Messiah on Perturbation Theory PhL 1.8.09

symbol grab: $\Sigma_{k \in \mathcal{L}_\perp}$ $\mathcal{H}^{(m)} \oplus \mathcal{H}^\perp$ $\mathcal{E}_a{}^0$ $(\mathcal{E}_a{}^0)^\perp$ $\in$ $\subset$

(1.8.09) I have been confused and unhappy about degenerate perturbation theory (stationary state) since I started working on it on Jan 2, so my unhappiness has continued for about 6 days. I read Schiff and Saxon and could not find happiness, so now I will see if Messiah can bail me out. This will of course require going into a whole new notation with all the overhead that involves, but that is the price of consultation.

[It is now Jan 27 and I have just finished my Round Two notes, and things are still not *crystal* clear, but I was at least able to answer every question I asked. So as of today I have invested 25 days of my life on this topic (but also did Stakgold Chapter 2). My notes here and in Round Two don't deal with the Kato resolvent method, just the Messiah stuff in earlier sections, as carried on by me. At least I was able to produce expressions for those mysterious $B'_{ij,1}$ coefficients from my personal exposition of SSPT. That happened for the first time today.]

[It is now Feb 8 and I just finished up all Kato notes, 5 documents!]

Overview: *All* of Chapter 16 is on SSPT, pp 685-721. I read only the "theory sections" of this chapter and skipped over "applications" like Stark Effect, at least for now. The degenerate problem is fully solved (in a sense) in the tour de force Topic III Kato section, which I will need to review in some meta notes. This was about the hardest stuff to understand that I have ever read, and is based on a 1949 paper of Kato with reference on page 712. That work was updated in 1958 by Messiah's friend C. Bloch and must have come out just as this book was being wrapped up. There are of course **newer treatments** of the complicated degeneracy problem. I found a saved a 1981 paper by Silverstone and Moats that seems pretty good and references Kato and Bloch. Newer papers still exist, with an eye to practical calculation in higher orders.

MESSIAH VOLUME II CHAPTER XVI (16): STATIONARY PERTURBATIONS. (P 685) 2

1. General introduction to Part Four (methods of approximation)..... 2

TOPIC I. THE NON-DEGENERATE CASE (P 686)..... 2

2. Expansion in Powers of the Perturbation (p 686) 2

Digression on the operators $Q_\circ$ and Q_{0a}. 3

3. First-order perturbations (688). 5

6. Higher order corrections. (694). 6

TOPIC II. THE DEGENERATE CASE 7

8. Elementary Degenerate Theory (698). 7

Rederive facts about Q_0 and Q_{0a} 10

Clarification of my Theorem B confusion. 17

Clarification of my B'_{ij} confusion. 19

Messiah Volume II Chapter XVI (16): Stationary Perturbations. (p 685)

1. General introduction to Part Four (methods of approximation). You approximate in the world of bound states and in the world of collisions, that latter being messier. In the bound state world, Messiah has already done the WKB method in Chapter 6. In this chapter we focus on SSPT.

Topic I. The Non-Degenerate Case (p 686)

2. Expansion in Powers of the Perturbation (p 686). He is going to refer to H' as V, but the smallness parameter is still called λ. The notation E_i^0 are the H_0 energy eigenvalues. States are $|E_i^0\ \alpha\rangle$ where the α is needed only if you have degeneracy. So in this section we have just $|E_i^0\rangle$ and he favors using "a" for the index, so we have $|E_a^0\rangle$. This "a" is like the "m" of Schiff, we are picking out some particular state to think about. As a shorthand notation, he will use $|0\rangle = |E_a^0\rangle$ since we know we are talking all the time about "a" and we want to reduce notational clutter. The perturbed version of |0> will be called |ψ> which again agrees with Schiff's choice of notation for this state. Messiah will be "ket oriented" which is fine by me.

Now we come to page 687 equation 16.4 (Roman numerals are horrible, Messiah should not have done that to my way of thinking). It says $\langle 0|\psi\rangle = 1$ which is Schiff's statement that $\langle u_m | \psi'^{(m)}\rangle = 1$. Oddly, Messiah says nothing about this and I have a little pencil "WHY?" from some earlier time I looked at this chapter. This is explained as Plan D in my Schiff Chapter 8 notes, and also as Theorem B of my Degenerate Perturbation Theory doc. I am quite happy with it.

Next we come to the power series shown in 16.5 and 16.6, so let's do more notational linkage:

Schiff	Messiah		
m	a	// selected state (or space of N degenerate states)	
$\mathcal{H}^{(m)}$	$\mathcal{E}_a^0$	// space containing N unperturbed state(s) of interest	
	$\mathcal{E}_a$	// space containing the N perturbed states	
H'	V	// perturbing Hamiltonian	
$W_m = \Sigma_{n=0}\ \lambda^n\ W_{m,n}$	$E = E_a^0 + \Sigma_{n=1}\ \lambda^n\ \varepsilon_n$		// energy expansion
W_m	E_m		
$W_{m,0}$	E_a^0		
$W_{m,n}$	ε_n	n = 1,2,3...	
$\psi_m = \Sigma_{n=0}\ \lambda^n\ \psi_{m,n}$	$\|\psi\rangle = \Sigma_{n=0}\ \lambda^n\ \|n\rangle$		// state expansion
$\psi_{m,n}$	$\|n\rangle$	n = 0,1,2....	// the tag "a" is implicit
$\langle u_m \| \psi_{m,n}\rangle = 0$ n>0	$\langle 0\|n\rangle = 0$ n>0		// special normalization
|r> (PhL)	$\|E_a^0\alpha\rangle$		// starting base states inside $\mathcal{E}_a^0$

We now come to Messiah's version of Schiff 31.4, but Messiah moves Schiff's RHS terms to the LHS. I notice Messiah's attempt to give a separate number for equations in a group using a little super.

Next, Schiff's 31.6 which said $\langle u_m | \psi_{m,n}\rangle = 0$ for n>0 appears as **Messiah 16.8** page 688. Of course it is very nice in Messiah notation : $\langle 0|n\rangle = 0$ for n>0.

Next, Schiff's 31.7 which said $W_{m,s} = \langle u_m|H'|\psi_{m,s-1}\rangle$ becomes $\varepsilon_s = \langle 0|V|s\text{-}1\rangle$ but Messiah uses n instead of s, very reasonable, and thus we have **Messiah 16.9**. In Messiah, this result comes from "closing with <0| "onto his 31.4 version.

Where is equation 688.A coming from? When he writes $|E^0_\alpha\rangle$, he just means some unperturbed eigenstate *other than* the |0> one which we are focusing on. If we close his 31.4 version with this, and then divide through by the energy difference that then appears, **we get A page 688**, fine.

Messiah now defines Q_0 to be the identity operator in our entire Hilbert Space minus the contribution of our unperturbed state of interest, just fine by me. [It is the "perp space projector" to this one state.]

Digression on the operators $Q_\circ$ and $Q_{\circ a}$.

Equation **16.10** is "less obvious" to me. Here is my derivation. First start with

$$Q_0 \equiv 1 - |0\rangle\langle 0| = \Sigma_{k\neq a} |E_k^0\rangle\langle E_k^0|$$

This thing is an operator, and on the right above I write it as a sum over all the *unperturbed eigenstates* (which I label by E_k^0) *except* the state a, which is the idea that we are subtracting out $|0\rangle\langle 0|$. Notice that I cannot express these other states by the notation |k> because that notation is already used to mean $|\psi_{m,k}\rangle = |\psi_{a,k}\rangle$.

This thing Q_0 is a "projection operator". Notice this fact:

$$Q_0^2 = [1 - |0\rangle\langle 0|][1 - |0\rangle\langle 0|] = 1 - |0\rangle\langle 0 - |0\rangle\langle 0| + |0\rangle\langle 0|0\rangle\langle 0|$$

$$= 1 - |0\rangle\langle 0| - |0\rangle\langle 0| + |0\rangle\langle 0| = 1 - |0\rangle\langle 0| = Q_0$$

as appropriate for any projection operator. Notice that this implies that

$$Q_0^{-1} = Q_0$$

What exactly does this operator Q_0 do? If you have some ket $|\varphi\rangle$ which is an arbitrary linear combination of all unperturbed kets in the full Hilbert Space, this projects out all pieces *except* the component along the unperturbed state Messiah calls |0> which is his state with E_a and would be Schiff's state u_m. It is sort of a deletion operator in this sense.

