the exact result, $E \sim -t^2/U$, obtained from second-order perturbation theory is found to be multiplied by a factor which is non-analytic in U . This explained why the ground state energy E_G for the Hubbard model is not very accurate, as noted earlier on the basis of numerical investigations of one-dimensional rings [42].

Does the result (19) automatically imply that $|\Psi_G\rangle$ is a poor wave function in the strong-coupling limit? The answer is quite subtle: while it is true that $|\Psi_G\rangle$ is not a very good wave function for the *Hubbard model* at $U \gg t$, it is nevertheless an excellent wavefunction in $d = 1$ for the *t-J model*, the effective model for large U , where doubly occupied sites have been projected out; for a more detailed discussion see Section 2.1 of Ref. [43]. This is demonstrated by the results for the spin-spin correlation function C_j^{SS} , with $j \equiv |\mathbf{R}_j|$. Florian Gebhard, also a diploma student at the Technical University of Munich at that time, to whom I had suggested to calculate correlation functions for the Hubbard model in terms of the Gutzwiller wave function using the technique developed in Refs. [40, 41], was able to analytically evaluate four different correlation functions in $d = 1$ [44, 45]. The result for the spin-spin correlation function explicitly showed that in the strong coupling limit ($U = \infty$) the Gutzwiller wave function describes spin correlations in the nearest-neighbor, isotropic Heisenberg chain extremely well. For $n = 1$ and $U = \infty$ we obtained [44, 45]

$$C_{j>0}^{SS} = (-1)^j \frac{\text{Si}(\pi j)}{\pi j} \underset{j \rightarrow \infty}{\sim} \frac{(-1)^j}{2j} \quad (20)$$

⁵For a more detailed account see Section 2.3 of Ref. [39].

where $Si(x)$ is the sine-integral. The asymptotic behavior implies a logarithmic divergence at momentum $2k_F$, signaling antiferromagnetic fluctuations. Comparison with the exact analytic result for the spin correlation function of the Heisenberg model for $j = 1, 2$ and for large j , where [46, 47] $C_j^{SS} \sim (-1)^j j^{-1} (\ln j)^{1/2}$, shows that the Gutzwiller wave function without doubly occupied sites ($U = \infty$) yields excellent results in $d = 1$ [44, 45]. The same is true for hole-hole correlations in the limit $n \lesssim 1$ and $U = \infty$. Shortly afterwards Haldane [48] and Shastry [49] independently proved that the Gutzwiller wave function for $U = \infty$ is the *exact* solution of the spin- $1/2$ antiferromagnetic Heisenberg chain for an exchange interaction J_{ij} which decreases as⁶ $J_{ij} \sim 1/|i - j|^2$. Thus the Gutzwiller wave function corresponds [48] to the one-dimensional version of Anderson's resonating valence bond (RVB) state [26].

4.1.1 Simplifications in the limit $d \rightarrow \infty$

Our results [40, 41, 44, 45] had demonstrated that in $d = 1$ it was possible to calculate expectation values in terms of the Gutzwiller wave function analytically for all interaction strengths. However, our attempts to generalize this to dimensions $d > 1$ failed. To gain insight into the density dependence of the coefficients of the power series in $g^2 - 1$ in dimensions $d > 1$ Walter Metzner computed the sums over the internal momenta of the diagrams of many-body perturbation theory by Monte-Carlo integration. The results for the lowest-order contribution to the correlation energy for $d = 1$ up to $d = 15$ led to a surprise. Namely, the plot of the results for the second-order diagram as a function of d (Fig. 2) showed that for large d the value of this diagram converged to a simple result which could also be obtained if one assumed that the momenta carried by the lines of a diagram are *independent*, i.e., that there is no momentum conservation at a vertex. When summed over all diagrams this approximation gave exactly the results of the Gutzwiller approximation [40, 41]. Thus we had re-derived the Gutzwiller approximation within conventional many-body perturbation theory! In view of the random generation of momenta in a typical Monte-Carlo integration over momenta we concluded that the assumed independence of momenta at a vertex is correct in the limit of infinite spatial dimensions ($d \rightarrow \infty$). The results of the Gutzwiller approximation thus correspond to the evaluation of expectation values in terms of the Gutzwiller wave function in the limit of infinite dimensions. This provided a straightforward explanation of the mean-field character of the Gutzwiller approximation. The drastic simplifications of diagrammatic calculations in the limit $d \rightarrow \infty$ allow one to calculate expectation values of the kinetic energy and the Hubbard interaction in terms of Gutzwiller-type wave functions exactly [50, 51]. However, these calculations become quite difficult or even untractable when it comes to calculating with Gutzwiller-correlated wave functions of the more general form

$$|\Psi_G\rangle = g^{\hat{D}} |\Psi_0\rangle \quad (21)$$

where $|\Psi_0\rangle$ is no longer the ground state of the Fermi gas, but a more complicated one-particle starting wave function. This has to do with the fact that, in spite of the simplifications arising

⁶This distance dependence of the exchange coupling leads to a partial frustration of the spin orientation whereby the antiferromagnetic correlations are weaker than in the original Heisenberg model.

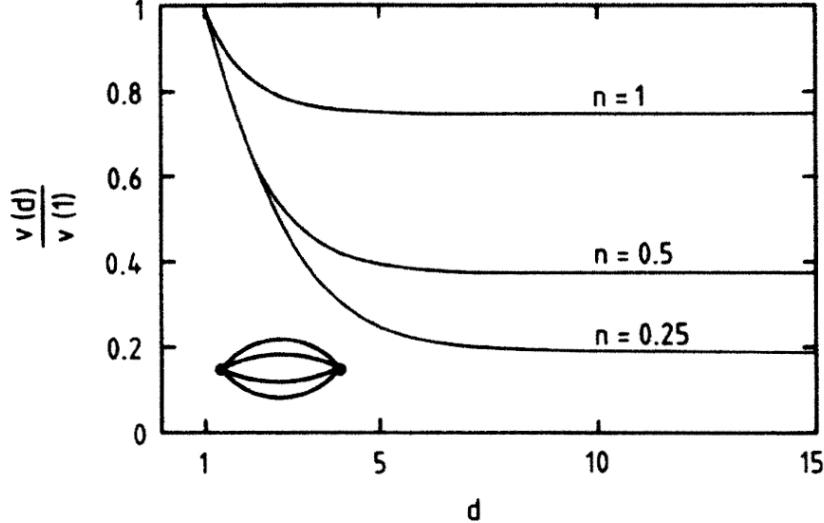


Fig. 2: Value of the second-order diagram for the ground state energy of the Hubbard model (see insert) as calculated with the Gutzwiller wave function for spatial dimensions $d = 1, \dots, 15$, and normalized by the value for $d = 1$, $v(1) = (2/3)(n/2)^3$, where n is the particle density. In the limit of high dimensions the normalized values $v(d)/v(1)$ approach the constant $3n/4$. As discussed in the text the same result is obtained within a diagrammatic approximation that yields the results of the Gutzwiller approximation; from Ref. [41].

from the diagrammatic collapse in $d = \infty$, the remaining diagrams have to be calculated in terms of $|\Psi_0\rangle$. Florian Gebhard [52] showed that this problem can be overcome when $|\Psi_0\rangle$ in (21) is written in the form

$$|\Psi_0\rangle = g^{-\sum_{i\sigma} \mu_{i\sigma} \hat{n}_{i\sigma}} |\tilde{\Psi}_0\rangle \quad (22)$$

where $|\tilde{\Psi}_0\rangle$ is again an arbitrary, normalized one-particle wave function and the local chemical potentials $\mu_{i\sigma}$ are explicit functions of g and the local densities $\tilde{n}_{i\sigma} = \langle \tilde{\Psi}_0 | \hat{n}_{i\sigma} | \tilde{\Psi}_0 \rangle$. The operator in (22) corresponds to a gauge-transformation by which the local chemical potentials can be chosen such that all Hartree bubbles disappear in $d = \infty$. With this re-interpretation all diagrammatic calculations remain identical to the earlier ones, but vertices are given a new value and lines correspond to

$$\tilde{g}_{ij,\sigma}^0 = g_{ij,\sigma}^0 (1 - \delta_{ij}), \quad (23)$$

where now $\tilde{g}_{ii,\sigma}^0 \equiv 0$, and hence $\sum_{i\sigma} \tilde{g}_{ii,\sigma}^0 \equiv 0$. Consequently, in $d = \infty$ results are obtained without the calculation of a single diagram. So what remains in $d = \infty$ at all? First of all one finds that the “law of mass action”, (14), is valid even locally and for arbitrary states $|\tilde{\Psi}_0\rangle$ (even for states with long-range order). Secondly, the expectation value of the Hubbard-Hamiltonian in terms of (21), (22) assumes the following general form for arbitrary $|\tilde{\Psi}_0\rangle$:

$$\langle \hat{H} \rangle = -t \sum_{\langle \mathbf{R}_i, \mathbf{R}_j \rangle} \sum_{\sigma} \sqrt{q_{i\sigma}} \sqrt{q_{j\sigma}} g_{ij,\sigma}^0 + U \sum_i \bar{d}_i, \quad (24)$$

where $\langle \mathbf{R}_i, \mathbf{R}_j \rangle$ denotes nearest-neighbor sites, $\bar{d}_i = \langle \hat{D}_i \rangle$, and $q_{i\sigma}$ is given by $q_{i\sigma}$ in (15) with $n_{i\sigma}$ replaced by $\tilde{n}_{i\sigma}$. In the translationally invariant case $|\Psi_0\rangle \equiv |\text{FG}\rangle$ the two wave-functions $|\Psi_0\rangle$ and $|\tilde{\Psi}_0\rangle$ are the same up to a trivial factor, and $q_{i\sigma} \equiv q_i$, whereby (15) is re-derived.

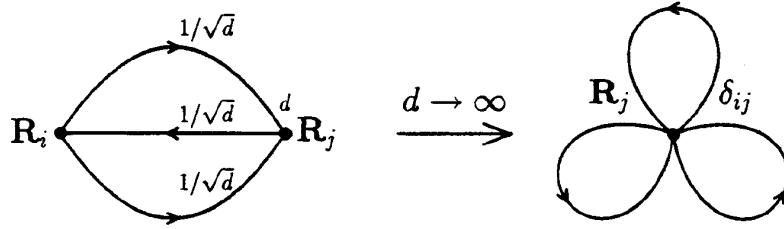


Fig. 3: Contribution to the irreducible self-energy for the Hubbard model in second-order perturbation theory in U and its collapse in the limit $d \rightarrow \infty$.

Interestingly, the result (24) is identical to the saddle-point solution of the slave-boson mean-field theory for the Hubbard model [38] mentioned at the beginning of Section 4. In fact, one finds that in $d = \infty$ the general set of Gutzwiller-correlated wave functions (21) with (22) reproduce the full set of static saddle-point equations of the slave-boson approach. This provides a direct connection between two seemingly different approaches, the slave-boson mean-field theory and the diagrammatic calculation of expectation values in terms of the Gutzwiller wave function in the limit $d = \infty$. It also shows that the slave-boson approach obeys the variational principle and is valid for an arbitrary starting wave function $|\tilde{\Psi}_0\rangle$; for a brief review see Ref. [53].

Calculations with the Gutzwiller wave function in $d = \infty$ are thus possible without the calculation of a single diagram. Later Gebhard and collaborators generalized this approach to multi-band Hubbard models. This led them to the formulation of a *Gutzwiller density-functional theory* which can be used to calculate, for example, the dispersion of quasi-particle excitations in the Fermi liquid state of transition metals and other materials [54, 55].

4.2 Lattice fermions in infinite spatial dimensions

By studying the Hubbard model with the Gutzwiller wave function Walter Metzner and I had found that in the limit $d \rightarrow \infty$ diagrammatic calculations greatly simplify. Apparently, this limit was not only useful for the investigation of spin models, but also in the case of lattice fermions. To better understand this point, we analyzed the diagrams involved in the calculation of expectation values with the Gutzwiller wave function in more detail. As mentioned earlier, the form of the diagrams is identical to that of the usual Feynman diagrams in many-body perturbation theory, but lines correspond to one-particle density matrices, $g_{ij,\sigma}^0 = \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} \rangle_0$. We showed that in the limit $d \rightarrow \infty$ diagrams collapse in position space [50, 51], such that only local contributions remain (Fig. 3). In other words, momentum conservation at a vertex of a skeleton diagram becomes irrelevant in the limit $d \rightarrow \infty$, implying that the momenta carried by the lines of a graph are indeed independent. In particular, the diagrams contributing to the proper self-energy are purely diagonal in $d = \infty$.

4.2.1 Diagrammatic collapse in $d = \infty$

The reason behind the diagrammatic collapse can be understood as follows. The one-particle density matrix may be interpreted as the amplitude for transitions between site \mathbf{R}_i and \mathbf{R}_j . The square of its absolute value is therefore proportional to the *probability* for a particle to hop from \mathbf{R}_j to a site \mathbf{R}_i . In the case of nearest-neighbor sites \mathbf{R}_i , \mathbf{R}_j on a lattice with coordination number Z , this implies $|g_{ij,\sigma}^0|^2 \sim \mathcal{O}(1/Z)$. For nearest-neighbor sites \mathbf{R}_i , \mathbf{R}_j on a hypercubic lattice (where $Z = 2d$), one therefore finds for large d

$$g_{ij,\sigma}^0 \sim \mathcal{O}\left(\frac{1}{\sqrt{d}}\right). \quad (25)$$

For general i, j one finds [56, 51]

$$g_{ij,\sigma}^0 \sim \mathcal{O}\left(1/d^{\|\mathbf{R}_i - \mathbf{R}_j\|/2}\right), \quad (26)$$

where $\|\mathbf{R}\| = \sum_{n=1}^d |R_n|$ is the length of \mathbf{R} in the Manhattan metric.

It is important to bear in mind that, although $g_{ij,\sigma}^0 \sim 1/\sqrt{d}$ vanishes for $d \rightarrow \infty$, the particles are not localized but are still mobile. Indeed, even in the limit $d \rightarrow \infty$ the off-diagonal elements of $g_{ij,\sigma}^0$ contribute, since a particle may hop to d nearest neighbors with reduced amplitude t^*/\sqrt{d} . For non-interacting electrons at $T = 0$ the expectation value of the kinetic energy is given by

$$E_{\text{kin}}^0 = -t \sum_{\langle \mathbf{R}_i, \mathbf{R}_j \rangle} \sum_{\sigma} g_{ij,\sigma}^0. \quad (27)$$

On a hypercubic lattice the sum over the nearest neighbors (NN) leads to a factor $\mathcal{O}(d)$. In view of the $1/\sqrt{d}$ dependence of $g_{ij,\sigma}^0$ it is therefore necessary to scale the NN-hopping amplitude t

$$t \rightarrow \frac{t^*}{\sqrt{d}}, \quad t^* = \text{const.}, \quad (28)$$

since only then the kinetic energy remains finite for $d \rightarrow \infty$. The same result is obtained in a momentum-space formulation.⁷

A rescaling of the microscopic parameters of the Hubbard model with d is only required in the kinetic energy. Namely, since the interaction term is purely local, it is independent of the spatial dimension. Altogether this implies that only the Hubbard Hamiltonian with a rescaled kinetic energy

$$\hat{H} = -\frac{t^*}{\sqrt{d}} \sum_{\langle \mathbf{R}_i, \mathbf{R}_j \rangle} \sum_{\sigma} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + U \sum_{\mathbf{R}_i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \quad (29)$$

has a non-trivial $d \rightarrow \infty$ limit where both terms, the kinetic energy and the interaction, are of the same order of magnitude in d .

⁷This can be seen by calculating the density of states (DOS) of non-interacting particles. For nearest-neighbor hopping on a d -dimensional hypercubic lattice $\varepsilon_{\mathbf{k}}$ has the form $\varepsilon_{\mathbf{k}} = -2t \sum_{i=1}^d \cos k_i$ (here and in the following we set Planck's constant \hbar , Boltzmann's constant k_B , and the lattice spacing equal to unity). The DOS corresponding to $\varepsilon_{\mathbf{k}}$ is given by $N_d(\omega) = \sum_{\mathbf{k}} \delta(\omega - \varepsilon_{\mathbf{k}})$, which is the probability density for finding the value $\omega = \varepsilon_{\mathbf{k}}$ for a random choice of $\mathbf{k} = (k_1, \dots, k_d)$. If the momenta k_i are chosen randomly, $\varepsilon_{\mathbf{k}}$ is the sum of d independent (random) numbers $-2t \cos k_i$. The central limit theorem then implies that in the limit $d \rightarrow \infty$ the DOS is given by a Gaussian, i.e., $N_d(\omega) \xrightarrow{d \rightarrow \infty} \frac{1}{2t\sqrt{\pi d}} \exp\left[-\left(\frac{\omega}{2t\sqrt{d}}\right)^2\right]$. Only if t is scaled with d as in (28) does one obtain a non-trivial DOS $N_\infty(\omega)$ in $d = \infty$ [57, 50] and thus a finite kinetic energy.

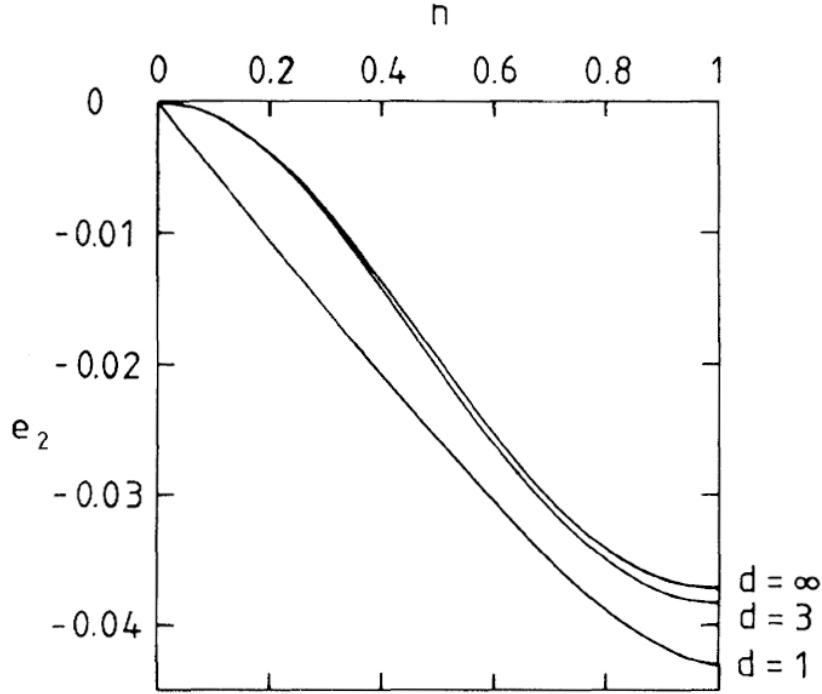


Fig. 4: Correlation energy $E_c^{(2)} = (2U^2/|\varepsilon_0|)e_2$ of the Hubbard model calculated in second-order Goldstone perturbation theory in U vs. density n for dimensions $d = 1, 3, \infty$. Here ε_0 is the kinetic energy for $U = 0$ and $n = 1$; from Ref. [50].

4.2.2 Simplifications of quantum many-body perturbation theory in the limit $d \rightarrow \infty$

Walter and I now wanted to understand to what extent the simplifications that occur in diagrammatic calculations with the Gutzwiller wave function in $d = \infty$ carry over to general many-body calculations for the Hubbard model. For this purpose, we evaluated the second-order diagram in Goldstone perturbation theory [58] that determines the correlation energy at weak coupling [50]. Due to the diagrammatic collapse in $d = \infty$, calculations were again found to be much simpler.⁸ Namely, the nine-dimensional integral in $d = 3$ over the three internal momenta reduces to a single integral in $d = \infty$, implying that in $d = \infty$ the calculation is simpler than in any other dimension. More importantly, the numerical value obtained in $d = \infty$ turned out to be very close to that in the physical dimension $d = 3$ and therefore provides an easily tractable, quantitatively reliable approximation (see Fig. 4).

These results clearly showed that microscopic calculations for correlated lattice fermions in $d = \infty$ dimensions were useful and very promising. Further insights were made quickly: Müller-Hartmann [59] showed that in infinite dimensions only on-site interactions remain dynamical,

⁸The one-particle Green function (propagator) $G_{ij,\sigma}^0(\omega)$ of the non-interacting system obeys the same $1/\sqrt{d}$ dependence as the one-particle density matrix $g_{ij,\sigma}^0$ (see (25)). This follows directly from $g_{ij,\sigma}^0 = \lim_{t \rightarrow 0^-} G_{ij,\sigma}^0(t)$ and the fact that the scaling properties do not depend on the time evolution and the quantum mechanical representation. The Fourier transform of $G_{ij,\sigma}^0(\omega)$ also preserves this property. For this reason the same results as those obtained in the calculation with the Gutzwiller wave function hold: all connected one-particle irreducible diagrams collapse in position space, i.e., they are purely diagonal in $d = \infty$.

that the proper self-energy becomes momentum-independent

$$\Sigma_\sigma(\mathbf{k}, \omega) \stackrel{d \rightarrow \infty}{\equiv} \Sigma_\sigma(\omega) \quad (30a)$$

and hence is purely local in position space

$$\Sigma_{ij,\sigma}(\omega) \stackrel{d \rightarrow \infty}{=} \Sigma_{ii,\sigma}(\omega) \delta_{ij}, \quad (30b)$$

as in the case of diagrams calculated with the Gutzwiller wave function [50, 51], and therefore typical Fermi liquid features are preserved (Sec. 4.2.5) [60]. Schweitzer and Czycholl [61] demonstrated that calculations for the periodic Anderson model also become much simpler in high dimensions.⁹ In particular, Brandt and Mielsch [65] derived the exact solution of the Falicov-Kimball model for infinite dimensions by mapping the lattice problem onto a solvable atomic problem in a generalized, time-dependent external field.¹⁰ They also noted that such a mapping is, in principle, also possible for the Hubbard model.

Due to the property (30), the most important obstacle for actual diagrammatic calculations in finite dimensions $d \geq 1$ – namely, the integration over intermediate momenta – is greatly simplified in $d = \infty$. Nevertheless, the limit $d \rightarrow \infty$ does not affect the *dynamics* of the system. Hence, in spite of the simplifications in position (or momentum) space, the problem retains its full dynamics in $d = \infty$.

4.2.3 Interactions beyond the on-site interaction

In the case of more general interactions than the Hubbard interaction, e.g., nearest neighbor interactions such as

$$\hat{H}_{nn} = \sum_{\langle \mathbf{R}_i, \mathbf{R}_j \rangle} \sum_{\sigma\sigma'} V_{\sigma\sigma'} \hat{n}_{i\sigma} \hat{n}_{j\sigma'} \quad (31)$$

the interaction constant has to be scaled, too, in the limit $d \rightarrow \infty$. In the case of (31), which has the form of a classical interaction, the “classical” scaling

$$V_{\sigma\sigma'} \rightarrow \frac{V_{\sigma\sigma}^*}{Z} \quad (32)$$

is required. Of course, the propagator still has the dependence (26). Due to (32), all contributions, except for the Hartree-term, are found to vanish in $d = \infty$ [59]. Hence, nonlocal interactions only contribute through their Hartree approximation, which is purely static. This gives the Hubbard interaction a unique role: of all interactions for fermionic lattice models only the Hubbard interaction remains dynamical in the limit $d \rightarrow \infty$ [59].

⁹For a more detailed discussion of the simplifications occurring in the investigation of Hubbard-type lattice models or the t - J model [62, 63] in high dimensions see Ref. [64].

¹⁰Alternatively, it can be shown that in the limit $Z \rightarrow \infty$ the dynamics of the Falicov-Kimball model reduces to that of a *non-interacting*, tight-binding model on a Bethe lattice with coordination number $Z = 3$, which can thus be solved exactly [66].

4.2.4 One-particle propagator

Due to the \mathbf{k} -independence of the irreducible self-energy, (30a), the one-particle propagator of an interacting lattice fermion system is given by

$$G_{\mathbf{k},\sigma}(\omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} + \mu - \Sigma_{\sigma}(\omega)}. \quad (33)$$

Most importantly, the \mathbf{k} dependence of $G_{\mathbf{k}}(\omega)$ comes entirely from the energy dispersion $\varepsilon_{\mathbf{k}}$ of the *non*-interacting particles. This means that in a homogeneous system described by the propagator

$$G_{ij,\sigma}(\omega) = \frac{1}{L} \sum_{\mathbf{k}} G_{\mathbf{k},\sigma}(\omega) e^{i\mathbf{k}\cdot(\mathbf{R}_i - \mathbf{R}_j)}, \quad (34)$$

its local part, $G_{ii,\sigma}$, is given by

$$G_{ii,\sigma}(\omega) = \frac{1}{L} \sum_{\mathbf{k}} G_{\mathbf{k},\sigma}(\omega) = \int_{-\infty}^{\infty} d\varepsilon \frac{N_0(\varepsilon)}{\omega - \varepsilon + \mu - \Sigma_{\sigma}(\omega)}, \quad (35)$$

where $N_0(\varepsilon)$ is the density of states of the non-interacting system. In the paramagnetic phase we can suppress site and spin indices and write $G_{ii,\sigma}(\omega) \equiv G(\omega)$. The spectral function of the interacting system (often referred to as the DOS as in the non-interacting case) is then given by

$$A(\omega) = -\frac{1}{\pi} \text{Im } G(\omega + i0^+). \quad (36)$$

4.2.5 Consequences of the \mathbf{k} -independence of the self-energy: Fermi liquid behavior

We now discuss some further consequences of the \mathbf{k} -independence of the self-energy in the paramagnetic phase as derived by Müller-Hartmann [60]. At $T = 0$, the one-particle propagator (33) takes the form (again we suppress the spin index)

$$G_{\mathbf{k}}(\omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} + E_F - \Sigma(\omega)}. \quad (37)$$

In general, even when $\Sigma(\omega)$ is \mathbf{k} -dependent, the Fermi surface is defined by the $\omega = 0$ limit of the denominator of (37) as

$$\varepsilon_{\mathbf{k}} + \Sigma_{\mathbf{k}}(0) = E_F. \quad (38a)$$

According to Luttinger and Ward [67], the volume within the Fermi surface is not changed by interactions, provided the latter can be treated in perturbation theory. This is expressed by

$$n = \sum_{\mathbf{k}\sigma} \Theta(E_F - \varepsilon_{\mathbf{k}} - \Sigma_{\mathbf{k}}(0)), \quad (38b)$$

where n is the particle density and $\Theta(x)$ is the step function. The \mathbf{k} -dependence of $\Sigma_{\mathbf{k}}(0)$ in (38a) implies that, in spite of (38b), the shape of the Fermi surface of the interacting system will be quite different from that of the non-interacting system (except for the rotationally invariant

case $\varepsilon_{\mathbf{k}} = f(|\mathbf{k}|)$. By contrast, for lattice fermion models in $d = \infty$, where $\Sigma_{\mathbf{k}}(\omega) \equiv \Sigma(\omega)$, the Fermi surface itself (and hence the volume enclosed) is not changed by interactions. The Fermi energy is simply shifted uniformly from its non-interacting value E_F^0 , i.e., $E_F = E_F^0 + \Sigma(0)$, to keep n in (38b) constant. Thus the $\omega = 0$ value of the local propagator, $G(0)$, and hence of the spectral function, $A(0) = -\frac{1}{\pi} \text{Im } G(i0^+)$, is not changed by interactions. This behavior is well-known from the single-impurity Anderson model [4]. Renormalizations of $N(0)$ can only come from a \mathbf{k} -dependence of Σ , i.e., if $\partial\Sigma/\partial\mathbf{k} \neq 0$.

For $\omega \rightarrow 0$ the self-energy has the property

$$\text{Im } \Sigma(\omega) \propto \omega^2, \quad (38c)$$

which implies Fermi liquid behavior. The effective mass of the quasiparticles

$$\frac{m^*}{m} = 1 - \frac{d\Sigma}{d\omega} \Big|_{\omega=0} = 1 + \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{\text{Im}\Sigma(\omega + i0^-)}{\omega^2} \geq 1 \quad (38d)$$

is seen to be enhanced. In particular, the momentum distribution

$$n_{\mathbf{k}} = \frac{1}{\pi} \int_{-\infty}^0 d\omega \text{Im } G_{\mathbf{k}}(\omega) \quad (39)$$

has a discontinuity at the Fermi surface, given by $n_{k_F^-} - n_{k_F^+} = (m^*/m)^{-1}$, where $k_F^\pm = k_F \pm 0^+$.

5 Dynamical mean-field theory for correlated lattice fermions

The diagrammatic simplifications of many-body perturbation theory in infinite spatial dimensions provide the basis for the construction of a comprehensive mean-field theory for lattice fermions that is diagrammatically controlled and whose free energy has no unphysical singularities. The construction is based on the scaled Hamiltonian (29). Since the self-energy is momentum independent but retains its frequency dependence, i.e., describes the full many-body dynamics of the interacting system,¹¹ the resulting theory is mean-field-like *and* dynamical and hence represents a *dynamical mean-field theory* (DMFT) for lattice fermions which is able to describe genuine correlation effects.

5.1 Derivation of the self-consistent DMFT equations

The DMFT equations can be derived in different ways. They all employ the fact that in $d = \infty$ lattice fermion models with a local interaction reduce to an effective many-body problem whose dynamics corresponds to that of correlated fermions on a single site embedded in a bath provided by the other fermions. This is illustrated in Fig. 5.

The single-site action and the DMFT equations were first derived by Václav Janiš [68] within a generalization of the self-consistent coherent potential approximation (CPA)¹² to lattice fermion

¹¹This is in contrast to Hartree-Fock theory where the self-energy is merely a static potential.

¹²The CPA is a well-known mean-field theory for non-interacting, disordered systems. It becomes exact in the limit $d, Z \rightarrow \infty$ [69].

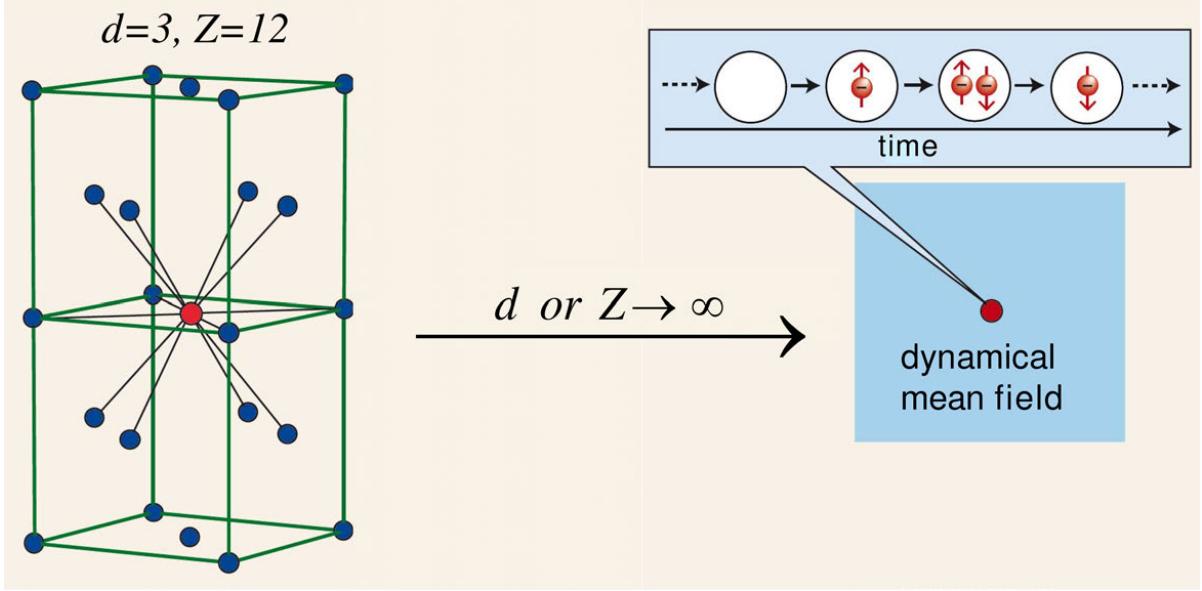


Fig. 5: In the limit $Z \rightarrow \infty$ the Hubbard model effectively reduces to a dynamical single-site problem, which may be viewed as a lattice site embedded in a dynamical mean field. Electrons may hop from the mean field onto this site and back, and interact on the site as in the original Hubbard model (see Fig. 1). The local propagator $G(\omega)$, i.e., the return amplitude, and the dynamical self-energy $\Sigma(\omega)$ of the surrounding mean field play the main role in this limit. The quantum dynamics of the interacting electrons is still described exactly.

models with local interaction and local self-energy, such as the Falicov-Kimball and Hubbard model in the limit $d = \infty$; for details see Refs. [68, 70, 18]. Shortly after that Václav joined me, then at the RWTH Aachen University, on an Alexander-von-Humboldt fellowship. Before we could start to solve the self-consistency equations [70], I received a preprint from Antoine Georges and Gabi Kotliar [71] in which they had formulated the DMFT by mapping the lattice problem onto a self-consistent single-impurity Anderson model. This mapping was also employed by Mark Jarrell [72].

Although the DMFT equations derived within the CPA approach and the single-impurity approach are identical, the latter was immediately adopted by the community since it is connected with the well-studied theory of quantum impurities and Kondo problems [4], for whose solution efficient numerical codes such as the quantum Monte-Carlo (QMC) method [73] had been developed and were readily available. For this reason the single-impurity based derivation of the DMFT immediately became the standard approach. For a detailed discussion see the review by Georges, Kotliar, Krauth, and Rozenberg [74]; for an introductory presentation see the article by Gabi Kotliar and myself [75].

The DMFT equations are given by

(A) the *local propagator* $G_\sigma(i\omega_n)$, which is expressed as a functional integral

$$G_\sigma(i\omega_n) = -\frac{1}{Z} \int \prod_\sigma Dc_\sigma^* Dc_\sigma [c_\sigma(i\omega_n)c_\sigma^*(i\omega_n)] e^{-S_{\text{loc}}} \quad (40)$$

with the partition function

$$\mathcal{Z} = \int \prod_{\sigma} Dc_{\sigma}^* Dc_{\sigma} e^{-S_{\text{loc}}} \quad (41)$$

and the local action

$$S_{\text{loc}} = - \int_0^{\beta} d\tau_1 \int_0^{\beta} d\tau_2 \sum_{\sigma} c_{\sigma}^*(\tau_1) \mathcal{G}_{\sigma}^{-1}(\tau_1 - \tau_2) c_{\sigma}(\tau_2) + U \int_0^{\beta} d\tau c_{\uparrow}^*(\tau) c_{\uparrow}(\tau) c_{\downarrow}^*(\tau) c_{\downarrow}(\tau). \quad (42)$$

Here \mathcal{G}_{σ} is the effective local propagator (also called *bath Green function*, or *Weiss mean field*)¹³ which is defined by a Dyson-type equation

$$\mathcal{G}_{\sigma}(i\omega_n) = \left([G_{\sigma}(i\omega_n)]^{-1} + \Sigma_{\sigma}(i\omega_n) \right)^{-1}. \quad (43)$$

Furthermore, by identifying the Hilbert transform of the *lattice Green function*

$$G_{\mathbf{k}\sigma}(i\omega_n) = \frac{1}{i\omega_n - \varepsilon_{\mathbf{k}} + \mu - \Sigma_{\sigma}(i\omega_n)} \quad (44)$$

with the local propagator (40) one obtains

(B) the *self-consistency condition*

$$G_{\sigma}(i\omega_n) = \frac{1}{L} \sum_{\mathbf{k}} G_{\mathbf{k}\sigma}(i\omega_n) = \int_{-\infty}^{\infty} d\varepsilon \frac{N(\varepsilon)}{i\omega_n - \varepsilon + \mu - \Sigma_{\sigma}(i\omega_n)} \quad (45)$$

$$= G_{\sigma}^0(i\omega_n - \Sigma_{\sigma}(i\omega_n)). \quad (46)$$

In (45) the ionic lattice on which the electrons live is seen to enter only through the DOS of the non-interacting electrons. Eq. (46) illustrates the mean-field character of the DMFT-equations particularly clearly: the local Green function of the interacting system is given by the *non-interacting* Green function G_{σ}^0 with a renormalized energy $i\omega_n - \Sigma_{\sigma}(i\omega_n)$, which corresponds to the energy $i\omega_n$ measured relative to the energy of the surrounding dynamical fermionic bath, i.e., the energy of the mean field $\Sigma_{\sigma}(i\omega_n)$.

The self-consistent DMFT equations can be solved iteratively: starting with an initial value for the self-energy $\Sigma_{\sigma}(i\omega_n)$ one obtains the local propagator $G_{\sigma}(i\omega_n)$ from (45) and thereby the bath Green function $\mathcal{G}_{\sigma}(i\omega_n)$ from (43). This determines the local action (42) that is needed to compute a new value for the local propagator $G_{\sigma}(i\omega_n)$ from (40) and, by employing the old self-energy, a new bath Green function \mathcal{G}_{σ} , and so on. In spite of the fact that the solution can be obtained self-consistently, there remains a complicated many-body problem which is generally not exactly solvable. A generalization of the DMFT equations for the Hubbard model in the presence of local disorder was derived in Ref. [76].

It should be stressed that although the DMFT corresponds to an effectively local problem, the propagator $G_{\mathbf{k}}(\omega)$ is a momentum-dependent quantity. Namely, it depends on the momentum through the dispersion $\varepsilon_{\mathbf{k}}$ of the non-interacting electrons. However, there is no additional momentum-dependence through the self-energy, since it is strictly local within the DMFT.

¹³In principle, the local functions $\mathcal{G}_{\sigma}(i\omega_n)$ and $\Sigma_{\sigma}(i\omega_n)$ can both be viewed as a “dynamical mean field” acting on particles on a site, since they all appear in the bilinear term of the local action (42).

5.1.1 Solution of the self-consistent DMFT equations

The dynamics of the Hubbard model remains complicated even in the limit $d = \infty$ due to the purely local nature of the interaction. Hence an exact, analytic evaluation of the self-consistent set of equations for the local propagator G_σ or the effective propagator $\mathcal{G}_\sigma(i\omega_n)$ is not possible. A valuable semi-analytic approximation is provided by the *iterated perturbation theory* (IPT) [71, 77]. Exact evaluations become feasible when there is no coupling between the frequencies. This is the case, for example, in the Falicov-Kimball model [65, 78].

Solutions of the self-consistent DMFT equations require extensive numerical methods, in particular quantum Monte Carlo simulations [72, 79, 80, 74, 81], the numerical renormalization group [82–84], exact diagonalization [85–87], and other techniques.

It quickly turned out that the DMFT is a powerful tool for the investigation of electronic systems with strong correlations [88, 74]. It provides a non-perturbative and thermodynamically consistent approximation scheme for finite-dimensional systems that is particularly valuable for the study of intermediate-coupling problems where perturbative techniques fail; for detailed discussions see Refs. [89, 75, 90, 91, 43].

5.2 The LDA+DMFT approach to correlated materials

The Hubbard model is able to explain important general features of correlated electrons, but it cannot describe the physics of real materials in any detail. Namely, realistic approaches must take into account the explicit electronic and lattice structure of the systems.

For a long time the electronic properties of solids were investigated by two essentially separate communities, one using model Hamiltonians in conjunction with many-body techniques, the other employing density functional theory (DFT) [92, 93]. DFT and its local-density approximation (LDA) are *ab initio* approaches that do not require empirical parameters as input. They proved to be highly successful techniques for the calculation of the electronic structure of real materials [94]. Still, it was soon recognized that DFT/LDA is severely restricted in its ability to describe strongly correlated materials such as *f*-electron systems and Mott insulators. For such systems the model Hamiltonian approach is more powerful since there exist systematic theoretical techniques to investigate the many-electron problem with increasing accuracy. Nevertheless, the uncertainty in the choice of the model parameters and the technical complexity of the correlation problem itself prevent the model Hamiltonian approach from being flexible enough to study real materials. The two approaches are therefore largely complementary. In view of the individual power of DFT/LDA and the model Hamiltonian approach, respectively, a combination of these techniques for *ab initio* investigations of real materials is clearly desirable. One of the first successful attempts in this direction was the LDA+U method [95, 96], which combines LDA with a Hartree-like, static mean-field approximation for a multi-band Anderson lattice model. This method turned out to be a very useful tool in the study of long-range ordered, insulating states of transition metals and rare-earth compounds. However, the *paramagnetic* metallic phase of correlated electron systems clearly requires a treatment which includes dynamical effects, i.e., the frequency dependence of the self-energy.

Here the so-called LDA+DMFT approach, whose foundations were laid in the papers by Anisimov, Poteryaev, Korotin, Anokhin, and Kotliar [97] as well as Lichtenstein and Katsnelson [98], has led to an enormous progress in our understanding of correlated electron materials [97–106, 75]. LDA+DMFT is a computational scheme that merges electronic band structure calculations in the local density approximation (LDA) or generalized-gradient approximations (GGA) with many-body physics originating from the local Hubbard interaction and Hund’s rule coupling terms, and then solves the corresponding correlation problem by DMFT. Sometimes this combined approach is also referred to as DFT+DMFT.

As in the case of the Hubbard model the many-body model constructed within the LDA+DMFT scheme consists of two parts: a kinetic energy which describes the specific band structure of the uncorrelated electrons, and the local interactions between the electrons in the same orbital as well as in different orbitals. It is then necessary to take into account a double counting of the interaction, since the LDA already includes some of the static contributions of the electronic interaction; for details, see Refs. [101–106]). This complicated many-particle problem with its numerous energy bands and local interactions is then solved within DMFT, usually by the application of quantum Monte-Carlo (QMC) techniques. By construction, LDA+DMFT includes the correct quasiparticle physics and the corresponding energetics. It also reproduces the LDA results in the limit of weak Coulomb interaction U . More importantly, LDA+DMFT correctly describes the correlation induced dynamics near a Mott-Hubbard metal-insulator transition and beyond. Thus, LDA+DMFT and related, material-specific dynamical mean-field approaches that are presently being developed [107–109] are, in principle, able to account for the physics at all values of the Coulomb interaction and doping.

6 Summary and outlook

By now the DMFT has developed into a powerful method for the investigation of electronic systems with strong correlations. It provides a comprehensive, non-perturbative and thermodynamically consistent approximation scheme for the investigation of finite-dimensional systems (in particular for dimension $d = 3$), and is particularly useful for the study of problems where perturbative approaches fail. For this reason, the DMFT has now become the standard mean-field theory for fermionic correlation problems. The generalization of this approach and its applications is currently a subject of active research. Here non-local generalizations of the DMFT play an important role [90, 81]. They make it possible to study and explain even short range correlation effects which occur on the scale of several lattice constants. Furthermore, investigations of inhomogeneous bulk systems and of internal and external inhomogeneities, such as surfaces and interfaces [110–116], lead to an improved understanding of correlation effects in thin films and multi-layered nanostructures. This is particularly desirable in view of the novel functionalities of these structures and their possible applications in electronic devices.

The investigation of correlation phenomena in the field of cold atoms in optical lattices is another intriguing field of current research. Within a short time it led to the development of a versatile instrument for the simulation and investigation of quantum mechanical many-particle

systems [117–121]. While for electrons in solids the Hubbard model with its purely local interaction is a rather strong assumption, it can describe cold atoms in optical lattices very accurately since the interaction between the atoms is indeed extremely short ranged. Here the DMFT has once again proved to be extremely useful. Experiments with cold atoms in optical lattices can even assess the quality of the results of the DMFT. The results obtained in this way show that the DMFT indeed leads to reliable results even for finite dimensional systems [120].

The study of correlated electrons out of equilibrium within non-equilibrium DMFT [122–128] has become yet another fascinating new research area. Non-equilibrium DMFT will be able to explain, and even predict, the results of time-resolved experiments; for an upcoming review, see Ref. [129].

The combination of the DMFT with methods for the computation of electronic band structures (LDA+DMFT) has led to a conceptually new theoretical framework for the realistic study of correlated materials. In 10 to 15 years from now, DMFT-based approaches can be expected to be as successful and standardized as the presently available density-functional methods. The development of a comprehensive theoretical approach which allows for a quantitative understanding and prediction of correlation effects in materials, ranging from complex inorganic materials all the way to biological systems, is one of the great challenges for modern theoretical physics. For details I refer to the scientific program of the Research Unit FOR 1346 *Dynamical Mean-Field Approach with Predictive Power for Strongly Correlated Materials* [130] which is being funded by the Deutsche Forschungsgemeinschaft since 2010. The Research Unit FOR 1346 initiated the series of Autumn Schools on correlated materials which are held at the Forschungszentrum Jülich since 2011. The lecture notes of these Autumn Schools provide an excellent introduction into this very active field of research [131–134].

Acknowledgment

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

References

- [1] N.F. Mott, Rev. Mod. Phys. **40**, 677 (1968)
- [2] N.F. Mott, *Metal-Insulator Transitions* (Taylor and Francis, London, 2nd ed., 1990)
- [3] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
- [4] A.C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, 1997)
- [5] N. Grewe and F. Steglich, in *Handbook on the Physics and Chemistry of Rare Earths*, Vol. 14, ed. by K.A. Gschneidner, L. Eyring (North Holland, Amsterdam, 1991), p. 343
- [6] K. Baberschke, M. Donath, and W. Nolting (eds.): *Band-Ferromagnetism*, Lecture Notes in Physics, Vol. 580, (Springer, Heidelberg, 2001)
- [7] J.R. Schrieffer (ed.): *Handbook of High-Temperature Superconductivity* (Springer, Berlin, 2007)
- [8] E. Dagotto: *Nanoscale Phase Separation and Colossal Magnetoresistance* (Springer, Berlin, 2002)
- [9] H. v. Löhneysen, A. Rosch, M. Vojta, and P. Wölfle, Rev. Mod. Phys. **79**, 1015 (2007)
- [10] Y. Tokura, Physics Today, July 2003, p. 50.
- [11] J.H. de Boer and E.J.W. Verwey, Proc. Phys. Soc. **49**, No. **4S**, 59 (1937)
- [12] N.F. Mott and R. Peierls, Proc. Phys. Soc. A **49**, 72 (1937)
- [13] M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
- [14] J. Hubbard, Proc. Roy. Soc. London A **276**, 238 (1963)
- [15] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963)
- [16] E. Lieb and F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
- [17] P. Fulde: *Electron Correlations in Molecules and Solids* (Springer, Berlin, 1995)
- [18] D. Vollhardt, p. 339 in *Lectures on the Physics of Strongly Correlated Systems XIV*, AIP Conference Proceedings vol. 1297, ed. by A. Avella and F. Mancini (American Institute of Physics, Melville, 2010); <http://arxiv.org/abs/1004.5069v3>
- [19] R.J. Baxter: *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1982)
- [20] E. Feenberg: *Theory of Quantum Fluids* (Academic, New York, 1969)

- [21] R.P. Feynman: *Statistical Physics* (Benjamin, Reading, 1972)
- [22] V.R. Pandharipande and R.B. Wiringa, Rev. Mod. Phys. **51**, 821 (1979)
- [23] R.B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983)
- [24] C.M. Varma and Y. Yafet, Phys. Rev. B **13**, 2950 (1976)
- [25] O. Gunnarsson and K. Schönhammer, Phys. Rev. B **28**, 4315 (1983)
- [26] P.W. Anderson, P.A. Lee, M. Randeria, T.M. Rice, N. Trivedi, and F.C. Zhang, J. Phys.: Condens. Matter **16**, R755 (2004)
- [27] J. Bardeen, L.N. Cooper and J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957)
- [28] M.C. Gutzwiller, Phys. Rev. **137**, A1726 (1965)
- [29] T. Ogawa, K. Kanda, and T. Matsubara, Prog. Theor. Phys. **53**, 614 (1975)
- [30] D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)
- [31] P.G.J. van Dongen and D. Vollhardt, Phys. Rev. B **40**, 7252 (1989)
- [32] W.F. Brinkman and T.M. Rice, Phys. Rev. B **2**, 4302 (1970)
- [33] T.M. Rice and W.F. Brinkman, in *Alloys, Magnets, and Superconductors*, ed. by R.E. Mills, E. Ascher, and R. Jaffee (McGraw-Hill, New York, 1971), p. 593
- [34] K. Levin and O.T. Valls, Phys. Rep. **98**, 1 (1983)
- [35] P.W. Anderson and W.F. Brinkman, in *The Physics of Liquid and Solid Helium*, Part II, ed. by K.H. Bennemann and J.B. Ketterson (Wiley, New York, 1978), p. 177
- [36] D. Vollhardt, P. Wölfle, and P.W. Anderson, Phys. Rev. B **35**, 6703 (1987)
- [37] For a personal account by Andrei Ruckenstein of the developments leading to the formulation of the Kotliar-Ruckenstein slave-boson mean-field theory see:
[http://www.aspenphys.org/aboutus/history/
presidentialessays/ruckenstein.html](http://www.aspenphys.org/aboutus/history/presidentialessays/ruckenstein.html)
- [38] G. Kotliar and A.E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986)
- [39] D. Vollhardt, Ann. Phys. (Berlin), **524**, 1 (2012) [Einstein Lecture]
<http://dx.doi.org/10.1002/andp.201100250>
- [40] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987)
- [41] W. Metzner and D. Vollhardt, Phys. Rev. B **37**, 7382 (1988)
- [42] T.A. Kaplan, P. Horsch and P. Fulde, Phys. Rev. Lett. **49**, 889 (1982)

- [43] D. Vollhardt, p. 31 in *Proceedings of the International School of Physics Enrico Fermi, Course CXXI*, ed. by R.A. Broglia and J.R. Schrieffer
(North-Holland, Amsterdam, 1994); http://www.physik.uni-augsburg.de/theo3/Research/research_varenn.vollha.de.shtml
- [44] F. Gebhard and D. Vollhardt, Phys. Rev. Lett. **59**, 1472 (1987)
- [45] F. Gebhard and D. Vollhardt, Phys. Rev. B **38**, 6911 (1988)
- [46] R.R.P. Singh, M.E. Fisher, and R. Shankar, Phys. Rev. B **39**, 2562 (1989)
- [47] T. Giamarchi and H.J. Schulz, Phys. Rev. B **39**, 4620 (1989)
- [48] F.D.M. Haldane, Phys. Rev. Lett. **60**, 635 (1988)
- [49] B.S. Shastry, Phys. Rev. Lett. **60**, 639 (1988)
- [50] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [51] W. Metzner, Z. Phys. B **77**, 253 (1989)
- [52] F. Gebhard, Phys. Rev. B **41**, 9452 (1990)
- [53] D. Vollhardt, P.G.J. van Dongen, F. Gebhard, and W. Metzner,
Mod. Phys. Lett. B **4**, 499 (1990)
- [54] J. Bünenmann, F. Gebhard, and W. Weber, Found. of Physics **30**, 2011 (2000)
- [55] J. Bünenmann, Chapt. 5 in [132]
- [56] P.G.J. van Dongen, F. Gebhard, and D. Vollhardt, Z. Phys. **76**, 199 (1989)
- [57] U. Wolff, Nucl. Phys. B **225**, 391 (1983)
- [58] W. Metzner and D. Vollhardt, Phys. Rev. B **39**, 4462 (1989)
- [59] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
- [60] E. Müller-Hartmann, Z. Phys. B **76**, 211 (1989)
- [61] H. Schweitzer and G. Czycholl, Solid State Comm. **69**, 171 (1989)
- [62] W. Metzner, P. Schmit, and D. Vollhardt, Phys. Rev. B **45**, 2237 (1992)
- [63] R. Strack and D. Vollhardt, Phys. Rev. B **46**, 13852 (1992)
- [64] D. Vollhardt, p. 57 in *Correlated Electron Systems*, ed. by V.J. Emery
(World Scientific, Singapore, 1993); http://www.physik.uni-augsburg.de/theo3/Research/research_jerusalem.vollha.en.shtml

- [65] U. Brandt and C. Mielsch, Z. Phys. B **75**, 365 (1989)
- [66] P.G.J. van Dongen and D. Vollhardt, Phys. Rev. Lett. **65**, 1663 (1990)
- [67] J.M. Luttinger and J.C. Ward, Phys. Rev. **118**, 1417 (1960)
- [68] V. Janiš, Z. Phys. B **83**, 227 (1991)
- [69] R. Vlaming and D. Vollhardt, Phys. Rev. B, **45**, 4637 (1992)
- [70] V. Janiš and D. Vollhardt, Int. J. Mod. Phys. B **6**, 731 (1992)
- [71] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [72] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
- [73] J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)
- [74] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [75] G. Kotliar and D. Vollhardt, Physics Today, March 2004, p. 53
- [76] V. Janiš and D. Vollhardt, Phys. Rev. B **46**, 15712 (1992)
- [77] X.Y. Zhang, M.J. Rozenberg, and G. Kotliar, Phys. Rev. Lett. **70**, 1666 (1993)
- [78] J.K. Freericks and V. Zlatić, Rev. Mod. Phys. **75**, 1333 (2003)
- [79] M.J. Rozenberg, X.Y. Zhang, and G. Kotliar, Phys. Rev. Lett. **69**, 1236 (1992)
- [80] A. Georges and W. Krauth, Phys. Rev. Lett. **69**, 1240 (1992)
- [81] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, P. Werner, Rev. Mod. Phys. **83**, 349 (2011)
- [82] R. Bulla, Phys. Rev. Lett. **83**, 136 (1999)
- [83] W. Hofstetter, Phys. Rev. Lett. **85**, 1508 (2000)
- [84] R. Bulla, T.A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008)
- [85] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
- [86] Q. Si, M.J. Rozenberg, G. Kotliar, and A.E. Ruckenstein, Phys. Rev. Lett. **72**, 2761 (1994)
- [87] M.J. Rozenberg, G. Moeller, and G. Kotliar, Mod. Phys. Lett. B **8**, 535 (1994)
- [88] T. Pruschke, M. Jarrell, and J.K. Freericks, Adv. Phys. **44**, 187 (1995)

- [89] A. Georges, in *Lectures on the Physics of highly correlated electron systems VIII*, ed. by A. Avella and F. Mancini, AIP Conference Proceedings, Vol. 715 (American Institute of Physics, Melville, 2004), p. 3
- [90] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
- [91] K. Byczuk, W. Hofstetter, and D. Vollhardt, in *Fifty Years of Anderson Localization*, ed. by E. Abrahams (World Scientific, Singapore, 2010), p. 473; reprinted in Int. J. Mod. Phys. B **24**, 1727 (2010)
- [92] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964)
- [93] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965)
- [94] R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
- [95] V.I. Anisimov, J. Zaanen, and O.K. Andersen, Phys. Rev. B **44**, 943 (1991)
- [96] V.I. Anisimov, F. Aryasetiawan, and A.I. Lichtenstein, J. Phys.: Condens. Matter **9**, 767 (1997)
- [97] V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin, and G. Kotliar, J. Phys.: Condens. Matter **9**, 7359 (1997)
- [98] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
- [99] I.A. Nekrasov, K. Held, N. Blümer, A.I. Poteryaev, V.I. Anisimov, and D. Vollhardt, Eur. Phys. J. B **18**, 55 (2000)
- [100] K. Held, I.A. Nekrasov, N. Blümer, V.I. Anisimov, and D. Vollhardt, Int. J. Mod. Phys. B **15**, 2611 (2001)
- [101] A.I. Lichtenstein, M.I. Katsnelson, and G. Kotliar, p. 428 in *Electron Correlations and Materials Properties*, ed. by A. Gonis, N. Kioussis, and M. Ciftan (Plenum, NY, 2002)
- [102] K. Held, I.A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A.K. McMahan, R.T. Scalettar, T. Pruschke, V.I. Anisimov, and D. Vollhardt, Psi-k Newsletter **56**, 65 (2003); reprinted in Phys. Status Solidi B **243**, 2599 (2006)
- [103] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
- [104] K. Held, Adv. Phys. **56**, 829 (2007)
- [105] M.I. Katsnelson, V.Yu. Irkhin, L. Chioncel, A.I. Lichtenstein, and R.A. de Groot, Rev. Mod. Phys. **80**, 315 (2008)

- [106] J. Kuneš, I. Leonov, M. Kollar, K. Byczuk, V.I. Anisimov, and D. Vollhardt, Eur. Phys. J. Special Topics **180**, 5 (2010)
- [107] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003)
- [108] J. Minár, L. Chioncel, A. Perlov, H. Ebert, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **72**, 45125 (2005)
- [109] C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, Phys. Rev. Lett. **112**, 196402 (2014)
- [110] M. Potthoff and W. Nolting, Phys. Rev. B **59**, 2549 (1999)
- [111] J.K. Freericks: *Transport in multilayered nanostructures: The dynamical mean-field approach* (Imperial College Press, London, 2006)
- [112] M. Takizawa, H. Wadati, K. Tanaka, M. Hashimoto, T. Yoshida, A. Fujimori, A. Chikamatsu, H. Kumigashira, M. Oshima, K. Shibuya, T. Mihara, T. Ohnishi, M. Lippmaa, M. Kawasaki, H. Koinuma, S. Okamoto, and A.J. Millis, Phys. Rev. Lett. **97**, 057601 (2006)
- [113] L. Chen and J.K. Freericks, Phys. Rev. B **75**, 1251141 (2007)
- [114] R.W. Helmes, T.A. Costi, and A. Rosch, Phys. Rev. Lett. **100**, 056403 (2008)
- [115] L. Chioncel, I. Leonov, H. Allmaier, F. Beiuseanu, E. Arrigoni, T. Jurcut, and W. Pötz, Phys. Rev. B **83**, 035307 (2011)
- [116] F. Lechermann, L. Boehnke, and D. Grieger, Phys. Rev. B **87**, 241101(R) (2013)
- [117] D. Jaksch, C. Bruder, J.I. Cirac, C.W. Gardiner, and P. Zoller, Phys. Rev. Lett. **81**, 3108 (1998)
- [118] M. Greiner, O. Mandel, T. Esslinger, T.W. Hänsch, and I. Bloch, Nature **415**, 39 (2002)
- [119] A. Rapp, G. Zarand, C. Honerkamp, W. Hofstetter, Phys. Rev. Lett. **98**, 160405 (2007)
- [120] U. Schneider, L. Hackermüller, S. Will, T. Best, I. Bloch, T.A. Costi, R.W. Helmes, D. Rasch, and A. Rosch, Science **322**, 1520 (2008)
- [121] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008)
- [122] V. Turkowski and J.K. Freericks, Phys. Rev. B **71**, 085104 (2005)
- [123] J.K. Freericks, Phys. Rev. B **77**, 075109 (2008)
- [124] N. Tsuji, T. Oka, and H. Aoki, Phys. Rev. B **78**, 235124 (2008)
- [125] M. Eckstein, and M. Kollar, Phys. Rev. B **78**, 205119 (2008)

- [126] J.K. Freericks, H.R. Krishnamurthy, and T. Pruschke,
Phys. Rev. Lett. **102**, 136401 (2009)
- [127] M. Eckstein, M. Kollar, and P. Werner, Phys. Rev. Lett. **103**, 056403 (2009)
- [128] M. Eckstein and P. Werner, Phys. Rev. B **82**, 115115 (2010)
- [129] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner,
Rev. Mod. Phys. **86**, 779 (2014)
- [130] See the homepage of the DFG Research Unit FOR 1346 *Dynamical Mean-Field Approach with Predictive Power for Strongly Correlated Materials*:
<http://www.physik.uni-augsburg.de/for1346/>
- [131] E. Pavarini, E. Koch, D. Vollhardt, and A.I. Lichtenstein (eds.):
The LDA+DMFT approach to strongly correlated materials
Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
<http://www.cond-mat.de/events/correl11>
- [132] E. Pavarini, E. Koch, F. Anders, and M. Jarrell (eds.):
Correlated Electrons: From Models to Materials
Reihe Modeling and Simulation, Vol. 2 (Forschungszentrum Jülich, 2012)
<http://www.cond-mat.de/events/correl12>
- [133] E. Pavarini, E. Koch, and U. Schollwöck (eds.):
Emergent Phenomena in Correlated Matter
Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
<http://www.cond-mat.de/events/correl13>
- [134] E. Pavarini, E. Koch, D. Vollhardt, and A.I. Lichtenstein (eds.):
DMFT at 25: Infinite Dimensions
Reihe Modeling and Simulation, Vol. 4 (Forschungszentrum Jülich, 2014)
<http://www.cond-mat.de/events/correl14>

2 Electronic Structure of Correlated Materials: Slave-Boson Methods and Dynamical Mean-Field Theory

Gabriel Kotliar
Rutgers University
Piscataway, NJ, USA

Contents

| | |
|---|-----------|
| 1 First-principles approaches and model Hamiltonians | 2 |
| 2 Slave-boson methods and emergence of local Fermi-liquids | 6 |
| 3 DMFT for model Hamiltonians: Embedding and truncation | 8 |
| 4 Correlations in the solid state, LDA, hybrids, LDA+DMFT | 10 |
| 5 Electronic structure methods from a diagrammatic many-body perspective | 13 |
| 6 Discretization of basis-sets and Coulomb integrals | 15 |
| 7 Bridging between first-principles and model Hamiltonian approaches | 18 |
| 8 Applications: Iron pnictides and Hund's metals | 20 |
| 9 Applications: Actinides | 23 |
| 10 Summary and outlook | 24 |

1 First-principles approaches and model Hamiltonians

The standard model of solid state physics, described in solid state textbooks, has been extraordinarily successful in describing the properties of simple metals, semiconductors, and insulators. It is firmly grounded in the Landau Fermi Liquid Theory and perturbative expansions around the, by now standard, implementations of the density-functional theory (DFT), such as the Local-Density Approximation (LDA) or Generalized-Gradient Approximations (GGA) by means of the GW method.

Strongly correlated electron systems are materials that fall outside the standard model. They display remarkable phenomena ranging from high-temperature superconductivity in iron pnictides and copper-oxides, huge volume collapses in the 4f and 5f elemental series, to metal-to-insulator transitions in many transition-metal oxides such as V_2O_3 and VO_2 , to name a few.

From a theoretical perspective, these systems display remarkable emergent phenomena that cannot be accessed by perturbation theory starting from the band limit. Strong correlation phenomena require a different reference frame for their description and a methodology that is quite different from what is learned in traditional solid-state or many-body physics courses. Forty years ago, the theoretical toolbox to treat strong correlations was very limited. The focus was on variational wave functions of the type written by Gutzwiller [1] as used in the mixed-valence problem by Varma and Yaffet [2] and on the decoupling of equations of motion used by Hubbard [3]. Development of renormalization group methods for simple condensed matter physics problems was just beginning [4]. Methods for treating the unusual excitation spectra and the finite-temperature properties of strongly correlated materials were badly needed. At that time, we could not even contemplate a realistic treatment nor even a system-specific study of actual strongly correlated materials. The situation is completely different today, and the methods that brought about this change are the subject of these introductory notes, compiled by Wenhui Xu from lectures delivered by the author. They are intended as an orientation for beginning students in the field of electronic-structure calculations of strongly correlated materials. Their goal is to motivate students to enter the field by highlighting a couple of research achievements, rather than provide a complete overview with a complete list of references which can be found in the excellent collection of reviews in *Reviews of Modern Physics* [5–7]. The focus is on methods that target not only total energies, but finite-temperature properties and, most important, correlation functions.

Modern electronic-structure methods that treat correlated materials have developed into theoretical spectroscopies. This allows detailed comparison with experiments, which in turn catalyzes further theoretical progress. This iterative feedback loop is one of the characteristic strengths of condensed matter physics. Hence, some comparisons to experimental results are included in this lecture. While they convey some sense of collective achievement, they should also be a reminder that the theory of strongly correlated electron systems is still in its infancy. The goal is to highlight, in broad strokes, some advances that have taken place while indicating some problems that remain to be tackled to pave the way for a predictive theory of strongly correlated materials.

Historically, there have been two approaches to understanding and describing the physical properties of strongly correlated materials. First-principles (also called *ab-initio*) methods begin from the full Hamiltonian of electrons in the solid. This *theory of everything* is given by

$$H = \sum_i \frac{\nabla_i^2}{2m_e} + \sum_{\alpha} \frac{\nabla_{\alpha}^2}{2m_{\alpha}} - \sum_{\alpha,i} \frac{Z_{\alpha}e^2}{|\vec{R}_{\alpha} - \vec{R}_i|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{R}_i - \vec{R}_j|} + \frac{1}{2} \sum_{\alpha \neq \beta} \frac{Z_{\alpha}^2 e^2}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|}$$

+ relativistic effects. (1)

Here i and j are indices of electrons; α and β are indices of nuclei. Relativistic effects include spin-orbit coupling and are actually very important and give rise to qualitatively new physics in strongly correlated materials. One term, $\sum_i \frac{\vec{l}_i \cdot \vec{s}_i}{R_i^3}$, is essential to have non zero magnetocrystalline anisotropy, which selects the magnetization axis in crystals. We will not consider these terms in the lectures. We treat ions as very heavy objects (*adiabatic approximation*). In this limit $\sum_{\alpha,\beta} \frac{Z_{\alpha}Z_{\beta}e^2}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|^2}$ becomes a number and $\sum_{\alpha,i} \frac{Z_{\alpha}e^2}{|\vec{R}_{\alpha} - \vec{R}_i|^2} \equiv \sum_i V_{\text{crystal}}(R_i)$ becomes an external potential for the electrons. Fluctuations around the equilibrium positions give rise to the lattice vibrations (phonons). Therefore, with these approximations Z_{α} and \vec{R}_{α} are the only input parameters, and approaches starting from Eq. (1) are referred to as first principles methods.

The standard model of solid state physics is grounded on two firm pillars. The first is the Fermi liquid theory, which justifies the use of free electrons as a reference system to describe the properties of an interacting Fermi system. In its renormalization group formulation [8], this can be understood by the statement that, in many instances, the interactions flow rapidly to zero as one approaches the Fermi surface. Then, below a certain scale, materials behave as non-interacting electrons, since the interactions have renormalized away. The only interactions that remain are Hartree-like terms that renormalize the responses to external fields (Landau Parameters). When this Fermi liquid scale is much larger than the temperatures of interest, the textbook picture of free fermions, in the presence of a periodic Bloch potential with renormalized parameters is thus justified.

The second pillar of the standard model enables the actual calculation of the quasiparticle dispersions and Fermi liquid parameters. It starts with the Kohn-Sham formulation [9] of density functional theory. It states the existence of a potential $V_{\text{KS}}(\vec{r})$, which is itself a functional of the density. One should write $V_{\text{KS}}(\vec{r})[\{\rho(\vec{r}')\}]$ to indicate this dependence, but we omit this in the following. The exact (but unknown) functional is such that the solution of the set of self-consistent equations,

$$[-\nabla^2 + V_{\text{KS}}(\vec{r})] \psi_{\vec{k}j}(\vec{r}) = \varepsilon_{\vec{k}j} \psi_{\vec{k}j}(\vec{r}). \quad (2)$$

$$\sum_{\vec{k}j} |\phi_{\vec{k}j}(\vec{r})|^2 f(\varepsilon_{\vec{k}j}) = \rho(\vec{r}) \quad (3)$$

reproduces the density of the solid. It is useful to divide the Kohn-Sham potential into several parts: $V_{\text{KS}} = V_{\text{Hartree}} + V_{\text{cryst}} + V_{xc}$, where one lumps into V_{xc} exchange and correlation effects beyond Hartree.

The eigenvalues $\varepsilon_{\vec{k}j}$ of the solution of the self-consistent set of Eq. (2) and (3) are not to be interpreted as excitation energies. The excitation spectra should be extracted from the poles of the one particle Green's function

$$G(\omega) = \frac{1}{[\omega + \nabla^2 + \mu - V_{\text{Hartree}} - V_{\text{cryst}}] - \Sigma(\omega)}. \quad (4)$$

Here μ is the chemical potential and we have singled out in Eq. (4) the Hartree potential expressed in terms of the exact density and the crystal potential, and lumped the rest of the effects of the correlation in the self-energy operator, which depends on frequency as well as on two space variables.

In a weakly correlated material, the one-particle excitation spectrum is perturbatively connected to the LDA Kohn-Sham spectrum, in the sense that the first-order correction in the screened Coulomb interactions for the self-energy Σ_{GW} (see the diagrams in figure 1) is such that $\Sigma = \Sigma_{GW} - V_{xc}$ is relatively small and able to bring the spectra sufficiently close to the experimental results. In fact, we can *define* weakly correlated materials as those solids for which the previous statement is true. Lowest order perturbation theory in the screened Coulomb interactions is called the *GW* method [10]. It has been very successful in predicting the trends of the gaps in semiconducting materials [11]

The *GW* method involves several steps, summarized in the diagrams shown in Fig. 1.

1. Computation of the polarization bubble

$$\Pi(t, t') = G_0(t, t') G_0(t', t). \quad (5)$$

2. Evaluation of the screened Coulomb potential W in random-phase approximation (RPA)

$$W^{-1} = v_{Coul}^{-1} - \Pi. \quad (6)$$

where v_{Coul} is the bare Coulomb potential.

3. Evaluation of the Σ_{GW} contribution to self-energy by lowest-order perturbation theory in W ; it is given in real space by (see Fig. 1)

$$\Sigma_{GW} = G_0 W. \quad (7)$$

4. From the self-energy one obtains the full Green's function using the Dyson equation where one removes the V_{xc} term from G_0 and adds the *GW* contribution to the self-energy to obtain an approximation to Eq. (4)

$$G^{-1} = G_0^{-1} - \Sigma. \quad (8)$$

We have not yet specified what one should take for G_0 in this algorithm. Various ideas have been discussed and implemented, leading to different variants of the *GW* method. In the “one-shot” *GW* method one uses the LDA Kohn-Sham Green's function

$$G_0(i\omega)^{-1} = i\omega + \mu + \nabla^2 - V_{\text{Hartree}} - V_{\text{cryst}} - V_{xc}^{LDA}. \quad (9)$$

and the self-energy is thus taken to be $\Sigma = \Sigma_{GW} - V_{xc}^{LDA}$.

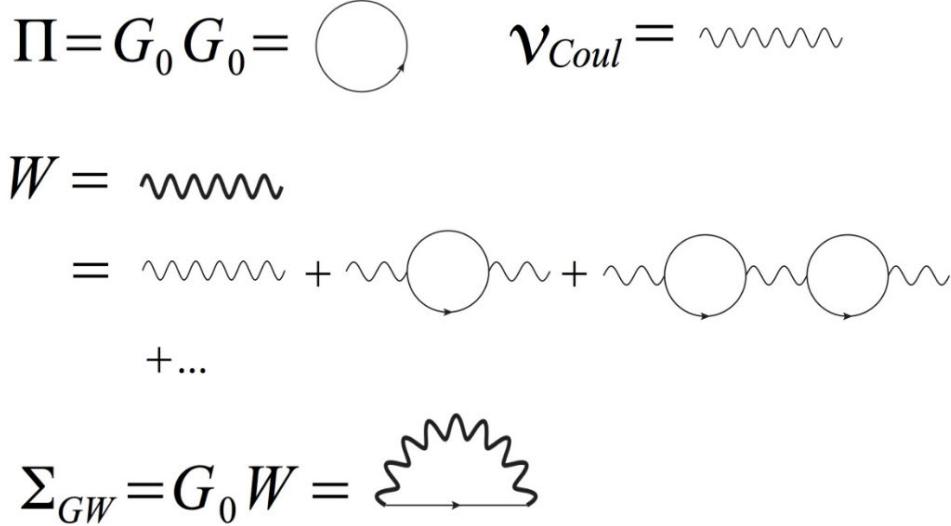


Fig. 1: Schematic diagrams for the *GW* method. Starting from some G_0 , a polarization bubble is constructed, which is used to screen the Coulomb interactions resulting in an interaction W . This W is then used to compute a self-energy Σ_{GW} using W and G_0 . To obtain the full Green's function G in Eq. (4), one goes from Σ_{GW} to Σ by subtracting the necessary single-particle potential and uses the Dyson equation $G^{-1} = G_0^{-1} - \Sigma$ as discussed in the text.

In the original self-consistent scheme proposed by Hedin [10] (the self-consistent *GW*) $G_0 = G$ is used and in this case $V_{xc} = 0$ is not needed and is not used in intermediate steps. There are numerous advantages, however, in using a non-interacting form for G_0 in the algorithm. In the quasi-particle self-consistent *GW* (QPGW) [11] the “best” non-interacting Green’s function is used for G_0 which uses an “exchange and correlation potential” V_{xc}^{QPGW} chosen to reproduce the same quasiparticle spectra as the full *GW* greens function.

$$G_0(i\omega)^{-1} = i\omega + \mu + \nabla^2 - V_{\text{Hartree}} - V_{\text{cryst}} - V_{xc}^{QPGW}. \quad (10)$$

The *GW* or RPA method captures an important physical effect. Electrons are charged objects that interact via the long range Coulomb interactions. Quasiparticles, on the other hand, are neutral. They are composed of electrons surrounded by screening charges, thus reducing the strength and the range of their interaction. For this reason, in many model Hamiltonians describing metals, only the short range repulsion is kept. To get a feeling for the screening effect, let’s evaluate Eqs. (5) and (6) for an effective interaction W in the case that there is only one band of electrons with dispersion ε_k

$$\begin{aligned} \Pi(i\Omega = 0, \vec{q}) &= T \sum_{\omega} \sum_{\vec{k}} \frac{1}{i\omega - \varepsilon_{\vec{k}+\vec{q}}} \frac{1}{i\omega - \varepsilon_{\vec{k}}} = \sum_{\vec{k}} \frac{f(\varepsilon_{\vec{k}+\vec{q}}) - f(\varepsilon_{\vec{k}})}{\varepsilon_{\vec{k}+\vec{q}} - \varepsilon_{\vec{k}}} \\ &\simeq - \sum_{\vec{k}} \left(\frac{\partial f(\varepsilon)}{\partial \varepsilon} \right)_{\varepsilon_{\vec{k}}} \simeq \sum_{\vec{k}} \delta(\varepsilon_{\vec{k}}) = \rho_{DOS}, \end{aligned} \quad (11)$$

where ρ_{DOS} is the density of states at the Fermi surface.

In momentum space, $v_{Coul}(\vec{q}) = 4\pi e^2/q^2$. Then

$$W(\vec{q}) = \frac{v_{Coul}(\vec{q})}{1 + v_{Coul}(\vec{q})\rho_{DOS}} \simeq \frac{1}{q^2 + 4\pi e^2\rho_{DOS}} \quad (12)$$

and its Fourier transform $W(\vec{r})$, which is now a function of one variable due to translation invariance, decays exponentially in space.

Model Hamiltonians are simplified Hamiltonians describing a reduced set of degrees of freedom, and involve a number of parameters. They are extremely useful for learning the qualitative physics exhibited by strongly correlated materials. Conceptually, we can obtain model Hamiltonians by selecting low-energy degrees of freedom (usually a few bands) and describing their interactions, which in metals at low energies are short-ranged due to the screening mechanism. The most famous example is the multi-orbital Hubbard model.

$$H = \sum_{i,j} c_\alpha^\dagger(i) t_{ij}^{\alpha\beta} c_\beta(j) + \sum_i U_{\alpha\beta\gamma\delta} c_\alpha^\dagger(i) c_\beta^\dagger(i) c_\gamma(i) c_\delta(i). \quad (13)$$

Even simplified model Hamiltonians have proved to be very difficult to solve exactly in the thermodynamical limit except for the cases of one dimension [12] and the limit of infinite dimensions [13], a limit where Dynamical Mean-Field Theory (DMFT) becomes exact.

Another celebrated model, the Anderson Impurity Model, was introduced by Anderson in the sixties to describe transition-metal impurities in metallic hosts [14]. It will play an important role in the DMFT analysis of the Hubbard model in Sec. 3.

2 Slave-boson methods and emergence of local Fermi-liquids

The spectra of strongly correlated electron materials are very far from those of free fermions. The one electron spectral function $A(\vec{k}, \omega)$ displays not only a dispersive quasiparticle peak but also other features commonly denoted as satellites. The collective excitation spectra, which appear in the spin and charge excitation spectra, do not resemble the particle-hole continuum of the free Fermi gas, with additional collective modes (zero sound, spin waves) produced by the residual interactions among them. Finally, the damping of the elementary excitations in many regimes does not resemble that of a Fermi liquid.

The key idea of the slave-boson method is to enlarge the Hilbert space so as to be able to more explicitly introduce operators that closely describe the physical excitations.

This is done by reformulation of the Hamiltonian in terms of additional slave variables, with additional Lagrange multipliers that impose constraints. We illustrate this idea with the multi-orbital Hubbard model following Refs. [15] and [16]. We first focus on one site and on the local interaction term

$$H_{loc} = \sum_\alpha \varepsilon_\alpha^0 \hat{n}_\alpha + \sum_{\alpha\beta} U_{\alpha\beta} \hat{n}_\alpha \hat{n}_\beta \quad (14)$$

acting on a Hilbert space

$$|n\rangle = \left(d_1^\dagger\right)^{n_1} \cdots \left(d_M^\dagger\right)^{n_M} |vac\rangle. \quad (15)$$

We notice that Eq. (14) is equivalent to another Hamiltonian which acts on a larger Hilbert space on which we will impose some constraints to retrieve the original problem

$$|\underline{n}\rangle \equiv \phi_n^\dagger |vac\rangle \otimes |n\rangle_f, \quad (16)$$

$$|n\rangle_f \equiv \left(f_1^\dagger\right)^{n_1} \cdots \left(f_M^\dagger\right)^{n_M} |vac\rangle. \quad (17)$$

The states in the original Hilbert space, denoted by a bar, are in one-to-one correspondence with the states of the enlarged Hilbert space once the constraints

$$\sum_n \phi_n^\dagger \phi_n = 1, \quad (18)$$

$$\sum_n n_\alpha \phi_n^\dagger \phi_n = f_\alpha^\dagger f_\alpha, \quad \forall \alpha \quad (19)$$

are imposed. In the enlarged Hilbert space, the physical electron is described by

$$\underline{d}_\alpha^\dagger = R_\alpha[\phi] f_\alpha^\dagger, \quad (20)$$

where

$$R_\alpha[\phi] = \sum_{nm} \langle n | f_\alpha^\dagger | m \rangle \left[\hat{\Delta}_\alpha \right]^{-1/2} \phi_n^\dagger \phi_m \left[1 - \hat{\Delta}_\alpha \right]^{-1/2} \quad (21)$$

with

$$\hat{\Delta}_\alpha [\phi] \equiv \sum_n n_\alpha \phi_n^\dagger \phi_n. \quad (22)$$

The kinetic energy is then

$$\underline{H} = \sum_{ij} R_\alpha[\phi] f_\alpha^\dagger(i) t_{ij}^{\alpha\beta} R_\beta[\phi] f_\beta(j). \quad (23)$$

while the local energy and interaction terms in the enlarged Hilbert space are reproduced by a quadratic Hamiltonian

$$H_{loc} = \sum_n \phi_n^\dagger \phi_n \varepsilon_n \quad (24)$$

where $\varepsilon_n = \sum_\alpha (n_\alpha + \sum_\beta U_{\alpha\beta} n_\alpha n_\beta)$.

The fact that the Hamiltonian is now quadratic in bosons and fermions suggests simple approximations for its treatment. The square root factors are largely arbitrary, in the sense that they only affect the degrees of freedom outside the physical Hilbert space, and they were chosen so as to give the same results as the Gutzwiller approximation and have a simple probabilistic interpretation.

The self-energy of the Green's function that results from the mean-field approximation (replacing ϕ 's and Lagrange multipliers λ by numbers) has the form

$$\Sigma_\alpha(\omega) = \Sigma_\alpha(0) + \omega \left(1 - \frac{1}{Z_\alpha} \right), \quad (25)$$

where

$$Z_\alpha = |R_\alpha|^2, \quad (26)$$

and

$$\Sigma_\alpha(0) = \lambda_\alpha/|r_\alpha|^2 - \varepsilon_\alpha^0. \quad (27)$$

Hence this theory describes the emergence of a local Fermi liquid. Non-local self-energies can be obtained with the significant extension introduced in Ref. [17].

This formulation explicitly exhibits the local collective modes (local charge, spin, and orbital fluctuations) in terms of the slave-boson operators. It has been extended [17–19] to make it manifestly rotationally invariant. For example, in the one-band Hubbard model in the simple slave-boson formulation, the state with one spin is described by the slave-boson ϕ_σ and transforms according to the fundamental representation of $SU(2)$. The spin fluctuations are naturally described by objects that transform according to the adjoint representation of $SU(2)$, which requires a matrix representation of the slave particles. Another advantage of the rotationally invariant formulation [17] is that it allows the treatment of realistic Hamiltonians including general multiplet interactions.

The physical electron operator is now represented in the enlarged Hilbert space by

$$\underline{d}_\alpha = \hat{R}_{\alpha\beta}[\phi] f_\beta. \quad (28)$$

At the mean-field level, \hat{R} has the interpretation of the quasiparticle residue, exhibiting the strong renormalizations induced by the electronic correlations. An important feature of the rotationally invariant formalism is that the basis that diagonalizes the quasiparticles represented by the operators f is not necessarily the same basis as that which would diagonalize the one electron density matrix expressed in terms of the operators d and d^\dagger . Strongly renormalized fermionic quasiparticles emerge in this treatment. This slave-boson formulation [15], reproduces at the saddle point level the results of the Guzwiller approximation. Fluctuations around the saddle point generate the Hubbard bands in the one-particle spectra [20]. This method can be applied to the Anderson impurity model. When supplemented by the DMFT self-consistency condition, it gives the same results as its direct application to the lattice [21]. We envision many synergistic applications of the slave-boson technique and exact implementations of the DMFT, and we will return to this perspective at the end of Sec. 9.

3 DMFT for model Hamiltonians: Embedding and truncation

Dynamical mean field theory [22] is the natural extension of the Weiss mean-field theory of spin systems to treat quantum mechanical model Hamiltonians. It involves two steps. The first step focuses on a single lattice site and describes the rest of the sites by a medium with which an electron at this site hybridizes. This *truncation* to a single site problem is common to all mean-field theories. In the Weiss mean-field theory one selects a spin at a given site and replaces the rest of the lattice by an effective magnetic field or Weiss field. In the dynamical mean-field theory, the local Hamiltonian at a given site is kept, and the kinetic energy is replaced

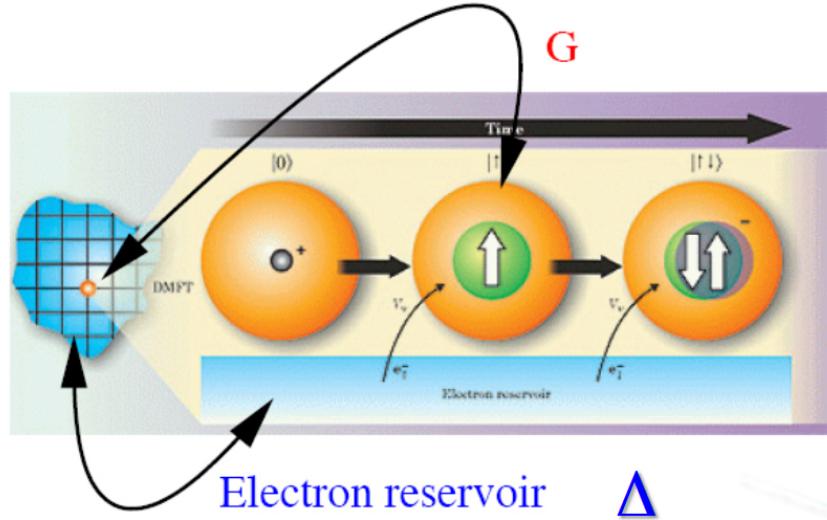


Fig. 2: *Dynamical Mean-Field Theory (DMFT) maps (or truncates) a lattice model to a single site embedded in a medium (impurity model) with a hybridization strength that is determined self-consistently. Adapted from Ref. [23]*

| Weiss mean field theory | dynamical mean-field theory |
|--|--|
| Ising model → single spin in an effective Weiss field | Hubbard-type model → impurity in an effective bath |
| Weiss field: h_{eff} | effective bath: $\Delta(i\omega_n)$ |
| local observable: $m = \langle s_i \rangle$ | local observable: $G_{loc}(i\omega_n)$ |
| self-consistent condition: $\tanh(\beta \sum_j J_{ij} s_j) = m$ | self-consistent condition: $i\omega_n - E_{imp} - \Delta(i\omega_n) - \Sigma(i\omega_n) = [\sum_{\vec{k}} G_{\vec{k}}(i\omega_n)]^{-1}$ |

Table 1: Corresponding quantities in dynamical mean-field theory (right) and Weiss or static mean-field theory in statistical mechanics (left).

by a hybridization term with a bath of non-interacting electrons, which allows the atom at the selected site to change its configuration. This is depicted in Fig. 2.

The second step involves the reconstruction of lattice observables by *embedding* the local impurity self-energy into a correlation function of the lattice. $G_{latt}(\vec{k}, i\omega)^{-1} = i\omega + \mu - t_{\vec{k}} - \Sigma_{imp}(i\omega)$. Here $\Sigma_{imp}(i\omega)$ are viewed as functionals of the Weiss field. The requirement $\sum_k G_{latt} = G_{loc}$ determines the Weiss field. Table 1 summarizes the analogies between Weiss mean-field theory and dynamical mean-field theory.

The DMFT mapping of a lattice model onto an impurity model gives a local picture of the solid in terms of an impurity model, which then can be used to generate lattice quantities, such as the Green's function of electrons and the magnetic susceptibility, by computing the corresponding irreducible quantities. This is illustrated in Fig. 3.

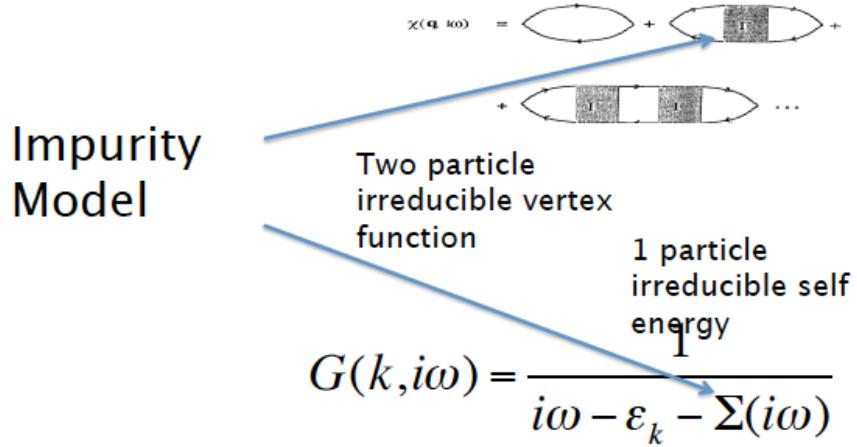
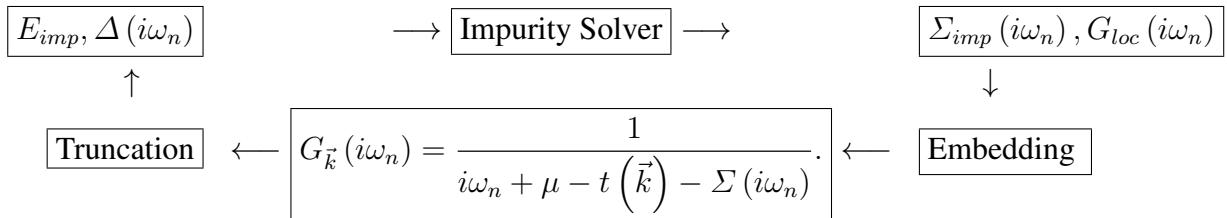


Fig. 3: The DMFT impurity model is used to generate irreducible quantities such as self-energies and two-particle vertices. These are then embedded in the lattice model to generate momentum dependent lattice quantities such as spectral functions or spin susceptibilities.

The self-consistent loop of DMFT is summarized in the following iterative cycle



The impurity model is the engine of a DMFT calculation. Multiple approaches have been used for its solution, and full reviews are needed to do this topic justice. Recent advances in the continuous-time quantum Monte Carlo method for impurity models [24] have provided numerically exact solutions at relatively low computational cost. Dynamical mean-field theory becomes exact in the limit of infinite dimensions that was introduced by Metzner and Vollhardt [13]. With suitable extensions, it plays an important role in realistically describing strongly correlated materials. This is the subject of the following sections.

4 Correlations in the solid state, LDA, hybrids, LDA+DMFT

In the context of the simple Hubbard model, a clear measure of the strength of the electronic correlations is the ratio U/t . But how do we quantify correlation-strength in an actual solid described by Eq. (1), which has no reference to a U or a t ?

To address this question we need to start from the exact one-particle Green's function in the solid that we introduced in Eq. (4) and focus on the self-energy Σ which should be viewed as an infinite-dimensional matrix in a specified basis set.

For a chemist, correlations mean large departures of Σ with respect to the Fock self-energy. Hence, a strongly correlated system is a system where $\Sigma - \Sigma_{Fock}$ is large. From this perspective,

even simple metals are strongly correlated since pure exchange is a poor approximation to the self-energy of even simple metals.

We adopt a different definition, the one used by physicists, and measure the strength of the correlation by the departure of the self-energy from the exchange-correlation potential of the LDA. A correlated material is one where

$$\Sigma(\omega) - V_{xc(LDA)} \quad (29)$$

is large in some *low* frequency range. Notice that at infinite frequencies Σ is given by just the Fock diagram and therefore at large frequencies the difference in Eq. (29) is large, but this usually occurs above the plasma frequency, a fairly large energy scale.

Sometimes the difference in Eq. (29) is local and restricted to a few orbitals, as will be explained in the following. In this case we can describe this difference using DMFT. This is the basis of the DFT+DMFT methodology, to be described below. Notice, however, that this methodology should be used as a description of the spectra below the plasma frequency.

Introducing a complete basis set of localized wavefunctions labeled by site and orbital index, we can expand the self-energy as

$$\Sigma(\vec{r}, \vec{r}', \omega) = \sum_{\alpha \vec{R}, \beta \vec{R}'} \chi_{\alpha \vec{R}}^*(\vec{r}) \Sigma(\omega)_{\alpha \vec{R}, \beta \vec{R}'} \chi_{\beta \vec{R}'}(\vec{r}') . \quad (30)$$

Eq. (30) allows us to introduce an approximate or simplified representation of the self-energy [25] involving a sum of a non-local but frequency independent term plus a frequency-dependent but local self-energy.

$$\Sigma(\vec{k}, \omega) \simeq \Sigma(\vec{k}) + \sum_{\vec{R}, \alpha, \beta \in L} |\vec{R}\alpha\rangle \Sigma(\omega)_{loc, \vec{R}\vec{R}} \langle \vec{R}\beta| . \quad (31)$$

Notice that the notion of locality is defined with reference to a basis set of orbitals. The self-energy is approximately local when the on-site term $\vec{R} = \vec{R}'$ in Eq. (30) is much larger than the rest, and the ansatz is useful when the sum over orbitals in Eq. (31) runs over a small set L (much smaller than the size of the basis set), for example over a single shell of d - or f -orbitals. The validity of the local ansatz for graphs beyond the GW approximation was tested for transition metals in an LMTO basis set by N. Zein *et al.* [26].

For semiconductors, non-local (but frequency-independent) correlation effects are needed to increase the gap from its LDA value. This admixture of exchange can be done within the GW method or using hybrid density functionals. It reflects the importance of exchange beyond the LDA, which is due to the long-range but static part of the Coulomb interaction. It has recently been shown that this type of correlation effect is important in materials near a metal-to-insulator transition such as BaBiO_3 or HfClN [27]. In these systems, $\Sigma(\vec{k})$ is much more important than the frequency dependence in the self-energy.

Frequency dependence implies non-locality in time and is important in materials governed by Mott or Hund's physics. This physics tends to be local in space and can be captured by DMFT.

Static mean-field theories such as the LDA do not capture this non-locality in time, and therefore fail to describe Mott or Hund's phenomena.

In the quantum chemistry jargon the frequency-independent self-energy is ascribed to *dynamical* correlation effects, while the frequency-dependent self-energy is ascribed to *static* correlation effects. This difference in terminologies among two communities that are describing similar effects has been a continuous source of confusion. In real materials, both effects are present to some degree thus motivating physically the ansatz of Eq. (31). Some examples discussed recently are CeO₃ (using hybrid DFT+DMFT) in Ref. [28] and the iron pnictides and chalcogenides in Ref. [25].

This discussion motivates the DFT+DMFT method, which was introduced in Ref. [29] (see also Ref. [30]). DFT here stands for density-functional theory, and refers to the standard practical implementations of this theory, such as LDA or GGA, which are used with similar frequency. However DFT could be replaced by another static mean-field theory like hybrid DFT or QPGW. In the following we will use the terminology LDA+DMFT.

Starting from the model Hamiltonian point of view, one divides the orbitals into two sets, the first set containing the large majority of the electrons, which are properly described by the LDA Kohn-Sham matrix. The second set contains the more localized orbitals (*d*-electrons in transition metals and *f*-electrons in rare earths and actinides), which require the addition of DMFT corrections. A subtraction (called the double-counting correction) takes into account that the Hartree and exchange correlation has been included in that orbital twice since it was treated both in LDA and in DMFT. The early LDA+DMFT calculations proceeded in two steps (one-shot LDA+DMFT). First an LDA calculation was performed for a given material. Then a model Hamiltonian was constructed from the resulting Kohn-Sham matrix corrected by E_{DC} written in a localized basis set. The values of the Coulomb matrix for the correlated orbitals were estimated or used as fitting parameters. Finally DMFT calculations were performed to improve on the one-particle Green's function of the solid.

In reality, the charge is also corrected by the DMFT self-energy, which in turn changes the exchange and correlation potential away from its LDA value. Therefore charge self-consistent LDA+DMFT is needed.

For this purpose, it is useful to notice that the LDA+DMFT equations can be derived as stationary points of an LDA+DMFT functional, which can be viewed as a functional of the density and local Green's function of correlated orbitals. This is a spectral density-functional

$$\begin{aligned}
 & \Gamma_{\text{DFT+DMFT}} \left[\rho(\vec{r}), G_{\alpha\beta, \vec{R}}, V_{\text{KS}}(\vec{r}), \Sigma_{\alpha\beta, \vec{R}} \right] \\
 = & -\text{Tr} \ln \left[i\omega_n + \frac{\nabla^2}{2} - V_{\text{KS}} - \sum_{R, \alpha\beta \in L} \chi_{\alpha R}^*(\vec{r}) \Sigma_{\alpha\beta R}(i\omega_n) \chi_{\beta R}(\vec{r}) \right] \\
 = & \int V_{\text{KS}}(\vec{r}) \rho(\vec{r}) d^3 r - \sum_n \text{Tr} [\Sigma(i\omega_n) G(i\omega_n)] + \int d^3 r V_{\text{ext}}(\vec{r}) \rho(\vec{r}) d^3 r \\
 = & +\frac{1}{2} \int \frac{\rho(\vec{r}) \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r d^3 r' + E_{xc}^{\text{DFT}}[\rho] + \sum_{\vec{R}} \Phi \left[G_{\alpha\beta, \vec{R}}, U \right] - \Phi_{\text{DC}} . \tag{32}
 \end{aligned}$$

Φ is the sum of two-particle irreducible diagrams written in terms of G and U . It was written down for the first time in Ref. [31] building on the earlier work of Chitra [32,33] and is essential for total energy calculations which require the implementation of charge self-consistency in the LDA+DMFT method. The first implementation of charge self-consistent LDA +DMFT was carried out in a full-potential LMTO basis set [31]. It was used to compute total energy and phonons of δ -plutonium.

The form of the LDA+DMFT functional makes it clear that the method is independent of the basis set used to implement the electronic structure calculation provided that the basis is complete enough. On the other hand, it is clearly dependent on the parameter U chosen, on the form of the double counting correction and the choice of the projector (i.e., the orbitals $\chi_\alpha(\vec{r})$ with $\alpha \in L$ that enter this definition). A projector of the form $P(r, r') = \sum_{\alpha\beta \in L} \chi_{\alpha\vec{R}}^*(\vec{r}) \chi_{\beta\vec{R}}(\vec{r}')$ was used to define a truncation from G to G_{loc} . The inverse of P is the embedding operator E defined by $P \cdot E = I_L$ where I_L is the identity operator in the correlated subspace. If one restricts $E \cdot P$ to the space L , one also obtains the identity operator in that space. E is used to define an embedding of the self-energy $\Sigma(r, r') = E^{\alpha,\beta}(r, r') \Sigma^{loc}_{\alpha,\beta}$.

However, more general projectors can be considered as long as causality of the DMFT equations is satisfied. Ideas for choosing an optimal projector for LDA+DMFT based on orbitals were presented in Ref. [34]. Choosing suitable projectors (and correspondingly a suitable value of the U matrix and a proper double counting correction) is crucial for the accuracy of an LDA+DMFT calculation as demonstrated recently in the context of the hydrogen molecule [35].

5 Electronic structure methods from a diagrammatic many-body perspective

The formulation of LDA+DMFT presented in the previous section is rooted in the model Hamiltonian approach, which contains parameters such as the screened Coulomb interaction matrix or hopping matrix elements. These elements are absent in the starting point of the first principles approaches Eq. (1). We now describe a route proposed by Chitra [32,33] to embed DMFT into a many-body approach of electronic structure within a purely diagrammatic approach formulated in the continuum.

The starting point once again is the theory of everything:

$$\begin{aligned} S = & \int dx \psi^\dagger(x) [\partial_\tau - \nabla^2 + V_{ext}(x)] \\ & + \frac{1}{2} \int dx dx' \psi^\dagger(x) \psi^\dagger(x') v_{Coul}(x - x') \psi(x) \psi(x'), \end{aligned} \quad (33)$$

which can be rewritten exactly in terms of a Hubbard-Stratonovich field $\phi(x)$ that represents the electric field present in the solid,

$$\begin{aligned} S = & \int dx \psi^\dagger(x) [\partial_\tau - \nabla^2 + V_{ext}(x)] \\ & + \frac{1}{2} \int dx dx' \phi(x) v_{Coul}^{-1}(x - x') \phi(x') + \int dx i\phi(x) \psi^\dagger(x) \psi(x). \end{aligned} \quad (34)$$

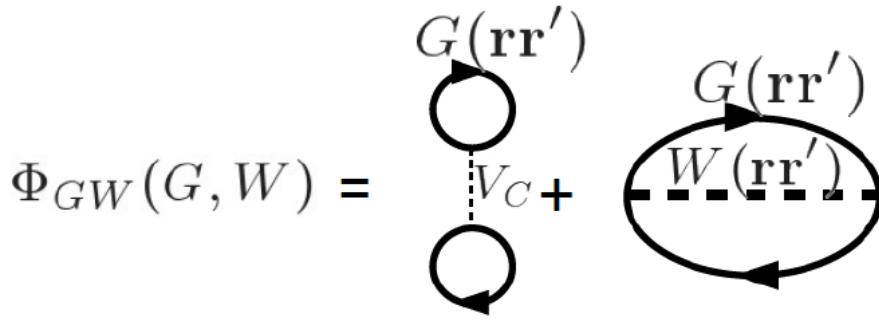


Fig. 4: Lowest order graphs in the Φ -functional of Eq. (37). They give rise to the fully self-consistent GW approximation.

From this action, one can compute the Green's function

$$G(x, x') = -\langle \psi(x) \psi^\dagger(x') \rangle \quad (35)$$

and

$$W(x, x') = \langle \phi(x) \phi(x') \rangle - \langle \phi(x) \rangle \langle \phi(x') \rangle, \quad (36)$$

which are the same symbols as used in the GW method [10].

The free energy of the solid can be written as an exact functional of $G(x, x')$ and $W(x, x')$ by means of a Legendre transformation and results in

$$\begin{aligned} \Gamma[G, W, \Sigma, \Pi] &= -\text{Tr} \ln [G_0^{-1} - \Sigma] - \text{Tr} [\Sigma G] + \frac{1}{2} \text{Tr} \ln [v_{\text{Coul}}^{-1} \Pi] \\ &\quad - \frac{1}{2} \text{Tr} [\Pi W] + E_{\text{Hartree}} + \Phi[G, W]. \end{aligned} \quad (37)$$

This reformulation is exact but not practical unless some approximations are made on the functional Φ , defined as sum of all two-particle irreducible diagrams. The lowest order graphs of this functional are shown in Fig. 4, which reproduce the self-consistent GW approximation.

If one selects a projector, which allows us to define a local Green's function, it was suggested in Refs. [32, 33, 36] that one can perform a local approximation and keep only the local higher order graphs in selected orbitals $\Phi[G, W] \simeq \Phi_{\text{EDMFT}}[G_{\text{loc}}, W_{\text{loc}}, G_{\text{nonlocal}} = 0, W_{\text{nonlocal}} = 0] + \Phi_{\text{GW}} - \Phi_{\text{GW}}[G_{\text{loc}}, W_{\text{loc}}, G_{\text{nonlocal}} = 0, W_{\text{nonlocal}} = 0]$. Since the lowest graph is contained in the GW approximation, one should start from the second order graph and higher order .

These ideas were formulated and fully implemented in the context of a simple extended Hubbard model by Ping Sun and the author [37, 38]. An open problem in this area, explored in Ref. [38], is the level of self-consistency that should be imposed. As discussed in Sec. 1, this important issue is already present in the implementation of the GW method, and the work of Ref. [38] should be revisited using the lessons from the QPGW method [25].

The functional Φ can be viewed as the functional of an impurity model which contains a frequency-dependent interaction U , obeying the self-consistency condition

$$U^{-1} = W_{\text{loc}}^{-1} + \Pi_{\text{loc}}. \quad (38)$$

One can understand the successes of LDA+DMFT from the GW+EDMFT perspective. Consider a system such as cerium, containing light *spd*-electrons and heavier, more correlated, *f*-electrons. We know that for very extended systems, the *GW* quasiparticle band structure is a good approximation to the LDA band structure. Therefore the self-energy of a diagrammatic treatment of the light electrons can be approximated by the exchange-correlation potential of the LDA (or by other improved static approximations if more admixture of exchange is needed). Diagrams of all orders, but in a local approximation, are used for the *f*-electrons. In the full many-body treatment, Σ_{ff} is computed using skeleton graphs with G_{loc} and W_{loc} . To reach the LDA+DMFT equations, one envisions that, at low energies, the effects of the frequency dependent interaction $U(\omega)$ can be taken into account by a static U , which should be close to (but slightly larger than) $U(\omega = 0)$. The *f-f*-block of the Green's function now approaches $\Sigma_{ff} - E_{DC}$.

We reach the LDA+DMFT equations, with some additional understanding on the origin of the approximations used to derive them from the EDMFT+GW approximation as summarized schematically in

$$\Sigma_{GW+DMFT}(\vec{k}, \omega) \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{ff} - E_{DC} \end{pmatrix} + \begin{pmatrix} V_{xc}[\vec{k}]_{spd,spd} & V_{xc}[\vec{k}]_{spd,f} \\ V_{xc}[\vec{k}]_{f,spd} & V_{xc}[\vec{k}]_{f,f} \end{pmatrix}. \quad (39)$$

Realistic implementations of combinations of *GW* and DMFT have not yet reached the maturity of LDA+DMFT implementations and are a subject of current research.

6 Discretization of basis-sets and Coulomb integrals

There are now a large number of implementations of LDA+DMFT in various electronic structure codes in progress and this is an active area of research. In this section we provide some background elementary material, to give the student a feeling for the various parameters that enter in these modern LDA+DMFT calculations. This involves one-electron ideas such as muffin-tin radii and augmentation spheres as well as atomic physics concepts such as Slater integrals. The applications described in Sec. 8 and Sec. 9 were carried out using the LDA+DMFT implementation of K. Haule described in Ref. [39]. Early studies mentioned in section Sec. 9 used the LMTO basis set and the implementation described in Ref. [31].

Eq. (2) is a partial differential equation. To solve it on a computer, a discretization is needed to reduce it to a finite matrix diagonalization problem. More generally, the Kohn-Sham matrix is infinitely dimensional and model Hamiltonians require some reduction to finite dimensional matrices to be used in conjunction with DMFT. This is generally done by introducing a basis set χ_i ,

$$\psi = \sum_i c_i \chi_i. \quad (40)$$

The Schrödinger equation becomes

$$\langle \chi_j | H | \chi_i \rangle = \sum_i c_i \langle \chi_j | H | \chi_i \rangle = \varepsilon \sum_i \langle \chi_j | \chi_i \rangle. \quad (41)$$

That is,

$$\sum_i h_{ji} c_i = \varepsilon \sum_i O_{ji} c_i. \quad (42)$$

$$\begin{cases} h_{ji} = \langle \chi_j | H | \chi_i \rangle = \langle \chi_j | -\nabla^2 + v_{\text{KS}}(\vec{r}) | \chi_i \rangle. \\ O_{ji} = \langle \chi_j | \chi_i \rangle \text{ (overlap matrix)} \end{cases} \quad (43)$$

The linear augmented plane wave (LAPW) [40, 41] method divides the space into two categories, the interstitial region (I) and the muffin-tin region (MT). The LAPW basis set is defined by

$$\chi_{\vec{k}, \vec{G}}(\vec{r}) = \begin{cases} e^{i(\vec{k} + \vec{G}) \cdot \vec{r}}, & \text{for } \vec{r} \in I; \\ \sum_{lm} a_{lm}(\vec{k}) \phi_{lm} + b_{lm}(\vec{k}) \dot{\phi}_{lm}, & \text{for } \vec{r} \in MT \end{cases} \quad (44)$$

In the interstitial region, plane waves constitute a natural basis. In the muffin-tin sphere, the basis set contains linear combinations of atomic-like wavefunctions $\phi_{lm}(\vec{r}, E_\nu)$ and their derivatives, $\dot{\phi}_{lm}(\vec{r}, E_\nu)$, with respect to the energy parameter E_ν , which is called the linearization energy. The key idea is to allow enough variational freedom to reproduce the exact solution of the one-particle Schrödinger equation in the sphere [40]. The basis functions are

$$\phi_{lm}(\vec{r}, \varepsilon) = \phi_{lm}(\vec{r}, E_\nu) + (\varepsilon - E_\nu) \dot{\phi}_{lm}(\vec{r}, E_\nu), \quad (45)$$

where the atomic-like wavefunctions satisfy

$$[-\nabla^2 + V_{av}(\vec{r})] \phi_{lm}(\vec{r}, E_\nu) = E_\nu \phi_{lm}(\vec{r}, E_\nu) \quad (46)$$

and their derivatives with respect to E_ν satisfy

$$[-\nabla^2 + V_{av}(\vec{r})] \dot{\phi}_{lm} = \phi_{lm}. \quad (47)$$

Besides the one-particle Hamiltonian, one also needs to discretize the Coulomb interaction part of the Hamiltonian. We explain how this is done in the context of a single atom. In a model Hamiltonian language, the two most important terms are the Hubbard U , which suppresses charge fluctuations, and the Hund's rule coupling J , which promotes locally large values of spin

$$H_{\text{int}} \sim U \hat{N}^2 + J \hat{S}^2. \quad (48)$$

To see the origin of these terms we start from the atomic Hamiltonian with the Coulomb interaction written in second quantized form

$$\frac{1}{2} \int d^3r d^3r' \psi_\sigma^\dagger(\vec{r}) \psi_{\sigma'}^\dagger(\vec{r}') \frac{1}{|\vec{r} - \vec{r}'|} \psi_{\sigma'}(\vec{r}') \psi_\sigma(\vec{r}) = \frac{1}{2} \sum_{\alpha\beta\gamma\delta} c_{\alpha\sigma}^\dagger c_{\beta\sigma'}^\dagger \langle \alpha\beta | V | \gamma\delta \rangle c_{\delta\sigma'} c_{\gamma\sigma} \quad (49)$$

with

$$\langle \alpha\beta | V | \gamma\delta \rangle = \int d^3r_1 d^3r_2 \phi_\alpha^*(\vec{r}_1) \phi_\beta^*(\vec{r}_2) \frac{1}{|\vec{r}_1 - \vec{r}_2|} \phi_\delta(\vec{r}_2) \phi_\gamma(\vec{r}_1). \quad (50)$$

Using $\phi_{lm}(r, \vartheta, \varphi) = R_l(r) Y_{lm}(\vartheta, \varphi)$ and

$$\frac{1}{|\vec{r} - \vec{r}'|} = 4\pi \sum_{k=0}^{\infty} \frac{r_<^k}{r_>^{k+1}} \frac{1}{2k+1} \sum_{q=-k}^k Y_{kq}(\hat{r}) Y_{kq}^*(\hat{r}'), \quad (51)$$

we restrict ourselves to the d -shell ($l = 2$) as an example. The interaction has the form we encountered previously in the context of the multi-band Hubbard model, Eq. (13)

$$\sum_{m_1 m_2 m_3 m_4} \sum_{\sigma \sigma'} U_{m_1 m_2 m_3 m_4} c_{m_1 \sigma}^\dagger c_{m_2 \sigma'}^\dagger c_{m_3 \sigma'} c_{m_4 \sigma}, \quad (52)$$

Here U is a four-index tensor. In the atom, its form is strongly constrained by symmetries and is parametrized in terms of a few parameters (Slater integrals) F^k .

$$\begin{aligned} U_{m_1 m_2 m_3 m_4} &= \sum_k \frac{4\pi}{2k+1} \sum_{q=-k}^k \int d^3 r_1 \phi_{2m_1}^*(\vec{r}_1) \phi_{2m_4}(\vec{r}_1) Y_{kq}^*(\vartheta_1, \varphi_1) \\ &\quad \times \int d^3 r_2 \phi_{2m_2}^*(\vec{r}_2) \phi_{2m_3}(\vec{r}_2) Y_{kq}(\vartheta_2, \varphi_2) \times \frac{r_<^k}{r_>^{k+1}} \\ &= \sum_k F^k \sum_{q=-k}^k \langle Y_{2m_1} | Y_{kq}^* | Y_{2m_4} \rangle \langle Y_{2m_2} | Y_{kq} | Y_{2m_3} \rangle, \end{aligned} \quad (53)$$

and

$$F^k = \frac{4\pi}{2k+1} \int r_1^2 dr_1 \int r_2^2 dr_2 \frac{r_<^k}{r_>^{k+1}} R_{l=2}^2(r_1) R_{l=2}^2(r_2). \quad (54)$$

When $k = 0$ and hence $q = 0$, $\langle Y_{2m_1} | Y_{00}^* | Y_{2m_4} \rangle \propto \delta_{m_1 m_4}$ and $\langle Y_{2m_2} | Y_{00}^* | Y_{2m_3} \rangle \propto \delta_{m_2 m_3}$. The $k = 0$ contribution in the Coulomb interaction is

$$H_{Coul}^{k=0} \simeq F^0 \sum_{m, m'} \sum_{\sigma, \sigma'} (c_{m\sigma}^\dagger c_{m\sigma}) (c_{m'\sigma'}^\dagger c_{m'\sigma'}). \quad (55)$$

Thus, F^0 defines the Hubbard U in the atom. It involves the *direct* Coulomb integral U_{nmnn} . We now turn to the exchange Coulomb integral

$$U_{mnmn} = \int d^3 r_1 d^3 r_2 \phi_m^*(\vec{r}_1) \phi_n(\vec{r}_1) \frac{1}{|\vec{r}_1 - \vec{r}_2|} \phi_n^*(\vec{r}_2) \phi_m(\vec{r}_2). \quad (56)$$

and its average $J = \frac{1}{2l+1} \sum_{m < n} U_{mnmn}$. To understand its physical meaning consider the case that the full interaction is replaced by its more symmetric (averaged) form $U_{m_1 m_2 m_3 m_4} = J \delta_{m_1, m_3} \delta_{m_2, m_4} + \frac{J}{2} \delta_{m_1, m_4} \delta_{m_2, m_3}$. Then it is easy to evaluate the form of the interaction Hamiltonian using the Fierz identity $\sum_a \sigma^a{}_{\alpha\beta} \sigma^a{}_{\gamma\delta} + \delta_{\alpha\beta} \delta_{\gamma\delta} = 2\delta_{\alpha\delta} \delta_{\beta\gamma}$ to obtain

$$H \sim -2JS^2 \quad (57)$$

with $S^a = \frac{1}{2} \sum_{m, \alpha\beta} c_{m\alpha}^\dagger c_{m\beta} \sigma^a{}_{\alpha\beta}$. Notice the sign in Eq. (57), which gives rise to the famous Hund's rule: to minimize the energy one has to maximize the spin. The Hund's J for d -electrons can be expressed in terms of the Slater integrals [42] by

$$J = \frac{1}{14} (F^2 + F^4), \quad (58)$$

A second parameter in addition to J is needed to parametrize the Slater integrals

$$C = \frac{1}{14} \left(\frac{9}{7} F^2 - \frac{5}{7} F^4 \right) \quad (59)$$

When $C = 0$, additional symmetries are present in the spectra [43]. Sometimes F^2 and F^4 are viewed as parametrized by the Hund's coupling, as $F^2 \simeq \frac{14}{1.6} J$ and $F^4 \simeq \frac{0.6}{1.6} J$. This relation between F^2 and F^4 is exact only for the hydrogen atom. Note that in the case of instantaneous interactions, the terms of odd l , (F^1, F^3, \dots) are absent due to parity symmetry. Finally we give the expression for the atomic energy averaged over all the configurations for d -electrons

$$U_{av} = F^0 - \frac{2}{63} (F^2 + F^4). \quad (60)$$

The Hund's interaction is very important in determining the physical properties of many solids and we will return to this point in Sec. 8.

7 Bridging between first-principles and model Hamiltonian approaches

The ideas that we have pursued in this lecture are rooted in the philosophy of the Anderson model [14]. Many-body correlations are applied to a small subset of orbitals and are kept relatively local in space. The rationale, is that the large majority of electronic states can be treated accurately by some static mean-field plus low-order perturbative corrections, while a summation to all orders is only needed for a small subset of orbitals. This should be contrasted with other methods, such as GW or variational Monte Carlo where all electrons are treated at the same level of approximation. The advantage of LDA+DMFT-like approaches is that they focus the available computational power on the orbitals or sectors that need it the most. The disadvantage is that, like all methods rooted in model Hamiltonians, there is some arbitrariness that has to be resolved, for example in the determination of the form of the projector and the value of the concomitant interaction. The derivation of model Hamiltonians and their parameters can be carried out along two different lines.

The first is a Wilsonian approach, where high-energy degrees of freedom are eliminated. It involves the following schematic steps which results in a model Hamiltonian with well defined parameters:

1. start with the theory of everything in the path integral formulation
2. in metallic systems, screen the long range part of the Coulomb interaction
3. eliminate (integrate-out, approximately) degrees of freedom which are not of interest and therefore outside the scope of the model

A different philosophy for deriving parameters of low-energy Hamiltonian involves matching observables:

1. choose an approximate method
2. apply the method to the first-principle theory and the model
3. match enough quantities to determine (or over-determine) a few physical observables

An good example of this approach is the constrained RPA [44] proposed by Aryasetiawan and developed with his collaborators [45].

The model Hamiltonian will only keep bands within a low energy window near the Fermi surface as degrees of freedom of interest. The observables to be matched are the screened Coulomb interaction among the electrons W . The screened Coulomb potential in first-principle theory is given by

$$W = \frac{v_{\text{Coul}}}{1 + v_{\text{Coul}}\Pi} = \frac{v_{\text{Coul}}}{1 + v_{\text{Coul}}(\Pi_{\text{model}} + \Pi_{\text{rest}})}. \quad (61)$$

Π_{model} is the polarization of the low-energy bands. Π_{rest} is the polarization due to excitations between low- and high-energy bands and within the high-energy bands. The screened Coulomb potential in the model is

$$W_{\text{model}} = \frac{U}{1 + U\Pi_{\text{model}}}. \quad (62)$$

U is the “bare” interaction in the model Hamiltonian. The matching between the low energy theory and the full theory is

$$W = W_{\text{model}}, \quad (63)$$

which results in

$$\frac{U}{1 + U\Pi_{\text{model}}} = \frac{v_{\text{Coul}}}{1 + v_{\text{Coul}}(\Pi_{\text{model}} + \Pi_{\text{rest}})}, \quad (64)$$

leading to

$$U = \frac{v_{\text{Coul}}}{1 + v_{\text{Coul}}\Pi_{\text{rest}}}. \quad (65)$$

This is constrained RPA [44]. In general, $U(\omega; \vec{r}, \vec{r}')$ is a function of ω , \vec{r} , and \vec{r}' . Hence this method delivers both local and non-local frequency dependent interactions. Evaluating the results at zero frequency and projecting U on the relevant orbitals gives rise to the parameters of the model Hamiltonian.

A variant of this approach, a constrained local GW , was proposed in Ref. [46]. Here the observable to be matched is the local W , defined by projecting W onto local orbitals using the same projector to be used in DFT+DMFT. The matching equations are

$$W_{\text{loc}} = (W_{\text{model}})_{\text{loc}}, \quad (66)$$

$$\Pi_{\text{loc}} = (\Pi_{\text{model}})_{\text{loc}}, \quad (67)$$

which result in a different definition of the frequency dependent U in a solid

$$W_{\text{loc}}^{-1} = U^{-1} - \Pi_{\text{loc}} \quad (68)$$

The constrained RPA method, as GW , depends significantly on the level of self-consistency (one-shot, full self-consistency, or QPGW). It was observed in Ref. [46] that the one-shot approximation used in all earlier studies considerably underestimates the values of U that should

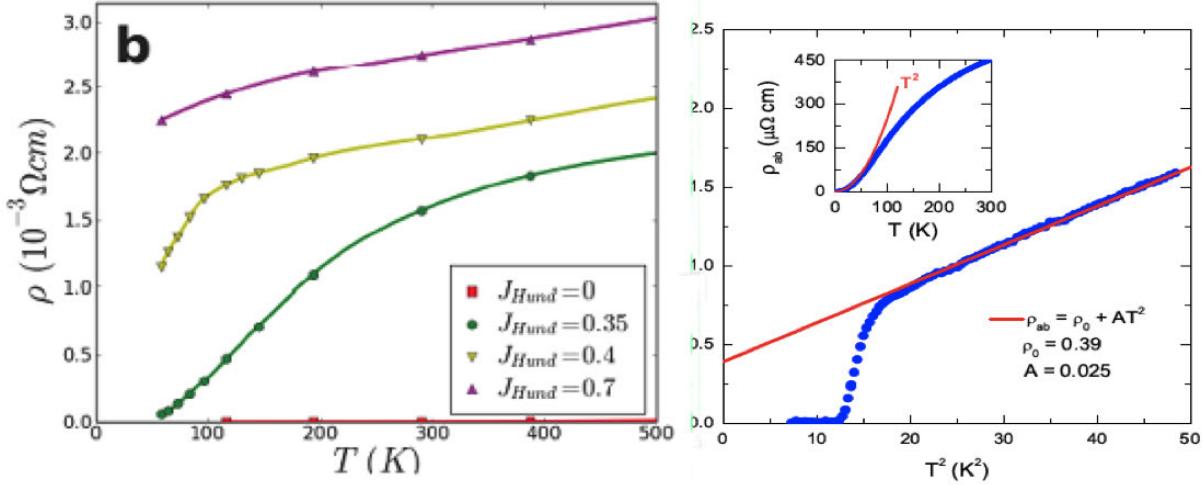


Fig. 5: (Left Panel) Resistivities as a function of Hund’s rule coupling J , Ref. [50]. Notice the extreme sensitivity to J . The effects of U on the correlation strength are small even for values of U comparable to the bandwidth. The correlations are induced very rapidly by the Hund’s J , and a coherence-incoherence crossover as a function of temperature was predicted. (Right Panel) Recent observation of this behavior in ultra-pure KFe_2As_2 [51].

be used together with a localized projector. On the other hand, the fully self-consistent GW which was implemented for solids in Ref. [47] gives instead fairly reasonable U values and total energies, while being less accurate than the QPGW and the one-shot GW for electronic spectra. Further investigations of this point in other materials are needed. Furthermore, the determination of the parameters and the type of projectors to be used in LDA+DMFT-like implementations remains a fundamental challenge in condensed matter physics.

8 Applications: Iron pnictides and Hund’s metals

The field of correlated-electron materials continues to periodically produce surprising discoveries. The latest in the series is the high-temperature superconductivity in materials containing iron pnictide layers [48]. This recent development provided a unique opportunity to confront electronic structure methods with rapidly developing experiments and assess the predictive power of current methodologies and implementations. We use this as a first example in these lectures.

Shortly after the experimental discovery of the iron pnictides, it was determined that the electron-phonon coupling was not responsible for their superconductivity and correlations in the form of a mass enhancement (m^*/m between 3 and 5) were predicted [49]. Even more surprising was the *origin* of the mass enhancement which was elucidated in Ref. [50]. The left panel of Fig. 5 shows the crossover from coherence (Fermi-liquid behavior) at low temperatures to incoherence (bad-metal behavior) at high temperatures predicted in Ref. [49]. The right panel of Fig. 5 describes the evolution of the resistivity as a function of J , for a reasonably large value of the Hubbard U ($U=5$, comparable to the bandwidth). For small J , the system behaves as a weakly interacting material, with very low resistivity and negligible mass enhancement [49]. The rea-

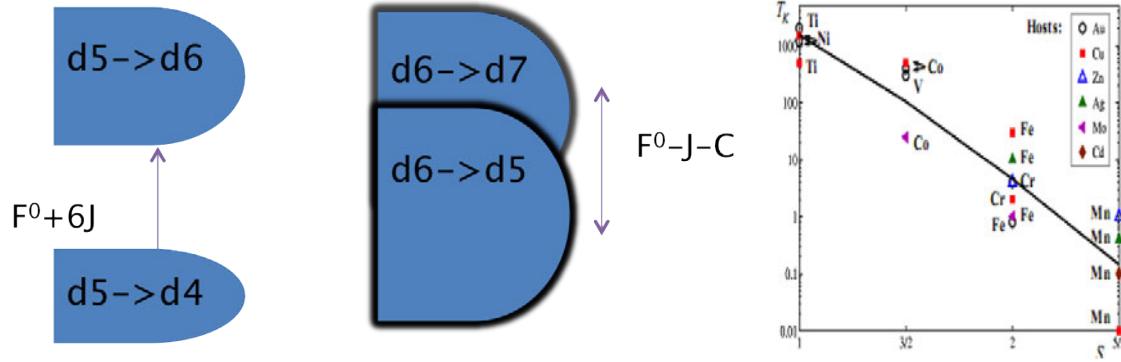


Fig. 6: J promotes metallicity in a d^6 configuration (middle) and insulating behavior in a d^5 configuration. The figure on the right shows the experimental dependence of the Kondo temperature on the d-valence [53]. It falls exponentially as the half filled shell (Mn) is approached.

son for this behavior is the very small value of the crystal-field splittings in the pnictides (of the order of ten meV as opposed to the value characteristic of oxides which is of the order of eV). The orbital degeneracy is then very large, with considerable room to move within the d -shell, rendering Mott blocking ineffective. In different words, the critical U for the Mott transition scales with N [52], ($U_{c2} \sim N^2$ and $U_{c1} \sim N$) and enormous values of U would be needed to induce Mott localization in this system.

The iron pnictide materials were therefore not just new high-temperature superconductors arising from a magnetic element. They were a new class of strongly correlated materials, where the correlations derive not from the blocking effect of the Mott Hubbard U , but from the effects of the Hund's rule coupling J . The theory of Hund's metals is not fully developed yet. Some basic understanding can be traced to the early work of van der Marel and Sawatzky [42], who observed that while in a half filled configuration such as Mn d^5 , Hund's rule J enhances the Mott Hubbard gap, in a d^6 configuration the Hund's rule J reduces the gap between the Hubbard bands, thus promoting metallicity, as shown in Fig. 6. Hundness is also clearly seen in the valence histogram, describing pictorially the diagonal elements of the local many body density matrix. Physically, it represents the fraction of the time that the shell spends in each different atomic eigenstate. It is shown in the right panel of Fig. 9 for BaFe_2As_2 . The material is clearly metallic with a very large number of configurations and several valences participating in the histogram. The Hund's J weights heavily the maximal spin states within each valence.

The Hund's coupling also has a dramatic impact at low energies. This has been known from the studies of magnetic impurities in transition-metals which were discussed intensively in the 1960s [54]. The Kondo temperature of transition metals decreases dramatically as it approaches the half-filled shell as shown in the right panel of Fig. 6. This can be understood as a result of the blocking of the orbitals which reduces the Kondo interaction to a diagonal form. The Kondo scale is exponential in the Kondo coupling. In a $SU(2N)$ -symmetric situation, containing only the interaction U , the Kondo scales as $\exp(-1/J\rho N)$ where N is the orbital degeneracy. Introduction of Hund's coupling removes the degeneracy from $SU(N)$ to $SU(2)$, and renormalizes the Kondo-coupling from J to J/N , resulting in a Kondo scale $\exp(-N/J\rho)$. These con-

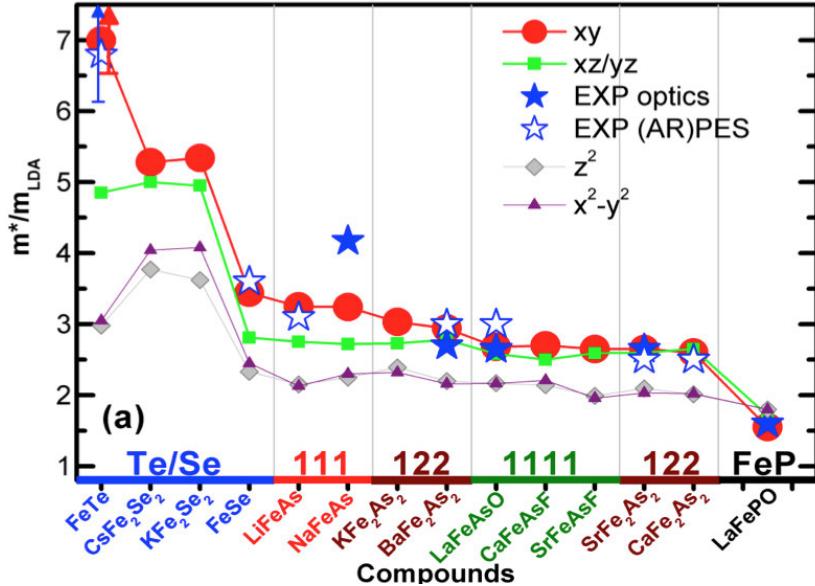


Fig. 7: Effective masses across the families of iron pnictides. All calculations were performed with fixed F^0 , F^2 , F^4 , and double-counting correction parameters. The variations in correlation strength agree reasonably well with experiment and can be traced to the position of the pnictogen height.

siderations are suitable for understanding the half-filled situation [55]. A full weak-coupling renormalization group treatment of the impurity model underlying the Hund's metal was only given very recently [56].

The question of whether the iron pnictides should be thought of as weakly correlated itinerant magnets, doped Mott insulators, or Hund's metals continues to be actively debated in the community. An important question is what controls the strength of the correlations within LDA+DMFT. At this point, technical advances in implementation finally enable the calculation of physical properties for whole families of compounds as illustrated in Fig. 7 from Ref. [57].

A big advantage of this type of calculation is that while absolute values of physical quantities are very sensitive to the strength of the Hund's coupling, this quantity is not expected to vary much from material to material and can be kept fixed as the chemical trends across similar materials are examined. This type of calculation clarified early confusion which classified some iron pnictides such as the 1111 system as weakly correlated, while placing others such as the 122 system in the strong correlation regime as a result of variations in the atomic parameters and double-counting corrections. This unified picture of the iron pnictide families was also confirmed by subsequent experimental optical studies.

The other factor that controls the strength of the interaction in the iron pnictides is the variation of the pnictide valence, with the correlation strength being an increasing function of decreasing valence. Arguments in favor of this point of view, in a very itinerant picture of Hund's metals, was advanced in Refs. [56] and [58].

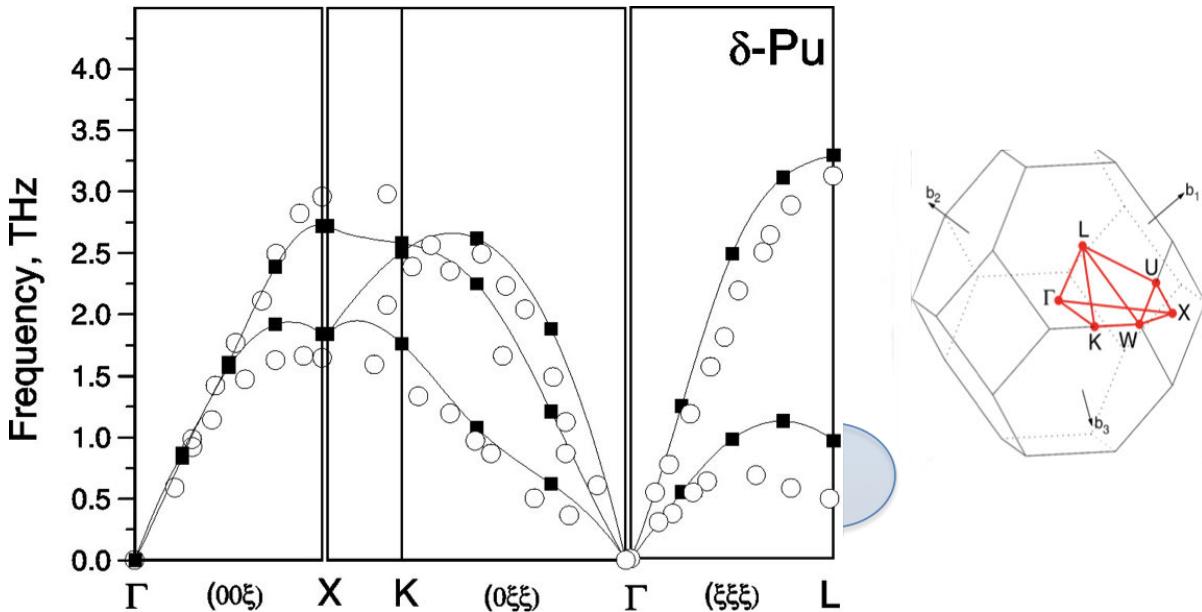


Fig. 8: Theoretical prediction for the phonon spectra of the δ -phase of Pu (empty circles) [60] and observation by inelastic X-ray scattering (black squares) [61]. The overall agreement is reasonable with maximum deviations near the L -point of the Brillouin-zone as indicated.

9 Applications: Actinides

Computation of total energies was almost exclusively the domain of density-functional theory. This can be rationalized by saying that density-functionals target the density and total energy of the material and therefore are more accurate for these quantities than for their corresponding excitation spectra. This is supported by the fact that even in materials as correlated as the high-temperature superconductors, LDA or GGA predict the structural properties with a few percent accuracy. A notable exception to this rule is provided by a $5f$ system, elemental plutonium. Nonmagnetic LDA or GGA underestimates the volume of the δ -phase by more than 30%, while allowing for magnetism gives a volume close to experiment but with a very large moment, of the order of $5\mu_B$, which is not observed experimentally. A similar problem arises in other electronic structure methods, ranging from GW to hybrid density-functionals.

The computation of total energies and phonon frequencies became possible with the introduction of the LDA+DMFT functional. The first application of charge self-consistent LDA+DMFT [59] pointed a path to solving the Pu conundrum by demonstrating that the correct volume of δ -Pu emerges from the paramagnetic LDA+DMFT calculation. Predictions for the phonon spectra [60] were largely confirmed by inelastic X-ray experiments at the ESRF.

The difference between theory and experiment highlighted in Fig. 8 focuses the research by raising interesting questions. In that early work, simplified impurity solvers were used. Hence, the calculations can be improved further. Also, temperature-dependent experimental studies should be performed, since the calculations were zero temperature calculations. Finally, effects of alloying and inhomogeneities could be investigated. This is clearly an area where theory

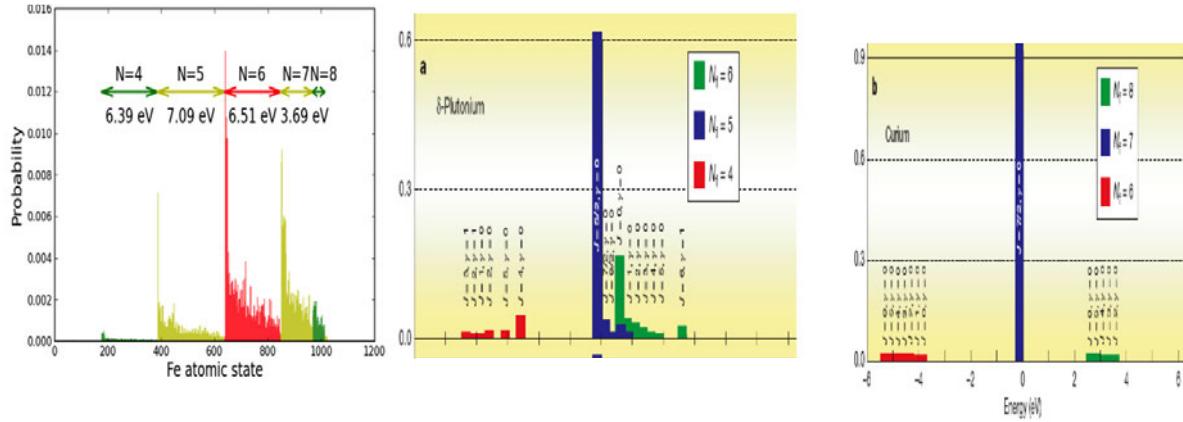


Fig. 9: LDA+DMFT valence histogram for different materials. In Cm (right) the ordinary notion of valence applies. Pu (middle) exhibits a clear mixed valence character [62]. The left panel displays the valence histogram of a Hund's metal BaFe_2As_2 [46].

of strongly correlated materials can continue to make important contributions to an area where experiments are very difficult.

Another important aspect of LDA+DMFT calculations is the qualitative insights they provide into the behavior of a material. Plutonium was shown to be a strongly mixed-valence system, and the absence of magnetism was explicitly demonstrated [62]. The LDA+DMFT valence histograms describe the fraction of the time that the atom spends in each atomic eigenstate. When the f -electron is very localized, there is only one atomic eigenstate that is important. This is illustrated in Fig. 9 for curium. Plutonium is very different, having appreciable fluctuations into the $5f^6$ configuration as shown in Fig. 9. The mixed valence of Pu is responsible for its unique physical properties.

The studies of Pu over the last decade illustrate very nicely the advances in the quality of the LDA+DMFT implementations. Very recently, calculations for the ground state of Pu, α -Pu, a complicated monoclinic structure with many atoms in the unit cell, were carried out, using CTQMC as an impurity solver [63].

Another recent development is the determination of the energy *vs.* volume for all the phases of Pu, carried out in Ref. [64]. It was shown that all their phases are mixed-valent. Furthermore, correcting the LDA energy with the mean-field slave-boson (or Gutzwiller) correction brings all the phases very close together in energy as highlighted in Fig. 10 (Ref. [64]). These calculations are outside the scope of what can be currently done using exact impurity solvers but are easily accessible to the formalism introduced in Sec. 2 or the equivalent (at zero temperature) Gutzwiller approximation introduced by Xi Dai and collaborators. This brings us back to the beginning of the lectures.

10 Summary and outlook

After completing the development of quantum mechanics Dirac stated that “the underlying laws necessary for the mathematical theory of the whole chemistry are thus completely known

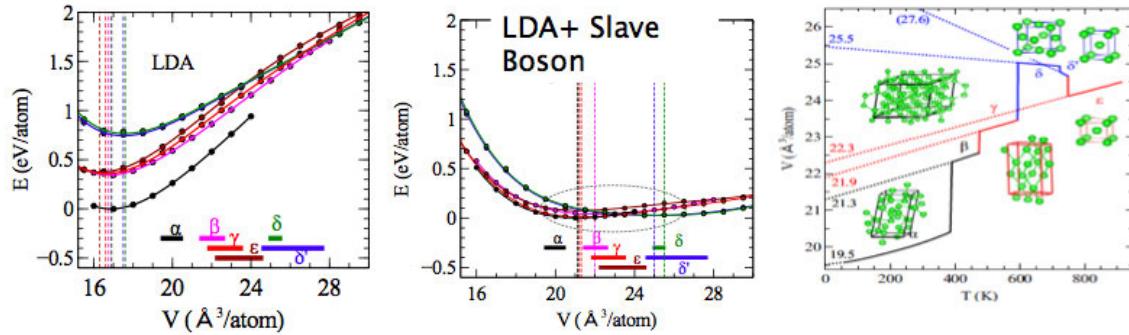


Fig. 10: *Left: Energy vs. volume of all the phases of Pu from the work of Lanatà et al. [64]. Notice that within the slave-boson method the energy differences are two orders of magnitude smaller, explaining the extreme sensitivity of the material to changes in concentration of impurities, temperature, etc. This sensitivity is highlighted in the right panel showing the experimental volume-temperature phase-diagram of Pu. The dotted lines indicate the zero-temperature equilibrium volumes extrapolated by linear interpolation. Correlations shift the region of volumes (highlighted) that becomes relevant at zero pressure. LDA already contains the relevant information about structural differences.*

and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.” But he proceeded to add that “approximate practical methods of applying quantum mechanics should be developed which can lead to an explanation of the main features of complex atomic systems without too much computation.”

In his famous article *More is Different* [65], P.W. Anderson, remarked that “the constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity.” He goes on to say that “at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other,” stressing that at each level of description new concepts are needed to describe nature and there are new laws to be discovered.

Dirac and Anderson’s dictums are sometimes viewed as contradictory, but in condensed matter physics and in particular in the field of strongly correlated electrons both are needed to make progress. The quest for ideas to describe the emergent phenomena as well as the invention of techniques to compute physical properties using the basic laws of quantum mechanics of matter both play a very important role.

The developments of methodologies to treat correlated materials are an illustration of Dirac’s vision and they have lead to concepts, abstractions, and physical pictures that enable us to understand the behavior of correlated materials that are useful and will guide further studies.

There is a close interaction between scientific advances and the development of new methodologies. New methodologies enable breakthroughs in challenging scientific problems, and in turn outstanding scientific problems spur the development of new methodologies.

The development of DMFT for model Hamiltonians led to a detailed understanding of the mechanism of the Mott transition. In turn, the need to understand the Mott transition, a transition that lacks an obvious order parameter, was an important driving force for the development of DMFT.

DMFT and its extensions to clusters enabled accurate solutions of model Hamiltonians. It also gave us useful concepts for thinking about strongly correlated materials. The DMFT Weiss field makes quantitative the notion of the degree of localization of the electron. The local Green's function, with its characteristic three peak structure, gives a precise formulation of the Mott transition in terms of the transfer of one-electron spectral weight. The existence of a finite-temperature Mott transition marks a sharp boundary beyond which perturbation theory in the interactions fails.

Condensed matter theory has a dual role. One one side, it provides tools for predicting the properties of materials. On the other side, it builds the conceptual framework in which to frame and understand the results of experiments. We used two classes of materials, the actinides and the iron pnictides to illustrate the rapid progress in the field. The Pu problem was intractable with the tools of band theory, but using a combination of slave-boson and DMFT methods the determination of its phase diagram appears within sight. The iron pnictides provided a real-time demonstration of the power of the LDA+DMFT methodology for their quantitative description and at the same time resulted in a surprising discovery of a new class of strongly correlated systems. There are many more that have been studied already, and even more to be discovered. The introduction of LDA+DMFT enabled the computation of the photoemission and inverse photoemission spectra of correlated materials starting from first principles. The intensity and position of the qualitative features, already present in the model Hamiltonian, are now made quantitative and system-specific in a way that allowed comparison with experiments. The calculated spectral properties (ARPES, optics, neutron scattering, etc.) of a large number of p , $3d$, $4d$, $4f$, $5d$, and $5f$ -based materials were in surprisingly good agreement with experiments. Through the study of a very large number of materials, the community has gained confidence that we have a zeroth-order picture of strongly correlated materials, with the Kohn-Sham Hamiltonian with a double-counting correction subtracted, as a one-particle Hamiltonian and a Coulomb interaction matrix parametrized by a few Slater-Racah parameters. We have a working practical approach that gives a zeroth order-picture of correlated solids, an important challenge is to quantify its accuracy and limitations. We can start looking for deviations from this framework as well as continue to improve its implementation and foundation to increase its accuracy.

Finally it is worth reminding students that all the remarkable discoveries in strongly correlated electron materials have been the result of serendipity, and there is no reason to doubt that this will continue in the foreseeable future. Still, at this point in time, theorists working on strongly correlated electrons have in their hand sufficiently powerful tools to participate more closely in the process of understanding of these fascinating materials and accelerate their discovery. These are very exciting times to enter the field of correlated electron material research.

Acknowledgments

Research in Sec. 9 and 8 was supported by the DOE and the NSF respectively. Wenhua Xu was instrumental in the development of these notes. I thank him and K. Haule, Z. Yin, A. Kuteepov, N. Lanatà, Y. Yao for collaborations and multiple discussions.

References

- [1] M.C. Gutzwiller, Phys. Rev. Lett **10**, 159 (1963)
- [2] C.M. Varma and Y. Yafet, Phys. Rev. **13**, 2950 (1976)
- [3] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963)
- [4] K. Wilson Rev. Mod. Phys. **47**, 773 (1975)
- [5] M. Imada, A. Fujimori and Y. Tokura Rev. Mod. Phys. **70**, 1039 (1998)
- [6] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1995)
- [7] G. Kotliar *et al.*, Rev. Mod. Phys. **78**, 865 (2006)
- [8] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994)
- [9] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965)
- [10] L. Hedin, Phys. Rev. **139**, A796 (1965)
- [11] S.V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004)
- [12] E. Lieb and F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
- [13] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [14] P.W. Anderson, Phys. Rev. **124**, 41 (1961)
- [15] G. Kotliar and A.E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986)
- [16] R. Fresard and G. Kotliar, Phys. Rev. B **56**, 12909 (1997)
- [17] F. Lechermann, A. Georges, G. Kotliar, and O. Parcollet, Phys. Rev. B **76**, 155102 (2007)
- [18] T. Li, P. Wölfle, and P.J. Hirschfeld, Phys. Rev. B **40**, 6817 (1989)
- [19] R. Fresard and P. Wölfle, Int. J. Mod. Phys. B **6**, 685 (1992)
- [20] R. Raimondi and C. Castellani, Phys. Rev. B. **48**, 11453(R) (1993)
- [21] N. Lanata, Y. Yao, C Wang K.M. Ho and G. Kotliar arXiv:1405.6934
- [22] A. Georges, and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [23] G Kotliar and D Vollhardt, Physics Today, March 2004, p. 53
- [24] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)

- [25] J.M. Tomczak, M. van Schilfgaarde, and G. Kotliar, Phys. Rev. Lett. **109**, 237010 (2012)
- [26] N.E. Zein, S.Y. Savrasov, and G. Kotliar, Phys. Rev. Lett. **96**, 226403 (2006)
- [27] Z.P. Yin, A. Kutepon, and G. Kotliar, Phys. Rev. X **3**, 021011 (2013)
- [28] D. Jacob, K. Haule, G. Kotliar, EPL **84**, 57009 (2008)
- [29] V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin, and G. Kotliar, J. Phys: Condens. Mat. **9**, 7359 (1997)
- [30] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
- [31] S.Y. Savrasov and G. Kotliar, Phys. Rev. B **69**, 245101 (2004)
- [32] R. Chitra and G. Kotliar, Phys. Rev. B **62**, 12715 (2000)
- [33] R. Chitra, and G. Kotliar, Phys. Rev. B **63**, 115110 (2001)
- [34] I. Paul and G. Kotliar I. Paul and G. Kotliar, Eur. Phys. Jour. B **51**, 189 (2006)
- [35] J. Lee and K. Haule arXiv:1403.2474
- [36] S. Biermann F. Aryasetiawan, A. Georges PRL **90**, 086402 (2003)
- [37] P. Sun, G. Kotliar, Phys. Rev. B **66**, 085120 (2002)
- [38] P. Sun, and G. Kotliar, Phys. Rev. Lett. **92**, 196402 (2004)
- [39] K. Haule, C. H. Yee, and K. Kim, Phys. Rev. B **81**, 195107 (2010).
- [40] O.K. Andersen, Phys. Rev. B **12**, 3060 (1975)
- [41] D. J. Singh and L. Nordström: *Planewaves, Pseudopotentials and the LAPW Method* (Springer, New York, 2006)
- [42] D. van der Marel and G.A. Sawatzky, Phys. Rev. B **37**, 10674 (1988)
- [43] O. Laporte and J.R. Platt, Phys. Rev. **61**, 305 (1942)
- [44] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Phys. Rev. B **70**, 195104
- [45] T Miyake, F. Aryasetiawan, and M. Imada, Phys. Rev. B **80**, 155134 (2009),
- [46] A. Kutepon, K. Haule, S.Y. Savrasov, and G. Kotliar, Phy. Rev. B **82**, 045105 (2010)
- [47] A. Kutepon S. Savrasov and G. Kotliar Phys. Rev. B **80**, 041103(R) (2009)
- [48] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, J. Am. Chem. Soc. **130**, 3296 (2008)

- [49] K. Haule, J. H. Shim, and G. Kotliar, Phys. Rev. Lett. **100**, 226402 (2008)
- [50] J. K. Haule, and G. Kotliar, New Journal of Physics **11**, 025021 (2009)
- [51] F. Hardy, A.E. Böhmer, D. Aoki, P. Burger, T. Wolf, P. Schweiss, R. Heid, P. Adelmann, Y.X. Yao, G. Kotliar, J. Schmalian, and C. Meingast, Phys. Rev. Lett. **111**, 027002 (2013)
- [52] S. Florens, A. Georges, G. Kotliar, and O. Parcollet, Phys. Rev. B **66**, 205102 (2002)
- [53] M. Daybell, in *Magnetism*, pp. 122-147, Vol. 5, ed. by G. Rado and H. Suhl (Academic Press, New York, 1973)
- [54] J. R. Schrieffer, J. App. Phys. **38**, 1143 (1967)
- [55] I. Okada and K. Yosida, Prog. Theor. Phys. **49**, 1483(1973)
- [56] C. Aron and G. Kotliar, arxiv:1401.0331
- [57] Z.P. Yin, K. Haule, and G. Kotliar, Nature Mat. **10**, 932 (2011)
- [58] Z.P. Yin and G. Kotliar, Phys. Rev. B **86**, 195141 (2012)
- [59] S.Y. Savrasov, G. Kotliar, and E. Abrahams, Nature **410**, 793 (2001)
- [60] X. Dai, S.Y. Savrasov, G. Kotliar, A. Migliori, H. Ledbetter, and E. Abrahams, Science **300**, 953 (2003)
- [61] J. Wong, M. Krisch, D.L. Farber, F. Occelli, A.J. Schwartz, T.-C. Chiang, M. Wall, C. Boro, and R. Xu, Science **301**, 1078 (2003)
- [62] J.H. Shim, K. Haule, and G. Kotliar, Nature **446**, 513 (2007)
- [63] J.-X. Zhu, R.C. Albers, K. Haule, G. Kotliar, J.M. Wills Nat. Commun. **4**, 2644 (2013)
- [64] N. Lanatà, Y.X. Yao, C.Z. Wang, K.M. Ho, and G. Kotliar, arXiv:1407.4862
- [65] P.W. Anderson, Science **177**, 393 (1972)

3 Dynamical Mean-Field Theory: Materials from an Atomic Viewpoint Beyond the Landau Paradigm

Antoine Georges
Collège de France
Place Marcelin Berthelot, Paris

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | Why DMFT? | 3 |
| 2.1 | Atomic physics in the solid-state: Mott insulators | 3 |
| 2.2 | Atomic physics in the solid-state: metals close to the Mott transition | 5 |
| 2.3 | Origin of the Mott phenomenon: blocking of charge, not magnetism | 7 |
| 3 | DMFT in a nutshell | 10 |
| 3.1 | DMFT: solids as self-consistently embedded atoms | 10 |
| 3.2 | When is DMFT exact or accurate? | 12 |
| 3.3 | From particles to waves | 13 |
| 4 | How good metals turn bad: quasiparticles beyond Landau theory and spectral weight transfers | 14 |
| 5 | Atomic physics in the solid-state: Hund's metals | 16 |
| 6 | Growing correlations: superexchange, pseudogap, and cluster extensions of DMFT | 17 |
| 7 | Hiking down the energy trail: DMFT as a compass | 18 |

1 Introduction

Materials are made of atoms. Kanada in ancient India and Democritus in ancient Greece already had this intuition, and by now this is not exactly a surprising or revolutionary statement. However, many standard solid-state physics textbooks do not emphasize this point of view very strongly, to say the least. Condensed matter physics is often presented there as the science of the electron gas, the underlying atoms being merely responsible for producing a periodic potential. Thanks to Bloch's theorem and independent particles pictures, invoking Walter Kohn and Lev Landau as tutelary figures, the emphasis is quickly put on electrons, or at best quasiparticles. So much for the atoms, which many students will then view as annoying curiosities that one should soon forget, and of no use for the final exam.

Chemists of course, know better. We condensed-matter physicists have a lot to learn from chemists, but we have trouble discussing with them when the first thing we show is a bunch of energy bands (notwithstanding that a colourful plate of spaghetti can be truly enjoyable food for thought). To enter a constructive dialogue with a chemist, better speak about atomic or molecular orbitals, bonding, hybridization, etc.

Physicists working on materials with strong electron correlations have learned the hard way that *atoms matter*. More precisely: an atom is a small many-body problem in itself, whose eigenstates are multiplets (not just Slater determinants built out of the hydrogen atom single-particle levels). In order to describe the physics of strongly correlated materials (especially the ones with the most localized orbitals such as *f*-electron compounds), it is a good idea to adopt a theoretical framework in which these atomic multiplets are correctly described, at least for a subset of the atomic shells.

Of course, how we perceive reality depends on the scale at which we look at it. This has been beautifully formalized by the renormalization group: to each scale (in terms of distance, energy, or time) corresponds an appropriate effective theory. While there is no doubt that at high-enough energy (short time-scales, short distances) localized atomic excitations are important, why should they matter at low-energy (long time-scales, long distances)?

For many materials (the weakly correlated ones), they do not – hence the validity of the *standard model* of condensed-matter theory textbooks alluded to above. The reason is that, in this case, the kinetic energy of the quasiparticles (which are only slightly renormalized from independent electrons in their periodic potential) is a very high-energy scale. Over all the range of energies or temperatures where most experiments are conducted, the quasiparticles remain long-lived and form a non-degenerate Fermi gas with a very high effective Fermi energy (or, equivalently, a high degeneracy temperature or quasiparticle coherence scale).

In materials with strong electronic correlations, this is no longer true. The relevant orbitals (typically: the *d*-orbitals of transition-metals and their oxides, the *f*-orbitals of rare-earths, actinides and their compounds, the molecular orbitals of organic conductors) are very localized and their bandwidth (just the bare one obtained from bandstructure) is comparable or even smaller than the typical matrix elements of the screened Coulomb interaction. In such a situation, there is not such a clear separation of energy scales, and we cannot dispose of atomic correlations so

easily. In Mott insulators [1], this has a dramatic consequence: the electrons remain localized because the cost in Coulomb energy is too high for them to move in comparison to the potential gain in kinetic energy.

In metallic systems, strong correlations also have drastic effects: the effective degeneracy temperature of the quasiparticles is often renormalized to very low values, for example down to a few Kelvin in heavy-fermion compounds (Kondo effect [2]), and still as low as ~ 25 K in a metallic oxide such as Sr_2RuO_4 [3]. As a consequence, Landau's Fermi liquid theory, which postulates a degenerate Fermi gas of very long-lived quasiparticles, only applies below a very low energy scale (when it applies at all). Dealing properly with atomic correlations is then crucial. In particular, these atomic correlations determine how, and by how much, the quasiparticle coherence scale is renormalized in comparison to the bare bandwidth. Furthermore, regimes where quasiparticles are no longer fully coherent become apparent in experiments even at not very high temperatures. The crossover corresponding to the gradual destruction of the quasiparticles as the system is heated up (or equivalently, how quasiparticles emerge as the system is cooled down) must be addressed in order to understand the experimental observations – often characterized by large transfers of spectral weight between low-energy and intermediate- or high-energy excitations.

In order to address these issues, we need a theoretical description that adapts itself to the energy scale at which we look at the system. At high-energy, it has to correctly describe atomic physics, correlations, and multiplet structures. At low energy, it has to account for the emergence of long-lived quasiparticles. And it has to describe how this process takes place as the energy scale is lowered. This is what Dynamical Mean-Field Theory (DMFT) does. The term *dynamical* is perhaps not ideally chosen, since we are not talking here about the out-of-equilibrium dynamics of the system. Instead, it indicates that the theory handles the different time-scales or energy-scales involved in the excitation spectrum of the system at equilibrium. In order to do so, DMFT introduces a generalization of the classical Weiss mean-field concept to that of a full function of energy (or time scale).

Let me end this introduction on a disclaimer. These are not standard lecture notes (such as the ones in [4]). Many good reviews exist by now which can be consulted for an introduction to DMFT and its extensions, as well as for detailed technical aspects. Rather, I would like to present a set of physical issues which in my view motivate the DMFT concept. I will do this with hindsight, not necessarily following the historical development of the theory.

2 Why DMFT?

2.1 Atomic physics in the solid-state: Mott insulators

In Fig. 1, I reproduce the early angular-integrated photoemission spectra of some transition metal oxides, from the pioneering work of Fujimori and coworkers [5]. All these oxides have a common feature: the d -shell of the transition-metal is nominally occupied by a single electron (d^1 configuration). As one moves from top (ReO_3) to bottom (YTiO_3), the degree of correlation

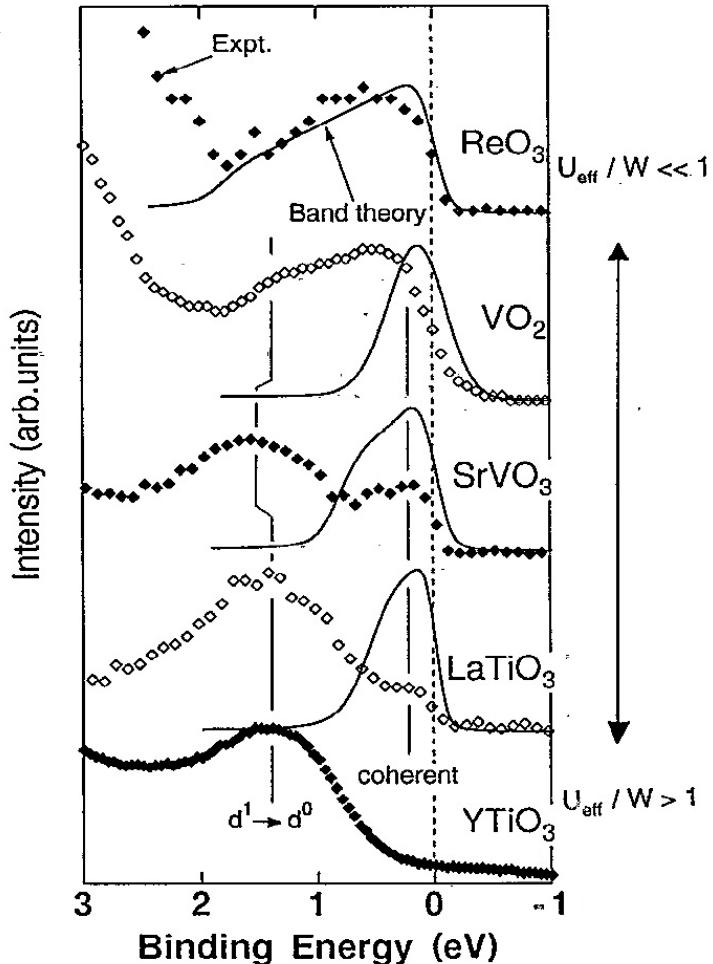


Fig. 1: Photoemission spectra of several d^1 transition metal oxides, as reported in Ref. [5]. The strength of electronic correlations increases from ReO_3 (a weakly correlated metal) to YTiO_3 (a Mott insulator). The plain lines are the densities-of-states obtained from band structure calculations. A lower Hubbard band around -1.5 eV is clearly visible for the most correlated materials, both for the metallic (SrVO_3) and the Mott insulating (LaTiO_3 , YTiO_3) materials.

increases (for reasons mentioned below). LaTiO_3 and YTiO_3 are Mott insulators. Their spectra display a clear peak around -1.5 eV binding energy. This feature cannot be reproduced from band-structure calculations (based, e.g., on DFT-LDA). It is, however, very easy to understand as an *atomic-like* transition. The photoelectron kicks an electron out of the d -shell, inducing a $d^1 \rightarrow d^0$ transition. In the jargon of correlated electrons physicists, this is called a *lower Hubbard band* (LHB), but it really has nothing to do with a band – at least not when we look at it from the point of view of single-particle spectroscopy: it is an atomic-like transition, broadened, of course, by the solid-state environment.

The simplest way to model this is to consider a caricature of an atom: a single atomic level at energy ε_d , which can be occupied by zero, one, or two electrons with opposite spin. Neglecting any orbital quantum number, the hamiltonian of such a *Hubbard atom* reads

$$H_{\text{at}} = \varepsilon_d(\hat{n}_\uparrow + \hat{n}_\downarrow) + U \hat{n}_\uparrow \hat{n}_\downarrow \quad (1)$$

in which U is the repulsion energy that the two electrons have to pay to sit in the same atomic level (In the solid, screening has to be taken into account when evaluating U). There are four eigenstates: $|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$, with energies $0, \varepsilon_d, \varepsilon_d, 2\varepsilon_d + U$, respectively. The energy of the transition $d^1 \rightarrow d^0$ is ε_d , hence the above measurement provides information on the effective position of the transition-metal atomic level in the solid. In such a simple model, one would also expect another peak in the electron addition spectrum (inverse photoemission) at an energy $\varepsilon_d + U$, corresponding to the transition $d^1 \rightarrow d^2$. Measuring both transitions provides information about U . This can be summarized in the spectral function of this simple isolated Hubbard atom, which reads ($n_d = \langle \hat{n}_\uparrow + \hat{n}_\downarrow \rangle$)

$$A_d(\omega) = \left(1 - \frac{n_d}{2}\right) \delta(\omega - \varepsilon_d) + \frac{n_d}{2} \delta(\omega - \varepsilon_d - U). \quad (2)$$

The message of this section is: the single-particle excitation spectrum of Mott insulators is most easily understood in terms of atomic transitions. Bandstructure calculations do not handle this properly. A proper description of many-body atomic eigenstates (multiplets) [6] is required to understand these spectra. Of course, the eigenstates of our oversimplified Hubbard atom are so simple that they hardly deserve to be called multiplets. Orbital degrees of freedom and additional matrix elements of the Coulomb interaction must be included for a realistic description.

2.2 Atomic physics in the solid-state: metals close to the Mott transition

Let us now turn to the spectrum of SrVO_3 shown in Fig. 1. This material is a metal, as indicated by the presence of low-energy spectral-weight and the absence of a gap. Nonetheless, the atomic-like transition (LHB) at ~ -1.5 eV is still visible. Hence, the high-energy spectrum of such a metal is not so different from that of related Mott insulators, indicating that atomic physics is relevant for correlated metals as well. This is not even a very correlated metallic state, rather one with an intermediate level of correlations still far away from the Mott transition. As compared to the band-theory mass, the measured quasiparticle effective mass is enhanced by approximately a factor of two.

But this is high-energy. What about low-energy excitations ? For this, we have to turn to more recent photoemission experiments, which have disentangled the relative contributions of the surface and the bulk [7, 8]. The spectra obtained in Ref. [8] are reproduced in Fig. 2. There, we see that the bandwidth of the low-energy part of the density of states, corresponding to quasiparticle excitations, is narrowed by roughly a factor of two as compared to the bandstructure (LDA) calculation. This is in line with the measured enhancement of the quasiparticle effective mass.

This phenomenon was understood early on by Brinkman and Rice [9] as being due to the Hubbard repulsion U . Their description was based on the simple Gutzwiller approximation. As the system approaches the metal-to-Mott-insulator transition, the spectral weight Z of the quasiparticle excitations diminishes, and vanishes at the transition point U_{BR} . In this simple description (which neglects the effect of the superexchange on quasiparticles), the effective mass

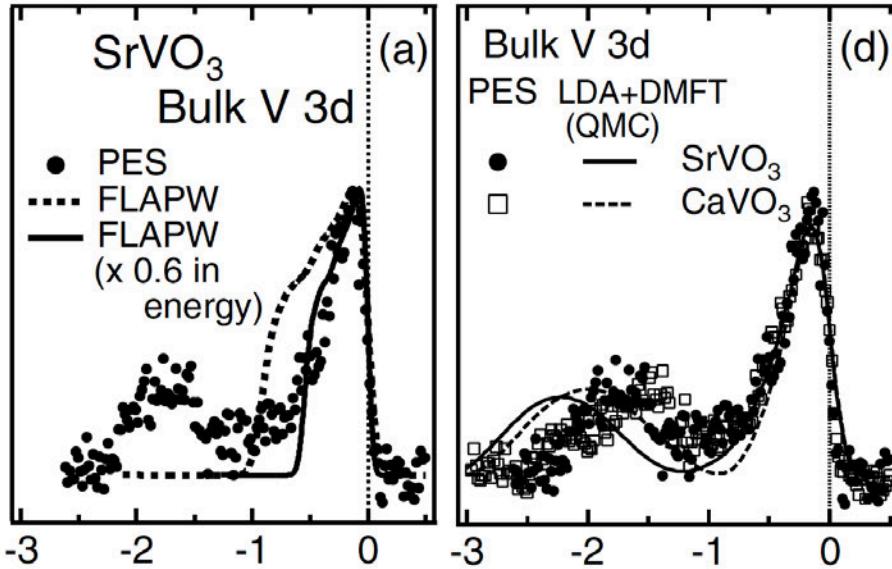


Fig. 2: Photoemission spectra of SrVO_3 and CaVO_3 . Left: comparison to an LDA calculation (dashed curve) and to the LDA spectrum narrowed by a factor 0.6 (plain curve). Right: comparison to an LDA+DMFT calculation. Reproduced from Ref. [8]. See also Ref. [11] for the LDA+DMFT spectrum of SrVO_3 and of the oxides in Fig. 1.

correspondingly diverges at the transition as $m^*/m = 1/Z$. The quasiparticle bandwidth is uniformly reduced by a factor Z .

The Brinkman-Rice description focuses solely on quasiparticles, however, and cannot address at the same time the high-energy part of the excitation spectrum. An important achievement of DMFT is the ability to describe both types of excitations on an equal footing. The actual spectrum of a Mott-correlated metal has three salient spectral features (Fig. 3): lower and upper *Hubbard-bands* at high energy, and a narrowed density of states corresponding to quasiparticle excitations at low energy. This three-peak structure has become some sort of icon of DMFT, and is to a large extent a prediction of the theory [10]. Confirmation and precise comparison to photoemission had to wait, in particular, for a proper identification of the photoemission signal associated with the bulk. As the Mott transition is reached, spectral weight is transferred from the quasiparticles to the Hubbard bands. DMFT allows for a detailed description of these spectral-weight transfers, as coupling or temperature is varied, and this is essential in comparing to experiments. The development of materials-realistic calculations with DMFT (the LDA+DMFT framework, which combines DFT-based electronic structure with a DMFT treatment of many-body correlations) makes quantitative comparison to experiments possible, as displayed in Fig. 2 for SrVO_3 . Importantly, it also allows for both a qualitative and precise answer to questions such as: why are SrVO_3 and CaVO_3 metallic, while LaTiO_3 and YTiO_3 are insulating? The answer [11] is that increased distortions of the structure (tilts of the octahedra) lead to a smaller bandwidth but also, importantly, to an increased splitting between the energy levels within the t_{2g} shell, hence reducing orbital degeneracy and lowering the critical coupling associated with the Mott transition.

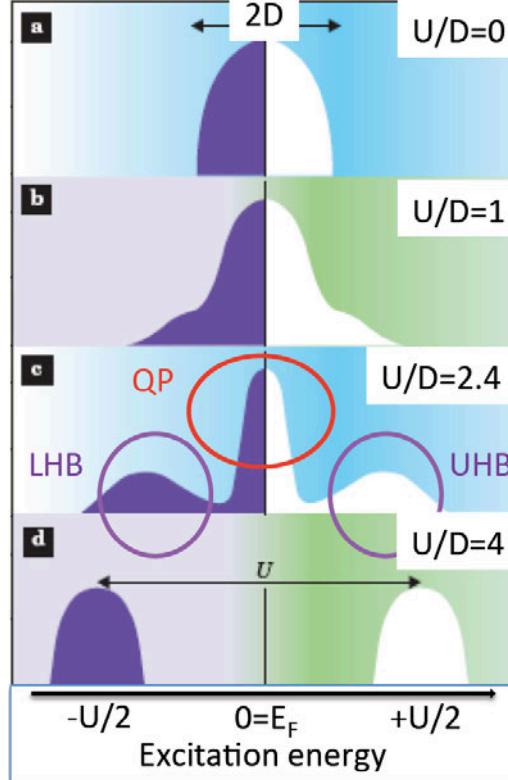


Fig. 3: Schematic evolution of the momentum-integrated spectral function (total density of states) as the coupling is increased, for the half-filled Hubbard model in its paramagnetic phase, according to single-site DMFT. The low-energy (quasiparticle) part of the spectrum gradually narrows down, while the corresponding spectral weight is transferred to the lower and upper Hubbard bands (atomic-like excitations). Adapted from Ref. [12].

2.3 Origin of the Mott phenomenon: blocking of charge, not magnetism

A material with a partially filled band may end up being an insulator because of interactions between electrons; this is the Mott phenomenon. Here, I want to address the following issue: is the cause of this phenomenon related to magnetism? This question has often caused confusion in the field and also provides a key motivation for DMFT.

At strong coupling (when, say, the Hubbard U is larger than the bandwidth $2D$), the answer to this question is quite clear: magnetism is not the cause of the Mott phenomenon. The driving force behind the Mott phenomenon in this regime is the blocking of translational (charge) degrees of freedom. The electrons would have to pay too much in repulsive Coulomb energy (U) to get delocalized in comparison to the potential gain in kinetic energy. In this regime, the Mott insulating gap is of order $\Delta_g \simeq U - 2D \sim U$. Of course, we have to worry about spin physics, but in this regime this physics involves a much smaller energy scale: the superexchange $J \sim D^2/U \ll U, D$. As long as $\Delta_g \gtrsim T \gtrsim J$, the system is in a paramagnetic state with fluctuating local moments signaled by a Curie-Weiss law for the magnetic susceptibility and a large spin-entropy. When the system is cooled down below $T \sim J$, this entropy starts to be quenched out, and the local moments usually order (or in more exotic situations, they might bind into singlets and form a spin liquid state). This is illustrated on Fig. 4. In many oxides,

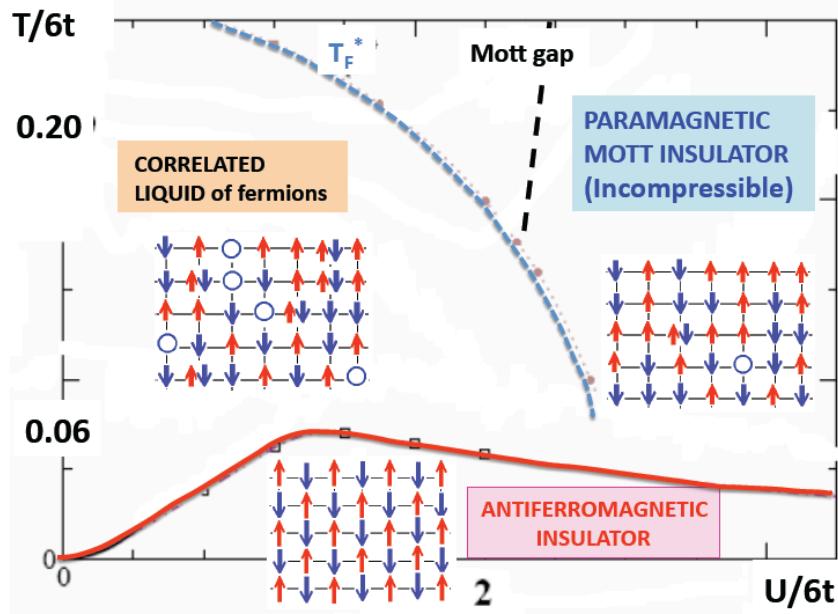


Fig. 4: Phase diagram of the Hubbard model for a three-dimensional cubic lattice with one particle per site on average. The red line denotes the phase transition into a long-range ordered antiferromagnet (Néel temperature). The black dashed line denotes the Mott gap; to the right of this line the paramagnetic phase behaves as an incompressible Mott insulator. The blue dashed line denotes the quasiparticle coherence scale. To the left of this line, the paramagnetic phase behaves as an itinerant fermionic liquid with long-lived quasiparticles. Typical snapshots of the wave-function in real space are displayed for each regime.

long-range spin ordering (and even the onset of spin correlations) occur at a temperature much lower than the insulating gap. For example [1], LaTiO₃ with $\Delta_g \simeq 0.2$ eV orders antiferromagnetically below $T_N \simeq 140$ K, and YTiO₃ with $\Delta_g \simeq 1$ eV orders ferromagnetically below $T_C \simeq 30$ K. Hence, there is an extended regime of temperature in which the system is a disordered insulating paramagnet with clearly insulating properties (apart from a small amount of thermal excitations in the gap). Cuprates are, in this respect, rather exceptional in view of their very high Néel temperature and large superexchange.

The weak-coupling regime is a different story. Magnetic long-range order and the opening of the insulating gap occur simultaneously (Fig. 4). This is the Slater regime. There are actually very few documented examples of antiferromagnetic Slater insulators, a recently investigated one being NaOsO₃ [13] – a rather weakly correlated oxide of a transition metal of the 5d series. This crossover between a weak-coupling (Slater) regime in which the insulating character is linked to magnetism and a strong-coupling (Mott) regime in which they become two completely distinct phenomena is the $U > 0$ analogue of the BCS-BEC crossover that applies to the attractive $U < 0$ case.

In a nutshell: the Mott phenomenon has nothing to do with magnetism at strong coupling. The reason I am emphasizing this is because many theoretical descriptions, such as LDA+U, can only describe Mott insulators by going into the ordered phase with magnetic long-range order. The problem with this is that the magnetism is then the cause of the gap opening, and this is not physically correct. A proper description must account for the fact that the system

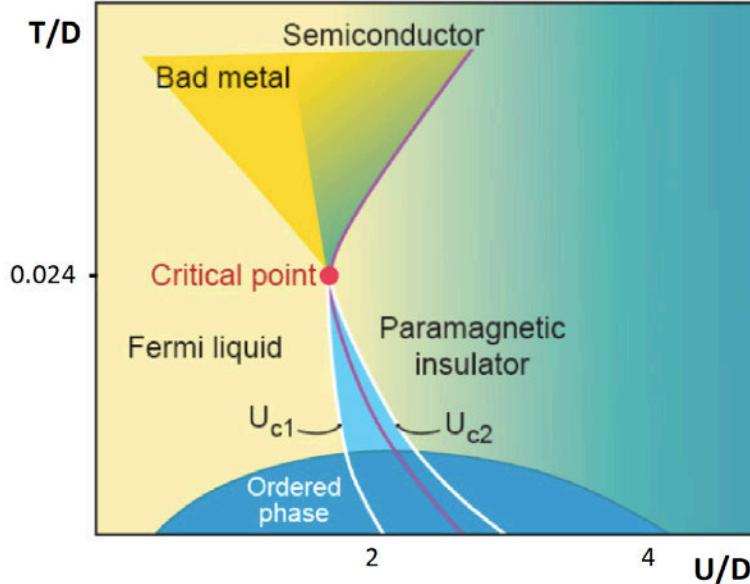


Fig. 5: Generic phase diagram of the half-filled fermionic Hubbard model, as obtained from DMFT. For a lattice with frustration (e.g., with next-nearest neighbour hopping), the transition temperature into phases with long-range spin ordering is reduced. Then, a first-order transition from a metal to a paramagnetic Mott insulator becomes apparent. Adapted from Ref. [17].

is insulating even when the local moments fluctuate, and that this is due to charge blocking. This is very difficult to achieve in any static (energy-independent) mean-field in which the insulating gap develops because of a rigid shift of spin-polarized bands. Instead, a proper theory must handle two widely separated energy scales: the gap in the charge sector and the much smaller superexchange scale in the spin sector. This is a key motivation for DMFT (and LDA+DMFT), which is able to describe a Mott insulator with fluctuating local moments and no broken symmetry in the regime $\Delta_g \gtrsim T \gtrsim J$.

