

Coulomb interaction written in terms of U, J, etc. and DFT+U

Sohrab Ismail-Beigi

November 19, 2013

The idea of these notes is to look at localized Hubbard-like models for electron correlation effects and see what one can do exactly (very little) and then more approximately within a mean-field approximation such as DFT or Hartree-Fock or the like. One of the main aims is to present background information as well as some derivations and explanations of the various outcomes that are used in DFT+U calculations.

1 Matrix elements of the Coulomb operator

Start with the electron-electron interaction or Coulomb operator which is usually written as

$$\hat{V}_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|r_i - r_j|}$$

for a fixed number of electrons and antisymmetry imposed on the wavefunction this \hat{V}_{ee} acts on. It will be more convenient to work in second quantized notation so the operator contains the antisymmetry and we can work in an orbital basis directly. The same operator written using field operators is

$$\hat{V}_{ee} = \frac{1}{2} \sum_{\sigma, \sigma'} \int dr \int dr' \frac{\hat{\psi}_{\sigma}^{\dagger}(r) \hat{\psi}_{\sigma'}^{\dagger}(r') \hat{\psi}_{\sigma'}(r') \hat{\psi}_{\sigma}(r)}{|r - r'|}$$

where σ, σ' go over spin up and down and the field operator $\hat{\psi}_{\sigma}(r)$ annihilates (*i.e.*, removes) an electron with spin σ at position r . The form looks simple in term of spin (particular matchings and only two spin indices) since the Coulomb interaction is spin-independent. We will be using some type of orthonormal basis to work in, so we write the field operator in

that basis. Let's assume as is usual that the basis respects spin: there are spin up and spin down basis orbitals which we can label as $\phi_{n\sigma}(r)$. So

$$\hat{\psi}_\sigma(r) = \sum_n \phi_{n\sigma}(r) \hat{c}_{n\sigma}.$$

The $\hat{c}_{n\sigma}$ is the usual fermion annihilation operator that removes an electron from state n, σ . When you plug this in, you get

$$\hat{V}_{ee} = \frac{1}{2} \sum_{\sigma, \sigma'} \sum_{i, j, k, l} (i\sigma k\sigma' | V | j\sigma l\sigma') \hat{c}_{i\sigma}^\dagger \hat{c}_{k\sigma'}^\dagger \hat{c}_{l\sigma'} \hat{c}_{j\sigma}$$

where the matrix elements or integrals of the Coulomb operator are

$$(i\sigma k\sigma' | V | j\sigma l\sigma') = \int dr \int dr' \phi_{i\sigma}^*(r) \phi_{j\sigma}(r) \frac{1}{|r - r'|} \phi_{k\sigma'}^*(r') \phi_{l\sigma'}(r'). \quad (1)$$

This is the Coulomb interaction between two “charge” density distributions given by the pair products: density $\phi_{i\sigma}^*(r) \phi_{j\sigma}(r)$ and density $\phi_{k\sigma'}^*(r) \phi_{l\sigma'}(r)$. When $i = j$ or $k = l$, each is actually an honest probability density that is positive and integrates to one. But for $i \neq j$ or $k \neq l$, the density is in general complex and has zero integral due to the orthonormality of the basis. So we expect the largest integrals when $i = j$ and $k = l$ as the interaction of two distributions each with net “charge” +1; next is $i = j$ but $k \neq l$ (and converse) which is interaction of a net charge +1 with a neutral distribution; and smallest is $i \neq j$ and $k \neq l$ which is the interaction of two neutral distributions. Note that if all the indices are the same ($i = j = k = l$ and $\sigma = \sigma'$) we get zero since $\hat{c}_{i\sigma}^2 = 0$ for any fermion operator: you can't remove two electrons from one single orbital since you had at most one electron there to start with! So there is no term with $(i\sigma i\sigma | V | i\sigma i\sigma)$. This just says electrons don't interact with themselves (see below for more on this).

2 Coulomb operator for localized basis sets

This is pretty much all we can say at this very high level of generalization. To make any progress, we need some simplifying assumptions. If we're interested in the Coulombic physics coming from states that are well localized in space around atoms or from bands formed from such states, then we can make some progress. We can choose our basis states to be highly localized around each atom, which in principle means Wannier functions, but this could be localized orthogonalized atomic basis functions as well. The localization means that V integrals for orbitals on different atoms will be much smaller than those where the orbitals are on the same atom. Namely, i and j should be on the same atom and so should k and l — this just follows from the locality assumption of the basis and should be a good one since the localized states decay something like exponentially with separation. In addition, if i and k refer to different atoms, the integrals will be smaller than when they refer to the same atom: so we are ignoring inter-atomic V in comparison to intra-atomic V values. The drop off here

is not exponential since for $i = j$ and $k = l$ we are talking about the Coulomb interaction between two +1 net charge distributions that could be on neighboring atoms. This is likely much smaller than when the two are on the same atoms but the falloff is like $1/r$. At any rate, we drop these, although one can try to include them and their extensions of DFT+U (or many-body calculations) where nearest neighbor V matrix elements are retained in some way; within DFT+U, it is called the DFT+U+V method.

Anyways, we will assume that only intra-site V values are relevant. Then \hat{V}_{ee} falls apart into a sum over atomic sites at , and we will from now on use i, j, k, l to label orbitals on each atomic site:

$$\hat{V}_{ee} = \sum_{at} \frac{1}{2} \sum_{\sigma, \sigma'} \sum_{i, j, k, l} (i\sigma k\sigma' | V | j\sigma l\sigma') \hat{c}_{i\sigma}^\dagger \hat{c}_{k\sigma'}^\dagger \hat{c}_{l\sigma'} \hat{c}_{j\sigma}. \quad (2)$$

3 General considerations in the many-body case

If we only have one orbital of interest per atomic site, then $i = j = k = l$ is forced. The four operator sequence is then $\hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma'}^\dagger \hat{c}_{i\sigma'} \hat{c}_{i\sigma}$. This is zero if $\sigma = \sigma'$. We can also move the operators around a bit to get

$$\hat{V}_{ee} = \sum_{at} \frac{1}{2} \sum_{\sigma} (i\sigma i\bar{\sigma} | V | i\sigma i\bar{\sigma}) \hat{n}_{i\sigma} \hat{n}_{i\bar{\sigma}} = \sum_{at} U \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$$

where $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger c_{i\sigma}$ counts electrons in that state and $\bar{\sigma} = -\sigma$ is the opposite spin. We've recovered the Coulombic sector of the single site Hubbard Hamiltonian. U here is shorthand for the on-site, same-orbital, opposite-spin Coulomb integral.

