Towards a full \textit{ab initio} theory of strong electronic correlations in nanoscale devices

David Jacob

Max-Planck-Institut für Mikrostrukturphysik, Weinberg 2, 06120 Halle, Germany
E-mail: djacob@mpi-halle.de

Received 16 April 2015
Accepted for publication 17 April 2015
Published 3 June 2015

Abstract
In this paper I give a detailed account of an \textit{ab initio} methodology for describing strong electronic correlations in nanoscale devices hosting transition metal atoms with open $d$- or $f$-shells. The method combines Kohn–Sham density functional theory for treating the weakly interacting electrons on a static mean-field level with non-perturbative many-body methods for the strongly interacting electrons in the open $d$- and $f$-shells. An effective description of the strongly interacting electrons in terms of a multi-orbital Anderson impurity model is obtained by projection onto the strongly correlated subspace properly taking into account the non-orthogonality of the atomic basis set. A special focus lies on the \textit{ab initio} calculation of the effective screened interaction matrix $U$ for the Anderson model. Solution of the effective Anderson model with the one-crossing approximation or other impurity solver techniques yields the dynamic correlations within the strongly correlated subspace giving rise e.g. to the Kondo effect. As an example the method is applied to the case of a Co adatom on the Cu(0 0 1) surface. The calculated low-bias tunnel spectra show Fano–Kondo lineshapes similar to those measured in experiments. The exact shape of the Fano–Kondo feature as well as its width depend quite strongly on the filling of the Co 3$d$-shell. Although this somewhat hampers accurate quantitative predictions regarding lineshapes and Kondo temperatures, the overall physical situation can be predicted quite reliably.

Keywords: molecular electronics, strong correlations, Kondo effect
(Some figures may appear in colour only in the online journal)

1. Introduction

Modern experimental techniques now allow one to reliably create, manipulate and control nanoscale devices with atomic precision in the lab, thus bringing the dream of molecular electronics or nanoelectronics to create ultimately miniaturized electronic devices from single molecules closer to reality [1–4]. Prospective building blocks for molecular electronic circuits such as molecular rectifiers [5–8] and field-effect transistors [9–11] have already been demonstrated in experiments. The use of magnetic atoms or molecules promises to further enhance the functionality of molecular devices by exploiting the spin-degree of freedom of the electron in addition to its charge. Such devices could serve e.g. as basic building blocks for nanoscale spintronics applications [12, 13] or as ultimately miniaturized magnetic information storage devices [14].

Naturally, quantum effects play a crucial role in electronic devices of such tiny dimensions. Consequently, experiments with atomic- and molecular-scale devices have produced a wealth of quantum phenomena such as conductance quantization [15], quantum interference [16–19], or quantum phase transitions [20]. On the other hand, details of the atomic structure also play an important role for determining the electronic properties of nanoscale devices, especially regarding the contact between molecule and metal leads [21–23]. Also the coupling to the leads can significantly alter the electronic and magnetic properties of nanoscale devices by broadening and shifting of energy levels, as well as screening effects. Hence a proper theoretical description of nanoelectronic devices needs to take into account all of the following: quantum effects, the actual atomic structure of the device and the coupling to the leads.
The now standard approach for the description of molecular electronic devices is to combine density functional theory (DFT) calculations with the Landauer transport theory or with the non-equilibrium Green’s function formalism (NEGF) [24–27]. The DFT based transport approach yields an effective mean-field description for the electronic structure and transport properties of molecular devices, taking into account quantum effects, as well as the actual atomic structure of the device, and the coupling of the device to the metallic leads. The approach works quite well for the description of metallic nanon contacts, nanowires and carbon nanotubes [21, 26, 28]. On the other hand, it was realized quite early on that this approach often overestimates conductances of molecules attached to metal leads by orders of magnitude. The origin of this was a matter of debate for about a decade [29–31]. It is now understood that this problem is largely due to DFT based transport calculations failing to produce the correct positions of the molecular energy levels by not properly capturing renormalization of molecular levels at the metal-molecule interface [32–37].

Moreover, nanoscale devices comprising magnetic atoms or molecules often display phenomena induced by so-called strong dynamic correlations that arise when the effective Coulomb interaction between the electrons exceeds their kinetic energies. Dynamic correlations can have a profound impact on the electronic and magnetic structure and the transport properties of the system. One of the most intriguing phenomena induced by dynamic correlations in nanoscale devices is probably the Kondo effect [38, 39]: Below a critical temperature characteristic of the system, the Kondo temperature \( T_K \), the atomic or molecular spin forms a many-body singlet state with the nearby conduction electrons, thereby screening the magnetic moment of the device. The correlations usually originate from the strongly interacting open 3d- or 4f-shells of transition metal atoms. But also molecular orbitals of purely organic molecules only weakly coupled to the leads can give rise to strong correlations. This is corroborated by the fact that the Kondo effect is not only frequently observed in molecular devices comprising transition metal atoms [40–51], but also for devices made from purely organic molecules [42, 52–56]. By construction the DFT based transport method being a static mean-field approach cannot capture the dynamic correlations that lead e.g. to the Kondo effect in nanoscale devices.

Recent efforts to go beyond the DFT based transport approach are to combine time-dependent DFT (TDDFT) with the NEGF [58, 59] or the GW approximation with NEGF [60–62]. A problem of the TDDFT approach is that the standard approximations for TDDFT functionals in connection with an adiabatic exchange correlation kernel does not yield an improvement for the description of correlation effects with respect to the static DFT approach. Some progress has been made recently in that direction by finding a non-adiabatic exchange-correlation kernel for strongly correlated systems but only in the context of simplified models such as the Hubbard or Anderson model [63, 64]. The GW based transport approach on the other hand has been implemented in a fully ab initio way and has been applied to realistic molecular devices. GW yields an energy-dependent self-energy for describing the electronic interactions and thus captures dynamic correlation effects to some extent. Importantly, the afore mentioned problem of overestimating molecular conductances in DFT based transport calculations is solved in GW as it takes into account dynamical screening at the metal-molecule interface which leads to a proper renormalization of molecular energy levels [35]. However, GW, being perturbative in nature, does not properly account for strong electronic correlations such as those leading to the Kondo effect or the Mott–Hubbard metal–insulator transition.

Here I give a detailed account of a different ab initio approach for the description of strongly correlated molecular conductors which has been developed, successively refined and extended in previous work [65–69]. In this approach only the strongly interacting part of the electronic spectrum is described by advanced many-body methods in order to capture dynamic correlations effects. The weakly to moderately interacting part of the electronic system is still treated on a static mean-field level by standard Kohn–Sham DFT (KSDFT). This approach is basically an adaption of the DFT+dynamical mean-field theory (DFT+DMFT) approach [70–73], which has been developed for the realistic description of strongly correlated solids, to the special situation of nanoscale conductors. Similar approaches for treating strong correlations in molecular devices have recently appeared in the literature [74–79].

This paper is organized as follows: in section 2 a detailed account of the so far developed methodology is presented. In section 3 the methodology is applied to the case of a Co adatom at the Cu(0 0 1) surface which has been studied extensively in the recent past [80–83] and thus presents an ideal testbed for the theory. In section 4 I draw conclusions from comparison of the results to the experiments and other theoretical methods. I also discuss some of the caveats of the developed theory and possible solutions to these problems as well as future directions.

2. Methodology

The typical situations encountered in experiments with atomic and molecular devices are depicted in figure 1: (a) a magnetic molecule suspended between the tips of a metal nanocontact and (b) a magnetic atom or molecule deposited on a metallic surface probed by an STM tip. The magnetism and hence the strong correlations of the molecule are here assumed to stem from a single transition metal atom at its center. But the approach can be easily generalized to the case of multiple magnetic atoms by adaption of the dynamical mean-field theory (DMFT) to the case of molecular conductors [67]. Both situations depicted in figures 1(a) and (b) can be described by the model depicted schematically in figure 1(c): the central region, called device region D, contains the molecule or atom

\[ 1 \text{Recently it has been shown by Bergfield et al that the exact exchange correlation functional yields the exact transmission at the Fermi level in the case of the simple Anderson impurity model. However, even the exact Kohn–Sham spectrum does not yield a correct description of the spectral function and transmission outside the Fermi level. Hence the renormalization of the Kondo peak by the interactions cannot be captured by Kohn–Sham DFT based transport calculations [57].} \]
and part of the leads, and will be described on the level of KSDFT. Within the atom or molecule the correlated subspace C yields the strongly interacting levels of the atom/molecule that will be treated by advanced many-body techniques in order to capture the strong dynamic correlations. The polarization region P where the polarizability is calculated in order to compute the screened interaction $U$ of the strongly correlated subspace on the other hand extends over that part of the molecule and/or leads in immediate vicinity of the correlated subspace C. This approach can in principle also be applied directly to the case of purely organic molecules. In that case one has to identify the molecular orbitals responsible for the strong correlations [84].