These considerations raise the following question. Imagine one is able to frustrate magnetic long-range order to a large degree, e.g., by considering the Hubbard model on a lattice with next-nearest neighbour hopping or geometric frustration. What would eventually happen at low-temperature to the crossover between the metallic state at small U/D and the paramagnetic Mott insulating state at large U/D ? DMFT answers this question in the following way (Fig. 5): the crossover becomes a first-order transition below a critical temperature T_{MIT} . The point $(T_{\text{MIT}}, U_{\text{MIT}})$ is a second-order (Ising) critical endpoint, and the situation is analogous to the liquid-gas transition between two phases which have the same symmetry. Such a first-order-metal to Mott insulator transition ending in a critical endpoint is observed in several materials, such as V_2O_3 and κ -BEDT organic compounds. In those materials, the transition is always accompanied by a discontinuous change in lattice parameters. In my view however, the driving force behind the transition is clearly of electronic origin, and the lattice just follows. From a theory point of view, however, it is still a somewhat open call whether such a purely electronic first-order transition occurs for finite-dimensional frustrated lattices. Cluster extensions of DMFT (for reviews, see e.g. [14–16]) have provided solid evidence that this is indeed the case, as have recent variational Monte Carlo studies by M. Ogata *et al.*

3 DMFT in a nutshell

3.1 DMFT: solids as self-consistently embedded atoms

How does DMFT manage to simultaneously describe atomic-like excitations at high energy and the formation of long-lived quasiparticles at low energy? The basic concept is illustrated in Fig. 6. The idea is to start from the atom (or from a specific atomic shell) and embed it into an effective medium with which it can exchange electrons [10] (for a review, see e.g. [18]). Focusing for simplicity on the single-level Hubbard atom introduced above, this embedding can be described by the following Hamiltonian (*single-impurity Anderson model*)

$$H_{\text{imp}} = H_{\text{at}} + \sum_{p\sigma} E_p a_{p\sigma}^\dagger a_{p\sigma} + \sum_{p\sigma} (V_p a_{p\sigma}^\dagger d_\sigma + V_p^* d_\sigma^\dagger a_{p\sigma}) . \quad (3)$$

This Hamiltonian was introduced many decades ago in order to describe a magnetic impurity atom embedded into a conduction electron gas [2]. It is important to note that the physics of the impurity atom does not depend on the details of the dispersion relation E_p and of the hybridisation matrix elements V_p . All the relevant information can be condensed in the energy-dependent hybridisation function

$$\Delta(i\omega_n) = \sum_p \frac{|V_p|^2}{i\omega_n - E_p} = \int d\omega \frac{-\text{Im}\Delta(\omega)/\pi}{i\omega_n - \omega} , \quad -\frac{1}{\pi} \text{Im}\Delta(\omega) = \sum_p |V_p|^2 \delta(\omega - E_p) . \quad (4)$$

The second equation expresses Fermi's golden rule: when coupled to the bath, the atomic level is broadened in an energy-dependent way, proportionally to the square of the hybridisation matrix element and to the available density of states in the bath. For path-integral aficionados, one can equivalently state that the physics of the embedded atom is described by the following effective action (obtained by performing the Gaussian integral over the bath degrees of freedom)

$$S_{\text{imp}} = - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_\sigma d_\sigma^\dagger(\tau) \mathcal{G}_0^{-1}(\tau - \tau') d_\sigma(\tau') + U \int_0^\beta d\tau n_\uparrow(\tau) n_\downarrow(\tau) \quad (5)$$

in which

$$\mathcal{G}_0^{-1}(i\omega_n) = i\omega_n - \varepsilon_d - \Delta(i\omega_n) \quad (6)$$

plays the role of an effective bare propagator for this action.

In the DMFT context, the Anderson impurity model is introduced in order to provide a *representation* of the local Green's function of the lattice problem. Denoting by $G(\mathbf{k}, i\omega_n)$ the lattice Green's function, one requires that

$$G_{\text{loc}}(i\omega_n) \equiv \sum_{\mathbf{k}} G(\mathbf{k}, i\omega_n) = G_{\text{imp}}[i\omega_n, \Delta] . \quad (7)$$

This equation should be understood in the following manner. Imagine one knows the exact local Green's function G_{loc} of the lattice model under consideration. One then requires that, when solving the impurity model (5), one obtains an impurity Green's function that coincides

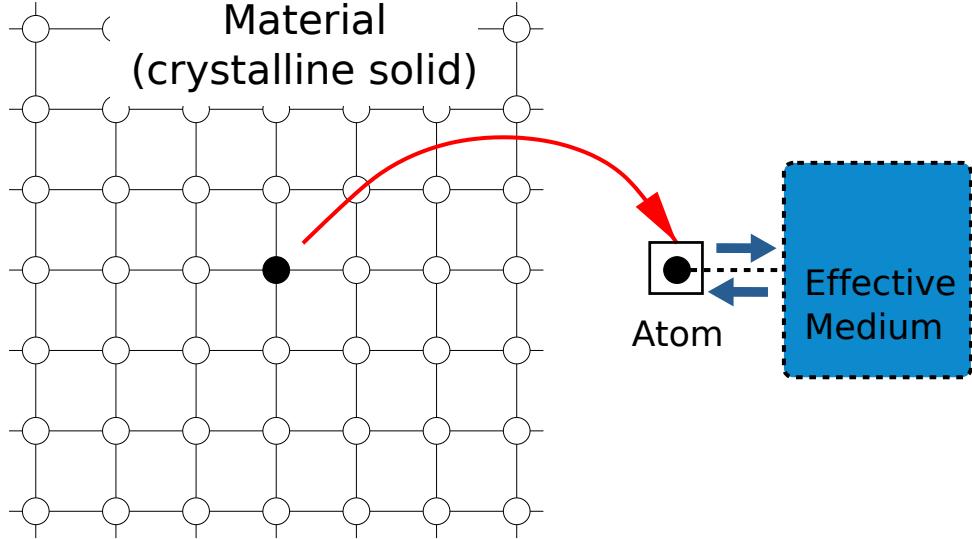


Fig. 6: The Dynamical Mean-Field Theory (DMFT) concept. A solid is viewed as an array of atoms exchanging electrons, rather than as a gas of interacting electrons moving in an periodic potential. DMFT replaces the solid by a single atom exchanging electrons with a self-consistent medium and takes into account local many-body correlations on each site.

with G_{loc} . This can be achieved by a proper choice of the energy-dependent hybridisation $\Delta(\omega)$. Hence, equation (7) should be viewed as a functional equation which determines the function Δ (and ε_d) given a known G_{loc} . Note that in this context, the quantum number p that appears in (3) is merely a label for the energy shell, which is only necessary when one insists on a Hamiltonian form such as (3) for which auxiliary degrees of freedom describing the bath explicitly must be introduced.

At this point, all we have done is to introduce a representation of the exact local Green's function by that of an embedded atom (impurity model). We now introduce an approximation that allows one to obtain a closed set of equations determining both Δ and G_{loc} . To do this, we need to be more specific about the lattice model under consideration. Let us consider the simplest case, that of the single-band Hubbard model. This simply describes a lattice of single-level Hubbard atoms, in which electrons hop from an atom on a given lattice site i to another one on site j with an amplitude t_{ij}

$$H_{\text{Hubbard}} = \sum_i H_{\text{at}}(i) - \sum_{ij} t_{ij} \left(d_{i\sigma}^\dagger d_{j\sigma} + d_{j\sigma}^\dagger d_{i\sigma} \right). \quad (8)$$

The single-electron Green's function of this model can be written as (assuming no translational or spin symmetry-breaking):

$$G(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n + \mu - \varepsilon_{\mathbf{k}} - \Sigma(\mathbf{k}, i\omega_n)} \quad (9)$$

in which μ is the chemical potential (simply related to the atomic level position by $\mu = -\varepsilon_d$) and $\varepsilon_{\mathbf{k}}$ is the dispersion relation of the tight-binding Bloch band (lattice Fourier transform of the hopping t_{ij}).

In (9), $\Sigma(\mathbf{k}, i\omega_n)$ is the single-particle self-energy, which is, in general, a function of both frequency and momentum. The DMFT approximation consists in neglecting this momentum dependence (i.e., ignoring all non-local terms of the self-energy, keeping the local term only) and approximating the local component by that of the impurity model (embedded atom) introduced above. That is, requiring that

$$\Sigma_{ij}(i\omega_n) \simeq \Sigma_{\text{imp}}(i\omega_n) \delta_{ij} , \quad \Sigma_{\text{imp}} \equiv \mathcal{G}_0^{-1} - G_{\text{imp}}^{-1} \quad (10)$$

Using $\mathcal{G}_0^{-1}(i\omega_n) = i\omega_n + \mu - \Delta(i\omega_n)$, this allows us to rewrite (7) in the following form:

$$G_{\text{imp}}[i\omega_n, \Delta] = \sum_{\mathbf{k}} \frac{1}{G_{\text{imp}}[i\omega_n; \Delta]^{-1} + \Delta(i\omega_n) - \varepsilon_{\mathbf{k}}} \quad (11)$$

In a nutshell, the DMFT construction involves the solution of a self-consistent local many-body problem: an atomic shell embedded in a self-consistent medium with which it exchanges electrons. In the simplest case (Hubbard model, no symmetry breaking), the embedded atom is defined by the effective action (5) or equivalently by the Hamiltonian (3). The impurity Green's function G_{imp} , obtained by solving this problem, and the energy-dependent hybridisation function (dynamical mean-field) Δ , which enters its definition, should obey the functional equation (11) (self-consistency condition). These requirements provide enough constraints to determine the two unknown functions G_{imp} and Δ . In practice, this is done by following an iterative scheme, as illustrated in Fig. 7. For this purpose, efficient algorithms must be used to calculate the impurity Green's function, self-energy and possibly two-particle response functions of the embedded atomic shell (*impurity solvers*). Remarkable progress on this front has been achieved in the past few years, thanks to continuous-time quantum Monte Carlo techniques (see [19] for a review). Code libraries are available on the web [20–22].

3.2 When is DMFT exact or accurate?

The single-site DMFT construction becomes exact in the following limits.

- In the atomic limit $t_{ij} = 0$, by construction (then, $\Delta = 0$).
- In the non-interacting limit $U = 0$. Indeed, in this case the self-energy $\Sigma = 0$, so that it is trivially \mathbf{k} -independent.
- Hence, both the limit of a non-interacting band and that of isolated atoms are correctly reproduced by DMFT, which provides an interpolating scheme between these extreme cases.
- In the limit of infinite lattice coordination (infinite number of spatial dimensions), first introduced for fermions in the pioneering work of Metzner and Vollhardt [23]. The hopping must be scaled as $t_{ij} = t/\sqrt{d}$ for this limit to be properly defined and non-trivial.
- Being an exact solution of Hubbard-like models in the limit of infinite dimensions, it is thus guaranteed that DMFT preserves all sum-rules and conservation laws.

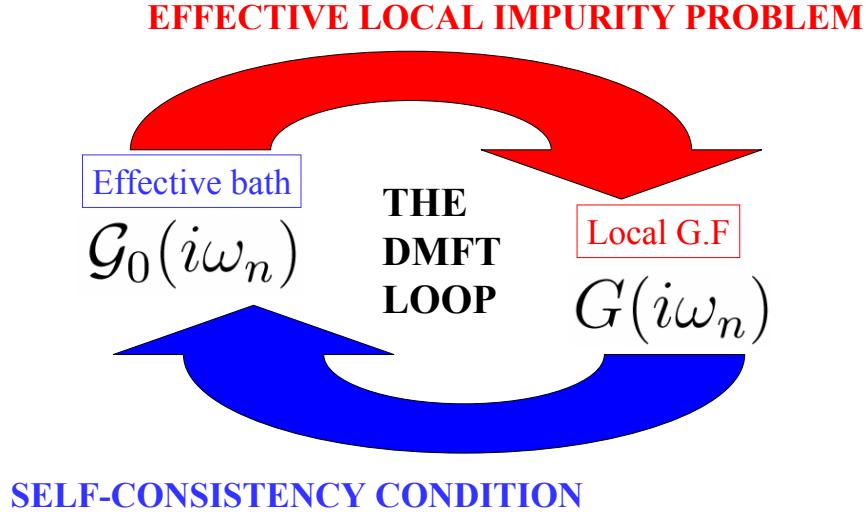


Fig. 7: The DMFT iterative loop. The following procedure is generally used in practice: starting from an initial guess for \mathcal{G}_0 , the impurity Green's function G_{imp} is calculated by using an appropriate solver for the impurity model (top arrow). The impurity self-energy is also calculated from $\Sigma_{imp} = \mathcal{G}_0^{-1}(i\omega_n) - G_{imp}^{-1}(i\omega_n)$. This is used in order to obtain the on-site Green's function of the lattice model by performing a \mathbf{k} -summation (or integration over the free density of states): $G_{loc} = \sum_{\mathbf{k}} [i\omega_n + \mu - \varepsilon_{\mathbf{k}} - \Sigma_{imp}(i\omega_n)]^{-1}$. An updated Weiss function is then obtained as $\mathcal{G}_{0,new}^{-1} = G_{loc}^{-1} + \Sigma_{imp}$, which is injected again into the impurity solver (bottom arrow). The procedure is iterated until convergence is reached.

Besides these formal considerations, it is important to emphasize when single-site DMFT is accurate and physically meaningful. Obviously, this is the case when inter-site correlations do not strongly affect single-particle properties. This is true when the correlation lengths for any kind of incipient ordering are small, i.e., sufficiently far away from critical boundaries. The local approximation (single-site DMFT) is a good starting point when spatial correlations are short-range, which is the case in any of the following regimes: high temperature, high energy, high doping, large number of fluctuating degrees-of-freedom competing with each other, large orbital degeneracy, large degree of frustration. For further considerations along these lines, especially in relation to the high-temperature regime, see [24].

3.3 From particles to waves

As emphasized above, in strongly-correlated materials, electrons are “hesitant” entities with a dual character. At high energy they behave as localized, and the relevant excitations are particle-like atomic excitations. At low energy in metallic compounds, they eventually form wave-like itinerant quasiparticles.

Having introduced the basic concept behind DMFT, we are now in a position to understand how this theory handles this dual nature of excitations, from particle-like at high energy to wave-like at low energy. The key point here is the energy- and temperature-dependence of the dynamical mean-field $\Delta(\omega)$.

At high temperature, $\text{Im}\Delta(\omega)$ has spectral weight mostly at high energy, in the range $\omega \sim -\mu$ and $\omega \sim U - \mu$, corresponding to the lower and upper Hubbard bands. The value of the hybridisation function in this energy range determines the broadening of the Hubbard bands by the solid-state environment. In this high-temperature regime, several atomic states compete with comparable spectral weight. Let us focus for simplicity on the half-filled case. There, the ground-state of the isolated atom is doubly degenerate ($|\uparrow\rangle$ or $|\downarrow\rangle$). At high temperature, the system fluctuates between these two states, leading to a fluctuating local moment when charge excitations are suppressed ($T \lesssim U$).

As temperature is lowered (or as hopping is turned on starting from the isolated atom), the key issue is whether this degeneracy is lifted or not. In the Mott-insulating paramagnetic phase, it is not. The spectral density of the dynamical mean-field $\text{Im}\Delta(\omega)$ self-consistently vanishes within the energy gap, and the local moment is unscreened. In contrast, in the metallic phase $U < U_{\text{BR}}$, $\text{Im}\Delta(\omega)$ is non-zero at low-energy, and grows as the temperature is lowered. This allows the screening of the local moment through the Kondo effect: spin-flip processes involving exchanges of electrons between the atom and the bath become more and more frequent at low energy and low temperature.

DMFT thus describes the formation of quasiparticles as a self-consistent Kondo effect. At low-enough temperature, a local Fermi liquid description applies below a scale T_{FL} , which is the self-consistent Kondo scale. Hence, how quasiparticles form, and most importantly the scale below which they form, depends on how the local atomic multiplet is screened by the solid-state environment. This is why starting from a proper description of the atomic physics and describing how this screening process takes place is essential for understanding strongly correlated metallic phases.

4 How good metals turn bad: quasiparticles beyond Landau theory and spectral weight transfers

Because DMFT is able to describe both the atomic-like excitations and the low-energy quasiparticles, it is also able to describe the full crossover between the Fermi-liquid regime at low temperature and the regime of bad-metallic transport at high temperature. This question has been the subject of several recent works [25, 26].

For $T \lesssim T_{\text{FL}}$, long-lived quasiparticles excitations exist and obey a local version of Landau's Fermi-liquid theory. They give rise to the central peak of the spectral function (Fig. 3), of spectral weight $\sim Z$ and width $\sim ZD$. The quasiparticle spectral weight Z vanishes at the Brinkman-Rice critical point. Because the self-energy is momentum-independent, the quasiparticle effective mass is large $\propto 1/Z$, corresponding to a low-energy Fermi velocity suppressed by Z . The inverse quasiparticle lifetime obeys $Z \text{Im}\Sigma \propto T^2, \omega^2$.

A key observation is that the Fermi-liquid scale T_{FL} is actually much smaller than the width of the quasiparticle peak, i.e., the Brinkman-Rice scale associated with the reduced quasiparticle bandwidth $\sim ZD$. This point was clarified in recent studies [25] and also explained from the

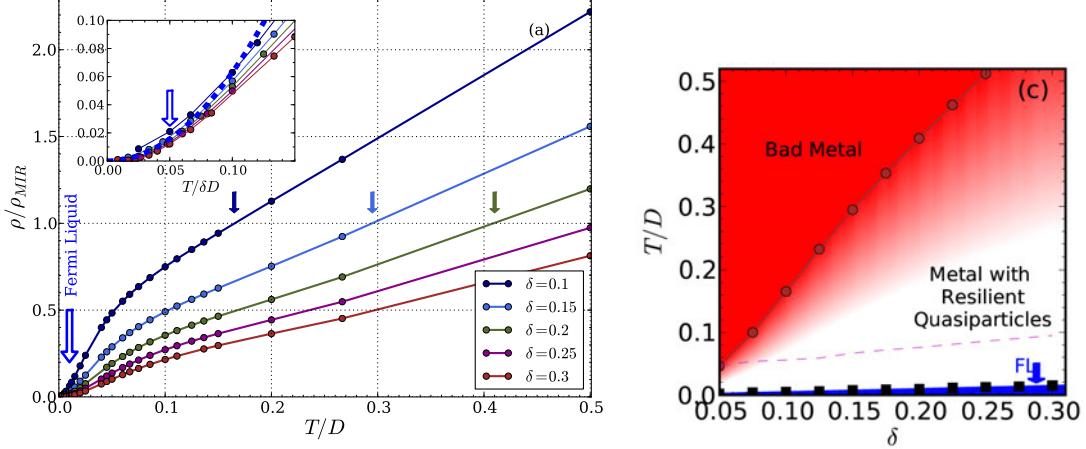


Fig. 8: Left: Temperature-dependence of the resistivity for the single-band Hubbard model ($U/D = 4$), as obtained from DMFT for several values of the hole-doping δ . The plain arrows indicate the temperature at which the Mott-Ioffe-Regel (MIR) value is reached, indicating bad-metal behaviour. Inset: resistivity at low temperatures vs. $T/\delta D$: Fermi liquid T^2 behavior applies up to $T_{FL} \simeq 0.05 \delta D$, indicated by the empty arrow. Right: The different regimes: Fermi liquid (blue), bad metal (red) and intermediate regime with resilient quasiparticles. The crossover into the bad metal is gradual: the onset of red shading corresponds to the optical spectroscopy signatures, while the red points indicate where the MIR value is reached. Reproduced from Ref. [25].

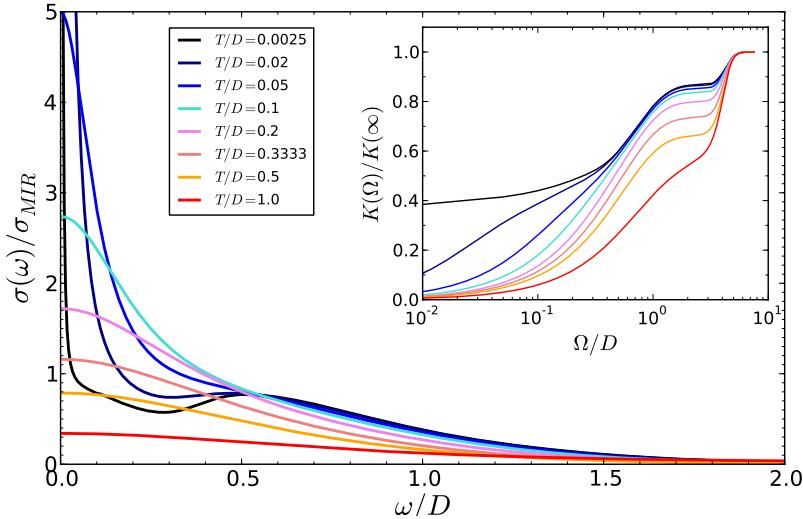


Fig. 9: Optical conductivity of the single-band Hubbard model with $\delta = 20\%$ hole-doping, as obtained from DMFT. Inset: optical spectral weight integrated up to Ω , normalized to the kinetic energy. Reproduced from Ref. [25].

point of view of local Kondo screening, the self-consistent Kondo scale being much smaller than the width of the effective Kondo resonance [27]. This separation of scale leads to three distinct regimes, summarized in Fig. 8: a Fermi-liquid regime with Landau quasiparticles for $T < T_{FL}$, an intermediate regime in which well-defined *resilient* quasiparticle excitations exist for $T_{FL} \lesssim T \lesssim ZD$, and a *bad metal* regime for $T \gtrsim ZD$. Accordingly, the optical conductivity, Fig. 9, displays transfers of spectral-weight that involve the Drude and mid-infrared regions only in the resilient quasiparticle regime, and a much broader energy range in the bad metal regime.

5 Atomic physics in the solid-state: Hund's metals

One of the most striking illustration of the relevance of atomic physics to strongly correlated but quite itinerant metals is to be found in the notion of *Hund's metals*. Indeed, it has been recently recognized through the work of several groups (see e.g. [28–32] and see [33] for a review) that materials that are not directly close to the Mott transition display strong electronic correlations because of the Hund's rule coupling (Fig. 10). The basic mechanism behind this observation is illustrated on Fig. 11, which displays the quasiparticle weight of a multi-orbital Kanamori-Hubbard model. It is seen that (i) the Mott critical coupling is increased by the Hund's coupling and that (ii) in the metallic phase, the Fermi-liquid (Kondo screening) scale is drastically suppressed by the Hund's coupling. These observations apply to a generic integer filling, except when the shell is half-filled or occupied by a single electron or hole. In the *Hund's metal* or *Janus* regime, the local atomic multiplet is difficult to screen, resulting in a low value of T_{FL} and a distinctive non Fermi-liquid behavior for $T > T_{\text{FL}}$. This regime has relevance to transition-metal oxides of the 4d series, as well as to iron superconductors.

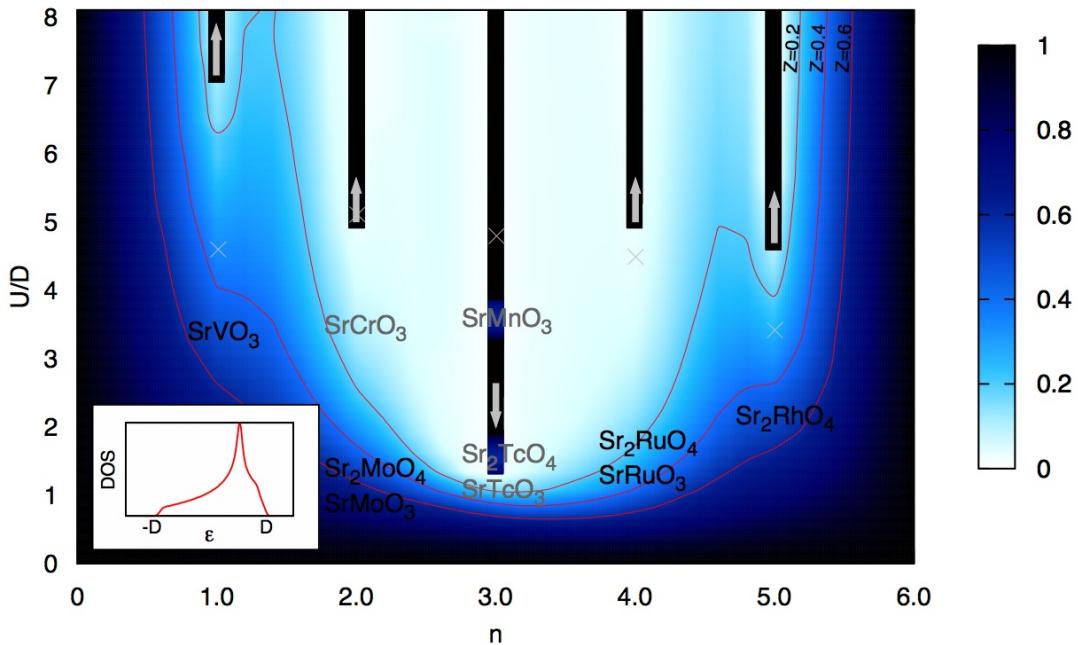


Fig. 10: Colour intensity map of the degree-of-correlation (as measured by the quasiparticle weight Z ; right scale) for a Hubbard-Kanamori model with 3 orbitals appropriate to the description of early transition-metal oxides with a partially occupied t_{2g} shell (bare DOS in inset). The vertical axis is the interaction strength U normalized to the half-bandwidth D ; a finite Hund's coupling $J = 0.15 U$ is taken into account. The horizontal axis is the number of electrons per site, from 0 (empty shell) to 6 (full shell). Darker regions correspond to good metals, lighter to correlated metals. The black bars signal the Mott-insulating phases for $U > U_c$. The arrows indicate the evolution of U_c upon further increasing J , and emphasize the opposite trend between half-filling and a generic filling. Crosses denote the values of U_c for $J = 0$. One notes that, among integer fillings, the case of 2 electrons (2 holes) displays correlated behaviour in an extended range of coupling, with “spin-freezing” above some low coherence scale. Specific materials are schematically placed on the diagram. Adapted from Refs. [32, 33]

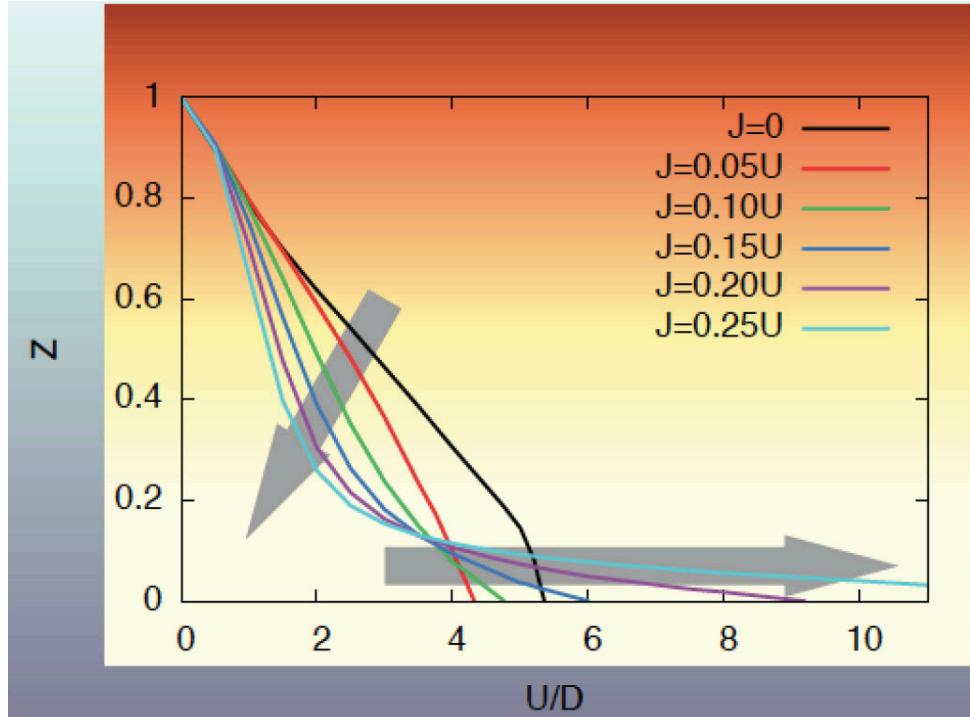


Fig. 11: Quasiparticle weight Z vs. U/D for 2 (or 4) electrons for a three-orbital Hubbard-Kanamori model. The grey arrows indicate the influence of an increasing Hund's rule coupling J/U and emphasize the Janus effect (see text). From Ref. [32]

6 Growing correlations: superexchange, pseudogap, and cluster extensions of DMFT

In spin-1/2 one-band systems with a large nearest-neighbour superexchange, inter-site magnetic correlations are particularly significant, because no orbital fluctuations or frustration compete with spin correlations. As a result, a strong momentum dependence is expected at low doping $\delta \lesssim J/t$. This is indeed what is observed in the normal state of cuprates, in the temperature range $T_c < T < T^*$ where a pseudogap opens up in the antinodal region of the Brillouin zone, while reasonably well-defined quasiparticles survive in the nodal region. As doping is reduced towards the insulator, an increasingly large fraction of the Fermi surface is eaten up by the pseudogap. This is a quite different route to the Mott transition than the uniform reduction of quasiparticle weight following from local theories, and one clearly needs to go beyond single-site DMFT to describe it.

Cluster extensions of DMFT have been quite successful at addressing this problem (for reviews of these approaches, and of the numerous works in the field over the past ten years, see e.g., [14–16]). As for single-site DMFT, those approaches should be viewed as approaching the problem from the high-energy/high-temperature/high-doping side. Very low energies or small doping levels require very good momentum resolution which is hard to reach currently within those approaches. Nevertheless, robust qualitative trends can be established which do

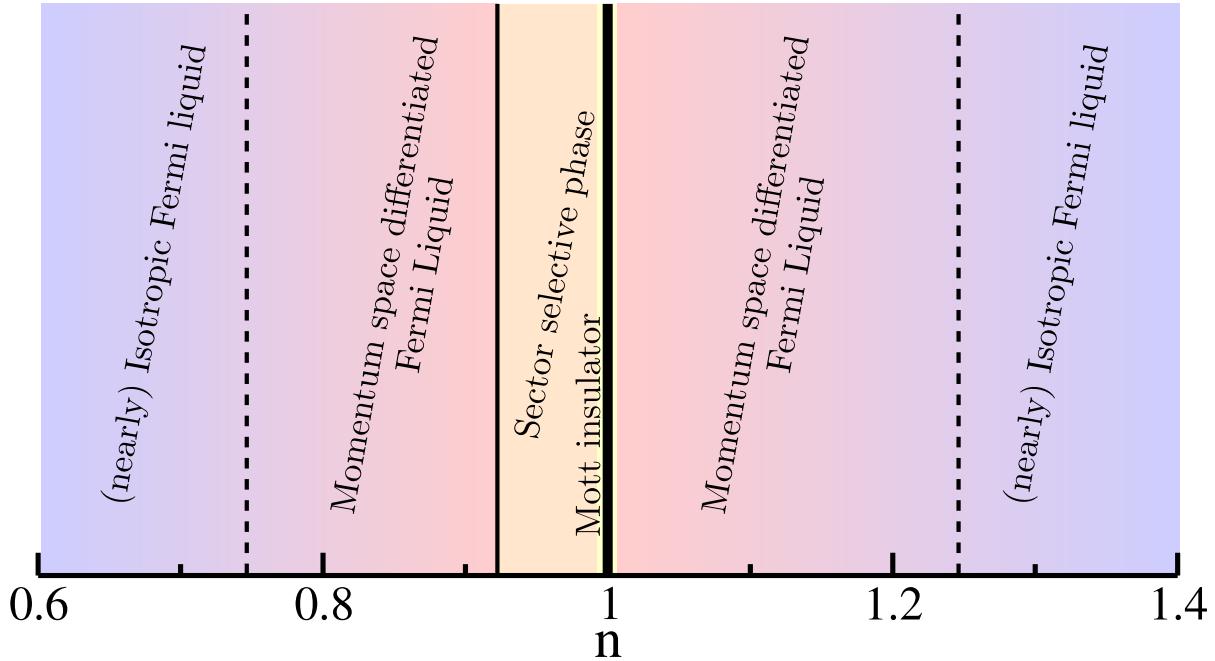


Fig. 12: From weak momentum dependence at high doping to strong momentum selectivity close to half-filling: the various doping regimes of the two-dimensional single-band Hubbard model, as found in cluster extensions of DMFT (reproduced from [34]).

not depend strongly on the specific scheme or cluster size. In [34], a comparative study of different clusters was performed in order to establish these robust qualitative trends which are summarized in Fig.12. At large doping levels, the momentum dependence is weak and single-site DMFT is quite accurate. At intermediate doping, momentum differentiation emerges: on the hole-doped side, antinodal quasiparticles acquire a shorter lifetime than nodal ones. Finally, on the hole doped side, these approaches produce a transition at a critical doping below which the antinodal quasiparticles become gapped. This momentum-selective gapping is the cluster-DMFT description of the pseudogap. It is clearly associated with the physics of the antiferromagnetic superexchange and the formation of inter-site singlets, as can be checked by investigating the statistical weights of the different local states. Hence, at a qualitative level, those approaches seem to support some of the resonating-valence-bond ideas, while extending them considerably by providing a theoretical framework in which the consequence of singlet formation can be studied in an energy- and momentum-dependent way.

7 Hiking down the energy trail: DMFT as a compass

In closing, let me comment on the title of this chapter. A well established and very successful way of thinking about condensed matter is to start from an understanding of the ground-state and, most importantly, of the low-energy excitations built on this ground-state. Crossovers encountered as temperature or energy are increased are then described as fluctuations of the $T = 0$ long-range order and proliferation of low-energy excitations.

DMFT, to a large extent, reverses this perspective, in a manner which is closer in spirit to the renormalization-group approach. The starting point is high-energy: atomic physics. The formation of long-lived coherent excitations is described as an emerging phenomenon as the energy scale is reduced. Crossovers between high-temperature incoherent regimes and low-temperature coherent ones, so important to the physics of strongly correlated materials, are encountered and described along the way.

While hiking down the energy trail, some correlation lengths may grow. Single-site DMFT then becomes insufficient, and cluster extensions of DMFT must be used in order to describe these tendencies to short-range ordering and their effect on quasiparticles. Other techniques may prove necessary as long-wavelength physics becomes relevant. DMFT can in this sense be viewed as a compass when hiking down the energy trail.

References

- [1] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
- [2] A. Hewson: *The Kondo Problem to Heavy Fermions* (Cambridge University Press, 1993)
- [3] A.P. Mackenzie and Y. Maeno, Rev. Mod. Phys. **75**, 657 (2003)
- [4] A. Georges in *Lectures on the physics of highly correlated electron systems VIII* ed. by A. Avella and F. Mancini (American Institute of Physics, 2004) cond-mat/0403123
- [5] A. Fujimori, I. Hase, H. Namatame, Y. Fujishima, Y. Tokura, H. Eisaki, S. Uchida, K. Takegahara, and F.M.F. de Groot, Phys. Rev. Lett. **69**, 1796 (1992)
- [6] S. Sugano, Y. Tanabe, and H. Kamimura: *Multiplets of transition-metal ions in crystals* (Academic Press, 1970)
- [7] K. Maiti, D.D. Sarma, M. Rozenberg, I. Inoue, H. Makino, O. Goto, M. Pedio, and R. Cimino, Europhys. Lett. **55**, 246 (2001)
- [8] A. Sekiyama, H. Fujiwara, S. Imada, S. Suga, H. Eisaki, S.I. Uchida, K. Takegahara, H. Harima, Y. Saitoh, I. A. Nekrasov, G. Keller, D. E. Kondakov, A.V. Kozhevnikov, T. Pruschke, K. Held, D. Vollhardt, and V.I. Anisimov, Phys. Rev. Lett. **93**, 156402 (2004)
- [9] W.F. Brinkman and T.M. Rice, Phys. Rev. B **2**, 4302 (1970)
- [10] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [11] E. Pavarini, S. Biermann, A. Poteryaev, A.I. Lichtenstein, A. Georges, and O.K. Andersen, Phys. Rev. Lett. **92**, 176403 (2004)
- [12] G. Kotliar and D. Vollhardt, Physics Today, March 2004, p. 53
- [13] I.L. Vecchio, A. Perucchi, P. di Pietro, O. Limaj, U. Schade, Y. Sun, M. Arai, K. Yamaura, and S. Lupi, Nature Scientific Reports **3**, 2990 (2013)
- [14] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
- [15] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, and C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
- [16] A. Tremblay, B. Kyung, and D. Sénéchal, Low Temperature Physics **32**, 424 (2006)
- [17] G. Kotliar, Science **302**, 67 (2003)
- [18] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [19] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)

- [20] M. Ferrero and O. Parcollet: *TRIQS: a Toolbox for Research on Interacting Quantum Systems*, <http://ipht.cea.fr/triqs>
- [21] The ALPS project (Algorithms and Libraries for Physics Simulations)
<http://alps.comp-phys.org/>
- [22] K. Haule: *Rutgers DFT and DMFT Materials Database*
<http://hauleweb.rutgers.edu/downloads/>
- [23] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [24] A. Georges, Annalen der Physik **523**, 672 (2011)
- [25] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, Phys. Rev. Lett. **110**, 086401 (2013)
- [26] W. Xu, K. Haule, and G. Kotliar, Phys. Rev. Lett. **111**, 036401 (2013)
- [27] K. Held, R. Peters, and A. Toschi, Phys. Rev. Lett. **110**, 246402 (2013)
- [28] K. Haule and G. Kotliar, New J. Phys. **11**, 025021 (2009)
- [29] P. Werner, E. Gull, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **101**, 166405 (2008)
- [30] J. Mravlje, M. Aichhorn, T. Miyake, K. Haule, G. Kotliar, and A. Georges, Phys. Rev. Lett. **106**, 096401 (2011)
- [31] L. de' Medici, Phys. Rev. B **83**, 205112 (2011)
- [32] L. de' Medici, J. Mravlje, and A. Georges, Phys. Rev. Lett. **107**, 256401 (2011)
- [33] A. Georges, L. de' Medici, and J. Mravlje, Annual Reviews of Condensed Matter Physics **4**, 137 (2013) arXiv:1207.3033
- [34] E. Gull, M. Ferrero, O. Parcollet, A. Georges, and A.J. Millis, Phys. Rev. B **82**, 155101 (2010)

4 Development of the LDA+DMFT Approach

Alexander Lichtenstein
I. Institut für Theoretische Physik
Universität Hamburg

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | Functional approach: from DFT to DMFT | 4 |
| 3 | Local correlations: DMFT and beyond | 7 |
| 4 | Multiorbital quantum impurity solvers | 13 |
| 5 | LDA+DMFT scheme for real materials | 17 |
| 6 | Problem of double counting | 20 |
| 7 | Conclusions | 25 |

1 Introduction

In this lecture, we give an introduction to the theoretical description of strongly correlated materials based on the dynamical mean-field theory (DMFT) and its extensions. The goal of this theoretical construction is to retain the many-body aspects of local atomic physics within the extended solid. The effects of short-range non-local correlations within cluster extensions of the DMFT scheme, as well as long-range fluctuations within the fully renormalized dual-fermion perturbation scheme, are discussed extensively. Recent progress in the numerical solution of the DMFT effective quantum impurity problem within the recently developed continuous-time quantum Monte Carlo schemes is reviewed. We then describe realistic extensions of this approach that combine the accuracy of first-principles density-functional theory (DFT) with the treatment of local many-body effects within DMFT.

Scientific progress in the last century was closely related with the design of silicon based materials for the semiconductor industry. Therefore, the theoretical developments of the last fifty years were associated with realistic electronic structure calculations of such weakly correlated materials. Density-functional theory (DFT) emerged from seminal works of Walter Kohn, Pierre Hohenberg, and Lu Sham [1, 2]. This first-principles scheme is based on an exact theorem, stating that the ground state of interacting electron systems can be found by minimizing a universal functional of the density in some additional external field. The main problem of the DFT scheme is related with the fact that this functional is not known in general and can be calculated numerically with a reasonable accuracy only for the simple case of the homogeneous electron gas. These calculations, which proved to be very useful for the DFT scheme, have been done by David Ceperley and Berni Alder [3] using a two-step quantum Monte Carlo procedure starting from the fixed-node approximation and then releasing the nodal constraint. Nevertheless, the accuracy of such a scheme is still limited and sensitive to the computational details [4]. The main restriction of density-functional theory is that in practice it is restricted to ground state properties, while spectral properties can be found only within the time-dependent DFT scheme [5]. Whereas the structural minimization of complex materials can be carried out very efficiently in the generalized gradient approximation of the DFT, due to the almost spherical properties of the exchange-correlation hole [6], the quality of the spectral properties crucially depends on the systems in question. The TDFT scheme has more problems than the static DFT approach, since there are no suitable time-dependent reference systems to find the exchange-correlation kernel. Understanding the properties of transition-metal systems is key to important materials like high-temperature superconductors, biological molecules like hemoglobin containing iron, and many others. The enormous progress of the last three decades in designing completely new materials for high- T_c superconductivity, giant and colossal magnetoresistance, or artificially created two-dimensional lattices brings new problems to the theory of transition-metal systems. It turns out that even the ground-state properties of antiferromagnetic oxides or orbitally-ordered compounds are not described well in the DFT scheme [7]. The accurate angle-resolved photoemission study of the cuprate superconductors clearly shows that the spectral properties of such systems with strong electron-electron interactions in the $3d$ -shell of transition metals have pro-

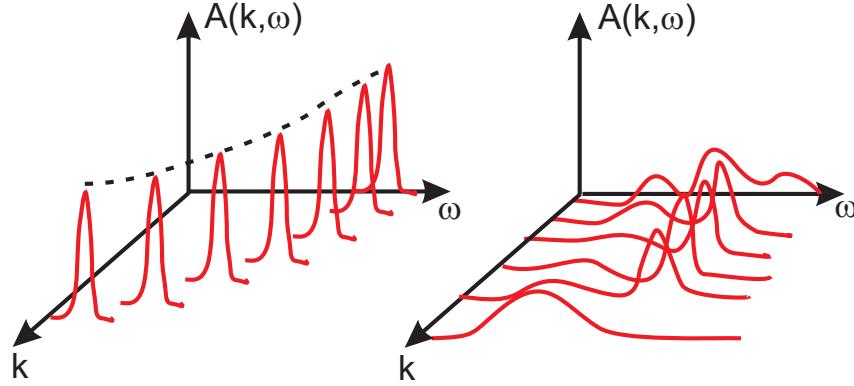


Fig. 1: Schematic view of ARPES spectra for normal (left) and correlated materials (right).

nounced incoherent features [8]. We present in Fig. 1 the qualitative difference between the spectral function of normal metals with well defined quasiparticle peaks at all momenta k and the strongly correlated case with an incoherent part and a non-quasiparticle spectrum in the Brillouin zone. The main source of the complex correlated behavior of electronic systems, related to strong fluctuations between different low-energy fermionic configurations, is shown schematically in Fig. 2. For example, if the free-energy of electronic systems has only one well defined minimum at zero local moment, then one can expect small electron fluctuations and normal paramagnetic quasiparticle behavior. In the case of two low-lying minima corresponding to singlet and triplet excitations, one can expect strong many body fluctuations and possibly non-quasiparticle behavior related to local Hund's rule physics [9]. In order to treat systems with such effective energy profiles, we need to use the path-integral approach and calculate the corresponding correlation functions using complicated quantum Monte Carlo schemes, which can handle many local minima in the free-energy functional on equal footing.

In this lecture, a general functional approach to the strongly correlated electron systems will be discussed. We will separate the local and non-local correlations and show that it is possible to solve the local correlation problem using the recently developed continuous time Monte Carlo (CT-QMC) scheme. Finally, we discuss an efficient way to go from simple model investigation of strongly correlated systems to realistic investigation of complex electronic materials.

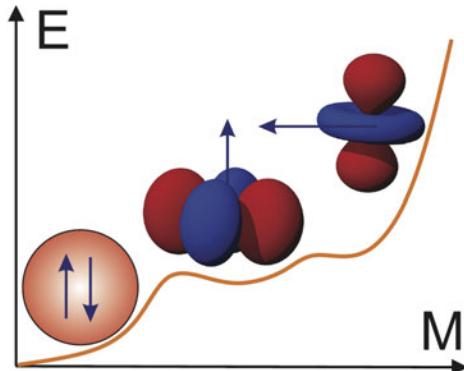


Fig. 2: Schematic representation of magnetic and orbital fluctuations in correlated systems.

2 Functional approach: from DFT to DMFT

We introduce a general functional approach which will cover the DFT, Dynamical Mean Field Theory (DMFT) and Baym-Kadanoff (BK) theories [10]. Let us start from the full many-body Hamiltonian describing electrons moving in a periodic external potential of ions $V(\mathbf{r})$ with chemical potential μ and interacting via the Coulomb law: $U(\mathbf{r} - \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$. We use atomic units $\hbar = m = e = 1$. In the field-operator representation the Hamiltonian has the following form:

$$\begin{aligned} H = & \sum_{\sigma} \int d\mathbf{r} \hat{\psi}_{\sigma}^{\dagger}(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V(\mathbf{r}) - \mu \right] \hat{\psi}_{\sigma}(\mathbf{r}) \\ & + \frac{1}{2} \sum_{\sigma\sigma'} \int d\mathbf{r} \int d\mathbf{r}' \hat{\psi}_{\sigma}^{\dagger}(\mathbf{r}) \hat{\psi}_{\sigma'}^{\dagger}(\mathbf{r}') U(\mathbf{r} - \mathbf{r}') \hat{\psi}_{\sigma'}(\mathbf{r}') \hat{\psi}_{\sigma}(\mathbf{r}). \end{aligned} \quad (1)$$

We can always use a single-particle orthonormal basis set $\phi_n(\mathbf{r})$, for example Wannier orbitals, with a full set of quantum numbers, e.g. site, orbital and spin index: $n = (i, m, \sigma)$ and expand the fields in creation and annihilation operators

$$\begin{aligned} \hat{\psi}(\mathbf{r}) &= \sum_n \phi_n(\mathbf{r}) \hat{c}_n \\ \hat{\psi}^+(\mathbf{r}) &= \sum_n \phi_n^*(\mathbf{r}) \hat{c}_n^+ \end{aligned} \quad (2)$$

Going from fermionic operators to the Grassmann variables $\{c_n^*, c_n\}$, we can write the functional integral representation for the partition function of the many-body Hamiltonian in the imaginary time domain using the Euclidean action S

$$Z = \int \mathcal{D}[c^*, c] e^{-S} \quad (3)$$

$$S = \sum_{12} c_1^* (\partial_{\tau} + t_{12}) c_2 + \frac{1}{2} \sum_{1234} c_1^* c_2^* U_{1234} c_4 c_3, \quad (4)$$

where the one- and two-electron matrix elements are defined as following:

$$\begin{aligned} t_{12} &= \int d\mathbf{r} \phi_1^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V(\mathbf{r}) - \mu \right] \phi_2(\mathbf{r}) \\ U_{1234} &= \int d\mathbf{r} \int d\mathbf{r}' \phi_1^*(\mathbf{r}) \phi_2^*(\mathbf{r}') U(\mathbf{r} - \mathbf{r}') \phi_3(\mathbf{r}) \phi_4(\mathbf{r}') \end{aligned} \quad (5)$$

and we use the following short definition of the sum

$$\sum_1 \dots \equiv \sum_{im} \int d\tau \dots \quad (6)$$

The one-electron Green function is defined via a simple correlation function for fermions

$$G_{12} = -\langle c_1 c_2^* \rangle_S = -\frac{1}{Z} \int \mathcal{D}[c^*, c] c_1 c_2^* \exp(-S) \quad (7)$$

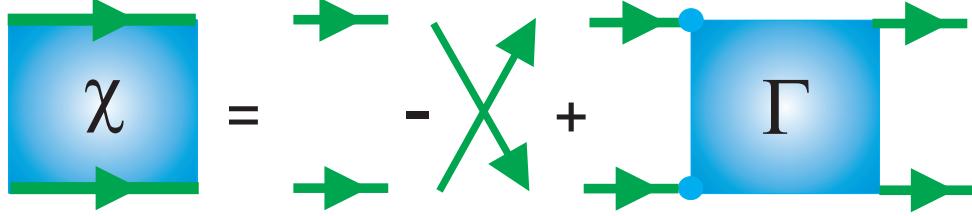


Fig. 3: Representation of full two-particle Green function in terms trivial products of single-particle Green function and the full vertex function Γ .

The main problem of strongly interacting electronic systems is related to the fact that the higher-order correlation functions are not separated into a product of lower-order correlation functions. For example, the two-particle Green function or generalized susceptibility (X) is defined in the following form [11]

$$X_{1234} = \langle c_1 c_2 c_3^* c_4^* \rangle_S = \frac{1}{Z} \int \mathcal{D}[c^*, c] c_1 c_2 c_3^* c_4^* \exp(-S), \quad (8)$$

and can be expressed graphically through the Green functions and the full vertex function Γ_{1234} [12] (see Fig (3))

$$X_{1234} = G_{14}G_{23} - G_{13}G_{24} + \sum_{1'2'3'4'} G_{11'}G_{22'}\Gamma_{1'2'3'4'}G_{3'3}G_{4'4} \quad (9)$$

In the case of non-interacting electron systems, the high-order correlations X are reduced to the antisymmetrized product of lower-order correlations G , which would correspond to the first two terms (Hartree- and Fock-like) with the vertex function Γ in Eq. (9) equal to zero. In strongly correlated electron systems, the part with the vertex is dominant and even diverges close to electronic phase transitions.

The Baym–Kadanoff functional [13] gives the one–particle Green function and the total free energy at its stationary point. In order to construct the exact functional of the Green function (Baym–Kadanoff) we modify the action by introducing the source term J in the following form

$$S[J] = S + \sum_{12} c_1^* J_{12} c_2. \quad (10)$$

The partition function Z , or equivalently the free energy of the system F , becomes a functional of the auxiliary source field

$$Z[J] = e^{-F[J]} = \int \mathcal{D}[c^*, c] e^{-S'[J]}. \quad (11)$$

Variation of this source functional gives all correlation functions, for example the Green function

$$G_{12} = \frac{\delta F[J]}{\delta J_{21}}|_{J=0}. \quad (12)$$

If we use the definition of the generalized susceptibility as a second variation of the $F[J]$ functional instead of $Z[J]$, one will get only the connected part of the X -function, which is represented by the last term in Eq. (9).

The Baym–Kadanoff functional can be obtained via Legendre transform from J to G

$$F[G] = F[J] - \text{Tr}(JG), \quad (13)$$

We can use the standard decomposition of the free energy F into the single-particle part and the correlated part

$$F[G] = \text{Tr} \ln G - \text{Tr}(\Sigma G) + \Phi[G], \quad (14)$$

where Σ_{12} is single-particle self-energy and $\Phi[G]$ is the correlated part of the Baym–Kadanoff functional and is equal to the sum of all two-particle irreducible diagrams. In the stationary point, this functional gives the free energy of the system. One can use different Legendre transforms and obtain functionals of the self-energy Σ [14] or complicated functionals of two variables G and Γ [15] or a more simple functional of G and screened Coulomb interactions W [10] useful in the GW theory.

In practice, $\Phi[G]$ is not known for interacting electron systems, which is similar to the problem of the functional in the density-functional theory. Moreover, this general functional approach reduces to the DFT if one uses the only the space- and time-diagonal part of the Green function, which corresponds to the one-electron density

$$n_1 = G_{12}\delta_{12} = \langle c_1^* c_1 \rangle_S, \quad (15)$$

with the Kohn–Sham potential $V_{KS} = V_{ext} + V_H + V_{xc}$ playing the role of the “constrained field” - J . Here, V_{ext} is the external potential, and V_H is the Hartree potential. In principle, the exchange-correlation potential V_{xc} is known only for the homogeneous electron gas; therefore, in all practical applications, one use a so-called local density approximation to DFT. In this case, the DFT functional is defined in the following way:

$$F_{DFT}[n] = T_0[n] + V_{ext}[n] + V_H[n] + V_{xc}[n], \quad (16)$$

where T_0 is the kinetic energy of the non-interacting system. Finally, if we define the total electron density as

$$n(\mathbf{r}) = \sum_i \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}),$$

the local density approximation to the DFT reads

$$T_0[n] + V_{ext}[n] = \sum_i \int d\mathbf{r} \phi_i^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V_{ext}(\mathbf{r}) - \mu \right] \phi_i(\mathbf{r}) \quad (17)$$

$$V_H[n] = \frac{1}{2} \int d\mathbf{r} n(\mathbf{r}) U(\mathbf{r} - \mathbf{r}') n(\mathbf{r}') \quad (18)$$

$$V_{xc}[n] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon(n(\mathbf{r})), \quad (19)$$

where $\varepsilon(n)$ is exchange-correlation density for the homogeneous electron gas, which can be calculated within a QMC-scheme [3].

In the DFT scheme, we lose information about the non-equal-time Green function, which gives the single-particle excitation spectrum as well as the \mathbf{k} -dependence of the spectral function, and restrict ourself only to the ground state energy of the many-electron system. Moreover, we also lose information about all collective excitations in solids, such as plasmons or magnons, which can be obtained from the generalized susceptibility or from the second variation of the free-energy.

One can probably find the Baym–Kadanoff interacting potential $\Phi[G]$ for simple lattice models using the quantum Monte Carlo scheme (QMC). Unfortunately, due to the sign problem in lattice simulations, this numerically exact solution of electronic correlations is not possible. On the other hand, one can obtain the solution of the local interacting quantum problem in a general fermionic bath using the QMC scheme, which has no sign problem if it is diagonal in spin- and orbital-space. Therefore, a reasonable approach to strongly correlated systems is to keep only the local part of the many-body fluctuations. In such a Dynamical Mean-Field Theory (DMFT), one can obtain numerically the correlated part of the local functional. In this scheme, we only use the local part of the many-electron vertex and obtain in a self-consistent way an effective functional of the local Green function. In the following section we discuss the general dual fermion (DF) transformations [16] that help us to separate the local fluctuations in many-body system and shows a perturbative way to go beyond the DMFT approximations.

3 Local correlations: DMFT and beyond

We will only consider the local but multiorbital interaction vertex $U_{mm'm''m'''}^i$. Sometimes we will omit all orbital indices for simplicity. All equations will be written in matrix form, giving an idea of how to generalize a DF scheme to the multiorbital case [17, 18]. The general strategy to separate the local and non-local correlation effects is associated with the introduction of auxiliary fermionic fields that will couple separated local correlated impurities models back to the lattice [16]. In order to include the smaller non-local part of the Coulomb interactions, one can use a more general approach consists of auxiliary fermionic and bosonic fields [19].

We rewrite the corresponding original action from Eq. (3) in Matsubara space as a sum of the non-local one-electron contribution with t_{12} and the local interaction part U :

$$S[c^*, c] = - \sum_{\omega \mathbf{k} \sigma mm'} c_{\omega \mathbf{k} \sigma m}^* \left[(i\omega + \mu) \mathbf{1} - t_{\mathbf{k} \sigma}^{mm'} \right] c_{\omega \mathbf{k} \sigma m'} + \sum_i S_U[c_i^*, c_i]. \quad (20)$$

where $\omega = (2n + 1)\pi/\beta$, ($\Omega = 2n\pi/\beta$), $n = 0, \pm 1, \dots$ are the fermionic (bosonic) Matsubara frequencies, β is the inverse temperature, μ is the chemical potential. The index i labels the lattice sites, m refers to different orbitals, σ is the spin projection, and the \mathbf{k} -vectors are quasi-momenta. In order to keep the notation simple, it is useful to introduce the combined index $\alpha \equiv \{m, \sigma\}$. Translational invariance is assumed for simplicity in the following, although a real-space formulation is straightforward. The local part of the action S_U may contain any type of local multi-orbital interaction.

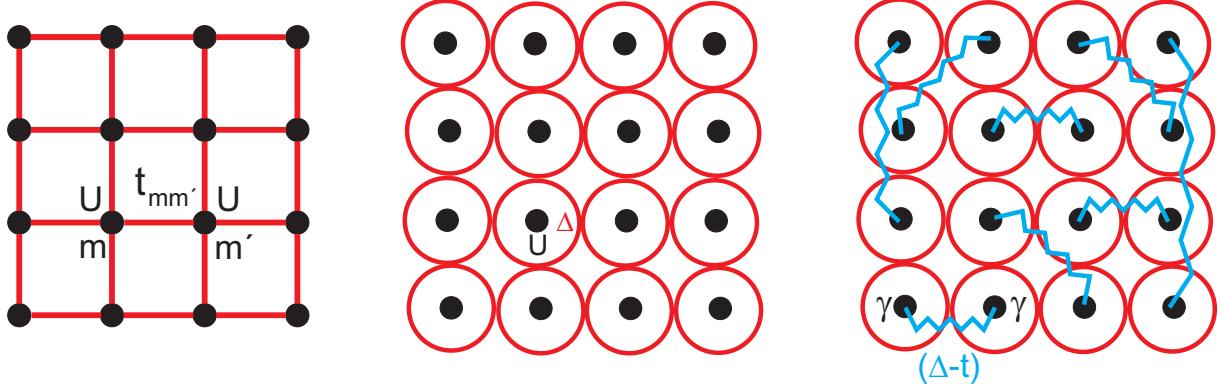


Fig. 4: From the lattice model (left) to the real-space DMFT (middle) following up with the non-local DF perturbation (right).

In order to formulate an expansion around the best possible auxiliary local action, a quantum impurity problem is introduced:

$$S_{\text{loc}}[c^*, c] = - \sum_{\omega \alpha \beta} c_{\omega \alpha}^* [(i\omega + \mu) \mathbf{1} - \Delta_{\omega}^{\alpha \beta}] c_{\omega \beta} + S_U[c^*, c], \quad (21)$$

where Δ_{ω} is the effective hybridization matrix describing the coupling of the impurity to an auxiliary fermionic bath. The main motivation for rewriting the lattice action in terms of a quantum impurity model is that such a reference system can be solved numerically exactly for an arbitrary hybridization function using the CT-QMC method [20]. Using the locality of the hybridization function Δ_{ω} , the lattice action (20) can be rewritten exactly in terms of individual impurity models and the effective one-electron coupling ($t_{ij} - \Delta_{\omega}$) between different impurities:

$$S[c^*, c] = \sum_i S_{\text{loc}}[c_i^*, c_i] + \sum_{\omega \mathbf{k} \alpha \beta} c_{\omega \mathbf{k} \alpha}^* \left(t_{\mathbf{k}}^{\alpha \beta} - \Delta_{\omega}^{\alpha \beta} \right) c_{\omega \mathbf{k} \beta}. \quad (22)$$

We will find the condition for the optimal choice of the hybridization function later. Although we can solve the individual impurity model exactly, the effect of spatial correlations due to the second term in Eq. (22) is very hard to treat even perturbatively, since the impurity action is non-Gaussian and one cannot use the Wick theorem. The main idea of a dual-fermion transformation is to change variables from (c^*, c) to weakly correlated Grassmann fields (f^*, f) in the path-integral representation for the partition function from Eq. (3), followed by a simple perturbation treatment. The new variables were introduced through the following Hubbard-Stratonovich transformation

$$\exp(c_{\alpha}^* b_{\alpha} (M^{-1})_{\alpha \beta} b_{\beta} c_{\beta}) = \frac{1}{\det M} \int \mathcal{D}[f^*, f] \exp(-f_{\alpha}^* M_{\alpha \beta} f_{\beta} - c_{\alpha}^* b_{\alpha} f_{\alpha} - f_{\beta}^* b_{\beta} c_{\beta}). \quad (23)$$

In order to transform the exponential of the bilinear term in (22), we choose the matrices $M_{\alpha \beta}$, and scaling function b_{α} (if we assume for simplicity that the local Green function is diagonal in orbital and spin space) in accordance with Refs. [16] as

$$M = g_{\omega}^{-1} (\Delta_{\omega} - t_{\mathbf{k}})^{-1} g_{\omega}^{-1}, \quad b = g_{\omega}^{-1}, \quad (24)$$

where g_ω is the local, interacting Green function of the impurity problem

$$g_{12} = -\langle c_1 c_2^* \rangle_{\text{loc}} = -\frac{1}{Z_{\text{loc}}} \int \mathcal{D}[c^*, c] c_1 c_2^* \exp \left(-S_{\text{loc}}[c^*, c] \right). \quad (25)$$

With this choice, the lattice action transforms to

$$S[c^*, c, f^*, f] = \sum_i S_{\text{site}}^i + \sum_{\omega \mathbf{k} \alpha \beta} f_{\omega \mathbf{k} \alpha}^* [g_\omega^{-1} (\Delta_\omega - t_{\mathbf{k}})^{-1} g_\omega^{-1}]_{\alpha \beta} f_{\omega \mathbf{k} \beta}. \quad (26)$$

Hence the coupling between sites is transferred to a local coupling to the auxiliary fermions

$$S_{\text{site}}^i[c_i^*, c_i, f_i^*, f_i] = S_{\text{loc}}[c_i^*, c_i] + \sum_{\alpha \beta} f_{\omega i \alpha}^* g_{\omega \alpha \beta}^{-1} c_{\omega i \beta} + c_{\omega i \alpha}^* g_{\omega \alpha \beta}^{-1} f_{\omega i \beta}. \quad (27)$$

Since g_ω is local, the sum over all states labeled by \mathbf{k} could be replaced by the equivalent summation over all sites by a change of basis in the second term. The crucial point is that the coupling to the auxiliary fermions is purely local and S_{site} decomposes into a sum of local terms. The lattice fermions can therefore be integrated out from S_{site} for each site i separately. This completes the change of variables

$$\int \mathcal{D}[c^*, c] \exp(-S_{\text{site}}[c_i^*, c_i, f_i^*, f_i]) = \mathcal{Z}_{\text{loc}} \exp \left(- \sum_{\omega \alpha \beta} f_{\omega i \alpha}^* g_{\omega \alpha \beta}^{-1} f_{\omega i \beta} - V_i[f_i^*, f_i] \right). \quad (28)$$

The above equation may be viewed as the defining equation for the dual potential $V[f^*, f]$. The choice of matrices (24) ensures a particularly simple form of this potential. An explicit expression is found by expanding both sides of Eq. (28) and equating the resulting expressions by order. Formally, this can be done to all orders, and in this sense the transformation to the dual fermions is exact. For most applications, the dual potential is approximated by the first non-trivial interaction vertex

$$V[f^*, f] = \frac{1}{4} \gamma_{1234} f_1^* f_2^* f_4 f_3, \quad (29)$$

where the combined index $1 \equiv \{\omega \alpha\}$ comprises frequency, spin, and orbital degrees of freedom. γ is the exact, fully antisymmetric, reducible two-particle vertex of the local quantum impurity problem. It is given by

$$\gamma_{1234} = g_{11'}^{-1} g_{22'}^{-1} [\chi_{1'2'3'4'} - \chi_{1'2'3'4'}^0] g_{3'3}^{-1} g_{4'4}^{-1}, \quad (30)$$

with the two-particle Green function of the impurity being defined as

$$\chi_{1234} = \langle c_1 c_2 c_3^* c_4^* \rangle_{\text{loc}} = \frac{1}{Z_{\text{loc}}} \int \mathcal{D}[c^*, c] c_1 c_2 c_3^* c_4^* \exp \left(-S_{\text{loc}}[c^*, c] \right). \quad (31)$$

The disconnected part reads

$$\chi_{1234}^0 = g_{14} g_{23} - g_{13} g_{24}. \quad (32)$$

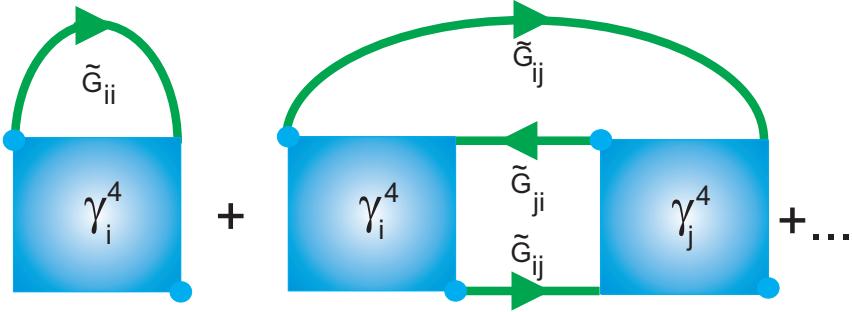


Fig. 5: Diagrams contributing to the dual self-energy $\tilde{\Sigma}$.

The single- and two-particle Green functions can be calculated using the CT-QMC Monte Carlo algorithms [20]. After integrating out the lattice fermions, the dual action depends on the new variables only and reads

$$\tilde{S}[f^*, f] = - \sum_{\omega \mathbf{k} \alpha \beta} f_{\omega \mathbf{k} \alpha}^* [\tilde{G}_\omega^0(\mathbf{k})]_{\alpha \beta}^{-1} f_{\omega \mathbf{k} \beta} + \sum_i V_i[f_i^*, f_i]. \quad (33)$$

and the bare dual Green function is found to be

$$\tilde{G}_\omega^0(\mathbf{k}) = [g_\omega^{-1} + \Delta_\omega - t_\mathbf{k}]^{-1} - g_\omega, \quad (34)$$

which involves the local Green function g_ω of the impurity model.

Up to now, Eqs. (33), (34) are merely a reformulation of the original problem. In practice, approximate solutions are constructed by treating the dual problem perturbatively. Several diagrams that contribute to the dual self-energy are shown in Fig. 5. These are constructed from the impurity vertices and dual Green functions as lines. The first diagram is purely local, while higher orders contain nonlocal contributions, e.g., the second diagram in Fig. 5. In practice, approximations to the self-energy are constructed in terms of skeleton diagrams. The lines shown in Fig. 5 are therefore understood to be fully dressed propagators. The use of skeleton diagrams is necessary to ensure that the resulting theory is conserving in the Baym-Kadanoff sense [13], i.e., that it fulfills the basic conservation laws for energy, momentum, spin, and particle number. The most useful property of such a dual perturbation theory is good convergence both in the weak-coupling limit, when the local vertex is small, and in the strong-coupling limit, when the dual Green function is small [21].

Finally, we can understand the general dual-fermion scheme (Fig. 6) as a two-step process for \mathbf{k} -dependent self-energy. First, we need to find an optimal hybridization function Δ_ω , which defines an effective impurity model. Using a numerically exact Monte-Carlo impurity solver, we can obtain the local Green function g_ω which, together with the hopping parameters, defines the non-local dual Green function $\tilde{G}_\omega^0(\mathbf{k})$ and the interaction vertex $\gamma_{\omega, \omega'}^\omega$, which can be used in renormalized dual perturbation theory [16].

The hybridization function Δ , which so far has not been specified, allows us to optimize the starting point of the perturbation theory and should be chosen in an optimal way. The condition that the first diagram (Fig. 5), as well as all local diagrams with higher order correlation functions in the expansion of the dual self-energy, must be equal to zero at all frequencies fixes the

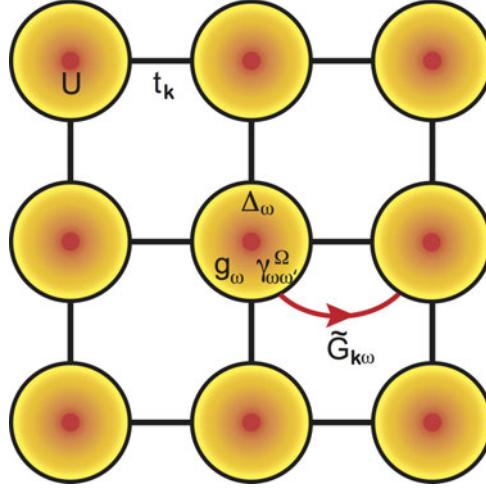


Fig. 6: General view on dual-fermion approach: effective impurity model defined by hybridization function Δ_ω . It can be exactly solved within CT-QMC, resulting in single-particle Green function g_ω and full connected vertex $\gamma_{\omega,\omega'}^\Omega$ with two fermionic (ω) and one bosonic (Ω) Matsubara frequencies. Based on this local information, one can perform an efficient lattice perturbation expansion for the dual Green function $\tilde{G}_\omega^0(\mathbf{k})$.

hybridization. This eliminates the leading-order diagrammatic correction to the self-energy and establishes a connection to DMFT, which can be seen as follows: since the γ vertex is local, this condition amounts to demanding that the local part of the dual Green function be zero:

$$\sum_{\mathbf{k}} \tilde{G}_\omega(\mathbf{k}) = 0. \quad (35)$$

The simplest nontrivial approximation is obtained by taking the leading-order correction, the first diagram in Fig. 5, evaluated with the bare dual propagator (34). Using the expression for the DMFT Green function [22]:

$$G_\omega^{\text{DMFT}}(\mathbf{k}) = [g_\omega^{-1} + \Delta_\omega - t_{\mathbf{k}}]^{-1}, \quad (36)$$

it immediately follows that (35) evaluated with the bare dual Green function is exactly equivalent to the DMFT self-consistency condition for Δ_ω :

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} G_\omega^{\text{DMFT}}(\mathbf{k}) = g_\omega. \quad (37)$$

In the limit of infinitely large lattice connectivity, the DMFT scheme becomes exact with local self-energy [23]. The DMFT approximation for real lattice models appears to be one of the most successful many-body schemes for realistic multi-orbital systems [10]. Since it involves the exact solution of the many-body multi-orbital impurity model Eq. (25), all local quantum fluctuations of different orbitals, spins, and charges (Fig. (7)) are included in this scheme.

In the DMFT approach, one can study paramagnetic correlated phases of complex crystals with strong spin and orbital fluctuations above transition temperatures of the spin- and orbital-ordered states [24]. Hence, DMFT appears as the zero-order approximation in this approach,

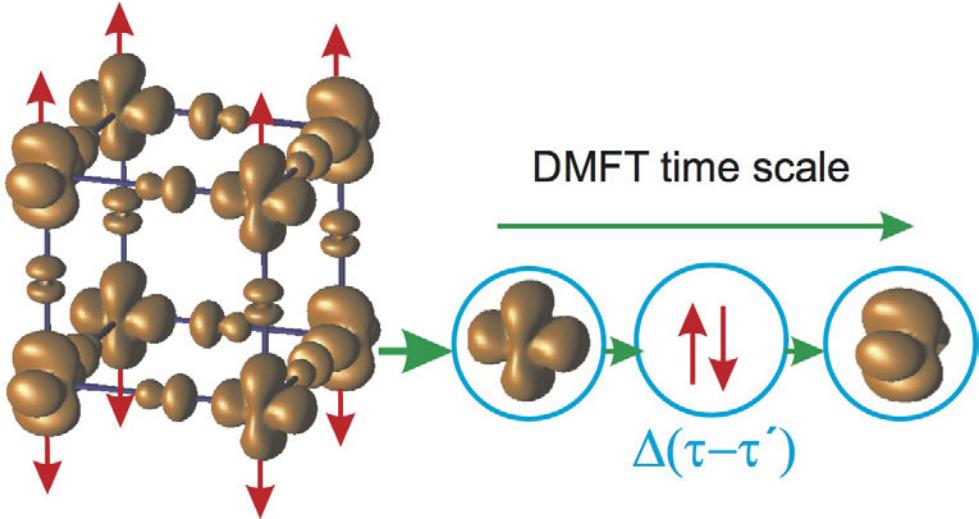


Fig. 7: Schematic representations of initial lattice model (left) and the local DMFT approach with orbital and spin fluctuations.

and corrections to DMFT are included perturbatively. A formal relation to DMFT can be established using the Feynman variational functional approach. In this context, DMFT appears as the optimal approximation to a Gaussian ensemble of dual fermions [25].

When diagrammatic corrections are taken into account and the first diagram is evaluated with the dressed propagator \tilde{G} , the condition (35) will in general be violated. It can be reinforced by adjusting the hybridization function iteratively. This corresponds to eliminating an infinite partial series of all local diagrams starting from the first term in Fig. 5. These contributions are effectively absorbed into the impurity problem. Note that such an expansion is not around DMFT, but rather around an optimized impurity problem.

The only difference between a DMFT and a DF calculation are the diagrammatic corrections that are included into the dual Green function. To this end, the local impurity vertex γ has to be calculated in addition to the Green function in the impurity solver step.

It is an important consequence of the exact transformation (23) that for a theory that is conserving in terms of dual fermions, the result is also conserving in terms of lattice fermions [25]. This allows the construction of general conserving approximations within the dual fermion approach. Numerically, the self-energy is obtained in terms of skeleton diagrams by performing a self-consistent renormalization as described below. Once an approximate dual self-energy is found, the result may be transformed back to a physical result in terms of lattice fermions using exact relations.

The action (33) allows for a Feynman-type diagrammatic expansion in powers of the dual potential V . The rules are similar to those of the antisymmetrized diagrammatic technique [26]. Extension of these rules to include generic n -particle interaction vertices is straightforward. Due to the use of an antisymmetrized interaction, the diagrams acquire a combinatorial prefactor. For a tuple of n equivalent lines, the expression has to be multiplied by a factor $1/n!$. As the simplest example we can write schematically the first self-energy correction of the diagram

in Fig. 5, which contains a single closed loop

$$\tilde{\Sigma}_{12}^{(1)} = -T \sum_{34} \gamma_{1324} \tilde{G}_{43}^{\text{loc}}, \quad (38)$$

where $\tilde{G}^{\text{loc}} = (1/N_k) \sum_{\mathbf{k}} \tilde{G}(\mathbf{k})$ denotes the local part of the dual Green function. The second-order contribution represented in Fig. 5 contains two equivalent lines and one closed loop and hence is \mathbf{k} -dependent

$$\tilde{\Sigma}_{12}^{(2)}(\mathbf{k}) = -\frac{1}{2} \left(\frac{T}{N_k} \right)^2 \sum_{\mathbf{k}_1 \mathbf{k}_2} \sum_{345678} \gamma_{1345} \tilde{G}_{57}(\mathbf{k}_1) \tilde{G}_{83}(\mathbf{k}_2) \tilde{G}_{46}(\mathbf{k} + \mathbf{k}_2 - \mathbf{k}_1) \gamma_{6728} \quad (39)$$

In practice, it is more efficient to evaluate the lowest order diagrams in real space and transform back to reciprocal space using the fast Fourier transform. After calculating the best possible series for the self-energy $\tilde{\Sigma}$ in the dual space, one can calculate the renormalized Green function matrix for original fermions using the following simple transformation [19]

$$G_{\omega}(\mathbf{k}) = \left[\left(g_{\omega} + g_{\omega} \tilde{\Sigma}_{\omega}(\mathbf{k}) g_{\omega} \right)^{-1} + \Delta_{\omega} - t_{\mathbf{k}} \right]^{-1}, \quad (40)$$

which is a useful generalization of the DMFT Green function (see Eq. (36)) to include non-local correlation effects. One can see that the dual self-energy plays the role of an effective T-matrix for the exactly solvable local problem.

The progress of the DMFT approach is strongly related to developments of efficient numerical solvers for an effective quantum impurity model.

4 Multiorbital quantum impurity solvers

Even though DMFT reduces the extended lattice problem to a single-site problem, the solution of the underlying Anderson impurity model remains a formidable quantum many-body problem that requires accurate solvers. Recently, a new class of solvers has emerged, the continuous-time quantum impurity solvers. These are based on stochastic Monte-Carlo methods and mainly come in two different flavors: The weak- and strong-coupling approach.

The weak-coupling or interaction expansion continuous-time (CT-INT) quantum Monte Carlo algorithm for fermions was originally introduced by Aleksei Rubtsov [27]. There are two main previous attempts: first, work by Nikolay Prokof'ev *et al.* [29], who devised a continuous-time scheme to sample the infinite series of Feynman diagrams for bosons, and second, work by Natalie Jachowicz and co-workers [30], who developed a continuous-time lattice Monte Carlo algorithm using the Hubbard-Stratonovich decomposition. The power of the new CT-QMC scheme is that it represents just the integration of the complex path integral without any transformation to effective non-interacting models and can be used for any complicated electron-electron vertex. We introduce the algorithm in the path integral formulation for the single-orbital Anderson impurity problem with a Hubbard-type interaction $U n_{\uparrow} n_{\downarrow}$. The generalization to the multiorbital case can be found in Ref. [20]. First, the action of the Anderson impurity model is divided

into a Gaussian part S_0 and an interaction part S_U as follows

$$S_0 = \sum_{\sigma} \int_0^{\beta} d\tau \int_0^{\beta} d\tau' c_{\sigma}^{*}(\tau) [(\partial_{\tau} - \mu + U\alpha_{-\sigma}(\tau))\delta(\tau - \tau') + \Delta(\tau - \tau')] c_{\sigma}(\tau') , \quad (41)$$

$$S_U = U \int_0^{\beta} d\tau [c_{\uparrow}^{*}(\tau)c_{\uparrow}(\tau) - \alpha_{\uparrow}(\tau)][c_{\downarrow}^{*}(\tau)c_{\downarrow}(\tau) - \alpha_{\downarrow}(\tau)] . \quad (42)$$

The parameters α are introduced to control the sign problem. A formal series expansion for the partition function is obtained by expanding the exponential in the interaction term,

$$\begin{aligned} \mathcal{Z} = & \int \mathcal{D}[c^{*}, c] e^{-S_0[c^{*}, c]} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} U^k \int_0^{\beta} d\tau_1 \dots \int_0^{\beta} d\tau_k [c_{\uparrow}^{*}(\tau_1)c_{\uparrow}(\tau_1) - \alpha_{\uparrow}(\tau_1)] \\ & [c_{\downarrow}^{*}(\tau_1)c_{\downarrow}(\tau_1) - \alpha_{\downarrow}(\tau_1)] \dots [c_{\uparrow}^{*}(\tau_k)c_{\uparrow}(\tau_k) - \alpha_{\uparrow}(\tau_k)][c_{\downarrow}^{*}(\tau_k)c_{\downarrow}(\tau_k) - \alpha_{\downarrow}(\tau_k)] . \end{aligned} \quad (43)$$

Using the definition of the average over the noninteracting action

$$\langle \dots \rangle_0 = \frac{1}{\mathcal{Z}_0} \int \mathcal{D}[c^{*}, c] \dots \exp(-S_0), \quad (44)$$

the partition function can be expressed in the following form

$$\mathcal{Z} = \mathcal{Z}_0 \sum_{k=0}^{\infty} \int_0^{\beta} d\tau_1 \dots \int_{\tau_{k-1}}^{\beta} d\tau_k \text{sgn}(\Omega_k) |\Omega_k| , \quad (45)$$

where the integrand is given by

$$\begin{aligned} \Omega_k = & (-1)^k U^k \langle [c_{\uparrow}^{*}(\tau_1)c_{\uparrow}(\tau_1) - \alpha_{\uparrow}(\tau_1)][c_{\downarrow}^{*}(\tau_1)c_{\downarrow}(\tau_1) - \alpha_{\downarrow}(\tau_1)] \dots \\ & \dots [c_{\uparrow}^{*}(\tau_k)c_{\uparrow}(\tau_k) - \alpha_{\uparrow}(\tau_k)][c_{\downarrow}^{*}(\tau_k)c_{\downarrow}(\tau_k) - \alpha_{\downarrow}(\tau_k)] \rangle_0 . \end{aligned} \quad (46)$$

Note that here the range of time integration has been changed such that time ordering is explicit: $\tau_1 < \dots < \tau_{k-1} < \tau_k$. For a given set of times, all $k!$ permutations of this sequence contribute to Eq. (43). These can be brought into the standard sequence by permuting quadruples of Grassmann numbers, and hence without gaining an additional sign. Since all terms are subject to time-ordering, their contribution to the integral is identical, so that the factor $1/k!$ in Eq. (43) cancels. A configuration can be fully characterized by specifying a perturbation order k and a set of k times: $C_k = \{\tau_1, \dots, \tau_k\}$.

The Monte Carlo algorithm performs importance sampling over this configuration space. The weight of a configuration is thereby taken to be equal to the modulus of the integrand, Eq. (46). Since S_0 is Gaussian, the average over the noninteracting system can be evaluated using Wick's theorem. Hence, the weight of a configuration is essentially given by a fermionic determinant of a matrix containing the bare Green functions

$$\Omega_k = (-1)^k U^k \prod_{\sigma} \det \hat{g}^{\sigma} , \quad (47)$$

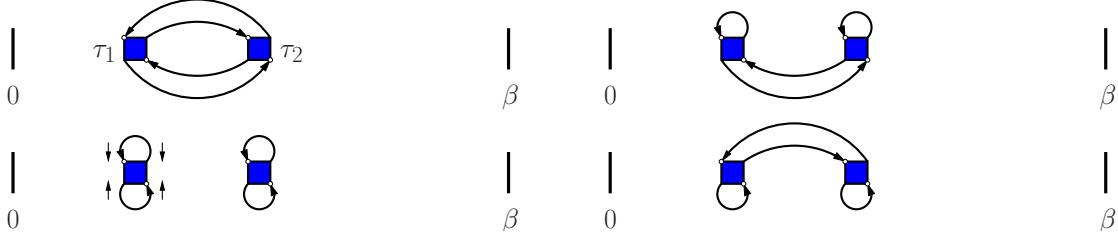


Fig. 8: The four contributions to the partition function for $k = 2$. The interaction vertices are depicted by squares. Bare Green functions are shown as lines.

where the local Green function in the α fields is equal to

$$(\hat{g}^\sigma)_{ij} = g_0^\sigma(\tau_i - \tau_j) - \alpha_\sigma(\tau_i)\delta_{ij}. \quad (48)$$

Note that determinants for different spin orientations factorize since the Green function is diagonal in spin-space.

The hybridization expansion (CT-HYB) or strong-coupling algorithm was initially introduced by Philipp Werner *et al.* [28] and has been generalized to multiorbital systems with general interactions [31, 32]. Here, the algorithm is discussed in the segment representation, which exploits the possibility of a very fast computation of the trace for a density-density type of interaction. The action is regrouped into the atomic part

$$S_{\text{at}} = \int_0^\beta d\tau \sum_\sigma c_\sigma^*(\tau)[\partial_\tau - \mu]c_\sigma(\tau) + U \int_0^\beta d\tau c_\uparrow^*(\tau)c_\uparrow(\tau)c_\downarrow^*(\tau)c_\downarrow(\tau) \quad (49)$$

and the part of the action S_Δ which contains the hybridization term:

$$S_\Delta = \int_0^\beta d\tau' \int_0^\beta d\tau \sum_\sigma c_\sigma^*(\tau)\Delta(\tau - \tau')c_\sigma(\tau'). \quad (50)$$

Here, the sign is taken out by reversing the original order of c and c^* to avoid an alternating sign in the expansion. To simplify the notation, consider first the spinless fermion model, which is obtained by disregarding the spin sums and interaction in Eqs. (49), (50). The series expansion for the partition function is generated by expanding in the hybridization term:

$$\begin{aligned} \mathcal{Z} = & \int \mathcal{D}[c^*, c] e^{-S_{\text{at}}} \sum_k \frac{1}{k!} \int_0^\beta d\tau'_1 \int_0^\beta d\tau_1 \dots \int_0^\beta d\tau'_k \int_0^\beta d\tau_k \times \\ & \times c(\tau_k)c^*(\tau'_k) \dots c(\tau_1)c^*(\tau'_1)\Delta(\tau_1 - \tau'_1) \dots \Delta(\tau_k - \tau'_k). \end{aligned} \quad (51)$$

The important observation now is that, at any order, the diagrams can be collected into a determinant of hybridization functions. The partition function then takes the form

$$\begin{aligned} \mathcal{Z} = & \mathcal{Z}_{\text{at}} \sum_k \int_0^\beta d\tau'_1 \int_{\tau'_1}^\beta d\tau_1 \dots \int_{\tau_{k-1}}^\beta d\tau'_k \int_{\tau'_k}^{\circ\tau'_k} d\tau_k \times \\ & \times \langle c(\tau_k)c^*(\tau'_k) \dots c(\tau_1)c^*(\tau'_1) \rangle_{\text{at}} \det \hat{\Delta}^{(k)}, \end{aligned} \quad (52)$$

where the average is over the states of the atomic problem described by S_{at} . Here $\det \hat{\Delta}^{(k)}$ denotes the determinant of the matrix of hybridizations $\hat{\Delta}_{ij} = \Delta(\tau_i - \tau'_j)$. The diagrams contributing to the partition function for $k = 3$ are shown in Fig. 9. A diagram is depicted by a

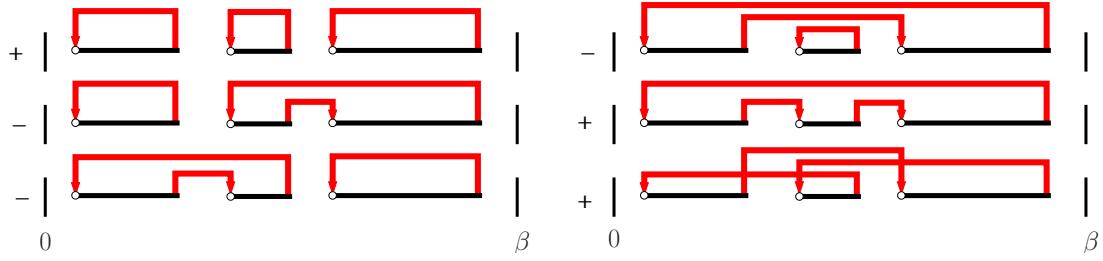


Fig. 9: Diagrammatic representation of the six contributions to the partition function for spinless fermions at $k = 3$. An electron is inserted at the start of a segment (marked by an open circle) and removed at the segment endpoint. The hybridization function lines $\Delta(\tau_i - \tau'_j)$ (shown in red) are connected to the segments in all possible ways. The sign of each diagram is given on the left. Reproduced from Ref. [28].

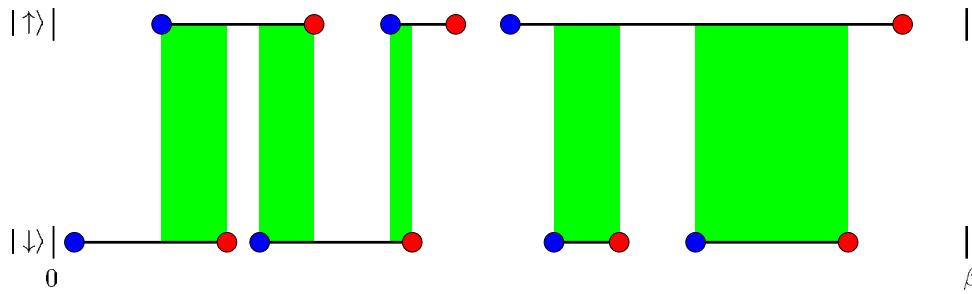


Fig. 10: Example one band CT-HYB in a segment picture: The blue dots illustrate an annihilation operator, the red ones a creation operator and the black line represent the hybridization function $\Delta(\tau_i - \tau'_j)$. The green region represents the time intervals at which two electrons are present on the impurity.

collection of segments, where a segment is symbolic for the time interval where the impurity is occupied. The collection of diagrams obtained by connecting the hybridization lines in all possible ways corresponds to the determinant. Collecting the diagrams into a determinant is essential to alleviate or completely suppress the sign problem. Note that the imaginary time interval in Eq. (52) is viewed as a circle denoted by $\circ\tau_k'$. The trajectories in the path integral are subject to antiperiodic boundary conditions, which is accommodated by an additional sign if a segment winds around the circle.

For the single-orbital Anderson impurity model with Hubbard interaction, the segment picture still holds and gives a very intuitive picture of the imaginary time dynamics. A configuration is visualized by two separate timelines, one for each spin. The additional sum over spins, $\sum_{\sigma_1 \dots \sigma_k}$, which enters in the first line of Eq. (52), generates contributions such as the one shown in Fig. 10. The only difference to the spinless fermion model is that when the impurity is doubly occupied the energy U has to be paid and the trace is $e^{\mu(l_\uparrow + l_\downarrow) - U l_d}$, where l_σ is the time spent on the impurity for an electron with spin σ and l_d is the time the impurity is doubly occupied.

In Fig. 11, we show comparison of the CT-INT and CT-HYB for the strong-coupling case $U \geq W$ of a single band model. The perfect agreement of these two complementary CT-QMC schemes supports the important statement about the possibility of a numerically exact solution of the quantum impurity problem.

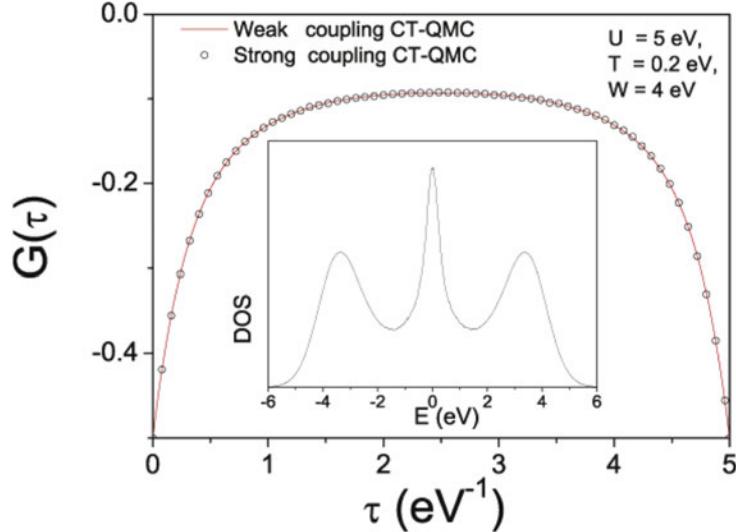


Fig. 11: Comparisson of the weak coupling (CT-INT) and strong coupling (CT-HYB) CT-QMC impurity solvers for one-band semicircular model with $U \geq W$. In the insert, the density of states obtained with maximum entropy scheme is shown.

5 LDA+DMFT scheme for real materials

In order to investigate real correlated systems with the local DMFT scheme, we need to have an efficient scheme of partitioning the space and orbital degrees of freedom. For example, in the high-temperature superconducting oxide $\text{YBa}_2\text{Cu}_3\text{O}_7$, the strongly correlated electrons are in the Cu 3d orbitals, and moreover there is only one per non-equivalent copper $d_{x^2-y^2}$ band that crosses the Fermi level with strong many-body fluctuations, i.e., just a few of the electronic states need to be included in the DMFT calculations. Therefore the simplest realistic correlated scheme would be a DFT+DMFT approach [33, 34] with partitioning of the orbital space into normal band electrons $|K\rangle$ described by the DFT Bloch basis and correlated local orbitals $|L\rangle$ described by some optimal Wannier basis (see Fig. 12 for an illustration).

The treatment of correlated electron systems requires the calculation of Green functions and hybridization functions in terms of local orbitals. This is readily achieved when using a basis set that is localized in real space, such as linear (or N-th order) muffin-tin orbitals (NMTO) [35] or Gaussian basis sets [37]. However, many implementations of density-functional theory use a delocalized plane wave basis set. This has the advantages that the basis set is simple and universal and that its convergence is controlled in principle by a single parameter, the energy cutoff. The projector augmented-wave method (PAW) [38], being representative of plane-wave-based methods, can be used as a simple example of the general projection scheme from the Bloch to the local basis: $\langle K|L\rangle$ (Fig. (12)).

Following the general projection scheme of Ref. [36,37], the desired quantity for the implementation of a DFT+DMFT method is a projection $\mathcal{P}^C = \sum_L |L\rangle \langle L|$ of the full DFT Kohn-Sham Green function $G_{\text{KS}}(\omega)$ on a set of localized orbitals $\{|L\rangle\}$

$$G^C(\omega) = \mathcal{P}^C G^{\text{KS}}(\omega) \mathcal{P}^C. \quad (53)$$

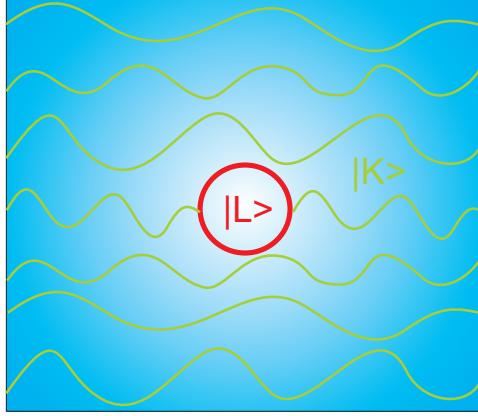


Fig. 12: Schematic representation of the projection scheme from the Bloch basis to a local Wannier correlated subset.

The subspace $\mathcal{C} = \text{span}(\{|L\rangle\})$ is usually called the correlated subspace. It is the subspace of orbitals in which many-body fluctuations play a major role and where the DMFT corrections to the DFT will be considered. In plane-wave-based calculations, $G^{\text{KS}}(\omega)$ in Matsubara space is available in terms of an almost complete set of Bloch states $|K\rangle$ that are eigenstates of the Kohn-Sham Hamiltonian $H_{\text{KS}}|K\rangle = \varepsilon_K|K\rangle$

$$G_{\text{KS}}(\omega) = \sum_K \frac{|K\rangle\langle K|}{i\omega + \mu - \varepsilon_K}. \quad (54)$$

Inserting equation (54) into equation (53) shows that one needs to evaluate projections of the type $\langle L|K\rangle$ in order to access the matrix elements $G_{LL'}^{\mathcal{C}}(\omega)$ of the local Green function. In most cases the correlated orbitals are d or f orbitals, which are localized inside the PAW augmentation spheres to a good approximation. For $|L\rangle$ within these spheres and given the PAW decomposition [38] of a Bloch state $|K\rangle$ one obtains

$$\langle L|K\rangle = \sum_i \langle L|\phi_i\rangle \langle \tilde{p}_i|\tilde{K}\rangle. \quad (55)$$

The index i of the augmentation functions $|\phi_i\rangle$ includes site s , angular momentum l and m as well as an index ν labeling the radial function: $i = (s, l, m, \nu)$, and $|\tilde{p}_i\rangle$ are the projectors of the PAW scheme.

In the described projection-scheme the $|L\rangle\langle L|$ matrix is not properly normalized for two reasons: (1) the Bloch basis is incomplete since only a limited number of Bloch bands is included and (2) the PAW augmentation functions are in general not orthonormal. The simplest way is to orthonormalize the projection matrices by the following Wannier-type construction: by definition, the localized states $|L\rangle$ are labeled by site and angular momentum indices: $L = (s, l, m)$. We split the site index $s = \mathbf{R} + \mathbf{T}$ such that \mathbf{R} labels the position within the unit cell and \mathbf{T} is the Bravais lattice vector of the unit cell in which s is located. This allows us to construct the Bloch transform of the localized states,

$$|L_{\mathbf{k}}\rangle = \sum_T e^{i\mathbf{k}\cdot\mathbf{T}} |L_{\mathbf{T}}\rangle, \quad (56)$$

where \mathbf{k} is from the first Brillouin zone and $|L_T\rangle \equiv |L\rangle = |s, l, m\rangle$. The sum in equation (56) runs over the Bravais lattice. Labeling the Bloch states $|K\rangle = |\mathbf{k}, n\rangle$ by their crystal momentum, \mathbf{k} , and band index, n , we normalize our projection matrices $\mathcal{P}_{Ln}^C(\mathbf{k}) = \langle L_{\mathbf{k}} | \mathbf{k}, n \rangle$ using the overlap operator

$$O_{LL'}(\mathbf{k}) = \sum_n \mathcal{P}_{Ln}^C(\mathbf{k}) \mathcal{P}_{L'n}^{*C}(\mathbf{k}) \quad (57)$$

in

$$\bar{\mathcal{P}}_{Ln}^C(\mathbf{k}) = \sum_{L'} O_{LL'}^{-1/2}(\mathbf{k}) P_{L'n}^C(\mathbf{k}). \quad (58)$$

These orthonormalized projection matrices are calculated once at the beginning of any calculation and can then be used to obtain the local Green function of the correlated orbitals from the full Bloch Green function $G_{nn'}^B$

$$G_{LL'}^C(\omega) = \sum_{\mathbf{k}, nn'} \bar{\mathcal{P}}_{Ln}^C(\mathbf{k}) G_{nn'}^B(\mathbf{k}, \omega) \bar{\mathcal{P}}_{L'n'}^{*C}(\mathbf{k}).$$

Similarly the hybridization function, $\Delta(\omega)$, is available. It is related to the local Green function by

$$G^{-1}(\omega) = i\omega - \epsilon_d - \Delta(\omega), \quad (59)$$

where ϵ_d is the static crystal field. Equation (59) is a matrix equation with G , Δ , and ϵ_d being $(\dim \mathcal{C}) \times (\dim \mathcal{C})$ matrices, in general. To separate the hybridization from the static DFT crystal field, we numerically evaluate the limit $\omega \rightarrow \infty$, where $\omega - G^{-1}(\omega) \rightarrow \epsilon_d$.

In a DFT+DMFT calculation the projection matrices $\bar{\mathcal{P}}_{Ln}^C(\mathbf{k})$ are used for up- and downfolding quantities like the Green function and the self-energy in the course of the iterative DMFT procedure in exactly the same way as shown for the local Green function above. For example, the self energy obtained by an impurity solver for the effective impurity model $\Sigma_{LL'}^C(\omega)$ can be upfolded to the Bloch basis as follows:

$$\Sigma_{nn'}^B(\mathbf{k}, \omega) = \sum_{LL'} \bar{P}_{Ln}^{*C}(\mathbf{k}) \Sigma_{LL'}^C(\omega) \bar{P}_{L'n'}^C(\mathbf{k}).$$

Since the self energy in DMFT is a purely local quantity, the index \mathbf{k} on $\Sigma_{nn'}^B(\mathbf{k}, \omega)$ reflects the momentum dependence brought about by the projection matrices. The presented projection scheme allows for the inclusion of both correlated and uncorrelated states in the procedure. Therefore, information about the interplay of correlated orbitals with their uncorrelated ligands can be obtained.

Figure 13 shows that the DFT+DMFT calculation commences with the solution of the Kohn-Sham equations by DFT. In a second step the projection onto the correlated subset is computed. The Kohn-Sham Green function is then computed and used as an initial guess for the mean-field \mathcal{G} of the DMFT cycle, which consists of the usual steps detailed before. In usual applications the DFT+DMFT loop will stop after DMFT self-consistency is obtained. It has, however, recently become possible to continue the cycle supplying the DFT code with an altered charge density that includes correlation effects. In such a unified approach changes in the charge density

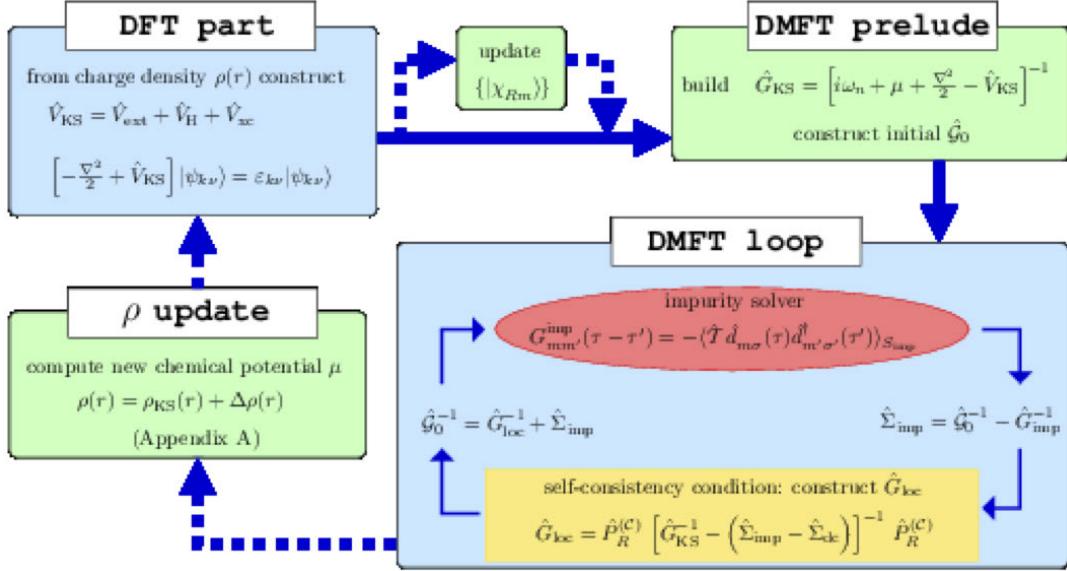


Fig. 13: Illustration of the DFT+DMFT procedure. As a first step, the Kohn-Sham (KS) equations, determining the Kohn-Sham potential and thus the Hamiltonian, are solved. Secondly, the KS Green function and from it the starting value for the bath Green function \mathcal{G} is constructed and passed on to the DMFT loop, which consists of the usual steps described before. A potential self-consistency over the charge density is also indicated [37].

induced by correlations can be studied [37, 40]. It allows furthermore for the accurate calculation of total energies, which allow the determination of crystal structures and other coupled electronic and structural effects.

As example, we show a realistic DFT+DMFT calculation of the SrVO₃ spectral function in the Fig.14, where one can see the renormalisation of the valence correlated V t_{2g} states as well as broadening of the Bloch O 2p states [39].

6 Problem of double counting

We will finally discuss the problem of double-counting corrections in the LDA+DMFT scheme [41]. We use the standard definitions of the parameters U and J

$$U = F^0 \quad \text{and} \quad J = \frac{F^2 + F^4}{14}. \quad (60)$$

Over the years different methods to fix μ_{DC} have been devised. Two main approaches are the around mean-field (AMF) approximation and the fully localized or atomic limit (FLL) [7]. Both methods use analytic arguments to devise a double counting correction. The AMF is based on the conjecture that DFT corresponds to a mean-field solution of the many-body problem. The resulting double counting potential can be written as

$$\mu_{DC}^{AMF} = U(N_{imp} - \bar{n}) - J(N_{imp}^\sigma - \bar{n}), \quad (61)$$

where N_{imp} is the total occupancy of the impurity, N_{imp}^σ the occupancy per spin ($N_{imp}^\sigma = N_{imp}/2$ for the paramagnetic case), and $\bar{n} = \frac{1}{2(2l+1)} \sum_{m,\sigma} n_{m\sigma}$ is the average occupancy. The AMF

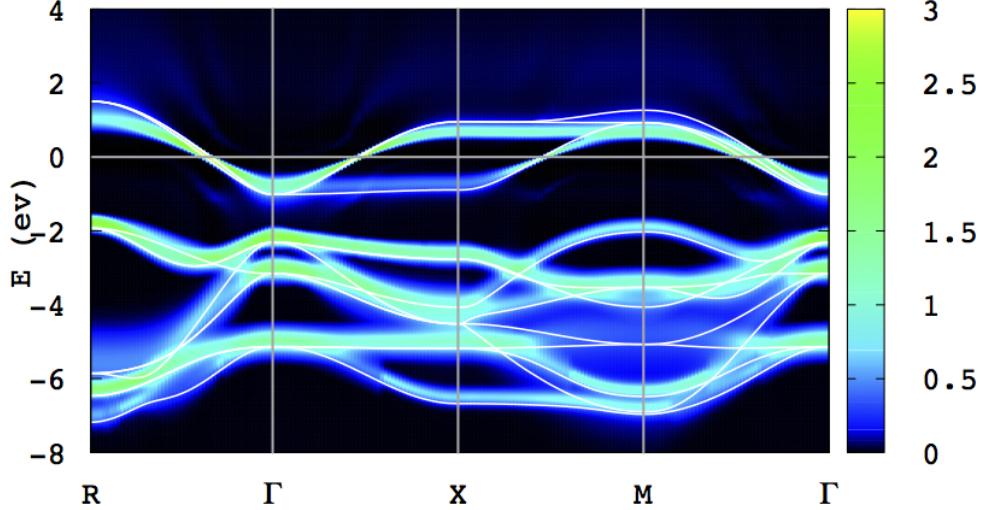


Fig. 14: Momentum resolved impurity spectral function of SrVO_3 obtained by DFT+DMFT. The LDA band structure of the $\text{V } t_{2g}$ and $\text{O } 2p$ Bloch states is shown for comparison.

functional is known to produce unsatisfactory results for strongly correlated systems, which led to the development of another method, the so-called FLL. The FLL functional takes the converse approach to the AMF and begins with the atomic limit. It has been shown that this new potential can also be written as a correction to the AMF solution [7], Eq. (61),

$$\mu_{\text{DC}}^{\text{FLL}} = U \left(N_{\text{imp}} - \frac{1}{2} \right) - J \left(N_{\text{imp}}^{\sigma} - \frac{1}{2} \right) = \mu_{\text{DC}}^{\text{AMF}} + (U - J) \left(\bar{n} - \frac{1}{2} \right).$$

This addition to the AMF potential has the effect of a shift of the centroid of the level depending on its occupation. An empty level is raised in energy by $\frac{1}{2}(U - J)$ and the converse happens to a fully occupied level. The form of the functional is based on the property of the *exact* density-functional that the one electron potential should jump discontinuously at integer electron number [7], a property which is not fulfilled in LDA or GGA. Ultimately the FLL leads to a stronger trend towards integer occupancies and localization. The general problem with analytic expressions like the ones presented above is that their scope is limited to certain classes of systems that fulfill the assumptions made in the derivation process. As an additional complication, both approaches can be used employing the average orbital occupancies obtained from DFT or in a self-consistent manner allowing the occupancies to be determined within the DFT+DMFT loop.

The obvious problems with analytical formulae make conceptually different approaches worth exploring. It would certainly be an improvement if the double counting could be found self-consistently along with the chemical potential in the DMFT self-consistency loop. One possible ansatz using the impurity self-energy $\Sigma_{mm'}^{\text{imp}}$ is to constrain the high energy tails in the real part of the self-energy to sum up to zero

$$\text{Re}[\text{Tr}(\Sigma_{mm'}^{\text{imp}}(i\omega_N))] = 0. \quad (62)$$

Here, ω_N is the highest Matsubara frequency included in the calculation. Physically this amounts to the requirement that the shift in the centroid of the impurity orbitals contains no static com-

ponent. Since the self-energy is a quantity computed self-consistently via the quantum Monte Carlo in our case, one has to ensure that a reasonably high number of Matsubara frequencies is included in the calculation and additionally that the Monte Carlo data is well converged. Double-counting corrections based on the self-energy have been applied successfully to metallic systems [42].

Another possible approach is to constrain the total charge in the impurity based on the Friedel sum rule [12]. The Friedel sum rule gives a zero-temperature relationship between the extra states induced below the Fermi level by a scattering center (an impurity) and the phase shift at the chemical potential. For the Anderson model the extra states induced are given by the occupation number of the impurity states, and the scattering potential is the hybridization that affects the conduction electrons. The simplest version requires the charges in the self-consistent bath and the impurity to be equal

$$\text{Tr } G_{mm'}^{\text{imp}}(\beta) \stackrel{!}{=} \text{Tr } \mathcal{G}_{mm'}(\beta). \quad (63)$$

The value of the imaginary time Green function at $\tau = \beta$ gives the orbital occupancy, thus the trace over them amounts to the total occupancy of the impurity. Both versions of the method give very similar results and work well in metallic systems, since in a metal the total particle number of the system N_{tot} and of the impurity N_{imp} are both very sensitive to small variations in μ and μ_{DC} . Also the likeness to the Friedel sum rule, which applies to metals, indicates that such a constraint will work for metals only. As NiO has a quite large gap, the charge is almost invariant with respect to the chemical and the double-counting potential in the gap. Over a region of several electron volts

$$\frac{\delta N_{\text{tot}}}{\delta \mu} \approx 0.$$

We will see below the consequences for this specific double-counting prescription and for double-countings in general that arise from the presence of the insulating gap.

Now that we have introduced the double-counting in the context of DFT+DMFT and discussed ways to fix the underlying parameter, we can elucidate our test case: NiO. The double-counting potential has been treated here as an adjustable parameter and has been varied between 55 eV and 60 eV. The most prominent effects of the double counting on the spectral properties are the shift of the oxygen $2p$ bands with respect to the nickel $3d$ bands, as well as the variation in gap size. The μ_{DC} controls the filling of the Ni $3d$ orbitals and since the total number of particles in the system is fixed at $N_{\text{tot}} = 14$, the chemical potential μ of the full system must be varied together with μ_{DC} . The filling of the Ni $3d$ orbitals can be varied from $N_{\text{imp}} \approx 8.1$ at $\mu_{\text{DC}} = 55$ eV to $N_{\text{imp}} \approx 8.3$ at $\mu_{\text{DC}} = 60$ eV. The double counting potential μ_{DC} has profound impact on the spectrum $A_m(\omega) = -\frac{1}{\pi}G_m(\omega)$ shown in Fig. 15 and the \mathbf{k} -resolved spectral function

$$A_m(\mathbf{k}, \omega) = -\frac{1}{\pi} \text{Im} \left(\omega + \mu - \varepsilon_m(\mathbf{k}) - \Sigma_m^{\text{imp}}(\mathbf{k}, \omega) \right)^{-1}$$

shown along the line $\Gamma - X$ in the Brillouin zone in Fig. 16. All spectral functions were obtained by the maximum entropy method from imaginary-time Green functions. With increasing μ_{DC} ,

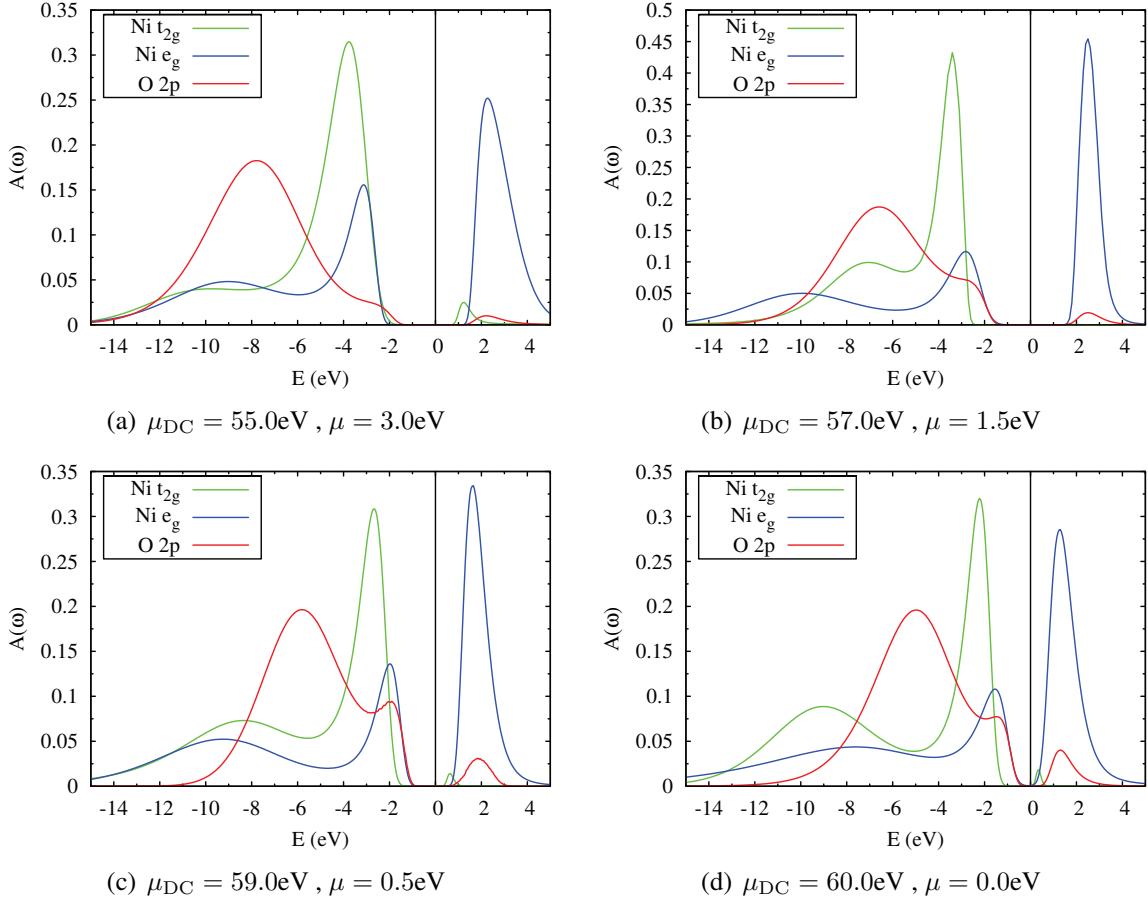


Fig. 15: LDA+DMFT spectral functions at $\beta = 5\text{eV}^{-1}$ for different values of the double counting correction μ_{DC} , illustrating its impact on the spectrum [41].

the system evolves from a large gap Mott insulator at $\mu_{DC} = 55$ eV (Fig. 15 (a)), with a gap of about 4 eV opening between the Ni 3d states and almost no oxygen at the valence band edge, towards an almost closed-gap at $\mu_{DC} = 60$ eV (Fig. 15 (d)). Plainly speaking, the double-counting correction allows for a tuning of the spectral properties from a large-gap Mott insulator in the region $\mu_{DC} \leq 55$ eV to a metal at $\mu_{DC} > 60$ eV. The regime of the charge transfer insulator, the expected physical state of NiO, lies somewhere in between at $\mu_{DC} \sim 59$ eV (Fig. 15 (c)). In this state the oxygen states make up about half the spectral weight at the valence band edge, as observed in experiments [43, 44]. The gap, however, is much smaller than 4 eV, thus if the only criterion of the quality of the calculation would be the gap, the double-counting of choice would be around $\mu_{DC} = 55$ eV, missing the physics of the system entirely.

Let us now turn to the k-resolved spectral functions shown in Fig. 16 and compare them with ARPES data [45]. We have superimposed the ARPES data on top of our calculated spectral functions; the data were aligned at the valence band edge to facilitate comparison. The uppermost band in Figs. 16 (a) and (b) at 2eV above the Fermi level is a Ni e_g band, while the other bands can be identified with the ones obtained by ARPES. The two lowest lying bands

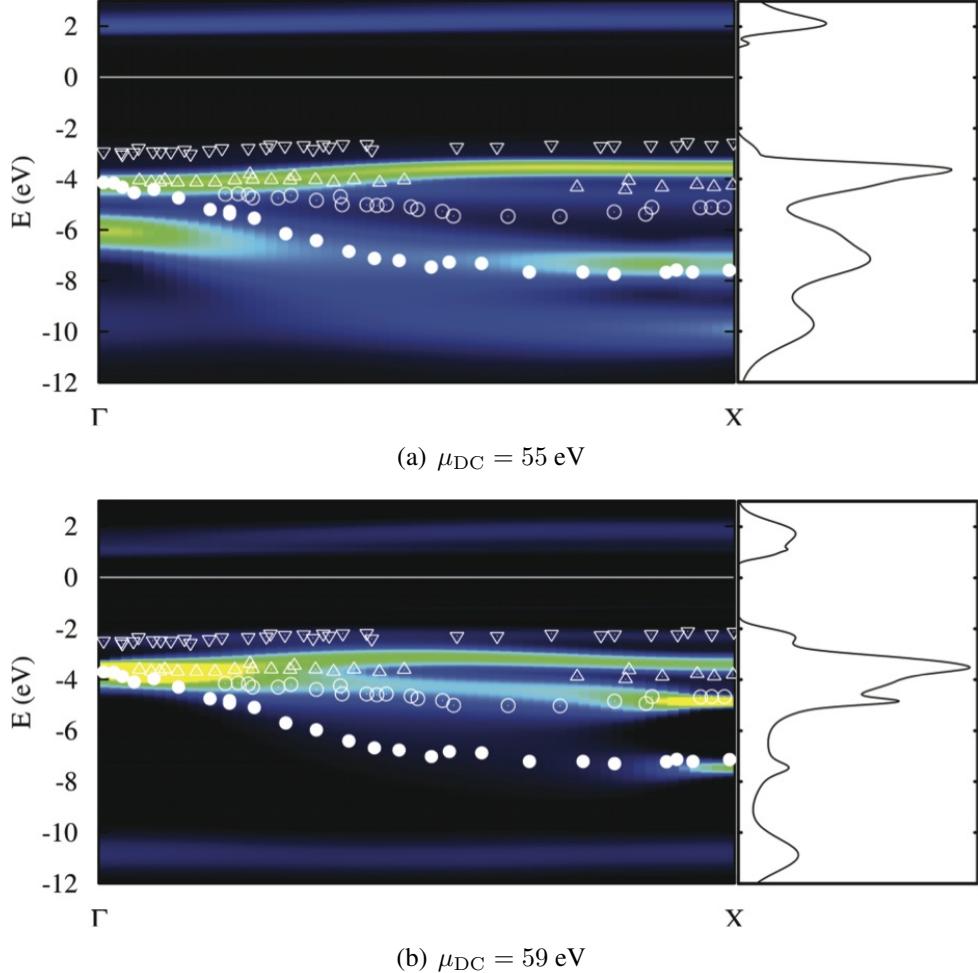


Fig. 16: *k*-resolved spectral functions $A(\mathbf{k}, \omega)$ along the line $\Gamma - X$ in the Brillouin zone for different values of the double counting μ_{DC} . On the right the spectra integrated along the shown crystal momentum direction are shown [41].

correspond to oxygen $2p$ states, the bands above are formed by Ni $3d$ states. The characteristic features seen in ARPES, like the broadening of the oxygen bands around the midpoint of the $\Gamma - X$ line, are clearly present. The quantitative features, especially the relative band energies can strongly differ, depending on the double counting chosen. The bands in Fig. 16 (a) ($\mu_{DC} = 55 \text{ eV}$) show a clear separation between the oxygen and the nickel part at the Γ -point as well as the X -point. At the increased value $\mu_{DC} = 59 \text{ eV}$, Fig. 16 (b), the oxygen bands are shifted towards the Fermi level, coming to overlap with the Ni $3d$ bands at the Γ -point as in the ARPES data. A detailed comparison of the calculated band structures with experiments shows that the bands calculated with $\mu_{DC} = 59 \text{ eV}$ agree very well with the experimental data. These calculations reproduce the flat bands at -4 eV , and another at about -2 eV becomes more prominently visible at $\mu_{DC} = 59 \text{ eV}$, while it is very faint at $\mu_{DC} = 55 \text{ eV}$. The dispersive bands in the region -4 eV to -8 eV also agree very well with experiment. Calculations with other values of the double counting can strongly differ from the experimental data, as shown by the example of $\mu_{DC} = 55 \text{ eV}$.

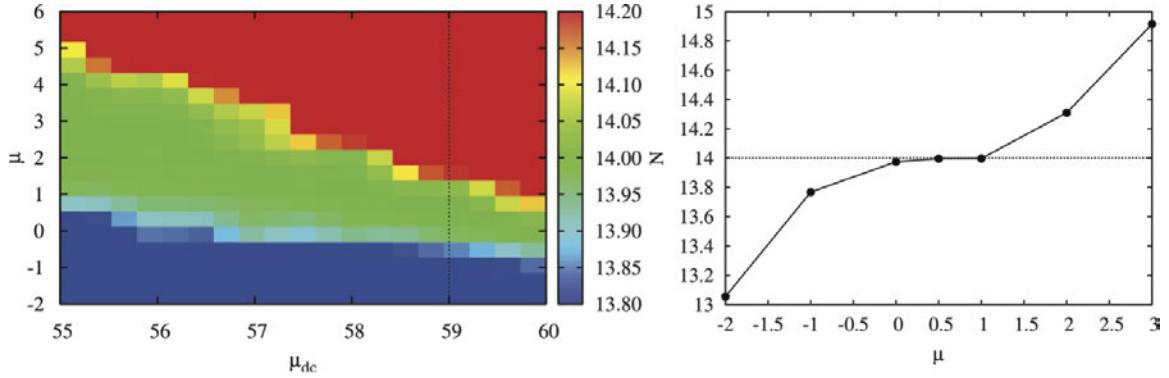


Fig. 17: Surface created by different combinations of the chemical potential μ and the double counting potential μ_{DC} . The particle number has been color coded: the green plateau corresponds to a particle number very close to the desired value of 14, values below are encoded in blue, values above in red. In addition the dotted line indicates the location of the N vs μ curve for $\mu_{DC} = 59$ eV that is shown on the right. The line is a guide to the eye only [41].

The dimension of the double-counting problem becomes apparent when the parameter space of the overall chemical potential μ and the double counting potential μ_{DC} versus the total particle number in the system N is examined. The result is shown in Fig. 17 on the left with the particle number color coded. The picture shows that in principle any combination of μ and μ_{DC} that yields a point in the green plateau, corresponding to the desired particle number $N \approx 14$ *a priori* describes the system equivalently well.

7 Conclusions

We have learnt how to treat electronic correlations in correlated materials within the local DMFT scheme. This knowledge can be used in realistic LDA+DMFT calculations for strongly correlated transition-metals and rare-earth systems, where spin, orbital, and charge fluctuations in the d - or f -shell play a crucial role in photoemission spectra and in magnetic and optical excitations. The numerically exact solution of the quantum impurity problem gives us an effective local exchange-correlation functional for given correlated materials in a specific external field. The combination of DMFT with first-principles approaches is a very useful tool for the investigation of correlated electronic materials.

Acknowledgment

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

References

- [1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964)
- [2] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965)
- [3] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
- [4] G. Ortiz and P. Ballone, Phys. Rev. B **50**, 1391 (1994)
- [5] E.K.U. Gross and W. Kohn, Adv. Quantum Chem. **21**, 255 (1990)
- [6] R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
- [7] V.I. Anisimov, F. Aryasetiawan, and A.I. Lichtenstein,
J. Phys.: Condens. Matter **9**, 767 (1997)
- [8] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003)
G. Kotliar and D. Vollhardt, Physics Today **57**, 53 (2004)
- [9] L. de Medici, J. Mravlje, and A. Georges, Phys. Rev. Lett. **107**, 256401 (2011)
- [10] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, and C.A. Marianetti,
Rev. Mod. Phys. **78**, 865 (2006)
- [11] A.B. Migdal: *Theory of finite Fermi Systems and applications to atomic nuclei*
(Interscience Publishers, New York, 1967)
- [12] P. Nozières: *Theory of interacting Fermi systems* (Benjamin, New York, 1964).
- [13] G. Baym and L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)
- [14] M. Potthoff, Eur. Phys. J. B **32**, 429 (2003)
- [15] R. van Leeuwen, N.E. Dahlen, and A. Stan, Phys. Rev. B **74**, 195105 (2006)
- [16] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)
- [17] H. Hafermann, F. Lechermann, A.N. Rubtsov, M.I. Katsnelson, A. Georges, and
A.I. Lichtenstein, Lecture Notes in Physics **843**, 145 (2012)
- [18] H. Hafermann, S. Brener, A.N. Rubtsov, M.I. Katsnelson, A.I. Lichtenstein,
JETP Lett. **86**, 677 (2007)
- [19] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Annals Phys. **327**, 1320 (2012)
- [20] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner,
Rev. Mod. Phys. **83**, 349 (2011)

- [21] H. Hafermann, G. Li, A.N. Rubtsov, M.I. Katsnelson, A.I. Lichtenstein, and H. Monien, Phys. Rev. Lett. **102**, 206401 (2009)
- [22] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [23] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [24] E. Pavarini, E. Koch, and A.I. Lichtenstein, Phys. Rev. Lett. **101**, 266405 (2008)
- [25] A.N. Rubtsov, M.I. Katsnelson, A.I. Lichtenstein, and A. Georges, Phys. Rev. B **79**, 045133 (2009)
- [26] A.A. Abrikosov, L.P. Gorkov, and I.E. Dzyaloshinskii: *Methods of Quantum Field Theory in Statistical Physics* (Pergamon Press, New York, 1965)
- [27] A.N. Rubtsov and A.I. Lichtenstein, JETP Lett. **80**, 61 (2004)
- [28] P. Werner, A. Comanac, L. de Medici, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **97**, 076405 (2006)
- [29] N.V. Prokofev, B.V. Svistunov, and I.S. Tupitsyn, JETP Sov. Phys. **87**, 310 (1998)
- [30] S.M.A. Rombouts, K. Heyde, and N. Jachowicz, Phys. Rev. Lett. **82**, 4155 (1999)
- [31] P. Werner and A.J. Millis, Phys. Rev. B **74**, 155107 (2006)
- [32] K. Haule, Phys. Rev. B **75**, 155113 (2007)
- [33] V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin, and G. Kotliar, J. Phys.: Condensed Matter **9**, 7359 (1997)
- [34] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
- [35] O.K. Andersen and T. Saha-Dasgupta, Phys. Rev. B **62**, R16219 (2000)
- [36] V.I. Anisimov, D.E. Kondakov, A.V. Kozhevnikov, I.A. Nekrasov, Z.V. Pchelkina, J.W. Allen, S.-K. Mo, H.-D. Kim, P. Metcalf, S. Suga, A. Sekiyama, G. Keller, I. Leonov, X. Ren and D. Vollhardt, Phys. Rev. B **71**, 125119 (2005)
- [37] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki and O.K. Andersen, Phys. Rev. B **74**, 125120 (2005)
- [38] P.E. Blöchl, Phys. Rev. B **50**, 17953 (1994)
- [39] M. Karolak, T.O. Wehling, F. Lechermann and A.I. Lichtenstein, J. Phys.: Condens. Matter **23**, 085601 (2011)
- [40] D. Grieger, C. Piefke, O.E. Peil, and F. Lechermann, Phys. Rev. B **86**, 155121 (2012)

- [41] M. Karolak, G. Ulm, T. Wehling, V. Mazurenko, A. Poteryaev, and A. Lichtenstein, *J. Electr. Spectr.* **181**, 11 (2010)
- [42] A.I. Lichtenstein, M.I. Katsnelson, and G. Kotliar, *Phys. Rev. Lett.* **87**, 067205 (2001).
- [43] G.A. Sawatzky and J.W. Allen, *Phys. Rev. Lett.* **53**, 2339 (1984)
- [44] D.E. Eastman and J.L. Freeouf, *Phys. Rev. Lett.* **34**, 395 (1975)
- [45] Z.X. Shen, C.K. Shih, O. Jepsen, W.E. Spicer, I. Lindau, and J.W. Allen, *Phys. Rev. Lett.* **64**, 2442 (1990)

5 Projectors, Hubbard U , Charge Self-Consistency, and Double-Counting

Tim Wehling

Institute for Theoretical Physics

Bremen Center for Computational Material Sciences

University of Bremen

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Correlated subspaces and projectors | 3 |
| 2.1 | Quantum impurity problems | 4 |
| 2.2 | Projector formalism in LDA+DMFT | 6 |
| 3 | Interaction terms: Hubbard U and beyond | 10 |
| 3.1 | The constrained random phase approximation | 12 |
| 3.2 | Non-local Coulomb interactions | 14 |
| 4 | Double-counting and charge self-consistency | 17 |
| 5 | Conclusions | 21 |

1 Introduction

The electronic structure problem in real materials involves two sources of complexity: First, on a single-particle level, a large Hilbert space can be required to describe how electronic wave functions adjust to a certain arrangement of nuclei. Second, electrons interact, and the dimension of the Fock space to describe a many-electron system grows exponentially with system size. There are two kinds of electronic structure approaches that circumvent either one of these problems.

So-called first-principles methods – most prominently density functional theory (DFT) – map the interacting-electron problem onto an auxiliary single-particle system but treat the full complexity of the single-particle wave functions. Thereby, DFT and related approaches provide material specific and atomistic descriptions of various extended systems (particularly *sp*-electron metals). These approaches fail, however, for strongly correlated materials such as transition-metal oxides or Kondo systems, where the electronic ground state requires a superposition of multiple Slater determinants and where excitations are governed by dynamic self-energy effects. The description of strongly correlated electron systems relies on model Hamiltonians, which operate on many-body Fock spaces constructed out of a reduced set of single-particle basis states. In this way, several correlation phenomena such as the Kondo effect, metal-insulator transitions, magnetism or unconventional superconductivity can be addressed but at the price that relations of models and materials are sometimes ambiguous due to *a priori* unknown model parameters.

Obviously, first-principles and model Hamiltonian approaches have complementary merits and shortcomings, so that their combination presents a promising route towards realistic, i.e., atomistic and material specific, descriptions of strongly correlated electron systems. The combination of density functional theory with dynamical mean field theory, termed LDA+DMFT, is one very successful example of such hybrid approaches and is introduced in the lecture of Alexander Lichtenstein. Generally, approaches which are based on combining DFT and model Hamiltonian approaches to correlated systems are referred to as DFT++ [1].

In this lecture, we will discuss how the realm of first-principles theories can be brought into model Hamiltonian approaches to strongly correlated electron systems. We will see in section 2 that a possible strategy to do so works as follows: A correlated subspace \mathcal{C} , i.e., a subset of single-particle states where electron correlations take place, is identified. Then, it will turn out that projectors from the full space of Kohn-Sham eigenstates onto the subspace \mathcal{C} provide a general way for linking DFT and model Hamiltonian approaches. Within the correlated subspace, the DFT band structure is augmented with interaction terms to generate the correlation effects missed in DFT. These interactions affect only a subset of orbitals and often only local interactions are considered, while clearly electrons in all orbitals and also at distant sites interact with each other. The question of how to determine meaningful interaction parameters entering the many-body models will be addressed in section 3. DFT includes already some (partly unknown portion of) interaction effects in a static mean field manner, which have to be accounted for. This leads to so-called double-counting corrections which are discussed in section 4. The inclu-

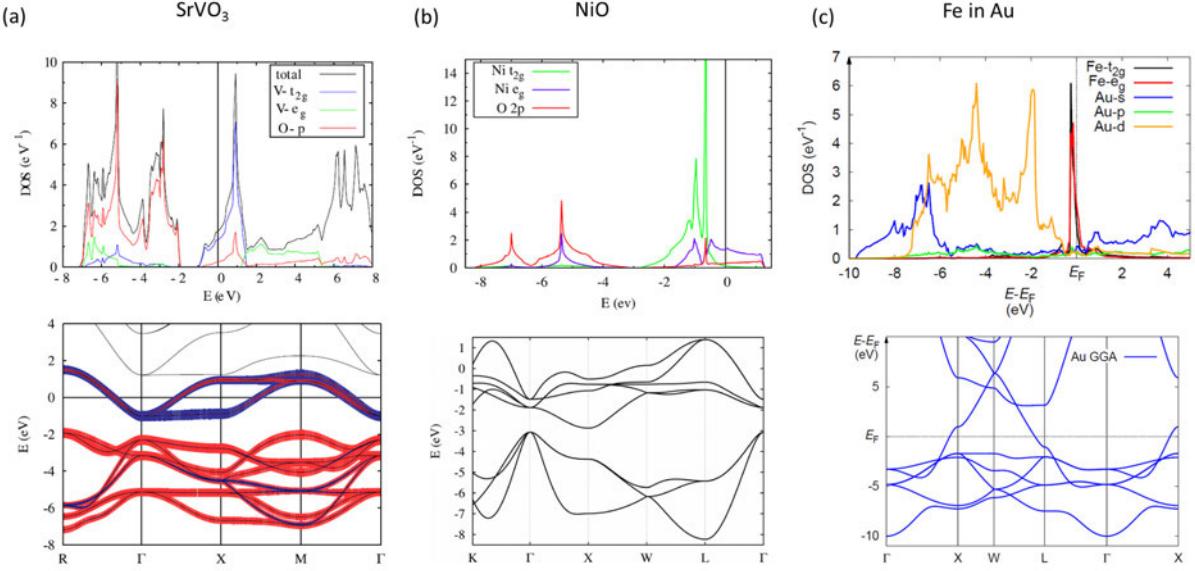


Fig. 1: Electronic structure of three example systems. (a) SrVO_3 , (b) NiO , and (c) Fe impurities in Au. The electronic density of states (DOS) and band structures as obtained from DFT are shown. For Fe in Au the band structure of the host is shown. The systems illustrate different levels of complexity on the single-particle level. While for SrVO_3 the block of three t_{2g} -bands around the Fermi level turns out to control the low energy physics, NiO is experimentally known to be a charge transfer insulator and a description of elementary electronic excitations requires to consider both Ni-3d and O-2p states. For Fe impurities in Au, there is a continuum of sp-like host states together with impurity 3d-states at the Fermi level. Crystal momentum is no more a good quantum number of single-particle states in the the impurity system, which makes its description already involved on the single-particle level.

sion of interactions in the correlated subspace can lead to a redistribution of charges between different orbitals of the system, which would in turn modify the mean field interaction terms contained in DFT and thereby included in the single-particle part of the DFT++ models. Resulting issues of charge self-consistency between the DFT and the many-body parts will finally be considered in section 4.

2 Correlated subspaces and projectors

We start our discussion with SrVO_3 , NiO , and magnetic Fe impurities in Au, which are examples of, respectively, correlated metals, charge transfer insulators and Kondo systems. Band structures and density of states as derived from DFT calculations are summarized in Fig. 1.

SrVO_3 has a block of three bands, the so-called t_{2g} bands, in the vicinity of the Fermi level. These bands have mainly V 3d character and it turns out that the low energy electronic structure can be understood in terms of these bands. In other words, a many-body Hamiltonian for the description of correlation effects in SrVO_3 could be obtained from these t_{2g} bands alone.

NiO is a so-called charge transfer insulator [2]. In contrast to Mott-Hubbard insulators, where the Hubbard U opens a charge gap within the transition-metal d bands, in charge-transfer systems the gap typically opens between ligand p bands and the upper Hubbard bands derived from

the transition-metal d states. Thus, the so-called charge transfer energy $|\varepsilon_p - \varepsilon_d|$ determines the size of the gap. Quite often, ligand p bands and transition-metal d states mix through hybridization and it is thus the interplay of transition-metal d states and oxygen $2p$ states that defines the fundamental electronic excitations. Paramagnetic DFT fails to reproduce this behavior; it predicts NiO to be metallic with Ni- e_g bands close to the Fermi level, $E_F = 0$, and O $2p$ -states more than 3eV below E_F . It reveals, nevertheless, hybridization between Ni and O-derived bands. A description of the electronic excitations of NiO should involve the intermixed O $2p$ and Ni $3d$ -bands, while we expect that electronic correlation effects result mostly from the partially filled transition-metal $3d$ states. A natural many-body model would thus involve 8 bands, i.e., 3 carrying mainly O $2p$ -weight and 5 bands derived from the Ni $3d$ states. We would then end up with a correlated subspace of dimension 5 embedded into larger 8 dimensional space of single-particle Bloch states.

Finally, Fe in Au is an impurity problem, and there are no well defined bands since the crystal momentum k is not a good quantum number of the single-particle states any more. In practice, impurity problems are often modeled using supercells containing the transition-metal impurity atom and on the order of hundred atoms to mimic the host. Assuming that electron correlations mainly take place in the impurity d -orbitals one still has a correlated subspace of dimension 5 but embedded into single-particle Hilbert space spanned by a few hundred supercell bands. We will below see that coupling between the correlated and uncorrelated parts of the single-particle Hilbert space can be elegantly formulated based on projection operators. The discussion follows mainly Refs. [3–8].

The first step of any DFT++ approach is to identify a correlated subspace $\{|m\rangle\}$, where the Kohn-Sham Hamiltonian H_K is augmented by interactions H_U and a double-counting correction H_{DC} . One thus arrives at a Hamiltonian

$$H = \underbrace{\sum_k \varepsilon_k c_k^\dagger c_k}_{H_K} - \underbrace{\mu_{DC} \sum_m d_m^\dagger d_m}_{H_{DC}} + \underbrace{\frac{1}{2} \sum_{m\dots m'''} U_{m\dots m'''} d_m^\dagger d_{m'}^\dagger d_{m''} d_{m'''}}_{H_U}, \quad (1)$$

where the Kohn-Sham energies ε_k and the Kohn-Sham eigenstates $|k\rangle$ obtained from DFT define the non-interacting starting point. To make calculations feasible one often assumes local interactions that couple only the correlated orbitals at the same site. In the general form of Eq. (1), H could be a multiband Hubbard or a multiorbital Anderson Impurity Model (AIM).

2.1 Quantum impurity problems

We start with the discussion of an impurity problem, as for instance realized by ad-atoms on surfaces or by magnetic dopants in bulk metals. We will show that projections $\langle k|m\rangle$ of the Kohn-Sham states onto the states of the correlated subspace are sufficient to connect the DFT real material simulations with the multi-orbital AIM. In general, states $|k\rangle$ and $|m\rangle$ have a finite overlap $\langle k|m\rangle \neq 0$. i.e., they are non-orthogonal. If one constructs an orthonormalized basis of single-particle states which includes the states $|m\rangle$ and an orthogonal set $\{|\tilde{k}\rangle\}$ with

corresponding Fermi operators $c_{\tilde{k}}$ one could rewrite the Hamiltonian in the form

$$H_{\text{AIM}} = \sum_{\tilde{k}} \varepsilon_{\tilde{k}} c_{\tilde{k}}^\dagger c_{\tilde{k}} + \sum_{\tilde{k}, m} (V_{\tilde{k}m} c_{\tilde{k}}^\dagger d_m + h.c.) + \sum_m (\varepsilon_m - \mu_{DC}) d_m^\dagger d_m + \frac{1}{2} \sum_{m \dots m'} U_{m \dots m'''} d_m^\dagger d_{m'}^\dagger d_{m''} d_{m'''}. \quad (2)$$

Equivalently, the problem can be characterized through the corresponding action

$$S_{\text{AIM}}(c^*, c, d^*, d) = \int_0^\beta d\tau \sum_{\tilde{k}} c_{\tilde{k}}^*(\tau) \partial_\tau c_{\tilde{k}}(\tau) + \sum_m d_m^*(\tau) \partial_\tau d_m(\tau) + H_{\text{AIM}}(c^*, c, d^*, d)(\tau), \quad (3)$$

where the Fermi operators are replaced by Grassmann numbers, which leads to the partition function \mathcal{Z} via the imaginary time path integral

$$\mathcal{Z} = \int \mathcal{D}[c^*, c, d^*, d] e^{-S_{\text{AIM}}(c^*, c, d^*, d)}. \quad (4)$$

This integral is Gaussian in the Grassmann numbers $c_{\tilde{k}}^*$ and $c_{\tilde{k}}$, i.e., the bath parts can be integrated out and we arrive at an effective action

$$S_{\text{eff}}(d^*, d) = - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_{m, m'} d_m^*(\tau) (\mathcal{G}_0^{-1}(\tau - \tau')) d_m(\tau') + \int_0^\beta d\tau H_U(d^*, d)(\tau), \quad (5)$$

which specifies all local electronic properties of the AIM. The bath enters indirectly through the non-interacting Green function \mathcal{G}_0 of the correlated orbitals. \mathcal{G}_0 can be easily obtained from our first-principles calculations: As the Kohn-Sham eigenvalues and eigenstates *define* the non-interacting starting point of our DFT++ impurity treatment, the Kohn-Sham Green function

$$G_{KS}(i\omega_n) = \sum_k \frac{|k\rangle\langle k|}{i\omega_n + \mu - \varepsilon_k} \quad (6)$$

plays the role of the non-interacting Green function and the matrix elements entering Eq. (5) read

$$\mathcal{G}_0(i\omega_n)_{mm'} = \langle m | G_{KS}(i\omega_n) | m' \rangle = \sum_k \frac{\langle m | k \rangle \langle k | m' \rangle}{i\omega_n + \mu - \varepsilon_k}. \quad (7)$$

The single-particle part of the AIM is thus fully defined, once the DFT Kohn-Sham problem of the system of interest is solved and the projections $\langle k | m \rangle$ of the Kohn-Sham eigenstates onto the basis states of the correlated subspace are known.

Defining the projector $\mathcal{P}_{\mathcal{C}} = \sum_m |m\rangle\langle m|$ onto the correlated subspace \mathcal{C} , $\mathcal{G}_0(i\omega_n)$ from Eq. (7) can be represented in a compact matrix notation

$$\mathcal{G}_0(i\omega_n) = \mathcal{P}_{\mathcal{C}} G_{KS}(i\omega_n) \mathcal{P}_{\mathcal{C}}. \quad (8)$$

In this way, the single-particle terms entering the DFT++ quantum impurity model for a complex system like a magnetic impurity in a metal or metal-organic molecules on metal surfaces are defined from first-principles. Once the interaction (see section 3) and double-counting terms (see section 4) are also specified, the impurity model, Eq. (5), can be solved, e.g., by quantum Monte Carlo methods as explained in the lecture by Fakher Assaad.

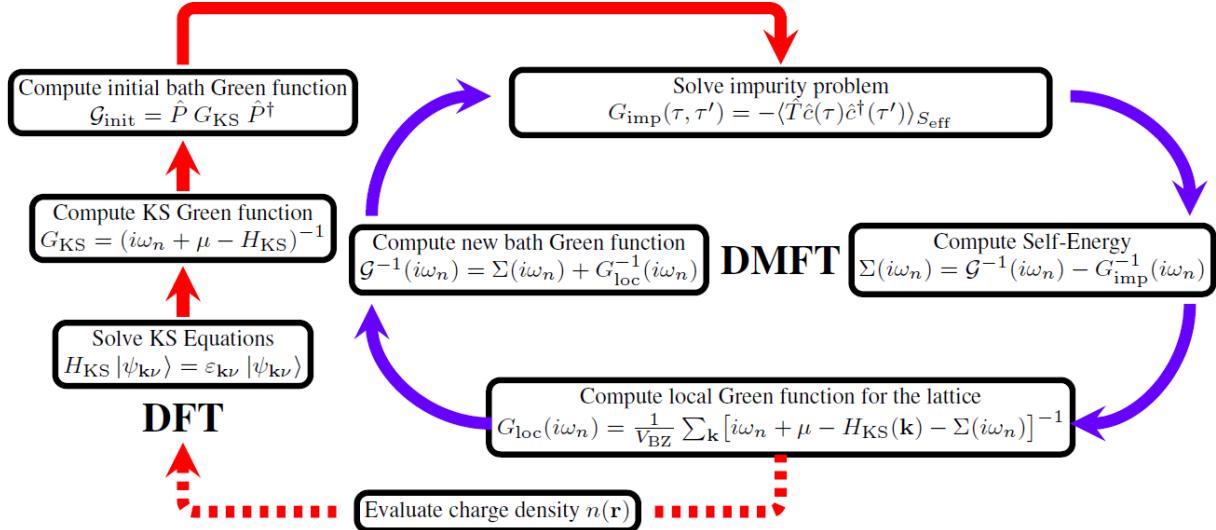


Fig. 2: DFT+DMFT self-consistency cycle. The algorithm starts with a DFT calculation, which yields the Kohn-Sham Hamiltonian and Kohn-Sham eigenstates $|k\rangle \equiv |\psi_{k\nu}\rangle$. Then, the KS Green function and from it the starting value for the bath Green function \mathcal{G}_0 are constructed and passed on to the DMFT loop. The DMFT loop involves upfolding of the self-energy, Eq. (15), and downfolding of the full Green function to the correlated subspace, Eq. (17). Both can be accomplished using projectors. A potential self-consistency over the charge density $n(\mathbf{r})$ is also indicated. From Ref. [8].

2.2 Projector formalism in LDA+DMFT

Dynamical mean field theory maps correlated lattice models such as the Hubbard model onto Anderson impurity models with a self-consistency condition, as explained in the lecture by Antoine Georges. In DMFT, the auxiliary impurity problems involve the correlated orbitals with local interaction, coupled to a self-consistent energy-dependent bath. In the effective action formulation, this takes again the form of Eq. (5), where only the meaning of \mathcal{G}_0 changes. It is no longer the bare (i.e. non-interacting) local Green function of the DFT++ Hamiltonian (1), but rather the dynamical mean-field, which is determined self-consistently and which encodes the coupling of the embedded atom to the effective bath. In other words, \mathcal{G}_0 is the analogue of the Weiss field in the mean field theory of classical magnets. Due to the self-consistency cycle of DMFT, \mathcal{G}_0 depends on many-body effects in the material under consideration.

The central quantity in the DMFT formalism is the local Green function $G_R^{\text{loc}}(i\omega_n)$, which contains simply those matrix elements of the full Green function $G(i\omega_n)$ which belong to correlated orbitals from the same site R . We can thus use the projection operators $\mathcal{P}_R^C = \sum_m |\mathbf{R}m\rangle\langle\mathbf{R}m|$ to correlated orbitals $|\mathbf{R}m\rangle$ at site R to write the local Green function

$$G_R^{\text{loc}}(i\omega_n) = \mathcal{P}_R^C G(i\omega_n) \mathcal{P}_R^C. \quad (9)$$

As before, we interpret the Green functions as operators acting on the space of single-particle states.

By definition of the self-energy $\Sigma(i\omega_n)$, the full Green function of the lattice problem is given

by

$$G(i\omega_n)^{-1} = G_{KS}(i\omega_n)^{-1} - \Sigma(i\omega_n). \quad (10)$$

Since the interaction terms of the DFT++ Hamiltonian are restricted to the correlated subspace, the self-energy can be non-zero only within \mathcal{C} . In the DMFT approximation, $\Sigma(i\omega_n)$ is local and obtained from the auxiliary impurity problem according to

$$\Sigma(i\omega_n) = \mathcal{G}_0^{-1}(i\omega_n) - G_{imp}^{-1}(i\omega_n), \quad (11)$$

where $G_{imp}(i\omega_n)$ is the Green function of the auxiliary impurity problem defined through S_{eff} . From the DMFT self-consistency condition

$$G_{imp}(i\omega_n) = G_{\mathbf{R}}^{\text{loc}}(i\omega_n), \quad (12)$$

we obtain a prescription on how to construct the bath Green function:

$$\mathcal{G}_0^{-1}(i\omega_n) = \Sigma(i\omega_n) + (G_{\mathbf{R}}^{\text{loc}})^{-1}(i\omega_n). \quad (13)$$

Eqs. (11) and (13) define the DMFT self-consistency cycle in a basis-independent way as illustrated in Fig. 2. For actual computations, a basis $\{|B_{\mathbf{k}\alpha}\rangle\}$ (often referred to as Bloch basis) has to be chosen to represent the Green functions explicitly as matrices. There are two natural choices: one could either use the Kohn-Sham eigenstates or any basis set (e.g. plane waves, projector augmented plane waves, or full potential linearized augmented plane waves) which is implemented in the DFT-code used.

To translate the equations defining the self-consistency cycle (Eqs. (9–13)) to this basis set dependent notation, we need the matrix-representations of the Kohn-Sham Hamiltonian

$$H_{\text{KS}}(\mathbf{k})_{\alpha\alpha'} = \sum_k \langle B_{\mathbf{k}\alpha}|k\rangle \varepsilon_k \langle k|B_{\mathbf{k}\alpha'}\rangle \quad (14)$$

and of the self-energy operator

$$\Sigma_{\alpha\alpha'}(\mathbf{k}, i\omega_n) = \langle B_{\mathbf{k}\alpha}|\Sigma(i\omega_n)|B_{\mathbf{k}\alpha'}\rangle = \langle B_{\mathbf{k}\alpha}|m\rangle \Sigma_{mm'}(i\omega_n) \langle m'|B_{\mathbf{k}\alpha'}\rangle. \quad (15)$$

These lead directly to the full Green function

$$G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) = \{i\omega_n + \mu - H_{\text{KS}}(\mathbf{k}) - \Sigma(\mathbf{k}, i\omega_n)\}_{\alpha\alpha'}^{-1} \quad (16)$$

in the Bloch basis. Eq. (15) obviously upfolds the self-energy, which is obtained from the solution of the impurity problem in the localized basis of \mathcal{C} , to the full space of Bloch basis functions. Although $\Sigma(i\omega_n)$ is purely local when expressed in the set of correlated orbitals, it acquires in general momentum dependence when expressed in an arbitrary basis set.

The bath Green function is provided to the impurity solver in the localized basis and thus requires (cf. Eq. (13)) the local Green function in the localized basis

$$G_{\mathbf{R}}^{\text{loc}}(i\omega_n)_{mm'} = \sum_{\mathbf{k}, \alpha, \alpha'} \langle \mathbf{R}m|B_{\mathbf{k}\alpha}\rangle G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) \langle B_{\mathbf{k}\alpha'}|\mathbf{R}m'\rangle. \quad (17)$$

That is indeed all we need to implement the DFT+DMFT cycle as depicted in Fig. 2. The projections $\langle \mathbf{R}m | B_{\mathbf{k}\alpha} \rangle$ facilitate both the upfolding of the self-energy to the Bloch basis and downfolding of the full Green function to the local Green function on the correlated subspace. To summarize, DFT+DMFT requires decisions on the following two issues [4]:

1. The local orbitals $|\mathbf{R}m\rangle$ spanning the correlated subspace \mathcal{C} have to be chosen. Different definitions of \mathcal{C} can lead to different results and the quality of the DMFT approximation will in general depend on the choice of \mathcal{C} . One might want to define \mathcal{C} such that the DMFT approximation is best justified, which is intuitively associated with well-localized orbitals.
2. To keep the DFT+DMFT computationally tractable, the basis of Bloch states $\{|B_{\mathbf{k}\alpha}\rangle\}$ should be chosen in such a way that the number of states that have non-zero overlap with the correlated subspace $\langle B_{\mathbf{k}\alpha} | \mathcal{P}_{\mathbf{R}}^{\mathcal{C}} | B_{\mathbf{k}\alpha} \rangle > 0$ remains sufficiently small. Taking $\{|B_{\mathbf{k}\alpha}\rangle\}$ to be the Kohn-Sham eigenstates is often a good choice in this respect while simple plane waves typically lead to too large Bloch spaces.

There are indeed several possibilities for the construction of local orbitals including different flavors of Wannier functions such as maximally localized Wannier functions or so-called N th order Muffin Tin Orbitals (NMTO). A very practical way to construct a basis for \mathcal{C} is to use entities that are already existing in most of the common band-structure codes, namely, the decomposition of local atomic-like orbitals $|\widetilde{\mathbf{R}m}\rangle$ in terms of Bloch basis functions [5]. Indeed, if the set of Bloch states (e.g. the Kohn-Sham eigenstates generated by the DFT code) were complete, we could simply take the set $\{|\widetilde{\mathbf{R}m}\rangle\}$ as basis of the correlated subspace \mathcal{C} .

However, independently of the particular Bloch basis set which is chosen, one has to restrict practical DFT+DMFT calculations always to a finite space of N_B Bloch states. Those states span a finite subspace \mathcal{W} of the total Hilbert space. The local atomic-like states $|\widetilde{\mathbf{R}m}\rangle$ will, in general, have a decomposition involving all Bloch bands. Projections of $\{|\widetilde{\mathbf{R}m}\rangle\}$ onto \mathcal{W} can thus lead to a non-orthonormal set of localized states. The obvious way out is to reorthonormalize, which is easiest done in the following way:

We consider the Bloch transform of the local atomic-like orbitals, $|\widetilde{\mathbf{k}m}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\mathbf{R}} |\widetilde{\mathbf{R}m}\rangle$, where N is the number of atoms in the crystal. The projections of $|\widetilde{\mathbf{k}m}\rangle$ onto \mathcal{W} reads

$$\widetilde{|\mathbf{k}m\rangle} = \sum_{\alpha \in \mathcal{W}} |B_{\mathbf{k}\alpha}\rangle \langle B_{\mathbf{k}\alpha}| \widetilde{|\mathbf{k}m\rangle}. \quad (18)$$

The $\{\widetilde{|\mathbf{k}m\rangle}\}$ are not true Wannier functions as they are not orthonormal, i.e., their overlap matrix $O_{mm'}(\mathbf{k}) = \langle \widetilde{|\mathbf{k}m\rangle} | \widetilde{|\mathbf{k}m'\rangle} \rangle$ is not the unit matrix. We arrive however at an orthonormal basis set of the correlated subspace by orthonormalizing according to

$$|\mathbf{k}m\rangle = \sum_{m'} O_{mm'}^{-1/2}(\mathbf{k}) \widetilde{|\mathbf{k}m'\rangle}. \quad (19)$$

Eqs. (18) and (19) indeed define a proper Wannier construction. We call the resulting Wannier functions "projector guided Wannier functions" (PWF). PWFs often serve as starting point of an MLWF localization procedure. The extent to which PWFs differ from MLWFs is system-specific. For rather localized states, like in transition-metal oxides, the differences are indeed small [6], while they are larger in materials with highly extended electronic states like p_z electrons in graphene.

More severe than differences between PWFs and MLWFs are differences associated with different choices of the Bloch space \mathcal{W} . This shall be illustrated with the example of SrVO₃, which is a metal with one electron in the t_{2g} bands and empty e_g bands (see Fig. 1). SrVO₃ is a good test-case for DFT+DMFT calculations because excitation spectra and thermodynamic properties hint at correlation effects taking place: First, direct comparison of the photoemission spectra with the one-particle band-structures, e.g. from DFT-LDA, yields poor agreement. Moreover, the linear coefficient in the temperature-dependent specific heat is twice larger than estimated from DFT-LDA, which suggests correlation induced mass enhancement [4].

For an DFT+DMFT description of SrVO₃, one could make two rather different choices of \mathcal{W} : First, we could focus on a very limited set of low-energy Bloch bands, such as the three t_{2g} bands in the vicinity of the Fermi level (cf. Fig. 1) and generate \mathcal{W} just from the three corresponding Bloch functions. In this case, we would have $\mathcal{W} = \mathcal{C}$. Since now the Bloch bands span a narrow energy window, the Wannier functions defining \mathcal{C} will be rather spatially extended: As Fig. 3b shows, they are centered on vanadium atoms but also have sizable weight on neighboring oxygen atoms, which reflects the hybrid character of the low energy t_{2g} -type states. This first choice of $\mathcal{W} = \mathcal{C}$ is of course appealing since it involves a minimal number of bands but comes at the expense that the investigation of the indirect effects of correlations on bands other than the t_{2g} ones are out of scope.

Alternatively, we could define \mathcal{W} from a larger energy window of Bloch bands including all bands associated with the O 2p and all V 3d states. Then, the indirect impact of electron correlations on largely O 2p-derived bands can also be addressed. Having an enlarged Bloch space \mathcal{W} means also that the basis orbitals $|\mathbf{R}m\rangle$ of the correlated subspace can involve Bloch states from a correspondingly wider energy range (cf. Eq. (19)) and will be more spatially localized. As Fig. 3c shows, they are now indeed much closer to vanadium atomic-like orbitals. A comparison of k-resolved spectral functions as obtained from DFT+DMFT simulations of SrVO₃ with a Bloch space made up by the t_{2g} bands only and involving additionally the oxygen 2p states is given in Fig. 3. In both cases, we find a narrowing of the t_{2g} -like quasi-particle states close to the Fermi level with renormalization factors $Z \approx 0.6$, which is in line with experimentally found band narrowing [6]. Also, independently of the choice of \mathcal{W} , we see spectral weight being transferred from the t_{2g} bands to lower and upper Hubbard bands. The major effect of different choices of \mathcal{W} concerns, as expected, the oxygen 2p derived states. For the larger Bloch space, hybridization with the correlated subspace leads to lifetime broadening of the oxygen 2p bands. While for SrVO₃ both models yield a reasonable description of the low energy physics, there are many materials for which the inclusion of bands beyond the correlated subspace is absolutely required. For example, in charge-transfer insulators like NiO

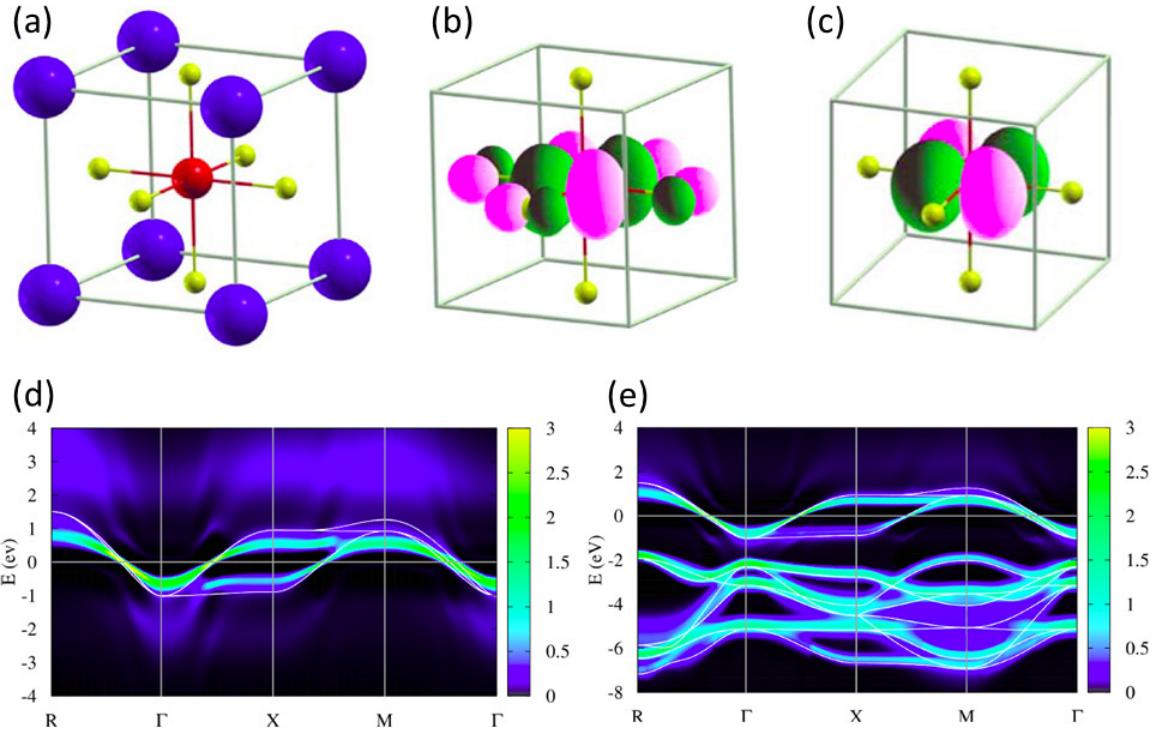


Fig. 3: (a) SrVO_3 structure with Sr large (blue), V (red), and O small (yellow). (b,c) Perspective view of Maximally Localized Wannier functions of the t_{2g} states as obtained from the three t_{2g} -like bands around the Fermi level (b) and from a 14 band calculation, i.e., involving three 2p-like states from three oxygen ions and five 3d-like states from the vanadium ion. From Ref. [4]. Momentum resolved spectral functions as calculated within DFT+DMFT using models involving three Bloch bands only (d) and a Bloch space involving also the oxygen derived states (e). The DFT bands are indicated as white lines. From Ref. [6].

(see section 4), the inclusion of oxygen 2p states in the DFT+DMFT procedure is necessary to describe the fundamental electronic excitations.

Using the projector formalism outlined so far, we can in principle fix all terms entering the DFT++ Hamiltonian, Eq. (1), apart from the double-counting shifts and the interaction matrices. We will discuss these terms in the following two sections.

3 Interaction terms: Hubbard U and beyond

Correlation effects are generated by the interaction terms in the DFT++ Hamiltonian, Eq. (1). What are these interaction terms? Naively, one might guess that they could be obtained as matrix elements of the *bare* Coulomb interaction

$$U_{m\dots m'''} \stackrel{?}{=} \langle \mathbf{R}m | \langle \mathbf{R}m' | \frac{e^2}{\hat{\mathbf{r}} - \hat{\mathbf{r}}'} | \mathbf{R}m'' \rangle | \mathbf{R}m''' \rangle.$$

Indeed, that is not the case, since the DFT++ Hamiltonian restricts interactions to the correlated subspace \mathcal{C} and also assumes that interactions are local. The *Gedankenexperiment* depicted in Fig. (4) illustrates the problem.

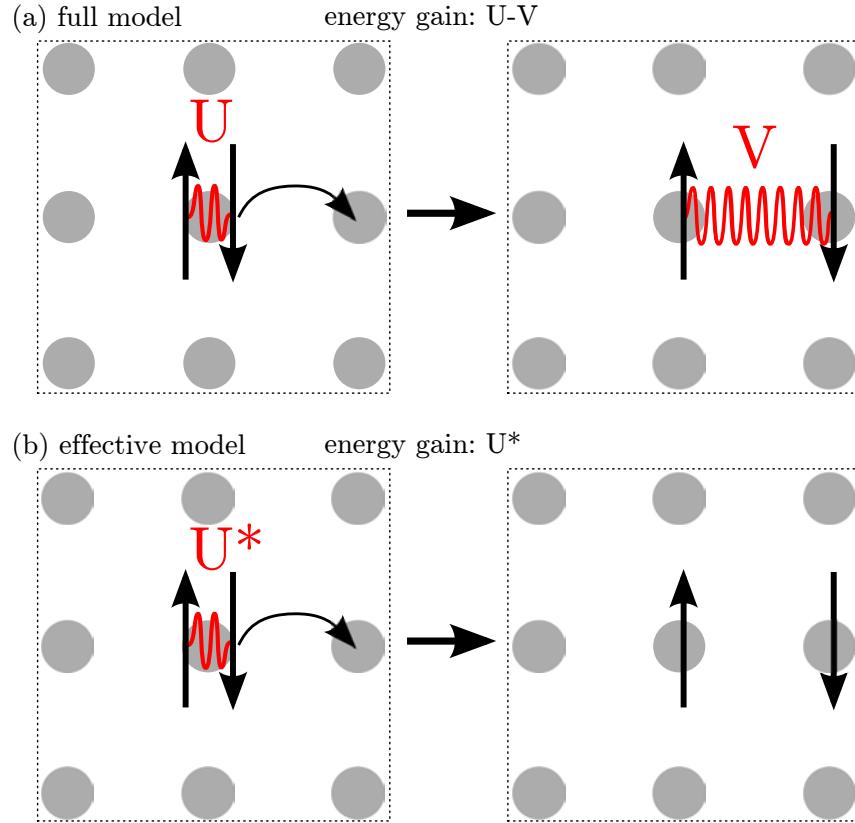


Fig. 4: Interaction energies in models with non-local (a) and purely on-site interactions (b). Wavy lines illustrate Coulomb interactions. (a) An electron in the extended Hubbard model hopping from a doubly occupied site to an empty one, gains an energy $U - V$. (b) The same situation in the model with purely local interactions leads to an energy gain of U^* . From Ref. [9].

In a model with on-site repulsion U and nearest neighbor interaction V , an electron hopping from a doubly occupied site to a neighboring empty site gains interaction energy $U - V$, while in a model with purely local Hubbard repulsion U^* the corresponding energy gain is U^* . A similar argument applies for electrons moving from a correlated orbital of an atom, e.g., the $3d$ -orbital of a transition-metal atom like Fe or Ni, to an uncorrelated $4p$ -orbital at the same site. Here, one would have to account for neglecting the repulsion U_{pd} between p and d electron when determining some Hubbard interaction which is restricted to the space of d -orbitals. Thus far, our discussion involved only two electrons. Obviously in a real solid, all surrounding electrons rearrange in response to the Coulomb potential generated by a charge moving through the material and Coulomb interactions will be screened. The DFT++ model Hamiltonian involves explicitly a limited set of Coulomb processes and associated screening effects. It thus accounts for some amount of electronic screening but many screening channels are indeed neglected. The Coulomb matrix elements entering the DFT++ model should therefore be derived from an appropriately partially screened interaction.

One approach to this problem is the so-called constrained local-density approximation (cLDA) [10]: The Hubbard interaction U , e.g. between $3d$ -electrons at the same site, contributes to the

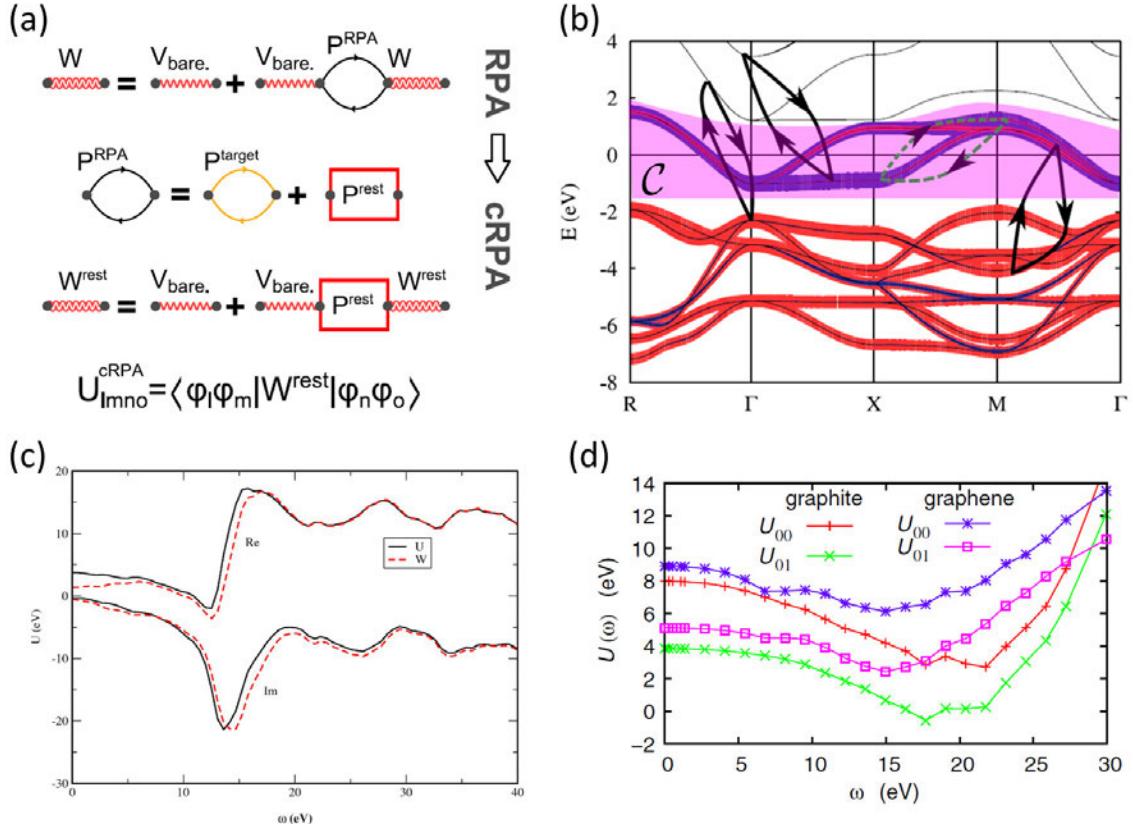


Fig. 5: Screened interaction in the random phase approximation (a). (From Ref. [12]) The constrained random phase approximation (cRPA) excludes bubbles which are entirely within the correlated subspace \mathcal{C} as illustrated here with the example of SrVO₃. If we choose \mathcal{C} to be the t_{2g} block around E_F , the bubbles (dashed) involving t_{2g} to t_{2g} transitions are excluded (b). Frequency dependent effective interactions in SrVO₃ (c) (from Ref. [13]) and graphene (d) (from Ref. [14]) as obtained from cRPA.

total energy a term of the form $E = \frac{1}{2}U n_{3d}(n_{3d} - 1)$. The second derivative of E with respect to the occupancy thus yields the Hubbard interaction

$$U = \frac{\partial^2 E}{\partial n_{3d}^2}. \quad (20)$$

This approach is within the DFT reasoning and has been implemented in several DFT codes.

3.1 The constrained random phase approximation

More recently, a method implementing the idea of partial screening in a diagrammatic language – the so-called constrained random phase approximation (cRPA) – has been proposed [11] and is widely used to date. Let us assume that we have a solid with a well-defined correlated subspace of flat bands near the Fermi level, e.g. transition-metal 3d bands, with the remainder including, for instance, 4s or 4p bands. For simplicity, we assume that there is no hybridization between 3d and 4sp states, i.e., we can span our Bloch space by Kohn-Sham eigenstates $|\psi_d\rangle$ and $|\psi_r\rangle$ referring to 3d-states and the rest, respectively.

The idea is now to construct a partially screened interaction, which accounts for all screening processes except for those involving 3d-to-3d-transitions since the latter processes will be contained in our DFT++ Hamiltonian and should not be counted twice. We can implement this idea within the random phase approximation if we include all bubble diagrams except for those involving 3d-to-3d-transitions in the expansion of the polarization function (cf. Fig. 5). The sum of bubble diagrams leads to the polarization operator

$$\hat{P}(i\Omega_m) = -\frac{1}{\beta} \sum_{i\omega_n} \hat{G}(i\omega_n + i\Omega_m) \hat{G}(i\omega_n). \quad (21)$$

The screened interaction W in RPA is given by

$$\hat{W} = \hat{v} + \hat{v}\hat{P}\hat{v} + \hat{v}\hat{P}\hat{v}\hat{P}\hat{v} + \dots = \left[1 - \hat{v}\hat{P} \right]^{-1} \hat{v}. \quad (22)$$

Taking G to be the Kohn-Sham Green function, the evaluation of the Matsubara sum, (21), leads to a sum over transitions between occupied and empty states which reads in position space representation as

$$P(\mathbf{r}, \mathbf{r}'; i\Omega_m) = \sum_i^{\text{occ}} \sum_j^{\text{empty}} \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \psi_j(\mathbf{r}') \psi_j^*(\mathbf{r}) \left\{ \frac{1}{i\Omega_n + \varepsilon_i - \varepsilon_j} - \frac{1}{i\Omega_n + \varepsilon_j - \varepsilon_i} \right\}. \quad (23)$$

We split \hat{P} into $\hat{P} = \hat{P}_d + \hat{P}_r$, where \hat{P}_d includes only 3d to 3d transitions (i.e, restricting the sums in Eq. (23) to $i, j \in \mathcal{C}$), and P_r be the rest of the polarization. The screened interaction in RPA can then be expressed as

$$\begin{aligned} \hat{W} &= \left[1 - \hat{v}\hat{P}_r - \hat{v}\hat{P}_d \right]^{-1} \hat{v} \\ &= \left[(1 - \hat{v}\hat{P}_r) \{ 1 - (1 - \hat{v}\hat{P}_r)^{-1} \hat{v}\hat{P}_d \} \right]^{-1} \hat{v} \\ &= \left[1 - (1 - \hat{v}\hat{P}_r)^{-1} \hat{v}\hat{P}_d \right]^{-1} (1 - \hat{v}\hat{P}_r)^{-1} \hat{v} \\ &= \left[1 - \hat{W}_r \hat{P}_d \right]^{-1} \hat{W}_r, \end{aligned} \quad (24)$$

where

$$\hat{W}_r(i\Omega_n) = (1 - \hat{v}\hat{P}_r(i\Omega_n))^{-1} \hat{v} \quad (25)$$

is the partially screened interaction we were searching for [11]. It describes the interaction between the electrons of the correlated subspace and accounts for screening by the rest of the system in RPA. The matrix elements of \hat{W}_r can therefore be used obtain the interaction terms entering the DFT++ model from first principles

$$U_{m\dots m'''}(i\Omega_n) = \langle \mathbf{R}m | \langle \mathbf{R}m' | \hat{W}_r(i\Omega_n) | \mathbf{R}m'' \rangle | \mathbf{R}m''' \rangle. \quad (26)$$

From Eq. (25) we see that, the frequency dependence in the polarization function leads to a frequency dependent, i.e. retarded, interaction which in general carries a real and an imaginary part. This effective DFT++ theory will thus not take a Hamiltonian form and the interaction

$\hat{W}_r(i\Omega_n)$ is, in general non-local. We analyze two example cases, SrVO₃ and graphene, to illustrate this point.

For SrVO₃, cRPA yields in the static limit a value of $U = 3.5$ eV [13], which is indeed close to the value of $U \approx 4$ eV used in many DFT+DMFT calculations which include only the t_{2g} bands in the Bloch space and the correlated subspace (cf. Refs. [4–6]). With this value, DFT+DMFT yields quasi-particle weights and the location of the Hubbard bands in reasonable agreement with experiments [4]. The interpretation of the static limit of the cRPA interaction as the interaction to include in the DFT++ Hamiltonian thus seems to be a reasonable approximation and that is indeed what is often done in practice. It should, however, be noted that this approximation neglects several physical processes that affect electronic excitation spectra in real materials. Plasmons, manifesting for instance as poles/resonances in the screened interaction, i.e., dynamic plasmon effects associated with the uncorrelated states, will not enter the DFT++ model if the static limit of the RPA interaction is chosen. For Ni, it has been shown that these high-energy plasmons can affect the low-energy spectra by spectral weight transfer to higher-energy plasmon satellites and a concomitant reduction of the quasi particle weight [13]. It has been suggested that this spectral weight transfer can be accounted for through a renormalization of the hopping and hybridization parameters [15].

3.2 Non-local Coulomb interactions

So far, we considered only local interaction terms. One reasoning behind this is that in transition-metals like Fe or Ni, the uncorrelated sp -electron bands provide efficient screening such that the non-local terms are small [16]. This is not necessarily always the case, as can be seen from the partially screened interaction \hat{W}_r of the p_z -electrons in graphene and graphite in Fig. 5. We see that graphene hosts both sizable on-site repulsion $U_{00} = 9.3$ eV $\approx 3.3 t$ and nearest neighbor interaction $U_{01} = 5.5$ eV $\approx 2 t$, which exceed the nearest neighbor hopping t and are both on the order of the electronic band width $D = 6 t$ [14]. This coexistence of local- and non-local interaction terms is typical for effective models of low dimensional materials and has also been found for two-dimensional superstructures of ad-atoms on semiconductor surfaces [17]. We will therefore discuss the example of graphene a bit closer.

Quantum Monte Carlo simulations of the Hubbard model on the honeycomb lattice have indicated many-body instabilities from a Dirac material towards gapped phases for interaction strengths $U \gtrsim 3.5 t$. At large interactions $U > 4.5 t$ the formation of an antiferromagnetic insulator appears well established, while there is controversy about intermediate interaction strengths $3.5 t < U < 4.5 t$, where Ref. [18] argues for the presence of a spin-liquid phase. More recent calculations question this formation of a spin-liquid but find an antiferromagnetic insulator for $U > 3.9 t$ [19]. Thus, taking the cRPA local Coulomb interaction $U_{00} \approx 3.3 t$ and neglecting all other terms would put graphene close to an instability towards an insulating phase driven by local correlations. This appears surprising, since graphene is indeed one of the best known electric conductors and electrons in graphene are generally assumed to be rather delocalized.