A slightly more general version of this is to assume multiple orbitals on an atomic-like site but to only keep the very largest V integrals: namely $i = j$ and $k = l$ and then further assume they are all equal to some number U (regardless of what i or k are). Then

$$\hat{V}_{ee} = \sum_{at} \frac{U}{2} \sum_{\sigma, \sigma'} \sum_{i, k} \hat{c}_{i\sigma}^\dagger \hat{c}_{k\sigma'}^\dagger \hat{c}_{k\sigma'} \hat{c}_{i\sigma}.$$

Rearranging the operators while using their anticommutation properties gives

$$\hat{V}_{ee} = \sum_{at} \frac{U}{2} \sum_{\sigma, \sigma'} \sum_{i, k} \hat{n}_{i\sigma} \hat{n}_{k\sigma'} - \hat{n}_{i\sigma} \delta_{\sigma, \sigma'} \delta_{i, k}.$$

Since $\hat{n}_{i\sigma}$ is the electron counting operator, it must have eigenvalues of only 0 and 1 so it is a projection operator and $\hat{n}_{i\sigma}^2 = \hat{n}_{i\sigma}$. Using this we can rearrange the sums a little to get

$$\hat{V}_{ee} = \sum_{at} \frac{U}{2} \sum_{i\sigma, k\sigma' | i\sigma \neq k\sigma'} \hat{n}_{i\sigma} \hat{n}_{k\sigma'}.$$

We get a sum over products of number operators between pairs of orbitals except if they are the same orbital which is called self-interaction correction: an electron in an orbital does not interact with itself but only with the other electrons in other orbitals. This is some type of a generalized Hubbard model. One can write it in a few equivalent ways using the total number of electron operators $\hat{N} = \sum_{i,\sigma} \hat{n}_{i\sigma}$ as

$$\hat{V}_{ee} = \sum_{at} \frac{U}{2} \hat{N}(\hat{N} - 1) = \sum_{at} \frac{U}{2} (\hat{N}^2 - \sum_{i\sigma} \hat{n}_{i\sigma}^2) = \sum_{at} \frac{U}{2} (\hat{N}^2 - \sum_{i\sigma} \hat{n}_{i\sigma}).$$

This form is particularly nice since it says that when U is the same for all orbital interactions on a site, the repulsion energy only depends on the total number of electrons on that site and not their specific distribution among orbitals. This is also sometimes called an $SU(n)$ Hubbard model where n is the number of orbitals on the site since rotating the orbitals by any unitary transformation among themselves is an invariant of the system. This invariance is very useful in many-body slave-boson or slave-rotor approaches to electron correlation (it greatly reduces the number of “fake” bosons that are added to try to model the correlated behavior).

4 Mean-field Hartree-Fock approach

Beyond these simple cases, it is hard to make any real progress: the four operators in general are a big mess and hard to rearrange into anything sensible. Separately, the many-body problem is impossible to solve anyways due to correlations once we add hopping terms (kinetic and electron-ion). So we now make the key approximation of mean field to get rid of operators and give us numbers. Another way to say the same thing is that when taking expectation of the four operators, we assume we can “split” it into the expectation of all the allowed pairings with appropriate signs for fermion behavior (*i.e.*, Wick’s theorem). This is easiest written in the following way for the expectation:

$$\langle \hat{c}_{i\sigma}^\dagger \hat{c}_{k\sigma'}^\dagger \hat{c}_{l\sigma'} \hat{c}_{j\sigma} \rangle \approx \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} \rangle \langle \hat{c}_{k\sigma'}^\dagger \hat{c}_{l\sigma'} \rangle - \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{l\sigma'} \rangle \langle \hat{c}_{k\sigma'}^\dagger \hat{c}_{j\sigma} \rangle.$$

So the two-body expectation with four operators has turned into a product of two one-body expectations each with two operators. This approximation is a mean-field one. If the wave function over which we are taking the expectation is a single Slater determinant (*i.e.*, we’re doing Hartree-Fock), then this expression is exact and not approximate. To the extent that it is good we are saying that the two particle correlations are not so strong that they invalidate the mean-field view.

We will generally assume that the spin is a good quantum number so that expectations like $\langle \hat{c}_{i\sigma}^\dagger \hat{c}_{k\sigma'} \rangle$ are zero unless $\sigma = \sigma'$. In this case, the two operator expectation is actually the one-body density matrix ρ :

$$\rho_{ji}^\sigma = \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} \rangle$$

which in the position representation is

$$\rho^\sigma(r, r') = \langle \hat{\psi}_\sigma^\dagger(r') \hat{\psi}_\sigma(r) \rangle.$$

In the position representation, the mean-field approximation says

$$\langle \hat{V}_{ee} \rangle = \frac{1}{2} \int dr \int dr' \frac{n(r)n(r') - \sum_\sigma |\rho^\sigma(r, r')|^2}{|r - r'|}$$

where $n(r) = \sum_\sigma \rho^\sigma(r, r)$ is the electron density. This expression is a standard way to write the Coulomb interaction energy for a Hartree-Fock mean-field theory: the interaction energy has been simplified and only depends on products of one-body expectations (density and density matrix). The first term is the Hartree energy and the second the Fock exchange energy as it only applies to parallel spins. If you are a purist, however, notice that the above form is true both if we have a single Slater determinant (*i.e.*, single-particle occupancies are either zero or one) or a statistical mixture of Slater determinants (occupancies between zero and one), so it is slightly more general than the traditional single determinant Hartree-Fock approach.

In the orbital representation in mean-field, the basic equation (2) becomes

$$\langle \hat{V}_{ee} \rangle = \sum_{at} \frac{1}{2} \sum_{\sigma, \sigma'} \sum_{i, j, k, l} (i\sigma k\sigma' | V | j\sigma l\sigma') (\rho_{ji}^\sigma \rho_{lk}^{\sigma'} - \rho_{li}^\sigma \rho_{jk}^\sigma \delta_{\sigma, \sigma'}). \quad (3)$$

We have the Hartree part and then the Fock part which requires parallel spins. This is a very general expression but hard to really digest in this form since the density matrix $\rho_{ij\sigma}$ is not diagonal. However, we can certainly choose the diagonal basis for purposes of analysis, but when we do an actual computation we will use the general non-diagonal basis.