Now obviously if you apply this to some "other unperturbed state" like u_k somewhere, $k \neq a$, then $Q_0|u_k\rangle = |u_k\rangle$. But we are going to apply this more significantly to a perturbed state which we should regard as a rotated ket in the space of unperturbed basis kets. We have for example,

$$|\psi\rangle = \Sigma_{n=0} \lambda^n |n\rangle$$

which I might enhance just a bit and write as

$$|\psi_a\rangle = \Sigma_{n=0} \lambda^n |n; a\rangle$$

so

$Q_0|\psi_a\rangle = \Sigma_{n=0}\, \lambda^n\, Q_0\, |n; a\rangle = \Sigma_{n=1}\, \lambda^n\, |n; a\rangle = |\psi_a\rangle - |0; a\rangle$

so in this case Q_0 removes the entire $|0;a\rangle$ part of $|\psi_a\rangle$. [It is a perp space projector.]

In particular, in Messiah notation we have

$Q_0|\, n\rangle = |\, n\rangle \qquad\qquad n > 0$

In my own notation where I had

$\psi'_{mi} = \{\ u_{mi} + \Sigma_{j\neq i} B'_{ij}\ u_{mj}\ \} \oplus \Sigma_{k\neq m}\, B'_{ik}\, u_k$

$Q_0\psi'_{mi} = \{\ \Sigma_{j\neq i} B'_{ij}\ u_{mj}\ \} \oplus \Sigma_{k\neq m}\, B'_{ik}\, u_k$

which is to say, it projects out ALL of the full perturbed state other than the unperturbed part. So this operator gives you access to the part of the perturbed state that is caused by the perturbation. It is an interesting operator.

Now consider (where all three objects are operators since H_0 is an operator).

$Q_0\, (E_a{}^0 1 - H_0)^{-1}\, Q_o = \Sigma_{k\neq a}\, |E_k{}^0\rangle\langle E_k{}^0|(E_a{}^0 1 - H_0)^{-1}\, \Sigma_{k'\neq a}\, |E_{k'}{}^0\rangle\langle E_{k'}{}^0|$

$= \Sigma_{k\neq a}\, \Sigma_{k'\neq a}\, |E_k{}^0\rangle\langle E_k{}^0|(E_a{}^0 1 - H_0)^{-1}|E_{k'}{}^0\rangle\langle E_{k'}{}^0|$

$= \Sigma_{k\neq a}\, \Sigma_{k'\neq a}\, |E_k{}^0\rangle\langle E_k{}^0|(E_a{}^0 1 - E_k{}^0 1)^{-1}|E_{k'}{}^0\rangle\langle E_{k'}{}^0|$ // H_0 acting to the left

$= \Sigma_{k\neq a}\, \Sigma_{k'\neq a}\, |E_k{}^0\rangle(E_a{}^0 - E_k{}^0)^{-1}\langle E_k{}^0|E_{k'}{}^0\rangle\langle E_{k'}{}^0|$ // the 1's are implicit

$= \Sigma_{k\neq a}\, \Sigma_{k'\neq a}\, |E_k{}^0\rangle(E_a{}^0 - E_k{}^0)^{-1}\delta_{k,k'}\langle E_{k'}{}^0|$ // all unperturbed states ortho

$= \Sigma_{k\neq a}\, |E_k{}^0\rangle(E_a{}^0 - E_k{}^0)^{-1}\langle E_k{}^0|$ **// which is 16.10**

so this verifies the equation shown in 16.10. His summation notation is a little strange, I will stick with mine, they are both OK.

He then makes this very strange notational definition:

$Q_0/a \equiv Q_0\, (E_a{}^0 - H_0)^{-1}\, Q_o = \Sigma_{k\neq a}\, |E_k{}^0\rangle(E_a{}^0 - E_k{}^0)^{-1}\langle E_k{}^0| \equiv Q_{0a}$

where I think Q_0/a is all one symbol, a certain operator. **I will call it Q_{0a}** .

So why is this a useful thing to have? Here is a **mini-theorem**:

Mini Theorem #1:

$$Q_{0a}(H_0 - E_a^0) = Q_0$$

Proof: Where we use the fact that $Q_0^{-1} = Q_0 = Q_o^2$:

$$Q_{0a} = Q_0 (E_a^0 - H_0)^{-1} Q_o$$
$$Q_{0a}^{-1} = Q_0 (E_a^0 - H_0) Q_o$$

so

$$Q_{0a} (E_a^0 - H_0) = \{ Q_0 (E_a^0 - H_0)^{-1} Q_o \} (E_a^0 - H_0)$$

$$= \{ Q_0 (E_a^0 - H_0)^{-1} Q_o \} \{ Q_o (E_a^0 - H_0) Q_o \} Q_o = Q_{0a} Q_{0a}^{-1} Q_o = Q_o \quad \text{QED.}$$

Now go back to 16.7^n which is this ($n > 0$) where $|0\rangle$ is assumed "in the perp space" : (a $|k\rangle$ ie)

$$(H_0 - E_a^0) |n\rangle + (V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle - \varepsilon_n|0\rangle = 0$$

$$(H_0 - E_a^0) |n\rangle - \varepsilon_n|0\rangle + [(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle] = 0$$

$$(E_a^0 - H_0) |n\rangle + \varepsilon_n|0\rangle = [(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle]$$

If we apply Q_{0a} to both sides of the last line above, the $\varepsilon_n|0\rangle$ is killed off by the right-side Q_0 in Q_{0a}, and then the remaining operator $(E_a^0 - H_0)$ is sort of "inverted" by Q_{0a} as shown above since $Q_0|n\rangle = |n\rangle$ and we get

$$|n\rangle = Q_{0a} [(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle] \qquad \text{// } \textbf{which is 16.11}$$

which is 16.11. This fancy operator Q_{0a} creates the next order of perturbed state when it acts on the square bracket ket which involves only lower order kets. So this is a pretty nice compact operator form for doing this operation.

Comment: The operator Q_{0a} "exposes" the perp-space state $|n\rangle$ on the LHS of 16.7 by inverting the operator $(H_0 - E_a^0)$ normally sitting there. Your thus get a formal equation for $|n\rangle$ in terms of lower state corrections, which is the key idea to this iterative perturbation theory.

3. First-order perturbations (688). He writes the first order energy correction 16.12 as expected. Then 16.13 shows the energy through order λ, and notice he has H in there and not just V = H' , so $\langle 0|H|0\rangle$ generates the first two terms $E_a^0 + \lambda \langle 0|V|0\rangle$.

Now we get to practice with his little Q operator stuff. Let's try it on the lowest order iteration right now:

$$|1\rangle = Q_{0a} [(V-\varepsilon_1)|0\rangle = Q_{0a} V |0\rangle \qquad \textbf{// which is 16.14}$$

so

$$\langle E_\alpha^0| 1\rangle = \langle E_\alpha^0| Q_{0a} V |0\rangle$$

$= \langle E_\alpha^0| \Sigma_{k\neq a} |E_k^0\rangle (E_a^0 - E_k^0)^{-1} \langle E_k^0| V |0\rangle$

$= (E_a^0 - E_\alpha^0)^{-1} \langle E_\alpha^0| V |0\rangle$ // **which is 16.16**

4. Ground state of Helium atom. skip for now

5. Coulomb energy of atomic nuclei. skip for now.

6. Higher order corrections. (694). So now we resume the theory flow. Equation 16.26 is the usual formula I think of as Schiff 31.7, but it is Messiah 16.9. This formula is always simple in any order.

$|n\rangle = Q_{0a} [(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle]$

$|2\rangle = Q_{0a} (V-\varepsilon_1)|1\rangle$

so this one has just a single term as |1> had. Now we know that

$|1\rangle = Q_{0a} V |0\rangle$

so we have

$|2\rangle = Q_{0a} (V-\varepsilon_1) Q_{0a} V |0\rangle$ with $\varepsilon_1 = \langle 0| V |0\rangle$

$= Q_{0a} V Q_{0a} V |0\rangle - \langle 0| V |0\rangle Q_{0a}^2 V |0\rangle$

$= [Q_{0a} V Q_{0a} - \langle 0| V |0\rangle Q_{0a}^2] V |0\rangle$

For some reason I do not comprehend, he writes this second term differently, using this fact

$Q_{0a}^2 \langle 0| V |0\rangle V |0\rangle = Q_{0a}^2 V |0\rangle\langle 0| V |0\rangle = Q_{0a}^2 V |0\rangle\langle 0| V |0\rangle$

to get

$|2\rangle = [Q_{0a} V Q_{0a} - Q_{0a}^2 V |0\rangle\langle 0|] V |0\rangle$ // **Messiah p 194 (27)**

and for the second time he uses the odd notation Q_0/a^2 to mean Q_{0a}^2.