So, what is wrong, here? Obviously, the non-local interaction terms have been neglected or equivalently have been only included on a Hartree level.¹ From the illustration of Fig. 4, it is however clear that non-local terms could indeed weaken the effective local interactions. If we are interested in thermodynamic instabilities (e.g. transitions between a Dirac material and an antiferromagnetic insulator), the following variational approach provides a connection between models with strictly local and non-local interactions [9]:

The starting point is the extended Hubbard model

$$H = - \sum_{i,j,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} + \frac{1}{2} \sum_{\substack{i \neq j \\ \sigma, \sigma'}} V_{ij} n_{i\sigma} n_{j\sigma'}, \quad (27)$$

where t_{ij} are the hopping matrix elements and U and V_{ij} are the local and nonlocal Coulomb matrix elements, respectively. The goal is to map the Hamiltonian (27) onto the effective model

$$H^* = - \sum_{i,j,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U^* \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (28)$$

The effective on-site interaction U^* shall be chosen such that the canonical density operator $\rho^* = 1/Z^* e^{-\beta H^*}$ of the auxiliary system, where $Z^* = \text{Tr} \{ e^{-\beta H^*} \}$ is the partition function, approximates the exact density operator ρ derived from H as close as possible. This requirement leads to the Peierls-Feynman-Bogoliubov variational principle [20–22] for the functional

$$\tilde{\Phi}[\rho^*] = \Phi^* + \langle H - H^* \rangle^*, \quad (29)$$

where $\Phi^* = -\frac{1}{\beta} \ln Z^*$ is the free energy of the auxiliary system. $\langle \dots \rangle^*$ denotes thermodynamic expectation values with respect to the auxiliary system: $\langle H - H^* \rangle^* = \text{Tr} \rho^* (H - H^*)$. In the case of $\rho^* = \rho$ the functional $\tilde{\Phi}[\rho^*]$ becomes minimal and coincides with the free energy. The optimal U^* is thus obtained for minimal $\tilde{\Phi}[\rho^*] = \tilde{\Phi}[U^*]$:

$$\partial_{U^*} \tilde{\Phi}[U^*] = 0. \quad (30)$$

By evaluating Eq. (30) one finds

$$U^* = U + \frac{1}{2} \sum_{\substack{i \neq j \\ \sigma, \sigma'}} V_{ij} \frac{\partial_{U^*} \langle n_{i\sigma} n_{j\sigma'} \rangle^*}{\sum_l \partial_{U^*} \langle n_{l\uparrow} n_{l\downarrow} \rangle^*}. \quad (31)$$

This rule quite closely resembles the *Gedankenexperiment* depicted in Fig. 4: Increasing the on-site term U^* reduces the double occupancy $\langle n_{i\uparrow} n_{i\downarrow} \rangle^*$ and pushes away electrons approaching an already occupied site $i = 0$ to neighboring sites. In the case of purely local Coulomb interactions, there is a Coulomb energy gain of U^* upon suppressing the double occupancy (Fig. 4b). However, when there are nonlocal Coulomb interactions with the surrounding lattice sites j , the displaced electrons raise the energy of the system by terms proportional to V_{0j} . This

¹Neglecting the non-local interactions or inclusion on a Hartree level are equivalent in the model of Eq. (27) if translation invariance is assumed due to cancellation of Hartree terms with the positive charge background stemming from the nuclei.

process is depicted in Fig. 4a for the simple case of two electrons on one site. In this case, it is obvious that the Coulomb energy gain due to the electron displacement in the full and the auxiliary model become energetically equivalent for $U^* = U - V$. In general, this energy gain depends both on the sites to which the charge density is displaced due to the local Coulomb interaction and on how strong the nonlocal Coulomb terms are.

For a translationally invariant system, the local part of the interaction U is reduced according to $U^* = U - \bar{V}$, where

$$\bar{V} = - \sum_{\substack{j \neq 0 \\ \sigma'}} V_{0j} \frac{\partial_{U^*} \langle n_{0\uparrow} n_{j\sigma'} \rangle^*}{\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^*}. \quad (32)$$

The conservation of the total electron number N leads to the sum rules $\sum_{j\sigma} \langle n_{0\uparrow} n_{j\sigma} \rangle^* = \text{const.}$ and $\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^* = - \sum_{j \neq 0, \sigma} \partial_{U^*} \langle n_{0\uparrow} n_{j\sigma} \rangle^*$. Thus, \bar{V} is a weighted average of the nonlocal Coulomb interactions. Under the assumption that an increasing U^* displaces electrons only to next neighbors, we find $\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^* = -N_n \partial_{U^*} \sum_{\sigma} \langle n_{0\uparrow} n_{1\sigma} \rangle^*$, where N_n is the coordination number. Equation (31) then yields

$$U^* = U - V_{01}, \quad (33)$$

which is exactly the situation depicted in Fig. 4.

It is reasonable that \bar{V} is positive (repulsive) in most situations that correspond to real materials. Then, the nonlocal Coulomb interaction reduces the effective on-site interaction and therefore stabilizes the Fermi sea against transitions e.g. to a Mott insulator. This is indeed what happens also in graphene where an evalution of Eq. (32) using correlation functions $\partial_{U^*} \langle n_{0\uparrow} n_{j\sigma'} \rangle^*$ obtained by means of lattice QMC calculations yields $U^* \approx 1.6t$ [9]. The effective local interaction is thus reduced due to the non-local Coulomb terms by more than a factor of two and the Dirac electron phase in graphene is correspondingly stabilized against transitions into an antiferromagnetic insulating phase. The example of graphene thus shows that treatments of non-local interactions beyond the Hartree approximation can be very important to assess phase transitions in strongly correlated electron systems.

The approach discussed here is variational and comes with the simplicity that the auxiliary system (which we solve numerically) involves only local interactions and can thus be treated e.g. by standard DMFT. It is also possible to account for non-local interactions diagrammatically, as for instance in the GW+DMFT approach [23]. In GW+DMFT non-local interactions and related electronic correlation effects are included on an RPA level. Regarding the effect of non-local interactions on boundaries between metallic and Mott insulating phases in low dimensional correlated materials, GW+DMFT also predicts that non-local interactions can stabilize the metallic phase [17]. The pictures emerging from GW+DMFT and the above explained variational approach are thus consistent.

4 Double-counting and charge self-consistency

The central idea of the DFT++ Hamiltonian, Eq. (1), is to introduce interaction terms $U_{m...m''}$ within the correlated subspace to account for dynamic correlation effects. However, the Kohn-Sham energies from DFT already include interaction effects through the Hartree and exchange-correlation terms. Without correction, some interaction contributions would thus be counted twice in DFT++. Thus, some double-counting correction H_{DC} has to be included. One often assumes a form like

$$H_{DC} = \mu_{DC} \sum_m d_m^\dagger d_m. \quad (34)$$

The major problem is that widely used exchange-correlation functionals such as LDA or GGA are non-linear, do not have a diagrammatic representation, and most importantly do not allow one to judge which portion of exchange- and correlation entering the Kohn-Sham eigenvalues is associated with the interactions added in DFT++ within the correlated subspace. Double-counting problems are typical for electronic structure methods where semilocal approximate DFT functionals are augmented with additional interaction terms and also occur in approaches like LDA+U. There is no universal solution to this problem, and the following discussion of practical ways to deal with double-counting will be quite empirical.

Several schemes to fix the double-counting terms have been put forward. All of them are based on some assumption either on how exchange and correlation effects within the correlated subspace are included in a functional like LDA or on some quantity which is assumed to be correctly obtained already from the DFT and which should not change when adding correlations within DFT++.

Since the double-counting correction is intrinsically an impurity quantity and not a global quantity (like the chemical potential μ) it appears natural to use intrinsic quantities of the impurity like the impurity self-energy or the impurity Green function to fix it. One physically intuitive assumption is to require that the electronic charge computed from the local noninteracting Green function and the one computed from the interacting impurity Green function are identical [5]

$$\text{Tr } G_{mm'}^{imp}(\beta) \stackrel{!}{=} \text{Tr } G_{mm'}^{0,loc}(\beta). \quad (35)$$

Alternatively one can also use the Weiss field $\mathcal{G}_{mm'}$ instead of the local noninteracting Green function in the above equation. Both versions of the method give very similar results and work very well in metallic systems [5], since in a metal the total particle number of the system N and of the impurity n_{imp} are both very sensitive to small variations in μ and μ_{dc} .

One possible ansatz using the impurity self-energy $\Sigma_{mm'}^{imp}$ is to constraint the high energy tails in the real part of the self-energy to sum up to zero

$$\text{Re Tr} (\Sigma_{mm'}^{imp}(i\omega_N)) \stackrel{!}{=} 0. \quad (36)$$

Here, ω_N is the highest Matsubara frequency included in the computation. Physically this amounts to the requirement that the shift in the centroid of the impurity orbitals contains no

static component, i.e., that static mean field components of the self-energy are correctly provided by LDA. This criterion is sometimes used in insulating materials [24]. In metals it is otherwise reasonable to assume that the exchange correlation potential yields a good approximation of the self-energy at the Fermi level and thus to require that $\text{Re} \text{Tr} (\Sigma_{mm'}^{imp}(i2\pi/\beta)) \stackrel{!}{=} 0$. The criteria (35) and (36) define the double-counting correction implicitly. There are also two widely used schemes on how to fix the double-counting explicitly from occupation numbers termed “around mean field” (AMF) [25] and “fully localized limit” (FLL) [26]. AMF bases on the idea that exchange and correlation effects are included in LDA but only in a spherically and thus orbitally averaged mean field manner. The resulting double-counting potential is

$$\mu_{dc}^{AMF} = \sum_{m'} U_{mm'} n^0 + \sum_{m', m' \neq m} (U_{mm'} - J_{mm'}) n^0, \quad (37)$$

where $n^0 = \frac{1}{2(2l+1)} \sum_{m,\sigma} n_{m\sigma}$ is the average occupancy. FLL is based on essentially the *opposite* idea. It assumes that total energies for fully localized atomic systems are rather well represented in functionals like LDA, i.e., that LDA (or LSDA) total energies are reliable if orbital occupation numbers $n_{m\sigma}$ are either 1 or 0. While the LDA total energies are assumed to be appropriate in this fully localized case, for non-integer occupations LDA is known to be problematic since it does not correctly reproduce the derivative discontinuity of the exact density functional: it is known that the Kohn-Sham energies (which are derivatives of the total energy with respect to orbital occupations) in LDA do not jump discontinuously as they should for the exact density functional. One can combine the observation of good total energies at integer occupancy but lacking derivative discontinuity into the following prescription for the double-counting potential:

$$\mu_{dc}^{FLL} = U(N_{\text{imp}} - 1/2) + J(N_{\text{imp}}^\sigma - 1/2), \quad (38)$$

with $N_{\text{imp}}^\sigma = \sum_m n_{m\sigma}$ being the total occupancy of the spin σ -component and $N_{\text{imp}} = \sum_\sigma N_{\text{imp}}^\sigma$ being the total occupancy.

Indeed, different prescriptions for the double-counting can lead to different predictions regarding material properties like excitation spectra, as can be seen for the example system of NiO. The double-counting potential μ_{dc} has profound impact on the density of states $N_i(\omega) = -\frac{1}{\pi} \text{Im } G_i(\omega)$ shown in Fig. 6. In the LDA+DMFT study of Ref. [24], the double-counting potential has been treated as an adjustable parameter and has been varied between 21 eV and 26 eV.² The most prominent effects of the double-counting on the spectral properties are the shift of the oxygen p bands with respect to the nickel d bands, as well as the variation in gap size. Plainly speaking, the double-counting correction allows for a tuning of the spectral properties from a large gap Mott-Hubbard insulator to a metal. The regime of the charge transfer insulator, the expected physical state of NiO, lies somewhere in between. The calculated LDA+DMFT(QMC) spectral functions shown in Fig. 6 reveal basically the two different physical situations of a Mott-Hubbard, Fig. 6a, and a charge-transfer insulator, Fig. 6b, mentioned

²These values already contain the intrinsic shift due to the energy of the particle-hole symmetry in the Hirsch-Fye QMC method that amounts to 34 eV with our values of U and J .

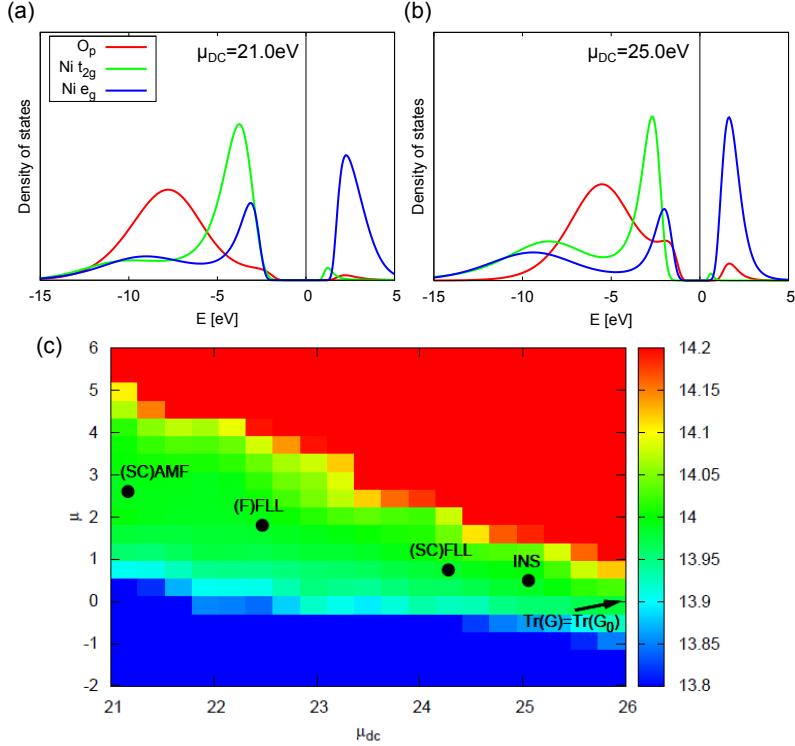


Fig. 6: (a,b) Spectral functions of NiO obtained with LDA+DMFT (QMC) at inverse temperature $\beta = 5\text{ eV}^{-1}$ for different values of the double-counting μ_{dc} . (c) Number of particles N per unit cell (color coded) as function of the chemical potential μ and the double-counting potential μ_{dc} as obtained with LDA+DMFT (QMC). Ni-d and O-p states are included in the calculation, which yields $N = 14$ electrons per unit cell. The green plateau corresponds to a particle number very close to the desired value of $N = 14$. Values below are encoded in blue, values above in red. Additionally the results produced by different prescriptions to fix the double-counting are indicated. For the AMF and FLL functionals SC or F in parentheses indicates, that the occupancies from the DMFT or the formal occupancies have been used, respectively. INS refers to the double correction of $\mu_{DC} = 25\text{ eV}$, where best agreement of ARPES spectra and LDA+DMFT simulations is achieved. From Ref. [24].

above, which are realized depending on the double-counting correction μ_{DC} . The characteristic feature of a charge-transfer system, the strongly hybridized ligand p and transition-metal d character of the low-energy charge excitations, is only present in the spectrum in Fig. 6b. The spectrum in Fig. 6a is missing this feature almost completely and shows Mott-Hubbard behavior. This difference underscores the importance of the double-counting correction. A detailed comparison of calculated bandstructures with experiments shows that the choice of $\mu_{dc} = 25\text{ eV}$ yields best agreement of LDA+DMFT and the experimental data [24].

The pronounced impact of the double-counting correction can be further seen from the plot of the total number of electrons per unit cell on the chemical potential μ and the double-counting correction μ_{DC} in Fig. 6c. μ_{DC} directly affects the pd -charge transfer energy and controls thereby the gap of the system which can be inferred from the $N = 14$ plateau region. Where would the above explained prescriptions for fixing the double-counting correction lead to? Both the AMF and the self-energy criterion, Eq. (36), would lead to $\mu_{DC} \approx 21\text{ eV}$ and thus predict NiO to be basically a Mott-Hubbard but not a charge transfer system. I.e., these criteria are

not in agreement with experiments. FLL yields $\mu_{DC} \approx 24$ eV, which comes closer to the experimental situation of a charge transfer insulator. Finally, criteria trying to fix the occupation of the correlated subspace to some value provided by LDA or the bath Green function, cf. Eq. (35), drives the system towards a metallic state at double-counting $\mu_{DC} = 26.5$ eV indicated by the arrow pointing out of Fig. 6c. Here, FLL appears to describe the system best, which turns out to be often the case for insulators. On the other hand, metals are often well described by double-counting corrections based on traces of the Green function as in Eq. (35). So, the choice of an appropriate double-counting is rather empirical. It is sometimes beneficial to treat the double-counting correction as adjustable parameter and to study the dependence of LDA+DMFT prediction on the choice of the double-counting correction.

A promising way to circumvent double-counting issues are fully diagrammatic approaches like GW+DMFT [23]. These are currently being under development and first GW+DMFT studies of example materials like SrVO₃ [27–29] have been reported. GW+DMFT comes, however, at the expense of considerably higher computational demands than DFT+DMFT.

The DFT++ Hamiltonian in the form of Eq. (1) implicitly includes interactions between electrons in the correlated subspace and the rest of the system through the Hartree as well as the exchange correlation potential from DFT. As soon as the many-body part of DFT++ redistributes electrons between correlated and uncorrelated orbitals or also between different sites there will be associated Hartree (as well as possible exchange or correlation) energies and the DFT++ Hamiltonian should be correspondingly updated. In general, it is obviously problematic to obtain the update of the DFT++ Hamiltonian simply from a double-counting correction applied to the correlated subspace only. This can be better achieved by including self-consistency over the charge-density in the DFT++ approach.

To this end, one calculates the electron density of the DFT++ system,

$$n(r) = \frac{1}{\beta} \sum_{\mathbf{k}, \alpha, \alpha', n} \langle r | B_{\mathbf{k}\alpha} \rangle G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) \langle B_{\mathbf{k}\alpha'} | r \rangle, \quad (39)$$

which includes corrections due to dynamic self-energy effects within the correlated subspace. With this density $n(r)$ one can recalculate the DFT potential and solve the resulting Kohn-Sham Hamiltonian, which then reenters the non-interacting part of the DFT++ Hamiltonian, Eq. (1). In this way, a charge self-consistent DFT++ scheme is obtained, see Fig. 2, which includes interactions between electrons of the correlated subspace and the rest in a fully self-consistent static mean-field manner. Several implementations of charge self-consistent of DFT+DMFT have been reported, e.g. Refs. [30–32], based on projector formalisms similar to Sec. 2.

It is intuitively clear that the Hartree terms occurring within DFT++ charge self-consistency counteract large charge redistributions. In other words, ambiguities stemming for instance from the unknown double-counting potential can be expected to be less severe in charge self-consistent DFT++ calculations as compared to one-shot calculations. This has been explicitly demonstrated, e.g., for the Matsubara self-energies in the iron pnictide superconductor LaFeAsO, where the discrepancy between FLL and AMF approaches is significantly reduced in the fully charge self-consistent scheme [30].

5 Conclusions

The combination of first-principles and model Hamiltonian approaches termed DFT++ presents a promising route towards realistic and material specific descriptions of strongly correlated electron systems. The number of adjustable parameters normally present in models of strongly correlated materials like transition-metal compounds or impurities on surfaces can be indeed largely reduced by deriving them from *ab-initio* calculations. Thereby, realistic studies of ever more complex correlated electron systems are coming into reach. At the same time, the model Hamiltonian level involved in DFT++ offers the chance to study how material properties depend, e.g., on the strength of Coulomb interactions by deliberately treating them as adjustable parameters. With the projector formalism, DFT++ can in principle be applied to arbitrarily complex systems. Many developments in this direction are being pursued throughout the last few years. Naturally, this lecture covered only a very limited amount of these activities, as readers familiar with the subject of DFT++ will have noticed and as becomes clear from a deeper look into the literature referenced here. As has already become clear in the discussions of double-counting issues or frequency dependent and non-local interactions a lot of method development at the interface of first-principles and model based approaches remains still to be done. This includes both the further development of, e.g., DFT+DMFT to a point where it can be as widely and routinely applied as LDA+U, developments in the combination of diagrammatic ab-initio approaches with model based approaches such as GW+DMFT, descriptions of non-local correlation effects or also the coupling of correlated electrons and bosonic modes such as plasmons, phonons or magnons.

The DFT++ model Hamiltonians discussed in this lecture have put an emphasis on correlation effects due to local Coulomb interactions, which are indeed essential for various magnetic phenomena or Mott metal insulator transitions. Other many-body phenomena can rely on different kinds of interactions. Wigner crystallization of electrons, exciton binding, or plasmon modes are often controlled by non-local Coulomb interaction terms, and appropriate models for such phenomena will naturally have to include different interaction terms. In other words: any DFT++ modeling requires an idea on which interactions form the basis of the many-body problem to be described. It can of course be very challenging to identify the essential interactions responsible for an unknown phenomenon or to determine whether some observed effect is a many-body phenomenon or not.

Acknowledgments

Support by Deutsche Forschungsgemeinschaft through FOR 1346 is gratefully acknowledged.

References

- [1] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
- [2] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
- [3] V.I. Anisimov, D.E. Kondakov, A.V. Kozhevnikov, I.A. Nekrasov, Z.V. Pchelkina, J.W. Allen, S.-K. Mo, H.-D. Kim, P. Metcalf, S. Suga, A. Sekiyama, G. Keller, I. Leonov, X. Ren, and D. Vollhardt, Phys. Rev. B **71**, 125119 (2005)
- [4] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and O.K. Andersen, Phys. Rev. B **74**, 125120 (2006)
- [5] B. Amadon, F. Lechermann, A. Georges, F. Jollet, T.O. Wehling, and A.I. Lichtenstein, Phys. Rev. B **77**, 205112 (2008)
- [6] M. Karolak, T.O. Wehling, F. Lechermann, and A.I. Lichtenstein, J. Phys.: Condens. Matter **23**, 085601 (2011)
- [7] T.O. Wehling: *Impurities and Inhomogeneities in Nanoelectronic Systems* (PhD thesis, University of Hamburg, 2010)
- [8] M. Karolak: *Electronic Correlation Effects in Transition Metal Systems: From Bulk Crystals to Nanostructures* (PhD thesis, University of Hamburg, 2013)
- [9] M. Schüler, M. Rösner, T.O. Wehling, A.I. Lichtenstein, and M.I. Katsnelson, Phys. Rev. Lett. **111**, 036601 (2013)
- [10] O. Gunnarsson, O.K. Andersen, O. Jepsen, and J. Zaanen, Phys. Rev. B **39**, 1708 (1989)
- [11] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A.I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004)
- [12] P. Hansmann, L. Vaugier, H. Jiang, and S. Biermann, J. Phys.: Condens. Matter **25**, 094005 (2013)
- [13] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, Phys. Rev. B **74**, 125106 (2006)
- [14] T.O. Wehling, E. Şaşioğlu, C. Friedrich, A.I. Lichtenstein, M.I. Katsnelson, and S. Blügel, Phys. Rev. Lett. **106**, 236805 (2011)
- [15] M. Casula, P. Werner, L. Vaugier, F. Aryasetiawan, T. Miyake, A.J. Millis, and S. Biermann, Phys. Rev. Lett. **109**, 126408 (2012)
- [16] J. Hubbard, Proc. Roy. Soc. (London) A **276**, 238 (1963)

- [17] P. Hansmann, T. Ayral, L. Vaugier, P. Werner, and S. Biermann, Phys. Rev. Lett. **110**, 166401 (2013)
- [18] Z.Y. Meng, T.C. Lang, S. Wessel, F.F. Assaad, and A. Muramatsu, Nature **464**, 847 (2010)
- [19] S. Sorella, Y. Otsuka, and S. Yunoki, Sci. Rep. **2**, 992 (2012)
- [20] R.E. Peierls, Phys. Rev. **54**, 918 (1938)
- [21] N.N. Bogoliubov., Dokl. Akad. Nauk SSSR **119**, 244 (1958)
- [22] R.P. Feynman: *Statistical Mechanics* (Benjamin, Reading Mass., 1972)
- [23] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003)
- [24] M. Karolak, G. Ulm, T. Wehling, V. Mazurenko, A. Poteryaev, and A.I. Lichtenstein, Journal of Electron Spectroscopy and Related Phenomena **181**, 11 (2010)
- [25] V.I. Anisimov, J. Zaanen, and O.K. Andersen, Phys. Rev. B **44**, 943 (1991)
- [26] M.T. Czyżyk and G.A. Sawatzky, Phys. Rev. B **49**, 14211 (1994)
- [27] J.M. Tomczak, M. Casula, T. Miyake, F. Aryasetiawan, and S. Biermann, Europhys. Lett. **100**, 67001 (2012)
- [28] R. Sakuma, P. Werner, and F. Aryasetiawan, Phys. Rev. B **88**, 235110 (2013)
- [29] C. Taranto, M. Kaltak, N. Parragh, G. Sangiovanni, G. Kresse, A. Toschi, and K. Held, Phys. Rev. B **88**, 165119 (2013)
- [30] M. Aichhorn, L. Pourovskii, and A. Georges, Phys. Rev. B **84**, 054529 (2011)
- [31] D. Grieger, C. Piefke, O.E. Peil, and F. Lechermann, Phys. Rev. B **86**, 155121 (2012)
- [32] B. Amadon, J. Physics: Condens. Matter **24**, 075604 (2012)

6 Linear Response Functions

Eva Pavarini

Institute for Advanced Simulation
Forschungszentrum Jülich

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | Linear response theory | 8 |
| 2.1 | The linear susceptibility $\chi(\mathbf{r}, \mathbf{r}'; t, t')$ | 8 |
| 2.2 | The Fourier transform $\chi(\mathbf{q}; \omega)$ | 10 |
| 2.3 | Analytic properties of $\chi(\mathbf{q}; \omega)$ | 13 |
| 2.4 | Kramers-Kronig relations and sum rules | 14 |
| 2.5 | Fluctuation-dissipation theorem | 16 |
| 2.6 | Single-particle Green function | 17 |
| 2.7 | Two-particle Green function | 24 |
| 3 | The dynamical susceptibility | 27 |
| 3.1 | The magnetic susceptibility | 27 |
| 3.2 | The generalized susceptibility | 30 |
| 3.3 | The generalized susceptibility in DMFT | 32 |
| 3.4 | The $\chi_0(\mathbf{q}; \omega)$ diagram | 34 |
| 3.5 | The local susceptibility $\chi(\omega)$ | 37 |
| 4 | Conclusion | 43 |

1 Introduction

All we know about a physical system stems either from its effects on other physical systems or from its response to external forces [1–3]. Making sense out of this information is a very difficult task. First of all, what do we actually mean by physical system? Typically we are interested in crystals. A sample of a crystal is an object with unique characteristics, a specific number of atoms or defects, a certain surface, a given weight. What we are interested in, however, are its general properties, i.e., those common to all possible samples. This idealized crystal is the one we have in mind [4, 5] with the word *system* or *material*. How do we unravel the mechanisms behind its properties? The first step is constructing the Hamiltonian. The latter should be abstract enough to indeed describe the ideal crystal only, but it should retain enough details to actually distinguish it from other systems, i.e., it should be *material-specific*. At a first glance, constructing such a material-specific Hamiltonian appears straightforward. All of solid-state physics stems from the Coulomb interaction, attractive between nuclei and electrons and repulsive between electrons; we can therefore, in principle, just choose a complete one-electron basis and, in the Born-Oppenheimer approximation, write down in second quantization the electronic Hamiltonian for a given ideal crystal \hat{H}_e . To make progress, we then have to solve the eigenvalue problem $\hat{H}_e\Psi = E\Psi$. Here trouble starts, since the Schrödinger equation defined by \hat{H}_e cannot, in general, be solved exactly, and even if we knew the exact solution, with $N_e \rightarrow \infty$ electrons, it would be very hard to make sense out of it. For the ground state, a very powerful tool to attack such a many-body problem is density-functional theory (DFT), which shifts the focus from finding the ground-state wavefunction to finding the ground-state electronic density. Remarkably, this can be achieved via the solution of a reference auxiliary one-electron problem, a much simpler task than solving the original problem. Although DFT is an exact ground-state theory, in practice only approximated forms of the DFT universal functional are known, the most popular of which is perhaps the local-density approximation (LDA). DFT, in the LDA or its simple extensions, is very successful in describing and even predicting the properties of various classes of materials, to the extent that it can be considered the standard model of solid-state physics [5–7]. For strongly correlated systems, those on which we focus in this lecture, the LDA and its simple extension fail even qualitatively, however. Thus, if we want to understand correlated materials we have to revisit the first step, the Hamiltonian. We can change the perspective; if the full many-body problem cannot be solved, the best approach is perhaps to reduce the number of degrees of freedom to the essential by integrating out high-energy ones in the spirit of the Wilson renormalization group. Hence, we have to construct low-energy *minimal* material-specific models. Systematically downfolding the high-energy states of the full many-body problem, although desirable in theory, is basically impossible in practice [8]. It turns out, however, that we can exploit the successes of DFT to build *ab-initio* Wannier functions spanning the low-energy bands [9, 10]; via these Wannier functions, we can construct effective many-body models, minimal and materials-specific, suited to describe the low-energy part of the spectrum. But we are still not at the end of the story; with few exceptions, even these minimal models cannot be solved exactly, and finding powerful and flexible solution techniques

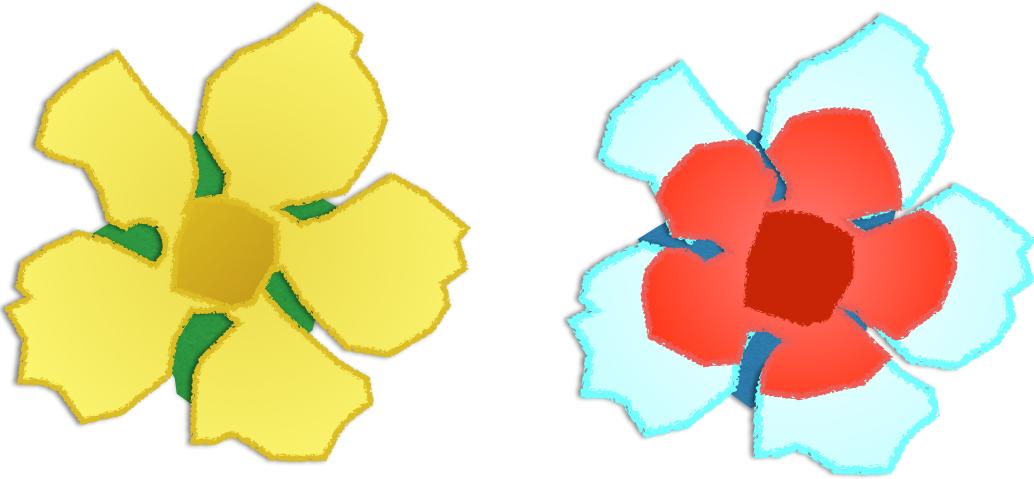


Fig. 1: A silverweed: Schematic view of color patterns as seen by the human eye (left) and as might be seen by the eyes of a bee (right, false colors).

is therefore crucial. Typically this involves making a series of approximations, which, together with the model, have at the end to be put to a test. To this end a big step forward was the development, started 25 years ago, of the dynamical mean-field theory (DMFT) [11–15] and shortly after of the LDA+DMFT method [16, 17], which combines density-functional theory and dynamical mean-field theory. The LDA+DMFT technique quickly proved very successful. Through this method, we have learned that details do matter; for example we understood that a crystal-field splitting much smaller than the band-width can play an important role in the metal-insulator transition of correlated transition-metal oxides [10]. The LDA+DMFT approach was key in identifying the nature of important phenomena such as, e.g., the origin of orbital ordering in paradigmatic correlated transition-metal systems [18]. Various of the striking LDA+DMFT success stories are told in the lecture notes in this book. Thanks to its successes, LDA+DMFT is nowadays the method of choice for strongly-correlated materials.

A central issue remains to be discussed at this point. Once we have built a material specific model and solved it within a given set of approximations, e.g., using the LDA+DMFT approach, how do we actually test our theory against experiments? To make this connection we have to calculate the quantity actually measured, the response of the system to external perturbations. This is a challenge on its own. Methods or approximations that work well for the ground state can, e.g., perform badly for the excitation spectrum, the knowledge of which is crucial for obtaining response functions – examples are the Hartree-Fock or static mean-field approximation or the LDA itself. Furthermore, experiments let us see a system only through a distorting glass, and our task is to reconstruct the original image. Perhaps some part of the energy spectrum is invisible to us, because the transition probability from the ground state to certain excited states is forbidden by symmetry. If the response of a system to a given perturbation is zero in a certain energy window, e.g., filtered away, we have no chance of seeing what is there. Our eyes filter the ultraviolet, and thus we cannot see the beautiful color patterns that attract insects to flowers

such as silverweeds, for us monochromatic (Fig. 1). Response theory therefore plays a crucial role in our comprehension of Nature. Typically, the external force used in experiments is small with respect to the internal ones in a crystal, so that the system is weakly perturbed. Thus, the dominant term is the *linear response function*. If we are able to disentangle it, the linear-response function returns us information on the ground state and the excitation spectrum, their symmetry properties, the strength of correlations. In this lecture, I will present the basics of linear response theory [1–3] together with some representative examples for strongly correlated materials, using for the latter LDA+DMFT as theoretical approach.

To set the stage, let us introduce the basics of the LDA+DMFT method, different aspects of which will be used in the whole lecture. Let us start with the electronic Hamiltonian. In the Born-Oppenheimer approximation, the non-relativistic electronic Hamiltonian for an ideal crystal \hat{H}_e can then be written as the sum of a one-electron part \hat{H}_0 and an interaction part \hat{H}_U

$$\hat{H}_e = \hat{H}_0 + \hat{H}_U. \quad (1)$$

In a complete basis of Wannier functions $\psi_{in\sigma}(\mathbf{r})$, the one-electron term is given by

$$\hat{H}_0 = - \sum_{\sigma} \sum_{ii'} \sum_{nn'} t_{n,n'}^{i,i'} c_{in\sigma}^{\dagger} c_{i'n'\sigma},$$

where $c_{in\sigma}^{\dagger}$ ($c_{in\sigma}$) creates (destroys) an electron with spin σ in orbital n at site i . The on-site ($i = i'$) terms yield the crystal-field matrix while the $i \neq i'$ contributions are the hopping integrals. This part of the Hamiltonian describes the attraction between electrons and nuclei, the latter forming an ideal lattice. The electron-electron repulsion \hat{H}_U is instead given by

$$\hat{H}_U = \frac{1}{2} \sum_{ii'jj'} \sum_{\sigma\sigma'} \sum_{nn'pp'} U_{npn'p'}^{iji'j'} c_{in\sigma}^{\dagger} c_{jpn\sigma'}^{\dagger} c_{j'p'\sigma'} c_{i'n'\sigma}.$$

Although the Hamiltonian (1) is very general, for a given system, the *material-specific* hopping and crystal-field parameters can be obtained *ab-initio* using, e.g., Wannier functions constructed from first principles via density-functional theory [19, 20]. Then

$$t_{n,n'}^{i,i'} = - \int d\mathbf{r} \overline{\psi_{in\sigma}}(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + v_R(\mathbf{r}) \right] \psi_{i'n'\sigma}(\mathbf{r}),$$

where $v_R(\mathbf{r})$ is the self-consistent one-electron LDA reference potential. The *bare* Coulomb integrals can be expressed in terms of Wannier functions as well

$$U_{npn'p'}^{iji'j'} = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \overline{\psi_{in\sigma}}(\mathbf{r}_1) \overline{\psi_{jp\sigma'}}(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_{j'p'\sigma'}(\mathbf{r}_2) \psi_{i'n'\sigma}(\mathbf{r}_1).$$

Here we have to be careful, however. The LDA includes in $v_R(\mathbf{r})$ also Coulomb effects, via the long-range Hartree term and the exchange-correlation contribution; if we use LDA Wannier functions as one-electron basis, to avoid double counting we have to subtract from \hat{H}_U the effects already included in the LDA. This means that we have to replace

$$\hat{H}_U \rightarrow \Delta \hat{H}_U = \hat{H}_U - \hat{H}_{DC},$$

where \hat{H}_{DC} is the double-counting correction. Unfortunately we do not know which correlation effects are exactly included in the LDA, and therefore the exact expression of \hat{H}_{DC} is also unknown.¹ The remarkable successes of the LDA suggest, however, that in many materials the LDA is overall a good approximation, and therefore, in those systems at least, the term $\Delta\hat{H}_U$ can be neglected. What about strongly correlated materials? Even in correlated systems, most likely the LDA works rather well for the delocalized electrons or in describing the average or the long-range Coulomb effects. Thus one can think of separating the electrons into *uncorrelated* and *correlated*; only for the latter we do take the correction $\Delta\hat{H}_U$ into account explicitly, assuming furthermore that $\Delta\hat{H}_U$ is local or almost local [19]. Typically, correlated electrons are those that partially retain their atomic character, e.g., those that originate from localized d and f shells; for convenience in this lecture we assume that in a given system they stem from a single atomic shell l (e.g., d for transition-metal oxides or f for heavy-fermion systems) and label their states with the atomic quantum numbers l and $m = -l, \dots, l$ of that shell. Thus

$$U_{np,n'p'}^{iji'j'} \sim \begin{cases} U_{m_\alpha m_\beta m'_\alpha m'_\beta}^l & iji'j' = iiii \quad npn'p' \in l \\ 0 & iji'j' \neq iiii \quad npn'p' \notin l \end{cases}$$

and $\Delta\hat{H}_U$ is replaced by $\Delta\hat{H}_U^l = \hat{H}_U^l - \hat{H}_{\text{DC}}^l$, where \hat{H}_{DC}^l is, e.g., given by the static mean-field contribution of \hat{H}_U^l . There is a drawback in this procedure, however. By splitting electrons into correlated and uncorrelated we implicitly assume that the main effect of the latter is the renormalization or *screening* of parameters for the former, in particular of the Coulomb interaction. The calculation of screening effects remains, unfortunately, a challenge to date. Approximate schemes are the constrained LDA and the constrained random-phase approximation (RPA) methods [6]. Nevertheless, we have now identified the general class of models for strongly-correlated systems, namely the generalized Hubbard model

$$\hat{H}_e = \hat{H}^{\text{LDA}} + \hat{H}_U^l - \hat{H}_{\text{DC}}^l. \quad (2)$$

It is often convenient to integrate out or downfold empty and occupied states and work directly with a set of Wannier functions spanning the correlated bands only. The LDA term in \hat{H}_e is then given by

$$\hat{H}^{\text{LDA}} = - \sum_{ii'} \sum_{\sigma} \sum_{m_\alpha m'_\alpha} t_{m_\alpha m'_\alpha}^{i,i'} c_{im_\alpha \sigma}^\dagger c_{i'm'_\alpha \sigma} = \sum_{\mathbf{k}} \sum_{\sigma} \sum_{m_\alpha m'_\alpha} [H_{\mathbf{k}}^{\text{LDA}}]_{m_\alpha m'_\alpha} c_{\mathbf{k}m_\alpha \sigma}^\dagger c_{\mathbf{k}m'_\alpha \sigma},$$

where the right-hand side is rewritten using as a one-electron basis Bloch functions $\psi_{\mathbf{k}m_\alpha \sigma}$ constructed from the Wannier functions $\psi_{im_\alpha \sigma}$. The local *screened* Coulomb interaction is instead given by

$$\hat{H}_U^l = \frac{1}{2} \sum_i \sum_{\sigma\sigma'} \sum_{m_\alpha m'_\alpha} \sum_{m_\beta m'_\beta} U_{m_\alpha m_\beta m'_\alpha m'_\beta} c_{im_\alpha \sigma}^\dagger c_{im_\beta \sigma'}^\dagger c_{im'_\beta \sigma'} c_{im'_\alpha \sigma}.$$

¹A more detailed discussion of the double-counting correction can be found in the lectures of Tim Wehling and Alexander Lichtenstein.

More details on building realistic models can be found, e.g., in the lecture notes of previous schools [5–7], in particular in the chapters listed in Refs. [19–22]. The simplest version of Hamiltonian (2) is the one-band Hubbard model

$$\hat{H}_{\text{Hubbard}} = - \underbrace{\sum_{ii'} \sum_{\sigma} t_{1,1}^{i,i'} c_{i\sigma}^{\dagger} c_{i'\sigma}}_{\hat{H}_0} + \underbrace{\varepsilon_d \sum_{i\sigma} n_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}}_{\hat{H}_U}, \quad (3)$$

where ε_d is the crystal-field level, $t_{1,1}^{i,i'}$ is the hopping integral between electrons at site i and i' , U the on-site Coulomb repulsion, and $n_{i\sigma} = c_{i\sigma}^{\dagger} c_{i\sigma}$. Since in this model only the correlated orbital ($m_{\alpha} = 1$) appears, the double-counting correction amounts to a mere shift of the chemical potential, and therefore does not have to be included explicitly. The one-band Hubbard model describes, at least in first approximation, the low-energy states of high-temperature superconducting cuprates (HTSCs), characterized by a partially filled Cu d $x^2 - y^2$ -like band at the Fermi level. In these systems the most relevant hopping integrals are the one between nearest neighbors, t , and the one between next-nearest neighbors, t' ; the ratio t'/t ranges from 0.1 to 0.4 [9]. At half-filling, the one-band Hubbard model describes the physics of the Mott metal-insulator transition. In the atomic limit ($U/t \rightarrow \infty$), it is a collection of decoupled one-electron atoms

$$\hat{H}_{\text{Hubbard}} \sim \hat{H}_U = \varepsilon_d \sum_{i\sigma} n_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (4)$$

In the non-interacting limit ($U/t \rightarrow 0$) it describes instead a metallic half-filled band

$$\hat{H}_{\text{Hubbard}} \sim \hat{H}_0 = - \sum_{ii'} \sum_{\sigma} t_{1,1}^{i,i'} c_{i\sigma}^{\dagger} c_{i'\sigma} = \sum_{\sigma} \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}\sigma}. \quad (5)$$

In this lecture we will use the half-filled Hubbard model (3) for most examples. In particular, we will discuss its magnetic linear-response function; the microscopic mechanisms leading to magnetism in the Hubbard model – and in correlated systems in general – are discussed in detail in my lecture of last year’s school [22], which is complementary to the present one.

Although apparently simple, even Hamiltonian (3) cannot be solved exactly except than in special cases. The state-of-the art method for solving Hubbard-like models is, as discussed, dynamical mean-field theory. The latter maps the correlated *lattice* problem described by the Hubbard model onto a correlated single-impurity problem [14, 15], i.e., onto an effective model describing a correlated site i_c coupled via hybridization to a bath of non-correlated electrons. This effective single impurity model has to be solved self-consistently in the spirit of mean-field theories. The DMFT self-consistency loop is shown in Fig. 2. The impurity Green function $G(\omega)$, in general a matrix in spin-orbital space, is obtained by solving the quantum-impurity problem for a given bath. Next, the Dyson equation yields the self-energy

$$\Sigma(\omega) = \mathcal{G}^{-1}(\omega) - G^{-1}(\omega),$$

where $\mathcal{G}(\omega)$ is the bath Green-function matrix; the self-energy matrix $\Sigma(\omega)$ is then used to calculate the local Green-function matrix

$$G_{i_c, i_c}(\omega) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} [\omega - H_{\mathbf{k}}^{\text{LDA}} - \Sigma(\omega)]_{i_c, i_c}^{-1},$$

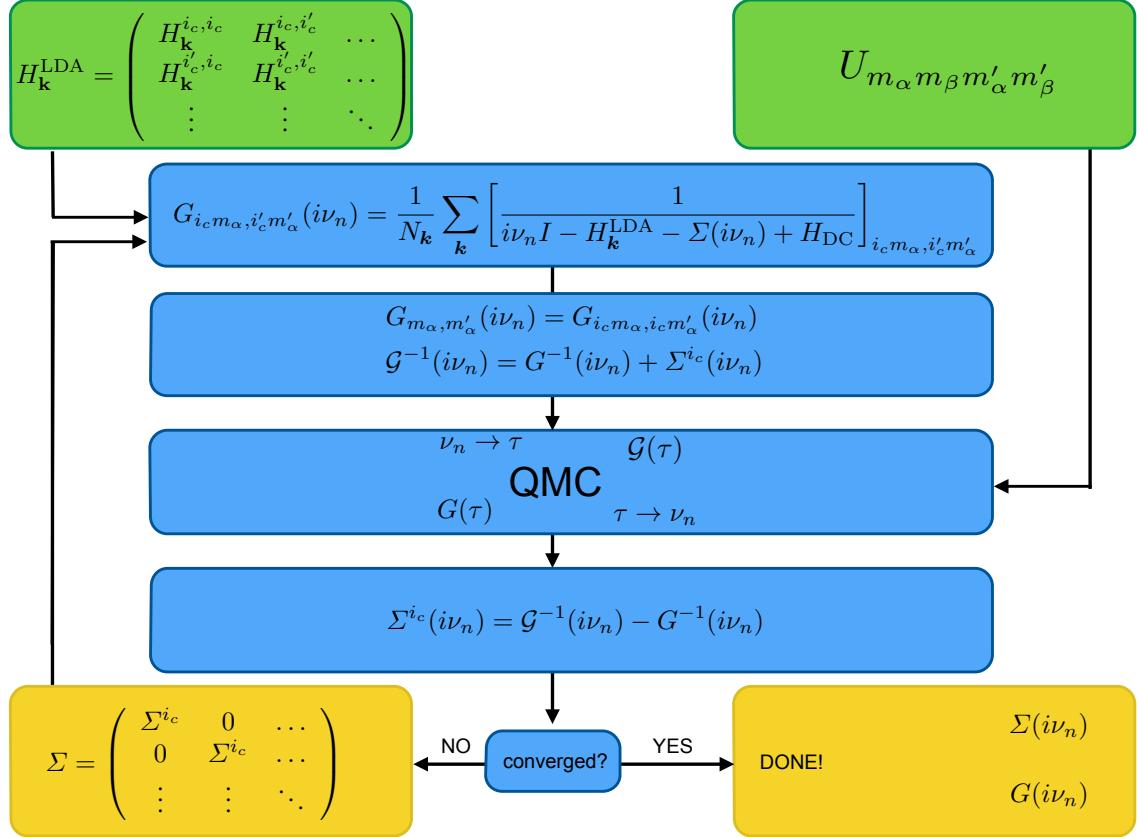


Fig. 2: The LDA+DMFT self-consistency loop. The LDA Hamiltonian $H_{\mathbf{k}}^{\text{LDA}}$ is built in the basis of Bloch states obtained from localized Wannier functions. The set $\{i_c\}$ labels the equivalent correlated sites inside the unit cell. The local Green-function matrix is at first calculated using an initial guess for the self-energy matrix. The bath Green-function matrix is then obtained via the Dyson equation and used to build an effective quantum-impurity model. The latter is solved via a quantum-impurity solver, here quantum Monte Carlo (QMC), yielding the impurity Green-function matrix. Through the Dyson equation the self-energy is then obtained, and the procedure is repeated till self-consistency is reached.

where $N_{\mathbf{k}}$ is the number of \mathbf{k} points. Self-consistency is reached when the impurity Green function $G(\omega)$ equals the actual local Green function $G_{i_c, i_c}(\omega)$ of the system. The main approximation adopted is that the self-energy is local; the self-energy becomes indeed local in the infinite-coordination-number limit [11, 12]. The combination of DMFT with density-functional theory, just described above in short, defines the LDA+DMFT approach [17]. In this lecture we will not further discuss this technique, except for the specific aspects related to the calculation of linear-response functions. These include the local-vertex approximation, used to obtain from DMFT calculations the \mathbf{q} -dependent linear-response function $\chi(\mathbf{q}; \omega)$ and the quantum-impurity model and its solution; the latter yields the local impurity Green-function matrix and the local linear response function $\chi(\omega)$. A more detailed description of the DMFT and LDA+DMFT methods and their development can be found in the lecture notes of Dieter Vollhardt, Gabriel Kotliar, Antoine Georges, and Alexander Lichtenstein.

2 Linear response theory

2.1 The linear susceptibility $\chi(\mathbf{r}, \mathbf{r}'; t, t')$

Let us consider a system described by the Hamiltonian \hat{H} and a space- and time-dependent perturbation $\mathbf{h}(\mathbf{r}, t)$, for example a magnetic field. How does the system react to the external perturbation? If the perturbation is weak and we can calculate the change in the Hamiltonian to linear order, the response of the system can be given in terms of retarded correlation functions calculated at equilibrium, even if the perturbation has brought the sample out of equilibrium. The linear correction to the Hamiltonian can be expressed as

$$\begin{aligned}\hat{H} &\rightarrow \hat{H} + \int d\mathbf{r} \hat{H}_1(\mathbf{r}; t) + \dots \\ \hat{H}_1(\mathbf{r}; t) &= -\sum_{\nu} \hat{O}_{\nu}(\mathbf{r}; t) h_{\nu}(\mathbf{r}; t),\end{aligned}\tag{6}$$

where $\hat{O}(\mathbf{r}; t)$ is an operator that describes the system property affected by the perturbation; often this operator is a vector, thus we indicate with $\nu = x, y, z$ its components along the Cartesian axes. If the perturbation is an external magnetic field, $\hat{O}(\mathbf{r}; t)$ could be the magnetic moment density, $\hat{\mathbf{M}}(\mathbf{r}; t)$. It is convenient to express $\hat{O}(\mathbf{r}; t)$ in the Heisenberg representation

$$\hat{O}_{\nu}(\mathbf{r}; t) = e^{i(\hat{H}-\mu\hat{N})t} \hat{O}_{\nu}(\mathbf{r}) e^{-i(\hat{H}-\mu\hat{N})t},$$

where μ is the chemical potential and \hat{N} the electron number operator. The perturbation $\mathbf{h}(\mathbf{r}; t)$ can display very different forms, depending on the experiment. It could have been, e.g., initially switched on adiabatically at $t = -\infty$; this can be expressed mathematically by multiplying the perturbation by the prefactor $e^{\epsilon t}$, where here ϵ is an infinitesimally small positive number, $\mathbf{h}(\mathbf{r}; t) \rightarrow e^{\epsilon t} \mathbf{h}(\mathbf{r}; t)$, and taking later the limit $\epsilon \rightarrow 0$. The perturbation could also be a sharp impulse at $t = t_0$ and therefore have the form $\mathbf{h}(\mathbf{r}; t) = \mathbf{h}(\mathbf{r}) \delta(t - t_0)$, or have been switched on or off suddenly at a certain time $t = t_0$.

Whatever its form, let us consider the effect of the time-dependent perturbation $\mathbf{h}(\mathbf{r}; t)$ on a specific system property, described by the operator $\hat{P}(\mathbf{r}; t)$, also expressed in the Heisenberg representation; in general $\hat{P}(\mathbf{r}; t)$ can be a different operator than $\hat{O}(\mathbf{r}; t)$, but in many common cases it is proportional to it. To linear order in the perturbation, at a given temperature the expectation value of $\hat{P}(\mathbf{r}; t)$ is modified as follows

$$\begin{aligned}\langle \hat{P}_{\nu}(\mathbf{r}; t) \rangle &= \langle \hat{P}_{\nu}(\mathbf{r}) \rangle_0 + \langle \delta \hat{P}_{\nu}(\mathbf{r}; t) \rangle_0, \\ \langle \delta \hat{P}_{\nu}(\mathbf{r}; t) \rangle_0 &= -i \int d\mathbf{r}' \int_{-\infty}^t dt' \left\langle \left[\Delta \hat{P}_{\nu}(\mathbf{r}; t), \Delta \hat{H}_1(\mathbf{r}'; t') \right] \right\rangle_0.\end{aligned}$$

Here $\langle \hat{P}_{\nu}(\mathbf{r}) \rangle_0$ is the (equilibrium) thermal average in the absence of the perturbation. For a given operator \hat{A} , the latter is defined as

$$\langle \hat{A} \rangle_0 = \frac{1}{Z} \text{Tr} \left[e^{-\beta(\hat{H}-\mu\hat{N})} \hat{A} \right],$$

where $\beta = 1/k_B T$, and Z is the partition function

$$Z = \text{Tr } e^{-\beta(\hat{H}-\mu\hat{N})}.$$

The difference

$$\Delta\hat{A}(\mathbf{r}; t) = \hat{A}(\mathbf{r}; t) - \langle \hat{A}(\mathbf{r}) \rangle_0,$$

measures the deviation with respect to the thermal average in the absence of perturbation; in most common cases $\langle \hat{A}(\mathbf{r}) \rangle_0 = 0$. Let us consider as an example the case in which the perturbation is a magnetic field and $\hat{P}_\nu(\mathbf{r}; t)$ the ν -component of the magnetic density operator $\hat{M}_\nu(\mathbf{r}; t)$; for a paramagnetic system the equilibrium expectation value $\langle \hat{M}_\nu(\mathbf{r}) \rangle_0$ is zero and $\Delta\hat{M}_\nu(\mathbf{r}; t) = \hat{M}_\nu(\mathbf{r}; t)$. Finally, for a given operator $\hat{A}(\mathbf{r})$, the expectation value $\langle \Delta\hat{A}^2(\mathbf{r}) \rangle_0$ yields the mean-square fluctuation of the quantity $\hat{A}(\mathbf{r})$.

By replacing the operator $\Delta\hat{H}_1(\mathbf{r}'; t')$ in the commutator with its expression obtained from Eq. (6), we can express the linear correction to $\langle \hat{P}_\nu(\mathbf{r}; t) \rangle_0$ as

$$\langle \delta\hat{P}_\nu(\mathbf{r}; t) \rangle_0 = i \sum_{\nu'} \int d\mathbf{r}' \int_{-\infty}^t dt' \left\langle \left[\Delta\hat{P}_\nu(\mathbf{r}; t), \Delta\hat{O}_{\nu'}(\mathbf{r}'; t') \right] \right\rangle_0 h_{\nu'}(\mathbf{r}'; t').$$

The linear response function or *linear susceptibility* is then given by

$$\chi_{\hat{P}_\nu\hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t, t') \equiv \lim_{h_{\nu'} \rightarrow 0} \frac{\partial \langle \hat{P}_\nu(\mathbf{r}; t) \rangle}{\partial h_{\nu'}(\mathbf{r}'; t')}.$$

The equation

$$\chi_{\hat{P}_\nu\hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t, t') = i \left\langle \left[\Delta\hat{P}_\nu(\mathbf{r}; t), \Delta\hat{O}_{\nu'}(\mathbf{r}'; t') \right] \right\rangle_0 \Theta(t - t'), \quad (7)$$

is known as the *Kubo formula*. In order to respect causality, a perturbation can only modify the system *after* it has been switched on. Thus, if the perturbation is switched on at time t' , the linear response function can only have finite value for $t > t'$, and it has to vanish for $t < t'$; in other words, the response function is *retarded*. This cause-and-effect principle is included in Eq. (7) through the Heaviside step function $\Theta(t - t')$, defined as

$$\Theta(t - t') = \begin{cases} 1 & \text{if } t - t' > 0 \\ 0 & \text{if } t - t' < 0. \end{cases}$$

It is worth pointing out that the Kubo formula Eq. (7) yields the response function in terms of the *correlation function*

$$\mathcal{S}_{\hat{P}_\nu\hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t, t') = \langle \Delta\hat{P}_\nu(\mathbf{r}; t) \Delta\hat{O}_{\nu'}(\mathbf{r}'; t') \rangle_0. \quad (8)$$

The latter expresses the joint probability of having a finite $\Delta\hat{P}_\nu$ at position \mathbf{r} and time t if there was a finite $\Delta\hat{O}_{\nu'}$ at position \mathbf{r}' and time t' . This relation between linear response function and correlation function will turn out to be very important, as we will see in Sec. 2.5.

2.2 The Fourier transform $\chi(\mathbf{q}; \omega)$

Let us consider the case in which the Hamiltonian \hat{H} of the system is time-independent and thus also invariant under time translations. Then the linear response function depends only on time differences, $t - t'$, and we can rewrite it as follows

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t, t') = \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t - t').$$

Thus the linear correction to the expectation value of $\hat{P}_\nu(\mathbf{r}; t)$ becomes

$$\langle \delta \hat{P}_\nu(\mathbf{r}; t) \rangle_0 = i \sum_{\nu'} \int d\mathbf{r}' \int_{-\infty}^{+\infty} dt' \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t - t') h_{\nu'}(\mathbf{r}'; t'). \quad (9)$$

Many perturbations are periodic in time after they have been switched on. It is therefore convenient to Fourier transform Eq. (9) with respect to time (see Appendix for definitions and conventions on Fourier transforms adopted in this lecture), obtaining

$$\langle \delta \hat{P}_\nu(\mathbf{r}; \omega) \rangle_0 = \sum_{\nu'} \int d\mathbf{r}' \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; \omega) h_{\nu'}(\mathbf{r}'; \omega),$$

where $\mathbf{h}(\mathbf{r}'; \omega)$ is the Fourier transform of the perturbation and $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; \omega)$ the Fourier transform of the susceptibility. The latter is given by

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; \omega) = \int_{-\infty}^{\infty} dt \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t) e^{i\omega t} = \int_0^{\infty} dt \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; t) e^{i\omega t}.$$

It is also convenient to Fourier transform Eq. (9) with respect to \mathbf{r} ; for a system with full spatial translational invariance symmetry, i.e., for which the momentum is conserved,

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; \omega) = \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r} - \mathbf{r}'; \omega),$$

and thus we have

$$\begin{aligned} \langle \delta \hat{P}_\nu(\mathbf{q}; \omega) \rangle_0 &= \sum_{\nu'} \int \frac{d\mathbf{q}'}{(2\pi)^3} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}, -\mathbf{q}'; \omega) h_{\nu'}(\mathbf{q}'; \omega) \\ &= \sum_{\nu'} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) h_{\nu'}(\mathbf{q}; \omega), \end{aligned} \quad (10)$$

where

$$\begin{aligned} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}, -\mathbf{q}'; \omega) &= \int d\mathbf{r} e^{i\mathbf{q} \cdot \mathbf{r}} \int d\mathbf{r}' e^{-i\mathbf{q}' \cdot \mathbf{r}'} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}, \mathbf{r}'; \omega) \\ &= \int d\mathbf{r}' \left[\int d\mathbf{r} e^{i\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r} - \mathbf{r}'; \omega) \right] e^{-i(\mathbf{q}' - \mathbf{q}) \cdot \mathbf{r}'} \\ &= (2\pi)^3 \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) \delta(\mathbf{q} - \mathbf{q}'), \end{aligned}$$

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = \int d\mathbf{r}'' e^{i\mathbf{q} \cdot \mathbf{r}''} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}''; \omega).$$

An ideal crystal has only lattice translational invariance, however. How does relation (10) change in this case? It turns out that it is still valid, but we have to express $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}, -\mathbf{q}'; \omega)$ differently. Let us see how. For simplicity, we consider a lattice of Bravais type (one atom per unit cell); we define as \mathbf{T}_i the lattice vector which identifies site i . Let us assume that $\Delta \hat{P}_\nu(\mathbf{r})$ is the one-body operator $\Phi^\dagger(\mathbf{r}) \Delta \hat{\mathcal{P}}_\nu \Phi(\mathbf{r})$, where $\Phi^\dagger(\mathbf{r})$ is the fermionic field creation operator. The term $\Delta \hat{P}_\nu(\mathbf{r})$ can then be expressed as follows

$$\Delta \hat{P}_\nu(\mathbf{r}) = \sum_{ii'} \sum_{\alpha\alpha'} \underbrace{\overline{\psi_{i\alpha}(\mathbf{r})} \psi_{i'\alpha}(\mathbf{r})}_{\rho_{\alpha'\alpha}^{i,i'}(\mathbf{r})} \underbrace{[\Delta \hat{\mathcal{P}}_\nu]_{\alpha\alpha'} c_{i'\alpha}}_{\Delta \hat{\mathcal{P}}_{\nu,\alpha\alpha'}^{i,i'}} = \sum_{ii'} \sum_{\alpha\alpha'} \rho_{\alpha'\alpha}^{i,i'}(\mathbf{r}) \Delta \hat{\mathcal{P}}_{\nu,\alpha\alpha'}^{i,i'},$$

where $\{\psi_{i\alpha}(\mathbf{r})\}$ is a complete set of orthonormal one-electron wavefunctions and α a collective index for its quantum numbers (for example $\alpha = m_\alpha \sigma$). If we now choose for $\{\psi_{i\alpha}(\mathbf{r})\}$ a set of localized Wannier functions, to first approximation the overlap of two $\psi_{i\alpha}(\mathbf{r})$ centered at different sites is small and can be neglected; this means that $\rho_{\alpha\alpha'}^{i,i'}(\mathbf{r})$ is only sizeable for $i = i'$, and therefore

$$\Delta \hat{P}_\nu(\mathbf{r}) \sim \sum_i \sum_{\alpha\alpha'} \rho_{\alpha'\alpha}^{i,i}(\mathbf{r}) \Delta \hat{\mathcal{P}}_{\nu,\alpha\alpha'}^i,$$

i.e., $\Delta \hat{P}_\nu(\mathbf{r})$ is approximatively a weighted sum of the site operators $\Delta \hat{\mathcal{P}}_{\nu,\alpha\alpha'}^i$. A similar approximation holds for $\Delta \hat{O}_\nu(\mathbf{r})$,

$$\Delta \hat{O}_\nu(\mathbf{r}) \sim \sum_i \sum_{\gamma\gamma'} \rho_{\gamma'\gamma}^{i,i}(\mathbf{r}) \Delta \hat{\mathcal{O}}_{\nu,\gamma\gamma'}^i.$$

Let us introduce the tensorial components of the linear-response function for the site operators

$$\chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^{i'}}^{\alpha\alpha' \gamma\gamma'}(t - t') = i \left\langle \left[\Delta \hat{\mathcal{P}}_{\nu,\alpha\alpha'}^i(t - t'), \Delta \hat{\mathcal{O}}_{\nu',\gamma\gamma'}^{i'}(0) \right] \right\rangle_0 \Theta(t - t').$$

The Fourier transform of $\chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^{i'}}^{\alpha\alpha' \gamma\gamma'}(t - t')$ in time is given by $\chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^{i'}}^{\alpha\alpha' \gamma\gamma'}(\omega)$; furthermore

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}^{\alpha\alpha' \gamma\gamma'}(\mathbf{q}, -\mathbf{q}'; \omega) = \sum_{ii'} e^{i\mathbf{q} \cdot \mathbf{T}_i - i\mathbf{q}' \cdot \mathbf{T}_{i'}} \chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^{i'}}^{\alpha\alpha' \gamma\gamma'}(\omega) = \underbrace{\sum_{i'} e^{i(\mathbf{q} - \mathbf{q}') \cdot \mathbf{T}_{i'}}}_{N_s \sum_{\mathbf{G}} \delta_{\mathbf{q}', \mathbf{q} + \mathbf{G}}} \underbrace{\sum_i e^{i\mathbf{q} \cdot \mathbf{T}_i} \chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^0}^{\alpha\alpha' \gamma\gamma'}(\omega)}_{\chi_{\hat{P}_\nu \hat{O}_{\nu'}}^{\alpha\alpha' \gamma\gamma'}(\mathbf{q}; \omega)},$$

where N_s is the number of lattice sites. In terms of these components, the term $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}, -\mathbf{q}'; \omega)$ in Eq. (10) is given by

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}, -\mathbf{q}'; \omega) = N_s \sum_{\alpha\alpha' \gamma\gamma'} \sum_{\mathbf{G}} \rho_{\alpha'\alpha}(\mathbf{q}) \rho_{\gamma'\gamma}(-\mathbf{q}') \delta_{\mathbf{q}', \mathbf{q} + \mathbf{G}} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}^{\alpha\alpha' \gamma\gamma'}(\mathbf{q}; \omega).$$

where $\rho_{\alpha'\alpha}(\mathbf{q}) = \int d\mathbf{r} e^{i\mathbf{q} \cdot \mathbf{r}} \rho_{\alpha'\alpha}^{i_0, i_0}(\mathbf{r})$. Long-range order instabilities are typically at \mathbf{q}_C vectors that correspond to deformations commensurate with the lattice. To study them, one can, e.g., perturb the system at $\mathbf{q}' = \mathbf{q}_C$, or possibly use experimental techniques that have access to the \mathbf{q} -dependent response function.

Let us now consider as an example the case in which the perturbation is a magnetic field oriented along z and the operators $\Delta\hat{P}_\nu(\mathbf{r})$ and $\Delta\hat{O}_\nu(\mathbf{r})$ are both equal to $\hat{M}_z(\mathbf{r}) = \Phi^\dagger(\mathbf{r})\hat{\sigma}_z\Phi(\mathbf{r})$, the magnetic density at position \mathbf{r} ; the operator $\hat{\sigma}_z$ is the z Pauli matrix. Let us derive the magnetic response for a strongly correlated system with a partially filled l shell; for simplicity, we focus on the l electrons and choose as a basis the set of Wannier functions $\psi_{im_\alpha\sigma}(\mathbf{r})$ spanning the corresponding l bands. Then $\hat{M}_z(\mathbf{r})$ can be expressed as follows

$$\hat{M}_z(\mathbf{r}) \sim -g\mu_B \sum_i \sum_{m_\alpha m'_\alpha} \rho_{m_\alpha m'_\alpha}(\mathbf{r}) \frac{1}{2} \sum_{\sigma\sigma'} c_{im_\alpha\sigma}^\dagger [\Delta\hat{M}_z]_{\sigma\sigma'} c_{im'_\alpha\sigma'},$$

where $[\Delta\hat{M}_z]_{\sigma\sigma'} = \langle\sigma|\hat{\sigma}_z|\sigma'\rangle$. If our low-energy model includes only a single orbital we can drop the indices $\{m_\alpha\}$; in this case we obtain the simple expression

$$\hat{M}_z(\mathbf{r}) \sim -g\mu_B \rho(\mathbf{r}) \sum_i \hat{S}_z^i. \quad (11)$$

The Fourier transform of $\hat{M}_z(\mathbf{r})$ is then given by $\hat{M}_z(\mathbf{q}) \sim -g\mu_B \rho(\mathbf{q}) \hat{S}_z(\mathbf{q})$ with

$$\hat{S}_z(\mathbf{q}; 0) = \frac{1}{2} \sum_{\mathbf{k}} \sum_{\sigma} c_{\mathbf{k}\sigma}^\dagger \langle\sigma|\hat{\sigma}_z|\sigma\rangle c_{\mathbf{k}+\mathbf{q}\sigma}.$$

Finally, to linear order the change in $\langle\hat{M}_z(\mathbf{q}; \omega)\rangle_0$ can be expressed as

$$\begin{aligned} \langle\delta\hat{M}_z(\mathbf{q}; \omega)\rangle_0 &\sim (g\mu_B)^2 |\rho(\mathbf{q})|^2 \sum_{ii'} e^{-i\mathbf{q}\cdot(\mathbf{T}_i - \mathbf{T}_{i'})} \sum_{\sigma\sigma'} \sigma\sigma' \chi_{\hat{S}_z^i \hat{S}_z^{i'}}^{\sigma\sigma\sigma'\sigma'}(\omega) h_z(\mathbf{q}; \omega) \\ &= (g\mu_B)^2 |\rho(\mathbf{q})|^2 \chi_{\hat{S}_z \hat{S}_z}(\mathbf{q}; \omega) h_z(\mathbf{q}; \omega), \end{aligned}$$

where $\sigma = 1$ for spin up and $\sigma = -1$ for spin down and

$$\chi_{\hat{S}_z \hat{S}_z}(\mathbf{q}; \omega) = i \int dt e^{i\omega t} \left\langle \left[\hat{S}_z(\mathbf{q}; t), \hat{S}_z(-\mathbf{q}; 0) \right] \right\rangle_0 \Theta(t).$$

Later in the lecture we will discuss the case of the one-band Hubbard model; we will focus on the response function for site operators, $\chi_{\hat{S}_z^i \hat{S}_z^{i'}}(\omega)$ and its Fourier transform $\chi_{\hat{S}_z \hat{S}_z}(\mathbf{q}; \omega)$ and do not further discuss the prefactor $\rho(\mathbf{q})$. For multi-orbital systems with well-defined localized spins but quenching of the angular momentum, e.g., $3d$ transition-metal oxides that are Mott insulators [22], the magnetization density can still be expressed via Eq. (11). We have, however, to replace $\rho(\mathbf{r})$ with the normalized spin density at the atomic site, $\rho_s(\mathbf{r})$, originating from the unpaired electrons in the l shell. Thus

$$\hat{M}_z(\mathbf{r}) \sim -g\mu_B \rho_s(\mathbf{r}) \sum_i \hat{S}_z^i.$$

Then we have

$$\begin{aligned} \langle\delta\hat{M}_z(\mathbf{q}; \omega)\rangle_0 &\sim (g\mu_B)^2 |\rho_s(\mathbf{q})|^2 \sum_{ii'} e^{-i\mathbf{q}\cdot(\mathbf{T}_i - \mathbf{T}_{i'})} \sum_{\sigma\sigma'} \sigma\sigma' \chi_{\hat{S}_z^i \hat{S}_z^{i'}}^{\sigma\sigma\sigma'\sigma'}(\omega) h_z(\mathbf{q}; \omega) \\ &= (g\mu_B)^2 |\rho_s(\mathbf{q})|^2 \chi_{\hat{S}_z \hat{S}_z}(\mathbf{q}; \omega) h_z(\mathbf{q}; \omega). \end{aligned}$$

The Fourier transform of the spin density $\rho_s(\mathbf{q})$ is the so-called atomic form factor and can be probed via, e.g., neutron scattering experiments.

2.3 Analytic properties of $\chi(\mathbf{q}; \omega)$

The Fourier transform of $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{r}; t)$ to momentum space can be written as

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t) = i \left\langle \left[\Delta \hat{P}_\nu(\mathbf{q}; t), \Delta \hat{O}_{\nu'}(-\mathbf{q}; 0) \right] \right\rangle_0 \Theta(t).$$

Let us assume that $\{\Psi_n^N\}$ is the full set of eigenvectors of the Hamiltonian \hat{H} for N electrons and that the corresponding eigenenergies are $\{E_n^N\}$. Let us also define the matrix elements

$$P_\nu^{nm}(\mathbf{q}) = \langle \Psi_n^N | \Delta \hat{P}_\nu(\mathbf{q}; 0) | \Psi_m^N \rangle,$$

$$O_{\nu'}^{mn}(\mathbf{q}) = \langle \Psi_m^N | \Delta \hat{O}_{\nu'}(\mathbf{q}; 0) | \Psi_n^N \rangle.$$

We can then rewrite $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t)$ as follows

$$\begin{aligned} \chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) &= \frac{i}{Z} \int_0^\infty dt e^{i\omega t} \text{Tr} \left\{ e^{-\beta(\hat{H}-\mu\hat{N})} \left[\Delta \hat{P}_\nu(\mathbf{q}; t), \Delta \hat{O}_{\nu'}(-\mathbf{q}; 0) \right] \right\} \\ &= \frac{i}{Z} \sum_{nm} \int_0^\infty dt e^{-\beta(E_n^N - \mu N)} e^{i(\omega + E_n^N - E_m^N)t} P_\nu^{nm}(\mathbf{q}) O_{\nu'}^{mn}(-\mathbf{q}) \\ &\quad - \frac{i}{Z} \sum_{nm} \int_0^\infty dt e^{-\beta(E_n^N - \mu N)} e^{i(\omega + E_m^N - E_n^N)t} O_{\nu'}^{mn}(-\mathbf{q}) P_\nu^{mn}(\mathbf{q}) \\ &= \frac{i}{Z} \sum_{nm} F_{nm} \int_0^\infty dt e^{i(\omega - E_m^N + E_n^N)t} \hat{P}_\nu^{nm}(\mathbf{q}) \hat{O}_{\nu'}^{mn}(-\mathbf{q}), \end{aligned}$$

where $F_{nm} = e^{-\beta(E_n^N - \mu N)} - e^{-\beta(E_m^N - \mu N)}$. The time integral can be obtained from the formula

$$I(x) = \int_0^\infty e^{ixt} dt = \frac{i}{x + i\delta},$$

where δ is a positive infinitesimal; we thus arrive to the final expression

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = \frac{1}{Z} \sum_{nm} \frac{e^{-\beta(E_n^N - \mu N)} - e^{-\beta(E_m^N - \mu N)}}{E_m^N - E_n^N - \omega - i\delta} P_\nu^{nm}(\mathbf{q}) O_{\nu'}^{mn}(-\mathbf{q}). \quad (12)$$

This equation shows that the complex function $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; z)$, obtained from $\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)$ replacing ω with the complex variable z , thanks to the positive infinitesimal δ , is analytic in the upper half and has poles in the lower half of the complex plane. The fact that the response function is analytic is a direct consequence of causality; the fact that it is analytic in the upper instead than in the lower half of the complex plane is a consequence of our conventions on the signs of the exponents in the Fourier transform.

Up to now we did not make any assumption on the properties of the operators \hat{O}_ν and \hat{P}_ν ; if, however, they are Hermitian, one can show, starting from Eq. (12), that

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = \left[\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(-\mathbf{q}; -\omega) \right]^*$$

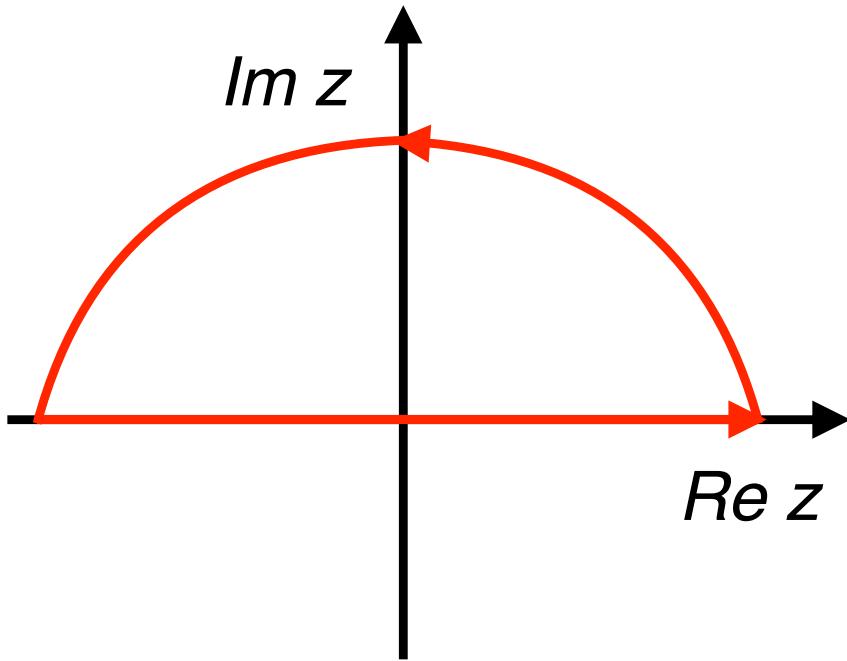


Fig. 3: The semicircular contour \mathcal{C} in the upper complex plane.

Therefore if we split the response function into a real and an imaginary part,

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = \operatorname{Re} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)] + i \operatorname{Im} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)],$$

the two components should satisfy the relations

$$\begin{aligned} \operatorname{Re} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)] &= \operatorname{Re} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(-\mathbf{q}; -\omega)], \\ \operatorname{Im} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)] &= -\operatorname{Im} [\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(-\mathbf{q}; -\omega)]. \end{aligned}$$

Thus the real part of the linear response function is even and the imaginary part is odd in ω .

2.4 Kramers-Kronig relations and sum rules

The Kramers-Kronig relations are valid for any retarded response function $\chi(\mathbf{q}; \omega)$. They follow from the fact that the complex function $\chi(\mathbf{q}; z)$ is analytic in the upper half of the complex plane, a property that we have just proved, and vanishes in the limit $|z| \rightarrow \infty$. Let us consider the integral on the real axis

$$I_{\mathcal{R}} = \int_{-\infty}^{+\infty} \frac{\chi(\mathbf{q}; \omega')}{\omega' - \omega + i\delta} d\omega',$$

where δ is an infinitesimally small positive number. The integrand is a complex analytic function in the upper part of the complex plane, and therefore, because of the Cauchy integral theorem, the integral on any closed contour \mathcal{C} in that half-plane has to vanish

$$I_{\mathcal{C}} = \oint_{\mathcal{C}} \frac{\chi(\mathbf{q}; z)}{z - \omega + i\delta} dz = 0.$$

Let us take as contour \mathcal{C} the semicircle shown in Fig. 3. If $\chi(\mathbf{q}; z)$ vanishes as $1/|z|$ or faster for $|z| \rightarrow \infty$, from $I_{\mathcal{C}} = 0$ it also follows that $I_{\mathcal{R}} = 0$. To ensure that indeed this condition is met, we subtract from the real part² of the susceptibility its infinite frequency limit $\text{Re}[\chi(\mathbf{q}; \infty)]$. Next, we rewrite $I_{\mathcal{R}}$ by using the Sokhotski-Plemelj formula

$$\frac{1}{\omega + i\delta} = \mathcal{P} \frac{1}{\omega} - i\pi\delta(\omega)$$

where \mathcal{P} is the Cauchy principal value

$$\int_{-\infty}^{\infty} \mathcal{P} \frac{1}{\omega} d\omega = \int_{-\infty}^{-\varepsilon} \frac{1}{\omega} d\omega + \int_{\varepsilon}^{\infty} \frac{1}{\omega} d\omega,$$

and $\delta(\omega)$ the Dirac delta function. Thus we obtain the Cauchy relation

$$I_{\mathcal{R}} = \mathcal{P} \int_{-\infty}^{+\infty} \frac{\chi(\mathbf{q}; \omega') - \text{Re}[\chi(\mathbf{q}; \infty)]}{\omega' - \omega} d\omega' - i\pi \{ \chi(\mathbf{q}; \omega) - \text{Re}[\chi(\mathbf{q}; \infty)] \}$$

As a consequence the real and imaginary part of the susceptibility are the Hilbert transform of each other, hence they satisfy the so-called Kramers-Kronig relations

$$\begin{aligned} \text{Re}[\chi(\mathbf{q}; \omega)] - \text{Re}[\chi(\mathbf{q}; \infty)] &= \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{+\infty} \frac{\text{Im}[\chi(\mathbf{q}; \omega')]}{\omega' - \omega} d\omega', \\ \text{Im}[\chi(\mathbf{q}; \omega)] &= -\frac{1}{\pi} \mathcal{P} \int_{-\infty}^{+\infty} \frac{\text{Re}[\chi(\mathbf{q}; \omega')] - \text{Re}[\chi(\mathbf{q}; \infty)]}{\omega' - \omega} d\omega'. \end{aligned}$$

The first Kramers-Kronig relation yields the sum rule

$$\text{Re}[\chi(\mathbf{q}; \omega = 0)] - \text{Re}[\chi(\mathbf{q}; \infty)] = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{+\infty} \frac{\text{Im}[\chi(\mathbf{q}; \omega')]}{\omega'} d\omega'. \quad (13)$$

In the $\mathbf{q} = \mathbf{0}$ limit, Eq. (13) is known as *thermodynamic sum rule* and

$$\chi_{\nu\nu'}(\mathbf{0}; 0) = \lim_{h_{\nu'} \rightarrow 0} \frac{\partial \langle P_{\nu} \rangle}{\partial h_{\nu'}},$$

is the response to a static and uniform perturbation, $h_{\nu'} = h_{\nu'}(\mathbf{0}; 0)$.

Finally, if $\hat{O}_{\nu'} \propto \hat{P}_{\nu}'$ and thus the product $P_{\nu}^{nm}(\mathbf{q}) O_{\nu'}^{mn}(-\mathbf{q})$ is real, by using (12) for the left-hand side and the invariant properties of the trace under cyclic permutations for the right-hand side, one can show that

$$\frac{2}{\pi} \int_0^{\infty} \omega \text{Im} \left[\chi_{\hat{P}_{\nu} \hat{O}_{\nu'}}(\mathbf{q}; \omega) \right] d\omega = \left\langle \left[[\hat{P}_{\nu}, \hat{H}], \hat{O}_{\nu'} \right] \right\rangle_0,$$

a relation known as Thomas-Reich-Kuhn or *f-sum rule*.

²For Hermitian operators the real part of the function $\chi(\mathbf{q}; \omega)$ is even in ω ; thus, the infinite frequency limit of $\chi(\mathbf{q}; \omega)$ could, in principle, be a constant.

2.5 Fluctuation-dissipation theorem

The fluctuation-dissipation theorem is a relation between the correlation function $\mathcal{S}(\mathbf{q}; \omega)$, which essentially describes fluctuations at equilibrium, and the linear response function $\chi(\mathbf{q}; \omega)$, which describes dissipative effects or relaxation phenomena as well. To derive it, let us start from the spatial Fourier transform of the correlation function

$$\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t) = \left\langle \Delta \hat{P}_\nu(\mathbf{q}; t) \Delta \hat{O}_{\nu'}(-\mathbf{q}) \right\rangle_0.$$

This equation, together with the definition of the susceptibility, yields the relation

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t) = i[\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t) - \mathcal{S}_{\hat{O}_{\nu'} \hat{P}_\nu}(-\mathbf{q}; -t)]\Theta(t). \quad (14)$$

Let us now take the time Fourier transform of the correlation function and express it in term of a full set of eigenvectors of the Hamiltonian, as we have previously done for the susceptibility

$$\begin{aligned} \mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) &= \int_{-\infty}^{\infty} dt e^{i\omega t} \langle \Delta \hat{P}_\nu(\mathbf{q}; t) \Delta \hat{O}_{\nu'}(-\mathbf{q}; 0) \rangle_0 \\ &= \frac{1}{Z} \sum_{nm} \int_{-\infty}^{\infty} dt e^{i(\omega + E_n^N - E_m^N)t} e^{-\beta(E_n^N - \mu^N)} P_\nu^{nm}(\mathbf{q}) O_{\nu'}^{mn}(-\mathbf{q}) \\ &= \frac{2\pi}{Z} \sum_{nm} e^{-\beta(E_n^N - \mu^N)} P_\nu^{nm}(\mathbf{q}) O_{\nu'}^{mn}(-\mathbf{q}) \delta(\omega - E_m^N + E_n^N). \end{aligned}$$

If we exchange first the order of the operators and later the indices n and m in the sum, we find

$$\mathcal{S}_{\hat{O}_{\nu'} \hat{P}_\nu}(\mathbf{q}; \omega) = \frac{2\pi}{Z} \sum_{nm} e^{-\beta(E_m^N - \mu^N)} P_\nu^{nm}(-\mathbf{q}) O_{\nu'}^{mn}(\mathbf{q}) \delta(\omega - E_n^N + E_m^N).$$

The correlation function therefore satisfies the *principle of detailed balance*

$$\mathcal{S}_{\hat{O}_{\nu'} \hat{P}_\nu}(-\mathbf{q}; -\omega) = e^{-\beta\omega} \mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega).$$

The relation above can be understood as follows. If $\omega > 0$, the correlation function $\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)$ describes the probability $P_{n \rightarrow m} \propto n(E_n)[1 - n(E_m)]$ that the system is excited from an initial state with energy E_n to a final state with higher energy $E_m = E_n + \omega$. Instead, $\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(-\mathbf{q}; -\omega)$, describes the probability $P_{m \rightarrow n} \propto n(E_m)[1 - n(E_n)]$ that the system goes from the initial state with energy E_m to a final state with lower energy $E_n = E_m - \omega$. The probability $P_{m \rightarrow n}$ is lower than $P_{n \rightarrow m}$ by the factor $e^{-\beta\omega}$.

We are now ready to Fourier transform Eq. (14). To do this, first we replace the step function with its Fourier representation

$$\Theta(t) = i \int_{-\infty}^{\infty} \frac{d\omega''}{2\pi} e^{-i\omega''t} \frac{1}{\omega'' + i\delta},$$

and do the same with the correlation function,

$$\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t) = \int \frac{d\omega'}{2\pi} e^{-i\omega't} \mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega');$$

next we Fourier transform in t the left- and right-hand side of Eq. (14), obtaining

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = - \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \frac{\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega') - \mathcal{S}_{\hat{O}_{\nu'} \hat{P}_\nu}(-\mathbf{q}; -\omega')}{\omega - \omega' + i\delta}.$$

Finally, via the principle of detailed balance we arrive at the expression

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = - \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega') \frac{1 - e^{-\beta\omega'}}{\omega - \omega' + i\delta}.$$

Thus, if $\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega')$ is real, as happens when $\hat{O}_{\nu'} \propto \hat{P}_\nu^\dagger$, the following relation holds

$$\mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega) = 2(1 + n_B)\text{Im}[\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega)], \quad n_B(\omega) = \frac{1}{e^{\beta\omega} - 1}.$$

This is the fluctuation-dissipation theorem. The left-hand side yields the spectrum of spontaneous fluctuations and the right-hand side the energy dissipation. When $k_B T$ is large, it follows from the first Kramers-Kronig relation, Eq. (13), and the fluctuation-dissipation theorem that

$$\text{Re}[\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \omega = 0)] - \text{Re}[\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; \infty)] \sim \frac{1}{k_B T} \mathcal{S}_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; t = 0).$$

2.6 Single-particle Green function

2.6.1 Definitions

In the non-interacting limit or within certain approximations, the susceptibility can be written in terms of single-particle Green functions. It is therefore important to introduce the latter and their properties. Let us first define the Green function or *propagator*

$$G_{\alpha\alpha'}(t, t') = -i \left\langle \mathcal{T} c_\alpha(t) c_{\alpha'}^\dagger(t') \right\rangle_0,$$

where \mathcal{T} is the time-ordering operator, which orders the operators in decreasing time from left to right. The indices α and α' are flavors, and $c_\alpha^\dagger(c_\alpha)$ is a fermionic creation (annihilation) operator. When the Hamiltonian is time-independent,

$$G_{\alpha\alpha'}(t, t') = G_{\alpha\alpha'}(t - t').$$

It is useful to express the Green function using a full set of eigenvectors $\{\Psi_n^N\}$ of the Hamiltonian, as we have done for the susceptibility and the correlation function. We obtain

$$G_{\alpha\alpha'}(t) = -\frac{i}{Z} \sum_{Nnm} e^{-\beta(E_n^N - \mu N)} \begin{cases} e^{i(E_n^N - E_m^{N+1} + \mu)t} \langle \Psi_n^N | c_\alpha | \Psi_m^{N+1} \rangle \langle \Psi_m^{N+1} | c_{\alpha'}^\dagger | \Psi_n^N \rangle & t > 0 \\ -e^{i(-E_n^N + E_m^{N-1} + \mu)t} \langle \Psi_n^N | c_{\alpha'}^\dagger | \Psi_m^{N-1} \rangle \langle \Psi_m^{N-1} | c_\alpha | \Psi_n^N \rangle & t < 0 \end{cases}$$

Let us define $C_\alpha^{nm} = \langle \Psi_n^N | c_\alpha | \Psi_m^{N+1} \rangle$ and $\overline{C_{\alpha'}^{nm}} = \langle \Psi_m^{N+1} | c_{\alpha'}^\dagger | \Psi_n^N \rangle$. If, for negative times, we replace N with $N + 1$ and exchange m and n , we obtain

$$G_{\alpha\alpha'}(t) = -\frac{i}{Z} \sum_{Nnm} e^{-\beta(E_n^N - \mu N)} \overline{C_{\alpha'}^{nm}} C_\alpha^{nm} e^{i(E_n^N - E_m^{N+1} + \mu)t} \left[\Theta(t) - e^{-\beta(E_m^{N+1} - E_n^N - \mu)} \Theta(-t) \right].$$

The Fourier transform with respect to time yields

$$G_{\alpha\alpha'}(\omega) = \frac{1}{Z} \sum_{Nnm} e^{-\beta(E_n^N - \mu N)} \overline{C_{\alpha'}^{nm}} C_{\alpha}^{nm} \\ \times \left\{ \frac{1}{\omega - E_m^{N+1} + E_n^N + \mu + i\delta} + \frac{e^{-\beta(E_m^{N+1} - E_n^N - \mu)}}{\omega - E_m^{N+1} + E_n^N + \mu - i\delta} \right\}.$$

Let us define the function

$$A_{\alpha\alpha'}(\omega) = \frac{2\pi}{Z} \sum_{Nnm} e^{-\beta(E_n^N - \mu N)} \overline{C_{\alpha'}^{nm}} C_{\alpha}^{nm} [1 + e^{-\beta\omega}] \delta(\omega - E_m^{N+1} + E_n^N + \mu).$$

In terms of $A_{\alpha\beta}(\omega)$, the Green function takes the simpler form

$$G_{\alpha\alpha'}(\omega) = \int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi} \left[\frac{1 - n(\omega')}{\omega - \omega' + i\delta} + \frac{n(\omega')}{\omega - \omega' - i\delta} \right] A_{\alpha\alpha'}(\omega').$$

This expression is known as *Lehmann representation*, and $A_{\alpha\beta}(\omega)$ is called *spectral function*. It is often useful to introduce the retarded and advanced Green functions $G_{\alpha\alpha'}^R(t, t')$ and $G_{\alpha\alpha'}^A(t, t')$. The first is given by

$$G_{\alpha\alpha'}^R(t, t') = -i\Theta(t - t') \left\langle \left[c_{\alpha}(t), c_{\alpha'}^{\dagger}(t') \right] \right\rangle_0,$$

and the latter by

$$G_{\alpha\alpha'}^A(t, t') = i\Theta(t' - t) \left\langle \left[c_{\alpha}(t), c_{\alpha'}^{\dagger}(t') \right] \right\rangle_0.$$

If we Fourier transform them in time we have

$$G_{\alpha\alpha'}^R(\omega) = \int_{-\infty}^{+\infty} d\omega' \frac{1}{2\pi} \left[\frac{1}{\omega - \omega' + i\delta} \right] A_{\alpha\alpha'}(\omega'), \\ G_{\alpha\alpha'}^A(\omega) = \int_{-\infty}^{+\infty} d\omega' \frac{1}{2\pi} \left[\frac{1}{\omega - \omega' - i\delta} \right] A_{\alpha\alpha'}(\omega'),$$

and therefore

$$A_{\alpha\alpha'}(\omega) = i [G_{\alpha\alpha'}^R(\omega) - G_{\alpha\alpha'}^A(\omega)].$$

2.6.2 Temperature Green function

To build a consistent many-body perturbation theory at finite temperature it is convenient to introduce an imaginary time variable, τ . The imaginary-time Green function, given by

$$G_{\alpha\alpha'}(\tau) = -\langle \mathcal{T} c_{\alpha}(\tau_1) c_{\alpha'}^{\dagger}(\tau_2) \rangle_0 = -\frac{1}{Z} \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} \mathcal{T} c_{\alpha}(\tau_1) c_{\alpha'}^{\dagger}(\tau_2) \right],$$

is known as the temperature or Matsubara Green function. In this expression, \mathcal{T} is again the time-ordering operator; the imaginary time fermionic operators $o(\tau) = c(\tau), c^{\dagger}(\tau)$ are given by

$$o(\tau) = e^{\tau(\hat{H} - \mu \hat{N})} o e^{-\tau(\hat{H} - \mu \hat{N})}.$$

The vector τ is defined as $\tau = (\tau_1, \tau_2)$. Writing explicitly the action of \mathcal{T} , we obtain

$$G_{\alpha\alpha'}(\tau) = -\Theta(\tau_1 - \tau_2)\langle c_\alpha(\tau_1)c_{\alpha'}^\dagger(\tau_2)\rangle + \Theta(\tau_2 - \tau_1)\langle c_{\alpha'}^\dagger(\tau_2)c_\alpha(\tau_1)\rangle.$$

If the Hamiltonian is time-independent,

$$G_{\alpha\alpha'}(\tau) = G_{\alpha\alpha'}(\tau_1 - \tau_2).$$

The imaginary-time Green function is well defined only in the interval $-\beta < \tau_1 - \tau_2 = \tau < \beta$. This can be seen by writing it explicitly in a basis of eigenvectors of the Hamiltonian, $\{\Psi_n^N\}$

$$G_{\alpha\alpha'}(\tau) = \frac{1}{Z} \sum_{Nnm} \langle \Psi_n^N | c_\alpha | \Psi_m^{N+1} \rangle \langle \Psi_m^{N+1} | c_{\alpha'}^\dagger | \Psi_n^N \rangle e^{-\beta(E_n^N - \mu N)} \begin{cases} -e^{(E_n^N - E_m^{N+1} + \mu)\tau} & \tau > 0 \\ e^{-(E_n^N - E_m^{N+1} + \mu)(-\beta - \tau)} & \tau < 0 \end{cases}$$

For imaginary times outside the interval $-\beta < \tau < \beta$, the high energy states would give divergent contributions. Either from the expression above or from the definition of $G_{\alpha\alpha}(\tau)$ and the invariance of the trace under cyclic permutation of operators, one can show that $G_{\alpha\alpha'}(\tau)$ has the following symmetry property

$$G_{\alpha\alpha'}(\tau) = -G_{\alpha\alpha'}(\tau + \beta) \quad \text{for } -\beta < \tau < 0.$$

Finally, if n_α is the number of electrons for flavor α , one can show that

$$G_{\alpha\alpha}(\tau \rightarrow 0^+) = -1 + n_\alpha, \quad G_{\alpha\alpha}(\tau \rightarrow \beta^-) = -n_\alpha.$$

For negative times, we have

$$G_{\alpha\alpha}(\tau \rightarrow 0^-) = n_\alpha, \quad G_{\alpha\alpha}(\tau \rightarrow -\beta^-) = 1 - n_\alpha.$$

Thus, $G_{\alpha\alpha}(\tau)$ is discontinuous at $\tau = 0$ because of time ordering. It is at this point convenient to introduce a generalized imaginary-time Green function $\tilde{G}_{\alpha\alpha'}(\tau)$ defined for any τ ,

$$\tilde{G}_{\alpha\alpha'}(\tau_1 \pm n_1\beta, \tau_2 \pm n_2\beta) \equiv (-1)^{n_1+n_2} G_{\alpha\alpha'}(\tau_1, \tau_2),$$

where n_1 and n_2 are integers. The Green function $\tilde{G}_{\alpha\alpha'}(\tau_1, \tau_2) = \tilde{G}_{\alpha\alpha'}(\tau_1 - \tau_2)$ is, by construction, antiperiodic with period β in both τ_1 and τ_2 and in the time difference $\tau = \tau_1 - \tau_2$. From now on we will work with $\tilde{G}_{\alpha\alpha'}(\tau)$, and therefore for simplicity we rename it $G_{\alpha\alpha'}(\tau)$. Thanks to its periodicity, $G_{\alpha\alpha'}(\tau)$ can be written as the Fourier series

$$G_{\alpha\alpha'}(\tau) = \frac{1}{\beta} \sum_{n=-\infty}^{+\infty} e^{-i\nu_n\tau} G_{\alpha\alpha'}(i\nu_n),$$

where the frequency ν_n is given by 2π multiplied by an integer multiple of the inverse period, $1/2\beta$. For fermionic Green functions only the odd Matsubara frequencies, $\nu_n = \pi(2n+1)/\beta$, for which $e^{\pm i\nu_n\beta} = -1$, yield finite Fourier coefficients, given by

$$G_{\alpha\alpha'}(i\nu_n) = \frac{1}{2} \int_{-\beta}^{\beta} d\tau e^{i\nu_n\tau} G_{\alpha\alpha'}(\tau) = \frac{1}{2}(1 - e^{-i\nu_n\beta}) \int_0^{\beta} d\tau e^{i\nu_n\tau} G_{\alpha\alpha'}(\tau) = \int_0^{\beta} d\tau e^{i\nu_n\tau} G_{\alpha\alpha'}(\tau).$$

| $g_r(\tau; x)$ | $g_r(\nu_n; x) = \int_0^\beta e^{i\nu_n\tau} g_r(\tau; x) d\tau$ |
|------------------------------------|---|
| $g_0(\tau; x) = e^{-x\tau}$ | $[i\nu_n - x]^{-1} [n_\sigma(x) - 1]^{-1}$ |
| $g_1(\tau; x) = \tau e^{-x\tau}$ | $\beta n_\sigma(x) [i\nu_n - x]^{-1} [n_\sigma(x) - 1]^{-1} - [i\nu_n - x]^{-2} [n_\sigma(x) - 1]^{-1}$ |
| $g_r(\tau; x) = \tau^r e^{-x\tau}$ | $\beta^r n_\sigma(x) g_0(i\nu_n; x) - r [i\nu_n - x]^{-1} g_{r-1}(i\nu_n; x)$ |

Table 1: Some of the most common Matsubara Fourier transforms (fermionic case). The function $n_\sigma(x)$ is the Fermi-Dirac distribution function $n_\sigma(x) = 1/(1 + e^{x\beta})$; the parameter x is a real number. For fermionic Matsubara frequencies $e^{i\nu_n\beta} = -1$.

It is interesting to point out the relation between the imaginary-time Green function and the actual retarded and advanced Green functions. One can show that they correspond to different analytic continuations to the real axis

$$\begin{aligned} G_{\alpha\alpha'}^R(\omega) &= G_{\alpha\alpha'}(i\nu_n)|_{i\nu_n \rightarrow \omega + i\delta} \\ G_{\alpha\alpha'}^A(\omega) &= G_{\alpha\alpha'}(i\nu_n)|_{i\nu_n \rightarrow \omega - i\delta}. \end{aligned}$$

The odd or fermionic Matsubara frequencies $i\nu_n$ are also the poles of the Fermi-Dirac distribution function. The even or bosonic Matsubara frequencies, $i\omega_m = \pi(2m)/\beta$, are, correspondingly, the poles of the Bose-Einstein distribution function.

2.6.3 One-band Hubbard model: Non-interacting limit

Let us now consider as a representative case the one-band Hubbard model. In the non-interacting limit the Hubbard Hamiltonian is given by Eq. (5), which we rewrite here for convenience

$$\hat{H}_0 = \sum_{\sigma} \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}\sigma}.$$

For high-temperature superconducting cuprates, the dispersion $\varepsilon_{\mathbf{k}}$ describing the Cu d $x^2 - y^2$ band crossing the Fermi level is

$$\varepsilon_{\mathbf{k}} = -2t(\cos k_x a + \cos k_y a) + 4t' \cos k_x a \cos k_y a + \dots, \quad (15)$$

where a is the lattice constant, t the nearest neighbor hopping integral and t' the next-nearest neighbor hopping integral [9].

The imaginary-time Green function for the Hamiltonian \hat{H}_0 can be calculated analytically

$$\begin{aligned} \mathcal{G}_{\mathbf{k}\sigma}(\tau) &= -\left\langle \mathcal{T} \left[c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma}^\dagger(0) \right] \right\rangle_0 \\ &= -\Theta(\tau) \frac{1}{Z} \text{Tr} \left[e^{-\beta(\hat{H}_0 - \mu\hat{N})} c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma}^\dagger(0) \right] + \Theta(-\tau) \frac{1}{Z} \text{Tr} \left[e^{-\beta(\hat{H}_0 - \mu\hat{N})} c_{\mathbf{k}\sigma}^\dagger(0) c_{\mathbf{k}\sigma}(\tau) \right] \\ &= -[\Theta(\tau)(1 - n_\sigma(\varepsilon_{\mathbf{k}})) - \Theta(-\tau)n_\sigma(\varepsilon_{\mathbf{k}})] e^{-(\varepsilon_{\mathbf{k}} - \mu)\tau}, \end{aligned} \quad (16)$$

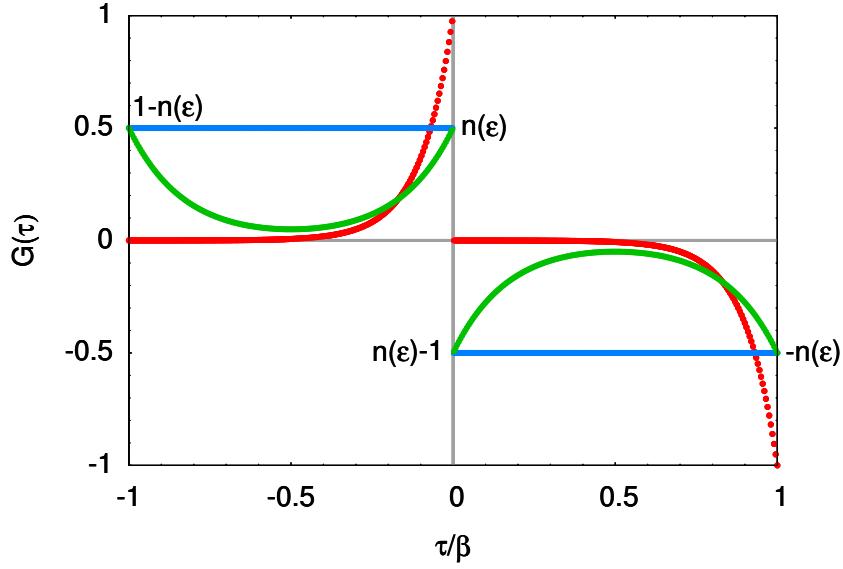


Fig. 4: The function $\mathcal{G}_{k\sigma}(\tau)$ defined in Eq. (16) for a state well below the Fermi level (red) and at the Fermi level (blue) and $\beta = 2 \text{ eV}^{-1}$. The green line shows the atomic Green function $G(\tau)$ from Eq. (19) calculated for $U = 6 \text{ eV}$.

where $n_\sigma(\varepsilon_k)$ is the Fermi-Dirac distribution function

$$n_\sigma(\varepsilon_k) = \frac{1}{1 + e^{\beta(\varepsilon_k - \mu)}}.$$

The Matsubara Fourier transform of $\mathcal{G}_{k\sigma}(\tau)$ is simple to obtain since

$$\int_0^\beta e^{(i\nu_n - x)\tau} d\tau = \frac{1}{i\nu_n - x} [e^{(i\nu_n - x)\beta} - 1] = -\frac{1}{i\nu_n - x} [e^{-x\beta} + 1] = \frac{1}{i\nu_n - x} \frac{1}{n_\sigma(x) - 1}.$$

This result, together with other useful Matsubara Fourier transforms, can be found in Tab. 1. Thus

$$\mathcal{G}_{k\sigma}(i\nu_n) = \frac{1}{i\nu_n - \varepsilon_k + \mu}.$$

It is often necessary to perform sums of Green functions or of product of Green functions over the Matsubara frequencies. To see how these can be calculated, let us consider the integral

$$I_C = \frac{1}{2\pi i} \oint_C \mathcal{F}_{k\sigma}(z) n_\sigma(z) e^{z\tau} dz, \quad (17)$$

where $0 < \tau < \beta$ and $n_\sigma(z)$ is the Fermi function, which has poles for $z = i\nu_n$. We assume that $\mathcal{F}_{k\sigma}(z)$ is a complex function, analytic everywhere except at some poles $\{z_p\}$, which differ from the Fermionic Matsubara frequencies; for example, $\mathcal{F}_{k\sigma}(z)$ could be the Green function $\mathcal{G}_{k\sigma}(z)$. We define the contour C (see Fig. 5) as a circle in the full complex plane, centered at the origin and including the poles of the integrand. The integral I_C is zero because the integrand vanishes exponentially for $|z| \rightarrow \infty$. Using Cauchy's integral theorem we then have

$$\frac{1}{\beta} \sum_n e^{i\nu_n \tau} \mathcal{F}_{k\sigma}(i\nu_n) = \sum_{z_p} \text{Res} [\mathcal{F}_{k\sigma}(z_p)] n_\sigma(z_p) e^{z_p \tau}, \quad (18)$$

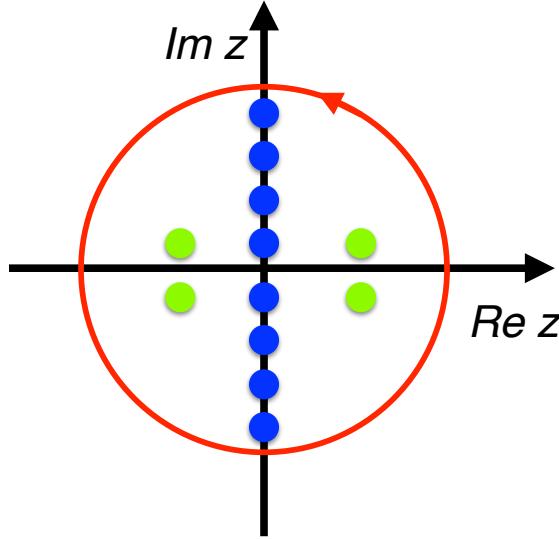


Fig. 5: The contour \mathcal{C} used to perform the integral I_C defined by Eq. (17). The blue circles represent poles of the Fermi function, i.e., Matsubara frequencies, and the green circles are representative poles of the function $\mathcal{F}_{k\sigma}(z)$. Since the integral vanishes for $|z| \rightarrow \infty$, the sum of the contributions from all the poles must add up to zero for an infinitely large contour \mathcal{C} .

where we used the fact that $\text{Res } [n_\sigma(i\nu_n)] = -\frac{1}{\beta}$. Let us now apply this result in some typical cases. If $\mathcal{F}_{k\sigma}(z) = \mathcal{G}_{k\sigma}(z)$, remembering that $G_{\alpha\alpha}(0^-) = n_\alpha$ and $G_{\alpha\alpha}(0^+) = n_\alpha - 1$, we obtain

$$\begin{aligned} \frac{1}{\beta} \sum_n e^{-i\nu_n 0^-} \mathcal{G}_{k\sigma}(i\nu_n) &= \mathcal{G}_{k\sigma}(0^-) = n_\sigma(\varepsilon_k), \\ \frac{1}{\beta} \sum_n e^{-i\nu_n 0^+} \mathcal{G}_{k\sigma}(i\nu_n) &= \mathcal{G}_{k\sigma}(0^+) = n_\sigma(\varepsilon_k) - 1. \end{aligned}$$

In a similar way we can show that

$$\frac{1}{\beta} \sum_n e^{i\nu_n 0^+} \mathcal{G}_{k\sigma}(i\nu_n) \mathcal{G}_{k\sigma}(i\nu_n) = \frac{dn_\sigma(\varepsilon_k)}{d\varepsilon_k} = \beta n_\sigma(\varepsilon_k) [-1 + n_\sigma(\varepsilon_k)],$$

$$\frac{1}{\beta} \sum_n e^{i\nu_n 0^+} \mathcal{G}_{k\sigma}(i\nu_n) \mathcal{G}_{k+q\sigma}(i\nu_n + i\omega_m) = \frac{n(\varepsilon_{k+q}) - n(\varepsilon_k)}{i\omega_m + \varepsilon_{k+q} - \varepsilon_k},$$

where in the last relation $\omega_m = 2m\pi/\beta$ is a bosonic Matsubara frequency. In Tab. 2 we display some of the inverse Fourier transforms involving one-particle Green functions for a non-interacting system.

It is important to point out that, using Matsubara frequencies, the Lehmann representation takes the simple form

$$\mathcal{G}_{k\sigma}(i\nu_n) = \int_{-\infty}^{+\infty} A_{k\sigma}(\omega') \frac{1}{i\nu_n - \omega'} d\omega',$$

where the spectral function satisfies the relation

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} A_{k\sigma}(\omega') d\omega' = 1.$$

| $g_\alpha(\nu_n; x, y)$ | $g_\alpha(\tau; x, y) = \frac{1}{\beta} \sum_n e^{-i\nu_n \tau} g_\alpha(\nu_n; x, y)$ |
|--|--|
| $g_a(\nu_n; x, y) = [i\nu_n - x]^{-1}$ | $[n_\sigma(x) - 1]e^{-x\tau}$ |
| $g_b(\nu_n; x, y) = [i\nu_n - x]^{-2}$ | $n_\sigma(x)(\tau - \beta n_\sigma(x))e^{-x(\tau-\beta)}$ |
| $g_c(\nu_n; x, y) = [i\nu_n - x]^{-1} [i\nu_n - y]^{-1}$ | $-[e^{-x(\tau-\beta)}n_\sigma(x) - e^{-y(\tau-\beta)}n_\sigma(y)] [x - y]^{-1}$ |
| $g_d(\nu_n; x, y) = [i\nu_n - x]^{-1} [i\nu_n + x]^{-1}$ | $[g_a(\tau; x, y) - g_a(\tau; -x, y)]/2x$ |

Table 2: Some of the most common Matsubara Fourier transforms (fermionic case), obtained from Eq. (18). The function $n_\sigma(x)$ is the Fermi-Dirac distribution function $n_\sigma(x) = 1/(1+e^{x\beta})$. The parameters x and y are real numbers. For τ we consider the interval $(0, \beta)$.

Furthermore the normalized spectral function, $\tilde{A}_{k\sigma}(\varepsilon) = A_{k\sigma}(\varepsilon)/2\pi = -\frac{1}{\pi}\text{Im}[\mathcal{G}_{k\sigma}(\varepsilon)]$ is related to the density of states as follows

$$\rho_\sigma(\varepsilon) = \frac{1}{N_k} \sum_k \tilde{A}_{k\sigma}(\varepsilon).$$

These relations between $\tilde{A}_{k\sigma}(\varepsilon)$ and the density of states or the Green function are also valid for the interacting Hubbard model. In the non-interacting case

$$\tilde{A}_{k\sigma}(\varepsilon) = \delta(\omega - \varepsilon_k).$$

2.6.4 One-band Hubbard model: Atomic limit

Let us consider now the half-filled one-band Hubbard model in the atomic ($t = 0$) limit. Since the lattice sites are decoupled, we can focus on a single site and rewrite the Coulomb interaction in terms of the spin operator $\hat{S}_z = \frac{1}{2}[n_\uparrow - n_\downarrow]$ and the electron number operator $\hat{N} = n_\uparrow + n_\downarrow$. We obtain

$$\hat{H}_U = \varepsilon_d \sum_\sigma n_\sigma + U \left(\frac{\hat{N}^2}{4} - \hat{S}_z^2 \right).$$

This Hamiltonian describes an idealized single-level atom. This system has four states, $|0\rangle$, $c_\uparrow^\dagger |0\rangle$, $c_\downarrow^\dagger |0\rangle$, $c_\uparrow^\dagger c_\downarrow^\dagger |0\rangle$, with expectation values of the operator $\hat{H}_U - \mu\hat{N}$ equal to $0, -U/2, -U/2$ and 0 respectively. We can calculate the imaginary-time Green function for such a system analytically; it is sufficient to perform the calculation in the interval $(0, \beta)$, since we can reconstruct the Green function in the interval $(-\beta, 0)$ by using the antiperiodic properties. For $0 < \tau < \beta$

$$G_\sigma(\tau) = -\frac{1}{2} \frac{1}{1 + e^{\beta U/2}} [e^{\tau U/2} + e^{(\beta-\tau)U/2}]. \quad (19)$$

The Matsubara Fourier coefficients can be obtained via the integrals in Tab. 1. We find

$$G_\sigma(i\nu_n) = \frac{1}{2} \left[\frac{1}{i\nu_n + U/2} + \frac{1}{i\nu_n - U/2} \right]. \quad (20)$$

Thus the atomic Green function, as the non-interacting Green function, is the sum of functions with first order poles; the corresponding retarded Green function on the real axis can be obtained by analytic continuation replacing $i\nu_n$ with $\omega + i\delta$, with $\delta > 0$. To obtain $G_\sigma(\tau)$ from $G_\sigma(i\nu_n)$ one can use the Matsubara sums in Tab. 2.

2.7 Two-particle Green function

2.7.1 Generalized imaginary time Green function

The temperature Green function can also be defined for quadratic operators; this generalization is relevant for calculating the elements of the linear-response tensor. Let us consider the operators $\Delta\hat{P}_{\alpha\alpha'}(\tau_1, \tau_2)$ and $\Delta\hat{O}_{\gamma\gamma'}(\tau_1, \tau_2)$, with

$$\begin{aligned} \Delta\hat{P}_{\alpha\alpha'}(\tau_1, \tau_2) &= c_{\alpha'}^\dagger(\tau_2)c_\alpha(\tau_1) - \langle \mathcal{T}c_{\alpha'}^\dagger(\tau_2)c_\alpha(\tau_1) \rangle, \\ \Delta\hat{O}_{\gamma\gamma'}(\tau_3, \tau_4) &= c_{\gamma'}^\dagger(\tau_4)c_\gamma(\tau_3) - \langle \mathcal{T}c_{\gamma'}^\dagger(\tau_4)c_\gamma(\tau_3) \rangle. \end{aligned}$$

where α, α' and γ, γ' are, as usual, flavors. We define the temperature Green function for these operators as the two-particle Green function

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = \langle \mathcal{T}\Delta\hat{P}_{\alpha\alpha'}(\tau_1, \tau_2)\Delta\hat{O}_{\gamma\gamma'}(\tau_3, \tau_4) \rangle, \quad (21)$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3, \tau_4)$. From the invariance of the trace under cyclic permutation of operators one can show that, for a time-independent Hamiltonian,

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14}, \tau_{24}, \tau_{34}, 0),$$

where $\tau_{ji} = \tau_j - \tau_i$. In analogy with what we have seen in the single-particle case, the two-particle Green function at negative times ($-\beta < \tau_{j4} < 0$) can be obtained from the two-particle Green function at positive times ($0 < \tau_{j4} < \beta$). For example, if $-\beta < \tau_{14} < 0$

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14} + \beta, \tau_{24}, \tau_{34}, 0) = -\chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14}, \tau_{24}, \tau_{34}, 0),$$

and similar relations hold for $-\beta < \tau_{24} < 0$ or $-\beta < \tau_{34} < 0$. This can be shown using, once more, the invariance of the trace under cyclic permutation of the operators. As for the one-particle Green function, we can extend the two-particle Green function to any time interval by defining it periodic in each τ_{j4} with period β .

It is often convenient to express $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau})$ as a function of the three independent variables τ_{12} , τ_{23} and τ_{34} . We then rewrite it as follows

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = \mathcal{T} \text{Tr} \frac{e^{-\beta(\hat{H}-\mu\hat{N})}}{Z} \left[e^{\tau_{14}(\hat{H}-\mu\hat{N})} c_\alpha e^{-\tau_{12}(\hat{H}-\mu\hat{N})} c_{\alpha'}^\dagger e^{-\tau_{23}(\hat{H}-\mu\hat{N})} c_\gamma e^{-\tau_{34}(\hat{H}-\mu\hat{N})} c_{\gamma'}^\dagger \right],$$

with $\tau_{14} = \tau_{12} + \tau_{23} + \tau_{34}$, and $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{12}, \tau_{34}; \tau_{23})$. From this expression and the symmetry properties discussed above, it follows that the two-particle Green function is antiperiodic with period β in τ_{12} and τ_{34} and periodic with period β in τ_{23} (see Fig. 6 as example).

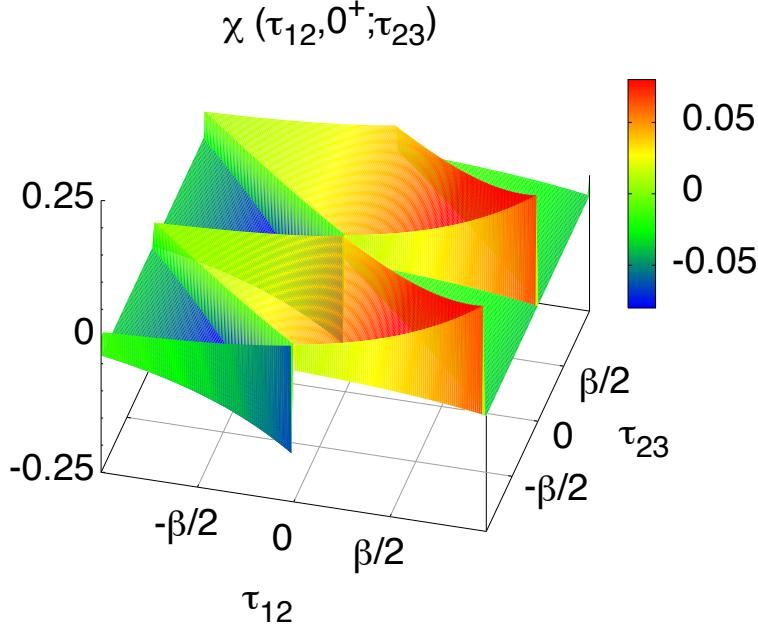


Fig. 6: The function $\chi(\tau_{12}, 0^+; \tau_{23}) = -G_{\alpha\alpha}(\tau_{12} + \tau_{23} + 0^+)G_{\alpha\alpha}(-\tau_{23})$ for a non-interacting system with $n \sim 0.8$ electrons; $\beta = 30 \text{ eV}^{-1}$. It is antiperiodic in τ_{12} and periodic in τ_{23} .

2.7.2 The Fourier transform $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\nu)$ and its symmetry properties

The Fourier transform of the imaginary-time two-particle Green function is given by

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\nu) = \frac{1}{16} \iiint d\tau e^{i\nu \cdot \tau} \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau), \quad (22)$$

where the frequency vector is $\nu = (\nu_1, \nu_2, \nu_3, \nu_4)$, and the imaginary times are chosen in the interval $(-\beta, \beta)$. From the fact that $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau) = \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14}, \tau_{24}, \tau_{34}, 0)$ we obtain $\nu_4 = -\nu_1 - \nu_2 - \nu_3$ (energy conservation), and thus only three frequencies are actually independent. Let us define $\nu = (\nu_n, -\nu_n - \omega_m, \nu_{n'} + \omega_m, -\nu_{n'})$ where ω_m is a bosonic frequency. Then

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\nu) = \chi_{n,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) = \frac{\beta}{8} \iiint d\tau e^{i[-\omega_m \tau_{23} + \nu_n \tau_{12} + \nu_{n'} \tau_{34}]} \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau),$$

where $\tau = (\tau_{14}, \tau_{24}, \tau_{34})$ and all integrals go from $-\beta$ to β . By using the antiperiodicity of the two-particle Green function in imaginary times we can further simplify this expression, obtaining

$$\chi_{n,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) = \beta \int_0^\beta d\tau_{14} \int_0^\beta d\tau_{24} \int_0^\beta d\tau_{34} e^{i[-\omega_m \tau_{23} + \nu_n \tau_{12} + \nu_{n'} \tau_{34}]} \chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14}, \tau_{24}, \tau_{34}, 0).$$

Let us now analyze the symmetry properties of $\chi_{n,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m)$. For simplicity, we consider here only the case in which the one-electron basis can be chosen such that $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau_{14}, \tau_{24}, \tau_{34}, 0)$ is real; indeed, this is the case for most strongly-correlated 3d transition metal-oxides, since their

Hamiltonian has typically time-reversal and even inversion symmetry and the spin-orbit interaction can be neglected. The complex conjugate is then given by

$$\left[\chi_{n,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) \right]^* = \chi_{-n-1,-n'-1}^{\alpha\alpha'\gamma\gamma'}(-i\omega_m),$$

where $\nu_{-n-1} = -\nu_n$, and $\nu_{-n'-1} = -\nu_{n'}$. Furthermore, if in the integral (22) we replace $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau})$ with its complex conjugate and then exchange $\tau_1 \leftrightarrow -\tau_4$ and $\tau_2 \leftrightarrow -\tau_3$, we find

$$\chi_{n,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) = \chi_{n',n}^{\gamma'\gamma\alpha'\alpha}(i\omega_m),$$

and hence, if $\alpha = \gamma'$, $\alpha' = \gamma$, $\nu_n = \nu'_n$ is a reflection axis for the absolute value

$$\left| \chi_{n,n'}^{\alpha\alpha'\alpha'\alpha}(i\omega_m) \right| = \left| \chi_{n',n}^{\alpha\alpha'\alpha'\alpha}(i\omega_m) \right|.$$

An additional reflection axis can be found by first shifting the frequency $\nu_n = \nu_l - \omega_m$

$$\chi_{l,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) = \frac{1}{16} \iiint d\boldsymbol{\tau} e^{i(-\omega_m\tau_{13} + \nu_l\tau_{12} + \nu_{n'}\tau_{34})} \chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}).$$

Since $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau})$ is invariant under particle exchange, if we exchange in the integrand $\tau_1 \leftrightarrow \tau_3$ and $\tau_2 \leftrightarrow \tau_4$, we have

$$\chi_{l,n'}^{\alpha\alpha'\gamma\gamma'}(i\omega_m) = \chi_{n',l}^{\gamma\gamma'\alpha\alpha'}(-i\omega_m),$$

so that, if $\alpha = \gamma$ and $\alpha' = \gamma'$, $\nu_{n+m} = -\nu_{n'}$ is a mirror line for the absolute value

$$\left| \chi_{n+m,n'}^{\alpha\alpha'\alpha\alpha'}(i\omega_m) \right| = \left| \chi_{-n'-1,-n-m-1}^{\alpha\alpha'\alpha\alpha'}(i\omega_m) \right|.$$

2.7.3 Non-interacting case: Wick's theorem

For a non-interacting system Wick's theorem holds. It states that high-order Green functions can be factorized into products of lower-order Green function. For the two-particle Green function Wick's factorization yields

$$\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = -\langle \mathcal{T}c_\alpha(\tau_1)c_{\gamma'}^\dagger(\tau_4) \rangle \langle \mathcal{T}c_\gamma(\tau_3)c_{\alpha'}^\dagger(\tau_2) \rangle = -G_{\alpha\gamma'}(\tau_{14})G_{\gamma\alpha'}(-\tau_{23}). \quad (23)$$

If the two-particle Green function is written in this form, its periodicity properties can be directly derived from those of the one-particle Green function. For example, using as independent variables τ_{12} , τ_{23} , and τ_{34} , since $\tau_{14} = \tau_{12} + \tau_{23} + \tau_{34}$, one can verify that the susceptibility is periodic in the time associated with the bosonic frequency, τ_{23} . As the single-particle Green function, the two-particle Green function has discontinuities associated with the jump coming from the time ordering operator. In Fig. 6 we see such jumps, e.g., along the line $\tau_{12} = -\tau_{23}$ (for which $\tau_{14} = 0^+$) and $\tau_{12} = -\tau_{23} \pm \beta$ (for which $\tau_{14} = \beta^+$ or $\tau_{14} = \beta^-$).

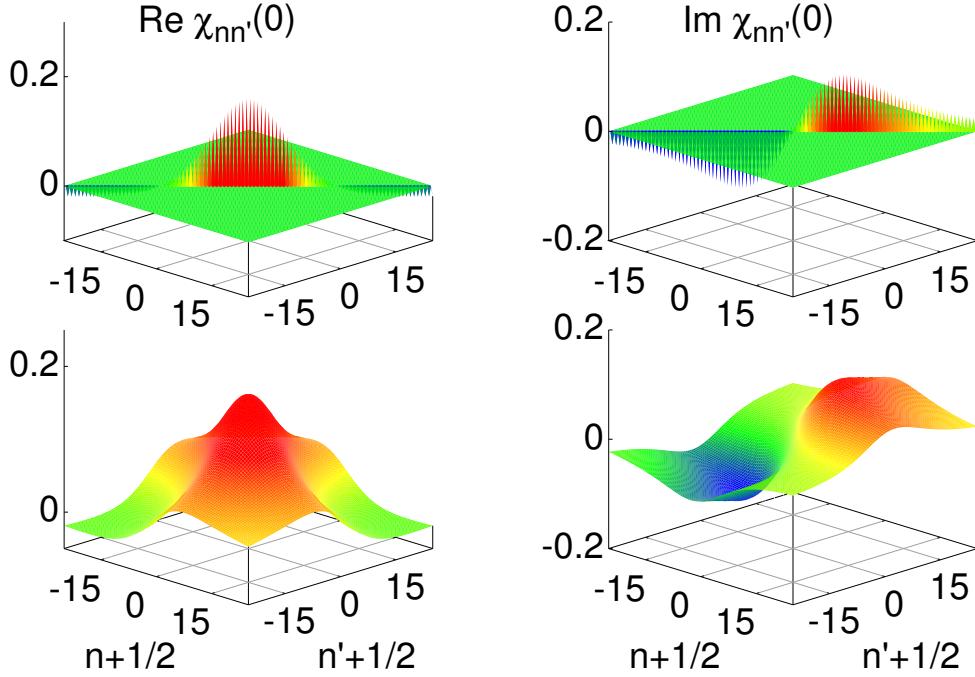


Fig. 7: Top: Real (left) and imaginary (right) part of $\chi_{n,n'}^{\alpha\alpha\alpha\alpha}(0)$ for a non-interacting system, calculated for an energy level ε well above the Fermi level and $\beta = 30 \text{ eV}^{-1}$. Bottom: Real (left) and imaginary (right) part of the mean-field contribution. To better show the symmetries the center is shifted by $(1/2, 1/2)$. In all the plots the colors range from blue (minimum negative value) through green and yellow to red (maximum positive value).

3 The dynamical susceptibility

3.1 The magnetic susceptibility

For a system made of well defined localized spins, as we have seen, the magnetic linear response function is proportional to the site susceptibility

$$\chi_{zz}^{i,i'}(\tau) = \chi_{zz}^{i,i'}(\tau) = \langle \mathcal{T} \hat{M}_z^i(\tau) \hat{M}_z^{i'}(0) \rangle_0 - \langle \hat{M}_z^i \rangle_0 \langle \hat{M}_z^{i'} \rangle_0,$$

where $\hat{M}_z^i = -g\mu_B \hat{S}_z^i$ is the magnetization for lattice site i . Its Fourier transform is

$$\begin{aligned} \chi_{zz}(\mathbf{q}; i\omega_m) &= \sum_{ii'} e^{i\mathbf{q} \cdot (\mathbf{T}_i - \mathbf{T}_{i'})} \int d\tau e^{i\omega_m \tau} \chi_{zz}^{i,i'}(\tau) \\ &= \langle \hat{M}_z(\mathbf{q}; \omega_m) \hat{M}_z(-\mathbf{q}; 0) \rangle_0 - \langle \hat{M}_z(\mathbf{q}) \rangle_0 \langle \hat{M}_z(-\mathbf{q}) \rangle_0, \end{aligned} \quad (24)$$

where ω_m is a bosonic Matsubara frequency. In actual calculations it is, however, often necessary to work with the full two-particle Green function tensor. Let us consider explicitly the case of the one-band Hubbard model; the tensor elements $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\tau)$ that are relevant for the magnetic susceptibility are those for which $\alpha = \mathbf{k}_1\sigma$, $\alpha' = \mathbf{k}_2\sigma$, $\gamma = \mathbf{k}_3\sigma'$, and $\gamma' = \mathbf{k}_4\sigma'$. Each

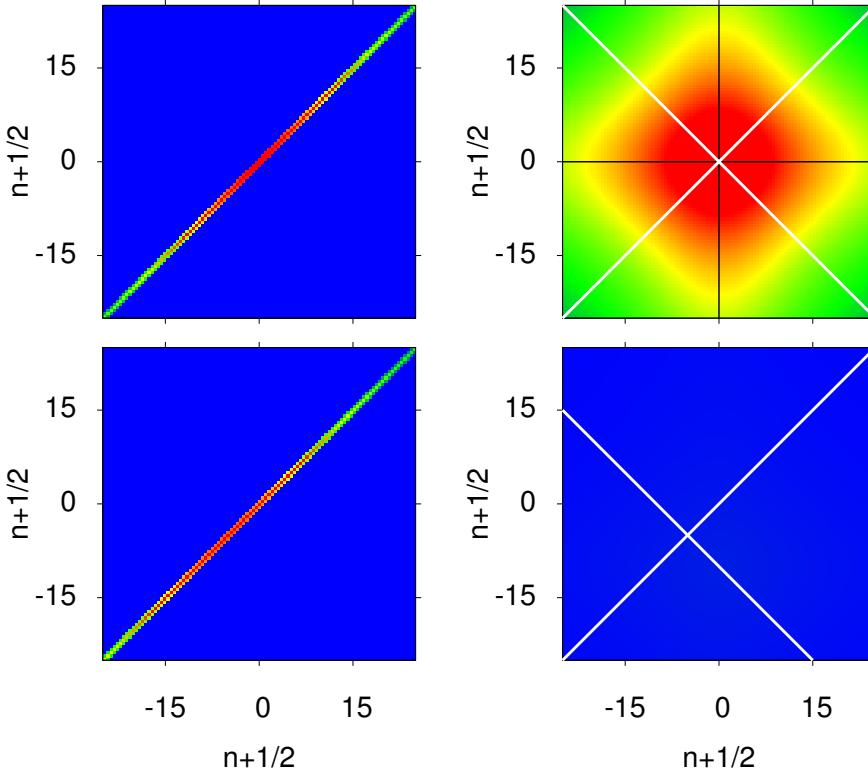


Fig. 8: Left: Absolute value of $\chi_{n,n'}^{\sigma\sigma}(0)$ (top) and $\chi_{n,n'}^{\sigma\sigma}(i\omega_{10})$ (bottom) for a non-interacting system for an energy level ε well above the Fermi level and $\beta = 30 \text{ eV}^{-1}$. Right: Absolute value of the corresponding mean-field contribution. The white lines show symmetry axes. To better show the symmetries the center is shifted by $(1/2, 1/2)$. The colors range from blue (minimum value, here zero) through green and yellow to red (maximum positive value).

flavor has a momentum \mathbf{k}_i associated, but, as we have seen for the frequencies, only three of the four \mathbf{k}_i vectors are independent (momentum conservation). Let us set $\alpha = \mathbf{k}\sigma$, $\alpha' = \mathbf{k} + \mathbf{q}\sigma$, $\gamma = \mathbf{k}' + \mathbf{q}\sigma'$ and $\gamma' = \mathbf{k}'\sigma'$ and write the tensor as a $2N_k \times 2N_k$ matrix whose elements are defined as

$$[\chi(\mathbf{q}; \boldsymbol{\tau})]_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} = \chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) = \langle \mathcal{T} c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q}\sigma}^\dagger(\tau_2) c_{\mathbf{k}'+\mathbf{q}\sigma'}(\tau_3) c_{\mathbf{k}'\sigma'}^\dagger(\tau_4) \rangle_0 - \langle \mathcal{T} c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q}\sigma}^\dagger(\tau_2) \rangle_0 \langle \mathcal{T} c_{\mathbf{k}'+\mathbf{q}\sigma'}(\tau_3) c_{\mathbf{k}'\sigma'}^\dagger(\tau_4) \rangle_0. \quad (25)$$

The magnetic susceptibility is then given by

$$\chi_{zz}(\mathbf{q}; \boldsymbol{\tau}) = (g\mu_B)^2 \frac{1}{4} \sum_{\sigma\sigma'} \sigma\sigma' \chi^{\mathbf{q}\sigma\sigma'}(\boldsymbol{\tau}), \quad (26)$$

where $\sigma = 1$ or -1 for up and down, respectively, and

$$\chi^{\mathbf{q}\sigma\sigma'}(\boldsymbol{\tau}) = \frac{1}{\beta} \frac{1}{N_k^2} \sum_{\mathbf{k}\mathbf{k}'} [\chi(\mathbf{q}; \boldsymbol{\tau})]_{\mathbf{k}\sigma, \mathbf{k}'\sigma'} . \quad (27)$$

After we Fourier transform with respect to imaginary time and sum over the fermionic Matsubara frequencies, we obtain the actual magnetic response function

$$\chi_{zz}(\mathbf{q}; i\omega_m) = (g\mu_B)^2 \frac{1}{4} \sum_{\sigma\sigma'} \sigma\sigma' \frac{1}{\beta^2} \sum_{nn'} \chi_{n,n'}^{q\sigma\sigma'}(i\omega_m), \quad (28)$$

where

$$\chi_{n,n'}^{q\sigma\sigma'}(i\omega_m) = \chi^{q\sigma\sigma'}(\boldsymbol{\nu}) = \frac{\beta}{8} \iiint d\boldsymbol{\tau} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\tau}} \chi^{q\sigma\sigma'}(\boldsymbol{\tau}).$$

3.1.1 One-band Hubbard model: Non interacting limit

In the non-interacting limit we can use Wick's theorem. It follows that the elements of the two-particle Green function tensor vanish if $\mathbf{k} \neq \mathbf{k}'$, and that, in the paramagnetic case, Eq. (26) becomes

$$\chi_{zz}(\mathbf{q}; \boldsymbol{\tau}) = -(g\mu_B)^2 \frac{1}{4} \frac{1}{\beta} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{\sigma} \mathcal{G}_{\mathbf{k}\sigma}(\tau_{14}) \mathcal{G}_{\mathbf{k}+\mathbf{q}\sigma}(-\tau_{23}).$$

For the frequency-dependent magnetic susceptibility Eq. (28) we have instead

$$\chi_{zz}(\mathbf{q}; i\omega_m) = (g\mu_B)^2 \frac{1}{4} \frac{1}{\beta^2} \sum_{nn'} \sum_{\sigma} \chi_{n,n'}^{q\sigma\sigma}(i\omega_m),$$

where

$$\sum_{\sigma} \chi_{n,n'}^{q\sigma\sigma}(i\omega_m) = -\beta \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{\sigma} \mathcal{G}_{\mathbf{k}\sigma}(i\nu_n) \mathcal{G}_{\mathbf{k}+\mathbf{q}\sigma}(i\nu_n + i\omega_m) \delta_{n,n'}. \quad (29)$$

The static susceptibility is given by

$$\chi_{zz}(\mathbf{q}; 0) = - (g\mu_B)^2 \frac{1}{4} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_{\sigma} \frac{n_{\sigma}(\varepsilon_{\mathbf{k}+\mathbf{q}}) - n_{\sigma}(\varepsilon_{\mathbf{k}})}{\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_{\mathbf{k}}}.$$

Finally, in the $\mathbf{q} \rightarrow 0$ and $T \rightarrow 0$ limit we recover as expected the Pauli susceptibility

$$\begin{aligned} \chi_{zz}(\mathbf{0}; 0) &= \frac{1}{4} (g\mu_B)^2 \rho(\varepsilon_F), \\ \rho(\varepsilon_F) &= - \sum_{\sigma} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \left. \frac{dn_{\sigma}(\varepsilon_{\mathbf{k}})}{d\varepsilon_{\mathbf{k}}} \right|_{T=0}. \end{aligned}$$

If we consider the HTSCs dispersion relation Eq. (15) and assume $t' = 0$, at half filling the non-interacting static susceptibility exhibits a divergence at $\mathbf{q} = \Gamma = (0, 0, 0)$; this is due to the van Hove singularity in the density of states at the Fermi level. It also diverges at, e.g., $X = (\pi/a, 0, 0)$ because of perfect nesting, $\varepsilon_{\mathbf{k}+X} = -\varepsilon_{\mathbf{k}}$. More details about the magnetic susceptibility of the non-interacting half-filled one-band Hubbard model with dispersion given by Eq. (15) can be found in Ref. [22].

3.1.2 One-band Hubbard model: Atomic limit

Let us now consider the opposite case, the atomic limit. First we use a simple approach, we directly calculate the right-hand side of Eq. (24) by summing up the contributions of the atomic states, $|0\rangle$, $c_{\uparrow}^{\dagger}|0\rangle$, $c_{\downarrow}^{\dagger}|0\rangle$, $c_{\uparrow}^{\dagger}c_{\downarrow}^{\dagger}|0\rangle$; since the atoms are decoupled, only on-site terms $i = i'$ contribute. The magnetic susceptibility, normalized to a single atom, is given by

$$\chi_{zz}(\mathbf{q}; 0) = (g\mu_B)^2 \frac{1}{4k_B T} \frac{e^{\beta U/2}}{1 + e^{\beta U/2}}. \quad (30)$$

The same expression can be obtained from the two-particle Green function tensor $\chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau})$, defined in Eq. (25) for the Hubbard model. In the atomic limit it is better to work directly in real space; since only $i = i'$ terms contribute, carrying out the \mathbf{k} sums in Eq. (27) we find

$$\chi^{q\sigma\sigma'}(\boldsymbol{\tau}) = \frac{1}{\beta} \sum_i \chi_{i\sigma'i\sigma'}^{i\sigma i\sigma'}(\boldsymbol{\tau}).$$

As we have seen in Sec. 2.7.2, it is sufficient to calculate $\chi_{i\sigma'i\sigma'}^{i\sigma i\sigma'}(\boldsymbol{\tau})$ for positive times $0 < \tau_{j4} < \beta$. Because of the time ordering operator we have, however, to distinguish the various imaginary-time sectors. Let us consider first the case $\tau_{j4} > \tau_{j+14}$ and label the corresponding $\boldsymbol{\tau}$ -vector as $\boldsymbol{\tau}^+$. Calculating the trace we obtain

$$\chi_{i\sigma'i\sigma'}^{i\sigma i\sigma'}(\boldsymbol{\tau}^+) = \frac{1}{2(1 + e^{\beta U/2})} (e^{\tau_{12}U/2 + \tau_{34}U/2} + \delta_{\sigma\sigma'} e^{(\beta - \tau_{12})U/2 - \tau_{34}U/2}).$$

For a paramagnetic system the mean-field terms $G_{\sigma}(\tau_{12})G_{\sigma'}(\tau_{34})$ cancel out in the actual susceptibility; thus we dropped them in the expression above. For a single atom, the imaginary-time magnetic susceptibility in the $\boldsymbol{\tau}^+$ sector is then given by

$$\chi_{zz}(\boldsymbol{\tau}^+) = (g\mu_B)^2 \frac{1}{4\beta} \sum_{\sigma\sigma'} \sigma\sigma' \chi_{i\sigma'i\sigma'}^{i\sigma i\sigma'}(\boldsymbol{\tau}) = \frac{(g\mu_B)^2}{4\beta} \frac{1}{(1 + e^{\beta U/2})} e^{(\beta - \tau_{12} - \tau_{34})U/2}.$$

The terms corresponding to the remaining imaginary-time sectors can be obtained in a similar way (see Appendix); summing up the various contribution to Eq. (28), i.e., to the Fourier transform $\chi_{n,n'}^{\sigma\sigma'}(i\omega_n)$ we recover the initial expression Eq. (30).

In the atomic limit, $\chi_{zz}(\mathbf{q}; 0)$ decreases for large temperatures as $1/k_B T$, i.e., it has a Curie behavior. This is very different from what we find in the $U = 0$ limit. The non-interacting Pauli susceptibility $\chi_{zz}(0; 0)$ is weakly temperature-dependent; for the HTSCs dispersion relation, at half-filling the temperature dependence of $\chi_{zz}(0; 0)$ is enhanced for $t' = 0$, i.e., when the logarithmic van-Hove singularity is at the Fermi level [22].

3.2 The generalized susceptibility

In this section, we generalize what we have seen in the previous one to the case of the multi-band Hubbard model defined in (2); furthermore, we consider the linear response to a non-specified external field, not necessarily a magnetic field. Let us start from the site susceptibility in imaginary time

$$\chi_{\hat{P}_{\nu}^i \hat{O}_{\nu'}^{i'}}(\boldsymbol{\tau}) = \langle \mathcal{T} \Delta \hat{P}_{\nu}^i(\tau_1, \tau_2) \Delta \hat{O}_{\nu'}^{i'}(\tau_3, \tau_4) \rangle_0,$$

where the site operators \hat{P}_ν^i and $\hat{O}_{\nu'}^{i'}$ are defined as

$$\begin{aligned}\hat{P}_\nu^i(\tau_1, \tau_2) &= \sum_{\alpha} p_\alpha^\nu c_{i\alpha'}^\dagger(\tau_2) c_{i\alpha}(\tau_1), \\ \hat{O}_{\nu'}^{i'}(\tau_3, \tau_4) &= \sum_{\gamma} o_\gamma^{\nu'} c_{i'\gamma'}^\dagger(\tau_4) c_{i'\gamma}(\tau_3).\end{aligned}$$

The labels $\boldsymbol{\alpha} = (\alpha, \alpha')$, $\boldsymbol{\gamma} = (\gamma, \gamma')$ are, as usual, collective flavors, and for the multi-band Hubbard model include spin (σ) and orbital (m) quantum number, plus a fractional vector identifying a correlated basis atom in the unit cell (i_c). The weight factors o_α^ν and $p_\gamma^{\nu'}$, in general complex numbers, identify the type of response we calculate. The imaginary-time site susceptibility is then given by

$$\chi_{\hat{P}_\nu^i \hat{O}_{\nu'}^{i'}}(\boldsymbol{\tau}) = \sum_{\alpha\gamma} v_{\alpha\gamma} \chi_{\gamma_{i'}}^{\alpha_i}(\boldsymbol{\tau}),$$

where $v_{\alpha\gamma} = p_\alpha^\nu o_\gamma^{\nu'}$, $\alpha_i = (i\alpha, i\alpha')$, and $\gamma_{i'} = (i'\gamma, i'\gamma')$. The function $\chi_{\gamma_{i'}}^{\alpha_i}(\boldsymbol{\tau})$ is defined in Eq. (21) and its Fourier transform in time, $\chi_{\gamma_{i'}}^{\alpha_i}(\boldsymbol{\nu})$, in Eq. (22). If we perform the Fourier transform in both time and lattice vectors we find

$$\begin{aligned}\chi(\mathbf{q}; \boldsymbol{\nu}) &= \sum_{\alpha\gamma} v_{\alpha\gamma} \sum_{ii'} e^{i(\mathbf{T}_i - \mathbf{T}_{i'}) \cdot \mathbf{q}} \chi_{\gamma_{i'}}^{\alpha_i}(\boldsymbol{\nu}) = \sum_{\alpha\gamma} v_{\alpha\gamma} \frac{1}{N_k^2} \sum_{kk'} \chi_{\gamma_{k'}}^{\alpha_k}(\boldsymbol{\nu}) \\ &= \sum_{\alpha\gamma} v_{\alpha\gamma} [\chi(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma},\end{aligned}$$

where $\boldsymbol{\alpha}_k = (\alpha\mathbf{k}_1, \alpha'\mathbf{k}_2)$ and $\boldsymbol{\gamma}_{k'} = (\gamma\mathbf{k}_3, \gamma'\mathbf{k}_4)$; as in the case of the magnetic susceptibility for the one-band Hubbard model, $\mathbf{k}_1 = \mathbf{k}$, $\mathbf{k}_2 = -\mathbf{k} - \mathbf{q}$, $\mathbf{k}_3 = \mathbf{k}' + \mathbf{q}$ and $\mathbf{k}_4 = -\mathbf{k}'$. The terms $\chi_{\gamma_{k'}}^{\alpha_k}(\boldsymbol{\nu})$ build a square matrix

$$\chi_{\gamma_{k'}}^{\alpha_k}(\boldsymbol{\nu}) \equiv [\chi(\mathbf{q}; i\omega_m)]_{\mathbf{k}L_\alpha, \mathbf{k}'L_\gamma}$$

whose elements are labeled for convenience by the collective indices $\mathbf{k}L_\alpha = (\alpha\mathbf{k}n, \alpha'\mathbf{k}n)$ and $\mathbf{k}'L_\gamma = (\gamma\mathbf{k}'n', \gamma'\mathbf{k}'n')$; by summing over \mathbf{k} and \mathbf{k}' we obtain $[\chi(\mathbf{q}; \omega_m)]_{L_\alpha, L_\gamma}$. Finally, the actual linear response function is given by the sum over the fermionic Matsubara frequencies

$$\chi_{\hat{P}_\nu \hat{O}_{\nu'}}(\mathbf{q}; i\omega_m) = \sum_{\alpha\gamma} v_{\alpha\gamma} \frac{1}{\beta^2} \sum_{nn'} [\chi(\mathbf{q}; \omega_m)]_{L_\alpha, L_\gamma}. \quad (31)$$

In the single-orbital case ($\alpha = \alpha' = \sigma$ and $\gamma = \gamma' = \sigma'$) when $\hat{P}_\nu^i = \hat{M}_z^i = \hat{O}_\nu^i$ we have

$$o_\alpha^z = -g\mu_B \langle \sigma | \hat{\sigma}_z | \sigma \rangle, \quad p_\alpha^z = -g\mu_B \langle \sigma' | \hat{\sigma}_z | \sigma' \rangle,$$

and we recover the magnetic susceptibility for the one-band Hubbard model.

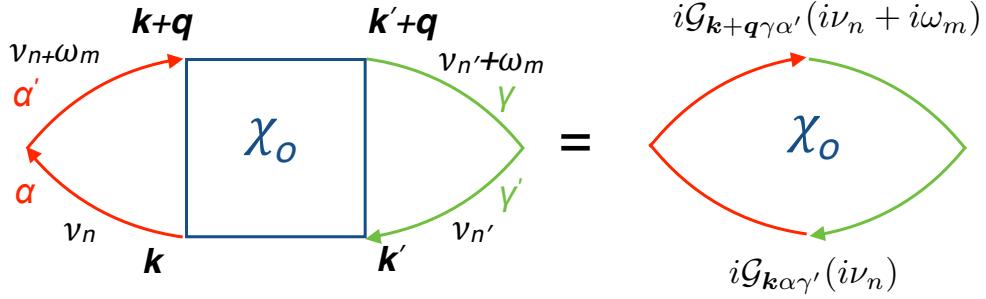


Fig. 9: Diagram contributing to the linear susceptibility for a non-interacting system. The red lines indicates that the creator/annihilator is originally from the operator \hat{P}_ν and the green lines indicate that the creator/annihilator is from the operator \hat{O}_ν . The corresponding frequencies and momenta are explicitly assigned.

3.3 The generalized susceptibility in DMFT

The linear response function is nothing more than a generalized many-particle retarded Green function. Thus we can use, in principle, all standard many-body techniques for deriving a perturbation series for it. Let us consider a system described by the multi-band Hubbard model (2), which we write here as $\hat{H}_e = \hat{H}_0 + \hat{H}_U$, where \hat{H}_0 is the non-interacting part. Let us now formally construct a perturbation series for $\chi_{\hat{P}_\nu \hat{O}_\nu}(\mathbf{q}; \omega_m)$ in the interaction \hat{H}_U . The first step is to calculate the zero-order contribution, i.e., the linear response function for the non-interacting term \hat{H}_0 . Since for \hat{H}_0 the Wick's theorem holds, we have

$$[\chi_0(\mathbf{q}; i\omega_m)]_{\mathbf{k}L_\alpha, \mathbf{k}'L_\gamma} = -\beta N_{\mathbf{k}} \mathcal{G}_{\mathbf{k}\alpha\gamma'}(i\nu_n) \mathcal{G}_{\mathbf{k}'+\mathbf{q}\alpha'\gamma}(i\nu_{n'} + i\omega_m) \delta_{n,n'} \delta_{\mathbf{k},\mathbf{k}'} \quad (32)$$

The Feynman diagram corresponding to $[\chi_0(\mathbf{q}; \omega_m)]_{L_\alpha, L_\gamma}$ is shown in Fig. 9. Once we switch on the interaction, many-body perturbation theory leads to the Bethe-Salpeter (BS) equation, pictorially shown in Fig. 10. Mathematically, it can be written in a matrix form as follows

$$[\chi(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma} = \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{k}\mathbf{k}'} \left[\chi_0(\mathbf{q}; i\omega_m) + \frac{1}{N_{\mathbf{k}}^2} \chi_0(\mathbf{q}; i\omega_m) \Gamma(\mathbf{q}; i\omega_m) \chi(\mathbf{q}; i\omega_m) \right]_{\mathbf{k}L_\alpha, \mathbf{k}'L_\gamma},$$

where the external sums on \mathbf{k} vectors are explicitly written. For systems for which dynamical mean-field is a good approximation, however, it is more convenient to construct a diagrammatic series starting from the DMFT linear response function rather than from the non-interacting term. If we do so, $\chi_0(\mathbf{q}; \omega_m)$ in the Bethe-Salpeter equation is given by Eq. (32) with \mathcal{G} replaced by the DMFT Green function matrices

$$[\chi_0(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma} = -\beta \delta_{nn'} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} G_{\alpha\gamma'}^{\text{DMFT}}(\mathbf{k}; i\nu_n) G_{\alpha'\gamma}^{\text{DMFT}}(\mathbf{k} + \mathbf{q}; i\nu_n + i\omega_m).$$

There is a catch, however. How do we calculate the vertex matrix $[\Gamma(\mathbf{q}; \omega_m)]_{\mathbf{k}L_\alpha, \mathbf{k}'L_\gamma}$? In the infinite dimension limit it has been shown that, in the BS equation, the vertex can be replaced

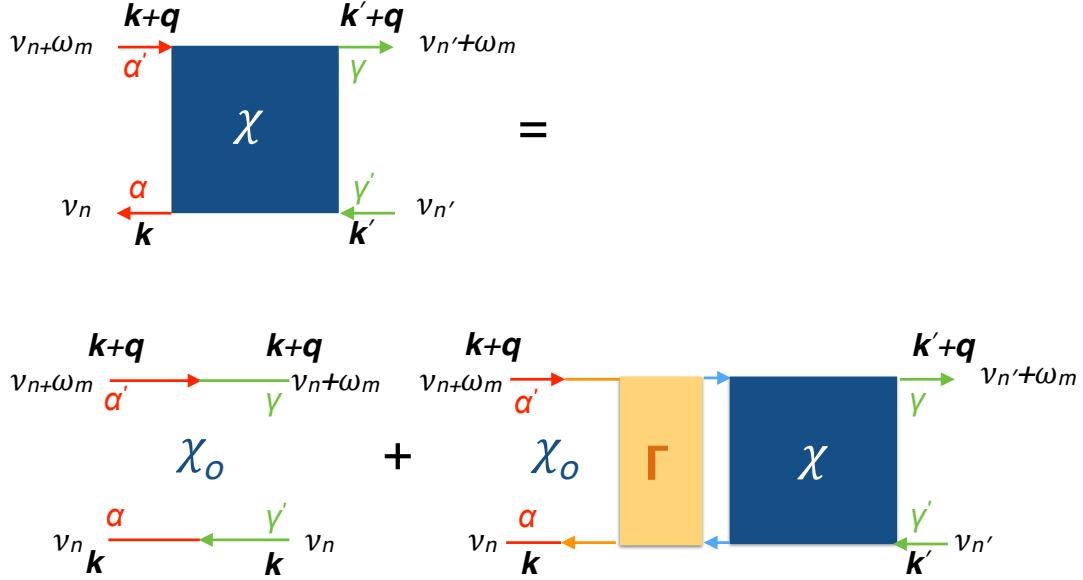


Fig. 10: Diagrammatic representation of the Bethe-Salpeter equation for the linear susceptibility. The red lines indicates a creator/annihilator stemming from the operator \hat{P}_ν and the green lines from the operator $\hat{O}_{\nu'}$. The box labeled with Γ is the vertex function, the one labeled with χ the full susceptibility, and χ_0 is the pair-bubble term.

by a local quantity [15, 23, 24]; assuming that, in the spirit of the dynamical mean-field approximation, for a real 3-dimensional system we can still do the same, the BS equation becomes

$$[\chi(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma} = [\chi_0(\mathbf{q}; \omega_m) + \chi_0(\mathbf{q}; i\omega_m)\Gamma(i\omega_m)\chi(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma}.$$

By solving it formally we find

$$[\chi^{-1}(\mathbf{q}; i\omega_m)]_{L_\alpha, L_\gamma} = [\chi_0^{-1}(\mathbf{q}; i\omega_m) - \Gamma(i\omega_m)]_{L_\alpha, L_\gamma}. \quad (33)$$

To actually obtain $\chi(\mathbf{q}; i\omega_m)$ from this equation we still need the local vertex. The latter can be calculated by means of a further approximation, i.e., assuming that (33) is also satisfied if we replace the \mathbf{q} -dependent susceptibilities with their local counterparts, defined as

$$[\chi_0(i\omega_m)]_{L_\alpha^{i_c}, L_\gamma^{i_c}} = \frac{1}{N_q} \sum_q [\chi_0(\mathbf{q}; i\omega_m)]_{L_\alpha^{i_c}, L_\gamma^{i_c}},$$

$$[\chi(i\omega_m)]_{L_\alpha^{i_c}, L_\gamma^{i_c}} = \frac{1}{N_q} \sum_q [\chi(\mathbf{q}; i\omega_m)]_{L_\alpha^{i_c}, L_\gamma^{i_c}}.$$

Since the local response function is the same for all equivalent correlated basis sites i_c we work with the matrix block of a given site; to make this explicit we renamed the corresponding elements L_α as $L_\alpha^{i_c}$. The local term $\chi(i\omega_m)$ is obtained via the quantum impurity solver in the final iteration of the DMFT self-consistency loop. By inverting the local BS equation we have the vertex

$$[\Gamma(i\omega_m)]_{L_\alpha, L_\gamma} = [\chi_0^{-1}(i\omega_m)]_{L_\alpha, L_\gamma} - [\chi^{-1}(i\omega_m)]_{L_\alpha, L_\gamma}. \quad (34)$$

Replacing $\Gamma(i\omega_m)$ obtained via Eq. (34) into Eq. (33) yields the q -dependent susceptibility. It has to be noticed that, although the two equations (33) and (34) look innocent, solving them numerically is a delicate task because the local susceptibility is in general not diagonal in n, n' and does not decay very fast with the frequencies. There are, however, various ways around based on extrapolations [25] or using auxiliary polynomials [26] or other methods.

3.4 The $\chi_0(\mathbf{q}; \omega)$ diagram

It is tempting to stop at the first term in the expansion, $\chi_0(\mathbf{q}; \omega)$. In the non-interacting case, $\chi_0(\mathbf{q}; \omega)$ is the exact solution by construction; for small U we can expect that $\chi_0(\mathbf{q}; \omega)$ is a reasonable approximation. Can we use it as an approximated linear response function more in general, i.e., also for intermediate or even large U ? Unfortunately the answer is no. In the large U limit $\chi_0(\mathbf{q}; \omega)$ is very different from the exact susceptibility. To understand this point let us calculate the static large U magnetic susceptibility for the half-filled one-band Hubbard model. We consider two cases, the atomic limit and the insulating regime (small t/U limit); for the latter we use an approximate expression for the self-energy.

3.4.1 One-band Hubbard model: Atomic limit

In the atomic limit, using the atomic Green function instead of \mathcal{G} in Eq. (32), we obtain

$$\chi_{n,n'}^{\sigma\sigma'}(0) = -\beta\delta_{nn'}\delta_{\sigma\sigma'}\frac{1}{4} \left[\frac{1}{i\nu_n + U/2} + \frac{1}{i\nu_n - U/2} \right] \left[\frac{1}{i\nu_{n'} + U/2} + \frac{1}{i\nu_{n'} - U/2} \right].$$

By performing the Matsubara sums

$$\chi_{zz}^0(0) = \frac{1}{4}(g\mu_B)^2 \sum_{\sigma} \frac{1}{\beta^2} \sum_n \chi_{n,n}^{\sigma\sigma}(0) = \frac{1}{4}(g\mu_B)^2 \frac{\beta e^{\beta U/2}}{1 + e^{\beta U/2}} \left[\frac{1}{1 + e^{\beta U/2}} + \frac{1}{U\beta} \left(\frac{1 - e^{-\beta U}}{1 + e^{-\beta U/2}} \right) \right]$$

If we assume that U is finite, for very large temperatures ($\beta U \rightarrow 0$) we find $\chi_{zz}^0(0) \sim \beta(g\mu_B)^2/8$. In realistic cases, however, the temperature is in the range 0-1000 K and U is of the order of few eV; under these conditions the atomic spin $S = 1/2$ is well defined. In such a large βU limit $\chi_{zz}^0(0) \sim (g\mu_B)^2/4U$, i.e., the term $\chi_{zz}^0(0)$ does not exhibit the Curie behavior. The approximated susceptibility $\chi_{zz}^0(0)$ should be compared with the actual magnetic susceptibility of our idealized atom which, as we have seen in Sec. 3.1, is given by

$$\chi_{zz}(0) = (g\mu_B)^2 \frac{1}{4} \left[\frac{\beta e^{\beta U/2}}{1 + e^{\beta U/2}} \right] \underset{\beta U \rightarrow \infty}{\sim} \frac{(g\mu_B S)^2}{k_B T}.$$

3.4.2 One-band Hubbard model: Mott-insulating regime

Let us now consider the half-filled Hubbard model with HTCSs dispersion; the latter is defined in Eq. (15). For simplicity, we assume that all hopping integrals except t are zero. In the atomic limit ($t = 0$) we can rewrite the atomic Green function as

$$G(i\nu_n) = \frac{1}{i\nu_n + \mu - \Sigma(i\nu_n)},$$

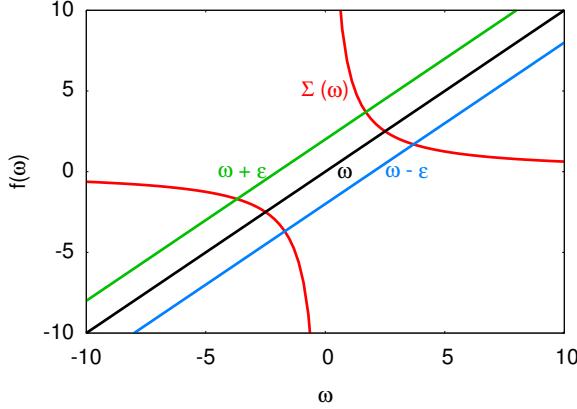


Fig. 11: Graphical solution of the equation $\omega + \varepsilon + \Sigma(\omega) = 0$ yielding the poles E^+ and E^- of the Green function defined in Eq. (36).

where $\mu = U/2$ and the self-energy is given by

$$\Sigma(i\nu_n) = \mu + \frac{U^2}{4} \frac{1}{i\nu_n}. \quad (35)$$

In the Mott insulating regime, i.e., for small but finite t/U , we can assume that the self-energy can be still written in the form given in Eq. (35), with μ replaced by the actual chemical potential and $U^2/4$ by a quantity which plays the role of a dimensionless *order parameter* [24] for the insulating phase $r_U U^2/4$, defined as

$$\frac{1}{r_U} \frac{4}{U^2} = \int_{-\infty}^{+\infty} d\varepsilon \frac{\rho(\varepsilon)}{\varepsilon^2},$$

where $\rho(\varepsilon)$ is the density of states per spin; the integral $4/r_U U^2$ diverges in the metallic phase. The Green function can then be rewritten as

$$G_{\mathbf{k}}(i\nu_n) = \frac{1}{i\nu_n - \Sigma(i\nu_n) - \varepsilon_{\mathbf{k}}} = \frac{1}{E_{\mathbf{k}}^+ - E_{\mathbf{k}}^-} \left[\frac{E_{\mathbf{k}}^+}{i\nu_n - E_{\mathbf{k}}^+} - \frac{E_{\mathbf{k}}^-}{i\nu_n - E_{\mathbf{k}}^-} \right] \quad (36)$$

where $E_{\mathbf{k}}^+$ and $E_{\mathbf{k}}^-$ are the two roots of the equation $\omega - \Sigma(\omega) - \varepsilon_{\mathbf{k}} = 0$,

$$E_{\mathbf{k}}^{\pm} = \frac{1}{2} \varepsilon_{\mathbf{k}} \pm \frac{1}{2} \sqrt{\varepsilon_{\mathbf{k}}^2 + r_U U^2}.$$

By performing the Matsubara sums, one finds

$$\begin{aligned} \chi_{zz}^0(\mathbf{q}; 0) &= (g\mu_B)^2 \frac{1}{4} \sum_{\sigma} \frac{1}{\beta^2} \sum_n \chi_{n,n}^{\sigma\sigma}(0) \\ &= (g\mu_B)^2 \frac{1}{2} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \left[\underbrace{-I_{\mathbf{k},\mathbf{q}}^{++} - I_{\mathbf{k},\mathbf{q}}^{--}}_{A_{\mathbf{k},\mathbf{q}}} + \underbrace{I_{\mathbf{k},\mathbf{q}}^{+-} + I_{\mathbf{k},\mathbf{q}}^{-+}}_{B_{\mathbf{k},\mathbf{q}}} \right], \end{aligned}$$

where, setting $\alpha = \pm$ and $\gamma = \pm$,

$$I_{\mathbf{k},\mathbf{q}}^{\alpha\gamma} = \frac{E_{\mathbf{k}}^{\alpha} E_{\mathbf{k}+\mathbf{q}}^{\gamma}}{(E_{\mathbf{k}}^+ - E_{\mathbf{k}}^-)(E_{\mathbf{k}+\mathbf{q}}^+ - E_{\mathbf{k}+\mathbf{q}}^-)} \frac{n(E_{\mathbf{k}}^{\alpha}) - n(E_{\mathbf{k}+\mathbf{q}}^{\gamma})}{E_{\mathbf{k}}^{\alpha} - E_{\mathbf{k}+\mathbf{q}}^{\gamma}}.$$

In the $\mathbf{q} \rightarrow \mathbf{0}$ limit we find

$$\begin{aligned} A_{\mathbf{k},0} &= \beta \left[\frac{(E_{\mathbf{k}}^+)^2}{\varepsilon_{\mathbf{k}}^2 + r_U U^2} n(E_{\mathbf{k}}^+) [1 - n(E_{\mathbf{k}}^+)] + \frac{(E_{\mathbf{k}}^-)^2}{\varepsilon_{\mathbf{k}}^2 + r_U U^2} n(E_{\mathbf{k}}^-) [1 - n(E_{\mathbf{k}}^-)] \right] \\ B_{\mathbf{k},0} &= \frac{r_U U^2}{2(\varepsilon_{\mathbf{k}}^2 + r_U U^2)^{3/2}} [n(E_{\mathbf{k}}^-) - n(E_{\mathbf{k}}^+)]. \end{aligned}$$

In the large βU limit, the $A_{\mathbf{k},0}$ term, proportional to the density of states at the Fermi level, vanishes exponentially; the $B_{\mathbf{k},0}$ term yields the dominant contribution

$$\chi_{zz}^0(\mathbf{0}; 0) \sim (g\mu_B)^2 \frac{1}{4} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \frac{r_U U^2}{[\varepsilon_{\mathbf{k}}^2 + r_U U^2]^{3/2}} \sim (g\mu_B)^2 \frac{1}{4\sqrt{r_U} U} \left[1 - \frac{3}{2} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \frac{\varepsilon_{\mathbf{k}}^2}{r_U U^2} + \dots \right].$$

The right-hand side is equal to the atomic $\chi_{zz}^0(0)$ minus a correction; hence the term $\chi_{zz}^0(\mathbf{0}; 0)$ for a Mott insulator has basically the same defects as $\chi_{zz}^0(0)$ for an idealized one-level atom. What about the \mathbf{q} -dependence of $\chi_{zz}^0(\mathbf{q}; 0)$? The one-band Hubbard model has an antiferromagnetic instability in the Mott-insulating regime due to superexchange [22, 27]. Let us therefore calculate $\chi_{zz}^0(\mathbf{q})$ at the \mathbf{q} vector associated with antiferromagnetic order, $\mathbf{q}_C = (\pi/a, \pi/a, 0)$. Since at \mathbf{q}_C the band dispersion satisfies the perfect nesting condition ($\varepsilon_{\mathbf{k}+\mathbf{q}_C} = -\varepsilon_{\mathbf{k}}$) we find

$$\begin{aligned} A_{\mathbf{k},\mathbf{q}_C} &= \frac{1}{2} \frac{r_U U^2}{\varepsilon_{\mathbf{k}}^2 + r_U U^2} \frac{n(E_{\mathbf{k}}^+ - \varepsilon_{\mathbf{k}}) - n(E_{\mathbf{k}}^+)}{\varepsilon_{\mathbf{k}}} \\ B_{\mathbf{k},\mathbf{q}_C} &= \frac{1}{2} \frac{\varepsilon_{\mathbf{k}}^2}{\varepsilon_{\mathbf{k}}^2 + r_U U^2} \frac{n(E_{\mathbf{k}}^+ - \varepsilon_{\mathbf{k}}) - n(E_{\mathbf{k}}^+)}{\varepsilon_{\mathbf{k}}} - \frac{1}{2} \frac{1}{\sqrt{\varepsilon_{\mathbf{k}}^2 + r_U U^2}} [n(E_{\mathbf{k}}^+) - n(E_{\mathbf{k}}^-)], \end{aligned}$$

and therefore

$$\chi_0(\mathbf{q}_C; 0) \sim (g\mu_B)^2 \frac{1}{4\sqrt{r_U} U} \left[1 - \frac{1}{2} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \frac{\varepsilon_{\mathbf{k}}^2}{r_U U^2} \right].$$

Thus $\chi_0(\mathbf{q}; 0)$ is larger at $\mathbf{q} = \mathbf{q}_C$ than at $\mathbf{q} = \mathbf{0}$; it does not exhibit, however, Curie-Weiss instabilities. The latter appear as soon as we take the vertex into consideration. In line with the results above, we express $\chi_0(\mathbf{q}; 0)$ as

$$\chi_0(\mathbf{q}; 0) \sim (g\mu_B)^2 \frac{1}{4\sqrt{r_U} U} \left[1 - \frac{1}{2} \frac{J_{\mathbf{0}}}{\sqrt{r_U} U} - \frac{1}{4} \frac{J_{\mathbf{q}}}{\sqrt{r_U} U} \right],$$

where $J_{\mathbf{q}} = 2J[\cos q_x + \cos q_y]$ with $J \propto t^2/U$. Next we take the exact atomic susceptibility in the large βU limit as local term and calculate the vertex as

$$\Gamma \sim \left[\frac{1}{\chi_{zz}^0(0)} - \frac{1}{\chi_{zz}(0)} \right] \sim \frac{1}{(g\mu_B)^2} \left[4\sqrt{r_U} U \left(1 + \frac{1}{2} \frac{J_{\mathbf{0}}}{\sqrt{r_U} U} \right) - 4k_B T \right].$$

Therefore,

$$\chi_{zz}(\mathbf{q}; 0) = \frac{1}{[\chi_{zz}^0(\mathbf{q}; 0)]^{-1} - \Gamma} \sim (g\mu_B)^2 \frac{1}{4} \frac{1}{k_B T + J_{\mathbf{q}}/4} = \frac{(g\mu_B)^2}{k_B} \frac{1}{4} \frac{1}{T - T_q}.$$

Thus, including the vertex correction we recover the Curie-Weiss behavior expected for a system with coupled localized spins; we also correctly find the antiferromagnetic instability, since \mathbf{q}_C is the vector for which T_q is largest. In conclusion, we have seen that Γ is essential to properly describe the magnetic response function of strongly-correlated systems.

3.5 The local susceptibility $\chi(\omega)$

The local susceptibility can be obtained in various ways. Here we briefly recollect the essential steps for calculating it via a quantum Monte Carlo approach. A more detailed description of quantum Monte Carlo approaches can be found in the lecture of Fakher Assaad.

3.5.1 Hirsch-Fye QMC quantum impurity solver

For the one-band Hubbard model, the typical quantum-impurity problem used in DMFT calculations is the Anderson Hamiltonian

$$\hat{H} = \underbrace{\sum_{\sigma} \varepsilon_d n_{d\sigma}}_{\hat{H}_{\text{loc}}} + U n_{d\uparrow} n_{d\downarrow} + \underbrace{\sum_{\sigma} \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}\sigma}}_{\hat{H}_{\text{bth}}} + \underbrace{\sum_{\sigma} \sum_{\mathbf{k}} [V_{\mathbf{k}} c_{\mathbf{k}\sigma}^{\dagger} c_{d\sigma} + \text{h.c.}]}_{\hat{H}_{\text{hyb}}}.$$

This Hamiltonian describes a correlated site, e.g., site $i = 0$ in the original lattice Hubbard model, whose states are labeled with d , coupled to a non-correlated bath, whose states have for quantum numbers the momentum \mathbf{k} and the spin σ , via the hybridization $V_{\mathbf{k}}$. The Anderson model was originally introduced in the context of the single-impurity Kondo problem [28].

The Hirsch-Fye quantum Monte Carlo approach [29] is based on *imaginary-time discretization*. To calculate a physical observable a crucial ingredient is the partition function

$$Z = \text{Tr } e^{-\beta(\hat{H}-\mu\hat{N})}.$$

If all eigenvalues and eigenvectors were known, the partition function would be of course also known. How do we calculate a physical observable without this information? If we split the interval $\tau = [0, \beta]$ in time steps $\Delta\tau = \beta/L$ we can rewrite the partition function as follows

$$Z = \text{Tr} \prod_{l=1}^L e^{-\Delta\tau(\hat{H}-\mu\hat{N})}. \quad (37)$$

In the Anderson model, the interaction term is \hat{H}_{loc} ; we therefore rewrite the Hamiltonian \hat{H} as $\hat{H} = \hat{H}_0 + \hat{H}_U$, where $\hat{H}_0 = \hat{H}_{\text{bth}} + \hat{H}_{\text{hyb}}$ and $\hat{H}_U = \hat{H}_{\text{loc}}$. If $\Delta\tau$ is small, we can approximate the partition function via the Trotter decomposition

$$Z = \text{Tr} \prod_{l=1}^L e^{-\Delta\tau(\hat{H}_0 - (\mu\hat{N} - \mu_d\hat{N}_d))} e^{-\Delta\tau(\hat{H}_U - \mu_d\hat{N}_d)} + O(\Delta\tau^2), \quad (38)$$

where $\hat{N}_d = n_{d\uparrow} + n_{d\downarrow}$ is the impurity electron number operator, and $\mu_d\hat{N}_d$ yields a shift. The equivalence between Eq. (37) and Eq. (38) up to first order can be verified, e.g., by performing a first-order Taylor expansion of both expressions and comparing the results. Remarkably, Eq. (38) can be rewritten in a simpler form using the Hubbard-Stratonovich transformation. Taking $\mu_d = \varepsilon_d + U/2$ we obtain, for a given $\Delta\tau$

$$e^{-\Delta\tau[\hat{H}_{\text{loc}} - \mu_d\hat{N}_d]} = e^{-\Delta\tau U[n_{d\uparrow}n_{d\downarrow} - \frac{1}{2}(n_{d\uparrow} + n_{d\downarrow})]} = \frac{1}{2} \sum_{s=\pm 1} e^{s\lambda(n_{d\uparrow} - n_{d\downarrow})}$$

where s is an auxiliary Ising variable and can take two values, $s = -1$ or $s = +1$, and $\cosh \lambda = e^{\Delta\tau U/2}$. One can verify this relation by applying the left and right operator to the basis states. The result is summarized in the table below

| | $e^{-\Delta\tau U[n_{d\uparrow}n_{d\downarrow}-\frac{1}{2}(n_{d\uparrow}+n_{d\downarrow})]}$ | $\frac{1}{2}\sum_{s=\pm 1} e^{s\lambda(n_{d\uparrow}-n_{d\downarrow})}$ |
|--|--|---|
| $ 0\rangle$ | 1 | 1 |
| $c_\sigma^\dagger 0\rangle$ | $e^{\Delta\tau U/2}$ | $\cosh \lambda$ |
| $c_{d\uparrow}^\dagger c_{d\downarrow}^\dagger 0\rangle$ | 1 | 1 |

Thus, via the auxiliary-field decoupling we replaced the two-particle term $U n_{d\uparrow} n_{d\downarrow}$ with two single-particle terms, much easier to handle. If we introduce one Ising spin per $\Delta\tau$ interval, i.e., L in total, $\{s_i\} = s_1, \dots, s_L$, the partition function becomes

$$Z = \frac{1}{2^L} \sum_{\{s_i\}} \text{Tr} \prod_{l=1}^L e^{-\Delta\tau(\hat{H}_0 - \mu \hat{N} + \mu_d \hat{N}_d) + \lambda s_l (n_{d\uparrow} - n_{d\downarrow})} + O(\Delta\tau^2).$$

For a specific configuration of Ising spins, i.e., for a given set of values for the variables $\{s_i\}$, the contribution to the partition function of the auxiliary model has the form

$$Z_{\{s_i\}} = \text{Det}[O_{\{s_i\}}^\uparrow] \text{Det}[O_{\{s_i\}}^\downarrow],$$

Taking the inverse of $O_{\{s_i\}}^\sigma$ yields the Green function matrix $G_{\{s_i\}}^\sigma$. The latter has dimension $NL \times NL$ where N is the total number of sites (bath plus impurity) and satisfies the Dyson equation

$$G_{\{s_j\}}^\sigma = [A^\sigma]^{-1} G_{\{s_i\}}^\sigma, \quad A^\sigma = 1 + [1 - G_{\{s_i\}}^\sigma] \left[e^{V_{\{s_i\}} - V_{\{s_j\}}} - 1 \right],$$

where $\{s_j\}$ and $\{s_i\}$ are two different configurations and the matrix $V_{\{s_i\}} = \lambda \sigma s_i(l) |d\rangle \langle d|$ is a potential acting only on the impurity site. Since the potential is local, the impurity Green function G_d satisfies the Dyson equation

$$[G_d^\sigma]_{\{s_j\}} = [A_d^\sigma]^{-1} [G_d^\sigma]_{\{s_i\}}, \quad A_d^\sigma = 1 + [1 - [G_d^\sigma]_{\{s_i\}}] \left[e^{V_{\{s_i\}} - V_{\{s_j\}}} - 1 \right].$$

By summing over all possible configurations, we have

$$Z = \sum_{\{s_i\}} Z_{\{s_i\}}, \quad G_d^\sigma = \sum_{\{s_i\}} w_{\{s_i\}} [G_d^\sigma]_{\{s_i\}}$$

where

$$w_{\{s_i\}} = \frac{Z_{\{s_i\}}}{\sum_{\{s_i\}} Z_{\{s_i\}}}.$$

Thus, if $w_{\{s_i\}}$ were positive definite, it could be used as Boltzmann weight for importance sampling. Unfortunately the ratio of determinants can be negative for some configurations, giving rise to the minus-sign problem. Still, we can define $|w_{\{s_i\}}|$ as Boltzmann weight and keep track of the sign of the product of determinants.

Remarkably, the ratio of the ‘Boltzmann weights’ of two different configurations

$$\frac{w_{\{s_j\}}}{w_{\{s_i\}}} = \frac{\text{Det}O_{\{s_j\}}^\uparrow \text{Det}O_{\{s_j\}}^\downarrow}{\text{Det}O_{\{s_i\}}^\uparrow \text{Det}O_{\{s_i\}}^\downarrow} = R_\uparrow R_\downarrow.$$

can be obtained via the Dyson equation; in particular, if we flip a spin at time slice l

$$R_\sigma = 1 + \left[1 - [[G_d^\sigma]_{\{s_i\}}]_{l,l} \right] [e^{-2\lambda s_i(l)} - 1].$$

Thus, we do not need to calculate the determinants. If a new configuration is accepted we have, however, to recalculate the Green function. In conclusion, the Green function can be obtained as

$$G_d^\sigma \sim \frac{\sum_{\{c\}} \langle G_d^\sigma \rangle_c \text{sign}(w_c)}{\sum_{\{c\}} \text{sign}(w_c)},$$

where $\{c\}$ are the visited configurations. The linear response function χ_α is obtained similarly

$$\chi_\alpha \sim \frac{\sum_{\{c\}} \langle \chi_\alpha \rangle_c \text{sign}(w_c)}{\sum_{\{c\}} \text{sign}(w_c)}.$$

Since the Wick theorem holds for a given configuration, we have $\langle \chi_\alpha \rangle_c = \langle G_{\alpha,\gamma'} \rangle_c \langle G_{\gamma,\alpha'} \rangle_c$. The Hirsch-Fye algorithm can be generalized to more complex local interactions, such as the density-density Coulomb interaction in the multi-orbital Hubbard model, by introducing additional Ising fields s and correspondingly additional parameters λ .