In the diagonal basis of the density matrix, things simplify quite bit. In this nice basis, $\rho_{ji\sigma}$ is zero unless $i = j$. The diagonal values (eigenvalues) are nothing other than the average occupancies $f_{i\sigma} = \langle \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma} \rangle = \langle \hat{n}_{i\sigma} \rangle$. We know $0 \leq f_{i\sigma} \leq 1$ is a constraint (since $\hat{n}_{i\sigma}$ has eigenvalues of only 0 and 1). In this basis, after minor relabeling,

$$\langle \hat{V}_{ee} \rangle = \sum_{at} \frac{1}{2} \sum_{i, j, \sigma} \left(\sum_{\sigma'} (i\sigma j\sigma' | V | i\sigma j\sigma') f_{i\sigma} f_{j\sigma'} \right) - (i\sigma j\sigma | V | j\sigma i\sigma) f_{i\sigma} f_{j\sigma}.$$

Things have simplified quite a bit: there are actually only two overall types of Coulomb matrix elements. In the Hartree term, we get positive integrals between two densities each of which is net +1; in the Fock term, all the integrals should be smaller since they represent neutral-neutral interactions except for $i = j$ where we get again a big interaction. We can call these Coulomb and exchange matrix elements in some general sense.

We notice that the $i = j$ term in the Fock actually exactly cancels the $i = j$ and $\sigma' = \sigma$ term in the Hartree. This is correct and is the self-interaction issue discussed above. The Hartree term itself has this spurious self-interaction, but the Fock zaps it. The fact that

Hartree-Fock removes self-interaction is general and not restricted to our specific choices of basis or approximation on what matrix elements we kept from the start; that Hartree-Fock does this generally is simply a statement of the fact that no matter how bad the Hartree-Fock single Slater determinant wave function might be, it is a valid many-body fermionic wave function that obeys antisymmetry so it never can place two electrons in the same state; or equivalently, we compute the exact expectation of the electron-electron repulsion operator with an approximate wave function, and the actual electron-electron repulsion operator never has the same electron interacting with itself. More practically, it seems that we should remove this term manually so that all the remaining Fock contributions are “small” and we can make more progress.

Removing the self-interaction term gives a key working expression

$$\langle \hat{V}_{ee} \rangle = \frac{1}{2} \sum_{at,\sigma} \left[\sum_{i \neq j} [(i\sigma j\sigma | V | i\sigma j\sigma) - (i\sigma j\sigma | V | j\sigma i\sigma)] f_{i\sigma} f_{j\sigma} + \sum_{i,j} (i\sigma j\bar{\sigma} | V | i\sigma j\bar{\sigma}) f_{i\sigma} f_{j\bar{\sigma}} \right] \quad (4)$$

which is split into parallel and anti-parallel spin terms ($\bar{\sigma}$ is the opposite spin to σ). In principle, one could base a mean-field DFT+U scheme on this formula. One would need to calculate (more likely choose as free parameters) the various Coulombic integrals. Then one has a simple function of the occupancies and one will optimize it (plus other energy terms) to get the ground state. Again, the physical content of this equation is identical to equation (3) but it is just easier to figure out what one is talking about.

5 Dudarev scheme

A most popular approach for DFT+U of Dudarev *et al.*¹ makes a lot of simplifications to get something straightforward. For example, one could argue or just posit that one will be doing a kind of spherical averaging: the sums over i and j are over atomic states, e.g. 5 components of the d states on a transition metal. If one assumes the Coulombic integrals are quite similar irrespective of the choice of orbitals, or one uses an average over the possible choices of orbitals, then one can simplify greatly. So for the Hartree integrals one says $U = (i\sigma j\sigma' | V | i\sigma j\sigma')$ and for the presumably smaller exchange integrals $J = (i\sigma j\sigma | V | j\sigma i\sigma)$ where $i \neq j$. Then we can simplify Eq. (4) to

$$\langle \hat{V}_{ee} \rangle = \sum_{at} \frac{1}{2} \sum_{\sigma} \left((U - J) \sum_{i \neq j} f_{i\sigma} f_{j\sigma} + U \sum_{i,j} f_{i\sigma} f_{j\bar{\sigma}} \right). \quad (5)$$

Here there are only two Coulombic parameters U and J , and we sum over all the orbital pairs with equal weight so we have a high degree of rotational invariance. Same spin electron pairs feel repulsion $U - J$ while opposite spins feel repulsion U . This type of expression doesn't

¹S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, A. P. Sutton, *Physical Review B* **57**, 1505 (1998); DOI:10.1103/PhysRevB.57.1505

assume that all the orbitals are equally occupied and thus will admit not only spin ordering (ferromagnetism or antiferromagnetism) but also orbital ordering (*i.e.* unequal occupancy of atomic-like orbitals on a given site). However, it doesn't know anything about orbital shapes: so the Coulombic repulsion between two electrons in the same spatial state and or different spatial states on a site are assumed identical.

The expression can be massaged a bit by using the total (average) number of electrons on the site for given spin

$$N_\sigma = \sum_i f_{i\sigma}$$

and doing some fiddling

$$\langle \hat{V}_{ee} \rangle = \sum_{at} \frac{1}{2} \sum_\sigma \left((U - J) \left(N_\sigma^2 - \sum_i f_{i\sigma}^2 \right) + U N_\sigma N_{\bar{\sigma}} \right).$$

This is simple enough to write out the spin sum explicitly

$$\langle \hat{V}_{ee} \rangle = \sum_{at} \frac{(U - J)}{2} \left(N_\uparrow^2 + N_\downarrow^2 - \sum_i (f_{i\uparrow}^2 + f_{i\downarrow}^2) \right) + U N_\uparrow N_\downarrow. \quad (6)$$

Most of the terms here only have the total number of electrons of each spin and are thus rather “averaged”. The exceptions are the $f_{i\sigma}^2$ terms which are coming from self-interaction correction (SIC). Thus DFT+U schemes can be thought of as SIC schemes or close cousins, a point made emphasized by Marzari and collaborators. This is quite visible if we contemplate the contribution of this type of energy to the one-electron eigenvalue (band energy). In the most simplified approach, the Janak theorem would lead to $\epsilon_{i\sigma} = \partial E_{tot} / \partial f_{i\sigma}$, so this Coulombic part contributes the following to the eigenvalue

$$\frac{\partial \langle \hat{V}_{ee} \rangle}{\partial f_{i\sigma}} = (U - J)(N_\sigma - f_{i\sigma}) + U N_{\bar{\sigma}}.$$

The interpretation is simple: the electron in question feels a repulsive energy from average number of electrons present on the site (by $U - J$ for like spin and U for unlike spin). The $-f_{i\sigma}$ term corrects this to make sure it doesn't repel itself. Another way to think about it is that filled states $f_{i,\sigma} = 1$ are lowered in energy while empty states are higher (by $U - J$); this is formation of Hubbard bands for broken symmetry (*i.e.*, when the one-electron occupancies are breaking equivalency). Finally, note that $N_\sigma - f_{i\sigma}$ is the number of “other” electrons of same spin: again SIC. But as we increase or decrease $f_{i\sigma}$ while holding all other occupancies fixed, this term doesn't change so it is SIC for any $f_{i\sigma}$ (and not only for some particular value) which is as it should be.