The approach has been implemented within the ANT.G package [85] which interfaces the Gaussian quantum chemistry code [86] in order to implement the DFT based ab initio transport methodology for molecular conductors. The Gaussian code makes use of Gaussian atomic orbitals as basis sets for performing quantum chemistry and DFT calculations of finite clusters and molecules. The ANT.G package embeds the finite cluster representing the device region into bulk electrodes in order to model the transport situation depicted in figure 1. However, the formalism developed below is not specific to Gaussian basis sets. It can directly be applied to any chemistry code [87]. Even more general, the formalism might in principle also be applied to any basis set as long as the different subspaces taking into account the non-orthogonality of the orbitals $\{\alpha\} \in H$, of $H$ spanned by a subset $M = \{|m\}\}$. Due to the non-orthogonality of the orbitals $\{\alpha\} \in H$, subspace M will in general have a finite overlap with the subspace R spanned by the rest of the orbitals $\{|r\} \in R \equiv H \setminus M$, i.e. $S_{mr} = \langle m|r \rangle \neq 0$ for $\{|m\} \in M$ and $\{|r\} \in R$. Hence the question arises how to define a proper projection $\hat{P}_M$ onto that subspace. We note that there has actually been some controversy about this question in the literature (see e.g. [89] and references therein).

It turns out that the proper choice for $\hat{P}_M$ is actually quite obvious: Let us first consider the simplest case of the subspace $M$ being spanned by a single orbital $|m\rangle$. By definition, the projection operator for a single state is simply $\hat{P}_m = |m\rangle\langle m|$. This definition is independent of how (in which basis) the Hilbert space of the entire system is defined; i.e. it does not matter whether $|m\rangle$ forms part of the basis set within the entire Hilbert space or not; or in case it does whether it has some overlap with the Hilbert space R spanned by the rest of the basis set.

Hence it is clear that the projection $\hat{P}_M$ for the subspace $M$ can be written in an orthonormal basis set $M^+ = \{|m^+\}\}$. Such an orthonormal set can always be found by Löwdin orthogonalization of the original non-orthogonal set spanning M: $|m^+\rangle = \sum_{m \in M} \langle m^+ | m \rangle_{mn} |m\rangle$ where $S_{mn}$ is the overlap matrix between the basis set elements of M only and $S_{mn}^{-1/2}$ is an abbreviation for $(S_{mn})^{-1/2}$, i.e. the matrix power $-1/2$ of the matrix $S_{mn}$. Hence we find for the projection operator:

$$
\hat{P}_M = \sum_{m^+ \in M^+} |m^+\rangle \langle m^+ |
= \sum_{m,n \in M} \sum_{m^+ \in M^+} (S_{mn})_{mn}^{-1/2} (S_{mn}^{-1/2})_{m^+n} |m\rangle \langle n |
= \sum_{m,n \in M} |m\rangle (S_{mn}^{-1/2})_{mn} |n\rangle
$$

(1)

which is nothing but the identity operator for the subspace $M$ written in the non-orthogonal basis set. It has been argued on more formal grounds that this choice for the projection is actually the only physical reasonable one as it is the only one that leads to a tensorial consistent occupancy matrix which generates a Hermitian potential [89]. Note that the subspace projection $\hat{P}_M$ defined here corresponds to the projector with regard to the $\Delta$ metric denoted by $\hat{P}_M^\Delta$ in [88].
Also note that in general we cannot write the identity operator for the entire system as the sum of the projection onto subspace \( M \) and subspace \( R \) spanned by the rest of the basis set if there is some overlap between the two subspaces, i.e. \( \hat I \neq \hat P_M + \hat P_R \). Rather we have to correct for the overlap between the two subspaces:

\[
\hat I = \sum_{a, \beta \in H} |a\rangle (S^{-1})_{\alpha \beta} \langle \beta| = \hat P_M + \hat P_R + \hat O
\]

(2)

where \( S^{-1} \) is the inverse of the overlap matrix for the entire system. \( \hat O \) is an operator correcting the sum of projections by the overlap between the two subspaces \( M \) and \( R \). \( \hat P_M \equiv \hat P_R + \hat O \) defines the projection onto a new subspace \( \bar M \) which is actually orthogonal to subspace \( M \). The projection \( \hat P_M \) thus defines an orthogonalization scheme which orthonormalizes \( R \) with respect to subspace \( M \) preserving the latter.

Now let us have a look at how an operator \( \hat A \) acting on the full Hilbert space is projected onto the subspace \( M \):

\[
\hat A_M = \hat P_M \hat A \hat P_M = \sum_{m, \alpha, n \in M} \hat P_M |m\rangle \langle m'|\hat (\hat A_M)_{n\alpha} |n\rangle \hat P_M
\]

\[
= \sum_{m, \alpha, n \in M} |m\rangle (S_{M}^{-1})_{mn} |n\rangle \langle m'| \hat (\hat A_M)_{n\alpha} = \sum_{m, \alpha, n \in M} |m\rangle (\hat A_M)_{mn} |n\rangle
\]

(3)

where \( \hat A_M = \langle (|m\rangle \hat A |n\rangle) \) is the direct matrix given by the matrix elements of \( \hat A \) with the basis \( \{|m\rangle\} \) of subspace \( M \) and \( \hat A_M = S_M^{-1} \hat A S_M^{-1} \) is the so-called nuclear matrix in that basis. Note that for an orthonormal basis of \( M \) we have \( \hat A_M = \hat A_M \).

Frequently, we will also have to project an operator \( \hat A \) given for some subspace \( M' \subset M \):

\[
\hat A_M = \hat P_M \hat A \hat P_M = \sum_{m, \alpha, n \in M'} \hat P_M |m\rangle \langle m'|\hat (\hat A_M)_{n\alpha} |n\rangle \hat P_M
\]

\[
= \sum_{m', \alpha, n \in M'} |m\rangle (S_{M'}^{-1})_{m'\alpha} |\alpha\rangle \langle \alpha'| \hat (\hat A_M)_{n\alpha} |n\rangle \langle n'| \hat P_M
\]

(4)

where \( S_{M'} \) is the overlap matrix between orbitals \( |m'\rangle \in M' \) and orbitals \( |m\rangle \) \( \in M \). Hence we obtain the following expression for the nuclear matrix of subspace \( M' \) in terms of the nuclear matrix for subspace \( M \):

\[
\hat A_M = S^{-1}_{M} S_{M'M} \hat A_{M'} S_{M'M}^{-1} S^{-1}_{M}
\]

(5)

On the other hand, we may also have the opposite situation where we have some operator \( \hat A_M \) only defined on subspace \( M \), and we want to know the direct matrix for the entire space \( H \), i.e.

\[
\langle \alpha | \hat A_M | \beta \rangle = \sum_{m, \alpha, n \in M} \langle \alpha | m \rangle (\hat A_M)_{mn} |n\rangle \langle n| \beta \rangle
\]

(6)

Hence the direct matrix of the operator \( \hat A_M \) is given by

\[
\hat A_M = S_{HM} \hat A_M S_{MH}
\]

(7)

### 2.2. Projected Green’s functions

The central quantities both in DFT based transport calculations of molecular electronics devices and in quantum many-body theory are Green’s functions (GF). The one-body GF is defined as the resolvent of the one-body Schrödinger equation [90]:

\[
\hat G(z + \mu - \hat H) = \hat I
\]

(8)

where \( z \) is complex, \( \mu \) is the chemical potential, and \( \hat H \) is the Hamiltonian of the system. \( \hat G(z) \) has poles at the eigen values \( \epsilon_k \) of \( \hat H \) for a finite system or a branch cut on the real axis at the energy bands for an infinite system. Its spectral representation in terms of the eigen states \( |\alpha\rangle \) of \( \hat H \) is given by:

\[
\hat G(z) = (z + \mu - \hat H)^{-1} = \sum_{\alpha} \frac{|\alpha\rangle \langle \alpha|}{z + \mu - \epsilon_k}
\]

(9)

The GF operator projected onto subspace \( M \) is given by:

\[
\hat G_M(z) = \hat P_M \hat G(z) \hat P_M = \sum_{a, \beta \in M} |a\rangle \langle \beta| \hat G_M(z)
\]

(10)

Defining the GF of the isolated subspace \( M \) as

\[
\hat g_M(z) = ((z + \mu) \hat P_M - \hat H_M)^{-1}
\]

(11)

and the self-energy operator \( \hat S_M \) associated with the coupling of the subspace \( M \) to the rest of the world as

\[
\hat S_M(z) = [\hat g_M(z)]^{-1} - [\hat G_M(z)]^{-1}
\]

(12)

it is possible to rewrite the projected GF as

\[
\hat G_M(z) = ((z + \mu) \hat P_M - \hat H_M - \hat S_M(z))^{-1}
\]

(13)

The self-energy \( \hat S_M(z) \) is not to be confused with the one describing electron-electron interactions in the many-body GF formalism. Note that in many-body physics in the context of the Anderson impurity model [91] \( \hat S_M(z) \) is often called hybridization function and is denoted by \( \Delta_M(z) \).