3.5.2 CT-HYB QMC quantum impurity solver

In continuous-time QMC [30] the partition function is expanded in either the hybridization (CT-HYB) or the interaction (CT-INT). Here we discuss shortly the first algorithm, in which the expansion series is in powers of \hat{H}_{hyb} . We follow the notation of Ref. [31]. Since for this algorithm we do not need to specify the form of the local interaction, let us consider the most general quantum-impurity Hamiltonian $\hat{H} = \hat{H}_{\text{loc}} + \hat{H}_{\text{bth}} + \hat{H}_{\text{hyb}}$, where

$$\begin{aligned} \hat{H}_{\text{loc}} &= \sum_{\alpha\bar{\alpha}} \tilde{\varepsilon}_{\alpha\bar{\alpha}} c_\alpha^\dagger c_{\bar{\alpha}} + \frac{1}{2} \sum_{\alpha\alpha'} \sum_{\bar{\alpha}\bar{\alpha}'} U_{\alpha\alpha'\bar{\alpha}\bar{\alpha}'} c_\alpha^\dagger c_{\alpha'}^\dagger c_{\bar{\alpha}'} c_{\bar{\alpha}}, \\ \hat{H}_{\text{bth}} &= \sum_\gamma \epsilon_\gamma b_\gamma^\dagger b_\gamma, \\ \hat{H}_{\text{hyb}} &= \sum_\gamma \sum_\alpha [V_{\gamma,\alpha} c_\alpha^\dagger b_\gamma + h.c.]. \end{aligned}$$

This is a generalized Anderson model that describes a multi-orbital correlated impurity, e.g., site $i = 0$ in the original multi-band Hubbard model (2), coupled to a bath; it is the typical local impurity model which we have to solve in a realistic DMFT calculation. The combined index $\alpha = m\sigma$ labels spin and orbital degrees of freedom. For the bath we can use, without loss of generality, the basis that diagonalizes \hat{H}_{bth} , with quantum numbers γ . Finally, we define

$\tilde{\varepsilon}_{\alpha\bar{\alpha}} = \varepsilon_{\alpha\bar{\alpha}} - \Delta\varepsilon_{\alpha\bar{\alpha}}^{\text{DC}}$, where $\varepsilon_{\alpha\bar{\alpha}}$ is the crystal-field matrix and $\Delta\varepsilon_{\alpha\bar{\alpha}}^{\text{DC}}$ is the double-counting correction. We work in the interaction picture, and therefore

$$\hat{H}_{\text{hyb}}(\tau) = e^{\tau(\hat{H}_{\text{bth}} + \hat{H}_{\text{loc}})} \hat{H}_{\text{hyb}} e^{-\tau(\hat{H}_{\text{bth}} + \hat{H}_{\text{loc}})}.$$

By expanding the partition function in powers of \hat{H}_{hyb} we obtain the series

$$Z = \text{Tr} \left[e^{-\beta(\hat{H}_{\text{bth}} + \hat{H}_{\text{loc}})} \mathcal{T} e^{-\int_0^\beta d\tau \hat{H}_{\text{hyb}}(\tau)} \right] = \sum_{m=0}^{\infty} (-1)^m \int^{(m)} d\boldsymbol{\tau} \text{Tr} \mathcal{T} \left[e^{-\beta(\hat{H}_{\text{bth}} + \hat{H}_{\text{loc}})} \prod_{i=m}^1 \hat{H}_{\text{hyb}}(\tau_i) \right],$$

where \mathcal{T} is again the time-ordering operator, $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_m)$ with $\tau_{i+1} \geq \tau_i$ and

$$\int^{(m)} d\boldsymbol{\tau} \equiv \int_0^\beta d\tau_1 \dots \int_{\tau_{m-1}}^\beta d\tau_m.$$

In the trace, only terms containing an equal number of creation and annihilation operators in both the bath and impurity sector, i.e., only even expansion orders $m = 2n$, yield a finite contribution. Introducing the bath partition function $Z_{\text{bth}} = \text{Tr} e^{-\beta\hat{H}_{\text{bth}}}$, the partition function can be factorized as

$$\frac{Z}{Z_{\text{bth}}} = \sum_{n=0}^{\infty} \int^{(n)} d\boldsymbol{\tau} \int^{(n)} d\bar{\boldsymbol{\tau}} \sum_{\alpha\bar{\alpha}} z_{\alpha,\bar{\alpha}}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}), \quad (39)$$

with $z_{\alpha,\bar{\alpha}}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = t_{\alpha,\bar{\alpha}}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) d_{\bar{\alpha},\alpha}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}})$. The first factor is the trace over the impurity states

$$t_{\alpha,\bar{\alpha}}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = \text{Tr} \mathcal{T} \left[e^{-\beta(\hat{H}_{\text{loc}} - \mu\hat{N})} \prod_{i=n}^1 c_{\alpha_i}(\tau_i) c_{\bar{\alpha}_i}^\dagger(\bar{\tau}_i) \right],$$

where $c_\alpha^{(\dagger)}(\tau) = e^{\tau(\hat{H}_{\text{loc}} - \mu\hat{N})} c_\alpha^{(\dagger)} e^{-\tau(\hat{H}_{\text{loc}} - \mu\hat{N})}$ and N is the total number of electrons on the impurity. For expansion order $m = 2n$, the vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ gives the flavors α_i associated with the n annihilation operators on the impurity at imaginary times τ_i , while the $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_n)$ are associated with the n creation operators at $\bar{\tau}_i$. The second factor is the trace over the non-interacting bath, which is given by the determinant

$$d_{\bar{\alpha},\alpha}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = \det[F_{\bar{\alpha},\alpha}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}})]$$

of the $n \times n$ square hybridization-function matrix with matrix elements $[F_{\bar{\alpha},\alpha}^{(n)}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}})]_{i',i} = F_{\bar{\alpha}_{i'},\alpha_i}(\bar{\tau}_{i'} - \tau_i)$ given by

$$F_{\bar{\alpha}\alpha}(\tau) = \sum_{\gamma} \frac{V_{\gamma,\bar{\alpha}} \bar{V}_{\gamma,\alpha}}{1 + e^{-\beta\epsilon_{\gamma}}} \times \begin{cases} -e^{-\epsilon_{\gamma}\tau} & \tau > 0 \\ e^{-\epsilon_{\gamma}(\beta+\tau)} & \tau < 0. \end{cases}$$

On the fermionic Matsubara frequencies, ω_n , its Fourier transform

$$F_{\bar{\alpha}\alpha}(\omega_n) = \sum_{\gamma} \frac{V_{\gamma,\bar{\alpha}} \bar{V}_{\gamma,\alpha}}{i\omega_n - \epsilon_{\gamma}}$$

is related to the bath Green-function matrix \mathcal{G} by

$$F_{\bar{\alpha}\alpha}(\omega_n) = i\omega_n \delta_{\bar{\alpha}\alpha} - \tilde{\varepsilon}_{\bar{\alpha}\alpha} - (\mathcal{G})_{\bar{\alpha}\alpha}^{-1}(\omega_n),$$

as can be shown by downfolding the original multiband Hubbard model, Eq. (2), to the impurity block (say, the $i = i_0$ site)

$$(\mathcal{G})^{-1}(\omega_n) = \left(\begin{array}{c|ccc} i\omega_n I_{i_0} - H_{i_0} & V_{1,i_0} & V_{2,i_0} & \dots \\ \hline \bar{V}_{1,i_0} & i\omega_n - \epsilon_1 & 0 & \dots \\ \bar{V}_{2,i_0} & 0 & i\omega_n - \epsilon_2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{array} \right).$$

Here the matrix elements of H_{i_0} and I_{i_0} are given by $(H_{i_0})_{\alpha\bar{\alpha}} = \tilde{\varepsilon}_{\alpha\bar{\alpha}}$ and $(I_{i_0})_{\alpha\bar{\alpha}} = \delta_{\alpha,\bar{\alpha}}$, while $(V_{i_0,i})_{\bar{\alpha}i} = V_{\bar{\alpha},i}$, and $(\bar{V}_{i,i_0})_{i\alpha} = \bar{V}_{i,\alpha}$. The partition function defined in Eq. (39) can be seen as the sum over all configurations $c = \{\alpha_i \tau_i, \bar{\alpha}_i \bar{\tau}_i, n\}$ in imaginary time and flavors. In a compact form, similar to the case of the Hirsch-Fye algorithm, we have

$$Z = \sum_c \langle Z \rangle_c = \sum_c w_c \sim \sum_{\{c\}} \text{sign}(w_{\{c\}}),$$

where in the last term the sum is over a sequence of configurations $\{c\}$ sampled by Monte Carlo using $|w_c|$ as the probability of configuration c .

Finally, a generic observable O can then be obtained as the Monte Carlo average

$$O \sim \frac{\sum_{\{c\}} \langle O \rangle_c \text{sign}(w_c)}{\sum_{\{c\}} \text{sign}(w_c)}$$

where $\langle O \rangle_c$ is the value of the observable for configuration c . The average expansion order increases linearly with the inverse temperature.

The Green function matrix can, e.g., be obtained as the Monte Carlo average with

$$\langle O \rangle_c = \langle G_{\alpha\bar{\alpha}} \rangle_c = \sum_{ij=1}^n \Delta(\tau, \tau_j - \bar{\tau}_i) [M^{(n)}]_{ji} \delta_{\alpha_j \alpha} \delta_{\bar{\alpha}_i \bar{\alpha}}.$$

Here $M^{(n)} = [F^{(n)}]^{-1}$ is the inverse of the hybridization-function matrix, updated at each accepted move, and Δ is given by

$$\Delta(\tau, \tau') = -\frac{1}{\beta} \begin{cases} \delta(\tau - \tau') & \tau' > 0 \\ -\delta(\tau - (\tau' + \beta)) & \tau' < 0, \end{cases}$$

where the δ -function is discretized. Alternatively, we can calculate the Green function matrix from its Legendre coefficients [26], i.e., use the Fourier-Legendre series. The Legendre polynomials $\{P_l(x)\}$ form a complete orthogonal system over the interval $[-1, 1]$ and therefore the Green function may be expanded in terms of them as

$$G_{\alpha\bar{\alpha}}(\tau) = \sum_{l=0}^{\infty} \frac{\sqrt{2l+1}}{\beta} P_l(x(\tau)) G_{\alpha\bar{\alpha}}^l, \quad (40)$$

with $x(\tau) = 2\tau/\beta - 1$. The first four Legendre polynomials are

$$\begin{array}{c|cccc} & l=0 & l=1 & l=2 & l=3 \\ \hline P_l(x) & 1 & x & \frac{1}{2}(3x^2 - 1) & \frac{1}{2}(5x^3 - 3x) \end{array}$$

The coefficients of the expansion are given by

$$G_{\alpha\bar{\alpha}}^l = \sqrt{2l+1} \int_0^\beta d\tau p_l(x(\tau)) G_{\alpha\bar{\alpha}}(\tau). \quad (41)$$

This can be verified by replacing (40) into (41) and then using the orthogonality property of the Legendre polynomials

$$\int_{-1}^1 dx P_l(x) P_{l'}(x) dx = \frac{2}{2l+1} \delta_{l,l'}.$$

We can then sample directly the Legendre coefficients; for a given configuration

$$\begin{aligned} \langle G_{\alpha\bar{\alpha}}^l \rangle_c &= \sum_{i,j=1}^n \tilde{P}_l(\tau_j - \bar{\tau}_i) [M^{(n)}]_{j,i} \delta_{\alpha_j \alpha} \delta_{\bar{\alpha}_i \bar{\alpha}} \\ \tilde{P}_l(\tau) &= -\frac{\sqrt{2l+1}}{\beta} \times \begin{cases} P_l(x(\tau)), & \tau > 0 \\ -P_l(x(\tau + \beta)), & \tau < 0 \end{cases} \end{aligned}$$

Since the coefficients G_l decay fast for large l , it is often convenient to work with the Legendre rather than with the Matsubara representation, or with a mixed representation. In particular, the Bethe-Salpeter equation can be rewritten in terms of Legendre coefficients for the susceptibility. Let us first start from the Green function matrix

$$\begin{aligned} G_{\alpha\gamma}(\nu_n) &= \int_0^\beta d\tau_{12} e^{i\nu_n \tau_{12}} G_{\alpha\gamma}(\tau_{12}) = \sum_l T_{n,l} G_{\alpha\gamma}^l \\ T_{n,l} &= \frac{\sqrt{2l+1}}{\beta} \int_0^\beta d\tau_{12} e^{i\nu_n \tau_{12}} p_l(x(\tau_{12})). \end{aligned}$$

We can now express the non-interacting susceptibility in terms of the transformation matrices $T_{n,l}$ and the Legendre coefficients $\chi_0^{l,l'}(\omega_m)$

$$\begin{aligned} [\chi_0(\omega_m)]_{n,n'} &= \frac{1}{16} \iiint d\boldsymbol{\tau} e^{i[\nu_n \tau_{12} + \nu_{n'} \tau_{34}]} \chi_{\gamma\gamma'}^{\alpha\alpha'}(\boldsymbol{\tau}) e^{-i\omega_m \tau_{23}} \\ &= -\frac{1}{16} \iiint d\boldsymbol{\tau} e^{i[\nu_n \tau_{12} + \nu_{n'} \tau_{34}]} G_{\alpha\gamma'}(\tau_{14}) G_{\gamma\alpha'}(-\tau_{23}) e^{-i\omega_m \tau_{23}} \\ &= \sum_{ll'} T_{n,l} \chi_0^{ll'}(\omega_m) T_{n',l}^*. \end{aligned}$$

A similar relation holds for the full susceptibility $[\chi(\omega_m)]_{n,n'} = \sum_{ll'} T_{n,l} \chi^{ll'}(\omega_m) T_{n',l}^*$.

4 Conclusion

In this lecture, I have introduced some of the fundamental aspects of linear-response theory, with focus on strongly correlated materials. Along the way, we have seen an important theorem connecting the linear susceptibility and the correlation function, the fluctuation-dissipation theorem; we have discussed the analytic and symmetry properties of the linear susceptibility tensor; we have introduced the thermodynamic- and the f-sum rule. In the second part of the lecture, we have seen how to calculate the susceptibility using dynamical mean-field theory, the state-of-the art approach for strongly correlated materials. Within this method, the local susceptibility is obtained via the quantum-impurity solver; the q -dependent susceptibility can, instead, be calculated solving the Bethe-Salpeter equation in the local-vertex approximation. We have seen that the vertex in the Bethe-Salpeter equation plays a crucial role. As representative examples we have used the one-band Hubbard model and when possible the generalized multi-band Hubbard model. Finally, we have presented two impurity solvers, the Hirsch-Fye QMC and the hybridization-expansion continuous-time QMC method.

Linear response functions are of fundamental importance when we want to compare our theory of a given phenomenon to experiments. In addition, the linear susceptibility tensor is a key ingredient of extensions of DMFT such as, e.g., the dual-fermion and the dual-boson approach.

Acknowledgment

I would like to thank Amin Kiani Sheikhabadi for many questions and discussions. Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

Appendix

Atomic units

In this lecture, formulas are expressed in atomic units unless specified otherwise. The unit of mass m_0 is the electron mass ($m_0 = m_e$), the unit of charge is the electron charge ($e_0 = e$), the unit of length is the Bohr radius ($a_0 = a_B \sim 0.52918 \text{ \AA}$), and the unit of time $t_0 = 4\pi\varepsilon_0\hbar a_0/e^2$. In these units, the numerical value of the Bohr radius a_B , of the electron charge e , of the electron mass m_e , of $1/4\pi\varepsilon_0$, and of \hbar is 1. Furthermore, the speed of light is $c = 1/\alpha \sim 137$, the Bohr magneton $\mu_B = 1/2$, and the unit of energy is the Hartree (1 Ha $\sim 27.211 \text{ eV}$).

Fourier transforms

We use the following conventions for the Fourier transforms. su For the direct and inverse Fourier transform in frequency and time

$$\begin{aligned} f(\omega) &= \int_{-\infty}^{\infty} dt f(t) e^{i\omega t} \\ f(t) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} f(\omega) e^{-i\omega t}. \end{aligned}$$

For the direct and inverse transform in spatial/momentum coordinates

$$\begin{aligned} g(\mathbf{q}) &= \int d\mathbf{r} g(\mathbf{r}) e^{i\mathbf{q} \cdot \mathbf{r}} \\ g(\mathbf{r}) &= \int \frac{d\mathbf{q}}{(2\pi)^3} g(\mathbf{q}) e^{-i\mathbf{q} \cdot \mathbf{r}}. \end{aligned}$$

For ideal lattices and \mathbf{k} vectors in the first Brillouin Zone

$$\begin{aligned} \frac{1}{N_s} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{T}} &= \delta_{\mathbf{T}, \mathbf{0}}, \\ \frac{1}{N_s} \sum_{\mathbf{T}} e^{i\mathbf{k} \cdot \mathbf{T}} &= \sum_{\mathbf{G}} \delta_{\mathbf{k}, \mathbf{G}}, \end{aligned}$$

where N_s is the number of lattice sites.

Dirac delta function

The Dirac delta function is defined as

$$\delta(x) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega x}.$$

Some of the properties of the delta function:

$$\begin{aligned}\int f(x)\delta(x-b)dx &= f(b), \\ \delta(x) &= \delta(-x), \\ \delta(ax) &= \frac{1}{|a|}\delta(x), \\ \delta(\mathbf{r}) &= \delta(x)\delta(y)\delta(z).\end{aligned}$$

Heisenberg representation

Real time:

$$\hat{A}(t) = e^{i(\hat{H}-\mu\hat{N})t} \hat{A} e^{-i(\hat{H}-\mu\hat{N})t}.$$

Imaginary time:

$$\hat{A}(\tau) = e^{(\hat{H}-\mu\hat{N})\tau} \hat{A} e^{-(\hat{H}-\mu\hat{N})\tau},$$

$$[\hat{A}(\tau)]^\dagger = e^{-(\hat{H}-\mu\hat{N})\tau} \hat{A}^\dagger e^{(\hat{H}-\mu\hat{N})\tau} = \hat{A}^\dagger(-\tau)$$

Fermi-Dirac distribution function

The Fermi-Dirac distribution function is defined as

$$n(\varepsilon) = \frac{1}{1 + e^{\beta\varepsilon}}.$$

In the lecture we used the following relations

$$\begin{aligned}1 - n(\varepsilon) &= \frac{e^{\beta\varepsilon}}{1 + e^{\beta\varepsilon}} = n(-\varepsilon), \\ n(\varepsilon)[1 - n(\varepsilon)] &= \frac{e^{\beta\varepsilon}}{(1 + e^{\beta\varepsilon})^2} = -\frac{1}{\beta} \frac{dn(\varepsilon)}{d\varepsilon}, \\ \frac{dn(\varepsilon)}{d\varepsilon} &= -\frac{dn(-\varepsilon)}{d\varepsilon}.\end{aligned}$$

Futhermore

$$\begin{aligned}\lim_{T \rightarrow 0} \beta n(\varepsilon)[1 - n(\varepsilon)] &= \delta(\varepsilon), \\ \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} [n(\varepsilon) - n(\varepsilon + \Delta)] &= \delta(\varepsilon).\end{aligned}$$

Analytic functions

Here, we shortly summarize the properties of complex functions which have been used in the lecture. A complex function is said to be *analytic* on a domain D in the complex plane if it is differentiable at every point inside D . If $f(z)$ is analytic at all points within and on a closed path \mathcal{C} (traversed in anti-clockwise direction) the Cauchy integral theorem holds

$$\oint_{\mathcal{C}} f(z) dz = 0. \quad (42)$$

The points on which a function $f(z)$ is not analytic are called *singularities*. There are two types of singularities, *isolated* (i.e., at some specific $z = z_0$) and *extended* singularities. If a function $f(z)$ has an isolated singularity in $z = z_0$ and there is an integer n such that for $m \geq n$ the function $(z - z_0)^m f(z)$ is analytic in z_0 , then z_0 is said to be a *pole of order n* of the function $f(z)$. Apart from poles, other types of isolated singularities are *essential singularities*, *logarithmic singularities* and *removable singularities*. Let us consider some examples of complex functions that we have used in this lecture. The function

$$f(z) = \frac{1}{z - z_0}$$

has a single pole in z_0 ; this pole is of first order. The Fermi-Dirac distribution function

$$f(z) = \frac{1}{1 + e^{\beta z}}$$

has instead infinite poles, all the fermionic Matsubara frequencies $z = i\nu_n = i(2n + 1)\pi/\beta$. In a similar way the Bose distribution function

$$f(z) = \frac{1}{e^{\beta z} - 1}$$

has infinite poles at the bosonic Matsubara frequencies $z = i\omega_m = i2m\pi/\beta$.

If within the contour \mathcal{C} a function $f(z)$ is analytic except for a set of poles $\{z_n\}$ of order one, the residue theorem holds

$$\begin{aligned} \oint_{\mathcal{C}} f(z) dz &= 2\pi i \sum_{\{z_n\}} \text{Res}[f(z_n)], \\ \text{Res}[f(z_n)] &= \lim_{z \rightarrow z_n} f(z)(z - z_n), \end{aligned}$$

where the term $\text{Res}[f(z_n)]$ is called residue of the function $f(z)$ at the point z_n .

Extended singularities are of two types, *natural boundaries* and *branch cuts*. The latter are curves (e.g., lines or segments) in the complex plane across which a multi-valued function is discontinuous. Let us consider as example the function $f(z) = \sqrt{z}$; this function is double-valued, for example, in $z = 1$ it can take both the values ± 1 . Let us define the principal value of the function \sqrt{z} as the positive square root (+1), and let us rewrite $z = |z|e^{i\phi}$; the function \sqrt{z} has then a branch cut at $\phi = 2\pi$, i.e., along the line $y \geq 0$.

Spin and magnetization operators

The spin operators \hat{S}_ν are defined as

$$\hat{S}_\nu = \frac{1}{2} \sum_{\sigma\sigma'} c_\sigma^\dagger \sigma_\nu c_{\sigma'},$$

where $\nu = x, y, z$ and $\hat{\sigma}_\nu$ are the Pauli matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The magnetization operators \hat{M}_ν are defined as $\hat{M}_\nu = -g\mu_B \hat{S}_\nu$.

Useful formulas

- Sokhotski-Plemelj formula

$$\frac{1}{\omega + i\delta} = \mathcal{P}\frac{1}{\omega} - i\pi\delta(\omega)$$

- Cauchy principal value

$$\int_{-\infty}^{\infty} \mathcal{P}\frac{1}{\omega} d\omega = \int_{-\infty}^{-\varepsilon} \frac{1}{\omega} d\omega + \int_{\varepsilon}^{\infty} \frac{1}{\omega} d\omega,$$

- Fourier representation of $\Theta(t)$ function

$$\Theta(t) = i \int_{-\infty}^{\infty} \frac{d\omega''}{2\pi} e^{-i\omega''t} \frac{1}{\omega'' + i\delta},$$

- Integral of imaginary exponential

$$I(x) = \int_0^{\infty} e^{ixt} dt = \frac{i}{x + i\delta}, \quad \delta = 0^+.$$

Atomic magnetic susceptibility

Let us consider an idealized single-level atom described by the Hamiltonian $\hat{H}_U = Un_\uparrow n_\downarrow$. The eigenstates of this system, $|\Psi_i^N\rangle$, as well as the expectation values $E_i = \langle \Psi_i^N | \hat{H}_U - \mu \hat{N} | \Psi_i^N \rangle$ at half-filling, are given in the table below

| $ \Psi_i^N\rangle$ | N | $E_i = \langle \Psi_i^N \hat{H}_U - \mu \hat{N} \Psi_i^N \rangle$ |
|---|-----|---|
| $ 0\rangle$ | 0 | 0 |
| $c_\sigma^\dagger 0\rangle$ | 1 | $-U/2$ |
| $c_\uparrow^\dagger c_\downarrow^\dagger 0\rangle$ | 2 | 0 |

The magnetic susceptibility in Matsubara space is given by

$$[\chi_{zz}]_{nn'}(i\omega_m) = \beta \frac{1}{4} (g\mu_B)^2 \sum_P \text{sign}(P) f_P$$

$$f_P(i\omega_{P_1}, i\omega_{P_2}, i\omega_{P_3}) = \int_0^\beta d\tau_{14} \int_0^{\tau_{14}} d\tau_{24} \int_0^{\tau_{24}} d\tau_{34} e^{i\omega_{P_1}\tau_{14} + i\omega_{P_2}\tau_{24} + i\omega_{P_3}\tau_{34}} f_P(\tau_{14}, \tau_{24}, \tau_{34})$$

where $P = A, B, \dots$ are the six possible permutations of the indices (123) and

$$f_P(\tau_{14}, \tau_{24}, \tau_{34}) = \frac{1}{Z} \sum_{\sigma\sigma'} \sigma\sigma' \text{Tr} e^{-\beta(\hat{H}-\mu\hat{N})} \left[\hat{o}_{P_1}(\tau_{14}) \hat{o}_{P_2}(\tau_{24}) \hat{o}_{P_3}(\tau_{34}) c_{\sigma'}^\dagger \right]$$

$$= \frac{1}{Z} \sum_{\sigma\sigma'} \sigma\sigma' \sum_{ijkl} e^{-\beta E_i} \langle i | \hat{o}_{P_1} | j \rangle \langle j | \hat{o}_{P_2} | k \rangle \langle k | \hat{o}_{P_3} | l \rangle \langle l | c_{\sigma'}^\dagger | i \rangle$$

$$\times [e^{\Delta E_{ij}\tau_{14} + \Delta E_{jk}\tau_{24} + \Delta E_{kl}\tau_{34}}],$$

where $\Delta E_{ij} = E_i - E_j$. For the identity permutation the operators are $\hat{o}_{P_1} = c_\sigma$, $\hat{o}_{P_2} = c_\sigma^\dagger$, and $\hat{o}_{P_3} = c_{\sigma'}$ and the frequencies are $\omega_1 = \nu_n$, $\omega_2 = -\omega_m - \nu_n$, $\omega_3 = \omega_m + \nu_{n'}$. This expression can be used to calculate the magnetic susceptibility of any one-band system whose eigenvalues and eigenvectors are known, e.g., via exact diagonalization. In the case of our idealized atom

$$f_E(\tau_{14}, \tau_{24}, \tau_{34}) = \frac{1}{(1 + e^{\beta U/2})} e^{\beta U/2} e^{-(\tau_{12} + \tau_{34})U/2} = \frac{1}{(1 + e^{\beta U/2})} g_E(\tau_{14}, \tau_{24}, \tau_{34}).$$

The frequencies and functions $f_P(\tau_{14}, \tau_{24}, \tau_{34})$ for all permutations are given in the table below

| | ω_{P_1} | ω_{P_2} | ω_{P_3} | $g_P(\tau_{14}, \tau_{24}, \tau_{34})$ | $\text{sign}(P)$ |
|----------|-----------------------|-----------------------|-----------------------|--|------------------|
| $E(123)$ | ν_n | $-\omega_m - \nu_n$ | $\omega_m + \nu_{n'}$ | $e^{\beta U/2} e^{-(\tau_{12} + \tau_{34})U/2}$ | + |
| $A(231)$ | $-\omega_m - \nu_n$ | $\omega_m + \nu_{n'}$ | ν_n | $-e^{\beta U/2} e^{-(\tau_{12} + \tau_{34})U/2}$ | + |
| $B(312)$ | $\omega_m + \nu_{n'}$ | ν_n | $-\omega_m - \nu_n$ | $-e^{+(\tau_{12} + \tau_{34})U/2}$ | + |
| $C(213)$ | $-\omega_m - \nu_n$ | ν_n | $\omega_m + \nu_{n'}$ | $-e^{\beta U/2} e^{-(\tau_{12} + \tau_{34})U/2}$ | - |
| $D(132)$ | ν_n | $\omega_m + \nu_{n'}$ | $-\omega_m - \nu_n$ | $-e^{+(\tau_{12} + \tau_{34})U/2}$ | - |
| $F(321)$ | $\omega_m + \nu_{n'}$ | $-\omega_m - \nu_n$ | ν_n | $e^{+(\tau_{12} + \tau_{34})U/2}$ | - |

The missing ingredient is the integral

$$I_P(x, -x, x; i\omega_{P_1}, i\omega_{P_2}, i\omega_{P_3}) = \int_0^\beta d\tau_{14} \int_0^{\tau_{14}} d\tau_{24} \int_0^{\tau_{24}} d\tau_{34} e^{i\omega_{P_1}\tau_{14} + i\omega_{P_2}\tau_{24} + i\omega_{P_3}\tau_{34}} e^{x(\tau_{14} - \tau_{24} + \tau_{34})}$$

$$= + \int_0^\beta d\tau_{14} \int_0^{\tau_{14}} d\tau \int_0^{\tau_{14}-\tau} d\tau' e^{(i\omega_{P_1} + i\omega_{P_2} + i\omega_{P_3} + x)\tau_{14} - i(\omega_{P_2} + \omega_{P_3})\tau} e^{-(i\omega_{P_3} + x)\tau'}$$

$$= + \frac{1}{i\omega_{P_3} + x} \frac{1}{-i\omega_{P_2} + x} \left[\frac{1}{i\omega_{P_1} + x} \frac{1}{n(x)} + \beta \delta_{\omega_{P_1} + \omega_{P_2}} \right]$$

$$+ \frac{1}{i\omega_{P_3} + x} \frac{1 - \delta_{\omega_{P_2} + \omega_{P_3}}}{i(\omega_{P_2} + \omega_{P_3})} \left[\frac{1}{i\omega_{P_1} + x} - \frac{1}{i(\omega_{P_1} + \omega_{P_2} + \omega_{P_3}) + x} \right] \frac{1}{n(x)}$$

$$+ \delta_{\omega_{P_2} + \omega_{P_3}} \frac{1}{i\omega_{P_3} + x} \left\{ \left[\frac{1}{(i\omega_{P_1} + x)} \right]^2 \frac{1}{n(x)} - \beta \left[\frac{1}{(i\omega_{P_1} + x)} \right] \frac{1 - n(x)}{n(x)} \right\}.$$

where $x = \pm U/2$, depending on the permutation. The general expression of the integral $I_P(\Delta E_{ij}, \Delta E_{jk}, \Delta E_{kl}; i\omega_1, i\omega_2, i\omega_3)$ can be found in Refs. [32, 33].

References

- [1] H. Bruus and K. Flensberg: *Many-Body Quantum Theory in Condensed Matter Physics* (Oxford University Press, 2004)
- [2] K. Yosida: *Theory of Magnetism* (Springer, Heidelberg, 1998)
- [3] J. Sóliom: *Fundamentals of the Physics of Solids* (Springer, Berlin Heidelberg, 2010)
- [4] E. Pavarini and E. Koch, *Introduction* in Ref. [5]
- [5] E. Pavarini, E. Koch and U. Schöllwöck:
Emergent Phenomena in Correlated Matter,
Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
<http://www.cond-mat.de/events/correl13>
- [6] E. Pavarini, E. Koch, A. Lichtenstein and D. Vollhardt:
The LDA+DMFT Approach to Strongly Correlated Materials,
Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
<http://www.cond-mat.de/events/correl11>
- [7] E. Pavarini, E. Koch, F. Anders and M. Jarrell:
Correlated Electrons: from Models to Materials,
Reihe Modeling and Simulation, Vol. 2 (Forschungszentrum Jülich, 2012)
<http://www.cond-mat.de/events/correl12>
- [8] O. Gunnarsson, *Strongly Correlated Electrons: Estimates of Model Parameters*, in Ref. [7]
- [9] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen and O.K. Andersen,
Phys. Rev. Lett. **87**, 047003 (2001)
- [10] E. Pavarini, S. Biermann, A. Poteryaev, A.I. Lichtenstein, A. Georges and O.K. Andersen,
Phys. Rev. Lett. **92**, 176403 (2004)
- [11] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [12] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989); Z. Phys. B **76**, 211 (1989);
Int. J. Mod. Phys. B **3**, 2169 (1989)
- [13] F. Ohkawa, J. Phys. Soc. Jap. **60**, 3218 (1991)
- [14] A. Georges and G. Kotliar, Phys. Rev. B **89**, 6479 (1992)
- [15] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
- [16] V. Anisimov, A. Poteryaev, M. Korotin, A. Anokhin and G. Kotliar
J. Phys.: Condens. Matter **9**, 7359 (1997)

- [17] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57** 6884 (1998)
- [18] E. Pavarini, E. Koch and A.I. Lichtenstein, Phys. Rev. Lett. **101**, 266405 (2008);
E. Pavarini and E. Koch, Phys. Rev. Lett. **104**, 086402 (2010)
- [19] E. Pavarini, *The LDA+DMFT Approach*, in Ref. [6]
- [20] J. Kunes, *Wannier Functions and Construction of Model Hamiltonians*, in Ref. [6]
- [21] E. Pavarini, *Crystal-field Theory, Tight-binding Method and Jahn-Teller Effect*,
in Ref. [7]
- [22] E. Pavarini, *Magnetism: Models and Mechanisms*, in Ref. [5]
- [23] V. Slatić and B. Horvatić, Sol. Stat. Comm. **75**, 263 (1990)
- [24] A. Georges, G. Kotliar, W. Krauth and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [25] A. Kiani Sheikhabadi, PhD thesis, RWTH Aachen and FZ Jülich (2014)
- [26] L. Boehnke, H. Hafermann, M. Ferrero, F. Lechermann and O. Parcollet,
Phys. Rev. B **84**, 075145 (2011)
- [27] E. Koch, *Exchange Mechanisms*, in Ref. [7]
- [28] A.C. Hewson: *The Kondo Problem to Heavy Fermions*
(Cambridge University Press, 1993)
- [29] J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)
- [30] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner,
Rev. Mod. Phys. **83**, 349 (2011)
- [31] A. Flesch, E. Gorelov, E. Koch and E. Pavarini, Phys. Rev. B **87**, 195141 (2013)
- [32] A. Dolfen, PhD thesis, RWTH Aachen and GRS Jülich (2010)
- [33] H. Hafermann, C. Jung, S. Brener, M.I. Katsnelson, A.N. Rubtsov and A.I. Lichtenstein,
Eur. Phys. Lett. **85**, 27007 (2009)

7 Continuous-time QMC Solvers for Electronic Systems in Fermionic and Bosonic Baths

Fakher F. Assaad

Institut für Theoretische Physik und Astrophysik
Universität Würzburg

Contents

| | | |
|----------|--|-----------|
| 1 | The single impurity Anderson model (SIAM) | 3 |
| 2 | CT-INT | 4 |
| 2.1 | The partition function | 4 |
| 2.2 | Observables and Wick's theorem | 6 |
| 2.3 | The negative sign problem | 8 |
| 2.4 | The Monte Carlo sampling | 10 |
| 2.5 | Fast updates | 11 |
| 2.6 | Average expansion parameter | 13 |
| 3 | CT-HYB | 13 |
| 3.1 | The partition function | 13 |
| 3.2 | The Monte Carlo sampling and evaluation of the trace | 16 |
| 3.3 | Selected applications | 17 |
| 4 | Application of CT-INT to the Hubbard-Holstein model | 18 |
| 4.1 | Integrating-out the phonons | 18 |
| 4.2 | Formulation of CT-INT for the Hubbard-Holstein model | 19 |
| 4.3 | The quarter-filled Holstein model from adiabatic to anti-adiabatic phonons | 21 |
| 5 | Concluding remarks | 23 |
| A | Basic principles of Monte Carlo methods in statistical mechanics | 24 |
| A.1 | The central limit theorem | 24 |
| A.2 | Jackknife and bootstrap methods for error evaluation | 27 |
| A.3 | Markov chains | 28 |
| A.4 | Construction of the transition matrix T | 31 |
| A.5 | One-dimensional Ising model | 33 |
| B | Proof of the determinant identity | 35 |

Continuous-time quantum Monte Carlo (CT-QMC) methods are a tool of choice to solve correlated electron problems embedded in bosonic or fermionic baths [1]. This is precisely the problem that is encountered in dynamical mean-field theories (DMFT), where in the limit of infinite coordination number the environment can be replaced by a fermionic bath [2]. In DMFT, the physics of the Hubbard model in the limit of infinite coordination number maps onto that of the single-impurity Anderson model (SIAM). This mapping provides invaluable insight into the important problem of the Mott transition [3]. The beauty of the CT-QMC algorithms lie in their flexibility. Rather than being Hamiltonian-based – like the auxiliary-field QMC method (see Ref. [4] for a review), they are action-based and allow the simulation of effective low-energy models after having integrated out high-energy degrees of freedom. This aspect of the method has spurred many applications. In the domain of model-building, screening effects by high energy bands, which can be taken into account within the constrained random phase approximation (cRPA) [5], naturally lead to a low-energy effective model with retarded interactions which only has an action-based formulation. Retarded interactions are also obtained in the context of electron-phonon interactions. Here, one can integrate out the bosonic phonon modes at the expense of a retarded interaction [6–9]. Other applications of the algorithm have been introduced in the realm of topological insulators [10, 11]. In this context, helical liquids can only be realized as the edge theory of a quantum spin Hall insulator [12]. In many cases, correlation effects can be neglected in the bulk but are dominant on the edge [13]. Thereby, one can retain interactions along the edge of the system and view the bulk as a bath, which one can readily integrate out [14–17]. Further applications of the CT-QMC include for example formulations along the Keldysh contour (see Ref. [18]) or applications within the realm of extensions to DMFT methods to include spatial fluctuations. Here one can mention cluster generalizations such as the dynamical cluster or cellular DMFT approximations [19], the dual fermion approach [20], the dynamical vertex approximation [21], or extended DMFT [22].

There is a price to the flexibility of the CT-QMC algorithms. In the best-case scenario – absence of a sign problem – the computational time scales as the third power of the Euclidean volume; to be more precise $(N_{imp}\beta)^3$, where β corresponds to the inverse temperature and N_{imp} to the number of impurities. This scaling has to be contrasted with the auxiliary-field methods [4], which have linear scaling in β . Such algorithms have recently been used in the context of DMFT [23]. In the worst case, all stochastic methods are prone to the so-called negative sign problem, which effectively leads to a signal-to-noise ratio that grows exponentially in the Euclidian volume. There is to date no solution to this problem. Different algorithms or different formulations of the same algorithm can lead to very different sign problems. Clever tricks such as the fermion-bag approach can sometimes solve the problem in special situations [24].

Another issue when opting for stochastic algorithms – as opposed to NCA or tensor-network based approaches – is the fact that QMC operates on the imaginary time axis. To produce spectral functions on the real-frequency axis so as to compare with experiments, analytical continuation is necessary. This is a numerically ill-conditioned problem which limits the precision for the calculation of spectral functions. This issue may be especially severe when considering multi-orbital problems with complicated spectral line-shapes. The only solution to this prob-

lem is to work directly on the real time axis, or to analytically continue spectral functions with *simple* line shapes.

The CT-QMC approach is the method of choice for thermodynamics and equal time correlation functions. It is unbiased. Given an adequate error analysis, exact results are reproducible within the cited error-bars. The CT-QMC methods have different formulations. The interaction expansion (CT-INT) [25] and auxiliary-field (CT-AUX) [26, 27] approaches turn out to be identical, with CT-INT being the more general. The CT-INT is a *weak coupling* method and to access the strong coupling it is more convenient to use the hybridization expansion CT-HYB [28].

The organization of this lecture is the following. In section 2, we concentrate on the CT-INT [25, 26] approach in the context of the SIAM. We will review this algorithm in detail, since it has the potential for tackling lattice problems. In section 3, we will cover the basic ideas of the CT-HYB algorithm [28]. The CT-HYB is certainly the method of choice for multi-orbital models in the strong coupling limit. In section 4, we provide a generalization of the CT-INT method to tackle problems with bosonic baths. Here, the bosonic bath corresponds to a phonon mode that after integration leads to a retarded interaction.

For completeness, we have included an appendix that reviews the basic ideas of the Monte Carlo method as well as a discussion of error analysis. This is an important aspect of the implementation of CT-INT algorithms.

1 The single impurity Anderson model (SIAM)

The Anderson impurity model [29] describes the formation and screening of local magnetic moments in a metallic host. The metallic host is described by a conduction electron band with dispersion relation $\epsilon(\mathbf{k})$. The impurity state is described by a localized Kramers doublet orbital. The localized nature of the impurity orbital obliges one to include the Coulomb repulsion in terms of a Hubbard U . Finally, a hybridization matrix element $V_{\mathbf{k}}$ allows for charge transfer between the localized orbital and extended Bloch states. In second quantization, the model is given by

$$\hat{H}_{\text{SIAM}} = \sum_{\mathbf{k},\sigma} \epsilon(\mathbf{k}) \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{c}_{\mathbf{k},\sigma} + \sum_{\mathbf{k},\sigma} \left(V_{\mathbf{k}} \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{f}_\sigma + \bar{V}_{\mathbf{k}} \hat{f}_\sigma^\dagger \hat{c}_{\mathbf{k},\sigma} \right) + \epsilon_f \sum_\sigma \hat{n}_\sigma + U \hat{n}_\uparrow \hat{n}_\downarrow. \quad (1)$$

Here, $\hat{c}_{\mathbf{k},\sigma}^\dagger$ creates a Bloch electron with z -component of spin σ , \hat{f}_σ^\dagger an electron on the Kramers doublet localized orbital and $\hat{n}_\sigma = \hat{f}_\sigma^\dagger \hat{f}_\sigma$. A discussion of the physics described by the single impurity Anderson model can be found in Ref. [30]. In the spirit of action based CT-QMC algorithms, it is convenient to integrate out the conduction electrons. To carry out this step, we introduce fermion coherent states $|c, f\rangle$ that satisfy

$$\hat{c}_{\mathbf{k},\sigma}|c, f\rangle = c_{\mathbf{k},\sigma}|c, f\rangle, \quad \hat{f}_\sigma|c, f\rangle = f_\sigma|c, f\rangle, \quad (2)$$

with $c_{\mathbf{k},\sigma}$ and f_σ being Grassmann variables. Using standard many body formalism, reviewed for example in [31], the partition function of the SIAM is given by a path integral over Grass-

mann variables

$$Z_{\text{SIAM}} = \text{Tr} \left[e^{-\beta \hat{H}_{\text{SIAM}}} \right] = \int D \{ c^\dagger c f^\dagger f \} e^{-\int_0^\beta d\tau [\sum_{\mathbf{k},\sigma} c_{\mathbf{k},\sigma}^\dagger(\tau) \frac{\partial}{\partial \tau} c_{\mathbf{k},\sigma}(\tau) + \sum_\sigma f_\sigma^\dagger(\tau) \frac{\partial}{\partial \tau} f_\sigma(\tau) + H_{\text{SIAM}}(c^\dagger, c, f^\dagger, f)]}. \quad (3)$$

Since the Grassmann variables satisfy anti-periodic boundary conditions in β , we can define the Fourier transform

$$f_\sigma(i\omega_m) = \frac{1}{\sqrt{\beta}} \int_0^\beta d\tau e^{i\omega_m \tau} f_\sigma(\tau) \quad (4)$$

with $\omega_m = (2m+1)\pi/\beta$ a fermionic Matsubra frequency. An equivalent equation holds for the conduction electrons.

Owing to the fact that the action is bilinear in the conduction electrons one can integrate them out with a Gaussian integration to obtain our final result

$$Z_{\text{SIAM}} = \int D \{ f^\dagger f \} e^{-S(f^\dagger, f)} \quad \text{with} \\ S(f^\dagger, f) = - \int_0^\beta d\tau d\tau' \sum_\sigma f_\sigma^\dagger(\tau) G_0^{-1}(\tau - \tau') f_\sigma(\tau') + U \int_0^\beta d\tau f_\uparrow^\dagger(\tau) f_\uparrow(\tau) f_\downarrow^\dagger(\tau) f_\downarrow(\tau). \quad (5)$$

Here $G_0(\tau - \tau') = -\langle T \hat{f}_\sigma(\tau) \hat{f}_\sigma^\dagger(\tau') \rangle_0$ corresponds to the non-interacting f -Green function. The Gaussian integration yields

$$G_0^{-1}(\tau - \tau') = -\delta(\tau - \tau') \left[\frac{\partial}{\partial \tau'} + \epsilon_f \right] + \Delta(\tau - \tau') \quad \text{with} \quad \Delta(i\omega_m) = \sum_{\mathbf{k}} \frac{|V_{\mathbf{k}}|^2}{i\omega_m - \epsilon(\mathbf{k})}. \quad (6)$$

The above equation is the starting point for both the CT-INT and CT-HYB algorithms. The CT-INT follows the idea of expanding in the interaction term, whereas the CT-HYB expands in the hybridization. In this lecture, we will assume that the fermionic bath possesses $U(1)$ gauge symmetry. Generalizations of the CT-INT to account for superconducting leads can be found in Refs. [32, 33].

2 CT-INT

In this section, we will describe in some detail the implementation of the CT-INT and show that it is equivalent to the CT-AUX

2.1 The partition function

Anticipating the elimination of the sign problem in some cases and to establish the equivalence between the CT-INT and CT-AUX algorithms, we will rewrite the Hubbard interaction as

$$\frac{U}{2} \sum_{s=\pm 1} (\hat{n}_\uparrow - \rho/2 - s\delta) (\hat{n}_\downarrow - \rho/2 + s\delta). \quad (7)$$

Here we have introduced a additional Ising variable s , and ρ corresponds to the average electronic density. We will discuss the additional parameter δ later on in this section. Starting from the action in Eq. (5), we can Taylor expand in the Hubbard interaction to obtain

$$\frac{Z_{\text{SIAM}}}{Z_0} = \sum_{n=0}^{\infty} \left(\frac{-U}{2} \right)^n \frac{1}{n!} \int_0^{\beta} d\tau_1 \sum_{s_1} \cdots \int_0^{\beta} d\tau_n \sum_{s_n} \prod_{\sigma} \langle [n_{\sigma}(\tau_1) - \alpha_{\sigma}(s_1)] \cdots [n_{\sigma}(\tau_n) - \alpha_{\sigma}(s_n)] \rangle_0 \quad (8)$$

The expectation value $\langle \bullet \rangle_0$ is taken with respect to the non-interacting SIAM with partition function Z_0 and we have used the short cut notation

$$\alpha_{\sigma}(s) = \rho/2 + \sigma s \delta . \quad (9)$$

The thermal expectation value is the sum over all diagrams, connected and disconnected, of a given order n . Using the general formulation of Wick's theorem, this sum can be expressed as a determinant where the entries are the Green functions of the non-interacting system

$$\langle [n_{\sigma}(\tau_1) - \alpha_{\sigma}(s_1)] \cdots [n_{\sigma}(\tau_n) - \alpha_{\sigma}(s_n)] \rangle_0 = \det \underbrace{\begin{pmatrix} g_0(\tau_1, \tau_1) - \alpha_{\sigma}(s_1) & g_0(\tau_1, \tau_2) & \cdots & g_0(\tau_1, \tau_n) \\ g_0(\tau_2, \tau_1) & g_0(\tau_2, \tau_2) - \alpha_{\sigma}(s_2) & \cdots & g_0(\tau_2, \tau_n) \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ g_0(\tau_n, \tau_1) & g_0(\tau_n, \tau_2) & \cdots & g_0(\tau_n, \tau_n) - \alpha_{\sigma}(s_n) \end{pmatrix}}_{\equiv \mathbf{M}_{\sigma}(C_n)}, \quad (10)$$

where we have defined the Green function:

$$g_0(\tau_1, \tau_2) = \langle T \hat{f}_{\sigma}^{\dagger}(\tau_1) \hat{f}_{\sigma}(\tau_2) \rangle_0 , \quad (11)$$

which we have assumed to be spin-independent. In the above, T corresponds to the time ordering. The product of the two determinants is nothing but the sum over connected and disconnected Feynman diagrams. The summation over individual Feynman diagrams reduces the negative sign problem and, as we will see later, eliminates it altogether for a class of problems. A configuration C_n is defined by the n Hubbard vertices and Ising spins introduced in Eq. (7)

$$C_n = \{[\tau_1, s_1] \cdots [\tau_n, s_n]\} . \quad (12)$$

With the short-hand notation

$$\sum_{C_n} = \sum_{n=0}^{\infty} \int_0^{\beta} d\tau_1 \sum_{s_1} \cdots \int_0^{\beta} d\tau_n \sum_{s_n} , \quad (13)$$

the partition function can conveniently be written as

$$\frac{Z_{\text{SIAM}}}{Z_0} = \sum_{C_n} W(C_n), \text{ with } W(C_n) = \left(-\frac{U}{2} \right)^n \frac{1}{n!} \prod_{\sigma} \det \mathbf{M}_{\sigma}(C_n) . \quad (14)$$

Here $\mathbf{M}_{\sigma}(C_n)$ is the $n \times n$ matrix of Eq. (10).

2.2 Observables and Wick's theorem

Observables $\hat{O}(\tau)$ can now be computed with

$$\langle \hat{O}(\tau) \rangle = \frac{\sum_{C_n} W(C_n) \langle \langle \hat{O}(\tau) \rangle \rangle_{C_n}}{\sum_{C_n} W(C_n)}, \quad (15)$$

where for $\hat{O}(\tau) = \prod_{\sigma} \hat{O}_{\sigma}(\tau)$ we have

$$\langle \langle \hat{O}(\tau) \rangle \rangle_{C_n} = \frac{\prod_{\sigma} \langle T [\hat{n}_{\sigma}(\tau_1) - \alpha_{\sigma}(s_1)] \cdots [\hat{n}_{\sigma}(\tau_n) - \alpha_{\sigma}(s_n)] \hat{O}_{\sigma}(\tau) \rangle_0}{\prod_{\sigma} \langle T [\hat{n}_{\sigma}(\tau_1) - \alpha_{\sigma}(s_1)] \cdots [\hat{n}_{\sigma}(\tau_n) - \alpha_{\sigma}(s_n)] \rangle_0}. \quad (16)$$

We will compute the single-particle Green function and then show that any many-particle Green function can be expressed in terms of this quantity. This statement corresponds to Wick's theorem, which holds when expanding around a Gaussian theory.

Using the determinant identity given by Eq. (10), one will readily see that the single-particle Green function is given by the ratio of two determinants:

$$\langle \langle T \hat{f}_{\sigma}^{\dagger}(\tau) \hat{f}_{\sigma}(\tau') \rangle \rangle_{C_n} = \frac{\det B_{\sigma}(C_n)}{\det M_{\sigma}(C_n)} \quad (17)$$

where

$$B_{\sigma}(C_n) = \begin{pmatrix} & & g_0(\tau_1, \tau') \\ & \vdots & \\ \mathbf{M}_{\sigma}(C_n) & & g_0(\tau_n, \tau') \\ g_0(\tau, \tau_1) & \dots g_0(\tau, \tau_n) & g_0(\tau, \tau') \end{pmatrix}. \quad (18)$$

To compute the ratio of the two determinants, we use the determinant identity

$$\det(\mathbf{A} + \mathbf{u} \otimes \mathbf{v}) = \det(\mathbf{A}) (1 + \mathbf{v} \cdot \mathbf{A}^{-1} \mathbf{u}) \quad (19)$$

as well as the Sherman-Morrison formula

$$(\mathbf{A} + \mathbf{u} \otimes \mathbf{v})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{u} \otimes \mathbf{v} \mathbf{A}^{-1}}{1 + \mathbf{v} \cdot \mathbf{A}^{-1} \mathbf{u}} \quad (20)$$

to obtain

$$\begin{aligned} \det(\mathbf{A} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2) = \\ \det(\mathbf{A}) [(1 + \mathbf{v}_1 \cdot \mathbf{A}^{-1} \mathbf{u}_1) (1 + \mathbf{v}_2 \cdot \mathbf{A}^{-1} \mathbf{u}_2) - (\mathbf{v}_2 \cdot \mathbf{A}^{-1} \mathbf{u}_1) (\mathbf{v}_1 \cdot \mathbf{A}^{-1} \mathbf{u}_2)], \end{aligned} \quad (21)$$

where the outer product is given by $(\mathbf{u} \otimes \mathbf{v})_{i,j} = \mathbf{u}_i \mathbf{v}_j$ and the scalar product by $\mathbf{u} \cdot \mathbf{v} = \sum_i \mathbf{u}_i \mathbf{v}_i$. Eq. (20) can be formally derived by Taylor-expanding $(1 + \mathbf{A}^{-1} \mathbf{u} \otimes \mathbf{v})^{-1}$. Eq. (19) can equally be formally demonstrated by using the fact that $\det(\mathbf{A}) = \exp \text{Tr} \log(\mathbf{A})$.

Decomposing the $\mathbf{B}_{\sigma}(C_n)$ matrix as

$$\mathbf{B}_{\sigma}(C_n) = \begin{pmatrix} & & 0 \\ & \vdots & \\ \mathbf{M}_{\sigma}(C_n) & & 0 \\ 0 & \dots 0 & 1 \end{pmatrix} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2 \quad (22)$$

with

$$\mathbf{u}_1 = (g_0(\tau_1, \tau'), \dots, g_0(\tau_n, \tau'), g_0(\tau, \tau') - 1), \quad (\mathbf{v}_1)_i = \delta_{i,n+1}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n+1}, \quad \mathbf{v}_2 = (g_0(\tau, \tau_1), \dots, g_0(\tau, \tau_n), 0)$$

yields the following expression for the single-particle Green function

$$g(\tau, \tau')_{C_n} \equiv \langle\langle T\hat{f}_\sigma^\dagger(\tau)\hat{f}_\sigma(\tau')\rangle\rangle_{C_n} = g_0(\tau, \tau') - \sum_{r,s=1}^n g_0(\tau, \tau_r) (\mathbf{M}_\sigma(C_n)^{-1})_{r,s} g_0(\tau_s, \tau'). \quad (23)$$

An important consequence of a continuous-time formulation is that one can compute the Green function directly in Matsubara frequencies. With the Fourier transformation of Eq. (4) one obtains:

$$g(i\omega_m, i\omega'_m)_{C_n} = \delta_{\omega_m, \omega'_m} g_0(i\omega_m) - g_0(i\omega_m) \left(\frac{1}{\beta} \sum_{r,s=1}^n e^{-i\omega_m \tau_r} (\mathbf{M}_\sigma(C_n)^{-1})_{r,s} e^{i\omega'_m \tau_s} \right) g_0(i\omega'_m). \quad (24)$$

For a given configuration of vertices C_n , translation symmetry in imaginary time is broken such that $g(i\omega_m, i\omega'_m)_{C_n}$ has to be a function of two Matsubara frequencies. Clearly, translation symmetry has to be restored after summation over the configurations C_n has been carried out. One can use this fact to define improved estimators for the Green function.

Higher-order Green functions may be computed by using the matrix identity demonstrated in App. B [32]. Here we consider Green functions of the form $\langle T\hat{f}_\sigma^\dagger(1)\hat{f}_\sigma(1') \dots \hat{f}_\sigma^\dagger(m)\hat{f}_\sigma(m')\rangle$. For every configuration C_n , a relation similar to Wick's theorem can be found, which greatly simplifies the calculation of higher-order Green functions. The application of the ordinary Wick theorem to the denominator and the numerator of Eq. (16) yields

$$\langle\langle T\hat{f}_\sigma^\dagger(1)\hat{f}_\sigma(1') \dots \hat{f}_\sigma^\dagger(m)\hat{f}_\sigma(m')\rangle\rangle_{C_n} = \frac{\det \mathbf{B}_\sigma(C_n)}{\det \mathbf{M}_\sigma(C_n)}, \quad (25)$$

where now $\mathbf{B}_\sigma(C_n)$ is a $\mathbb{C}^{(n+m) \times (n+m)}$ matrix given by

$$\mathbf{B}_\sigma(C_n) = \begin{pmatrix} & g_0(\tau_1, 1') & \dots & g_0(\tau_1, m') \\ \mathbf{M}_\sigma(C_n) & \vdots & \ddots & \vdots \\ & g_0(\tau_n, 1') & \dots & g_0(\tau_n, m') \\ g_0(1, \tau_1) & \dots & g_0(1, \tau_n) & g_0(1, 1') & \dots & g_0(1, m') \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_0(m, \tau_1) & \dots & g_0(m, \tau_n) & g_0(m, 1') & \dots & g_0(m, m') \end{pmatrix}. \quad (26)$$

Defining the matrices $\mathbf{B}_\sigma^{ij}(C_n) \in \mathbb{C}^{(n+1) \times (n+1)}$,

$$\mathbf{B}_\sigma^{ij}(C_n) = \begin{pmatrix} & g_0(\tau_1, i') \\ \mathbf{M}_\sigma(C_n) & \vdots \\ & g_0(\tau_n, i') \\ g_0(j, \tau_1) & \dots & g_0(j, \tau_n) & g_0(j, i') \end{pmatrix}, \quad (27)$$

we can make use of the determinant identity (145) yielding

$$\frac{\det \mathbf{B}_{C_n}}{\det \mathbf{M}_{C_n}} = \frac{1}{(\det \mathbf{M}_{C_n})^n} \det \begin{pmatrix} \det \mathbf{B}_{C_n}^{11} & \dots & \det \mathbf{B}_{C_n}^{1m} \\ \vdots & \ddots & \vdots \\ \det \mathbf{B}_{C_n}^{m1} & \dots & \det \mathbf{B}_{C_n}^{mm} \end{pmatrix}. \quad (28)$$

From Eq. (17) and (18), it is obvious that $\det \mathbf{B}_\sigma^{ij}(C_n) / \det \mathbf{M}_\sigma(C_n)$ is identical to the contribution of the configuration C_n to the one particle Green's function $\langle \langle T \hat{f}_\sigma^\dagger(j) \hat{f}_\sigma(i') \rangle \rangle$. Hence, Wick's theorem holds for every configuration C_n and is given by

$$\langle \langle T \hat{f}_\sigma^\dagger(1) \hat{f}_\sigma(1') \dots \hat{f}_\sigma^\dagger(m) \hat{f}_\sigma(m') \rangle \rangle_{C_n} = \det \begin{pmatrix} \langle \langle T \hat{f}_\sigma^\dagger(1) \hat{f}_\sigma(1') \rangle \rangle_{C_n} & \dots & \langle \langle T \hat{f}_\sigma^\dagger(1) \hat{f}_\sigma(m') \rangle \rangle_{C_n} \\ \vdots & \ddots & \vdots \\ \langle \langle T \hat{f}_\sigma^\dagger(m) \hat{f}_\sigma(1') \rangle \rangle_{C_n} & \dots & \langle \langle T \hat{f}_\sigma^\dagger(m) \hat{f}_\sigma(m') \rangle \rangle_{C_n} \end{pmatrix}. \quad (29)$$

This demonstrates that knowing the single particle Green function $\langle \langle T \hat{f}_\sigma^\dagger(\tau) \hat{f}_\sigma(\tau') \rangle \rangle_{C_n}$ suffices to compute any observable.

2.3 The negative sign problem

For each configuration of vertices C_n , we are able to compute arbitrary correlation functions. Due to the dimension of the configuration space, it is prohibitively expensive to carry out the summation over C_n exactly. One will thus opt for a stochastic Monte Carlo approach which, for completeness sake, is reviewed in Appendix A.

A prerequisite for applying the Monte Carlo approach is that the *weight*, $W(C_n)$, of Eq. (14) is positive. For quantum systems, a positive formulation is not always possible, and one will decide to sample $|W(C_n)|$. Thereby, we will carry out the Monte Carlo evaluation with

$$\langle \hat{O}(\tau) \rangle = \frac{\sum_{C_n} W(C_n) \langle \langle \hat{O}(\tau) \rangle \rangle_{C_n}}{\sum_{C_n} W(C_n)} = \frac{\frac{\sum_{C_n} |W(C_n)| \text{sign}(W(C_n)) \langle \langle \hat{O}(\tau) \rangle \rangle_{C_n}}{\sum_{C_n} |W(C_n)|}}{\frac{\sum_{C_n} |W(C_n)| \text{sign}(W(C_n))}{\sum_{C_n} |W(C_n)|}} \quad (30)$$

and separately compute the numerator and denominator. The denominator corresponds to the average sign. On general grounds, one can argue that it is given by the ratio of two partition functions and thereby decays exponentially with the inverse temperature and the number of impurities N_{imp}

$$\langle \text{sign} \rangle = \frac{\sum_{C_n} W(C_n)}{\sum_{C_n} |W(C_n)|} \simeq e^{-\Delta \beta N_{\text{imp}}}. \quad (31)$$

In other words, at low temperatures there is a next to perfect cancellation of positive and negative weights. The sign problem is a consequence of the law of large numbers which states that the error on the average sign scales as

$$\sigma_{\text{sign}} \sim \frac{1}{\sqrt{T_{\text{CPU}}}}, \quad (32)$$

where T_{CPU} corresponds to the computational time. Obviously, to obtain sensible results, we will require that

$$\frac{\sigma_{\text{Sign}}}{\langle \text{sign} \rangle} \ll 1 \quad (33)$$

such that

$$T_{\text{CPU}} \gg e^{2\Delta\beta N_{\text{imp}}} . \quad (34)$$

Hence the required CPU time will scale exponentially with inverse temperature and number of impurities. Note that one can counter the sign problem if one can define an improved estimator for the average sign such that the fluctuations are greatly suppressed! Let us furthermore note that the pre-factor Δ is formulation dependent. One can for instance mention recent work of Huffman *et al.* [24], who have found a CT-INT formulation for spin-polarized electron problems at half-band filling that is free of the sign problem. Hence in this case $\Delta = 0$. There are other problems which are free of the negative sign problem. To show this, for the special case of attractive and repulsive Hubbard interactions, we will consider the mapping of the CT-INT to the CT-AUX [26, 27, 34]. This mapping allows the use of results derived in the framework of the Hirsch-Fye [35] and auxiliary-field QMC (for a review see Ref. [4]) algorithms to argue for the absence of the sign problem.

Let us start with the repulsive Hubbard interaction, which we will write as

$$H_U = \frac{U}{2} \sum_{s=\pm 1} [(\hat{n}_{\uparrow} - 1/2) + s\delta] [(\hat{n}_{\uparrow} - 1/2) - s\delta] = -\frac{U(\delta^2 - 1/4)}{2} \sum_{s=\pm 1} e^{\alpha s \hat{m}} \quad (35)$$

with

$$\cosh(\alpha) - 1 = \frac{1}{2} \frac{1}{\delta^2 - 1/4}$$

and magnetization $\hat{m} = \hat{n}_{\uparrow} - \hat{n}_{\downarrow}$. This identity relies on the fact that $\hat{m}^4 = \hat{m}^2$ and requires $\delta > 1/2$. Using this identity, the weight of a vertex configuration is given by

$$W(C_n) = \left(\frac{U(\delta^2 - 1/4)}{2} \right)^n \frac{1}{n!} \langle T e^{\alpha s_1 \hat{m}(\tau_1)} \dots e^{\alpha s_n \hat{m}(\tau_n)} \rangle_0 \quad (36)$$

where $\langle \bullet \rangle_0$ corresponds to the thermal expectation value with respect to the non-interacting model. Several comments are in order. (i) The notation $K = U\beta(\delta^2 - 1/4)$ makes the mapping to the CT-AUX explicit [27]. (ii) For $\delta > 1/2$, $U(\delta^2 - 1/4)$ is positive for the repulsive case. (iii) For a given set of Ising fields, the thermal expectation value has precisely the same structure as in the Hirsch-Fye and auxiliary-field QMC algorithms [4]. It hence follows that the CT-INT, CT-AUX, Hirsch-Fye and auxiliary-field QMC algorithms have the same sign problem for repulsive Hubbard interactions. Thus, as shown in Ref. [36], the SIAM is free of the negative sign problem. By the same token, one can argue that a class of one-dimensional problems [6] and problems with particle-hole symmetry such as the Kane-Mele-Hubbard model [37, 13] are sign problem free if formulated within the CT-INT. Studies of correlation effects in one-dimensional helical liquids [15] hinge on this observation.

For attractive interactions, one can use a similar identity as above. In this case, we have to adopt a different convention for the δ -shift

$$H_U = \frac{U}{2} \sum_{s=\pm 1} [(\hat{n}_\uparrow - 1/2) + s\delta] [(\hat{n}_\uparrow - 1/2) + s\delta] = \frac{U(\delta^2 - 1/4)}{2} \sum_{s=\pm 1} e^{\alpha s(\hat{n}-1)} \quad (37)$$

with $\hat{n} = \hat{n}_\uparrow + \hat{n}_\downarrow$. Again, the above equation relies on the fact that $(\hat{n} - 1)^4 = (\hat{n} - 1)^2$ and the same equation as above holds for α . Thus, for the attractive case, the weight reads

$$W(C_n) = \left(-\frac{U(\delta^2 - 1/4)}{2} \right)^n \frac{1}{n!} \langle T e^{\alpha s_1(\hat{n}(\tau_1)-1)} \dots e^{\alpha s_n(\hat{n}(\tau_n)-1)} \rangle_0. \quad (38)$$

Since $U < 0$ and $\delta > 1/2$, $-U(\delta^2 - 1/4)$ is positive. Furthermore, provided that the non-interacting model factorizes into identical spin-up and spin-down real representable Hamiltonians, the thermal expectation value reads

$$\langle T e^{\alpha s_1(\hat{n}(\tau_1)-1)} \dots e^{\alpha s_n(\hat{n}(\tau_n)-1)} \rangle_0 = \left[\langle T e^{\alpha s_1(\hat{n}_\uparrow(\tau_1)-1/2)} \dots e^{\alpha s_n(\hat{n}_\uparrow(\tau_n)-1/2)} \rangle_{0,\uparrow} \right]^2, \quad (39)$$

which is manifestly positive. Here $\langle \bullet \rangle_{0,\uparrow}$ corresponds to the thermal expectation value in the spin-up sector. Factorization is not necessarily required for the absence of the sign problem in the presence of attractive interactions. In general, time-reversal-symmetric fermionic problems where time reversal symmetry is present for *every configuration* C_n , are free of the minus sign problem. This follows essentially from Kramers theorem and is proven in Ref. [38].

2.4 The Monte Carlo sampling

At this point, we will assume that the weight is positive such that we can carry out Monte Carlo importance sampling. Only two moves are required to guarantee ergodicity: addition and subtraction of vertices. Vertex addition corresponds to the proposal

$$C_n = \{[\tau_1, s_1] \dots [\tau_n, s_n]\} \rightarrow C_{n+1} = \{[\tau_1, s_1] \dots [\tau_i, s_i], [\tau', s'], [\tau_{i+1}, s_{i+1}] \dots [\tau_n, s_n]\} \quad (40)$$

where we add the vertex τ', s' at position i in the string. The proposal probability reads

$$T_{C_n \rightarrow C_{n+1}}^0 = \underbrace{\frac{1}{n+1}}_{\text{Position in string}} \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau'} \underbrace{\frac{1}{2}}_{\text{Value of } s'} . \quad (41)$$

Vertex removal corresponds to

$$C_n = \{[\tau_1, s_1] \dots [\tau_n, s_n]\} \rightarrow C_{n-1} = \{[\tau_1, s_1] \dots [\tau_i, s_i], [\tau_{i+2}, s_{i+2}], \dots [\tau_n, s_n]\}, \quad (42)$$

where vertex i has been removed. The probability to propose this move reads

$$T_{C_n \rightarrow C_{n-1}}^0 = 1/n, \quad (43)$$

which corresponds to the probability of choosing vertex i under the assumption that each vertex is equally probable. As shown in Appendix A (see Eq. (132)), the Metropolis acceptance reads

$$P_{C \rightarrow C'} = \min \left(\frac{T_{C' \rightarrow C}^0 W(C')}{T_{C \rightarrow C'}^0 W(C)}, 1 \right). \quad (44)$$

Thus

$$\begin{aligned} P_{C_n \rightarrow C_{n+1}} &= \min \left(-\frac{U\beta}{(n+1)} \frac{\prod_\sigma \det \mathbf{M}_\sigma(C_{n+1})}{\prod_\sigma \det \mathbf{M}_\sigma(C_n)}, 1 \right) \\ P_{C_{n+1} \rightarrow C_n} &= \min \left(-\frac{(n+1)}{U\beta} \frac{\prod_\sigma \det \mathbf{M}_\sigma(C_n)}{\prod_\sigma \det \mathbf{M}_\sigma(C_{n+1})}, 1 \right). \end{aligned}$$

Note that in our formulation the ordering of the vertices is important since we have defined the integration without time-ordering $\int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_n$ as opposed to a time-ordered formulation $\int_0^\beta d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{n-1}} d\tau_n$. The reader is encouraged to show that both formulations lead to the same acceptance/rejection ratios. In practical implementations one will also include a move that keeps the vertex number constant but flips the value of the Ising spin. Strictly speaking, this move is not necessary but has the potential of improving the autocorrelation time. For repulsive interactions, this statement follows from the notion that summing over the Ising fields will restore the broken $SU(2)$ spin-symmetry.

2.5 Fast updates

The Monte Carlo dynamics relies on the calculation of ratios of determinants. Such ratios can be computed using the determinant identities of Eq. (21). For instance, under vertex addition we will have to compute for each spin sector

$$\begin{aligned} \frac{\det \mathbf{M}_\sigma(C_{n+1})}{\det \mathbf{M}_\sigma(C_n)} &= \frac{\det \begin{pmatrix} & & g_0(\tau_1, \tau') \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ & & g_0(\tau_n, \tau') \\ g_0(\tau', \tau_1) & \dots g_0(\tau', \tau_n) & g_0(\tau', \tau') - \alpha_\sigma(s') \end{pmatrix}}{\det \mathbf{M}_\sigma(C_n)} \\ &= \frac{\det \begin{pmatrix} & & 0 \\ \mathbf{M}_\sigma(C_n) & & \vdots \\ 0 & \dots 0 & 1 \end{pmatrix} + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2}{\det \mathbf{M}_\sigma(C_n)} \\ &= (1 + \mathbf{v}_1 \cdot \mathbf{M}_\sigma^{-1}(C_n) \mathbf{u}_1) (1 + \mathbf{v}_2 \cdot \mathbf{M}_\sigma^{-1}(C_n) \mathbf{u}_2) - (\mathbf{v}_2 \cdot \mathbf{M}_\sigma^{-1}(C_n) \mathbf{u}_1) (\mathbf{v}_1 \cdot \mathbf{M}_\sigma^{-1}(C_n) \mathbf{u}_2). \end{aligned}$$

with

$$\mathbf{u}_1 = (g_0(\tau_1, \tau'), \dots, g_0(\tau_n, \tau'), g_0(\tau', \tau') - \alpha_\sigma(s') - 1), \quad (\mathbf{v}_1)_i = \delta_{i,n+1}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n+1}, \quad \mathbf{v}_2 = (g_0(\tau', \tau_1), \dots, g_0(\tau', \tau_n), 0).$$

Carrying out the calculation yields

$$\frac{\det \mathbf{M}_\sigma(C_{n+1})}{\det \mathbf{M}_\sigma(C_n)} = g_0(\tau', \tau') - \alpha_\sigma(s') - \sum_{i,j=1}^n g_0(\tau', \tau_i) [M_\sigma^{-1}(C_n)]_{i,j} g_0(\tau_j, \tau'). \quad (45)$$

Hence, provided that the matrix $\mathbf{M}_\sigma^{-1}(C_n)$ is known, computing the ratio involves n^2 operations. The vertex removal takes a very simple form. Assume that we remove the n^{th} vertex of the configuration C_n . Then for a given spin sector, we will have to compute

$$\begin{aligned} \frac{\det \mathbf{M}_\sigma(C_{n-1})}{\det \mathbf{M}_\sigma(C_n)} &= \frac{\det \begin{pmatrix} g_0(\tau_1, \tau_1) - \alpha_\sigma(s_1) & \dots & g_0(\tau_1, \tau_{n-1}) & 0 \\ \vdots & & \vdots & \vdots \\ g_0(\tau_{n-1}, \tau_1) & \dots & g_0(\tau_{n-1}, \tau_{n-1}) - \alpha_\sigma(s_{n-1}) & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}}{\det \mathbf{M}_\sigma(C_n)} \\ &= \frac{\det [\mathbf{M}_\sigma(C_n) + \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2]}{\det \mathbf{M}_\sigma(C_n)} \end{aligned} \quad (46)$$

with

$$\mathbf{u}_1 = -([M_\sigma(C_n)]_{1,n}, \dots, [M_\sigma(C_n)]_{n,n} - 1), \quad (\mathbf{v}_1)_i = \delta_{i,n}$$

and

$$(\mathbf{u}_2)_i = \delta_{i,n}, \quad \mathbf{v}_2 = -([M_\sigma(C_n)]_{n,1}, \dots, [M_\sigma(C_n)]_{n,n-1}, 0).$$

Evaluating the above gives

$$\frac{\det \mathbf{M}_\sigma(C_{n-1})}{\det \mathbf{M}_\sigma(C_n)} = [M_\sigma^{-1}(C_n)]_{n,n}. \quad (47)$$

Again, provided that we have the matrix $\mathbf{M}_\sigma^{-1}(C_n)$ at hand, the computational cost for computing the ratio for vertex removal is negligible.

Having computed the ratio of determinants, we can compute the acceptance probability, draw a pseudo-random number, and accept or reject the move. If accepted, we will have to upgrade the matrix $\mathbf{M}_\sigma^{-1}(C_n)$. This is readily done with the use of the Sherman-Morrison formula of Eq. (20) and involves n^2 operations.

In some cases, it is desirable to add more than one vertex at a time. For this purpose, it is more useful to use the Woodbury formula

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} \quad (48)$$

with $A \in \mathbb{C}^{n \times n}$, $U \in \mathbb{C}^{n \times k}$, $C \in \mathbb{C}^{k \times k}$ and $V \in \mathbb{C}^{k \times n}$. The Woodbury identity reduces to the Sherman-Morrison formula of Eq. (20) at $k = 1$ and $C = 1$. A discussion of block updates as well as a demonstration of various matrix identities can be found in Ref. [39].

2.6 Average expansion parameter

A crucial issue concerns the average expansion parameter $\langle n \rangle$ since it will determine the average size of the matrix $\mathbf{M}_\sigma(C_n)$. The computational effort to visit each vertex – a sweep – will then scale as $\langle n \rangle^3$. For a general interaction term \hat{H}_1 , the average expansion parameter is [6]

$$\begin{aligned} \langle n \rangle &= \frac{1}{Z} \sum_n \frac{(-1)^n n}{n!} \int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_n \langle T \hat{H}_1(\tau_1) \cdots \hat{H}_1(\tau_n) \rangle_0 \\ &= -\frac{1}{Z} \sum_m \frac{(-1)^m}{m!} \int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_m \int_0^\beta d\tau \langle T \hat{H}_1(\tau_1) \cdots \hat{H}_1(\tau_m) \hat{H}_1(\tau) \rangle_0 \\ &= - \int_0^\beta d\tau \langle \hat{H}_1(\tau) \rangle . \end{aligned} \quad (49)$$

For the Hubbard model, replacing \hat{H}_1 by the form of Eq. (7), we obtain

$$\langle n \rangle = -\beta U [\langle (\hat{n}_\uparrow - 1/2)(\hat{n}_\downarrow - 1/2) \rangle - \delta^2] , \quad (50)$$

where we have set $\rho = 1/2$. Thus, the computational time for a sweep scales as in the Hirsch-Fye approach, namely as $(\beta U)^3$. The algorithm can be used for lattice models with N correlated sites. In this case, the computational time for a sweep scales as $(N\beta U)^3$, which is more expensive than the auxiliary-field approach, which scales as $\beta U N^3$. As mentioned in the introduction, the advantage of the CT-INT method lies in the fact that it is action-based such that fermionic and bosonic baths can be easily implemented.

3 CT-HYB

In this section we will provide a very succinct overview of the basic formulation of the CT-HYB. For a detailed discussion of the algorithm, the reader is referred to the review article [1] and references therein.

3.1 The partition function

In contrast to the CT-INT, the CT-HYB carries out the expansion in the hybridization matrix $\Delta(\tau - \tau')$. The action of the SIAM is decomposed into local

$$S_{\text{loc}}(f^\dagger, f) = \int_0^\beta d\tau \sum_\sigma f_\sigma^\dagger(\tau) \left[\frac{\partial}{\partial \tau} + \epsilon_f \right] f_\sigma(\tau) + U \int_0^\beta d\tau f_\uparrow^\dagger(\tau) f_\uparrow(\tau) f_\downarrow^\dagger(\tau) f_\downarrow(\tau) \quad (51)$$

and hybridization

$$S_{\text{hyb}}(f^\dagger, f) = - \int_0^\beta d\tau d\tau' \sum_\sigma f_\sigma^\dagger(\tau) \Delta(\tau - \tau') f_\sigma(\tau') \quad (52)$$

parts. Taylor expanding in S_{hyb} gives

$$\begin{aligned} Z_{\text{SIAM}} = & \int D\{f^\dagger f\} e^{-S_{\text{loc}}(f^\dagger, f)} \sum_n \frac{1}{n!} \sum_{\sigma_1 \dots \sigma_n} \int_0^\beta d\tau_1 d\tau'_1 \dots d\tau_n d\tau'_n \\ & \times f_{\sigma_1}^\dagger(\tau_1) f_{\sigma_1}(\tau'_1) \dots f_{\sigma_n}^\dagger(\tau_n) f_{\sigma_n}(\tau'_n) \Delta(\tau_1 - \tau'_1) \dots \Delta(\tau_n - \tau'_n). \end{aligned} \quad (53)$$

The insight of Ref. [28] is to sum up a set of configurations corresponding to the permutations of the Grassmann variables f . As we will see, the weight of this sum of configurations is given by an $n \times n$ determinant of the hybridization function. This is the crucial step in the basic formulation of the algorithm that leads to the absence of a sign problem for the SIAM. To achieve this, it is convenient to introduce the notation

$$x = (\tau, \sigma), \quad \int dx = \sum_\sigma \int_0^\beta d\tau, \quad \Delta_{x,x'} = \Delta(\tau - \tau') \delta_{\sigma,\sigma'} \quad (54)$$

such that

$$\begin{aligned} & \sum_{\sigma_1 \dots \sigma_n} \int_0^\beta d\tau_1 d\tau'_1 \dots d\tau_n d\tau'_n f_{\sigma_1}^\dagger(\tau_1) f_{\sigma_1}(\tau'_1) \dots f_{\sigma_n}^\dagger(\tau_n) f_{\sigma_n}(\tau'_n) \Delta(\tau_1 - \tau'_1) \dots \Delta(\tau_n - \tau'_n) \\ &= \int dx_1 dx'_1 \dots dx_n dx'_n f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n} \Delta_{x_1, x'_1} \dots \Delta_{x_n, x'_n} \\ &= \frac{1}{n!} \sum_{P \in S_n} \int dx_1 dx'_{P(1)} \dots dx_n dx'_{P(n)} f_{x_1}^\dagger f_{x'_{P(1)}} \dots f_{x_n}^\dagger f_{x'_{P(n)}} \Delta_{x_1, x'_{P(1)}} \dots \Delta_{x_n, x'_{P(n)}}. \end{aligned} \quad (55)$$

In the above P is a permutation of n objects, and we have merely replicated the result $n!$ times. Using the anti-commuting property of the Grassmann algebra, one can show that

$$f_{x_1}^\dagger f_{x'_{P(1)}} \dots f_{x_n}^\dagger f_{x'_{P(n)}} = (-1)^P f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n}$$

where $(-1)^P$ is the sign of the permutation. Since $dx_1 dx'_{P(1)} \dots dx_n dx'_{P(n)} = dx_1 dx'_1 \dots dx_n dx'_n$ Eq. (55) transforms as

$$\begin{aligned} & \frac{1}{n!} \int dx_1 dx'_1 \dots dx_n dx'_n f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n} \sum_{P \in S_n} (-1)^P \Delta_{x_1, x'_{P(1)}} \dots \Delta_{x_n, x'_{P(n)}} \\ &= \frac{1}{n!} \int dx_1 dx'_1 \dots dx_n dx'_n f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n} \det \begin{pmatrix} \Delta_{x_1, x'_1} & \dots & \Delta_{x_1, x'_n} \\ \vdots & & \vdots \\ \Delta_{x_n, x'_1} & \dots & \Delta_{x_n, x'_n} \end{pmatrix} \end{aligned} \quad (56)$$

Using the above, the partition function is written as

$$\frac{Z_{\text{SIAM}}}{Z_{\text{loc}}} = \sum_n \frac{1}{n!^2} \int dx_1 dx'_1 \dots dx_n dx'_n \langle f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n} \rangle_{\text{loc}} \det \begin{pmatrix} \Delta_{x_1, x'_1} & \dots & \Delta_{x_1, x'_n} \\ \vdots & & \vdots \\ \Delta_{x_n, x'_1} & \dots & \Delta_{x_n, x'_n} \end{pmatrix} \quad (57)$$

where

$$Z_{\text{loc}} = \int D\{f^\dagger f\} e^{-S_{\text{loc}}(f^\dagger, f)}$$

and

$$\langle \bullet \rangle_{\text{loc}} = \frac{1}{Z_{\text{loc}}} \int D\{f^\dagger f\} e^{-S_{\text{loc}}(f^\dagger, f)} \bullet.$$

For the special case where the hybridization function is spin diagonal, one can simplify the above equation. In this case, the spin variables $\{\sigma'_1 \dots \sigma'_n\}$ have to be a permutation of $\{\sigma_1 \dots \sigma_n\}$ such that

$$\sigma'_i = \sigma_{P(i)} \text{ with } P \in \mathcal{S}_n. \quad (58)$$

Hence,

$$\begin{aligned} & \int dx_1 dx'_1 \dots dx_n dx'_n \langle f_{x_1}^\dagger f_{x'_1} \dots f_{x_n}^\dagger f_{x'_n} \rangle_{\text{loc}} \det \begin{pmatrix} \Delta_{x_1, x'_1} & \dots & \Delta_{x_1, x'_n} \\ \vdots & & \vdots \\ \Delta_{x_n, x'_1} & \dots & \Delta_{x_n, x'_n} \end{pmatrix} \\ &= \sum_{P \in \mathcal{S}_n} \sum_{\sigma_1 \dots \sigma_n} \int_0^\beta d\tau_1 d\tau'_1 \dots d\tau_n d\tau'_n \langle f_{\tau_1, \sigma_1}^\dagger f_{\tau'_1, \sigma_{P(1)}} \dots f_{\tau_n, \sigma_n}^\dagger f_{\tau'_n, \sigma_{P(n)}} \rangle_{\text{loc}} \\ & \quad \times \det \begin{pmatrix} \Delta_{(\tau_1, \sigma_1), (\tau'_1, \sigma_{P(1)})} & \dots & \Delta_{(\tau_1, \sigma_1), (\tau'_n, \sigma_{P(n)})} \\ \vdots & & \vdots \\ \Delta_{(\tau_n, \sigma_n), (\tau'_1, \sigma_{P(1)})} & \dots & \Delta_{(\tau_n, \sigma_n), (\tau'_n, \sigma_{P(n)})} \end{pmatrix}. \end{aligned}$$

For a given permutation P one can carry out the substitution $\tau'_i = \tau''_{P(i)}$. This substitution leaves the integration measure invariant such that the above reads

$$\begin{aligned} & \sum_{P \in \mathcal{S}_n} \sum_{\sigma_1 \dots \sigma_n} \int_0^\beta d\tau_1 d\tau''_1 \dots d\tau_n d\tau''_n \langle f_{\tau_1, \sigma_1}^\dagger f_{\tau''_{P(1)}, \sigma_{P(1)}} \dots f_{\tau_n, \sigma_n}^\dagger f_{\tau''_{P(n)}, \sigma_{P(n)}} \rangle_{\text{loc}} \\ & \quad \times \det \begin{pmatrix} \Delta_{(\tau_1, \sigma_1), (\tau''_{P(1)}, \sigma_{P(1)})} & \dots & \Delta_{(\tau_1, \sigma_1), (\tau''_{P(n)}, \sigma_{P(n)})} \\ \vdots & & \vdots \\ \Delta_{(\tau_n, \sigma_n), (\tau''_{P(1)}, \sigma_{P(1)})} & \dots & \Delta_{(\tau_n, \sigma_n), (\tau''_{P(n)}, \sigma_{P(n)})} \end{pmatrix}. \end{aligned}$$

One can get rid of the permutation under the integral by reordering the Grassmann variables $f_{\tau''_{P(i)}, \sigma_{P(i)}}$ as well as the columns of the matrix. In this process, the minus signs cancel and the sum over the permutations gives a factor $n!$. Hence, the partition function reads

$$\begin{aligned} \frac{Z_{\text{SIAM}}}{Z_{\text{loc}}} &= \sum_n \frac{1}{n!} \sum_{\sigma_1 \dots \sigma_n} \int_0^\beta d\tau_1 d\tau''_1 \dots d\tau_n d\tau''_n \langle f_{\tau_1, \sigma_1}^\dagger f_{\tau''_1, \sigma_1} \dots f_{\tau_n, \sigma_n}^\dagger f_{\tau''_n, \sigma_n} \rangle_{\text{loc}} \\ & \quad \times \det \begin{pmatrix} \Delta_{(\tau_1, \sigma_1), (\tau''_1, \sigma_1)} & \dots & \Delta_{(\tau_1, \sigma_1), (\tau''_n, \sigma_n)} \\ \vdots & & \vdots \\ \Delta_{(\tau_n, \sigma_n), (\tau''_1, \sigma_1)} & \dots & \Delta_{(\tau_n, \sigma_n), (\tau''_n, \sigma_n)} \end{pmatrix}. \quad (59) \end{aligned}$$

3.2 The Monte Carlo sampling and evaluation of the trace

For the SIAM where the hybridization matrix is spin-diagonal, we can define a configuration as

$$C_n = \{[\tau_1, \tau'_1, \sigma_1] \cdots [\tau_n, \tau'_n, \sigma_n]\} \quad (60)$$

and the weight of a configuration as

$$W(C_n) = \frac{1}{n!} \langle f_{\tau_1, \sigma_1}^\dagger f_{\tau''_1, \sigma_1} \cdots f_{\tau_n, \sigma_n}^\dagger f_{\tau''_n, \sigma_n} \rangle_{\text{loc}} \det \begin{pmatrix} \Delta_{(\tau_1, \sigma_1), (\tau''_1, \sigma_1)} & \cdots & \Delta_{(\tau_1, \sigma_1), (\tau''_n, \sigma_n)} \\ \vdots & & \vdots \\ \Delta_{(\tau_n, \sigma_n), (\tau''_1, \sigma_1)} & \cdots & \Delta_{(\tau_n, \sigma_n), (\tau''_n, \sigma_n)} \end{pmatrix} \quad (61)$$

such that $\frac{Z_{\text{SIAM}}}{Z_{\text{loc}}} = \sum_{C_n} W(C_n)$. The simplest possible proposal matrix to add a vertex reads,

$$T_{C_n \rightarrow C_{n+1}}^0 = \underbrace{\frac{1}{n+1}}_{\text{Position in string}} \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau} \underbrace{\frac{1}{\beta}}_{\text{Value of } \tau'} \underbrace{\frac{1}{2}}_{\text{Spin up or spin down}}. \quad (62)$$

To remove a vertex

$$T_{C_n \rightarrow C_{n-1}}^0 = \frac{1}{n}. \quad (63)$$

With the above we are now in a position to compute the Metropolis acceptance ratio given in Eq. (132). The update of the hybridization matrix follows the same ideas as in the CT-INT.

The computationally expensive part of the CT-HYB algorithm is the evaluation of the trace $\langle \cdot \rangle_{\text{loc}}$. In general, as the expectation value is taken with respect to the local Hamiltonian, the operators can be represented by matrices in a basis of the local Hilbert space.

One particular basis that is often used is the eigenbasis of the local Hamiltonian. This makes the time evolution of the operators trivial but leads to dense operator matrices, which have to be multiplied. Here the main drawback of the CT-HYB becomes apparent: the size of the local Hilbert space D , and thus of the operator matrices, grows exponentially with the number of local degrees of freedom. Thus, simulation of larger systems become unfeasible quite rapidly. Nevertheless, there are strategies to simulate multi-orbital systems as well as small clusters for Cluster-DMFT applications.

To reduce the size of the matrices that have to be multiplied, one can exploit certain symmetries of the Hamiltonian. This allows splitting the matrices into blocks of size $d_i \ll D$ [40], such that $\sum_i d_i = D$. Each block contains all the states associated with a certain value of a conserved quantity. The most obvious conserved quantities are the total f -particle number N_f and the z -component of the total spin S_z . There are well-defined rules governing how an operator connects the different blocks corresponding to different quantum numbers. As an example, the creation operator \hat{f}_\uparrow^\dagger connects the blocks (N_f, S_z) and $(N_f + 1, S_z + 1/2)$.

In Ref. [41], another set of conserved quantities was identified, which leads to even smaller matrix blocks. For each orbital or site, the projection onto single occupation

$$PS_a = (n_{a,\uparrow} - n_{a,\downarrow})^2, \quad a = 1 \dots M \quad (64)$$

| Considered symmetries | None | N_f | N_f, S_z | N_f, S_z, PS |
|-----------------------|--------------|-------|------------|----------------|
| $N = 3$ | $64 = 4^3$ | 20 | 9 | 3 |
| $N = 5$ | $1024 = 4^4$ | 252 | 100 | 10 |

Table 1: Largest matrix block size, with different sets of conserved quantities taken into account.

commutes with the local Hamiltonian. This makes the whole sequence of $PS = (PS_1, \dots, PS_M)$ a good quantum number, so that the Hamiltonian can be written in a block-diagonal form with respect to PS . The resulting reduction of the matrix block sizes is shown in table 1.

Besides the matrix code described above, there are other possibilities for calculating the local trace. An algorithm based on a Krylov-subspace method was brought forward in Ref. [42]. Here the particle-number basis is used, so that applying operators to states is trivial. However, now the time evolution becomes more involved. Another recent proposal uses matrix product states for the propagation in time [43].

3.3 Selected applications

The CT-HYB is clearly the method of choice to tackle complex impurity problems at strong coupling and with the full Coulomb repulsion. In the context of the DMFT, extracting the local Green function is essential. In this context it is important to point out the work of Hafermann *et al.* [44], which describes an improved estimator for computing the self-energy.

In a two-orbital model, the inclusion of the Hund's coupling J greatly influences the critical Hubbard interaction U_c for the Mott transition [45]. A three-orbital model studied in Ref. [46] exhibits not only the Mott phase at integer fillings, but also a non-Fermi-liquid frozen-moment phase. In Ref. [47], models with only d -orbitals and with d - and additional oxygen p -orbitals were compared. With the p -orbitals included, the filling of the d -orbitals changes significantly, which thereby leads to a very different low-energy behavior.

Another recent application of the CT-HYB method in the context of DMFT is the study of models for topological Kondo insulators conducted in Ref. [48]. It was found that, starting from the non-interacting case, switching on the Hubbard interaction can drive the system through a series of transitions. In particular, a transition was observed between different topological states that are distinct due to the point-group symmetry of the lattice considered. Meanwhile, in Ref. [49] it was found that the edge spectrum of topological Kondo insulators is governed by the same scale as the bulk heavy-fermion state, namely the coherence scale T_{coh} . This makes it possible to infer information about the bulk coherence from the topological properties of the system.

4 Application of CT-INT to the Hubbard-Holstein model

The CT-INT allows for a very simple and efficient inclusion of phonon degrees of freedom. The path we follow here is to integrate out the phonons in favor of a retarded interaction and then solve the purely electronic model with the CT-INT approach. Starting from the Hubbard-Holstein model with Einstein phonons we show how to integrate out the phonons, describe some details of the algorithm, and then present results for the crossover from adiabatic to anti-adiabatic phonons in the one-dimensional Holstein model.

4.1 Integrating-out the phonons

The Hubbard-Holstein Hamiltonian we consider reads

$$\hat{H} = -\sum_{i,j,\sigma} t_{i,j} \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + U \sum_i (\hat{n}_{i,\uparrow} - 1/2) (\hat{n}_{i,\downarrow} - 1/2) + g \sum_i \hat{Q}_i (\hat{n}_i - 1) + \sum_i \left(\frac{\hat{P}_i^2}{2M} + \frac{k}{2} \hat{Q}_i^2 \right) \quad (65)$$

Here, $\hat{n}_i = \sum_\sigma \hat{n}_{i,\sigma}$ and the last two terms correspond respectively to the electron-phonon coupling, g , and the phonon energy. The Hamiltonian is written such that for a particle-hole symmetric band, half-filling corresponds to chemical potential $\mu = 0$. Opting for fermion coherent states

$$\hat{c}_{i,\sigma} |c\rangle = c_{i,\sigma} |c\rangle , \quad (66)$$

$c_{i,\sigma}$ being a Grassmann variable, and a real space representation for the phonon coordinates

$$\hat{Q}_i |q\rangle = q_i |q\rangle , \quad (67)$$

the path integral formulation of the partition function reads

$$Z = \int D\{q\} D\{c^\dagger c\} e^{-(S_U + S_{ep})} , \quad (68)$$

with

$$\begin{aligned} S_U &= \int_0^\beta d\tau \sum_{i,j,\sigma} c_{i,\sigma}^\dagger(\tau) \left(\delta_{i,j} \frac{\partial}{\partial \tau} - t_{i,j} \right) c_{j,\sigma}(\tau) + U \sum_i (n_{i,\uparrow}(\tau) - 1/2) (n_{i,\downarrow}(\tau) - 1/2) \\ S_{ep} &= \int_0^\beta d\tau \sum_i \left(\frac{M \dot{q}_i^2(\tau)}{2} + \frac{k}{2} q_i^2(\tau) + g q_i(\tau) (n_i(\tau) - 1) \right) . \end{aligned} \quad (69)$$

In Fourier space,

$$q_j(\tau) = \frac{1}{\sqrt{\beta N}} \sum_{k,\Omega_m} e^{-i(\Omega_m \tau - kj)} q_{k,m} , \quad (70)$$

where Ω_m is a bosonic Matsubara frequency, the electron-phonon part of the action reads

$$\begin{aligned} S_{ep} &= \sum_{\Omega_m, k} \frac{M}{2} (\Omega_m^2 + \omega_0^2) q_{k,m}^\dagger q_{k,m} + g q_{k,m} \rho_{k,m}^\dagger , \\ \rho_{k,m}^\dagger &= \frac{1}{\sqrt{\beta N}} \int d\tau \sum_j e^{-i(\Omega_m \tau - kj)} (n_j(\tau) - 1) . \end{aligned} \quad (71)$$

Gaussian integration over the phonon degrees of freedom leads to a retarded density-density interaction

$$\int D\{\mathbf{q}\} e^{-S_{\text{ep}}} = e^{\int_0^\beta d\tau \int_0^\beta d\tau' \sum_{i,j} [n_i(\tau) - 1] D^0(i-j, \tau - \tau') [n_j(\tau') - 1]} . \quad (72)$$

For Einstein phonons the phonon propagator is diagonal in real space,

$$\begin{aligned} D^0(i - j, \tau - \tau') &= \delta_{i,j} \frac{g^2}{2k} P(\tau - \tau') \text{ with} \\ P(\tau) &= \frac{\omega_0}{2(1 - e^{-\beta\omega_0})} (e^{-|\tau|\omega_0} + e^{-(\beta - |\tau|)\omega_0}) . \end{aligned} \quad (73)$$

Hence the partition function of the Hubbard-Holstein model takes the form

$$Z = \int D\{c^\dagger c\} e^{-(S_U - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_{i,j} [n_i(\tau) - 1] D^0(i-j, \tau - \tau') [n_j(\tau') - 1])} .$$

In the anti-adiabatic limit, $\lim_{\omega_0 \rightarrow \infty} P(\tau) = \delta(\tau)$ such that the phonon interaction maps onto an attractive Hubbard interaction of magnitude g^2/k . We are now in a position to apply the CT-INT algorithm by expanding in both the retarded and the Hubbard interactions.

4.2 Formulation of CT-INT for the Hubbard-Holstein model

To avoid the minus-sign problem at least for one-dimensional chains with nearest neighbor hopping matrix element t , we rewrite the phonon retarded interaction as

$$H_P(\tau) = -\frac{g^2}{4k} \int_0^\beta d\tau' \sum_{i,\sigma,\sigma'} \sum_{s=\pm 1} P(\tau - \tau') [n_{i,\sigma}(\tau) - \alpha_+(s)] [n_{i,\sigma'}(\tau') - \alpha_+(s)] . \quad (74)$$

For each phonon vertex, we have introduced an Ising variable s . Summation over this Ising field reproduces, up to a constant, the original interaction. Since the phonon term is attractive, the adequate choice of signs is $\alpha_+(s) \equiv 1/2 + s\delta$, irrespective of the spin σ and σ' . Following Eq. (7), we rewrite the Hubbard term as

$$H_U(\tau) = \frac{U}{2} \sum_{i,s} \prod_{\sigma} (n_{i,\sigma}(\tau) - \alpha_{\sigma}(s)) . \quad (75)$$

To proceed with a description of the implementation of the algorithm, it is useful to define a general vertex

$$V(\tau) = \{i, \tau, \sigma, \tau', \sigma', s, b\} , \quad (76)$$

where b defines the type of vertex at hand, Hubbard ($b = 0$) or phonon ($b = 1$). For this vertex, we define a sum over the available phase space

$$\sum_{V(\tau)} = \sum_{i,\sigma,\sigma',s,b} \int_0^\beta d\tau' , \quad (77)$$

a weight

$$w[V(\tau)] = \delta_{b,0} \frac{U}{2} - \delta_{b,1} P(\tau - \tau') \frac{g^2}{4k}, \quad (78)$$

as well as

$$\begin{aligned} H[V(\tau)] = & \delta_{b,0} \delta_{\sigma,\uparrow} \delta_{\sigma',\downarrow} \delta(\tau - \tau') [n_{i,\uparrow}(\tau) - \alpha_+(s)] [n_{i,\downarrow}(\tau) - \alpha_-(s)] \\ & + \delta_{b,1} [n_{i,\sigma}(\tau) - \alpha_+(s)] [n_{i,\sigma'}(\tau') - \alpha_+(s)]. \end{aligned} \quad (79)$$

With the above definitions, the partition function can now be written as

$$\frac{Z}{Z_0} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^\beta d\tau_1 \sum_{V_1(\tau_1)} w[V_1(\tau_1)] \cdots \int_0^\beta d\tau_n \sum_{V_n(\tau_n)} w[V_n(\tau_n)] \langle T \hat{H}[V_1(\tau_1)] \cdots \hat{H}[V_n(\tau_n)] \rangle_0. \quad (80)$$

As for the Hubbard model, a configuration consists of a set of vertices $C_n = \{V_1(\tau_1), \dots, V_n(\tau_n)\}$. With the short-hand notation

$$\sum_{C_n} = \sum_n \int_0^\beta d\tau_1 \sum_{V(\tau_1)} \cdots \int_0^\beta d\tau_n \sum_{V(\tau_n)} \quad (81)$$

and

$$W(C_n) = \frac{(-1)^n}{n!} w[V_1(\tau_1)] \cdots w[V_n(\tau_n)] \langle T \hat{H}[V_1(\tau_1)] \cdots \hat{H}[V_n(\tau_n)] \rangle_0 \quad (82)$$

the partition function takes a form amenable to Monte Carlo sampling

$$\frac{Z}{Z_0} = \sum_{C_n} W(C_n). \quad (83)$$

The Monte Carlo sampling follows precisely the scheme presented in Sec. 2.4, namely the addition and removal of vertices. To be more specific, we consider the following form for vertex addition

$$T_{C_n \rightarrow C_{n+1}}^0 = P_U \frac{1}{n+1} \frac{1}{L} \frac{1}{\beta} \frac{1}{2} + (1 - P_U) \frac{1}{n+1} \frac{1}{L} \frac{1}{\beta} P(\Delta\tau) \left(\frac{1}{2}\right)^3. \quad (84)$$

The first term corresponds to the addition of a Hubbard vertex and has the same form as in Eq. (41). Note the additional factor $1/L$, which corresponds to the choice of the lattice site. This move is carried out with probability P_U . The second term corresponds to the addition of the phonon vertex and is carried out with probability $1 - P_U$. The factor $P(\Delta\tau)$ allows a direct sampling of the phonon propagator. In particular, for a randomly chosen value of τ , $\tau' = \tau + \Delta\tau$. The factor $(1/2)^3$ accounts for the choice of the Ising field, as well as for the choice of the spin variables σ, σ' entering into the density-density two-body term. The vertex removal retains the same form as for the SIAM

$$T_{C_n \rightarrow C_{n-1}}^0 = \frac{1}{n}. \quad (85)$$

4.3 The quarter-filled Holstein model from adiabatic to anti-adiabatic phonons

The concept of pre-formed fermion pairs or bosonic degrees of freedom that condense to form a superfluid can be found in many domains of correlated quantum many-body systems. Examples include the resonating valence bond theory of high-temperature superconductivity [50], Mott metal-insulator transitions in cold atoms [51], or transitions between charge-density-wave and superconducting states in the family of dichalcogenides [52]. As an interesting application of the CT-INT to the Holstein model, we show that these concepts can be carried over to one dimension where, in the absence of continuous symmetry breaking, the phase transition is replaced by a crossover.

In the anti-adiabatic limit $\omega_0 \rightarrow \infty$, $P(\tau)$ in equation (73) reduces to a Dirac δ -function, facilitating the above-mentioned mapping of the Holstein model onto the attractive Hubbard model, with $U = g^2/k$. The ratio of this binding energy and the bandwidth $W = 4t$ gives the dimensionless electron-phonon coupling

$$\lambda = \frac{g^2}{kW}. \quad (86)$$

In this section, we set $\lambda = 0.35$ and concentrate on the quarter-filled band with $k_F = \pi/4$.

Figure 1 shows equal-time correlation functions

$$C_O(q) = \sum_r e^{iqr} \left(\langle \hat{O}_r^\dagger \hat{O}_0 \rangle - \langle \hat{O}_r^\dagger \rangle \langle \hat{O}_0 \rangle \right), \quad (87)$$

for charge, $\hat{O}_r = \hat{n}_{r,\uparrow} + \hat{n}_{r,\downarrow}$; spin, $\hat{O}_r = \hat{n}_{r,\uparrow} - \hat{n}_{r,\downarrow}$; pairing, $\hat{O}_r = \hat{c}_{r,\uparrow}^\dagger \hat{c}_{r,\downarrow}^\dagger$; and single-particle, $\hat{O}_r = \hat{c}_{r,\uparrow} + \hat{c}_{r,\downarrow}$, degrees of freedom at various phonon frequencies. In the adiabatic limit $\omega_0/t = 0$, any $\lambda > 0$ leads to an insulating state with $2k_F$ long-range charge order corresponding to the Peierls instability. We choose the coupling strength $\lambda = 0.35$ such that we have a metallic Luther-Emery liquid with dominant $2k_F$ charge correlations for low phonon frequencies and then study the evolution as a function of increasing ω_0/t . In particular, we have verified that for $\omega_0/t = 0.1$, the lowest nonzero frequency considered in the following, there is no long-range order; this can be seen from the finite-size dependence of the charge susceptibility [53].

The density (or charge) structure factor is plotted in Fig. 1(a). For classical phonons ($\omega_0 = 0$) the Peierls instability leads to long-range $2k_F$ charge order at zero temperature. As discussed above, quantum lattice fluctuations (occurring for $\omega_0 > 0$) can melt this order and lead to a state with dominant but power-law $2k_F$ charge correlations [8], as confirmed by the cusp at $2k_F = \pi/2$ in Fig. 1(a). The magnitude of the peak at $q = 2k_F$ initially decreases and then saturates upon increasing the phonon frequency, signalling competing ordering mechanisms as well as enhanced lattice fluctuations. The linear form of the charge structure factor at long wavelengths [see figure 1(a)] indicates a $1/r^2$ power-law decay of the real-space charge correlations and hence a metallic state.

In Fig. 1(b), we present the pair correlation function in the onsite s -wave channel. In contrast to the density correlator, which picks up diagonal order, $P(r)$ detects off-diagonal order charac-

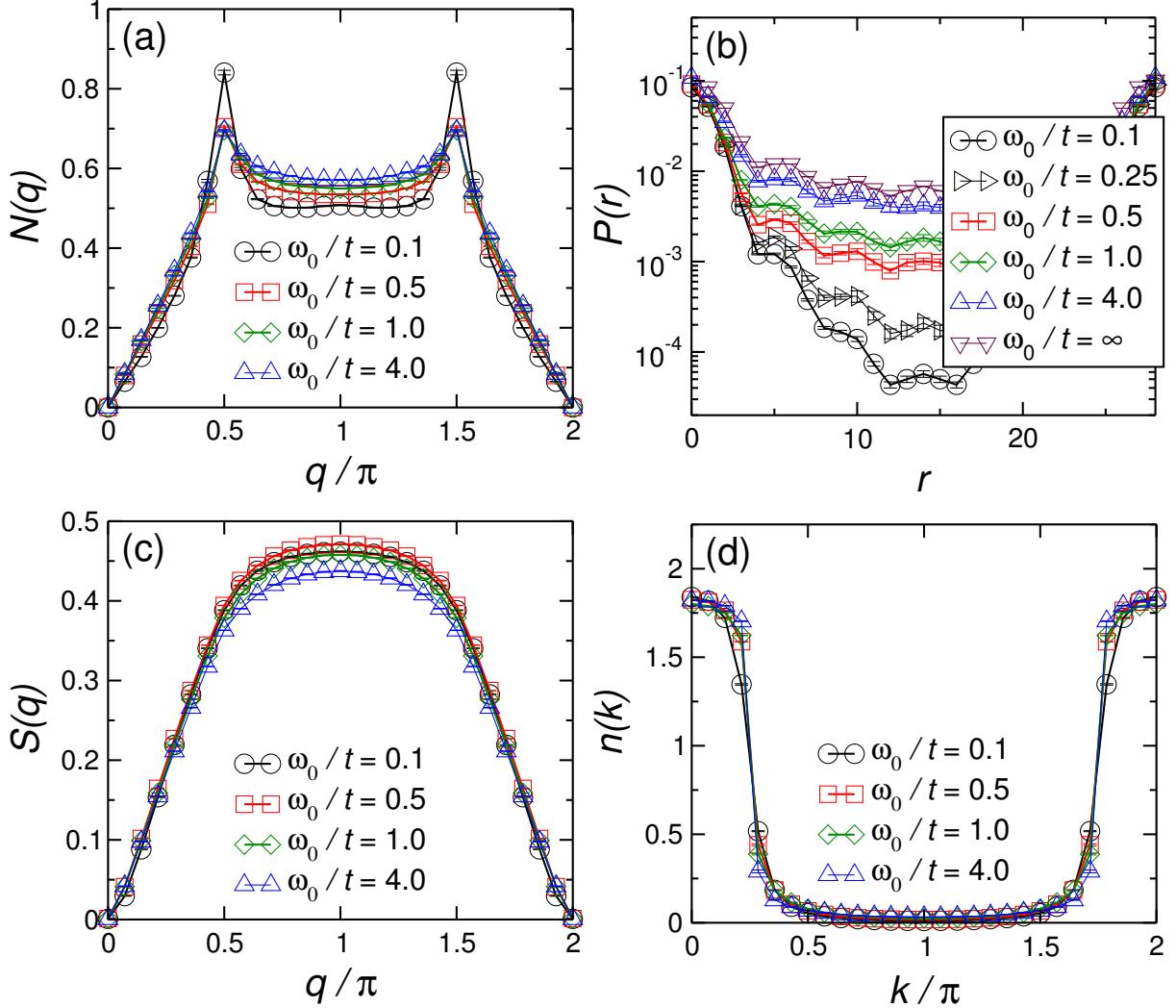


Fig. 1: Static correlation functions for different values of the phonon frequency ω_0/t at $\lambda = 0.35$ for a quarter-filled band ($n = 0.5$). The panels show (a) the charge structure factor, (b) the pairing correlator, (c) the spin structure factor, and (d) the momentum distribution function. Here $L = 28$ and $\beta t = 40$.

teristic of a superconducting state. In the Peierls state obtained for classical phonons, diagonal long-range charge order leads to an exponential decay of pairing correlations at long distances. The fluctuations resulting from a finite phonon frequency close the charge gap and render the pairing correlations critical. Comparing figures 1(a) and 1(b), we see that the suppression of the $2k_F$ charge correlations is accompanied by an increase of the pairing correlations, especially at large distances. A possible interpretation is that with increasing phonon frequency, the trapping of bipolarons in the $2k_F$ lattice modulation gives way to a “condensation” (in the usual sense of superfluidity in one dimension) of those preformed pairs. A detailed study of the dynamics of the observed crossover can be found in Ref. [54].

5 Concluding remarks

The aim of this lecture was to give a detailed theoretical overview of the CT-INT and to comment on the general formulation of the CT-HYB. For the SIAM, the CT-QMC methods are easy to implement and the interested reader is encouraged to try. It is essential to point out once again that the CT-QMC methods are action based such that the pool of potential applications is extremely large.

Acknowledgments

I would like to thank T. Lang, D. Luitz, M. Aulbach, F. Goth, M. Hohenadler, J. Werner, M. Weber and J. Hofmann who have implemented and generalized CT-QMC methods thus triggering many questions and discussions. Support of the Deutsche Forschungsgemeinschaft through FOR 1346 is gratefully acknowledged.

A Basic principles of Monte Carlo methods in statistical mechanics

In this section, we show how to use the Monte Carlo method to compute integrals of the form

$$\langle O \rangle_P = \int_{\Omega} dx P(x) O(x). \quad (88)$$

Ω is a discrete or continuous configuration space with elements x , $P(x)$ is a probability distribution on this space,

$$\int_{\Omega} dx P(x) = 1 \text{ and } P(x) \geq 0 \quad \forall x \in \Omega, \quad (89)$$

and O is an observable or random variable.

To illustrate why stochastic methods are useful, let us assume Ω to be a subspace of \mathbb{R}^d with volume $V = L^d$. One can break up Ω into hypercubes of linear dimension h and approximate the integral by a Riemann sum. The required number of function evaluations N then scales as $V/h^d = (L/h)^d = e^{d \ln(L/h)}$. Hence, the numerical effort – which is nothing but the number of function evaluations – grows exponentially with the dimension d . In contrast, stochastic methods provide an estimate of the integral with statistical uncertainty scaling as the inverse square root of the number of function evaluations, irrespective of the dimensionality d . Hence, in the large- d limit, stochastic methods become attractive. This result stems from the central limit theorem.

A.1 The central limit theorem

Before proceeding in illustrating and proving the central limit theorem, let us introduce some notation. We will denote by $P_O(\mathcal{O})$ the probability that the observable O takes the value \mathcal{O} . Hence in terms of integrals over the configuration space,

$$P_O(\mathcal{O}) = \int_{\Omega} dx P(x) \delta(O(x) - \mathcal{O}) \quad (90)$$

such that

$$\langle O \rangle_p = \int d\mathcal{O} P_O(\mathcal{O}) \mathcal{O}, \text{ and } \langle O^2 \rangle_p = \int d\mathcal{O} P_O(\mathcal{O}) \mathcal{O}^2. \quad (91)$$

Suppose that we have a set of configurations $\{x_i \mid i \in 1 \dots N\}$ distributed according to the probability distribution $P(x)$, then we can approximate $\langle O \rangle_P$ by

$$\langle O \rangle_P \sim \frac{1}{N} \sum_{i=1}^N \underbrace{O(x_i)}_{=\mathcal{O}_i} = X. \quad (92)$$

Clearly, X will depend upon the chosen set $\{x_i \mid i \in 1 \dots N\}$. Hence the relevant question is the distribution of X , $\mathcal{P}(X)$, for a given value of N . The central limit theorem tells us that

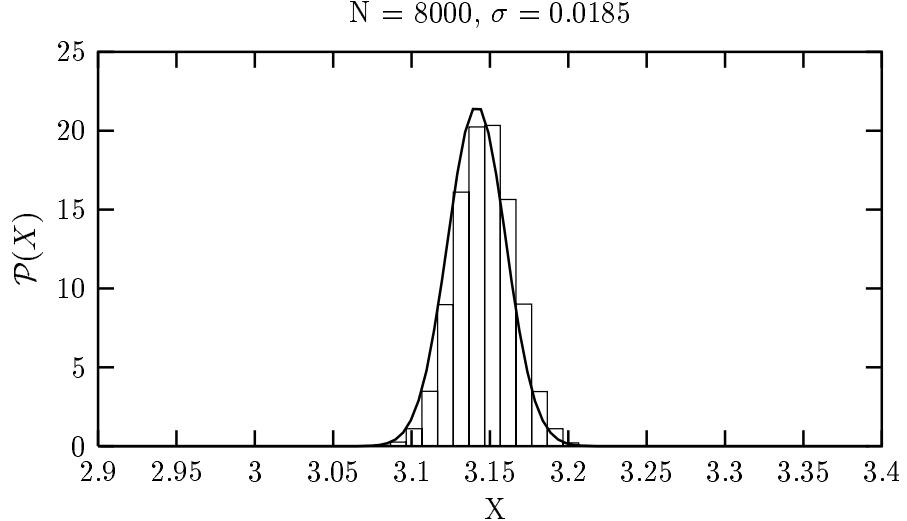


Fig. 2: Boxes correspond to the distribution results obtained after 10000 simulations. For each simulation we draw $N = 8000$ points. For a single simulation, we obtain $\sigma = 0.0185$. The heavy line corresponds to the result of central limit theorem with above value of σ .

provided that the $\mathcal{O}_1 \cdots \mathcal{O}_N$ are statistically independent and that N is large enough $\mathcal{P}(X)$ reads

$$\begin{aligned} \mathcal{P}(X) &= \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(X - \langle O \rangle_P)^2}{2\sigma^2} \right] \\ \text{with } \sigma^2 &= \frac{1}{N} (\langle O^2 \rangle_P - \langle O \rangle_P^2). \end{aligned} \quad (93)$$

Thus irrespective of the dimension d , the convergence to the exact result scales as $1/\sqrt{N}$, and the width of the above normal distribution, σ , corresponds to the statistical error. For practical purposes, one estimates σ by

$$\sigma^2 \approx \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^N O(x_i)^2 - \left(\frac{1}{N} \sum_{i=1}^N O(x_i) \right)^2 \right). \quad (94)$$

More general methods for estimating the error are discussed in section A.2.

Before demonstrating the central limit theorem we give a simple example, the evaluation of the number π obtained with

$$\pi = 4 \int_0^1 dx \int_0^1 dy \Theta(1 - x^2 - y^2). \quad (95)$$

Here Θ is the Heaviside function, $\Theta(x) = 1$ for $x > 0$ and vanishes otherwise. In this example we have $P(x, y) \equiv 1$. To generate a sequence of N points $(x, y)_i$ from this probability distribution, we draw random numbers, x, y , in the interval $[0, 1]$. For $N = 8000$ we obtain an error $\sigma = 0.0185$ with the use of Eq. (94). To check the central limit theorem, we repeat the simulation 10000 times with different random numbers. Fig. (2) shows the thus obtained distribution which compares well to the result of the central limit theorem.

We now demonstrate the central limit theorem.

Proof: The probability of obtaining the result X reads

$$\mathcal{P}(X) = \int d\mathcal{O}_1 \cdots d\mathcal{O}_N P_O(\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_N) \delta \left(X - \frac{1}{N} \sum_{i=1}^N \mathcal{O}_i \right). \quad (96)$$

The assumption that the \mathcal{O}_i are statistically independent means that the combined probability factorizes

$$P_O(\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_N) = P_O(\mathcal{O}_1)P_O(\mathcal{O}_2) \cdots P_O(\mathcal{O}_N).$$

Furthermore, using the representation $\delta(x) = \frac{1}{2\pi} \int d\lambda e^{i\lambda x}$ of the Dirac- δ function, we can reduce the above expression of $\mathcal{P}(X)$ to

$$\begin{aligned} \mathcal{P}(X) &= \frac{1}{2\pi} \int d\lambda d\mathcal{O}_1 \cdots d\mathcal{O}_N P_O(\mathcal{O}_1) \cdots P_O(\mathcal{O}_N) e^{i\lambda(X - \frac{1}{N} \sum_i \mathcal{O}_i)} \\ &= \frac{1}{2\pi} \int d\lambda e^{i\lambda X} \left(\int d\mathcal{O} P_O(\mathcal{O}) e^{-i\frac{\lambda}{N}\mathcal{O}} \right)^N \\ &= \frac{N}{2\pi} \int d\lambda e^{-NS(\lambda, X)} \text{ with } S(\lambda, X) = -i\lambda X - \ln \int d\mathcal{O} P_O(\mathcal{O}) e^{-i\lambda\mathcal{O}} \end{aligned} \quad (97)$$

As $N \rightarrow \infty$ the saddle point approximation becomes exact

$$\begin{aligned} \mathcal{P}(X) &\simeq \frac{N}{2\pi} \int d\lambda e^{-N \left(S(\lambda^*(X), X) + \frac{(\lambda^*(X) - \lambda)^2}{2} \frac{\partial^2}{\partial \lambda^2} S(\lambda, X) \Big|_{\lambda=\lambda^*(X)} \right)} \\ &= \sqrt{\frac{N}{2\pi \frac{\partial^2}{\partial \lambda^2} S(\lambda, X) \Big|_{\lambda=\lambda^*(X)}}} e^{-NS(\lambda^*(X), X)} \text{ with } \frac{\partial}{\partial \lambda} S(\lambda, X) \Big|_{\lambda=\lambda^*(X)} = 0 \end{aligned} \quad (98)$$

such that $\lambda^*(X)$ is determined by:

$$X = \frac{\int d\mathcal{O} P_O(\mathcal{O}) \mathcal{O} e^{-i\lambda^*(X)\mathcal{O}}}{\int d\mathcal{O} P_O(\mathcal{O}) e^{-i\lambda^*(X)\mathcal{O}}}. \quad (99)$$

To proceed, we again use that for large values of N we can expand around the saddle point

$$\begin{aligned} \frac{d}{dX} S(\lambda^*(X), X) \Big|_{X=X^*} &= \left[\underbrace{\frac{\partial}{\partial \lambda} S(\lambda, X) \Big|_{\lambda=\lambda^*(X)}}_{=0} \frac{d}{dX} \lambda^*(X) + \frac{\partial}{\partial X} S(\lambda^*(X), X) \right]_{X=X^*} \\ &= -i\lambda^*(X^*) = 0. \end{aligned} \quad (100)$$

Hence with Eq. (99) we have $X^* = \langle O \rangle_P$. Expanding around X^* (note that $S(\lambda^*(X^*), X^*) = 0$) yields

$$\mathcal{P}(X) = \sqrt{\frac{N}{2\pi \frac{\partial^2}{\partial \lambda^2} S(\lambda, X^*) \Big|_{\lambda=\lambda^*(X^*)}}} e^{-N \frac{(X - \langle O \rangle_P)^2}{2} \frac{d^2}{dX^2} S(\lambda^*(X), X) \Big|_{X=X^*}}. \quad (101)$$

Using Eqn. (100) and (99) we obtain

$$\begin{aligned} \frac{d^2}{dX^2} S(\lambda^*(X), X) \Big|_{X=X^*} &= -i \frac{d\lambda^*(X)}{dX} \Big|_{X=X^*} = \frac{1}{\langle O^2 \rangle_P - \langle O \rangle_P^2} \\ \frac{\partial^2}{\partial \lambda^2} S(\lambda, X^*) \Big|_{\lambda=\lambda^*(X^*)} &= \langle O^2 \rangle_P - \langle O \rangle_P^2 \end{aligned} \quad (102)$$

such that

$$\mathcal{P}(X) \simeq \sqrt{\frac{N}{2\pi(\langle O^2 \rangle_P - \langle O \rangle_P^2)}} e^{-N \frac{(X - \langle O \rangle_P)^2}{\langle O^2 \rangle_P - \langle O \rangle_P^2}}. \quad (103)$$

This completes the demonstration of the central limit theorem. The two important conditions for the validity of the theorem are that (i) the $\mathcal{O}_1 \cdots \mathcal{O}_N$ are statistically independent and that (ii) N is large.

A.2 Jackknife and bootstrap methods for error evaluation

The jackknife and bootstrap methods [55] provide alternative ways of estimating the error (94). These methods become particularly useful, if not essential, when one wishes to estimate the error on $f(\langle O_1 \rangle, \dots, \langle O_n \rangle)$, where f is an arbitrary function of n variables. For a given sequence of configurations $\{x_1 \cdots x_N\}$ drawn from the probability distribution $P(x)$, the jackknife focuses on the samples that leave out one configuration at a time

$$f_i^J = f \left(\frac{1}{N-1} \sum_{j \neq i} O_1(x_j), \dots, \frac{1}{N-1} \sum_{j \neq i} O_n(x_j) \right). \quad (104)$$

The error estimate on f is then given by

$$(\sigma_f^J)^2 \approx N \left(\frac{1}{N} \sum_{i=1}^N (f_i^J)^2 - \left(\frac{1}{N} \sum_{i=1}^N f_i^J \right)^2 \right). \quad (105)$$

One may verify explicitly that for $n = 1$ and $f(x) = x$ Eq. (105) reduces to Eq. (94) up to a factor $(N/(N-1))^2$, which tends to unity in the large N limit.

An alternative method for determining errors of f is the bootstrap algorithm. For a given sample of N configurations $\{x_1 \cdots x_N\}$ drawn from the probability distribution $P(x)$, we can construct N^N sets of N configurations, $\{x_{i_1} \cdots x_{i_N}\}$ with $i_1 \in 1 \cdots N$, $i_2 \in 1 \cdots N$, \dots , $i_N \in 1 \cdots N$, which correspond to the ideal bootstrap samples. For a given bootstrap sample, defined by the vector $\mathbf{i} = (i_1, \dots, i_N)$,

$$f_{\mathbf{i}}^B = f \left(\frac{1}{N} \sum_{k=1}^N O_1(x_{i_k}), \dots, \frac{1}{N} \sum_{k=1}^N O_n(x_{i_k}) \right). \quad (106)$$

The bootstrap estimate of the error is given by

$$(\sigma_f^B)^2 \approx \frac{1}{N^N} \sum_{\mathbf{i}}^{N^N} (f_{\mathbf{i}}^B)^2 - \left(\frac{1}{N^N} \sum_{\mathbf{i}}^{N^N} f_{\mathbf{i}}^B \right)^2. \quad (107)$$

Again, one may check that for the special case $n = 1$ and $f(x) = x$, Eq. (107) reduces to Eq. (94). Clearly, when N is large, it is numerically out of reach to generate all of the N^N bootstrap samples. Typically, to estimate the right-hand side of Eq. (107), 200 or more bootstrap samples are generated stochastically. Since each bootstrap sample is equally probable we can generate them with: $i_k = \text{trunc}(N * \xi_k + 1)$ where ξ_k is a random number in the interval $[0, 1)$ and the function `trunc` returns an integer by truncating the numbers after the decimal point.

A.3 Markov chains

Our task is now to generate a set of states distributed according to a given probability distribution. Here we will consider a discrete space Ω with N_s states; x runs over all the states $1 \dots N_s$, and $P(x)$ denotes the probability of the occurrence of the state x . We introduce a Monte-Carlo time t and a time dependent probability distribution $P_t(x)$ that evolves according to a Markov process: the future ($t + 1$) depends only on the present (t). To define the Markov process, we introduce a matrix $T_{y,x}$, which corresponds to the transition probability from state x to state y . The time evolution of $P_t(x)$ is given by

$$P_{t+1}(y) = \sum_x T_{y,x} P_t(x). \quad (108)$$

T has to satisfy the properties

$$\sum_y T_{y,x} = 1 \text{ and } T_{y,x} \geq 0. \quad (109)$$

Hence, if $P_t(x)$ is a probability distribution then $P_{t+1}(x)$ is also a probability distribution.

T has to be ergodic

$$\forall x, y \in \Omega \exists s \mid (T^s)_{y,x} > 0. \quad (110)$$

Thus, we are assured to sample the whole phase space provided the above is satisfied. Last, we have the requirement of stationarity

$$\sum_x T_{y,x} P(x) = P(y). \quad (111)$$

Once we have reached the desired distribution $P(x)$ we wish to stay there. Stationarity is automatically satisfied if

$$T_{y,x} P(x) = T_{x,y} P(y) \quad (112)$$

as may be seen by summing on both sides over x . This relation is referred to as *detailed balance* or microreversibility. However, one has to keep in mind that only stationarity and not detailed balance is essential.

Given the above, in the Monte Carlo simulation we will generate the Markov Chain

$$x_1, x_2, \dots, x_n,$$

where the conditional probability of sampling the state x_{t+1} given the state x_t reads

$$P(x_{t+1} | x_t) = T_{x_{t+1}, x_t}. \quad (113)$$

Our aim is now to show that when $n \rightarrow \infty$ the fraction of the time one can expect the Markov process to be in state x is $P(x)$, independent of the initial state x_1 . In other words as $n \rightarrow \infty$ the set of states $x_1 \dots x_n$ are distributed according to $P(x)$.

At step t in the Markov chain, the state x will occur on average with probability, $[T^t]_{x, x_1}$. Hence, we have to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n [T^t]_{x, x_1} = P(x). \quad (114)$$

We will first show the above under the assumption that T is regular. That is, there is an integer N such that T^N has only positive, non-zero, entries. If T is regular then T is ergodic. However the inverse is not true so that the condition of regularity is more stringent than that of ergodicity. After the demonstration, we will argue on the basis of a simple example that Eq. (114) is equally valid for ergodic but not regular transition matrices.

We now demonstrate Eq. (114) for regular transition matrices.

Proof: We introduce the set of vectors of real numbers, \mathbf{a}_t

$$\mathbf{a}_{t+1} = \mathbf{a}_t T \quad \text{and} \quad d_t = \max(\mathbf{a}_t) - \min(\mathbf{a}_t) \quad (115)$$

where $\max(\mathbf{a}_t)$ corresponds to the largest element of the vector \mathbf{a} . Since

$$\begin{aligned} [\mathbf{a}_{t+1}]_x &= \sum_y [\mathbf{a}_t]_y T_{y,x} \leq \max(\mathbf{a}_t) \sum_y T_{y,x} = \max(\mathbf{a}_t) \\ \text{and} \quad [\mathbf{a}_{t+1}]_x &= \sum_y [\mathbf{a}_t]_y T_{y,x} \geq \min(\mathbf{a}_t) \sum_y T_{y,x} = \min(\mathbf{a}_t) \end{aligned} \quad (116)$$

the sequence d_t satisfies

$$d_{t+1} \leq d_t. \quad (117)$$

We now show that there is a subsequence that is strictly decreasing. Let us consider vector \mathbf{a}_t with $\max(\mathbf{a}_t) = [\mathbf{a}_t]_M = a_M$ and $\min(\mathbf{a}_t) = [\mathbf{a}_t]_m = a_m$. We define the vector \mathbf{a}^M (\mathbf{a}^m) by replacing all the elements of \mathbf{a}_t apart from the minimal (maximal) one by a_M (a_m). Hence $\mathbf{a}^m \leq \mathbf{a}_t \leq \mathbf{a}^M$ where the inequalities hold element-wise. Furthermore let ϵ be the minimal entry in T^N . In light of the assumption that T is regular, $\epsilon > 0$. With those definitions, we have

$$\begin{aligned} [\mathbf{a}_{t+N}]_x &= [\mathbf{a}_t T^N]_x \leq [\mathbf{a}^M T^N]_x = a_M \sum_{y \neq m} T_{y,x}^N + a_m T_{m,x}^N \\ &= a_M (1 - T_{m,x}^N) + a_m T_{m,x}^N = a_M - (a_M - a_m) T_{m,x}^N \\ &\leq a_M - (a_M - a_m) \epsilon \end{aligned}$$

$$\text{and} \quad [\mathbf{a}_{t+N}]_x \geq [\mathbf{a}_t^m T^N]_x = a_m + (a_M - a_m) T_{M,x}^N \geq a_m + (a_M - a_m) \epsilon. \quad (118)$$

With the above inequalities,

$$d_{t+N} = \max(\mathbf{a}_{t+N}) - \min(\mathbf{a}_{t+N}) \leq a_M - (a_M - a_m)\epsilon - a_m - (a_M - a_m)\epsilon = d_t(1 - 2\epsilon) \quad (119)$$

Hence, $d_{t+mN} \leq d_t(1 - 2\epsilon)^m$ such that the series d_{t+mN} decreases at least exponentially with increasing m . Recalling that $d_{t+1} \leq d_t$, we can find positive numbers τ and b such that

$$d_t \leq be^{-t/\tau}. \quad (120)$$

In particular, we can set $\mathbf{a}_1 = \mathbf{e}_i$ such that $\mathbf{e}_i T$ corresponds to the i^{th} row of the matrix T . Hence, since d_t tends towards zero in the limit $t \rightarrow \infty$, the i^{th} row of matrix T^t tends towards a constant when $t \rightarrow \infty$. Hence, we have

$$\lim_{t \rightarrow \infty} T^t = \begin{pmatrix} \alpha_1 & \cdots & \alpha_1 \\ \alpha_2 & \cdots & \alpha_2 \\ \vdots & \ddots & \vdots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \alpha_{N_s} & \cdots & \alpha_{N_s} \end{pmatrix} \quad (121)$$

where N_s denotes the total number of states. Many comments are now in order.

- (i) Since T satisfies Eq. (109), so does T^t for all values of t . Thus $\sum_{x=1}^{N_s} \alpha_x = 1$ and $\alpha_x > 0$. In other words $\boldsymbol{\alpha}$ is a probability vector.
- (ii) Due to Eq. (120), $T_{x,y}^t = \alpha_x + \Delta_{x,y}^{(t)}$ with $|\Delta_{x,y}^{(t)}| \leq be^{-t/\tau}$.
- (iii) For any probability vector \mathbf{v} , $\lim_{t \rightarrow \infty} T^t \mathbf{v} = \boldsymbol{\alpha}$.
Hence, there is a unique asymptotic distribution: $\boldsymbol{\alpha}$.
- (iv) Due to the stationarity condition, we have $T^t \mathbf{P} = \mathbf{P}$ for all values of t and hence also for $t \rightarrow \infty$. Since there is a unique asymptotic distribution, $\boldsymbol{\alpha} = \mathbf{P}$.

The validity of Eq. (114) now follows from

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n [T^t]_{x,x_1} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n (P(x) + \Delta_{x,x_1}^{(t)}) = P(x) + \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \Delta_{x,x_1}^{(t)}. \quad (122)$$

However since

$$\left| \sum_{t=1}^n \Delta_{x,x_1}^{(t)} \right| \leq \sum_{t=1}^n |\Delta_{x,x_1}^{(t)}| \leq b \sum_{t=1}^n e^{-t/\tau} \quad (123)$$

and $b \sum_{t=1}^n e^{-t/\tau}$ is a convergent series as $n \rightarrow \infty$, Eq. (114) is satisfied. **QED**

We now show on the basis of a simple example that if T is ergodic but not regular, Eq. (114) is still valid. Consider $N_s = 2$ and

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T^{2n+1} = T. \quad (124)$$

T is ergodic but not regular. It is cyclic and in contrast to the regular transition matrices the $\lim_{t \rightarrow \infty} T^t$ does not exist. However, since

$$TP = P \text{ with } P = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \quad (125)$$

and for even and odd values of n

$$\frac{1}{2n} \sum_{t=1}^{2n} T^t = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad \frac{1}{2n+1} \sum_{t=1}^{2n} T^t = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} + \frac{1}{2n+1} T, \quad (126)$$

Eq. (114) holds. For further reading and a more precise and mathematical oriented discussion of Markov chains, the reader is referred to [56].

A.4 Construction of the transition matrix T

Having defined T , we now have to construct it explicitly. Let $T_{y,x}^0$ be the probability of proposing a move from x to y and $a_{y,x}$ the probability of accepting it. $1 - a_{y,x}$ corresponds to the probability of rejecting the move. T_0 is required to satisfy Eq. (109). Since in general we want to propose moves that change the initial configuration, $T_{x,x}^0 = 0$. With $a_{y,x}$ and $T_{y,x}^0$ we build $T_{y,x}$ with

$$T_{y,x} = \begin{cases} T_{y,x}^0 a_{y,x} & \text{if } y \neq x \\ \sum_{z \neq x} T_{z,x}^0 (1 - a_{z,x}) & \text{if } y = x \end{cases} \quad (127)$$

Clearly $T_{y,x}$ satisfies Eq. (109). To satisfy the stationarity, we impose the detailed balance condition to obtain the equality

$$T_{y,x}^0 a_{y,x} P_x = T_{x,y}^0 a_{x,y} P_y. \quad (128)$$

Let us set

$$a_{y,x} = \mathcal{F} \left(\frac{T_{x,y}^0 P_y}{T_{y,x}^0 P_x} \right) \quad (129)$$

with $\mathcal{F} :]0 : \infty[\rightarrow [0, 1]$. Since

$$a_{x,y} = \mathcal{F} \left(\frac{T_{y,x}^0 P_x}{T_{x,y}^0 P_y} \right) = \mathcal{F} \left(\frac{1}{\frac{T_{x,y}^0 P_y}{T_{y,x}^0 P_x}} \right), \quad (130)$$

the detailed balance condition reduces to

$$\frac{\mathcal{F}(Z)}{\mathcal{F}(1/Z)} = Z \text{ where } Z = \frac{T_{x,y}^0 P_y}{T_{y,x}^0 P_x}. \quad (131)$$

There are many possible choices. The *Metropolis algorithm* is based on the choice

$$\mathcal{F}(Z) = \min(Z, 1). \quad (132)$$

Thus, one proposes a move from \mathbf{x} to \mathbf{y} and accepts it with probability $Z = \frac{T_{\mathbf{x},\mathbf{y}}^0 P_{\mathbf{y}}}{T_{\mathbf{y},\mathbf{x}}^0 P_{\mathbf{x}}}$. In the practical implementation, one picks a random number r in the interval $[0 : 1]$. If $r < Z$ ($r > Z$) one accepts (rejects) the move. Alternative choices of $\mathcal{F}(Z)$ are for example

$$\mathcal{F}(Z) = \frac{Z}{1+Z} \quad (133)$$

which is referred to as the heat-bath method.

That the so-constructed T matrix is ergodic depends upon the choice of T^0 . In many cases, one will wish to combine different types of moves to achieve ergodicity. For a specific move i , we construct $T^{(i)}$ as shown above so that $T^{(i)}$ satisfies conditions (109) and (112). The moves may be combined in two ways

$$T = \sum_i \lambda_i T^{(i)}, \quad \sum_i \lambda_i = 1 \quad (134)$$

which is referred to as random updating since one picks with probability λ_i the move $T^{(i)}$. Clearly, T equally satisfies (109) and (112), and the moves have to be chosen appropriately to satisfy the ergodicity condition. Another choice is sequential upgrading. A deterministic ordering of the moves is chosen to obtain

$$T = \prod_i T^{(i)}. \quad (135)$$

This choice does not satisfy detailed balance but does satisfy stationarity (111) as well as (109). Again ergodicity has to be *checked* on a case to case basis.

The observable O may now be estimated with

$$\langle O \rangle_P \approx \frac{1}{N} \sum_{t=1}^N O(\mathbf{x}_t). \quad (136)$$

The required value of N depends on the autocorrelation time of the observable O

$$C_O(t) = \frac{\frac{1}{N} \sum_{s=1}^N O(\mathbf{x}_s)O(\mathbf{x}_{s+t}) - \left(\frac{1}{N} \sum_{s=1}^N O(\mathbf{x}_s)\right)^2}{\frac{1}{N} \sum_{s=1}^N O(\mathbf{x}_s)^2 - \left(\frac{1}{N} \sum_{s=1}^N O(\mathbf{x}_s)\right)^2}. \quad (137)$$

One expects $C_O(t) \sim e^{-t/\tau_O}$, where τ_O corresponds to the MC time scale on which memory of the initial configuration is lost. Hence, to obtain meaningful results, $N \gg \tau_O$. Note that one should also take into account a *warm up* time by discarding at least the first τ_O configurations in the MC sequence. Naively, one would expect $\tau_O = \tau$. However, this depends on the overlap of the observable with the slowest mode in the MC dynamics, which relaxes as $e^{-t/\tau}$. In particular in a model with spin rotation symmetry the slowest mode may correspond to the rotation of the total spin. In this case, observables that are invariant under a spin rotation will not be affected by the slowest mode of the MC dynamics. Hence, in this case $\tau_O < \tau$.

We now consider the error estimation. To apply the central limit theorem, we need a set of independent estimates of $\langle O \rangle_P$. This may be done by regrouping the data into *bins* of size $n\tau_O$.

$$\tilde{O}_n(t) = \frac{1}{n\tau_O} \sum_{s=1}^{n\tau_O} O(\mathbf{x}_{(t-1)n\tau_O+s}) \quad (138)$$

with $t = 1 \dots N/(n\tau_O)$. If n is large enough (i.e. $n \approx 10 - 20$) then $\tilde{O}_n(t)$ may be considered as an independent estimate, and the error is given by

$$\sigma_n = \sqrt{\frac{1}{M} \left(\frac{1}{M} \sum_{t=1}^M \tilde{O}_n(t)^2 - \left(\frac{1}{M} \sum_{i=1}^M \tilde{O}_n(t) \right)^2 \right)} \quad (139)$$

where $M = N/(n\tau_O)$. If n is large enough, the error σ_n should be independent of n .

A.5 One-dimensional Ising model

We conclude this appendix with an example of a Monte Carlo simulation with error analysis for the one-dimensional Ising model

$$H(\{\sigma\}) = -J \sum_{i=1}^L \sigma_i \sigma_{i+1}, \quad \sigma_{L+1} = \sigma_1 \quad (140)$$

where $\sigma_i = \pm 1$. This model is easily solved exactly with the transfer matrix method and thus produces a useful testing ground for the MC approach. In particular at zero temperature, a phase transition to a ferromagnetically ordered phase ($J > 0$) occurs [57]. Spin-spin correlations are given by

$$g(r) = \frac{1}{L} \sum_{i=1}^L \langle \sigma_i \sigma_{i+r} \rangle \quad \text{with} \quad \langle \sigma_i \sigma_{i+r} \rangle = \frac{\sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})} \sigma_i \sigma_{i+r}}{\sum_{\{\sigma\}} e^{-\beta H(\{\sigma\})}} \quad (141)$$

where β corresponds to the inverse temperature.

We now construct the transition matrix T corresponding to a random single site updating scheme. For an L -site chain, we denote the spin configuration by

$$x = (\sigma_{x,1}, \sigma_{x,2}, \dots, \sigma_{x,L}). \quad (142)$$

The transition matrix reads

$$T = \frac{1}{L} \sum_{i=1}^L T^{(i)} \quad \text{with} \quad T_{x,y}^{(i)} = T_{x,y}^{(i),0} a_{x,y} \quad \text{and}$$

$$T_{x,y}^{(i),0} = \begin{cases} 1 & \text{if } x = (\sigma_{y,1}, \dots, -\sigma_{y,i}, \dots, \sigma_{y,L}) \\ 0 & \text{otherwise} \end{cases} \quad (143)$$

The above corresponds to choosing a site i randomly, changing the orientation of the spin ($\sigma_i \rightarrow -\sigma_i$), and accepting the move with probability $a_{x,y}$ corresponding to a heat-bath or Metropolis algorithm. The MC time unit corresponds to a single sweep, meaning that L sites are randomly chosen before a measurement is carried out.

Fig. 3 plots the autocorrelation time for $g(r = L/2)$ on an $L = 24$ site lattice at $\beta J = 1$ and $\beta J = 2$. From Fig. 3a one can extract the autocorrelation time: $\tau_O \approx 11, 54$ for $\beta J = 1, 2$ respectively. Fig. 3b plots the error as a function of bin size in units of the τ_O (see Eq. (139)). As one can see, $n \approx 10$ is sufficient to get a reliable estimate of the error.

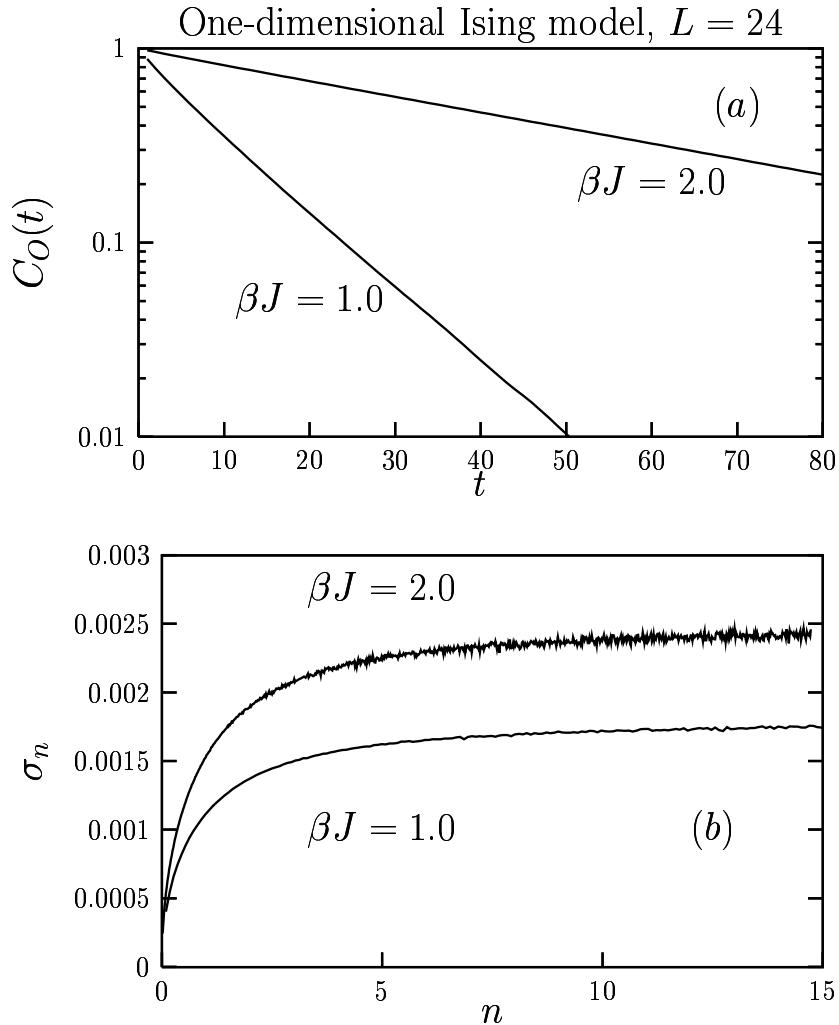


Fig. 3: One dimensional Ising model on an $L=24$ site lattice. (a) Autocorrelation time (see Eq. (137)) for $g(r = L/2)$. The time unit corresponds to a single sweep. (b) Estimate of the error (see Eq. (139)). Here, n corresponds to the size of the bins in units of the autocorrelation time. $n \sim 10$ is sufficient to obtain a reliable estimate of the error. After 2×10^6 sweeps, our results yield $g(r = L/2) = 0.076 \pm 0.0018$ and 0.909 ± 0.0025 for $\beta t = 1$ and 2 respectively. The exact result is $g(r = L/2) = 0.0760$ and 0.9106 at $\beta t = 1$ and 2 respectively.

B Proof of the determinant identity

In this section, a general determinant identity is proven [32] that can be used to derive Wick's theorem for contributions of a configuration C_n to physical observables. Let us define the vectors $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{C}^m$ and the numbers $\alpha_{ij} \in \mathbb{C}$. Further, let $A \in \mathbb{C}^{m \times m}$ be a matrix of rank m . We define the non-singular matrices $M_n \in \mathbb{C}^{(m+n) \times (m+n)}$ and $A_{ij} \in \mathbb{C}^{(m+1) \times (m+1)}$ by

$$M_n = \begin{pmatrix} A & \mathbf{u}_1 & \dots & \mathbf{u}_n \\ \mathbf{v}_1^T & \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_n^T & \alpha_{n1} & \dots & \alpha_{nn} \end{pmatrix}, \quad A_{ij} = \begin{pmatrix} A & \mathbf{u}_j \\ \mathbf{v}_i^T & \alpha_{ij} \end{pmatrix}. \quad (144)$$

With these definitions, the following determinant identity holds

$$\det M_n (\det A)^{n-1} = \det \begin{pmatrix} \det A_{11} & \dots & \det A_{1n} \\ \vdots & \ddots & \vdots \\ \det A_{n1} & \dots & \det A_{nn} \end{pmatrix}. \quad (145)$$

The identity can be proven by induction in n . It is trivial for $n = 1$, so we have to start with $n = 2$, where we have to show

$$\frac{\det M_2}{\det A} = \frac{\det A_{11}}{\det A} \frac{\det A_{22}}{\det A} - \frac{\det A_{12}}{\det A} \frac{\det A_{21}}{\det A}. \quad (146)$$

For the following calculations, we introduce several vectors

$$\mathbf{u}_{ij}^1 = \begin{pmatrix} \mathbf{u}_j \\ \alpha_{ij} - 1 \end{pmatrix}, \quad \mathbf{v}_{ij}^2 = \begin{pmatrix} \mathbf{v}_i \\ 0 \end{pmatrix}, \quad \mathbf{u}^2 = \mathbf{v}^1 = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in \mathbb{C}^{m+1}. \quad (147)$$

$$\mathbf{u}_M^1 = \begin{pmatrix} \mathbf{u}_2 \\ \alpha_{12} \\ \alpha_{22} - 1 \end{pmatrix}, \quad \mathbf{v}_M^2 = \begin{pmatrix} \mathbf{v}_2 \\ \alpha_{21} \\ 0 \end{pmatrix}, \quad \mathbf{u}_M^2 = \mathbf{v}_M^1 = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in \mathbb{C}^{m+2}. \quad (148)$$

Let us define the *expanded* matrix C_{ex} of a square matrix C as the matrix C expanded by one row and one column containing a unit vector

$$C_{ex} = \begin{pmatrix} C & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}. \quad (149)$$

As a last definition, we introduce the abbreviation $b_{ij} = \mathbf{v}_i^T \mathbf{A}^{-1} \mathbf{u}_j$. Using these notations, we can write the matrices A_{ij} as

$$A_{ij} = A_{ex} + \mathbf{u}_{ij}^1 \mathbf{v}^{1T} + \mathbf{u}^2 \mathbf{v}_{ij}^{2T}. \quad (150)$$

To calculate the determinant $\det A_{ij}$, we use the matrix determinant lemma $\det(A + uv^T) = (1 + v^T A^{-1} u) \det A$, yielding

$$\frac{\det A_{ij}}{\det A_{ex}} = \left[1 + \mathbf{v}_{ij}^{2T} (A_{ex} + \mathbf{u}_{ij}^1 \mathbf{v}^{1T})^{-1} \mathbf{u}^2 \right] (1 + \mathbf{v}^{1T} A_{ex}^{-1} \mathbf{u}_{ij}^1). \quad (151)$$

The inverse matrix of $(\mathbf{A}_{\text{ex}} + \mathbf{u}_{ij}^1 \mathbf{v}^1{}^T)$ can be obtained from the Sherman-Morrison formula, and a tedious calculation making use of the special form of the vectors and matrices gives the result

$$\frac{\det \mathbf{A}_{ij}}{\det \mathbf{A}} = \alpha_{ij} - b_{ij}. \quad (152)$$

From this, the right-hand side of Eq. (146) can be easily obtained. For the left-hand side, we have to perform an analogous calculation using the decomposition of the matrix \mathbf{M}_2

$$\mathbf{M}_2 = \mathbf{A}_{11\text{ex}} + \mathbf{u}_{\mathbf{M}}^1 \mathbf{v}_{\mathbf{M}}^1{}^T + \mathbf{u}_{\mathbf{M}}^2 \mathbf{v}_{\mathbf{M}}^2{}^T. \quad (153)$$

Again, we apply the matrix determinant lemma two times and insert the Sherman-Morrison formula to calculate the inverse matrix of $(\mathbf{A}_{11\text{ex}} + \mathbf{u}_{\mathbf{M}}^1 \mathbf{v}_{\mathbf{M}}^1{}^T)$. Simplifying the result as far as possible, we finally arrive at

$$\frac{\det \mathbf{M}_2}{\det \mathbf{A}} = (\alpha_{11} - b_{11})(\alpha_{22} - b_{22}) - (\alpha_{12} - b_{12})(\alpha_{21} - b_{21}). \quad (154)$$

If we compare (154) with (152), it is clear that Eq. (146) holds.

We now assume that for a certain value $n \in \mathbb{N}$ Eq. (145) holds. For $n+1$, we can cast the matrix \mathbf{M}_{n+1} in a form where we can make use of Eq. (145) holding for n

$$\mathbf{M}_{n+1} = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{u}}_2 & \dots & \tilde{\mathbf{u}}_{n+1} \\ \tilde{\mathbf{v}}_2^T & \alpha_{2,2} & \dots & \alpha_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{v}}_n^T & \alpha_{n,2} & \dots & \alpha_{n,n+1} \\ \tilde{\mathbf{v}}_{n+1}^T & \alpha_{n+1,2} & \dots & \alpha_{n+1,n+1} \end{pmatrix}, \quad (155)$$

where we have introduced the new matrix $\tilde{\mathbf{A}}$ and the vectors $\tilde{\mathbf{u}}_i$ and $\tilde{\mathbf{v}}_i$ with

$$\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A} & \mathbf{u}_1 \\ \mathbf{v}_1^T & \alpha_{11} \end{pmatrix}, \quad \tilde{\mathbf{u}}_i = \begin{pmatrix} \mathbf{u}_i \\ \alpha_{1i} \end{pmatrix}, \quad \tilde{\mathbf{v}}_i = \begin{pmatrix} \mathbf{v}_i \\ \alpha_{i1} \end{pmatrix}. \quad (156)$$

Further, we need the matrices $\tilde{\mathbf{A}}_{ij}$ defined analogously to (144)

$$\tilde{\mathbf{A}}_{ij} = \begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{u}}_j \\ \tilde{\mathbf{v}}_i^T & \alpha_{ij} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{u}_1 & \mathbf{u}_j \\ \mathbf{v}_1^T & \alpha_{11} & \alpha_{1j} \\ \mathbf{v}_i^T & \alpha_{i1} & \alpha_{ij} \end{pmatrix}. \quad (157)$$

With these definitions, and with the abbreviations $a_{ij} = \det \mathbf{A}_{ij}$ and $\tilde{a}_{ij} = \det \tilde{\mathbf{A}}_{ij}$, we are now able to apply Eq. (145) holding for n

$$\det \mathbf{M}_{n+1} (\det \tilde{\mathbf{A}})^{(n-1)} = \det \begin{pmatrix} \tilde{a}_{2,2} & \dots & \tilde{a}_{2,n+1} \\ \vdots & \ddots & \vdots \\ \tilde{a}_{n+1,2} & \dots & \tilde{a}_{n+1,n+1} \end{pmatrix}. \quad (158)$$

For \tilde{a}_{ij} , we make use of Eq. (145) with $n = 2$, which we have proved above

$$\tilde{a}_{ij} = \frac{1}{\det \mathbf{A}} (a_{11}a_{ij} - a_{i1}a_{1j}). \quad (159)$$

Inserting this result in (158) yields a determinant with entries of the form $a_{11}a_{ij} - a_{i1}a_{1j}$. We make use of the multi linearity of the determinant to decompose this expression and we obtain a sum of determinants with prefactors of the form a_{ij} . Eliminating zero contributions, the resulting expression corresponds precisely to the Laplace expansion of a larger determinant, and we finally obtain

$$\det \mathbf{M}_{n+1} \det \mathbf{A}^n = \det \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n+1} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,n+1} \end{pmatrix}. \quad (160)$$

This is the identity (145) for $n + 1$. Hence, we have derived the determinant identity for $n + 1$ using only the identity for n and $n = 2$. By induction, the identity (145) therefore holds for every $n \in \mathbb{N}$, as it is trivial for $n = 1$.

References

- [1] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)
- [2] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [3] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [4] F. Assaad and H. Evertz in *Computational Many-Particle Physics* ed. by H. Fehske, R. Schneider, and A. Weiße, Lecture Notes in Physics, Vol. 739 (Springer, 2008) p. 277
- [5] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A.I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004)
- [6] F.F. Assaad and T.C. Lang, Phys. Rev. B **76**, 035116 (2007)
- [7] P. Werner and A.J. Millis, Phys. Rev. Lett. **99**, 146404 (2007)
- [8] F.F. Assaad, Phys. Rev. B **78**, 155124 (2008)
- [9] M. Hohenadler, F.F. Assaad, and H. Fehske, Phys. Rev. Lett. **109**, 116407 (2012)
- [10] M.Z. Hasan and C.L. Kane, Rev. Mod. Phys. **82**, 3045 (2010)
- [11] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011)
- [12] C.L. Kane and E.J. Mele, Phys. Rev. Lett. **95**, 146802 (2005)
- [13] M. Hohenadler and F.F. Assaad, J. Phys.: Condens. Matter **25**, 143201 (2013)
- [14] M. Hohenadler, T.C. Lang, and F.F. Assaad, Phys. Rev. Lett. **106**, 100403 (2011)
- [15] M. Hohenadler and F.F. Assaad, Phys. Rev. B **85**, 081106 (2012)
- [16] F. Goth, D.J. Luitz, and F.F. Assaad, Phys. Rev. B **88**, 075110 (2013)
- [17] M. Weber, M. Hohenadler, and F.F. Assaad, Phys. Rev. B **89**, 205125 (2014)
- [18] F. Goth and F.F. Assaad, Phys. Rev. B **85**, 085129 (2012)
- [19] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
- [20] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)
- [21] A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. B **75**, 045118 (2007)
- [22] J.L. Smith and Q. Si, Phys. Rev. B **61**, 5184 (2000)
- [23] D. Rost, F. Assaad, and N. Blümer, Phys. Rev. E **87**, 053305 (2013)

- [24] E.F. Huffman and S. Chandrasekharan, Phys. Rev. B **89**, 111101 (2014)
- [25] A.N. Rubtsov, V.V. Savkin, and A.I. Lichtenstein, Phys. Rev. B **72**, 035122 (2005)
- [26] S.M.A. Rombouts, K. Heyde, and N. Jachowicz, Phys. Rev. Lett. **82**, 4155 (1999)
- [27] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Europhys. Lett. **82**, 57003 (2008)
- [28] P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **97**, 076405 (2006)
- [29] P.W. Anderson, Phys. Rev. **124**, 41 (1961)
- [30] A.C. Hewson: *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, 1997)
- [31] J.W. Negele and H. Orland: *Quantum Many body systems* (Addison-Wesley, Redwood City, 1988)
- [32] D.J. Luitz and F.F. Assaad, Phys. Rev. B **81**, 024509 (2010)
- [33] D.J. Luitz, F.F. Assaad, T. Novotný, C. Karrasch, and V. Meden, Phys. Rev. Lett. **108**, 227001 (2012)
- [34] K. Mikelsons, A. Macridin, and M. Jarrell, Phys. Rev. E **79**, 057701 (2009)
- [35] J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)
- [36] J. Yoo, S. Chandrasekharan, R.K. Kaul, D. Ullmo, and H.U. Baranger, J. Phys. A **38**, 10307 (2005)
- [37] M. Hohenadler, Z.Y. Meng, T.C. Lang, S. Wessel, A. Muramatsu, and F.F. Assaad, Phys. Rev. B **85**, 115132 (2012)
- [38] C. Wu and S.-C. Zhang, Phys. Rev. B **71**, 155115 (2005)
- [39] D.J. Luitz: *Numerical methods and applications in many fermion systems*, PhD thesis, Universität Würzburg Astronomie (2013) <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:20-opus-75927>
- [40] K. Haule, Phys. Rev. B **75**, 155113 (2007)
- [41] N. Parragh, A. Toschi, K. Held, and G. Sangiovanni, Phys. Rev. B **86**, 155158 (2012)
- [42] A.M. Läuchli and P. Werner, Phys. Rev. B **80**, 235117 (2009)
- [43] H. Shinaoka, M. Dolfi, M. Troyer, and P. Werner, arxiv:1404.1259 (2014)
- [44] H. Hafermann, K.R. Patton, and P. Werner, Phys. Rev. B **85**, 205106 (2012)

- [45] P. Werner and A.J. Millis, Phys. Rev. Lett. **99**, 126405 (2007)
- [46] P. Werner, E. Gull, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **101**, 166405 (2008)
- [47] N. Parragh, G. Sangiovanni, P. Hansmann, S. Hummel, K. Held, and A. Toschi, Phys. Rev. B **88**, 195116 (2013)
- [48] J. Werner and F.F. Assaad, Phys. Rev. B **88**, 035113 (2013)
- [49] J. Werner and F.F. Assaad, Phys. Rev. B **89**, 245119 (2014)
- [50] P.W. Anderson, P.A. Lee, M. Randeria, T.M. Rice, N. Trivedi, and F.C. Zhang, Journal of Physics: Condens. Matter **16**, R755 (2004)
- [51] M. Greiner, O. Mandel, T. Esslinger, T.W. Hansch, and I. Bloch, Nature **415**, 39 (2002)
- [52] B. Sipos, A.F. Kusmartseva, A. Akrap, H. Berger, L. Forro, and E. Tutis, Nature Mater. **7**, 960 (2008)
- [53] R.P. Hardikar and R.T. Clay, Phys. Rev. B **75**, 245103 (2007)
- [54] M. Hohenadler and F.F. Assaad, J. Phys.: Condens. Matter **25**, 014005 (2013)
- [55] B. Efron: *The jackknife, the bootstrap and other resampling plans*, CBMS-NSF conference series in applied mathematics (J. W. Arrowsmith, Bristol, 1982)
- [56] G.J. Kemeny and J. Snell: *Finite Markov Chains*, The university series in undergraduate mathematics (Van Nostrand, New York, 1960)
- [57] R.J. Baxter: *Exactly solved models in statistical mechanics* (Academic Press Limited, London, 1989)

8 Quantum Cluster Methods

Erik Koch

Computational Materials Science

German Research School for Simulation Sciences

Jülich

Contents

| | |
|--|-----------|
| 1 Periodic systems | 4 |
| 1.1 Lattices | 4 |
| 1.2 Superlattices | 8 |
| 2 Variational methods | 10 |
| 2.1 Variational Monte Carlo | 14 |
| 2.2 Correlated sampling | 16 |
| 2.3 Gutzwiller approximation | 18 |
| 2.4 Brinkman-Rice transition | 20 |
| 3 Projection techniques | 23 |
| 3.1 Importance sampling | 24 |
| 3.2 Fixed-node Monte Carlo | 25 |
| 3.3 Optimization of the trial function | 26 |
| 3.4 Quasi-particle energies | 30 |
| 4 Conclusion | 35 |

The central challenge of electronic structure theory is the solution of the many-electron Hamiltonian in the Born-Oppenheimer approximation (in atomic units)

$$H = -\frac{1}{2} \sum_i \vec{\nabla}_i^2 + \sum_i V_{\text{ext}}(\vec{r}_i) + \sum_{i < j} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \quad (1)$$

where the external potential is, e.g., the Coulomb potential of the nuclei of charge Z_I at position \vec{R}_I , shifted by their mutual Coulomb interaction

$$V_{\text{ext}}(\vec{r}) = \sum_I \frac{Z_I}{|\vec{r} - \vec{R}_I|} + \sum_{I < J} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}. \quad (2)$$

Introducing a basis-set $\{\varphi_\alpha(x)\}$ of spin-orbitals, we can rewrite H in second quantization

$$H = -\sum_{\alpha\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} U_{\beta\gamma} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta \quad (3)$$

where the creation/annihilation operators fulfill the anticommutation relations $\{c_\alpha^\dagger, c_\beta^\dagger\} = 0 = \{c_\alpha, c_\beta\}$ and $\{c_\alpha, c_\beta^\dagger\} = \langle \alpha | \beta \rangle = S_{\alpha\beta}$, and S is the overlap matrix [1]. The matrix elements are given by integrating over the orbital degrees of freedom $x = (\vec{r}, \sigma)$

$$t_{\alpha\beta} = \sum_{\alpha'\beta'} (S^{-1})_{\alpha\alpha'} \int dx \overline{\varphi_{\alpha'}(x)} \left(\frac{1}{2} \vec{\nabla}^2 - V_{\text{ext}}(\vec{r}) \right) \varphi_{\beta'}(x) (S^{-1})_{\beta'\beta} \quad (4)$$

and

$$U_{\beta\gamma} = \sum_{\alpha'\delta'} S_{\alpha\alpha'}^{-1} S_{\beta\beta'}^{-1} \int dx \int dx' \overline{\varphi_{\alpha'}(x)} \overline{\varphi_{\beta'}(x')} \frac{1}{|\vec{r} - \vec{r}'|} \varphi_{\gamma'}(x') \varphi_{\delta'}(x) S_{\gamma'\gamma}^{-1} S_{\delta'\delta}^{-1}. \quad (5)$$

This representation of the Hamiltonian is suited for introducing approximations. By truncating the basis-set to only K functions, the Hilbert space \mathcal{H} of H for a system with N electrons is restricted to a finite variational subspace $\mathcal{H}(\{\varphi_{\alpha_n}(x); |n = 1 \dots K\})$ of dimension $\binom{K}{N}$. Working with a finite basis set introduces a basis-set error. To keep it small, the basis functions are chosen such that the eigenstates of interest are represented well on $\mathcal{H}(\{\varphi_{\alpha_n}(x) | n = 1 \dots K\})$, using, e.g., low-energy orbitals to represent the ground state. The basis-set error can be estimated by comparing results calculated with basis sets of increasing size and extrapolating to the complete-basis-set-limit. Such calculations are computationally demanding, as the dimension of the Hilbert space increases for $K \gg N \gg 1$ (using Stirling's approximation) with a high power, given by the (fixed) number of electrons in the system, as $\mathcal{O}(K^N)$.

For extended systems the problem becomes even harder. To ensure size-consistency [2], meaning that the basis-set error for extensive observables, e.g., the total energy, scales at most with the number of electrons, we have to increase K along with N , leading to an exponential scaling of the variational space with system size. Practical simulations are therefore restricted to quite small clusters. These have a large fraction of surface atoms. As a simple example, for a $10 \times 10 \times 10$ cluster 488 of the 1000 atoms are on the surface. The surface effects can be

removed by putting the system in a simulation cell spanned by three vectors \vec{R}_i and assuming that the system is periodically repeated [3]. Instead of leaving the system, an electron passing through a face of the simulation cell continues into the neighboring cell, while one of its images enters through the opposite face. Thus, we are dealing with an extended system with an infinite number of electrons. Still, by virtue of the periodicity, only the N electrons inside the simulation cell are independent degrees of freedom. We can then restrict the calculation to the simulation cell $\mathcal{C} = \{\sum_i x_i \vec{R}_i \mid x_i \in [0, 1]\}$ by including the periodic images of the external potential (created inside the simulation cell) and the interaction with the electrons outside the simulation box in the Hamiltonian [4]

$$H_{\text{pbc}} = -\frac{1}{2} \sum_{i=1}^N \vec{\nabla}_i^2 + \sum_{\vec{n} \in \mathbb{Z}^3} \sum_i V_{\text{ext}}^{\mathcal{C}}(\vec{r}_i - \vec{R}_{\vec{n}}) + \frac{1}{2} \sum_{\vec{n} \in \mathbb{Z}^3} \sum'_{i,j} \frac{1}{|\vec{r}_i - \vec{r}_j - \vec{R}_{\vec{n}}|}. \quad (6)$$

where $\vec{R}_{\vec{n}} = \sum_{i=1}^3 n_i \vec{R}_i$ and the prime on the last sum indicates that $i \neq j$ when $\vec{n} = 0$. The eigenfunctions of (6) on the simulation cell \mathcal{C} represent a system of average electron density $N/V_{\mathcal{C}}$. We see that, while removing surface effects, the introduction of periodic boundary conditions not only modifies interactions of ranges longer than the radius of the simulation cell but also suppress fluctuations of the number of electrons between simulation cells. Moreover, the average electron density can only be a multiple of $1/V_{\mathcal{C}}$ as the simulation cell must contain an integer number of electrons. Obviously, these finite-size errors vanish in the limit of infinite simulation cell volume $V_{\mathcal{C}} \rightarrow \infty$. Fig. 1 shows a comparison of the finite-size scaling for the ground-state energy with open and with periodic boundary conditions. For systems that develop long-range correlations, care has to be taken in the finite-size extrapolation for correlation functions $C(\vec{r}, \vec{r}')$, as imposing periodic boundary conditions $C(\vec{r}, \vec{r} + \vec{R}_i) = C(\vec{r}, \vec{r}')$ can frustrate correlations that are not commensurate with the simulation cell. This becomes particularly evident for a crystal in which the external potential of the infinite system is periodic $V_{\text{ext}}(\vec{r} + \vec{a}_i) = V_{\text{ext}}(\vec{r})$ with the periodicity of the lattice spanned by the vectors \vec{a}_i . Only when the vectors \vec{R}_i spanning \mathcal{C} are chosen as integer linear combinations $\vec{R}_i = \sum_j n_{ij} \vec{a}_j$ will the external potential in H_{pbc} agree with V_{ext} :

$$\sum_{\vec{n} \in \mathbb{Z}^3} V_{\text{ext}}^{\mathcal{C}}(\vec{r} - \vec{R}_{\vec{n}}) = V_{\text{ext}}(\vec{r}), \quad (7)$$

where $V_{\text{ext}}^{\mathcal{C}}$ is the external potential originating, e.g., from the nuclei inside \mathcal{C} .

For such periodic systems the nature of the many-body problem becomes apparent. The interaction term by itself is not particularly complicated. It is diagonal in real space, so finding the arrangement of electrons that minimizes their mutual Coulomb repulsion is a straightforward classical optimization problem, the solution being a Wigner crystal [5]. The kinetic energy, on the other hand, is diagonal in k -space. For a lattice-periodic potential, V_{ext} couples only a discrete set of k -vectors, so that the single-electron part of H can be solved in terms of Bloch waves. Solutions of the full problem thus have to balance the extended Bloch waves (kinetic energy) against the localized Wigner crystal (electron-electron repulsion).

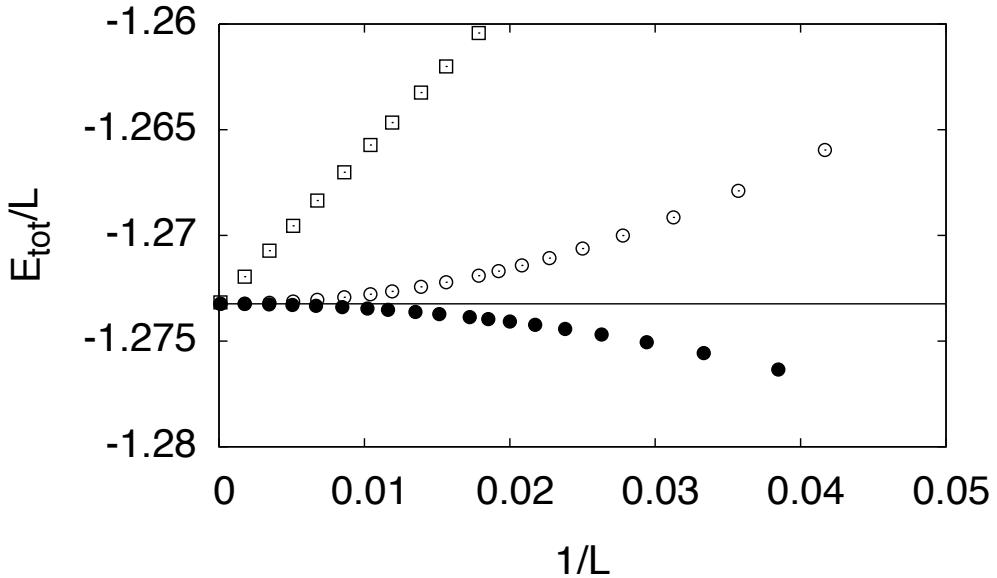


Fig. 1: Energy per site for Hubbard chains of L sites with open and periodic boundary conditions. Calculations are for the non-interacting half-filled Hubbard model with nearest-neighbor hopping t . For open boundary conditions (open squares) the energy converges linearly with $1/L$ to the result for the infinite chain, $\varepsilon_\infty = -4t/\pi$. For periodic boundary conditions (circles) convergence is quadratic, both for open- and closed-shell chains.

1 Periodic systems

1.1 Lattices

A d -dimensional lattice is the collection of points

$$\mathcal{L} = \left\{ \mathbf{r}_{n_1, \dots, n_d} = \sum_i n_i \mathbf{a}_i \mid n_i \in \mathbb{Z} \right\} \quad (8)$$

defined by the integer linear combinations of a set of d linearly independent vectors \mathbf{a}_i . Arranging the vectors \mathbf{a}_i into a matrix $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots)$, we see that the points in the lattice are given by $\mathbf{r}_n = \mathbf{An}$ with $n \in \mathbb{Z}^d$; i.e., \mathbf{A} maps the d -dimensional cubic lattice into \mathcal{L} . Likewise, we can write an arbitrary point $\mathbf{r} = \mathbf{Af}$. In general, the coordinates $f = \mathbf{A}^{-1}\mathbf{r}$ in the basis \mathbf{A} will not be integers. They are called the fractional coordinates of \mathbf{r} . The primitive cell defined by \mathbf{A} is the set of all points \mathbf{r} whose fractional coordinates lie in the unit cube $f \in [0, 1)^d$. Its volume is $V_c = |\det(\mathbf{A})|$. It is convenient to arrange the vectors \mathbf{a}_i such that the determinant is positive. The vectors \mathbf{a}_i are called the primitive vectors of the lattice \mathcal{L} . They are not unique: we can construct an equivalent set of primitive vectors $\tilde{\mathbf{A}}$ by adding to $\pm \mathbf{a}_i$ any integer multiple of the other $\mathbf{a}_{j \neq i}$, such that $|\det(\tilde{\mathbf{A}})| = |\det(\mathbf{A})|$. Since the transformation $\tilde{\mathbf{A}} = \mathbf{AM}$ is then given by an integer matrix \mathbf{M} with $|\det(\mathbf{M})| = 1$, by Cramer's rule, its inverse is also an integer matrix. Thus any point in the lattice \mathcal{L} can be written in terms of either set of primitive vectors: $\mathbf{An} = \tilde{\mathbf{A}}\tilde{n}$, with the integer indices related by $\tilde{n} = \mathbf{M}^{-1}\mathbf{n}$. The canonical choice is to make the primitive vectors as short as possible, $\tilde{\mathbf{a}}_i = \sum_{j \neq i} \text{round}(\mathbf{a}_i \cdot \mathbf{a}_j / \|\mathbf{a}_j\|^2) \mathbf{a}_j$, so they provide the notion of nearest-neighbor and give a compact unit cell.

The reciprocal lattice $\mathcal{R}_{\mathcal{L}}$ associated with \mathcal{L} arises naturally when considering the Fourier transform of lattice periodic functions, i.e., functions $V(\mathbf{r} + \mathbf{A}\mathbf{n}) = V(\mathbf{r})$. The Fourier expansion of a general function $V(\mathbf{r})$ is given by

$$V(\mathbf{r}) = \int d^d k \hat{V}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (9)$$

When $V(\mathbf{r})$ is periodic on \mathcal{L} , we have

$$V(\mathbf{r} + \mathbf{A}\mathbf{n}) = \int d^d k \hat{V}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{A}\mathbf{n}} = V(\mathbf{r}). \quad (10)$$

By the linear independence of Fourier modes it follows that only terms with $\exp(i\mathbf{k}\cdot\mathbf{A}\mathbf{n}) = 1$ for all $\mathbf{n} \in \mathbb{Z}^d$ can contribute. The k -vectors fulfilling this condition form the reciprocal lattice

$$\mathcal{R}_{\mathcal{L}} = \left\{ \mathbf{Gm} \mid \mathbf{m} \in \mathbb{Z}^d \right\} \quad (11)$$

with primitive vectors $\mathbf{G} = (2\pi\mathbf{A}^{-1})^T$. Since $(2\pi\mathbf{G}^{-1})^T = \mathbf{A}$, the reciprocal lattice of $\mathcal{R}_{\mathcal{L}}$ is \mathcal{L} . By construction, the reciprocal lattice vectors $\mathbf{g} \in \mathcal{R}_{\mathcal{L}}$ define plane waves $\exp(i\mathbf{g}\cdot\mathbf{r})$ for which all lattice points $\mathbf{r}_n \in \mathcal{L}$ fall on planes of phase = 1. The reciprocal lattice vectors, except the gamma-point $\mathbf{g} = \mathbf{0}$, are thus orthogonal to planes containing an infinite number of lattice points. A given set of lattice planes can be characterized by the shortest reciprocal lattice vector $\mathbf{g}_{\min} = \mathbf{Gm}$ perpendicular to it. The expansion coefficients \mathbf{m} in terms of the primitive reciprocal vectors are the Miller indices. As for the real lattice, the primitive vectors are not unique: $\tilde{\mathbf{A}} = \mathbf{AM}$ gives $\tilde{\mathbf{G}} = (2\pi\tilde{\mathbf{A}}^{-1})^T = \mathbf{G}(\mathbf{M}^{-1})^T$, which also span $\mathcal{R}_{\mathcal{L}}$. The canonical choice for the primitive reciprocal cell is $\mathbf{k} \in \mathbf{G} (-1/2, 1/2]^d$. A momentum \mathbf{k} from a primitive reciprocal cell (first Brillouin zone) is called a crystal momentum.