The above derivative also shows that the minimization of the Coulombic part will drive the occupancies to extremes. Consider holding the number of electrons of a given spin N_σ fixed while optimizing over the $f_{i\sigma}$. The derivative expression shows that one lowers the energy more by increasing the larger occupancy states at the expense of the smaller ones. This

will push as many of the occupancies to zero as possible and try to lump the most number of electrons into the smallest number of orbitals. This is nothing really amazing as many mean-field theories work like this, but it is a danger of this type of method: by making U and/or J large enough, one is guaranteed to create symmetry breaking of the occupancies.

A remaining question is how to construct a DFT+U total energy from all of this that one can use in an actual calculation, or in other words, how to deal with double counting. We want to paste this $\langle \hat{V}_{ee} \rangle$ onto the DFT total energy but don't want to double count Hartree energies and possible exchange ones already there. So we want to use the following form of the total energy functional

$$E_{tot} = E_{DFT} + \langle \hat{V}_{ee} \rangle - E_{dc}$$

where E_{dc} is some double-counting energy. There are probably lots of ways to wave one's hands to get something reasonable. One way is to assert that the DFT total energy has big problems for partial occupancies of orbitals since it is not SIC but is probably doing an okay job for the total energy for integer fillings. So one picks the E_{dc} that cancels the $\langle \hat{V}_{ee} \rangle$ under such a condition. If the occupancies of each orbital are either 1 or 0 then $f_{i\sigma}^2 = f_{i\sigma}$ and we would get

$$\begin{aligned} \langle \hat{V}_{ee} \rangle \Big|_{\text{integer fillings}} &= \sum_{at} \frac{(U - J)}{2} \left[N_{\uparrow}(N_{\uparrow} - 1) + N_{\downarrow}(N_{\downarrow} - 1) \right] + UN_{\uparrow}N_{\downarrow} \\ &= \sum_{at} \frac{U}{2} N(N - 1) - \frac{J}{2} \left(N_{\uparrow}(N_{\uparrow} - 1) + N_{\downarrow}(N_{\downarrow} - 1) \right). \end{aligned}$$

Choosing the above expression to be E_{dc} gives the spherically averaged Dudarev DFT+U approach. As there are many common terms that cancel, it simplifies to

$$E_{tot} = E_{DFT} + \frac{(U - J)}{2} \sum_{at} \sum_{i\sigma} (f_{i\sigma} - f_{i\sigma}^2) = E_{DFT} + \frac{(U - J)}{2} \sum_{at,\sigma} \text{trace}(\rho - \rho^2) \quad (7)$$

where we have noticed in the second form that we can easily go back to the non-diagonal basis for a general expression. The eigenvalue corrections in a simplified view looks like

$$\epsilon_{i\sigma} = \frac{\partial E_{tot}}{\partial f_{i\sigma}} = \epsilon_{i\sigma}^{DFT} + (U - J) \left(\frac{1}{2} - f_{i\sigma} \right).$$

So it shifts the LDA bands by a constant as well as splitting occupied and unoccupied states by $U - J$.

The only technical criticism of what is written above is the cavalier method of doing the eigenvalue correction. Technically, the Janak theorem says that for a Bloch state $\psi_{nk\sigma}$ with occupancy $f_{nk\sigma}$ we have

$$\epsilon_{nk\sigma} = \frac{\partial E_{tot}}{\partial f_{nk\sigma}} = \frac{\partial E_{DFT}}{\partial f_{nk\sigma}} + \frac{\partial (E_U - E_{dc})}{\partial f_{nk\sigma}}$$

The density matrix is given by

$$\rho_{ij\sigma} = \langle i\sigma | \hat{\rho}_{\sigma} | j\sigma \rangle = \sum_{nk} f_{nk\sigma} \langle i\sigma | \psi_{nk\sigma} \rangle \langle \psi_{nk\sigma} | j\sigma \rangle$$

so the more proper way to describe the eigenvalue correction is to go back to the expression of Eq. (7) and to take derivatives and end up with (in the diagonal eigenbasis)

$$\begin{aligned}\epsilon_{nk\sigma} &= \epsilon_{nk\sigma}^{DFT} + \sum_{at,i\sigma} (U - J) \left(\frac{1}{2} - f_{i\sigma} \right) \frac{\partial f_{i\sigma}}{\partial f_{nk}} \\ &= \epsilon_{nk\sigma}^{DFT} + \sum_{at,i\sigma} (U - J) \left(\frac{1}{2} - f_{i\sigma} \right) |\langle i\sigma | \psi_{nk\sigma} \rangle|^2.\end{aligned}$$

The naïve eigenvalue correction needs an extra factor saying how large the projection of the Bloch band $nk\sigma$ is onto the localized states.

6 Liechtenstein scheme

The more sophisticated approach is to not make such a dramatic assumption of equality of all types of Coulombic matrix elements as in the Dudarev scheme. In general, this takes us back to an expression like Eq. (3), which is fine but quite general. To make some progress, we need to make some assumptions. We will be assuming that the localized orbitals are atomic-like of the form of a radial function (with no spin dependence) times a spherical harmonic $R_{nl}(r)Y_{lm}(\theta, \phi)$, and using the particularly simple form of the bare Coulomb operator, we can handle the angular terms explicitly and exactly. Then for the radial parts we will have to wave hands and parameterize. To make things feasible, we will be assuming the atomic orbitals come from a single shell of an atom on each site. This results in a so called rotationally invariant DFT+U form of Liechtenstein *et al.*²

We take the general expression of Eq. (3) and reduce to the case that the localized orbitals have the same atomic nl indices so that only the magnetic quantum number m and the spin are summed over. We have in general form and in diagonal basis

$$\begin{aligned}\langle \hat{V}_{ee} \rangle &= \frac{1}{2} \sum_{\substack{at,\sigma,\sigma' \\ m,m',m'',m'''}} (m\sigma m''\sigma' | V | m'\sigma m'''\sigma') (\rho_{m'm}^\sigma \rho_{m''m'''}^{\sigma'} - \rho_{m''m}^{\sigma''} \rho_{m'\sigma}^{\sigma'} \delta_{\sigma\sigma'}) \\ &= \frac{1}{2} \sum_{\substack{at,\sigma,\sigma' \\ i,j}} (i\sigma j\sigma' | V | i\sigma j\sigma') f_{i\sigma} f_{j\sigma'} - (i\sigma j\sigma | V | j\sigma i\sigma) f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'}\end{aligned}\quad (8)$$

We now start to massage the Coulomb matrix elements based on our assumptions. The first thing we need is to express the Coulomb operator in terms of something compatible with radial and angular functions, so we do a multipole expansion and use the spherical harmonic addition theorem:

$$\frac{1}{|\vec{r}_1 - \vec{r}_2|} = \sum_{k=0}^{\infty} P_k(\hat{r}_1 \cdot \hat{r}_2) \frac{\min(r_1, r_2)^k}{\max(r_1, r_2)^{k+1}} = \sum_{k=0}^{\infty} \frac{4\pi}{2k+1} \sum_{q=-k}^k Y_{kq}(\hat{r}_1) Y_{kq}(\hat{r}_2)^* \frac{\min(r_1, r_2)^k}{\max(r_1, r_2)^{k+1}}$$

²A. I. Liechtenstein, V. I. Anisimov, J. Zaanen, *Physical Review B* **52**, R5467 (1995).