One can easily write \( \hat S_M(z) \) in terms of the GF for the isolated (i.e. not coupled to \( M \)) complementary space \( \bar M \) defined by \( \hat P_M \).

\[
\hat g_M(z) = ((z + \mu) \hat P_M - \hat H_M)^{-1}
\]

where \( \hat S_M(z) = \hat H_M \hat g_M(z) \hat H_M \) where \( \hat H_{M,M} = \hat P_M \hat H \hat P_M = (\hat H_{MM})^{-1} \). In order to find the matrix representations \( \hat G_M \) and \( \hat G_M \) of the projected GF \( \hat G_M(z) \), equation (13), is multiplied with the denominator of the rhs, the matrix elements are taken and the subspace identity \( \hat P_M \) is inserted between the two factors of the lhs:

\[
\langle \alpha | \hat G_M(z) \sum_{a', \beta \in M} |a'\rangle \langle S_{M}^{-1}_{\alpha \beta} \rangle \beta (|z + \mu\rangle \hat P_M - \hat S_M(z) |\beta\rangle = \langle \alpha | \hat P_M | \beta \rangle
\]

(14)

Hence we find for the direct GF matrix

\[
\hat G_M(z) = \hat S_{MM} ((z + \mu) \hat S_{MM} - \hat H_M - \hat S_M(z))^{-1} \hat S_M
\]

(15)

and for the corresponding nuclear matrix

\[
\hat G_M(z) = ((z + \mu) \hat S_{MM} - \hat H_M - \hat S_M(z))^{-1}
\]

(16)
2.3. Many-body Green’s functions and Feynman diagrams in an atomic basis set

The generalization of the one-body Green’s function for an (effectively) non-interacting system to the case of interacting electrons are the single-particle Green’s function or single-particle propagators. The single-particle Matsubara GF [92] for atomic states \( \alpha \) and \( \alpha' \) is defined as

\[
G_{\alpha\alpha}(\tau, \tau') = \langle c_{\alpha}(\tau) c_{\alpha}^+(\tau') \rangle
\]

where \( \tau \) is imaginary time and the creation and annihilation operators obey the generalized anti-commutation rules for non-orthogonal basis sets [93]:

\[
\{c_{\alpha}, c_{\beta}^+\} = S_{\alpha\beta}
\]

The Fourier transform with respect to imaginary time \( \tau \) yields the Matsubara GF for imaginary frequencies (called Matsubara frequencies):

\[
G_{\alpha\alpha}(i\omega) = \int_0^\beta d\tau e^{i\omega\tau} G_{\alpha\alpha}(\tau, 0)
\]

By analytic continuation to the real frequency axis one obtains the retarded single-particle GF:

\[
\tilde{G}_{\alpha\alpha}(z) = \int_0^\beta d\tau e^{iz\tau} G_{\alpha\alpha}(\tau, 0)
\]

For the development of a diagrammatic expansion for the interacting GF in a non-orthogonal basis in terms of the Coulomb interaction [93], one may either use the nuclear GF matrix or the direct GF matrix in combination with the nuclear matrix of the Coulomb interaction [93]. Here we will work with the nuclear GF matrix for the interacting single-particle GF as \( \tilde{G} = S^{-1} G S^{-1} \).

For the development of a diagrammatic expansion for the interacting GF in a non-orthogonal basis in terms of the Coulomb interaction [93], one may either use the nuclear GF matrix or the direct GF matrix in combination with the nuclear matrix of the Coulomb interaction [93]. Here we will work with the nuclear GF matrix for the Green’s functions and the direct matrix for the interactions.

The bare Coulomb interaction in an atomic basis set is given by:

\[
V_{\alpha\beta, \alpha'\beta'} = \sum_{\sigma, \sigma'} \tilde{V}_{\alpha\beta, \alpha'\beta'} c_{\alpha\sigma}^+ c_{\beta\sigma} c_{\alpha'\sigma'} c_{\beta'\sigma'}
\]

where \( \tilde{V}_{\alpha\beta, \alpha'\beta'} \) is the nuclear matrix of the Coulomb interaction [93], i.e. \( \tilde{V}(1, 2) = S(1)^{-1} S(2)^{-1} \tilde{V}(1, 2) S(2)^{-1} S(1)^{-1} \) and the direct matrix elements are given by

\[
V_{\alpha\beta, \alpha'\beta'} = e^2 \int dr_1 dr_2 \phi_\alpha^*(r_1) \phi_\beta(r_1) \phi_\alpha^*(r_2) \phi_\beta(r_2) / \|r_1 - r_2\|
\]

The Feynman diagrams for the GF and the Coulomb interaction in an atomic basis set are shown in figure 2.

2.4. DFT based transport calculations

We consider the situation schematically depicted in figure 1(c).

The central device region \( D \) containing a molecule is coupled to two electrodes \( L \) and \( R \). This situation can be realized in a number of ways as shown in figure 1(a) and (b): (a) A molecule bridging the tips of a nanocontact or (b) a molecule deposited on a metal substrate and coupled to an STM tip. In addition to the molecule the device region \( D \) contains those parts of the two electrodes which are in close proximity to the molecule and whose electronic structure is modified by the presence of the molecule and vice versa. In the case of the molecular bridge (a) the tips of the nanocontact are included in the device region while in the case of the molecule on the substrate (b), part of the surface and of the STM tip are included in the device region.

The electronic structure of the central device region is calculated ab initio on the level of DFT in the Kohn–Sham (KS) framework, taking into account the coupling to the electrodes \( L \) and \( R \). The Kohn–Sham Green’s function of the device region \( D \) is given by:

\[
\tilde{G}_D^0(z) = ((z + \mu) S_D - \Sigma_0^D(z) - \Sigma_0^L(z) - \Sigma_0^R(z))^{-1}
\]

where \( \Sigma_0^D(z) \) is the KS Hamiltonian of the device region which yields an effective mean-field description of the electronic structure of the device region. \( \Sigma_0^L(z) \) and \( \Sigma_0^R(z) \) are the lead self-energies associated with the coupling of the device region to the bulk electrodes.

From the device GF the electronic density can easily be calculated by integration up to \( \omega = 0 \) (corresponding to the chemical potential \( \mu \))

\[
\tilde{D}_D^0 = -\text{Im} \frac{1}{\pi} \int_{-\infty}^{0} d\omega \tilde{G}_D^0(\omega + i\eta)
\]

The density matrix yields a new KS Hamiltonian for the device region thus closing the self-consistency cycle of the KS calculation. Hence we can self-consistently calculate the electronic structure of the device region taking into account the coupling to the electrodes (open system).

In contrast to \( D \), the electronic structure (Hamiltonian) of the electrodes \( L \) and \( R \), and hence the self-energies are kept fixed during the self-consistent calculation of the electronic structure of \( D \). Depending on the situation, different models for the bulk electrodes can be employed. One can for example choose nanowires [94], embed the cluster into a perfect crystalline surface calculated ab initio [26], or use so-called absorbing boundary conditions (ABC) [95]. Here we choose...
a tight-binding Bethe lattice model [96] with realistic tight-binding parameters obtained from DFT calculations [97]. The actual choice of the electrode model is not crucial for calculations as long as the bulk electrodes are far enough away from the central scattering region, i.e. the device region is chosen big enough and contains a sufficiently big part of the electrodes [98].

Once the KS calculation is converged the transport properties can be calculated within the Landauer approach from the transmission function which is given by:

\[ T^0(\omega) = \text{Tr}[\Gamma_L(\omega)\tilde{G}^0(\omega)\Gamma_R(\omega)\tilde{G}^0(\omega)] \quad (24) \]

where \( \Gamma_L \equiv i(\Sigma_L - \Sigma_L^\dagger) \) and \( \Gamma_R \equiv i(\Sigma_R - \Sigma_R^\dagger) \) are the so-called coupling matrices which yield the broadening of the device region due to the coupling to the leads. From the transmission function the current and conductance can be calculated using the Landauer formula

\[ I(V) = \frac{2e}{\hbar} \int d\omega T^0(\omega) \left( f(\omega - \mu_L) - f(\omega - \mu_R) \right) \quad (25) \]

where \( \mu_L \) and \( \mu_R \) are the electrochemical potentials of the left and right, respectively, defined by the applied bias voltage \( eV = \mu_L - \mu_R \). Note that in general the transmission function \( T^0(\omega) \) also depends on the applied voltage \( V \), i.e. \( T^0 = T^0(\omega, V) \), and actually has to be calculated out of equilibrium by combining the KSDFT with the NEGF [24–26]. However, within the mean-field like KSDFT based NEGF approach the transmission is often not so strongly voltage dependent, and within the mean-field like KSDFT based NEGF approach the fixed to the equilibrium one tip, i.e. the electrochemical potential of the substrate remains from the transmission function the current and conductance can be calculated using the Landauer formula

\[ G(V) = \frac{\partial I}{\partial V} = \frac{2e^2}{\hbar} \int_0^{eV} d\omega T^0(\omega) = \frac{2e^2}{\hbar} x T^0(eV) \quad (26) \]

In contrast, for the situation of a molecule undergoing a symmetrically to two leads (figure 1(a)), the voltage will drop more or less symmetrically across the junction, i.e. \( \mu_L = \mu - eV/2 \) and \( \mu_R = \mu + eV/2 \). Hence for the conductance we obtain now

\[ G(V) = \frac{2e^2}{\hbar} \int_{-eV/2}^{eV/2} d\omega T^0(\omega) = \frac{e^2}{\hbar} \left[ T^0\left(\frac{eV}{2}\right) + T^0\left(-\frac{eV}{2}\right) \right] \quad (27) \]

In a more general situation where the coupling is neither completely symmetric nor completely asymmetric, more sophisticated modelling of the electrostatics or even a KS-NEGF calculation would be necessary in order to find the actual voltage drop.