He suggests that you can keep going with this idea to get complex formulas. Let's just try one more:

$|n\rangle = Q_{0a} [(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle]$

$|3\rangle = Q_{0a} [(V-\varepsilon_1)|2\rangle - \varepsilon_2|1\rangle]$ where $\varepsilon_2 = \langle 0|V|1\rangle = \langle 0|V Q_{0a} V |0\rangle$

so get

$|3\rangle = Q_{0a} [(V- \langle 0| V |0\rangle)\{ [Q_{0a} V Q_{0a} - Q_{0a}^2 V |0\rangle\langle 0|] V |0\rangle\} - \langle 0|V Q_{0a} V |0\rangle Q_{0a} V |0\rangle]$

$= Q_{0a} V Q_{0a} V Q_{0a} V |0\rangle$ + lots of other terms

It looks a little like scattering theory and I think we are just picking up more messy energy denominators.

Now start over and suppose ε_1 was zero. Then we have

$|2\rangle = Q_{0a} V Q_{0a} V |0\rangle = (Q_{0a} V)^2 |0\rangle$

$\varepsilon_2 = \langle 0|V|1\rangle = \langle 0|V (Q_{0a} V)^1 |0\rangle$

And now suppose ε_2 is also 0. Then

$|3\rangle = Q_{0a} [(V-\varepsilon_1)|2\rangle - \varepsilon_2|1\rangle] = Q_{0a} V|2\rangle = Q_{0a} V Q_{0a} V Q_{0a} V |0\rangle = (Q_{0a} V)^3 |0\rangle$

$\varepsilon_3 = \langle 0|V|2\rangle = \langle 0|V(Q_{0a} V)^2 |0\rangle$

So it seems pretty reasonable that this pattern can continue, and in general

$|n\rangle = (Q_{0a} V)^n |0\rangle$ which means of course $|n\rangle = (Q_{0a} V) |n-1\rangle$ // **Messiah p 694 (30)**

$\varepsilon_n = \langle 0|V(Q_{0a} V)^{n-1} |0\rangle$ both provided $\varepsilon_1 = \varepsilon_2 = ... \varepsilon_{n-1} = 0$ // **Messiah p 694 (29)**

In some sense, $(Q_{0a} V)$ is a "raising operator" taking you from the n-1st perturbation contribution to the nth perturbation contribution. This really looks like some kind of iterative Foldy-Wouthuysen type deal going on here. The point is that you always want to carry the thing until you get some non-zero energy correction, and in this case that first correction has a somewhat simple form in that those "lots of other terms" won't be present.

I am now on page 695 and I don't follow his convergence comment where I have just put a pencil ? in the margin. And I am skipping his bound comments which finish out this section. Come back some day if interested. I want to get on to the degenerate stuff which is why I am here in the first place.

7. Stark Effect for a Rigid Rotor (695). This is a rather long 2-page example of computing the second order energy correction for a certain problem. As 16.31 shows, this really requires an infinite summation so one wonders how that is done, or perhaps how approximated. I would guess you ignore distant weakly coupled states somehow. I am skipping this section for the reason just noted.

Topic II. The Degenerate Case

OK, I had to do a lot of work and notational switching just to get to this point, see all of the above notes in this doc.

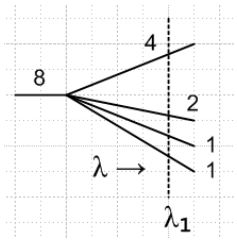
8. Elementary Degenerate Theory (698). We have some pretty solid setting up going on here. Here are some preliminary comparisons of notation. First, here is the degenerate subspace where we have a bunch of states of H_0 having the same eigenvalue

$\mathcal{H}^{(m)}$ of dimension N $\mathcal{E}_a{}^0$ of dimension g_a m = a

Happily he uses the word "subspace".

P acting on a general vector in $\mathcal{H} = \mathcal{H}^{(m)} \oplus \mathcal{H}^{\perp}$ projects out that portion in $\mathcal{H}^{(m)}$

But now he is going to handle another complication. Suppose the perturbation H' = V does not fully split all the states in $\mathcal{H}^{(m)}$. So as λ comes on, the degenerate states wonder off in groups. Here is a picture suggesting this idea, where we are plotting let us say the first order energy correction ε_1 versus λ.



and I am sure this situation is very common. When λ > 0, you can regard the group of 4 states as a subspace. So when λ >0 at some value like λ_1 shown, we have

$\mathcal{H}^{(m)} \rightarrow \mathcal{H}^{(4)} \oplus \mathcal{H}^{(2)} \oplus \mathcal{H}^{(1)} \oplus \mathcal{H}^{(1)}$

where I am using a sloppy notation just to make the point. Messiah is going to assume that there are n such groups that fan out as λ is turned on, and that they have dimensions:

$g_1, g_2 \ldots g_n$ = degeneracies of the groups

$g_1 + g_2 + g_3 + \ldots + g_n = g_a$ Subspaces are called $\mathcal{E}_1$, $\mathcal{E}_2$, ... $\mathcal{E}_n$

E_1 E_2 E_3 E_n are the full energy eigenvalues for these smaller subspaces

I like how we are going.

In this context, he says to let ψ be one of those groups on the right with energy E (= some E_i).

Equation 16.35: P_0 keeps only states that are in $\mathcal{H}^{(m)} = \mathcal{E}_a{}^0$. Certainly an unperturbed state |0> of energy $E_a{}^0$ survives this projector, QED.

Equation 16.36. How does P_0 act on this equation 17.7^1 ? Why does the first term vanish? (and then the other term gives 16.36 trivially)

he seems to be claiming that: $P_0 (H_0 - E_a{}^0)|1> = 0$

How can we write this projector using kets?

$P_0 = \Sigma_k |E_k^0\rangle\langle E_k^0|$? sum over all states in $\mathcal{E}_a^0$

I think this is right. It is just "this part of unity". Then we would have

$P_0 (H_0 - E_a^0)|1\rangle = \Sigma_k |E_k^0\rangle\langle E_k^0|(H_0 - E_a^0)|1\rangle = \Sigma_k |E_k^0\rangle\langle E_k^0|(E_k^0 - E_a^0)|1\rangle = 0$

So we can add this

Mini Theorem #2: $P_0 (H_0 - E_a^0) = 0$ proven above

How should I interpret this theorem? Consider:

[I am temporarily reverting to notation $|n\rangle$ meaning $|u_n\rangle$, an unperturbed eigenstate, not the nth correction to some particular state.]

$H_0 = \Sigma_n |n\rangle\langle n|H_o$ // sum over the entire full Hilbert Space

$= \Sigma_n |n\rangle\langle n|E_n$ // since states $|n\rangle$ are eigenstates of H_0

$= \Sigma_n E_n |n\rangle\langle n|$ // sum still over all states.

$P_0 H_0 = \Sigma_j E_j |j\rangle\langle j|$ // sum is now only over subspace $\mathcal{E}_a^0$ $E_j = E_a^0$

Meanwhile we know that

$1^{(full)} = \Sigma_n |n\rangle\langle n|$ // we used this above

$1^{(Ea0)} = \Sigma_j |j\rangle\langle j| = P_0$ // unity's portion in subspace $\mathcal{E}_a^0$

So then we know that

$P_0 E_a^0 = E_a^0 \Sigma_j |j\rangle\langle j|$

Thus $P_0 H_0 = P_0 E_a^0$. So here are some words:

Question: Why is $P_0 (H_0 - E_a^0) = 0$?

Answer: When you act on the ANY operator Ω with the operator P_0, whatever you end up with must be entirely within $\mathcal{E}_a^0$. Thus, if Ω is an operator whose eigenvalue is ω_a *for every* basis state of $\mathcal{E}_a^0$, then it follows that $P_0 (\Omega - \omega_a) = 0$. Thus, if you restrict your interest to the space $\mathcal{E}_a^0$, you can always replace operator Ω by number ω_a. H_0 is just an example of such an operator, so $P_0 H_0 = P_0 E_a^0$. I can see this could be a powerful tool that I have not been using.

