Kato Round One Notes PhL 1.8.09

This is the first set of notes I took on the incredibly complicated Kato sections 15,16,17 in Messiah's Chapter 16. I now call this set of notes "round one" because I then did two more "rounds" and then finally some meta notes. There is so much sub-detail in this subject, you cannot see the forest for the very dense trees. It is hard to even remember what is generally going on here as you get buried in some piece of it. This round one is 30 pages long! By the way, I had just read Stakgold Chapter 2 and was interested in the Hilbert Space *operator* aspect of things when I started off, but that never was essential to Kato.

TOPIC III. EXPLICIT FORMS FOR THE PERTURBATION EXPANSION IN ALL ORDERS 1

15. The Hamiltonian and its Resolvent G(z). (712). 1

16. Expansion of G(z), P and HP into power series in λV. (714). 3

Derivation of page 716 A. 4

Review of derivation of page 716 A. [Again, skip this and see Kato's Bag in Round Three] 8

Derive (67) and (68) 10

Derive (70) and (71) 12

Summary: 14

17. Calculation of Eigenvalues and Eigenvectors (717). 15

Case $g_a = 1$. 15

Case $g_a > 1$. 22

Comments on the Generalized Eigenvalue Problem. 26

Topic III. Explicit Forms for the Perturbation Expansion in all Orders

15. The Hamiltonian and its Resolvent G(z). (712).

So here we go, starting 1.20.09 after having read Chap 2 Stakgold, and then having reviewed pretty completely all the perturbation notes I have made up to this point, including the notes above.

Recall that you might want to solve the equation $Lu - \lambda u = f$, you write it as $(L-\lambda)u = f$, and then $u = (L-\lambda)^{-1}f$ is your formal solution. In our case, I think we want to use $L = H$ for our differential operator, the Hamiltonian of some problem. Then $Hu - \lambda u = f \Rightarrow u = (H-\lambda)^{-1}f$ and we can call $(H-\lambda)^{-1}$ by the name "resolvent" and this is what appears in (52) except $\lambda = z$.

Footnote 1:

The bound is $||A|| = \max ||Ax||^2 / ||x||^2$ as I now well know , so $||A|| = M$. Messiah claims that the various notions of the space E_n pass to L_2 only if we restrict to bounded = continuous linear operators. He does not say d/dx is a bounded operator (it is not), he says that the usual algebraic things we are used to doing, such as d/dx and limits and convergence can only be *applied* to a bounded operator. The claim is that $(H-\lambda)^{-1}$ is bounded!

Question: Is H a self-adjoint operator, or is it just symmetric in Stakgold lingo? Recall this theorem:

Theorem 4: If A symmetric on *all* of H, then A is self-adjoint. (this is the case in E_n)

I am betting that H acts on our entire L_2 Hilbert Space, so it really is self-adjoint, and therefore also symmetric. Then here is another theorem we can have:

Theorem 4: The entire spectrum of a *self-adjoint* operator lies on the real axis, and there is no residual spectrum, meaning that case 4 never occurs.

So our H Hamiltonian will have only a point spectrum and a continuous spectrum, at most. I think the conditions of a bound state problem which we are studying right now make our eigenvalue range of interest rule out the continuous spectrum. [but figure on page 713 shows some continuous spectrum.] Therefore, for our problem, we have only the point spectrum of H, and all other λ are in the resolvent set. In this set, $B = H-\lambda I$ is regular, and our definition of regular means that B^{-1} is bounded! If it were unbounded, we would be talking the continuous spectrum. But how would you prove directly that B^{-1} was bounded?

$$\max \|B^{-1}x\|^2/\|x\|^2 = \text{some finite number.}$$

I have to admit, I am right now not sure how I would show this for our $B = H-\lambda$, even knowing that H is self-adjoint.

Is H bounded in our problem? For only bound state solutions, I guess we could show that $\|H\psi\|/\|\psi\|$ was less than the most negative eigenvalue which we know exists for a QM problem (the ground state). So even though d/dx is an unbounded operator in general, I suppose H is bounded for our problem. And then H* is also bounded. And then B is bounded. But I am still hard pressed to show that B^{-1} is bounded as Messiah claims.

I will presume that $B^{-1}(\lambda)$ = his G(z) is analytic except at the eigenvalues where we know H is singular so we are in the nullspace of B. From my Chap 2 Stak notes:

Note: Is Stakgold's continuous spectrum associated with the scattering type solutions in QM where the eigenvalue E is continuous? I assume this is true. Is the term eigenvalue then appropriate in this case?
[In any event, we avoid the continuous part of the spectrum when we draw our special contour Γ_a.]

Fact: If we write $B(\lambda) = (A - \lambda I)$, then we know that $B(\lambda)$ has zeros at the eigenvalues of A. We know that for any eigenvector ψ, we have $B\psi = 0$ since ψ is in the nullspace of B. For the point spectra, these things really are zeros, they are isolated points. Perhaps if there are degenerate eigenvalues, we have multiple zeros. In any event, if we define the resolvent as $G(\lambda) = B^{-1}(\lambda)$, then the $G(\lambda)$ has **poles** at the point spectra eigenvalues.