Transforming the single-electron Hamiltonian

$$H_{\text{single}} = -\frac{1}{2} \nabla_{\mathbf{r}}^2 + V_{\text{ext}}(\mathbf{r}) \quad (12)$$

with lattice-periodic potential $V_{\text{ext}}(\mathbf{r}) = \sum_{\mathbf{m} \in \mathbb{Z}^d} \hat{V}_{\mathbf{Gm}} e^{i\mathbf{Gm}\cdot\mathbf{r}}$ to k -space

$$\langle \mathbf{k} | H_{\text{single}} | \mathbf{k}' \rangle = \frac{\mathbf{k}^2}{2} \delta(\mathbf{k} - \mathbf{k}') + \hat{V}_{\mathbf{Gm}} \delta(\mathbf{k} - \mathbf{k}' - \mathbf{Gm}) \quad (13)$$

or, more elegantly,

$$H_{\text{single}} = \sum_{\mathbf{k}} \frac{\mathbf{k}^2}{2} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k},\sigma} + \sum_{\mathbf{k}} \sum_{\mathbf{m} \in \mathbb{Z}^d} \hat{V}_{\mathbf{Gm}} c_{\mathbf{k}+\mathbf{Gm},\sigma}^\dagger c_{\mathbf{k},\sigma}, \quad (14)$$

we see that the Hamiltonian only couples states whose wave-vectors differ by reciprocal lattice vectors (for $\mathbf{m} \neq \mathbf{0}$ they are Umklapp processes). Thus, H_{single} is block-diagonal in k -space so that its eigenstates are of the form

$$\varphi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{m} \in \mathbb{Z}^d} c_{n,\mathbf{m}} e^{i(\mathbf{k}+\mathbf{Gm})\cdot\mathbf{r}}, \quad (15)$$

where \mathbf{k} , now restricted to the primitive reciprocal cell, is the crystal momentum of the state and n its band index. Under translations by a lattice vector $\mathbf{A}\mathbf{n}$ they transform as

$$\varphi_{n,\mathbf{k}}(\mathbf{r} + \mathbf{A}\mathbf{n}) = e^{i\mathbf{k}\cdot\mathbf{A}\mathbf{n}} \varphi_{n,\mathbf{k}}(\mathbf{r}), \quad (16)$$

i.e., as irreducible representations of the abelian translation group. This is the Bloch theorem. Transforming back to real space, we could determine the eigenfunctions $\varphi_{n,\mathbf{k}}(\mathbf{r})$ by solving the eigenvalue problem for H_{single}

$$\left(\frac{1}{2} \nabla_{\mathbf{r}}^2 + V_{\text{ext}}(\mathbf{r}) \right) \varphi_{n,\mathbf{k}}(\mathbf{r}) = \varepsilon_{n,\mathbf{k}} \varphi_{n,\mathbf{k}}(\mathbf{r}) \quad (17)$$

not on the entire space \mathbb{R}^d , but on a single primitive lattice cell imposing \mathbf{k} -boundary conditions $\varphi_{n,\mathbf{k}}(\mathbf{a}_i) = e^{i\mathbf{k}\cdot\mathbf{a}_i} \varphi_{n,\mathbf{k}}(\mathbf{0})$. It is, however, more common to rewrite (15) as

$$\varphi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{m}} c_{n,\mathbf{m}} e^{i\mathbf{G}\mathbf{m}\cdot\mathbf{r}} = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}). \quad (18)$$

with the lattice-periodic Bloch function $u_{n,\mathbf{k}}(\mathbf{r})$. Using this form as an ansatz in (17), we see that the Bloch functions can be obtained from the eigenvalue problem

$$\left(\frac{1}{2} (-i\nabla_{\mathbf{r}} + \mathbf{k})^2 + V(\mathbf{r}) \right) u_{n,\mathbf{k}}(\mathbf{r}) = \varepsilon_{n,\mathbf{k}} u_{n,\mathbf{k}}(\mathbf{r}), \quad (19)$$

on a primitive cell of \mathcal{L} with periodic boundary conditions. We note that in this equation \mathbf{k} plays the role of a constant vector potential.

For a general time-reversal-symmetric Hamiltonian, every eigenfunction $\varphi_{\alpha}(\mathbf{r})$ is degenerate with its complex conjugate, so that we can choose real eigenfunctions. Taking the complex conjugate of (19) shows that for a real potential we can choose $\overline{u_{n,\mathbf{k}}(\mathbf{r})} = u_{n,-\mathbf{k}}(\mathbf{r})$. When the potential is inversion-symmetric, $V(-\mathbf{r}) = V(\mathbf{r})$, then $\varphi_{\alpha}(-\mathbf{r})$ is degenerate with $\varphi_{\alpha}(\mathbf{r})$. For (19) it implies $u_{n,\mathbf{k}}(-\mathbf{r}) = u_{n,-\mathbf{k}}(\mathbf{r})$. In the presence of both symmetries we obtain $\overline{u_{n,\mathbf{k}}(\mathbf{r})} = u_{n,-\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(-\mathbf{r})$.

A Bloch theorem also holds for the eigenstates of a many-body Hamiltonian that is invariant under lattice translations. Translating *all* electrons by the *same* lattice vector $\mathbf{A}\mathbf{n}$ will multiply the wave-function by a phase $e^{i\mathbf{k}_{\text{tot}}\cdot\mathbf{A}\mathbf{n}}$, where \mathbf{k}_{tot} is the *total* crystal momentum of the many-body state. Thus, in k -space the Hamiltonian block-diagonalizes into sectors with a given total crystal momentum. Writing H in k -space

$$H = \sum_{\mathbf{k},\sigma} \left(\frac{\mathbf{k}^2}{2} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k},\sigma} + \sum_{\mathbf{m}} \hat{V}_{\mathbf{G}\mathbf{m}} c_{\mathbf{k}+\mathbf{G}\mathbf{m},\sigma}^\dagger c_{\mathbf{k},\sigma} + \frac{1}{2} \sum_{\mathbf{k}',\sigma';\mathbf{q}} c_{\mathbf{k}+\mathbf{q},\sigma}^\dagger c_{\mathbf{k}'-\mathbf{q},\sigma'}^\dagger \frac{1}{|\mathbf{q}|^2} c_{\mathbf{k}',\sigma'} c_{\mathbf{k},\sigma} \right) \quad (20)$$

we see that acting on a Slater determinant of plane waves with momenta \mathbf{k}_i (or Bloch waves of crystal momenta \mathbf{k}_i), the Hamiltonian does not change total crystal momentum $\mathbf{k}_{\text{tot}} = \sum \mathbf{k}_i$. However, while the kinetic energy is diagonal and the external potential scatters only between plane waves differing by a reciprocal lattice vector, the electron-electron interaction scatters

plane waves of arbitrary single-electron momentum. Thus, for the eigenvalue problem we have to consider Slater determinants of plane waves with arbitrary wave-vectors \mathbf{k}_i .

As the simplest example, let us consider the Slater determinant of two plane waves

$$\Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \left(\frac{1}{(2\pi)^{d/2}} \right)^2 \begin{vmatrix} e^{i\mathbf{k}_1 \cdot \mathbf{r}_1} & e^{i\mathbf{k}_2 \cdot \mathbf{r}_1} \\ e^{i\mathbf{k}_1 \cdot \mathbf{r}_2} & e^{i\mathbf{k}_2 \cdot \mathbf{r}_2} \end{vmatrix} \propto e^{i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2)} - e^{i(\mathbf{k}_2 \cdot \mathbf{r}_1 + \mathbf{k}_1 \cdot \mathbf{r}_2)} \quad (21)$$

By construction it transforms as desired under a shift of *all* electrons by a lattice vector

$$\Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1 + \mathbf{A}\mathbf{n}, \mathbf{r}_2 + \mathbf{A}\mathbf{n}) = e^{(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{A}\mathbf{n}} \Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2). \quad (22)$$

The situation is, however, markedly different from the single-electron case (15): there, we can always translate a single electron coordinate into a primitive cell, allowing us to consider the single-electron Bloch functions on a finite volume. For more electrons, however, their relative distance is unchanged under the collective translation, so that we cannot bring all coordinates into a finite volume. To make this possible, we would need a Bloch-type theorem for translations of *individual* electrons:

$$\Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1 + \mathbf{c}, \mathbf{r}_2) = e^{i\tilde{\mathbf{k}} \cdot \mathbf{c}} \Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{r}_1, \mathbf{r}_2) \quad (23)$$

(the equivalent equation for translations of \mathbf{r}_2 follows from the antisymmetry). Inserting (21) we see that such an individual-electron Bloch-condition puts constraints on the allowed momenta: $\exp(i(\mathbf{k}_i - \tilde{\mathbf{k}}) \cdot \mathbf{c}) = 1$. Thus, to be able to restrict the many-electron wavefunction to a primitive cell spanned by three vectors \mathbf{C} with boundary conditions

$$\Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{c}_i, \mathbf{r}_2, \dots) = e^{i\tilde{\mathbf{k}} \cdot \mathbf{c}_i} \Phi_{\mathbf{k}_1, \mathbf{k}_2}(\mathbf{0}, \mathbf{r}_2, \dots), \quad (24)$$

we can only allow Slater determinants constructed from plane waves with wave vectors \mathbf{k}_i such that the $\mathbf{k}_i - \tilde{\mathbf{k}}$ are reciprocal lattice vectors of \mathbf{C} . For $\tilde{\mathbf{k}} = \mathbf{0}$ we obtain the simulation cell with periodic boundary conditions discussed in the introduction. The eigenfunctions of H with boundary conditions (24) can be written in the Bloch-like form

$$\Psi_{n, \tilde{\mathbf{k}}}^{\mathbf{C}}(\mathbf{r}_1, \mathbf{r}_2, \dots) = e^{i\tilde{\mathbf{k}} \cdot \sum_i \mathbf{r}_i} U_{n, \tilde{\mathbf{k}}}^{\mathbf{C}}(\mathbf{r}_1, \mathbf{r}_2, \dots), \quad (25)$$

where $U_{n, \tilde{\mathbf{k}}}^{\mathbf{C}}(\mathbf{r}_1, \mathbf{r}_2, \dots)$ is invariant under translations of a single electron by a vector from \mathbf{C} and antisymmetric under particle exchange. While $U_{n, \tilde{\mathbf{k}}}^{\mathbf{C}}$ apparently is the many-body generalization of the single-electron Bloch function $u_{n, \mathbf{k}}(\mathbf{r})$, we have to keep in mind that its construction is based on the artificial boundary conditions (24), which depend on the choice of the volume \mathbf{C} . When the volume is chosen small, calculations are simple but the wave functions $\Psi_{n, \tilde{\mathbf{k}}}^{\mathbf{C}}$ will give poor approximations to the actual ground state, while increasing the cell improves the accuracy but also makes calculations increasingly difficult.

Physically, as in the single-electron case (19), the twisted boundary conditions (24) correspond to a constant vector potential. The dependence of the ground state energy $E_0^{\mathbf{C}}(\tilde{\mathbf{k}})$ can be used to distinguish metallic from (Mott) insulating systems: the second derivative at $\tilde{\mathbf{k}} = 0$ of the energy with respect to the current driving vector potential gives the static response. For metals it stays finite while for insulators it vanishes in the thermodynamic limit [6].

1.2 Superlattices

For a system with lattice periodicity the single- and many-electron boundary conditions (16) and (24) should be consistent. This implies, in particular, that the vectors spanning the cell should be integer linear-combinations of the primitive lattice vectors, $\mathbf{c}_i = \sum \mathbf{a}_i l_{ij}$, or, in matrix notation, $\mathbf{C} = \mathbf{A}\mathbf{L}$, where the columns of \mathbf{L} are the primitive cell vectors in units of the primitive lattice vectors. The vectors \mathbf{C} span a lattice $\mathcal{S} \subseteq \mathcal{L}$, called a superlattice. The volume of the primitive unit cell of \mathbf{C} is $|\det(\mathbf{L})|$ times the volume of the primitive lattice cell. Since \mathbf{L} is an integer matrix, its determinant is also an integer.

As the choice of the primitive lattice vectors \mathbf{A} for a given lattice \mathcal{L} is not unique, so is the choice of \mathbf{L} for a given superlattice \mathcal{S} . We can, however, easily check whether, for given primitive lattice vectors \mathbf{A} , two integer matrices span the same superlattice by reducing them to their Hermite normal form [7] and checking if they agree. The reduction of a non-singular integer matrix \mathbf{L} to its *Hermite normal form* (HNF)

$$\Lambda = \begin{pmatrix} \lambda_{11} & 0 & 0 & \cdots \\ \lambda_{21} & \lambda_{22} & 0 & \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \\ \vdots & & & \ddots \end{pmatrix} \quad (26)$$

with $\lambda_{ii} \geq 1$, and $\lambda_{ij} > \lambda_{ij} \geq 0$ can be done recursively. Allowed operations that leave the superlattice spanned by the transformed matrix unchanged are (i) multiplying a column by ± 1 , (ii) exchanging columns, and (iii) adding an integer multiple of another column. The reduction algorithm [8] is based on the Euclidean algorithm for finding the greatest common divisor

$$\gcd(a, b) = \begin{cases} |a| & \text{if } b = 0 \quad (\text{operation of type (i)}) \\ \gcd(b, a) & \text{if } |a| < |b| \quad (\text{operation of type (ii)}) \\ \gcd(a - \lfloor a/b \rfloor b, b) & \text{otherwise} \quad (\text{operation of type (iii)}) \end{cases} \quad (27)$$

where a and b are the matrix elements in a given row of the matrix, and we perform the operations not just on these matrix elements but on their entire column vectors. To reduce the first row of \mathbf{L} to the required form, we apply the Euclidean algorithm on the last two columns, reducing the coefficient in the last column to zero. In this way we reduce all matrix elements except the first to zero, obtaining $\lambda_{11} = \gcd(l_{11}, \dots, l_{1d})$. We iterate this procedure for the submatrix obtained by removing the first row and column to bring the matrix to lower triangular form. Finally, we use column operations to replace the off-diagonal elements in a row by their remainders on division by the corresponding diagonal element.

Besides giving a criterion for determining equivalent primitive cell vectors, the Hermite normal form gives a prescription for enumerating all non-equivalent periodic clusters of a given size. We have to be careful, however, when the lattice \mathcal{L} spanned by \mathbf{A} has point symmetries besides simple inversion. Since the construction of Λ contains no information on the underlying lattice, such additional symmetries can render superlattices with different HNF equivalent.

Fig. 2 gives an example of how different equivalent primitive cells can appear. For the square lattice with point-symmetry C_{4v} , we see that primitive vectors spanning the same superlattice

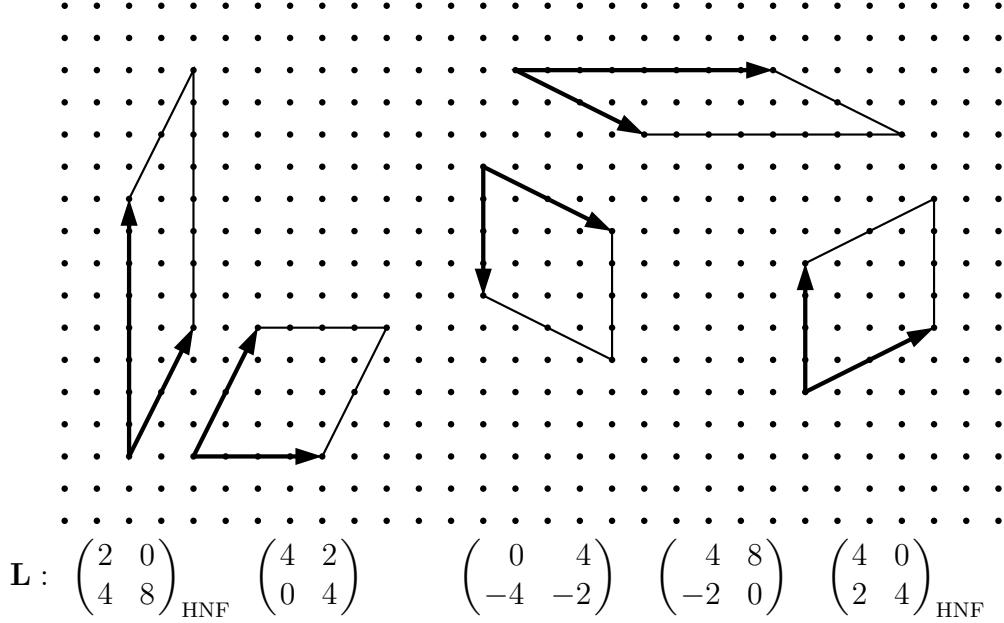


Fig. 2: Primitive superlattice cells on a square lattice. All primitive superlattice vectors span the same superlattice, i.e., calculations for any of the shown cells will give the same results. The primitive superlattice vectors \mathbf{L} are shown below the plot. The first two plots show a set of primitive vectors and the corresponding Hermite normal form. The next two are obtained by rotating the vectors by $-\pi/2$, $(x, y) \rightarrow (-y, x)$. Because of the point-symmetry of the underlying square lattice, they span equivalent superlattices. Their Hermite normal form, shown on the far right, is, however, different from the unrotated HNF on the far left.

can have different Hermite normal form when they are related by a non-trivial point-symmetry operation. This happens when the primitive cells break the point symmetry of the lattice. It is often advantageous to work with cells that retain symmetries of the underlying lattice. Well known examples are the conventional cells of cubic lattices, e.g., the face-centered cubic lattice

$$\mathbf{A} = \frac{a}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \mathbf{C} = a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}. \quad (28)$$

While the primitive lattice vectors \mathbf{A} do not exhibit the cubic symmetry of the lattice, the vectors of the conventional unit cell \mathbf{C} do. Symmetry is one of the criteria used for selecting cells that best represent the infinite system [9].

The reciprocal lattice \mathcal{R}_S of S is spanned by the primitive vectors $\mathbf{K}_s = (2\pi\mathbf{C}^{-1})^T$. They can be written in terms of the reciprocal lattice vectors of \mathcal{L} as

$$\mathbf{K}_s = (2\pi\mathbf{C}^{-1})^T = \mathbf{K} (\mathbf{L}^{-1})^T \quad (29)$$

The primitive cell spanned by \mathbf{K}_s is smaller than that spanned by \mathbf{K} by a factor of $1/|\det(\mathbf{L})|$. The Slater determinants in a simulation on the primitive cell spanned by \mathbf{C} with periodic boundary conditions can then contain Bloch waves of wave vector $\mathbf{k} = \mathbf{K}_s \mathbf{m}$ with $\mathbf{m} \in \mathbb{Z}^d$ that fall in the Brillouin zone of the original lattice. With twisted boundary conditions (24), the allowed

wave vectors are shifted by $\tilde{\mathbf{k}}$. The choice $\tilde{\mathbf{k}} = \mathbf{k}_i/2$ corresponds to a sign-change under a translation by \mathbf{c}_i , i.e., antiperiodic boundary conditions in that direction.

When the primitive superlattice vectors are chosen as integer multiples of the primitive lattice vectors, $\mathbf{c}_i = n_i \mathbf{a}_i$, the reciprocal lattice shifted by $\tilde{\mathbf{k}} = \sum_i (n_i - 1) \mathbf{k}_i / 2n_i$ forms a Monkhorst-Pack grid of special k -points that are popular for Brillouin-zone integrations [10].

2 Variational methods

Conceptually, the variational approach is straightforward: to find the ground state of a Hamiltonian H , just minimize the energy expectation value

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (30)$$

The practical problem is, of course, the choice of a suitable variational space. The systematic approach is to write the trial wave function as a linear combination of Slater determinants $\Psi(\mathbf{r}) = \sum_{\alpha} c_{\alpha} \Phi_{\alpha}(\mathbf{r})$ and allow all amplitudes c_{α} to vary. For a finite system with N electrons and a finite basis set of K orbitals there will be $\binom{K}{N}$ Slater determinants. Minimizing $E[\Psi] = E(c_1, c_2, \dots)$ amounts then to a high-dimensional optimization problem. As $E(c_1, c_2, \dots)$ has no local minima, this can be done using a steepest descent method, e.g., the Lanczos method [11]. It involves the repeated application of the Hamiltonian to the trial function. When working with a basis set, a Hamiltonian (2) with pair interaction only couples Slater determinants that differ in at most two orbitals. Thus, the matrix representation of H in Slater-determinant space is reasonably sparse so that the matrix-vector product can be efficiently calculated. Nevertheless this method, called *configuration interaction (CI)* as it describes the interplay of Slater determinants (electron configurations), is limited to quite small systems by the sheer number of Slater determinants spanning the Hilbert space, or, equivalently, by the number of parameters c_{α} that need to be simultaneously optimized: For a system with 25 electrons and just 50 basis functions, the number of parameter is already above 10^{14} , i.e., requiring a peta byte of memory just for storing the parameters c_{α} . A way out might be to consider only “important” Slater determinants. It turns out that the variational energy converges, however, only slowly with the number of determinants included in the calculation. Moreover, when we want to study systems of increasing size a truncated CI easily leads to size-consistency problems [2].

An alternative to the full-CI ansatz are wave functions that capture the strongest effects of electron correlation with only a small number of parameters. To identify the major effect of electron correlation on the wave function, we return to the energy expectation value (30). Considered as a wave function functional, the stationarity condition

$$0 = \frac{\delta E}{\delta \Psi} = \frac{H|\Psi\rangle}{\langle\Psi|\Psi\rangle} - \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle^2} |\Psi\rangle \quad (31)$$

is equivalent to the Schrödinger equation $H\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$. Dividing by the wave function we obtain the *local energy*

$$E_{\text{loc}}(\mathbf{r}) = \frac{H\Psi(\mathbf{r})}{\Psi(\mathbf{r})}, \quad (32)$$

which is constant for eigenstates of H , i.e., its variance is zero (*zero variance property*). We can thus find eigenstates by minimizing the variance of the local energy (*variance minimization*)

$$\sigma^2[\Psi] = \int |E_{\text{loc}}(\mathbf{r})|^2 |\Psi(\mathbf{r})|^2 d\mathbf{r} - \left(\int E_{\text{loc}}(\mathbf{r}) |\Psi(\mathbf{r})|^2 d\mathbf{r} \right)^2 = \langle \Psi | H^2 | \Psi \rangle - \langle \Psi | H | \Psi \rangle^2. \quad (33)$$

This approach can also be employed for constructing good trial wave functions.

We might think that, as the solution of a second-order differential equation, wave functions are smooth with continuous first derivative. This is, however, not true at singularities in the potential. A well-known example is the hydrogen atom. Its ground state fulfills

$$H\varphi_{1s}(\vec{r}) = -\frac{1}{2} \vec{\nabla}^2 \varphi_{1s}(\vec{r}) - \frac{1}{r} \varphi_{1s}(\vec{r}) = E_{1s} \varphi_{1s}(\vec{r}). \quad (34)$$

For $r \rightarrow 0$, the potential energy diverges while $H\varphi_{1s}(\vec{r})$ remains finite. For this reason, the $1s$ function goes to a finite value at the position of the nucleus, producing a *cusp* \wedge , i.e., a discontinuity in the first derivative, which gives rise to the canceling divergence in the kinetic energy. The cancellation condition determining the cusp in the wave function, in the case of hydrogen $\varphi_{1s} \sim \exp(-\sqrt{x^2 + y^2 + z^2})$, is called the *cusp condition* [12]. Removing divergences in the local energy by implementing the cusp condition is the most important step towards reducing the variance of $E_{\text{loc}}(\mathbf{r})$, i.e., constructing good variational wave functions.

The cusps at the position of the nuclei are built into the single-particle orbitals obtained from a mean-field solution of (1). The many-body eigenstates will, however, also have cusps when two electrons meet ($\vec{r}_i \rightarrow \vec{r}_j$). These cusps are, of course, not easily reproduced by a linear combination of Slater determinants, which explains the slow convergence of CI expansions.

To derive the cusp conditions, we start with the electron-nucleus cusp. Following the example of hydrogen, we write the wave function close to a nucleus of charge Z as $\exp(-u_Z(r))$, where r is the distance of the electron from the nucleus. For r close to zero the local energy is

$$E_{\text{loc}}(r) = -\frac{1}{2} \frac{\left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) e^{-u_Z(r)}}{e^{-u_Z(r)}} - \frac{Z}{r} = -\frac{-u_Z''(r) + u_Z'^2(r) - \frac{2u_Z'(r)}{r}}{2} - \frac{Z}{r}. \quad (35)$$

Thus, it stays finite for $r \rightarrow 0$ when

$$\left. \frac{du_Z}{dr} \right|_{r=0} = +Z. \quad (36)$$

For electrons, the cusp condition will depend on their relative spin orientation. Electrons with opposite spin need not be antisymmetrized so that we can write the wave function when the electrons are close to each other as $\exp(-u_{\sigma,-\sigma}(r))$ with $r = |\mathbf{r}_1 - \mathbf{r}_2|$. The situation is almost the same as for the electron-nucleon cusp, except that (i) electrons repel each other and (ii) we now have two electronic degrees of freedom, i.e., we get contributions from the kinetic energy operator for both electrons, resulting in

$$\left. \frac{du_{\sigma,-\sigma}}{dr} \right|_{r=0} = -\frac{1}{2}. \quad (37)$$

For electrons with parallel spin, the wave function must be antisymmetric in the electron coordinates, i.e., there must be a *nodal surface* separating the region where the wave function is positive from that where it is negative. For $\mathbf{r}_1 \approx \mathbf{r}_2$ we can approximate the nodal surface by $\mathbf{a} \cdot \mathbf{r} = 0$, which is a plane when we keep one of the electron coordinates fixed. We can then write the antisymmetric wave function close to $r = 0$ as $\mathbf{a} \cdot \mathbf{r} \exp(-u_{\sigma,\sigma}(r))$. Removing the singularity in the local energy now requires [4]

$$\left. \frac{du_{\sigma,\sigma}}{dr} \right|_{r=0} = -\frac{1}{4}. \quad (38)$$

Thus, the correlation cusp for opposite-spin electrons \swarrow affects the wave function more than that for parallel spins \smile since electrons with the same spin already tend to avoid each other as a consequence of exchange.

The cusp conditions just tell us the form of the wave function right at the singularity. To put this information into a usable wave function, we have to parametrize the electron-electron functions $u_{\sigma,\sigma'}(r)$ for finite r . This is typically done by writing $u(r)$ as a rational function that fulfills the cusp condition for $r \rightarrow 0$ and goes to a constant for $r \rightarrow \infty$. To ensure antisymmetry and the electron-nucleus cusps, the electron-electron correlators are multiplied onto a Slater determinant. This gives the *Jastrow wave function* [14]

$$\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N) = \Phi(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N) \prod_{i < j} e^{-u_{\sigma_i, \sigma_j}(r_{ij})}. \quad (39)$$

The product of pair functions is called the *Jastrow factor*. It will tend to reduce the amplitude of the Slater determinant when electrons come close to each other, i.e., it introduces a correlation hole. For systems with inhomogeneous charge density this means that the Jastrow factor pushes electrons away from regions of high charge-density, where the probability of two electrons approaching each other is largest. This can be compensated by introducing single-electron terms in the Jastrow factor [15]. That is particularly important when the Slater determinant used in (39) accurately describes the charge density of the system, e.g., from density-functional theory.

Having chosen parametrizations for $u_{\sigma,\sigma'}(r)$ and the single-electron term, the variational approach looks straightforward: just minimize the energy expectation value $\langle \Psi_J | H | \Psi_J \rangle / \langle \Psi_J | \Psi_J \rangle$ with respect to the (relatively few) Jastrow parameters. The pair functions, however, make it impossible to evaluate the expectation value other than by integrating over all electron configurations

$$\frac{\langle \Psi_J | H | \Psi_J \rangle}{\langle \Psi_J | \Psi_J \rangle} = \frac{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N \overline{\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)} H \Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)}{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N |\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)|^2} \quad (40)$$

This $3N$ -dimensional integral is best done using stochastic sampling – variational Monte Carlo. What improvements in energy can we expect? Typically, optimizing a good trial function will lower the energy expectation value by only a few percent of the energy calculated with just the mean-field Slater determinant. This might seem little reward for the considerable effort. We

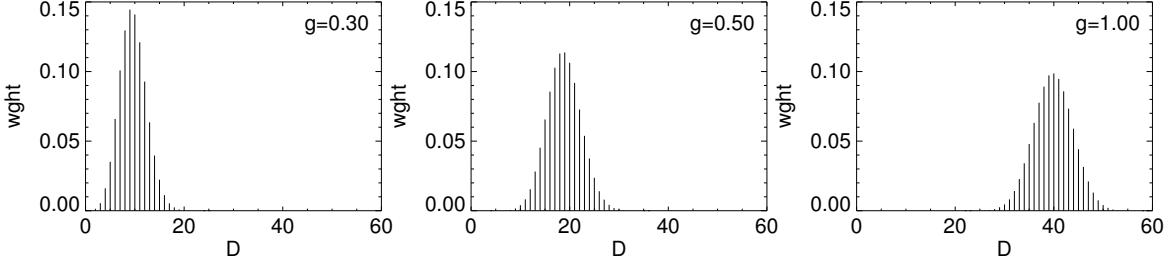


Fig. 3: Weight of configurations with given number D of double occupancies for Gutzwiller wave functions $\Psi_T(R) = g^{D(R)} \Phi(R)$. Reducing the Gutzwiller factor g suppresses configurations with high Coulomb energy $E_{\text{Coul}}(R) = U D(R)$ at the expense of increasing the kinetic energy. The results shown are for a Hubbard model on a square lattice of 16×16 sites with periodic boundary conditions and 101 electrons of each spin. For the uncorrelated Slater determinant ($g = 1$) the distribution is centered around $D = N_{\text{site}} n_{\uparrow} n_{\downarrow} = 16^2 (101/16^2)^2 \approx 40$.

have, however, to keep in mind that the effects of correlation, essentially the integral over the correlation-hole, are very small compared to the Coulomb energy of the uncorrelated charge density (Hartree energy) and the kinetic term. So correlation effects are barely noticeable on the scale of the total energy. The dominating role of the Hartree energy also becomes apparent when comparing the charge density of a solid to a simple superposition of atomic charge densities: the bonding induces barely noticeable changes, see, e.g., Fig. 1 in [16]. A more sensible benchmark than the change in total energy is how much of the *correlation energy*, i.e., the difference between the Hartree-Fock and the exact energy, is captured. On this count, variational wave functions fare much better: they typically recover roughly 90% of the correlation energy. Still, correlations are a subtle effect also on this energy scale, and variational methods are usually not sufficient for reaching chemical accuracy.

When working in second quantization with a finite (and therefore incomplete) basis set, we cannot describe two electrons coming arbitrarily close to each other. So there are no cusp conditions here. Still, we can correlate the electrons in the orbitals of the basis set by introducing correlation factors of the form $\exp(-\eta D)$, where D is a two-body operator. Such trial wave functions, with $D = \sum_i n_{i\uparrow} n_{i\downarrow}$ the number of doubly occupied sites, were introduced by Gutzwiller [17] as variational states for the Hubbard model

$$H = -t \sum_{ij,\sigma} c_{j\sigma}^\dagger c_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (41)$$

The *Gutzwiller wave function (GWF)*, with $|\Phi\rangle$ a mean-field solution of H , can be written as

$$|\Psi_T\rangle = e^{-\eta \sum_i n_{i\uparrow} n_{i\downarrow}} |\Phi\rangle = g^{\sum_i n_{i\uparrow} n_{i\downarrow}} |\Phi\rangle = \prod_i (1 - (1-g) n_{i\uparrow} n_{i\downarrow}) |\Phi\rangle, \quad (42)$$

where $g = \exp(-\eta)$ and the final equality arises from the fact that $n_{i\uparrow} n_{i\downarrow}$ can only take the values 0 or 1. The role of the Gutzwiller parameter $g \in [0, 1]$ is to reduce the number of doubly occupied sites relative to the mean-field solution $|\Phi\rangle$, thus reducing the Coulomb repulsion at the expense of increasing the kinetic energy. This is illustrated in Fig. 3. Introducing more general two-body operators D , Gutzwiller wave functions can be devised for realistic multi-band models. See [18] for a nice introduction.

2.1 Variational Monte Carlo

As seen in (40), evaluating the energy expectation value for a Jastrow wave function involves the integration over the $3N$ -dimensional configuration space of the electrons. The key for doing this using stochastic sampling is again the local energy, which allows us to rewrite (40) as

$$\frac{\langle \Psi_J | H | \Psi_J \rangle}{\langle \Psi_J | \Psi_J \rangle} = \frac{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N E_{\text{loc}}(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N) |\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)|^2}{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N |\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)|^2}. \quad (43)$$

As it is non-negative and normalized,

$$p(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N) = \frac{|\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)|^2}{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N |\Psi_J(\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)|^2} \quad (44)$$

is a probability distribution function on the configuration space, so that we can evaluate (43) by sampling configurations $R = (\mathbf{r}_1, \sigma_1; \dots; \mathbf{r}_N, \sigma_N)$ with probability $p(R)$ and average the corresponding local energy $E_{\text{loc}}(R)$.

The same approach works for Hamiltonians written in second quantization, the main difference being that in this case the electron configurations are discrete, specifying the occupation of the orbitals used in second quantization. In the following, we specialize to the case of the simple Hubbard model (41) with one orbital per site. Denoting by R an electron configuration, specifying on which site the electrons are located as well as their spin, we can write the energy expectation value of a trial function Ψ_T as

$$E_T = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \frac{\sum_R E_{\text{loc}}(R) \Psi_T^2(R)}{\sum_R \Psi_T^2(R)}, \quad (45)$$

with the local energy

$$E_{\text{loc}}(R) = \sum_{R'} \frac{\langle \Psi_T | R' \rangle \langle R' | H | R \rangle}{\langle \Psi_T | R \rangle} = \sum_{R' \neq R} t \frac{\Psi_T(R')}{\Psi_T(R)} + U D(R). \quad (46)$$

If the Hamiltonian allows only hopping to near neighbors, the sum over R' in the local energy scales with the number of near neighbors times the number of electrons in the system. In contrast, the sum over R in (45) is over *all* configurations, i.e., of the order of the dimension of the Hilbert space. With increasing system size this rapidly becomes extremely large. To give an impression, the dimension of the Hilbert space for the model shown in Fig. 3 is $\binom{16^2}{101} \times \binom{16^2}{101}$, which is larger than 10^{146} . So it seems quite impossible to do the sum in (45). Even generating configurations at a rate of 3.3 GHz, we could visit just 10^{17} configurations per year. It is the magic of stochastic methods that sums over such spaces can still be done to an astonishing accuracy.

The idea of *variational Monte Carlo* [19, 20] is to perform a random walk in the space of configurations, with transition probabilities $p(R \rightarrow R')$ chosen such that the configurations R_{VMC} in the random walk have the probability distribution function $\Psi_T^2(R)$. Then

$$E_{\text{VMC}} = \frac{\sum_{R_{\text{VMC}}} E_{\text{loc}}(R_{\text{VMC}})}{\sum_{R_{\text{VMC}}} 1} \approx \frac{\sum_R E_{\text{loc}}(R) \Psi_T^2(R)}{\sum_R \Psi_T^2(R)} = E_T. \quad (47)$$

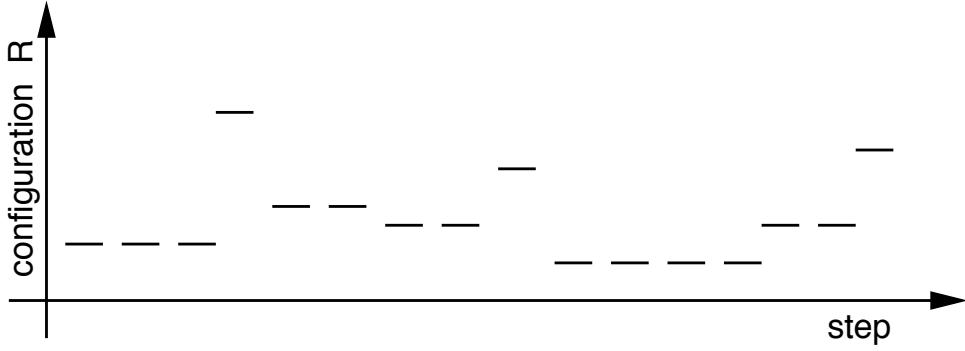


Fig. 4: Illustration of the random walk in configuration space.

The transition probabilities can be determined from detailed balance

$$\Psi_T^2(R) p(R \rightarrow R') = \Psi_T^2(R') p(R' \rightarrow R), \quad (48)$$

which gives $p(R \rightarrow R') = 1/N \min(1, \Psi_T^2(R')/\Psi_T^2(R))$, with N being the maximum number of possible transitions. It is sufficient to consider only transitions between configurations that are connected by the Hamiltonian, i.e., transitions in which one electron hops to a neighboring site. The standard prescription is then to propose a transition $R \rightarrow R'$ with probability $1/N$ and accept it with probability $\min(1, \Psi_T^2(R')/\Psi_T^2(R))$. This works well when U is not too large. For strongly correlated systems, however, the random walk will stay for long times in configurations with a small number of double occupancies $D(R)$, since most of the proposed moves will increase D and hence be rejected with probability $\approx 1 - g^{D(R')-D(R)}$.

Fortunately, there is a way to integrate-out the time the walk stays in a given configuration [21]. To see how, we first observe that for the local energy (46) the ratio of the wave functions for all transitions induced by the Hamiltonian have to be calculated. This in turn means that we also know all transition probabilities $p(R \rightarrow R')$. We can therefore eliminate any rejection, i.e., accept with probability one, by proposing moves $R \rightarrow R'$, ($R' \neq R$), with probabilities

$$\tilde{p}(R \rightarrow R') = \frac{p(R \rightarrow R')}{\sum_{R'} p(R \rightarrow R')} = \frac{p(R \rightarrow R')}{1 - p_{\text{stay}}(R)}. \quad (49)$$

Checking detailed balance (48) we find that now we are sampling configurations \tilde{R}_{VMC} from the probability distribution function $\Psi_T^2(R) (1 - p_{\text{stay}}(R))$. To compensate for this, we assign a weight $w(R) = 1/(1 - p_{\text{stay}}(R))$ to each configuration R . The energy expectation value is then given by

$$E_T \approx \frac{\sum_{\tilde{R}_{\text{VMC}}} w(\tilde{R}_{\text{VMC}}) E_{\text{loc}}(\tilde{R}_{\text{VMC}})}{\sum_{\tilde{R}_{\text{VMC}}} w(\tilde{R}_{\text{VMC}})}. \quad (50)$$

The above method is very efficient since it ensures that in every Monte Carlo step a new configuration is created. Instead of staying in a configuration where Ψ_T is large, this configuration is weighted with the expectation value of the number of steps the simple Metropolis algorithm would stay there. This is particularly convenient for simulations of systems with strong correlations: instead of having to do longer and longer runs as U is increased, the above method produces, for a fixed number of Monte Carlo steps, results with comparable error estimates.

2.2 Correlated sampling

The essence of the variational method is the minimization of the energy expectation value (45) as a function of the variational parameters in the trial function. To this end, we could simply perform independent VMC calculations for a set of different parameters. It is, however, difficult to compare the energies from independent calculations since each VMC result comes with its own statistical errors. This problem can be avoided with *correlated sampling* [19, 22]. The idea is to use the same random walk in calculating the expectation value for different trial functions. This reduces the relative errors and hence makes it easier to find the minimum.

Let us assume that we have generated a random walk $\{R_{\text{VMC}}\}$ for the trial function Ψ_T . Using the *same* random walk, we can also estimate the energy expectation value (47) for a different trial function $\tilde{\Psi}_T$. To do so we have to compensate for the fact that the configurations have the probability distribution Ψ_T^2 instead of $\tilde{\Psi}_T^2$ by introducing reweighting factors

$$\tilde{E}_T \approx \frac{\sum_{R_{\text{VMC}}} \tilde{E}_{\text{loc}}(R) \tilde{\Psi}_T^2(R) / \Psi_T^2(R)}{\sum_{R_{\text{VMC}}} \tilde{\Psi}_T^2(R) / \Psi_T^2(R)}. \quad (51)$$

Likewise, (50) is reweighted into

$$\tilde{E}_T \approx \frac{\sum_{\bar{R}_{\text{VMC}}} w(\bar{R}) \tilde{E}_{\text{loc}}(\bar{R}) \tilde{\Psi}_T^2(\bar{R}) / \Psi_T^2(\bar{R})}{\sum_{\bar{R}_{\text{VMC}}} w(\bar{R}) \tilde{\Psi}_T^2(\bar{R}) / \Psi_T^2(\bar{R})}. \quad (52)$$

Also, the local energy $\tilde{E}_{\text{loc}}(R)$ can be rewritten such that the new trial function appears only in ratios with the old one. For Gutzwiller functions this implies a drastic simplification. Since they differ only in the Gutzwiller factor, the Slater determinants cancel, leaving only powers $(\tilde{g}/g)^{D(R)}$

$$E_T(\tilde{g}) \approx \frac{\sum_{R_{\text{VMC}}} \tilde{E}_{\text{loc}}(R) (\tilde{g}/g)^{2D(R)}}{\sum_{R_{\text{VMC}}} (\tilde{g}/g)^{2D(R)}} \quad (53)$$

and

$$\tilde{E}_{\text{loc}}(R) = -t \sum_{R' \neq R} (\tilde{g}/g)^{D(R')-D(R)} \frac{\Psi_T(R')}{\Psi_T(R)} + U D(R). \quad (54)$$

As the number of doubly occupied sites $D(R)$ for a configuration R is an *integer*, we can rearrange the sums in (53) and (54) into polynomials in \tilde{g}/g . The energy expectation value for any Gutzwiller parameter \tilde{g} is then given by a rational function in the variable \tilde{g}/g , where the coefficients only depend on the fixed trial function $|\Psi(g)\rangle$.

It is then clear how we proceed to optimize the Gutzwiller parameter in variational Monte Carlo [21]: we first pick a reasonable g and perform a VMC run for $|\Psi(g)\rangle$ during which we also estimate the coefficients of the above polynomials. We can then easily calculate $E_T(\tilde{g})$ by evaluating the rational function in \tilde{g}/g . Since the number of non-vanishing coefficients typically is only of the order of a few tens (see the distribution of weights shown in Figure 3), this is a very efficient process.

Figure 5 shows how the method works in practice. Although we deliberately picked a bad starting point, we still find the correct minimum. Of course, this will not be true for the whole

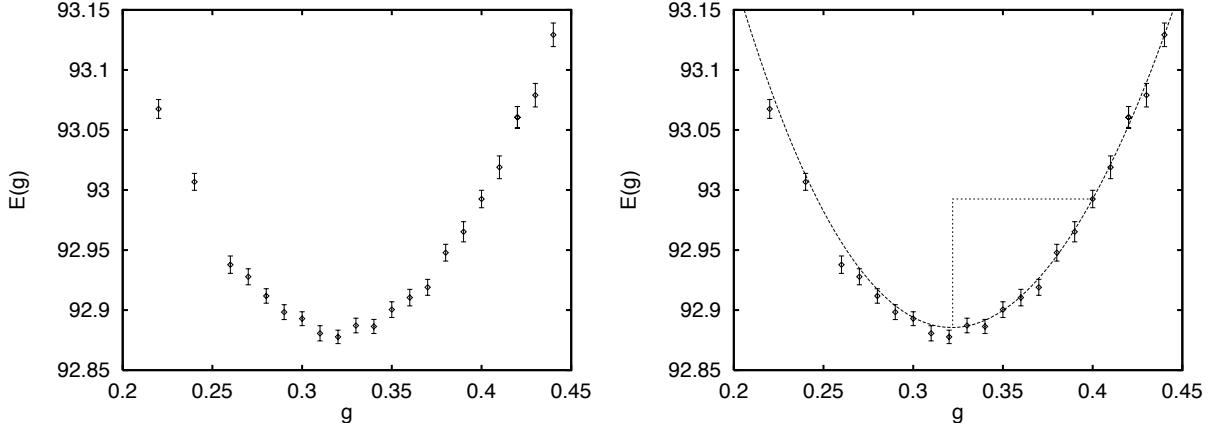


Fig. 5: Optimizing the Gutzwiller parameter g : The left-hand panel shows the straightforward approach of calculating the variational energy for a number of different values of g in separate VMC runs. The curve in the right-hand panel shows the result of a single correlated sampling run, calculated at $g = 0.4$. The predicted minimum is indicated by the dotted line and corresponds to the actual minimum. The calculations are for a cluster of 32 C_{60} molecules with 48+48 electrons (half-filling) and $U = 1.0$ eV [23].

range of Gutzwiller parameters. When \tilde{g} differs too much from g , the method breaks down. To understand this we again turn to Figure 3. We see that most configurations in a random walk generated with, say, $g = 0.50$ will have about 20 doubly occupied sites. In the Monte Carlo run we therefore sample the coefficients for $(\tilde{g}/g)^{2 \times 20}$ best, while the statistics for much larger or smaller powers is poor. But it is exactly these poorly sampled coefficients that we need for calculating the energy expectation value of trial functions with \tilde{g} much different from g . We can thus use the overlap of the wave functions $\langle \Psi(\tilde{g}) | \Psi(g) \rangle$ as a measure of the reliability of the calculated energy $E_T(\tilde{g})$. Like the energy expectation value itself, it can be recast in the form of polynomials, the coefficients of which can be sampled during the VMC run

$$\langle \Psi(\tilde{g}) | \Psi(g) \rangle = \frac{\sum_R \tilde{\Psi}(R) \Psi(R)}{\sqrt{\sum_R \tilde{\Psi}^2(R) \sum_R \Psi^2(R)}} = \frac{\sum_{R_{\text{VMC}}} (\tilde{g}/g)^{D(R)}}{\sqrt{\sum_{R_{\text{VMC}}} (\tilde{g}/g)^{2D(R)} \sum_{R_{\text{VMC}}} 1}}. \quad (55)$$

Figure 6 shows how the reliability of correlated sampling results depends on the overlap of with the trial function that is used in the VMC run.

There are some straightforward modifications of the scheme we have described above. Often, it is more appropriate to minimize the variance in the local energy $\sigma^2(g)$ rather than the energy $E(g)$ [22]. Since the variance can also be rewritten in terms of a rational function in \tilde{g}/g , variance minimization can be implemented in much the same way as the energy minimization that we have described here. Furthermore, it is clear that the method is not restricted to the plain Gutzwiller wave function but can be generalized to trial functions with more correlation factors of the type $r^{c(R)}$. As long as the correlation function $c(R)$ is integer-valued on the space of configurations, expectation values for such trial functions can still be rewritten as rational functions. The only difference to the simpler case described above is that now the rational function is multivariate, reflecting the fact that there is more than one variational parameter.

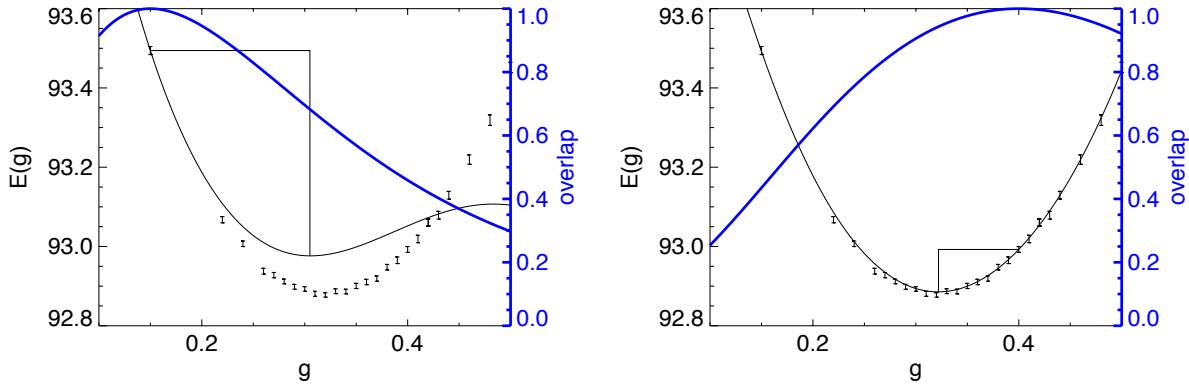


Fig. 6: The reliability of a correlated sampling run depends on the overlap of the trial function with the wave function at the predicted minimum. Left: inefficient sampling for small overlap. Right: reliable sampling for sufficiently large overlap. The thin line gives the energy curve obtained from the correlated sampling run, the thick line shows the overlap. The calculations are for the same system as in Figure 5.

2.3 Gutzwiller approximation

In Fig. 3 we found that we can estimate the number of doubly occupied sites in a non-interacting Slater determinant by simply assuming that the electrons of different spins are distributed uniformly over the lattice: $D \approx N_{\text{site}} n_{\uparrow} n_{\downarrow}$. In terms of electron configurations, this can be rephrased: all electron configurations R have the same weight in $|\Phi\rangle$, i.e., $|\Phi\rangle = \sum_R c_R |R\rangle$ with $|c_R|^2 = \text{const}$. This is the basic assumption of the Gutzwiller approximation (GA) [17, 24]. It provides surprisingly reliable estimates of the properties of the Gutzwiller wave function using simple combinatorics.

As electron correlation in the Hubbard model arises from the doubly occupied sites, it is reasonable to use the number of doubly occupied sites to characterize an electron configuration R . More specifically, we introduce the notation

- M : number of lattice sites ($= N_{\text{site}}$)
- N_σ : number of electrons with spin σ
- E : number of empty sites
- L_σ : number of sites with a single electron of spin σ
- D : number of doubly occupied sites

While the number of lattice sites and electrons is fixed for a given system, the other quantities have to fulfill physical constraints. A site is either empty, singly, or doubly occupied, i.e., $M = E + L_\uparrow + L_\downarrow + D$, and the electrons are on singly or doubly occupied sites, i.e., $N_\sigma = L_\sigma + D$. Given this notation, the number of configurations with a given D is obtained by distributing the empty, singly, and double occupied sites over the lattice

$$G(D) = \binom{M}{E} \binom{M-E}{L_\uparrow} \binom{M-E-L_\uparrow}{L_\downarrow} \binom{M-E-L_\uparrow-L_\downarrow}{D} = \frac{M!}{E! L_\uparrow! L_\downarrow! D!}. \quad (56)$$

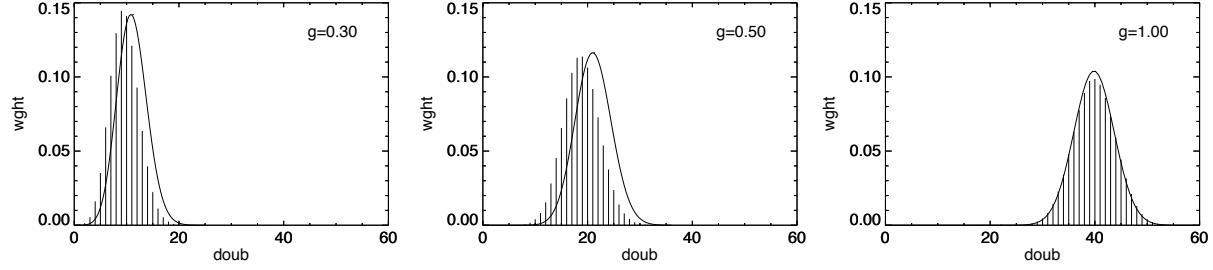


Fig. 7: Comparison of the weight of doubly occupied configurations for Gutzwiller wave functions (histograms) and as calculated in the Gutzwiller approximation (curves). The system is the same as in Figure 3.

Figure 7 shows that the distribution (56) matches that calculated for an uncorrelated Slater determinant almost perfectly. For a Gutzwiller wave function, the Gutzwiller parameter changes the GA-weights to $\sum_{R \text{ with } D \text{ double occs}} |c_R|^2 \propto G(D) g^{2D}$. Even for quite large correlation, i.e., small g , the agreement with the actual weights is surprisingly good. As we are dealing with quite large factorials, we can use Stirling's approximation $N! \sim \sqrt{2\pi N} \exp(N \ln(N) - N)$ to simplify (56)

$$\frac{G(D)}{G_{\text{tot}}} g^{2D} = \sqrt{\frac{n_\uparrow(1-n_\uparrow) n_\downarrow(1-n_\downarrow)}{2\pi M e l_\uparrow l_\downarrow d}} \left(\frac{n_\uparrow^{n_\uparrow} (1-n_\uparrow)^{1-n_\uparrow} n_\downarrow^{n_\downarrow} (1-n_\downarrow)^{1-n_\downarrow}}{e^e l_\uparrow^{l_\uparrow} l_\downarrow^{l_\downarrow} d^d} \right)^M \quad (57)$$

where $n_\sigma = N_\sigma/M$ and $d = D/M$, etc. From this we see that the distribution of double occupancies, as a function of d , narrows with increasing M . Thus, for general $g \in [0, 1]$, the density of double occupancies in the thermodynamic limit is given by the position of the maximum of the distribution. Using the asymptotics of the derivative of the gamma function $\Gamma(z)' = \Gamma(z)\Psi_0(z) \sim \Gamma(z) \ln(z)$, we find $g^2 = e d / (l_\uparrow l_\downarrow)$, or, explicitly,

$$d(n_\uparrow, n_\downarrow; g) = -\frac{1}{2} \left(\frac{1}{1-g^2} - n_\uparrow - n_\downarrow \right) + \sqrt{\frac{1}{4} \left(\frac{1}{1-g^2} - n_\uparrow - n_\downarrow \right)^2 + \frac{n_\uparrow n_\downarrow g^2}{1-g^2}}, \quad (58)$$

i.e., in the thermodynamic limit the Gutzwiller parameter determines d . In the uncorrelated case it gives the familiar $d(g=1) = n_\uparrow n_\downarrow$ from Fig. 3, while in the opposite limit there are only double occupancies above half-filling: $d(g=0) = \max(0, n - 1)$.

Similarly, we can estimate the overlap of wave functions with different Gutzwiller parameters, which is a measure of the efficiency of correlated sampling,

$$\langle \Psi(\tilde{g}) | \Psi(g) \rangle \sim \sum_D \frac{M!}{E! L_\uparrow! L_\downarrow! D!} (g\tilde{g})^D. \quad (59)$$

Expanding around $\tilde{g} = g$ we find that the overlap looks like a Gaussian: $\exp[-M(\tilde{g} - g)^2/\sigma_0^2]$, with M the number of lattice sites. As expected, for fixed $\tilde{g} \neq g$, the overlap goes to zero exponentially with system size (orthogonality catastrophe). σ_0 is a function of g and the filling and generally decreases with g . This can be seen in Figure 7: for small g the weights are peaked more sharply than for larger Gutzwiller parameters. For half-filling, $\sigma_0 = \sqrt{2g} 2(1+g)$. The relation between the overlap and the reliability of correlated sampling (53) is illustrated in Figure 6.

For the energy expectation value

$$E(g) = \frac{\langle \Psi(g) | H | \Psi(g) \rangle}{\langle \Psi(g) | \Psi(g) \rangle} = -2t \sum_{ij,\sigma} \frac{\langle \Psi(g) | c_{j\sigma}^\dagger c_{i\sigma} | \Psi(g) \rangle}{\langle \Psi(g) | \Psi(g) \rangle} + U D(g) \quad (60)$$

in the Gutzwiller approximation, the Hubbard energy is explicitly given through the relation (58) between g and the density of doubly occupied sites. For estimating the kinetic energy, we first observe that $\langle \Psi(g) | c_{j\sigma}^\dagger c_{i\sigma} | \Psi(g) \rangle / \langle \Psi(g) | \Psi(g) \rangle$ is the probability for an electron of spin σ to hop from site i to site j . The probability for a hop being allowed by the Pauli principle is $n_\sigma(1 - n_\sigma)$. In the Gutzwiller approximation there are more severe constraints on the hopping processes coming from the condition that the density of doubly occupied sites is fixed at (58). Thus, only hops from a singly occupied to an empty site (probability $l_\sigma e$) or from a doubly to a singly occupied site (probability $d l_{-\sigma}$) are allowed. Thus, replacing the Pauli constraint with the more severe Gutzwiller constraints reduces the hopping matrix elements of the uncorrelated Slater determinant by the *hopping reduction factor*

$$\gamma_\sigma(n_\sigma, g) = \frac{(\sqrt{l_\sigma e} + \sqrt{d l_{-\sigma}})^2}{n_\sigma(1 - n_\sigma)}, \quad (61)$$

where we have added the amplitudes for the two allowed hopping processes. Using again the basic assumption of the Gutzwiller approximation that all configurations contribute the same, we find for the energy per site

$$\varepsilon_{GA}(g) = \sum_\sigma \gamma_\sigma(n_\sigma, g) \varepsilon_\sigma^{(0)}(n_\sigma) + U d(n_\uparrow, n_\downarrow; g), \quad (62)$$

where $\varepsilon_\sigma^{(0)}(n_\sigma)$ is the kinetic energy for the Slater determinant of the Gutzwiller wave functions. Optimizing the Gutzwiller parameter g thus means finding the best trade-off between lowering the Hubbard energy by reducing the density of double occupancies and the simultaneous increase in the (negative) kinetic energy due to the band narrowing proportional to the hopping reduction γ_σ . Because of the relation (58) between g and d we actually need not consider g but can minimize the energy using d as the parameter.

2.4 Brinkman-Rice transition

At half-filling ($n_\sigma = 1/2$) the expressions from the Gutzwiller approximation simplify significantly. The hopping reduction factor becomes $\gamma = 16d(1/2 - d)$, so we can write the energy expectation value per site as

$$\varepsilon(d) = 16d(1/2 - d) \varepsilon^{(0)} + U d, \quad (63)$$

where $\varepsilon^{(0)}$ is the kinetic energy density of the uncorrelated system (both spins). Minimizing gives

$$d_{\min}(U) = \frac{1}{4} + \frac{U}{32\varepsilon^{(0)}}. \quad (64)$$

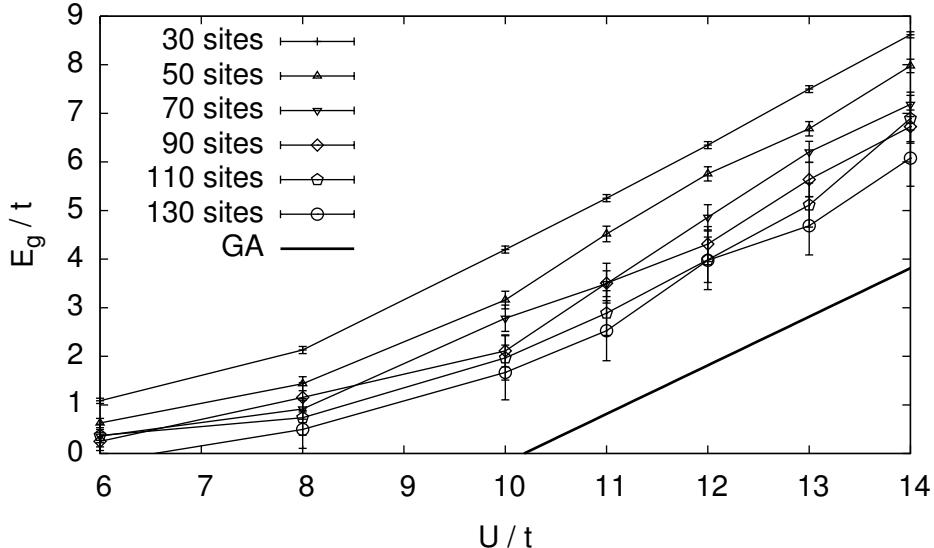


Fig. 8: Gap calculated by variational Monte Carlo for the Gutzwiller wave function on the finite one-dimensional Hubbard model with periodic boundary conditions (Hubbard rings). For 30-site rings a gap appears to open at $U \approx 6$. For larger rings the gap opens at slightly increasing values of U , and it looks as if the result of the Gutzwiller approximation is reached for large enough systems. But actually the opening of a gap is a pure finite-size effect, as a Gutzwiller wave function with metallic Slater determinant is always metallic, unless $g = 0$.

For the uncorrelated system this gives the familiar $d_{\min}(U = 0) = n_\uparrow n_\downarrow$. With increasing U , double occupancies are reduced until they vanish entirely at $U_c = 8|\varepsilon^{(0)}|$. From that point on, the system has no doubly occupied sites; the hopping reduction becomes $\gamma = 0$ suppressing all hopping, i.e., making the system an insulator. This is the Brinkman-Rice scenario for a half-filled band turning insulating [25].

We might wonder if we could see the metal-insulator transition in VMC calculations. The insulating state can be determined by calculating the gap

$$E_g(M) = E(N + 1) - 2E(N) + E(N - 1) \quad (65)$$

$$= \frac{\varepsilon(n + 1/M) - \varepsilon(n)}{1/M} - \frac{\varepsilon(n) - \varepsilon(n - 1/M)}{1/M} \quad (66)$$

$$\rightarrow \left. \frac{dE(n)}{dn} \right|_{n^+} - \left. \frac{dE(n)}{dn} \right|_{n^-} = \mu^+ - \mu^- . \quad (67)$$

Opening of the gap $E_g > 0$ indicates a jump in the chemical potential, i.e., an insulator. For an insulating half-filled system $N = M$ at large U we would expect $E_g \sim U$, since the system cannot avoid double occupancy with $M + 1$ electrons. Results for a simple Hubbard chain are shown in Figure 8. It appears that around $U_c = 32t/\pi \approx 10t$, the value predicted by the Gutzwiller approximation for the one-dimensional Hubbard model, the gap indeed starts to open proportionally to U . As the size of the Hubbard chain used in the simulation increases, the gap is slightly reduced, seemingly approaching the value predicted by the Gutzwiller approximation. This is, however, a pure finite-size effect.

For the one-dimensional Hubbard model, the ground state properties were actually calculated analytically by Metzner and Vollhardt [26]. For any finite U , the ground state energy $\varepsilon_{GWF}(n)$

has a continuous derivative at half-filling. So the Gutzwiller wave function always describes a metal, except for $g = 0$. This example should serve as a warning that finite-size extrapolations can be quite tricky. Here, even though the energy per site converges quickly to the exact result, having to take finite differences instead of derivatives in the evaluation of the gap for finite systems can create the appearance of a gapped system.

There is an elegant argument using the response of the energy to twisted boundary conditions that shows that, quite generally, Gutzwiller-type wave functions with a metallic Slater determinant are always metallic [27]. Consider a variational wave function

$$|\Psi\rangle = \prod_{\alpha} g_{\alpha}^{C_{\alpha}} |\Phi\rangle, \quad (68)$$

where $|\Phi\rangle$ is a Slater determinant and C_{α} a set of correlation functions (the simple Gutzwiller function uses only one correlation function, $\sum_i n_{i\uparrow} n_{i\downarrow}$). For twisted boundary conditions in direction c on a finite simulation cell, moving an electron by the cell vector $\mathbf{c} = \mathbf{A}\mathbf{n}$ introduces a phase $\exp(i\mathbf{k} \cdot \mathbf{c})$. This phase can be absorbed into the Hamiltonian by transforming the creation/annihilation operators at site $\mathbf{A}\mathbf{j}$ into $c_{\mathbf{j}\sigma} \rightarrow e^{i\mathbf{k} \cdot \mathbf{A}\mathbf{j}} c_{\mathbf{j}\sigma}$ and introducing periodic boundary conditions. The Hamiltonian thus becomes dependent on \mathbf{k} with the hopping terms picking up a phase from the twisted boundary conditions, while in the Hubbard interaction the phases cancel

$$H(\mathbf{k}) = -t \sum_{ij,\sigma} e^{i\mathbf{k} \cdot \mathbf{A}(\mathbf{i}-\mathbf{j})} c_{\mathbf{j}\sigma}^{\dagger} c_{\mathbf{i}\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (69)$$

The energy expectation value for the Gutzwiller wave function depends on k and the g_{α} , where the Gutzwiller parameters change with the boundary conditions as

$$E_G(k+dk, \{g_{\alpha}(k+dk)\}) = E_G(k, \{g_{\alpha}(k)\}) + \left(\frac{\partial E_G}{\partial k} + \sum_{\alpha} \frac{\partial E_G}{\partial g_{\alpha}} \frac{dg_{\alpha}}{dk} \right) dk + \mathcal{O}(dk^2). \quad (70)$$

The Gutzwiller parameters minimize E_G , i.e., the variations of the energy expectation value with respect to the g_{α} vanish. Solving the resulting linear system for the first-order term gives the dependence of the Gutzwiller parameters on the boundary conditions

$$\frac{dg_{\alpha}}{dk} = - \sum_{\beta} \left(\frac{\partial^2 E_G}{\partial g_{\beta} \partial g_{\alpha}} \right)^{-1} \left(\frac{\partial^2 E_G}{\partial g_{\beta} \partial k} \right), \quad (71)$$

while the second derivative of the energy with respect to the boundary conditions is

$$\frac{d^2 E_G}{dk^2} = \frac{d}{dk} \left(\frac{\partial E_G}{\partial k} + \sum_{\alpha} \frac{\partial E_G}{\partial g_{\alpha}} \frac{dg_{\alpha}}{dk} \right) = \frac{\partial^2 E_G}{\partial k^2} + \sum_{\alpha} \frac{\partial^2 E_G}{\partial k \partial g_{\alpha}} \frac{dg_{\alpha}}{dk}. \quad (72)$$

When the Gutzwiller factors in (68) are independent of the boundary conditions, e.g., the C_i are density or spin correlation functions, the explicit dependence of E_G on the boundary conditions k is only through the kinetic energy T . For a metallic Slater determinant the first term will then produce a non-vanishing conductivity for any U , except in the atomic limit $U \rightarrow \infty$.

The Brinkman-Rice transition is thus produced by the Gutzwiller approximation, although it is not present in the underlying Gutzwiller wave function, except in the limit $d \rightarrow \infty$, where the Gutzwiller approximation becomes exact [26].

3 Projection techniques

We can systematically improve on the variational results by using projection techniques [28]. The basic idea is surprisingly simple: when we operate with $\exp(-\tau H)$ on a wave function $|\Psi_T\rangle$ then, for large τ , the ground state $|\psi_0\rangle$ will dominate in the projected function, provided that the initial function had non-zero overlap with it. For a finite-dimensional Hamiltonian, where the spectrum is bounded not only from below but also from above, this imaginary-time propagation can be simplified to a matrix vector product

$$|\Psi^{(n+1)}\rangle = [1 - \tau(H - \bar{E}^{(n)})] |\Psi^{(n)}\rangle ; \quad |\Psi^{(0)}\rangle = |\Psi_T\rangle , \quad (73)$$

where τ has to be small enough and $\bar{E}^{(n)}$ is chosen to ensure normalization of the projected functions. To see under what conditions this converges to the ground state, we expand the starting function $|\Psi_T\rangle = \sum_i c_i |\Psi_i\rangle$ in eigenstates $H|\Psi_i\rangle = E_i |\Psi_i\rangle$. Then

$$|\Psi^{(n)}\rangle = \sum_i c_i \prod_n [1 - \tau(E_i - \bar{E}^{(n)})] |\Psi_i\rangle . \quad (74)$$

Convergence to $|\Psi_0\rangle$, up to normalization, is ensured if $c_i \neq 0$ and

$$|1 - \tau(E_0 - \bar{E}^{(n)})| > |1 - \tau(E_i - \bar{E}^{(n)})| \quad \forall i \neq 0 . \quad (75)$$

For $\tau > 0$ we distinguish two cases

- $1 - \tau(E_0 - \bar{E}^{(n)}) > 1 - \tau(E_i - \bar{E}^{(n)})$, which leads to the trivial $E_0 < E_i$, and
- $1 - \tau(E_0 - \bar{E}^{(n)}) > -[1 - \tau(E_i - \bar{E}^{(n)})]$, from which follows that $2 > \tau(E_i + E_0 - 2\bar{E}^{(n)})$.

Thus, to secure convergence, one has to choose

$$0 < \tau < \frac{2}{E_{\max} + E_0 - 2\bar{E}^{(n)}} \quad (76)$$

which implies that $\bar{E}^{(n)} \in [E_0, E_{\max}]$ must lie inside the spectrum of H . In fact, for large n it will approach the ground state energy.

Because of the prohibitively large dimension of the many-body Hilbert space, the matrix vector product in (73) cannot be done exactly. Instead, we rewrite the equation in configuration space

$$\sum_{R'} |R'\rangle \langle R' | \Psi^{(n+1)} \rangle = \sum_{R, R'} |R'\rangle \underbrace{\langle R' | 1 - \tau(H - E_0) | R \rangle}_{=: F(R', R)} \langle R | \Psi^{(n)} \rangle \quad (77)$$

and perform the propagation in a stochastic sense: $|\Psi^{(n)}\rangle$ is represented by an ensemble of configurations R with weights $w(R)$. The transition matrix element $F(R', R)$ is rewritten as a transition probability $p(R \rightarrow R')$ times a normalization factor $m(R', R)$. The iteration (77) is then stochastically performed as follows: for each R we pick, out of the set of all allowed configurations, *one* new configuration R' with probability $p(R \rightarrow R')$ and multiply its weight by $m(R', R)$. Then the new ensemble of configurations R' with their respective weights represents the new function $|\Psi^{(n+1)}\rangle$.

3.1 Importance sampling

Importance sampling introduces a guiding function $|\Psi_G\rangle$ to decisively improve the efficiency of the stochastic projection by enhancing transitions from configurations where the trial function is small to configurations with large trial function, i.e., by replacing the transition matrix element $F(R', R)$ with $G(R', R) = \langle R'|\Psi_G\rangle F(R', R) / \langle R|\Psi_G\rangle$. The propagation is then given by

$$\sum_{R'} |R'\rangle \langle R'|\Psi_G\rangle \langle R'| \Psi^{(n+1)} \rangle = \sum_{R, R'} |R'\rangle G(R', R) \langle R|\Psi_G\rangle \langle R| \Psi^{(n)} \rangle \quad (78)$$

and the ensemble of configurations now represents the product $\Psi_G \Psi^{(n)}$. This means that the probability distribution function $P^{(n)}(w, R) dw$ of configurations R with weight w is such that

$$\Psi_G(R) \Psi^{(n)}(R) = \int w P^{(n)}(w, R) dw. \quad (79)$$

To see this, we rewrite the matrix element of the propagation as

$$G(R', R) = p(R \rightarrow R') m(R', R), \quad (80)$$

where $p(R \rightarrow R')$ is the probability for the random walk to move from configuration R to R' and the weight $m(R', R)$ takes care of the normalization. For the probability distribution function this implies

$$P^{(n+1)}(w', R') dw' = \sum_R p(R \rightarrow R') P^{(n)}\left(\frac{w'}{m(R', R)}, R\right) \frac{dw'}{m(R', R)} \quad (81)$$

and hence

$$\begin{aligned} \int w' P^{(n+1)}(w', R') dw' &= \sum_R p(R \rightarrow R') \int w' P^{(n)}\left(\frac{w'}{m(R', R)}, R\right) \frac{dw'}{m(R', R)} \\ &= \sum_R p(R \rightarrow R') m(R', R) \int w P^{(n)}(w, R) dw \\ &= \sum_R G(R', R) \Psi_G(R) \Psi^{(n)}(R) \\ &= \Psi_G(R') \Psi^{(n+1)}(R'). \end{aligned}$$

After a large number n of iterations, the ground-state energy is given by the *mixed estimator*

$$E_0^{(n)} = \frac{\langle \Psi_G | H | \Psi^{(n)} \rangle}{\langle \Psi_G | \Psi^{(n)} \rangle} \approx \frac{\sum_R E_{\text{loc}}(R) w^{(n)}(R)}{\sum_R w^{(n)}(R)}. \quad (82)$$

When we start the iteration from the guiding function, we can generate the configurations for the initial state $\Psi_G(R) \Psi^{(0)}(R)$ by a variational Monte Carlo run for Ψ_G . For this practical reason one usually chooses the guiding function to be the VMC trial function. In the following we therefore use Ψ_G and Ψ_T synonymously.

| Hamiltonian | Hilbert space | | |
|------------------|---------------|------------|---|
| | full | restricted | system |
| H | -1.2238 | -1.2203 | 4×4 -Hubbard model, 5 + 5 electrons |
| H_{eff} | -1.2236 | -1.2201 | $U = 4t$, free electron nodes |
| H | 8.46487 | 8.46490 | 4 C ₆₀ molecules, 5 + 6 electrons |
| H_{eff} | 8.46557 | 8.46728 | $U = 1.0 \text{ eV}$, $U_0 = 0.1 \text{ eV}$ |
| H | 10.69937 | 10.69937 | 4 C ₆₀ molecules, 6 + 6 electrons |
| H_{eff} | 10.70045 | 10.70045 | $U = 1.0 \text{ eV}$, $U_0 = 0.1 \text{ eV}$ |
| H | 19.50941 | 19.54933 | 4 C ₆₀ molecules, 7 + 8 electrons |
| H_{eff} | 19.51073 | 19.55372 | $U = 1.0 \text{ eV}$, $U_0 = 0 \text{ eV}$ |
| H | 19.50941 | 19.50941 | 4 C ₆₀ molecules, 7 + 8 electrons |
| H_{eff} | 19.50941 | 19.51068 | $U = 1.0 \text{ eV}$, $U_0 = 0.1 \text{ eV}$ |

Table 1: Approximations in FNMC: results of exact diagonalization of the true/effective Hamiltonian on the full/restricted Hilbert space for the Hubbard model and a model of solid K₃C₆₀ [23]. For the meaning of U_0 see section 3.3.2.

3.2 Fixed-node Monte Carlo

As long as the evolution operator has only non-negative matrix elements $G(R', R)$, all weights $w(R)$ will be positive. If, however, G has negative matrix elements there will be configurations with negative as well as positive weight. Their contributions to the estimator (82) tend to cancel so that eventually the statistical error dominates, rendering the simulation useless. This is the infamous *sign problem*. A straightforward way to get rid of the sign problem is to remove the offending matrix elements from the Hamiltonian, thus defining a new Hamiltonian H_{eff} by

$$\langle R' | H_{\text{eff}} | R \rangle = \begin{cases} 0 & \text{if } G(R', R) < 0 \\ \langle R' | H | R \rangle & \text{else} \end{cases} \quad (83)$$

for $R' \neq R$. For each off-diagonal element $\langle R' | H | R \rangle$ that has been removed, a term is added to the diagonal

$$\langle R | H_{\text{eff}} | R \rangle = \langle R | H | R \rangle + \sum_{R'} \Psi_G(R') \langle R' | H | R \rangle / \Psi_G(R). \quad (84)$$

This is the *fixed-node approximation* for lattice Hamiltonians [29]. By construction, H_{eff} is free of the sign-problem, and it gives a variational energy, i.e. $E_0^{\text{eff}} \geq E_0$. The exact ground-state energy is obtained if $\Psi_G(R')/\Psi_G(R) = \Psi_0(R')/\Psi_0(R)$ for all R, R' with $G(R', R) < 0$.

It is important to realize that fixed-node Monte Carlo involves two distinct approximations. The obvious approximation is replacing the true Hamiltonian by the effective Hamiltonian H_{eff} . Somewhat less obvious is the fact that importance sampling amounts to restricting the accessible region of the Hilbert space to configurations where $\Psi_G(R) \neq 0$. For small systems these two approximations can be considered separately by performing exact diagonalization for the effective Hamiltonian on the full Hilbert space and by diagonalizing the true Hamiltonian on the restricted Hilbert space. Finally, the fixed-node Monte Carlo energy can be determined by diagonalizing H_{eff} on the restricted Hilbert space. Examples are given in Table 1.

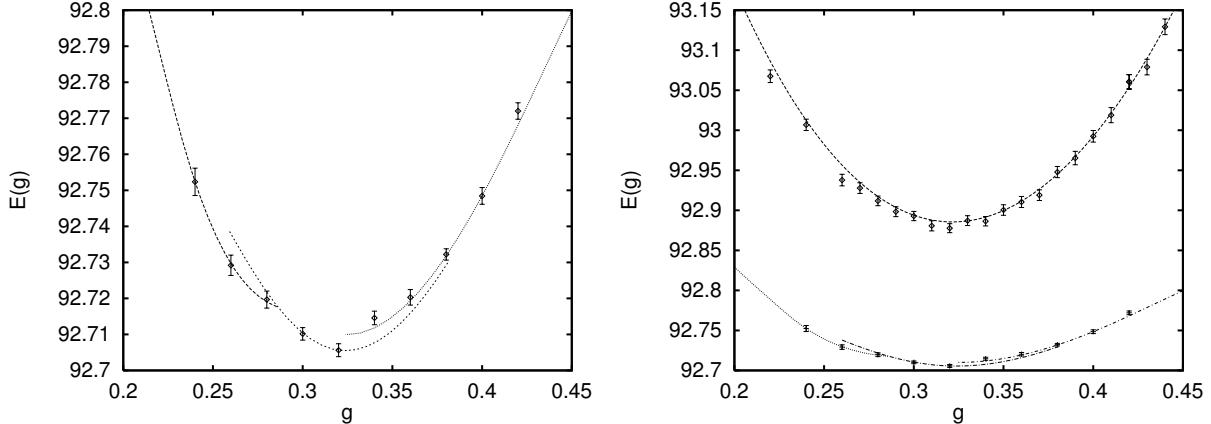


Fig. 9: Optimizing the Gutzwiller parameter g in fixed-node Monte Carlo: the left-hand panel shows how the correlated sampling approach for the Gutzwiller parameter g works for FNMC. The error bars give the results of independent Monte Carlo calculations, while the curves show the results of different correlated sampling runs. The right-hand panel compares variational and fixed-node Monte Carlo. The calculations are for the same system as in Figure 5.

Fixed-node Monte Carlo for a lattice Hamiltonian thus means that we choose a guiding/trial function from which we construct an effective Hamiltonian and determine its ground-state by Monte Carlo. Because of the variational property, we want to pick the Ψ_T such that E_0^{eff} is minimized; i.e., we want to optimize the trial function or, equivalently, the effective Hamiltonian. As in variational Monte Carlo we can use the concept of correlated sampling [21]. For optimizing the Gutzwiller parameter g we can even exploit the idea of rewriting the correlated sampling sums into polynomials in \tilde{g}/g . There is, however, a problem arising from the fact that the weight of a given configuration $R^{(n)}$ in iteration n is given by the product $w(R^{(n)}) = \prod_{i=1}^n m(R^{(i)}, R^{(i-1)})$: each individual normalization factor $m(R', R)$ can be written as a finite polynomial, but the order of the polynomial for $w(R^{(n)})$ keeps increasing with the number of iterations. It is therefore not practical to try to calculate the ever-increasing number of coefficients for the correlated sampling function $E^{(n)}(\tilde{g})$. But since we still can easily calculate the coefficients for the $m(R', R)$, we may use them to evaluate $E^{(n)}(\tilde{g})$ in each iteration on a set of *predefined* values \tilde{g}_i of the Gutzwiller parameter. Figure 9 shows an example. We find that the FNMC energy depends much less on the trial function than in VMC. This is not unexpected: while in variational Monte Carlo the whole trial function is fixed, only the values of the trial function next to a node enter the fixed-node Hamiltonian H_{eff} , which in FNMC is then treated exactly. To realize the higher accuracy of the FNMC method, it is nevertheless important to carefully optimize the trial function. Finally, it is interesting to note that the Gutzwiller factor that minimizes E_{VMC} is usually not quite the optimum Gutzwiller factor for fixed-node MC.

3.3 Optimization of the trial function

As mentioned before, typical trial functions for quantum Monte Carlo calculations are of the type $\Psi_T(R) = g^{D(R)} \Phi(R)$, with g the Gutzwiller parameter and Φ a Slater determinant. Along with explaining the Monte Carlo approaches, we have already described how g can be opti-

mized. The fundamental idea was that, for the reweighting in correlated sampling, only ratios of the new and old trial functions are needed so that the weights and energies appearing in the Monte Carlo calculation can be recast in the form of polynomials in the ratio of the Gutzwiller parameters. In the following, we discuss generalizations of this approach to trial functions with several Gutzwiller parameters. After that, we address the optimization of the other part of a Gutzwiller wave function: the Slater determinant. In particular, we demonstrate how the character of the Slater determinant affects the result of the Monte Carlo calculation.

3.3.1 More Gutzwiller parameters

To study the static dielectric screening [30], we have to determine the response of the charge density to the introduction of a test charge q placed on molecule i_q . To describe the test charge, the term

$$H_1(q) = qU \sum_{m\sigma} n_{i_q m\sigma} \quad (85)$$

is added to the Hamiltonian. In the spirit of the Gutzwiller ansatz, we correspondingly add a second Gutzwiller factor to the wave function that reflects the additional interaction term qUN_{i_q}

$$|\Psi_T(g, h)\rangle = g^D h^{N_{i_q}} |\Phi\rangle. \quad (86)$$

Finding the best Gutzwiller parameters is now a two-dimensional optimization problem. Dealing with polynomials in the two variables g and h , the method of correlated sampling works as straightforwardly as described above for the case of a plain Gutzwiller wave function. As an example, Fig. 10 shows the result of the optimization, both in variational and in fixed-node Monte Carlo, for a cluster of 64 C₆₀ molecules in an fcc arrangement (periodic boundary conditions) resembling K₃C₆₀ with a test charge $q = 1/4$. In practice, we first optimize the parameters in variational Monte Carlo. We then use the optimum VMC parameters as starting points for the optimization in the more time-consuming fixed-node Monte Carlo calculations.

3.3.2 Variation of the Slater determinant

In the traditional Gutzwiller ansatz, the Slater determinant Φ is the ground-state wave function of the non-interacting Hamiltonian. This is, however, not necessarily the best choice. An alternative would be to use the Slater determinant $\Phi(U)$ obtained by solving the interacting problem in the Hartree-Fock approximation. We can even interpolate between the two extremes by doing a Hartree-Fock calculation with a fictitious Hubbard interaction U_0 to obtain Slater determinants $\Phi(U_0)$. Yet another family of Slater determinants $\Phi(H_{\text{stag}})$ can be obtained from solving the non-interacting Hamiltonian with an added staggered magnetic field, which lets us control the antiferromagnetic character of the trial function. Although optimizing parameters in the Slater determinant is not possible with the method described in the preceding sections, an efficient optimization of the Gutzwiller factors makes it possible to optimize the overall trial function without too much effort.

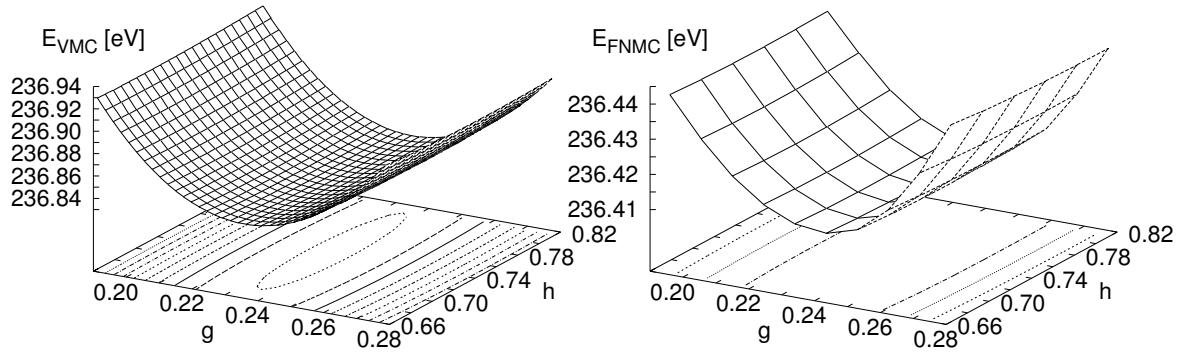


Fig. 10: Correlated sampling for the parameters g and h of the generalized Gutzwiller wave function $|\Psi_T\rangle = g^D h^{N_c} |\Phi\rangle$, in variational (left) and fixed-node Monte Carlo (right). The plots show the energy as a function of the Gutzwiller parameters g and h , both as surfaces and contours. The calculations are for an fcc cluster of 64 molecules with $96 + 96$ electrons (half-filled t_{1u} -band), an on-site Hubbard interaction $U = 1.25$ eV, and a test charge of $q = 1/4$ (in units of the electron charge).

Staggered magnetic field Introducing a staggered magnetic field, we can construct Slater determinants by solving the non-interacting Hamiltonian with an added Zeeman term. To be specific, we consider K_3C_{60} , which has a half-filled t_{1u} -band. Since K_3C_{60} crystallizes in an fcc lattice, antiferromagnetism is frustrated and the definition of a staggered magnetic field is not unique. We split the fcc lattice into two sublattices A and B such that frustration is minimized. The Zeeman term is then given by

$$H_m = H_{\text{stag}} \sum_i \text{sign}(i) [n_{i\uparrow} - n_{i\downarrow}] \quad (87)$$

with $\text{sign}(i) = +1$ if $i \in A$ and -1 if $i \in B$. It effectively introduces an on-site energy that has opposite sign for the two spin orientations on the same site, and for the same spin orientation, has opposite sign on the two sublattices. Therefore, hopping to neighboring sites on different sublattices involves an energy cost of twice the Zeeman energy. The staggered magnetic field thus not only induces antiferromagnetic order in the Slater determinant but also serves to localize the electrons. This is reflected in the fact that the optimum Gutzwiller parameter is much larger for Slater determinants constructed from a Hamiltonian with large H_{stag} than for paramagnetic Slater determinants. Varying H_{stag} then interpolates between paramagnetic/itinerant and antiferromagnetic/localized wave functions.

The energy expectation values for such trial functions as calculated in variational Monte Carlo are shown in Fig. 11. It shows E_{VMC} as a function of the antiferromagnetic correlation

$$\langle s_i s_{i+1} \rangle = \frac{1}{N} \sum_{\langle ij \rangle} (n_{i\uparrow} - n_{i\downarrow}) (n_{j\uparrow} - n_{j\downarrow}), \quad (88)$$

where the sum is over the N nearest neighbors. $\langle s_i s_{i+1} \rangle$ is a monotonic function of H_{stag} . For each different value of the Hubbard interaction U we find a curve with two minima. One

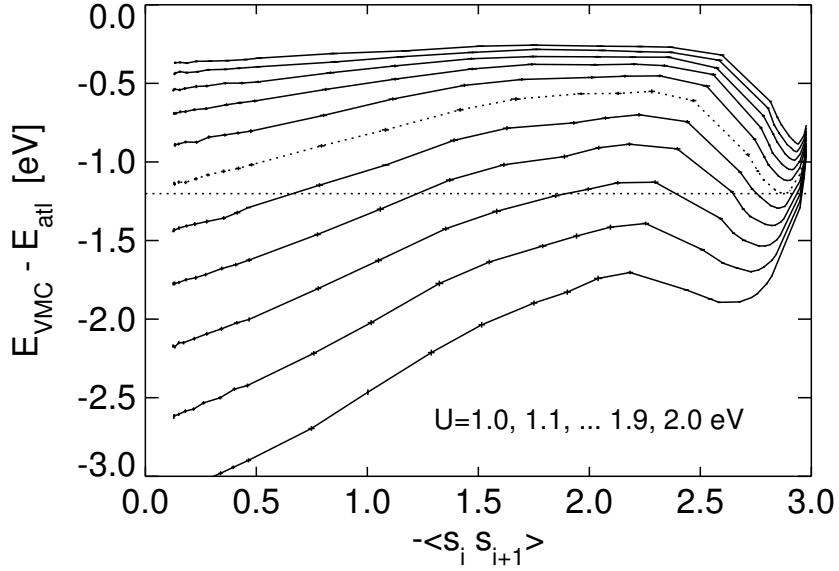


Fig. 11: Variational energy E_{VMC} for trial functions with different character. Plotted are the energies (error bars, lines are to guide the eye) for a Hamiltonian describing K_3C_{60} (periodic fcc cluster of 32 molecules) with Hubbard interaction $U = 1.0, 1.1 \dots 1.9, 2.0 \text{ eV}$ ($U = 1.0 \text{ eV}$ corresponding to the lowest curve). Instead of the total energies E_{tot} , we plot the difference of E_{tot} and the energy in the atomic limit (each site occupied by three electrons) so that the results for different U can be readily compared. The trial functions are of the Gutzwiller type. The Slater determinants were determined from diagonalizing the non-interacting Hamiltonian (i.e. setting $U = 0$) with a staggered magnetic field H_{stag} . This field gives rise to an antiferromagnetic correlation of neighboring spins, which is plotted on the abscissa. For $U = 1.5 \text{ eV}$ (dotted curve) the minima in the paramagnetic and the antiferromagnetic regions have about the same energy.

minimum is realized for the non-magnetic ($H_{\text{stag}} = 0$) trial function. The energy as a function of U scales roughly like $E_{\text{para}} \propto -(1-U/U_c)^2$, as predicted by the Gutzwiller approximation. The second minimum is in the antiferromagnetic/localized region and scales roughly like $E_{\text{AF}} \propto -t^2/U$, as expected. For small U , the non-magnetic state is more favorable, while for large U the localized Slater determinant gives lower variational energies. The crossover is at $U_c \approx 1.50 \text{ eV}$ (dotted line) and resembles a first-order phase transition.

Hartree-Fock An alternative method for constructing Slater determinants is to use the interacting Hamiltonian with the physical Hubbard interaction U replaced by a parameter U_0 and solve it in the Hartree-Fock approximation. In practice this is done simply by means of an unrestricted self-consistent calculation for the finite, periodic clusters under consideration, starting from some charge and spin density that breaks the symmetry of the Hamiltonian. It is well known that Hartree-Fock favors the antiferromagnetic Mott insulator, predicting a Mott transition at much too small values U_c^{HF} . It is therefore not surprising that good trial functions are obtained for values of U_0 considerably smaller than U . For U_0 close to zero, the Slater determinant has metallic character, while for somewhat larger U_0 there is a metal-insulator transition. Figure 12 shows the energy as a function of U_0 for the model of K_3C_{60} . We find that the re-

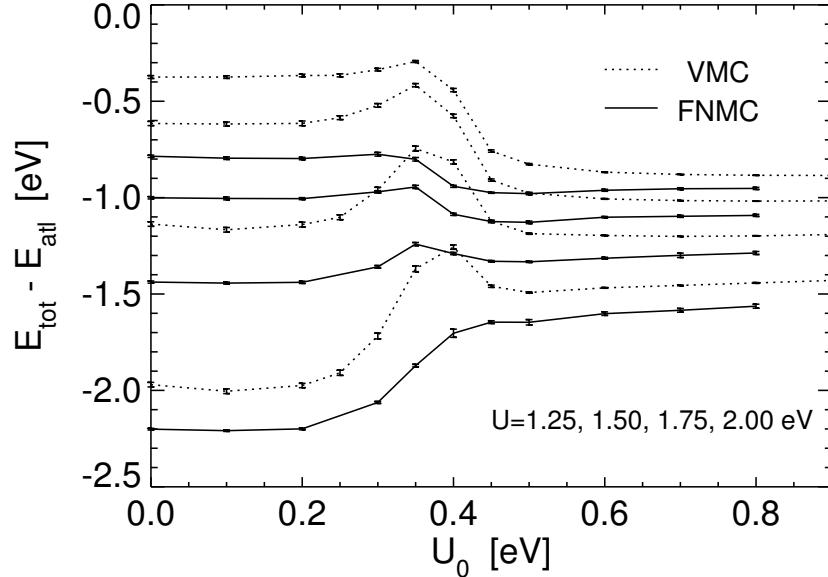


Fig. 12: Optimizing U_0 : Dependence of variational (VMC) and fixed-node Monte Carlo (FNMC) on the trial function. U_0 is the Hubbard interaction that was used for the Slater determinant in the Gutzwiller wave function $\Psi_T(R) = g^{D(R)} \Phi(U_0)$. The Gutzwiller parameter has always been optimized. The results shown here are the energies (relative to the atomic limit) for a Hamiltonian that describes K_3C_{60} (32 molecules), with U being varied from 1.25 (lowest curve) to 2.00 eV (highest curve).

sults of variational Monte Carlo depend quite strongly on the parameter U_0 . As expected, for a given Hubbard interaction U there is a transition from the paramagnetic region for small U_0 to a region where the trial function is antiferromagnetic. In fixed-node Monte Carlo, energies are consistently lowered and the dependence on the trial function is weaker. It seems that here it is mainly the character (paramagnetic/antiferromagnetic) of the trial function that matters. For small U , trial functions with small U_0 give lower energy, while for large U trial functions with larger U_0 are favorable. The crossover coincides with the Mott transition, which takes place between $U = 1.50$ and 1.75 eV.

3.4 Quasi-particle energies

As discussed above, the sign-problem in quantum Monte Carlo simulations for fermions makes it necessary to introduce an approximation. Such an approximation is of course undesirable when we are interested in the ground state. Surprisingly, however, we can use the fixed-node approximation to our advantage to calculate excited states and quasi-particle energies. The basic idea is that by the proper choice of the trial function, an excited state can be stabilized in the Monte Carlo simulation that otherwise would decay to a state with lower energy.

The fundamental concept for understanding this is what we call the *nodal sets*, defined as the sets of configurations where the trial function is positive/negative:

$$\mathcal{N}_\Psi^\pm := \left\{ R \mid \langle R | \Psi \rangle \gtrless 0 \right\}. \quad (89)$$

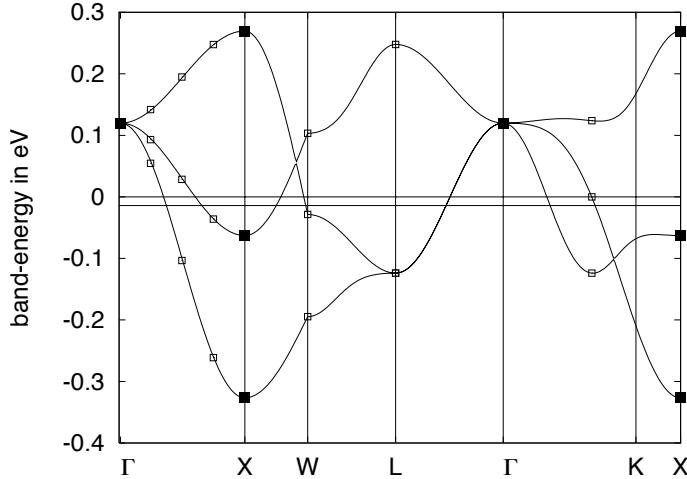


Fig. 13: C_{60} bands: Full symbols show the k -states that are present in a 4-molecule cluster with periodic boundary conditions: 3 (degenerate states) with $k = 0$ (Γ -point) and 9 states [3 band states with $k = (2\pi/a, 0, 0)$, $(0, 2\pi/a, 0)$, and $(0, 0, 2\pi/a)$] (X -point).

The first observation we make is that under the importance-sampled FNMC iterations (78) (with H replaced by H_{eff}) the nodal sets do not change. To see this, we write $|\Psi^{(n)}\rangle$ in terms of the trial function

$$\langle R|\Psi^{(n)}\rangle = \alpha^{(n)}(R) \langle R|\Psi_T\rangle. \quad (90)$$

Then the iterations (78) take the form

$$\begin{aligned} \langle \Psi_T | R' \rangle \langle R' | \Psi^{(n)} \rangle &= \sum G_{FN}(R', R) \langle \Psi_T | R \rangle \langle R | \Psi^{(n-1)} \rangle \\ &= \alpha^{(n)}(R') \underbrace{|\Psi_T(R')|^2}_{>0} = \sum_R G_{FN}(R', R) \underbrace{|\Psi_T(R)|^2}_{\geq 0} \underbrace{\alpha^{(n-1)}(R)}_{>0}. \end{aligned}$$

Since $\alpha^{(0)}(R) = 1$ and never changes sign, $\alpha^{(n)}(R) \geq 0$ for all n .

On the other hand, each eigenstate Ψ_n of H is characterized by its nodal sets. If there were two eigenstates with the same nodal sets, then

$$\langle \Psi_n | \Psi_m \rangle = \sum_{R \in \mathcal{N}^+} \Psi_n(R) \Psi_m(R) + \sum_{R \in \mathcal{N}^-} \Psi_n(R) \Psi_m(R) > 0 \quad (91)$$

in contradiction to the orthogonality of the eigenstates.

It thus seems reasonable to try to calculate excited-state energies with fixed-node Monte Carlo, where the trial function is chosen such as to have nodal sets similar to that of the desired state. As we have seen above, the most straightforward guess for the ground-state is to use the non-interacting Slater determinant $|\Phi_0\rangle$ (remember that the Gutzwiller factors do not change the sign of the wave function). Then the simplest trial functions for the nodes of excited states are the non-interacting wave functions

- quasi-electron: $c_k^\dagger |\Phi_0\rangle$
- quasi-hole: $c_k |\Phi_0\rangle$.

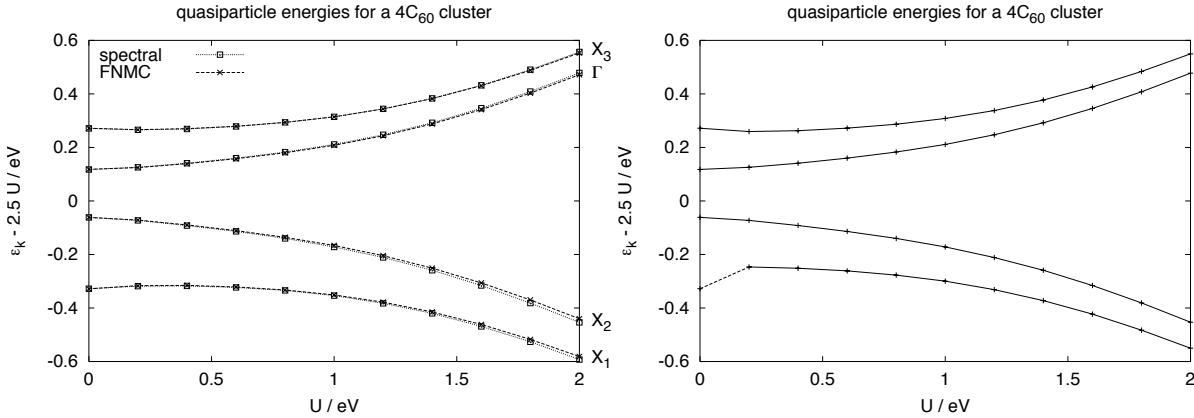


Fig. 14: Quasi-particle energies for the half-filled, 4-molecule C₆₀ cluster: The left-hand panel shows the fixed-node energies ($\varepsilon_n = E_n^{N+1} - E_0^N$ for quasi-electrons in states X_3 and Γ , $\varepsilon_n = E_0^N - E_n^{N-1}$ for quasi-holes in X_1 and X_2). For comparison, the open squares give the position of the peak in the corresponding spectral function calculated by exact diagonalization, as shown in Figure 15. The right-hand panel shows the result of Lanczos calculations starting from the quasi-particle trial functions used in fixed-node Monte Carlo. It is clear that for $U > 0$ the quasi-hole at X_1 is not stable.

To see how this simple approach works, we consider the model for describing the t_{1u} -electrons in C₆₀ (cf. Table 1) and compare to results of exact diagonalization for a cluster of 4 molecules. The t_{1u} -bands are reproduced in Figure 13, where the filled squares represent the reciprocal lattice vectors of the 4-molecule simulation cell: three degenerate states at the Γ point, and three different levels at the X point, each of which is three-fold degenerate. For the half-filled system, the lower two levels at the X point are filled (below we will refer to them as X_1 and X_2), while the highest level at X (referred to as X_3) and Γ are empty. To calculate the quasi-particle energies, we first perform a calculation for the ground-state energy of the half-filled system E_0^N using a Gutzwiller wave function based on the non-interacting Slater determinant $|\Phi_0\rangle$. Then, we calculate the fixed-node energy for the system with an extra electron/hole with a trial function based on $c_k^\dagger |\Phi_0\rangle$ or $c_k |\Phi_0\rangle$, respectively. To keep the wave functions real we make use of the inversion symmetry, setting $c_k = \cos(k r_i) c_i$. The resulting quasi-particle energies are plotted in Figure 14. To compare to the true quasi-particle energies, we have calculated the corresponding spectral functions

$$A(\mathbf{k}, \omega) = \begin{cases} \sum_n \left| \langle \Psi_n^{N+1} | c_k^\dagger | \Psi_0^N \rangle \right|^2 \delta(\omega - (E_n^{N+1} - E_0^N)) & (\text{quasi-electron}) \\ \sum_n \left| \langle \Psi_n^{N-1} | c_k | \Psi_0^N \rangle \right|^2 \delta(\omega - (E_0^N - E_n^{N-1})) & (\text{quasi-hole}) \end{cases}$$

by exact diagonalization (see Figure 15). The position of the peak in $A(\mathbf{k}, \omega)$ is plotted in Figure 14. We find a remarkable agreement between fixed-node and exact energies. This is not too unexpected. Looking again at the band-structure, we see that the quasi-electron at X_3 is the lowest energy state with that k -vector, while the quasi-electron at Γ is the lowest with $k = 0$. So these states are ground-states in their respective symmetry sectors and should thus be accessible by a ground-state method like fixed-node Monte Carlo. The same is true for the quasi-hole at

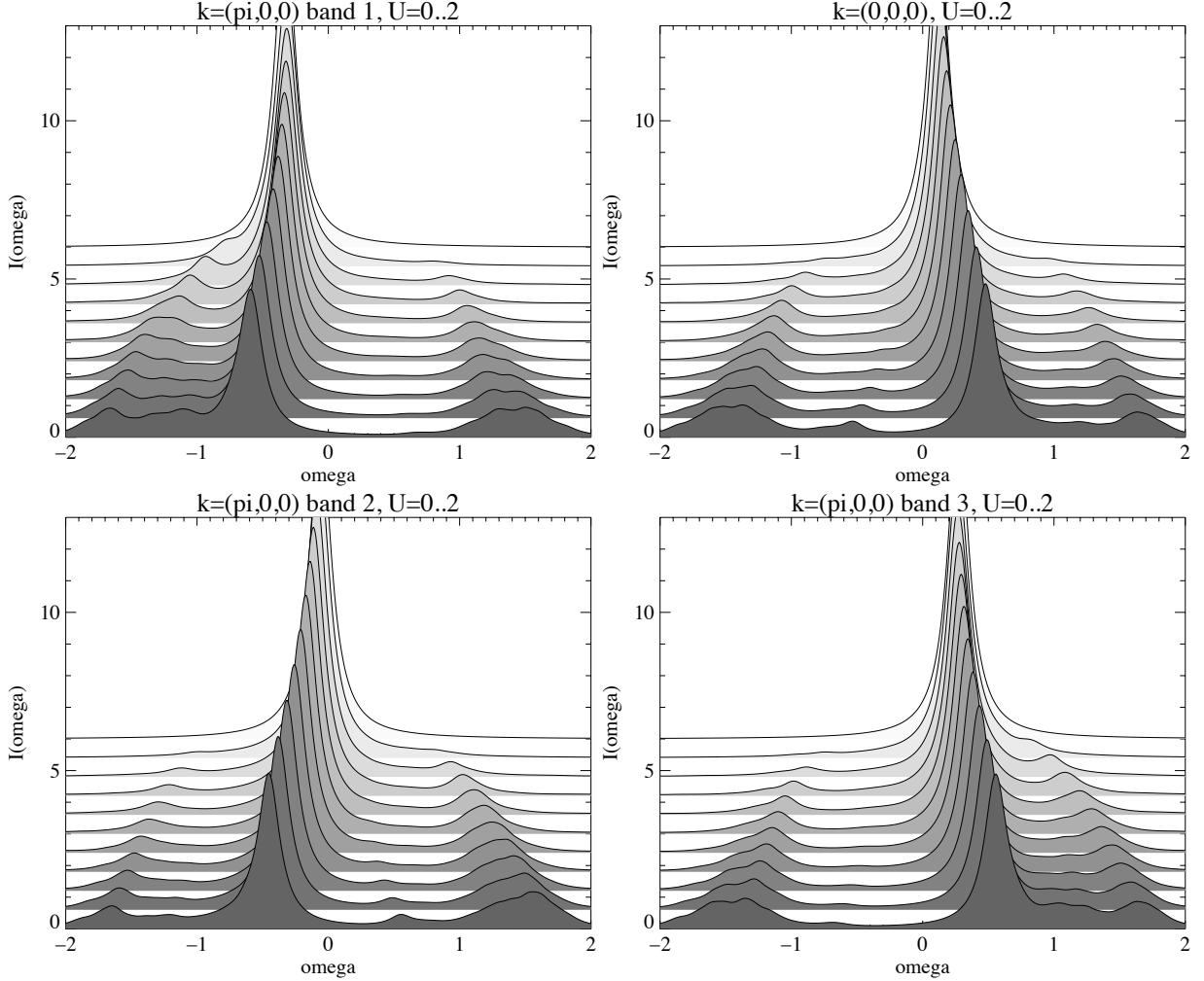


Fig. 15: Spectral function for a 4 C_{60} molecule cluster calculated by exact diagonalization. From back to front the curves show the spectral function for U from 0 to 2 eV in steps of 0.2 eV.

X_2 . The quasi-hole at X_1 , however, has the same k as the one at X_2 and thus cannot be the ground-state in that symmetry sector. To check this, we have performed Lanczos runs starting from the determinants used in the respective fixed-node calculations. The results are shown in the right panel of Figure 14. We indeed find that all states are stable, except for X_1 , which for $U > 0$ decays into a lower-energy quasi-hole. This decay is induced by the interaction term. This can be understood by writing it in k -space

$$\frac{U}{M} \sum_{k,k',q} c_{k\uparrow}^\dagger c_{k-q\uparrow} c_{k'\downarrow}^\dagger c_{k'+q\downarrow} .$$

What happens for the X_1 state is visualized in Figure 16: The lower-energy quasi-hole is obtained by filling the X_1 hole with an X_2 electron, while exciting another X_2 electron into Γ .

To verify that this picture is indeed correct, we have calculated the spectral function for the decayed state

$$A = \sum \left| \langle \Psi_n^{N-1} | c_{X_2} c_\Gamma^\dagger c_{X_2} | \Psi_0^N \rangle \right|^2 \delta(\omega - (E_0^N - E_n^{N-1})) , \quad (92)$$

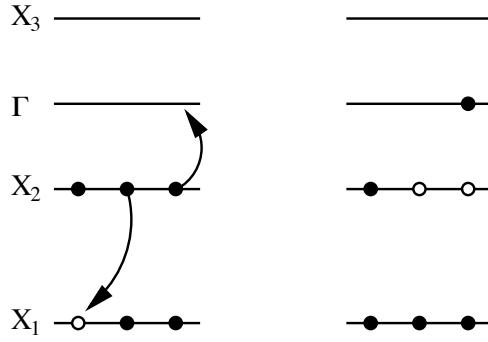


Fig. 16: Quasi-particle decay induced by interaction term for a quasi-hole at X_1 . Both states have the same k , so the transition is allowed by symmetry. (Notice that there are three different X -points in the three different directions. Choosing appropriate directions for the quasi-holes, the momentum for the original and decayed state are the same, up to a reciprocal lattice vector.)

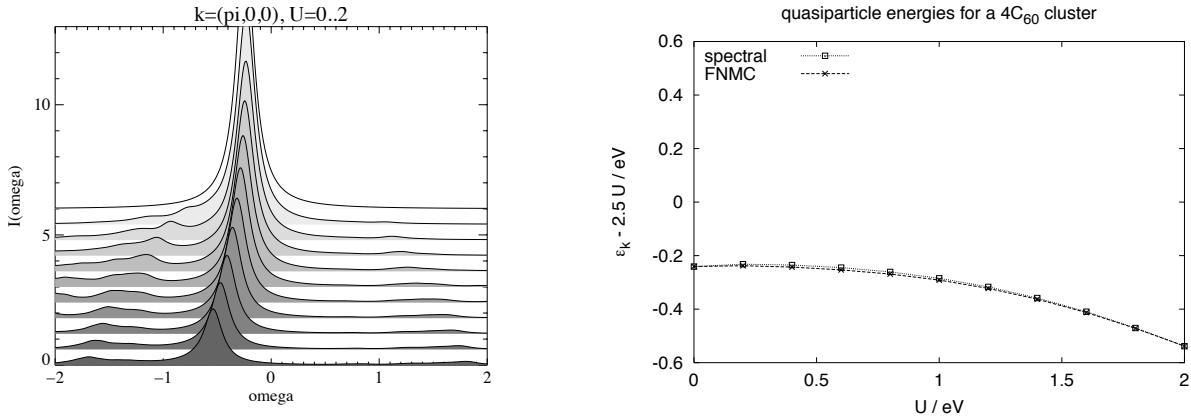


Fig. 17: Three-body quasi-hole $c_{X_2} c_G^\dagger c_{X_2}$: spectral function and fixed-node energies.

shown in Figure 17. As expected, the position of the peak corresponds closely to the energy of the decayed state found in the Lanczos calculation of Figure 14. Moreover, performing a fixed-node calculation with the trial function $c_{X_2} c_G^\dagger c_{X_2} |\Phi_0\rangle$ again gives excellent agreement with the exact quasi-particle energy (see right panel of Figure 17). Working with such more complex quasi-particles, we could extract quasi-particle interactions. A more straightforward application is to look at how the quasi-particle dispersion changes with U , as shown in Figure 18. This can be related to the effective mass m^* defined as

$$\frac{k_F}{m^*} = \left. \frac{d\varepsilon_k}{dk} \right|_{k_F} . \quad (93)$$

Rewriting the derivative as a finite difference, for the ratio of the effective mass to the bare mass at $U = 0$ we obtain

$$\frac{m^*}{m_0} \approx \frac{\Delta\varepsilon_0}{\Delta\varepsilon} . \quad (94)$$

The ratio on the right-hand side is also plotted in Figure 18. Even though the k -points for the four molecule cluster are quite far apart, so that the finite difference is not a good approximation to the derivative, we get consistent results when comparing the results derived from the quasi-electron and the quasi-hole states.

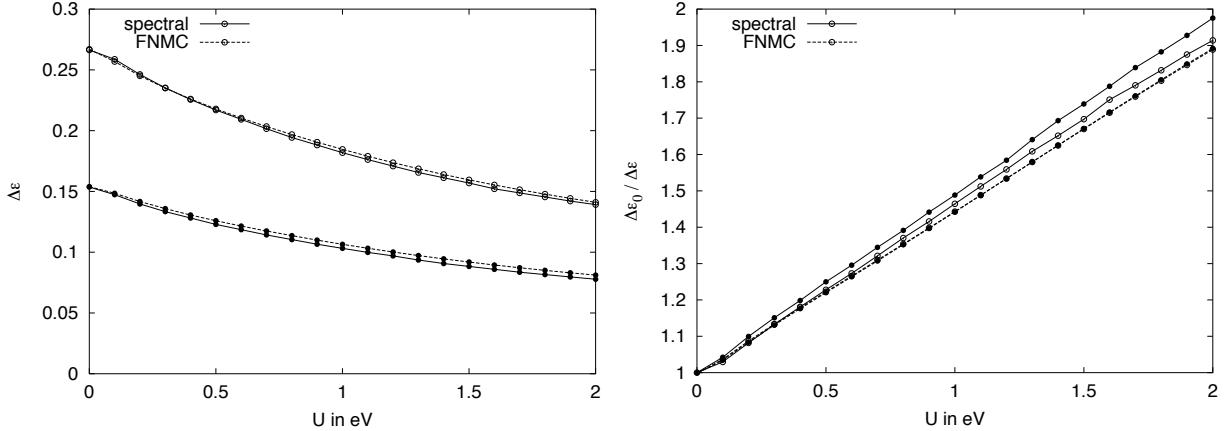


Fig. 18: Change in quasi-particle dispersion with U . The left-hand panel shows the energy difference between the two quasi-electron (filled circles) and the two simple quasi-hole (open circles) states. The full line connects the exact results, the full dashed line the fixed-node results. The right-hand panel shows the inverse of the energy difference normalized by its value at $U = 0$.

4 Conclusion

Practical many-body calculations need to be done on finite systems. A standard approach for reducing a solid to a finite system is to introduce periodic boundary conditions on finite clusters. How well such a periodic cluster represents the infinite solid depends on its shape. We have seen how we can use the Hermite normal form to systematically enumerate all distinct clusters. Introducing more general twisted boundary conditions, we can study the dependence of the system on the boundary conditions, providing an elegant criterion for distinguishing metals from (Mott) insulators.

We have demonstrated how to use the local energy to derive compact variational wave functions that incorporate a correlation-hole based on the cusp conditions. In contrast to a CI-expansion, such Jastrow/Gutzwiller functions need only few parameters to efficiently describe correlation effects. Still, simple correlation factors are usually not sufficient to describe a metal-insulator transition. To improve on the wave function, we can use projection techniques. In principle, they converge to the ground state, in practice their stochastic implementation is, however, hampered by the appearance of negative matrix elements – the fermion sign problem. The fixed-node approximation is a practical way for eliminating the sign problem without compromising accuracy. Moreover, fixing the nodes can be used to stabilize quasiparticle states.

Acknowledgment

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

References

- [1] E. Koch: *Many-body states*, in [33]
- [2] A. Szabo and N.S. Ostlund: *Modern Quantum Chemistry* (Dover Publications, 1996)
- [3] M. Born and T. von Kármán, Physikal. Z. **13**, 297 (1912)
- [4] W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001)
- [5] E. Wigner, Phys. Rev. **46**, 1002 (1934)
- [6] W. Kohn, Phys. Rev. **133**, A171 (1964)
- [7] M. Newman: *Integral Matrices* (Academic Press, New York, 1972)
- [8] H. Cohen: *A Course in Computational Algebraic Number Theory* (Springer, Heidelberg, 1993)
- [9] D.D. Betts, H.Q. Lin, and J.S. Flynn, Can. J. Phys. **77**, 353 (1999)
- [10] H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976)
- [11] E. Koch: *The Lanczos Method*, in [31]
- [12] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957)
- [13] B.L. Hammond, W.A. Lester, P.J. Reynolds: *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, 1994)
- [14] R. Jastrow, Phys. Rev. **98**, 1479 (1955)
- [15] S. Fahy, X.W. Wang, and S.G. Louie, Phys. Rev. B **42**, 3502 (1990)
- [16] E. Pavarini and E. Koch: *Introduction*, in [31]
- [17] M.C. Gutzwiller, Phys. Rev. **137**, A1726 (1965)
- [18] J. Bünenmann: *The Gutzwiller Density Functional Theory* in [32]
- [19] D.M. Ceperley, G.V. Chester, and M.H. Kalos, Phys. Rev. B **16**, 3081 (1977)
- [20] P. Horsch and T.A. Kaplan, J. Phys. C **16**, L1203 (1983);
H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 1490 (1987)
- [21] E. Koch, O. Gunnarsson, and R.M. Martin, Phys. Rev. B **59**, 15632 (1999)
- [22] C. Umrigar, K. Wilson, and J. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)
- [23] E. Koch, O. Gunnarsson, and R.M. Martin, Phys. Rev. B **60**, 15714 (1999)

- [24] D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)
- [25] W.F. Brinkman and T.M. Rice, Phys. Rev. B **2**, 4302 (1970);
T.M. Rice and W.F. Brinkman, in *Critical Phenomena in Alloys, Magnets, and Superconductors*, ed. by R.E. Mills, E. Ascher, R. Jaffee (McGraw Hill, New York, 1971)
- [26] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987)
and Phys. Rev. B **37**, 7382 (1988)
- [27] A.J. Millis and S.N. Coppersmith, Phys. Rev. B **43**, 13770 (1991)
- [28] N. Trivedi and D.M. Ceperley, Phys. Rev. B **40**, 2737 (1989)
- [29] D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, and W. van Saarloos, Phys. Rev. Lett. **72**, 2442 (1994) and D.F.B. ten Haaf, H.J.M van Bemmel, J.M.J. van Leeuwen, W. van Saarloos, and D.M. Ceperley, Phys. Rev. B **51**, 13039 (1995)
- [30] E. Koch, O. Gunnarsson, and R.M. Martin, Phys. Rev. Lett. **83**, 620 (1999)
- [31] E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein (eds.):
The LDA+DMFT approach to strongly correlated materials
Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
<http://www.cond-mat.de/events/correl11>
- [32] E. Pavarini, E. Koch, F. Anders, and M. Jarrell (eds.):
Correlated Electrons: From Models to Materials
Reihe Modeling and Simulation, Vol. 2 (Forschungszentrum Jülich, 2012)
<http://www.cond-mat.de/events/correl12>
- [33] E. Pavarini, E. Koch, and U. Schollwöck (eds.):
Emergent Phenomena in Correlated matter
Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
<http://www.cond-mat.de/events/correl13>

9 Making Use of Self-Energy Functionals: The Variational Cluster Approximation

Michael Potthoff

I. Institut für Theoretische Physik
Universität Hamburg

Contents

| | |
|---|-----------|
| 1 Motivation | 2 |
| 2 The cluster approach | 4 |
| 2.1 Tiling the lattice into small clusters | 4 |
| 2.2 Cluster perturbation theory | 5 |
| 2.3 Green's function and exact diagonalization | 5 |
| 2.4 Freedom in the CPT construction | 6 |
| 2.5 The Ritz principle? | 7 |
| 3 Diagrammatic perturbation theory | 9 |
| 3.1 S -matrix and Green's function | 9 |
| 3.2 Scattering at the inter-cluster potential, diagrammatically | 11 |
| 3.3 Diagram language for systems with Coulomb interaction | 12 |
| 3.4 Diagrammatic derivation of the CPT | 14 |
| 4 Self-energy functional theory | 16 |
| 4.1 Luttinger-Ward generating functional | 16 |
| 4.2 Self-energy functional | 17 |
| 4.3 Evaluation of the self-energy functional | 18 |
| 4.4 The variational cluster approximation | 20 |
| 5 Implementation of the variational cluster approximation | 21 |
| 6 Selected results | 25 |
| 6.1 One-dimensional Hubbard model | 25 |
| 6.2 Antiferromagnetism | 26 |
| 6.3 Mott metal-insulator transition | 28 |
| 7 Relation to other methods and conclusions | 30 |

1 Motivation

Self-energy-functional theory (SFT) [1–4] is a general theoretical framework that can be used to construct various approximate approaches by which the thermal properties and the spectrum of one-particle excitations of a certain class of correlated electron systems can be studied. The prototype considered here is the single-band Hubbard model [5–7], but quite generally, the SFT applies to models of strongly correlated fermions on three or lower-dimensional lattices with local interactions.

There are several extensions of the theory, e.g. to systems with non-local interactions [8], to bosonic systems [9, 10] and the Jaynes-Cummings lattice [11, 12], to electron-phonon systems [13], to systems with quenched disorder [14], as well as for the study of the real-time dynamics of systems far from thermal equilibrium [15]. To be concise, those extensions will not be covered here.

The prime example of an approximation that can be constructed within the SFT is the variational cluster approximation (VCA) [2, 16]. Roughly, one of the main ideas of the VCA is to adopt a *divide and conquer* strategy: A tiling of the original lattice into disconnected small clusters is considered, as shown in Fig. 1, for example. While the Hubbard model on the infinite square lattice cannot be solved exactly, there are no serious practical problems in solving the same model for an isolated cluster or for a set of disconnected clusters. The VCA constructs an approximate solution for the infinite lattice from the solution of the individual clusters by means of all-order perturbation theory for those terms in the Hamiltonian that connect the clusters.

This is actually the concept of the so-called cluster perturbation theory (CPT) [17, 18]. However, it is not sufficient in most cases, and we would like to go beyond the CPT. The essential problem becomes apparent, e.g., for a system with spontaneously broken symmetry such as an antiferromagnet. The antiferromagnetic state is characterized by a finite value for the sublattice magnetization which serves as an order parameter. On the other hand, quite generally, the order parameter must be zero for a system of finite size and thus for a small cluster in particular. Coupling finite (and thus necessarily non-magnetic) clusters by means of the CPT, however, one never gets to an antiferromagnetic solution for the infinite lattice. Divide and conquer is not sufficient to describe the emergence of new phases with broken symmetries.

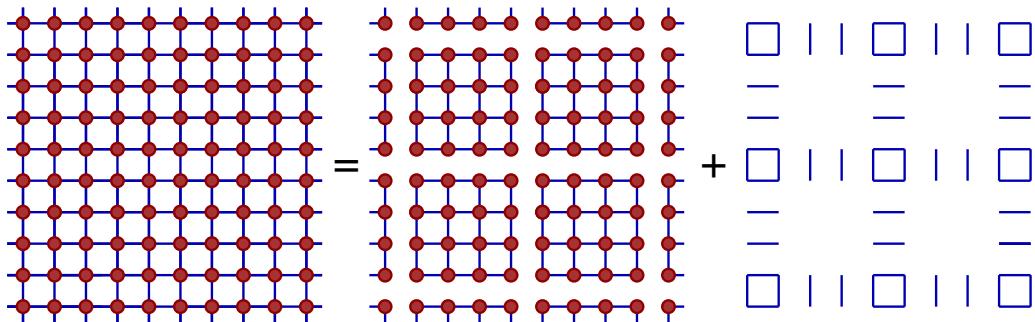


Fig. 1: Sketch of the decomposition of the original system $H = H_0(\mathbf{t}) + H_1$ into a reference system $H' = H_0(\mathbf{t}') + H_1$ and the inter-cluster hopping $H_0(\mathbf{V})$ for a square lattice and cluster size $L_c = 16$. Blue lines: nearest-neighbor hopping t . Red dots: on-site Hubbard interaction.

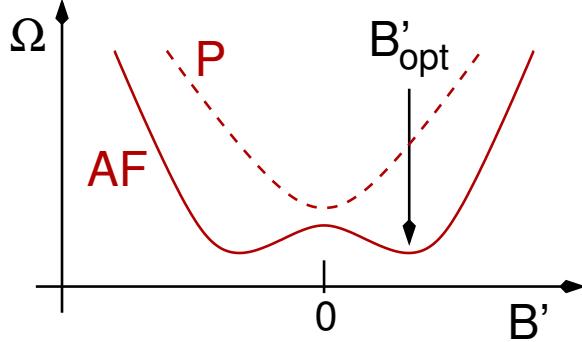


Fig. 2: Grand potential Ω as a function of a Weiss field B' in the case of a paramagnet (P) and in the case of an antiferromagnet (AF). B' is a fictitious staggered field, the optimal value of which (B'_{opt}) must be determined by minimization of Ω . As there is no physically applied staggered field, i.e. $B = 0$, a finite B'_{opt} indicates spontaneous symmetry breaking.

An obvious way out is to *enforce* a finite antiferromagnetic order parameter within each of the isolated clusters by applying a (staggered) magnetic field B' . Coupling those antiferromagnetic clusters may then result in an antiferromagnetic solution for the entire lattice.

However, what determines the strength of this magnetic field? As we are aiming at a description of *spontaneous* antiferromagnetic order, there is no external *physical* field B that is applied to the original system ($B = 0$). The field B' is actually a *Weiss field*, i.e. a fictitious field or *mean field* that is produced by the system itself. We are seeking for a formalism that allows for the formation of a finite Weiss field if this is favorable, i.e. if a thermodynamical potential can be lowered in this way.

Self-energy-functional theory provides a relation $\Omega(B')$ between the grand potential of the system Ω and the Weiss field B' that can be used to fix the optimal value B'_{opt} of the staggered magnetic field by minimization (see Fig. 2):

$$\left. \frac{\partial \Omega(B')}{\partial B'} \right|_{B'=B'_{\text{opt}}} \stackrel{!}{=} 0 \quad (1)$$

The purpose of this lecture is to show how this can be achieved in practice. To this end we have to answer the following *how to* questions:

- How can we solve the problem for an isolated cluster?
- With this at hand, how can we construct a solution for the problem on the infinite lattice?
- How can we construct the relation $\Omega(B')$ such that Eq. (1) determines B'_{opt} ?

Actually, there is no reason to consider only a staggered magnetic field as a Weiss field. Another goal is therefore to generalize the idea to arbitrary Weiss fields or to an arbitrary set of variational parameters λ' that characterize the isolated cluster and that are optimized via $\partial \Omega(\lambda'_{\text{opt}})/\partial \lambda' \stackrel{!}{=} 0$. Finally, the VCA should be compared with other theories available, and its practical as well as fundamental limitations have to be discussed.

2 The cluster approach

2.1 Tiling the lattice into small clusters

We start with the second question and consider a simple non-interacting system given by

$$H_0 = \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} = H_0(\mathbf{t}) . \quad (2)$$

Here, $c_{i\sigma}^\dagger$ creates an electron with spin $\sigma = \uparrow, \downarrow$ at the site i of a D -dimensional lattice, and t_{ij} are the (spin-independent) hopping parameters, which are also considered as the elements of the hopping matrix \mathbf{t} . Furthermore,

$$H'_0 = \sum_{ij\sigma} t'_{ij} c_{i\sigma}^\dagger c_{j\sigma} = H_0(\mathbf{t}') , \quad (3)$$

denotes the Hamiltonian of the system with decoupled clusters (see Fig. 1 and take $H_1 = 0$). If L is the number of lattice sites in the original lattice model H_0 and L_c is the number of sites in an individual cluster, there are L/L_c decoupled clusters. We assume that all clusters are identical. In terms of hopping matrices, we have

$$\mathbf{t} = \mathbf{t}' + \mathbf{V} \quad (4)$$

where \mathbf{V} is the inter-cluster hopping.

Consider the resolvent of the hopping matrix, i.e. the Green's function

$$\mathbf{G}_0(\omega) = \frac{1}{\omega + \mu - \mathbf{t}} . \quad (5)$$

Here, ω is a complex frequency (units with $\hbar = 1$ are used). We have also introduced the chemical potential μ (which is not important here but used later). Furthermore, we employ a matrix notation and write ω rather than $\omega \mathbf{1}$ etc. for short. Note that $(\dots)^{-1}$ and $1/(\dots)$ both mean matrix inversion.

Having the Green's function of the reference system at hand,

$$\mathbf{G}'_0(\omega) = \frac{1}{\omega + \mu - \mathbf{t}'} , \quad (6)$$

how can we get the Green's function of the original model? With some algebra, one easily derives the equation

$$\mathbf{G}_0(\omega) = \mathbf{G}'_0(\omega) + \mathbf{G}'_0(\omega) \mathbf{V} \mathbf{G}_0(\omega) \quad (7)$$

which is solved by

$$\mathbf{G}_0(\omega) = \frac{1}{\mathbf{G}'_0(\omega)^{-1} - \mathbf{V}} . \quad (8)$$

We see that using Green's functions it is formally rather easy to couple a system of isolated clusters.

2.2 Cluster perturbation theory

Actually, we are interested in interacting systems. For the single-orbital model H_0 , the only possible local interaction is a Hubbard interaction of the form

$$H_1 = \frac{U}{2} \sum_{i\sigma} n_{i\sigma} n_{i-\sigma} \quad (9)$$

with $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$, where U is the interaction strength. Since H_1 is completely local, the Hamiltonian of the so-called *reference system* $H' = H_0(\mathbf{t}') + H_1$ is obtained from the Hamiltonian of the original system $H = H_0(\mathbf{t}) + H_1$ by switching off the inter-cluster hopping \mathbf{V} .

For a small cluster and likewise for a system of disconnected clusters, even for the interacting case, it is comparatively simple to solve the problem exactly (by numerical means if necessary), while for the original lattice model this is a hard problem. One therefore cannot expect a simple relation between the original and the reference system like Eq. (7). Nevertheless, as it is too tempting, we will write down

$$\mathbf{G}(\omega) = \mathbf{G}'(\omega) + \mathbf{G}'(\omega) \mathbf{V} \mathbf{G}(\omega) \quad (10)$$

where now \mathbf{G} and \mathbf{G}' are interacting Green's functions. This is an equation that constitutes the cluster-perturbation theory [17, 18]. It must be seen as an approximate way to compute the Green's function of the interacting model from the exact cluster Green's function. In a way the approximation is controlled by the size L_c of the clusters in the reference system since for $L_c \rightarrow \infty$ one can expect the approximation to become exact. In fact, the CPT is not too bad and has been successfully applied in a couple of problems, see Ref. [19] and references therein.

2.3 Green's function and exact diagonalization

Before proceeding with the interpretation of the CPT equation (10), which provides an approximate expression for $\mathbf{G}(\omega)$, let us give the *exact* definition of the Green's function for the interacting case. Its elements are defined as

$$G_{ij\sigma}(\omega) = \int_{-\infty}^{\infty} dz \frac{A_{ij\sigma}(z)}{\omega - z}, \quad (11)$$

where ω is an arbitrary complex frequency and where

$$A_{ij\sigma}(z) = \int_{-\infty}^{\infty} dt e^{izt} A_{ij\sigma}(t) \quad (12)$$

is the single-particle spectral density whose Fourier transform

$$A_{ij\sigma}(t) = \frac{1}{2\pi} \langle [c_{i\sigma}(t), c_{j\sigma}^\dagger(0)]_+ \rangle \quad (13)$$

is given as the thermal expectation value of the anti-commutator of the annihilator with the creator in the (grand-canonical) Heisenberg picture, e.g.

$$c_{i\sigma}(t) = e^{i(H-\mu N)t} c_{i\sigma} e^{-i(H-\mu N)t} \quad (14)$$

with $N = \sum_{i\sigma} n_{i\sigma} = \sum_{i\sigma} c_{i\sigma}^\dagger c_{i\sigma}$. The thermal average is a grand-canonical average $\langle \dots \rangle = Z^{-1} \text{Tr}(e^{-\beta(H-\mu N)} \dots)$, where $Z = Z(\beta, \mu) = \text{Tr } e^{-\beta(H-\mu N)}$ is the partition function at chemical potential μ and inverse temperature β .

In the case of a non-interacting system, one may use the Baker-Campbell-Hausdorff formula to get the simple result (\mathbf{t} : hopping, t : time):

$$c_{i\sigma}(t) = \sum_j (e^{-i(\mathbf{t}-\mu)t})_{ij} c_{j\sigma} \quad (15)$$

which can be used in Eq. (13), and then via the Fourier transformation (12) and finally the Hilbert transformation (11) one arrives at the result given by Eq. (5) above.

In the interacting case ($U > 0$), one may compute the Green's function from the eigenvalues E_n and eigenstates $|n\rangle$ of the (grand-canonical) Hamiltonian:

$$(H - \mu N)|n\rangle = E_n|n\rangle. \quad (16)$$

Using a resolution of the unity $\mathbf{1} = \sum_n |n\rangle\langle n|$ in Eq. (13), one can easily do the calculation and arrives at

$$G_{ij\sigma}(\omega) = \frac{1}{Z} \sum_{mn} \frac{(e^{-\beta E_m} + e^{-\beta E_n}) \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle}{\omega - (E_n - E_m)}. \quad (17)$$

However, as one must solve the many-body energy eigenvalue problem (16), this way of calculating the Green's function is obviously impossible in practice for the Hamiltonian of the original system – the Hilbert-space dimension exponentially increases with L . On the other hand, for the reference system and if the size of the cluster L_c is not too large, this can be done numerically. For a half-filled system ($N = L_c$), up to $L_c = 8$ sites can be managed in this way easily. At zero temperature, using the Lanczos algorithm [20], the Green's function for somewhat larger clusters can be computed, typically $L_c \leq 12$ at half-filling. This already answers the first question posed in the introduction.

2.4 Freedom in the CPT construction

The CPT gives a preliminary answer to the second question. However, it is easily seen that the answer is not unique: consider a *modified* reference system with a Hamiltonian

$$H_0(\tilde{\mathbf{t}}') + H_1 = H_0(\mathbf{t}') + H_0(\Delta\mathbf{t}) + H_1 = \sum_{ij\sigma} (t'_{ij} + \Delta t_{ij}) c_{i\sigma}^\dagger c_{j\sigma} + H_1, \quad (18)$$

i.e. a reference system where $\mathbf{t}' \mapsto \tilde{\mathbf{t}}' = \mathbf{t}' + \Delta\mathbf{t}'$. The new reference system shall still describe the same set of decoupled clusters but with different intra-cluster hoppings $\tilde{\mathbf{t}}'$. The modified non-interacting Green's function of the reference system is $\tilde{\mathbf{G}}'_0(\omega) = 1/(\omega + \mu - \tilde{\mathbf{t}}')$. Now, the non-interacting Green's function of the original model is obtained from the equation

$$\mathbf{G}_0(\omega) = \tilde{\mathbf{G}}'_0(\omega) + \tilde{\mathbf{G}}'_0(\omega) \tilde{\mathbf{V}} \mathbf{G}_0(\omega) \quad (19)$$

with the modified inter-cluster hopping $\tilde{\mathbf{V}} = \mathbf{t} - \tilde{\mathbf{t}'} = \mathbf{V} - \Delta\mathbf{t}'$. $\mathbf{G}_0(\omega)$ can be considered as the limit of a geometrical series that is found by iterating equation (19):

$$\mathbf{G}_0(\omega) = \tilde{\mathbf{G}}'(\omega) + \tilde{\mathbf{G}}'(\omega)\tilde{\mathbf{V}}\tilde{\mathbf{G}}'(\omega) + \dots . \quad (20)$$

We infer that $\mathbf{G}_0(\omega)$ can be obtained by (all-order) perturbation theory in $\tilde{\mathbf{V}}$ when expanding around the Green's function of the modified reference system given by the hopping matrix $\tilde{\mathbf{t}'}$. Obviously, the same result is obtained by perturbation theory in \mathbf{V} around the Green's function of the modified reference system with hopping matrix \mathbf{t}' . This freedom in choosing the starting point for perturbation theory that we have in the non-interacting case turns into a real problem for the interacting case. Namely, since the CPT equation (10) is approximate, we generally have:

$$\tilde{\mathbf{G}}(\omega) \equiv \tilde{\mathbf{G}}'(\omega) + \tilde{\mathbf{G}}'(\omega)\tilde{\mathbf{V}}\tilde{\mathbf{G}}'(\omega) + \dots \neq \mathbf{G}'(\omega) + \mathbf{G}'(\omega)\mathbf{V}\mathbf{G}'(\omega) + \dots \equiv \mathbf{G}(\omega) . \quad (21)$$

Concluding, different starting points \mathbf{t}' and $\tilde{\mathbf{t}'}$ for the all-order cluster perturbation theory in \mathbf{V} and $\tilde{\mathbf{V}}$ lead to different results $\mathbf{G}(\omega)$ and $\tilde{\mathbf{G}}(\omega)$, respectively.

But which is the *right* starting point? The idea is to turn the problem into an advantage by *optimizing* the starting point: this can be done by making use of a variational principle, i.e. by expressing a thermodynamical potential, e.g. the grand potential Ω , as a function of \mathbf{t}' and by subsequent minimization. The optimal \mathbf{t}'_{opt} shall be obtained by

$$\left. \frac{\partial \Omega(\mathbf{t}')}{\partial \mathbf{t}'} \right|_{\mathbf{t}'=\mathbf{t}'_{\text{opt}}} \stackrel{!}{=} 0 . \quad (22)$$

We see that the set of variational parameters is just the set of hopping parameters of the reference systems or, in the case of multi-orbital models, simply the set of *all* one-particle parameters except for those, of course, that would couple the different clusters. This set also includes a staggered magnetic field

$$H_0(\tilde{\mathbf{t}'}) = H_0(\mathbf{t}') - B' \sum_i z_i(n_{i\uparrow} - n_{i\downarrow}) , \quad (23)$$

where $z_i = \pm 1$ alternates between the sites of a bipartite lattice.

2.5 The Ritz principle?

The most popular variational principle is the Ritz variational principle. It states that

$$E[|\Psi\rangle] = \langle \Psi | H | \Psi \rangle = \min. \quad (24)$$

for the ground state of H when the search extends over all normalized trial states $\langle \Psi | \Psi \rangle = 1$. Evaluated at the ground state $|\Psi_0\rangle$, the functional yields the ground-state energy $E[|\Psi_0\rangle] = E_0$.

Hence, a straightforward idea that suggests itself is to compute the normalized ground state $|\Psi(\mathbf{t}')\rangle$ of a reference system with hopping matrix \mathbf{t}' and to use this as a trial state. The trial state can be varied by varying the parameters \mathbf{t}' , and the optimal parameters are given by

$$\frac{\partial E[|\Psi(\mathbf{t}')\rangle]}{\partial \mathbf{t}'} \Bigg|_{\mathbf{t}'=\mathbf{t}'_{\text{opt}}} \stackrel{!}{=} 0. \quad (25)$$

To test this idea, let

$$|\Psi(\mathbf{t}')\rangle = |\Psi_1(\mathbf{t}'_1)\rangle \otimes |\Psi_2(\mathbf{t}'_2)\rangle \otimes \cdots \otimes |\Psi_{L/L_c}(\mathbf{t}'_{L/L_c})\rangle \quad (26)$$

be the ground state of $H' = H_0(\mathbf{t}') + H_1$. It is given as a product of the ground states of the L/L_c individual clusters where the ground state of the I -th cluster with hopping matrix \mathbf{t}'_I is $|\Psi_I(\mathbf{t}'_I)\rangle$. Now, if $E_0(\mathbf{t}')$ denotes the ground-state energy of the reference system,

$$E[|\Psi(\mathbf{t}')\rangle] = \langle \Psi(\mathbf{t}') | (H_0(\mathbf{t}') + H_1) | \Psi(\mathbf{t}') \rangle = E_0(\mathbf{t}') + \langle \Psi(\mathbf{t}') | H_0(\mathbf{V}) | \Psi(\mathbf{t}') \rangle. \quad (27)$$

However, the inter-cluster hopping Hamiltonian $H_0(\mathbf{V})$ only contains terms like $c_{i\sigma}^\dagger c_{j\sigma}$ where the sites i and j belong to *different* clusters, say I and J . Hence, $\langle \Psi(\mathbf{t}') | c_{i\sigma}^\dagger c_{j\sigma} | \Psi(\mathbf{t}') \rangle = \langle \Psi_I(\mathbf{t}'_I) | \otimes \langle \Psi_J(\mathbf{t}'_J) | c_{i\sigma}^\dagger c_{j\sigma} | \Psi_J(\mathbf{t}'_J) \rangle \otimes \langle \Psi_I(\mathbf{t}'_I) | = \langle \Psi_I(\mathbf{t}'_I) | c_{i\sigma}^\dagger | \Psi_I(\mathbf{t}'_I) \rangle \langle \Psi_J(\mathbf{t}'_J) | c_{j\sigma} | \Psi_J(\mathbf{t}'_J) \rangle = 0$ as enforced by the conservation of the total particle number. This means that we are left with $E[|\Psi(\mathbf{t}')\rangle] = E_0(\mathbf{t}')$. As this implies that the optimal parameters \mathbf{t}'_{opt} do not at all depend on \mathbf{V} , the result is trivial and useless, unfortunately. Even worse, the Hellmann-Feynman theorem [21] tells us that

$$\frac{\partial}{\partial \mathbf{t}'} E[|\Psi(\mathbf{t}')\rangle] = \frac{\partial}{\partial \mathbf{t}'} E_0(\mathbf{t}') = \frac{\partial}{\partial \mathbf{t}'} \langle \Psi(\mathbf{t}') | (H_0(\mathbf{t}') + H_1) | \Psi(\mathbf{t}') \rangle = \langle \Psi(\mathbf{t}') | \frac{\partial H_0(\mathbf{t}')}{\partial \mathbf{t}'} | \Psi(\mathbf{t}') \rangle. \quad (28)$$

This means that, using the Ritz principle, the variational parameters should be determined such that all one-particle intra-cluster correlation functions $\langle c^\dagger c \rangle$, in addition to the inter-cluster correlation functions, vanish.

Concluding, optimizing cluster-perturbation theory cannot be done with the help of the Ritz principle. We mention in passing that this also holds for its finite-temperature and mixed state generalization [22, 23]

$$\Omega[\rho] = \text{Tr}\left(\rho(H - \mu N + T \ln \rho)\right) \stackrel{!}{=} \min., \quad (29)$$

where the grand potential, expressed as a functional of the density matrix, is at a minimum for the thermal density matrix $\rho = \exp(-\beta(H - \mu N)) / \text{Tr} \exp(-\beta(H - \mu N))$. While this is an extremely useful variational principle, it cannot be used here: a trial density matrix $\rho(\mathbf{t}')$, defined as the thermal density matrix of a reference system with a hopping matrix \mathbf{t}' that describes decoupled clusters, is a simple product of individual cluster density matrices only. As for the standard Ritz principle, this implies that inter-cluster one-particle correlations are neglected altogether.

3 Diagrammatic perturbation theory

3.1 *S*-matrix and Green's function

As we have already seen, Green's functions, as opposed to wave functions or density matrices, can be used to couple isolated clusters. All-order perturbation theory in the inter-cluster hopping \mathbf{V} yields the exact Green's function in the non-interacting ($U = 0$) case and an approximate (CPT) Green's function for $U > 0$. For the necessary optimization of the starting point, i.e. of the intra-cluster one-particle parameters \mathbf{t}' , we should therefore try to formulate a variational principle based on Green's functions, i.e. a principle of the form $\delta\Omega[\mathbf{G}(\omega)]/\delta\mathbf{G}(\omega) \stackrel{!}{=} 0$, and try *test Green's functions* $\mathbf{G}'(\omega)$ taken from the reference system. In fact, a variational principle of this type can be constructed with the help of all-order perturbation theory in U [24, 25]. *Vice versa*, a systematic and general perturbation theory in U (and also in \mathbf{V}) requires putting Green's functions at the focus of the theory. Here, only a brief sketch is given, details can be found in Refs. [25–27], for example. Our goal is to use diagrammatic perturbation theory as a language that can be used to formulate a Green's-function-based variational principle.

We decompose the (grand-canonical) Hamiltonian $\mathcal{H} \equiv H - \mu N$ into a *free* part $\mathcal{H}_0 = H_0 - \mu N$ and the interaction $H_1 \equiv \mathcal{H} - \mathcal{H}_0$. Next we define, for $0 \leq \tau, \tau' \leq \beta$, the so-called *S*-matrix as

$$S(\tau, \tau') = e^{\mathcal{H}_0\tau} e^{-\mathcal{H}(\tau-\tau')} e^{-\mathcal{H}_0\tau'} , \quad (30)$$

One may interpret $\tau = it$ as an *imaginary-time* variable (where t is real). This *Wick rotation* in the complex time plane has the formal advantage that the thermal density matrix, $\propto e^{-\beta\mathcal{H}}$, is just given by the time-evolution operator, $e^{-i\mathcal{H}t} = e^{-\mathcal{H}\tau}$ at $\tau = \beta$.

There are two main purposes of the *S*-matrix. First, it can be used to rewrite the partition function in the following way:

$$Z = \text{Tr } e^{-\beta\mathcal{H}} = \text{Tr} (e^{-\beta\mathcal{H}_0} e^{\beta\mathcal{H}_0} e^{-\beta\mathcal{H}}) = \text{Tr} (e^{-\beta\mathcal{H}_0} S(\beta, 0)) = Z_0 \langle S(\beta, 0) \rangle^{(0)} . \quad (31)$$

The partition function of the interacting system is thereby given in terms of the partition function of the *free* system, which is known, and a *free* thermal expectation value of the *S*-matrix. The second main purpose is related to the imaginary-time Green's function which, for $-\beta < \tau < \beta$, is defined via

$$G_{ij\sigma}(\tau) = -\langle \mathcal{T} c_{i\sigma}(\tau) c_{j\sigma}^\dagger(0) \rangle \quad (32)$$

in terms of an annihilator and a creator with imaginary Heisenberg time dependence:

$$c_{i\sigma}(\tau) = e^{\mathcal{H}\tau} c_{i\sigma} e^{-\mathcal{H}\tau} , \quad c_{j\sigma}^\dagger(\tau) = e^{\mathcal{H}\tau} c_{j\sigma}^\dagger e^{-\mathcal{H}\tau} . \quad (33)$$

Furthermore, \mathcal{T} is the (imaginary) time-ordering operator. With the help of the *S*-matrix, the interacting time dependence can be transformed into a *free* time dependence, namely:

$$c_{i\sigma}(\tau) = S(0, \tau) c_{I,i\sigma}(\tau) S(\tau, 0) , \quad c_{j\sigma}^\dagger(\tau) = S(0, \tau) c_{I,j\sigma}^\dagger(\tau) S(\tau, 0) . \quad (34)$$

Here, the index I (*interaction picture*) indicates that the time dependence is due to \mathcal{H}_0 only. This time dependence is simple and can, again, be derived with the Baker-Campbell-Hausdorff formula:

$$c_{I,i\sigma}(\tau) = \sum_j (e^{-(\mathbf{t}-\mu)\tau})_{ij} c_{j\sigma}, \quad c_{I,i\sigma}^\dagger(\tau) = \sum_j (e^{+(\mathbf{t}-\mu)\tau})_{ij} c_{j\sigma}^\dagger. \quad (35)$$

Outside the imaginary-time interval $-\beta < \tau < \beta$, the Green's function is defined as the periodic continuation: $G_{ij\sigma}(\tau + k \cdot 2\beta) = G_{ij\sigma}(\tau)$ for any integer k . This function has a discrete Fourier representation:

$$G_{ij\sigma}(\tau) = \frac{1}{\beta} \sum_{n=-\infty}^{\infty} G_{ij\sigma}(i\omega_n) e^{-i\omega_n \tau}, \quad (36)$$

where the Fourier coefficients $G_{ij\sigma}(i\omega_n)$ are defined at the so-called fermionic Matsubara frequencies $i\omega_n = i(2n+1)\pi/\beta$ for integer n and can be computed from $G_{ij\sigma}(\tau)$ as

$$G_{ij\sigma}(i\omega_n) = \int_0^\beta d\tau G_{ij\sigma}(\tau) e^{i\omega_n \tau}. \quad (37)$$

The Green's function $G_{ij\sigma}(\tau)$ is just a different representation of the Green's function $G_{ij\sigma}(\omega)$ introduced with Eq. (11) as its Fourier coefficients are given by $G_{ij\sigma}(i\omega_n) = G_{ij\sigma}(\omega)|_{\omega=i\omega_n}$. The remaining problem consists in finding a much more suitable representation of the S -matrix. From its definition, one straightforwardly derives the following equation of motion:

$$-\frac{\partial}{\partial \tau} S(\tau, \tau') = H_{1,I}(\tau) S(\tau, \tau'). \quad (38)$$

Here, the time dependence of $H_{1,I}(\tau)$ is due to H_0 only. A formal solution of this differential equation with the initial condition $S(\tau, \tau) = 1$ can be derived easily using the time-ordering operator \mathcal{T} again:

$$S(\tau, \tau') = \mathcal{T} \exp \left(- \int_{\tau'}^{\tau} d\tau'' H_{1,I}(\tau'') \right). \quad (39)$$

Note that, if all quantities were commuting, the solution of Eq. (38) would trivially be given by Eq. (39) without \mathcal{T} . The appearance of \mathcal{T} can therefore be understood as necessary to enforce commutativity.

Using this S -matrix representation, the partition function and the Green's function can be written as:

$$\frac{Z}{Z_0} = \left\langle \mathcal{T} \exp \left(- \int_0^\beta d\tau'' H_{1,I}(\tau'') \right) \right\rangle^{(0)} \quad (40)$$

and

$$G_{ij\sigma}(\tau) = - \frac{\left\langle \mathcal{T} \exp \left(- \int_0^\beta d\tau H_{1,I}(\tau) \right) c_{I,i\sigma}(\tau) c_{I,j\sigma}^\dagger(0) \right\rangle^{(0)}}{\left\langle \mathcal{T} \exp \left(- \int_0^\beta d\tau H_{1,I}(\tau) \right) \right\rangle^{(0)}}. \quad (41)$$

The important point is that the expectation values and time dependencies appearing here are free and thus known. Therefore, expanding the exponentials in Eq. (40) and Eq. (41) provides

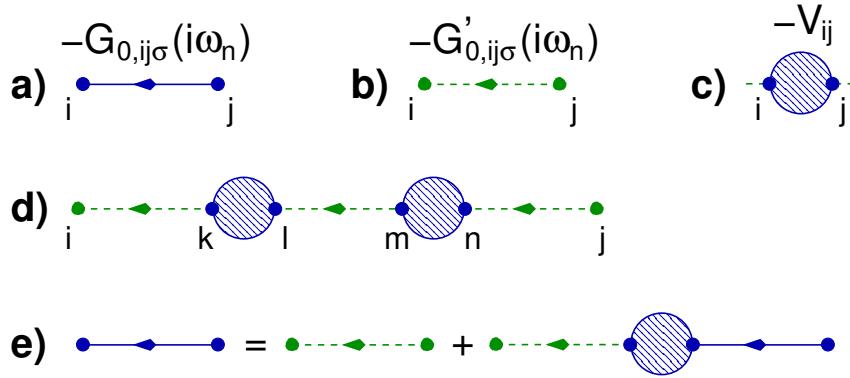


Fig. 3: Diagrams for potential scattering. See text for discussion.

an expansion of the partition function and of the Green's function in powers of the interaction strength. The coefficients of this expansion are given as free expectation values of time-ordered products of annihilators and creators with free time dependencies. In k th order, this is a k -particle free correlation function that can be simplified by using Wick's theorem. This is the central theorem of diagrammatic perturbation theory and applies to *free* higher-order correlation functions.

Consider the case of the partition function, as an example. At k -th order, the coefficient is given by a sum of $(2k)!$ terms, each of which factorizes into an k -fold product of terms of the form $\langle \mathcal{T} c_{i\sigma}(\tau) c_{j\sigma}^\dagger(\tau') \rangle^{(0)}$ called propagators. Apart from a sign, a propagator is nothing but the free Green's function. The summation of the $(2k)!$ terms is organized by means of a diagrammatic technique where *vertices* are linked via *propagators*. Wick's theorem and the details of the technique can be found in Refs. [25–27], for example.

3.2 Scattering at the inter-cluster potential, diagrammatically

Here, it is sufficient to illustrate the technique. To this end, we first consider the simple and exactly solvable system that is given by the Hamiltonian $\mathcal{H}_0(\mathbf{t}) = H_0(\mathbf{t}) - \mu N$ (see Eq. (2)). We decompose the Hamiltonian into a *free* part $\mathcal{H}_0(\mathbf{t}') = H_0(\mathbf{t}') - \mu N$ (see Eq. (3)) and an “interaction” $H_1 \equiv H_0(\mathbf{V})$ (see Eq. (4)). The “fully interacting” propagator, which we are interested in, is $-G_{0,ij\sigma}(i\omega_n)$ and is represented by an oriented line which starts at site j where the electron is created ($c_{j\sigma}^\dagger$) and ends at site i (see Fig. 3a). The *free* propagator $-G'_{0,ij\sigma}(i\omega_n)$ is represented by a dashed line (see Fig. 3b). Propagators carry a frequency $i\omega_n$ and a spin σ . A circle with two links, one for an incoming and one for an outgoing propagator, is called a vertex and stands for the “interaction” $-V_{ij}$ itself (see Fig. 3c). According to Wick's theorem, the contribution of order k to $-G_{0,ij\sigma}$ is obtained by drawing all topologically different diagrams where all links at k vertices are connected by free propagators, except for two external links at the sites i and j (see Fig. 3d). This contribution is calculated by performing the sums over *internal* variables (such as k, l, m, n in Fig. 3d) and respecting frequency and spin conservation at each vertex. These diagram rules can be derived strictly by expanding Eq. (41) and applying Wick's theo-

rem. Together with Eq. (40) this also leads to the important so-called linked-cluster theorem which allows us to concentrate on *connected* diagrams only. The disconnected diagrams for the Green's function (i.e. with external links) exactly cancel diagrammatic contributions from the denominator in Eq. (41). As concerns closed diagrams (no external links) contributing to the partition function, Eq. (40), the sum of only the connected closed diagrams yields, apart from a constant, $\ln Z$, i.e. the grand potential. For the simple case of scattering at the inter-cluster potential discussed at the moment, there is a single connected diagram at each order k only, and thus the “interacting” Green's function is given by

$$-G_{0,ij\sigma}(i\omega_n) = -G'_{0,ij\sigma}(i\omega_n) + \sum_{kl} [-G'_{0,ik\sigma}(i\omega_n)] [-V_{kl}] [-G'_{0,kj\sigma}(i\omega_n)] + \dots \quad (42)$$

or, using a matrix formulation and after elimination of the signs,

$$\begin{aligned} \mathbf{G}_0 &= \mathbf{G}'_0 + \mathbf{G}'_0 \mathbf{V} \mathbf{G}'_0 + \mathbf{G}'_0 \mathbf{V} \mathbf{G}'_0 \mathbf{V} \mathbf{G}'_0 + \dots \\ &= \mathbf{G}'_0 + \mathbf{G}'_0 \mathbf{V} (\mathbf{G}'_0 + \mathbf{G}'_0 \mathbf{V} \mathbf{G}'_0 + \dots) = \mathbf{G}'_0 + \mathbf{G}'_0 \mathbf{V} \mathbf{G}_0 . \end{aligned} \quad (43)$$

In this way we have simply re-derived Eq. (7) diagrammatically. This is not yet the CPT equation as the Hubbard interaction has been disregarded.

3.3 Diagram language for systems with Coulomb interaction

Next, let us consider the system given by the Hamiltonian $H_0(\mathbf{t}) + H_1$, see Eqs. (2) and (9), and treat the Hubbard (or Coulomb) term H_1 as the interaction, as usual. Also in this case, the free propagator is given by $-\mathbf{G}_0$ (Fig. 3a). To represent the interaction $-U$, we need a symbol (red dotted line) with four links, two for *outgoing* and two for *incoming* propagators (Fig. 4a) corresponding to the two creators and the two annihilators in the Hubbard interaction term. Note that the interaction is local and labeled by a site index and that there is energy and spin conservation at a vertex. A diagram contributing to the *interacting propagator* $-G_{ij\sigma}(i\omega_n)$ at order k consists of $2k+1$ propagators fully connecting the k vertices among each other and with the two external links at the sites i and j . Opposed to the potential-scattering problem discussed above, there are many more diagrams at a given order k , namely $(2k+1)!$, one of which, for $k=3$, is shown in Fig. 3b. The exact Green's function $\mathbf{G}(i\omega_n)$ is obtained by summing the algebraic expressions corresponding to those diagrams and summing over all k . The detailed rules necessary for the evaluation of diagrams (see Refs. [25–27]) are not needed here as we do *not* intend to construct a diagrammatically defined approximation by summing a certain subclass of diagrams. While this would be the standard procedure of many-body perturbation theory, here we just want to *speak diagrammatically*.

One can identify so-called self-energy insertions in the diagrammatic series, i.e. parts of diagrams that have links to two external propagators. Examples are given in Fig. 4c where we also distinguish between reducible and irreducible self-energy insertions. The reducible ones can be split into two disconnected parts by removal of a single propagator line. The self-energy

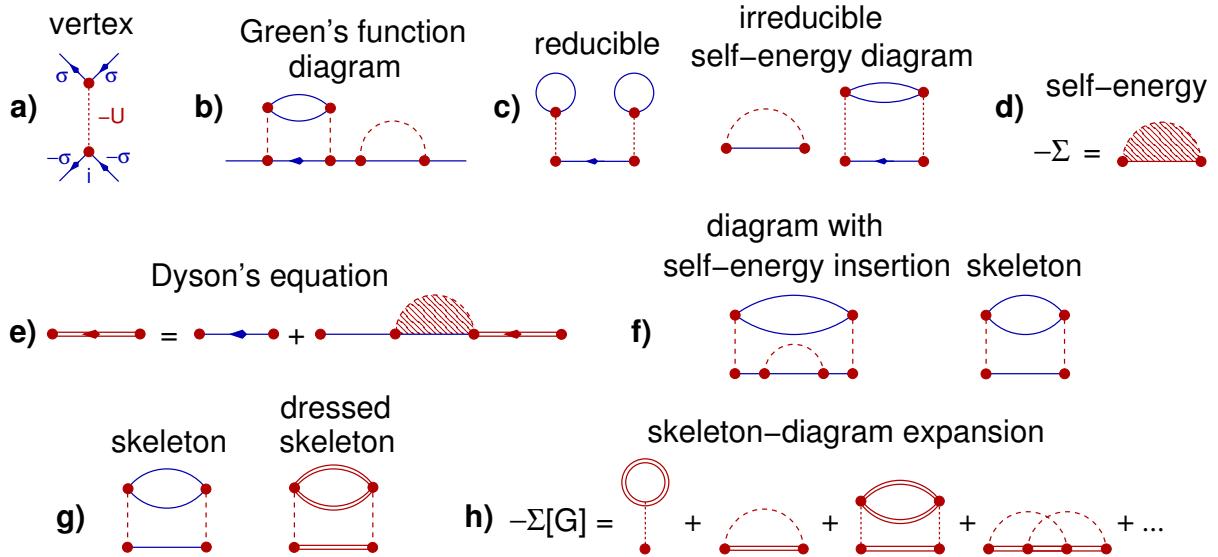


Fig. 4: Diagram language for systems with Hubbard interaction. See text for discussion.

$\Sigma_{ij\sigma}(i\omega_n)$ is then defined diagrammatically as the sum over all irreducible self-energy insertions, see Fig. 4d. With this we can derive Dyson's equation

$$\mathbf{G}(i\omega_n) = \mathbf{G}_0(i\omega_n) + \mathbf{G}_0(i\omega_n)\Sigma(i\omega_n)\mathbf{G}(i\omega_n) \quad (44)$$

corresponding to Fig. 4e. The double line stands for the interacting propagator $-\mathbf{G}(i\omega_n)$. Note that the self-energy plays the same role for the Coulomb interacting system as \mathbf{V} does for the scattering problem.

As the first diagram in Fig. 4f shows, there are irreducible self-energy diagrams that contain self-energy insertions. Diagrams without any self-energy insertion are called skeleton diagrams. Skeleton diagrams can be *dressed* by replacing in the diagram the free propagators with interacting propagators (double lines), see Fig. 4g. It is easy to see that the self-energy is given by the sum of the skeleton diagrams only, provided that these are dressed, see Fig. 4h. Therewith, the self-energy is given in terms of the interacting Green's function, $\Sigma = \Sigma[\mathbf{G}]$. It is only through diagrammatic language that this very important functional relationship, called skeleton-diagram expansion, can be defined rigorously. If combined with Dyson's equation (44), it provides us with a closed equation

$$\mathbf{G}(i\omega_n) = \frac{1}{\mathbf{G}_0(i\omega_n)^{-1} - \Sigma[\mathbf{G}](i\omega_n)} \quad (45)$$

the solution of which is the *exact* Green's function. It is clear, however, that the functional $\Sigma[\mathbf{G}]$ is extremely complicated and actually cannot be given in an explicit form, even for the most simple models such as the Hubbard model, and even in cases like small isolated Hubbard cluster, where a numerical computation of the self-energy and the Green's function is easily possible.

3.4 Diagrammatic derivation of the CPT

Equipped with the diagrammatic language, let us come back to the central topic. We have $H = H_0(\mathbf{t}') + H_0(\mathbf{V}) + H_1$ where the reference system $H' = H_0(\mathbf{t}') + H_1$ is easily solvable since it consists of decoupled small clusters, and where $H_0(\mathbf{V})$ is the inter-cluster hopping. Ideally, one would start from the solution of H' and perform a perturbative treatment of $H_0(\mathbf{V})$. This, however, is not possible (within the above-described standard perturbation theory) as the starting point H' is an interacting system and, therefore, Wick's theorem does not apply. On the other hand, nothing prevents us from starting with $H_0(\mathbf{t}')$ and treating both the inter-cluster hopping and the Hubbard interaction, $H_0(\mathbf{V})$ and H_1 , as the perturbation.

There are two ways to do this: (i) we start from the free ($U = V = 0$) propagator \mathbf{G}'_0 of $H_0(\mathbf{t}')$ and, in a first step, sum the diagrams of all orders in V but for $U = 0$ (see the first line in Fig. 5a). One must merely sum a geometrical series, which can be done exactly. This step has been discussed already in Sec. 3.2. In a subsequent step, the resulting propagator \mathbf{G}_0 is dressed by taking into account the Hubbard interaction to all orders (see second line in Fig. 5a). This summation would yield the full Green's function \mathbf{G} but obviously cannot be done in practice. We therefore consider an alternative and reverse the order of the two steps: first, the free ($U = V = 0$) propagator \mathbf{G}'_0 is renormalized by the electron-electron interaction U to all orders but at $V = 0$ (first line in Fig. 5b). This yields the fully interacting cluster Green's function \mathbf{G}' . While, of course, \mathbf{G}' cannot be computed by the extremely complicated summation of individual U diagrams, it is nevertheless easily accessible via a direct (numerical) calculation if the cluster size is sufficiently small (see Sec. 2.3). In the second step, the $V = 0$ propagator \mathbf{G}' is renormalized due to inter-cluster potential scattering. Again, this is easily done by summing a geometrical series but only yields an approximation \mathbf{G}_{CPT} to the exact Green's function \mathbf{G} . In fact, as the second line in Fig. 5b demonstrates, this is just the cluster-perturbation theory, see Eq. (10).

Note that the CPT equation (10) has been introduced in an *ad hoc* way. In contrast, the diagram approach enables understanding of the CPT as an approximation that is given by summing a certain subclass of diagrams. Fig. 5c displays a low-order self-energy diagram that is neglected in this summation. This clearly shows that the CPT cannot be exact and suggests two different routes for improvement, namely (i) taking into account missing diagrams and (ii) using the freedom in the CPT construction to optimize the starting point. The first idea is related to the attempt to perform a systematic perturbative expansion around the disconnected-cluster limit and is notoriously complicated (as Wick's theorem does not apply) [29].

Another *ad hoc* way to derive the CPT follows the idea of the so-called Hubbard-I approximation [5]: The main idea is to employ the Dyson equation (44) of the reference system to compute the reference system's self-energy,

$$\Sigma'(i\omega_n) = \mathbf{G}'_0(i\omega_n)^{-1} - \mathbf{G}'(i\omega_n)^{-1}, \quad (46)$$

and to consider this as an approximation for the self-energy of the original system: $\Sigma(i\omega_n) \approx \Sigma'(i\omega_n)$. The motivation for this step is that the self-energy, as opposed to the Green's function, is a much more local object, as is well known at least for the weak-coupling regime from

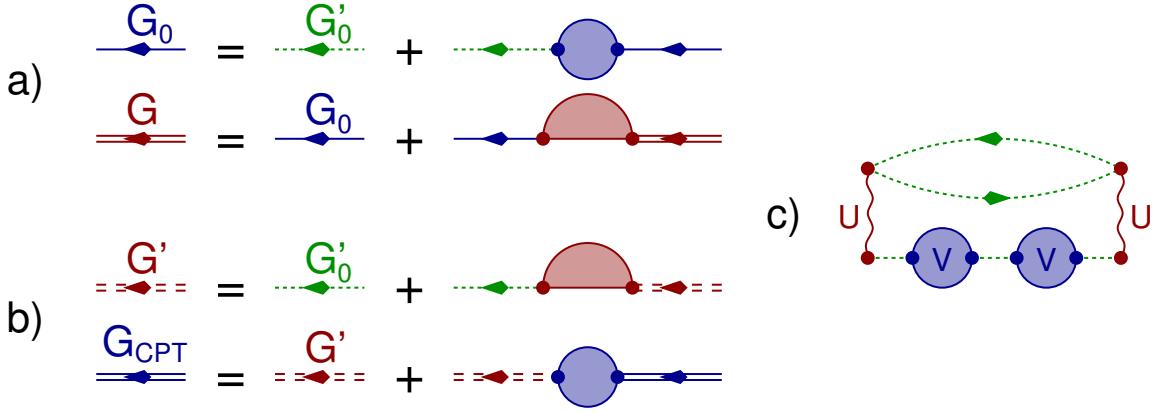


Fig. 5: Diagrammatic derivation of the CPT, see text for discussion and Ref. [28].

standard perturbation theory [30–32], and by the fact that the self-energy becomes purely local in the limit of lattices with infinite spatial dimensions [33,34]. Furthermore, the idea is reminiscent of dynamical mean-field theory (DMFT) [33,35,36] where the self-energy of an impurity model approximates the self-energy of the lattice model. Using Eq. (46) in Dyson’s equation for the original model, we find

$$\begin{aligned} \mathbf{G}(i\omega_n) &= \frac{1}{\mathbf{G}_0(i\omega_n)^{-1} - (\mathbf{G}'_0(i\omega_n)^{-1} - \mathbf{G}'(i\omega_n)^{-1})} \\ &= \frac{1}{\mathbf{G}'(i\omega_n)^{-1} - \mathbf{V}} , \end{aligned} \quad (47)$$

i.e., the CPT equation (10) is recovered. We note in passing that the Hubbard-I approach is obtained for $L_c = 1$ and with some *ad hoc* element of self-consistency [5].

This way to construct the CPT suggests to use, rather than the Ritz principle, a variational principle of the form

$$\frac{\delta \Omega[\Sigma]}{\delta \Sigma(i\omega_n)} = 0 , \quad (48)$$

where the trial self-energy is taken from the reference system and varied by varying the parameters of the reference system. Ideally, this *self-energy-functional approach* should also cure the different defects of the CPT, i.e. besides the arbitrariness of the CPT construction, the non-self-consistent nature of the approach, and its inability to describe spontaneous symmetry breaking, as well as different thermodynamical inconsistencies that show up in the computation of a thermodynamical potential from the Green’s function [25–27]. Furthermore, one may ask whether both the CPT and the DMFT can be understood in a single unifying theoretical framework.

4 Self-energy functional theory

4.1 Luttinger-Ward generating functional

The construction of a variational principle based on the self-energy is in fact possible with the help of the so-called Luttinger-Ward functional [24] (see [37] for a pedagogical introduction). This is a scalar functional Φ of the Green's function \mathbf{G} that was originally defined by all-order perturbation theory (a construction that uses the path integral can be found in Ref. [38]). More specifically, $\Phi[\mathbf{G}]$ is defined as the sum of all closed, connected, and fully dressed skeleton diagrams of any order k . Fig. 6 shows the lowest-order diagrams. Closed diagrams without links to external propagators are diagrams contributing to the partition function, see Eq. (40). The Luttinger-Ward series is given by dressed skeleton diagrams, i.e. diagrams without self-energy insertions where the free propagators are replaced by the fully interacting ones. One easily verifies that, due to dressing of the diagrams, some diagrams in the expansion of Z/Z_0 are counted twice or more. This is done on purpose. The most important property of the Luttinger-Ward functional constructed in this way is that its functional derivative just yields the skeleton-diagram expansion of the self-energy:

$$\frac{\delta\Phi[\mathbf{G}]}{\delta\mathbf{G}(i\omega_n)} = \frac{1}{\beta} \boldsymbol{\Sigma}[\mathbf{G}](i\omega_n) . \quad (49)$$

This can be verified, diagram by diagram: the functional derivative of a dressed skeleton just corresponds to the removal of a dressed propagator and results in a dressed skeleton diagram with two links for external propagators that contributes to the self-energy. When carefully taking into account the coefficients of the two different expansions Eq. (40) and Eq. (41), one easily derives Eq. (49). The equation is remarkable as it shows that the different components of the self-energy $\Sigma_{ij\sigma}(i\omega)$ can be obtained from the *scalar* functional. In fact, the existence of the Luttinger-Ward functional can also be proven by verifying a vanishing-curl condition as has been done by Baym and Kadanoff [39, 40].

The value Φ of the Luttinger-Ward functional has no direct physical meaning. Summing all closed diagrams (not only connected skeletons) yields, by construction, the partition function Z/Z_0 . Summing connected diagrams only yields $\ln Z$ as is known from the linked-cluster theorem [24, 25]. The sum of dressed connected skeletons, however, cannot provide the grand potential $\propto \ln Z$ because of the above-mentioned double counting.

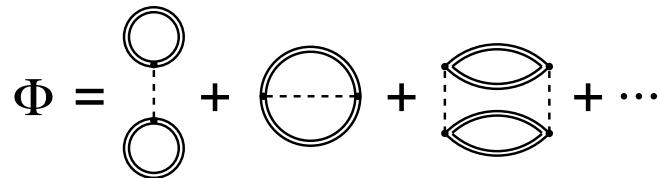


Fig. 6: Diagrammatic construction of the Luttinger-Ward functional $\Phi[\mathbf{G}]$. Double lines stand for fully interacting propagators, dashed lines for the Hubbard interaction.

4.2 Self-energy functional

We will make use of $\Phi[\mathbf{G}]$ by defining the following functional of the self-energy:

$$\Omega[\Sigma] = \text{Tr} \ln \frac{1}{\mathbf{G}_0^{-1} - \Sigma} + \Phi[\mathbf{G}[\Sigma]] - \text{Tr}(\Sigma \mathbf{G}[\Sigma]). \quad (50)$$

Here, the frequency dependencies are suppressed in the notation and

$$\text{Tr } \mathbf{A} \equiv \frac{1}{\beta} \sum_n \sum_{i\sigma} e^{i\omega_n 0^+} A_{ii\sigma}(i\omega_n) \quad (51)$$

is used where 0^+ is a positive infinitesimal. Furthermore, $\mathbf{G}[\Sigma]$ is the inverse of the functional $\Sigma[\mathbf{G}]$, i.e. $\mathbf{G}[\Sigma[\mathbf{G}]] = \mathbf{G}$. We assume that this inverse of the skeleton-diagram expansion of the self-energy exists at least *locally*. The second part of the self-energy functional,

$$F[\Sigma] \equiv \Phi[\mathbf{G}[\Sigma]] - \text{Tr}(\Sigma \mathbf{G}[\Sigma]), \quad (52)$$

is just the Legendre transform of the Luttinger-Ward functional. With $\Sigma[\mathbf{G}[\Sigma]] = \Sigma$ and Eq. (49) we immediately have

$$\frac{\delta F[\Sigma]}{\delta \Sigma} = -\frac{1}{\beta} \mathbf{G}[\Sigma]. \quad (53)$$

Therewith, we can also calculate the functional derivative of $\Omega[\Sigma]$:

$$\frac{\delta \Omega[\Sigma]}{\delta \Sigma} = \frac{1}{\beta} \left(\frac{1}{\mathbf{G}_0^{-1} - \Sigma} - \mathbf{G}[\Sigma] \right). \quad (54)$$

The equation

$$\mathbf{G}[\Sigma] = \frac{1}{\mathbf{G}_0^{-1} - \Sigma} \quad (55)$$

is a (highly non-linear) conditional equation for the self-energy of the system $H = H_0(\mathbf{t}) + H_1$. Inserting $\Sigma = \Sigma[\mathbf{G}]$ shows that it is (locally) equivalent to Eq. (45). It is satisfied by the exact self-energy of the system. Therefore, solving Eq. (55) is equivalent to a search for the stationary point of the self-energy functional:

$$\frac{\delta \Omega[\Sigma]}{\delta \Sigma} = 0. \quad (56)$$

This represents the dynamical variational principle we have been looking for. The exact self-energy of the system makes the self-energy functional $\Omega[\Sigma]$, Eq. (50), stationary.

The definition of the self-energy functional given with Eq. (50) is a formal one only. The argument of the \ln is not dimensionless, and furthermore, since $\mathbf{G}(i\omega_n) \propto 1/\omega_n \propto 1/(2n+1)$ for large n , the sum over the Matsubara frequencies $\sum_n \ln(2n+1)$ does not converge. This problem can be solved, however, by replacing $\Omega[\Sigma] \mapsto \Omega[\Sigma] - \text{Tr} \ln \mathbf{G}_{\text{reg}}$ with $G_{\text{reg},ij\sigma}^{-1}(i\omega_n) = \delta_{ij}(i\omega_n - \varepsilon_{\text{reg}})$ and taking the limit $\varepsilon_{\text{reg}} \rightarrow \infty$ after all calculations are done. As the constant $\text{Tr} \ln \mathbf{G}_{\text{reg}}$ does not depend on Σ , the variational principle is unaffected but now the Matsubara sum over both logarithms is well defined and convergent. One can show [24, 25, 3] that, if evaluated at the physical (exact) self-energy, the regularized $\Omega[\Sigma] - \text{Tr} \ln \mathbf{G}_{\text{reg}}$ is just the

grand potential of the system. This provides us with a physical interpretation of the self-energy functional. In the following this regularization is always implicit.

As a remark, we note that at $U = 0$ the self-energy functional reduces to the expression $\Omega_0 \equiv \text{Tr} \ln \mathbf{G}_0$ as becomes obvious from the diagrammatic definition of $\Phi[\mathbf{G}]$ and of Σ since there are simply no diagrams left at zeroth order in the interaction strength:

$$\Phi[\mathbf{G}] \equiv 0, \quad \Sigma(i\omega_n) = 0 \quad \text{for } U = 0. \quad (57)$$

If regularized properly, $\Omega_0 \mapsto \Omega_0 - \text{Tr} \ln \mathbf{G}_{\text{reg}}$, this exactly yields the grand potential of the non-interacting system.

4.3 Evaluation of the self-energy functional

The diagrammatic definition of the Luttinger-Ward functional (Fig. 6) uncovers another remarkable property: since any diagram contributing to Φ consists of vertices and dressed propagators only, the functional relation $\Phi[\dots]$ is completely determined by the interaction U but does not depend on t . Clearly, this universality also holds for its Legendre transform $F[\Sigma]$: two systems (at the same chemical potential μ and inverse temperature β) with the same interaction H_1 but different one-particle parameters t and t' are described by the same functional $F[\Sigma]$. In contrast, the first part of the self-energy functional,

$$\Omega[\Sigma] = \text{Tr} \ln \frac{1}{\mathbf{G}_0^{-1} - \Sigma} + F[\Sigma] \quad (58)$$

does depend on the hopping, namely via $\mathbf{G}_0^{-1}(i\omega_n) = i\omega_n + \mu - t$, but not on the interaction strength U .

