Getting an effective Hamiltonian on a sublattice; effective hoppings

Sohrab Ismail-Beigi

June 13, 2023

The idea is to start with a microscopic Hamiltonian describing the "actual" system and to find a reduced description on as subspace. For example, starting with a set of sites with various energies and hopping elements, we might want an effective description between a subset of them — which will make sense if they different energies — and to find the effective on-site and hopping elements for this subsystem. An example of the use would be for a metal oxide where one is interested in the bands and hoppings between the transition metal sites when they are mediated via the oxygen atoms in between (the direct hoppings between metal sites being negligible). This type of modeling is important for understanding superexchange and double-exchange (see below).

The main tool will be the non-interacting Green's function for one-particle systems. In the process we'll see how elimination of degrees of freedom gives nonlocality and frequency dependence to the self-energy or effective Hamiltonian — with no electron-electron *interactions* needed at all. Thus these properties of the self-energy are coming from eliminating degrees of freedom per se.

First, the plan of what we want to do. We have the Hamiltonian $\mathcal{H}$ on a full basis of state labeled by Greek letters so it has matrix elements $\mathcal{H}_{\alpha,\beta}$. We choose a subset (subspace) of these states which we call the d states labeled by d, d' (or other Latin letters) that we are interested in describing. The remaining states are the p states we want to eliminate. The block structure of $\mathcal{H}$ is

$$\mathcal{H} = \begin{pmatrix} H_d & T \\ T^\dagger & H_p \end{pmatrix} \tag{1}$$

The first block of entries are the d ones. What we want is an effective Hamiltonian H^{eff} on this d subspace which only has matrix elements in the subspace $H^{eff}_{d,d'}$ that gives us the same physics for the bands "belonging" to that subspace — of course, this idea only makes sense perturbatively in the hopping between the sites. We do this by matching matrix elements of the Green's functions.

The full Green's function $\mathcal{G}(\omega)$ is

$$\mathcal{G}(\omega) = (\omega I - \mathcal{H})^{-1} \tag{2}$$

and the matrix elements $\mathcal{G}(\omega)_{\alpha,\beta}$ tell us the propagation amplitude $\beta \to \alpha$ at frequency ω. We then extract the matrix elements of $\mathcal{G}$ only between the subspace sites

$$G(\omega)_{d,d'} \equiv \mathcal{G}(\omega)_{d,d'} \tag{3}$$

so in terms of blocks we have

$$\mathcal{G}(\omega) = \begin{pmatrix} G(\omega) & X(\omega) \\ Y(\omega) & Z(\omega) \end{pmatrix}. \tag{4}$$

We then *define* the effective Hamiltonian $H^{eff}(\omega)$ to give this subspace Green's function

$$G(\omega) \equiv (\omega I - H^{eff}(\omega))^{-1}. \tag{5}$$

In this way, $H^{eff}(\omega)$ is designed to give the correct propagation behavior in the subspace.

To do this in practice, we use perturbation theory. We know this obey's the Dyson equation

$$\mathcal{G}(\omega) = \mathcal{G}^0(\omega) + \mathcal{G}^0(\omega)\mathcal{T}\mathcal{G}(\omega)$$

where we have an unperturbed $\mathcal{G}^0(\omega) = (\omega I - \mathcal{H}^0)^{-1}$ where $\mathcal{H}^0$ is the Hamiltonian with no hoppings between the two subspaces

$$\mathcal{H}^0 = \begin{pmatrix} H_d & 0 \\ 0^\dagger & H_p \end{pmatrix} \tag{6}$$

and $\mathcal{T}$ is the strictly off-diagonal matrix of hopping elements

$$\mathcal{T} = \begin{pmatrix} 0 & T \\ T^\dagger & 0 \end{pmatrix} \tag{7}$$

and $\mathcal{H} = \mathcal{H}^0 + \mathcal{T}$. Thus $\mathcal{G}^0(\omega)$ is also block diagonal

$$\mathcal{G}^0(\omega) = \begin{pmatrix} G_d^0(\omega) & 0 \\ 0 & G_p^0(\omega) \end{pmatrix}. \tag{8}$$

The Dyson series for $\mathcal{G}$ is

$$\mathcal{G}(\omega) = \mathcal{G}^0(\omega) + \mathcal{G}^0(\omega)\mathcal{T}\mathcal{G}^0(\omega) + \mathcal{G}^0(\omega)\mathcal{T}\mathcal{G}^0(\omega)\mathcal{T}\mathcal{G}^0(\omega) + \cdots \tag{9}$$

We compute G based on this.