In fact, almost all the derivations below follow for a more general form of any interaction $V(\vec{r}_1, \vec{r}_2)$ which has rotational symmetry and interchange symmetry so that it can only depend on the lengths r_1 and r_2 and the angle between them parametrized by $\cos \theta_{12} = \hat{r}_1 \cdot \hat{r}_2$: we expand in the variable $\cos \theta_{12}$ using a Legendre polynomial expansion (instead of more traditional Taylor)

$$V(r_1, r_2, \cos \theta_{12}) = \sum_{k=0}^{\infty} f_k(r_1, r_2) P_k(\cos \theta_{12})$$

and here we have some series of functions $f_k(r_1, r_2)$ that are symmetric under interchange of r_1 and r_2 . For the bare Coulomb case $f_k(r_1, r_2) = \min(r_1, r_2)^k / \max(r_1, r_2)^{k+1}$.

A matrix element like $(m\sigma m''\sigma' | V | m'\sigma m'''\sigma')$ is then given by

$$\begin{aligned} (m\sigma m''\sigma' | V | m'\sigma m'''\sigma') &= \int d^3\vec{r}_1 \int d^3\vec{r}_2 \frac{R_{nl}(r_1)^2 R_{nl}(r_2)^2 Y_{lm}^*(\hat{r}_1) Y_{lm'}(\hat{r}_1) Y_{lm''}^*(\hat{r}_2) Y_{lm'''}(\hat{r}_2)}{|\vec{r}_1 - \vec{r}_2|} \\ &= \sum_{k=0}^{\infty} \sum_{q=-k}^k \frac{4\pi}{2k+1} \langle Y_{lm} | Y_{kq} Y_{lm'} \rangle \langle Y_{lm'''} | Y_{kq} Y_{lm''} \rangle^* F^k \end{aligned}$$

where the radial integrals are defined from Slater's time as

$$F^k = \int_0^{\infty} dr_1 \int_0^{\infty} dr_2 R_{nl}(r_1)^2 R_{nl}(r_2)^2 f_k(r_1, r_2).$$

The angular parts enforce a simple angular momentum conservation rule that $m - m' = m'' - m'''$. A stronger condition comes from the angular addition rules which say that $k > 2l$ always will give zero, so the sum actually only runs over $k = 0$ to $k = 2l$. Even more examination of the angular momentum rules for this particular type of product show that actually the allowed values of k skip in two $k = 0, 2, 4, \dots, 2l$. So for a typical application with $l = 2$ d orbitals, $k = 0, 2, 4$ are the only three values appearing.

The expression can be further compactified using Slater's angular integrals c^k

$$(m\sigma m''\sigma' | V | m'\sigma m'''\sigma') = \delta_{m-m', m''-m'''} \sum_{k=0}^{2l} c^k(lm, lm') c^k(lm''', lm'') F^k \quad (9)$$

where

$$c^k(lm, l'm') = \sqrt{\frac{4\pi}{2k+1}} \int d\Omega Y_{lm}(\theta, \phi)^* Y_{l'm'}(\theta, \phi) Y_{k, m-m'}(\theta, \phi).$$

These angular integrals can be tabulated, so we now have a working expression for all the Coulomb matrix elements. The question is instead what to use for the F^k .

To get to these, let's connect the F^k to simpler objects we could get from some first principles calculation where we would vary occupancies of the orbitals and see how the energy or eigenvalues change. However, it is not going to be easy to do this in the general form where the density matrices are not diagonal since the number of V matrix elements are very large

and also there are selection rules that are awkward to enforce. In the diagonal basis, however, we have more sensible Coulomb and exchange matrix elements which we can make sense of. Therefore, let us formalize the notion of the diagonal basis for the Hermitian density matrices ρ^σ . The eigenvalues $f_{i\sigma}$ and orthonormal eigenvectors V_{mi}^σ are related to ρ_σ in the usual way:

$$\rho_{mm'}^\sigma = \sum_i V_{mi}^\sigma f_{i\sigma} (V^\sigma)_{im'}^\dagger$$

and plugging this into the expression for the matrix element of V gives

$$\begin{aligned} \langle \hat{V}_{ee} \rangle &= \frac{1}{2} \sum_{\substack{at,\sigma,\sigma' \\ m,m',m'',m'''}} (m\sigma m''\sigma' | V | m'\sigma m'''\sigma') (\rho_{m'm}^\sigma \rho_{m''m'''}^{\sigma'} - \rho_{m''m'''}^\sigma \rho_{m'm}^{\sigma'} \delta_{\sigma\sigma'}) \\ &= \frac{1}{2} \sum_{\substack{at,\sigma,\sigma',k \\ m,m',m'',m'''}} F^k \delta_{m-m',m''-m'''} c^k(lm, lm') c^k(lm''', lm'') (\rho_{m'm}^\sigma \rho_{m''m'''}^{\sigma'} - \rho_{m''m'''}^\sigma \rho_{m'm}^{\sigma'} \delta_{\sigma\sigma'}) \\ &= \frac{1}{2} \sum_{at,\sigma,\sigma',k} F^k \sum_{i,j} f_{i\sigma} f_{j\sigma'} \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} c^k(lm, lm') c^k(lm''', lm'') \times \\ &\quad \left[V_{m'i}^\sigma (V^\sigma)_{im}^\dagger V_{m''j}^{\sigma'} (V^{\sigma'})_{j m''}^\dagger - V_{m''i}^\sigma (V^\sigma)_{im}^\dagger V_{m'j}^{\sigma'} (V^{\sigma'})_{j m''}^\dagger \delta_{\sigma\sigma'} \right] \\ &= \frac{1}{2} \sum_{at,\sigma,\sigma',k} F^k \sum_{i,j} f_{i\sigma} f_{j\sigma'} \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} \times \\ &\quad \left[V_{mi}^{\sigma*} c^k(lm, lm') V_{m'i}^\sigma V_{m''j}^{\sigma'} c^k(lm''', lm'') V_{m''j}^{\sigma'*} \right. \\ &\quad \left. - V_{mi}^{\sigma*} c^k(lm, lm') V_{m'j}^\sigma V_{m''i}^{\sigma'} c^k(lm''', lm'') V_{m''j}^{\sigma'*} \delta_{\sigma\sigma'} \right] \end{aligned}$$