### 2.5. Projection onto the correlated subspace: Anderson impurity model

Next we have to identify the strongly correlated subspace C. Usually C will be formed by the open d- or f-shells of transition metal atoms. However, also molecular orbitals of purely organic molecules such as C₆₀ or carbon nanotubes weakly coupled to electrodes can show strong correlations if the effective interaction in these levels is big in comparison with the broadening due to the coupling to the leads. Our approach is completely general in this respect.

From now on we assume that the orbitals \( \phi \) forming the subspace C are mutually orthogonal (but not necessarily orthogonal to the other orbitals in the device region). This can always be achieved by simple Löwdin orthogonalization of subspace C. However, note that often the orbitals spanning C are already mutually orthogonal. For example in the case of the atomic orbitals forming the open d- or f-shell of a transition metal atom, or in the case of molecular orbitals which are the eigenstates of the KS Hamiltonian of the molecule and hence by construction are orthogonal. In order to account for the strong correlations in subspace C an effective Coulomb interaction term

\[ \hat{V}^{\text{ee}}_C = \frac{1}{2} \sum_{i,j}^{C} U_{ik,jl} c_{i\sigma}^\dagger c_{j\sigma} c_{k\sigma} c_{l\sigma} \quad (28) \]

is added acting on the orbitals in C. Note that \( U_{ik,jl} \) is not the bare Coulomb interaction but an effective interaction which is usually much lower than the bare one due to screening processes by the conduction electrons. In the next section it is shown how to calculate \( U_{ik,jl} \) \textit{ab initio} from the DFT electronic structure. The full many-body Hamiltonian of the strongly interacting subspace C now reads:

\[ \hat{\mathcal{H}}_C = \hat{\mathcal{H}}_C^0 + \hat{V}^{\text{ee}}_C \quad (29) \]

where the one-body part \( \hat{\mathcal{H}}_C^0 = \sum_{i,j}^{C} \left( \phi_i | \hat{\mathcal{H}}^0_{\text{KS}} | \phi_j \right) c_{i\sigma}^\dagger c_{j\sigma} \) is given by projection of the KS Hamiltonian \( \hat{H}^0_{\text{KS}} \) onto C. However, since the Coulomb interaction has been taken into account already on a mean-field level in the Kohn–Sham Hamiltonian, we also need to subtract a double-counting correction (DCC) term:

\[ \hat{\mathcal{H}}_C = \hat{\mathcal{P}}_C \hat{H}^0_{\text{KS}} \hat{\mathcal{P}}_C - \hat{V}^{\text{dc}}_C \quad (30) \]

Unfortunately, the DCC term \( \hat{V}^{\text{dc}}_C \) is not exactly known for DFT, and several approximation schemes are used in practice [99]. Here the so-called atomic limit or fully localized limit (PLL) is employed [100], but generalized to the case of an anisotropic Coulomb repulsion \( U_{ii,jj} \) [69]:

\[ (V^{\text{dc}}_C)_{ii} = \sum_j U_{ii,jj} \left( n_j - \frac{1}{2M_C} \right) - J_{ii} N_C - \frac{1}{2} \quad (31) \]

where \( n_j = (c_{j\sigma}^\dagger c_{j\sigma}) \) is the electronic occupation of orbital \( \phi_j \), \( M_C \) is the dimension of subspace C, \( J_{ii} \) is the Hund’s rule coupling given by the orbital-averaged exchange matrix
elements $U_{ij,j'}$ and $N_C = \sum_j n_j$ is the total electronic occupation of subspace C.

According to (16) the self-energy (a.k.a. the hybridization function) associated with the coupling of the subspace C to the rest of the system is given by

$$\Delta_C(\omega) = (\omega + \mu) I_C - \mathbf{H}_C^0 - [\hat{G}_C(\omega)]^{-1}$$

where the projected GF of the correlated subspace $C \subset D$ can be calculated from the device GF according to (5) as

$$\hat{G}_C(\omega) = S_{CD} \hat{G}_D(\omega) S_{DC}$$

As customary in many-body physics we will call $\Delta_C(\omega)$ the hybridization function from now on. The many-body Hamiltonian $\hat{H}_C$ of subspace C together with the hybridization function $\Delta_C(\omega)$ define a multi-orbital Anderson impurity model (AIM). Solution of the AIM yields the self-energy $\Sigma_C(\omega)$ describing the strong electronic correlations within the C subspace which is fed-back to the DFT calculation in order to obtain electronic spectra and transport properties of the molecular device (see section 2.8).

### 2.6. Computation of the effective interaction in the correlated subspace

The effective interaction $\hat{V}_C^{\alpha\beta}$ between the electrons in the correlated subspace C is not the bare Coulomb interaction because of screening processes by formation of electron–hole (e–h) pairs in the rest of the system. Therefore the screened Coulomb matrix elements $U_{ik,j}$ are considerably lower than the bare Coulomb interaction $V_{ik,j}$. The screening of the bare interaction by formation of e–h pairs can be calculated within the so-called Random Phase Approximation (RPA) (see e.g. the book by Mahan [92] or any other textbook on quantum many-body theory). However, screening of the electrons within the C subspace will already be taken into account by the impurity solver. Hence the contribution of the impurity subspace C to the screening needs to be subtracted out. By doing so one arrives at the so-called constrained random phase approximation (cRPA) [101].

In order to calculate the effective screened interaction $U_{ik,j}$ of subspace C within cRPA we first define the so-called polarizability region P in which screening processes due to formation of e–h pairs are taken into account for calculating the screened interaction. P comprises the strongly correlated subspace C and a sufficient portion of the surrounding atoms of subspace C as is schematically indicated in figure 1. In principle, the whole D region could be chosen as P. However, in practice this is often not feasible because of computational limitations if the device region is reasonably big. Also as it turns out the screening of the correlated subspace C by the surrounding conduction electrons is relatively localized due to the usually localized nature of the strongly correlated orbitals making up C.

Within RPA the screened interaction $W$ is given by the Dyson equation shown diagrammatically in figure 3 which in an atomic orbital basis set can be written algebraically as

$$W_{\alpha_1\beta_1_1;\alpha_2\beta_2_2}(\tau_1, \tau_2) = -W_{\alpha_1\beta_1_1;\alpha_2\beta_2_2}(\tau_1, \tau_2)$$

$$- \sum_{\mu_1\nu_1\mu_2\nu_2} V_{\alpha_1\beta_1_1;\mu_1\nu_1} \int_0^\beta d\tau' (\tilde{\Pi}_P)_{\mu_1\nu_1;\mu_2\nu_2}(\tau_1, \tau') \times W_{\mu_2\nu_2;\alpha_2\beta_2_2}(\tau, \tau')$$

For the screening of the bare Coulomb interaction $V$ only screening processes within region P are taken into account. Hence we have to calculate the polarizability (i.e. the bubble diagram in figure 3) projected onto the P region:

$$(\tilde{\Pi}_P)_{\alpha_1\beta_1;\alpha_2\beta_2}(\tau, \tau') = \sum_\sigma (\tilde{G}_D^{\sigma\sigma}(\tau, \tau')) (\tilde{G}_D^{\sigma\sigma}(\tau', \tau')$$

where the projected GF for the P region can be obtained from the device GF according to (5) as

$$\tilde{G}_P = S_p^{-1} S_{PD} \hat{G}_D S_{DP} S_p^{-1}$$

For a stationary Hamiltonian we can replace the two times in the screened interaction and polarizability by time differences: $\Pi(\tau_1, \tau_2) \rightarrow \Pi(\tau_1 - \tau_2)$ and $W(\tau_1, \tau_2) \rightarrow W(\tau_1 - \tau_2)$. Hence (by setting $\tau_2 = 0$ and after some renaming), we can write the Dyson equation for the RPA screened interaction as:

$$W_{\alpha_1\beta_1_1;\alpha_2\beta_2_2}(\tau) = V_{\alpha_1\beta_1_1;\alpha_2\beta_2_2}(\tau) \times \delta(\tau)$$

$$+ \sum_{\mu_1\nu_1\mu_2\nu_2} V_{\alpha_1\beta_1_1;\mu_1\nu_1} \int_0^\beta d\tau' (\tilde{\Pi}_P)_{\mu_1\nu_1;\mu_2\nu_2}(\tau - \tau') \times W_{\mu_2\nu_2;\alpha_2\beta_2_2}(\tau')$$

Here we will only consider the static limit of the screened interaction, i.e. $W^0 \equiv W(\omega = 0) = \int d\tau W(\tau)$. Because of the $\beta$-periodicity of $\Pi(\tau)$ we also have $\int_0^\beta d\tau' \Pi(\tau - \tau') = \int_0^\beta d\tau' \Pi(\tau) = \Pi^0$. Hence we obtain the following Dyson equation for the static screened interaction $W^0$:

$$W^0_{\alpha_1\beta_1_1;\alpha_2\beta_2_2} = V_{\alpha_1\beta_1_1;\alpha_2\beta_2_2} + \sum_{\mu_1\nu_1\mu_2\nu_2} V_{\alpha_1\beta_1_1;\mu_1\nu_1} \tilde{\Pi}^{0}_{\mu_1\nu_1;\mu_2\nu_2}(\tau_1, \tau_2) \times W^0_{\mu_2\nu_2;\alpha_2\beta_2_2}$$

In the diagrammatic representation the bare Coulomb interaction $V$ is shown by wiggly lines, while the RPA screened interaction $W$ is shown by double wiggly lines. The hybridization function $\Delta$ is shown by the shaded area. The coupling of the subspace $C$ to the rest of the system is shown by the solid line.