Extracting G as the d sunblock of $\mathcal{G}$ and the block diagonal nature of $\mathcal{G}^0$, we find

$$G(\omega) = G_d^0(\omega) + G_d^0(\omega)\mathcal{T}_{dd}G_d^0(\omega) + G_d^0(\omega)\mathcal{T}_{dp}G_p^0(\omega)\mathcal{T}_{pd}G_d^0(\omega) + \ldots \tag{10}$$

Now, $\mathcal{T}$ is off diagonal so a block diagonal entry for it like $\mathcal{T}_{dd} = 0$. In this way, we find that all terms in the series with odd numbers of $\mathcal{T}$ vanish. We cleanup by using the definition of the off diagonal block T to see that non-zero terms in the Dyson series look like

$$G(\omega) = G_d^0(\omega) + G_d^0(\omega) T G_p^0(\omega) T^\dagger G_d^0(\omega) + G_d^0(\omega) T G_p^0(\omega) T^\dagger G_d^0(\omega) T G_p^0(\omega) T^\dagger G_d^0(\omega) + \dots \quad (11)$$

A short examination of this series shows that the d-block self-energy is $\Sigma(\omega) = T G_p^0(\omega) T^\dagger$ since

$$G(\omega) = G_d^0(\omega) + G_d^0(\omega)\Sigma(\omega) G_d^0(\omega) + G_d^0(\omega)\Sigma(\omega) G_d^0(\omega)\Sigma(\omega) G_d^0(\omega) + \dots \quad (12)$$

This we end up with the expression

$$G(\omega) = \left([G^0(\omega)]^{-1} - \Sigma(\omega)\right)^{-1} = \left(\omega I - H_d^0 - \Sigma(\omega)\right)^{-1} \quad (13)$$

so our effective Hamiltonian is

$$H^{eff}(\omega) = H_d^0 + \Sigma(\omega) = H_d^0 + T G_p^0(\omega) T^\dagger \,. \quad (14)$$

This is the main result we've been looking for. Written explicitly as matrix elements

$$\Sigma(\omega)_{d,d'} = \sum_{p,p'} T_{d,p} \left(\omega I - H_p\right)^{-1}_{p,p'} T^*_{d',p'} \quad (15)$$

We have a second-order virtual hopping through non-subspace p states. At first, it may seem confusing that there is an explicit ω dependence here. But there must be since the subspace Green's function is designed to reproduce the exact behavior including its full frequency dependence. Within the perturbative limit, however, we will be setting ω at or close to the unperturbed d energies and then we will have actual numbers to compute.

We now work out some examples below.

First consider a one dimensional lattice of sites where even sites have energy zero, odd sites have energy $-\epsilon$ and there is a hopping element t between even and odd sites. This could represent identical transition metal atoms separated by oxygen atoms in an oxide. The microscopic Hamiltonian is thus

$$H = \begin{pmatrix} t & 0 & t & & & \\ & t & -\epsilon & t & & \\ & & t & 0 & t & \\ & & & t & -\epsilon & t \end{pmatrix} .$$