If the Kronecker delta enforcing angular momentum conservation were not there, we could freely sum over the m, m', m'', m''' to get nice simple matrix expressions of i and j , but this is not meant to be. However, for $k = 0$ we can make progress because $c^0(lm, lm') = \delta_{mm'}$. We peel off the $k = 0$ term and use the unitary nature of the V_{mi}^σ matrices to get

$$\begin{aligned} \langle \hat{V}_{ee} \rangle &= \frac{1}{2} \sum_{at,\sigma,\sigma',i,j} f_{i\sigma} f_{j\sigma'} \left\{ F^0 (1 - \delta_{ij} \delta_{\sigma\sigma'}) + \sum_{k=2}^{2l} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} \times \right. \\ &\quad \left. \left[(V^\sigma)_{im}^\dagger c^k(lm, lm') V_{m'i}^\sigma (V^{\sigma'})_{j m''}^\dagger c^k(lm''', lm'') V_{m''j}^{\sigma'} \right. \right. \\ &\quad \left. \left. - (V^\sigma)_{im}^\dagger c^k(lm, lm') V_{m'j}^\sigma (V^{\sigma'})_{j m''}^\dagger c^k(lm''', lm'') V_{m''i}^{\sigma'} \delta_{\sigma\sigma'} \right] \right\} \\ &= \frac{F^0}{2} \sum_{at} \left(N^2 - \sum_{i\sigma} f_{i\sigma}^2 \right) + \frac{1}{2} \sum_{at,\sigma,\sigma',i,j} f_{i\sigma} f_{j\sigma'} \sum_{k=2}^{2l} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} \times \\ &\quad \left[(V^\sigma)_{im}^\dagger c^k(lm, lm') V_{m'i}^\sigma (V^{\sigma'})_{j m''}^\dagger c^k(lm''', lm'') V_{m''j}^{\sigma'} \right. \\ &\quad \left. - (V^\sigma)_{im}^\dagger c^k(lm, lm') V_{m'j}^\sigma (V^{\sigma'})_{j m''}^\dagger c^k(lm''', lm'') V_{m''i}^{\sigma'} \delta_{\sigma\sigma'} \right] \end{aligned}$$

The leading part is quite simple and so we collect all the remaining complicated stuff into

specialized Coulomb and exchange matrices defined by the expressions

$$C_{ij}^{\sigma\sigma'} = \sum_{k=2}^{2l} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} (V^\sigma)^\dagger_{im} c^k(lm, lm') V_{m'i}^\sigma (V^{\sigma'})^\dagger_{jm'''} c^k(lm''', lm'') V_{m''j}^{\sigma'}$$

and

$$X_{ij}^\sigma = \sum_{k=2}^{2l} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} (V^\sigma)^\dagger_{im} c^k(lm, lm') V_{m'j}^\sigma (V^\sigma)^\dagger_{im'''} c^k(lm''', lm'') V_{m''j}^\sigma$$

so the energy expression looks nice and neat

$$\langle \hat{V}_{ee} \rangle = \frac{1}{2} \sum_{at} \left[F^0 \left(N^2 - \sum_{i\sigma} f_{i\sigma}^2 \right) + \sum_{\sigma,\sigma',i,j} C_{ij}^{\sigma\sigma'} f_{i\sigma} f_{j\sigma'} - X_{ij}^\sigma f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \right] \quad (10)$$

Having done all of this, we can now try to figure out what to do with the F^k . For example, let's sum over all the Coulomb matrix elements:

$$\sum_{i,j} C_{ij}^{\sigma,\sigma'} = \sum_{k=2}^{2l} \sum_{i,j} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} (V^\sigma)^\dagger_{im} c^k(lm, lm') V_{m'i}^\sigma (V^{\sigma'})^\dagger_{jm'''} c^k(lm''', lm'') V_{m''j}^{\sigma'}$$

Using the unitary nature of the V_{mi}^σ matrices, $m = m'$ and $m'' = m'''$ becomes enforced which automatically satisfies the Kronecker delta, so this simplifies to

$$\sum_{i,j} C_{ij}^{\sigma,\sigma'} = \sum_{k=2}^{2l} \sum_{m,m''} F^k c^k(lm, lm) c^k(lm'', lm'') = 0$$

where we used the sum rule obeyed by the c^k integrals

$$\sum_m c^k(lm, lm) = (2l+1) \delta_{k,0}.$$

This means that the the average of Coulomb matrix elements as we have defined starting at $k = 2$ is zero so the only average that does survive is the $k = 0$ term F^0 we peeled off. We call this average Coulomb element the number U .

For the exchange element averages, we have

$$\sum_{i,j} X_{ij}^\sigma = \sum_{i,j,k=2}^{2l} F^k \sum_{m,m',m'',m'''} \delta_{m-m',m''-m'''} (V^\sigma)^\dagger_{im} c^k(lm, lm') V_{m'j}^\sigma (V^\sigma)^\dagger_{im'''} c^k(lm''', lm'') V_{m''j}^\sigma$$

and the sums over i, j enforce $m = m'''$ and $m' = m''$ which also satisfies the Kronecker delta. We end up with

$$\sum_{i,j} X_{ij}^\sigma = \sum_{k=2, m, m''}^{2l} F^k c^k(lm, lm'')^2 = \sum_{k=2}^{2l} F^k (2l+1) c^k(l0, l0)$$

where we used the sum rule

$$\sum_{m,m'} c^k(lm, l'm')^2 = \sqrt{(2l+1)(2l'+1)} c^k(l0, l'0).$$

There are $(2l+1)$ choices of magnetic quantum number m (or i or j). Note that when we sum over all pairs, we include “diagonal” pairs of same spin which actually don’t contribute to the energy since they cancel off corresponding terms from the Coulomb part. So really, we should be averaging only $i \neq j$ and $(2l+1)2l$ pairs instead of $(2l+1)^2$ pairs and should drop the leading F^0 term. We will call this averaged exchange integral J and it equals

$$J = \sum_{k=2}^{2l} \frac{c^k(l0, l0)}{2l} F^k.$$

For d electrons with $l = 2$, $c^2(20, 20) = c^4(20, 20) = 2/7$ and $2l = 4$ so $J = (F^2 + F^4)/14$. To close the loop for d electrons, Liechtenstein *et al.* say that $F^4/F^2 \sim 0.625$ for 3d elements and then using U , J and $0.625=5/8$ we can determine F^0 , F^2 and F^4 and we have a complete practical expression. For reference, $J = 13F^2/112$ and so $F^2 = (112/13)J$ while $F^4 = (70/13)J$.