Now, suppose we were to **define** a new Q_0 operator:

$Q_0 = 1 - P_0$

It seems pretty clear that if we think of $\mathcal{H} = \mathcal{H}^{(m)} \oplus \mathcal{H}^{\perp}$ then it follows that

$Q_0 = \Sigma_k |k><k|$ where k spans the perp space $\mathcal{H}^{\perp}$

This is similar to the Q_0 operator we had on page 688, but there we subtracted out just one state's contribution to unity which we wrote as |0><0|, whereas here we are subtracting out all states in the subspace $\mathcal{E}_a^0$. Probably all the facts we showed about Q_0 are still true but I want to rederive them here:
[P_0 is the subspace projector, Q_0 is the perp (to that subspace) projector]

Rederive facts about Q_0 and Q_{0a} .

Define:

$Q_0 = 1 - \Sigma_{j \in \mathcal{E}_0} |j><j| = \Sigma_{k \in \mathcal{E}\perp} |k><k|$ = "the perp space projector"

Consider: $F_{0a} \equiv Q_0 \; f(H_0) \; Q_0$

Theorem: You can remove either of the Q_0 operators and the result is still the same.

Proof: Let's just evaluate all three objects and show they are the same:

(1) $Q_0 \; f(H_0) = \Sigma_{k \in \mathcal{E}\perp} |k><k| \; f(H_0) = \Sigma_{k \in \mathcal{E}\perp} |k> f(E_k) <k|$

(2) $f(H_0) \; Q_0 = f(H_0) \; \Sigma_{k \in \mathcal{E}\perp} |k><k| = \Sigma_{k \in \mathcal{E}\perp} f(E_k) |k><k| = \Sigma_{k \in \mathcal{E}\perp} |k> f(E_k) <k|$

(3) $Q_0 \; f(H_0) \; Q_0 = Q_0 \; [f(H_0) \; Q_0] = Q_0 \; \Sigma_{k \in \mathcal{E}\perp} |k> f(E_k) <k| = \Sigma_{k \in \mathcal{E}\perp} |k> f(E_k) <k|$

Application: Let $f(H_0) = (E_a^0 - H_0)^{-1}$. Then we know that:

$Q_{0a} \equiv Q_0(E_a^0 - H_0)^{-1}Q_0 = Q_0(E_a^0 - H_0)^{-1} = (E_a^0 - H_0)^{-1}Q_0$ // **see warning!**

WARNING: Two operators Ax=f and Bx=f are equal if they have the same domain and range, and if on that common domain they have the same values in the range. If we look at the operators above acting on the normal ket space, we find that

$Q_0(E_a^0 - H_0)^{-1}Q_0 : (\mathcal{E}_a^0)^{\perp} => (\mathcal{E}_a^0)^{\perp}$ (1)

using my special domain/range notation. Any point outside this domain is mapped to 0. So we can consider

$Q_0(E_a^0 - H_0)^{-1}Q_0 : \mathcal{H} => (\mathcal{E}_a^0)^{\perp}$ (2)

as a Stakgold "extension" of the first mentioned operator, where the mapped value is 0 for points outside $(\mathcal{E}_a^0)^\perp$. If we compare this last operator to

$$Q_0(E_a^0 - H_0)^{-1}: \mathcal{H} \Rightarrow (\mathcal{E}_a^0)^\perp \qquad (3)$$

we see that operators (2) and (3) are *not the same*. In the extension space $\mathcal{E}_a^0$, the operator (2) gives 0, but operator (3) diverges! So our proof of equality shown above is only valid for $(\mathcal{E}_a^0)^\perp \Rightarrow (\mathcal{E}_a^0)^\perp$.

From the last two forms we then know that

$$Q_{0a}(E_a^0 - H_0) = (E_a^0 - H_0)Q_{0a} = Q_0$$

Now we can, as before, **apply this to 16.7^n** which says

$$(E_a^0 - H_0)\,|n\rangle = [+ (V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle - \varepsilon_n|0\rangle]$$

where the state here being corrected could be either in the perp space or the subspace. We know that *in either case* we can apply Q_{0a} to both sides and we get

$$Q_0\,|n\rangle = Q_{0a}[(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle - \varepsilon_n|0\rangle]$$

(1) Suppose $|0\rangle$ lies in the perp space. Then $Q_0|0\rangle = |0\rangle$ and we just get

$$Q_0\,|n\rangle = Q_{0a}[(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle]$$

Now we don't know which space the various corrections lie in, so we cannot say $Q_0\,|n\rangle = |n\rangle$ as we could before when we know that all corrections to $|0\rangle$ were in the "perp space due to the $\langle n|0\rangle = 0$ normalization convention. So here we have to just let Q_0 sit there.

(2) Suppose $|0\rangle$ lies in the sub space. Then $Q_0|0\rangle = 0$ and we just get

$$Q_0\,|n\rangle = Q_{0a}[(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - .... - \varepsilon_{n-1}|1\rangle - \varepsilon_n|0\rangle]$$

so our only change is that we can throw out the last term, since $|0\rangle$ lies in the subspace for sure. We might as well then not bother with these two special cases and just include the $\varepsilon_n|0\rangle$ term.

Suppose n = 1. We then get (think of this as a version of equation 16.7^1)

$$Q_0\,|1\rangle = Q_{0a}(V-\varepsilon_1)|0\rangle = Q_{0a}\,V|0\rangle \qquad \text{// } \textbf{agrees with 16.37}$$

After much fiddling, we are now starting the second paragraph on page 699. Just as a reminder, from 16.36 we also know that

$P_0(V-\varepsilon_1)|0\rangle = 0$

Rewrite this as (we are now assuming that $|0\rangle$ is in the subspace somewhere)

$P_0V|0\rangle = \varepsilon_1|0\rangle$

$P_0VP_0|0\rangle = \varepsilon_1|0\rangle$

He points out that this latter (either really) is an eigenvalue equation. If we close with some arbitrary state *in the subspace*, $\langle E_a^0\alpha|$, we get (using $P_o = \Sigma_{\alpha'} |E_a\alpha'\rangle\langle E_a\alpha'|$)

$\langle E_a^0\alpha|P_0VP_0|0\rangle = \varepsilon_1 \langle E_a^0\alpha|0\rangle$

$\langle E_a^0\alpha|V P_0|0\rangle = \varepsilon_1 \langle E_a^0\alpha|0\rangle$

$\langle E_a^0\alpha|V \Sigma_{\alpha'} |E_a\alpha'\rangle\langle E_a\alpha'| 0\rangle = \varepsilon_1 \langle E_a^0\alpha|0\rangle$

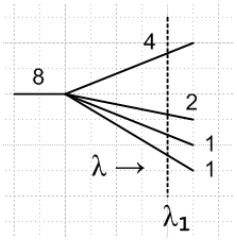
$= \Sigma_{\alpha'} \langle E_a^0\alpha|V|E_a\alpha'\rangle\langle E_a\alpha'| 0\rangle = \varepsilon_1 \langle E_a^0\alpha|0\rangle$ **// which is page 699 A**

which I might abbreviate as

$H'_{\alpha\alpha'}\psi_{0\alpha'} = \varepsilon_1 \psi_{0\alpha'} = V_{\alpha\alpha'}\psi_{0\alpha'\ell}$ // again, this is "a version of" 16.7^1

which is my famous subspace diagonalization problem! In Messiah, we have $N = g_a$ = the subspace dimension. So there will be g_a eigenvalues for ε_1. It is likely that a *particular* eigenvalue ε_1 will have some multiplicity, but it could also be non-degenerate.

Suppose a particular ε_1 is non-degenerate. I think he is saying that this is a "singleton" in my picture above (copied below), and it will "come in for a landing" on some well defined state in $\mathcal{E}_a^0$.



He claims that the state $|0\rangle$ is "fully determined to the zeroth order". He says this follows from $16.7^{0,1}$. Why is this?? I would have thought that 16.7^0 would be enough, since it just says $H_0|0\rangle = E_a^0|0\rangle$. Let's put this question on hold for a moment. // Well, as I have commented above, the matrix equation itself comes from 16.7^1 and so that is where ε_1 and its non-degeneracy (if true) comes from.