This problem can actually be solved exactly by hand due to the resulting 2×2 Bloch problem which gives the bands

$$E_\pm(k) \;=\; -\frac{\epsilon}{2} \pm \sqrt{\left(\frac{\epsilon}{2}\right)^2 + 4t^2 \cos(ka/2)^2} \;\approx\; 2t\cos(ka/2) \text{ for } t/\epsilon \gg 1$$

$$
\begin{aligned}
&= \frac{\epsilon}{2} \cdot \left(-1 \pm \sqrt{1 + \left(\frac{4t}{\epsilon}\right)^2 \cos(ka/2)^2} \right) \\
&\approx \left\{ \begin{array}{c} 0 \\ -\epsilon \end{array} \right\} \pm \left(\frac{4t^2}{\epsilon}\right) \cos(ka/2)^2 = \left\{ \begin{array}{c} 0 \\ -\epsilon \end{array} \right\} \pm \frac{2t^2}{\epsilon} \left[1 + \cos(ka)\right] \text{ for } t/\epsilon \ll 1
\end{aligned}
$$

where a is the lattice constant which contains two sites. Our objective is to describe one of these bands approximately by creating a Hamiltonian that only exists on one set of site, say the even high energy ones (zero energy even sites). This only really will work well for $t/\epsilon \ll 1$.

We now apply out above formalism to the high energy sites with $\epsilon_d = 0$. Since all the high energy sites have the same zero energy, the diagonal block $H_d = 0$ so in this special case $H^{eff}(\omega) = \Sigma(\omega)$. We get for $H^{eff}(\omega)$ written as a row vector (one row of the matrix which is then periodically replicated)

$$
H^{eff}(\omega) = \left(O(t^6), O(t^4), \frac{t^2}{\omega + \epsilon}, \frac{2t^2}{\omega + \epsilon}, \frac{t^2}{\omega + \epsilon}, O(t^4), O(t^6) \right)
$$

This type of Hamiltonian is trivial to diagonalize with plane waves and gives a single band

$$
E^{eff}(k, \omega) = \frac{2t^2}{\omega + \epsilon} \left(1 + \cos(ka)\right) + O(t^4)
$$

which, as expected, agrees with the exact result for the upper band for $t/\epsilon \ll 1$ and for $\omega = 0$. Notice how $\Sigma(\omega)$ is nonlocal: first of all it is in principle dense and not sparse (although with weaker and weaker off diagonal elements for farther hoppings), and second, it has intra-site matrix elements that don't exist in the full Hamiltonian. The nonlocality comes directly from the eliminated states since it was through multiple hoppings through the low-energy sites that the high-energy sites are connected. Furthermore, it is explicitly ω dependent: the dependence is centered around $\omega = -\epsilon$ which are the (unperturbed) energies of the states we eliminated.

At $\omega = 0$, we can recognize the various parts of $H^{eff}(\omega)$ from perturbation theory. The diagonal shift $2t^2/\epsilon$ is the just the second order level repulsion of a high-energy site from its neighboring low-energy sites. And t^2/ϵ is the effective hopping (through virtual processes one could say) between high-energy sites mediated through low-energy sites. The rule for these hoppings is to square the microscopic hopping element and divide by the site energy differences. Notice that t and ϵ could be individually very large, but that is not the relevant for the effective picture: the effective picture removes some of these large numbers to give us a low-energy description with different parameters.

The next more complex example is a four site unit cell where the site energies are (starting with site zero) 0, $-\epsilon$, 0, and v (repeating). Only nearest neighbor hoppings t exist again. This

could represent a chain of transition metal atoms separated by oxygen atoms where either the transition metals are different elemental species or where they are the same species but have an exchange splitting due to antiferromagnetic ordering (or anything giving an alternating energy shift). This is clearly a generalization of the above case which corresponds to $v = 0$: the $v = 0$ case can describe ferromagnetic ordering where all the metal sites are identical (and degenerate).

The microscopic Hamiltonian is

$$H = \begin{pmatrix} t & 0 & t & & & \\ & t & -\epsilon & t & & \\ & & t & v & t & \\ & & & t & -\epsilon & t \end{pmatrix} .$$

We want to compute the effective Hamiltonian on the transition metal sublattice (sites with site energies 0 or v). We just turn the crank on the above formalism. The effective Hamiltonian now has two inequivalent rows for the $\epsilon = 0$ and $\epsilon = v$ sites, and they are to lowest order given by

$$H^{eff}(\omega) = \begin{pmatrix} \tilde{t} & \Delta & \tilde{t} & & & \\ & \tilde{t} & v+\Delta & \tilde{t} & & \\ & & \tilde{t} & \Delta & \tilde{t} & \\ & & & \tilde{t} & v+\Delta & \tilde{t} \end{pmatrix}$$

where

$$\Delta = \frac{2t^2}{\omega + \epsilon} \quad , \quad \tilde{t} = \frac{t^2}{\omega + \epsilon}$$

These energy shifts and hoppings are actually unchanged from the simpler example above.