Figure 3. Dyson equation for RPA screened interaction for atomic basis set. Wiggly lines correspond to the bare Coulomb interaction $V$, double wiggly lines to the RPA screened interaction $W$.
The static Polarizability $\Pi^0$ is now found easily by integrating a Green’s function product over the frequency domain:

$$ (\tilde{\Pi}^0_p)_{\mu_1\nu_1\mu_2\nu_2} = \int_0^\beta d\tau (\tilde{\Pi}_p)_{\mu_1\nu_1\mu_2\nu_2}(\tau) $$

$$ = \int_0^\beta d\tau \sum_\sigma (\tilde{G}^0_p)^{\sigma}_{\nu_1\mu_1}(\tau) (\tilde{G}^0_p)^{\sigma}_{\nu_2\mu_2}(\tau) $$

$$ = \sum_\sigma \sum_\nu \sum_\mu \int_0^\beta d\tau (\tilde{G}^0_p)^{\sigma}_{\nu\mu}(\tau) (\tilde{G}^0_p)^{\sigma}_{\nu\mu}(\tau) $$

$$ \rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \sum_\sigma (\tilde{G}^0_p)^{\sigma}_{\nu\mu}(\omega) (\tilde{G}^0_p)^{\sigma}_{\nu\mu}(\omega) $$

where in the last step we have taken the zero temperature limit ($\beta \rightarrow \infty$) rendering the discrete Matsubara frequencies continuous.

We now define superindices $I := (i_1, i_2)$ in order to rewrite the Dyson equation in form of a matrix equation. Hence we have $W^0 = (W^0_{ij})$ etc., and the Dyson equation can be written in matrix form as:

$$ W^0 = V + V \tilde{\Pi}^0_p W^0 $$

Solving for the static screened interaction $W^0$ we find:

$$ W^0 = (I - V \tilde{\Pi}^0_p)^{-1} V $$

Projection to the correlated subspace $C$ then yields the RPA screened interaction for the correlated electrons $W^0_C$. However, since the screening within the correlated subspace will already be taken into account by the impurity solver in a more or less exact way, the screening of the correlated electrons by themselves has to be subtracted out in order to obtain the effective interaction $U$. Hence the effective interaction $U$ is partially screened interaction that results in the fully RPA screened interaction $W^0_C$ when taking into account only the polarizability $\Pi_C$ within the $C$ subspace. The corresponding Dyson equation is shown diagrammatically in figure 4. Solving for the effective interaction $U$ we arrive at the following ‘unscreening’ equation [102–104] computing $U$ from $W^0_C$:

$$ U = W^0_C (I_C + \Pi^0_C W^0_C)^{-1} $$

In order to determine the screening within subspace $C$, we have to calculate the polarizability corresponding to subspace $C$.

$$ \Pi^0_C = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \sum_\sigma (\tilde{G}_C)^{\sigma}_{\nu\mu}(\omega) (\tilde{G}_C)^{\sigma}_{\nu\mu}(\omega) $$

As the fully screened interaction $W^0_C$ is usually quite small (compared to the bare Coulomb interaction) and positive, $[W^0_C]^{-1}$ is big and positive. On the other hand the screening of the correlated electrons by themselves is usually strong, and therefore $\tilde{\Pi}^0_C$ is big and negative. Hence in order to obtain $U$ we are basically subtracting two relatively big numbers and inverting the resulting small number so that relatively small errors in calculating $W^0_C$ or $\Pi^0_C$ can result in quite large errors for $U$. It should be noted here that in the case of a semiconducting or insulating substrate or host material, as well as in the case of insulating compounds this issue is less problematic since then at low energies around the Fermi level, the two subspaces are completely decoupled, leading to weaker ‘self-screening’ of the correlated electrons, and hence smaller numbers for $[W^0_C]^{-1}$ and $\Pi^0_C$.

However, in the case of a metallic host or substrate considered here, it is thus crucial to correctly perform the projections of the different quantities involved in the calculation of the effective interaction (Green’s functions, polarizability) to the $P$ and $C$ subspaces in order to reliably calculate $U$. Also the usual neglect of certain product basis states in the computation of the screened interaction [105] might be problematic in this context. One way to stabilize the numerical evaluation of (42) is to decouple the correlated subspace from the rest of the system both in the calculation of $W$ and of $U$ as proposed by Miyake et al. [106]. This way the self-screening of the correlated electrons is reduced considerably, leading to smaller values of $[W^0_C]^{-1}$ and $\Pi^0_C$, and thus enhancing the numerical stability. However, this can lead to far too high matrix elements for the direct Coulomb interaction as will be shown in section 3. Apparently, ‘mixed propagators’ between the correlated subspace and the rest of the system (which vanish when the subspaces are decoupled) can be quite important for the screening of the effective interaction.

2.7. Solution of the Anderson impurity model: one-crossing approximation

Since the interaction $U_{ij\bar{k}l}$ is strong in comparison with the single-particle broadening (given by the imaginary part of

\[ \text{Figure 4. Dyson equation for fully screened RPA interaction } W^0_C \text{ of subspace } C \text{ in terms of effective interaction } U. \text{ Orbital indexes have been suppressed here.} \]
\( \Delta_C(\omega) \), the AIM problem cannot be solved by standard perturbation theory in the Coulomb interaction. Instead more advanced many-body methods usually starting from an exact diagonalization of the full impurity Hamiltonian \( \hat{H}_C \) have to be employed in order to properly take into account the strong correlations within subspace C. Here I use the one-crossing approximation (OCA) \([107]\) which is an improvement over the non-crossing approximation (NCA) \([108–110]\). However, it should be emphasized that the methodology presented so far can in principle be combined with any other method for solving the AIM, as e.g. continuous time quantum Monte-Carlo (CTQMC) \([111]\), or numerical renormalization group (NRG) \([112]\), or the Lanczos diagonalization scheme \([76]\).

One advantage of OCA over other schemes is that spectral data can be calculated directly on the real frequency axis. Hence in contrast to the numerically exact CTQMC, for example, it does not suffer from artifacts introduced by numerical analytic continuation of the spectra from the Matsubara axis to the real axis. Also spurious features in the spectra coming from the approximation of the infinite and continuous conduction electron bath in the Anderson model by a finite and discrete one as in direct diagonalization schemes such as Lanczos, are not a problem for OCA since the bath is not truncated or discretized. On the other hand, in contrast to the basically exact but computationally very demanding NRG, OCA can actually be applied to realistic Anderson models of 3d- and 4f-impurities with 5 and 7 impurity-levels, respectively.

However, being an approximate method, OCA also suffers from some deficiencies that one should be aware of. First, as in the case of the simpler NCA, spurious non-Fermi liquid behaviour is obtained in the zero-temperature limit, resulting in artifacts in the spectral density for low temperatures. While in NCA these artifacts already appear below \( T_K \), in OCA the critical temperature below which the artifacts appear is significantly lower (1–2 orders below \( T_K \)). Another problem of NCA and OCA is the violation of certain sum rules especially in the case of multi-orbital Anderson models that lead to errors in the high frequency expansion of the electronic self-energy \([113]\). Again, these errors are much less pronounced in OCA than in NCA.

The basic idea of both NCA and OCA methods is to treat the coupling of the correlated subspace C to the rest of the system given by the hybridization function \( \Delta_C(\omega) \) as a perturbation to the dynamics within the subspace induced by the strong electron-electron interactions which is treated exactly. Hence the starting point is an exact diagonalization of the many-body Hamiltonian of the correlated subspace:

\[
\hat{\mathcal{H}}_C = \sum_m E_m |m\rangle\langle m|
\]

where \(|m\rangle \) are the many-body eigenstates of \( \mathcal{H}_C \) and \( E_m \) the corresponding eigen-energies.

It is now convenient to represent the many-body eigenstates of \(|m\rangle \), in terms of auxiliary fields or pseudo-particles (PPs) \( \hat{a}_m, \hat{a}_m^\dagger \) which obey (anti-)commutation rules depending on the number of electrons represented by the corresponding many-body state \(|m\rangle \). The physical electron operators \( c_\sigma, c_\sigma^\dagger \) are related to the PP operators by:

\[
c_\sigma = \sum_{m,n} F_{\sigma m} F_{\sigma n}^\dagger a_m a_n
\]

where \( F_{\sigma m} \) are the matrix elements of the electron annihilation operator with the many-body eigenstates of \( \mathcal{C} \): \( F_{\sigma m} = \langle m|c_\sigma|n\rangle \). Since the PPs obey (anti-)commutation rules a diagrammatic expansion of PP propagators in terms of the coupling to the rest of the system is possible. The full PP propagator corresponding to a many-body state \(|m\rangle \) is then given by

\[
G_m(\omega) = \frac{1}{\omega - \lambda - E_m - \Sigma_m(\omega)}
\]

where \( \Sigma_m(\omega) \) is the PP self-energy describing the dynamic interaction with the other PPs induced by the hybridization with the rest of the system (bath).