The universality property of $F[\Sigma]$ is more important, however, as this functional is basically unknown. The central idea of self-energy-functional theory is to compare the self-energy functional of two systems, the original system with $H = H_0(t) + H_1$ and the reference system with $H' = H_0(t') + H_1$, i.e. Eq. (58) and

$$\Omega'[\Sigma] = \text{Tr} \ln \frac{1}{\mathbf{G}'_0^{-1} - \Sigma} + F[\Sigma]. \quad (59)$$

Due to its universality, one can eliminate the unknown functional $F[\Sigma]$ by combining both equations:

$$\Omega[\Sigma] = \Omega'[\Sigma] + \text{Tr} \ln \frac{1}{\mathbf{G}_0^{-1} - \Sigma} - \text{Tr} \ln \frac{1}{\mathbf{G}'_0^{-1} - \Sigma}. \quad (60)$$

This equation is still exact. Since the functional dependence of $\Omega'[\Sigma]$ is also unknown in the case of a simple reference system with decoupled clusters, it appears that this step amounts to a mere shift of the problem. The great advantage of Eq. (60) becomes manifest, however, when inserting the exact self-energy of the reference system $\Sigma_{t'}(i\omega_n)$ as a trial self-energy. (We use the notation $\Sigma_{t'}$ for the exact self-energy of the system with hopping parameters t' and interaction H_1 .) In this case, the first term on the right-hand side of Eq. (60) just reduces to the

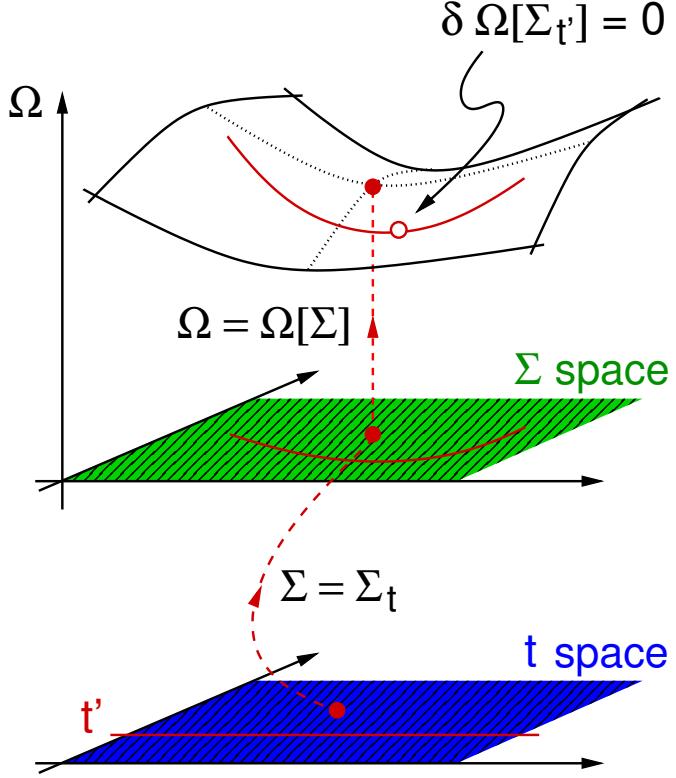


Fig. 7: The construction of consistent approximations within the self-energy-functional theory. The grand potential is considered as a functional of the self-energy that is stationary at the physical (exact) self-energy Σ_t (filled red circles). The functional dependence of $\Omega[\Sigma]$ is not accessible on the entire space of self-energies (Σ space). However, $\Omega[\Sigma]$ can be evaluated exactly on a restricted subspace of trial self-energies $\Sigma_{t'}$ parametrized by a subset of one-particle parameters t' (solid red lines). These t' define an exactly solvable reference system, i.e. a manifold of systems with the same interaction part but a modified one-particle part given by t' . Typically, the reference system consists of a set of decoupled clusters. A self-energy at which the grand potential is stationary on this sub-manifold represents the approximate self-energy of the original system and the grand potential at this self-energy represents the approximate grand potential (open circle).

grand potential of the reference system Ω' , which can be computed easily if, as we assume, the reference is amenable to an exact numerical solution. The same holds for the second and the third term. We find:

$$\Omega[\Sigma_{t'}] = \Omega' + \text{Tr} \ln \frac{1}{G_0^{-1} - \Sigma_{t'}} - \text{Tr} \ln \frac{1}{G_0'^{-1} - \Sigma_{t'}} . \quad (61)$$

This is a remarkable result. It shows that an *exact* evaluation of the self-energy functional of a non-trivial interacting system is possible, at least for trial self-energies that are taken from an exactly solvable reference system with the same interaction part (see Fig. 7).

4.4 The variational cluster approximation

We recall that the cluster perturbation theory approximates the self-energy of the original lattice-fermion model by the self-energy of a reference system of disconnected clusters. As one may choose the intra-cluster parameters of the reference system different from the corresponding parameters of the original system, there is a certain arbitrariness in the CPT construction. Usually, one simply assumes that e.g. the intra-cluster nearest-neighbor hopping of the reference system is the same as the physical hopping. There are, however, good reasons not to do so. Symmetry breaking Weiss fields are one example, as already mentioned above. Another one becomes obvious from Fig. 5c, where the CPT is seen to neglect the effect of the scattering at the inter-cluster potential on the self-energy. Therefore, an enhanced *intra*-cluster hopping could, at least partially, compensate for the missing feedback of the *inter*-cluster hopping on the approximate self-energy.

With the self-energy-functional framework at hand, we can now remove the arbitrariness of the CPT approach and determine the *optimal* self-energy from Eq. (61). This optimal self-energy is the exact self-energy of an optimized reference system that is specified by a set of one-particle (intra-cluster) parameters t' . Note that to derive Eq. (61) it was necessary to assume that the interaction part H_1 of the reference system cannot be optimized and must be the same as the interaction of the original system. Therefore, the role of the reference system is to generate a manifold of trial self-energies $\Sigma_{t'}$ that are parameterized by the one-particle parameters t' . As the self-energy functional Eq. (58) can be evaluated *exactly* on this manifold via Eq. (61), the optimal self-energy $\Sigma_{t'_{\text{opt}}}$ is given as the solution of the SFT Euler equation

$$\left. \frac{\partial \Omega[\Sigma_{t'}]}{\partial t'} \right|_{t'=t'_{\text{opt}}} = 0 . \quad (62)$$

For a cluster reference system, this constitutes the variational cluster approximation (VCA). The VCA represents an approximation as it provides the stationary point of the self-energy functional on a restricted manifold of trial self-energies only rather than on the entire *self-energy space* (see Fig. 7). The latter could be defined as the set of the self-energies of all models with the interaction part fixed at H_1 but with a completely arbitrary one-particle part (that also may connect the clusters). This space, of course, contains the exact self-energy Σ_t of our lattice model $H_0(t) + H_1$ while the optimized VCA self-energy $\Sigma_{t'_{\text{opt}}}$ is constrained to the manifold of cluster trial self-energies. Approximations generated in this way have a number of advantageous properties: first of all, although we have employed the language of perturbation theory, the VCA is non-perturbative. Formally, the diagram series has not been cut at any level, no subclass of diagrams is neglected, etc. The approximation rather results from a restricted domain of a self-energy functional. Second, the VCA is an internally consistent approximation in the sense that all observables derive from an explicit (though approximate) expression for a thermodynamical potential, namely from the self-energy functional evaluated at the optimal self-energy $\Omega[\Sigma_{t'_{\text{opt}}}]$. Third, the VCA can be improved in a systematic way by increasing the cluster size L_c , as one has to approach the exact solution for $L_c \rightarrow \infty$. Here, the reference system is basically identical

with the original system, and t is basically *within* the space of the variational parameters t' . One may interpret $1/L_c$ as the small parameter that controls the quality of the approximation. It is clear, however, that the numerical effort to solve Eq. (62) also increases with L_c . This increase is even exponential if an exact-diagonalization solver is used to compute the cluster self-energy, Green's function, and grand potential that enter Eq. (61). Unfortunately, one cannot *a priori* give a distance by which the optimal VCA self-energy differs from the exact self-energy. In practice, the quality of the approximation must therefore be controlled by comparing the results obtained for different cluster sizes L_c . For small clusters, also the cluster geometry and the imposed cluster boundary conditions matter, and must be checked.

Although the VCA derives from a general variational principle, it is not variational in the sense that the approximate VCA grand potential $\Omega[\Sigma_{t'_{\text{opt}}}]$ must always be larger than the exact grand potential Ω . As opposed to the Ritz principle and the state functional $E[|\Psi\rangle]$ or, at finite temperatures, the density-matrix functional $\Omega[\rho]$, the self-energy functional $\Omega[\Sigma]$ is not convex and hence there is no reason to assume that the VCA provides an upper bound to Ω .

Concluding, the VCA must be seen as a cluster mean-field approximation that focusses on one-particle correlations, the one-particle excitation spectrum (e.g. photoemission) and thermodynamics. It treats short-range one-particle correlations within the cluster in an explicit way while inter-cluster one-particle correlations are accounted for via Dyson's equation. The feedback of local and intra-cluster *two*-particle (and even higher) correlations on the one-particle self-energy is explicitly and non-perturbatively taken into account while the feedback of non-local two-particle, e.g. magnetic, correlations on the one-particle spectrum is neglected altogether. This is typical for cluster mean-field theories [41] and should be kept in mind when studying, e.g., systems close to a second-order phase transition, where non-local correlations play an important role.

5 Implementation of the variational cluster approximation

***Q*-matrices**

The bottleneck of a practical VCA calculation consists in the computation of the Green's function of the reference system. Using an exact-diagonalization technique, the Green's function for an individual cluster can be obtained in its Lehmann representation, see Eq. (17). Let $\alpha = (i, \sigma)$ be an index referring to the elements of the localized orbitals forming an orthonormal basis of the one-particle Hilbert space. Therewith, the elements of the cluster Green's function can be written in the form

$$G'_{\alpha\beta}(\omega) = \sum_m Q'_{\alpha m} \frac{1}{\omega - \omega'_m} Q'^{\dagger}_{m\beta}. \quad (63)$$

Here, $m = (r, s)$ refers to a single-particle excitation between two energy eigenstates $|s\rangle$ and $|r\rangle$ of the (grand-canonical) Hamiltonian of the reference system $H' - \mu N$, and $\omega'_m = E'_r - E'_s$ is the excitation energy. $Q'_{\alpha m}$ are the elements of the so-called *Q'*-matrix [42], which is a rectangular matrix with a small number of rows but a large number of columns (dimension of

the one-particle Hilbert space \times number of many-body excitations):

$$Q'_{\alpha m} = \langle r | c_\alpha | s \rangle \sqrt{\frac{\exp(-\beta E'_r) + \exp(-\beta E'_s)}{Z'}}, \quad Z' = \sum_r e^{-\beta E'_r}, \quad (64)$$

as is readily read off from the Lehmann representation Eq. (17). One also verifies that $\mathbf{Q}'\mathbf{Q}'^\dagger = \mathbf{1} \neq \mathbf{Q}'^\dagger\mathbf{Q}'$. Using the Q' -matrix, we can write the reference system's Green's function in a compact form as

$$\mathbf{G}'(\omega) = \mathbf{Q}' \frac{1}{\omega - \Lambda'} \mathbf{Q}'^\dagger, \quad (65)$$

where Λ' is the diagonal matrix with elements $\Lambda'_{mn} = \omega'_m \delta_{mn}$. With $\mathbf{V} = \mathbf{t} - \mathbf{t}'$, the Green's function of the original system is obtained as:

$$\begin{aligned} \mathbf{G}(\omega) &= \frac{1}{\mathbf{G}'(\omega)^{-1} - \mathbf{V}} = \mathbf{G}'(\omega) + \mathbf{G}'(\omega) \mathbf{V} \mathbf{G}'(\omega) + \dots \\ &= \mathbf{Q}' \left(\frac{1}{\omega - \Lambda'} + \frac{1}{\omega - \Lambda'} \mathbf{Q}'^\dagger \mathbf{V} \mathbf{Q}' \frac{1}{\omega - \Lambda'} + \dots \right) \mathbf{Q}'^\dagger = \mathbf{Q}' \frac{1}{\omega - \mathbf{M}} \mathbf{Q}'^\dagger, \end{aligned} \quad (66)$$

where $\mathbf{M} = \Lambda' + \mathbf{Q}'^\dagger \mathbf{V} \mathbf{Q}'$ is a (large) square Hermitian matrix that can be diagonalized by a unitary transformation, $\mathbf{M} = \mathbf{S} \Lambda \mathbf{S}^\dagger$. Here, $\Lambda_{mn} = \omega_m \delta_{mn}$ with the poles ω_m of $\mathbf{G}(\omega)$. We find

$$\mathbf{G}(\omega) = \mathbf{Q} \frac{1}{\omega - \Lambda} \mathbf{Q}^\dagger \quad (67)$$

with $\mathbf{Q} = \mathbf{Q}' \mathbf{S}$. The representations Eq. (65) and Eq. (67) are particularly useful to evaluate the self-energy functional Eq. (61) in practice. The trace contains a Matsubara-frequency summation which can be carried out analytically [43] such that one is left with a simple algebraic expression [42],

$$\text{Tr} \ln \frac{1}{\mathbf{G}_0^{-1} - \Sigma_{\mathbf{t}'}} - \text{Tr} \ln \frac{1}{\mathbf{G}'_0^{-1} - \Sigma_{\mathbf{t}'}} = - \sum_m \frac{1}{\beta} \ln(1 + e^{-\beta \omega_m}) + \sum_m \frac{1}{\beta} \ln(1 + e^{-\beta \omega'_m}), \quad (68)$$

which involves the poles of $\mathbf{G}(\omega)$ and $\mathbf{G}'(\omega)$ only. Finally, the grand potential of the reference system in Eq. (61) is easily computed as $\Omega' = -(1/\beta) \ln \sum_r e^{-\beta E'_r}$.

Recipe for practical calculations

A typical VCA calculation is carried out as follows:

- Construct a reference system by tiling the original lattice into identical clusters.
- Choose a set of one-particle parameters \mathbf{t}' of the reference system and compute $\mathbf{V} = \mathbf{t} - \mathbf{t}'$.
- Solve the problem for the reference system (U is fixed), i.e. compute the Green's function \mathbf{G}' and find the poles ω'_m and the Q' -matrix.
- Get the poles ω_m of the approximate Green's function of the original system by diagonalization of the matrix $\mathbf{M} = \Lambda' + \mathbf{Q}'^\dagger \mathbf{V} \mathbf{Q}'$.

- Calculate the value of the SFT grand potential via Eq. (61) and Eq. (68) and by calculating the grand potential of the reference system Ω' from the eigenvalues of H' .
- Iterate this scheme for different t' , such that one can solve

$$\left. \frac{\partial \Omega[\Sigma_{t'}]}{\partial t'} \right|_{t'=t'_{\text{opt}}} \stackrel{!}{=} 0 \quad (69)$$

for t'_{opt} .

- Evaluate observables, such as $\Omega[\Sigma_{t'_{\text{opt}}}]$, $G(\omega)$ and static expectation values derived from the SFT grand potential by differentiation, at the stationary point t'_{opt} .
- Redo the calculations for different parameters of the *original* system, e.g. a different U , filling or β to scan the interesting parameter space.

Tips and tricks

For a given topology of the reference system, i.e. for a given cluster geometry, one may in principle consider all one-particle parameters t' as variational parameters. However, besides an exponentially increasing Hilbert-space dimension, a larger cluster also implies an increasing numerical complexity for the search of the stationary point since $\Omega[\Sigma_{t'}]$ is a function of a multi-component variable t' . It is therefore advisable to restrict the search to a small number of physically important parameters. In most cases, a few variational parameters suggest themselves.

An overall shift $\Delta\varepsilon'$ of the on-site energies in the cluster (like the chemical potential), $t'_{ii} \mapsto t'_{ii} + \Delta\varepsilon'$, should be among the variational parameters to ensure thermodynamical consistency with respect to the total particle number as has been pointed out in Ref. [44]. This ensures that both ways to compute the total particle number, $\langle N \rangle = -\partial\Omega/\partial\mu$ and $\langle N \rangle = \sum_{i\sigma} \int dz A_{i\sigma}(z)/(e^{\beta z} + 1)$, must yield the same result. Analogously, in case of a (ferro- or antiferro-) magnetic system, one should include a (homogeneous or staggered) Weiss field B' in the set of variational parameters. For a paramagnetic system and for a system with manifest particle-hole symmetry, however, symmetry considerations *a priori* fix those variational parameters to $B' = 0$ and $\Delta\varepsilon' = 0$, i.e. $t'_{ii} = t_{ii}$. This can also be verified by a practical VCA calculation.

For the setup of self-energy-functional theory it is inevitable that the original and the reference system have the same interaction H_1 . Conversely, the one-particle part of the reference system can be designed at will. One very interesting option in this context is to add additional fictitious sites to the cluster. These *bath sites* have to be non-interacting ($U = 0$), contrary to the *correlated sites* ($U > 0$), which correspond to the physical sites (with the same U) of the original system. Adding the bath sites does not change the interaction part H_1 of the Hamiltonian and therefore leaves the Luttinger-Ward functional as well as its Legendre transform $F[\Sigma]$ unaffected. Bath sites can be coupled via one-particle hopping terms to the correlated sites in the

cluster. This construction has the appealing advantage of increasing the space of variational parameters and thereby improving the quality of the approximation *locally*. Adding bath sites and optimizing the additional associated parameters will improve the description of local temporal correlations while increasing the cluster improves the theory with respect to non-local spatial correlations.

If one decides to consider a reference system with bath sites, it is advisable to formally include the same bath sites also in the original system. Here, of course, they are completely decoupled from the correlated sites (the respective parts of the hopping matrix \mathbf{t} have to be set to zero) such that all physical quantities remain unchanged. The advantage of this trick is that \mathbf{t} and \mathbf{t}' have the same matrix dimension, and that the Hamiltonians H and H' operate on the same Hilbert space. The *inter-cluster hopping* $\mathbf{V} = \mathbf{t} - \mathbf{t}'$ includes the hopping terms between correlated and bath sites in the reference system only.

Rather than employing the above-mentioned Q -matrix technique, one may also perform the traces in Eq. (61), i.e. the trace of the spatial and orbital degrees of freedom and the implicit Matsubara-frequency summation, numerically. This is recommended if the dimension of M , given by the number of poles of G' with non-vanishing spectral weight, becomes too large.

If a full diagonalization of the cluster problem is not feasible and Krylov-space methods shall be applied, one has to make sure that the different elements $G'_{ij\sigma}(\omega)$ have the same set of poles: ω'_m should be independent of i, j . This can be achieved by the band Lanczos method [45]. The dimension of the matrix M is given by the number of iteration steps in the Lanczos procedure. Typically, about 100 steps are sufficient for reasonably well converged results. This should be checked regularly.

The SFT grand potential may exhibit more than a single stationary point. A minimal grand potential among the grand potentials at the different stationary points distinguishes the thermodynamically stable phase in most situations [46]. Often, the occurrence of several stationary points is welcome from a physical point of view. For example, scanning a physical parameter, e.g. U , a second-order magnetic phase transition is characterized by a bifurcation of a non-magnetic solution into a non-magnetic and a magnetic one (or even more magnetic ones).

There are different numerical strategies to determine a stationary point of the self-energy functional, see Ref. [47] for examples. If there is only a single variational parameter to be optimized, iterative bracketing of maxima and minima can be employed efficiently. In the case of more than one variational parameter, the SFT grand potential usually exhibits a saddle point rather than a minimum or maximum. A strategy that has been found to be useful for two or three parameters is to assume (and verify) a certain characteristic of the saddle point and to apply iterated one-dimensional optimizations. The downhill simplex method can be used for higher-dimensional parameter spaces to find the local minima of $|\partial\Omega[\Sigma_{\mathbf{t}}]/\partial\mathbf{t}'|^2$. If there is more than one, only those must be retained for which $\Omega[\Sigma(\mathbf{t}')]$ has a vanishing gradient.

6 Selected results

The VCA is not restricted to the single-band Hubbard model but has also been applied to a variety of multi-orbital systems. The necessary generalization to the multi-orbital case is straightforward. In this way, the VCA has contributed to the study of the correlated electronic structure of real materials such as NiO [48], CoO and MnO [49], CrO₂ [50], LaCoO₃ [51], TiOCl [52], TiN [53], and NiMnSb [54]. Here, however, we will focus on the single-band model and discuss a few and very simple examples to illustrate the theory.

6.1 One-dimensional Hubbard model

In the first example [55], we will consider the one-dimensional Hubbard model at zero temperature and half-filling with hopping $t = 1$ between nearest neighbors, see Fig. 8a. A tiling of the one-dimensional lattice into “clusters” is particularly simple: Each cluster is a finite chain of L_c sites. We treat the intra-cluster nearest-neighbor hopping t' as the only variational parameter. This is the most obvious choice. Nevertheless, one may numerically check that the optimal on-site hopping $t'_{ii,\text{opt}} = t_{ii} = 0$. The same holds for the hopping between second nearest neighbors: $t'_{2-\text{nd},\text{opt}} = 0$. Again this is predicted by particle-hole symmetry. On the other hand, if a third nearest-neighbor hopping is introduced as a variational parameter, it acquires a small finite value at the stationary point. Interestingly, one also finds $t'_{\text{pbc},\text{opt}} = 0$ [2], where t'_{pbc} is a hopping parameter that links the two edge sites of the cluster with each other. $t'_{\text{pbc}} = t'$ would be a realization of periodic boundary conditions, but the calculation shows that open boundaries, $t'_{\text{pbc},\text{opt}} = 0$, are preferred. Furthermore, one may also relax the constraint that the hopping t' be the same for all pairs of nearest neighbors. In this case one finds the strongest deviations close to the edges of the reference systems [55].

Fig. 8b shows the dependence of the SFT grand potential $\Omega[\Sigma_{t'}]$ on the *single* variational parameter t' . Actually, $(\Omega + \mu\langle N \rangle)/L$ is plotted. At zero temperature and at the stationary point, this is the (approximate) ground-state energy of the Hubbard model per site. There is a stationary point, a minimum in this case, with the optimal value for the intra-cluster hopping t'_{opt} being close to but different from the physical value $t = 1$ for strong Coulomb interaction U . Note that the CPT is given by $t' = t$ and that there is a gain in binding energy due to the optimization of t' , namely $\Omega(t'_{\text{opt}}) < \Omega(t)$ which implies that the VCA improves on the CPT result.

It is physically reasonable that in the case of a stronger interaction and thus more localized electrons, switching off the inter-cluster hopping is less significant and must therefore be offset to a lesser degree by an increase of the intra-cluster hopping. A considerably large deviation from the physical hopping, $t'_{\text{opt}} > t$, is only found for the weakly interacting system. However, even a “strong” approximation of the self-energy (measured as a strong deviation of t'_{opt} from t) becomes irrelevant in the weak-coupling limit as the self-energy becomes small. With decreasing U , the self-energy functional becomes flatter and flatter until at $U = 0$ the t' dependence is completely irrelevant. Note that not only the non-interacting limit but also the atomic limit ($t = 0$) is exactly reproduced by the VCA. In the latter case, the reference system becomes

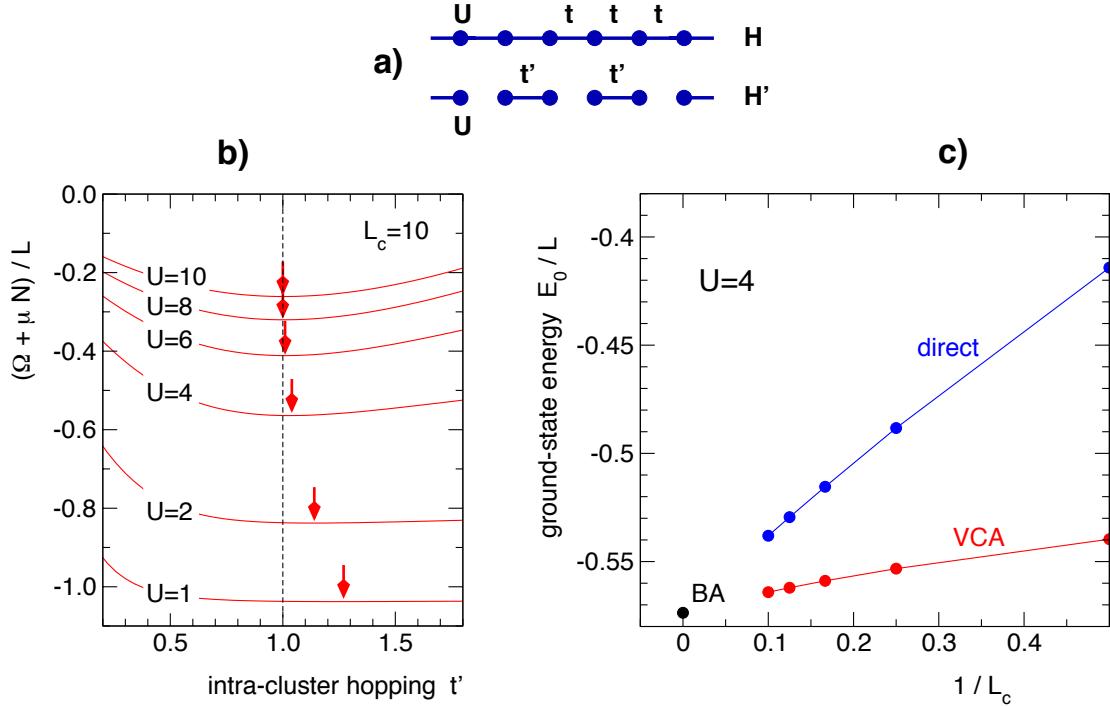


Fig. 8: a) Original system: one-dimensional Hubbard model. Reference system: decoupled clusters (for $L_c = 2$). b) SFT grand potential per site and shifted by μN as a function of the intra-cluster nearest-neighbor hopping t' . VCA calculation (with $L_c = 10$) for the one-dimensional Hubbard model at zero temperature, half-filling and different U as indicated. The nearest-neighbor hopping $t = 1$ sets the energy scale. Arrows mark the stationary points. c) VCA ground-state energy per site as a function of $1/L_c$ for $U = 4$ at the respective stationary points compared with the corresponding results for an isolated cluster and the exact results known from the Bethe ansatz (BA) [56]. (adapted from Ref. [55])

identical to the original system at $t' = 0$.

Fig. 8c shows the VCA ground-state energy (per site) at a fixed interaction strength $U = 4$ as a function of the inverse cluster size $1/L_c$. By extrapolation to $1/L_c = 0$ one recovers the exact Bethe-Ansatz result (BA) [56]. Furthermore, the VCA is seen to improve the ground-state energy as compared to calculations done for an isolated Hubbard chain with open boundaries. Convergence to the BA result is clearly faster within the VCA. Note that, as opposed to the VCA, the direct cluster approach is not exact for $U = 0$.

6.2 Antiferromagnetism

With the second example [16], we return to our original motivation, see Fig. 2: one of the main drawbacks of the CPT consists in its inability to describe spontaneous symmetry breaking. Consider SU(2) transformations in spin space and antiferromagnetic order, for example. As the exact solution of a finite Hubbard cluster is necessarily spin-symmetric, i.e., invariant under SU(2) transformations, and as the CPT equation proliferates this symmetry, the antiferromagnetic order parameter, the staggered magnetization m , must always be zero if there is

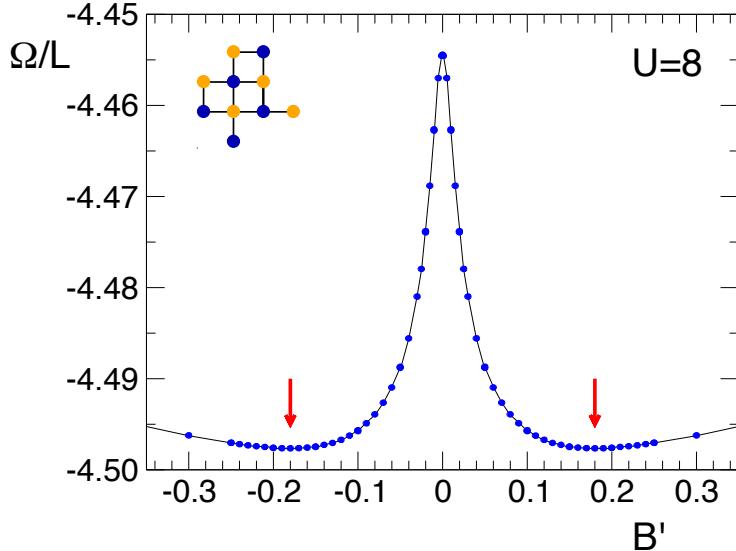


Fig. 9: SFT grand potential per site as a function of the strength of a fictitious staggered magnetic field B' . Calculation for the two-dimensional half-filled Hubbard model on the square lattice at zero temperature and $U = 8$, $t = 1$. The reference system consists of disconnected clusters with $L_c = 10$ sites each, see the inset for the cluster geometry. Arrows mark the two equivalent stationary points. (adapted from Ref. [16])

no physically applied staggered magnetic field that would explicitly break the symmetry, i.e. if $B = 0$. The VCA, on the other hand, in principle allows for a spontaneous SU(2) symmetry breaking. Namely, treating an intra-cluster fictitious staggered magnetic field of strength B' as a variational parameter offers the possibility for a symmetry-broken stationary point with $B'_{\text{opt}} \neq 0$. The reference-system Hamiltonian is given by

$$H' = H'|_{B'=0} - B' \sum_{i\sigma} z_i(n_{i\uparrow} - n_{i\downarrow}), \quad (70)$$

where $z_i = +1$ for sites on sublattice A and $z_i = -1$ for sublattice B (yellow and blue sites in the inset of Fig. 9)

The main part of Fig. 9 displays VCA results for the half-filled two-dimensional Hubbard model on the square lattice at zero temperature. Decoupled clusters with $L_c = 10$ sites are considered as a reference system, and the staggered field B' is the only variational parameter considered. There is a stationary point at $B' = 0$ that corresponds to the paramagnetic phase and to the CPT. In addition, however, there are two equivalent stationary points at *finite* B' corresponding to a phase with antiferromagnetic order. Comparing the ground-state energies of both the antiferromagnetic and the paramagnetic phase shows that the former is thermodynamically stable.

One should be aware, however, that the VCA, like any cluster mean-field approach, tends to overestimate the tendency towards magnetic order. Furthermore, a finite-temperature calculation is expected to produce a finite order parameter for the two-dimensional but also for the one-dimensional case, which would be at variance with the Mermin-Wagner theorem [57]. What is missing physically in the VCA is the effect of long wavelength spin excitations. The VCA is therefore restricted to cases where the physical properties are dominated by short-range correlations on a scale accessible by an exactly solvable finite cluster.

6.3 Mott metal-insulator transition

The third example [58] addresses a first-order (discontinuous) phase transition. This type of phase transition can be studied conveniently within the SFT framework as there is an explicit expression for a thermodynamical potential available from the very beginning. We again consider the two-dimensional Hubbard model on the square lattice at half-filling and zero temperature but disregard the antiferromagnetic phase and enforce a paramagnetic state by choosing $B' = 0$. The paramagnetic system is expected to undergo a transition from a correlated metal at weak U to a Mott insulator at strong U . This Mott transition is first of all interesting from a fundamental point of view as it is driven by electronic correlations opposed to other types of metal-insulator transitions [59]. The Mott insulator is characterized by a gap of the order of U in the single-particle excitation spectrum which is only weakly dependent on temperature. One therefore expects that a possible metal-insulator quantum phase transition at zero temperature is of relevance for the high-temperature state of the system as well, where it should give rise to a smooth crossover between a more metallic and a more insulating state. The crossover takes place at temperatures that may be well above the Néel temperature where the system is paramagnetic. This is the motivation to ignore the magnetic phase for the zero-temperature calculation.

Fig. 10 (right) shows the building block of the reference system. This is a cluster with $L_c = 4$ correlated sites (filled blue dots) but with four bath sites (open red dots) in addition, i.e. there are $n_s = 2$ local degrees of freedom (one additional bath site per correlated site). As mentioned above, including bath sites in a VCA calculation improves the description of local correlations. This is an important ingredient in understanding the Mott transition: a paradigmatic picture of the Mott transition could be worked out in the limit of infinite spatial dimensions with the help of the dynamical mean-field theory [33, 35]. In this limit the self-energy becomes a completely local quantity, $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij}\Sigma_{i\sigma}(i\omega_n)$ [34], and therefore the local temporal degrees of freedom (n_s) dominate the spatial ones (L_c). For two dimensions, the considered reference system with $L_c = 4$ and $n_s = 2$ is expected to represent a good compromise between the importance of local and non-local correlations and to result in a reasonable approximation.

VCA calculations with the full set of variational parameters indicated in Fig. 10 (right) have shown [58] that the hopping between the correlated and the bath sites V is the most important parameter to be optimized while $t'_{\text{opt}} \approx t$ and $t''_{\text{opt}} \approx 0$ can safely be ignored, i.e. set to the *a priori* plausible values $t'_{\text{opt}} = t$ and $t''_{\text{opt}} = 0$. The on-site energies of the correlated and the bath sites are fixed by particle-hole symmetry anyway. This drastically simplifies the study, as the SFT grand potential $\Omega[\Sigma_V]$ can be regarded as a function of a single variational parameter V .

Fig. 10 (left) displays the SFT grand potential, shifted by μN , per correlated site as a function of V for different U . For weak interactions, $U < U_{c2} \approx 6.35$, there is a stationary point (a minimum) at a comparatively large V_{opt} that describes a metallic phase (blue dots). The metallic character of the phase can be inferred from the finite value of the imaginary part of the local Green's function $\text{Im } G_{i\sigma}(i\omega)$ for $\omega \rightarrow 0$ (see Ref. [58]). Above the critical value U_{c2} , no metallic solution can be found. There is, however, an insulating phase for strong U (red

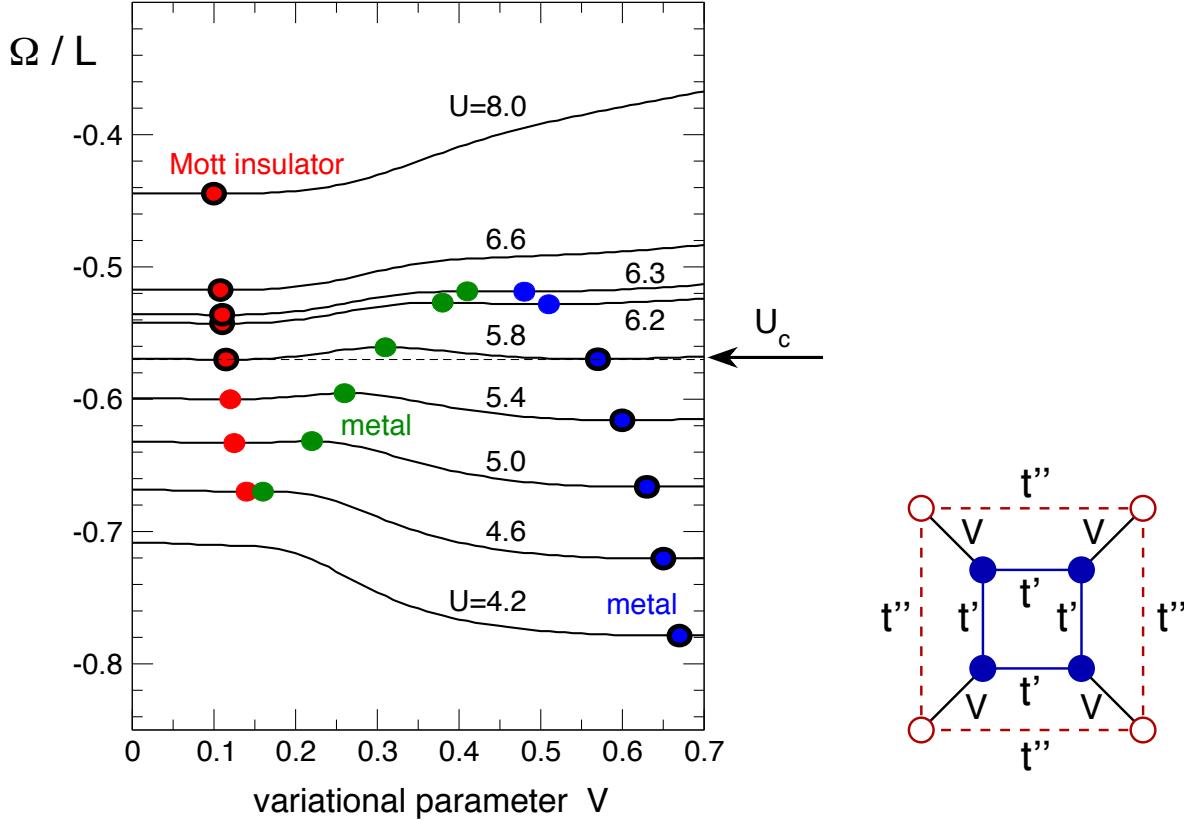


Fig. 10: Left: SFT grand potential per site as a function of the variational parameter V . VCA calculation for the two-dimensional Hubbard model on the square lattice at zero temperature, half-filling, and different U as indicated. The nearest-neighbor hopping $t = 1$ sets the energy scale. Symbols: stationary points. Red: Mott insulator, green and blue: metal. Fat symbols: thermodynamically stable phase. The first-order Mott transition is marked by an arrow. Right: Sketch of the building block of the reference system. Blue filled dots: correlated sites with $U > 0$. Red open dots: bath sites with $U = 0$. Calculations for $t' = t$, $t'' = 0$, arbitrary V . (adapted from Ref. [58]).

dots). The respective stationary point (a minimum) of the SFT grand potential is found at a comparatively low value of V_{opt} and can be traced with decreasing U down to another critical value $U_{c1} \approx 4.6$. For $U < U_{c1}$, there is no insulating phase.

It is interesting to observe that in the regime $U_{c1} < U < U_{c2}$ the metallic and the insulating phase are coexisting, i.e. that there are two stationary points of the grand potential. There is actually a third stationary point in the coexistence region, indicated by the green dots, where the SFT grand potential is at a maximum. Note that any stationary point, minimum, maximum or saddle point (in higher-dimensional parameter spaces), must be considered as an admissible solution within the SFT. However, the grand potential of the third phase is always higher than the grand potentials of the other phases. It therefore describes a physically irrelevant metastable phase but mathematically explains why the other two phases cease to exist above (below) a certain critical interaction U_{c2} (U_{c1}).

For a given U , the phase with the lowest grand potential is thermodynamically stable (see fat symbols). This means that the system is a correlated metal for $U < U_c$ and a Mott insulator for $U > U_c$ where the critical value for the Mott transition $U_c \approx 5.8$ is given by the interaction strength for which the metal and insulator have the same grand potential (the same ground-state energy at zero temperature), see the arrow in Fig. 10. Therefore, at U_c the optimal hopping parameter V_{opt} jumps between the large metallic and the small insulating value. Consequently, the ground state and thus the self-energy of the reference system changes abruptly at U_c . This leads to a discontinuous change of the SFT Green's function as well as of all observables that are computed as derivatives of the (optimized) SFT grand potential. The phase transition is of *first order* or *discontinuous*. This is interesting since the Mott transition in the Hubbard model on an infinite-dimensional, e.g. hyper-cubic, lattice is known [35] to be of second order or continuous. The VCA calculation discussed here actually corrects a mean-field artifact that is due to the neglect of non-local short-range antiferromagnetic correlations (see Ref. [58] for an extended discussion). For a discussion of more recent developments see Ref. [60], for example.

7 Relation to other methods and conclusions

Concluding, it is an appealing idea to divide a correlated lattice-fermion problem into small isolated clusters for which the problem can be solved easily and in a second step to employ the decoupled-cluster solution to construct the solution for the original lattice model. This construction must be approximate and has the spirit of a cluster mean-field theory, where the intra-cluster correlations are treated in a much better and more explicit way than the inter-cluster correlations. We have learned that this construction cannot be based on a technique that uses many-body wave-functions, rather, one has to employ Green's functions, namely the single-particle Green's function or, equivalently, the self-energy, as is done with the cluster-perturbation theory.

The self-energy-functional theory conceptually improves the CPT in several respects: first, it removes the arbitrariness that is inherent to the CPT regarding the choice of the cluster parameters. Second, it introduces an element of self-consistency or variational character by which it becomes possible to study phases that possess a symmetry different from the symmetry of the isolated cluster, i.e. one can address spontaneously symmetry-broken phases (collective magnetism, superconductivity etc.). Third, the SFT provides us with an explicit, though approximate, expression for a thermodynamical potential from which all observables have to be derived. This ensures that the approach is consistent in itself and obeys general thermodynamical relations, an important point that is missing in the plain CPT as well.

The self-energy-functional theory should actually be seen as a theoretical frame that allows the construction of different approximations. Each approximation is characterized by the choice of a corresponding reference system. Typically, this consists of decoupled clusters with L_c correlated sites each and an additional $n_s - 1$ uncorrelated bath sites. Large clusters are necessary in order to describe short-range correlations as well as possible, and a large number of local degrees of freedom n_s is recommended to improve the description of local temporal correlations.

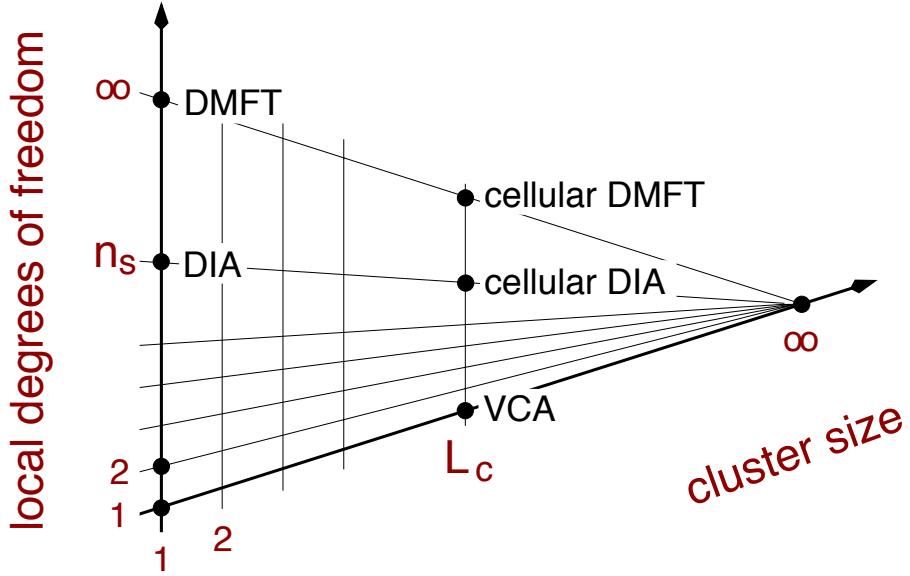


Fig. 11: Schematic picture for the relation between different approximations that can be constructed within the self-energy-functional theory. See text for discussion.

There is the ubiquitous tradeoff between the quality of the approximation on the one hand and the numerical effort on the other as the problem must be exactly solved for the isolated cluster. Using an exact-diagonalization solver the effort increases roughly exponentially with L_c and n_s . The space of possible approximations spanned by L_c and n_s is sketched in Fig. 11.

The most simple approximation is given by $L_c = 1$ and $n_s = 1$. Here, one approximates the self-energy of the lattice model by the self-energy of the atomic problem. This is in the spirit of the Hubbard-I approximation [5]. For $L_c > 1$ we find the variational cluster approximation that we have discussed at length here. Obviously, one would recover the exact solution of the lattice model in the limit $L_c \rightarrow \infty$.

Choosing a “cluster” with a single correlated site only, $L_c = 1$, but introducing a number of bath sites $n_s - 1 \geq 1$ specifies another approximation which is called dynamical impurity approximation (DIA). This is a true mean-field approximation as all non-local two-particle spatial correlations are neglected in this case.

Obviously, an ideal embedding of a single site into the lattice or an ideal mean-field theory is realized with an infinite number of bath sites $n_s \rightarrow \infty$. In this case, all local, temporal correlations are treated exactly – as opposed to a *static* mean-field theory like the Hartree-Fock approach. This optimal mean-field theory turns out to be identical with the well-known dynamical mean-field theory (DMFT) [33, 35]. In fact, one immediately recognizes that the reference system in this case is just the single-impurity Anderson model (if one starts out from a single-band Hubbard model, for example). Within DMFT the parameters of this impurity model are fixed by imposing the so-called DMFT self-consistency condition, stating that the impurity Green’s function of the impurity model should be equal to the local Green’s function of the lattice model. This is realized by setting up a self-consistency scheme that requires an iterated solution of the impurity model. Within the SFT, on the other hand, the parameters of

the reference impurity system are fixed by the SFT Euler equation which is the condition for the stationarity of the SFT grand potential. In fact both the DMFT self-consistency equation and the SFT Euler equation are basically equivalent in this case. From Eq. (58), we have

$$\frac{\partial}{\partial \mathbf{t}'} \Omega[\Sigma_{\mathbf{t}'}] = \frac{1}{\beta} \sum_{n,ij\sigma} \left(\frac{1}{\mathbf{G}_0^{-1}(i\omega_n) - \Sigma(i\omega_n)} - \mathbf{G}'(i\omega_n) \right)_{ij\sigma} \frac{\partial \Sigma_{ji\sigma}(i\omega_n)}{\partial \mathbf{t}'} = 0, \quad (71)$$

where \mathbf{G}' denotes the Green's function of the reference system. As the self-energy of a single-impurity Anderson model is local, $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij} \Sigma_{i\sigma}(i\omega_n)$ and non-zero on the correlated impurity site only. This SFT Euler equation is satisfied if the *impurity* Green's function $G'_{\text{imp}}(i\omega_n)$ is equal to the (approximate) *local* Green's function of the lattice model, i.e. if the local elements of the bracket vanish. This is the DMFT self-consistency equation.

What happens if $n_s < \infty$? For a small number of bath sites, an exact solution of the reference system by means of exact-diagonalization techniques becomes feasible in practice. The resulting approximation, the DIA, differs from DMFT as the bracket in Eq. (71) will never vanish in this case. This is easily seen by noting that the Green's function of an $n_s < \infty$ impurity model has a finite number of poles on the real frequency axis while the approximate lattice Green's function inherits its analytical structure from the non-interacting lattice Green's function \mathbf{G}_0 , which, for an infinite lattice, may exhibit isolated poles but must have branch cuts as well. The DIA is different and actually inferior compared with the DMFT but does not need an advanced solver (note that solving the single-impurity Anderson model with $n_s = \infty$ is still a demanding many-body problem). It has turned out, however, that with a few bath sites only, one often has a rather reliable approach to study the thermodynamics of a lattice model (see Ref. [43], for an example). This is also known from the exact-diagonalization approach to DMFT [61]. DMFT-ED considers the impurity model with a small n_s , as is done in the DIA, but employs another, actually more *ad hoc* condition to fix the parameters of the impurity model, namely one minimizes a suitably defined distance between the two local elements of the Green's function in the DMFT self-consistency condition to satisfy this at least approximately. The approach is able to yield similar results as the DIA in practice and is more easily implemented numerically but lacks internal consistency.

The third example discussed in the previous section has shown that one can favorably make use of approximations where a small number of bath degrees of freedom $n_s > 1$ are combined with a cluster approach $L_c > 1$. This approach is a VCA with additional bath sites and may be termed “cellular DIA” (see Fig. 11) since it is related to the cellular DMFT [62, 63] in the same way as the DIA is related to the DMFT. Note that with increasing cluster size $L_c \rightarrow \infty$, all approaches, the VCA, the cellular DIA as well as the cellular DMFT must recover the exact solution of the lattice problem in principle.

Another prominent and widely used cluster mean-field theory is the dynamical cluster approximation (DCA) [64]. As compared to the cellular DMFT, this is a cluster extension of dynamical mean-field theory that avoids one of the main drawbacks of various cluster approaches, namely the artificial breaking of the translational symmetries. Already the CPT yields an approximate Green's function that merely reflects the translational symmetries of the “superlattice,” which

periodically repeats the basic cluster, rather than the symmetries of the underlying physical lattice. In contrast, the DCA provides a Green's function and a self-energy with the correct symmetries but, on the other hand, must tolerate that the self-energy is discontinuous as a function of \mathbf{k} in reciprocal space. Here, we briefly mention that it is possible to re-derive the DCA within in the framework of the SFT as well. This is carried out in detail in Ref. [14]. It is based on the idea that, with a proper modification of the hopping parameters of the *original* lattice model, $t \mapsto \tilde{t}$, which becomes irrelevant for $L_c \rightarrow \infty$, the DCA becomes equivalent to the cellular DMFT. In a similar way [9] another cluster-mean-field variant, the periodized cellular DMFT [65] can be re-derived within the SFT.

To summarize, the self-energy-functional approach not only recovers a number of well-known mean-field and cluster mean-field concepts and provides a unified theoretical framework to classify the different approaches but has also initiated the construction of new non-perturbative and consistent approximations, the most prominent example of which is the variational cluster approximation. The challenges for future developments are manifold; let us mention only two directions here. The first consists in the generalization of the SFT to many-body lattice models far away from thermal equilibrium [15]. This requires a reformulation of the theory in terms of non-equilibrium Green's functions but offers the exciting perspective of studying the real-time dynamics of strongly correlated systems in a non-perturbative and consistent way. Another equally important direction of future work consists in an extension of the theory to correlated lattice models with non-local and even long-ranged interactions. First promising steps have already been made [8]. The restriction to local Hubbard-type interactions, inherent to all of the approaches mentioned here (see Fig. 11), represents an eventually unacceptable model assumption that must be abandoned.

Acknowledgement

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

References

- [1] M. Potthoff, Euro. Phys. J. B **32**, 429 (2003)
- [2] M. Potthoff, M. Aichhorn, and C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003)
- [3] M. Potthoff, AIP Conf. Proc. **1419**, 199 (2011)
- [4] M. Potthoff in *Strongly Correlated Systems: Theoretical Methods*, p. 303, ed. by A. Avella and F. Mancini, Springer Series in Solid-State Sciences, Vol. 171 (Springer, Berlin, 2012)
- [5] J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963)
- [6] M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
- [7] J. Kanamori, Prog. Theor. Phys. (Kyoto) **30**, 275 (1963)
- [8] N.-H. Tong, Phys. Rev. B **72**, 115104 (2005)
- [9] W. Koller and N. Dupuis, J. Phys.: Condens. Matter **18**, 9525 (2005)
- [10] M. Knap, E. Arrigoni, and W. von der Linden, Phys. Rev. B **81**, 235122 (2010)
- [11] M. Aichhorn, M. Hohenadler, C. Tahan, and P. B. Littlewood, Phys. Rev. Lett. **100**, 216401 (2008)
- [12] M. Knap, E. Arrigoni, and W. von der Linden, Phys. Rev. B **81**, 104303 (2010)
- [13] W. Koller, D. Meyer, Y. Ono, and A. C. Hewson, Europhys. Lett. **66**, 559 (2004)
- [14] M. Potthoff and M. Balzer, Phys. Rev. B **75**, 125112 (2007)
- [15] F. Hofmann, M. Eckstein, E. Arrigoni, and M. Potthoff, Phys. Rev. B **88**, 165124 (2013)
- [16] C. Dahnken, M. Aichhorn, W. Hanke, E. Arrigoni, and M. Potthoff, Phys. Rev. B **70**, 245110 (2004)
- [17] C. Gros and R. Valenti, Phys. Rev. B **48**, 418 (1993)
- [18] D. Sénéchal, D. Pérez, and M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000)
- [19] D. Sénéchal in *Strongly Correlated Systems: Theoretical Methods*, p. 237, ed. by A. Avella and F. Mancini, Springer Series in Solid-State Sciences, Vol. 171 (Springer, Berlin, 2012)
- [20] E. Koch: *The Lanczos Method* in
E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein (eds.):
The LDA+DMFT approach to strongly correlated materials
Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
<http://www.cond-mat.de/events/corre11>

- [21] R.P. Feynman, Phys. Rev. **56**, 340 (1939)
- [22] J.W. Gibbs: in: *The collected works of J. Willard Gibbs*, Vol. II, Chap. XI, pp. 129 (Yale University Press, New Haven, 1948)
- [23] R.P. Feynman, Phys. Rev. **97**, 660 (1955)
- [24] J.M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960)
- [25] A.A. Abrikosow, L.P. Gorkov, and I.E. Dzyaloshinski: *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey, 1964)
- [26] A.L. Fetter and J.D. Walecka: *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971)
- [27] J.W. Negele and H. Orland: *Quantum Many-Particle Systems* (Addison-Wesley, Redwood City, 1988)
- [28] M. Balzer and M. Potthoff, Phys. Rev. B **83**, 195132 (2011)
- [29] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)
- [30] H. Schweitzer and G. Czycholl, Solid State Commun. **74**, 735 (1990)
- [31] H. Schweitzer and G. Czycholl, Z. Phys. B **83**, 93 (1991)
- [32] M. Potthoff and W. Nolting, Z. Phys. B **104**, 265 (1997)
- [33] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [34] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
- [35] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [36] G. Kotliar and D. Vollhardt, Physics Today, March 2004, p. 53
- [37] R. Eder: *The Variational Cluster Approximation* in
E. Pavarini, E. Koch, and U. Schollwöck (eds.):
Emergent Phenomena in Correlated Matter
Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
<http://www.cond-mat.de/events/correl13>
- [38] M. Potthoff, Condens. Mat. Phys. **9**, 557 (2006)
- [39] G. Baym and L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)
- [40] G. Baym, Phys. Rev. **127**, 1391 (1962)
- [41] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)

- [42] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 235117 (2006)
- [43] M. Potthoff, Euro. Phys. J. B **36**, 335 (2003)
- [44] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, Phys. Rev. B **74**, 024508 (2006)
- [45] R. Freund: *Band Lanczos method*, in *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, ed. by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst (SIAM, Philadelphia, 2000)
- [46] M. Potthoff in *Effective models for low-dimensional strongly correlated systems*, ed. by G. Batrouni and D. Poilblanc (AIP Conf. Proc., Melville, 2006)
- [47] W. Press, S.A. Teukolsky, W.T. Vetterling, and B. Flannery: *Numerical Recipes* (Cambridge University Press, 3rd. ed., 2007)
- [48] R. Eder, Phys. Rev. B **76**, 241103(R) (2007)
- [49] R. Eder, Phys. Rev. B **78**, 115111 (2008)
- [50] L. Chioncel, H. Allmaier, E. Arrigoni, A. Yamasaki, M. Daghofer, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **75**, 140406 (2007)
- [51] R. Eder, Phys. Rev. B **81**, 035101 (2010)
- [52] M. Aichhorn, T. Saha-Dasgupta, R. Valenti, S. Glawion, M. Sing, and R. Claessen, Phys. Rev. B **80**, 115129 (2009)
- [53] H. Allmaier, L. Chioncel, and E. Arrigoni, Phys. Rev. B **79**, 235126 (2009)
- [54] H. Allmaier, L. Chioncel, E. Arrigoni, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **81**, 054422 (2010)
- [55] M. Balzer, W. Hanke, and M. Potthoff, Phys. Rev. B **77**, 045133 (2008)
- [56] E.H. Lieb and F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
- [57] N.D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966)
- [58] M. Balzer, B. Kyung, D. Sénéchal, A.-M.S. Tremblay, and M. Potthoff, Europhys. Lett. **85**, 17002 (2009)
- [59] F. Gebhard: *The Mott Metal-Insulator Transition* (Springer, Berlin, 1997)
- [60] T. Schäfer, F. Geles, D. Rost, G. Rohringer, E. Arrigoni, K. Held, N. Blümer, M. Aichhorn, and A. Toschi, arXiv:1405.7250
- [61] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)

- [62] G. Kotliar, S.Y. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
- [63] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **62**, R9283 (2000)
- [64] M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H.R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998)
- [65] G. Biroli, O. Parcollet, and G. Kotliar, Phys. Rev. B **69**, 205108 (2004)

10 Dynamical Vertex Approximation

Karsten Held

Institute for Solid State Physics

Vienna University of Technology

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Feynman diagrammatics | 3 |
| 2.1 | Parquet equations | 3 |
| 2.2 | Dynamical vertex approximation (DVA) | 10 |
| 3 | Two highlights | 16 |
| 3.1 | Critical exponents of the Hubbard model | 16 |
| 3.2 | Fate of the false Mott-Hubbard transition in two dimensions | 18 |
| 4 | Conclusion and outlook | 19 |

1 Introduction

The theoretical description and understanding of strongly correlated systems is particularly challenging since perturbation theory in terms of the Coulomb interaction is no longer possible and standard mean-field theory does not work. Also, bandstructure calculations in the local-density approximation (LDA) [1], which had been so successful for the calculations of many materials, do not work properly as electronic correlations are only rudimentarily taken into account. A big step forward in this respect is dynamical mean-field theory (DMFT) [2–5], which is a mean-field theory in the spatial coordinates but fully accounts for the local correlations in time (quantum fluctuations). In comparison, standard Hartree-Fock is mean-field in space *and* time. DMFT is non-perturbative since all Feynman diagrams are taken into account, albeit only their local contribution to the self-energy. If one is dealing with well localized d - or f -orbitals, these local DMFT correlations often provide the major part of the electronic correlations, which not only give rise to mass renormalizations [3, 4], metal-insulator transitions [3, 4] and magnetic ordering [6, 7], but also to unexpected new physics such as kinks [8, 9] or the filling of the Mott-Hubbard gap with increasing temperature [10]. More aspects of DMFT are discussed in other contributions to this Jülich Autumn School on DMFT at 25.

The question we would like to address here is: Can we do (or do we need to do) better than DMFT? Indeed, going beyond DMFT is necessary since many of the most fascinating and least understood physical phenomena such as quantum criticality and superconductivity originate from *non-local* correlations – which are by construction beyond DMFT. And we can: first steps to include non-local correlations beyond DMFT have been cluster approaches such as the dynamical cluster approximation (DCA) [11–13] and cluster DMFT (CDMFT) [12–14]. Here, instead of considering a single site embedded in a mean-field (as in DMFT) one considers a cluster of sites in a mean-field medium. Numerical limitations, however, restrict the DCA and CDMFT calculations to about 100 sites. This allows for studying short-range correlations, particularly for two-dimensional lattices, but severely restricts the approach for three dimensions, for multi-orbitals, and for long-range correlations.

Because of that, in recent years *diagrammatic extensions* of DMFT were at the focus of the methodological development. An early such extension was the $1/d$ (d : dimension) approach [15]; also, the combination of the non-local spin fermion self-energy with the local quantum fluctuations of DMFT has been proposed [16]. Currently most actively pursued are diagrammatic approaches based on the local two-particle vertex. The first such approach has been the dynamical vertex approximation (D Γ A) [17], followed by the dual fermion approach [18], the one-particle irreducible approach (1PI) [19], and DMFT to functional renormalization group (DMF 2 RG) [20].

The very idea of these approaches is to extend the DMFT concept of taking all (local) Feynman diagrams for the one-particle irreducible vertex (i.e., the self-energy) to the next, i.e., two-particle level. In these approaches, one calculates the local two-particle vertex, and from this, non-local correlations beyond DMFT are obtained diagrammatically. Indeed, we understand most (if not all) physical phenomena either on the one-particle level [e.g. the quasiparticle renor-

| Method | Local two-particle vertex | Feynman diagrams |
|--------------------------|---|---------------------------|
| DF [18] | one-particle reducible vertex, here* F_{loc} | 2nd order, ladder parquet |
| 1PI [19] | one-particle irreducible vertex F_{loc} | ladder |
| DMF ² RG [20] | one-particle irreducible vertex F_{loc} | RG flow |
| ladder D Γ A [17] | two-particle irreducible vertex in channel r $\Gamma_{r\text{loc}}$ | ladder |
| full D Γ A [17] | two-particle fully irreducible vertex Λ_{loc} | parquet |

Table 1: Summary of the different diagrammatic extensions of DMFT based on the two-particle vertex. All methods are based on the local part of the two-particle vertex named in the table; for a definition of the different vertex functions, see Section 2.1.

* Note that at the two-particle level every two-particle vertex is one-particle irreducible; third or higher order vertices can be, however, one-particle reducible which has consequences for the diagrammatics if truncated at the two-particle level, see [19].

malization and the Mott-Hubbard transition] or on the two-particle level [e.g. (para)magnons and (quantum) critical fluctuations]. Non-local correlations and associated physics on this two-particle level are included in these diagrammatic extensions of DMFT, which however still include the local DMFT one-particle physics such as the formation of Hubbard bands and the metal-insulator transition or, more precisely, a renormalized version thereof. The concept of all these approaches is similar, but they differ in which two-particle vertex is taken and which diagrams are constructed, see Table 1. Depending on the approach, Feynman diagrams are constructed from full Green function lines $G(\nu, \mathbf{k})$ or from the difference between $G(\nu, \mathbf{k})$ and the local Green function $G_{\text{loc}}(\nu)$ [ν : (Matsubara) frequency; \mathbf{k} : wave vector]. The DF, 1PI and DMF²RG approach are also based on a generating functional integral.

In these lecture notes, we will concentrate on D Γ A. Section 2.1 recapitulates the concept of reducible and irreducible diagrams as well as the parquet and Bethe-Salpeter equation. On this basis, we introduce in Section 2.2 the D Γ A approach. In Section 3 we have chosen two exemplary highlights that demonstrate what can be calculated by D Γ A and related approaches. These are the calculation of the critical exponents for the three dimensional Hubbard model (Section 3.1) and the effect of long-range correlations on the Mott-Hubbard transition for the two dimensional Hubbard model (Section 3.2): at zero temperature antiferromagnetic fluctuations always open a gap at any interaction $U > 0$.

2 Feynman diagrammatics

2.1 Parquet equations

The very idea of D Γ A is a resummation of Feynman diagrams, not in orders of the interaction U as in perturbation theory but in terms of the locality of diagrams. In this sense, DMFT is the first, one-particle level since it approximates the one-particle fully irreducible vertex, i.e., the self-energy Σ , to be local, see Table 2.

| | |
|------------------------|--|
| $n = 1$ | DMFT: local self-energy |
| $n = 2$ | DΓA: local fully irreducible two-particle vertex ⇒ non-local self-energy/correlations |
| ... | |
| $n \rightarrow \infty$ | exact solution |

Table 2: DΓA generalizes the DMFT concept of the local self-energy (i.e., the local fully irreducible one-particle vertex) to the fully irreducible n -particle vertex. It is hence a resummation of Feynman diagrams in terms of their locality. On the right-hand side, the different levels of approximation are indicated.

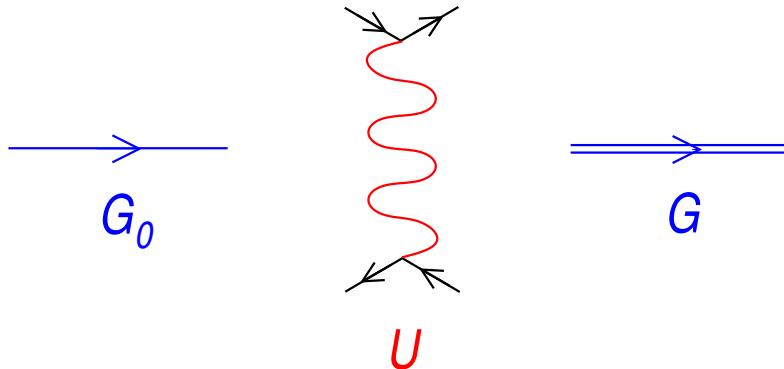


Fig. 1: Basic objects of Feynman diagrams: from the non-interacting Green function G_0 (left) and the bare interaction U (middle) we construct all topologically distinct diagrams for calculating the interacting Green function G (right).

For a better understanding of reducibility and irreducibility as well as of DΓA later on, let us recall that in quantum field theory we calculate the interacting Green function G by drawing all topologically distinct Feynman diagrams that consist of n interactions U and that are connected by non-interacting Green function lines G_0 , keeping one incoming and one outgoing G_0 line, see Fig. 1. Each G_0 line contributes a factor $G_0(\nu, \mathbf{k}) = 1/(\nu + \mu - \epsilon_{\mathbf{k}})$ [where ν denotes the (Matubara) frequency, μ the chemical potential and $\epsilon_{\mathbf{k}}$ the energy-momentum dispersion relation of the non-interacting problem] and each interaction (wiggled line) contributes a factor U .¹ Here and in the following, we assume a one-band model for the sake of simplicity. For an introduction to Feynman diagrams, more details, and how to evaluate Feynman diagrams including the proper prefactor, we refer the reader to textbooks of quantum field theory such as [21]; a more detailed presentation including DΓA can also be found in [22].

Dyson equation and self-energy

Instead of focusing on the Green function, we can consider a more compact object, the self-energy Σ , which is related to G through the Dyson equation, see Fig. 2. The Dyson equation

¹For a \mathbf{k} -dependent, i.e. non-local interaction the factor $U(\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}''' = \mathbf{k} + \mathbf{k}' - \mathbf{k}'')$ would be \mathbf{k} -dependent.

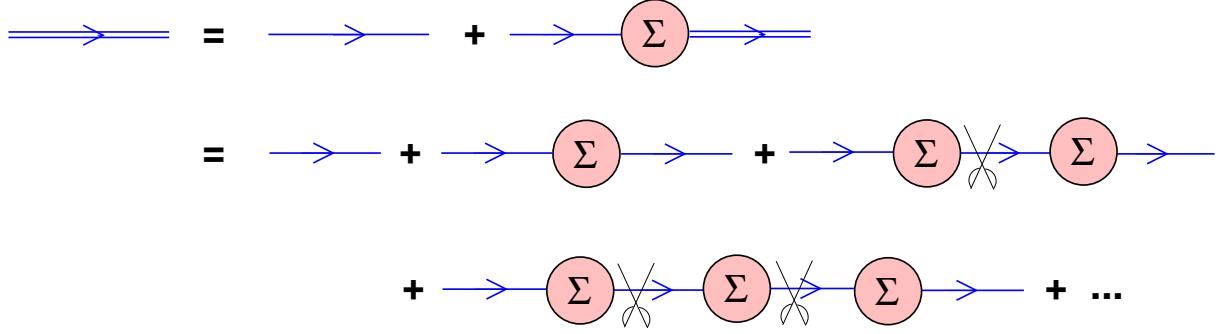


Fig. 2: Dyson equation connecting the Green function and self-energy. The pair of scissors indicates that these diagrams are one-particle reducible (i.e., cutting one G_0 line separates the Feynman diagram into two parts)

can be resolved for the interacting Green function:

$$G(\nu, \mathbf{k}) = [1/G_0(\nu, \mathbf{k}) - \Sigma(\nu, \mathbf{k})]^{-1}. \quad (1)$$

Since the geometric series of the the Dyson equation generates a series of Feynman diagrams, we can only include a reduced subset of Feynman diagrams when evaluating the self-energy. One obvious point is that the two outer “legs” (incoming and outgoing G_0 lines) are explicitly added when going from the self-energy to the Green function, see Fig. 2. Hence we have to “amputate” (omit) these outer “legs” for self-energy diagrams. More importantly, the self-energy can only include one-particle irreducible diagrams. Here, *one*-particle (ir)reducible means that by cutting *one* Green function line one can(not) separate the diagram into two parts. This is since, otherwise, we would generate Feynman diagrams twice: any one-particle reducible diagram can be constructed from two (or more) irreducible building blocks connected by one (or more) single G_0 lines. This is exactly what the Dyson equation does, see Fig. 2. For example, in the last line, we have three irreducible self-energy blocks connected by two single G_0 lines. This shows that, by construction, the self-energy has to include all one-particle irreducible Feynman diagrams and no one-particle reducible diagrams. Since the self-energy has *one* (amputated) incoming leg, it is a *one* particle vertex. It is also *one*-particle irreducible as explained above. Hence, the self-energy is the one-particle irreducible one-particle vertex. In Fig. 3 we show some diagrams that are part of the self-energy, i.e., cutting one-line does not separate the diagram into two parts, and some diagrams that are not.

Two-particle irreducibility

Let us now turn to the two-particle level. As illustrated in Fig. 4, the Feynman diagrams for the susceptibility χ (or similarly for the two-particle Green function) [21] consist of (i) an unconnected part (two G lines, as in the non-interacting case) and (ii) all connected Feynman diagrams (coined vertex corrections). Mathematically this yields [β : inverse temperature; σ :

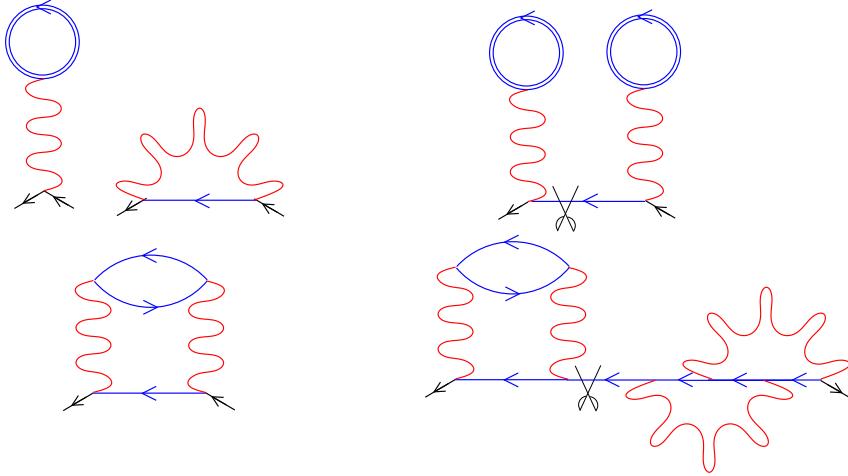


Fig. 3: Left: some examples of (one-particle irreducible) self-energy diagrams. Right: Diagrams that do not contribute to the self-energy since they are one-particle reducible (cutting the line indicated by the pair of scissors separates the diagram into two pieces).

spin]

$$\begin{aligned} \chi_{\sigma\sigma'}(\nu\nu'\omega; \mathbf{k}, \mathbf{k}', \mathbf{q}) = & -\beta G(\nu, \mathbf{k}) G(\nu + \omega, \mathbf{k} + \mathbf{q}) \delta_{\nu\nu'} \delta(\mathbf{k} - \mathbf{k}') \delta_{\sigma\sigma'} \\ & + G(\nu, \mathbf{k}) G(\nu + \omega, \mathbf{k} + \mathbf{q}) F_{\sigma\sigma'}(\nu\nu'\omega; \mathbf{k}, \mathbf{k}', \mathbf{q}) G(\nu', \mathbf{k}') G(\nu' + \omega, \mathbf{k}' + \mathbf{q}). \end{aligned} \quad (2)$$

Here F denotes the full, *reducible vertex*. In the following, let us introduce a short-hand notation for the sake of simplicity, where 1 represents a momentum-frequency-spin coordinate $1 \equiv (\mathbf{k}, \nu, \sigma)$, $2 \equiv (\mathbf{k} + \mathbf{q}, \nu + \omega, \sigma)$, $3 \equiv (\mathbf{k}' + \mathbf{q}, \nu' + \omega, \sigma')$, and $4 \equiv (\mathbf{k}', \nu', \sigma')$. In this notation we have

$$\chi(1234) = -G(14)G(23) - G(11')G(22')F(1'2'3'4')G(33')G(4'4), \quad (3)$$

as visualized in Fig. 4. Since the Green function is diagonal in spin, momentum, and frequency, i.e., $G(11') = G(11') \delta_{11'}$, some indices are the same, which yields Eq. (2).

Let us now again introduce the concept of irreducibility, this time for the two-particle vertex. In this case, we consider *two-particle irreducibility*.² In analogy to the self-energy, we define the *fully irreducible vertex* Λ , defined as the set of all Feynman diagrams that do not split into two parts by cutting two G lines. Let us remark that here and in the following, we construct the Feynman diagrams in terms of G instead of G_0 . This means that we have to exclude all diagrams that contain a structure as generated by the Dyson equation in Fig. 2 since otherwise these diagrams would be counted twice. The diagrams belonging to this reduced set with G instead of G_0 are called skeleton diagrams, see [21].

The *reducible diagrams* of F can be further classified according to how the Feynman diagram separates when cutting two internal Green functions. Since F has four (“amputated”) legs to the outside, there are actually *three* possibilities to split F into two parts by cutting two G lines.

²Note, that one-particle irreducibility is somehow trivial since one can show that there are no one-particle irreducible diagrams for the two-particle vertex (in terms of the interacting G /skeleton diagrams).

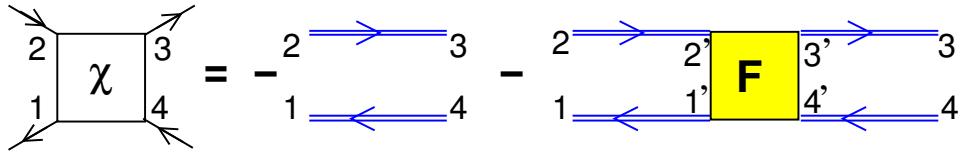


Fig. 4: The susceptibility χ consists of two unconnected Green function G lines (aka “bubble”) and vertex corrections F .

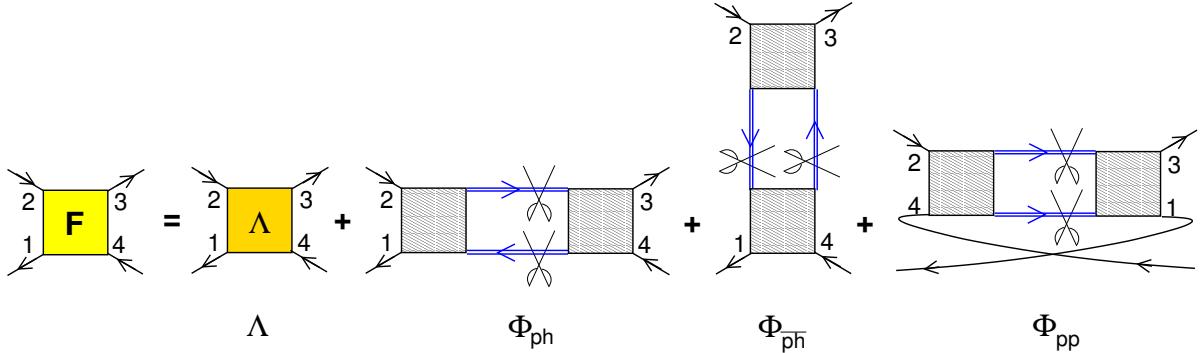


Fig. 5: The full (reducible) vertex F consists of the fully irreducible vertex Λ and two-particle reducible diagrams. These can be classified into three channels depending on which parts are disconnected by cutting two Green function lines. These are commonly denoted as the particle-hole reducible channel Φ_{ph} separating 12 from 34, the transversal particle-hole reducible channel $\Phi_{\bar{ph}}$ separating 14 from 23, and the particle-particle reducible channel Φ_{pp} separating 13 from 24. Each two-particle reducible diagram is reducible in one (and only one) of these three channels. The hatched blocks themselves can be irreducible, reducible in the same channel,* or reducible in the other two channels (in this last case the full diagram remains however reducible only in the scissors-indicated channel).

* Note, this is only possible for one hatched side, since otherwise the same diagram might be counted twice.

That is, an external leg, say 1, stays connected with one out of the three remaining external legs but is disconnected from the other two, see Fig. 5 for an illustration. One can show by means of the diagrammatic topology that each diagram is either fully irreducible or reducible in *exactly* one channel, so that

$$F(1234) = \Lambda(1234) + \Phi_{ph}(1234) + \Phi_{\bar{ph}}(1234) + \Phi_{pp}(1234). \quad (4)$$

Bethe-Salpeter equation

We have defined the reducible diagrams Φ_r in channel $r \in \{ph, \bar{ph}, pp\}$ as a subset of Feynman diagrams for F . The rest, i.e., $F - \Phi_r$, is called the vertex Γ_r *irreducible in r* so that

$$F(1234) = \Gamma_r(1234) + \Phi_r(1234). \quad (5)$$

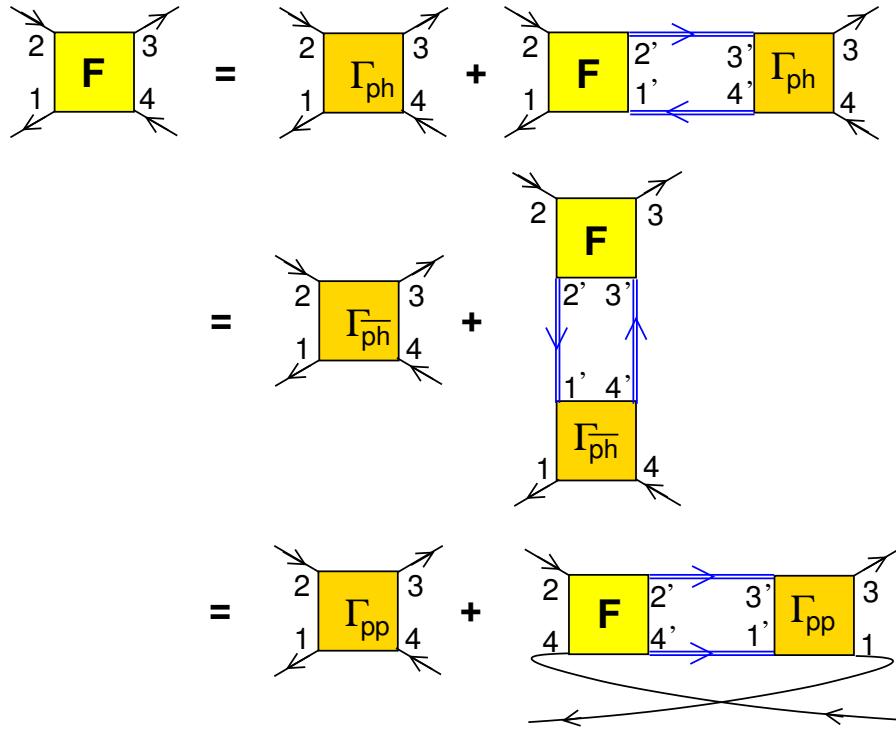


Fig. 6: Bethe-Salpeter equation for the particle-hole channel (ph , top) the transversal particle-hole channel (\bar{ph} , middle) and the particle-particle channel (pp , bottom).

In analogy to the Dyson equation, Fig. 2, the reducible vertices Φ_r in turn can be constructed from Γ_r . One Γ_r can be connected by two G 's with another Γ_r (which makes this diagram two-particle reducible in the channel r). This can be connected again by two G 's with a third Γ_r , etc. (allowing us to cut the two G 's at two or more different positions). This gives rise to a geometric ladder series, the so-called Bethe-Salpeter equation, see Fig. 6. Mathematically, these Bethe-Salpeter equations read in the three channels (with Einstein's summation convention):

$$F(1234) = \Gamma_{ph}(1234) + F(122'1') G(3'2') G(1'4') \Gamma_{ph}(4'3'34) \quad (6)$$

$$= \Gamma_{\bar{ph}}(1234) + F(2'233') G(2'1') G(3'4') \Gamma_{\bar{ph}}(11'4'4) \quad (7)$$

$$= \Gamma_{pp}(1234) + F(4'22'4) G(2'3') G(1'4') \Gamma_{pp}(13'31'). \quad (8)$$

Parquet equations

Since an irreducible Γ_r diagram in a channel r is either fully irreducible (Λ) or reducible in one of the two other channels $r' \neq r$ ($\Phi_{r'}$), we can express Γ_r as [this also follows directly from Eqs. (4) and (5)]:

$$\Gamma_r(1234) = \Lambda(1234) + \sum_{r' \neq r} \Phi_{r'}(1234) . \quad (9)$$

We can use this Eq. (9) to substitute the last Γ_r 's in Eq. (6) [or the Γ_r box in Fig. 6] by Λ and

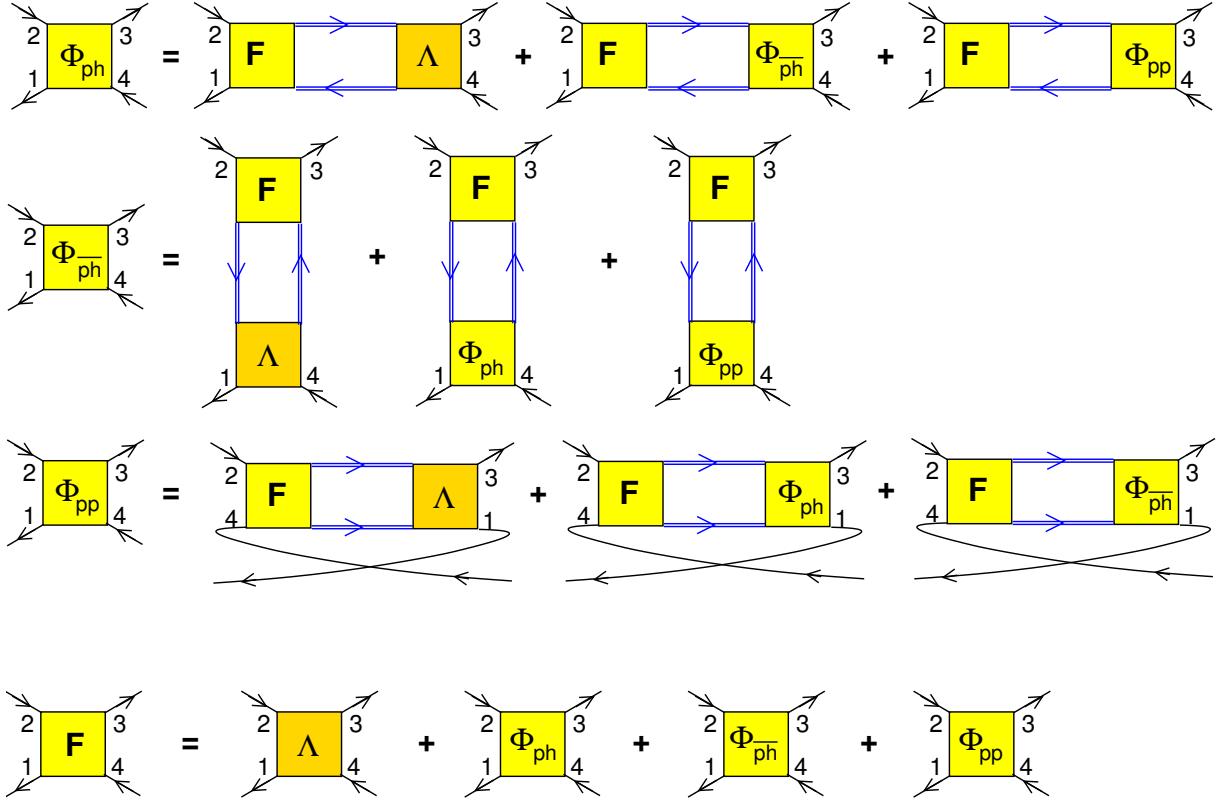


Fig. 7: Parquet equations. From the Bethe-Salpeter equations in the three channels we obtain three corresponding equations connecting the reducible vertex F , the fully irreducible vertex Λ and the reducible vertices Φ_r in the three channels r (first three lines). Together with the classification of F into Λ and the three Φ_r (last line) we have the four so-called parquet equations.

Φ_r 's. Bringing the first Γ_r on the left hand side, then yields

$$\Phi_{ph}(1234) = F(1234) - \Gamma_{ph}(1234) \quad (10)$$

$$= F(122'1') G(3'2') G(1'4') \Lambda(4'3'34) + \sum_{r' \neq ph} F(122'1') G(3'2') G(1'4') \Phi_{r'}(4'3'34)$$

and corresponding equations for the other two channels. The corresponding Feynman diagrams are shown in Fig. 7. If the fully irreducible vertex Λ is known, these three equations together with Eq. (4) allow us to calculate the four unknown vertex functions Φ_r and F , see Fig. 7. This set of equations is called the *parquet equations*.³ The solution can be done numerically by iterating these four *parquet equations*. Reflecting how we arrived at the parquet equations, the reader will realize that the parquet equations are nothing but a classification of Feynman diagrams into fully irreducible diagrams and diagrams reducible in the three channels r .

³Sometimes, only Eq. (5) is called the parquet equation and the equations of type Eq. (10) remain under the name Bethe-Salpeter equations.

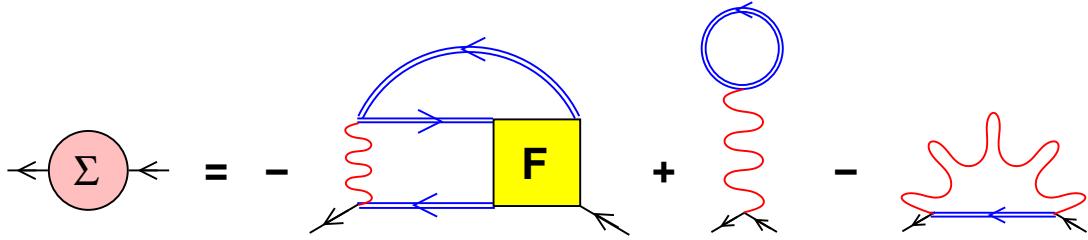


Fig. 8: Equation of motion (Schwinger-Dyson equation) for calculating the self-energy Σ from the bare interaction U , the reducible vertex F and the Green function lines G . The second and third diagram on the right-hand side are the Hartree and Fock diagrams, respectively, which are not included in the first term.

The thoughtful reader will have also noticed that the interacting Green function enters Eq. (10). This G can be calculated by one additional layer of self consistency. If the reducible vertex F is known, the self-energy follows from the Heisenberg equation of motion (also called Schwinger-Dyson equation in this context). This is illustrated in Fig. 8 and mathematically reads

$$\Sigma(14) = -U(1'2'3'1')G(1'4')G(23')G(2'3)F(4'234) + U(1234)G(23) - U(1432)G(23) \quad (11)$$

That is, the numerical solution of the four parquet equations has to be supplemented by the Schwinger-Dyson Eq. (11) and the Dyson Eq. (1), so that also G and Σ are calculated self-consistently.

Let us also note that these general equations, while having a simple structure in the 1234 notation, can be further reduced for practical calculations: the Green functions are diagonal $G(3'2') = G(3'3')\delta_{2'3'}$, there is a severe restriction in spin, there is SU(2) symmetry and one can decouple the equations into charge(spin) channels $\Gamma_{ph\uparrow\uparrow} + (-)\Gamma_{ph\uparrow\downarrow}$. A detailed discussion is beyond the scope of these lecture notes. For more details on the parquet equations see [23], for a derivation of the equation of motion also see [24].

2.2 Dynamical vertex approximation (D Γ A)

Hitherto, everything has been exact. If we know the exact fully irreducible vertex Λ , we can calculate through the parquet equations, Fig. 7 [Eqs. (5) and (10)], the full vertex F ; from this, through the Schwinger-Dyson equation of motion (11), the self-energy Σ ; and through the Dyson Eq. (1), the Green function G . With a new G we can (at fixed Λ) recalculate F , etc. until convergence. Likewise, if we know the exact irreducible vertex Γ_r in one channel r , we can calculate F through the corresponding Bethe-Salpeter Eq. (6-8) and from this (self consistently) obtain Σ and G .

But Λ (or Γ_r) still consists of an infinite set of Feynman diagrams that we usually do not know. Since the parquet (or Bethe-Salpeter) equations generate many additional diagrams, there are, however, many fewer (albeit still infinitely many) diagrams for Λ (or Γ_r) than for F . In the case of Λ , the bare interaction U is included but the next term is already a diagram of fourth(!) order in U (the so-called envelope diagram), see Fig. 9. There are no two-particle fully irreducible diagrams of second or third order in U .

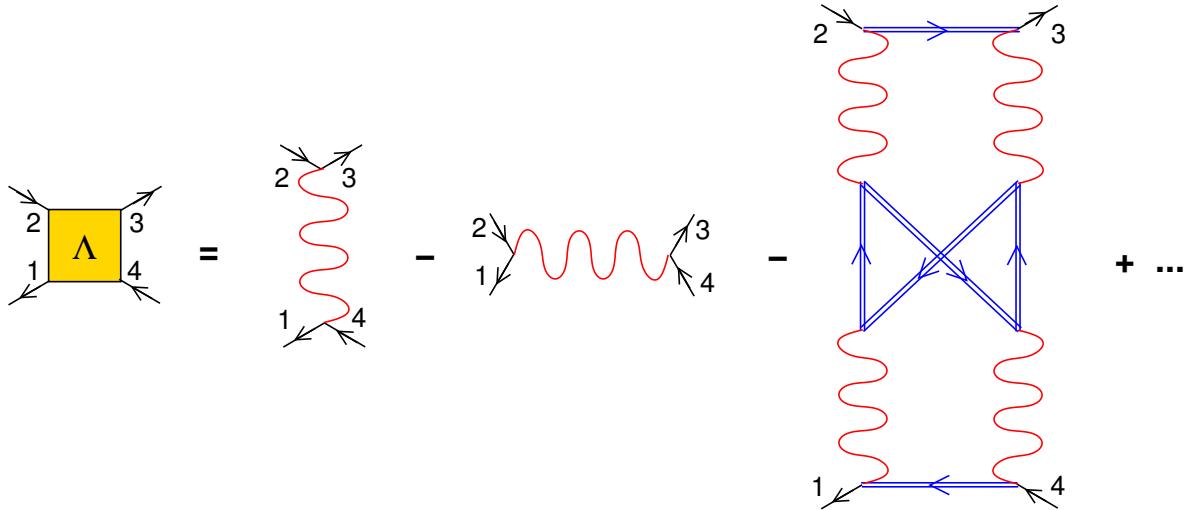


Fig. 9: Lowest order Feynman diagrams for the fully irreducible vertex Λ .

One approach is hence to approximate Λ by the bare interaction U only, i.e., the first two terms in Fig. 9. This is the so-called parquet approximation [23]. For strongly correlated electrons this will not be enough. A deficiency is, for example, that the parquet approximation does not yield Hubbard bands.

In DΓA, we hence take instead all Feynman diagrams for Λ but restrict ourselves to their local contribution, Λ_{loc} . This approach is non-perturbative in the local interaction U . It is putting the DMFT concept of locality to the next, i.e., to the two-particle, level. We can extend this concept to the n -particle fully irreducible vertex, so that by increasing n systematically more and more Feynman diagrammatic contributions are generated; and for $n \rightarrow \infty$ the exact solution is recovered, see Table 2 above.

In practice, one has to truncate this scheme at some n , hitherto at the two-particle-vertex level ($n = 2$). The local fully irreducible two-particle vertex Λ can be calculated by solving an Anderson impurity model that has the same local U and the same Green function G . This is because such an Anderson impurity model yields exactly the same (local) Feynman diagrams Λ_{loc} . It is important to note that the locality for Λ is much better fulfilled than that for Σ . Even in two dimensions, Λ is essentially k -independent, i.e., local. This has been demonstrated by numerical calculations for the two-band Hubbard model, see [25]. In contrast, for the same set of parameters Σ is strongly k -dependent, i.e., non-local. Also Γ_r and F are much less local than Λ , see [25]. There might be parameter regions in two dimensions or one-dimensional models where Λ also exhibits a sizable non-local contribution. One should keep in mind that DΓA at the $n = 2$ level is still an approximation. This approximation includes however not only DMFT but on top of that also non-local correlations on all length scales so that important physical phenomena can be described, and even in two dimensions substituting Λ by its local contribution Λ_{loc} is a good approximation, better than replacing Σ by its local contribution Σ_{loc} as in DMFT.

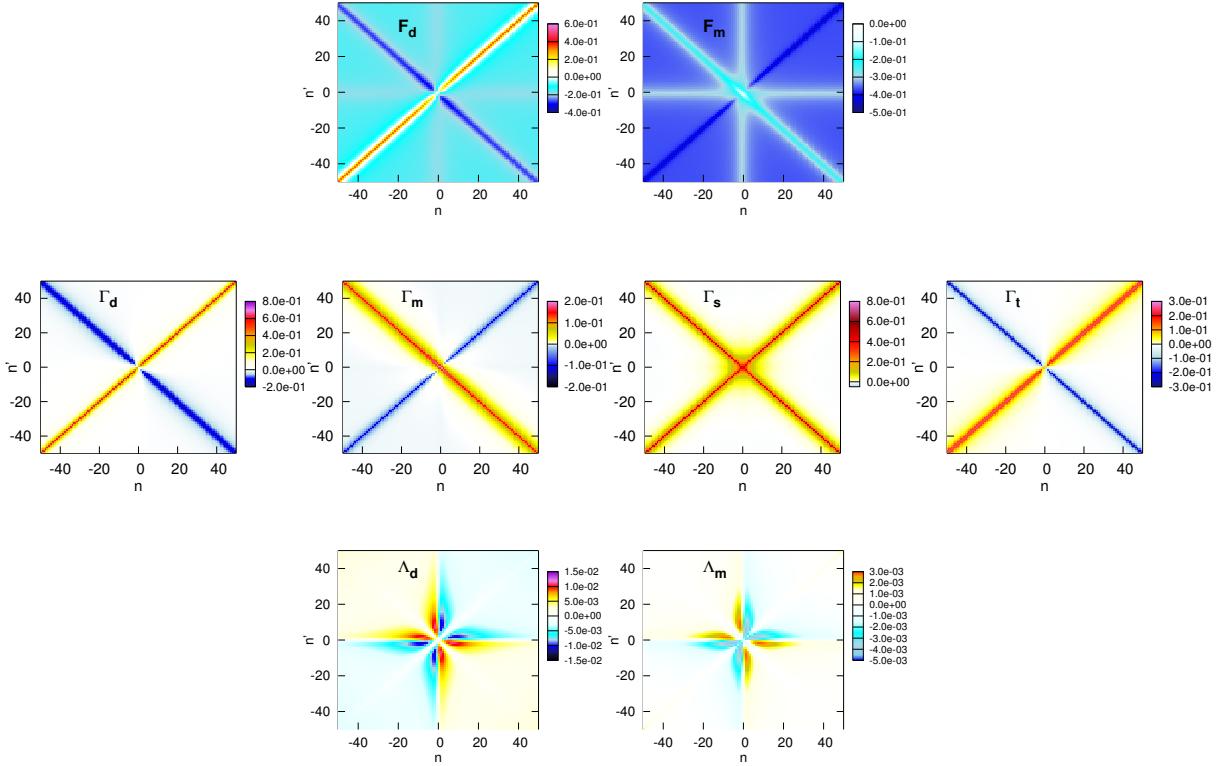


Fig. 10: Top: Local full vertex F in the magnetic (m) and charge or density (d) channel, i.e., $F_{\uparrow\uparrow} \pm F_{\uparrow\downarrow}$, as a function of the two incoming fermionic frequencies $\nu = (2n + 1)\pi T$ and $\nu' = (2n' + 1)\pi T$ at transferred bosonic frequency $\omega = 0$ for the three dimensional Hubbard model at $U = 0.5$, $T = 1/26$ (in units of nearest neighbor hopping $2\sqrt{6}t \equiv 1$). Middle: Corresponding particle-hole irreducible vertex Γ_{ph} (two left panels) and particle-particle vertex in the singlet (s) and triplet (t) spin combination. The transversal particle-hole channel follows from Γ_{ph} by (crossing) symmetry. Bottom: Fully irreducible vertex $\Lambda_{m(d)}$ for the two spin combinations. For all figures, the bare interaction U has been subtracted from the vertices (reproduced from [26]).

When solving the Anderson impurity model numerically, one does not obtain Λ_{loc} directly but first the local susceptibilities χ_{loc} (or the two-particle Green function). Going from here to Λ_{loc} is possible as follows: from χ_{loc} and G_{loc} , we obtain the local reducible vertex F_{loc} [via the local version of Eq. (2)], from this in turn we get Γ_{rloc} [via inverting the local version of the Bethe-Salpeter Eqs. (6-8)], Φ_{rloc} [via the local version of Eq. (5)], and finally Λ_{loc} [via the local version of Eq. (4)].

Fig. 10 shows the reducible vertex F , the irreducible vertex in the particle-hole channel Γ_{ph} and the fully irreducible vertex Λ for the three-dimensional Hubbard model on a simple cubic lattice. For more details on the calculation of Λ_{loc} , see [26] and [22]. Also note that Γ_{rloc} diverges at an interaction strength U below the Mott-Hubbard metal-insulator transition, which signals the breakdown of perturbation theory [27].

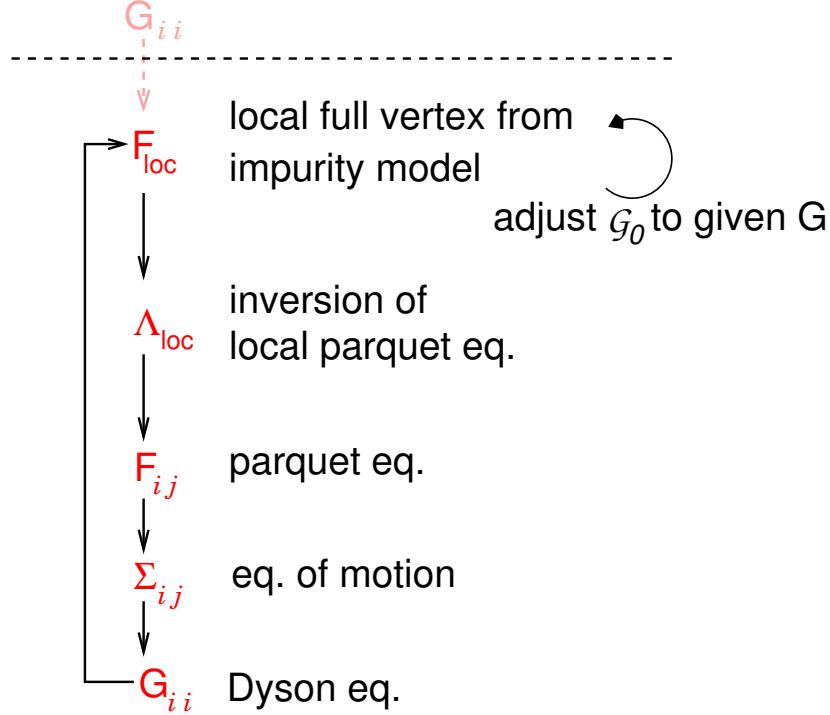


Fig. 11: Flow diagram of the DΓA approach. Starting from a test Green function G (e.g., that of DMFT), the local susceptibility and local full vertex F_{loc} is calculated by solving an Anderson impurity model. To this end, the non-local Green function \mathcal{G}_0 of an Anderson impurity is adjusted by Eq. (12) until this \mathcal{G}_0 impurity model has the given interacting G . From F_{loc} in turn, the inversion of the parquet and Bethe-Salpeter equations allow the calculation of the local fully irreducible vertex Λ_{loc} . This is the input of the parquet equations for calculating the non-local vertex F and through the equation of motion and the Dyson equation, the DΓA self-energy Σ and Green function G . In a self-consistent calculation a new local vertex is calculated from G etc. until convergence.

2.2.1 Self consistency

After calculating Λ_{loc} , we can calculate the full vertex F through the parquet equations, Fig. 7; and through the Schwinger-Dyson equation, Fig. 8, the non-local self-energy Σ and Green function G , as discussed in Section 2.1. Hitherto, all DΓA calculations have stopped at this point. That is, F and G are determined self-consistently but Λ_{loc} is not recalculated.

However, in principle, one can self-consistently iterate the approach. From the new G we can calculate a new G_{loc} . From this and U we obtain a new vertex, etc. until convergence, see Fig. 11. This self-consistency cycle is similar to that of the DMFT but now includes self-consistency on the two-particle (Λ) level.

Please note that the Anderson impurity model has now to be calculated with the interacting G_{loc} from DΓA, which is different from the G_{loc} of DMFT. As numerical approaches solve the Anderson impurity model for a given non-interacting Green function \mathcal{G}_0 , we need to adjust this \mathcal{G}_0 until the Anderson impurity model's Green function G agrees with the DΓA G_{loc} . This is

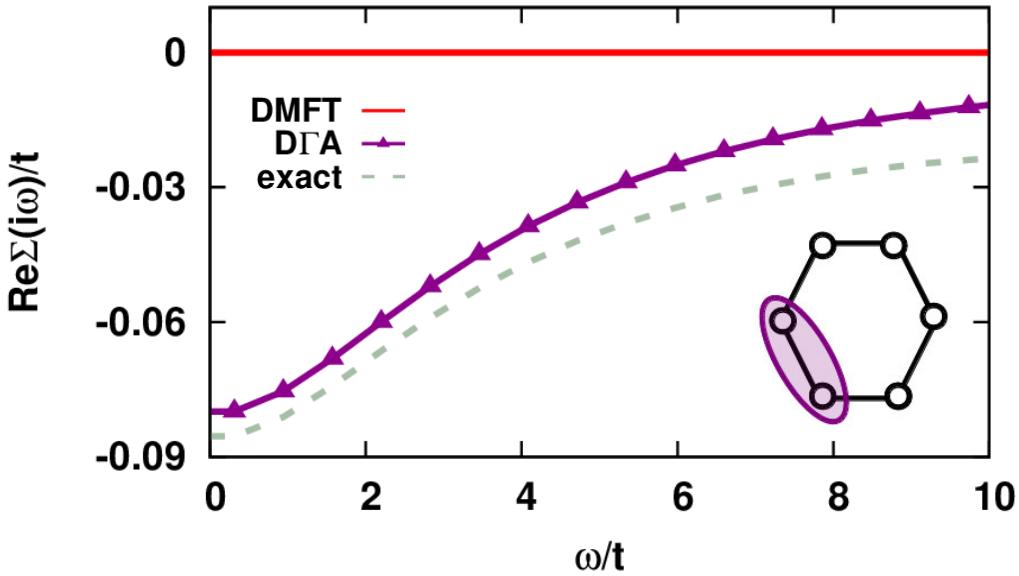


Fig. 12: Self-energy Σ_{ii+1} between neighboring sites on a six-site Hubbard ring (see inset), comparing DMFT (zero non-local Σ_{ii+1}), D Γ A and the exact solution. Parameters: $U = 2t$, $T = 0.1t$ with t being the nearest-neighbor hopping on the ring (reproduced from [29]).

possible by iterating \mathcal{G}_0 as follows:

$$[\mathcal{G}_0^{\text{new}}(\nu)]^{-1} = [\mathcal{G}_0^{\text{old}}(\nu)]^{-1} + [G_{\text{loc}}(\nu)]^{-1} - [G^{\text{old}}(\nu)]^{-1}, \quad (12)$$

until convergence. Here, $\mathcal{G}_0^{\text{old}}$ and G^{old} denote the non-interacting and interacting Green function of the Anderson impurity model from the previous iteration. This \mathcal{G}_0 -adjustment is indicated in Fig. 11 by the secondary cycle.

Fig. 12 compares the D Γ A self energies calculated this way, i.e., the D Γ A full parquet solution, with DMFT and the exact solution. The results are for a simple one-dimensional Hubbard model with nearest neighbor hopping, six sites and periodic boundary conditions so that the exact solution is still possible by an exact diagonalization of the Hamiltonian. This can be considered as a simple model for a benzene molecule. The D Γ A results have been obtained in a “one-shot” calculation with the DMFT G as a starting point. The good agreement between D Γ A and the exact solution shows that D Γ A can be employed for quantum chemistry calculations of correlations in molecules, at least if there is a gap [HOMO-LUMO gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO)]. For molecules with degenerate ground states (i.e., a peak in the spectral function at the Fermi level), the agreement is somewhat less impressive; note that one dimension is the worst possible case for D Γ A.

In Fig. 12, the first parquet D Γ A results have been shown. However, most D Γ A calculations hitherto employed a simplified scheme based on ladder diagrams, see Fig. 13. These calculations neglect one of the three channels in the parquet equations: the particle-particle channel. Both the particle-hole (ph) and the transversal particle-hole channel (\bar{ph}) are taken into account. These channels decouple for the spin and charge vertex and can hence be calculated by solving

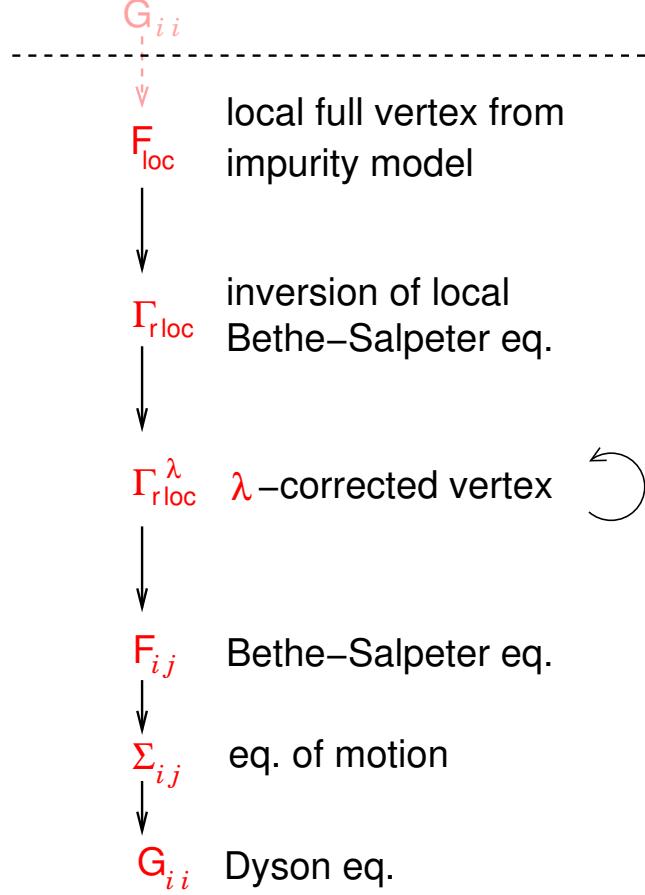


Fig. 13: Flow diagram of the DGA ladder approach. Same as Fig. 11 but solving the Bethe-Salpeter equation(s) with the local irreducible vertex $\Gamma_{r\text{loc}}$ in (two) channel(s) r instead of the parquet equations with the fully irreducible vertex Λ . Instead of the self-consistency, it is better to employ a so-called Moriya λ -correction in this case.

the simpler Bethe-Salpeter equations instead of the full parquet equations. If one neglects non-local contributions in one of the channels, it is better to restrict oneself to non-self-consistent calculations since part of the neglected diagrams cancel with diagrams generated by the self consistency. Instead one better does a “one shot” calculation mimicking the self-consistency by a so-called λ correction, see [30] for details. Physically, neglecting the particle-particle channel is justified if non-local fluctuations of particle-particle type are not relevant. Whether this is the case or not depends on the model and parameter range studied. Such particle-particle fluctuations are, e.g., relevant in the vicinity of superconducting instabilities, where they need to be considered. In the vicinity of antiferromagnetic order on the other hand, the two particle-hole channels are the relevant ones (and only their magnetic spin combination). These describe antiferromagnetic fluctuations, coined paramagnons in the paramagnetic phase above the antiferromagnetic transition temperature. The DGA results in the next section are for the half-filled Hubbard model. Here we are not only away from any superconducting instability, but at half-filling the interaction also suppresses particle-particle fluctuations. Hence, for the results presented below using the ladder instead of the full parquet approximation is most reasonable.

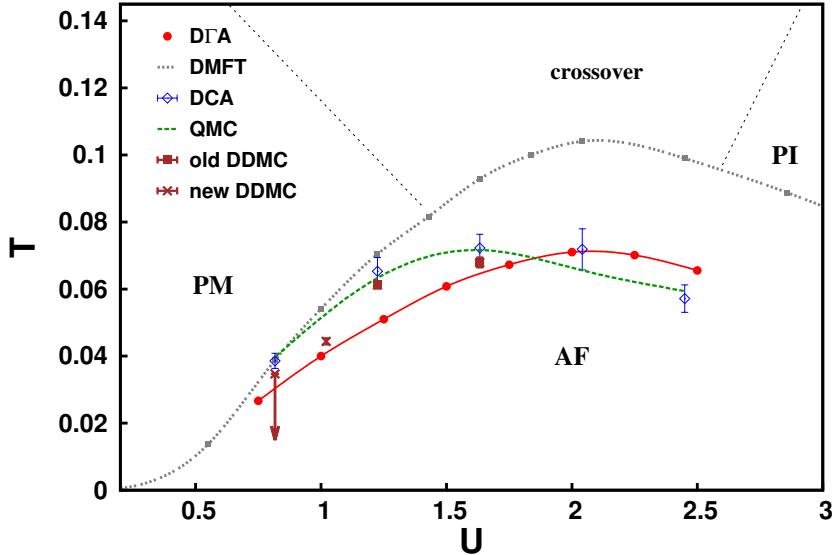


Fig. 14: Phase diagram of the Hubbard model on a cubic lattice with nearest neighbor hopping $2\sqrt{6}t \equiv 1$. The dashed black lines show the DMFT Néel temperature T_N and the DMFT crossover region from a paramagnetic metal (PM) to a paramagnetic insulator (PI). Non-local correlations reduce T_N with good agreement between D Γ A [31], DCA [32], lattice quantum Monte Carlo (QMC) [33], as well as determinantal diagrammatic Monte Carlo (DDMC) before [34] and after [35] our D Γ A results. Note, for the lowest U value, Ref. [35] could only give an upper bound for T_N which according to DDMC could be much smaller as indicated by the arrow.

3 Two highlights

3.1 Critical exponents of the Hubbard model

The Hubbard model is the prototypical model for strong electronic correlations. At half filling it shows an antiferromagnetic ordering at low enough temperatures T – for all interaction strengths $U > 0$ if the lattice has perfect nesting. Fig. 14 shows the phase diagram of the half-filled three-dimensional Hubbard model (all energies are in units of $D = 2\sqrt{6}t \equiv 1$, which has the same standard deviation as a Bethe lattice with half bandwidth D). At weak interaction strength, we have a Slater antiferromagnet which can be described by Hartree-Fock theory yielding an exponential increase of the Néel temperature with U . At strong interactions, we have preformed spins with a Heisenberg interaction $J = 4t^2/U$ yielding a Heisenberg antiferromagnet with $T_N \sim J$. In-between these limits, T_N is maximal.

All of this can be described by DMFT. However, since the DMFT is mean-field with respect to the spatial dimensions, it overestimates T_N . This can be overcome by including non-local correlations, i.e., spatial (here antiferromagnetic) fluctuations. These reduce T_N . In this respect, there is a good agreement between D Γ A, DCA and lattice QMC, see Fig. 14. The biggest deviations are observed for the smaller interaction strength. In principle, these differences might originate from the fact that the D Γ A calculations are not yet self-consistent and only use the Bethe-Salpeter Eqs. (6) and (7) in the two particle-hole channels instead of the full parquet

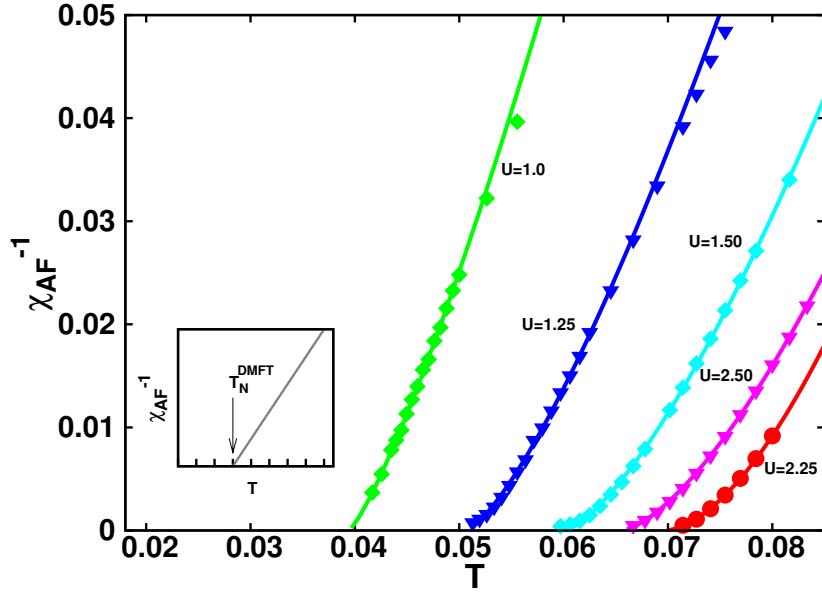


Fig. 15: Inverse antiferromagnetic spin susceptibility as a function of T for different interactions U as obtained by D Γ A (lower inset: DMFT), showing in the vicinity of the phase transition a $\chi_{AF} \sim (T - T_N)^{-2\nu}$ behavior, with a critical exponent $\nu \sim 0.7$ in agreement with the three dimensional Heisenberg model [36]. (reproduced from [31] © by the American Physical Society).

equations (5) and (10). On the other hand, however, we observe that long-range correlations are particularly important at weak coupling, cf. Section 3.2. Such long-ranged correlations cannot be captured by the cluster extensions of DMFT or lattice QMC since these are restricted to maximally ~ 10 sites in all three directions. More recent and accurate DDMC calculations on larger clusters [35] indeed show a smaller T_N and, in contrast to DCA and lattice QMC, better agree with D Γ A, see Fig. 14.

Even more important are long-range correlations in the immediate vicinity of the phase transition and for calculating the critical exponents. Each finite cluster will eventually show a mean-field exponent. In this respect, we could calculate for the first time the critical exponents of the Hubbard model [31]. In Fig. 15 we show the antiferromagnetic spin susceptibility, which shows a mean-field-like behavior $\chi_{AF} \sim (T - T_N)^{-1}$ at high temperature (and in DMFT). In the vicinity of the phase transition, however, long-range correlations become important and yield another critical exponent $\chi_{AF} \sim (T - T_N)^{-2 \times 0.7}$, which agrees with that of the three dimensional Heisenberg model [36] (as is to be expected from universality). In contrast, for the Falikov-Kimball model, the critical exponents calculated by the related Dual-Fermion approach [37] agree with those of the Ising model.

Except for the immediate vicinity of the phase transition, DMFT nonetheless yields a reliable description of the paramagnetic phase, at least for one-particle quantities such as the self-energy and spectral function. The susceptibility [31] and entropy [34] show deviations in a larger T -interval above T_N .

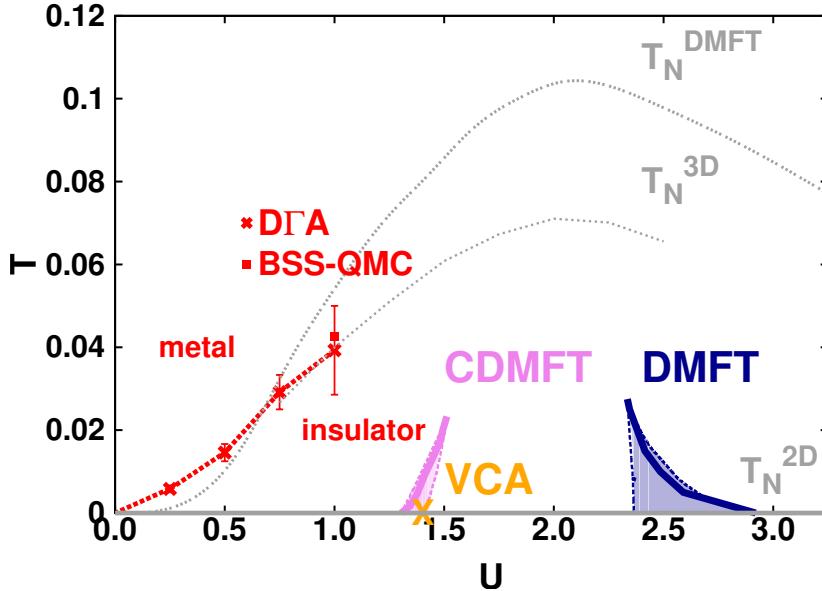


Fig. 16: Temperature T vs. interaction U phase diagram of the two dimensional Hubbard model on a square lattice with nearest-neighbor hopping (all energies are in units of $D = 4t \equiv 1$, yielding the same standard deviation as for the 3D phase diagram). From local correlations (DMFT [39]), via short-range correlations (CDMFT [40] and variational cluster approximation (VCA) [38]), to long-range correlations ($D\Gamma A$ [38] and BSSQMC [38]), the critical interaction strength for the metal-insulator transition is reduced to $U_c = 0$. The light gray lines denote the DMFT [32] and $D\Gamma A$ [31] T_N (reproduced from [38]).

3.2 Fate of the false Mott-Hubbard transition in two dimensions

As a second example, we review results for the interplay between antiferromagnetic fluctuations and the Mott-Hubbard transition in two dimensions. Even though the Hubbard model has been studied for 50 years, astonishingly little is known exactly. In one dimension it can be solved by the Bethe ansatz, and there is no Mott-Hubbard transition for the half-filled Hubbard model: for any finite interaction it is insulating. In infinite dimensions, on the other hand, DMFT provides for an exact (numerical) solution. It has been one of the big achievements of DMFT to clarify the nature of the Mott-Hubbard transition, which is of first order at a finite interaction strength [3,4], see Fig. 16. From cluster extensions of DMFT, it has been concluded that the Mott-Hubbard transition is actually at somewhat smaller U values and the coexistence region where two solutions can be stabilized is smaller, see Fig. 16. However, again, these cluster extensions are restricted to *short-range* correlations. In particular at low temperatures, there are strong *long-range* antiferromagnetic spin fluctuations, which for example at $U = 0.5$ and $T = 0.01$ exceed 300 lattice sites [38]. The physical origin are antiferromagnetic fluctuations emerging above the antiferromagnetically ordered phase. In two dimensions, this antiferromagnetic phase is restricted to $T = 0$ due to the Mermin-Wagner theorem, but antiferromagnetic fluctuations remain strong even beyond the immediate vicinity of the ordered phase (at $T = 0$). These long-range antiferromagnetic spin fluctuations (paramagnons) give rise to pseudogap physics, where first only part of the Fermi surface becomes gapped but at lower temperatures the entire Fermi sur-

face is gapped so that we have an insulating phase for any $U > 0$. Fig. 16 shows the development from local correlations (which only yield an insulating phase for relatively large U) to additional short-range correlations (which reduce the critical U_c for the Mott-Hubbard transition) to long-range correlations (which reduce U_c to zero). At large U , we have localized spins that can be described by a spin model or the Mott-insulating phase of DMFT. At smaller U , we also have an insulator caused by antiferromagnetic spin fluctuations. These smoothly go over into the $T = 0$ antiferromagnetic phase, which is of Slater type for small U . Since the correlations are exceedingly long-ranged, the nature of the low-temperature gap is the same as the Slater antiferromagnetic gap, even though there is no true antiferromagnetic order yet.

4 Conclusion and outlook

In recent years, we have seen the emergence of diagrammatic extensions of DMFT. All these approaches have in common that they calculate a local vertex and construct diagrammatically non-local correlations from this vertex. In regions of the phase diagram where non-local correlations are short-range, results are similar as for cluster extensions of DMFT. However, the diagrammatic extensions also offer the opportunity to include long-range correlations on an equal footing. This allowed us to study critical phenomena and to resign the Mott-Hubbard transition in the two-dimensional Hubbard model to its fate (there is no Mott-Hubbard transition).

These were just the first steps. Indeed, the diagrammatic extensions offer a new opportunity to address the hard problems of solid state physics, from superconductivity and quantum criticality to quantum phenomena in nano- and heterostructures. Besides a better physical understanding by means of model systems, also realistic materials calculations are possible – by *Ab initioDΓA* [41]. Taking the bare Coulomb interaction and all local vertex corrections as a starting point, *Ab initioDΓA* includes DMFT, *GW* and non-local correlations beyond within a common underlying framework. Both on the model level and for realistic materials calculations, there is plenty of physics to explore.

Acknowledgment

I sincerely thank my coworkers S. Andergassen, A. Katanin, W. Metzner, T. Schäfer, G. Rohringer, C. Taranto, A. Toschi, and A. Valli for the fruitful cooperations on diagrammatic extensions of DMFT in the last years, and G. Rohringer also for reading the manuscript. This work has been financially supported in part by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC through grant agreement n. 306447 (*Ab initioDΓA*) and in part by the Research Unit FOR 1346 of the Deutsche Forschungsgemeinschaft and the Austrian Science Fund (project ID I597).

References

- [1] R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
- [2] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [3] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [4] A. Georges, G. Kotliar, W. Krauth and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [5] G. Kotliar and D. Vollhardt, Physics Today, March 2004, p. 53
- [6] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
- [7] K. Held and D. Vollhardt, Eur. Phys. J. B **5**, 473 (1998)
- [8] K. Byczuk, M. Kollar, K. Held, Y.-F. Yang, I. A. Nekrasov, T. Pruschke and D. Vollhard, Nature Physics **3** 168 (2007);
A. Toschi, M. Capone, C. Castellani, and K. Held, Phys. Rev. Lett. **102**, 076402 (2009);
C. Raas, P. Grete, and G.S. Uhrig, Phys. Rev. Lett. **102**, 076406 (2009)
- [9] K. Held, R. Peters, A. Toschi, Phys. Rev. Lett. **110**, 246402 (2013)
- [10] S.-K. Mo, H.-D. Kim, J. W. Allen, G.-H. Gweon, J.D. Denlinger, J. -H. Park, A. Sekiyama, A. Yamasaki, S. Suga, P. Metcalf, and K. Held, Phys. Rev. Lett. **93**, 076404 (2004)
- [11] T. Maier, M. Jarrell, T. Pruschke and M. H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
- [12] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **62**, 9283 (R) (2000)
- [13] E. Koch: *Quantum cluster methods* in
E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein, (eds.):
DMFT at 25: Infinite Dimensions,
Reihe Modeling and Simulation, Vol. 4 (Forschungszentrum Jülich, 2014)
<http://www.cond-mat.de/events/correl14>
- [14] G. Kotliar, S. Y. Savrasov, G. Pálsson and G. Biroli, Phys. Rev. Lett. **87** 186401 (2001)
- [15] A. Schiller and K. Ingersent, Phys. Rev. Lett. **75**, 113 (1995)
- [16] E.Z. Kuchinskii, I.A. Nekrasov and M.V. Sadovskii, Sov. Phys. JETP Lett. **82**, 98 (2005)
- [17] A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. B **75**, 045118 (2007);
Prog. Theor. Phys. Suppl. **176**, 117 (2008)
- [18] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008)
- [19] G. Rohringer, A. Toschi, H. Hafermann, K. Held, V.I. Anisimov, and A.A. Katanin,
Phys. Rev. B **88**, 115112 (2013)

- [20] C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, Phys. Rev. Lett. **112**, 196402 (2014)
- [21] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski: *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1963)
- [22] G. Rohringer, PhD thesis, Vienna University of Technology (2013)
- [23] N.E. Bickers, *Self-Consistent Many-Body Theory of Condensed Matter in Theoretical Methods for Strongly Correlated Electrons*, CRM Series in Mathematical Physics Part III, ed. by D. Sénéchal, A.M. Tremblay, and C. Bourbonnais (Springer, New York 2004)
- [24] K. Held, C. Taranto, G. Rohringer, A. Toschi: *Hedin Equations, GW, GW+DMFT, and All That* in E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein, (eds.): *The LDA+DMFT approach to strongly correlated materials*, Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011) <http://www.cond-mat.de/events/correl11>
- [25] T.A. Maier, M.S. Jarrell, and D.J. Scalapino, Phys. Rev. Lett. **96**, 047005 (2006)
- [26] G. Rohringer, A. Valli, and A. Toschi, Phys. Rev. B **86**, 125114 (2012)
- [27] T. Schäfer, G. Rohringer, O. Gunnarsson, S. Ciuchi, G. Sangiovanni, and A. Toschi, Phys. Rev. Lett. **110**, 246405 (2013)
- [28] For details on the numerical solution of the parquet equation and how to stabilize it enforcing among others the crossing symmetry, see S.X. Yang, H. Fotso, J. Liu, T.A. Maier, K. Tomko, E.F. D'Azevedo, R.T. Scalettar, T. Pruschke, and M. Jarrell, Phys. Rev. E **80**, 046706 (2009); K.-M. Tam, H. Fotso, S.-X. Yang, T.-W. Lee, J. Moreno, J. Ramanujam, and M. Jarrell, Phys. Rev. E **87**, 013311 (2013)
- [29] A. Valli, PhD thesis, Vienna University of Technology (2013)
- [30] A.A. Katanin, A. Toschi, and K. Held, Phys. Rev. B **80**, 075104 (2009)
- [31] G. Rohringer, A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. Lett. **107**, 256402 (2011)
- [32] P.R.C. Kent, M. Jarrell, T.A. Maier, and T. Pruschke, Phys. Rev. B **72**, 060411 (2005)
- [33] R. Staudt, M. Dzierzawa, and A. Muramatsu, Eur. Phys. J. B **17**, 411 (2000)
- [34] S. Fuchs, E. Gull, L. Pollet, E. Burovski, E. Kozik, T. Pruschke, and M. Troyer, Phys. Rev. Lett. **106**, 030401 (2011)
- [35] E. Kozik, E. Burovski, V. W. Scarola, and M. Troyer, Phys. Rev. B **87**, 2051102 (2013)
- [36] M.F. Collins: *Magnetic Critical Scattering* (Oxford University Press, 1989)

- [37] A.E. Antipov, E. Gull, and S. Kirchner, Phys. Rev. Lett. **112**, 226401 (2014)
- [38] T. Schäfer, F. Geles, D. Rost, G. Rohringer, E. Arrigoni, K. Held, N. Blümer, M. Aichhorn, and A. Toschi, arXiv:1405.7250
- [39] N. Blümer, PhD Thesis, University Augsburg (2003)
- [40] H. Park, K. Haule and G. Kotliar, Phys. Rev. Lett. **101**, 186403 (2008)
- [41] A. Toschi, G. Rohringer, A.A. Katanin, K. Held, Annalen der Physik **523**, 698 (2011); cf. [24]

11 Functional Renormalization Group Approach to Interacting Fermi Systems: DMFT as a Booster Rocket

Walter Metzner

Max Planck Institute for Solid State Research
Stuttgart

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | Functional RG for Fermi systems | 3 |
| 2.1 | Generating functionals | 4 |
| 2.2 | Exact flow equation | 6 |
| 2.3 | Expansion in the fields | 8 |
| 3 | Two-dimensional Hubbard model | 10 |
| 3.1 | Stability analysis at weak coupling | 12 |
| 3.2 | Spontaneous symmetry breaking | 17 |
| 4 | Leap to strong coupling: DMFT as a booster rocket | 18 |
| 4.1 | Dynamical mean-field theory | 20 |
| 4.2 | From infinite to finite dimensions | 20 |
| 4.3 | Application to the 2D Hubbard model | 22 |
| 5 | Conclusion | 24 |
| A | Derivation of functional flow equation | 25 |

1 Introduction

Correlated electron compounds exhibit very distinct behavior on different energy and length scales. Collective phenomena emerge at scales far below the bare energy scales of the microscopic Hamiltonian. For example, in high-temperature superconductors one bridges three orders of magnitude from the highest scale, the bare Coulomb interaction, via the intermediate scale of short-range magnetic correlations, down to the lowest scale of d -wave superconductivity and other ordering phenomena (see Fig. 1). This diversity of scales is a major obstacle to a direct numerical solution of microscopic models, since the most interesting phenomena emerge only in large systems at low temperatures. It is also hard to tackle by conventional many-body methods if one tries to treat all scales at once and within the same approximation, for example by summing a subclass of Feynman diagrams. Perturbative approaches which do not separate different scales are plagued by infrared divergences and are therefore often inapplicable even at weak coupling.

It is thus natural to treat degrees of freedom with different energy or length scales step by step. This is the main idea behind all *renormalization group* (RG) schemes. Using a functional integral representation this strategy can be implemented by integrating out degrees of freedom (bosonic or fermionic fields) successively, following a suitable order of scales. One thus generates a one-parameter family of effective actions that interpolates smoothly between the bare action of the system, as given by the microscopic Hamiltonian, and the final effective action from which all physical properties can be extracted. The Green or vertex functions corresponding to the effective action at scale Λ obey a hierarchy of differential flow equations. This hierarchy is exact and involves the flow of functions of generally continuous variables. For these reasons this approach is frequently referred to as “*exact*” or “*functional*” RG [1–3].

The exact hierarchy can be solved exactly only in special cases, where the underlying model can also be solved exactly (and more easily) by other means. However, the functional RG is a valuable source for devising powerful new approximation schemes, which can be obtained by truncating the hierarchy and/or by a simplified parametrization of the Green or vertex functions. These approximations have several distinctive advantages: i) they have a renormalization group structure built in; that is, scales are handled successively and infrared singularities are thus treated properly; ii) they can be applied directly to microscopic models, not only to effective field theories that capture only some asymptotic behavior; iii) they are physically transparent; for example, one can see directly how and why new correlations form upon lowering the scale; iv) one can use different approximations at different scales. Small steps from a scale Λ to a slightly smaller scale Λ' are much easier to control than an integration over all degrees of freedom in one shot.

This lecture provides a concise introduction to the functional RG in the context of interacting Fermi systems. To illustrate the method at work, I review applications of the functional RG to the two-dimensional Hubbard model. A more detailed presentation and many other applications can be found in the recent review article Ref. [3]. In the last part I present the very recent idea [4] of using the DMFT solution as a non-perturbative starting point for a functional RG flow.

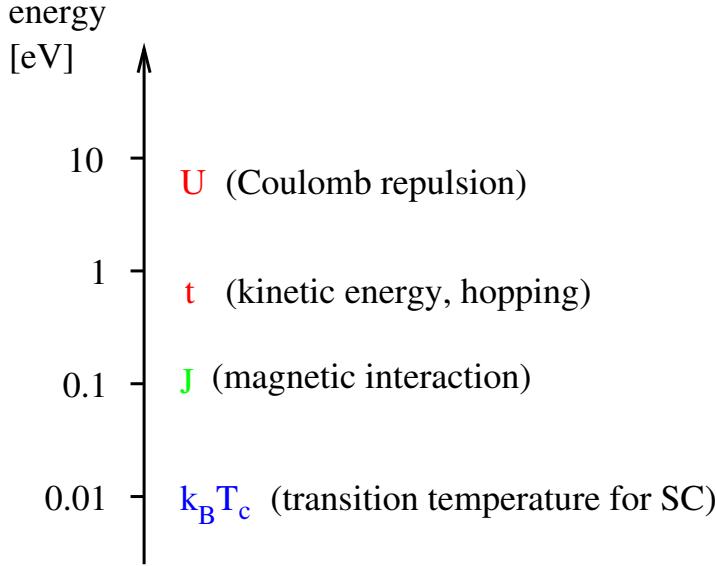


Fig. 1: Energy scales in cuprate high-temperature superconductors. Magnetic interactions and superconductivity are generated from the kinetic energy and the Coulomb repulsion. Figure taken from Ref. [3].

2 Functional RG for Fermi systems

Already in the 1970s, various RG methods were used to deal with infrared singularities arising in one-dimensional Fermi systems [5]. Renormalization group approaches dealing with interacting fermions in higher dimensions were developed much later. Due to the extended (not point-like) geometry of the Fermi surface singularity in dimensions $d > 1$, the renormalization group flow cannot be reduced to a small number of running couplings, even if irrelevant interactions are discarded. Aiming at a mathematical control of interacting Fermi systems, Feldman and Trubowitz [6], and independently Benfatto and Gallavotti [7], formulated a rigorous fermionic version of Wilson's momentum-shell RG [8] for interacting fermions in dimensions $d > 1$. Important rigorous results for two-dimensional systems have indeed been obtained [9, 10]. An essential message from these results is that no hitherto unknown instabilities or non-perturbative effects occur in Fermi systems with sufficiently weak interactions, at least in the absence of special features such as van Hove singularities at the Fermi level.

The Wilsonian RG for interacting Fermi systems was popularized among physicists by Shankar [11] and Polchinski [12], who presented some of the main ideas in a pedagogical style. In particular, they discussed an intuitive RG perspective of Fermi liquid theory.

The Wilsonian RG is not only useful for a rigorous understanding of interacting fermion systems. A specific variant of Wilson's RG known as *exact* or *functional* RG turned out to provide a valuable framework for computational purposes. Approximations derived from exact functional flow equations play an increasingly important role in the theory of interacting Fermi systems [3]. Exact flow equations describe the evolution of a generating functional for all many-particle correlation or vertex functions as a function of a flow parameter Λ , usually a cutoff. They can be easily derived from a functional integral representation.

For computational purposes, the exact flow equation for the effective action Γ^Λ , first derived in the context of bosonic field theories by Wetterich [13], turned out to be most convenient. The effective action is the generating functional for one-particle irreducible vertex functions, which are obtained by taking derivatives with respect to the source fields. The flow parameter Λ describes a regularization of the bare action, which regularizes infrared divergencies in perturbation theory. The regularization is removed at the end of the flow, say for $\Lambda \rightarrow 0$. The initial regulator (for $\Lambda = \Lambda_0$) can be chosen such that Γ^{Λ_0} is given by the bare action. The flow of Γ^Λ then provides a smooth interpolation between the bare action of the system and the final effective action Γ , from which any desired information can be extracted. This flow is determined by an exact functional differential equation [13]. Expanding in the fields one obtains a hierarchy of flow equations for the one-particle irreducible vertex functions.

The expression *functional RG* stems from the feature that the exact flow equations describe the flow of a functional or (equivalently) of a hierarchy of functions. An important difference compared to Wilson's original formulation is that a complete set of source fields is kept in the flowing generating functionals, not only those corresponding to scales below Λ . Hence, the full information on the properties of the system remains accessible, not only the low energy or long wavelength behavior.

In the remainder of this section, I present the exact flow equations and their expansion in the source fields. The first subsection summarizes the standard functional integral formalism as described, for example, in the excellent textbook by Negele and Orland [14].

2.1 Generating functionals

A system of interacting fermions can be represented by Grassmann fields $\psi, \bar{\psi}$, and an action of the form

$$\mathcal{S}[\psi, \bar{\psi}] = -(\bar{\psi}, G_0^{-1}\psi) + V[\psi, \bar{\psi}], \quad (1)$$

where $V[\psi, \bar{\psi}]$ is an arbitrary many-body interaction, and G_0 is the propagator of the non-interacting system. The bracket $(.,.)$ is a shorthand notation for the sum $\sum_K \bar{\psi}_K (G_0^{-1}\psi)_K$, where $(G_0^{-1}\psi)_K = \sum_{K'} G_0^{-1}(K, K') \psi_{K'}$. The Grassmann field index K collects the quantum numbers of a suitable single-particle basis set and the Matsubara frequency. The K -sums include integrals over continuous variables, and normalization factors such as temperature and volume. In particular, for spin-1/2 fermions with a single-particle basis labeled by momentum \mathbf{k} and spin orientation σ , one has $K = (k_0, \mathbf{k}, \sigma)$, where k_0 is the fermionic Matsubara frequency. If the bare part of the action is translation and spin-rotation invariant, the bare propagator has the diagonal and spin-independent form $G_0(K, K') = \delta_{KK'} G_0(K)$ with

$$G_0(K) = \frac{1}{ik_0 - \xi_{\mathbf{k}}}, \quad (2)$$

where $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$ is the single-particle energy relative to the chemical potential. A two-particle interaction has the general form

$$V[\psi, \bar{\psi}] = \frac{1}{4} \sum_{K_1, K_2, K'_1, K'_2} V(K'_1, K'_2; K_1, K_2) \bar{\psi}_{K'_1} \bar{\psi}_{K'_2} \psi_{K_2} \psi_{K_1}. \quad (3)$$

The generating functional

$$\mathcal{G}[\eta, \bar{\eta}] = -\log \int \prod_K d\psi_K d\bar{\psi}_K e^{-\mathcal{S}[\psi, \bar{\psi}]} e^{(\bar{\eta}, \psi) + (\bar{\psi}, \eta)}, \quad (4)$$

yields connected m -particle *Green functions* via derivatives with respect to the source fields, that is,

$$\begin{aligned} G^{(2m)}(K_1, \dots, K_m; K'_1, \dots, K'_m) &= -\langle \psi_{K_1} \dots \psi_{K_m} \bar{\psi}_{K'_m} \dots \bar{\psi}_{K'_1} \rangle_c \\ &= (-1)^m \frac{\partial^m}{\partial \bar{\eta}_{K_1} \dots \partial \bar{\eta}_{K_m}} \frac{\partial^m}{\partial \eta_{K'_m} \dots \partial \eta_{K'_1}} \mathcal{G}[\eta, \bar{\eta}] \Big|_{\eta=\bar{\eta}=0}, \end{aligned} \quad (5)$$

where $\langle \dots \rangle_c$ is the connected average of the product of Grassmann variables between the brackets. The one-particle Green function $G^{(2)}$ is the propagator of the interacting system, which is usually denoted without the superscript as G .

Legendre transforming $\mathcal{G}[\eta, \bar{\eta}]$ yields the *effective action*

$$\Gamma[\psi, \bar{\psi}] = (\bar{\eta}, \psi) + (\bar{\psi}, \eta) + \mathcal{G}[\eta, \bar{\eta}], \quad (6)$$

where $\psi = -\partial \mathcal{G} / \partial \bar{\eta}$ and $\bar{\psi} = \partial \mathcal{G} / \partial \eta$. The effective action is the generating functional for one-particle irreducible *vertex functions*,

$$\Gamma^{(2m)}(K'_1, \dots, K'_m; K_1, \dots, K_m) = \frac{\partial^{2m} \Gamma[\psi, \bar{\psi}]}{\partial \bar{\psi}_{K'_1} \dots \partial \bar{\psi}_{K'_m} \partial \psi_{K_m} \dots \partial \psi_{K_1}} \Big|_{\psi, \bar{\psi}=0}. \quad (7)$$

The Legendre correspondence between the functionals \mathcal{G} and Γ yields relations between the Green functions $G^{(2m)}$ and the vertex functions $\Gamma^{(2m)}$. In particular,

$$\Gamma^{(2)} = G^{-1} = G_0^{-1} - \Sigma, \quad (8)$$

where Σ is the self-energy. The two-particle Green function is related to the two-particle vertex by

$$\begin{aligned} G^{(4)}(K_1, K_2; K'_1, K'_2) &= \sum_{P_1, P_2, P'_1, P'_2} G(K_1, P'_1) G(K_2, P'_2) \\ &\quad \times \Gamma^{(4)}(P'_1, P'_2; P_1, P_2) G(P_1, K'_1) G(P_2, K'_2). \end{aligned} \quad (9)$$

Generally, the m -particle Green functions are obtained by adding all trees that can be formed with vertex functions of equal or lower order and G -lines [14].

The effective action obeys the reciprocity relations

$$\frac{\partial \Gamma}{\partial \psi} = -\bar{\eta}, \quad \frac{\partial \Gamma}{\partial \bar{\psi}} = \eta. \quad (10)$$

The second functional derivatives of \mathcal{G} and Γ with respect to the fields are also reciprocal. We define the matrices of second derivatives at *finite* fields

$$\mathbf{G}^{(2)}[\eta, \bar{\eta}] = \begin{pmatrix} -\frac{\partial^2 \mathcal{G}}{\partial \bar{\eta}_K \partial \eta_{K'}} & \frac{\partial^2 \mathcal{G}}{\partial \bar{\eta}_K \partial \bar{\eta}_{K'}} \\ \frac{\partial^2 \mathcal{G}}{\partial \eta_K \partial \eta_{K'}} & -\frac{\partial^2 \mathcal{G}}{\partial \eta_K \partial \bar{\eta}_{K'}} \end{pmatrix} = - \begin{pmatrix} \langle \psi_K \bar{\psi}_{K'} \rangle & \langle \psi_K \psi_{K'} \rangle \\ \langle \bar{\psi}_K \bar{\psi}_{K'} \rangle & \langle \bar{\psi}_K \psi_{K'} \rangle \end{pmatrix}, \quad (11)$$

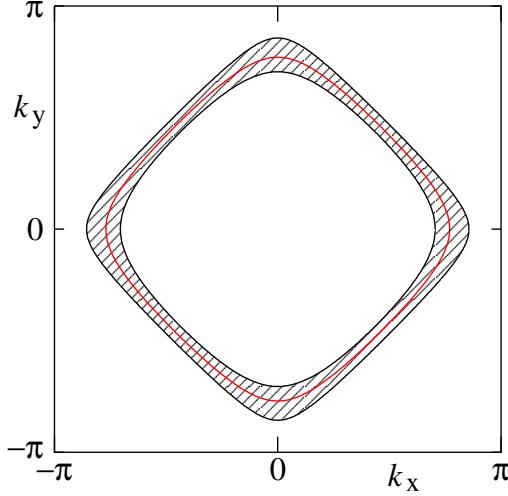


Fig. 2: Momentum space region around the Fermi surface excluded by a sharp momentum cutoff in a two-dimensional lattice fermion system. Figure taken from Ref. [3].

and

$$\boldsymbol{\Gamma}^{(2)}[\psi, \bar{\psi}] = \begin{pmatrix} \frac{\partial^2 \Gamma}{\partial \bar{\psi}_{K'} \partial \psi_K} & \frac{\partial^2 \Gamma}{\partial \bar{\psi}_{K'} \partial \bar{\psi}_K} \\ \frac{\partial^2 \Gamma}{\partial \psi_{K'} \partial \psi_K} & \frac{\partial^2 \Gamma}{\partial \psi_{K'} \partial \bar{\psi}_K} \end{pmatrix}, \quad (12)$$

where the matrix elements in the second matrix of the last equation are just a more convenient notation for those in the first matrix. The reciprocity relation for the second derivatives reads

$$\boldsymbol{\Gamma}^{(2)}[\psi, \bar{\psi}] = (\mathbf{G}^{(2)}[\eta, \bar{\eta}])^{-1}. \quad (13)$$

Note that *anomalous* components are involved as long as the source fields are finite. Only at $\eta = \bar{\eta} = 0$ and $\psi = \bar{\psi} = 0$, and in the absence of $U(1)$ charge symmetry breaking one has the simple relation $\boldsymbol{\Gamma}^{(2)} = \mathbf{G}^{-1}$.

2.2 Exact flow equation

We now endow the bare propagator G_0 with a dependence on a *flow parameter* Λ . Usually Λ is a cutoff suppressing contributions from fields with a single-particle energy or a Matsubara frequency below the scale Λ in the functional integral. For example, in a translation invariant system this may be done by modifying $G_0(K)$ to

$$G_0^\Lambda(K) = \frac{\Theta^\Lambda(\mathbf{k})}{ik_0 - \xi_\mathbf{k}}, \quad (14)$$

where $\Theta^\Lambda(\mathbf{k})$ is a function that vanishes for $|\xi_\mathbf{k}| \ll \Lambda$ and tends to one for $|\xi_\mathbf{k}| \gg \Lambda$. In this way the infrared singularity of the propagator at $k_0 = 0$ and $\xi_\mathbf{k} = 0$ is cut off at the scale Λ . The simplest choice for Θ^Λ is a step function, $\Theta^\Lambda(\mathbf{k}) = \Theta(|\xi_\mathbf{k}| - \Lambda)$, such that momenta in a shell around the Fermi surface are strictly excluded (see Fig. 2).

In the first applications of the functional RG to interacting Fermi systems a momentum cutoff was used, but later a frequency cutoff became more popular, since the latter does not interfere with Fermi surface shifts, and particle-hole excitations with a small momentum transfer are captured smoothly by the flow [15]. Moreover, a frequency cutoff can also be used in systems without translation invariance, such as systems with impurities [16].

The generating functionals constructed with G_0^A instead of G_0 depend on the flow parameter and will be denoted by $\mathcal{G}^A[\eta, \bar{\eta}]$ and $\Gamma^A[\psi, \bar{\psi}]$. The original functionals \mathcal{G} and Γ are recovered in the limit $A \rightarrow 0$. The evolution of $\Gamma^A[\psi, \bar{\psi}]$ as a function of A is described by an *exact functional flow equation*,

$$\frac{d}{dA} \Gamma^A[\psi, \bar{\psi}] = -(\bar{\psi}, \partial_A Q_0^A \psi) - \frac{1}{2} \text{Tr} \left[(\partial_A \mathbf{Q}_0^A) (\Gamma^{(2)A}[\psi, \bar{\psi}])^{-1} \right], \quad (15)$$

where

$$\mathbf{Q}_0^A = \begin{pmatrix} Q_0^A(K, K') & 0 \\ 0 & -Q_0^A(K', K) \end{pmatrix} \quad (16)$$

with $Q_0^A = (G_0^A)^{-1}$, and $\Gamma^{(2)A}[\psi, \bar{\psi}]$ is the matrix of second derivatives of Γ^A at finite fields,

$$\Gamma^{(2)A}[\psi, \bar{\psi}] = \begin{pmatrix} \frac{\partial^2 \Gamma^A}{\partial \bar{\psi}_K \partial \psi_{K'}} & \frac{\partial^2 \Gamma^A}{\partial \bar{\psi}_K \partial \bar{\psi}_{K'}} \\ \frac{\partial^2 \Gamma^A}{\partial \psi_K \partial \psi_{K'}} & \frac{\partial^2 \Gamma^A}{\partial \psi_K \partial \bar{\psi}_{K'}} \end{pmatrix}. \quad (17)$$

The trace on the right hand side of the flow equations includes a sum over the Grassmann field index K . Note that the inversion of $\Gamma^{(2)A}[\psi, \bar{\psi}]$ in Eq. (15) is not merely an inversion of a 2×2 matrix, since it involves also the additional matrix structure coming from the dependence on the Grassmann field indices K and K' . A derivation of the functional flow equation (15) is presented in Appendix A.

Alternative definitions of the effective action Γ^A , differing by interaction-independent terms, have also been used. A frequently used variant is [1]

$$\Gamma_R^A[\psi, \bar{\psi}] = \Gamma^A[\psi, \bar{\psi}] + (\bar{\psi}, R^A \psi), \quad (18)$$

where $R^A = Q_0^A - Q_0$. The additional quadratic term cancels the first term in the flow equation (15) for Γ^A , and one obtains the equivalent flow equation

$$\frac{d}{dA} \Gamma_R^A[\psi, \bar{\psi}] = -\frac{1}{2} \text{Tr} \left[(\partial_A \mathbf{R}^A) (\Gamma_R^{(2)A}[\psi, \bar{\psi}] + \mathbf{R}^A)^{-1} \right], \quad (19)$$

where $\mathbf{R}^A = \text{diag}(R^A(K, K'), -R^A(K', K))$. The functional Γ_R^A and its analogue for bosonic fields is known as *effective average action* in the literature [1]. Both Γ_R^A and Γ^A tend to the same effective action Γ in the limit $A \rightarrow 0$, where R^A vanishes.

Choosing the initial cutoff A_0 such that $G_0^{A_0}$ is identically zero, all contributions to the functional integral are suppressed. For a sharp momentum cutoff this is achieved by choosing A_0 larger than $\max_{\mathbf{k}} |\xi_{\mathbf{k}}|$, while for a frequency cutoff one has to start with $A_0 = \infty$ to eliminate all

modes. At the initial scale Λ_0 , one then has the simple initial condition $\Gamma_R^{\Lambda_0}[\psi, \bar{\psi}] = \mathcal{S}[\psi, \bar{\psi}]$, while

$$\Gamma^{\Lambda_0}[\psi, \bar{\psi}] = -(\bar{\psi}, Q_0^{\Lambda_0}\psi) + V[\psi, \bar{\psi}] = \mathcal{S}^{\Lambda_0}[\psi, \bar{\psi}] = \mathcal{S}[\psi, \bar{\psi}] - (\bar{\psi}, R^A\psi). \quad (20)$$

Hence, Γ_R^A interpolates smoothly between the (unregularized) bare action \mathcal{S} and the final effective action Γ , while Γ^A interpolates between the *regularized* bare action \mathcal{S}^{Λ_0} and Γ . The initial “regularization” with $Q_0^{\Lambda_0} = \infty$ amounts to a complete suppression of all contributions to the functional integral, not just a regularization of divergent contributions.

2.3 Expansion in the fields

Expanding the functional flow equation (15) for the effective action in powers of the fields yields a hierarchy of flow equations for the m -particle vertex functions. To expand the inverse of $\Gamma^{(2)\Lambda}[\psi, \bar{\psi}]$, we split

$$\Gamma^{(2)\Lambda}[\psi, \bar{\psi}] = (\mathbf{G}^A)^{-1} - \tilde{\Sigma}^A[\psi, \bar{\psi}], \quad (21)$$

where

$$\mathbf{G}^A = \left(\Gamma^{(2)\Lambda}[\psi, \bar{\psi}] \Big|_{\psi=\bar{\psi}=0} \right)^{-1} = \begin{pmatrix} G^A(K, K') & 0 \\ 0 & -G^A(K', K) \end{pmatrix}, \quad (22)$$

and $\tilde{\Sigma}^A[\psi, \bar{\psi}]$ contains all contributions which are at least quadratic in the fields. Now the inverse of $\Gamma^{(2)\Lambda}[\psi, \bar{\psi}]$ can be expanded in a geometric series,

$$\begin{aligned} (\Gamma^{(2)\Lambda}[\psi, \bar{\psi}])^{-1} &= (1 - \mathbf{G}^A \tilde{\Sigma}^A[\psi, \bar{\psi}])^{-1} \mathbf{G}^A \\ &= [1 + \mathbf{G}^A \tilde{\Sigma}^A[\psi, \bar{\psi}] + (\mathbf{G}^A \tilde{\Sigma}^A[\psi, \bar{\psi}])^2 + \dots] \mathbf{G}^A. \end{aligned} \quad (23)$$

Inserting this into the functional flow equation yields

$$\begin{aligned} \frac{d}{d\Lambda} \Gamma^A[\psi, \bar{\psi}] &= -\text{Tr}[(\partial_A Q_0^A) G^A] - (\bar{\psi}, \partial_A Q_0^A \psi) \\ &\quad + \frac{1}{2} \text{Tr}[\mathbf{S}^A (\tilde{\Sigma}^A[\psi, \bar{\psi}] + \tilde{\Sigma}^A[\psi, \bar{\psi}] \mathbf{G}^A \tilde{\Sigma}^A[\psi, \bar{\psi}] + \dots)], \end{aligned} \quad (24)$$

with the *single scale propagator*

$$\mathbf{S}^A = -\mathbf{G}^A (\partial_A Q_0^A) \mathbf{G}^A = \frac{d}{d\Lambda} \mathbf{G}^A \Big|_{\Sigma^A \text{ fixed}}. \quad (25)$$

The latter usually has its main support *at* the scale Λ . For example, for a sharp momentum cutoff acting on ξ_k as in Eq. (14), \mathbf{S}^A has a delta-peak at $|\xi_k| = \Lambda$ and vanishes elsewhere.

Expanding $\Gamma^A[\psi, \bar{\psi}]$ and $\tilde{\Sigma}^A[\psi, \bar{\psi}]$ in powers of ψ and $\bar{\psi}$, and comparing coefficients in Eq. (24), one obtains the flow equations for the self-energy $\Sigma^A = Q_0^A - \Gamma^{(2)\Lambda}$, the two-particle vertex $\Gamma^{(4)\Lambda}$, and all other m -particle vertices. The first three equations of this hierarchy are represented diagrammatically in Fig. 3. Note that only one-particle irreducible one-loop diagrams contribute, and internal lines are dressed by the self-energy. Contributions with several loops

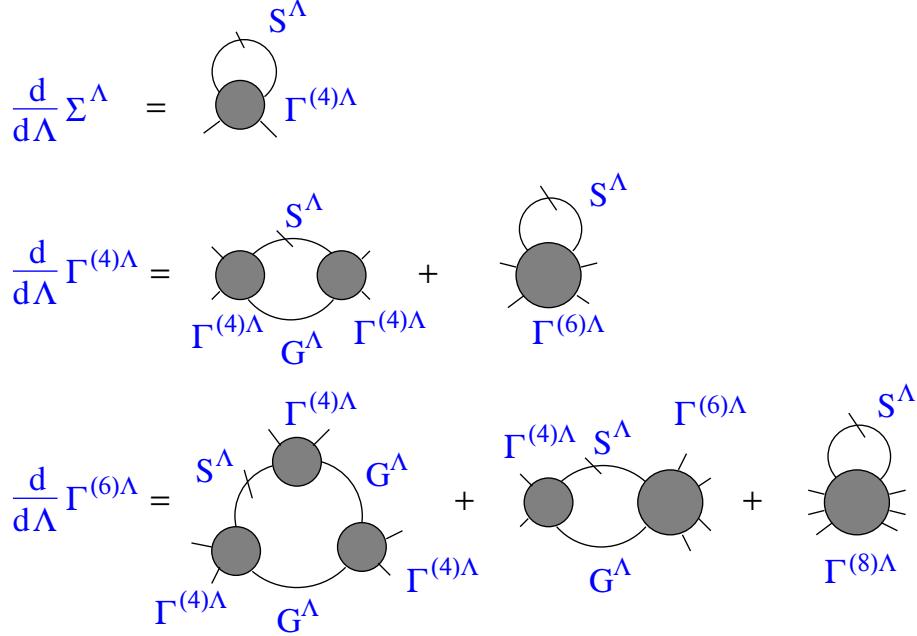


Fig. 3: Diagrammatic representation of the flow equations for the self-energy, the two-particle vertex, and the three-particle vertex. Lines with a dash correspond to the single-scale propagator S^Λ , and the other lines to the full propagator G^Λ . Figure taken from Ref. [3].

are generated only indirectly upon inserting the flow of higher order vertices into lower order flow equations. The flow equation for the self-energy has the simple form

$$\frac{d}{d\Lambda} \Sigma^\Lambda(K', K) = \sum_{P, P'} S^\Lambda(P, P') \Gamma^{(4)\Lambda}(K', P'; K, P) . \quad (26)$$

The flow equation for the two-particle vertex reads

$$\begin{aligned} \frac{d}{d\Lambda} \Gamma^{(4)\Lambda}(K'_1, K'_2; K_1, K_2) &= - \sum_{P_1, P'_1} \sum_{P_2, P'_2} G^\Lambda(P_1, P'_1) S^\Lambda(P_2, P'_2) \\ &\times \left\{ \Gamma^{(4)\Lambda}(K'_1, K'_2; P_1, P_2) \Gamma^{(4)\Lambda}(P'_1, P'_2; K_1, K_2) \right. \\ &- \left[\Gamma^{(4)\Lambda}(K'_1, P'_2; K_1, P_1) \Gamma^{(4)\Lambda}(P'_1, K'_2; P_2, K_2) + (P_1 \leftrightarrow P_2, P'_1 \leftrightarrow P'_2) \right] \\ &+ \left. \left[\Gamma^{(4)\Lambda}(K'_2, P'_2; K_1, P_1) \Gamma^{(4)\Lambda}(P'_1, K'_1; P_2, K_2) + (P_1 \leftrightarrow P_2, P'_1 \leftrightarrow P'_2) \right] \right\} \\ &- \sum_{P, P'} S^\Lambda(P, P') \Gamma^{(6)\Lambda}(K'_1, K'_2, P'; K_1, K_2, P) . \end{aligned} \quad (27)$$

There are several contributions that are quadratic in $\Gamma^{(4)\Lambda}$, corresponding to the particle-particle, direct particle-hole, and crossed particle-hole channel, respectively.

The one-particle irreducibility is a convenient feature of flow equations derived from the effective action $\Gamma^\Lambda[\psi, \bar{\psi}]$. Flow equations derived from other functionals, such as the Polchinski equations [17] and their Wick ordered variant [9], contain also one-particle reducible contributions.

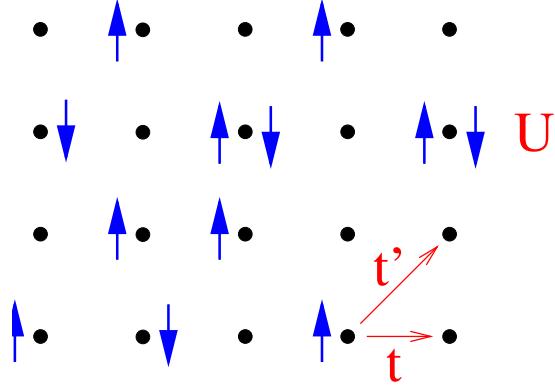


Fig. 4: Hubbard model with nearest and next-to-nearest neighbor hopping on a square lattice.

The full hierarchy of flow equations does not close at any finite order, since the flow of each $\Gamma^{(2m)\Lambda}$ receives a contribution from a tadpole diagram with $\Gamma^{(2m+2)\Lambda}$. Exact solutions of the flow equation hierarchy are possible only for relatively simple models, such as the Luttinger model [18], the reduced BCS model [19], or other mean-field models [20, 21], which can also be solved more directly by other methods. Usually the hierarchy of flow equations has to be truncated by neglecting effective interactions of higher order, and by using a simplified parametrization of the functional dependence of the remaining interactions on momenta, frequencies, et cetera. Truncations of the hierarchy at some finite order can be justified in case of sufficiently weak interactions, or if higher order terms are suppressed due to small phase space volumina [22]. Geometrical phase space restrictions are typically stronger in multi-loop integrals. A simplified parametrization of effective interactions can be obtained by neglecting dependences which become irrelevant in the low-energy limit.

3 Two-dimensional Hubbard model

Shortly after the discovery of high-temperature superconductivity in several cuprate compounds, Anderson [23] pointed out that the essential physics of the electrons in the copper-oxide planes of these materials could be described by the two-dimensional Hubbard model. The model describes tight-binding electrons with inter-site hopping amplitudes t_{ij} and a local repulsion $U > 0$, as specified by the Hamiltonian (see also Fig. 4)

$$H = \sum_{i,j} \sum_{\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_j n_{j\uparrow} n_{j\downarrow}. \quad (28)$$

Here $c_{i\sigma}^\dagger$ and $c_{i\sigma}$ are creation and annihilation operators for spin- $1/2$ fermions with spin orientation σ on a lattice site i , and $n_{j\sigma} = c_{j\sigma}^\dagger c_{j\sigma}$. The number of lattice sites will be denoted by L . A hopping amplitude $-t$ between nearest neighbors and an amplitude $-t'$ between next-nearest neighbors on a square lattice leads to the dispersion relation

$$\epsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y) - 4t'(\cos k_x \cos k_y) \quad (29)$$

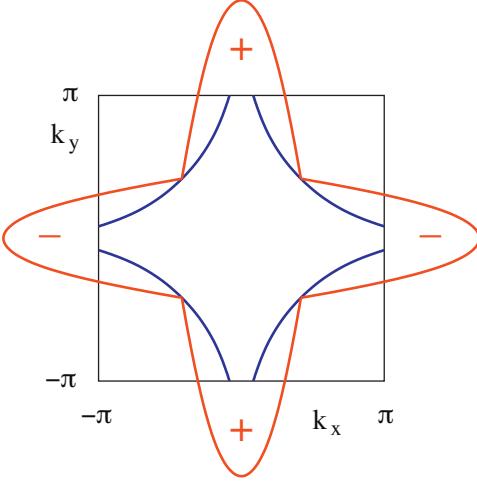


Fig. 5: Schematic shape of the gap function $\Delta_{\mathbf{k}}$ with $d_{x^2-y^2}$ -wave symmetry, for \mathbf{k} tracing the Fermi surface.

for single-particle states. This dispersion relation has saddle points at $\mathbf{k} = (0, \pi)$ and $(\pi, 0)$, which entail logarithmic van Hove singularities in the non-interacting density of states at the energy $\epsilon_{vH} = 4t'$.

In agreement with the generic phase diagram of the cuprates, the Hubbard model exhibits anti-ferromagnetic order at half-filling, and has been expected to become a d -wave superconductor away from half-filling in two dimensions for quite some time [24]. The exchange of antiferromagnetic spin fluctuations has been proposed as a plausible mechanism leading to d -wave pairing [25–27]. In this picture, the BCS effective interaction $V_{\mathbf{kk}'}$ is roughly proportional to the spin correlation function $\chi_s(\mathbf{k} - \mathbf{k}')$. Close to half-filling, $\chi_s(\mathbf{q})$ has a pronounced maximum at or near the antiferromagnetic wave vector (π, π) . As a consequence, the BCS gap equation

$$\Delta_{\mathbf{k}} = - \int \frac{d^2 \mathbf{k}'}{(2\pi)^2} V_{\mathbf{kk}'} \frac{\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}} \quad (30)$$

with $E_{\mathbf{k}} = (\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2)^{1/2}$ has a solution with $d_{x^2-y^2}$ -wave symmetry such that the gap $\Delta_{\mathbf{k}}$ has maximal modulus but opposite sign near the points $(\pi, 0)$ and $(0, \pi)$ in the Brillouin zone [24], as illustrated in Fig. 5. This intuitive argument has been corroborated by a self-consistent perturbative solution of the two-dimensional Hubbard model within the so-called fluctuation-exchange approximation [28, 29].

The spin-fluctuation mechanism for pairing might be spoiled by other contributions to the BCS interactions and also by spin density wave instabilities. It turned out to be very hard to detect superconductivity in the Hubbard model by exact numerical computation [24, 30] as a consequence of finite size and/or temperature limitations. Fortunately, the tendency toward antiferromagnetism and d -wave pairing is already captured by the two-dimensional Hubbard model at *weak* coupling. Conventional perturbation theory breaks down for densities close to half-filling, since competing infrared divergences appear as a consequence of Fermi surface nesting and van Hove singularities. A controlled and unbiased treatment of these divergences is achieved by a renormalization group analysis, which takes into account the particle-particle and particle-hole channels on equal footing.

Early RG studies of the two-dimensional Hubbard model started with simple scaling approaches, very shortly after the discovery of high- T_c superconductivity [31–33]. These studies focused on dominant scattering processes between the van Hove points in \mathbf{k} -space, for which a small number of running couplings could be defined and computed on a one-loop level. Spin-density and superconducting instabilities were identified from divergences of the corresponding correlation functions.

A complete treatment of all scattering processes in the Brillouin zone is more complicated since the effective interactions cannot be parametrized accurately by a small number of variables, even if irrelevant momentum and energy dependences are neglected. The *tangential* momentum dependence of effective interactions along the Fermi surface is strong and important in the low-energy limit. Hence, one has to deal with the renormalization of *functions*. This problem is treated most naturally by the functional RG.

3.1 Stability analysis at weak coupling

To detect instabilities in the weak-coupling limit, one can truncate the infinite hierarchy of flow equations at second order in the effective two-particle interaction and discard all vertices of higher order, and also self-energy corrections. The flow of the two-particle vertex is then fully determined by the first contribution in the second line of Fig. 3. In the absence of self-energy corrections, the internal lines are given by the bare propagator G_0^A and its scale derivative $S_0^A = \frac{d}{d\Lambda} G_0^A$.

Due to translation invariance on the lattice, G_0^A is diagonal in the momentum representation, that is, $G_0^A(K, K') = \delta_{KK'} G_0^A(K)$, where $K = (k, \sigma)$ with $k = (k_0, \mathbf{k})$. Hence, the truncated (without $\Gamma^{(6)A}$) flow equation (27) for the two-particle vertex can be written as

$$\begin{aligned} \frac{d}{d\Lambda} \Gamma^A(K'_1, K'_2; K_1, K_2) &= -\frac{1}{\beta L} \sum_{P, P'} \frac{d}{d\Lambda} [G_0^A(P) G_0^A(P')] \\ &\times \left[\frac{1}{2} \Gamma^A(K'_1, K'_2; P, P') \Gamma^A(P, P'; K_1, K_2) \right. \\ &\quad - \Gamma^A(K'_1, P'; K_1, P) \Gamma^A(P, K'_2; P', K_2) \\ &\quad \left. + \Gamma^A(K'_2, P'; K_1, P) \Gamma^A(P, K'_1; P', K_2) \right]. \end{aligned} \quad (31)$$

Here and in the following, we denote the two-particle vertex as Γ^A , and we now write temperature ($\beta = 1/T$) and volume factors explicitly. The three terms on the right-hand side correspond to the particle-particle (PP), the direct particle-hole (PH) and the crossed particle-hole (PH') channel, respectively, see Fig. 6.

Due to spin rotation invariance, the spin structure of the two-particle vertex is relatively simple. One can express the vertex by a single function V^A depending only on momenta and Matsubara frequencies [22],

$$\Gamma^A(K'_1, K'_2; K_1, K_2) = V^A(k'_1, k'_2; k_1, k_2) \delta_{\sigma_1 \sigma_1'} \delta_{\sigma_2 \sigma_2'} - V^A(k'_2, k'_1; k_1, k_2) \delta_{\sigma_1 \sigma_2'} \delta_{\sigma_2 \sigma_1'}. \quad (32)$$

Alternatively, one may decompose the vertex in singlet and triplet components [34],

$$\Gamma^A(K'_1, K'_2; K_1, K_2) = \Gamma_s^A(k'_1, k'_2; k_1, k_2) S_{\sigma'_1, \sigma'_2; \sigma_1, \sigma_2} + \Gamma_t^A(k'_1, k'_2; k_1, k_2) T_{\sigma'_1, \sigma'_2; \sigma_1, \sigma_2}, \quad (33)$$

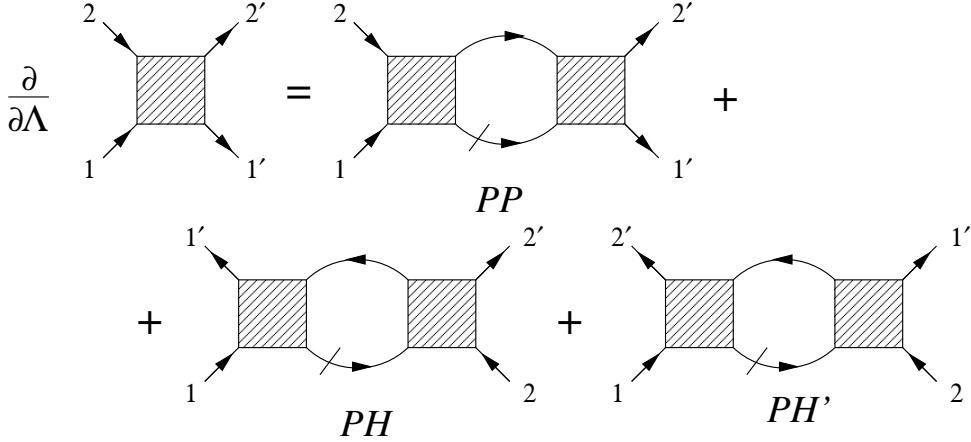


Fig. 6: One-loop flow equation for the two-particle vertex Γ^A with the particle-particle channel (PP) and the two particle-hole channels (PH and PH') written explicitly.

where

$$\begin{aligned} S_{\sigma'_1, \sigma'_2; \sigma_1, \sigma_2} &= \frac{1}{2} (\delta_{\sigma_1 \sigma'_1} \delta_{\sigma_2 \sigma'_2} - \delta_{\sigma_1 \sigma'_2} \delta_{\sigma_2 \sigma'_1}), \\ T_{\sigma'_1, \sigma'_2; \sigma_1, \sigma_2} &= \frac{1}{2} (\delta_{\sigma_1 \sigma'_1} \delta_{\sigma_2 \sigma'_2} + \delta_{\sigma_1 \sigma'_2} \delta_{\sigma_2 \sigma'_1}). \end{aligned} \quad (34)$$

Of course, one can easily compute Γ_s^A and Γ_t^A from V^A and vice versa [21]. The singlet vertex is symmetric, and the triplet vertex is antisymmetric under exchanges $k_1 \leftrightarrow k_2$ or $k'_1 \leftrightarrow k'_2$. We now outline the steps made to cast the one-loop flow equations in a form amenable to a numerical solution. To be specific, we use the singlet-triplet decomposition of the vertex and a momentum cutoff. Carrying out the spin sum in the flow equation one obtains

$$\begin{aligned} \partial_A \Gamma_\alpha^A(k'_1, k'_2; k_1, k_2) = - \sum_{i=s,t} \sum_{j=s,t} & [C_{\alpha ij}^{\text{PP}} \beta_{ij}^{\text{PP}}(k'_1, k'_2; k_1, k_2) \\ & + C_{\alpha ij}^{\text{PH}} \beta_{ij}^{\text{PH}}(k'_1, k'_2; k_1, k_2) + C_{\alpha ij}^{\text{PH'}} \beta_{ij}^{\text{PH'}}(k'_1, k'_2; k_1, k_2)] \end{aligned} \quad (35)$$

for $\alpha = s, t$, where $C_{\alpha ij}^{\text{PP}}$ etc. are simple numerical coefficients [34] and

$$\begin{aligned} \beta_{ij}^{\text{PP}}(k'_1, k'_2; k_1, k_2) &= \frac{1}{2\beta L} \sum_{k,k'} \partial_A [G_0^A(k) G_0^A(k')] \Gamma_i^A(k'_1, k'_2; k, k') \Gamma_j^A(k, k'; k_1, k_2), \\ \beta_{ij}^{\text{PH}}(k'_1, k'_2; k_1, k_2) &= -\frac{1}{\beta L} \sum_{k,k'} \partial_A [G_0^A(k) G_0^A(k')] \Gamma_i^A(k'_1, k; k_1, k') \Gamma_j^A(k', k'_2; k, k_2), \\ \beta_{ij}^{\text{PH'}}(k'_1, k'_2; k_1, k_2) &= -\beta_{ij}^{\text{PH}}(k'_2, k'_1; k_1, k_2). \end{aligned} \quad (36)$$

It is clearly impossible to solve the flow equations with the full energy and momentum dependence of the vertex function, since Γ^A has three independent energy and momentum variables. However, the flow equations can be simplified considerably by ignoring dependences which are irrelevant (in the RG sense) in the low-energy limit, that is, the energy dependence and the momentum dependence normal to the Fermi surface [3, 11].