When we solve this effective problem so get two bands with a gap instead of a single band. (It is actually isomorphic to the simpler problem we first solved but with orbitals renamed.) Solving this effective problem gives

$$\begin{aligned} E_{\pm}^{eff}(k) &= \Delta + \frac{v}{2} \pm \sqrt{\left(\frac{v}{2}\right)^2 + 4\tilde{t}^2 \cos(ka/2)^2} \approx \Delta + 2\tilde{t}\cos(ka) \text{ for } \tilde{t}/v \gg 1 \\ &\approx + \left\{ \begin{array}{c} \Delta \\ \Delta + v \end{array} \right\} \pm \left(\frac{4\tilde{t}^2}{v}\right) \cos(ka/2)^2 = \left\{ \begin{array}{c} \Delta \\ \Delta + v \end{array} \right\} \pm \frac{2\tilde{t}^2}{v} \left[1 + \cos(ka)\right] \text{ for } \tilde{t}/v \ll 1 . \end{aligned}$$

This simple model gives two bands with a gap between them (at $ka = \pi$ with band gap being v). But what is interesting to note here is the band widths of the two bands and how they depend on $\tilde{t}$ and v. When v is small, the two bands basically overlap: we have essentially two identical sites so we get two bands since we chose a unit cell twice as large as needed; the band width is given by the effective hopping $\tilde{t}$. However, for large v, the band widths are much reduced: they are now $\sim \tilde{t}^2/v \ll \tilde{t}$: the large splitting v prohibits efficient hopping between the two different (effective) sites.

We end with some comments on the relation of these results to Wannier functions. Returning to the simpler case with $v = 0$, for $t/\epsilon \ll 1$ we have to narrow band centered around 0 and $-\epsilon$ and we are focusing on the upper band. In the effective description for the upper band, we have one site per unit cell and they are connected by hopping $\tilde{t}$ and if one wanted to get a Wannier function for the upper band in this effective language, it would be trivial and just equal to the sites themselves (absolute localization on the site). But this is misleading since we are talking about the effective states: we are just saying we have a set of identical effective states on the upper energy sites and so obviously the Wannier function will be the sites themselves.

More interesting is to relate the effective description to the microscopic one to understand what is going on. Namely, we ask what the true (microscopic) Wannier function corresponding to the upper bands looks like. The reason this helps us understand things is that if we let $W(2n)$ be the Wannier function of the upper band on site $2n$, we first know that the other Wannier states are just translations of this. But in addition, this Wannier function must give the effective parameters:

$$\langle W(0)|H|W(0)\rangle = \Delta \quad , \quad \langle W(0)|H|W(2)\rangle = \tilde{t}\,.$$

So the spatial extent of the Wannier function must be responsible for giving $\tilde{t}$: the actual high energy (even) sites are not microscopically connected. One can compute the Wannier function perturbativley by writing an expansion of the Wannier function in the microscopic sites (and using the fact that the Wannier function will be symmetric here): going to second nearest neighbors, we write

$$W(0) = (\cdots, 0, \gamma, \beta, \alpha, \beta, \gamma, 0, \cdots)$$

where the center site with α is site zero, those with β are low energy sites ± 1, etc. Orthonormality gives

$$1 = \langle W(0)|W(0)\rangle = \alpha^2 + 2\beta^2 + 2\gamma^2 \quad , \quad 0 = \langle W(0)|W(2)\rangle = 2\alpha\gamma + \beta^2\,.$$