NCA consists in an infinite resummation of self-energy diagrams where conduction electron lines do not cross (hence the name). These are the diagrams shown in the left box in figure 5 for a certain PP \( m \) representing a many-body state of \( N \) electrons. The NCA diagrams describe processes where a single electron (hole) jumps from the bath to subspace C and back thereby temporarily creating a PP with \( N + 1 \) \((N - 1)\) electrons. Hence the NCA self-energy is given by a convolution of the hybridization function \( \Delta_C(\omega) \) with the PP propagators \( G_m \) of the PPs \( m' \) coupled to \( m \). OCA additionally...
takes into account diagrams where two bath electron lines cross as shown in the right box of figure 5. The algebraic expressions for the OCA self-energies involve double convolutions of two hybridization functions with three PP propagators. The exact algebraic expressions can be found in the literature [103, 107]. Since the self-energy of a PP propagator depends on the dressed propagators of the other PP propagators that interact via V lob with m, the NCA/OCA equations have to be solved self-consistently.

Once the NCA/OCA equations have been solved, the real and imaginary electronic quantities can be calculated from the PP spectral functions as

$$\rho_{\sigma}(\omega) = \frac{1}{Q} \sum_{mm'} \int d\varepsilon \, e^{-\beta\varepsilon} [1 + e^{-\beta\varepsilon}] |F_{mm'}|^{2} \times A_{m}(\varepsilon) A_{m'}(\omega + \varepsilon)$$

where $A_{m}(\varepsilon) = -\text{Im} G_{m}(\varepsilon)/\pi$ is the PP spectral function for PP propagator $m$ and $Q$ is the PP charge which is obtained by integration of the PP spectral functions. Again, in OCA the expression for calculating the electronic density $\rho_{\sigma}(\omega)$ is more complicated, involving double convolutions of PP spectral functions. From the electron spectral density $\rho_{\sigma}(\omega)$ being the imaginary part (modulo $\pi$) of the electron Green’s function $G_{\sigma}(\omega)$ we can calculate the real part of $G_{\sigma}(\omega)$ by Kramers-Kronig. Finally, from the GF $\hat{G}_{C}(\omega)$ the electronic self-energy describing the dynamic correlations within C is obtained by $\Sigma_{C}(\omega) = [\hat{G}_{C}(\omega)]^{-1} - [\hat{G}^{0}_{C}(\omega)]^{-1}$ where $\hat{G}^{0}_{C}(\omega) = (\omega + \mu)\hat{P}_{C} - \hat{H}_{C}^{0} - \Delta_{C}(\omega))^{-1}$ is the bare propagator of subspace C. For a more detailed account of the NCA, OCA and other methods based on a hybridization expansion of atomic states see e.g. [71, 103].

2.8. Feedback of the self-energy: correlated electronic structure and transport properties

Once we have solved the Anderson impurity model for the strongly interacting subspace C coupled to the rest of the system, we obtain the electronic self-energy describing the strong dynamic correlations within subspace C:

$$\Sigma_{C}(\omega) = \sum_{i,j \in C} |i\rangle \langle j| \Sigma_{C}(\omega)

Note that $\Sigma_{C} = \Sigma_{C}$ since we have assumed the basis to be orthonormal within subspace C. This self-energy is now fed back to the DFT part in order to obtain the correlated electronic structure and transport properties of the system. More specifically, we obtain the correlated device GF

$$\hat{G}_{D}(\omega) = \left( [\hat{G}^{0}_{D}(\omega)]^{-1} - [\Sigma_{C}(\omega) - \hat{V}_{C}^{\text{DC}}]^{-1} \right)$$

where $\hat{V}_{C}^{\text{DC}}$ is the DCC operator which like $\Sigma_{C}(\omega)$ only acts on C. According to (7) the corresponding nuclear matrix of the device GF is

$$\hat{G}_{D}(\omega) = ([\hat{G}^{0}_{D}(\omega)]^{-1} - S_{\text{DC}} [\Sigma_{C}(\omega) - \hat{V}_{C}^{\text{DC}} S_{\text{CD}}]^{-1}$$

where the overlap matrices $S_{\text{DC}}$ and $S_{\text{CD}}$ sandwiching $\Sigma_{C}(\omega) - \hat{V}_{C}^{\text{DC}}$ account for the overlap between the correlated subspace C and the rest of the system (see equation (7)).

From the correlated device GF $\hat{G}_{D}(\omega)$ we can calculate the correlated electronic density analogously to (23) by integration of $\hat{G}_{D}(\omega)$ up to 0 energy:

$$\tilde{\rho}_{D} = -\text{Im} \frac{1}{\pi} \int_{-\infty}^{0} d\omega \tilde{G}_{D}(\omega + i\eta)$$

From the correlated density in turn a KS Hamiltonian for the device region can be calculated, from which a new correlated density is obtained and so forth until self-consistency is reached. Hence we can calculate the effect of the correlation within the C subspace onto the charge distribution of the device region. This part corresponds to the so-called ‘charge self-consistency’ loop within the DFT+DMFT scheme [73].

Following Meir–Wingreen [114], the low-bias transport properties can be obtained in complete analogy to the case of KS-DFT transport equations (24)–(27) even in the presence of strong correlations from the correlated transmission function

$$T(\omega) = \text{Tr}[\Gamma_{L}(\omega)\hat{G}_{D}(\omega)\Gamma_{R}(\omega)\hat{G}_{D}(\omega)]$$

Note that the strong correlations giving rise e.g. to the Kondo effect are actually contained in $T(\omega)$ via the correlated GF $\hat{G}_{D}(\omega)$.

In the next section we will see that the Fano–Kondo lineshapes measured by STM spectroscopy of magnetic atoms and molecules on metal substrates can indeed be reproduced by calculating the conductance from the (zero-bias) transmission function. This is due to the fact that the Kondo effect is a low-energy phenomenon, i.e. the Kondo peak is observed for very small bias voltages so that finite-bias effects only play a minor role. For the description of actual non-equilibrium phenomena the formalism has to be generalized to include the effect of finite bias voltages. As shown by Meir and Wingreen in their landmark papers [114] this can be achieved by generalization of the formalism to the Keldysh contour. However, in this case the Anderson impurity problem has to be solved out of equilibrium which is computationally extremely demanding. So far it has only been achieved in the context of the single-level AIM [115–117], but not for realistic cases.

3. Results: Co adatom at the Cu(001) surface

Now the developed methodology is applied to the case of a Co adatom deposited on the Cu(001) surface. This system is an ideal testbed for the theory as it has been measured extensively in the recent past [80, 82, 83, 118–120]. Figure 6(a) shows the atomic structure of the device region. The device contains the Co atom on three layers of the Cu(001) surface and an STM tip consisting of a Cu pyramid grown in the (001) direction. The Co atom and its four nearest neighbour Cu atoms have been relaxed with Gaussian09 [86] using the local spin density approximation (LSDA) and the LANL2DZ double-zeta valence plus outer core electron basis set with core pseudo potentials [121] while the rest of the device atoms have
been kept fixed. The interlayer and intralayer distances for the fixed Cu atoms are those of a perfect Cu surface taken from [122]. In good agreement with [120, 123], I find that the Co atom relaxes at a height of about 1.5 Å above the four nearest neighbour Cu atoms while these in turn are pushed by 0.1 Å into the substrate.

Using ANT.G and the LANL2MB minimal basis set including valence and outer core electrons with pseudo potentials [121] the electronic and magnetic structure of the device coupled to the tip and substrate electrodes is calculated within DFT on the level of LSDA. The Co atom is essentially a $4s^23d^8$ configuration with the two holes in the $3d$-shell giving rise to an approximate spin-1 state of the Co atom (see table 1) again in good agreement with [123]. LSDA basically predicts a mixed valence situation for all the orbitals with the individual occupations around 1.6 with the exception of the $xy$-orbital which is nearly full.

From the LSDA electronic structure the effective Coulomb interaction $U_{ij,k\bar{k}}$ for the Co $3d$-shell is calculated as described in section 2.6. For the P region we take into account substrate atoms up to the 3rd nearest neighbour, i.e. the 9 Cu atoms closest to the Co adatom. The change in $U$ from taking into account 2nd nearest neighbours to 3rd nearest neighbours is about 2%. For 4th nearest neighbours (18 atoms in total) the super matrices in the RPA equation (41) become too big (linear matrix dimension $234^2 = 54756$) to be handled.