Neglecting the energy dependence, we approximate

$$\Gamma_\alpha^A(k'_1, k'_2; k_1, k_2) \approx \Gamma_\alpha^A(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) . \quad (37)$$

Choosing an energy independent cutoff function $\Theta^A(\mathbf{k})$ as in Eq. (14), the Matsubara sums on the right hand side of the flow equations can be performed analytically, yielding

$$\begin{aligned} \partial_A \Gamma_\alpha^A(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) = & - \sum_{i=s,t} \sum_{j=s,t} [C_{\alpha ij}^{\text{PP}} \beta_{ij}^{\text{PP}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) \\ & + C_{\alpha ij}^{\text{PH}} \beta_{ij}^{\text{PH}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) + C_{\alpha ij}^{\text{PH'}} \beta_{ij}^{\text{PH'}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2)] \end{aligned} \quad (38)$$

for $\alpha = s, t$, where the β -functions are now energy independent and read

$$\begin{aligned} \beta_{ij}^{\text{PP}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) &= \frac{1}{2L} \sum_{\mathbf{k}, \mathbf{k}'} \partial_A [\Theta^A(\mathbf{k}) \Theta^A(\mathbf{k}')] \\ &\times \frac{f(-\xi_{\mathbf{k}}) - f(\xi_{\mathbf{k}'})}{\xi_{\mathbf{k}} + \xi_{\mathbf{k}'}} \Gamma_i^A(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}, \mathbf{k}') \Gamma_j^A(\mathbf{k}, \mathbf{k}'; \mathbf{k}_1, \mathbf{k}_2) , \\ \beta_{ij}^{\text{PH}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) &= - \frac{1}{L} \sum_{\mathbf{k}, \mathbf{k}'} \partial_A [\Theta^A(\mathbf{k}) \Theta^A(\mathbf{k}')] \\ &\times \frac{f(\xi_{\mathbf{k}}) - f(\xi_{\mathbf{k}'})}{\xi_{\mathbf{k}} - \xi_{\mathbf{k}'}} \Gamma_i^A(\mathbf{k}'_1, \mathbf{k}; \mathbf{k}_1, \mathbf{k}') \Gamma_j^A(\mathbf{k}', \mathbf{k}'_2; \mathbf{k}, \mathbf{k}_2) , \\ \beta_{ij}^{\text{PH'}}(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) &= - \beta_{ij}^{\text{PH}}(\mathbf{k}'_2, \mathbf{k}'_1; \mathbf{k}_1, \mathbf{k}_2) , \end{aligned} \quad (39)$$

with the Fermi function $f(\xi) = [e^{\beta\xi} + 1]^{-1}$. Note that momentum conservation implies that \mathbf{k} and \mathbf{k}' are related by $\mathbf{k} + \mathbf{k}' = \mathbf{k}_1 + \mathbf{k}_2$ in the particle-particle channel and by $\mathbf{k} + \mathbf{k}' = \mathbf{k}' + \mathbf{k}_1$ in the direct particle-hole channel. Hence, only one independent momentum variable needs to be summed. For a sharp momentum cutoff $\Theta^A(\mathbf{k}) = \Theta(|\xi_{\mathbf{k}}| - A)$ one has $\partial_A \Theta^A(\mathbf{k}) = -\delta(|\xi_{\mathbf{k}}| - A)$, so that the two-dimensional momentum integral can be reduced to a one-dimensional integral. The flow equation can be solved only if the momentum dependence of the vertex function is simplified. At least for weak coupling (in practice also for moderate ones), the vertex function acquires strong momentum dependences only for momenta close to the Fermi surface. Note that for the Hubbard model the bare vertex function Γ^{A_0} does not depend on momentum at all. Weak coupling instabilities are signalled by divergencies of the vertex function Γ^A , which are driven by momenta close to the Fermi surface. Hence, we will focus on the flow of the vertex function with momenta close to the Fermi surface. For arbitrary momenta, we approximate the vertex by

$$\Gamma_\alpha^A(\mathbf{k}'_1, \mathbf{k}'_2; \mathbf{k}_1, \mathbf{k}_2) \approx \Gamma_\alpha^A(\mathbf{k}'_{F1}, \mathbf{k}_{F1} + \mathbf{k}_{F2} - \mathbf{k}'_{F1}; \mathbf{k}_{F1}, \mathbf{k}_{F2}) \quad (40)$$

where \mathbf{k}_{F1} etc. are projections of \mathbf{k}_1 etc. on the Fermi surface (see Fig. 7). Strong momentum dependences of the effective vertex are built up only by contributions with intermediate momenta \mathbf{k} and \mathbf{k}' (on the right hand side of the flow equations) which are close to the Fermi surface, because for such momenta the ratios $\frac{f(\pm\xi_{\mathbf{k}}) - f(\xi_{\mathbf{k}'})}{\xi_{\mathbf{k}} \pm \xi_{\mathbf{k}'}}$ in Eq. (39) can be big. Hence, for the most important momenta, the error made by the projection is relatively small (even if A

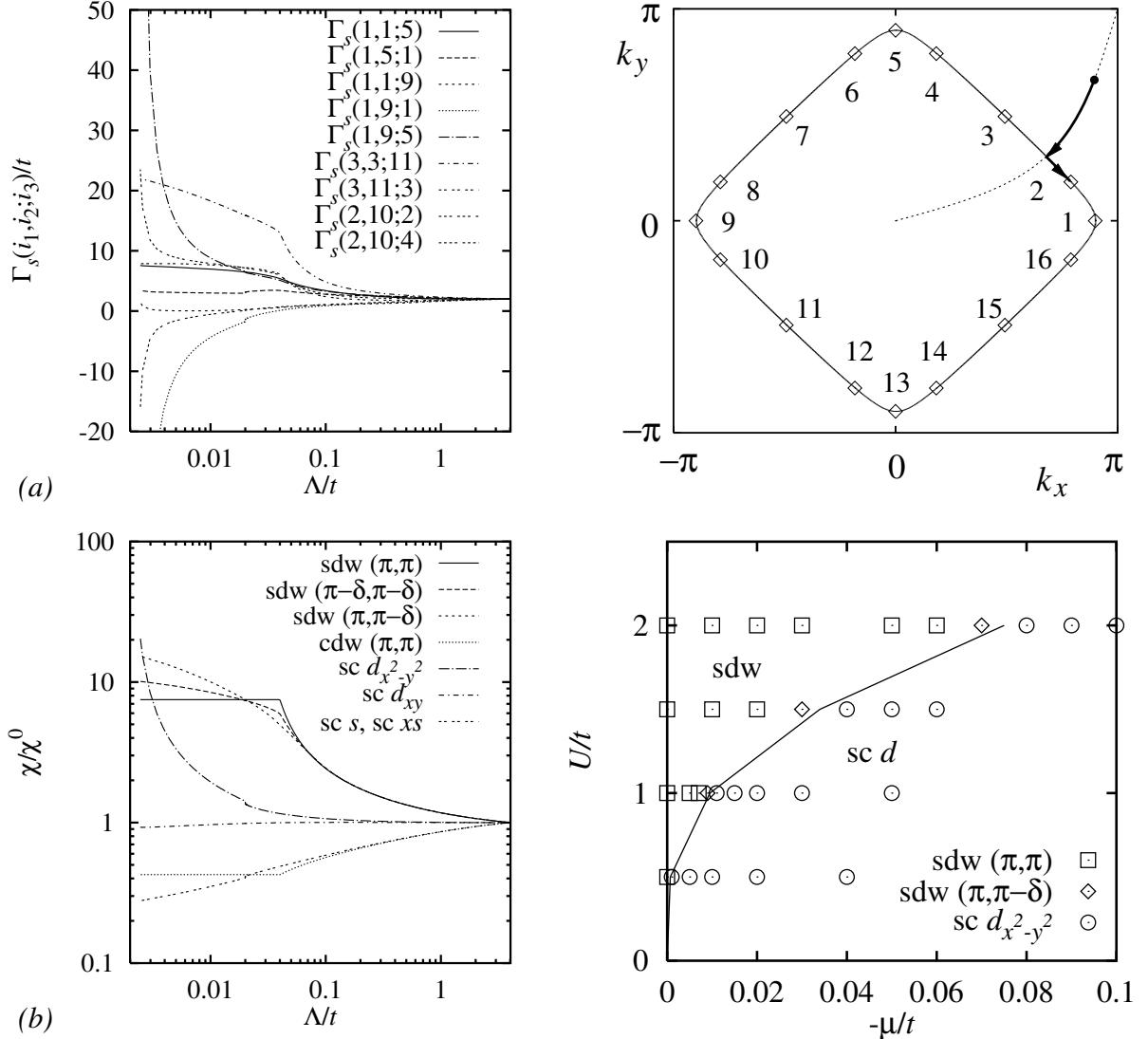


Fig. 7: (a) Flow of the singlet component of the two-particle vertex Γ_s^A as a function of Λ for several choices of the Fermi momenta \mathbf{k}_{F1} , \mathbf{k}_{F2} and \mathbf{k}'_{F1} , which are labeled according to the projection in the figure on the right. The model parameters are $U = t$ and $t' = 0$, the chemical potential is $\mu = -0.02t$, and the temperature is zero. (b) Left: Flow of the ratio of interacting and non-interacting susceptibilities, χ^A/χ_0^A , for the same parameters as in (a). Right: Ground state phase diagram for $t' = 0$ and $\mu \leq 0$ (at and below half-filling), as obtained from divergent susceptibilities. Figures taken from Ref. [34].

is not small), because these momenta are close to their projected counterparts. The projected vertex function can be parameterized by three angles ϕ_1, ϕ_2, ϕ_3 associated with \mathbf{k}_{F1} , \mathbf{k}_{F2} and \mathbf{k}'_{F1} , respectively, i.e.

$$\Gamma_\alpha^A(\mathbf{k}'_{F1}, \mathbf{k}_{F1} + \mathbf{k}_{F2} - \mathbf{k}'_{F1}; \mathbf{k}_{F1}, \mathbf{k}_{F2}) = \Gamma_\alpha^A(\phi_1, \phi_2, \phi_3) \quad (41)$$

The angular dependence turns out to be strong for small Λ and cannot be neglected. The remaining tangential momentum dependence is discretized (see again Fig. 7). Equivalently, one can view this parametrization as a discretization of momentum dependences corresponding to a partition of the Brillouin zone in "patches" or "sectors" [35, 36].

The flow of the two-particle vertex in the ground state of the two-dimensional Hubbard model has been computed for many different model parameters t'/t and U/t (t just fixes the absolute energy scale) and densities near half-filling [35, 34, 36, 37]. In all cases the vertex develops a strong momentum dependence for small Λ with divergences for several momenta at some critical scale $\Lambda_c > 0$, which vanishes exponentially for $U \rightarrow 0$. To see which physical instability is associated with the diverging vertex function, several susceptibilities have been computed, in particular commensurate and incommensurate spin susceptibilities $\chi_s(\mathbf{q})$ with $\mathbf{q} = (\pi, \pi)$, $\mathbf{q} = (\pi - \delta, \pi)$ and $\mathbf{q} = (1 - \delta)(\pi, \pi)$, where δ is a function of density [38]; the commensurate charge susceptibility $\chi_C((\pi, \pi))$; and singlet pair susceptibilities with s -wave and d -wave form factors. Some of these susceptibilities diverge together with the vertex function at the scale Λ_c . Depending on the choice of U , t' and μ , the strongest divergence is found either for the commensurate or incommensurate spin susceptibility, or for the pairing susceptibility with $d_{x^2-y^2}$ symmetry.

Fig. 7 shows a typical result for the flow of the two-particle interactions and susceptibilities in the ground state of the two-dimensional Hubbard model, as obtained from the Wick ordered version of the functional RG [34]. Within the lowest order truncation for the two-particle vertex, the results obtained from different functional RG versions do not deviate significantly from each other. Only the singlet part of the vertex is plotted, for various choices of two incoming momenta and one outgoing momentum on the Fermi surface. The triplet part of the vertex flows generally more weakly than the singlet part. Note the threshold at $\Lambda = 2|\mu|$ below which the amplitudes for various scattering processes, especially umklapp scattering, renormalize only very slowly. The flow of the antiferromagnetic spin susceptibility is cut off at the same scale. The infinite slope singularity in some of the flows at scale $\Lambda = |\mu|$ is due to the van Hove singularity being crossed at that scale. The *pairing* susceptibility with $d_{x^2-y^2}$ -symmetry diverges at the scale Λ_c , at which also the two-particle interaction diverges in the Cooper channel.

Following the flow of the vertex function and susceptibilities, one can see that those interaction processes which enhance the antiferromagnetic spin susceptibility (especially umklapp scattering) also build up an attractive interaction in the $d_{x^2-y^2}$ pairing channel. This confirms the spin-fluctuation route to d -wave superconductivity. Running the flow for various choices of μ/t and U/t one obtains an educated guess for the ground state phase diagram from the dominant divergences of the vertex and susceptibilities.

The static approximation of the vertex with a discretized momentum dependence as described above is sufficient for a weak coupling stability analysis. However, near the scale Λ_c , where the vertex diverges, this crude parametrization does not capture the momentum and energy dependence of the emerging singularities, and the power-counting argument invoked for its justification breaks down. Recently, an improved parametrization of the two-particle vertex based on an additive decomposition in charge, magnetic, and pairing channels has been established [15, 39, 40]. In this channel decomposition, singular momentum and energy dependences are isolated in a single bosonic momentum and energy variable, corresponding to a sum or difference of fermionic variables, which can then be parametrized much more accurately.

3.2 Spontaneous symmetry breaking

Within the one-loop truncation the effective two-particle interaction Γ^A always diverges in some momentum channels at a finite energy scale Λ_c , even for a small bare interaction U . Hence, one is always running into a strong coupling problem in the low-energy limit. The one-loop truncation breaks down, and also the simplified parametrization of the two-particle vertex described above cannot be justified in the presence of singular momentum and energy dependences.

If the vertex function diverges only in the Cooper channel, driven by the particle-particle contribution to the flow, the strong coupling problem emerging in the low-energy region can be controlled by exploiting Λ_c as a small parameter [41]. The formation of a superconducting ground state can then be described essentially by a BCS theory with renormalized input parameters. In the Hubbard model, the pure Cooper channel instability is always realized for $\mu \neq \epsilon_{vH}$ at sufficiently small U . In that regime, one can safely infer superconductivity with a d -wave order parameter from the divergence of the one-loop pairing susceptibility. At finite temperature, the off-diagonal long-range order will of course turn into the quasi long-range order of a Kosterlitz-Thouless phase.

In general, the one-loop calculation can produce divergences of the vertex function in various momentum channels, with large contributions from both particle-particle and particle-hole diagrams. This can happen even in the weak coupling limit $U \rightarrow 0$, namely when the chemical potential approaches the van Hove singularity. In that case, different possible instabilities can compete in a complicated way. Besides spin density wave and pairing instabilities, one has to deal with ferromagnetism (at moderate $|t'/t|$ in the Hubbard model) [42–44] and a d -wave Pomeranchuk instability of the Fermi surface [45] as alternative or coexisting candidates.

A complete theory of the effective strong coupling problem emerging from strong particle-particle and particle-hole fluctuations has not yet been achieved. For weak bare coupling, one may again try to exploit the smallness of the scale Λ_c where strong fluctuations appear to construct a tractable effective low-energy theory. Spontaneous symmetry breaking can be handled within the functional RG framework by adding an infinitesimal symmetry breaking term at the beginning of the flow, which is then promoted to a finite order parameter at the scale Λ_c [19, 20]. The calculations are complicated by the appearance of anomalous interaction vertices, and, in case of continuous symmetries, by singularities associated with Goldstone modes. Nevertheless, fermionic functional RG flows with spontaneous symmetry breaking were computed for the superconducting ground state of the attractive [46, 47] and repulsive [48] Hubbard model. In systems with a metastable phase, e.g., near a first order transition, a shift of the initial conditions by a counter term is needed to drive the flow into the stable symmetry broken phase [49]. Order parameter fluctuations are most conveniently treated by introducing appropriate bosonic fields, as first discussed for antiferromagnetic order in the half-filled Hubbard model [50].

In case of competing order parameters, such as antiferromagnetism and d -wave superconductivity near half-filling, a full RG treatment of spontaneous symmetry-breaking and order parameter fluctuations is a rather ambitious long-term goal [51]. As a simpler alternative one may neglect low-energy fluctuations and combine the RG with a mean-field (MF) theory of symmetry-

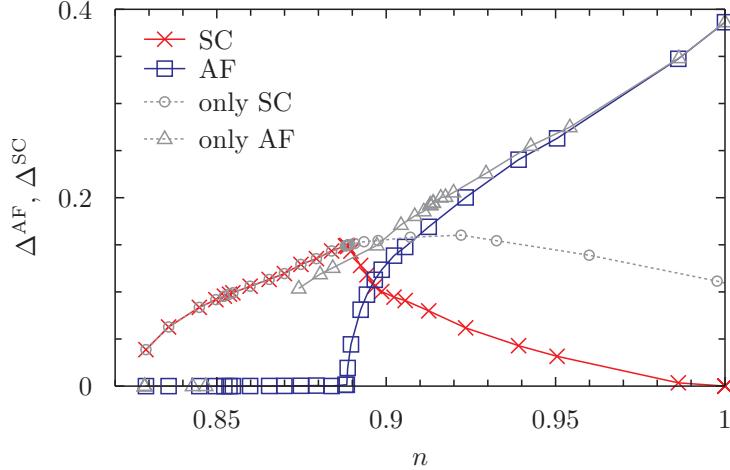


Fig. 8: Amplitudes of antiferromagnetic and superconducting gap functions in the ground state of the two-dimensional Hubbard model as a function of density, for $U/t = 3$ and $t'/t = -0.15$. Results from a coupled solution of the magnetic and superconducting gap equations with partial coexistence of orders are compared to purely magnetic and purely superconducting solutions. The amplitudes are plotted in units of t . Figure taken from Ref. [53].

breaking [52, 53]. In such a RG+MF approach, the one-loop flow is stopped at a scale Λ_{MF} that is small compared to the band width but still safely above the scale Λ_c where the two-particle vertex diverges. At this point the vertex has already developed a pronounced momentum dependence, reflecting in particular magnetic and superconducting correlations. The integration over the remaining modes, below Λ_{MF} , is treated in a mean-field approximation allowing, in particular, antiferromagnetic and superconducting order, where the effective interactions entering the mean-field equations are extracted from $\Gamma^{\Lambda_{\text{MF}}}$. At zero temperature, this approach should yield a decent approximation for the order parameters, since order parameter fluctuations usually do not play a crucial role for the gross features of the ground state. Results obtained for the amplitudes of the antiferromagnetic and superconducting gap functions in the ground state of the 2D Hubbard model, as obtained from a functional RG + MF calculation [53], are shown in Fig. 8. The amplitudes are defined as the maxima $\Delta^{\text{AF}} = \max_{\mathbf{k}} \Delta_{\mathbf{k}}^{\text{AF}}$ and $\Delta^{\text{SC}} = \max_{\mathbf{k}} \Delta_{\mathbf{k}}^{\text{SC}}$ of the gap functions. There is an extended region of *coexistence* of magnetic and superconducting order, which occurs naturally due to pairing of electrons near the reconstructed Fermi surface (pockets) in the antiferromagnetic state. In Fig. 9, the momentum dependence of the gap functions is shown for four distinct densities at and below half-filling. The superconducting gap function has the expected *d*-wave symmetry, while the antiferromagnetic gap has *s*-wave symmetry with a relatively weak momentum dependence.

4 Leap to strong coupling: DMFT as a booster rocket

A truncation of the functional RG hierarchy of flow equations can be justified only for weak interactions, with the exception of mean-field models where phase space restrictions suppress higher order contributions [19]. Although bare interactions are usually two-particle interactions, m -particle effective interactions with $m > 2$ are generated by the flow and affect the

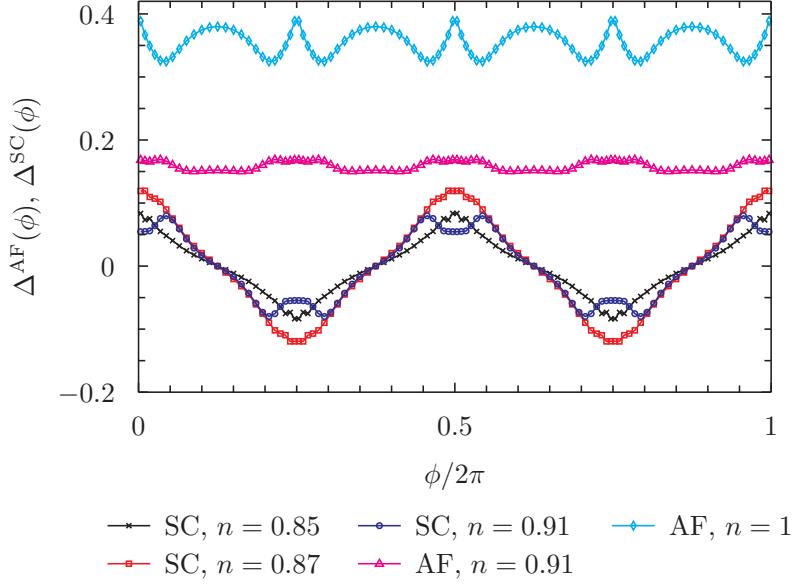


Fig. 9: Momentum dependence of the antiferromagnetic and superconducting gap functions for various choices of the density. The momentum dependence is parametrized by the angle ϕ between \mathbf{k} and the k_x -axis, where $\Delta_{\mathbf{k}}^{\text{AF}}$ is evaluated with \mathbf{k} on the umklapp surface ($|k_x \pm k_y| = \pi$) and $\Delta_{\mathbf{k}}^{\text{SC}}$ with \mathbf{k} on the Fermi surface. The model parameters are the same as in Fig. 8. Figure taken from Ref. [53].

effective two-particle interaction and the self-energy. For example, an effective three-particle interaction $\Gamma^{(6)\text{A}}$ is generated by a contribution of third order in the two-particle vertex $\Gamma^{(4)\text{A}}$, which then feeds back into the flow of $\Gamma^{(4)\text{A}}$, as can be seen in Fig. 3. For strong bare interactions, these contributions from effective interactions beyond the two-particle level become important already at relatively high energy scales, that is, at scales well above the critical scales for instabilities.

For systems with short-range interactions such as the Hubbard model, the correlations at high and intermediate energy scales are well described by the dynamical mean-field theory (DMFT) and its quantum cluster extensions, since long-range correlations emerge only at low energy scales [54]. In particular, the DMFT captures non-perturbative phenomena such as the Mott-Hubbard metal-insulator transition, which is a consequence of strong local correlations [55].

The DMFT is based on a local approximation for the fermionic self-energy, which is exact in the limit of infinite lattice dimensionality [56]. In that limit, irreducible m -particle vertices are local, too. It is thus natural to use the DMFT solution as a starting point, and include non-local correlations subsequently by expanding around the DMFT vertices and self-energy [57]. Several such extensions of the DMFT, involving various types of resummed perturbation expansions for non-local corrections, have already been proposed [58–60].

Most recently, it was shown that the DMFT can be used as a non-perturbative starting point for a functional RG flow [4]. The DMFT vertices and self-energy set the initial condition of the flow. In the remainder of this section, I will describe the fusion of DMFT and functional RG to the non-perturbative DMF²RG, and discuss first results for the two-dimensional Hubbard model.

4.1 Dynamical mean-field theory

The DMFT was developed in two steps. First, it was shown that models of interacting lattice fermions have a non-trivial infinite dimensionality $d \rightarrow \infty$ limit [56], where local correlations survive, while non-local contributions to one-particle irreducible quantities such as the self-energy vanish [56, 61]. Second, it was shown that a lattice fermion system with a local self-energy can be mapped to a quantum impurity problem with a self-consistency condition [62, 63]. The resulting equations are a dynamical variant of the static Weiss mean-field theory for magnets [62] – hence the name “dynamical mean-field theory” commonly used since 1995.

For the Hubbard model, the DMFT equations can be derived on the back of an envelope as follows. The propagator G is related to the bare propagator G_0 and the self-energy Σ by Dyson’s equation, $G^{-1}(k_0, \mathbf{k}) = G_0^{-1}(k_0, \mathbf{k}) - \Sigma(k_0)$. Non-local (in real space) contributions to the self-energy are discarded in the DMFT, so that $\Sigma(k)$ depends only on frequency. This approximation becomes exact in the limit $d \rightarrow \infty$ [56, 61]. In a one-particle irreducible skeleton expansion, the local part of the self-energy is a functional of the local propagator

$$G_{\text{loc}}(k_0) = \int_{\mathbf{k}} G(k_0, \mathbf{k}) , \quad (42)$$

where $\int_{\mathbf{k}} = \int \frac{d^d \mathbf{k}}{(2\pi)^d}$. The same skeleton expansion yields the self-energy of the purely local action

$$\mathcal{S}_{\text{loc}}[\psi, \bar{\psi}] = - \sum_{k_0, \sigma} \bar{\psi}_{k_0, \sigma} \mathcal{G}_0^{-1}(k_0) \psi_{k_0, \sigma} + U \int_0^\beta d\tau \bar{\psi}_\uparrow(\tau) \psi_\uparrow(\tau) \bar{\psi}_\downarrow(\tau) \psi_\downarrow(\tau) , \quad (43)$$

where τ denotes imaginary time and $\psi_\sigma(\tau)$ and $\bar{\psi}_\sigma(\tau)$ are the Fourier transforms of $\psi_{k_0, \sigma}$ and $\bar{\psi}_{k_0, \sigma}$, respectively. The dynamical Weiss field \mathcal{G}_0^{-1} is determined by the self-consistency condition requiring that the local propagator and the self-energy of the lattice electrons coincide with the propagator and the self-energy of \mathcal{S}_{loc} , that is,

$$G_{\text{loc}}^{-1}(k_0) = \mathcal{G}_0^{-1}(k_0) - \Sigma(k_0) . \quad (44)$$

The main difficulty is to compute the self-energy as a functional of \mathcal{G}_0 from the action \mathcal{S}_{loc} . Introducing an auxiliary bath of non-interacting conduction electrons, this problem can be mapped to a well-known quantum impurity problem, the single-impurity Anderson model [62, 63], for which efficient non-perturbative numerical algorithms exist. Once a self-consistent \mathcal{G}_0 has been determined, one can also compute DMFT vertices from \mathcal{S}_{loc} [64].

4.2 From infinite to finite dimensions

We now set up a flow which starts from the local self-energy and local vertices as given by the DMFT, and builds up non-local correlations successively [4]. Without truncations, the exact non-local quantities would be obtained at the end of the flow. There is no double-counting of contributions. One way of defining such a flow, which is associated with an intuitive picture, is

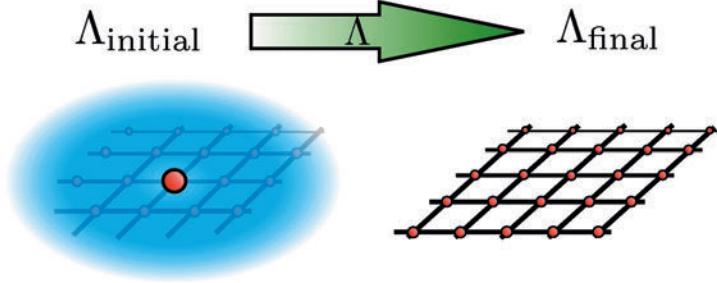


Fig. 10: Illustration of the flow from a DMFT solution to the solution of the d -dimensional (here $d = 2$) lattice system. Figure taken from Ref. [4].

to flow from infinite dimensions, where the DMFT is exact, to the actual dimensionality of the system.

To be specific, let us consider a d -dimensional hypercubic lattice with nearest neighbor hopping, supplemented by \bar{d} additional auxiliary dimensions with $\bar{d} \rightarrow \infty$. The flow equations derived in Sec. 2 are valid for an arbitrary scale dependence of the bare propagator. To flow from infinite to d dimensions, we introduce scale dependent hopping amplitudes in the physical and auxiliary directions, denoted by t^A and \bar{t}^A , respectively. The initial hopping amplitudes are $t^{A_0} = 0$ and $\bar{t}^{A_0} = (d/\bar{d})^{1/2} t$. The scaling with $\bar{d}^{-1/2}$ is required for a proper limit $\bar{d} \rightarrow \infty$ [56]. At the end of the flow, for $\Lambda \rightarrow 0$, the hopping amplitude in the auxiliary dimensions are switched off, $\bar{t}^{A \rightarrow 0} = 0$, while the real hopping amplitude is turned on, $t^{A \rightarrow 0} = t$.

The corresponding scale dependent bare propagator has the form

$$G_0^A(k_0, \mathbf{k}, \bar{\mathbf{k}}) = \frac{1}{ik_0 + \mu - \epsilon_{\mathbf{k}}^A - \bar{\epsilon}_{\bar{\mathbf{k}}}^A} , \quad (45)$$

with the scale dependent dispersion relations

$$\begin{aligned} \epsilon_{\mathbf{k}}^A &= -2t^A (\cos k_1 + \dots + \cos k_d) , \\ \bar{\epsilon}_{\bar{\mathbf{k}}}^A &= -2\bar{t}^A (\cos \bar{k}_1 + \dots + \cos \bar{k}_{\bar{d}}) . \end{aligned} \quad (46)$$

For $\bar{d} \rightarrow \infty$, only the part of the propagator which is local in the auxiliary dimensions contributes to the self-energy and vertices. That local part is given by a $\bar{\mathbf{k}}$ -integration,

$$G_0^A(k_0, \mathbf{k}) = \int_{\bar{\mathbf{k}}} G_0^A(k_0, \mathbf{k}, \bar{\mathbf{k}}) = \int d\bar{\epsilon} \frac{\bar{\rho}^A(\bar{\epsilon})}{ik_0 + \mu - \epsilon_{\mathbf{k}}^A - \bar{\epsilon}} , \quad (47)$$

where $\bar{\rho}^A(\bar{\epsilon})$ is the density of states for $\bar{\epsilon}_{\bar{\mathbf{k}}}^A$. The latter is a normalized Gaussian distribution with a width that shrinks to zero for $\Lambda \rightarrow 0$, so that $G_0^{A \rightarrow 0}(k_0, \mathbf{k}) = (ik_0 + \mu - \epsilon_{\mathbf{k}})^{-1}$.

G_0^A defines a flow that interpolates smoothly between the DMFT solution for the infinite dimensional model at Λ_0 and the solution of the d -dimensional system for $\Lambda \rightarrow 0$, as illustrated schematically in Fig. 10. The flow equations derived in Sec. 2 apply without any modification. However, the initial condition for the flow is not given by bare quantities, but by the self-energy

$$\frac{d}{d\Lambda} \Sigma^\Lambda = \text{loop with } \Gamma^\Lambda \quad \frac{d}{d\Lambda} \Gamma^\Lambda = \text{loop with } G^\Lambda$$

Fig. 11: Truncation of flow equations in DMF²RG applied to the 2D Hubbard model.

and vertices as obtained from the DMFT, that is, $\Sigma^{A_0} = \Sigma_{\text{DMFT}}$ and $\Gamma^{(2m)A_0} = \Gamma_{\text{DMFT}}^{(2m)}$. Hence, non-perturbative local correlations effects, such as the Mott-Hubbard transition, are built in already at the starting point. The DMFT thus becomes a ‘‘booster rocket’’ for the flow.

The infinite dimensionality limit served as a guide to develop the idea of setting up a flow that starts with the DMFT. However, the DMFT is usually applied as an approximation for a finite dimensional system, without introducing extra dimensions. In the same spirit, one can also define a flow with the DMFT as a starting point, while staying entirely in the d -dimensional physical space. A particularly simple choice for the scale dependence of the bare propagator is given by [4]

$$[G_0^A(k_0, \mathbf{k})]^{-1} = \Lambda \mathcal{G}_0^{-1}(k_0) + (1 - \Lambda) G_0^{-1}(k_0, \mathbf{k}), \quad (48)$$

with $\Lambda_0 = 1$. This choice is very simple but certainly not optimal. In particular, it does not regularize infrared singularities. Better choices obeying additional requirements besides the correct boundary conditions can be constructed on demand.

4.3 Application to the 2D Hubbard model

Since the DMF²RG has been proposed very recently, there is only one concrete application, for the two-dimensional Hubbard model with pure nearest neighbor hopping at half-filling [4]. Due to perfect nesting, the ground state is antiferromagnetically ordered for any $U > 0$ in that case. In that first application, relatively crude approximations were made to compute the flow. The calculation was done directly in two dimensions, with the DMFT solution as starting point, and the flow was defined by G_0^A as in Eq. (48). The flow was truncated by discarding $\Gamma^{(6)A}$, such that $\Gamma^A = \Gamma^{(4)A}$ and Σ^A are determined by a closed system of one-loop flow equations, see Fig. 11, with the DMFT self-energy Σ_{DMFT} and two-particle vertex Γ_{DMFT} as initial condition. The rationale behind this truncation is that the strongest local correlations are already included by the DMFT starting point, and that the effect of three-particle correlations on non-local correlations can be expected to be less important. Clearly, this needs to be checked in the future. The frequency dependence of the vertex was parametrized by using a channel decomposition as in Refs. [15, 40] with only one important frequency dependence in each channel. The latter can be discretized accurately, as can the frequency dependence of the self-energy. However, the channel decomposition does not capture certain structures appearing in the DMFT vertex

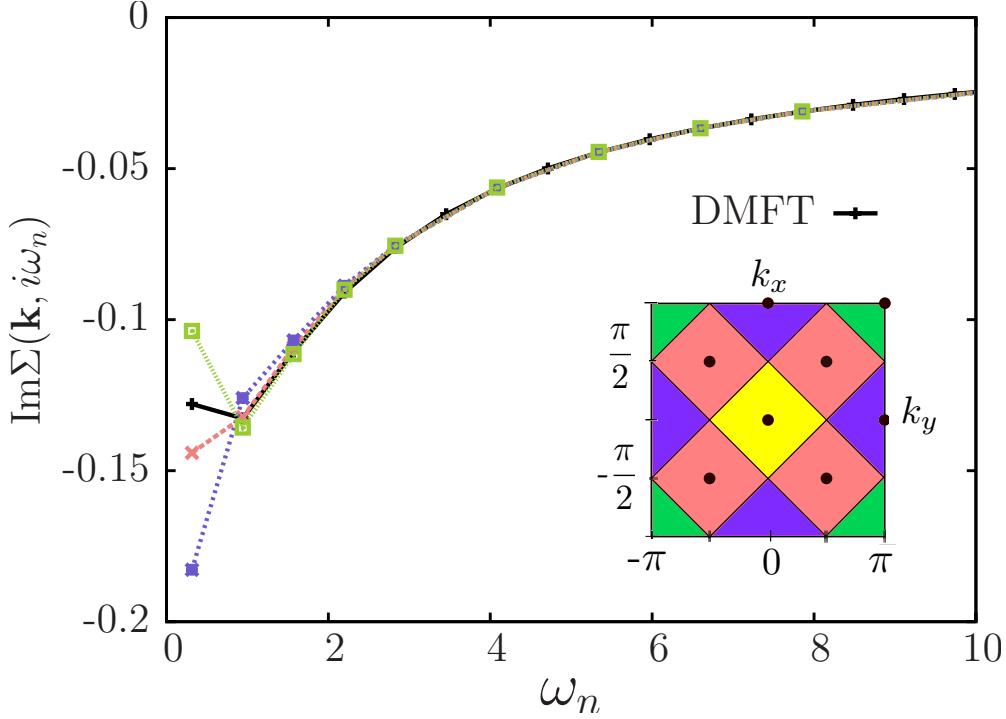


Fig. 12: Imaginary part of the fermionic self-energy as a function of Matsubara frequency for the half-filled 2D Hubbard model at $U = 4t$ and $T = 0.4t$. The DMFT result (\mathbf{k} -independent) is compared to the result from the DMF²RG, where the color coding of \mathbf{k} is defined in the inset. Note that the results for $\mathbf{k} = (0, 0)$ and $\mathbf{k} = (\pi, \pi)$ coincide due to particle-hole symmetry. Figure taken from Ref. [4].

at strong coupling [64], which limited the calculations to interaction strengths below the Mott insulator regime. The momentum dependence was discretized with a few patches.

A result for the self-energy of the half-filled 2D Hubbard model obtained from the DMF²RG as described above is shown in Fig. 12, and compared to the DMFT solution. The discretization of momentum space is shown in the inset. A relatively high temperature $T = 0.4t$ was chosen so that the flow could be carried out until $\Lambda = 0$ without running into divergences associated with the antiferromagnetic instability. The DMF²RG result deviates significantly from the momentum independent DMFT solution only at the lowest Matsubara frequencies. There, a pronounced momentum dependence emerges. The trend of a large enhancement of the self-energy at momenta near $(\pi, 0)$ and $(0, \pi)$, related to a pseudogap formation, agrees with results from cluster extensions of the DMFT [65]. The results also agree qualitatively with those obtained from a plain one-loop functional RG calculation, confirming that a weak coupling expansion is still reasonable at moderate coupling strengths as $U = 4t$. Clearly, qualitative differences between weak-coupling truncations of the functional RG and the non-perturbative DMF²RG are to be expected at interaction strengths comparable to the band width and beyond. To apply the DMF²RG at large U , a suitable parametrization of the complex frequency and momentum dependence of the two-particle vertex at strong coupling needs to be developed.

5 Conclusion

The functional RG has become a valuable source of new approximation schemes for interacting Fermi systems [3]. The method is based on an exact flow equation, which describes the flow of the effective action as a function of a suitable flow parameter. From the final effective action, at the end of the flow, any desired information about the system can be obtained. Approximations are constructed by truncating the effective action. In many cases, rather simple truncations turned out to capture rather complex many-body phenomena. Compared to the traditional resummations of perturbation theory, these approximations have the advantage that infrared singularities are treated properly due to the built-in RG structure. Unlike other RG methods, approximations derived in the functional RG framework can be applied directly to microscopic models, not only to renormalizable effective field theories. Remarkably, the functional RG reviewed here as a computational tool is very similar to RG approaches used by mathematicians to derive general rigorous results for interacting Fermi systems.

Applications of the functional RG to the two-dimensional Hubbard model have improved our understanding of its instabilities. In particular, the existence of d -wave superconductivity in that model was conclusively established.

With the very recent fusion of DMFT and functional RG [4], where the DMFT is used as a non-perturbative starting point for the functional RG flow, a promising route has been opened to capture all aspects of strongly interacting Fermi systems, such as the 2D Hubbard model at large U , over all energy scales in one framework.

Acknowledgment

I would like to thank my coworkers S. Andergassen, J. Bauer, A. Eberlein, C. Halboth, K. Held, C. Honerkamp, A. Katanin, V. Meden, J. Reiss, D. Rohe, G. Rohringer, M. Salmhofer, K. Schönhammer, C. Taranto, A. Toschi, D. Vollhardt, and J. Wang for our fruitful and enjoyable collaboration on the work presented in this lecture. Support from the DFG through research unit FOR 723 is also gratefully acknowledged.

Appendix

A Derivation of functional flow equation

Here we describe the derivation of the functional flow equation (15) for the effective action Γ^A . Introducing a scale-dependent bare propagator $G_0^A = (Q_0^A)^{-1}$ in the functional integral representation (4) of the generating functional for connected Green functions, one can write

$$e^{-\mathcal{G}^A[\eta, \bar{\eta}]} = \int \prod_K d\psi_K d\bar{\psi}_K e^{(\bar{\psi}, Q_0^A \psi)} e^{-V[\psi, \bar{\psi}]} e^{(\bar{\eta}, \psi) + (\bar{\psi}, \eta)} . \quad (49)$$

Taking a Λ -derivative on both sides yields

$$\begin{aligned} -(\partial_\Lambda \mathcal{G}^A[\eta, \bar{\eta}]) e^{-\mathcal{G}^A[\eta, \bar{\eta}]} &= \int \prod_K d\psi_K d\bar{\psi}_K (\bar{\psi}, \partial_\Lambda Q_0^A \psi) e^{(\bar{\psi}, Q_0^A \psi)} e^{-V[\psi, \bar{\psi}]} e^{(\bar{\eta}, \psi) + (\bar{\psi}, \eta)} \\ &= -(\partial_\eta, \partial_\Lambda Q_0^A \partial_{\bar{\eta}}) e^{-\mathcal{G}^A[\eta, \bar{\eta}]} , \end{aligned} \quad (50)$$

which leads directly to a flow equation for \mathcal{G}^A ,

$$\frac{d}{d\Lambda} \mathcal{G}^A[\eta, \bar{\eta}] = \left(\frac{\partial \mathcal{G}^A}{\partial \eta}, \frac{dQ_0^A}{d\Lambda} \frac{\partial \mathcal{G}^A}{\partial \bar{\eta}} \right) + \text{Tr} \left(\frac{dQ_0^A}{d\Lambda} \frac{\partial^2 \mathcal{G}^A}{\partial \bar{\eta} \partial \eta} \right) . \quad (51)$$

The effective action is the Legendre transform

$$\Gamma^A[\psi, \bar{\psi}] = \mathcal{G}^A[\eta^A, \bar{\eta}^A] + (\bar{\psi}, \eta^A) + (\bar{\eta}^A, \psi) . \quad (52)$$

Note that η^A and $\bar{\eta}^A$ are Λ -dependent functions of ψ and $\bar{\psi}$, so that

$$\frac{d}{d\Lambda} \Gamma^A[\psi, \bar{\psi}] = \frac{d}{d\Lambda} \mathcal{G}^A[\eta^A, \bar{\eta}^A] + (\bar{\psi}, \partial_\Lambda \eta^A) + (\partial_\Lambda \bar{\eta}^A, \psi) . \quad (53)$$

The total derivative acts also on the Λ -dependence of η^A and $\bar{\eta}^A$. Using the relations $\partial \mathcal{G}^A / \partial \bar{\eta} = -\psi$, and $\partial \mathcal{G}^A / \partial \eta = \bar{\psi}$, all terms arising from the Λ -dependence of η^A and $\bar{\eta}^A$ cancel, yielding

$$\frac{d}{d\Lambda} \Gamma^A[\psi, \bar{\psi}] = \frac{d}{d\Lambda} \mathcal{G}^A[\eta^A, \bar{\eta}^A] \Big|_{\eta^A, \bar{\eta}^A \text{ fixed}} . \quad (54)$$

The flow equation (15) for Γ^A now follows directly from Eq. (51) and the reciprocity relation $\Gamma^{(2)A}[\psi, \bar{\psi}] = (\mathbf{G}^{(2)A}[\eta^A, \bar{\eta}^A])^{-1}$.

References

- [1] J. Berges, N. Tetradis, and C. Wetterich, Phys. Rep. **363**, 223 (2002)
- [2] P. Kopietz, L. Bartosch, and F. Schütz:
Introduction to the Functional Renormalization Group (Springer, Berlin, 2010).
- [3] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Rev. Mod. Phys. **84**, 299 (2012)
- [4] C. Taranto, S. Andergassen, J. Bauer, K. Held, A.A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, Phys. Rev. Lett. **112**, 196402 (2014)
- [5] J. Solyom, Adv. Phys. **28**, 201 (1979)
- [6] J. Feldman and E. Trubowitz, Helv. Phys. Acta **63**, 156 (1990); *ibid* **64**, 214 (1991)
- [7] G. Benfatto and G. Gallavotti, J. Stat. Phys. **59**, 541 (1990)
- [8] K.G. Wilson and J. Kogut, Phys. Rep. C **12**, 75 (1974)
- [9] M. Salmhofer: *Renormalization: An Introduction* (Springer, Heidelberg, 1999)
- [10] J. Feldman, H. Knörrer, and E. Trubowitz, Rev. Math. Phys. **15**, 949-1169 (2003); Commun. Math. Phys. **247**, 1-320 (2004)
- [11] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994)
- [12] J. Polchinski, in *Proceedings of 1993 Theoretical Advanced Studies Institute in Elementary Particle Physics*, ed. by J. Harvey, J. Polchinski (World Scientific, Singapore, 1993)
- [13] C. Wetterich, Phys. Lett. B **301**, 90 (1993)
- [14] J.W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Reading, 1987)
- [15] C. Husemann and M. Salmhofer, Phys. Rev. B **79**, 195125 (2009)
- [16] S. Andergassen, T. Enss, V. Meden, W. Metzner, U. Schollwöck, and K. Schönhammer, Phys. Rev. B **70**, 075102 (2004)
- [17] J. Polchinski, Nucl. Phys. **B231**, 269 (1984)
- [18] F. Schütz, L. Bartosch, and P. Kopietz, Phys. Rev. B **72**, 035107 (2005)
- [19] M. Salmhofer, C. Honerkamp, W. Metzner, and O. Lauscher, Prog. Theor. Phys. **112**, 943 (2004)
- [20] R. Gersch, C. Honerkamp, D. Rohe, and W. Metzner, Eur. Phys. J. B **48**, 349 (2005)

- [21] A. Eberlein and W. Metzner, Prog. Theor. Phys. **124**, 471 (2010)
- [22] M. Salmhofer and C. Honerkamp, Prog. Theor. Phys. **105**, 1 (2001)
- [23] P.W. Anderson, Science **235**, 1196 (1987)
- [24] For an early review on *d*-wave superconductivity from spin fluctuations, see D.J. Scalapino, Phys. Rep. **250**, 329 (1995)
- [25] K. Miyake, S. Schmitt-Rink, and C.M. Varma, Phys. Rev. B **34**, 6554 (1986)
- [26] D.J. Scalapino, E. Loh, and J.E. Hirsch, Phys. Rev. B **34**, 8190 (1986)
- [27] N.E. Bickers, D.J. Scalapino, and R.T. Scalettar, Int. J. Mod. Phys. B **1**, 687 (1987)
- [28] N.E. Bickers, D.J. Scalapino, and S.R. White, Phys. Rev. Lett. **62**, 961 (1989)
- [29] D. Manske: *Theory of Unconventional Superconductors*, Springer Tracts in Modern Physics, vol. 202 (Springer, Berlin, 2004)
- [30] For reviews of numerical work, see E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994); N. Bulut, Adv. Phys. **51**, 1587 (2002)
- [31] H.J. Schulz, Europhys. Lett. **4**, 609 (1987)
- [32] I. Dzyaloshinskii, Sov. Phys. JETP **66**, 848 (1987)
- [33] P. Lederer, G. Montambaux, and D. Poilblanc, J. Phys. (Paris) **48**, 1613 (1987)
- [34] C.J. Halboth and W. Metzner, Phys. Rev. B **61**, 7364 (2000)
- [35] D. Zanchi and H.J. Schulz, Europhys. Lett. **44**, 235 (1998); Phys. Rev. B **61**, 13609 (2000)
- [36] C. Honerkamp, M. Salmhofer, N. Furukawa, and T.M. Rice, Phys. Rev. B **63**, 035109 (2001)
- [37] A.P. Kampf and A.A. Katanin, Phys. Rev. B **67**, 125104 (2003)
- [38] See, for example, H.J. Schulz, Phys. Rev. Lett. **64**, 1445 (1990), where also the choice of δ as a function of doping is discussed.
- [39] C. Husemann, K.-U. Giering, and M. Salmhofer, Phys. Rev. B **85**, 075121 (2012)
- [40] A similar channel decomposition was derived and applied also for the single-impurity Anderson model by C. Karrasch, R. Hedden, R. Peters, T. Pruschke, K. Schönhammer, and V. Meden, J. Phys. Condens. Matter **20**, 345205 (2008)
- [41] J. Feldman, J. Magnen, V. Rivasseau, and E. Trubowitz, Europhys. Lett. **24**, 437 (1993)
- [42] R. Hlubina, S. Sorella, and F. Guinea, Phys. Rev. Lett. **78**, 1343 (1997)

- [43] C. Honerkamp and M. Salmhofer, Phys. Rev. Lett. **87**, 187004 (2001)
- [44] A.A. Katanin and A.P. Kampf, Phys. Rev. B **68**, 195101 (2003)
- [45] C.J. Halboth and W. Metzner, Phys. Rev. Lett. **85**, 5162 (2000)
- [46] R. Gersch, C. Honerkamp, and W. Metzner, New J. Phys. **10**, 045003 (2008)
- [47] A. Eberlein and W. Metzner, Phys. Rev. B **87**, 174523 (2013)
- [48] A. Eberlein and W. Metzner, Phys. Rev. B **89**, 035126 (2014)
- [49] R. Gersch, J. Reiss, and C. Honerkamp, New J. Phys. **8**, 320 (2006)
- [50] T. Baier, E. Bick, and C. Wetterich, Phys. Rev. B **70**, 125111 (2004)
- [51] S. Friederich, H. C. Krahl, and C. Wetterich, Phys. Rev. B **81**, 235108 (2010);
ibid. **83**, 155125 (2011)
- [52] J. Reiss, D. Rohe, and W. Metzner, Phys. Rev. B **75**, 075110 (2007)
- [53] J. Wang, A. Eberlein, and W. Metzner, Phys. Rev. B **89**, 121116(R) (2014)
- [54] A. Georges, Ann. Phys. (Berlin) **523**, 672 (2011)
- [55] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [56] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [57] Within the Gutzwiller variational ansatz, a perturbative expansion for non-local corrections around the local $d = \infty$ solution was shown to be very promising already long ago, see W. Metzner, Z. Phys. **77**, 253 (1989)
- [58] J.P. Hague, M. Jarrell, and T.C. Schulthess, Phys. Rev. B **69**, 165113 (2004)
- [59] A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. B **75**, 045118 (2007); A. Valli, G. Sangiovanni, O. Gunnarsson, A. Toschi, and K. Held, Phys. Rev. Lett. **104**, 246402 (2010); G. Rohringer, A. Toschi, A.A. Katanin, and K. Held, Phys. Rev. Lett. **107**, 256402 (2011)
- [60] A.N. Rubtsov, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **77**, 033101 (2008); H. Hafermann, G. Li, A.N. Rubtsov, M.I. Katsnelson, A.I. Lichtenstein, and H. Monien, Phys. Rev. Lett. **102**, 206401 (2009)
- [61] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989); *ibid.* **76**, 211 (1989).
- [62] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [63] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)

- [64] For a comprehensive recent analysis of the two-particle vertex obtained by DMFT, see G. Rohringer, A. Valli, and A. Toschi, Phys. Rev. B **86**, 125114 (2012)
- [65] T.A. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)

12 Correlated Electron Dynamics and Nonequilibrium Dynamical Mean-Field Theory

Marcus Kollar

Center for Electronic Correlations and Magnetism
University of Augsburg

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Nonequilibrium Green functions | 3 |
| 2.1 | Time contour | 3 |
| 2.2 | Contour Green functions | 5 |
| 2.3 | Equilibrium case | 7 |
| 2.4 | Noninteracting case | 8 |
| 2.5 | Self-energy | 8 |
| 3 | Nonequilibrium DMFT | 9 |
| 3.1 | Cavity method | 9 |
| 3.2 | DMFT action for an infinite-dimensional lattice | 10 |
| 3.3 | Local self-energy | 11 |
| 3.4 | Self-consistency condition | 13 |
| 3.5 | Bethe lattice | 13 |
| 3.6 | Numerical methods | 14 |
| 4 | Correlated electrons in nonequilibrium | 15 |
| 4.1 | Relaxation and thermalization | 15 |
| 4.2 | Interaction quench in the Hubbard model | 16 |

1 Introduction

In Chapters 1 and 3, it was discussed how the limit of infinite lattice dimension $d \rightarrow \infty$ [1] leads to drastic simplifications for many-body theory, which describes interacting electrons in their ground state or in thermodynamic equilibrium. In particular, Hubbard-type models in the thermodynamic limit are mapped exactly onto effective single-site problems with a local self-energy, which in turn may be represented as self-consistent single-impurity Anderson models that can be solved numerically. For systems in dimension $d = 1, 2, 3$ this approach corresponds to a mean-field approximation, i.e. to dynamical mean-field theory (DMFT) [2], which can be further improved by including corrections for finite dimensions, see e.g., Chapters 9 and 10. DMFT can also be applied to nonequilibrium problems, i.e., the single-band Hubbard model with time-dependent hopping amplitudes and interaction parameter,

$$H_{\text{Hubbard}}(t) = \sum_{ij\sigma} t_{ij}(t) c_{i\sigma}^\dagger c_{j\sigma} + U(t) \sum_i (n_{i\uparrow} - \frac{1}{2})(n_{i\downarrow} - \frac{1}{2}). \quad (1)$$

Due to recent experimental advances, theoretical methods to study such systems are of great interest. Correlated materials can be excited and their relaxation monitored using pump-probe spectroscopy with femtosecond laser pulses [3,4]. In a suitable gauge, the electric field couples to the band energies via the time-dependent vector potential according to the Peierls substitution [5],

$$t_{ij}(t) = t_{ij} \exp\left(-\frac{ie}{\hbar} \int_{\mathbf{R}_i}^{\mathbf{R}_j} d\mathbf{r} \cdot \mathbf{A}(\mathbf{r}, t)\right). \quad (2)$$

However, after a few hundred femtoseconds the electronic degrees of freedom will typically have relaxed and their coupling to the slower vibrational lattice degrees of freedom will come into play. From a quite different perspective, it is also possible to study the real-time behavior of many-body states using ultracold atomic gases in optical lattices, which can be kept in excellent isolation from the environment and for which kinetic and interaction energies can be controlled very precisely for many hundreds of microseconds [6, 7].

In general, many-body theory for nonequilibrium is numerically even more demanding than for equilibrium. For systems in equilibrium, the main task is to evaluate expectation values (such as Green functions) for a grand-canonical density matrix (such as a thermal state for an interacting Hamiltonian or possibly its ground state). In nonequilibrium, additionally, the time evolution under a time-dependent Hamiltonian must be taken into account. However, using nonequilibrium Green functions according to the Keldysh formalism, the limit of infinite lattice dimensions provides similar benefits as in the equilibrium case, i.e., the problem is again reduced to a single site by integrating out the rest of the lattice, although the remaining nonequilibrium impurity problem is more complicated than in equilibrium. The nonequilibrium DMFT approach was first developed and applied for the Falicov-Kimball model [8–15] (following an earlier incomplete attempt in Ref. [16]). Since then, many more applications and extensions have appeared, including for time-resolved spectroscopy [17–19], abrupt and slow

changes in parameters [20–26], response to DC and AC fields and pulses [27–43], antiferromagnetic phases [44–47], coupling to phonons [48, 49], inhomogeneous systems [50], extensions for lower dimensions [51–53], and dynamics of lattice bosons [54]. A recent review of the nonequilibrium DMFT and its applications can be found in Ref. [55].

Below, the main ingredients for nonequilibrium DMFT are discussed. First, nonequilibrium Green functions are defined according to the Keldysh formalism (Sec. 2). Then, the mapping of a Hubbard model to a single-site problem with a dynamic bath is discussed (Sec. 3). Finally, some aspects of the single-band Hubbard model in nonequilibrium (Sec. 4) are reviewed. We follow mostly the setup and notation of Refs. [55] and [26].

2 Nonequilibrium Green functions

2.1 Time contour

We consider the time evolution of a quantum many-body system with density matrix $\rho(t)$ that starts at time $t = 0$ from thermal equilibrium, i.e., from a grand-canonical Gibbs state of the Hamiltonian $H(0)$,

$$\rho(0) = \frac{1}{Z} e^{-\beta \mathcal{H}(0)}, \quad (3)$$

i.e., the classical superposition of all eigenstates of the Hamiltonian depending on their Boltzmann weights. Here $\beta = 1/(k_B T)$ is the inverse temperature, $k_B = 1$, $\mathcal{H}(t) = H(t) - \mu N(t)$, with μ the chemical potential, $N(t)$ the particle number operator, and $Z = \text{Tr } e^{-\beta \mathcal{H}(0)}$ the equilibrium partition function. At time $t = 0$ the Hamiltonian changes, either continuously or abruptly, e.g., by switching on an electric field. The time-dependent Schrödinger equation then determines the evolution of the wave function, which is continuous in t . For a density matrix, this yields the von Neumann equation ($\hbar = 1$) and its formal solution,

$$i \frac{d}{dt} \rho(t) = [\mathcal{H}(t), \rho(t)], \quad \rho(t) = U(t, 0) \rho(0) U(0, t). \quad (4)$$

Our goal is to obtain the time-dependent expectation value of a (time-independent) Schrödinger operator A ,

$$\langle A \rangle_t = \text{Tr}[\rho(t)A]. \quad (5)$$

The propagator obeys $\frac{d}{dt}U(t, t') = -i\mathcal{H}(t)U(t, t')$, hence it is unitary, $U(t, t')U(t, t')^\dagger = U(t, t')U(t', t) = 1$, and fulfills $U(t, t')U(t', t'') = U(t, t'')$. Because $\mathcal{H}(t)$ and $\mathcal{H}(t')$ do not in general commute at different times, the formal solution of the differential equation for $U(t, t')$ is

$$U(t, t') = \begin{cases} \mathbf{T} \exp \left(-i \int_{t'}^t d\bar{t} \mathcal{H}(\bar{t}) \right) & \text{for } t > t', \\ \bar{\mathbf{T}} \exp \left(-i \int_{t'}^t d\bar{t} \mathcal{H}(\bar{t}) \right) & \text{for } t < t', \end{cases} \quad (6)$$

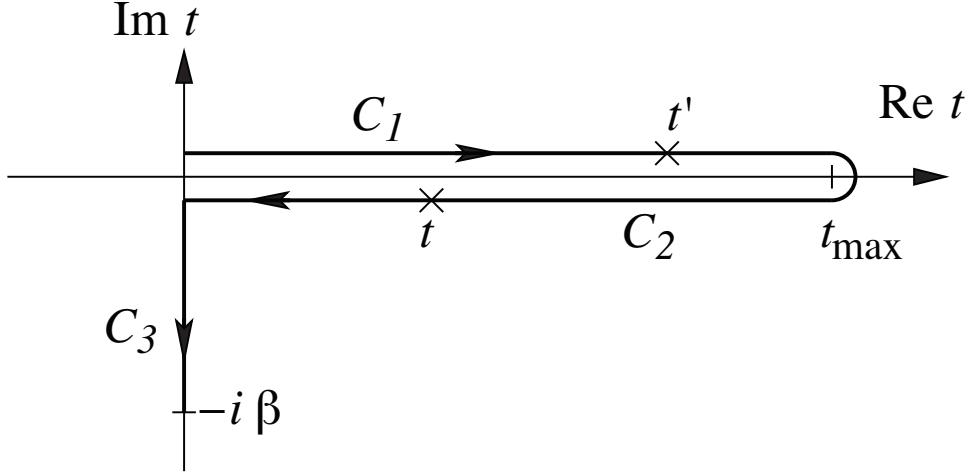


Fig. 1: The integration along the L-shaped integration contour C runs along C_1 from 0 to t_{\max} on the real axis, goes back along C_2 , and proceeds along C_3 to $-i\beta$. In the figure, the times $t \in C_2$ and $t' \in C_1$ are located such that $t >_C t'$, i.e., t is later than t' in the sense of contour ordering.

where T and \bar{T} denote time-ordering and anti-time-ordering operators, respectively; i.e., T re-orders the operators $\mathcal{H}(\bar{t})$ that occur in the expansion of the exponential such that the time arguments \bar{t} increase from right to left (and from left to right for \bar{T}). The density matrix $\rho(t)$ in the time-dependent expectation value (5) then involves one exponential with a forward integration along the time axis due to $U(t, t')$, one with a backwards integration due to $U(t', t)$, and in between sits $\exp(-\beta\mathcal{H}(0))$ representing the initial state, the exponent of which can be rewritten as an integral with respect to t from 0 to $-i\beta$ of $\mathcal{H}(0)$. The time ordering of operators $\mathcal{H}(\bar{t})$ in (5) thus involves three parts, C_1 : $0 \dots t_{\max}$, C_2 : $t_{\max} \dots 0$, and C_3 : $0 \dots -i\beta$ (where t_{\max} is the maximal time of interest)

$$\begin{aligned} \langle A \rangle_t &= \frac{1}{Z} \text{Tr}[U(-i\beta, 0)U(0, t) A U(t, 0)] \\ &= \frac{\text{Tr } T_C A(t) \exp[-i \int_C d\bar{t} \mathcal{H}(\bar{t})]}{\text{Tr } T_C \exp[-i \int_C d\bar{t} \mathcal{H}(\bar{t})]}. \end{aligned} \quad (7)$$

The integrals in the last expression are now along an L-shaped contour C that runs along C_1 , C_2 , and then C_3 ; T_C is the contour-ordering operator that arranges operators on C in the direction of the arrows in Fig. 1. The time argument t that has been attached to $A(t)$ merely indicates the time at which the (Schrödinger) operator A must be inserted in the contour time ordering. In the denominator of (7), no operator is inserted and hence the contributions from C_1 and C_2 cancel, yielding the partition function. The time parametrization along a contour allows to carry over many techniques from equilibrium many-body theory [56] (such as Feynman diagrams etc.), although it does not mean that integrations are actually performed in the complex plane. Rather, the contributions of the contour parts C_1 , C_2 , C_3 are evaluated separately. In particular, the following definitions for contour integrals, contour convolutions, time derivative, contour

theta and delta functions are useful:

$$g(t \in C) = \begin{cases} g^+(t) & \text{if } t \in [0, t_{\max}] \text{ on } C_1, \\ g^-(t) & \text{if } t \in [0, t_{\max}] \text{ on } C_2, \\ g^\dagger(-i\tau) & \text{if } t = -i\tau \text{ on } C_3, \tau \in [0, \beta], \end{cases} \quad (8)$$

$$\int_C dt g(t) = \int_0^{t_{\max}} dt g^+(t) - \int_0^{t_{\max}} dt g^-(t) - i \int_0^\beta d\tau g^\dagger(-i\tau), \quad (9)$$

$$[a * b](t, t') = \int_C d\bar{t} a(t, \bar{t}) b(\bar{t}, t'), \quad (10)$$

$$\partial_t g(t) = \begin{cases} \partial_t g(t^\pm) & t \in C_{1,2} \\ i\partial_\tau g(-i\tau) & t = -i\tau \in C_3 \end{cases}, \quad (11)$$

$$\theta_C(t, t') = \begin{cases} 1 & \text{for } t >_C t', \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

$$\delta_C(t, t') = \partial_t \theta_C(t, t'), \quad (13)$$

$$\int_C d\bar{t} \delta_C(t, \bar{t}) g(\bar{t}) = g(t). \quad (14)$$

Here $t >_C t'$ means that t appears later on the contour (as shown in Fig. 1).

The representation of time-dependent expectation values (7), in particular for nonequilibrium Green functions, is called the Keldysh formalism and is based on Refs. [57–59]. Modern introductions to the subject can be found, e.g., in Refs. [60–62], of which in particular [62] is very detailed, pedagogical, and complete. Note that, depending on the physical situation, other time contours are used in the literature. In particular, the so-called Keldysh contour that extends C_1 and C_2 to $-\infty$ without C_3 is useful to describe transport and nonequilibrium steady states with currents. However, for lattice models like (1) without reservoirs, the L-shaped contour is best suited, as it requires no further assumptions on the form of the nonequilibrium state, switching-on of interactions, etc. Further discussions of various contour shapes can be found, e.g., in Refs. [61–63].

2.2 Contour Green functions

The study of single-particle Green functions is a standard method to characterize the spectrum and state in many-particle systems. For later use the action S is defined for a given Hamiltonian as

$$S = -i \int_C dt \mathcal{H}(t), \quad (15)$$

so that $\langle A \rangle_t$ is now written as $\langle A(t) \rangle_S = \text{Tr}[T_C \exp(S) A(t)]/Z_S$ with $Z_S = \text{Tr}[T_C \exp(S)]$. If there is no ambiguity we will omit the subscript S . The definition of the time-ordering operator

is changed to

$$T_C A(t) B(t') = \begin{cases} AB & \text{if } t >_C t', \\ \pm BA & \text{if } t <_C t', \end{cases} \quad (16)$$

where the negative sign is used only if both A and B contain an odd number of fermionic annihilation or creation operators. Furthermore, if t and t' are equal and on the same branch of the contour, T_C performs a normal ordering by convention, moving all creators to the left. Single-particle contour Green functions for fermions are then defined as contour-ordered expectation values in analogy to the equilibrium case,

$$G(t, t') = -i\langle c(t)c^\dagger(t') \rangle = -\frac{i}{Z} \text{Tr} [T_C \{\exp(S)c(t)c^\dagger(t')\}], \quad (17)$$

where the basis state indices (e.g., site, spin, and orbital indices) have been omitted on the fermionic creation and annihilation operators for now, so all Green function should be regarded as matrices in these indices. The time arguments t and t' can each lie on one of the three parts of the contour so that G has nine entries with different physical meanings. A subscript $a, b = 1, 2, 3$ on $G_{ab}(t, t')$ then expresses whether a time argument is on the upper (1), lower (2), or imaginary (3) part of the contour C . There is some redundancy because one can shift the operator with the largest real-time argument between C_1 and C_2 because the time evolution along these paths cancels on its right. It follows that

$$G_{11}(t, t') = G_{12}(t, t') \quad \text{for } t \leq t', \quad (18a)$$

$$G_{11}(t, t') = G_{21}(t, t') \quad \text{for } t > t', \quad (18b)$$

$$G_{22}(t, t') = G_{21}(t, t') \quad \text{for } t < t', \quad (18c)$$

$$G_{22}(t, t') = G_{12}(t, t') \quad \text{for } t \geq t', \quad (18d)$$

$$G_{13}(t, \tau') = G_{23}(t, \tau'), \quad (18e)$$

$$G_{31}(\tau, t') = G_{32}(\tau, t'). \quad (18f)$$

It is then customary to define the following independent components: retarded (G^R), advanced (G^A), Keldysh (G^K), left-mixing (G^-), right-mixing (G^-), and Matsubara (G^M) Green function. For t and t' from C_1 or C_2 , τ from C_3 , and T_τ the imaginary-time ordering operator, they are given by

$$G^R(t, t') = \frac{1}{2}(G_{11} - G_{12} + G_{21} - G_{22}) = -i\theta(t - t')\langle\{c(t), c^\dagger(t')\}\rangle, \quad (19a)$$

$$G^A(t, t') = \frac{1}{2}(G_{11} + G_{12} - G_{21} - G_{22}) = i\theta(t' - t)\langle\{c(t), c^\dagger(t')\}\rangle, \quad (19b)$$

$$G^K(t, t') = \frac{1}{2}(G_{11} + G_{12} + G_{21} + G_{22}) = -i\langle[c(t), c^\dagger(t')]\rangle, \quad (19c)$$

$$G^-(t, \tau') = \frac{1}{2}(G_{13} + G_{23}) = i\langle c^\dagger(\tau')c(t)\rangle, \quad (19d)$$

$$G^-(\tau, t') = \frac{1}{2}(G_{31} + G_{32}) = -i\langle c(\tau)c^\dagger(t')\rangle, \quad (19e)$$

$$G^M(\tau, \tau') = -iG_{33} = -\langle T_\tau c(\tau)c^\dagger(\tau')\rangle. \quad (19f)$$

Furthermore, the so-called lesser and greater Green functions are defined as

$$G^<(t, t') = G_{12} = i\langle c^\dagger(t')c(t)\rangle, \quad G^>(t, t') = G_{21} = -i\langle c(t)c^\dagger(t')\rangle. \quad (20)$$

They fulfill $G^< = \frac{1}{2}(G^K - G^R + G^A)$ and $G^> = \frac{1}{2}(G^K + G^R - G^A)$. Hermitian conjugation entails the relations

$$G^{<,>,K}(t, t')^* = -G^{<,>,K}(t', t), \quad (21a)$$

$$G^R(t, t')^* = G^A(t', t), \quad (21b)$$

$$G^>(t, \tau)^* = G^>(\beta - \tau, t). \quad (21c)$$

The Matsubara Green function is translationally invariant, as \mathcal{H} does not depend on τ , i.e., $G^M(\tau, \tau') = G^M(\tau - \tau')$. It is also real and antiperiodic, $G^M(\tau) = -G^M(\tau + \beta)$, so it can be Fourier transformed to Matsubara frequencies as in equilibrium. Finally, the following boundary conditions hold because the trace is cyclic,

$$G(0^+, t) = -G(-i\beta, t), \quad G(t, 0^+) = -G(t, -i\beta). \quad (22)$$

Here and throughout, $G(t, t')$ (without superscript denoting a component) is a contour Green function for which the time arguments can be on any one of C_1, C_2, C_3 . By contrast, a Green function with superscript contains (real or imaginary) time arguments for which the contour part need not be specified. A convolution of contour functions $f * g$ involves contributions of several components. The procedures that separate the components of $f * g$ and express them in terms of the components of f and g are called Langreth rules [64]. They are discussed in detail, e.g., in Ref. [62].

2.3 Equilibrium case

For a time-independent Hamiltonian the formalism reduces of course to that for the equilibrium case [56], and only dependencies on time differences remain. In particular, the single-particle spectral function that characterizes the excitation spectrum is given in terms of Fourier transforms as

$$A(\omega) = -\frac{1}{\pi} \text{Im } G^R(\omega) = \frac{1}{\pi} \text{Im } G^A(\omega). \quad (23)$$

In equilibrium, all components of G can be recovered from it,

$$G(t, t') = -i \int d\omega e^{i\omega(t'-t)} A(\omega) [\theta_C(t, t') - f(\omega)], \quad (24)$$

where $f(\omega) = 1/(e^{\beta\omega} + 1)$ is the Fermi function.

In nonequilibrium, one can introduce average and relative times, $t_{\text{av}} = (t + t')/2$, $t_{\text{rel}} = t - t'$, in terms of which a partial Fourier transformation leads to the definition

$$A(\omega, t_{\text{av}}) = -\frac{1}{\pi} \text{Im} \int dt_{\text{rel}} e^{i\omega t_{\text{rel}}} G^R(t, t'), \quad (25)$$

which satisfies the sum rule $\int d\omega A(\omega, t_{\text{av}}) = 1$.

2.4 Noninteracting case

Consider now the simplest case of noninteracting fermions with a single time-dependent energy band, $\mathcal{H}_0(t) = \sum_{\mathbf{k}} [\varepsilon_{\mathbf{k}}(t) - \mu] c_{\mathbf{k}}^\dagger c_{\mathbf{k}}$. The time derivative of the corresponding noninteracting contour Green function $G_{0,\mathbf{k}}(t, t') = -i \langle T_C c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger(t') \rangle$ yields the equations of motion

$$[i\partial_t + \mu - \varepsilon_{\mathbf{k}}(t)] G_{0,\mathbf{k}}(t, t') = \delta_C(t, t'), \quad (26a)$$

$$G_{0,\mathbf{k}}(t, t') \left[-i \overleftarrow{\partial}_{t'} + \mu - \varepsilon_{\mathbf{k}}(t') \right] = \delta_C(t, t'). \quad (26b)$$

Here we let the derivative $\overleftarrow{\partial}_{t'}$ act to the left, i.e., $f(t) \overleftarrow{\partial}_t = \partial_t f(t)$, which makes the equations more symmetric. The inverse of $G_{0,\mathbf{k}}$ is then defined as the following differential operator,

$$G_{0,\mathbf{k}}^{-1}(t, t') = [i\partial_t + \mu - \varepsilon_{\mathbf{k}}(t)] \delta_C(t, t'). \quad (27)$$

The equations of motion thus correspond to the convolutions,

$$G_{0,\mathbf{k}}^{-1} * G_{0,\mathbf{k}} = G_{0,\mathbf{k}} * G_{0,\mathbf{k}}^{-1} = \delta_C. \quad (28)$$

Together with (22), either of these gives the unique solution [8],

$$G_{0,\mathbf{k}}(t, t') = -i[\theta_C(t, t') - f(\varepsilon_{\mathbf{k}}(0) - \mu)] e^{-i \int_{t'}^t d\bar{t} [\varepsilon_{\mathbf{k}}(\bar{t}) - \mu]}. \quad (29)$$

2.5 Self-energy

For an interacting Hamiltonian $\mathcal{H}(t)$, one usually has to resort to approximations to obtain the Green function G , e.g., by using perturbation expansions in terms of Feynman diagrams [62]. The self-energy $\Sigma(t, t')$ is then defined as a contour function (with boundary conditions as in (22)) in terms of one-particle irreducible diagrams, which are the same as those for finite-temperature equilibrium perturbation theory, but with the imaginary-time integrations in the diagram rules being replaced by time-contour integrations. The full Green function is then given in terms of the noninteracting Green function and self-energy insertions (on the contour), $G = G_0 + G_0 * \Sigma * G_0 + G_0 * \Sigma * G_0 * \Sigma * G_0 + \dots$. The Dyson equation therefore reads

$$G = G_0 + G_0 * \Sigma * G = G_0 + G * \Sigma * G_0. \quad (30)$$

After convoluting these equations with G_0^{-1} from either side this becomes

$$[G_0^{-1} - \Sigma] * G = G * [G_0^{-1} - \Sigma] = \delta_C. \quad (31)$$

This suggests the definition $G^{-1} = G_0^{-1} - \Sigma$, which is reminiscent of the equilibrium Dyson equation [56]. However, in the present case the equations for G , even if Σ is known, correspond to integral-differential equations along the contour. Their form will be discussed in more detail in the next section.

3 Nonequilibrium DMFT

The hallmark of DMFT, whether in equilibrium or nonequilibrium, is its formulation in terms of a dynamical effective single-site problem in the thermodynamic limit. The numerical solution of this single-site problem then provides the self-energy (and thus Green functions), often in a nonperturbative way. In this section, we show (following Ref. [26]) how the nonequilibrium DMFT equations are obtained in the limit of infinite lattice dimensions by means of the so-called cavity method, which was already employed in the equilibrium case in Ref. [2]. Alternatively, one can base the derivation on the identical diagrammatic skeleton expansions for the self-energies of the single-impurity Anderson model (SIAM) and the Hubbard model in infinite dimensions [9, 55].

3.1 Cavity method

We start from the Hubbard Hamiltonian (1) and, in the spirit of the cavity method, pick out one single site with the purpose of tracing out the remaining lattice. The action is thus split into

$$S = -i \int_C dt \mathcal{H}(t) = S_0 + \Delta S + S^{(0)}, \quad (32)$$

with

$$S_0 = -i \int_C dt \left[U(t) \left(n_{0\uparrow}(t) - \frac{1}{2} \right) \left(n_{0\downarrow}(t) - \frac{1}{2} \right) - \mu \sum_{\sigma} n_{0\sigma}(t) \right], \quad (33)$$

$$\begin{aligned} \Delta S &= -i \int_C dt \left[\sum_{i \neq 0, \sigma} t_{i0}^{\sigma}(t) c_{i\sigma}^{\dagger}(t) c_{0\sigma}(t) + \text{H.c.} \right], \\ S^{(0)} &= -i \int_C dt \mathcal{H}^{(0)}(t). \end{aligned} \quad (34)$$

Here $\mathcal{H}^{(0)}(t)$ is the Hamiltonian $\mathcal{H}(t)$ with the cavity site 0 removed. The effective single-site action for this site 0 will consist of S_0 for the local Hamiltonian at site 0 and a connection into the effective environment that comes from integrating out $\Delta S + S^{(0)}$. Splitting the trace for the states at site 0 and the other sites, the partition function takes the form

$$\begin{aligned} Z_S &= \text{Tr}_0 \left[\text{T}_C \left\{ \exp(S_0) \text{Tr}_{\text{rest}} \left(\exp(\Delta S + S^{(0)}) \right) \right\} \right] \\ &= Z_{S^{(0)}} \text{Tr}_0 \left[\text{T}_C \left\{ \exp(S_0 + \tilde{S}) \right\} \right], \quad Z_{S^{(0)}} = \text{Tr}_{\text{rest}} \left(\text{T}_C \left\{ \exp(S^{(0)}) \right\} \right). \end{aligned} \quad (35a)$$

Here \tilde{S} is the effective action connecting the cavity site and the environment:

$$\exp(\tilde{S}) = \sum_{n=0}^{\infty} \frac{1}{n!} \langle (\Delta S)^n \rangle_{S^{(0)}}, \quad \langle A(t) \rangle_{S^{(0)}} \equiv \frac{\text{Tr}_{\text{rest}} \left(\text{T}_C \left\{ \exp(S^{(0)}) A(t) \right\} \right)}{Z_{S^{(0)}}} \quad (36)$$

Note that here ΔS contains operators at site 0 that are not traced over; since they anticommute with those at other sites, the correct sign and time ordering must be kept when tracing over

the latter. From the definition of ΔS , we see that only terms with an equal number of $c_{i\sigma}^\dagger, c_{j\sigma}$ with $i, j \neq 0$ contribute, i.e., only terms with even powers of ΔS . After some combinatorial considerations and reorderings, the result is [26]

$$\exp(\tilde{S}) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} \langle (\Delta S)^{2n} \rangle_{S^{(0)}} = \sum_{n=0}^{\infty} \int_C dt_1 \cdots \int_C dt'_n \sum_{\substack{i_1, \dots, j_n \\ \sigma_1, \dots, \sigma'_n}} (-i)^n \times \frac{t_{0i_1}(t_1) \cdots t_{j_n 0}(t'_n)}{n!^2} G_{i_1 \sigma_1, \dots, j_n \sigma'_n}^{(0)}(t_1, \dots, t'_n) c_{0\sigma_1}^\dagger(t_1) \cdots c_{0\sigma'_n}^\dagger(t'_n), \quad (37)$$

with the n -particle contour-ordered Green function for the rest of the lattice (without site 0) defined as

$$G_{i_1 \sigma_1, \dots, j_n \sigma'_n}^{(0)}(t_1, \dots, t'_n) = (-i)^n \langle c_{i_1 \sigma_1}(t_1) \cdots c_{j_n \sigma'_n}^\dagger(t'_n) \rangle_{S^{(0)}}. \quad (38)$$

Next, the right-hand side of (37) must be re-exponentiated using connected (with respect to the interaction in $\mathcal{H}^{(0)}(t)$) contour-ordered Green functions $G^{(0),c}$ [26]. The result is

$$\tilde{S} = -i \sum_{n=1}^{\infty} \sum_{\sigma_1 \dots \sigma'_n} \int_C dt_1 \cdots \int_C dt'_n \Lambda_{\sigma_1 \dots \sigma'_n}(t_1, \dots, t'_n) c_{0\sigma_1}^\dagger(t_1) \cdots c_{0\sigma'_n}^\dagger(t'_n), \quad (39)$$

where we defined the n th-order hybridization functions

$$\Lambda_{\sigma_1 \dots \sigma'_n}(t_1, \dots, t'_n) \equiv \frac{(-i)^{n-1}}{n!^2} \sum_{i_1, \dots, j_n} t_{0i_1}(t_1) \cdots t_{j_n 0}(t'_n) G_{i_1 \sigma_1, \dots, j_n \sigma'_n}^{(0),c}(t_1, \dots, t'_n), \quad (40)$$

which involve connected cavity Green functions $G^{(0),c}$ that are obtained with $S^{(0)}$ only.

The effective action and its partition function, which only involve the degrees of freedom at the cavity site 0, are thus given by

$$S_{\text{eff}} = S_0 + \tilde{S}, \quad Z_{\text{eff}} = \frac{Z}{Z_{S^{(0)}}} = \text{Tr}_0 \left(\text{T}_C \{ \exp(S_{\text{eff}}) \} \right). \quad (41)$$

No approximation has been made yet, but of course the higher-order hybridization functions (40) are not easily accessible in general and also couple to correspondingly complicated cavity source terms. In this formulation, the limit of infinite dimensions lets only the hybridization functions with $n = 1$ contribute, leading to a quadratic coupling between cavity and environment, which we now discuss.

3.2 DMFT action for an infinite-dimensional lattice

As in equilibrium [2], the hybridization functions (40) simplify drastically in the limit $d \rightarrow \infty$. As a consequence of the quantum scaling [1] $t_{ij} \propto \mathcal{Z}_{ij}^{-\frac{1}{2}}$, where \mathcal{Z}_{ij} is the number of sites j connected to site i by hopping of type t_{ij} , only first-order terms (i.e., one-particle Green functions) contribute to the effective action. These power counting arguments are entirely analogous to

the equilibrium case. For the case of nearest-neighbor hopping on a hypercubic lattice, they proceed as follows.

The contributions to (40) from n th order Green functions contain lattice summations that yield a factor d^{2n} , $2n$ factors of hopping amplitudes $t_{0i} \propto d^{-1/2}$ giving d^{-n} , and a factor from the connected Green functions. The latter connects $2n$ nearest neighbors of 0, with the shortest path between them requiring 2 lattice steps, of which at least $2n - 1$ are needed, and in the best case (when all sites are different) this gives a factor $(\sqrt{d})^{2(2n-1)} = d^{2n-1}$. Hence

$$\Lambda_{\sigma_1 \dots \sigma_n}(t_1, \dots, t'_n) \propto \underbrace{\sum_{i_1, \dots, j_n} t_{0i_1}(t_1) \dots t_{j_n 0}(t'_n)}_{\propto d^{2n}} \underbrace{G_{(i_1 \sigma_1), \dots, (j_n \sigma_n)}^{(0),c}(t_1, \dots, t'_n)}_{\propto (\sqrt{d})^{-2(2n-1)}} \propto \frac{1}{d^{n-1}}, \quad (42)$$

so that for $d \rightarrow \infty$ only the quadratic term $n = 1$ survives. Furthermore, Λ is spin-diagonal as (1) does not contain spin-flip terms. Dropping the index 0 of the cavity site, the DMFT action is therefore

$$S_{\text{loc}} = -i \int_C dt \left[U(t) \left(n_\uparrow(t) - \frac{1}{2} \right) \left(n_\downarrow(t) - \frac{1}{2} \right) - \mu \sum_\sigma n_\sigma(t) \right] \\ - i \int_C dt_1 \int_C dt_2 \sum_\sigma \Lambda_\sigma(t_1, t_2) c_\sigma^\dagger(t_1) c_\sigma(t_2).$$

(43)

It remains to determine the hybridization Λ defined as

$$\Lambda_\sigma(t, t') = \sum_{i,j} t_{0i}(t) G_{ij\sigma}^{(0),c}(t, t') t_{j0}(t'), \quad (44)$$

such that the action indeed describes the original interacting lattice system. This requires linking the Green function $G_{ij\sigma}^{(0),c}$ to local quantities (using the self-energy, which turns out to be local). Physically, the hybridization characterizes the “dynamical mean-field”, i.e., the effective host into and out of which the particles on the impurity site can move.

Comparing the nonequilibrium DMFT action (43) with the equilibrium case, we note of course the appearance of time-contour integrals and ordering. Furthermore, the action is not time-translationally invariant because Λ depends explicitly on t and t' . This complicates the task of numerically obtaining G from S for a given Λ .

3.3 Local self-energy

The local nature of the self-energy in infinite dimensions can be obtained from the cavity method itself, as we now describe. It enters into the lattice and impurity Dyson equations, which determine the corresponding lattice and impurity Green functions. In general, the self-energy is also needed to obtain the self-consistency relation for Λ . For brevity, we now drop the spin indices. The hopping $t_{ij}(t)$ is again arbitrary.

We consider the full lattice Green function $G_{ij}(t, t') = -i \langle c_i(t) c_j^\dagger(t') \rangle_S$ on the one hand and the impurity Green function $G(t, t') = G_{00}(t, t')$ on the other. Their inverses are given by $G_{\text{lat}}^{-1}(t, t')$

and $G^{-1}(t, t')$ respectively,

$$\sum_l \int_C dt_1 (G_{\text{lat}}^{-1})_{il}(t, t_1) G_{lj}(t_1, t') = \delta_{ij} \delta_C(t, t') , \quad (45a)$$

$$\int_C dt_1 G^{-1}(t, t_1) G(t_1, t') = \delta_C(t, t') . \quad (45b)$$

The corresponding impurity and lattice self-energies are then determined by the relations

$$(G_{\text{lat}}^{-1})_{ij}(t, t') = [\delta_{ij}(i\partial_t + \mu) - t_{ij}(t)] \delta_C(t, t') - (\Sigma_{\text{lat}})_{ij}(t, t') , \quad (46)$$

$$G^{-1}(t, t') = (i\partial_t + \mu) \delta_C(t, t') - \Lambda(t, t') - \Sigma(t, t') . \quad (47)$$

Taking functional derivatives of $G_{ij}(t, t')$ with respect to the annihilation and creation operators at site 0 (and hopping matrix elements that connect them) provides the relations [26]

$$G_{ij}(t, t') = G_{ij}^{(0),c}(t, t') + \int_C dt_1 \int_C dt_2 \sum_{lm} G_{il}^{(0),c}(t, t_1) t_{l0}(t_1) G(t_1, t_2) t_{0m}(t_2) G_{mj}^{(0),c}(t_2, t') . \quad (48)$$

$$G_{0j}(t, t') = \int_C dt_1 \sum_i G(t, t_1) t_{0i}(t_1) G_{ij}^{(0),c}(t_1, t') , \quad j \neq 0 . \quad (49)$$

Putting (49) into (48) then gives us

$$G_{ij}^{(0),c}(t, t') = G_{ij}(t, t') - \int_C dt_1 \int_C dt_2 G_{i0}(t, t_1) G^{-1}(t_1, t_2) G_{0j}(t_2, t') , \quad (50)$$

which we recognize as an analogue of the relation in equilibrium (i.e., Eq. (36) in Ref. [2]): $G_{ij}^{(0),c}(i\omega_n) = G_{ij}(i\omega_n) - G_{i0}(i\omega_n) G_{0j}(i\omega_n)/G(i\omega_n)$. Furthermore, a conjugated equation similar to (49) can be derived for $G_{i0}(t, t')$ and summed,

$$\sum_i t_{0i}(t) G_{i0}(t, t') = \int_C dt_1 \Lambda(t, t_1) G(t_1, t') , \quad (51)$$

while (49) itself can be rewritten as

$$\int_C dt_1 G^{-1}(t, t_1) G_{0j}(t_1, t') = \sum_i t_{0i}(t) G_{ij}^{(0),c}(t, t') , \quad j \neq 0 . \quad (52)$$

Summing over (50), these equations can be used to obtain

$$\sum_i t_{0i}(t) G_{ij}(t, t') = \int_C dt_1 [G^{-1}(t, t_1) + \Lambda(t, t_1)] G_{0j}(t_1, t') , \quad j \neq 0 . \quad (53)$$

Finally, we decompose

$$\begin{aligned} \delta_{0j} \delta_C(t, t') &= \sum_l \int_C dt_1 (G_{\text{lat}}^{-1})_{0l}(t, t_1) G_{lj}(t_1, t') \\ &= (i\partial_t + \mu) G_{0j}(t, t') - \sum_l t_{0l}(t) G_{lj}(t, t') - \int_C dt_1 \sum_l (\Sigma_{\text{lat}})_{0l}(t, t_1) G_{lj}(t_1, t') , \end{aligned} \quad (54)$$

and use (51) and (53) to arrive at

$$\int_C dt_1 \Sigma(t, t_1) G_{0j}(t_1, t') = \int_C dt_1 \sum_l (\Sigma_{\text{lat}})_{0l}(t, t_1) G_{lj}(t_1, t'). \quad (55)$$

Here the cavity site 0 can be replaced by an arbitrary i (for a translationally invariant system),

$$\int_C dt_1 \Sigma(t, t_1) G_{ij}(t_1, t') = \int_C dt_1 \sum_l (\Sigma_{\text{lat}})_{il}(t, t_1) G_{lj}(t_1, t'). \quad (56)$$

Acting on this equation with the inverse of the lattice Green function from the right shows that indeed

$$(\Sigma_{\text{lat}})_{ij}(t, t') = \delta_{ij} \Sigma(t, t') \quad (57)$$

i.e., the lattice self-energy is local in the limit of infinite dimensions and given by the impurity self-energy.

3.4 Self-consistency condition

The DMFT self-consistency condition then corresponds to the Dyson equations for the lattice and impurity Green function. For the lattice Green function it is given by

$$\int_C dt_1 \sum_l \left[[\delta_{il} (i\partial_t + \mu) - t_{il}(t)] \delta_C(t, t_1) - \Sigma(t, t_1) \right] G_{lj}(t_1, t') = \delta_{ij} \delta_C(t, t'), \quad (58)$$

while for the impurity Green function,

$$\int_C dt_1 \left[(i\partial_t + \mu) \delta_C(t, t_1) - \Lambda(t, t_1) - \Sigma(t, t_1) \right] G(t_1, t') = \delta_C(t, t'). \quad (59)$$

Suppose G has been obtained for given Λ . Then in principle a new Σ can be found from (59) and a new G follows from (58), and then again a new Λ from (59). For the actual strategies regarding the nontrivial numerical solution of the Dyson equations (including Fourier transformation to momentum space) we refer to Refs. [9, 33, 65] as well as [55] and references therein.

3.5 Bethe lattice

As in equilibrium, (44) can directly be evaluated for nearest-neighbor hopping on a Bethe lattice with $\mathcal{Z} \rightarrow \infty$ nearest neighbors, $t_{ij} = v/\sqrt{\mathcal{Z}}$, and semielliptic density of states,

$$\rho(\varepsilon) = \frac{\sqrt{4v^2 - \varepsilon^2}}{2\pi v^2}. \quad (60)$$

For neighboring sites i, j of the cavity site 0, $G_{ij}^{(0),c}(t, t')$ is nonzero only for $i = j$, since there is no path from i to j other than through the removed cavity site 0. Furthermore, for $\mathcal{Z} \rightarrow \infty$ we have $G_{i\sigma}^{(0),c}(t, t') = G_\sigma(t, t')$. The quantum scaling ensures that the summation over all nearest neighbors of 0 stays finite. This yields the action (43) with the hybridization [66, 24],

$$\Lambda_\sigma(t, t') = v(t) G_\sigma(t, t') v(t'), \quad (61)$$

i.e., after obtaining a new G from S for given hybridization Λ , the new Λ can be obtained at once. The self-energy is nevertheless needed to calculate lattice quantities (such as the lattice Green function or momentum distribution).

3.6 Numerical methods

Several methods have been developed to calculate the contour Green function $G(t, t')$ from the single-site DMFT action (43) for given hybridization $\Lambda(t, t')$ [55]. Generally speaking, these methods are more involved as contour functions on the different branches must be obtained, and not only the initial many-body state (3) must be represented but also its time evolution during which small errors may grow substantially. The methods in the following list all have parameter regimes for which they are well-controlled and therefore accurate, and checking them against each other provides an important benchmark.

Many-body perturbation theory

Diagrammatic perturbation theory [62] is limited to either sufficiently small or large interaction. For weak coupling the self-energy is expanded in terms of Feynman diagrams to a certain finite order in U ; see, e.g., Refs. [11, 67, 21, 34, 68, 38, 44, 36, 46]. The Green function lines can be taken as bare or interacting Green functions (as in the equilibrium case [2]), although it is not *a priori* clear which choice is more accurate for a given problem. Strong-coupling perturbation theory is based on a representation in terms of auxiliary particles or an expansion in the hybridization [21, 55]; see, e.g., Refs. [33, 40, 45, 41, 50, 48, 47, 49]. For both small and strong coupling the perturbation expansions are asymptotic in the sense that they will be more accurate if the controlling parameter is smaller. For long times, however, it is never *a priori* clear up to which time they will remain accurate.

Continuous-time quantum Monte Carlo (CT-QMC)

CT-QMC [69–71] also comes in two versions, an interaction and hybridization expansion. Feynman diagrams are generated and sampled stochastically with appropriate weights. In addition to the fermionic sign problem, the imaginary exponents that appear in the contour Green function lead to a dynamic sign problem, so that only comparatively short times can be studied reliably (see, Refs. [20, 21] and Sec. 4). Also, for finite temperatures, an initial thermal state becomes more costly to obtain at low temperatures.

Hamiltonian-based methods

A single-impurity Anderson model (SIAM) with time-dependent couplings and bath energies yields the same nonequilibrium action as (43) with a specific hybridization $\Lambda(t, t')$ upon integrating out the bath. In order to use such a time-dependent SIAM representation this hybridization function must be matched with that obtained from the DMFT self-consistency condition (e.g. (61)), as discussed in detail in [26]. Then, the time evolution of the SIAM Hamiltonian may be obtained by exact diagonalization methods. In general, the accessible time is limited because many bath sites are needed to represent both the correlations in the initial state (which typically decrease with time) and the build-up of correlations in the time-evolved state [26].

Falicov-Kimball model

In the Falicov-Kimball model [72], only one electron species hops between lattice sites while the other is immobile. As in the equilibrium case, it has played an important role for the development of nonequilibrium DMFT [9–11, 13, 15, 18, 19, 29] because it reduces to a quadratic action that is partially solvable. However, its nonequilibrium (and equilibrium) properties are quite different from that of the Hubbard model.

4 Correlated electrons in nonequilibrium

As mentioned in Sec. 1, many aspects of correlated electrons in nonequilibrium have been studied with DMFT. Here we discuss only one of the simplest situations, namely an abrupt change in the Hubbard interaction U .

4.1 Relaxation and thermalization

The evolution of a quantum many-body system in real time raises interesting questions about the connection to equilibrium statistical mechanics. Suppose that an isolated system undergoes some experimental protocol with a Hamiltonian $H(t)$ that no longer changes after a certain time t_1 . How does the system behave at large times (during which $H(t \geq t_1) = \text{const}$)? Does it relax to the equilibrium state that is predicted by statistical mechanics for this Hamiltonian $H(t_1)$ for the average energy $E = \text{Tr}[\rho(t_1)H(t_1)]$? If it does, the system is said to *thermalize*. However, the density matrix $\rho(t)$ in (4), when regarded in the eigenbasis of $H(t_1)$, will contain many oscillating components that by themselves will in general not converge. Rather, expectation values (such as Green functions or other short-range correlation functions) will relax to stationary values because they average over many states and degrees of freedom.

In general, the coupling of the system to an environment is needed to prepare a mixed state such as the canonical or grand-canonical Gibbs ensemble (3). This is used in the usual derivation of the Gibbs state in statistical mechanics, based on Boltzmann’s concept of entropy $S = k_B \ln \Omega$. The system is coupled to a much larger external thermostat, and the number of available states for the system at energy E is then proportional to the number Ω of microstates of the thermostat at energy $\mathcal{E} - E$, where \mathcal{E} is the fixed total energy of system plus thermostat. Expanding Ω in the vicinity of \mathcal{E} one finds that the probability for the system to have energy E is proportional to $e^{-E/(k_B T)}$, where $T = \partial S / \partial \mathcal{E}$ is the temperature of the thermostat.

On the other hand, during the time evolution of an isolated system there is no environment to assist with the thermalization, i.e., the system must in some sense act as its own environment. *A priori* it is not obvious how the details of the initial state (or $\rho(t_1)$ in the above example) should be irrelevant in the long-time limit so that only a dependence on the average energy remains. The so-called eigenstate thermalization hypothesis [73–76] proposes that this is due to the fact that the expectation value of an observable A in an energy eigenstates with energy E_n usually depends only on the eigenenergy E_n and not on the details of the eigenstate $|n\rangle$. This property can be observed for many generic many-body systems and short-range observables A ,

although it is difficult to give precise criteria for its validity. From this property it follows at once that, after relaxation, thermal expectation values are attained for such observables. Integrable systems, on the other hand, are characterized by a large number of conserved quantities. These lead to a dependence of expectation values not only on E_n but also on the individual eigenstates $|n\rangle$. As a consequence, integrable systems usually do not thermalize, a behavior that has also been observed experimentally with cold atoms [77]. Nevertheless, a statistical prediction can often be made using generalized Gibbs ensembles (GGEs), which take the conserved constants of motion (in addition to the Hamiltonian) into account [78, 79, 75]; for reviews see [76, 80]. In general, however, there is still much debate how to even define thermalization or integrability properly for quantum many-body systems in general.

4.2 Interaction quench in the Hubbard model

One of the simplest situations that can be studied in this context is a so-called *quench*, i.e., a sudden switch of Hamiltonians. Here the quench is performed in the Hubbard model (1) at half-filling in the paramagnetic phase with semielliptic density of states (60) with bandwidth $4v \equiv 4$. The system is prepared in the zero-temperature ground state of the noninteracting Hamiltonian, i.e., $U(t < 0) = 0$. At $t = 0$ the Hubbard interaction is switched to a finite value, $U(t \geq 0) = U$. The Green function is obtained with CT-QMC (weak-coupling expansion) from the action (43), and the selfconsistency condition (61) applies [20, 21]. The noninteracting initial state makes things simpler because the imaginary branch of the contour does not enter the CT-QMC calculation.

In Fig. 2 the momentum distribution $n(\varepsilon_{\mathbf{k}}, t) = \langle c_{\mathbf{k}\sigma}^\dagger(t)c_{\mathbf{k}\sigma} \rangle$ is plotted as a function of the band energy $\varepsilon \equiv \varepsilon_{\mathbf{k}}$ for different final values of U . The initial Fermi sea evolves from a step function into a continuous function of ε . Fig. 3 shows the jump in the momentum distribution at the Fermi surface and the double occupation as a function of time. Three different parameter regimes can be observed: small and large values of U , separated by a sharp crossover or transition near the intermediate scale $U \approx 3.2 = U_c^{\text{dyn}}$. Near U_c^{dyn} , the momentum distribution relaxes quickly to the thermal distribution for all energies ε (solid blue line in Fig. 2b, obtained from a grand-canonical DMFT equilibrium calculation for the temperature that gives the same total energy E). Relaxation to thermal values is also found for dynamical observables like the retarded Green function $G^R(t+s, t)$ (as a function of time difference s) and the two-time optical conductivity $\sigma(t, t+s)$ [21]. For quenches close to U_c , the system hence thermalizes on short timescales.

For quenches to small or large values of U (away from U_c), thermalization is nevertheless expected on general grounds but cannot be observed on the short time scales that are available with CT-QMC. Away from U_c , the relaxation does not reach a thermal state quickly but instead passes through metastable states on intermediate time scales. For quenches to weak coupling, $U \leq 3$, the double occupation $d(t)$ relaxes from its initial uncorrelated value $d(0) = 1/4$ almost to its thermal value d_{th} , whereas the Fermi surface discontinuity $\Delta n(t)$ remains finite for $t \leq 5$. This behavior is called *prethermalization* and was predicted for a quenched

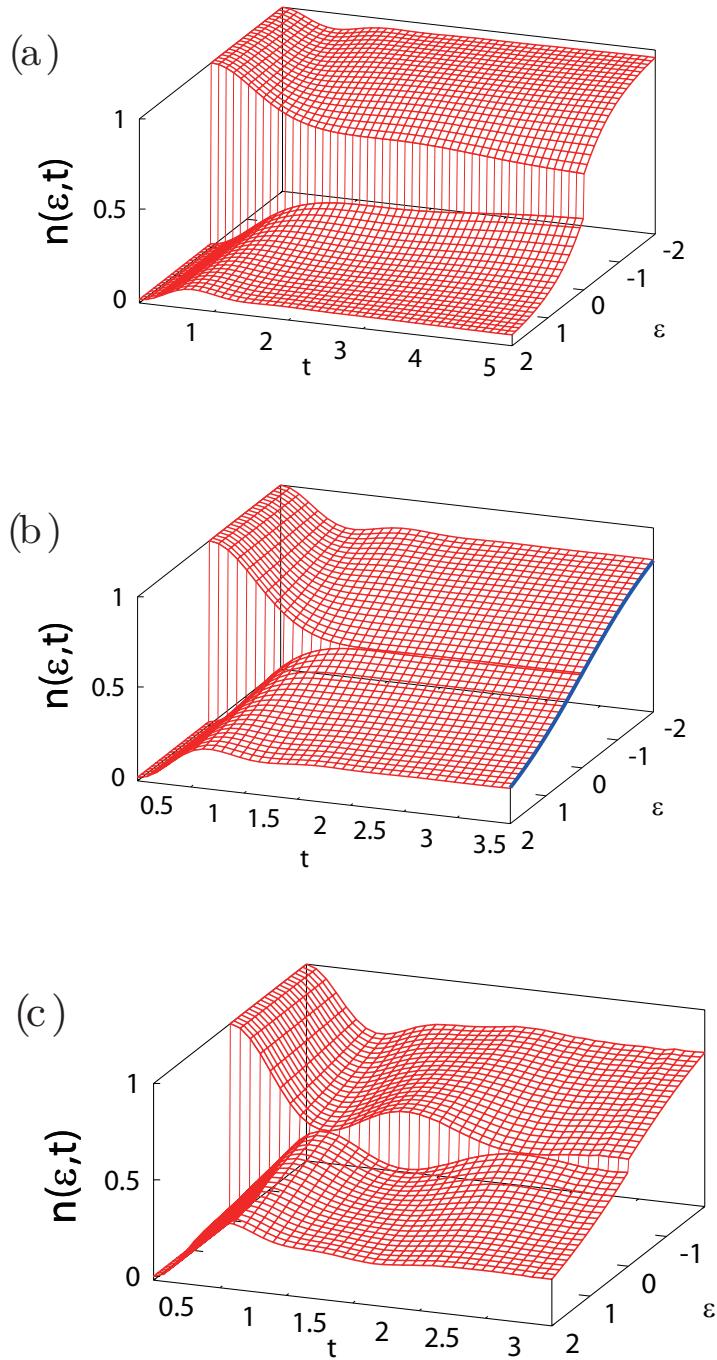


Fig. 2: Momentum distribution $n(\varepsilon, t)$ after an interaction quench in the Hubbard model with bandwidth 4 in DMFT [20, 21], starting from the noninteracting ground state ($U = 0$) to interaction (a) $U = 2$, (b) $U = 3.3$, (c) $U = 5$. The blue line in (b) is the equilibrium expectation value for the momentum distribution at the same total energy (temperature $T = 0.84$) as the time-evolved state.

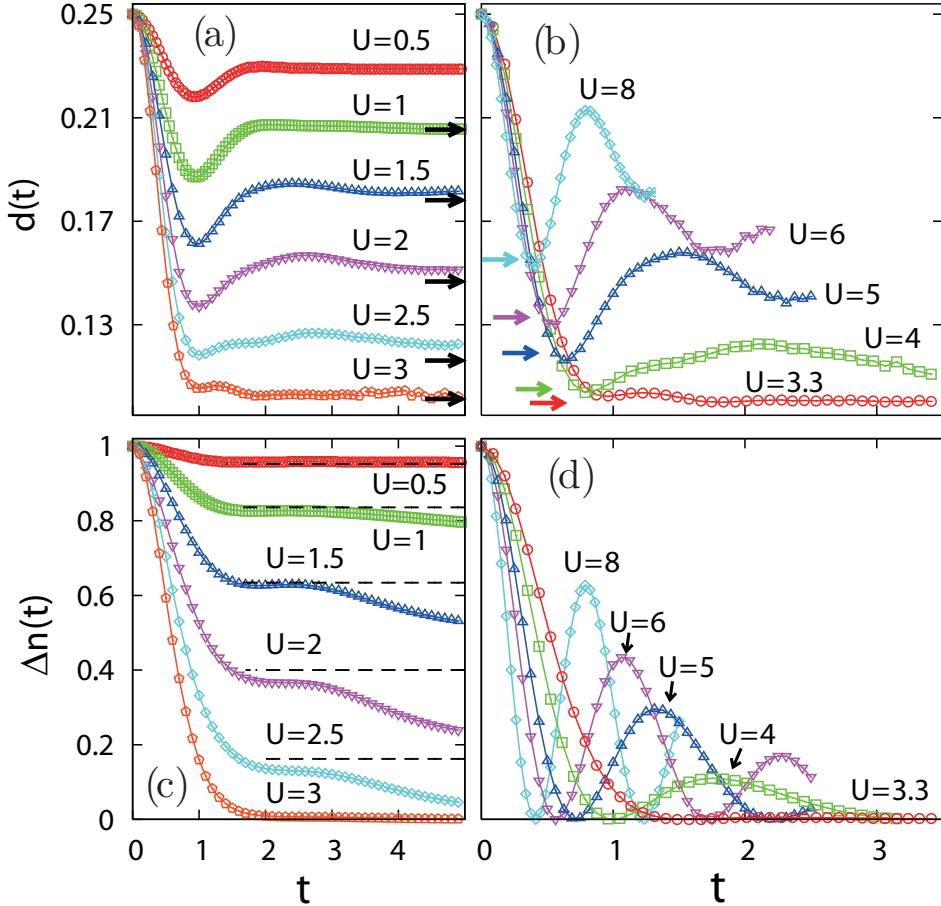


Fig. 3: Double occupation $d(t)$ and Fermi surface discontinuity Δn after interaction quenches to $U \leq 3$ (left panels) and $U \geq 3.5$ (right panels) [20]. Horizontal arrows: thermal values of the double occupation. Horizontal dashed lines in the lower left panel are the expected prethermalization plateaus [81].

Fermi liquid [81] on the basis of a weak-coupling calculation. Characteristically, the kinetic and interaction energy thermalize on time scales $1/U^2$ while the Fermi surface discontinuity only reaches a plateau that is located $\Delta n_{\text{stat}} = 1 - 2Z$, where Z is the quasiparticle weight in equilibrium at zero temperature. During this early phase the quasiparticles are formed, and during their subsequent scattering the momentum occupations are further redistributed. The weak-coupling result for the transient [81] towards the prethermalization plateau describes the DMFT data well for $U \lesssim 1.5$ [20], even though at the larger U values the timescales $1/U^2$ and $1/U^4$ are no longer well separated. A weakly interacting system may be regarded as nearly integrable, and indeed prethermalization plateaus after an interaction quench are quite generally predicted correctly by a generalized Gibbs ensemble that is built from approximate constants of motion [82]. Physically, the subsequent crossover from the prethermalization plateau to the thermal state is expected due to the scattering of quasiparticles, which can be described by a kinetic equation [83]. Note also that a short-time prethermalization regime has been observed for interaction quenches in the one- and two-dimensional Hubbard model [84, 85, 52], albeit with a less pronounced plateau in the momentum distribution.

For quenches to strong coupling ($U \geq 3.3$ in Fig. 3b,d), the relaxation shows so-called ‘collapse-and-revival’ oscillations with the approximate periodicity $2\pi/U$, which are due to the exact periodicity of the propagator e^{-iHt} without hopping [6]. For large values of U , both $d(t)$ and $n(\varepsilon, t)$ oscillate around nonthermal values. Strong-coupling perturbation theory [20] shows that the mean value of $d(t)$ for these oscillations is $d_{\text{stat}} = d(0) - \Delta d$ with $\Delta d = (1/2U)\langle H_{\text{kin}}/L \rangle_{t=0}$. By contrast, the thermal value is obtained as $d_{\text{th}} = d(0) + (1/U)\langle H_{\text{kin}}/L \rangle_0$ from a high-temperature expansion. Hence, during the initial stage of the relaxation the double occupation relaxes only halfway towards d_{th} . Although longer times cannot be accessed with the weak-coupling CT-QMC method, a relaxation to the thermal state is expected after the oscillations have decayed, as in the case of a pump-excited Mott insulator [33]. In general, this crossover will set in only on times scales that are exponentially large in the interaction U [86].

The rapid thermalization at $U \approx U_c^{\text{dyn}}$ occurs at the border between the delayed thermalization either due to weak-coupling prethermalization plateaus or strong-coupling oscillations around nonthermal values. Indeed, no finite width was detected for the width of this crossover region, so that the behavior at U_c^{dyn} might signal a dynamical phase transition. A similarly strong dependence on the quenched interaction was observed in Heisenberg chains [87] and the one-dimensional Hubbard model [84]. Several possible origins for nonequilibrium phase transitions of this type have been proposed [88–91]. For the DMFT data, the corresponding equilibrium temperature T_{eff} after the quench is much higher than the critical endpoint of the Mott metal-insulator transition in equilibrium ($T_c \approx 0.055$ [2], but $T_{\text{eff}} = 0.84$ for $U = 3.3$). Interestingly, a good approximation for the critical interaction, $U_c^{\text{dyn}} \approx 3.4$, is obtained from a time-dependent variational theory using the Gutzwiller approximation [92, 93]. Variational results also suggest that the difference in U_c for equilibrium and nonequilibrium is due to the rapid change in interaction and corresponding high excitation energy: if the Hubbard interaction is instead changed very slowly, this lessens the effective temperature and the dynamical critical value U_c approaches that of the equilibrium transition [94].

Acknowledgment

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

References

- [1] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [2] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [3] A.L. Cavalieri, N. Müller, T. Uphues, V.S. Yakovlev, A. Baltuška, B. Horvath, B. Schmidt, L. Blümel, R. Holzwarth, S. Hendel, M. Drescher, U. Kleineberg, P.M. Echenique, R. Kienberger, F. Krausz, and U. Heinzmann, Nature **449**, 1029 (2007)
- [4] S. Wall, D. Brida, S.R. Clark, H.P. Ehrke, D. Jaksch, A. Ardavan, S. Bonora, H. Uemura, Y. Takahashi, T. Hasegawa, H. Okamoto, G. Cerullo, and A. Cavalleri, Nature Phys. **7**, 114 (2011)
- [5] R. Peierls, Z. Physik **80**, 763 (1933)
- [6] M. Greiner, O. Mandel, T.W. Hänsch, and I. Bloch, Nature **419**, 51 (2002)
- [7] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008)
- [8] V. Turkowski and J.K. Freericks, Phys. Rev. B **71**, 085104 (2005)
- [9] J.K. Freericks, V.M. Turkowski, and V. Zlatić, Phys. Rev. Lett. **97**, 266408 (2006)
- [10] J.K. Freericks, Phys. Rev. B **77**, 075109 (2008)
- [11] V. Turkowski and J.K. Freericks, Phys. Rev. B **75**, 125110 (2007)
- [12] V. Turkowski and J.K. Freericks, Phys. Rev. B **77**, 205102 (2008);
Erratum: *ibid.* **82**, 119904 (2010)
- [13] M.-T. Tran, Phys. Rev. B **78**, 125103 (2008)
- [14] A.V. Joura, J.K. Freericks, and T. Pruschke, Phys. Rev. Lett. **101**, 196401 (2008)
- [15] M. Eckstein and M. Kollar, Phys. Rev. Lett. **100**, 120404 (2008)
- [16] P. Schmidt and H. Monien: *Nonequilibrium dynamical mean-field theory of a strongly correlated system*. Preprint arXiv:cond-mat/0202046
- [17] J.K. Freericks, H.R. Krishnamurthy, and T. Pruschke, Phys. Rev. Lett. **102**, 136401 (2009)
- [18] M. Eckstein and M. Kollar, Phys. Rev. B **78**, 205119 (2008)
- [19] M. Eckstein and M. Kollar, Phys. Rev. B **78**, 245113 (2008)
- [20] M. Eckstein, M. Kollar, and P. Werner, Phys. Rev. Lett. **103**, 056403 (2009)
- [21] M. Eckstein, M. Kollar, and P. Werner, Phys. Rev. B **81**, 115131 (2010)

- [22] M. Eckstein and P. Werner, Phys. Rev. B **82**, 115115 (2010)
- [23] P. Werner and M. Eckstein, Phys. Rev. B **86**, 045119 (2012)
- [24] M. Eckstein and M. Kollar, New J. Phys. **12**, 055012 (2010)
- [25] N. Eurich, M. Eckstein, and P. Werner, Phys. Rev. B **83**, 155122 (2011)
- [26] C. Gramsch, K. Balzer, M. Eckstein, and M. Kollar, Phys. Rev. B **88**, 235106 (2013)
- [27] A. Lubatsch and J. Kroha, Ann. Phys. **18**, 863 (2009)
- [28] N. Tsuji, T. Oka, and H. Aoki, Phys. Rev. B **78**, 235124 (2008)
- [29] N. Tsuji, T. Oka, and H. Aoki, Phys. Rev. Lett. **103**, 047403 (2009)
- [30] B. Moritz, T.P. Devereaux, and J.K. Freericks, Phys. Rev. B **81**, 165112 (2010)
- [31] B. Moritz, A.F. Kemper, M. Sentef, T.P. Devereaux, and J.K. Freericks, Phys. Rev. Lett. **111**, 077401 (2013)
- [32] M. Eckstein, T. Oka, and P. Werner, Phys. Rev. Lett. **105**, 146404 (2010)
- [33] M. Eckstein and P. Werner, Phys. Rev. B **84**, 035122 (2011)
- [34] M. Eckstein and P. Werner, Phys. Rev. Lett. **107**, 186406 (2011)
- [35] N. Tsuji, T. Oka, P. Werner, and H. Aoki, Phys. Rev. Lett. **106**, 236401 (2011)
- [36] N. Tsuji, T. Oka, H. Aoki, and P. Werner, Phys. Rev. B **85**, 155124 (2012)
- [37] K. Mikelsons, J.K. Freericks, and H.R. Krishnamurthy, Phys. Rev. Lett. **109**, 260402 (2012)
- [38] A. Amaricci, C. Weber, M. Capone, and G. Kotliar, Phys. Rev. B **86**, 085110 (2012)
- [39] C. Aron, Phys. Rev. B **86**, 085127 (2012)
- [40] M. Eckstein and P. Werner, Phys. Rev. Lett. **110**, 126401 (2013)
- [41] M. Eckstein and P. Werner, J. Phys.: Conf. Series **427**, 012005 (2012)
- [42] C. Aron, C. Weber, and G. Kotliar, Phys. Rev. B **87**, 125113 (2013)
- [43] J. Mentink and M. Eckstein: *Ultrafast quenching of the exchange interaction in a Mott insulator*. Preprint arXiv:1401.5308
- [44] N. Tsuji, M. Eckstein, and P. Werner, Phys. Rev. Lett. **110**, 136404 (2013)
- [45] P. Werner, N. Tsuji, and M. Eckstein, Phys. Rev. B **86**, 205101 (2012)

- [46] N. Tsuji and P. Werner, Phys. Rev. B **88**, 165115 (2013)
- [47] M. Eckstein and P. Werner: *Ultrafast separation of photo-doped carriers in Mott antiferromagnets*. Preprint arXiv:1403.1461
- [48] P. Werner and M. Eckstein, Phys. Rev. B **88**, 165108 (2013)
- [49] P. Werner and M. Eckstein: *Field-induced polaron formation in the Holstein-Hubbard model*. Preprint arXiv:1403.7376
- [50] M. Eckstein and P. Werner, Phys. Rev. B **88**, 075135 (2013)
- [51] C. Jung, A. Lieder, S. Brener, H. Hafermann, B. Baxevanis, A. Chudnovskiy, A. Rubtsov, M. Katsnelson, and A. Lichtenstein, Ann. Phys. (Berlin) **524**, 49 (2012)
- [52] N. Tsuji, P. Barmettler, H. Aoki, and P. Werner: *Nonequilibrium dynamical cluster theory*. Preprint arXiv:1307.5946
- [53] F. Hofmann, M. Eckstein, E. Arrigoni, and M. Potthoff, Phys. Rev. B **88**, 165124 (2013)
- [54] H.U.R. Strand, M. Eckstein, and P. Werner: *Nonequilibrium dynamical mean-field theory for bosonic lattice models*. Preprint arXiv:1405.6941
- [55] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Rev. Mod. Phys. **86**, 779 (2014)
- [56] A.L. Fetter and J.D. Walecka: *Quantum Theory of Many-Particle Systems* (Dover, New York, 2003)
- [57] J. Schwinger, J. Math. Phys. **2**, 407 (1961)
- [58] L.P. Kadanoff and G. Baym: *Quantum Statistical Mechanics* (W.A. Benjamin, New York, 1962)
- [59] L.V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)]
- [60] A. Kamenev: *Field Theory of Non-Equilibrium Systems* (Cambridge Univ. Press, 2011)
- [61] R. van Leeuwen, N.E. Dahlen, G. Stefanucci, C.-O. Almlödah, and U. von Barth: *Introduction to the Keldysh formalism*, Lecture Notes in Physics, Vol. 706 (Springer, Heidelberg, 2006)
- [62] G. Stefanucci and R. van Leeuwen: *Non-Equilibrium Many-Body Theory of Quantum Systems* (Cambridge University Press, 2013)
- [63] R. van Leeuwen and G. Stefanucci, J. Phys.: Conf. Series **427**, 012001 (2013)
- [64] D.C. Langreth in *Linear and Nonlinear Electron Transport in Solids* ed. by J.T. Devreese and V.E. van Doren (Plenum Press, New York, 1976)

- [65] K. Balzer and M. Eckstein, Phys. Rev. B **89**, 035148 (2014)
- [66] M. Eckstein, A. Hackl, S. Kehrein, M. Kollar, M. Moeckel, P. Werner, and F.A. Wolf, Eur. Phys. J. Special Topics **180**, 217 (2010)
- [67] R.J. Heary and J.E. Han, Phys. Rev. B **80**, 035102 (2009)
- [68] C. Aron, G. Kotliar, and C. Weber, Phys. Rev. Lett. **108**, 086401 (2012)
- [69] L. Mühlbacher and E. Rabani, Phys. Rev. Lett. **100**, 176403 (2008)
- [70] P. Werner, T. Oka, and A.J. Millis, Phys. Rev. B **79**, 035320 (2009)
- [71] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011)
- [72] J.K. Freericks and V. Zlatić, Rev. Mod. Phys. **75**, 1333 (2003)
- [73] M. Srednicki, Phys. Rev. E **50**, 888 (1994)
- [74] J.M. Deutsch, Phys. Rev. A **43**, 2046 (1991)
- [75] M. Rigol, V. Dunjko, and M. Olshanii, Nature **452**, 854 (2008)
- [76] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, Rev. Mod. Phys. **83**, 863 (2011)
- [77] T. Kinoshita, T. Wenger, and D.S. Weiss, Nature **440**, 900 (2006)
- [78] E.T. Jaynes, Phys. Rev. **106**, 620 (1957)
- [79] E.T. Jaynes, Phys. Rev. **108**, 171 (1957)
- [80] J. Dziarmaga, Adv. Phys. **59**, 1063 (2010)
- [81] M. Moeckel and S. Kehrein, Phys. Rev. Lett. **100**, 175702 (2008)
- [82] M. Kollar, F.A. Wolf, and M. Eckstein, Phys. Rev. B **84**, 054304 (2011)
- [83] M. Stark and M. Kollar: *Kinetic description of thermalization dynamics in weakly interacting quantum systems*. Preprint arXiv:1308.1610
- [84] S.A. Hamerla and G.S. Uhrig, Phys. Rev. B **87**, 064304 (2013)
- [85] S.A. Hamerla and G.S. Uhrig, Phys. Rev. B **89**, 104301 (2014)
- [86] R. Sensarma, D. Pekker, E. Altman, E. Demler, N. Strohmaier, D. Greif, R. Jördens, L. Tarruell, H. Moritz, and T. Esslinger, Phys. Rev. B **82**, 224302 (2010)

- [87] P. Barmettler, M. Punk, V. Gritsev, E. Demler, and E. Altman, Phys. Rev. Lett. **102**, 130603 (2009)
- [88] B. Sciolla and G. Biroli, Phys. Rev. Lett. **105**, 220401 (2010)
- [89] A. Gambassi and P. Calabrese, EPL **95**, 66007 (2011)
- [90] M. Heyl, A. Polkovnikov, and S. Kehrein, Phys. Rev. Lett. **110**, 135704 (2013)
- [91] C. Karrasch and D. Schuricht, Phys. Rev. B **87**, 195104 (2013)
- [92] M. Schiró and M. Fabrizio, Phys. Rev. Lett. **105**, 076401 (2010)
- [93] M. Schiró and M. Fabrizio, Phys. Rev. B **83**, 165105 (2011)
- [94] M. Sandri, M. Schiró, and M. Fabrizio, Phys. Rev. B **86**, 075122 (2012)

13 Theoretical Description of ARPES: The One-Step Model

Ján Minár

Department Chemie, Physikalische Chemie
Ludwig-Maximilians-Universität München
New Technologies-Research Center
University of West Bohemia, Plzen

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Combination of the LSDA+DMFT with the KKR method | 3 |
| 2.1 | LSDA+DMFT treatment of disordered alloys | 7 |
| 3 | One-step model of photoemission | 9 |
| 3.1 | General theory of photoemission | 10 |
| 3.2 | Fully relativistic one-step model of photoemission for alloys | 12 |
| 4 | LSDA+DMFT for calculations of spectroscopic properties | 16 |
| 4.1 | Angle-integrated valence band photoemission: Fano effect | 16 |
| 4.2 | ARPES: Correlation effects in transition-metals and their surfaces | 17 |
| 4.3 | ARPES of disordered correlated alloys: $\text{Ni}_x\text{Pd}_{1-x}$ (001) | 23 |
| 5 | Angle resolved soft and hard X-ray photomission | 24 |
| 5.1 | Photon momentum effects: Ag(001) | 26 |
| 5.2 | Thermal effects and XPS limit | 29 |
| 6 | Summary | 32 |

1 Introduction

Angle resolved photoemission (ARPES) and bremsstrahlung isochromat spectroscopy (ARBIS) have developed over several decades into the experimental techniques for directly determining the electronic structure of any new material [1, 2]. These experimental techniques allow to measure the dispersion of occupied bands as well as unoccupied bands and therefore reveal the electronic structure around the Fermi level with a high amount of accuracy. In particular, in recent years many improvements on the experimental side have lead to an increase of the resolution of ARPES down to the meV-regime. These improvements are primarily due to the use of synchrotron radiation, and laser sources, and to developments on the detector side (e.g. spin resolution). More details on the foundations of ARPES can be found in the lecture of M. Sing. On the theory side, about 50 years ago photoemission theory appeared to be an intractable many-body problem [3–7]. The first and most simple version of a one-electron approximation for the photocurrent was given by Berglund and Spicer [8], the so called three-step model of photoemission. In the framework of this model the photoemission process is divided into three independent steps: the excitation of the photoelectron, its transport through the crystal and its escape into the vacuum. Self-energy corrections, which represent, among others, damping processes and energetic shifts in the quasi-particle spectrum, are completely neglected. This means that the initial and final states in the photoemission process are assumed to be Bloch-states with an infinite lifetime. It should be mentioned that the assumption of an infinite electron lifetime does not allow for transitions into evanescent bandgap states, e.g. states that decay exponentially into the solid. Similarly, the assumption of an infinite lifetime for the initial state does in practice not allow to calculate photoemission spectra that involve surface states. To overcome the deficiencies of the three-step model, a dynamic approach has been suggested first for the final state by Liebsch [9] and Spanjaard *et al.* [10]. Later-on multiple scattering effects were properly included for both initial and final states by Pendry and coworkers [11, 12] in order to treat self-energy corrections on an equal footing. Pendry’s one-step approach to ARPES [11, 13] led to a numerically solvable scheme by replacing the retarded one-electron Green function for the initial state by the one-particle Green function determined within density functional theory (DFT) [14]. Within this scheme, electronic correlation effects are typically considered in photoemission theory making use in practice of the local (spin) density approximation (L(S)DA) [15, 16]. Life-time effects in the initial state are accounted for by an imaginary potential term V_{0ii} which is added to the single-particle cell potential. The final state is constructed within the formalism for spin-polarized low-energy electron diffraction (SPLEED) as a so-called time-reversed SPLEED state [13, 17]. The finite imaginary part V_{0fi} effectively simulates the inelastic mean free path (IMFP). As a consequence the amplitude of the high-energy photoelectron state inside the solid can be neglected beyond a certain distance from the surface [11].

After the one-step model had been established in the seventies, it was extended in many aspects. For example, the quantitative analysis of spin-orbit induced dichroic phenomena was worked out by several groups [17–24]. Furthermore, the so called full-potential formulation of photoemission was developed in order to achieve an accurate description of spectroscopic data even

for complex surface systems [17, 23, 25]. The treatment of disordered systems was worked out by Durham *et al.* [26, 20]. Nowadays, the one-step model allows for photocurrent calculations for photon energies ranging from a few eV to more than 10 keV [27–34], for finite temperatures and for arbitrarily ordered [35] and disordered systems [36], and considering in addition strong correlation effects within the dynamical mean-field theory (DMFT) [37–42]. The aim of this lecture is to present these recent developments in the theory of photoemission. The main emphasis will be given to its LSDA+DMFT extention and to the relatively new techniques of soft- and hard x-ray angle resolved photoemission (HARPES).

The LSDA+DMFT implementation within the multiple-scattering Korringa-Kohn-Rostoker (KKR) method [37, 43] is reviewed in Sec. 2. In section 3.2, we briefly review the one-step model of photomeission. The second part of this lecture is devoted to calculations of angle-resolved photoemission within the one-step model including more or less all relevant spectroscopy issues like matrix elements and surface effects. Recent technical developments allow one to perform calculations for ordered as well as for chemically disordered systems including electronic correlation effects. This topic is presented in Sec. 4. Finally, several aspects of soft and hard x-ray photoemission, which are relatively new experimental methods, are discussed in Sec. 5.

2 Combination of the LSDA+DMFT with the KKR method

In the following section we shortly review a fully self-consistent (with respect to charge density and self-energy) LSDA+DMFT implementation within the full-potential fully relativistic multiple scattering Korringa Kohn Rostoker method [37]. This method is used to solve the multiple scattering formalism for semi-infinite solids which in turn is a basis of the one-step model of photomeission as presented in next section. The KKR offers a number of advantages compared to other band structure methods due to the fact that the KKR represents the electronic structure by the corresponding single-particle Green's function (For a recent review of the KKR method see [44] and references therein). This allows one to combine the KKR method with the DMFT straightforwardly. Another important consequence is the possible use of the Dyson equation which relates the Green's function of a perturbed system with the Green's function of the corresponding unperturbed reference system. Using the Dyson equation allows in particular to calculate the properties of low dimensional systems like, e.g., semi-infinite 2D-surfaces, nano-structures or embedded 3D- or 2D- systems without using an artificial super cell construction. Finally, the KKR Green's function method allows one to deal with substitutional disordered alloys in combination with the coherent potential approximation (CPA) [45].

The central idea of the KKR-based implementation of the LSDA+DMFT is to account for the general non-local, site-diagonal, complex and energy-dependent self-energy Σ^{DMFT} already when calculating the basis functions, i.e., when solving the single-site Schrödinger (or Dirac) equation. This allows one to exploit directly all advantageous features of the KKR Green's function method when performing LSDA+DMFT calculations and consequently to account for correlation effects for a wide range of systems.

There are nowadays various approaches available to combine the LSDA with the DMFT method [46,47]. In contrast to the KKR-scheme, the corresponding LSDA problem is in general solved variationally using a given basis set (e.g. LMTO) as a first step. The corresponding local Green's function is determined by the spectral representation of the Kohn-Sham Hamiltonian. Solving subsequently the DMFT problem, the resulting local self-energy Σ^{DMFT} and local Green's function can in turn be used to calculate a new charge density and an effective LSDA potential. However, in order to combine coherently the LSDA with the DMFT method (in the spirit of spectral density functional theory [48]) one has to solve self-consistently the following Dyson equation

$$G(\vec{r}, \vec{r}', E) = G_0(\vec{r}, \vec{r}', E) + \int d^3r'' \int d^3r''' G_0(\vec{r}, \vec{r}'', E) \left(V_{\text{eff}}(\vec{r}'') \delta(\vec{r}'' - \vec{r''''}) + \Sigma(\vec{r}'', \vec{r}''', E) \right) G(\vec{r}''', \vec{r}', E), \quad (1)$$

where $G_0(\vec{r}, \vec{r}'', E)$ is the free electron Green's function. The potential $V_{\text{eff}}(\vec{r})$ denotes the (effective) potential. Within the relativistic version of spin DFT used here this is usually defined as $V_{\text{eff}}(\vec{r}) = [\bar{V}_{\text{eff}}(\vec{r}) + \beta \boldsymbol{\sigma} \mathbf{B}_{\text{eff}}(\vec{r})]$ where $\bar{V}_{\text{eff}}(\vec{r})$ denotes the spin-independent potential, and $\mathbf{B}_{\text{eff}}(\vec{r})$ is the magnetic field [49]. Correspondingly, the matrices β and α_k used below are the standard Dirac matrices with $\beta = \sigma_z \otimes 1_2$ and $\alpha_k = \sigma_x \otimes \sigma_k$ ($k = x, y, z$) in terms of the 2×2 Pauli-matrices σ_k .

A very efficient way of solving Eq. (1) is offered by the multiple scattering KKR method. Having decomposed the system into atomic regions (Wigner-Seitz-cells) and considering that Σ^{DMFT} is an on-site quantity, the equation can be solved using the standard KKR formalism. This implies that one first has to solve the single-site scattering problem to obtain the wave function $\Psi(\vec{r})$ and the corresponding single-site scattering t -matrix inside an atomic cell. In the relativistic spin density functional theory [50, 51] the corresponding single-site Dirac equation reads

$$\left[\frac{\hbar}{i} c \vec{\alpha} \cdot \vec{\nabla} + \beta mc^2 + V_{\text{eff}}(\vec{r}) + \int d^3r' \Sigma(\vec{r}, \vec{r}', E) \right] \Psi(\vec{r}) = E \Psi(\vec{r}). \quad (2)$$

Here, the $\Psi(\vec{r})$ are energy-dependent four-component spinor functions for energy E . To be able to solve Eq. (2) one makes the following ansatz for the wave function $\Psi = \sum_{\Lambda} \Psi_{\Lambda}$, with the combined relativistic quantum number $\Lambda = (\kappa, \mu)$, where κ and μ are the spin-orbit and magnetic quantum numbers, respectively. In addition, in the spirit of the DMFT one has to project $\Sigma(\vec{r}, \vec{r}', E)$ onto the localized set of orbitals $\phi_{\Lambda}^n(\vec{r})$. The corresponding matrix $\Sigma_{\Lambda\Lambda'}(E)$ is obtained as an output from the DMFT solver. In practice, $\Sigma_{\Lambda\Lambda'}(E)$ is used only for correlated d - or f -orbitals. It is worth noting that even in the case of the spherical muffin-tin or atomic-sphere approximation to the potential, the full-potential-like coupled Eqs. (2) have to be solved. This implies that the full-potential version of the KKR has to be used. After having solved the set of coupled equations for the wave functions $\Psi(\vec{r})$ one gets the corresponding single-site t matrix by standard matching to the Hankel and Bessel functions as free-electron solutions. When solving the single-site problem, obviously the entire complexity of the underlying complex non-local

potential within LSDA+DMFT is accounted for. Accordingly, the resulting regular and irregular scattering wave functions $Z_\Lambda(\vec{r}, E)$ and $J_\Lambda(\vec{r}, E)$ as well as the corresponding single-site t -matrix carry all information of the underlying LSDA+DMFT Hamiltonian. This means that in contrast to other LSDA+DMFT implementations, the effect of the self-energy is also reflected in the wave functions Ψ . This becomes important, for example, in a total energy calculation and for the photoemission matrix elements (see Sec. 3.2).

With the single-site t matrix available the next step of the KKR calculation is to solve the multiple scattering problem. This task can be done by using the scattering path operator τ [52] and it is independent from the DMFT. For a finite system this can be done straightforwardly by inverting the so called KKR-matrix, $\underline{\underline{\tau}}(E) = [\underline{\underline{t}}(E)^{-1} - \underline{\underline{G}}_0(E)]^{-1}$ with the double underline indicating matrices with respect to site and spin-angular (Λ) character. Dealing with a three-dimensional periodic system this equation can also be solved exactly by Fourier transformation. As a result the retarded site-diagonal Green's function $G(\vec{r}, \vec{r}', E)$ can be written as [53, 45]

$$\begin{aligned} G(\vec{r}, \vec{r}', E) &= \sum_{\Lambda, \Lambda'} Z_\Lambda(\vec{r}, E) \tau_{\Lambda, \Lambda'}^{nn}(E) Z_{\Lambda'}^\times(\vec{r}', E) \\ &- \sum_{\Lambda} \{ Z_\Lambda(\vec{r}, E) J_\Lambda^\times(\vec{r}', E) \Theta(\vec{r}' - \vec{r}) \\ &+ J_\Lambda(\vec{r}, E) Z_\Lambda^\times(\vec{r}', E) \Theta(\vec{r} - \vec{r}') \}, \end{aligned} \quad (3)$$

where \vec{r} (\vec{r}') lies in the atomic cell n representing cell-centered coordinates and \times indicates a so-called left-hand side solution [54]. With the Green's function $G(\vec{r}, \vec{r}', E)$ available all properties of interest, e.g., the charge density, can be calculated straightforwardly and in this way the calculated Green's function G includes all effects of the self-energy Σ^{DMFT} .

The definition of the Green's function and the expressions given above are not restricted to real energies E but also hold for arbitrary complex energies z . The fact that $G(\vec{r}, \vec{r}', E)$ is analytical [55] allows, in particular, to perform the energy integration for the charge density on a contour in the complex energy plane [56] with typically around 30 energy mesh points. On the other hand the self-energy Σ^{DMFT} is often calculated for a mesh of Matsubara frequencies. This implies that it is necessary to use analytical continuation techniques to transform Σ^{DMFT} from Matsubara frequencies ω onto the KKR complex energy contour. It is worth noting that in general Σ^{DMFT} is not Hermitian and for low-symmetry systems one has to consider right- and left-handed solutions of (2) when constructing the Green's function $G(\vec{r}, \vec{r}', E)$ [54].

In order to construct the bath Green's function needed as the input of the DMFT solver, the localized Green's function is calculated by projecting the Green's function given by Eq. (3) onto the correlated atomic orbitals $\phi_\Lambda(\vec{r})$

$$G_{\Lambda \Lambda'}(E) = \int d^3 r \int d^3 r' \phi_\Lambda(\vec{r}) G(\vec{r}, \vec{r}', E) \phi_{\Lambda'}(\vec{r}'). \quad (4)$$

A natural choice for the projection functions $\phi_\Lambda(\vec{r})$ are the regular single-site solutions of the Kohn-Sham-Dirac equations that are normalized to 1 and also are used to represent the self energy Σ . For transition-metal systems, only the d - d sub-block of the $\underline{\underline{G}}^{nn}(E)$ is considered,

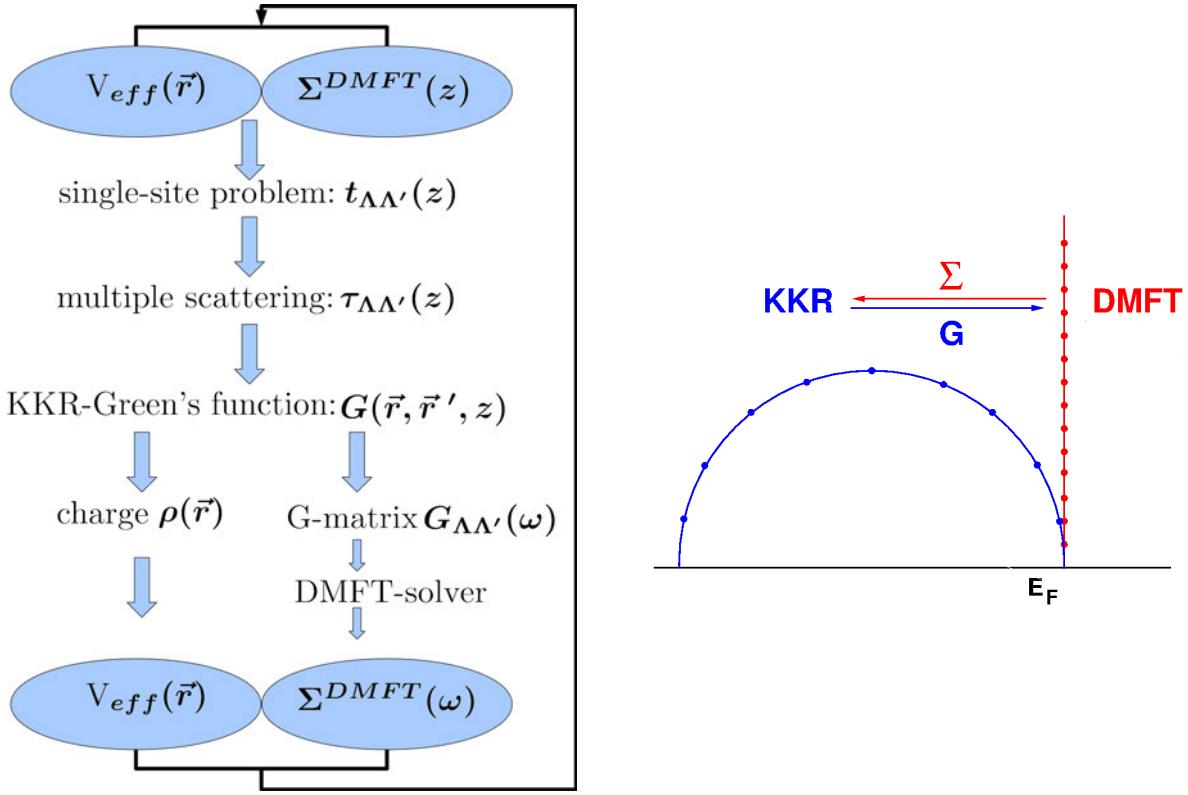


Fig. 1: Left: Schematic overview of the KKR-based LSDA+DMFT scheme. Right: Illustration of the energy paths involved. The blue semicircle is the complex energy path (with complex energies z) used by KKR to calculate the charge density. After the bath Green's function G is obtained, it is analytically continued onto the imaginary axis (red) to calculate the self-energy Σ^{DMFT} via the DMFT impurity solver. The latter is analytically extrapolated back to the semicircle. Figure taken from [43].

using $\phi_\Lambda(\vec{r})$ wave functions with $l = 2$. In principle, the choice of the $\phi_\Lambda(\vec{r})$ is arbitrary as long as $\phi_\Lambda(\vec{r})$ is a complete set of functions. This implies that a localized basis set is calculated at a given reference energy E_{ref} (set to be the center of gravity of the occupied d - or f -band) with the magnetic field set to zero in the relativistic case. In the full-potential case couplings to the other l -channels as a consequence of crystal symmetry have to be suppressed.

A flow diagram describing the resulting KKR-based self-consistent LSDA+DMFT scheme is presented in Fig. 1. Eq. (2) provides the set of regular (Z) and irregular (J) solutions of the single-site problem accounting for the LSDA potential as well as the DMFT self-energy Σ . Together with the t -matrix and the scattering-path operator τ the KKR Green's function is constructed from Eq. (3). To solve the many-body problem, the projected Green's function matrix is constructed according to Eq. (4). The LSDA Green's function $G_{\Lambda\Lambda'}(E)$ is calculated on the complex contour which encloses the valence band one-electron energy poles. The Padé analytical continuation scheme is used to map the complex local Green's function $G_{\Lambda\Lambda'}(E)$ onto the set of Matsubara frequencies or the real axis, which is used when dealing with the many-

body problem. In the current fully relativistic implementation, the perturbative SPTF (spin-polarized T -matrix + FLEX) [57] as well as $T = 0$ K spin-polarized T -matrix [58] solvers of the DMFT problem are used. In fact any DMFT solver could be included which supplies the self-energy $\Sigma(E)$ when solving of the many-body problem. The Padé analytical continuation is used once more to map the self-energy from the Matsubara axis back onto the complex plane, where the new KKR Green's function is calculated. As was described in the previous sections, the key role is played by the scattering path operator $\tau_{AA'}^{nn}(E)$, which allows us to calculate the charge in each SCF iteration and the new potentials that are used to generate the new single particle Green's function.

Finally, the double-counting corrections H_{DC} have to be considered. This problem is definitely one of the main challenges towards first-principles calculations within LSDA+DMFT. Until now various schemes for double-counting correction have been suggested [47]. The simplest choice, i.e., the idea of the static LSDA+ U scheme has been used here. We apply the double-counting corrections to the self-energy when solving the many-body problem. First of all we remove the static part of the self-energy coming from the SPTF solver and add the mean-field (AMF) LSDA+ U like static part [59]. In the case of pure transition metals, as well as their metallic compounds and alloys, the so-called AMF double-counting correction seems to be most appropriate [59, 38, 40, 60]. It is worth noting here that currently an exact analytical equation for the double-counting correction is not known. Alternatively, it might be possible to get an exact solution of the double-counting problem on the level of the GW+DMFT scheme [61]. Therefore, it is important to perform direct comparisons, for example, to ARPES experiments as a stringent test for the choice of the optimal H_{DC} . However, to be able to make a decision between various suggestions for H_{DC} it is helpful not only to calculate the bare spectral function, i.e. $\text{Im } G$, but also to perform a complete calculation of photoemission spectra (for example with the one-step model of photoemission, see Sec. 3.2). In fact, using the one-step model of photoemission, one can clearly see that the AMF H_{DC} is an appropriate choice at least for transition-metal systems [38, 40, 60].

2.1 LSDA+DMFT treatment of disordered alloys

It is an outstanding feature of the KKR method that it supplies the one-electron Green's function of the considered system directly without relying on Bloch's theorem. Because of this property, the KKR Green's function method allows one to deal with substitutional disorder including both diluted impurities and concentrated alloys in the framework of the CPA [62, 63]. Within this approach (KKR-CPA) the propagation of an electron in an alloy is regarded as a succession of elementary scattering processes due to random atomic scatterers, with an average taken over all configurations of the atoms. This problem can be solved assuming that a given scattering center is embedded in an effective medium whose choice is open and can be determined in a self-consistent way. The physical condition corresponding to the CPA is simply that a single scatterer embedded in the effective CPA medium should produce no further scattering on the average as illustrated by Fig. 2.

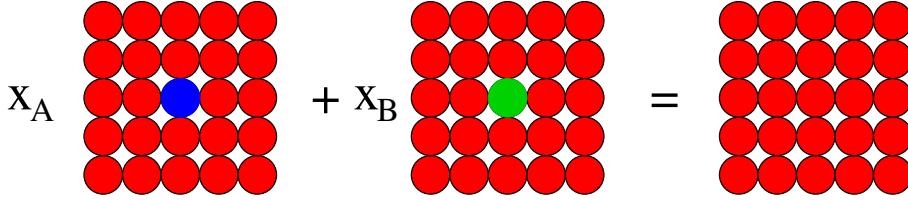


Fig. 2: The major ideas of the CPA: The configurational average over all configurations of a disordered alloy A_xB_{1-x} is represented by an auxiliary CPA medium. Embedding of an A or B atom should not give rise to additional scattering with respect to the CPA medium.

A similar philosophy is applied also when dealing with many-body problems for crystals in the framework of DMFT. Thus it is rather straightforward to combine the DMFT and KKR-CPA method as both schemes are used on a single-site level, i.e., any correlation in the occupation (e.g. short range order) is ignored and the DMFT self-energy Σ is taken to be on-site only. In fact, the combination of the KKR-CPA for disordered alloys and the DMFT scheme is based on the same arguments as used by Drchal *et al.* [64] when combining the LMTO Green's function method for alloys [65] with the DMFT.

The combination of the CPA and LSDA+DMFT turned out to be a rather powerful technique for calculating electronic structure properties of substitutionally disordered correlated materials [37, 59, 66, 36]. As mentioned, within the CPA the configurationally averaged properties of a disordered alloy are represented by a hypothetical ordered CPA-medium, which in turn may be described by a corresponding site-diagonal scattering path operator τ^{CPA} , which in turn is closely connected with the electronic Green's function. For example for a binary system A_xB_{1-x} composed of components A and B with relative concentrations x and $1 - x$ the corresponding single-site t -matrix t^{CPA} and the multiple scattering path operator τ^{CPA} are determined by the so called CPA-condition:

$$x\tau^A + (1 - x)\tau^B = \tau^{\text{CPA}}. \quad (5)$$

The CPA equation represents the requirement that substitutionally embedding an atom (of type A or B) into the CPA medium should not cause additional scattering. The scattering properties of an A atom embedded in the CPA medium are represented by the site-diagonal component-projected scattering path operator τ^A (angular momentum index omitted here)

$$\tau^A = \tau^{\text{CPA}} [1 + (t_A^{-1} - t_{\text{CPA}}^{-1}) \tau^{\text{CPA}}]^{-1} = \tau^{\text{CPA}} \mathcal{D}^A, \quad (6)$$

where t_A and t_{CPA} are the single-site matrices of the A component and of the CPA effective medium; the factor $\mathcal{D}^A = [1 + (t_A^{-1} - \tau_{\text{CPA}}^{-1})]^{-1}$ in Eq. (6) is called the CPA-projector. A corresponding equation holds also for the B component in the CPA medium. The coupled set of equations for τ^{CPA} and t_{CPA} has to be solved iteratively within the CPA cycle. For example when a binary system A_xB_{1-x} composed of components A and B with relative concentrations $x_A = x$ and $x_B = (1 - x)$ is considered, the above equation represents the requirement that embedding substitutionally an atom (of type A or B) into the CPA medium should not cause additional scattering.

The above scheme can straightforwardly be extended to include many-body correlation effects for disordered alloys [37]. Within the KKR approach the local multi-orbital and energy dependent self-energies ($\Sigma_A^{DMFT}(E)$ and $\Sigma_B^{DMFT}(E)$) are directly included into the single-site matrices t_A and t_B , respectively when solving the corresponding Dirac equation (2). Consequently, all the relevant physical quantities connected with the Green's function, for example the charge density, contain the electronic correlations beyond the LSDA scheme.

3 One-step model of photoemission

Spectroscopy is an extremely important experimental tool providing information on the electronic structure of the probed system that has to be seen as a stringent benchmark for the success of any electron structure theory. Photoemission spectroscopy (PES) or its inverse – the Bremsstrahlung isochromate spectroscopy (BIS) – in their angle-integrated form should reflect the density of states (DOS) rather directly, in particular in the high photon-energy regime (XPS). For that reason it is quite common to check the DMFT-based calculations by comparing the calculated DOS directly to the PES spectra (see the reviews [67, 46, 47] for example).

However, this approach ignores the influence of the specific PES matrix elements that in general will introduce an element- and energy-dependent weight to the partial DOS. In ARPES, the situation is even more severe as the surface as well as dipole selection rules may have a pronounced impact on the spectra [68] demanding a coherent description on the basis of the one-step model of photoemission [17]. To achieve a reliable interpretation of experiments it is inevitable to deal with so-called matrix-element effects that considerably modify the raw spectrum. In particular, the wave-vector and energy dependence of the transition-matrix elements has to be accounted for. These issues are known to be important and cannot be neglected. They arise from strong multiple-scattering processes in the final PES state that dominate the electron dynamics in the low-energy regime of typically 1-200 eV [13]. The transition-matrix elements also include the effects of selection rules which are not accounted for in the raw spectrum. Loosely speaking, it can be said that the main task of a theory of photoemission is to close the gap between the raw spectrum obtained by LSDA+DMFT electronic-structure calculations and the experiment. The most successful theoretical approach concerning this is the one-step model of photoemission as originally proposed by Pendry and co-workers [11–13]. In the following a short overview will be given on the recent extensions of this model which are connected with correlation effects and disordered alloys.

The main idea of the one-step model is to describe the actual excitation process, the transport of the photoelectron to the crystal surface as well as the escape into the vacuum [8] as a single quantum-mechanically coherent process including all multiple-scattering events. Within this model self-energy corrections, which give rise to damping in the quasi-particle spectrum, are properly included in both the initial and the final states. This, for example, allows for transitions into evanescent band gap states decaying exponentially into the solid. Similarly, the assumption of a finite lifetime for the initial states gives the opportunity to calculate photoemission intensities from surface states and resonances. Treating the initial and final states within the

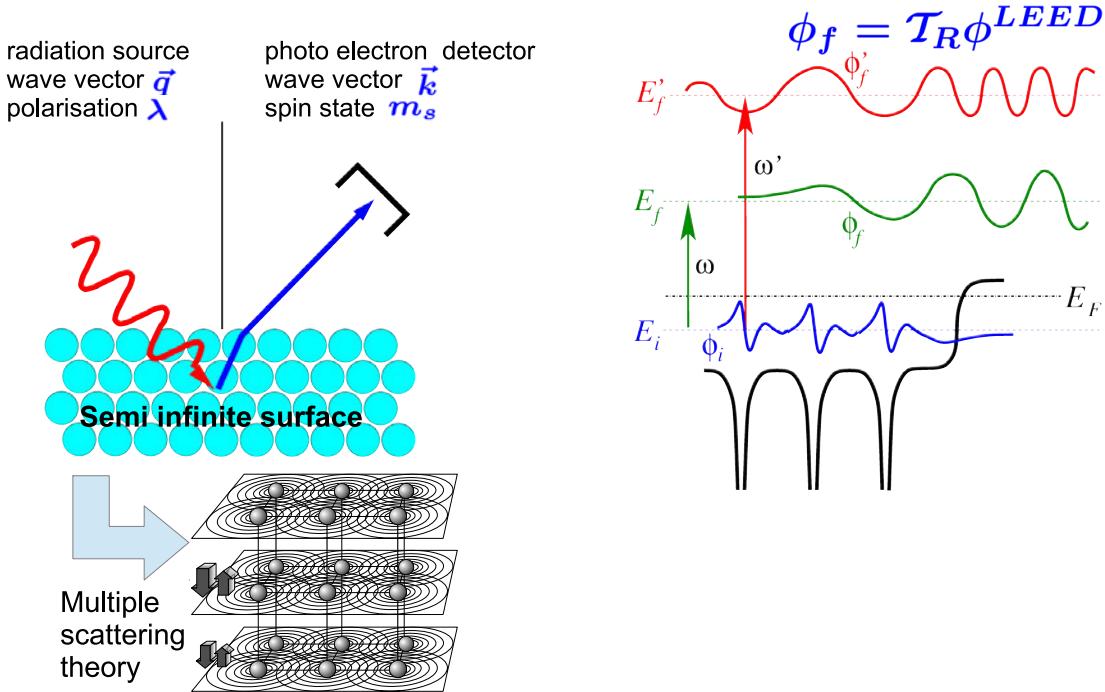


Fig. 3: Left: Schematic overview of the one-step model of photoemission. The whole photoemission process is solved within the multiple scattering theory for a semi-infinite surface. Right: Electron from initial state ϕ_i at energy E_i is exited to the final state ϕ_f (time-reversed SPLEED) which decays into the solid due to the inelastic processes (modeled by imaginary part of potential). By increasing the inelastic mean free path of the time-reversed SPLEED state (E'_f) the photoemission process becomes more bulk sensitive (See Sec. 5).

fully relativistic layer version of the KKR [69, 70], it is a straightforward task to describe the photoemission from complex layered structures like thin films and multilayers. Furthermore, the surface described by a barrier potential can be easily included into the multiple-scattering formalism as an additional layer. A realistic surface barrier model that shows the correct asymptotic behavior has been introduced, for example, by Rundgren and Malmström [71].

3.1 General theory of photoemission

In this section, the main features of general photoemission theory will be elucidated. The calculation of the photocurrent starts from first order time-dependent perturbation theory. Assuming a small perturbation Δ , the transition probability per unit time w between two N -electron states $|\Psi_F\rangle$ and $|\Psi_I\rangle$ of the same Hamiltonian H , is given by Fermi's golden rule:

$$w = \frac{2\pi}{\hbar} |\langle \Psi_F | \Delta | \Psi_I \rangle|^2 \delta(E_F - E_I - \hbar\omega), \quad (7)$$

where E_F and E_I denote the energies of the N -electron states and $\hbar\omega$ the excitation energy. This equation can be also derived within the Keldysh Green's function approach and can be

represented in the lowest order as a triangular like skeleton diagrams (See e.g. Fujikawa and Arai [72, 73]). In second quantization the interaction operator Δ is defined as follows

$$\Delta = \sum_{k,m} \Delta_{k,m} \mathbf{a}_k^\dagger \mathbf{a}_m , \quad (8)$$

where $\Delta_{k,m}$ denotes a one-particle matrix element between two single-particle states ϕ_k and ϕ_m . The initial and final states are then defined as $|\Psi_I\rangle = |\Psi_N^0\rangle$ and $|\Psi_F\rangle = \mathbf{a}_f^\dagger |\Psi_{N-1}^s\rangle$ where $|\Psi_{N-1}^s\rangle$ denotes an excited $N - 1$ particle state and $|\Psi_N^0\rangle$ defines the ground state of the many-particle system. For the explicit formulation of the initial state the so called sudden approximation is used. This means the photoelectron is described by a single-particle state and the interaction with the excited $N - 1$ state $|\Psi_{N-1}^s\rangle$ has been completely neglected. In other words $\mathbf{a}_f |\Psi_N^0\rangle = 0$. Using these approaches for the initial and final states the transition probability is given by

$$w_s = \frac{2\pi}{\hbar} |\langle \Psi_{N-1}^s | \sum_{k,m} \Delta_{k,m} \mathbf{a}_f \mathbf{a}_k \mathbf{a}_m | \Psi_N^0 \rangle|^2 \delta(E_N - E_{N-1} - \hbar\omega) , \quad (9)$$

where the delta-function describes the energy conservation in the photo-excitation process generated by a certain photon energy $\hbar\omega$. Performing some standard manipulations on Eq. (9) it follows for $w = \sum_s w_s$

$$w = \frac{2\pi}{\hbar} \sum_{m,m'} \Delta_{f,m}^\dagger A_{m,m'}(E_n) \Delta_{f,m'} , \quad (10)$$

where

$$A_{m,m'}(E_n) \frac{2}{\hbar} \sum_s \langle \Psi_N^0 | \mathbf{a}_m^\dagger | \Psi_{N-1}^s \rangle \delta(E_N - E_{N-1} - \hbar\omega) \langle \Psi_{N-1}^s | \mathbf{a}_{m'} | \Psi_N^0 \rangle , \quad (11)$$

represents the one-electron spectral function of the initial state. Using further the relation

$$A_{m,m'}(E_n) = -\frac{1}{\pi} \text{Im } G_{m,m'}^R(E_n) \quad (12)$$

between the spectral function and the one-electron retarded Green's function, the intensity of the photocurrent follows

$$I(\hbar\omega) = -\frac{1}{\pi} \text{Im} \sum_{m,m'} \langle \phi_f | \Delta^\dagger | \phi_m \rangle G_{m,m'}^R(E_n) \langle \phi_m | \Delta | \phi_f \rangle . \quad (13)$$

With the help of the operator representation for G^R

$$G^R(E_n) = \sum_{m,m'} |\phi_m\rangle G_{m,m'}^R(E_n) \langle \phi_{m'}| \quad (14)$$

we arrive at the final expression

$$I(\hbar\omega) = -\frac{1}{\pi} \text{Im} \langle \phi_f | \Delta^\dagger G^R(E_n) \Delta | \phi_f \rangle \quad (15)$$

for the photocurrent $I(\hbar\omega)$. Replacing the retarded one-electron Green's function by the one-particle Green function and reformulating Eq. (15) in the space representation one arrives at the one-step model description of the photocurrent as derived among others by Pendry [11] (see Eq. (16) and [74]).

3.2 Fully relativistic one-step model of photoemission for alloys

In this section we will show some basic ideas concerning the implementation of Eq. (15) as derived in Sec. 3.1. As a detailed description of the one-step model of photoemission can be found in various reviews (e.g. [17]), here we would like to show only the main steps with an emphasis on the calculations of the photomeission for correlated alloys.

An implementation of the one-step model of PES can be based on Pendry's expression for the photocurrent [11], which is nothing but the space representation of Eq. (15)

$$I^{\text{PES}} \propto \text{Im} \langle \varepsilon_f, \vec{k}_{\parallel} | G_2^+ \Delta G_1^+ \Delta^\dagger G_2^- | \varepsilon_f, \vec{k}_{\parallel} \rangle . \quad (16)$$

I^{PES} denotes the elastic part of the photocurrent with vertex renormalizations being neglected. This excludes inelastic energy losses and corresponding quantum-mechanical interference terms [11, 74, 6]. Furthermore, the interaction of the outgoing photoelectron with the rest of the system is not taken into account. This *sudden approximation* is expected to be justified for not too small photon energies. Considering an energy-, angle- and spin-resolved photoemission experiment the state of the photoelectron at the detector is written as $|\varepsilon_f, \vec{k}_{\parallel}\rangle$, where \vec{k}_{\parallel} is the component of the wave vector parallel to the surface, and ε_f is the kinetic energy of the photoelectron. The spin character of the photoelectron is implicitly included in $|\varepsilon_f, \vec{k}_{\parallel}\rangle$ which is understood as a four-component Dirac spinor. The advanced Green's function G_2^- in Eq. (16) characterizes the scattering properties of the material at the final-state energy $E_2 \equiv \varepsilon_f$. Via $|\Psi_f\rangle = G_2^- |\varepsilon_f, \vec{k}_{\parallel}\rangle$ all multiple-scattering corrections are formally included. For an appropriate description of the photoemission process we must ensure the correct asymptotic behavior of $\Psi_f(\vec{r})$ beyond the crystal surface, i.e., a single outgoing plane wave characterized by ε_f and \vec{k}_{\parallel} . Furthermore, the damping of the final state due to the imaginary part of the inner potential $iV_{0i}(E_2)$ must be taken into account. We thus construct the final state within spin-polarized low-energy electron diffraction (SPLEED) theory considering a single plane wave $|\varepsilon_f, \vec{k}_{\parallel}\rangle$ advancing onto the crystal surface. Using the standard layer-KKR method generalized for the relativistic case [17, 75], we first obtain the SPLEED state $-T\Psi_f(\vec{r})$. The final state is then given as the time-reversed SPLEED state ($T = -i\sigma_y K$ is the relativistic time inversion operator). Many-body effects are included phenomenologically in the SPLEED calculation by using a parametrized weakly energy-dependent and complex inner potential $V_0(E_2) = V_{0r}(E_2) + iV_{0i}(E_2)$ [13]. This generalized inner potential takes into account inelastic corrections to the elastic photocurrent [74] as well as the actual (real) inner potential, which serves as a reference energy inside the solid with respect to the vacuum level [76]. Due to the finite imaginary part $iV_{0i}(E_2)$, the flux of elastically scattered electrons is continuously reduced, and thus the amplitude of the high-energy wave field $\Psi_f(\vec{r})$ can be neglected beyond a certain distance from the surface (see right panel of Fig. 3).

In the last part of this section we would like to explicitly evaluate Eq. (16) for the CPA photocurrent. A more detailed description of the generalized one-step model for disordered magnetic alloys can be found in Braun *et al.* and Durham *et al.* [36, 26, 20]. The first step in an explicit calculation of the photocurrent consists in the setup of the relativistic spin-polarized

low energy electron diffraction (SPLEED)-formalism within the CPA theory. The coherent scattering matrix t_n^{CPA} for the n th atomic site together with the crystal geometry determines the scattering matrix M for a certain layer of the semi-infinite half-space

$$\begin{aligned} M_{\mathbf{g}\mathbf{g}'}^{\tau\tau'ss'} &= \delta_{\mathbf{g}\mathbf{g}'}^{\tau\tau'ss'} \\ &+ \frac{8\pi^2}{kk_{\mathbf{g}z}^+} \sum_{\Lambda\Lambda'\Lambda''} i^{-l} C_{\Lambda}^{m_s} Y_l^{\mu-m_s}(\hat{\mathbf{k}}_{\mathbf{g}}^\tau) t_{\Lambda\Lambda''}^{CPA}(1-X)^{-1}_{\Lambda''\Lambda'} i^{l'} C_{\Lambda'}^{m'_s} Y_{l'}^{\mu'-m'_s}(\hat{\mathbf{k}}_{\mathbf{g}'}^{\tau'}) e^{-i(\mathbf{k}_{\mathbf{g}}^\tau \mathbf{R}_n + \mathbf{k}_{\mathbf{g}'}^{\tau'} \mathbf{R}_{n'})}, \end{aligned} \quad (17)$$

where the X -matrix represents the Kambe lattice sum, which in turn is directly connected with the multiple scattering path operator $\tau = tX$. All quantities are indexed by the reciprocal lattice vectors of the 2D layer \mathbf{g} , relativistic quantum numbers Λ and site index n and index $\tau = \pm$ (+ for transmission, and - for reflection). C and Y are standard Clebsch-Gordan coefficients and the spherical harmonics, respectively. By means of the layer-doubling technique the so called bulk-reflection matrix can be calculated, which gives the scattering properties of a semi-infinite stack of layers. Finally, applying SPLEED theory [44, 77] we are able to derive the final state for the semi-infinite crystal. The quantity Δ in Eq. (16) is the dipole operator in the electric dipole approximation. It mediates the coupling of the high-energy final state with the low-energy initial states. In a fully relativistic theory the dipole interaction of an electron with the electromagnetic field is given by the dipole operator $\Delta = -\alpha \mathbf{A}_0$ where \mathbf{A}_0 is the spatially constant vector potential inside the crystal. The three components α_k of the vector α are defined through the tensor product $\alpha_k = \sigma_1 \otimes \sigma_k$, $k = z, y, z$, where σ_k denote the Pauli spin matrices. Dealing with the matrix element $\langle \Psi_f | \Delta | \Psi_i \rangle$ between eigenspinors $|\Psi_f\rangle$ and $|\Psi_i\rangle$ of the Dirac Hamiltonian with energies E_f and E_i , respectively, it is numerically more stable to transform Δ in the so called ∇V form of matrix elements. This is derived by making use of commutator and anticommutator rules analogously to the nonrelativistic case in Ref. [78, 79].

According to Pendry [11] the calculation of G_1^+ , and in consequence the calculation of the photocurrent, can be divided into four different steps. The first contribution I^{at} , the so called *atomic contribution* results from the replacement of G_1^+ in Eq. (16) by $G_{1,a}^+$. The second contribution I^{ms} , describes the multiple scattering of the initial state. The third contribution I^{s} to the photocurrent takes care of the surface. When dealing with the disorder in the alloys, an additional I^{inc} , the so called *incoherent* term, appears. Following Durham *et al.* [26, 20] the configurational average can be written as

$$\langle I^{\text{AR-PES}} \rangle = \frac{1}{\pi} \text{Im} \sum_{ij} \langle M_i \tau^{ij} M_j^* \rangle = \frac{1}{\pi} \text{Im} \sum_i \langle M_i^a \rangle + \langle I^{\text{s}} \rangle. \quad (18)$$

Herein τ^{ij} denotes the scattering path operator between the sites i and j . M_i^a represents an atomic-type matrix element containing the irregular solutions which appear as a part of the retarded Green function $G_{1,a}^+$. M_i indicates a conventional matrix element between regular solutions of the initial and final states. The first term can be decomposed in on-site and off-site contributions

$$\sum_{ij} \langle M_i \tau^{ij} M_j^* \rangle = \sum_{ij, i \neq j} \langle M_i \tau^{ij} M_j^* \rangle + \sum_i \langle M_i \tau^{ii} M_i^* \rangle. \quad (19)$$

The on-site term is called incoherent part of the photocurrent since this term reveals density-of-states (DOS)-like behavior by definition. The off-site contribution which contains all dispersing features represents the so called coherent part of the photocurrent. Together with the surface part that remains unchanged by the averaging procedure $\langle I^s \rangle = I^s$ the total one-step current can be written as

$$\begin{aligned} \langle I^{\text{AR-PES}} \rangle &= -\frac{1}{\pi} \text{Im} \sum_i \langle M_i^a \rangle + I^s \\ &+ \frac{1}{\pi} \text{Im} \sum_{ij, i \neq j} \langle M_i \tau^{ij} M_j^* \rangle \\ &+ \frac{1}{\pi} \text{Im} \sum_i \langle M_i \tau^{ii} M_i^* \rangle . \end{aligned} \quad (20)$$

Using Pendry's notation it follows

$$\langle I^{\text{AR-PES}}(\varepsilon_f, \mathbf{k}) \rangle = \langle I^{\text{at}}(\varepsilon_f, \mathbf{k}) \rangle + \langle I^{\text{ms}}(\varepsilon_f, \mathbf{k}) \rangle + \langle I^{\text{inc}}(\varepsilon_f, \mathbf{k}) \rangle + I^s(\varepsilon_f, \mathbf{k}), \quad (21)$$

where I^{ms} can be identified with the coherent contribution that describes all band-like features of the initial state and I^{inc} with the incoherent contribution that describes the corresponding DOS-like features. Because of this clear-cut separation in contributions that describe dispersing or non-dispersing features one may easily define the angle-integrated photocurrent by use of the CPA-formalism. The ordered case is then defined by a binary alloy with two identical species at each atomic site. Therefore, it follows

$$\langle I^{\text{AI-PES}}(\varepsilon_f, \mathbf{k}) \rangle = \langle I^{\text{at}}(\varepsilon_f, \mathbf{k}) \rangle + \langle I^{\text{inc}}(\varepsilon_f, \mathbf{k}) \rangle + I^s(\varepsilon_f, \mathbf{k}) . \quad (22)$$

For the atomic contribution the averaging procedure is trivial, since $\langle I^{\text{at}}(\varepsilon_f, \mathbf{k}) \rangle$ is a single-site quantity. In detail, the atomic contribution is build up by a product between the matrix $\mathcal{Z}_{jn\alpha_n}^{\text{at}}$ and the coherent multiple scattering coefficients $A_{jn\Lambda}^c$ of the final state. Herein n denotes the n th cell of the j th layer and Λ denotes the combined relativistic quantum numbers (κ, μ) . It follows

$$\langle I^{\text{at}}(\varepsilon_f, \mathbf{k}) \rangle \propto \text{Im} \sum_{\substack{j n \alpha_n \\ \Lambda \Lambda'}} x_{jn\alpha_n} A_{jn\Lambda}^c \mathcal{Z}_{jn\alpha_n}^{\text{at}} A_{jn\Lambda'}^{c*}, \quad (23)$$

where α_n denotes the different atomic species located at a given atomic site n of the j th layer. The corresponding concentration is given by $x_{jn\alpha_n}$.

For an explicit calculation \mathcal{Z}^{at} must be separated into angular matrix elements and radial double matrix elements. A detailed description of the matrix \mathcal{Z}^{at} and of the multiple scattering coefficients $A_{jn\Lambda}^c$ for the different atomic species is given in Refs. [17, 75].

The intra(inter)-layer contributions $\langle I^{\text{ms}}(\varepsilon_f, \mathbf{k}) \rangle$ to the photocurrent describe the multiple scattering corrections of the initial state G_1^+ between and within the layers of the single crystal. They can be written in a similar form

$$\langle I^{\text{ms}}(\varepsilon_f, \mathbf{k}) \rangle \propto \text{Im} \sum_{\substack{j n \\ \Lambda \Lambda'}} A_{jn\Lambda}^c \mathcal{Z}_{jn}^{c(2)} C_{jn\Lambda'}^{B,G} . \quad (24)$$

In analogy to the atomic contribution, the coherent matrix $\mathcal{Z}^{c(2)}$ can be separated into angular- and radial parts. The difference to the atomic contribution is that the radial part of the matrix $\mathcal{Z}^{c(2)}$ consists of radial single matrix elements instead of radial double integrals. In the alloy case this matrix results in the following expression

$$\mathcal{Z}_{jn}^{c(2)} = \sum_{\substack{\alpha_n \\ \Lambda \Lambda' \\ \Lambda_1 \Lambda_2 \Lambda_3}} x_{jn\alpha_n} D_{\Lambda_1 \Lambda_2} R_{\Lambda_1 \Lambda_2 \Lambda_3}^{(2)} \mathcal{D}_{jn\alpha_n} \cdot \quad (25)$$

The radial and angular parts of the matrix element are denoted by $R^{(2)}$ and D . The CPA-average procedure explicitly is represented in terms of the CPA-projector $\mathcal{D}_{jn\alpha_n}$ representing the α -species at site n for layer j . C^B and C^G denote the coherent multiple scattering coefficients of the initial state within a layer and between different layers. They have the form

$$C_{jn\Lambda}^B = \sum_{n' \Lambda' \Lambda''} B_{jn' \Lambda'}^{(o)c} (t^{CPA})_{\Lambda' \Lambda''}^{-1} \left((1 - X)_{\Lambda'' \Lambda}^{-1} - \delta_{nn' \Lambda'' \Lambda} \right), \quad (26)$$

with the coherent bare amplitudes $B_{jn' \Lambda'}^{(o)c}$

$$B_{jn' \Lambda'}^{(o)c} = \sum_{\Lambda''} \mathcal{Z}_{jn'}^{c(1)} A_{jn' \Lambda''}^{c*}. \quad (27)$$

and

$$\mathcal{Z}_{jn}^{c(1)} = \sum_{\substack{\alpha_n \\ \Lambda \Lambda' \\ \Lambda_1 \Lambda_2 \Lambda_3}} x_{jn\alpha_n} \mathcal{D}_{\Lambda \Lambda_3}^\dagger R_{\Lambda_1 \Lambda_3 \Lambda_2 \Lambda'}^{(1)} D_{\Lambda_1 \Lambda_2}^\dagger. \quad (28)$$

Finally, the coherent scattering coefficients C^G for the inter-layer contribution take the form

$$C_{jn\Lambda}^G = \sum_{n' \Lambda'} G_{jn' \Lambda'}^{(o)c} (1 - X)_{\Lambda' \Lambda}^{-1} \quad (29)$$

and the coherent bare amplitudes $G_{jn' \Lambda'}^{(o)c}$ are given by

$$G_{jn' \Lambda'}^{(o)c} = \sum_{\mathbf{g} m_s} 4\pi i^{l'} (-)^{\mu' - s} C_{\Lambda'}^{m_s} \left(d_{j\mathbf{g} m_s}^+ Y_{l'}^{m_s - \mu'} (\widehat{k_{1\mathbf{g}}^+}) e^{i\mathbf{k}_{1\mathbf{g}}^+ \cdot \mathbf{r}_{n'}} + d_{j\mathbf{g} m_s}^- Y_{l'}^{m_s - \mu'} (\widehat{k_{1\mathbf{g}}^-}) e^{i\mathbf{k}_{1\mathbf{g}}^- \cdot \mathbf{r}_{n'}} \right). \quad (30)$$

The coefficients $d_{j\mathbf{g} m_s}^\pm$ in Eq. (30) represent the plane-wave expansion of the initial state between the different layers of the semi-infinite stack of layers. For a detailed description of the matrices \mathcal{Z}^{at} , $\mathcal{Z}^{(1)}$ and $\mathcal{Z}^{(2)}$ and of the multiple scattering coefficients $d_{j\mathbf{g} m_s}^\pm$ the reader again is referred to Refs. [17, 75].

The last contribution to the alloy photocurrent is the so called incoherent part $\langle I^{inc}(\varepsilon_f, \mathbf{k}) \rangle$, which appears because the spectral function of an disordered alloy [45] is defined as a non single-site quantity. In fact this contribution is closely connected with the presence of the irregular wave functions well-known from the spherical representation of the Green function G_1^+ .

The incoherent term is defined as

$$\begin{aligned} < I^{\text{inc}}(\varepsilon_f, \mathbf{k}) > \propto & \text{Im} \sum_{\substack{j n \alpha_n \\ \Lambda \Lambda' \Lambda''}} x_{jn\alpha_n} A_{jn\Lambda}^c \mathcal{Z}_{jn\alpha_n}^{(1)} (\tau_{jn\alpha_n}^{00} - t_{jn\alpha_n})_{\Lambda' \Lambda''} \mathcal{Z}_{jn\alpha_n}^{(2)}_{\Lambda'' \Lambda'''} A_{jn\Lambda'''}^{c*} \\ & + \text{Im} \sum_{\substack{j n \\ \Lambda \Lambda'}} A_{jn\Lambda}^c \mathcal{Z}_{jn}^{c(1)} \tau_{c j n}^{00} \mathcal{Z}_{jn}^{c(2)} A_{jn\Lambda'''}^{c*}, \end{aligned} \quad (31)$$

where $\tau_{jn\alpha_n}^{00}$ denotes the one-site restricted average CPA-matrix for species α_n at atomic site n for layer j . $\tau_{c j n}^{00}$ represents the corresponding matrix for the coherent medium. This completes the CPA-averaged photocurrent within the fully-relativistic one-step model.

4 LSDA+DMFT for calculations of spectroscopic properties

In the following section we would like to show a couple of recent applications of the formalism presented in Secs. 2 and 3. The propose of the detailed discussion at each example is to show which additional information one can gain from the one-step model calculations in contrast to the standard interpretation of ARPES-spectra-based comparisons between ground state spectral functions or densities of states with experimental data.

4.1 Angle-integrated valence band photoemission: Fano effect

Spin-orbit coupling gives rise to many interesting phenomena in the electron spectroscopy of magnetic solids. A rather straightforward access to the understanding of these phenomena is provided by the study of the Fano effect. This effect was predicted by Fano at the end of the sixties and denotes the fact that one obtains a spin-polarized photoelectron current even for non-magnetic systems if the excitation is done using circularly polarized light [80]. Reversing the helicity in non-magnetic samples reverses the spin polarization of the photocurrent. This symmetry is in general broken for magnetically ordered systems leading to magnetic circular dichroism. As a consequence, in the case of magnetic materials, the spin polarization is usually due to the interplay between spin-orbit coupling and exchange splitting. Recently, we demonstrated by investigations on Fe, Co and Ni that the pure Fano effect can also be observed in angle-integrated valence band XPS (VB-XPS) for ferromagnets if circularly polarized light impinges perpendicular to the magnetization direction and if a subsequent spin analysis is done with respect to the direction of the photon beam [81]. This is demonstrated in Fig. 4 where the VB-XPS of Fe, Co and Ni at a photon energy of 600 eV is shown. The photon energy of 600 eV has been used in order to increase the bulk sensitivity of the photoemission process. In the upper panel of Fig. 4 we compare experimental data with corresponding LSDA and LSDA+DMFT VB-XPS data based on the one-step model of photoemission. In all three cases the LSDA+DMFT considerably improves the agreement with experiment. In particular, in the case of Ni LSDA+DMFT leads to a shrinking of the d band width and to a pronounced increase of the intensity in the regime of the 6 eV satellite. Also, for the total intensity of the Fe and

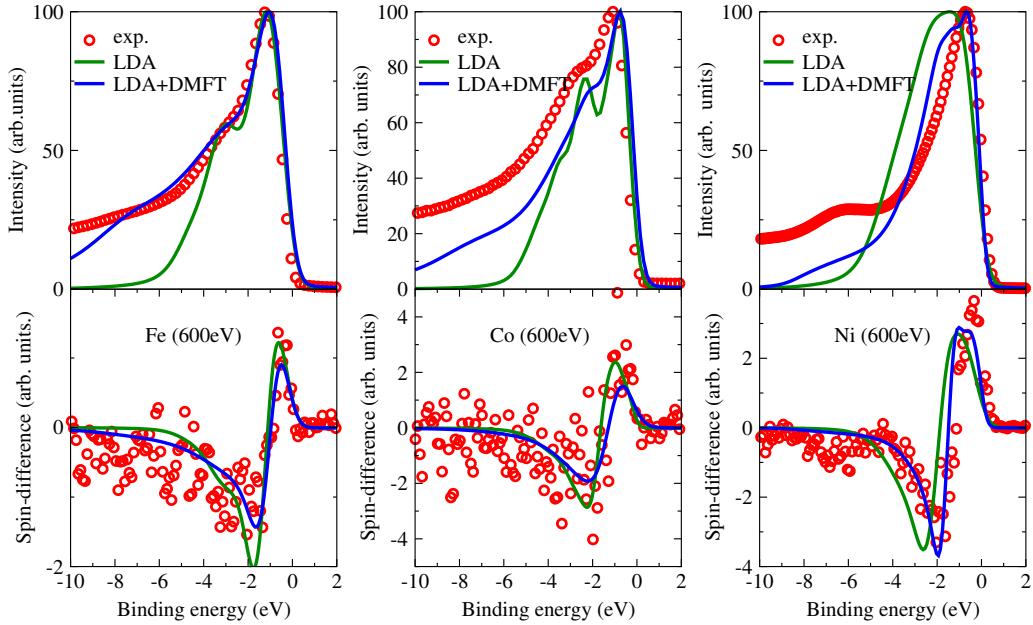


Fig. 4: Top panel: The experimental (dots), LSDA (green line) and LSDA+DMFT (blue line) angle integrated valence band XPS spectra of bcc Fe, hcp Co and fcc Ni for a photon energy of 600 eV. Lower panel: Spin difference $\Delta I^+ = I_\uparrow^+ - I_\downarrow^+$ of the photocurrent for excitation with left circularly polarized light. Figure reproduced from [81].

Co spectra we observed a pronounced improvement in the energy region from -2 to -8 eV. A decomposition of the theoretical spectrum according to the angular momentum character of the initial state shows that the d -contribution is by far dominating and that the spectrum essentially maps the corresponding DOS. This, in some sense, supports the common practice of comparing experimental XPS directly with the DOS. In the lower panel of Fig. 4 the corresponding spin difference $\Delta I^+ = I_\uparrow^+ - I_\downarrow^+$ (i.e., the difference of the currents of photoelectrons with spin-up and spin-down electrons, for excitation with left circularly polarized radiation) is shown. The occurrence of this spin current is a pure matrix element effect induced by spin-orbit coupling. In fact, one finds that the shape of the ΔI^+ curves are very similar to those that can be found for non-magnetic noble metals [82, 83]. In fact, the amplitudes scale with the spin-orbit coupling parameter of the Fe, Co and Ni d-states. To achieve this rather good agreement with the experimental data for the ΔI^+ intensity distribution the fully self-consistent LSDA+DMFT approach is obviously needed.

4.2 ARPES: Correlation effects in transition-metals and their surfaces

In the previous section, we showed angle-integrated XPS spectra which can be directly compared to the DOS ignoring to some extent matrix element effects (upper part of Fig. 4). However, the most detailed mapping of the band structure of correlated materials can be obtained by spin- and angle-resolved valence band photoemission. In the following section we present various examples of angle-resolved photoemission calculations done within the one-step model. These examples clearly demonstrate the need to include matrix-elements in corresponding calculations in order to obtain a quantitative understanding of the experimental data.