The next thing to deal with is the double-counting term. If we simply insist that for integer occupancies $f_{i\sigma} \in \{0, 1\}$ the double counting term should completely cancel the above energy term and leave us only with the DFT energy, then this is actually too weak: the expression of Eq. (8) with integer occupancies still depends on exactly which orbitals are being occupied since we have the matrix elements of the V between detailed angular momentum functions that have different shapes, etc. To simplify, either we can have a given number of electrons integer distributed over various $i\sigma$ and then we average over all possible such distributions, or equivalently we simply average the V matrix elements appearing in the expression over the i indices. But this second averaging ends up giving us U for the Coulomb interactions and J for exchange so we are just back to the Dudarev scheme with integer occupancies! So the double-counting is the same as the Dudarev.

So, the Liechtenstein scheme then has us write

$$E_{tot} = E_{DFT} + E_U - E_{dc}$$

with

$$\begin{aligned} E_U &= \frac{1}{2} \sum_{\substack{at, \sigma, \sigma' \\ m, m', m'', m'''}} (m\sigma m''\sigma' | V | m'\sigma m'''\sigma') (\rho_{m'm}^\sigma \rho_{m''m'''}^{\sigma'} - \rho_{m''m}^\sigma \rho_{m'm''}^{\sigma'} \delta_{\sigma\sigma'}) \\ &= \frac{1}{2} \sum_{\substack{at, \sigma, \sigma' \\ i, j}} (i\sigma j\sigma' | V | i\sigma j\sigma') f_{i\sigma} f_{j\sigma'} - (i\sigma j\sigma | V | j\sigma i\sigma) f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \\ &= \frac{1}{2} \sum_{at} \left[U \left(N^2 - \sum_{i\sigma} f_{i\sigma}^2 \right) + \sum_{\sigma, \sigma', i, j} C_{ij}^{\sigma\sigma'} f_{i\sigma} f_{j\sigma'} - X_{ij}^\sigma f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \right] \end{aligned}$$

with the V matrix elements given by Eq. (9) and

$$E_{dc} = \sum_{at} \frac{(U - J)}{2} \left[N_{\uparrow}(N_{\uparrow} - 1) + N_{\downarrow}(N_{\downarrow} - 1) \right] + U N_{\uparrow} N_{\downarrow}$$

To make progress into the meaning of this scheme and its relation to Dudarev, we will need to substitute in the angular momentum expansions into the energy and see what we get. The first order of business is to pull out a J term from the exchange element X_{ij}^{σ} to pair off with the U term just like in the double-counting term. We do this by writing

$$X_{ij}^{\sigma} = \Delta X_{ij}^{\sigma} + J(1 - \delta_{ij})$$

(this equality defines ΔX^{σ}). The sum over the J term is easy and we get

$$E_U = \frac{1}{2} \sum_{at} \left[U N^2 - J(N_{\uparrow}^2 + N_{\downarrow}^2) - (U - J) \sum_{i\sigma} f_{i\sigma}^2 + \sum_{\sigma, \sigma', i, j} C_{ij}^{\sigma\sigma'} f_{i\sigma} f_{j\sigma'} - \Delta X_{ij}^{\sigma} f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \right]$$

We are now ready to subtract off the double counting term to get

$$E_U - E_{dc} = \frac{1}{2} \sum_{at} \left[(U - J) \sum_{i\sigma} (f_{i\sigma} - f_{i\sigma}^2) + \sum_{\sigma, \sigma', i, j} C_{ij}^{\sigma\sigma'} f_{i\sigma} f_{j\sigma'} - \Delta X_{ij}^{\sigma} f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \right]$$

Amazingly, the first term is the Dudarev term so the rest is the new contribution of the Liechtenstein scheme. For the total energy we then have DFT plus Dudarev plus correction

$$E_{tot} = E_{DFT} + \frac{(U - J)}{2} \sum_{at, i, \sigma} (f_{i\sigma} - f_{i\sigma}^2) + \frac{1}{2} \sum_{at, \sigma, \sigma', i, j} C_{ij}^{\sigma\sigma'} f_{i\sigma} f_{j\sigma'} - \Delta X_{ij}^{\sigma} f_{i\sigma} f_{j\sigma} \delta_{\sigma\sigma'} \quad (11)$$

The correction term should be smaller than the Dudarev term. The Coulomb elements C^{σ} average to zero so we get nothing if the $f_{i\sigma}$ have no i dependence. In addition, we have explicitly separated off the average exchange term so ΔX^{σ} also averages to zero. In addition, the corrections involve F^k for $k \geq 2$ so this means for d electrons that they are all some type of exchange like terms because F^2 and F^4 are both constants times J . Basically, the correction term accounts for actual angular character of the interactions: the orbitals have actual lobes with shapes and point in different directions, and this is missing if we simply replace by orbital-independent interactions U and J . This shape dependence is something that is inside of J (and not U as defined) because exchange J is about different orbitals of same spin being integrated against each other in the matrix element of V .

Let's also compute the contribution to the eigenvalues in Liechtenstein DFT+U:

$$\frac{\partial E_U - E_{dc}}{\partial f_{i\sigma}} = (U - J) \left(\frac{1}{2} - f_{i\sigma} \right) + \sum_{\sigma', j} C_{ij}^{\sigma\sigma'} f_{j\sigma'} - \Delta X_{ij}^{\sigma} f_{j\sigma} \delta_{\sigma\sigma'}$$

To get a physical idea of what this can do, we need to specialize and work out some specific cases. For d electrons, both C and ΔX are strictly proportional to J so we can write this as a vector matrix equation

$$\nabla_{f_{\sigma}} (E_U - E_{dc}) = (U - J) \left(\frac{1}{2} - f_{\sigma} \right) + J [A^{\sigma} f_{\sigma} + B^{\sigma} f_{\bar{\sigma}}]$$

where f_σ is a column vector of the values $f_{m\sigma}$ and the same spin and we have defined same spin and opposite spin matrices A^σ and B^σ as

$$A^\sigma = C^{\sigma,\sigma} - \Delta X^\sigma \quad , \quad B^\sigma = C^{\sigma,\bar{\sigma}}$$

A first simple case is to assume that the density matrices ρ^σ are diagonal in the original Y_{lm} basis so $f_{i\sigma} = f_{m\sigma}$ and $V^\sigma = I$. Then one gets numerically that

$$A = \begin{pmatrix} 0.00 & -0.52 & -0.52 & 0.17 & 0.86 \\ -0.52 & 0.00 & 0.52 & -0.17 & 0.17 \\ -0.52 & 0.52 & 0.00 & 0.52 & -0.52 \\ 0.17 & -0.17 & 0.52 & 0.00 & -0.52 \\ 0.86 & 0.17 & -0.52 & -0.52 & 0.00 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 0.72 & -0.40 & -0.63 & -0.40 & 0.72 \\ -0.40 & 0.37 & 0.06 & 0.37 & -0.40 \\ -0.63 & 0.06 & 1.14 & 0.06 & -0.63 \\ -0.40 & 0.37 & 0.06 & 0.37 & -0.40 \\ 0.72 & -0.40 & -0.63 & -0.40 & 0.72 \end{pmatrix}$$

So, for example, occupation of the Y_{20} orbital causes the same spin $Y_{2,\pm 2}$ orbitals to drop in energy and the $Y_{2,\pm 1}$ to rise in energy and pushes the opposite spin Y_{20} orbital high up in energy, doesn't do much to opposite spin $Y_{2,\pm 1}$ and lowers $Y_{2,\pm 2}$ of opposite spin. (The difference between the two matrices is just ΔX here.)