Table 2 shows the matrix elements of the effective Coulomb interaction, namely the direct Coulomb repulsion matrix elements (density-density interaction) $U_{ij,k\bar{k}}$ and the exchange interaction matrix elements (Hund’s rule coupling) $U_{k\bar{k},i}$. The average density-density interaction is $\overline{U} = 4.14$ eV. It is strongly screened by the conduction electrons, resulting in a reduction of over 80% compared to the bare value of 22.9 eV for the Co $3d$-shell. On the other hand, the Hund’s rule coupling is much less affected by the screening: it is only reduced by about 10% from its bare value of 0.85 eV, resulting in an average Hund’s coupling of $J_H = 0.77$ eV. Note that the inter-orbital Coulomb repulsion ($U_{ij,k\bar{k}}$ for $i \neq k$) is related to the average intra-orbital Coulomb repulsion for both orbitals and the Hund’s rule coupling via $U_{ij,k\bar{k}} = (U_{ij,i} + U_{k\bar{k},k})/2 - 2U_{i\bar{k},i}$.

Both density-density interactions and Hund’s rule coupling are somewhat anisotropic (i.e. orbital-dependent). The intra-orbital Coulomb repulsion $U_{ij,i}$ deviates only by up to 0.17 eV or by to 3% from its mean value of $U = 5.4$ eV. The variation is stronger for the inter-orbital Coulomb repulsion $U_{ij,k\bar{k}}$, deviating by up to 0.43 eV or by up to 11% from its mean value of $U' = 3.85$ eV. The Hund’s rule coupling $U_{k\bar{k},i}$ deviates even stronger by up to 0.29 eV or by up to 38% from its mean value of $J_H = 0.77$ eV. It is worth noting at this point that the complete decoupling of the correlated subspace from the rest of the system as proposed in [106] in order to achieve a stable computation of the effective interaction in the case of ‘entangled bands’ produces a much higher density-density interaction of about 12 eV. Apparently the screening effects of ‘mixed propagators’ between the correlated subspace and the rest of the system are actually quite important and cannot be neglected.

Next, the electronic structure of the system is calculated for the paramagnetic case on the level of the local density-approximation (LDA) in order to obtain the KS energy levels of the Co $3d$-shell $\epsilon^0_d$ and hybridization functions $\Delta_d(\omega)$ in the absence of spin-polarization. Figure 6(b) shows the imaginary parts of the hybridization functions $\Delta_d(\omega)$ for each of the Co $3d$-orbitals which yields the (dynamic) broadening of the orbitals due to the coupling to the substrate. We see that the broadening near the Fermi level is basically featureless indicating coupling to the delocalized Cu $4s$-states of the substrate. As can be seen the degenerate $xz$- and $yz$-levels couple most strongly to the these states. Because of their shape these two orbitals couple very well to the $4s$-states of the four Cu atoms directly underneath the Co adatom. On the other hand, the coupling of the $xy$-orbital to the substrate is the weakest for all five orbitals since the direct coupling to the

| Table 1. Total and orbital resolved occupations and spin of Co $3d$-shell within DFT on the level of LSDA and LDA and DFT+OCA calculations. |
|---|---|---|---|---|---|---|
| | $N_d$ | $z^2$ | $xz$ | $yz$ | $x^2 - y^2$ | $xy$ | $\delta_4$ |
| LSDA | 8.13 | 1.59 | 1.66 | 1.66 | 1.33 | 1.89 | 0.82 |
| LDA | 8.25 | 1.66 | 1.63 | 1.52 | 1.52 | 1.81 | 0.86 |
| OCA | 8.26 | 1.34 | 1.94 | 1.94 | 1.08 | 1.97 | 0.86 |
| OCA (+0.4 eV) | 8.06 | 1.11 | 1.96 | 1.96 | 1.06 | 1.96 | 0.86 |

Note: In the last line we show the DFT+OCA results for the Co $3d$-levels $\epsilon_d$ shifted by 0.4 eV upwards in energy with respect to the FLL.

| Table 2. Direct Coulomb repulsion matrix elements $U_{i\bar{k},k\bar{k}}$ (density-density interaction) and exchange matrix elements $U_{i\bar{k},k\bar{k}}$ (Hund’s rule coupling) of effective Coulomb interaction for Co $3d$-shell. |
|---|---|---|---|---|---|---|---|
| density-density interaction (eV) | Hund’s coupling (eV) |
| $z^2$ | xz | yz | $x^2 - y^2$ | $xy$ | $z^2$ | xz | yz | $x^2 - y^2$ |
| $z^2$ | 5.38 | 4.27 | 4.27 | 3.45 | 3.46 |
| xz | 4.27 | 5.56 | 3.86 | 3.73 | 3.74 | 0.60 |
| yz | 4.27 | 3.86 | 5.56 | 3.73 | 3.74 | 0.60 | 0.83 |
| $x^2 - y^2$ | 3.45 | 3.73 | 3.73 | 5.23 | 4.28 | 0.94 | 0.82 | 0.82 |
| $xy$ | 3.46 | 3.74 | 3.74 | 4.28 | 5.26 | 0.92 | 0.83 | 0.83 | 0.48 |
Figure 6. Results for Co adatom at Cu(001) surface. (a) Atomic structure of device part. The Co adatom is shown in grey. (b) Orbitally resolved imaginary part of hybridization functions for Co 3d-shell. (c) Orbitally resolved OCA spectral functions for Co 3d-shell at T ∼ 10 K. (d) Total occupation of Co 3d-shell as a function of energy shift ∆ε. (e) Half-width of Kondo feature in z² spectral function as a function of the total shift ∆ε of Co 3d-levels with respect to energy levels given by FLL DCC. (f) Spectral functions of Co z²-orbital for different energy shifts ∆ε at T ∼ 10 K. (g) DFT+OCA transmission functions for different energy shifts ∆ε (line colours as in (f)).

underneath Cu atoms is strongly suppressed due to symmetry reasons. The coupling of the z²- and the x²−y²-orbitals to the substrate is intermediate between these two cases. At negative energies, the coupling to the localized Cu 3d-states of the substrate leads to strong peaks in the hybridization functions at energies between −5 and −2 eV. Less pronounced peaks at positive energies above 4 eV indicate coupling to the Cu 4p-orbitals of the substrate.

The bare energies εd of the Co 3d-levels constituting the impurity shell in the Anderson impurity model are obtained from their KS energies εd = PdH0Pd corrected by a DCC term, as explained earlier in section 2.5. The so-called FLL generalized to an anisotropic (i.e. orbital-dependent) density-density interaction Ulij,kk is employed (31). The values for the direct Coulomb repulsion are the ones shown in table 2. For the Hund’s rule coupling the orbital averaged exchange interaction is taken, i.e. JH = 0.77 eV.

The Anderson impurity problem presented by the interacting Co 3d-shell coupled to the substrate is now solved within OCA as described in section 2.7. For the effective Coulomb interaction of the Co 3d-shell we take into account the density-density interactions Ulij,kk as well as the exchange interactions Uik,kj as given in table 2. At the energy levels for the Co 3d-orbitals given by the FLL-DCC (31), the total occupancy for the Co 3d-shell is about 8.2 electrons similar to the ones of the LDA and LSDA calculations (see table 1).
However, the individual occupancies of the 3d-orbitals are now quite different from the DFT ones, namely they are now closer to integer occupancies, as opposed to the mixed-valence situations obtained in the DFT calculations. In particular, the \( x^2 - y^2 \)-orbital is now basically half-filled, and the \( xz, yz \)-and \( xy \)-orbitals are nearly full now. The \( z^2 \)-orbital is now also closer to half-filling than before but still has strong charge fluctuations (occupancy \( \sim 1.3 \)). Similar to LSDA, the spin of the Co 3d-shell is found to be \( S_d \sim 0.87 \), close to a spin-1 configuration.

In figure 6(c) the calculated spectral functions of the Co 3d-orbitals \( \rho_d(\omega) \) (at \( T \sim 10 \text{ K} \)) are shown. We can see a very strong Kondo peak at the Fermi level in the \( z^2 \)-orbital. The upper Hubbard peak is here quite close to the Kondo peak at the Fermi level due to the strong charge fluctuations. This orbital is still quite close to a mixed-valence situation. The \( x^2 - y^2 \)-orbital despite being half-filled and thus bearing a spin-1/2 does not yield a Kondo peak. We are dealing here essentially with a so-called underscreened Kondo effect [124, 125, 126]: Despite the relatively similar hybridization of the \( z^2 \)-channel and the \( x^2 - y^2 \)-channel, the Kondo temperature \( T_{K,z^2} \) of the \( z^2 \)-channel is much higher than that of the \( x^2 - y^2 \)-channel, \( T_{K,x^2-y^2} \), due to its stronger charge fluctuations. Hence at finite temperature \( T \) with \( T_{K,x^2-y^2} < T < T_{K,z^2} \), only the spin-1/2 in the \( z^2 \)-channel is Kondo-screened, while the spin-1/2 in the \( x^2 - y^2 \) channel remains unscreened.