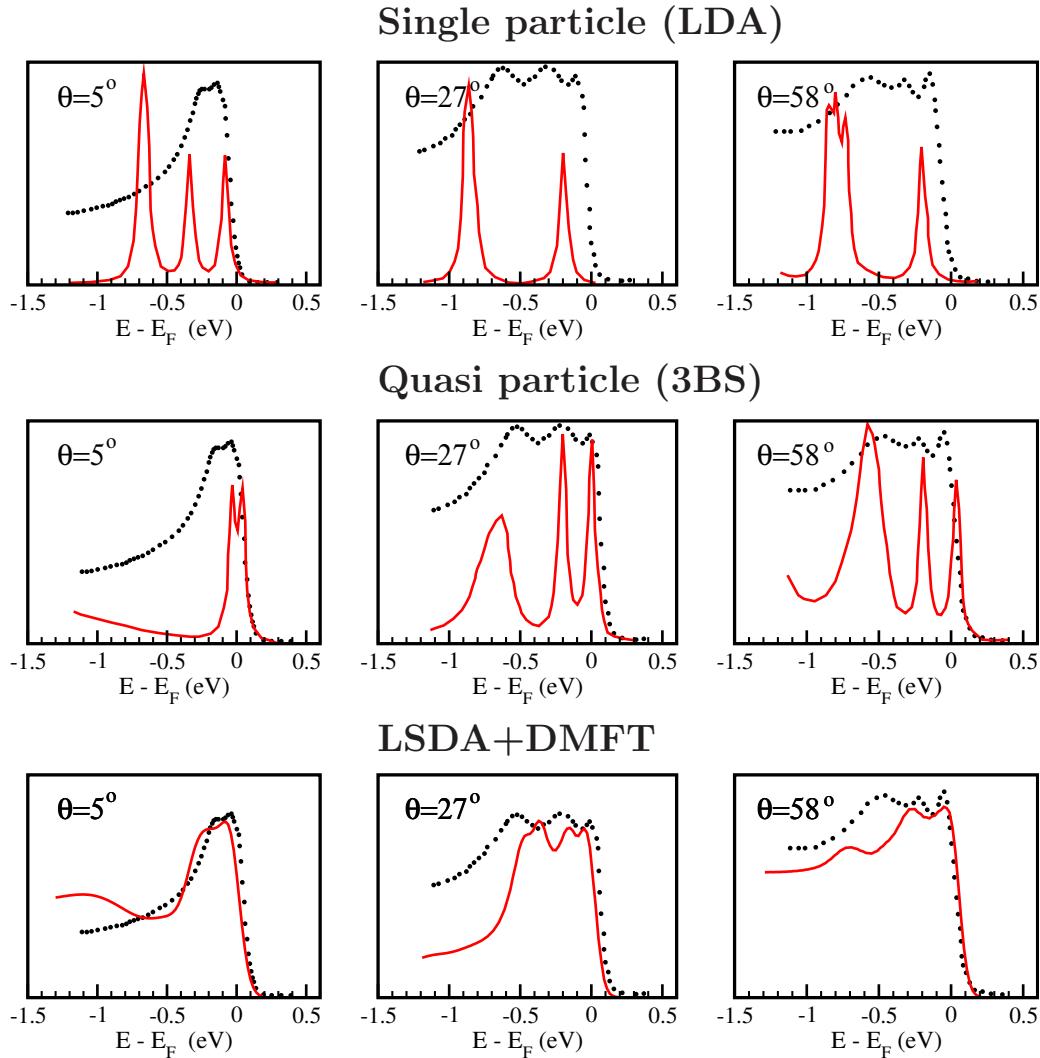


Fig. 5: Spin-integrated ARPES spectra from Ni(011) along $\bar{\Gamma}Y$ for three different angles of emission. Upper row: comparison between LSDA-based calculation and experiment [68]; middle row: comparison between experiment and non self-consistent quasi-particle calculations neglecting matrix element and surface effects [68]; lower row: spin-integrated LSDA+DMFT spectra including photoemission matrix elements (this work). Theory: solid red line, experiment: black dots. Figure reproduced from [38].

The following examples concern the ferromagnetic transition-metal systems Ni and Fe as prototype materials to study electronic correlations and magnetism beyond the LSDA scheme. In particular, the electronic structure of fcc Ni has been subject of numerous experimental [84–90] and theoretical studies [91–93] as a prototype of an itinerant electron ferromagnet, since shortcomings of simple one-electron theory are obvious. LSDA calculations for fcc Ni cannot reproduce various features of the electronic structure of Ni which had been observed experimentally. Besides the fact that valence band photoemission spectra of Ni [94–96] show a reduced 3d band width compared to LSDA calculations [97] the spectra show a dispersionless feature at a

binding energy (BE) of about 6 eV, the so-called 6 eV satellite [84, 85, 98–101], which is not reproducible within the LSDA approach. On the other hand, an improved description of correlation effects for the $3d$ electrons using many-body techniques [91, 92, 102] or in a more modern view applying the LSDA+DMFT scheme [43, 38] results more or less in the experimental width of the $3d$ -band complex and furthermore is able to reproduce the 6 eV satellite structure in the valence band region.

In Fig. 5 we present a comparison between experimental photoemission data [68] and calculated spectra using different theoretical approaches [38]. In the upper row spin-integrated ARPES spectra from Ni(011) along \overline{IY} for different angles of emission are shown. The dotted lines represent the experimental data, whereas the solid lines denote a single-particle representation of the measured spectral function. Obviously, the LSDA-based calculation completely fails to describe the experimental data. The energetic positions of the theoretical peaks deviate strongly from the measured ones. Furthermore, the complicated intensity distributions that appear for higher emission angles are not accounted for by the LSDA-based calculations. In contrast, the non self-consistent quasi-particle 3BS calculation provides a significant improvement when compared to the measured spectra. For the complete range of emission angles the energetic peak positions coincide with the experiment within about 0.1 eV. Only the overall shape of the measured spectral intensities deviate from the calculations because of the neglect of multiple scattering and surface-related as well as matrix-element effects. In the experiment the various peaks seem to be more broadened and the spectral weight especially for nearly normal emission is shifted by about 0.1 eV to higher binding energies. In addition it seems that for very high emission angles like 60° an even more complicated peak structure is hidden due to limited experimental resolution. The intensity distributions resulting from the corresponding photoemission calculation are shown in the lower row of Fig. 5. A first inspection reveals a very satisfying quantitative agreement between experiment and theory for all emission angles. Let us concentrate first on the excitation spectrum calculated for the emission angle $\Theta = 5^\circ$. The spin-integrated spectrum exhibits a pronounced double-peak structure with binding energies of 0.1 eV and 0.3 eV. The second peak is slightly reduced in intensity which is also in accordance with the experimental findings. Furthermore, the width of the spectral distribution is quantitatively reproduced. The calculated binding energies are related to the real part of the self-energy that corrects the peak positions due to a dynamical renormalization of the quasi-particles which is missing in a typical LSDA-based calculation. The relative intensities of the different peaks, on the other hand, must be attributed to the matrix-element effects which enter our calculations coherently via the one-step model of photoemission. The observed double-peak structure originates from excitation of the spin-split d -bands in combination with a significant amount of surface-state emission [103]. The two spectra calculated for high emission angles show the spectral distributions more broadened than observed in experiment. An explanation can be given in terms of matrix-element effects, due to the dominating dipole selection rules. The spin-resolved spectra reveal a variety of d -band excitations in both spin channels, which in consequence lead to the complicated shape of the spectral distributions hardly to be identified in the spin-integrated mode.

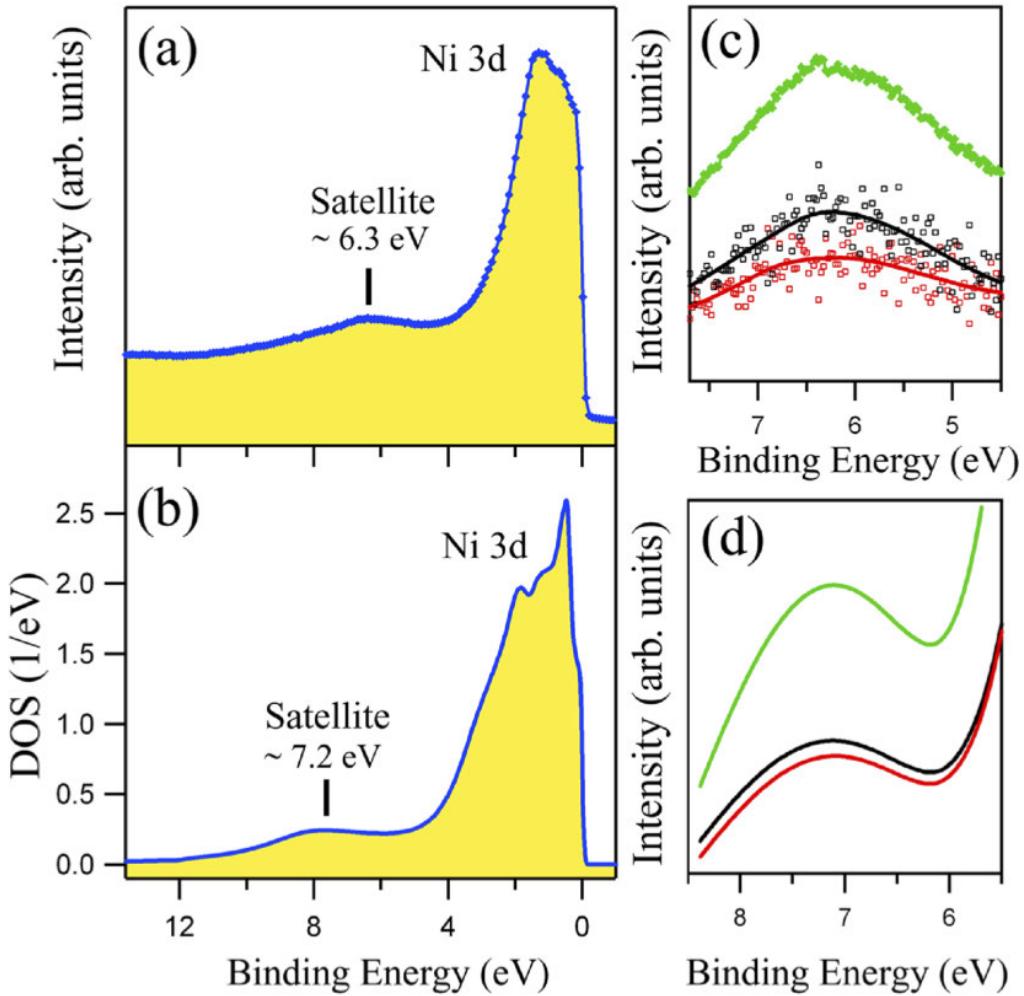


Fig. 6: (a) Experimental XPS-spectrum taken at $h\nu=150$ eV. The “Ni 6 eV satellite” structure appears at about 6.3 eV binding energy. (c) Spin- and angle-resolved photoemission spectra taken in normal emission at $h\nu=66$ eV with s-polarized light. Open black squares: majority spin states, open red squares: minority spin states, solid black and red lines serve as guides for the eyes. Spin-integrated intensity: green thick dotted line. (b) LSDA+DMFT calculation of the spin-integrated DOS. The satellite feature appears at about 7.2 eV binding energy. (d) LSDA+DMFT spin-resolved photoemission calculation in normal emission at $h\nu=66$ eV for a U value of 3.0 eV: solid black and red lines indicate majority and minority spin states, green line shows the spin-integrated intensity. Figure reproduced from [42].

The second example within this section concerns a spectroscopic study of the 6 eV satellite of Ni. As was shown by earlier calculations [102] and confirmed by photoemission experiments [87], the 6 eV satellite is spin-polarized. In a recent experimental study the XPS intensity at $h\nu=150$ eV as well as the spin- and angle-resolved photoemission spectra at $h\nu=66$ eV have been measured. Results for the second experiment are shown in Fig. 6(a) and Fig. 6(c).

The satellite feature is clearly visible at a BE of about 6.3 eV, fully in agreement with all former investigations. Furthermore, Fig. 6(c) shows the non-zero spin-polarization of the satellite in the spin-resolved experiment, again in full agreement with earlier studies [87]. After back-

ground subtraction, the spin polarization amounts to about 15%. In Fig. 6(b) we compare the experimental results with a DOS-calculation which is based on the LSDA+DMFT approach. The parametrization for $U=2.8$ eV and $J=0.9$ eV is identical with values that we used for the spin- and angle-resolved photoemission calculations. The satellite appears at a binding energy (BE) of ~ 7.2 eV. This is about 1 eV higher in BE than the experimental result. The explanation for this is found in terms of the many-body solver. The so called FLEX-solver [57] is based on perturbation theory while a more accurate quantum Monte Carlo solver is able to consider the complete diagrammatic expansion of the self-energy in a statistical way. As a consequence the energy dependence of the self-energy is less pronounced and this causes the observed shift of about 1 eV in the BE. Nevertheless, the satellite is observable in the calculated DOS and therefore one would expect its appearance in the theoretical photoemission intensity as well. The corresponding spin- and angle-resolved photoemission calculation is shown in Fig. 6(d). A weak intensity distribution in the vicinity of the *sp* band transition is present at a BE of ~ 7.2 eV. The green curve shows the spin-integrated intensity, whereas the black and red lines indicate the majority and minority spin related intensities. The calculated spin-polarization amounts to about 10% slightly smaller than the experimental one. Besides these small deviations between experiment and theory the agreement is very satisfying. Thus we show the first angle-resolved photoemission calculation for Ni metal in which this spectral feature appears.

In conclusion, we have presented a spectroscopical analysis for ferromagnetic Ni and Co, which coherently combine an improved description of electronic correlations with multiple-scattering, surface emission, dipole selection rules and other matrix-element related effects that lead to a modification of the relative photoemission intensities. As has been demonstrated, this approach allows on the one hand side a detailed and reliable interpretation of angle-resolved photoemission spectra of 3d-ferromagnets. On the other hand, it also allows for a very stringent test of new developments in the field of DMFT and similar many-body techniques.

The third example within this section concerns a spectroscopic study of ferromagnetic Fe [40]. In the left panel of Fig. 7 we compare the experimental peak positions from bulk-like transitions with spin-resolved LSDA+DMFT spectral functions. In addition to these investigations correlation effects were also accounted for within the 3BS approach [104]. Within the 3BS approach the self-energy is calculated using a configuration interaction-like expansion. In particular three-particle configurations like one hole plus one electron-hole pair are explicitly taken into account within 3BS-based calculations. The corresponding output can be directly related to the photoemission process and allows for a detailed analysis of various contributions to the self-energy (e.g., electron-hole lifetime). A more detailed quantitative comparison is shown in right panel of Fig. 7. Here we display a comparison between spin-integrated ARPES data and theoretical LSDA+DMFT based one-step photoemission calculations of Fe(110) along the Γ N direction of the bulk Brillouin zone (BZ) with *p*-polarized radiation. In our LSDA+DMFT investigation underlying the ARPES calculations we use for the averaged on-site Coulomb interaction U a value $U=1.5$ eV which lies between $U \approx 1$ eV deduced from experiment [105] and the value $U \approx 2$ eV obtained within other theoretical studies [106, 59]. From an analysis of the spectra, the k values associated with the observed transitions were determined using photon

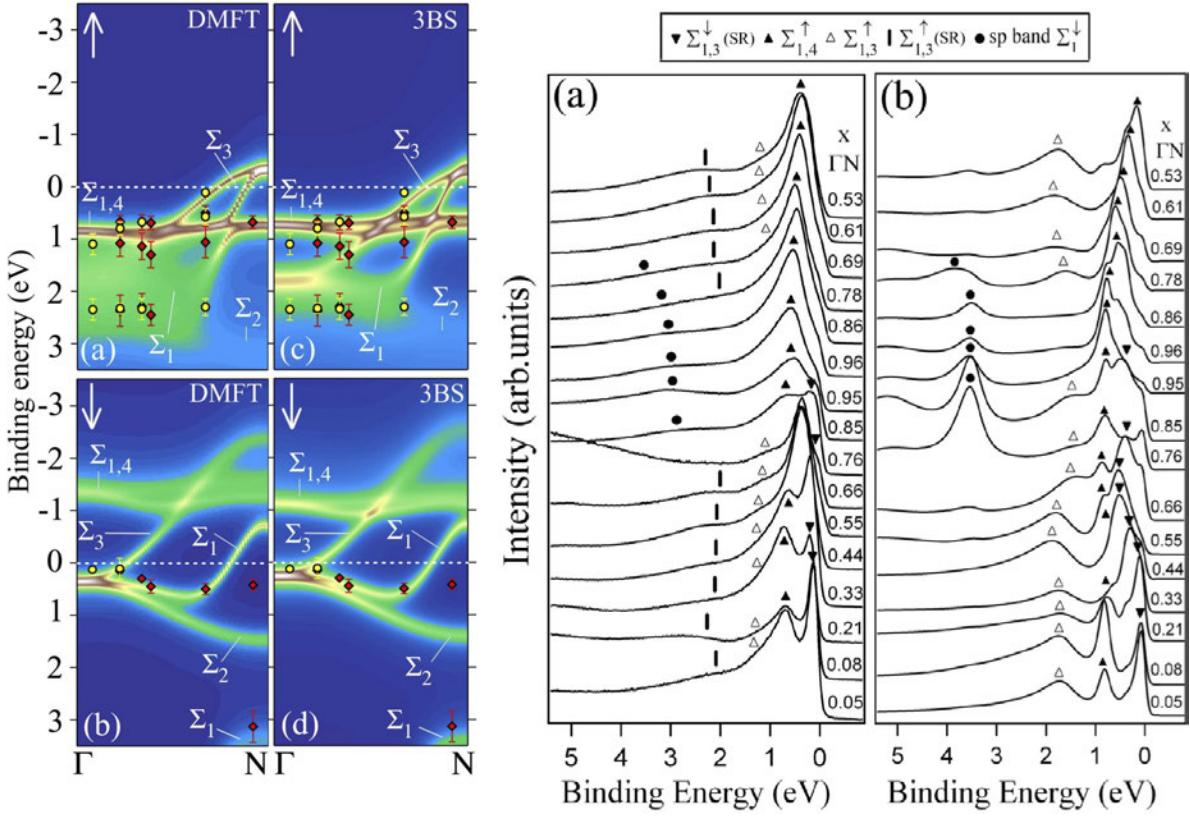


Fig. 7: Left panel: Spin resolved Bloch spectral functions calculated within LSDA+DMFT and 3BS formalism . Corresponding experimental data points have been deduced from the normal emission spectra along the ΓN direction. Right panel: (a) Experimental spin-integrated photoemission spectra of the Fe(110) surface measured with p -polarization in normal emission along the ΓN direction of the bulk Brillouin zone. The curves are labeled by the wave vectors in units of $\Gamma N=1.55 \text{ \AA}^{-1}$. (b) Corresponding one-step model calculations based on the LSDA+DMFT method which include correlations, matrix elements and surface effects. Figure reproduced from [40].

energies ranging from 25 to 100 eV. Near the Γ point ($k \sim 0.06 \Gamma N$), the intense peak close to the Fermi level corresponds to a $\Sigma_{1,3}^\downarrow$ minority surface resonance, as indicated in the top of Fig. 7. Experimentally, its Σ_3^\downarrow bulk component crosses the Fermi level at $k \sim 0.33 \Gamma N$, leading to a reversal of the measured spin-polarization and to a strong reduction of the intensity at $k = 0.68 \Gamma N$ in the minority channel. The peak at the binding energy $BE \sim 0.7 \text{ eV}$, visible mainly for p -polarization in a large range of wave vectors between Γ and N , can be assigned to almost degenerate $\Sigma_{1,4}^\uparrow$ bulk-like majority states. A Σ_3^\uparrow feature at $BE \sim 1.1 \text{ eV}$ dominates the spectrum close to the Γ -point. Depending on the polarization the degenerate Σ_1^\uparrow states form a shoulder around the same BE. The broad feature around 2.2 eV, visible at various k -points, but not at the N -point, is related to a majority $\Sigma_{1,3}^\uparrow$ surface state. Around the N -point ($0.76 \leq k \leq 1.0$) and at $BE \geq 3 \text{ eV}$ we observe a Σ_1^\downarrow band having strong sp character. The pronounced difference between its theoretical and experimental intensity distributions can be attributed to the fact that

in the present calculations only the local Coulomb repulsion between d electrons is considered, without additional lifetime effects for the sp bands. Finally, we notice that the background intensity of the spectrum at $k=0.66 \Gamma N$, corresponding to a photon energy of 55 eV, is strongly increased due to the appearance of the Fe $3p$ resonance. The direct comparison of the calculated and experimental spectra turned out to be a very stringent check for the Coulomb parameter U used in the calculations. This also applies to the DMFT self-energy, which was compared to its counterpart deduced from the experimental band dispersion and line width.

In summary, spectral function calculations for ferromagnetic Ni and Fe could be performed that coherently combine an improved description of electronic correlations, multiple-scattering, surface emission, dipole selection rules and other matrix-element related effects that lead to a modification of the relative photoemission intensities. A similar study has been performed recently for hcp Co(0001) [60] and fcc Co(001) [107]. The combined approach allows on the one hand side a detailed and reliable interpretation of high-resolution angle-resolved photoemission spectra of $3d$ -ferromagnets. On the other hand, it also allows for a stringent test of new developments in the field of DMFT and related many-body techniques.

4.3 ARPES of disordered correlated alloys: $\mathbf{Ni}_x\mathbf{Pd}_{1-x}$ (001)

In this section, alloying effects in combination with electronic correlations are considered [36]. Fig. 8 shows a series of spectra of $\mathbf{Ni}_x\mathbf{Pd}_{1-x}$ as a function of the concentration x calculated for a photon energy $h\nu = 40$ eV with linearly polarized light. The experimental data are shown in the left panel and the corresponding LSDA+DMFT-based photoemission calculations are presented in the right one. Our theoretical analysis shows that starting from the pure Ni, the agreement is fully quantitative with deviations less than 0.1 eV binding energy, as expected on the basis of the studies presented above. Going to the $\mathbf{Ni}_{0.80}\mathbf{Pd}_{0.20}$ alloy the agreement is comparably good for binding energies between the Fermi energy and 2 eV. Inspecting the density of states (DOS) for the $\mathbf{Ni}_{0.80}\mathbf{Pd}_{0.20}$ alloy this fact becomes explainable, because this energy interval represents the Ni-dominated region. The Pd derived states start to appear at about 2 eV below E_F next to the small dip at the Fermi level. For higher binding energies the agreement is also very good, although a bit more structure is observable in the theoretical spectra especially around 3.5 eV. An explanation for this behavior can be found in terms of lifetime effects. However, it should be mentioned here that the background in the experimental spectra due to secondary electrons was not considered for the theoretical spectra. From the results for $\mathbf{Ni}_{0.70}\mathbf{Pd}_{0.30}$ it becomes clearly visible that an increasing deviation between theory and experiment occurs with increasing Pd concentration. This can be seen from the spectra for $\mathbf{Ni}_{0.50}\mathbf{Pd}_{0.50}$ and $\mathbf{Ni}_{0.30}\mathbf{Pd}_{0.70}$ alloys shown next in the series. This is caused by the Pd d -states that seem to be slightly shifted to higher binding energies. This is well known from other paramagnetic metals like Ag and can be explained in terms of static correlations in the Pd-states not explicitly considered here. In addition, the spectra of $\mathbf{Ni}_{0.30}\mathbf{Pd}_{0.70}$ reveal some deviations near the Fermi level. Also, the spectral intensity of the Ni surface resonance, that appears at about 0.5 eV binding energy is underestimated in the calculation when compared to experiment.

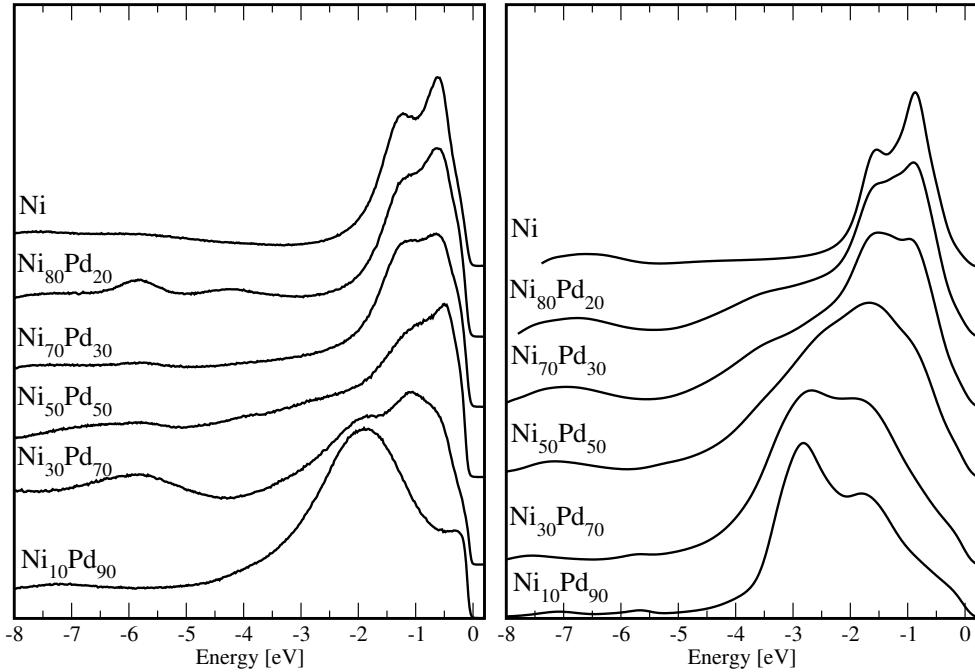


Fig. 8: ARPES spectra taken from the $\text{Ni}_x\text{Pd}_{1-x}(001)$ alloy surfaces as a function of the concentration x for a fixed photon energy of $h\nu = 40.0$ eV along ΓX in normal emission. Experimental data shown in the left panel calculated spectra presented in the right panel. Depending on the concentration x a pronounced shift in spectral weight towards the Fermi level is visible. Figure reproduced from [36].

Our spectroscopic analysis clearly demonstrates that the electronic properties of the $\text{Ni}_x\text{Pd}_{1-x}$ alloy system depend very sensitively on the interplay of alloying and electronic correlation. A description within the LSDA approach in combination with the CPA results in a quantitative description of the electronic structure of $\text{Ni}_x\text{Pd}_{1-x}$ [36]. This example may illustrate that the use of the CPA alloy theory self-consistently combined with the LSDA+DMFT approach serves as a powerful tool for electronic structure calculations, whereas the application of the fully relativistic one-step model of photoemission, which takes into account chemical disorder and electronic correlation on equal footing, guarantees a quantitative analysis of corresponding experimental spectroscopic data.

5 Angle resolved soft and hard X-ray photomission

It has always been realized, that the results obtained in UV ARPES are restricted in sensitivity to the near-surface region of the systems studied due to the short inelastic mean free paths (IMFPs) of $\sim 5\text{-}10$ Å of the low energy photoelectrons, which are typically in the range from 10-150 eV [108]. To overcome this limitation of surface sensitivity, there is now considerable interest in using x-rays in the soft x-ray sub-keV or even hard x-ray multi-keV regime to access deeper-lying layers in a sample, thus sampling more bulk-like properties [109–119, 27, 29, 30, 32]. One can thus think of soft x-ray or hard x-ray ARPES (HARPES), respectively. These techniques have

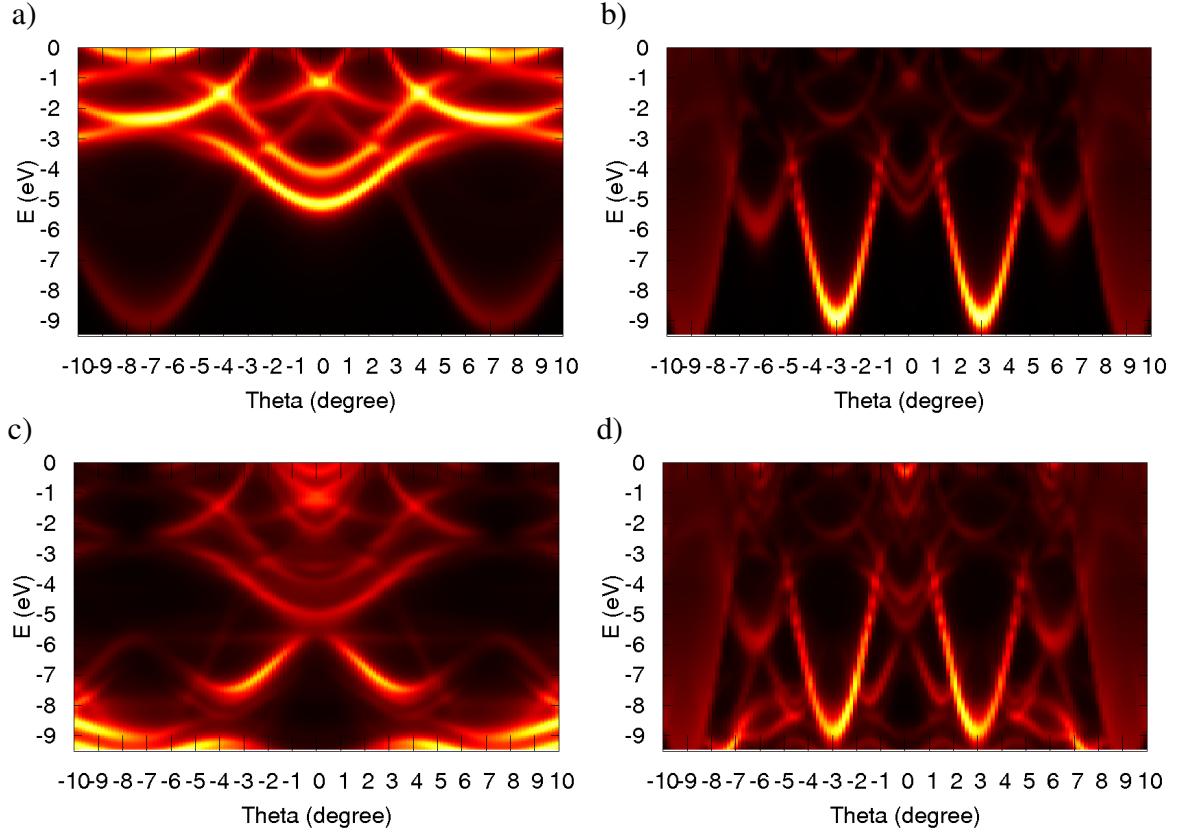


Fig. 9: Photoemission spectra calculated for clean Fe(001) (upper panel, a and b). Left side shows the intensity distribution obtained for 1 keV, right side represents the corresponding spectrum at 6 keV. The lower panels show theoretical spectra for the overlayer system 8MgO/Fe(001) at 1 keV (left side, c) and 6 keV (right side, d). Figure reproduced from [120].

to date been applied to a wide variety of materials, including free-electron like and transition-metals [109, 116], strongly correlated oxides and high T_C materials [111, 112], heavy fermion systems [111], mixed-valent Ce compounds [114], dilute magnetic semiconductors [29, 30, 121], layered transition-metal dichalcogenides [119]. Additional advantages in such experiments include being able to tune to core-level resonances so as to identify the atomic-orbital makeup of ARPES features [121], to map three-dimensional Fermi surfaces [119], and to take advantage of the longer IMFPs, which translate into less smearing of dispersive features along the emission direction (usually near the surface normal) [118]. Increasing the photon energy means increasing the bulk sensitivity of the corresponding photo emission data. We demonstrate this effect by comparing spectra obtained by photoelectron excitation with 1 keV and 6 keV radiation from the clean Fe(001) surface and from the overlayer system MgO/Fe(100) with 8 ML MgO on Fe. Fig. 9a presents the spectrum for the clean Fe(001) surface and a photon energy of 1 keV. As expected for this photon energy regime, the bulk sensitivity is enhanced and the surface emission is reduced to a negligible extent. Also the relative intensity fraction of the sp -bands is obviously increased. Going to a much higher photon energy of 6 keV the d -band intensity is strongly reduced when compared with sp -band related features. This is shown in Fig. 9b. In a next step we put 8 ML of MgO on the Fe(001) surface and repeat our calculations for both

photon energies. This is shown in the lower panel of Fig. 9. At 1 keV mostly the MgO bulk band-structure is visible in Fig. 9c. This is due to the thickness of the MgO film which consists of 8 ML. Increasing the photon energy to 6 keV we expect due to the much larger mean free path length of the photo-electron that Fe-related features will reappear, namely the Fe *sp*-states. This is clearly demonstrated by Fig. 9d which represents the corresponding photo emission spectrum calculated for the 8ML MgO/Fe(001) system at a photon energy of 6 keV. Besides the effect that more than one Brillouin zone is visible for a fixed escape angle regime due to the very high photon energy it is undoubtly observable that nearly all MgO related features have vanished in the intensity plot. This, for example, should give access to buried interfaces.

Going higher in energy, however, comes with some additional challenges for interpretation of the data [109, 115, 118]. Deviations from the dipole approximation in photoelectron excitations mean that the momentum of the photon can result in a non-negligible shift of the position of the initial-state wave vector in the reduced Brillouin zone (BZ) [116]. Also, phonon creation and annihilation during photoemission hinders the unambiguous identification of the initial state in the BZ via wave vector conservation [109, 110, 115, 116, 27, 118, 29].

5.1 Photon momentum effects: Ag(001)

Here we want to discuss the effect of the photon momentum \mathbf{q} on the intensity distribution of the photocurrent. The impact of the photon momentum on the initial state \mathbf{k} -vector is expressed by the following equation

$$\mathbf{k}_i = \left(\mathbf{k}_{||} - \mathbf{q}_{||} + \mathbf{g}, \sqrt{2(E - iV_{i1}) - |\mathbf{k}_{||} - \mathbf{q}_{||} + \mathbf{g}|^2} \right), \quad (32)$$

where V_{i1} denotes the imaginary part of the intial state energy E . Due to the relatively high photon energies the role played by the photon momentum is no longer negligible, with the importance of its influence depending on the chosen experimental geometry. Considering as an example photo emission from Ag(001), for $\phi_{ph} = 45^\circ$ the photon momentum has no component along the [110] crystallographic direction. However, the effect of the photon momentum along the [110] direction is to kick the photo-electron in a direction that is perpendicular to the probed high symmetry direction. This effect can experimentally be corrected for by rotating the crystal surface by a small amount, $\Theta = 0.7^\circ$, with respect to the entrance plane of the electron analyzer, thereby minimizing the effect of the photon momentum transfer along the [110] direction. However, one has to pay the price that the high symmetry plane, in this case the Γ XUL plane, is no longer the plane from which the emission takes place. A careful analysis reveals that an angle $\Theta = 0.7^\circ$ corresponds, for photon energies in the 500 eV to 600 eV range, to an effective variation of the angle $\pm 7^\circ$ for the given range of the emission angle θ ($\mathbf{q}_{||}$ effect). Fortunately, the largest deviation appears for nearly normal emission and for higher emission angles it approaches zero. Therefore, the total average deviation from the high symmetry plane is small and the experiment mainly represents emission from the Γ LUX plane with the added advantage of a minimized $\mathbf{q}_{||}$ effect. Fig. 10 shows the effect of both the correction angle Θ and the photon momentum $\mathbf{q}_{||}$. In Fig. 10(a) the intensity distribution has been calculated for

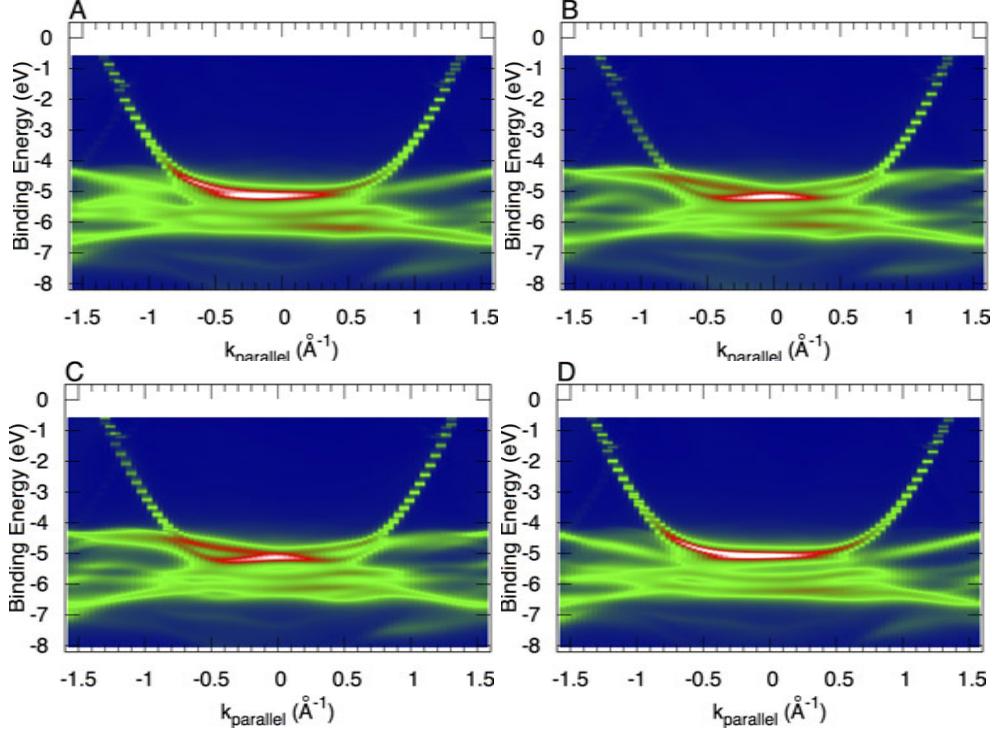


Fig. 10: Theoretical photo emission intensities for the Γ -K high symmetry direction Σ calculated for $h\nu=552$ eV, $\phi = 45^\circ$. To allow for a one-to-one comparison between experimental and theoretical data, all the theoretical data have been shifted by 1.2 eV. (a) Intensity distribution calculated for $\mathbf{q}_{||} = 0$ and $\Theta = 0^\circ$. (b) Intensity distribution calculated for a nonzero $\mathbf{q}_{||}$ vector and $\Theta = 0^\circ$. (c) Intensity distribution calculated for $\mathbf{q}_{||} = 0$ and $\Theta = 0.7^\circ$. (d) Intensity distribution calculated for a nonzero value of $\mathbf{q}_{||}$ and $\Theta = 0.7^\circ$. This corresponds to the experimental geometry setup. Figures taken from [116].

$\Theta = |\mathbf{q}_{||}| = 0$, whereas in Figs. 10(b) and 10(c) nonzero values for $\mathbf{q}_{||}$ and Θ have been used. It is observable that in both panels 1(b) and 1(c) the band dispersion is more pronounced than in Fig. 10(a), where the bands appear. This effect, together with an additional asymmetric intensity distribution around Γ , is caused by the small deviation from the desired high symmetry plane mentioned above.

Fig. 10(d) represents the experimental situation with $\Theta = 0.7^\circ$ and a nonzero $\mathbf{q}_{||}$ value. Observable is the similarity between panels 10(a) and 10(d). Therefore, we can conclude that the experimental procedure works in a satisfactory way. The measurements performed by Venturini *et al.* [116] were taken at $T = 20$ K for a photon energy of $h\nu = 552$ eV, that corresponds to the Γ and the X symmetry points along the direction that is perpendicular to the samples surface. The data sets are measured with right circularly polarized light. The results are shown in the left panels of Fig. 11 for $\phi = 0^\circ$ and $\phi = 45^\circ$, respectively. For $\phi = 0^\circ$, the parallel component of the initial-state wave vector $\mathbf{k}_{||}$ varies along the Δ direction, whereas for $\phi = 45^\circ$ the Σ direction is probed. The BZ boundaries along these two directions are found at $\mathbf{k}_{||} \approx 1.54 \text{ \AA}^{-1}$ and $\mathbf{k}_{||} \approx 1.63 \text{ \AA}^{-1}$, respectively, and a photon beam of energy $h\nu = 552$ eV allows to probe almost the entire BZ along these directions. The experimental results presented in Fig. 11 are in good qualitative agreement with our fully relativistic one-step model photo emission calculations shown in the right panels of the respective figures.

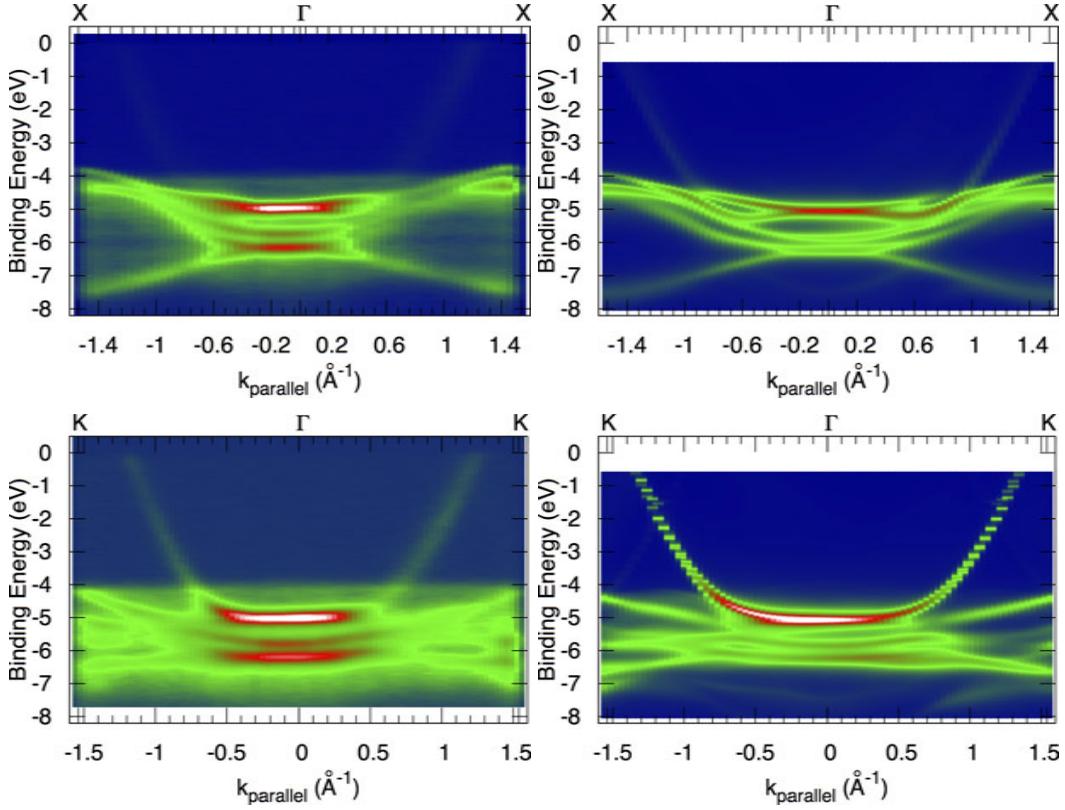


Fig. 11: Γ -X high symmetry direction Δ probed with $h\nu = 552$ eV, $\phi = 0^\circ$ (top) and 45° (bottom), at $T = 20$ K. The energy refers to the Fermi level E_F . Left: experimental data. Right: theoretical results. Figures taken from [116].

Along the Δ and Σ high symmetry directions, direct transitions originating from all the allowed initial states are visible close to the BZ center, except for the two deeper lying bands along both directions. In particular, for both investigated orientations the $\Gamma'_{25} \rightarrow \Gamma_8^+ + \Gamma_7^+$ spin-orbit splitting is observed at Γ . The agreement between our calculated binding energies of the high symmetry points along the Γ -X direction [116] and results previously determined in Ref. [122] is very good. Also the agreement with the corresponding experimental values found in literature [123, 122, 124] is good, with a maximum deviation of 0.22 eV for the most tightly bound X_6^+ energy level. The spin-orbit split level with X_7^+ symmetry is barely visible at the BZ boundary along the Δ direction (Fig. 11). Its binding energy along this direction is about 4.3 eV (see Ref. [116]). The ARPES data presented so far show evidence of the fact that, for well-defined combinations of $h\nu$ and temperature, direct transitions in the soft x-ray regime can indeed be observed. If compared to the low energy ARPES, the combination of a larger k -space sampling and a reduced curvature of the investigated path, together with the use of a two-dimensional position sensitive detection system, allow measuring the band structure along specific high symmetry directions with a single measurement. This has been done by Venturini *et al.* [116] for four different high symmetry directions in the BZ of fcc Ag. The corresponding results, which were obtained for $T = 20$ K, are also in good agreement with our fully relativistic one-step model photo emission calculations.

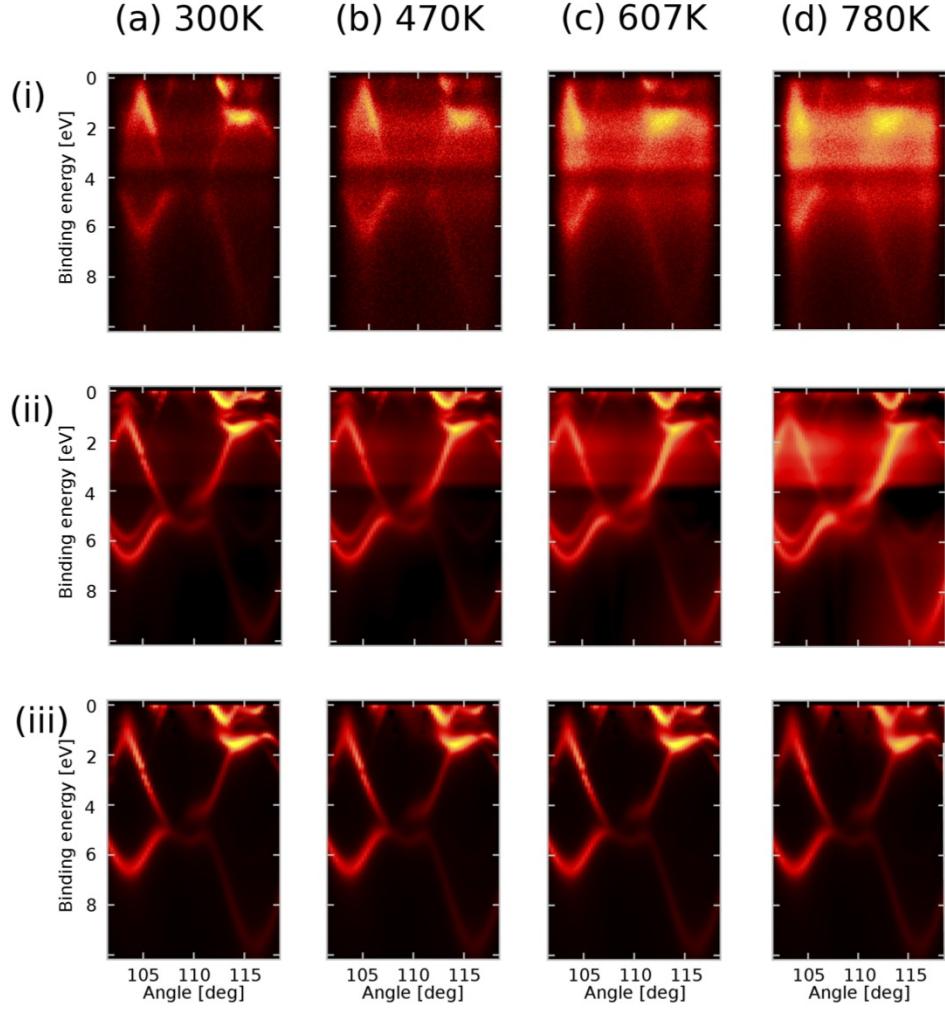


Fig. 12: (i) Plots of measured intensity versus angle of emission for 870 eV excitation from the valence bands of W(110) approximately along the Γ -N direction for four temperatures of (a) 300 K, (b) 470 K, (c) 607 K, and (d) 870 K (from [115]), where 90 deg corresponds to normal emission. (ii) Corresponding intensity distributions calculated from temperature-dependent one-step theory based on the CPA formulation. (iii) Conventional ARPES calculations of the direct contribution $I^{\text{DT}}(E, \mathbf{k})$ by use of complex scattering phase shifts and the Debye-Waller model. Figure reproduced from [35].

5.2 Thermal effects and XPS limit

Going higher in energy, however, comes with some additional challenges for the interpretation of the data [109, 115, 118]. Deviations from the dipole approximation in photoelectron excitation mean that the momentum of the photon can result in a non-negligible shift of the position of the initial-state wave vector in the reduced Brillouin zone (BZ), as first pointed out some time ago. Phonon creation and annihilation during photoemission also hinders the unambiguous specification of the initial state in the BZ via wave vector conservation [118, 115, 116, 27, 29]. Following Shevchik [125], the photoemission intensities at a given energy E and vector \mathbf{k} can be approximately divided into zero-phonon direct transitions $I^{\text{DT}}(E, \mathbf{k})$ and phonon-assisted

non-direct transitions $I^{\text{NDT}}(E, \mathbf{k})$. As a rough guide to the degree of direct-transition behavior expected in an ARPES experiment, one can use a temperature-dependent Debye-Waller factor $W(T)$ which qualitatively represents the fraction of direct transitions [118]. As a first approach that aimed to go beyond this simple scheme for temperature-dependent ARPES, Larsson and Pendry [126] introduced a model called Debye-Waller model later on that incorporates the effect of lattice vibrations on the photoemission matrix elements. More than 15 years later Zampieri *et al.* [127] introduced a cluster approach to model the temperature-dependent excitation of valence band electrons for photon energies of about 1 keV. More recently Fujikawa and Arai [73] discussed phonon effects on ARPES spectra on the basis of nonequilibrium Green's function theory. Recently, we presented a new approach which accurately models phonon effects over the full energy range from normal low-energy ARPES to HARPES. More importantly it converges for high temperatures and/or photon energies to the so called XPS-limit in photoemission, in particular the development of matrix-element weighted density-of-states (MEW-DOS)-like features in the intensity distribution [27, 28]. Our alloy analogy model includes vibrational atomic displacements via the coherent potential approximation (CPA), where vibrations of different lattice sites are assumed to be uncorrelated and averaged in the sense of CPA over various possible displacements which are calculated within Debye theory. Using the CPA-formulation of the one-step model [26, 20, 36] provides a self-consistent temperature-dependent averaging of the photoemission matrix elements. In other words, we describe in a quantitative sense the breakdown of the \mathbf{k} -conserving rules due to phonon-assisted transitions, the driving mechanism that leads finally to the XPS-limit. In order to demonstrate this effect below we present an example of soft-x-ray ARPES calculations for W(110). In Fig. 12, we compare results of our calculations directly to experimental data for W(110) with soft x-ray excitation at 870 eV [115]. W has a Debye temperature of 400 K and a atomic mass of 183.84 u, close to Au and Pt. In Fig. 12(i), we show experimental results for four different sample temperatures: (a) 300 K, (b) 470 K, (c) 607 K and (d) 780 K [115]. For all four temperatures, dispersive features are clearly seen but with significant smearing and an increase of MEW-DOS-like intensity features as temperature is raised. Also shown in Fig. 13(a), (b) are vertical and horizontal cuts, respectively, through the 2D data of Fig. 12(i). These cuts yield Fig. 13(a): energy distribution curves (EDCs) and Fig. 13(b): momentum distribution curves (MDCs) to illustrate more directly the changes in both types of distributions with temperature. Also, various spectral features are labeled by the numbers 1-6 in these figures. Fig. 12(ii) again presents fully relativistic one-step calculations which are done with our new alloy analogy model, whereas Fig. 12(iii) shows conventional one-step calculations in which phonon excitations are considered in a simplified way through a temperature-dependent single-site scattering matrix [77]. Although at the lowest temperature of 300 K the two different theoretical approaches yield very similar results, as expected for a Debye-Waller factor of 0.70, the temperature dependence of the experimental data is much better described by our temperature-dependent one-step calculations. The simpler calculation based on the single-site scattering matrix predicts neither the smearing of dispersing features nor the growth of MEW-DOS features for higher binding energies, but shows instead only the monotonous decrease of direct transition intensities with increasing temperature [115].

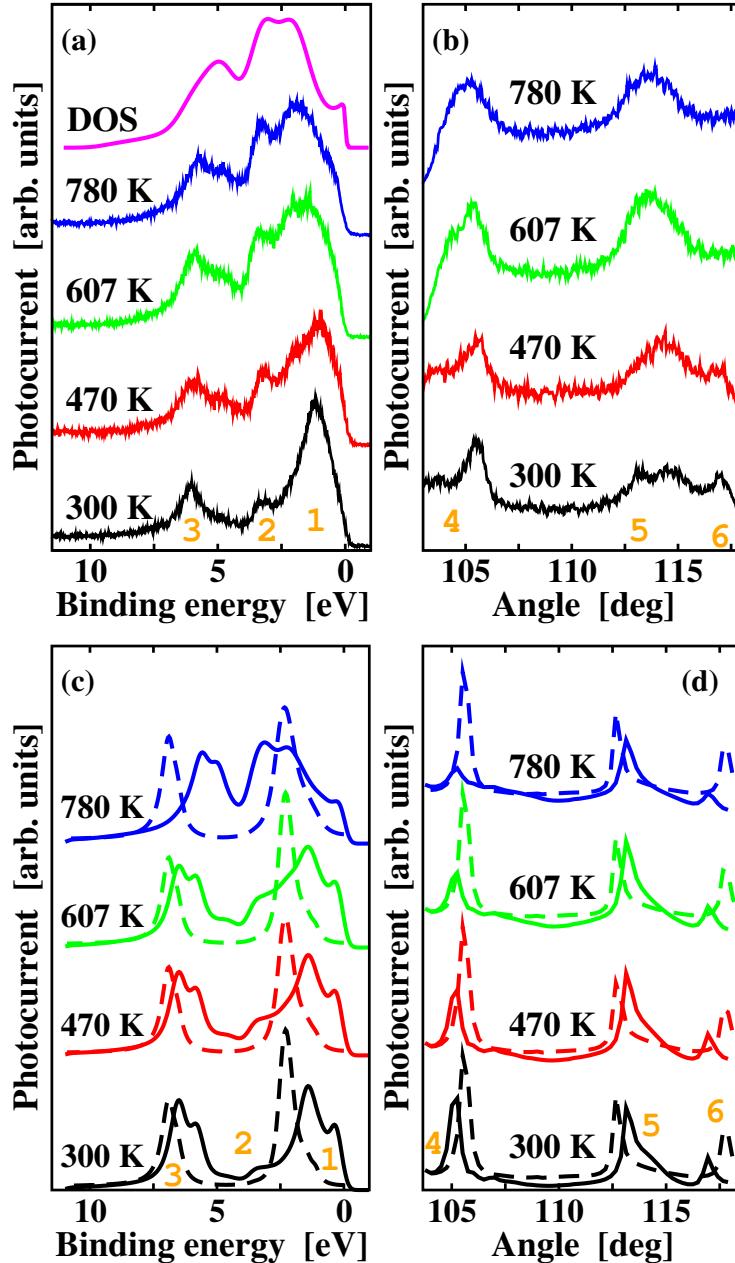


Fig. 13: (a) Measured temperature-dependent energy distribution curves (EDCs). A comparison to the W DOS (the topmost curve) is also given. (b) Measured temperature-dependent momentum distribution curves (MDCs). (c), (d) Corresponding theoretical results for (c) EDCs and (d) MDCs. Dashed lines indicate conventional one-step calculations, solid lines indicate calculations within the new alloy analogy model. Figure reproduced from [35].

Phonon induced smearing only appears via temperature-dependent matrix elements which cause a decrease of the direct part of the photocurrent due to a redistribution of spectral weight. Although for 780 K and a photon energy of 870 eV the XPS-limit is not fully established for W, the indirect contribution of the temperature-dependent CPA-like photocurrent dominates the corresponding angle-resolved soft x-ray spectra. This is clearly observable from both the experimental and theoretical data, which are nearly in quantitative agreement.

Figs. 13(a) and (c) compare in more detail the temperature dependence of experimental and theoretical spectra in the form of EDCs for a fixed angle of ≈ 104 deg which is 14 deg from the surface normal. Figs. 13(b) and (d) present the same comparison for MDCs at a fixed binding energy of ≈ 2 eV. The points labeled 1, 2, 4, 5 and 6 denote *d*-like electronic states, whereas point 3 labels bands that are more free-electron like and a mixture of *s* and *d* states. The experimental and theoretical data in Fig. 12 show pronounced smearing of features in both EDCs and MDCs as the temperature is raised, but some remnant direct-transition behavior is clearly still present, even at 780 K. The dashed lines shown in Fig. 13(c) and (d) indicate conventional one-step calculations using the one-site scattering matrix approach. As expected, only slight changes appear in the form of the EDCs and MDCs as a function of temperature. In contrast, the EDCs and MDCs strongly depend on temperature when using the alloy analogy approach, although in general the MDCs in experiment and those of the conventional and CPA approaches change less than the EDCs. Significant broadening of spectral features and shift of spectral weight, not at all present in the conventional one-site scattering matrix approach, can be observed. The EDC at the highest temperature has not converged to a MEW-DOS-like curve and the corresponding MDC still has structure in it. In the XPS-limit all MDCs would exhibit only x-ray photoelectron diffraction (XPD), with a different type of angular distribution [115, 29, 32, 127, 28]. This is obviously not the case. This is expected, because the Debye-Waller factor of 0.41 at 780 K indicates that a certain number of transitions should still be direct. Our calculations thus correctly predict a diminuation of the features expected due to direct transitions, and also a significant broadening of features in the EDCs or MDCs. The additional weak and smooth background observed in the experimental data thus must be ascribed to additional phonon effects, perhaps through multiple phonon excitations.

6 Summary

The implementation of the LSDA+DMFT on the basis of the KKR method has been described in some detail. The appealing feature of this approach is that the KKR delivers the one-electron Green's function directly. It therefore allows to combine the treatment of correlations via the DMFT with calculation of a great variety of physical properties for, in principle, any type of system. Within the one-step model of photoemission, this method can be applied to directly calculate ARPES in the wide photon energy range. As was demonstrated by results for photo emission spectra of various transition-metal systems, this allows in particular a direct comparison with experiment.

Acknowledgement

I would like to thank Jürgen Braun and Hubert Ebert for their stimulating discussions and detailed comments and suggestions to this manuscript. This work was supported by the Deutsche Forschungsgemeinschaft through FOR 1346, by the BMBF (Project 05K13WMA) and CEN-TEM (CZ.1.05/2.1.00/03.0088).

References

- [1] F. Reinert and S. Hüfner: *Photoemission Spectroscopy with Very High Energy Resolution* Lecture Notes in Physics, Vol. 715 (Springer, Berlin, 2007) p. 13
- [2] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003)
- [3] I. Adawi, Phys. Rev. **134**, A788 (1964)
- [4] G.D. Mahan, Phys. Rev. B **2**, 4334 (1970)
- [5] W. Schaich and N. Ashcroft, Solid State Commun. **8**, 1959 (1970)
- [6] C. Caroli, D. Lederer-Rozenblatt, B. Roulet, and D. Saint-James, Phys. Rev. B **8**, 4552 (1973)
- [7] P.J. Feibelman and D.E. Eastman, Phys. Rev. B **10**, 4932 (1974)
- [8] C.N. Berglund and W.E. Spicer, Phys. Rev. **136**, A1030 (1964)
- [9] A. Liebsch, Phys. Rev. Lett. **32**, 1203 (1974)
- [10] D.J. Spanjaard, D.W. Jepsen, and P.M. Marcus, Phys. Rev. B **15**, 1728 (1977)
- [11] J.B. Pendry, Surf. Sci. **57**, 679 (1976)
- [12] J.F.L. Hopkinson, J.B. Pendry, and D.J. Titterington, Comp. Phys. Commun. **19**, 69 (1980)
- [13] J.B. Pendry: *Low energy electron diffraction* (Academic Press, London, 1974)
- [14] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B 864 (1964)
- [15] W. Kohn and L.J. Sham, Phys. Rev. **140**, A 1133 (1965)
- [16] L.J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1966)
- [17] J. Braun, Rep. Prog. Phys. **59**, 1267 (1996)
- [18] B. Ackermann and R. Feder, J. Phys. C: Solid State Phys. **18**, 1093 (1985)
- [19] J. Braun, G. Thörner, and G. Borstel, phys. stat. sol. (b) **130**, 643 (1985)
- [20] B. Ginatempo, P.J. Durham, and B.I. Gyorffy, J. Phys.: Condens. Matter **1**, 6483 (1989)
- [21] S.V. Halilov, E. Tamura, D. Meinert, H. Gollisch, and R. Feder, J. Phys.: Condens. Matter **5**, 3859 (1993)
- [22] J. Braun and G. Borstel, Phys. Rev. B **48**, 14373 (1993).

- [23] M. Fluchtmann, M. Grass, J. Braun, and G. Borstel, Phys. Rev. B **52**, 9564 (1995)
- [24] J. Henk, T. Scheunemann, S.V. Halilov, and R. Feder, J. Phys.: Condens. Matter **8**, 47 (1996)
- [25] M. Grass, J. Braun, and G. Borstel, Prog. Surf. Sci. **46**, 107 (1994)
- [26] P.J. Durham, J. Phys. F: Met. Phys. **11**, 2475 (1981)
- [27] A.X. Gray, C. Papp, S. Ueda, B. Balke, Y. Yamashita, L. Plucinski, J. Minár, J. Braun, E.R. Ylvisaker, C.M. Schneider, W.E. Pickett, H. Ebert, K. Kobayashi, and C.S. Fadley, Nature Materials **10**, 759 (2011)
- [28] C. Papp, L. Plucinski, J. Minár, J. Braun, H. Ebert, C.M. Schneider, and C.S. Fadley, Phys. Rev. B **84**, 045433 (2011)
- [29] A.X. Gray, J. Minár, S. Ueda, P.R. Stone, Y. Yamashita, J. Fujii, J. Braun, L. Plucinski, C.M. Schneider, G. Panaccione, H. Ebert, O.D. Dubon, K. Kobayashi, and C.S. Fadley, Nature Materials **11**, 957 (2012)
- [30] J. Fujii, B.R. Salles, M. Sperl, S. Ueda, M. Kobata, K. Kobayashi, Y. Yamashita, P. Torelli, M. Utz, C.S. Fadley, A.X. Gray, J. Braun, H. Ebert, I. Di Marco, O. Eriksson, P. Thunström, G.H. Fecher, H. Stryhanyuk, E. Ikenaga, J. Minár, C.H. Back, G. van der Laan, and G. Panaccione, Phys. Rev. Lett. **111**, 097201 (2013)
- [31] J. Minár, J. Braun, and H. Ebert, J. Electron. Spectrosc. Relat. Phenom. **189**, 129 (2013)
- [32] J. Minár, J. Braun, and H. Ebert, J. Electron. Spectrosc. Relat. Phenom. **190**, 159 (2013)
- [33] A.X. Gray, J. Minár, L. Plucinski, M. Huijben, A. Bostwick, E. Rotenberg, S. Yang, J. Braun, A. Winkelmann, G. Conti, D. Eiteneer, A. Rattanachata, A.A. Greer, J. Ciston, C. Ophus, G. Rijnders, D.H.A. Blank, D. Doennig, R. Pentcheva, J.B. Kortright, C.M. Schneider, H. Ebert, and C.S. Fadley, Europhys. Lett. **104**, 17004 (2013)
- [34] J. Braun, K. Miyamoto, A. Kimura, T. Okuda, M. Donath, H. Ebert, and J. Minár, New J. Phys. **16**, 015005 (2014)
- [35] J. Braun, J. Minár, S. Mankovsky, V.N. Strocov, N.B. Brookes, L. Plucinski, C.M. Schneider, C.S. Fadley, and H. Ebert, Phys. Rev. B **88**, 205409 (2013)
- [36] J. Braun, J. Minár, F. Matthes, C.M. Schneider, and H. Ebert, Phys. Rev. B **82**, 024411 (2010)
- [37] J. Minár, L. Chioncel, A. Perlov, H. Ebert, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **72**, 045125 (2005)
- [38] J. Braun, J. Minár, H. Ebert, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. Lett. **97**, 227601 (2006)

- [39] M. Pickel, A.B. Schmidt, F. Giesen, J. Braun, J. Minár, H. Ebert, M. Donath, and M. Weinelt, Phys. Rev. Lett. **101**, 066402 (2008)
- [40] J. Sánchez-Barriga, J. Fink, V. Boni, I. Di Marco, J. Braun, J. Minár, A. Varykhalov, O. Rader, V. Bellini, F. Manghi, H. Ebert, M.I. Katsnelson, A.I. Lichtenstein, O. Eriksson, W. Eberhardt, and H.A. Dürr, Phys. Rev. Lett. **103**, 267203 (2009)
- [41] J. Braun, J. Minár, H. Ebert, A. Chainani, J. Miyawaki, Y. Takata, M. Taguchi, M. Oura, and S. Shin, Phys. Rev. B **85**, 165105 (2012)
- [42] J. Sánchez-Barriga, J. Braun, J. Minár, I. Di Marco, A. Varykhalov, O. Rader, V. Boni, V. Bellini, F. Manghi, H. Ebert, M.I. Katsnelson, A.I. Lichtenstein, O. Eriksson, W. Eberhardt, H.A. Dürr, and J. Fink, Phys. Rev. B **85**, 205109 (2012)
- [43] J. Minár, J. Phys.: Condens. Matter **23**, 253201 (2011)
- [44] H. Ebert, D. Ködderitzsch, and J. Minár, Rep. Prog. Phys. **74**, 096501 (2011)
- [45] J.S. Faulkner and G.M. Stocks, Phys. Rev. B **21**, 3222 (1980)
- [46] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, and C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
- [47] K. Held, Adv. Phys. **56**, 829 (2007)
- [48] S.Y. Savrasov and G. Kotliar, Phys. Rev. B **69**, 245101 (2004)
- [49] P. Strange, H. Ebert, J.B. Staunton, and B.L. Gyorffy, J. Phys.: Condens. Matter **1**, 2959 (1989)
- [50] A.H. MacDonald and S.H. Vosko, J. Phys. C: Solid State Phys. **12**, 2977 (1979)
- [51] R.M. Dreizler and E.K.U. Gross: *Density Functional Theory* (Springer-Verlag, Heidelberg, 1990)
- [52] H. Ebert, S. Bornemann, J. Braun, D. Ködderitzsch, S. Lowitzer, S. Mankovskyy, J. Minár, M. Offenberger, S. Polesya, , and V. Popescu: *Recent Developments in KKR Theory* http://www.psi-k.org/newsletters/News_97/Highlight_97.pdf
- [53] A. Gonis and W.H. Butler: *Multiple scattering in solids*
- [54] E. Tamura, Phys. Rev. B **45**, 3271 (1992)
- [55] E.N. Economou: *Green's Functions in Quantum Physics*, Springer Series in Solid-state Sciences, Vol. 7 (Springer, Berlin, 1990)
- [56] R. Zeller, J. Deutz, and P.H. Dederichs, Solid State Commun. **44**, 993 (1982)

- [57] L.V. Pourovskii, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. B **72**, 115106 (2005)
- [58] S. Chadov: *Application of Many-Body Perturbation Theory to the Description of Correlated Metals*
- [59] S. Chadov, J. Minár, M.I. Katsnelson, H. Ebert, D. Ködderitzsch, and A.I. Lichtenstein, Europhys. Lett. **82**, 37001 (2008)
- [60] J. Sánchez-Barriga, J. Minár, J. Braun, A. Varykhalov, V. Boni, I. Di Marco, O. Rader, V. Bellini, F. Manghi, H. Ebert, M.I. Katsnelson, A.I. Lichtenstein, O. Eriksson, W. Eberhardt, H.A. Dürr, and J. Fink, Phys. Rev. B **82**, 104414 (2010)
- [61] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003)
- [62] P. Soven, Phys. Rev. **156**, 809 (1967)
- [63] J.S. Faulkner, Prog. Mater. Sci. **27**, 1 (1982)
- [64] V. Drchal, V. Janiš, and J. Kudrnovský, Phys. Rev. B **60**, 15664 (1999)
- [65] I. Turek, V. Drchal, J. Kudrnovský, M. Sob, and P. Weinberger: *Electronic structure of disordered alloys, surfaces and interfaces* (Kluwer Academic Publ., Boston, 1997)
- [66] O. Šipr, J. Minár, S. Mankovsky, and H. Ebert, Phys. Rev. B **78**, 144403 (2008)
- [67] K. Held, I.A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A.K. McMahan, R.T. Scalettar, T. Pruschke, V.I. Anisimov, and D. Vollhardt, phys. stat. sol. (b) **243**, 2599 (2006)
- [68] F. Manghi, V. Bellini, J. Osterwalder, T.J. Kreutz, P. Aebi, and C. Arcangeli, Phys. Rev. B **59**, R10409 (1999)
- [69] K. Kambe, Z. Naturf. **22a**, 322 (1967)
- [70] J.M. MacLaren, S. Crampin, D.D. Vvedensky, and J.B. Pendry, Phys. Rev. B **40**, 12164 (1989)
- [71] G. Malmström and J. Rundgren, Comp. Phys. Commun. **19**, 263 (1980)
- [72] T. Fujikawa and H. Arai, J. Elec. Spec. Rel. Phenom. **123**, 19 (2002)
- [73] T. Fujikawa and H. Arai, J. Electron. Spectrosc. Relat. Phenom. **174**, 85 (2009)
- [74] G. Borstel, Appl. Physics A **38**, 193 (1985)
- [75] J. Braun: *New developments in UP- and XP-spectroscopy from ferromagnetic materials, in Band-Ferromagnetism* (Springer, Berlin, 2001), p. 341
- [76] G. Hilgers, M. Potthoff, N. Muller, U. Heinzmann, L. Haunert, J. Braun, and G. Borstel, Phys. Rev. B **52**, 14859 (1995)

- [77] R. Feder, J. Phys. C: Solid State Phys. **14**, 2049 (1981)
- [78] M. Potthoff, J. Lachnitt, W. Nolting, and J. Braun, phys. stat. sol. (b) **203**, 441 (1997)
- [79] H. Ebert, Rep. Prog. Phys. **59**, 1665 (1996)
- [80] U. Fano, Phys. Rev. **178**, 131 (1969)
- [81] J. Minár, H. Ebert, C. De Nadaï, N.B. Brookes, F. Venturini, G. Ghiringhelli, L. Chioncel, M.I. Katsnelson, and A.I. Lichtenstein, Phys. Rev. Lett. **95**, 166401 (2005)
- [82] J. Minár, H. Ebert, G. Ghiringhelli, O. Tjernberg, N.B. Brookes, and L.H. Tjeng, Phys. Rev. B **63**, 144421 (2001)
- [83] C. De Nadaï, J. Minár, H. Ebert, G. Ghiringhelli, A. Tagliaferri, and N.B. Brookes, Phys. Rev. B **70**, 134409 (2004)
- [84] S. Hüfner and G.K. Wertheim, Phys. Letters A **51A**, 299 (1975)
- [85] C. Guillot, Y. Ballu, J. Paigné, J. Lecante, K.P. Jain, P. Thiry, R. Pinchaux, Y. Pétroff, and L.M. Falicov, Phys. Rev. Lett. **39**, 1632 (1977)
- [86] A. Kakizaki, K. Ono, K. Tanaka, K. Shimada, and T. Sendohda, Phys. Rev. B **55**, 6678 (1997)
- [87] K.N. Altmann, D.Y. Petrovykh, G.J. Mankey, N. Shannon, N. Gilman, M. Hochstrasser, R.F. Willis, and F.J. Himpsel, Phys. Rev. B **61**, 15661 (2000)
- [88] J. Osterwalder, J. Elec. Spec. Rel. Phenom. **117-118**, 71 (2001)
- [89] N. Kamakura, Y. Takata, T. Tokushima, Y. Harada, A. Chainani, K. Kobayashi, and S. Shin, Phys. Rev. B **74**, 045127 (2006).
- [90] O. Karis, S. Svensson, J. Rusz, J. Oppeneer, M. Gorgoi, F. Schäfers, W. Braun, W. Eberhardt, M. Mårtensson, Phys. Rev. B **78**, 233105 (2008)
- [91] A. Liebsch, Phys. Rev. Lett. **43**, 1431 (1979)
- [92] A. Liebsch, Phys. Rev. B **23**, 5203 (1981)
- [93] O. Miura and T. Fujiwara, Phys. Rev. B **77**, 195124 (2008)
- [94] D.E. Eastman, F.J. Himpsel, and J.A. Knapp, Phys. Rev. Lett. **40**, 1514 (1978)
- [95] F.J. Himpsel, J.A. Knapp, and D.E. Eastman, Phys. Rev. B **19**, 2919 (1979)
- [96] W. Eberhardt and E.W. Plummer, Phys. Rev. B **21**, 3245 (1980)
- [97] C.S. Wand and J. Callaway, Phys. Rev. B **15**, 298 (1977)

- [98] J. Kanski, P.O. Nilsson, and C.G. Larsson, Solid State Commun. **35**, 397 (1980)
- [99] Y. Sakisaka, T. Kommeda, M. Onchi, H. Kato, S. Masuda, and K. Yagi, Phys. Rev. B **36**, 6383 (1987)
- [100] T. Kinoshita, T. Ikoma, A. Kakizaki, T. Ishii, J. Fujii, H. Fukutani, K. Shimada, A. Fujimori, T. Okane, and S. Sato, Phys. Rev. B **47**, 6787 (1993)
- [101] N. Nakajima, S. Hatta, J. Odagiri, H. Kato, and Y. Sakisaka, Phys. Rev. B **70**, 233103, (2004)
- [102] A.I. Lichtenstein, M.I. Katsnelson, and G. Kotliar, Phys. Rev. Lett. **87**, 067205 (2001)
- [103] W. Eberhardt, E.W. Plummer, K. Horn, and J. Erskine, Phys. Rev. Lett. **45**, 273 (1980)
- [104] C. Calandra and F. Manghi, Phys. Rev. B **50**, 2061 (1994)
- [105] M.M. Steiner, R.C. Albers, and L.J. Sham, Phys. Rev. B **45**, 13272 (1992)
- [106] M. Cococcioni and S. de Gironcoli, Phys. Rev. B **71**, 035105 (2005)
- [107] T. Allmers, M. Donath, J. Braun, J. Minár, and H. Ebert, Phys. Rev. B **84**, 245426 (2011)
- [108] S. Tanuma, C.J. Powell, and D.R. Penn, Surface and Interface Analysis **43**, 689 (2011)
- [109] Z. Hussain, C.S. Fadley, S. Kono, and L.F. Wagner, Phys. Rev. B **22**, 3750 (1980)
- [110] E.H. Sondheimer, Adv. Phys. **50**, 499 (2001)
- [111] A. Sekiyama and S. Suga, J. Electron Spectrosc. Relat. Phenom. **137**, 681 (2004)
- [112] A. Yamasaki, A. Sekiyama, S. Imada, M. Tsunekawa, A. Higashiya, A. Shigemoto, S. Suga, Nucl. Instr. and Meth. Phys. Res. A **547**, 136 (2005)
- [113] T. Yokoya, T. Nakamura, T. Matsushita, T. Muro, Y. Takano, M. Nagao, T. Takenouchi, H. Kawarada, and T. Oguchi, Nature **438**, 647 (2005)
- [114] M. Yano, A. Sekiyama, H. Fujiwara, T. Saita, S. Imada, T. Muro, Y. Onuki, and S. Suga, Phys. Rev. Lett. **98**, 036405 (2007)
- [115] L. Plucinski, J. Minár, B.C. Sell, J. Braun, H. Ebert, C.M. Schneider, and C.S. Fadley, Phys. Rev. B **78**, 035108 (2008)
- [116] F. Venturini, J. Minár, J. Braun, H. Ebert, and N.B. Brookes, Phys. Rev. B **77**, 045126 (2008)
- [117] S. Suga and A. Sekiyama, J. Electron Spectrosc. Relat. Phenom. **181**, 48 (2010)
- [118] C.S. Fadley, Synchrotron Radiation News **25**, 26 (2012)

- [119] V.N. Strocov, M. Shi, M. Kobayashi, C. Monney, X. Wang, J. Krempasky, T. Schmitt, L. Patthey, H. Berger, and P. Blaha, Phys. Rev. Lett. **109**, 086401 (2012)
- [120] J. Minár, J. Braun, S. Mankovsky and H. Ebert, J. Electron Spectrosc. Relat. Phenom. **184**, 91 (2011)
- [121] M. Kobayashi, I. Muneta, Y. Takeda, Y. Harada, A. Fujimori, J. Krempasky, T. Schmitt, S. Ohya, M. Tanaka, M. Oshima, and V.N. Strocov, arxiv.1302.0063v1
- [122] R. Courts, V. Bachelier, and S. Hüfner, Solid State Commun. **38**, 887 (1981)
- [123] P.S. Wehner, R.S. Williams, S.D. Kevan, D. Denley, and D.A. Shirley, Phys. Rev. B **19**, 6164 (1979)
- [124] J.G. Nelson, S. Kim, W.J. Gignac, R.S. Williams, J.G. Tobin, S.W. Robey, and D.A. Shirley, Phys. Rev. B **32**, 3465 (1985)
- [125] N.J. Shevchik, Phys. Rev. B **16**, 3428 (1977)
- [126] C.G. Larsson and J.B. Pendry, J. Phys. C: Solid State Phys. **14**, 3089 (1981)
- [127] M.A.V. Alvarez, H. Ascolani, and G. Zampieri, Phys. Rev. B **54**, 14703 (1996)

14 Introduction to Photoemission Spectroscopy

Michael Sing

Physikalisches Institut and

Röntgen Center for Complex Material Systems

Universität Würzburg

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | The method | 2 |
| 2.1 | Basics | 2 |
| 2.2 | Many-body picture | 6 |
| 3 | Case studies | 8 |
| 3.1 | Low-energy photoemission: Doping a one-dimensional Mott insulator | 8 |
| 3.2 | Hard x-ray photoemission: Profiling the buried two-dimensional electron system in an oxide heterostructure | 12 |
| 3.3 | Resonant angle-resolved soft x-ray photoemission: Direct k -space mapping of the electronic structure in an oxide-oxide interface | 14 |
| 4 | Outlook | 16 |

1 Introduction

Complex quantum materials are distinguished by their astonishingly huge variety of interesting and often peculiar electronic and magnetic properties which arise from the interplay of charge, spin, lattice, and orbital degrees of freedom. Such phenomena comprise, e.g., ferro-, pyro- and piezoelectricity, all kinds of magnetic order, colossal magnetoresistance, high-temperature superconductivity, and metal-insulator transitions. In order to arrive at a microscopic understanding of such diverse behavior the leading low-energy scales of the material under consideration have to be explored. To those, spectroscopic methods grant *direct* access by probing either low-lying single-particle or charge-neutral particle-hole and collective excitations. The former is realized, e.g., in photoelectron spectroscopy (PES) – lying at the core of this chapter – while the latter typically is implemented in scattering techniques. On the theory side, these two types of spectroscopic information correspond to the physical content of the one-particle and two-particle Green's functions, respectively. Since PES is related to the simpler one-particle Green's function and extremely versatile in that it can be applied to almost all kinds of solids it has assumed a prominent role among solid-state spectroscopies over the years, in particular whenever many-particle physics is important.

As a well established method PES is the subject of numerous monographs and review articles dealing with all kind of related aspects such as instrumentation, application to atoms, molecules, and solids, and the theoretical description [1–12]. However, it would be entirely wrong to believe that photoemission spectroscopy its theoretical understanding and implementation, is completely developed. The full calculation of photoemission spectra still represents a challenging task (cf. the chapter by J. Minar in this book) and necessitates progressively advanced methods from theory while technological evolution and innovation have made it possible, *inter alia*, to partially overcome the notorious surface sensitivity of photoemission with acceptable concessions to resolution and acquisition times.

In this chapter, after an introduction to the basics of photoemission spectroscopy the present potential of PES, in particular with regard to some modern techniques with enhanced volume sensitivity, shall be illustrated based on selected examples of complex material systems whose quantitative theoretical description often demands the advanced methods that are presented in most of the other chapters of this book. For convenience, the examples are taken from our own work.

2 The method

2.1 Basics

In a PES experiment monochromatic light is directed onto a sample. The emitted photoelectrons are discriminated with respect to their kinetic energy and, depending on the information desired, other observables like emission direction or spin, before they are detected and counted. The principle is sketched in Fig. 1, left, for an angle-resolved experiment on a single crystal.

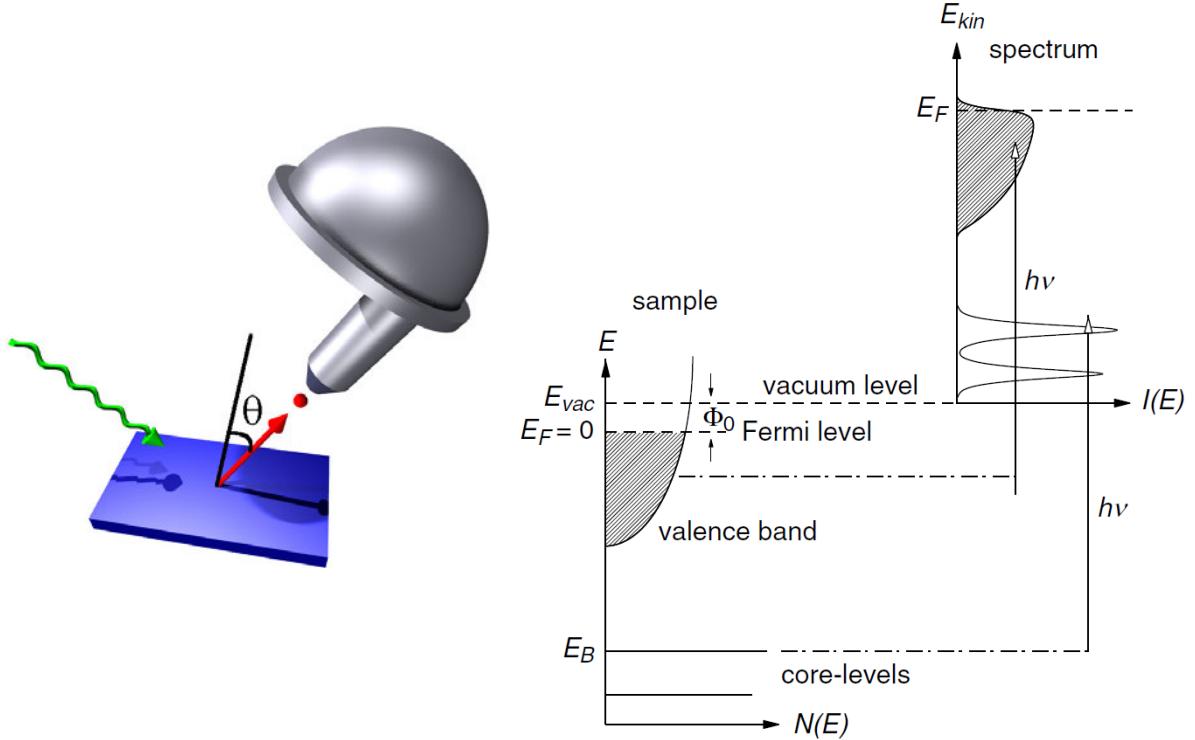


Fig. 1: Left: Geometry for an angle-resolved photoemission experiment. Right: Energy diagram of photoemission in a one-particle picture (from [9]).

An energy diagram of photoemission in a one-particle picture is sketched in Fig. 1, on the right. Electrons with a binding energy E_B are excited above the vacuum level E_{vac} by the absorption of a single photon with energy $h\nu$. Their kinetic energy is given by

$$E_{kin} = h\nu - E_B - \Phi_0 , \quad (1)$$

where Φ_0 denotes the work function. The kinetic energy distribution of the photoelectrons in vacuum $I(E_{kin})$ then reflects essentially the occupied part of the electronic structure, i.e., the density of states in the solid (weighted by the corresponding single-particle transition matrix elements, cf. section 2.2).

If, in addition, the photoelectrons are discriminated with respect to their emission direction relative to the surface of a single-crystalline sample, the momentum \mathbf{p} is completely determined in terms of its components parallel and perpendicular to the surface. For the parallel component we have

$$\mathbf{p}_{||} = \hbar \mathbf{k}_{||} = \sqrt{2mE_{kin}} \sin \theta . \quad (2)$$

Note that $\mathbf{p}_{||}$ – due to momentum conservation – equals the parallel component of the *crystal momentum* of the electron inside the solid *in the extended zone scheme*. In contrast, due to the lack of translational symmetry perpendicular to the sample surface, the perpendicular component of crystal momentum is *not* conserved. Hence, without additional information the crystal momentum cannot be determined completely. To do so, knowledge or assumptions about the dispersion of the photoemission final states are needed (cf. Fig. 2(a)), or one resorts to advanced

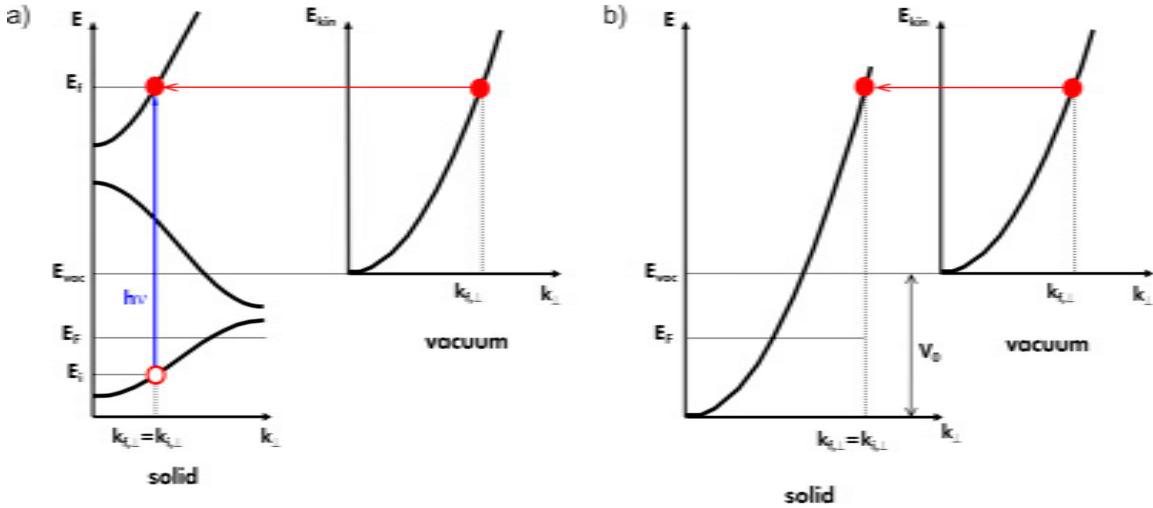


Fig. 2: Kinematics of the photoemission process. (a) Direct optical transition of an electron within the solid into a certain final state and energy of the corresponding photoelectron in vacuum. (b) Free-electron approximation for the final states in the solid and inner potential V_0 .

experimental methods which, however, are only feasible in certain cases. An often used approximation is the assumption of free electron final states. The situation is depicted in Fig. 2(b). The vertex of the parabola, describing the dispersion of the free electron final states inside the solid, is shifted along the energy axis to an energy V_0 below the vacuum level. Thus, V_0 is a measure of the depth of the potential well in a “particle-in-a-box” picture. The inner potential is a phenomenological parameter and has to be adjusted for a given material, e.g., such that the periodicity of the electron dispersions in reciprocal space with respect to k_\perp are reproduced for a series of PES spectra, taken in normal emission geometry (i.e., at $k_{||} = 0$) at various photon energies (thereby changing k_\perp). Outside the solid, the reference energy for the vertex of the free-electron parabola is simply the vacuum level. With a proper choice of V_0 one can simply read off the perpendicular component of the crystal momentum as shown in Fig. 2. If one takes the inner potential into account, the equation for the perpendicular component of the crystal momentum reads

$$\mathbf{p}_\perp = \hbar \mathbf{k}_\perp = \sqrt{2m(E_{kin} \cos^2 \theta + V_0)}. \quad (3)$$

Note that for angle-integrated measurements or in the case of one- or two-dimensional systems the determination of k_\perp becomes irrelevant.

Using Eq. (1), Eq. (2) and, if appropriate, Eq. (3), the energy-momentum relations of electronic excitations can be inferred from the angle- or momentum-resolved energy distribution curves (EDCs) of a PES experiment by simply tracing the energy positions of marked features as is illustrated in Fig. 3(a) and (b). In the case of a non-interacting electron system these correspond to the electronic band dispersions of Bloch states (see Fig. 3(a)), in an interacting system, e.g., in a Fermi liquid, to the dispersions of quasiparticle excitations (see Fig. 3(b)). In addition, the width of the spectral features in the interacting case reflects the finite lifetime of the respective excitations and also the coupling of other degrees of freedom to the electron system which can simultaneously be excited when a photoelectron is kicked off. This is illustrated in Fig. 3(c) for

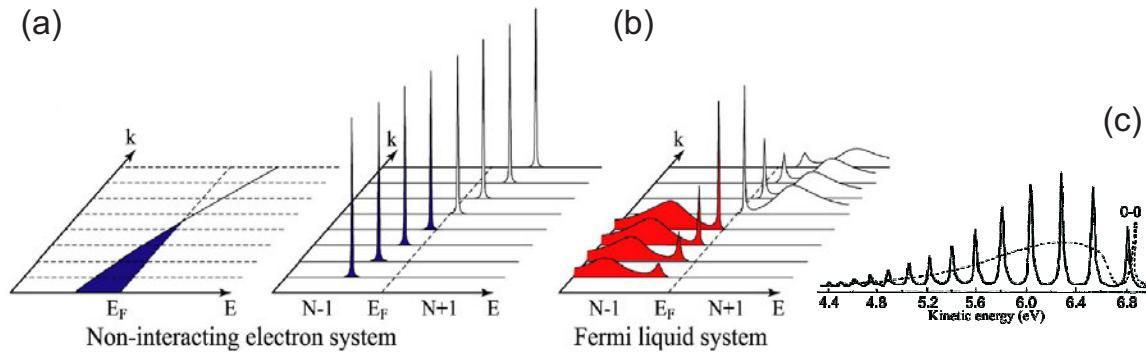


Fig. 3: Angle-resolved photoelectron spectroscopy: (a) Momentum-resolved energy distribution curves for a non-interacting electron system with a single band, crossing the Fermi energy E_F . (b) Same situation but for an interacting Fermi liquid (figure adapted from Ref. [8]) together with (c) the PES spectrum of hydrogen in the gas phase and the hypothetical spectrum for solid hydrogen (dashed line, figure adapted from Ref. [13]).

the PES spectrum of gaseous hydrogen. By kicking off a photoelectron, oscillations of the H_2 molecule are excited at the same time. Hence, besides the sharp line at about 6.8 eV as expected for the excitation from the groundstate of the rigid molecule a series of satellite lines are seen which correspond to excited vibrational states. If this spectrum is broadened, one arrives at a hypothetical spectrum (dashed line) which may serve as a paradigm of a PES spectrum for an interacting solid. It consists of a sharp, coherent quasiparticle excitation and a broad, incoherent contribution, representing the interaction of the kicked out electron with all possible excitations of the system such as phonons, magnons, spin-fluctuations, electron-hole pairs, etc.

To study the valence band of solids in the lab, usually gas discharge lamps are used whose line spectra cover the spectral range of about 10–50 eV ((AR)UPS – (angle-resolved) ultraviolet photoelectron spectroscopy). In this energy range the inelastic mean free path (IMFP) of electrons in solids, λ_{INFP} , amounts – due to the large cross section regarding plasmon excitations – only to a few Ångströms. For this reason, PES is extremely surface-sensitive.

For larger energies (>50 eV) for all materials the inelastic mean free path roughly behaves as $\lambda_{\text{INFP}} \propto \sqrt{E}$. Hence, the information depth of PES can be significantly enhanced if excitation energies in the soft ($\lambda_{\text{INFP}} \sim 15$ Å) or hard x-ray regime ($\lambda_{\text{INFP}} \sim 40$ –100 Å) are used.¹ However, the trade-off is a strong decrease of the photoemission signal since the photoionisation cross sections within the dipole approximation scale roughly like E^{-3} . Nevertheless, during the last years high-resolution PES with soft (SXPES – soft x-ray photoelectron spectroscopy) and hard x-rays (HAXPES – hard x-ray photoelectron spectroscopy) became feasible at third generation synchrotrons. Some of the case studies presented in the following rely on the use of high photon energies beyond the standard range. If appropriate, special variants of the application of PES like resonant PES will be discussed in the respective context.

¹An enhancement of the information depth can also be achieved for excitation energies <10 eV. In this regime the IMFP is determined mainly by interband transitions which are strongly material-dependent. Hence, an increase of the information depth is not universal for this energy range.

2.2 Many-body picture

The simplest starting point for a theoretical description of photoemission is Fermi's Golden Rule. The photocurrent I results from the photoexcitation (with energy $h\nu$), described by the appropriate perturbation operator (\mathcal{H}_{int}), of the N particle system in its ground state $|\Phi_0(N)\rangle$ into all possible final states $|\Phi_{\kappa,n}(N)\rangle$. The final states also describe a system with N particles, one of them being the photoelectron with momentum $\hbar\kappa$ and energy $(\hbar\kappa)^2/2m$. The index n denotes a complete set of quantum numbers defining all possible excitations in the final state. Hence the expression of the photocurrent is given by

$$I_{\kappa}(h\nu) = \frac{2\pi}{\hbar} \sum_n |\langle \Phi_{\kappa,n}(N) | \mathcal{H}_{\text{int}} | \Phi_0(N) \rangle|^2 \delta(E_{\kappa,n}(N) - E_0(N) - h\nu). \quad (4)$$

The operator \mathcal{H}_{int} describes the interaction of the photon field with a single electron within first-order perturbation theory. In second quantization it reads

$$\mathcal{H}_{\text{int}} = \sum_{i,j} \langle \mathbf{k}_i | \mathcal{H}_{\text{int}} | \mathbf{k}_j \rangle c_i^\dagger c_j = \sum_{i,j} M_{ij} c_i^\dagger c_j. \quad (5)$$

In the corresponding description based on the Schrödinger equation the explicit representation of \mathcal{H}_{int} is obtained by the canonical replacement of momentum according to $\mathbf{p} \rightarrow \mathbf{p} - e\mathbf{A}$, where \mathbf{A} is the vector potential of the photon field

$$\mathcal{H}_{\text{int}} = -\frac{e}{2m}(\mathbf{A} \cdot \mathbf{p} + \mathbf{p} \cdot \mathbf{A}) + \frac{e^2}{2m}\mathbf{A}^2. \quad (6)$$

The term quadratic in \mathbf{A} is only important in the case of very high photon intensities and can usually be neglected even when employing highly brilliant synchrotron radiation. Using the commutator relation $[\mathbf{p}, \mathbf{A}] = -i\hbar\nabla\mathbf{A}$ and under the assumption $\nabla\mathbf{A} \approx 0$, which corresponds to the dipole approximation valid for typical photon energies in the vacuum ultraviolet,² the perturbation operator can be simplified further, yielding

$$\mathcal{H}_{\text{int}} = -\frac{e}{m}\mathbf{A} \cdot \mathbf{p}. \quad (7)$$

The matrix element in Eq. (4) can be further evaluated in the case of a factorized final state. It then can be written as a product of the state of the photoelectron and the state of the remaining $(N - 1)$ particle system

$$|\Phi_{\kappa,n}(N)\rangle = c_{\kappa}^\dagger |\Phi_n(N-1)\rangle. \quad (8)$$

This approximation, known as *sudden approximation* is all but trivial. From a physical point of view, it means that the photoelectron is removed instantaneously, i.e., relaxation processes during photoemission are neglected. In general, it is assumed that this assumption is well justified for photon energies of several tens of eV and beyond [14]. From Eq. (4) with Eq. (5) one gets

$$I_{\kappa}(h\nu) = \frac{2\pi}{\hbar} \sum_j |M_{\kappa j}|^2 \sum_n |\langle \Phi_n(N-1) | c_j | \Phi_0(N) \rangle|^2 \delta(\varepsilon_{\kappa} + E_n(N-1) - E_0(N) - h\nu). \quad (9)$$

²At surfaces, due to the discontinuity of the dielectric function, this assumption is not necessarily fulfilled, which can lead to special surface effects.

Introducing the chemical potential, $\mu = E_0(N) - E_0(N-1)$, and with $\varepsilon = \varepsilon_{\kappa} - h\nu - \mu$ (i.e., energies are negative with respect to the chemical potential) the final expression for the photocurrent reads

$$\begin{aligned} I_{\kappa}(h\nu) &= \frac{2\pi}{\hbar} \sum_j |M_{\kappa j}|^2 \sum_n |\langle \Phi_n(N-1) | c_j | \Phi_0(N) \rangle|^2 \delta(\varepsilon + E_n(N-1) - E_0(N-1)) \\ &= \frac{2\pi}{\hbar} \sum_j |M_{\kappa j}|^2 A^{<}(\mathbf{k}_j, \varepsilon), \end{aligned} \quad (10)$$

where $A^{<}(\mathbf{k}_j, \varepsilon)$ is the electron removal spectral function at $T = 0$. Under the assumption that the one-particle matrix elements $M_{\kappa j}$ conserve momentum and are constant, angle-resolved photoemission thus essentially measures the momentum-resolved spectral function $A^{<}(\kappa, \varepsilon)$. The spectral function has a simple and instructive interpretation in terms of a quantum mechanical measurement. Kicking out an electron from the groundstate prepares the system in an excited $(N-1)$ particle state. This hole state, in general, is not an eigenstate of the $(N-1)$ particle system. Hence, when predicting the result of an energy measurement in photoemission this state has to be projected onto the eigenstates $|\Phi_0(N-1)\rangle$. Their relative contribution to the prepared hole state is then sampled in a PES experiment. The δ function simply ensures energy conservation.

In other words, the spectral function describes the probability to remove an electron from the system with wavevector \mathbf{k} and energy ε . In the case of independent electrons $c_j |\Phi_0(N)\rangle$ in Eq. (10) is exactly an eigenstate of the $(N-1)$ particle system, and $A^{<}(\mathbf{k}, \varepsilon)$ becomes a δ function. Integration over \mathbf{k} yields the one-particle density of states. For integration over ε the sum rule $\int_{-\infty}^{\infty} d\varepsilon A(\mathbf{k}, \varepsilon) = 1$ applies, when the spectral function is generalized also to the case that a particle is added to the system.

For a theoretical treatment the spectral function is in most cases only of limited use since a direct calculation in principle requires the knowledge of all excited states of the $(N-1)$ particle system. The major role of photoemission for the investigation of interacting electron systems stems from the fact that the spectral function is simply related to the already mentioned one-particle Green's function $G(\mathbf{k}, \varepsilon)$, for the calculation of which there exist many powerful methods. The relation is

$$A^{<}(\mathbf{k}, \varepsilon) = -\frac{1}{\pi} \text{Im } G(\mathbf{k}, \varepsilon - i0^+) f(\varepsilon, T), \quad (11)$$

where $f(\varepsilon, T)$ is the Fermi-Dirac distribution.

The Green's function that contains the full dynamics of an interacting many-particle system can be expressed in terms of the single-particle energy $\varepsilon_{\mathbf{k}}^0$ and a complex self-energy $\Sigma(\mathbf{k}, \varepsilon) = \Sigma'(\mathbf{k}, \varepsilon) + i\Sigma''(\mathbf{k}, \varepsilon)$

$$G(\mathbf{k}, \varepsilon) = \frac{1}{\varepsilon - \varepsilon_{\mathbf{k}}^0 - \Sigma(\mathbf{k}, \varepsilon)}. \quad (12)$$

The self-energy contains all interaction and correlation effects. With the self-energy Eq. (11) can be rewritten as

$$A(\mathbf{k}, \varepsilon) = -\frac{1}{\pi} \frac{\Sigma''(\mathbf{k}, \varepsilon)}{[\varepsilon - \varepsilon_{\mathbf{k}}^0 - \Sigma'(\mathbf{k}, \varepsilon)]^2 + [\Sigma''(\mathbf{k}, \varepsilon)]^2}, \quad (13)$$

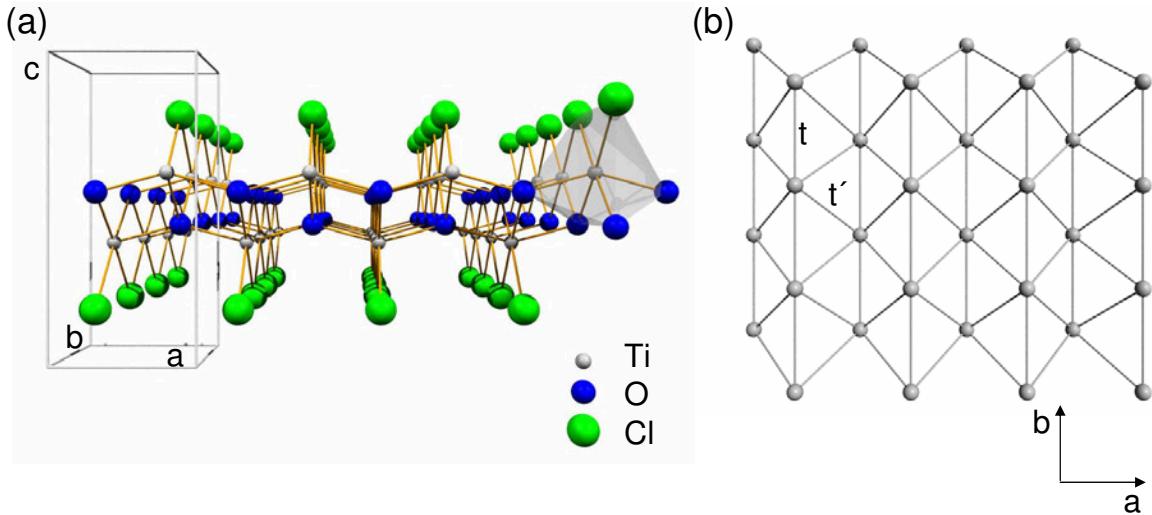


Fig. 4: (a) Crystal structure of $TiOCl$ along the b -axis. Ti - O bilayers, sandwiched by chlorine ions, are stacked along the c -axis with sizable van der Waals gaps in between. (b) Triangular-type lattice of the Ti ions viewed along the c -axis (from [10]).

which again has a simple interpretation. For fixed ε and with $\Sigma''(\mathbf{k}, \varepsilon)$ only weakly varying with \mathbf{k} , Eq. (13) represents a Lorentzian whose finite width is determined by $\Sigma''(\mathbf{k}, \varepsilon)$ and whose renormalized energy position is given by $\Sigma'(\mathbf{k}, \varepsilon)$. Hence, the full information about the finite lifetime of the excitations and the renormalization of the electronic dispersions due to interactions as contained in the self-energy can be determined experimentally from an analysis of ARPES spectra.

While the above expression for the photocurrent in Eq. (10) is general, one should be aware of the fact that it is usually evaluated using *bulk* Bloch states, thereby neglecting the existence of a surface and the concomitant special nature of the states in a real PES experiment. To heal this shortcoming, the resulting spectral function is considered only the first step in a so-called three-step-model in which the photoemission process is described by the initial optical excitation in the solid, the propagation of the photoelectron to the surface including scattering events, and the transmission into the vacuum across the surface potential barrier. A proper treatment of photoemission as a one-step process clearly is much more challenging to theory (cf. J. Minar's contribution in this book).

3 Case studies

3.1 Low-energy photoemission: Doping a one-dimensional Mott insulator

In the nineties, $TiOCl$ was discussed as a material in which a resonating valence bond state, if driven metallic, might result in exotic superconductivity [15]. Indeed, more recent results reported on an unusual spin-Peierls scenario due to the geometric frustration of magnetic interactions and a spin dimerized ground state [16–18], thus reviving the idea of superconductivity induced by some kind of bond-dimer fluctuations if it is possible to introduce additional charge carriers into the system.

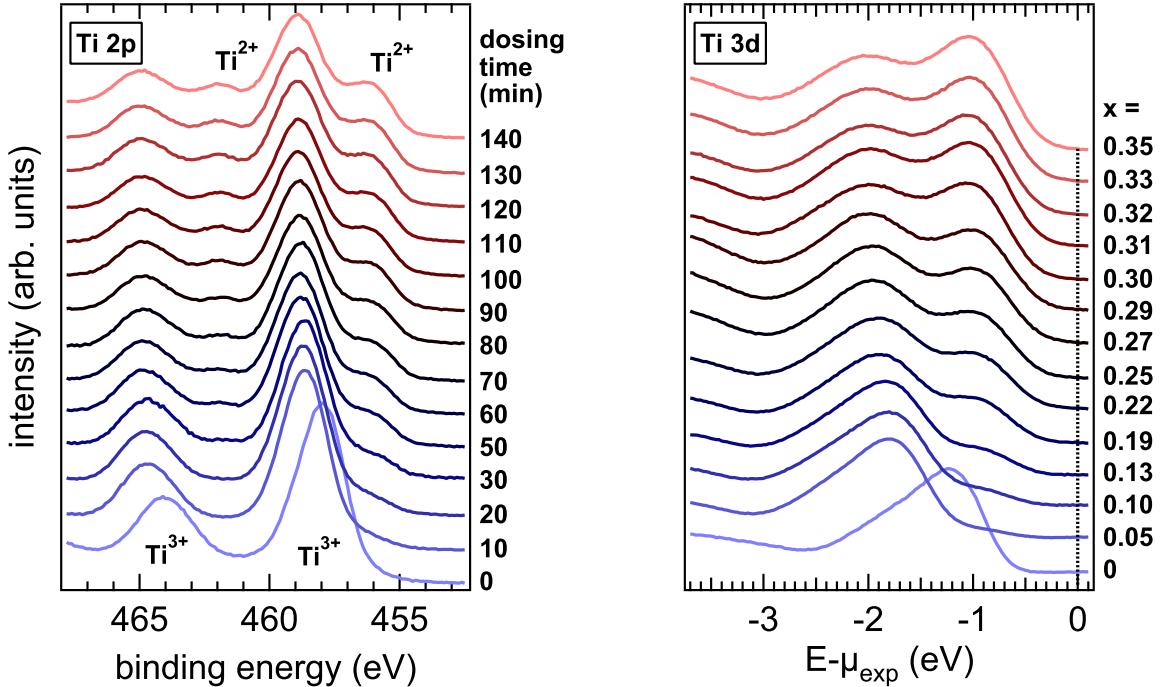


Fig. 5: Left: $Ti\ 2p$ spectra of $TiOCl$ after various durations of Na dosing. Right: Corresponding $Ti\ 3d$ spectra. The electron doping concentration x can be inferred from the $Ti\ 2p$ spectra (from [20]).

The structure of $TiOCl$ is depicted in Fig. 4. It is a layered material [19] with one-dimensional Ti chains running along the crystallographic b -axis (see Fig. 4). Also note that the Ti sites, if projected onto the (a, b) -plane, form a triangular lattice giving rise to the geometric frustration of magnetic interactions which, however, is weak since the hopping amplitudes along and across the chains differ by one order of magnitude.

Beside the interest in this compound as a spin- $\frac{1}{2}$ quantum magnet with unusual properties and the perspective of exotic superconductivity, $TiOCl$ deserves also attention since at room temperature it is a clean realization of a quasi-one-dimensional Mott insulator with the Ti ions in a $3d^1$ configuration and the orbital degrees of freedom quenched (the occupied d_{xy} orbital is slightly split off from the next higher $d_{yz,zx}$ orbitals).

One way of doping the system is by intercalation of alkali metal atoms into the layered structure. This can be achieved *in situ* by the evaporation of alkali metal atoms onto the surface of a freshly cleaved crystal. That the occupancy of the d shell can indeed be enhanced in this way is shown in the left panel of Fig. 5, which shows XPS spectra of the $Ti\ 2p$ core level [20]. With increasing duration of Na dosing, additional spectral weight appears at lower binding energies with respect to the two main lines. This can be attributed to the emission from the $2p$ shell of Ti ions that are in a $2+$ instead of a $3+$ oxidation state. The extra electron in the valence shell leads to a more effective screening of the core potential and hence to a shift of the core levels to lower binding energies (so-called chemical shift). From an analysis of the areas of the respective $Ti\ 2p$ lines the amount of electron doping into the $Ti\ 3d$ states can be determined quite accurately according to $x = A(2+)/A(2+) + A(3+))$.

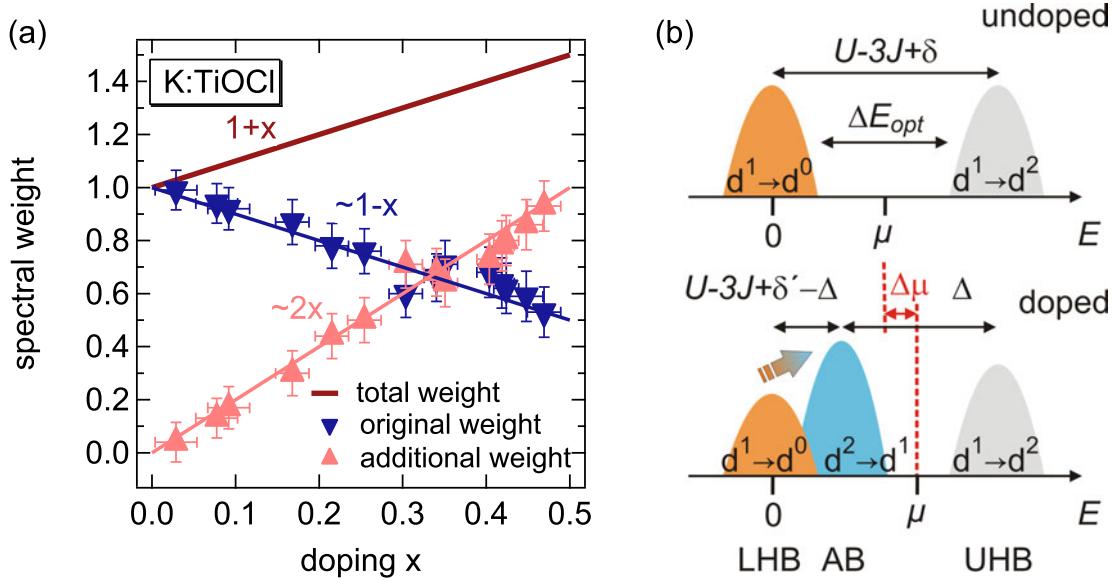


Fig. 6: (a) Quantitative analysis of the transfer of Ti 3d spectral weight upon doping. (b) Sketch of the spectral function of TiOCl for the undoped and doped case. The arrow indicates the transfer of spectral weight (from [20]).

The change in the Ti 3d weight at the chemical potential upon electron doping can be seen in the angle-integrated PES spectra of the right panel of Fig. 5 [20]. For $x = 0$ only the lower Hubbard band (LHB) and a clear gap is seen. Already for the lowest doping level the whole spectrum is shifted by about 0.6 eV to higher binding energies. With increasing x additional spectral weight develops near the chemical potential which overlaps with the lower Hubbard band. Surprisingly, no metallic phase, which would be signalled by a sharp quasiparticle peak at the chemical potential, is observed for any doping level as one would expect in a Mott Hubbard scenario. If the initial shift of the spectrum was interpreted as jump of the chemical potential from the middle of the Mott gap to the lower edge of the upper Hubbard band (UHB) it should amount to about half the optical gap of about 2 eV. However, it is significantly smaller. Further, upon closer inspection it is recognized that the spectral weight of the LHB decreases in favor of the additional weight upon doping.

A quantitative analysis of the data as shown in Fig. 6(a) – here for the case of K intercalation – reveals the following relationship: If the total Ti 3d weight, for which a sum rule applies (cf. section 2.2), is normalized to $1 + x$ and the integral intensities of both components are plotted as a function of doping x , it is seen that the LHB weight decreases as $1 - x$, while the additional weight upon doping increases as $2x$ [20]. This so-called transfer of spectral weight is exactly what is expected in the atomic limit ($t = 0$), i.e., when there is no quasiparticle, for the lower and upper Hubbard bands of a Mott insulator. For each extra electron on a Ti site there is one possibility lost to remove by PES an electron at the orbital energy ε while there are now two possibilities to take out an electron from a doubly occupied site at the single particle energy $\varepsilon + U$. Thus it is spectroscopically proven that both bands are correlated, i.e., they behave as LHB and UHB.

To understand the puzzling observations that no metallic phase is induced despite significant electron doping and that the size of the jump of the chemical potential does not match half the optical gap one has to consider unwanted side effects induced by chemical intercalation apart from the addition of electrons to the $3d$ shell. The salient point is that the alkali metal ions are not randomly distributed between the layers but assume only well-defined positions close to certain Ti sites [21, 22]. The consequences are two-fold: Firstly, the outer electron of the alkali metal is donated exactly to the Ti ion situated next to it. Secondly, the now positively charged alkali metal ion induces a sizable Coulomb potential at the next Ti site leading to a shift of the Ti $3d$ orbital energies at this site towards higher binding energies. In this way, a second sort of Ti site with renormalized orbital energies is created electrostatically, a mechanism that one could call electrostatic alloying. A further consequence of this scenario is that an electron double occupancy can only occur at such alloy sites. In PES these give rise to the observed additional spectral weight, viz. the alloy band (AB) within the original Mott gap between LHB and UHB. Hence, the system remains insulating for all doping concentrations. This qualitative picture is sketched in Fig. 6(b) [20]. For each spectral structure, the transitions in terms of the local electronic configuration of the d shell upon removal of an electron as in PES are also noted. This scenario bears some resemblance to ionic or alloy Hubbard models. However, within these models doping at some stage always leads to metallic phases. The difference to the situation encountered in experiment is that in these models the potential varies in an alternating manner or statistically for a fixed number of sites. In contrast, by the very method of doping through chemical intercalation the number of alloy sites is changed dynamically.

This qualitative picture can also be corroborated quantitatively. The energy for adding a second electron to the d shell of a Ti ion amounts to $U - 3J + \delta$ since for parallel spin orientation of the electrons in two different orbitals an exchange energy $3J$ is gained [23]. Here the energy splitting of the two lowest lying orbitals is taken into account by $\delta \approx 0.3$ eV. Upon intercalation this splitting is reduced to $\delta' \approx 0.1$ eV at the alloy sites [22]. The electrostatic potential at the alloy sites due to adjacent alkali metal ions is denoted by Δ . This yields an energy separation $U - 3J + \delta' - \Delta$ of LHB and AB. Note that the energy separation of LHB and UHB remains unchanged and is given by $U - 3J + \delta$. Experimentally, it amounts to about 1 eV. The chemical potential jumps by half this value – exactly as observed in the PES spectra – assuming that it is lying in mid-gap position, now between AB and UHB. Inserting feasible values for $U - 3J$ between 2.5 and 3.5 eV [23, 24], yields $\Delta \approx 2$ eV. This value can be reconciled within a simple, local model of point charges. If for all crystallographically inequivalent Ti sites the Coulomb potential due to a single, intercalated, positively charged alkali metal ion and an extra electron residing on the Ti site next to it is calculated, taking screening into account by a dielectric constant of 3.5 [25], one obtains a value of about 2 eV for a doubly occupied Ti site while the potential for the other Ti ions remains essentially unchanged. TiOCl, if n -doped by intercalation with alkali metal ions, hence turns out to be a special case of an alloy-Mott insulator where the alloy sites are created dynamically by the dopants themselves due to electrostatics.

3.2 Hard x-ray photoemission: Profiling the buried two-dimensional electron system in an oxide heterostructure

Transition metal oxides are well-known for their huge variety of intrinsic functionalities such as ferroelectricity, magnetism, superconductivity, or multiferroic behavior. Recent achievements in the atomic-scale fabrication of oxide heterostructures by means of pulsed laser deposition (PLD) and molecular beam epitaxy (MBE) made it feasible to combine these intrinsic functionalities with those specific for interfaces and thin films as is known from semiconductor technology. This paves the way not only for tuning the interactions in transition metal oxides to stabilize known phases but also for the creation of novel quantum states, which are not present in the bulk constituents [26]. Such control can be realized by design, the materials choice, applying a gate voltage, or strain. The common scheme thereby is to realize a re-balancing of the interactions of the charge, spin, orbital, or lattice degrees of freedom.

The paradigm for this new class of hybrid materials is the two-dimensional electron system (2DES) which forms at the interface of $\text{LaAlO}_3/\text{SrTiO}_3$ (LAO/STO), although both oxides are band insulators [27]. Intriguingly, the 2DES exists only at a LAO thickness of 4 unit cells (uc) or bigger if the LAO film has been grown on a TiO_2 -terminated STO(001) substrate [28]. Meanwhile, a vast array of intriguing properties has been found for this interface system, among them magnetic Kondo scattering [29], two-dimensional (2D) superconductivity [30], and a large electric-field enabling switching of the interface (super)conductivity by application of suitable gate voltages [28, 31]. More recently, phase separation or even coexistence of magnetism and superconductivity have been discovered at low temperatures [32–35].

The physical origin of the 2DES formation, however, is still debated. The observation that both interface-conductivity as well as ferromagnetism only appear for a critical LAO thickness of 4 unit cells (uc) and beyond [28, 36] has been related to electronic reconstruction. In this scenario electrons are transferred from the surface to the interface in order to minimize the electrostatic energy resulting from the polar discontinuity between LAO and STO [37]. While STO consists of charge neutral SrO and TiO_2 layers, LAO exhibits alternating $(\text{LaO})^+$ and $(\text{AlO}_2)^-$ lattice planes, which act like a series of parallel-plate capacitors [38] resulting in a polarization field across the LAO film. If the potential difference across the LAO gets large enough with increasing thickness, electrons are transferred to the interface to (partially) neutralize the electrostatic gradient. Alternative explanations involve doping by oxygen vacancies and/or cation intermixing (cf. also section 3.3).

The challenge for the notoriously surface-sensitive PES consists in accessing the interface buried below an overlayer of about 2 nm. This became possible only several years ago at 3rd generation synchrotron radiation sources that provide sufficiently high photon fluxes for measurements in the hard x-ray regime with high energy resolution and count rates [39]. Obviously, to gain a microscopic understanding of the origin and nature of the 2DES it is key to know about its vertical extension and the charge carrier concentration. As opposed to most other methods, HAXPES can provide this information non-destructively on as-is samples.

The simple idea is to exploit the angle-dependence of the effective inelastic mean free path

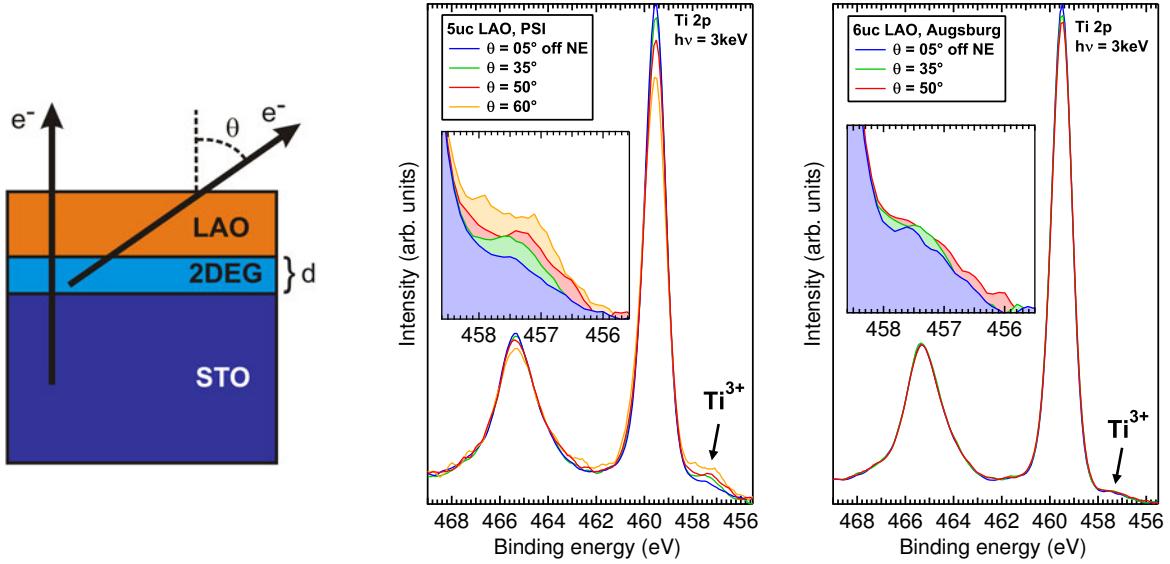


Fig. 7: Left: Sketch of the measuring geometry for angle-dependent HAXPES experiments on a LAO/STO heterostructure with an interface 2DES. Middle and right: Ti 2p spectra of two different LAO/STO heterostructures with varying emission angle θ (from [40]).

$\lambda_{\text{eff}} = \lambda_{\text{INFP}} \cos \theta$ to perform depth profiling, e.g., of the vertical distribution of an element with a certain oxidation state in the geometry as sketched in Fig. 7, left [40]. Figure 7 depicts in the middle and on the right Ti 2p spectra of two different LAO/STO samples with a conducting interface, recorded at a photon energy of 3 keV and with varying detection angle θ relative to the surface normal [40]. The spectra are normalized to integral intensity. The interesting feature is the tiny spectral weight at lower binding energies with respect to the main line, which is shown as close-up in the respective insets.

Due to its chemical shift it can be attributed to the emission from the 2p core level of Ti^{3+} ions, whereas the main line originates from that of Ti^{4+} sites. An additional electron in the d shell results in a more effective screening of the core-potential and hence a shift to lower binding energies. This feature thus provides indirect evidence of the existence of extra electrons in STO in which Ti otherwise is in a 4+ valence state. It is now interesting to realize that in the spectra of Fig. 7 the emission from Ti ions in the oxidation state 3+ increases with larger detection angle, i.e., with smaller information depth and hence higher sensitivity to the interface. This means that the region in STO with extra electrons is indeed confined to the interface and has a vertical extension of the order of the inelastic mean free path of the photoelectrons. Otherwise no angular dependence would be observed.

A more detailed, quantitative analysis of the angle dependence of the ratio of the 2p emission from Ti^{3+} and Ti^{4+} ions, based on a simple model assuming a homogeneous 2DES of thickness d as in Fig. 7, left, allows for an estimate of the sheet carrier density and the thickness of the 2DES. To do so one just has to integrate the contributions of the respective Ti species to the photoemission signal according to their vertical distribution, taking into account the damping factor $e^{-z/\lambda_{\text{eff}}}$ for photoelectrons created at a depth z below the interface [40]. The results are the following: (i) The sheet carrier densities are about one order of magnitude lower than

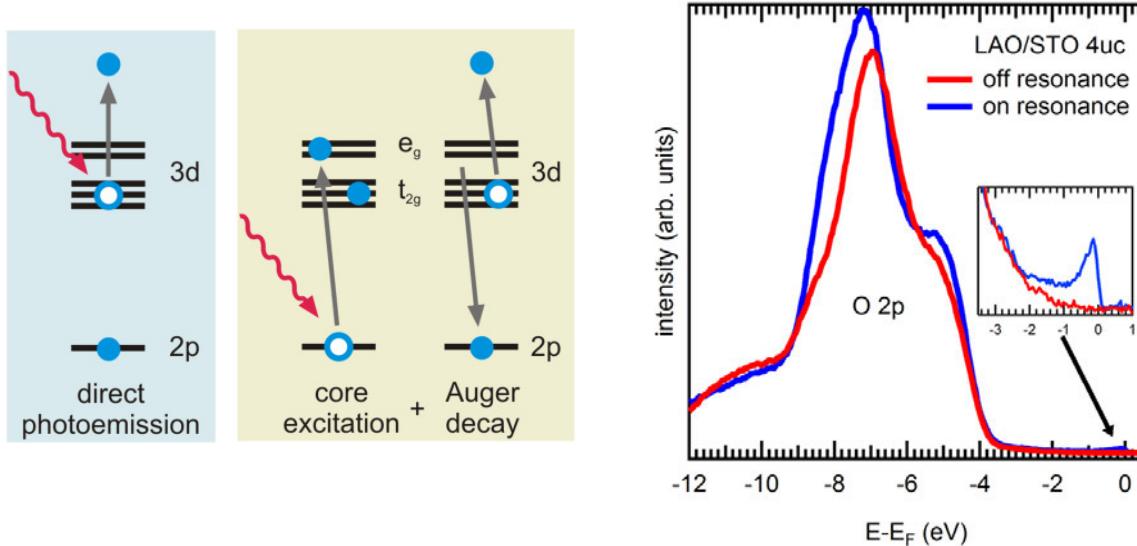


Fig. 8: Left: Sketch of the direct PES and the Auger-like channel which quantum mechanically interfere in ResPES. Right: Angle-integrated on- and off-resonance spectra of the valence band in LAO/STO. The resonance enhancement in the region of O 2p emission (below ≈ 3.5 eV) is due to O 2p-Ti 3d hybridization.

expected in an ideal electronic reconstruction scenario as also is inferred from Hall data [28]. (ii) At variance with the latter, already in samples with a non-conducting interface a finite charge carrier concentration is observed by HAXPES. Moreover, the charge carrier concentrations as inferred from the HAXPES data do not exhibit a jump with increasing LAO thickness but rather a continuous increase. (iii) The vertical extension of the 2DES amounts to only a few unit cells at room temperature.

It is difficult to reconcile these results, in particular (i) and (ii), with the standard electronic reconstruction scenario as mechanism for the formation of the 2DES. In the next section, we will see that in view of the findings by resonant soft x-ray PES the ideal electronic reconstruction actually has to be discarded.

3.3 Resonant angle-resolved soft x-ray photoemission: Direct k -space mapping of the electronic structure in an oxide-oxide interface

In the previous section we saw that HAXPES on the Ti 2p core-levels provides indirect evidence for the interface 2DES in LAO/STO. A direct observation of the 3d states contributing to the 2DES is hindered by the low photoionisation cross section at these high photon energies. Even for energies in the soft x-ray regime around 500 eV which would just grant a sufficiently high probing depth to access the buried interface the Ti 3d photoemission signal is too low for measurements to be feasible. However, there is a special technique called resonant PES (ResPES) which allows for a selective enhancement of the photoemission signal from specific orbitals by tuning the photon energy to an appropriate absorption threshold. According to the dipole selection rules, in our case the apt absorption threshold in the soft x-ray regime is the Ti L edge ($2p \rightarrow 3d$ transitions). There, an additional Auger-like channel opens which leads to the same

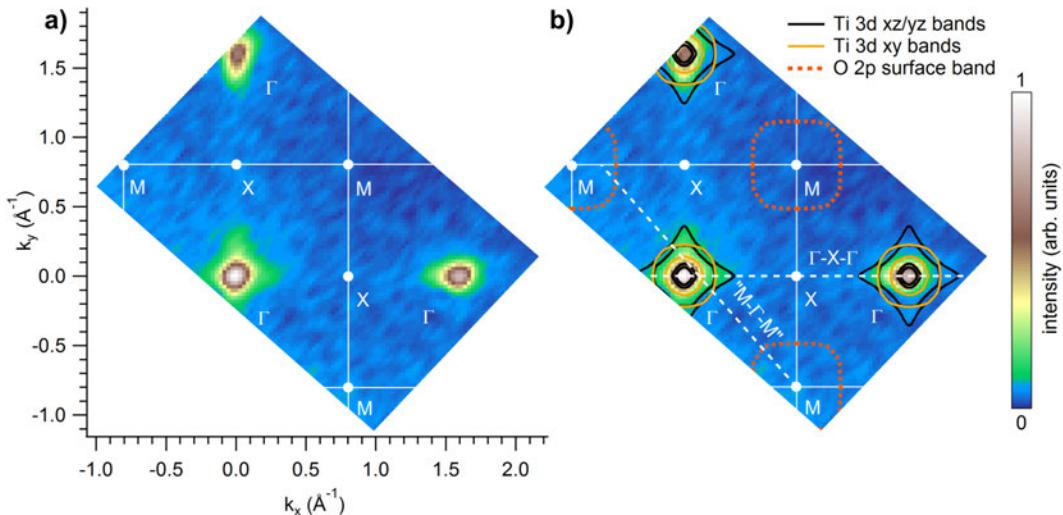


Fig. 9: (a) Fermi surface map recorded at 460.20 eV. (b) Same map as (a) but with the Fermi surface sheets from density-functional calculations overlaid (from [41]).

final state as in the direct PES process (see Fig. 8)



where ϵ denotes the ejected photoelectron. The probability amplitudes of both channels interfere quantum mechanically and thus give rise to an enhanced $3d$ spectral weight. How effectively the resonant enhancement indeed works can be judged from the comparison of angle-integrated on- and off-resonance spectra in Fig. 8. Off-resonance, no hint to the $Ti\ 3d$ states is discernable at the chemical potential, while on-resonance, a sharp quasiparticle peak is observed.

Since momentum information is still preserved using soft x-rays one now can exploit the resonance enhancement to perform even k -space mapping of the 2DES interface states. By integration of the recorded EDCs for each k point over an interval of 0.3 eV centered around the Fermi energy, one obtains the Fermi surface map in Fig. 9 [41]. One finds an isotropic contribution of high intensity around the Γ points of the Brillouin zone of STO. As is best seen for the lower left Γ point in Fig. 9, the spectral weight also extends towards the X points, forming a flower-like intensity distribution. These observations are consistent with the results of density-functional calculations of the Fermi surface sheets, which are overlaid on top of the PES data in Fig. 9(b) [41].

However, there is also a striking discrepancy: In the calculations, a hole-like Fermi surface is predicted around the M points as marked by the red dashed lines. It originates from O 2p states at the valence band maximum of LAO. This simply reflects the standard electronic reconstruction scenario in which the valence band maximum follows the built-in potential across the LAO overlayer towards the Fermi level. At the critical thickness, the valence band crosses the Fermi level at the very surface of the LAO film and gives rise to the hole pockets around M while the released electrons are transferred to the STO conduction band minimum.

The absence of such hole pockets thus leads to the conclusion that the built-in potential is basically screened out in the LAO film. This possibly could be caused by the separation of photogenerated electron-hole pairs in the initial polarization field. However, there are several indications against it. Samples with an insulating interface do charge up in HAXPES which should not happen if there was a significant amount of photogenerated charge carriers. Varying the photon flux by several orders of magnitude does not result in any noticeable changes of the spectra [42]. In a related heterostructure with a polar/non-polar interface a built-in potential has been observed by x-ray photoelectron spectroscopy [43].

Alternatively, the experimental observations can be reconciled by proposals, based mainly on density-functional calculations, that O vacancies at the very surface of the LAO film can serve as a charge reservoir for the electronic reconstruction [44–47]. Such O vacancies induce unoccupied, localized in-gap states. The two released electrons per vacancy are transferred to the interface and thereby the polarization field is efficiently reduced and finally screened out almost completely. Note that we recover the critical thickness in this picture since, with increasing thickness of the LAO film considered as a parallel-plate capacitor, at some stage the electrostatic energy gain for discharging due to the transfer of two electrons per oxygen vacancy outweighs its formation energy.

4 Outlook

In this brief introduction into photoelectron spectroscopy of complex quantum materials only a very small portion of the entire field could be covered. And even there only a few essential aspects could be touched upon. Nevertheless, it should have become clear that photoelectron spectroscopy in all its variants is a very versatile tool to study basically all kind of complex materials exhibiting interesting many-body phenomena. Although photoelectron spectroscopy is a mature, well-established technique the frontiers of its capabilities have been pushed back by technological progress regarding light sources, electron analyzers, and detectors during the last twenty years and this evolution still continues. Current developments are directed towards a further improvement of fast detectors with parallel data acquisition, high brilliance tunable light sources (also with well-defined time structure for time-resolved experiments), electron analyzers with respect to energy and momentum resolution, and spin detection.

Acknowledgments

Part of the work described in this contribution is the result of the author's fruitful collaborations with many researchers. Their names may be found in the respective original papers listed below. Financial support by BMBF (05 KS7WW3, 05 K10WW1) and DFG (CL124/6-1, CL124/6-2, FOR 1162, FOR 1346) is gratefully acknowledged.

References

- [1] M. Cardona and L. Ley (eds.): *Photoemission in solids I* (Springer, Berlin, 1978)
- [2] B. Feuerbacher, B. Fitton, and R.F. Willis (eds.):
Photoemission and the electronic properties of surfaces (Wiley, New York, 1978)
- [3] S.D. Kevan (Ed.): *Angle resolved photoemission – theory and current applications*,
Studies in surface science and catalysis, Vol. 74 (Elsevier, Amsterdam, 1992)
- [4] S. Hüfner: *Photoelectron spectroscopy. Principles and applications*
(Springer, Berlin, 2003)
- [5] W. Schattke and M.A. van Hove (Eds.): *Solid-state photoemission and related methods: theory and experiment* (Wiley-VCH, Weinheim, 2003)
- [6] L. Hedin and S. Lundqvist: *Effects of electron-electron and electron-phonon interactions on the one-electron states of solids*, in: H. Ehrenreich, D. Turnbull (eds.): *Solid State Physics*, Vol. 23 (Academic Press, New York, 1969)
- [7] C.O. Almbladh and L. Hedin in E.E. Koch (Ed.) *Handbook on synchrotron radiation* Vol. 1
(North-Holland, Amsterdam, 1983), p. 607
- [8] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003)
- [9] F. Reinert and S. Hüfner, New J. Phys. **7**, 97 (2005)
- [10] R. Claessen, J. Schäfer, and M. Sing: *Photoemission on quasi-one-dimensional solids: Peierls, Luttinger & Co.* in: *Very High Resolution Photoelectron Spectroscopy*, ed. by S. Hüfner, Lecture Notes in Physics, Vol. 715 (Springer, Berlin, 2007)
- [11] S. Suga and A. Sekiyama (eds.): *Photoelectron Spectroscopy: Bulk and Surface Electronic Structures*, Springer Series in Optical Sciences, Vol. 176 (Springer, Berlin, 2014)
- [12] A. Avella and F. Mancini (eds.): *Strongly Correlated Systems: Experimental Techniques*, Springer Series in Solid-State Sciences, Vol. 180 (Springer, Berlin, 2014)
- [13] G.A. Sawatzky, Nature **342**, 480 (1989)
- [14] L. Hedin and J.D. Lee, J. Electron Spectrosc. Relat. Phenom. **124**, 289 (2002)
- [15] R.J. Beynon and J.A. Wilson, J. Phys.: Condens. Matter **5**, 1983 (1993)
- [16] A. Seidel, C.A. Marianetti, F. C. Chou, G. Ceder, and P.A. Lee,
Phys. Rev. B **67**, 020405(R) (2003)
- [17] M. Shaz, S. van Smaalen, L. Palatinus, M. Hoinkis, M. Klemm, S. Horn, and R. Claessen,
Phys. Rev. B **71**, 100405(R) (2005)

- [18] R. Rückamp, J. Baier, M. Kriener, M.W. Haverkort, T. Lorenz, G.S. Uhrig, L. Jongen, A. Möller, G. Meyer, and M. Grüninger, Phys. Rev. Lett. **95**, 097203 (2005)
- [19] H. Schäfer, F. Wartenpfuhl, and E. Weise, Z. Anorg. Allg. Chem. **295**, 268 (1958)
- [20] M. Sing, S. Glawion, M. Schlachter, M.R. Scholz, K. Goß, J. Heidler, G. Berner, and R. Claessen, Phys. Rev. Lett. **106**, 056403 (2011)
- [21] S. Yamanaka, T. Yasunaga, K. Yamaguchi, and M. Tagawa, J. Mater. Chem. **19**, 2573 (2009)
- [22] Y.-Z. Zhang, K. Foyevtsova, H.O. Jeschke, M.U. Schmidt, and R. Valentí, Phys. Rev. Lett. **104**, 146402 (2010)
- [23] J.S. Lee, M.W. Kim, and T W. Noh, New J. Phys. **7**, 147 (2005)
- [24] M. Hoinkis, M. Sing, J. Schäfer, M. Klemm, S. Horn, H. Benthien, E. Jeckelmann, T. Saha-Dasgupta, L. Pisani, R. Valentí, and R. Claessen, Phys. Rev. B **72**, 125127 (2005)
- [25] C.A. Kuntscher, S. Frank, A. Pashkin, M. Hoinkis, M. Klemm, M. Sing, S. Horn, and R. Claessen, Phys. Rev. B **74**, 184402 (2006)
- [26] H.Y. Hwang, Y. Iwasa, M. Kawasaki, B. Keimer, N. Nagaosa, and Y. Tokura, Nature Mater. **11**, 103 (2012)
- [27] A. Ohtomo and H.Y. Hwang, Nature **427**, 423 (2004)
- [28] S. Thiel, G. Hammerl, A. Schmehl, C.W. Schneider, and J. Mannhart, Science **313**, 1942 (2006)
- [29] A. Brinkman, M. Huijben, M. van Zalk, J. Huijben, U. Zeitler, J.C. Maan, W.G. van der Wiel, G. Rijnders, D.H.A. Blank, and H. Hilgenkamp, Nature Mater. **6**, 493 (2007)
- [30] N. Reyren, S. Thiel, A.D-Caviglia, L.F. Kourkoutis, G. Hammerl, C. Richter, C.W. Schneider, T. Kopp, A.-S. Ruetschi, D. Jaccard, M. Gabay, D.A. Müller, J.-M. Triscone, and J. Mannhart, Science **317**, 1196 (2007)
- [31] A.D. Caviglia, S. Gariglio, N. Reyren, D. Jaccard, T. Schneider, M. Gabay, S. Thiel, G. Hammerl, J. Mannhart, and J.-M. Triscone, Nature **456**, 624 (2008)
- [32] Ariando, X. Wang, G. Baskaran, Z.Q. Liu, J. Huijben, J.B. Yi, A. Annadi, A.R. Barman, A. Rusydi, S. Dhar, Y.P. Feng, J. Ding, H. Hilgenkamp, and T. Venkatesan, Nature Commun. **2**, 188 (2011)
- [33] D.A. Dikin, M. Mehta, C.W. Bark, C.M. Folkman, C.B. Eom, and V. Chandrasekhar, Phys. Rev. Lett. **107**, 056802 (2011)
- [34] L. Li, C. Richter, J. Mannhart, and R. Ashoori, Nat. Phys. **7**, 762 (2011)

- [35] J.A. Bert, B. Kalisky, C. Bell, M. Kim, Y. Hikita, H.Y. Hwang, and K.A. Moler, *Nature Phys.* **7**, 767 (2011)
- [36] B. Kalisky, J.A. Bert, B.B. Klopfer, C. Bell, H.K. Sato, M. Hosoda, Y. Hikita, H.Y. Hwang, and K.A. Moler, *Nature Commun.* **3**, 922 (2012)
- [37] N. Nakagawa, H.Y. Hwang, and D.A. Müller, *Nature Mater.* **5**, 204 (2006)
- [38] C. Noguera, *J. Phys.: Condens. Matter* **12**, R367 (2000)
- [39] K. Kobayashi, *Nucl. Instrum. Methods Phys. Res. A* **601**, 32 (2009)
- [40] M. Sing, G. Berner, K. Goß, A. Müller, A. Ruff, A. Wetscherek, S. Thiel, J. Mannhart, S.A. Pauli, C.W. Schneider, P.R. Willmott, M. Gorgoi, F. Schäfers, and R. Claessen, *Phys. Rev. Lett.* **102**, 176805 (2009)
- [41] G. Berner, M. Sing, H. Fujiwara, A. Yasui, Y. Saitoh, A. Yamasaki, Y. Nishitani, A. Sekiyama, N. Pavlenko, T. Kopp, C. Richter, J. Mannhart, S. Suga, and R. Claessen, *Phys. Rev. Lett.* **110**, 247601 (2013)
- [42] G. Berner, A. Müller, F. Pfaff, J. Walde, C. Richter, J. Mannhart, S. Thiess, A. Gloskovskii, W. Drube, M. Sing, and R. Claessen, *Phys. Rev. B* **88**, 115111 (2013)
- [43] S.A. Chambers, L. Qiao, T.C. Droubay, T.C. Kaspar, B.W. Arey, and P.V. Sushko, *Phys. Rev. Lett.* **107**, 206802 (2011)
- [44] Z. Zhong, P.X. Xu, and P.J. Kelly, *Phys. Rev. B* **82**, 165127 (2010)
- [45] N.C. Bristowe, P.B. Littlewood, and E. Artacho, *Phys. Rev. B* **83**, 205405 (2011)
- [46] Y. Li, S.N. Phattalung, S. Limpijumnong, J. Kim, and J. Yu, *Phys. Rev. B* **84**, 245307 (2011)
- [47] N. Pavlenko, T. Kopp, E.Y. Tsymbal, J. Mannhart, and G.A. Sawatzky, *Phys. Rev. B* **86**, 064431 (2012)

15 Challenges from Experiment: Correlation Effects and Electronic Dimer Formation in Ti_2O_3

T. C. Koethe and L. H. Tjeng

University of Cologne

Max Planck Institute for Chemical Physics of Solids
Dresden

Contents

| | | |
|----------|--|----------|
| 1 | Introduction | 2 |
| 2 | Experimental | 3 |
| 3 | Results | 3 |
| 4 | Discussion: the Hubbard model for the hydrogen molecule | 4 |
| 4.1 | Ground state | 5 |
| 4.2 | Photoemission final states | 5 |
| 4.3 | Photoemission spectral weights | 6 |
| 5 | Conclusions | 8 |

1 Introduction

The spectacular physical properties often observed in materials containing transition-metal and rare-earth elements challenge our comprehension of solid state physics [1, 2]. These properties include superconductivity, unusually large magnetoresistance, metal-insulator transitions, and multiferroicity. We would like to understand how the electrons in such materials interact with each other as to generate those unusual quantum phenomena. From a theoretical viewpoint, it turns out that the equations we have to solve are so complicated that we will not be able to obtain exact solutions. To make things worse and more fascinating at the same time, tiny changes in temperature, pressure, or in the material composition may cause large changes of the properties so that it appears that there are many solutions available that lie very close together in energy. With exact solutions out of reach, the objective is to find smart approximations by which we can capture the essential physics to describe the correlated motion of the electrons in such materials. It may very well be that we need to develop and use different approximations for different materials or properties.

In order to test the validity and accuracy of the theoretical approximations, one can utilize photoelectron spectroscopy [3, 4] as a powerful tool to unravel the electronic structure of the material. One can use the extremely large dynamic range in energy: by studying excitation spectra in the energy range from several eV up to several hundreds of eV, one can obtain direct information about the “bare” electrons, e.g. the charge, spin, and orbital state of the ions that make up the correlated material. By measuring the excitation spectra in the vicinity of the chemical potential with ultra-high resolution, one can directly find the momentum-dependent behavior of the “dressed” electrons, i.e. quasiparticles.

In the following, we will address a long-standing issue concerning the electronic structure of Ti_2O_3 . The basic crystal structure and relevant orbitals are shown in Fig. 1. This material is a non-magnetic insulator at low temperatures and shows a gradual insulator-to-metal transition with metal-like conductivity for temperatures above 500 K [6–10]. The transition is not accompanied by a change in symmetry of the crystal structure [11]. Band structure calculations have great difficulties explaining the properties of this compound: the low temperature phase is calculated to be a metal rather than an insulator [12, 13]. Cluster methods were used to explain the insulating state and the transition by involving, among others, an on-site (intra $Ti\ 3d$) correlation energy [14–16].

Several photoelectron spectroscopic experiments have been carried out on the Ti core levels and the valence band [17–23], thereby identifying the $O\ 2p$ and $Ti\ 3d$ derived character of the valence band and establishing the insulating character of the compound in the low temperature phase. Strangely enough, none of these reported photoemission studies have revealed the presence of satellite structures that otherwise would be consistent with the formation of the c -axis dimers in the low-temperature phase as proposed by the cluster studies [14–16, 24]. The structural presence of the c -axis dimers can be seen in Fig. 1.

Here we will revisit the valence band photoemission experiment with the emphasis on obtaining spectra that are representative for the bulk material. We will analyze the spectra on the basis

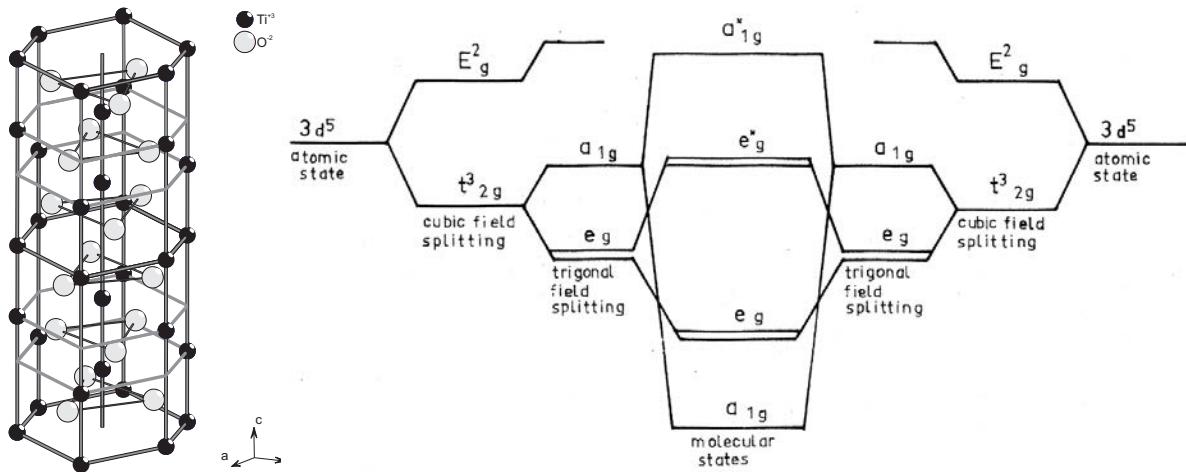


Fig. 1: (Left) Corundum structure of Ti_2O_3 at $T = 300\text{ K}$ in the insulating phase. Titanium (oxygen) sites are indicated by black (white) spheres. (Right) Energy level diagram adapted from Castellani et al. [5] for a c -axis dimer of titanium sites. One-electron orbitals of each site are shown on the left and right sides, with the resulting molecular orbitals in the center.

of the hydrogen molecule model, thereby establishing not only that the c -axis dimer is present electronically, but also quantifying the importance of correlation effects relative to the *intradimer hopping* for the gap formation in the low temperature phase.

2 Experimental

Single crystals of Ti_2O_3 were grown by H. Roth (University of Cologne) using a floating-zone mirror furnace and subsequently characterized by x-ray diffraction and thermogravimetric analysis. No impurities or foreign phases in the samples were detected, which was also confirmed by the photoemission measurements. These measurements were taken at room temperature on *in-situ* cleaved samples. The photoemission spectra (in the following referred to as XPS) were collected using a VG twin-crystal monochromatized $Al\ K_\alpha$ x-ray source with $h\nu = 1486.6\text{ eV}$ and a Scienta SES 100 hemispherical electron energy analyzer. The overall experimental energy resolution was set to 350 meV. The pressure in the spectrometer chamber was below $2 \times 10^{-10}\text{ mbar}$ during the measurement. The possible process of aging of the sample was monitored by repeated $O\ 1s$ spectra alternating with the measurement of the valence band region. No significant changes during the entire experiment (48 h) were observed.

3 Results

The room temperature XPS valence band spectrum is shown in Fig. 2. The large spectral weight at $4\text{--}10\text{ eV}$ binding energies can be attributed to bands with mainly $O\ 2p$ character. The lower energy region from the Fermi level up to 4 eV consists of mainly $Ti\ 3d$ contributions. This part of the spectrum is characterized by two distinct structures. The main line (M) is a sym-

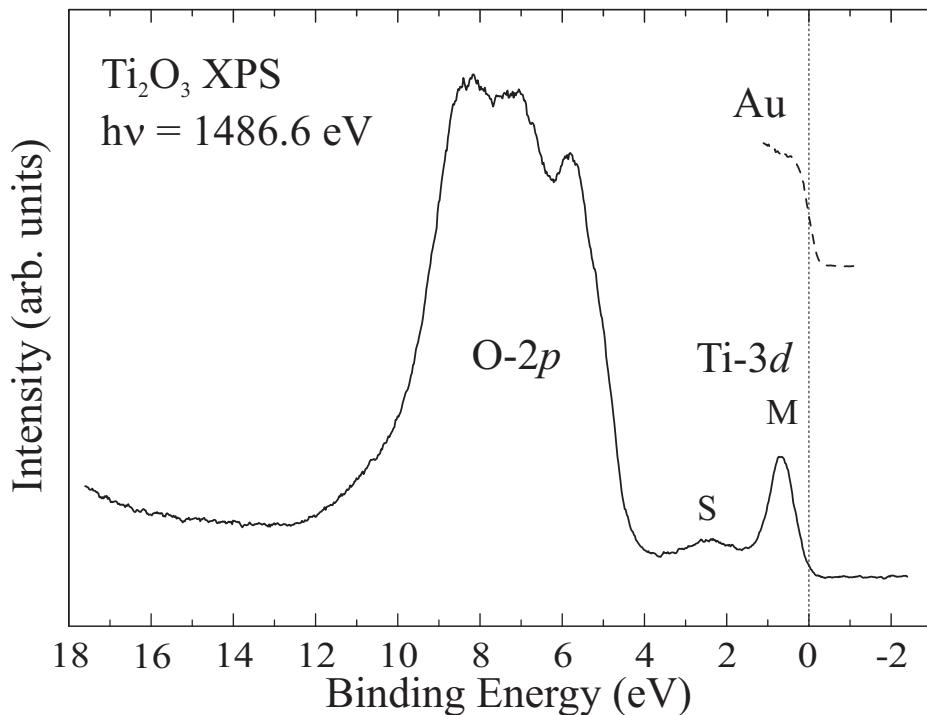


Fig. 2: Room temperature valence band photoemission spectrum of Ti_2O_3 taken at 1486.6 eV photons. The Fermi level was calibrated using a polycrystalline gold sample shown by the dashed line.

metric peak with a width of ≈ 0.8 eV FWHM centered at about 0.7 eV. The semiconducting or insulating nature of the system at room temperature [7–9] is reflected by the fact that the spectral weight vanishes close to the Fermi level, in agreement with earlier photoemission reports [17–22]. The second feature is a weak but clearly noticeable satellite (S) at around 2.4 eV binding energy. The existence of this satellite has not been reported in the literature before [17–22].

One of the concerns with respect to the results reported in the literature is related to the surface sensitivity of photoelectron spectroscopy when using low photon energies. One can estimate that the probing depth is no more than 7–10 Å for ultraviolet photons. At 1486.6 eV, the probing depth is estimated to be around 20 Å, and apparently, this is needed to observe the existence of the satellite (S).

4 Discussion: the Hubbard model for the hydrogen molecule

Establishing the presence of the satellite (S) structure is essential for identifying the appropriate model for Ti_2O_3 . Band theory not only fails to produce a gap at the Fermi level, but also does not reveal any structure in the calculated density of states in the 1–3 eV binding energy region [12, 13]. Only the main line (M) is reproduced. It is also important to note that single-site cluster approaches that include correlation effects do not produce a satellite structure in

the 1–3 eV region, only the main line appears in the calculations [21, 22, 16]. We therefore infer that the satellite (S) structure can only be explained if two aspects are taken into account simultaneously: correlation effects and strong intra-dimer hopping [15, 16]. To demonstrate this point, we now will simulate the Ti 3d derived features using the Hubbard model for the hydrogen molecule [25], where the H 1s orbital is representing the occupied Ti 3d a_{1g} orbital as illustrated in Fig. 1 and indicated experimentally by XAS measurements [24].

4.1 Ground state

The model consists of two electrons that can be distributed over two sites denoted by $i = 1, 2$. The ground state $|GS\rangle$ of the system can be described as a linear combination of a state $|\varphi_0\rangle$ in which the two electrons are on different sites coupled to a singlet, and another singlet state $|\varphi_1\rangle$ in which both electrons are on the same site,

$$\begin{aligned} |GS\rangle &= \alpha|\varphi_0\rangle + \beta|\varphi_1\rangle \\ |\varphi_0\rangle &= \frac{1}{\sqrt{2}} \left(c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger + c_{2\uparrow}^\dagger c_{1\downarrow}^\dagger \right) |0\rangle \\ |\varphi_1\rangle &= \frac{1}{\sqrt{2}} \left(c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger + c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger \right) |0\rangle. \end{aligned}$$

Here, $|0\rangle$ denotes the vacuum state out of which the operators $c_{i\sigma}^\dagger$ create an electron at site i with spin $\sigma = \uparrow, \downarrow$. The triplet states are not considered since they do not hybridize with each other to allow for the formation of a lower energy state [25]. The coefficients α and β are determined by diagonalizing the ground state Hamiltonian,

$$\mathcal{H}_{GS} = \begin{pmatrix} 0 & 2t \\ 2t & U \end{pmatrix},$$

in which U denotes the on-site Coulomb repulsion between two electrons at the same site, and t the hopping integral of the electron between the two sites. The energy of the ground state is then given by

$$E_0 = \frac{1}{2} \left(U - \sqrt{U^2 + 16t^2} \right).$$

4.2 Photoemission final states

The photoemission process, in which an electron is removed, is represented by the annihilation operators $c_{i\sigma}$. In a basis of states in which the remaining electron is localized at one of the two sites, $c_{i\sigma}^\dagger |0\rangle$, $i = 1, 2$, the final state Hamiltonian and the corresponding eigenstates are given by

$$\begin{aligned} \mathcal{H}_{FS} &= \begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix}, \\ |FS^\pm\rangle &= \frac{1}{\sqrt{2}} \left(c_{1\sigma}^\dagger \pm c_{2\sigma}^\dagger \right) |0\rangle. \end{aligned}$$

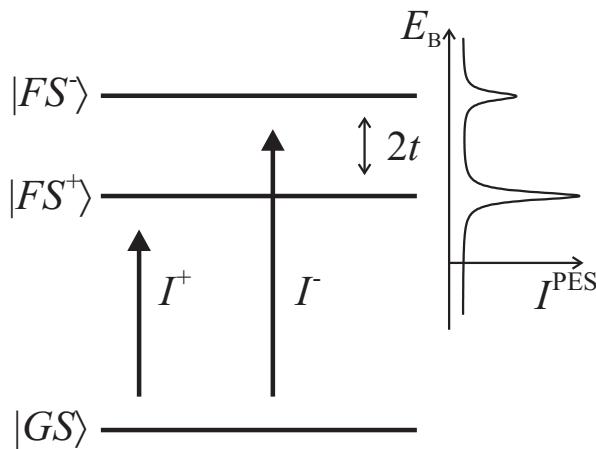


Fig. 3: Total energy level diagram for the photoemission process in a hydrogen molecule model. The two accessible final states $|FS^+\rangle$ and $|FS^-\rangle$ yield two lines in the spectrum of intensities I^+ and I^- , respectively, according to the initial state coefficients α and β . The corresponding spectrum is indicated on the right.

These final states are the well known bonding and anti-bonding states of the H_2^+ ion which are separated in energy by $2t$. See Fig. 3.

4.3 Photoemission spectral weights

Depending on the spin of the removed photoelectron, the final states can take two spin orientations σ of equal energy. One can see immediately that the photoemission spectrum of this system consists of two lines associated with the two final states. The separation of the peaks in the spectrum is given by the final state splitting $2t$. Their intensities depend only on the initial state coefficients α and β . A qualitative spectrum is sketched in Fig. 3 on the right with a lifetime broadening of the photoemission lines taken into account.

We calculate the spectrum as the intensity proportional to the square of the transition matrix elements for the photoemission process,¹

$$\begin{aligned} I^\pm &\propto \|\langle FS^\pm | c_{i\sigma} | GS \rangle\|^2 \\ &= \frac{1}{4} |\alpha \pm \beta|^2 \\ &= \frac{1}{4} \left(1 \pm \sqrt{16t^2 / (U^2 + 16t^2)} \right). \end{aligned}$$

We thus find that the intensities only depend on the ratio between U and t , and we can immediately evaluate the relative intensities for two limiting cases. Firstly, for $U = 0$, I^- vanishes, and the spectrum will be given by a single line only. In other words, although there are two final

¹Note that the sum of I^+ and I^- equals $1/4$ instead of 1 . Since there are four photoemission operators $c_{i\sigma}$ for $i = 1, 2$ and $\sigma = \uparrow, \downarrow$, which naturally all yield the same intensities, the total intensity equals 1 .

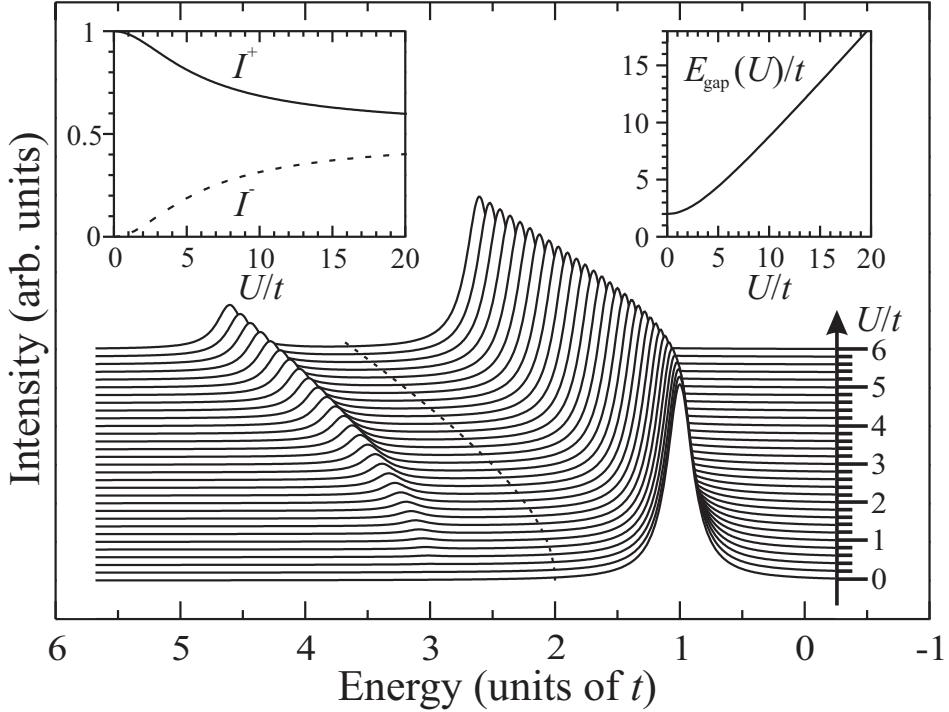


Fig. 4: Calculated photoemission spectra of the hydrogen molecule model in energy units of t for $0 \leq U/t \leq 6$ with a lifetime broadening of $0.25t$. Left inset: the variation of I^+ (solid line) and I^- (dashed) with U/t . Right inset: the energy gap in units of t as function of U/t .

states, we can only reach the lowest of them due to the fully constructive and fully destructive interference ($\alpha = \beta$ when $U = 0$) in the expression for the transition matrix elements. We are thus back to the one-electron approximation and reproduce essentially the results of band structure calculations. Secondly, for $U/t \rightarrow \infty$, the intensities for the two final states become equal: the satellite intensity I^- is as large as that of the main line I^+ . In this limit, there is no double occupation in the ground state and the two electrons reside on one site each ($\alpha = 1$ and $\beta = 0$).

In Fig. 4, we show the intensities for a range of U/t values. The energy scale is in units of t . The energy scale is chosen such that the zero, given by the chemical potential μ , lies in the center of the energy gap shown in the right inset. The gap is given by

$$E_{\text{gap}} = \sqrt{U^2 + 16t^2} - 2|t| .$$

One can clearly see how the satellite intensity I^- grows with increasing U/t . Comparing the H_2 model calculations as shown in Fig. 4 with the experimental spectrum for Ti_2O_3 as displayed in Fig. 2, we can make the estimate from the experimental satellite to main line intensity ratio that $U/t \approx 3 - 4$. The experimental energy separation between the satellite and the main line is $2.4 - 0.7 = 1.7$ eV, which corresponds with $2t$ in the hydrogen model. We thus obtain the estimate that $t \approx 0.85$ eV and $U \approx 2.5 - 3.4$ eV for Ti_2O_3 .

5 Conclusions

We have investigated the low temperature phase of Ti_2O_3 with bulk sensitive photoelectron spectroscopy. We find a distinct two-peak structure in the Ti $3d$ contribution to the valence band spectrum. We attribute this to correlation effects at the Ti $3d$ and to the presence of inter-site hopping within the c -axis dimer. On the basis of the Hubbard model for the hydrogen molecule, we can make realistic estimates for the magnitude of the Hubbard U and the intra-dimer hopping integral t . We can conclude that the c -axis dimers exist not only structurally but also electronically. This finding is relevant for a better understanding of the insulator-to-metal transition, in which an electronic break-up of the c -axis dimers for the high temperature phase can be hypothesized.

Acknowledgements

The research is supported by the Deutsche Forschungsgemeinschaft through SFB608 and FOR 1346. We would like to thank L. Hamdan for her skillful technical and organizational assistance and J. Weinen for his assistance in preparing this manuscript.

References

- [1] for a review see: Electronic Conduction in Oxides by N. Tsuda, K. Nasu, A. Yanase, and K. Siratori, Springer Series in Solid-State Sciences 94, (Springer Verlag, Berlin 1991)
- [2] for a review see: M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
- [3] C.-O. Almbaldh and L. Hedin, *Handbook on Synchrotron Radiation*, Vol. 1b, ed. by E.E. Koch, (North-Holland, Amsterdam, 1983) p. 607
- [4] for a review see: S. Hüfner: *Photoelectron Spectroscopy*, Springer Series in Solid-State Sciences, Vol. 82, (Springer Verlag, Berlin 1996)
- [5] C. Castellani, C.R. Natoli, and J. Ranninger, Phys. Rev. B **18**, 4945 (1978)
- [6] L.K. Keys and L.N. Mulay, Phys. Rev. **154**, 453 (1967)
- [7] L.L. van Zandt, J.M. Honig, and J.B. Goodenough, J. Appl. Phys. **39**, 594 (1968)
- [8] J.M. Honig and T.B. Reed, Phys. Rev. **174**, 1020 (1968)
- [9] J.M. Honig, Rev. Mod. Phys. **40**, 748 (1968)
- [10] R.M. Moon, T. Riste, W.C. Koehler, and S.C. Abrahams, J. Appl. Phys. **40**, 1445 (1969)
- [11] C.E. Rice and W.R. Robinson, Acta Cryst. B **33**, 1342 (1977)
- [12] L.F. Mattheiss, J. Phys.: Condens. Matter **8**, 5987 (1996)
- [13] V. Eyert, U. Schwingenschlögl, and U. Eckern, Europhys. Lett. **70**, 782 (2005)
- [14] H. Nakatsugawa and E. Iguchi, Phys. Rev. B **56**, 12931 (1997)
- [15] A. Tanaka, J. Phys. Soc. Jpn. **73**, 152 (2004)
- [16] A.I. Poteryaev, A.I. Lichtenstein, and G. Kotliar, Phys. Rev. Lett. **93**, 086401 (2004)
- [17] R.L. Kurtz and V.E. Henrich, Phys. Rev. B **25**, 3563 (1982)
- [18] J.M. McKay, M.H. Mohamed, and V.E. Henrich, Phys. Rev. B **35**, 4304 (1987)
- [19] K.E. Smith and V.E. Henrich, Phys. Rev. B **38**, 5965 (1988)
- [20] K.E. Smith and V.E. Henrich, Phys. Rev. B **38**, 9571 (1988)
- [21] T. Uozumi, K. Okada, and A. Kotani, J. Phys. Soc. Jpn. **62**, 2595 (1993)
- [22] T. Uozumi, K. Okada, A. Kotani, Y. Tezuka, and S. Shin, J. Phys. Soc. Jpn. **65**, 1150 (1996)

- [23] R.L. Kurtz and V.E. Henrich, *Surface Science Spectra* **5**, 179 (1998)
- [24] H. Sato, A. Tanaka, M. Sawada, F. Iga, K. Tsuji, M. Tsubota, M. Takemura, K. Yaji, M. Nagira, A. Kimura, T. Takabatake, H. Namatame, and M. Taniguchi, *J. Phys. Soc. Jpn.* **75**, 053702 (2006)
- [25] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, Ch. 32, Problem 5
(Holt-Saunders, Tokyo, 1981)

Index

A

actinides, 2.23
alloy, 13.23
alloy-Mott insulator, 14.11
angle-resolved photoemission spectroscopy (ARPES), 13.2, 14.14
antiferromagnetism, 9.2, 9.26
atomic form factor, 6.12

B

basis set, 2.15, 8.2
Bethe lattice, 12.13
Bethe-Salpeter equation, 6.32, 10.7
bootstrap method, 7.27
boundary conditions
 periodic, 8.3
 twisted, 8.7
Brinkman-Rice transition, 1.7, 1.9, 8.20

C

cavity method, 12.9
cellular dynamical mean-field theory (CDMFT), 9.30, 9.32
central limit theorem, 7.24
charge self-consistency, 5.20
chemical shift, 14.9, 14.13
cluster, 8.8
cluster dynamical mean-field theory (CDMFT), 10.2, 10.18
cluster mean-field theories, 9.30, 9.32
cluster perturbation theory (CPT), 9.2, 9.5
coherent potential approximation (CPA), 1.18, 13.7
configuration interaction, 8.10
constrained random phase approximation, 2.19, 5.12
correlated electron systems, 4.3
correlated materials, 2.2
correlated sampling, 8.16
correlated subspace, 5.4
correlation energy, 1.11, 8.13
correlation function, 6.9
Coulomb integrals, 2.15
Coulomb interaction, 6.4

critical exponents, 10.17
crystal field, 6.4
crystal momentum, 8.5
CT-AUX, 7.3, 7.4
CT-HYB, 6.39, 7.3, 7.13
CT-INT, 7.3, 7.4
CT-QMC, 7.2
cusp condition, 8.11

D

d-wave superconductivity, 11.10, 11.12
density matrix, 12.3
density-functional theory (DFT), 1.21, 2.2, 4.2
detailed balance, 6.16, 7.28
diagram technique, 9.9, 9.11
diagrammatic collapse, 1.14
DMFT
 and functional RG (DMF²RG), 10.2, 11.18
 cellular (CDMFT), 9.30, 9.32
 dynamical mean-field theory, 1.18, 11.20
 KKR+DMFT, 13.3
 LDA+DMFT, 1.21
 non-equilibrium, 1.23
 self-consistent equations, 1.18
double-counting corrections, 4.20, 5.17
downfolding, 5.8
dual fermion (DF) approach, 10.2
dynamical cluster approximation (DCA), 9.32, 10.2, 10.16
dynamical impurity approximation (DIA), 9.30, 9.32
dynamical mean-field theory (DMFT), 2.6, 9.31, 10.2, 10.18, 11.20
dynamical mean-field theory and functional renormalization group approach, 10.2, 11.18
dynamical variational principle, 9.17
dynamical vertex approximation (DVA), 10.1, 10.10, 10.18
Dyson equation, 1.20, 9.12, 10.4
 in nonequilibrium, 12.8

E

effective action, 11.4
 eigenstate thermalization hypothesis, 12.15
 embedding, 2.9
 energy distribution curves, 14.4
 exact diagonalization, 9.5
 exact flow equations, 11.6
 exact renormalization group, 11.2

F

f-sum rule, 6.15
 Falicov-Kimball model, 1.16, 12.2, 12.15
 Fano effect, 13.16
 fast updates, 7.11
 Fermi liquid theory, 1.8
 Fermi's golden rule, 13.10, 14.6
 Feynman diagrams, 9.11, 10.3
 fixed-node Monte Carlo (FNMC), 8.25
 fluctuation-dissipation theorem, 6.16
 fractional coordinates, 8.4
 functional flow equations, 11.6
 functional renormalization group (fRG),
 10.2, 11.2, 11.3

G

generalized-gradient approximation (GGA),
 1.22
 Green function, 9.4, 9.5
 lattice, 1.20
 on time contour, 12.6
 Gutzwiller
 approximation, 1.6, 1.9, 8.18
 projection, 1.6
 wave function, 1.5, 1.9, 8.13
 GW method, 2.4, 10.19

H

HARPES, 13.24
 heat-bath algorithm, 7.32
 Hermite normal form (HNF), 8.8
 Hirsch-Fye QMC, 6.37
 Holstein model, 7.18
 Hubbard model, 1.2, 1.14, 10.16, 10.18,
 11.10
 in nonequilibrium, 12.2, 12.16
 Hund's metal, 2.24
 hybridization
 in nonequilibrium, 12.11

hybridization function, 4.8

I

imaginary time Green function, 6.18
 importance sampling, 8.24
 impurity problem, 4.8
 inelastic mean free path, 14.5
 infinite spatial dimensions, 1.9, 1.13, 1.18,
 11.20
 inner potential, 14.4
 interaction quench, 12.16
 interfaces, 1.22
 iron pnictides, 2.20

J

jackknife method, 7.27
 Jastrow wave function, 8.12

K

Keldysh formalism, 12.3
 KKR Method, 13.3
 KKR+DMFT, 13.3
 Kramers-Kronig relations, 6.14, 6.15
 Kubo formula, 6.9

L

$\text{LaAlO}_3/\text{SrTiO}_3$, 14.12, 14.14
 Langreth rules, 12.7
 lattice, 8.4
 lattice Green function, 1.20
 LDA+DMFT, 1.21, 6.4
 linear response
 function, 6.9, 6.28
 theory, 6.8
 linear susceptibility, 6.9
 local energy, 8.10
 local propagator, 1.18, 1.19
 local-density approximation (LDA), 1.21
 local-vertex approximation, 6.33
 Luttinger-Ward functional, 9.16

M

magnetic susceptibility, 6.27
 atomic limit, 6.30, 6.34
 Mott-insulators, 6.34
 non-interacting, 6.29
 many-body Bloch theorem, 8.6
 many-body perturbation theory, 9.9
 Markov Chains, 7.28

Matsubara Green function, 6.18
Matsubara sums, 6.21
mean-field theory, 1.4
dynamical (DMFT), 1.18, 2.6
slave-boson, 1.9
Weiss, 2.8
Metropolis algorithm, 7.31, 8.15
mixed estimator, 8.24
model
 Falicov-Kimball, 1.16, 12.2
 Holstein, 7.18
 Hubbard, 1.2, 1.14
 single-impurity Anderson, 1.18, 1.19,
 7.3
Mott insulator, 14.9
Mott transition, 9.28
Mott-Hubbard transition, 1.8, 1.22, 10.18

N
nodal set, 8.30
nodal surface, 8.12
non-local correlations, 10.2, 10.16

O
one-band Hubbard model, 6.6
one-dimensional Hubbard model, 9.25
one-particle irreducible, 10.5
optical lattices, 1.22
oxide heterostructure, 14.12

P
parquet equations, 10.8, 10.9
partition function, 4.4
photocurrent, 13.11, 14.6
photoemission spectroscopy, 14.2
 angle-resolved (ARPES), 14.14
 angular-resoled (ARPES), 13.2
 depth profiling, 14.13
 energy conservation, 14.3
 hard x-ray, 14.12
 low-energy, 14.8
 many-body picture, 14.6
 momentum conservation, 14.3
 one-particle picture, 14.3
 perturbation operator, 14.6
 resonant, 14.14
 soft x-ray, 14.14
 three-step model, 13.2, 14.8

photoionisation cross section, 14.5, 14.14
polar discontinuity, 14.12
potential scattering, 9.14
prethermalization, 12.16
primitive cell, 8.4
projection methods, 8.23
projection scheme, 4.17
projectors, 5.5
propagator, 12.3

Q
quantum Monte Carlo (QMC), 4.7, 10.18
 fixed-node (FNMC), 8.25
 variational (VMC), 8.14
quasi particles, 8.30

R
random walk, 8.14
reciprocal lattice, 8.5
reciprocity relations, 11.4
reconstruction
 electronic, 14.12, 14.15
resonating valence bond (RVB) state,
 1.6, 1.11
Ritz variational principle, 9.7

S
second quantization, 8.2
self-consistency
 in nonequilibrium, 12.13
self-energy, 9.12, 10.4, 14.7
 in nonequilibrium, 12.8
self-energy functional theory, 9.2, 9.16
sign problem, 7.8, 8.25
simulation cell, 8.3
single-impurity Anderson model (SIAM),
 1.18, 1.19, 7.3
single-particle Green function, 6.17
slave-boson mean-field theory, 1.9
slave-boson method, 2.6
spectral density functional theory, 2.12
spectral function, 12.7, 13.11, 14.7
spontaneous symmetry breaking, 9.2, 11.17
strong-coupling CT-QMC, 6.39
sudden approximation, 14.6
superlattice, 8.8
surfaces, 1.22, 13.17
susceptibility, 6.28

T

- thermalization, 12.15
- thermodynamic sum rule, 6.15
- TiOCl, 14.8
- transfer of spectral weight, 14.10
- transition
 - Brinkman-Rice, 1.7, 8.20
 - Mott-Hubbard, 1.8, 1.22
- trial wave function, 8.10
- truncation, 2.8
- twisted boundary conditions, 8.7
- two-dimensional electron system, 14.12,
14.14
- two-particle Green function, 6.24

U

- upfolding, 5.7

V

- vacancy
 - oxygen, 14.16
- variance minimization, 8.11
- variational cluster approximation (VCA),
9.2, 9.20, 10.18
- variational Monte Carlo (VMC), 8.14
- vertex, 10.2, 10.6, 10.10
- vertex functions, 11.4

W

- Weiss mean-field theory, 2.8

Z

- zero variance property, 8.11

1. The LDA+DMFT approach to strongly correlated materials

Lecture Notes of the Autumn School 2011 Hands-on LDA+DMFT
edited by E. Pavarini, E. Koch, D. Vollhardt, A. Lichtenstein (2011), 420 pages
ISBN: 978-3-89336-734-4

2. Correlated Electrons: From Models to Materials

Lecture Notes of the Autumn School on Correlated Electrons 2012
edited by E. Pavarini, E. Koch, F. Anders, M. Jarrell (2012), 450 pages
ISBN: 978-3-89336-796-2

3. Emergent Phenomena in Correlated Matter

Lecture Notes of the Autumn School on Correlated Electrons 2013
edited by E. Pavarini, E. Koch, U. Schollwöck (2013), 520 pages
ISBN: 978-3-89336-884-6

4. DMFT at 25: Infinite Dimensions

Lecture Notes of the Autumn School on Correlated Electrons 2014
edited by E. Pavarini, E. Koch, D. Vollhardt, A. Lichtenstein (2014), 450 pages
ISBN: 978-3-89336-953-9