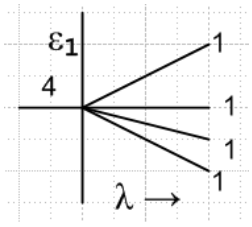
Comment: This text is at the nub of my various problems and I have to study each phrase *very* carefully.

$Q_0|1\rangle$ = "the perp space component of the correction $|1\rangle$ " = determined by 16.37 // I agree

But, we don't know what $P_0|1>$ is, the *subspace* projection. It is "undetermined" so far simply because we have not seen any expression for this. We do know that $<1|0> = 0$ from our norm convention.

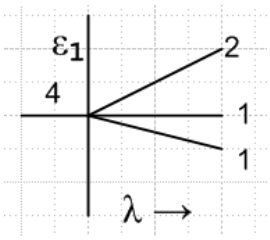
Now suppose instead that ε_1 has multiplicity g_1. In this case, we know that the landing state $|0>$ is in the subspace $\mathcal{E}_1 = \mathcal{E}_a^{(1)}$ (as were defined on page 698), and in this case $|0>$ is not known! He claims that if you want to know what $|0>$ is, you have to "go to higher orders". I don't really know what he means by this. // But see below. At some order you hope to get an EV equation which you can solve for the actual states so you know what they are.

Review: I took equation 7^1 and applied P_0 to it and used my now-obvious fact that $P_0 (H_0 - E_a^0) = 0$ to get (3) which said $P_0(V-\varepsilon_1)|0> = 0$. I first wrote this as an EV equation $(P_0V)|0> = \varepsilon_1|0>$. I then made the assumption that $|0>$ was some state in $\mathcal{E}_a^0$ -- up to now that fact was not needed. This then allowed me to write the new EV equation $P_0VP_0|0> = \varepsilon_1|0>$. I then closed this with an arbitrary $\mathcal{E}_a^0$ state to get $<E_a^0\alpha|V P_0|0> = \varepsilon_1 <E_a^0\alpha|0>$, where now on the left you see that the "right side P_0" was necessary in the triplet P_0VP_0. I could have just "pulled P_o out of the vacuum" so to speak, knowing that $|0> \in \mathcal{E}_a^0$. I then used the sum expansion for P_0 which is $P_o = \Sigma_{\alpha'} |E_a\alpha'><E_a\alpha'|$ and I ended up with a matrix equation $V_{\alpha\alpha'}\psi_{0\alpha'} = \varepsilon_1 \psi_{0\alpha}$ and this then applied to any state $|0>$ in $\mathcal{E}_a^0$. This thing is an NxN matrix EV problem involving all of the $\mathcal{E}_a^0$ space, dim = N = g_a. $\psi_{0\alpha}$ is the *component* of our $|0>$ state onto the base state called α or $<E_a^0\alpha|$ in longer notation. Our *matrix* equation is then $V\psi_0 = \varepsilon_1 \psi_0$. In our notation, we don't have any label that says which of the $\mathcal{E}_a$ states that $|0>$ is, we could say $V\psi_0^i = \varepsilon_1^i \psi_0^i$ to indicate it is the ith base state (and maybe $|0; i>$). Then at least we have a way to enumerate the eigenvalues ε_1^i. In the simplest case to consider, all N of the ε_1^i are different, and then we could solve this EV equation for the unique eigenvectors ψ_0^i. This is thing I like to refer to as the "pre-diagonalization" process. In this case that all ε_1^i are different for i = 1..N, our picture would show that all states in $\mathcal{E}_a$ are "breaking out into singletons in the first order energy correction", and here is the picture:



The solution of equation $V\psi_0^i = \varepsilon_1^i \psi_0^i$ will determine the "correct" 0th order wavefunctions and the 1st order energy correction. If we solve this equation and find the four unique ψ_0^i eigenvectors, then those are the four states that are going to fan out as shown when $\lambda > 0$. This is an energy plot, ε_1 on the vertical axis. So the point is that we know what the correct fanning out states are! So, if at the very start of your problem, you pre-diagonalize H' = V, then you will have the correct fanning out states if the eigenvalues are all different. I think all is well here.

Now, suppose we instead find something like this:



so that two of the ψ_0^i have the same value for ε_1^i. Then I suppose at this point, any linear combination of the two eigenstates ψ_0^i with this eigenvalue would be acceptable as your starting states.

Now comes the big point: In the above picture, perhaps the 2 states stay together at the ε_1 level. They may however split apart at the ε_2 correction level! Or if not there, at some higher level. This then explains the comment I underlined in red (p 699) saying you then must "proceed to higher orders". In my example above, we have $g_1 = 2$. We want to call this sub-sub space by the name $\mathcal{E}_a^{(1)}$ with dimension g_1. Clearly we have that $\mathcal{E}_a^{(1)} \subset \mathcal{E}_a$.

So let's try to mess around at the ε_2 level a bit. We obviously will deal with equation 7^2 which is the lowest on the ladder which shows ε_2 somewhere. Messiah then suggests that we consider a projection operator $P_0^{(1)}$ which projects out only components in $\mathcal{E}_a^{(1)}$. I like it. Then let

$P_0 = P^{(1)} + P'$ where P' projects onto the states in the perp space of $\mathcal{E}_a^{(1)}$ within $\mathcal{E}_a$

These projection operators are maybe better thought of as "filters" which block the wrong stuff and only pass the right stuff. So I now agree finally with **equations B on page 699** and we can finally turn the page!

Now we are going to do a little algebra top of page 700. Recall our little theorem above which said this:

Consider: $F_{0a} \equiv Q_0\ f(H_0)\ Q_0 = \Sigma_{k\in\mathcal{E}\perp} |k\rangle\ f(E_k)\ \langle k|$ $Q_0 = \Sigma_{k\in\mathcal{E}\perp} |k\rangle\langle k|$

Theorem: $F_{0a} \equiv Q_0\ f(H_0)\ Q_0 = f(H_0)\ Q_0 = Q_0\ f(H_0)$

where Q_0 was the perp space projector.

New Theorem. I think we can make a similar theorem for P_0 like this:

$G_{0a} \equiv P_0\ f(H_0)\ P_0 = f(H_0)\ P_0 = P_0\ f(H_0)$ // similar warning as noted above!

As soon we you write $P_0 = \Sigma_{k\in\varepsilon a0} |k\rangle\langle k|$, the H_0 converts to E_a^0 and then $f(H_0)$ becomes a number and you can put it anywhere you want.

This theorem would of course apply in any subspace $\mathcal{E}_a^{(1)}$ of $\mathcal{E}_a^0$ where you might use $P^{(1)}$ as the projector onto that subspace. So we can repeat the above claims with any superscript on P_0 we like. In particular, we can say

$P^{(1)} H_0 = E_a^0 P^{(1)}$ **B**

Another obvious fact is this one:

$P^{(1)} P_o = P^{(1)}$ **A**

since if you first project something into $\mathcal{E}_a^0$ and thence into $\mathcal{E}_a^{(1)}$, it is the same as if you go in one shot.

So **I have now justified all equations A,B,C on page 700,** the last arising because (..) = 1 from 699 B. How then do we get D? We have three terms to look at. The third term passes through. The first two terms are this:

$$P^{(1)}V (P^{(1)} + P') = P^{(1)}VP_0$$

He is claiming that

$$P^{(1)}VP_0 = \varepsilon_1 P^{(1)}$$

but I don't offhand see why that is true, another roadblock of many, which is why I move one sentence per hour in Messiah. We are not applying to a state here as we did in 16.36. We could apply this to an arbitrary state in our $\mathcal{H}$ to check its validity. We get 0=0 for any state not in $\mathcal{E}_a^0$, so we can restrict to that. We did know from earlier that $P_0VP_0 |0> = \varepsilon_1|0>$ for $|0>$ being any state in $\mathcal{E}_a^0$. So to test the validity of the above claimed equation, we can start with this fact we know:

$P_0VP_0 |0> = \varepsilon_1|0>$ $\forall$ $|0>$ in $\mathcal{E}_a^0$

We don't have a general *operator* equation $P_0VP_0 = \varepsilon_1$ here because it would not be true for a ket in the perp space. But we could fix that by saying $P_0VP_0 = \varepsilon_1 P_0$. So this IS a correct operator equation. So here we go:

$P^{(1)}VP_0 = P^{(1)} P_0VP_0$ // by our "obvious fact" quoted above

$= P^{(1)} \varepsilon_1 P_0$ // using our operator equation $P_0VP_0 = \varepsilon_1 P_0$ just quoted

$= \varepsilon_1 P^{(1)}$ // using the "obvious fact" again **D'**

and **thus we have shown p 700 D**. We can rewrite this equation D as:

$P^{(1)}V - \varepsilon_1 P^{(1)} = P^{(1)}(V - \varepsilon_1) = P^{(1)}VQ_0$ **D**

Now finally we are more prepared to mess with the 7^2 equation which starts out saying

$(H_0 - E_a^0) |2> + (V-\varepsilon_1)|1> - \varepsilon_2|0> = 0$

Now apply $P^{(1)}$ from the left. Our usual state expansion idea for $P^{(1)} = \Sigma_{k \in \varepsilon 1} |k><k|$ kills off the first term because it sets $H_0 = E_a^0$. We then get

$$P^{(1)} (V-\varepsilon_1)|1> - \varepsilon_2 P^{(1)} |0> = 0$$

But we now recognize $P^{(1)} (V-\varepsilon_1) = P^{(1)} V Q_0$ from our alternate D form above, so we have

$$P^{(1)} V Q_0 |1> - \varepsilon_2 P^{(1)} |0> = 0 \qquad \textbf{E}$$

which is **result E**. BUT now look back at our earlier result 37 which said

$$Q_0 |1> = Q_{0a} V|0>$$

so we then have

$$P^{(1)} V Q_{0a} V|0> - \varepsilon_2 P^{(1)} |0> = 0$$

$$P^{(1)} [V Q_{0a} V - \varepsilon_2] |0> = 0 \qquad \textbf{// which is 16.39} \qquad \text{version of } 16.7^2$$

and this was certainly a long time coming. This equation resembles 16.36. So as we did before, we can write this as **an EV equation** where ε_2 is the EV:

$$P^{(1)} V Q_{0a} V |0> = \varepsilon_2 P^{(1)} |0>$$

Now let's assume that |0> is in our $\mathcal{E}^{(1)}$ part of $\mathcal{E}_a^0$. Then we are free to insert $P^{(1)}$ on the left and remove it on the right to get

$$P^{(1)} (V Q_{0a} V) P^{(1)} |0> = \varepsilon_2 |0> \qquad \mathbf{E}_1$$

which reminds us of our earlier equation

$$P_0 (V) P_0 |0> = \varepsilon_1 |0>$$

So here is what we are supposed now to do. We now have an eigenvalue problem within $\mathcal{E}^{(1)}$ and we are going to get a matrix equation like this:

$$(V Q_{0a} V) \psi_0^i = \varepsilon_2^i \psi_0^i \qquad \text{where } i = 1,2.....g_i \qquad \text{within } \mathcal{E}^{(1)}$$

Suppose now that these ε_2^i are all different. Then we get unique ψ_0^i and these are then the correct states back in $\mathcal{E}_a^0$ for us to "start with". We could not find these states here until we went to second order in the energy correction!

So I presume in his section to come, he will show that you just keep doing this process to as many orders as you need until you get a splitting of the energies.

Now we backtrack to say this: suppose our original ε_1 was non-degenerate. Then we were able to identify |0> cleanly. Then consider:

$$(V\ Q_{0a}\ V)|0\rangle = \varepsilon_2\ |0\rangle$$

$\varepsilon_2 = \langle 0|\ V\ Q_{0a}\ V|0\rangle$ **// which is result F on page 700**

so in that case (ε_1 non-degenerate), we have a formula for the ε_2 correction. And if ε_1 was degenerate with multiplicity g_1, then we solve the above matrix equation $(V\ Q_{0a}\ V)\ \psi_0{}^i = \varepsilon_2{}^i\ \psi_0{}^i$ to get the ε_2 energies, and if they are all different, we are done and we know the states $\psi_0{}^i$. Otherwise, we have to continue on to third order!

He comments that you may retain degeneracy through all orders of perturbation theory.

Note added 1.28.09. When I first was reading this, I did not correctly interpret the meaning of a state |0>. One must understand that this is not "just any state" in $\mathcal{E}_a{}^0$ or $\mathcal{E}^{(1)}$ or what have you. The state |0> is the "zeroth order component" of a state that solves the full SE to all orders in λ. So only certain linear combinations of the "starting base states" (which I like to call |r> and he calls $|E^0{}_a\alpha\rangle$) are "legal" for |0>. To find the legal linear combinations, you have to process things out to a point where degeneracy is broken! So when you solve the EV problem at some level, such as $(V\ Q_{0a}\ V)|0\rangle = \varepsilon_2\ |0$ above, you not only discover the eigenvalues ε_2 , but you also discover the correct states |0> that are "legal". You might say that these are the shadows or projections into the ε_0 level of full SE solutions.

Comment: In my personal work on this problem, I just assumed that, when I did my pre-diagonalization, all the eigenvalues were different in the $\mathcal{H}^{(m)}$ space, and thus I was "done" with the problem of identifying the correct states to start with. Here Messiah has made the point that if this is not the case, you try to get the right states by going to higher and higher order just as he described for second order. I think he is going to present a fancy formalism showing exactly how you do this.

More Comments:

Messiah always writes an unperturbed state as |0>. Although you might then have formulas involving this state, he points out that you cannot *use* those formulas unless you actually know what the state |0> *is*. When you have a degeneracy situation, you may have to go to something like fourth order ε_4 before you can know what your states *are*. Otherwise all you know is that the state |0> is some as-yet-undetermined vector in a degenerate subspace, and then you cannot do a calculation with a formula involving |0>, such as $\varepsilon_1 = \langle 0|V|0\rangle$. // good, 1.28.09

Clarification of my Theorem B confusion.

One of my main motivations for even looking at Messiah was to understand my **Theorem B** which was this:

You cannot "do" a perturbation analysis on the states u_{mi} unless they diagonalize H' within $\mathcal{H}^{(m)}$.

In the proof of that theorem I start with the ladder equations

$(H_0 - E_m)\,\psi_{mi,1} = (W_{mi,1} - H')\,u_{mi}$ i = 1,2....N " level 1"

$(E_a - H_0)\,|1,i\rangle = (V-\varepsilon_1{}^i)|0,i\rangle$ // Messiah notation for same equation

where I just assume that the m subspace is spanned by some orthonormal basis u_{mi} i = 1,2..N. I assume this without regard to whether or not that basis diagonalizes H'. I know from Hilbert Space theory that in any finite subspace I can always construct such an orthonormal basis by doing GSO.

Now, just staring at either equation above, you don't "see" a matrix eigenvalue equation. In Messiah, we make this matrix equation appear by applying P_0 to both sides. This kills the LHS and gives $P_0\,(V-\varepsilon_1)|0\rangle = 0$ which we convert to our desired EV equation. In Schiff, we instead "close" from the left to get this

$\langle u_{mj}|\,(W_{mi,1} - H')\,|u_{mi}\rangle = 0 \quad \Leftrightarrow \quad H'_{ij} = W_{mi,1}\,(1)_{ij}$

and this also "kills" the LHS. What we now "see" is a *diagonal* EV equation! How come I don't ever "see" the non-diagonalized equation first? That is my complaint. [a good complaint I think]

OK, go back to the starting point again,

$(H_0 - E_m)\,\psi_{mi,1} = (W_{mi,1} - H')\,u_{mi}$

Suppose we "restrict our interest" to the m subspace. Then can I set the LHS = 0? Well, if by "restrict our interest" we mean that "we promise only to close from the left with subspace states", then I guess you can set LHS = 0. We then have within $\mathcal{H}^{(m)}$ this set of equations, for i = 1,2,3....

$0 = (W_{mi,1} - H')\,u_{mi} \quad \Rightarrow \quad H'\,u_{mi} = W_{mi,1}\,u_{mi} \quad H'\,|u_{mi}\rangle = W_{mi,1}\,|u_{mi}\rangle$

This says that the $|u_{mi}\rangle$ are solutions to the EV equation shown, in addition to being an arbitrary set of orthonormal basis functions.

So, when we start, we just say the $|u_{mi}\rangle$ span $\mathcal{H}^{(m)}$ and are arbitrary but orthonormal. But then when we examine our order-λ^1 ladder equations restricting to $\mathcal{H}^{(m)}$, we learn that in fact these ladder equations say that the kets $|u_{mi}\rangle$ are eigenkets of H' so they cannot be completely "arbitrary" after all.