However, this is not really very realistic since very few materials will have diagonal ρ^σ in the Y_{lm} basis. Instead, high symmetry materials with d shell states will have t_{2g} (i.e., xy, xz, yz) and e_g (i.e., $3z^2 - r^2, x^2 - y^2$) orbitals as eigenvectors of ρ^σ . Here are the relations of orthonormal such orbitals to Y_{lm} :

$$\begin{aligned} |3z^2 - r^2\rangle &= |Y_{20}\rangle \\ |x^2 - y^2\rangle &= \frac{1}{\sqrt{2}} (|Y_{2,2}\rangle + |Y_{2,-2}\rangle) \\ |xy\rangle &= \frac{-i}{\sqrt{2}} (|Y_{2,2}\rangle - |Y_{2,-2}\rangle) \\ |yz\rangle &= \frac{i}{\sqrt{2}} (|Y_{2,1}\rangle - |Y_{2,-1}\rangle) \\ |xz\rangle &= \frac{-1}{\sqrt{2}} (|Y_{2,1}\rangle + |Y_{2,-1}\rangle) \end{aligned}$$

So the matrix V^σ is

$$V^\sigma = \begin{pmatrix} 0 & 1/\sqrt{2} & i/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & -i/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -i/\sqrt{2} & 0 & 0 \end{pmatrix}$$

We can then recompute A and B in this basis to get

$$A = \left(\begin{array}{c|ccccc} & 3z^2 - r^2 & x^2 - y^2 & xy & yz & xz \\ \hline 3z^2 - r^2 & 0 & -0.52 & -0.52 & 0.52 & 0.52 \\ x^2 - y^2 & -0.52 & 0 & 0.86 & -0.17 & -0.17 \\ xy & -0.52 & 0.86 & 0 & -0.17 & -0.17 \\ yz & 0.52 & -0.17 & -0.17 & 0 & -0.17 \\ xz & 0.52 & -0.17 & -0.17 & -0.17 & 0 \end{array} \right)$$

and

$$B = \left(\begin{array}{c|ccccc} & 3z^2 - r^2 & x^2 - y^2 & xy & yz & xz \\ \hline 3z^2 - r^2 & 1.14 & -0.63 & -0.63 & 0.06 & 0.06 \\ x^2 - y^2 & -0.63 & 1.14 & 0.29 & -0.40 & -0.40 \\ xy & -0.63 & 0.29 & 1.14 & -0.40 & -0.40 \\ yz & 0.06 & -0.40 & -0.40 & 1.14 & -0.40 \\ xz & 0.06 & -0.40 & -0.40 & -0.40 & 1.14 \end{array} \right)$$

A number of interesting facts to illustrate what we get here

- Occupying any orbital causes the opposite spin same orbital to be pushed up in energy by $1.14J$. This effect is completely missing from Dudarev where the only change to the orbital energy depends on its own occupancy via $(U - J)(1/2 - f_{i\sigma})$. The reason is that an orbital overlaps best with itself spatially so we get the largest Coulombic repulsion from the Hartree/Coulomb term — it is larger than the average repulsion U over all possible orbital pairings.
- $3z^2 - r^2$ and $x^2 - y^2$ are not at all similar in terms of interaction with the other t_{2g} orbitals. This is because $x^2 - y^2$ and xy are actually just 45° rotates of each other while $3z^2 - r^2$ is something quite different and has more overlap with xz and yz and its interaction with xy and $x^2 - y^2$ are exactly the same.
- Let's say we have something like Mn^{4+} where there are three t_{2g} electrons of up spin and no other electrons. Then our spin up occupancy is $(0, 0, 1, 1, 1)$ and we get same-spin eigenvalue contributions of $(0.52, 0.52, -0.34, -0.34, -0.34)J$ and opposite spin eigenvalue contributions of $(-0.52, -0.52, 0.34, 0.34, 0.34)J$. So for same spin t_{2g} occupied states we get additional stabilization beyond Dudarev of $-0.34J$ and unoccupied same spin e_g are destabilized by $0.52J$. But opposite spin is exactly the other way: t_{2g} pushed up by $0.34J$ and e_g pulled down by $0.52J$.
- If we have Mn^{3+} in configuration $e_g^1 t_{2g}^3$ and spread the e_g electron evenly between $3z^2 - r^2$ and $x^2 - y^2$ (i.e., no orbital polarization) so occupancy $(0.5, 0.5, 1, 1, 1)$ then the energy shifts are actually the same as the Mn^{4+} case but just reduced in magnitude: $(0.26 \times 2, -0.17 \times 3)$ for same spin and $(-0.26 \times 2, 0.17 \times 3)$ for opposite spin.
- Continuing on with Mn^{3+} but with full orbital polarization: it depends which we fill. If we fill $3z^2 - r^2$ so occupancy is $(1, 0, 1, 1, 1)$, then same spin eigenvalue shift is $(0.52, 0.00, -0.86, 0.17, 0.17)J$ and opposite spin is $(0.63, -1.14, -0.29, 0.40, 0.40)J$.

But for occupancy $(0, 1, 1, 1, 1)$ we get $(0.00, 0.52, 0.52, -0.52, -0.52)J$ and $(-1.14, 0.63, 0.63, -0.06, -0.06)J$. So the action on the t_{2g} manifold does depend on which of the e_g we occupy.

- A closed-shell configuration like t_{2g}^6 does nothing: the up and down spin contributions exactly cancel. So all we have is Dudarev.
- A configuration like $e_g^1 t_{2g}^6$ where the e_g electron is spread evenly in spin and in e_g orbitals also gives zero. This is something like Ni^{3+} in a non-magnetic completely dull electronic configuration.
- A configuration like $e_g^1 t_{2g}^6$ where the e_g electron is purely spin up but split 1/2-1/2 between the two up spin e_g orbitals gives for spin up $(-0.25 \times 2, 0.17 \times 3)J$ and for spin down $(0.26 \times 2, -0.17 \times 3)J$.