In a perturbative expansion, β should go to zero as $t/\epsilon \to 0$ while $\alpha \to 1$. So from both these relations we realize that

$$\alpha = 1 + \bar{\alpha} t^2/\epsilon^2 \quad , \quad \beta = \bar{\beta} t/\epsilon \quad , \quad \gamma = \bar{\gamma} t^2/\epsilon^2$$

to leading order where the barred variables are constants of order unity that do not depend on t/ϵ. Plugging in the expansion for $W(0)$ into the two relations that give us Δ and $\tilde{t}$, after some algebra, gives $\bar{\alpha} = -1, \bar{\beta} = 1, \bar{\gamma} = -1/2$ so the Wannier function looks like

$$|W(0)\rangle = (\cdots, 0,\ -t^2/(2\epsilon^2)\ ,\ t/\epsilon\ ,\ 1 - t^2/\epsilon^2\ ,\ t/\epsilon\ ,\ -t^2/(2\epsilon^2)\ , 0, \cdots)\,.$$

Clearly, the Wannier function is centered at the site with large amplitude on site and rapidly decaying amplitudes on successive sites. (In general, it will be exponentially localized.) It is the t/ϵ coefficient of the odd site (the eliminated ones for the effective description) that lead to the couplings of neighboring even sites and make their Wannier functions overlap to have

a transfer element $\tilde{t}$. (Note: there are better ways of getting the Wannier function here since we can solve the 2×2 problem to get the Bloch states explicitly and just write the Fourier integral for the Wannier function and do things perturbatively there. See the end of these notes for the math.)

For the more complex example with $v \neq 0$, we can make some general observations. When $\tilde{t}/v \gg 1$, we are just in the case we analyzed above: sites with energy 0 and v are basically the same and they have very similar Wannier functions that are given above with hopping $\tilde{t}$. However, when $\tilde{t}/v \ll 1$ and we get the two very narrow bands, we have different Wannier functions for the 0 and v sites that are very localized on their respective sites; the hoppings between same energy sites is $\tilde{t}^2/v$ which is small (but then again it didn't exist in the effective model in the first place).

Here we do some math to compute the Wannier function for the two-site-per-cell-problem from its definition perturbatively. We will choose Block states with phases defined via

$$\psi_k = (\cdots, ae^{-ika}, b^{-ika/2}, a, be^{ika/2}, ae^{ika}, be^{3ika/2}, \cdots)$$

where even sites have a and odd sites b and the central site with energy zero corresponds to the a entry. These Bloch states have he 2×2 Hamiltonian

$$H_k = \begin{pmatrix} 0 & 2t\cos(ka/2) \\ 2t\cos(ka/2) & -\epsilon \end{pmatrix} .$$

The eigenvalues of this are given at the beginning of these notes as $E_\pm(k)$. The eigenvectors follow from writing

$$\begin{aligned} H_k &= -(\epsilon/2)I + \begin{pmatrix} \epsilon/2 & 2t\cos(ka/2) \\ 2t\cos(ka/2) & -\epsilon/2 \end{pmatrix} \\ &= -(\epsilon/2)I + \sqrt{(\epsilon/2)^2 + (2t\cos(ka/2))^2} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} . \end{aligned}$$

Perturbatively, the angle θ is

$$\theta \approx \sin(\theta) = \frac{2t\cos(ka/2)}{\sqrt{(\epsilon/2)^2 + (2t\cos(ka/2))^2}} \approx \frac{4t\cos(ka/2)}{\epsilon} .$$

The upper 2×2 eigenvector (for the upper we care about) is

$$(\cos(\theta/2), \sin(\theta/2)) \approx (1 - \theta^2/8, \theta/2) = (1 - 2t^2/\epsilon^2 \cos(ka/2)^2, (2t/\epsilon)\cos(ka/2)) .$$

The Wannier function for site n is

$$W_n = \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi/a} e^{-ikan} \psi_k$$

so the Wannier function for site 0 is

$$W_0 = \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi/a} (\cdots, \cos(\theta/2)e^{-ika}, \sin(\theta/2)e^{-ika/2}, \cos(\theta/2), \sin(\theta/2)e^{ika/2} \cos(\theta/2)e^{ika}, \cdots) \,.$$

One can plug in the above expansions for $\cos(\theta/2)$ and $\sin(\theta/2)$ and do the integrals using double angle formula and get the same results as before. This approach is more direct but assumes we can have explicit expressions for the Bloch states.