You can certainly construct an orthonormal basis $|v_{mi}\rangle$ which spans $\mathcal{H}^{(m)}$ and which does NOT diagonalize H'. But these states then do not satisfy our ladder equations. We will have

$(W_{mi,1} - H')\,v_{mi} \neq 0$ at least for one or more i in 1..N

Let's try saying it again and again until it sinks in:

Fact: If you are looking for an orthonormal basis of kets in $\mathcal{H}^{(m)}$ which *solve* the level-1 ladder equations in $\mathcal{H}^{(m)}$, then those kets must be eigenkets of H' and so H' will be diagonal in that basis.

In my notes, once again, I said this for Theorem B:

You cannot "do" a perturbation analysis on the states u_{mi} unless they diagonalize H' within $\mathcal{H}^{(m)}$.

and I used the term "inconsistent". I would now state this differently. I would say that if you want a set of $\mathcal{H}^{(m)}$ basis states that *solve* the ladder equations through level 1, then those basis states must diagonalize H'. So the first order of business is going to be to get H' diagonalized within $\mathcal{H}^{(m)}$, and of course in so doing we learn the spectrum for $\varepsilon_1 = W_{mi,1}$.

And we can go further. Suppose we find that there is still some degenerate subspace after we diagonalize H' such that some ε_1 has multiplicity g_1. Then our claim will be this: If you want a set of basis states within $\mathcal{H}^{(1)} \subset \mathcal{H}^{(m)}$ that *solve* the ladder equations *through level* 2, then those basis states must diagonalize the operator $H'Q_{mo}H' = VQ_{a0}V$ as Messiah shows page 700 (39). Here, the operator Q_{0m} is given by

$$Q_{0m} \equiv Q_0(E_m - H_0)^{-1}Q_0$$

where Q_0 is the perp space projector relative to $\mathcal{H}^{(m)}$ and of course E_m is the eigenenergy of $\mathcal{H}^{(m)}$.

And presumably if we still have some degeneracy after this ε_2 analysis, we use the level three equations to resolve things.

Clarification of my B'$_{ij}$ confusion.

In my own "personal" development of perturbation theory, I would write things like so

$$\psi_{mi} = \{ u_{mi} + \Sigma_{j\neq i} B'_{ij} u_{mj} \} \oplus \Sigma_{k\neq m} B'_{ik} u_k$$

$$\psi_{mi,n} = \{ \Sigma_{j\neq i} B'_{ij,n} u_{mj} \} \oplus \Sigma_{k\neq m} B'_{ik,n} u_k \qquad n > 0$$

and I would ask "how do I compute the coefficients B'$_{ij}$ " ? These tell you how the perturbed state picks up components of basis states within $\mathcal{H}^{(m)}$. I kept finding that these coefficients were sort of undeterminable. In Messiah, we see a similar thing happening. When we consider the level 1 ladder equations, we find page 699 16.37 that

$$Q_0 |1\rangle = Q_{0a} H |0\rangle$$

but this is telling us information only about the part of the first order state correction $B'_{ik,1}$ that is in the perp space to $\mathcal{H}^{(m)}$! We learn nothing about $B'_{ij,n}$. So at least both Messiah and I have "the same problem" at this ladder equations level.

Flailing Away with no real success:

Now you would think that the level 2 ladder equations would tell you something about the $B'_{ij,n}$. Suppose that we diagonalize H' and this removes all degeneracy, the simplest case. Then "if we restrict our interest" to $\mathcal{H}^{(m)}$, our level 2 ladder equation set reads:

$$(V-\varepsilon_1)|1\rangle - \varepsilon_2|0\rangle = 0 \qquad \text{// "level 2"}$$

which in my notation is more like this:

$$(H' - W_{im,1})\psi_{mi,1} - W_{im,2}\, u_{mi} = 0$$

which becomes

$$(H' - W_{im,1})[\ \{\ \Sigma_{j\neq i} B'_{ij,1}\ u_{mj}\ \} \oplus \Sigma_{k\neq m} B'_{ik,1}\, u_k\ ] = W_{im,2}\, u_{mi}$$

At this point, I "closed" with some u_{ms} inside $\mathcal{H}^{(m)}$

$$u_{ms}\,(H' - W_{im,1})[\ \{\ \Sigma_{j\neq i} B'_{ij,1}\ u_{mj}\ \} \oplus \Sigma_{k\neq m} B'_{ik,1}\, u_k\ ] = W_{im,2}\ \delta_{s,i}$$

$$\Sigma_{j\neq i} B'_{ij,1}\, u_{ms}\,(H' - W_{im,1})\, u_{mj} + \Sigma_{k\neq m} B'_{ik,1}\, u_{ms}\,(H' - W_{im,1})\, u_k = W_{im,2}\ \delta_{s,i}$$

But $(H' - W_{im,1})$ is a diagonal matrix within $\mathcal{H}^{(m)}$ (only!) so this all becomes

$$\Sigma_{j\neq i} B'_{ij,1}\,(H'_{ss} - W_{im,1})\,\delta_{s,j} + \Sigma_{k\neq m} B'_{ik,1}\,(H'_{sk} - W_{im,1}\,\delta_{s,k}) = W_{im,2}\ \delta_{s,i}$$

$$\Sigma_{j\neq i} B'_{ij,1}\,(W_{sm,1} - W_{im,1})\,\delta_{s,j} + \Sigma_{k\neq m} B'_{ik,1}\, H'_{sk} = W_{im,2}\ \delta_{s,i}$$

We then looked first at s = i and got then no hit from the first term and

$$\Sigma_{k\neq m} B'_{sk,1}\, H'_{sk} = W_{sm,2}$$

and then we looked with $s \neq i$ and got

$$B'_{is,1}\,(W_{sm,1} - W_{im,1}) + \Sigma_{k\neq m} B'_{ik,1}\, H'_{sk} = 0$$

and then since we assumed that all the H' diagonal values are different, we get our answer:

$$B'_{is,1} = [\Sigma_{k\neq m} B'_{ik,1}\, H'_{sk}]\,/(W_{mi,1} - W_{ms,1}) \qquad i \neq s \quad W_{mi,1} \neq W_{ms,1}$$

But we learned earlier in my notes that

$$B'_{ik,1} = H'_{ki}/(E_m - E_k)$$

so we end up then with

$$B'_{is,1} = 1/(W_{mi,1} - W_{ms,1}) * [\Sigma_{k\neq m} H'_{sk}\ H'_{ki}/(E_m - E_k)] \qquad i \neq s \quad W_{mi,1} \neq W_{ms,1}$$

and of course we know that $B'_{ss,1} = 0$ from the way we normalized. So this result says that you get some non-zero $B'_{is,1}$ only if $H'_{sk}\ H'_{ki} \neq 0$ for at least some k in the perp space (and even in this case it could in theory still be zero). Let's write this out a little:

$B'_{is,1} = 1/(W_{mi,1} - W_{ms,1}) * [\Sigma_{k\neq m} \langle s\,H'|k\rangle (E_m - E_k)^{-1} \langle k|H'|i\rangle]$

Now recall from Messiah above that

$$Q_{0m} = \Sigma_{k\neq a} |E_k^0\rangle (E_m^0 - E_k^0)^{-1} \langle E_k^0|$$

Thus, we can write our result more compactly as

$B'_{is,1} = 1/(\varepsilon_1^i - \varepsilon_1^s) * \langle s | H' Q_{0m} H' | i\rangle$

Also, from above we just had

$\Sigma_{k\neq m} B'_{sk,1} H'_{sk} = W_{sm,2}$

$\varepsilon_2^i = \Sigma_{k\neq m} B'_{sk,1} H'_{sk} = \Sigma_{k\neq m} H'_{sk} B'_{sk,1} H'_{ki}/(E_m - E_k) = \langle s| H' Q_{0m} H'|s\rangle$ // see M 700F

because that same structure appears. So we can summarize all these results compactly as:

$\varepsilon_1^i = \langle i | H' | i\rangle$
$B'_{ik,1} = H'_{ki}/(E_m - E_k)$ // level 1 ladder equations

$\varepsilon_2^i = \langle i| H' Q_{0m} H'|i\rangle$
$B'_{is,1} = 1/(\varepsilon_1^i - \varepsilon_1^s) * \langle s | H' Q_{0m} H' | i\rangle$ // level 2 ladder equations

which only applies to the case that $\varepsilon_1^i \neq \varepsilon_1^s$ for all i,s of interest.