The half-width of the Kondo peak is about 90 K, in very good agreement with the experimentally observed values [80, 82, 83]. However, the lineshape of the calculated transmission function (red curve in figure 6(g)) is rather peak-like, different from the experimentally observed asymmetric Fano-lineshapes. As the DCC for DFT is not exactly known and equation (31) is only an approximation, we now shift the Co 3d-levels upwards in energy by an amount \( \Delta \varepsilon \) thus emptying the Co 3d-shell as can be seen in figure 6(d). Emptying the Co 3d-shell mainly lowers the occupancy of the \( z^2 \)-orbital reducing the charge fluctuations for that orbital, while the occupancies of the other orbitals are quite stable. Figure 6(f) shows the effect of shifting the Co 3d-levels and the concomitant reduction of charge fluctuations on the the spectral function of the \( z^2 \)-orbital (at \( T \sim 10 \text{ K} \)): As the \( z^2 \)-orbital is emptied, its occupation approaches 1, the Kondo peak becomes smaller, and the upper Hubbard peak moves away from the Fermi level. The width of the Kondo peak decreases at first and then starts to grow again for shifts \( \gtrsim 0.2 \text{ eV} \), as can be seen in figure 6(e). Note that the non-monotonic behaviour of the width of the Kondo peak is actually a finite temperature effect: As the actual Kondo temperature decreases with decreasing charge fluctuations, the Kondo peak in the finite temperature spectra (here \( T \sim 10 \text{ K} \)) does not attain its full (zero-temperature) height anymore. As the height of the (finite-\( T \)) Kondo peak decreases, its width starts to grow again at some point. Hence the half-width of the Kondo peak measured at some finite temperature really only yields an apparent Kondo temperature.

Figure 6(g) shows the effect of shifting the Co 3d-levels on the low-energy transmission spectra. As said above, for the 3d-levels at the values given by the FLL-DCC the transmission function near zero energy is rather peak-like, unlike the ones observed experimentally. But when shifting the 3d-levels upwards in energy the lineshapes become more asymmetric Fano-like. Good agreement between theoretical and experimental Fano-lineshapes is achieved for shifts between 0.4–0.5 eV. In this regime the half-width of the Kondo peak is between 67 K and 86 K, in good agreement with the experimentally observed ones between 70 K and 100 K for the Co on Cu(001) system [80, 82, 83]. These results are quite different from those obtained recently for the case of Co on Cu(1 1 1) with a similar approach [127] where all the Co 3d-orbitals give rise to Kondo-like resonances at the Fermi level. The reason could be the altogether quite different geometric situation at the (1 1 1) surface leading to a decidedly different symmetry and occupancy for the Co 3d-shell.

4. Conclusions

In conclusion, an \textit{ab initio} methodology has been developed for describing the impact of strong electronic correlations on the electronic structure and transport properties of nanoscale devices. Starting from the DFT electronic structure of an embedded nanoscale device, an Anderson impurity model is constructed by projection of the Kohn–Sham Hamiltonian onto the correlated subspace. The effective Coulomb interaction \( U \) for the correlated subspace (impurity) is calculated \textit{ab initio} from the DFT electronic structure by making use of the constrained RPA approach. The solution of the Anderson impurity model yields the dynamic correlations originating from strong interactions within the correlated subspace in form of a self-energy which is fed back to the DFT calculation in order to obtain the correlated electronic structure and transport properties.

The methodology has been tested for the case of a single Co adatom on Cu(001) substrate. On a qualitative level the results are in good agreement with experiments: A Fano–Kondo feature with the width in good agreement with experiments is obtained in the calculated low-energy tunnelling spectra. However, the lineshape of the Fano–Kondo feature is not correctly reproduced at the energies for the Co 3d-levels given by the double-counting correction. Only when shifting the Co 3d-levels slightly upwards in energy good agreement with the experimentally observed lineshapes is achieved. It is a well known problem of DFT+U and DFT+DMFT approaches that the double-counting correction is not exactly known and in general does not yield the correct position (and thus charge) of the correlated levels. Nevertheless, the so-called fully-localized limit employed here, is actually not too far off as only moderate shifts are necessary to achieve good quantitative agreement with experiments. Importantly, the physics is actually not affected by the shifting of the Co 3d-levels: Independent of the shift (in that energy range) the Co 3d-shell constitutes essentially a spin-1 system that experiences an underscreened Kondo effect. The shifting only affects the weight of the Kondo peak by lowering the charge fluctuations in the Kondo-screened orbital.

Hence the developed methodology is capable of qualitative predictions of strong correlation phenomena. But
accurate quantitative predictions for example of Kondo temperatures and the exact shapes of Fano–Kondo features are difficult as these are dependent on the exact occupancy of the correlated subspace which cannot be calculated accurately because of the approximate nature of the double-counting correction in our approach. One possibility to overcome these difficulties is to make use of the GW approach instead of DFT for the description of the weakly interacting part of the system, similar to the GW+DMFT approach for strongly correlated materials [128–130] since for GW the double-counting correction term is exactly known.

Acknowledgments

I would like to thank J Palacios and M Soriano for fruitful discussions about projections in non-orthogonal basis sets, and S Biermann and K Haule for stimulating discussions on the calculation of the effective interaction U. I am also grateful to Hardy Gross for his support and encouragement.

References

[1] Heath J R and Ratner M A 2003 Molecular electronics Phys. Today 56 43
[2] Joachim C and Ratner M A 2005 Molecular electronics: some views on transport junctions and beyond Proc. Natl Acad. Sci. USA 102 8801
[3] Cuniberti G, Fagas G and Richter K 2005 Introducing Molecular Electronics (Berlin: Springer)
[4] Cuevas J C and Scheer E 2010 Molecular Electronics (Singapore: World Scientific)
[5] Aviram A and Ratner M A 19764 Molecular rectifiers Chem. Phys. Lett. 29 277
[6] Elbing M, Ochs R, Koentopp M, Fischer M, von Hnisch C, Weigend F, Evers F, Weber H B and Mayor M 2005 A single-electron diode Proc. Natl Acad. Sci. 102 8815
[7] Diez-Pérez I, Hithath J, Lee Y, Yu L, Adamska L, Kozhushner M A, Olynyk I I and Tao N 2009 Rectification and stability of a single molecular diode with controlled orientation Nat. Chem. 1 635
[8] Batra A, Darancet P, Chen Q, Meisner J S, Widawsky J R, Neaton J B, Nuckolls C and Venkataraman L 2013 Tuning rectification in single-molecular diodes Nanolett. 13 6233
[9] Tans S J, Verscheren A R M and Dekker C 1998 Room-temperature transistor based on single carbon nanotube Nature 393 49
[10] Martel R, Schmidt T, Shea H R, Hertel T and Avouris Ph 1998 Single- and multi-wall carbon nanotube field-effect transistors Appl. Phys. Lett. 73 2447
[11] Song H, Kim Y, Jang Y H, Jeong H, Reed M A and Lee H 2009 Observation of molecular orbital gating Nature 462 1039
[12] Žutić I and Das Sarma S 2004 Spintronics: fundamentals and applications Rev. Mod. Phys. 76 323
[13] Bogani L and Wernsdorfer W 2008 Molecular spintronics using single-molecule magnets Nat. Mater. 7 179
[14] Gatteschi D, Sessoli R and Villain J 2006 Molecular Nanomagnets (Oxford: Oxford University Press)
[15] Agraït N, Yegati A L and van Ruitenbeek J M 2003 Quantum properties of atomic-sized conductors Phys. Rep. 377 81
[16] Liang W, Bockrath M, Bozovic D, Hafner J H, Tinkham M and Park H 2001 Fábr–Peró interference in a tunnelot electron waveguide Nature 411 665
[17] Kong J, Yenilmez E, Tombler T W, Kim W, Dai H, Laughlin R B, Liu L, Jayanthi C S and Wu S Y 2001
[122] Da Silva J, Schroeder K and Blügel S 2004 First-principles investigation of the multilayer relaxation of stepped Cu surfaces Phys. Rev. B 69 245411
[123] Pick Š, Stepanyuk V, Baranov A, Hergert W and Bruno P 2003 Effect of atomic relaxations on magnetic properties of adatoms and small clusters Phys. Rev. B 68 104410
[124] Nozières Ph and Blandin A 1980 Kondo effect in real metals J. Phys. 41 193
[125] Coleman P and Pépin C 2003 Singular Fermi liquid behavior in the underscreened Kondo model Phys. Rev. B 68 220405
[126] Posazhennikova A, Bayani B and Coleman P 2007 Conductance of a spin-1 quantum dot: the two-stage Kondo effect Phys. Rev. B 75 245329
[127] Surer B, Troyer M, Werner Ph, Wehling T O, Läuchli A M, Wilhelm A and Lichtenstein A I 2012 Multiorbital Kondo physics of Co in Cu hosts Phys. Rev. B 85 085114
[128] Biermann S, Aryasetiawan F and Georges A 2003 First principles approach to the electronic structure of strongly correlated systems: combining GW with DMFT Phys. Rev. Lett. 90 086402
[129] Taranto C, Kaltak M, Parragh N, Sangiovanni G, Kresse G, Toschi A and Held K 2013 Comparing quasiparticle GW+DMFT and LDA+DMFT for the test bed material SrVO$_3$ Phys. Rev. B 88 165119
[130] Hansmann P, Ayral T, Vaugier L, Werner P and Biermann S 2013 Long-range coulomb interactions in surface systems: a first-principles description within self-consistently combined GW and dynamical mean-field theory Phys. Rev. Lett. 110 166401