So I have replicated my work in my on deg doc, and have put it into more compact messiah notation, and I have a complete answer to the knowing all the $B'_{ab,1}$!

I then realized that we had a problem Houston if $\varepsilon_1^i = \varepsilon_1^s$ and of course this is the very problem that Messiah attacks! Let's see now if I have really learned anything. Assume we have degeneracy g_1 at the ε_1 level, so we are dealing with a subspace $\mathcal{H}^{(1)}$ say of dimension g_1. Let's just try to run through the exact sequence above and see what happens. Cut, paste and edit:

Then "if we restrict our interest" to $\mathcal{H}^{(m)}$, our level 2 ladder equation set reads:

$(V-\varepsilon_1)|1\rangle - \varepsilon_2|0\rangle = 0$ // "level 2"

which in my notation is more like this: [at this point, we know ε_1]

$(H' - W_{im,1})\psi_{mi,1} - W_{im,2} u_{mi} = 0$

which becomes [but we don't yet know exactly what $\psi_{mi,1}$ is, it is somewhere in $\mathcal{H}^{(1)}$]

$$(H' - W_{im,1})[\ \{\ \Sigma_{j\neq i} B'_{ij,1}\ u_{mj}\ \} \oplus \Sigma_{k\neq m} B'_{ik,1}\ u_k\ ]\ =\ W_{im,2}\ u_{mi}$$

[Thus, we are saying we don't know what any of these B' coefficients are yet]

At this point, I "closed" with some u_{ms} inside $\mathcal{H}^{(1)}$

$$u_{ms}\ (H' - W_{im,1})[\ \{\ \Sigma_{j\neq i} B'_{ij,1}\ u_{mj}\ \} \oplus \Sigma_{k\neq m} B'_{ik,1}\ u_k\ ] = W_{im,2}\ \delta_{s,i}$$

$$\Sigma_{j\neq i} B'_{ij,1}\ u_{ms}\ (H' - W_{im,1})\ u_{mj}\ + \Sigma_{k\neq m} B'_{ik,1}\ u_{ms}\ (H' - W_{im,1})\ u_k\ = W_{im,2}\ \delta_{s,i}$$

But $(H' - W_{im,1})$ is a diagonal matrix within $\mathcal{H}^{(1)}$ (only!) so this all becomes

$$\Sigma_{j\neq i} B'_{ij,1}\ (H'_{ss} - W_{im,1})\ \delta_{s,j}\ + \Sigma_{k\neq m} B'_{ik,1}\ (H'_{sk} - W_{im,1}\ \delta_{s,k})\ = W_{im,2}\ \delta_{s,i}$$

$$\Sigma_{j\neq i} B'_{ij,1}\ (W_{sm,1} - W_{im,1})\ \delta_{s,j}\ + \Sigma_{k\neq m} B'_{ik,1}\ H'_{sk}\ = W_{im,2}\ \delta_{s,i}$$

We then looked first at s = i and got then no hit from the first term and

$$\Sigma_{k\neq m} B'_{sk,1}\ H'_{sk}\ = W_{sm,2}\ = \varepsilon_2{}^{s}\ =\ \langle s |\ H'\ Q_{0m}\ H'\ | s\rangle\ \ [\text{ is likely different for different s }]$$

and then we looked with $s \neq i$ and get

$$B'_{is,1}\ (W_{sm,1} - W_{im,1})\ + \Sigma_{k\neq m} B'_{ik,1}\ H'_{sk}\ = 0$$

$$\Rightarrow \qquad \Sigma_{k\neq m} B'_{ik,1}\ H'_{sk}\ = 0$$

$$\Rightarrow \qquad [\Sigma_{k\neq m} \langle s\ H'|k\rangle\ (E_m - E_k)^{-1}\ \langle k|H'|i\rangle\ ]\ = \langle s |\ H'\ Q_{0m}\ H'\ | i\rangle\ = 0$$

So what we learn at this level is that the states in $\mathcal{H}^{(1)}$ which I am calling $|i\rangle$ [a completely different notation than the $|n\rangle$ of Messiah!!!] must diagonalize the operator $H'\ Q_{0m}\ H'$ if they are to solve the level 2 ladder equations (and all lower level ones). Once we actually carry out this diagonalization, we then know the various $\varepsilon_2{}^{s} = \langle s |\ H'\ Q_{0m}\ H'\ | s\rangle$ which are just the diagonal elements of our diagonalized matrix. Since we now know exactly what these states are, we can actually compute something like this. Of course since Q_{0m} involves the perp space $\mathcal{H}^{\perp}$ which is infinite dimensional, we have a rough calculation to do, since there are an infinite number of terms to add up, as $\Sigma_{k\neq m}$ above shows. We might eliminate certain terms in the sum by using symmetry to show that certain $\langle k|H'|i\rangle = 0$. Or we might approximate by only including "nearby" states $|k\rangle$ which have therefore small energy denominators which dominate the sum.

Now, never letting go like a good bulldog, I want to compute $B'_{is,1}$ for these states I just supposedly found by diagonalizing $H'\ Q_{0m}\ H'$. Obviously we are not going to get an answer from the above level 2

equations, so we go to the level 3 equations, and now I am a bit on my own. Let's try this in Messiah notation first. From 16.7^3 we have

$$(H_0 - E_a^0)\,|3\rangle + (V-\varepsilon_1)|2\rangle - \varepsilon_2|1\rangle - \varepsilon_3|0\rangle = 0$$

where $|0\rangle$ is a state in $\mathcal{H}^{(1)}$. Now I can either "go formal" or "go informal" at this point. Doing the second, I just say "if we restrict our interest closing from the left only with states in $\mathcal{H}^{(1)}$", this equation becomes

$$(V-\varepsilon_1)|2\rangle - \varepsilon_2|1\rangle - \varepsilon_3|0\rangle = 0 \qquad \text{or I could be formal:} \qquad P^{(1)}(V-\varepsilon_1)|2\rangle - \varepsilon_2\, P^{(1)}|1\rangle - \varepsilon_3|0\rangle = 0$$

What do I do next??? Well, this seems to say that

$$\varepsilon_2\, P^{(1)}|1\rangle = P^{(1)}(V-\varepsilon_1)|2\rangle - \varepsilon_3|0\rangle$$

and $P^{(1)}|1\rangle$ is basically what I am looking for. So I guess I need to learn both $|2\rangle$ and ε_3 before I have an answer to my question. But if I close the above with $\langle 0|$, the ε_2 term vanishes and I will get

$$\langle 0|\, P^{(1)}(V-\varepsilon_1)|2\rangle = \varepsilon_3 = \langle 0|\, P^{(1)} V|2\rangle \qquad \text{// because } \langle 0|2\rangle = 0$$

so if I can just find $|2\rangle$, then I can compute ε_3, so we really have *one unknown* $|2\rangle$ that we need to find.

Now go back to the level 2 equations which say

$$(H_0 - E_a^0)\,|2\rangle + (V-\varepsilon_1)|1\rangle - \varepsilon_2|0\rangle = 0$$

If I apply Q_{0a} to both sides this becomes

$$Q_0\,|2\rangle = Q_{0a}\,[\; -(V-\varepsilon_1)|1\rangle + \varepsilon_2|0\rangle\;] = -Q_{0a}\,(V-\varepsilon_1)|1\rangle$$

so this tells me for sure the perp space part of $|2\rangle$

Well, <u>now I am just back to my big confusion</u> as I was in degen doc working on this same problem. I need to thread a path somehow to get to the solution, but I don't know how to do it, and of course now I am caught halfway between two totally different notations.

So now I think is the perfect time to read the systematic solution starting on Messiah page 712, because the author he quotes there has done this threading, so I should not spend 6 weeks trying to reinvent this wheel.

9. Atomic Levels in the Absence of Spin-Orbit Forces (700) skip for now.

10. Spin-Orbit Forces. LS and jj coupling (703) skip for now.

11. The Atom in LS coupling. Splitting due to spin-orbit coupling (705). skip for now.

12. The Zeeman and Paschen-Back Effects (706). skip for now.

13. Symmetry of H and Removal of Degeneracy (709). skip for now.

14. Quasi-Degeneracy (711). skip for now.

review pause here