So this Messiah set-up is the following: For the full H we have a set of energy poles E_0 and corresponding eigenfunctions which we want to find. Probably as we turn off H' = V, we will see poles move together and merge in the E plane. [yup]

Now his H is the full H, and I guess even this full H has some degeneracy left, so we can talk about P_i for the final problem to all orders, where P_i goes with the subspace of energy E_i.

Page 714 A: First, recall our theorem from above:

$$P_0\ f(H_0)\ P_0 = f(H_0)\ P_0 = P_0\ f(H_0) = \text{replace } H_o \text{ in any of these by } E_a{}^0.$$

Now just replace P_0 by P_i and we can say

$$P_i\ f(H)\ P_i = f(H)\ P_i = P_i\ f(H) = \text{replace H in any of these by } E_i$$

[See Round Two notes for more detail why the above is true.]

As an example, we say $G(H)\ P_i = G(E_i)\ P_i$ where G is our resolvent, and **this is page 714 A**. Then just sum both sides on i and get 1 on the left, so interestingly we now have an expansion for the resolvent:

$G(z) = \Sigma_i\ (z\text{-}E_i)^{-1}\ P_i$ // this is **page 714 (56)**

This says that the projector is the residue of the pole, fascinating, and we have **(57)**.

What about degeneracy and these poles? Suppose we have $G(z) = P_1'/(z\text{-}E_1) + P_1''/(z\text{-}E_1)$ for degenerate eigenvalues where we give each one its own projector. We can write this as $G(z) = (P_1' + P_1'')/(z\text{-}E_1)$. And we would just combine the two projectors and write $G(z) = P_1/(z\text{-}E_1)$, and this P_1 would be the projector onto this little subspace of 2 eigenvalues. So I might write this:

$$G(z) = \Sigma_s\ (\ \Sigma_i\ P_{si}/(z\text{-}E_s)\ ) = \Sigma_s\ (1/(z\text{-}E_s))\ \Sigma_i\ P_{si} = \Sigma_s\ (1/(z\text{-}E_s))\ P_s = \Sigma_s\ (P_s/(z\text{-}E_s))$$

where Σ_s is a sum over multiple degenerate subspaces "s", and within each subspace we sum over the degenerate set of poles as shown, they all have energy E_s so no i index needed there. We then process the result as shown above and define P_s as the entire subspace projector. So this then is how I think we should be interpreting result (56) (a subspace sum as shown above). So now all the derivation so far is justified for degenerate eigenvalues. He should have said something about that I think.

The contour picking up a single eigenvalue pole is called Γ_i. He then allows as you could make a contour Γ which encircles some set of poles, and then the integral is the sum of the projectors for those poles, so we then have **(58)**. I agree with **B** from the definition of G (since H commutes with H), and in pencil at the bottom of page 714 I derive **(59)** which is simple.

16. Expansion of G(z), P and HP into power series in λV. (714).

This sounds very heavy duty! We first define a separate G_o that goes with H_o. Page 715 **A** is seen to be the identity 1 = 1 if you right-mult through by the inverse of the LHS. And then **B** follows at once, very simple. And sure enough, we get **(61)** which is an operator equation which has G on both sides. If we regard these operators as being in coordinate space, say, then the operator products are spatial integrals, and this is an integral equation! This looks just like scattering theory eg BD vol 1. Obviously if λ=0, we get $G = G_o$. Here is the usual iterative solution:

$G = G_o(1+\lambda V[G_o(1+\lambda V\{ G_o(1+\lambda V G_o\} \}) = G_o + \lambda\, G_o\, V G_o + \lambda^2\, G_o\, V G_o V G_o + \lambda^3\, G_o\, V G_o V G_o V G_o$

so yes $G = \Sigma_{n=0}\, \lambda^n\, G_o (V G_o)^n$

At this point, he suddenly starts talking about an **operator P** without defining it. But we are referred to Section 8 which does have such a P. This is the section on the degenerate theory stuff. **P is the projector onto the union of subspaces i=1,2....a which subspaces become $\mathcal{E}_a^0$ when λ=0.** When λ=0, the projector P approaches P_0, the projector onto $\mathcal{E}_a^0$ which has degeneracy $g_1 + g_2 + ... + g_a$. We are supposed to imagine a contour that captures all these eigenvalues both when λ=0 and when λ = small where they have moved a little. This contour is called Γ_a. So page **715 C** is our version of (58).

OK, now take C and insert into it the (62) expansion for G. The integral over G_o is of course the projector with λ=0 and that already has the name P_0.

Now if we restrict our interest to subspace $\mathcal{E}_a^0$, we can say that $G_0 = 1/(z-E_a^0)$ so not an operator, and then we can say that $G_o(VG_o)^n = V^n\, G_0^{n+1}$ and yes, this has a pole of order n+1 at $z = E_a^0$. Suppose we wrote $X = (1/2\pi i) \oint dz\, f(z)/(z-z_o)^{n+1}$. The residue theorem then says $X = 1/n!\ * \partial_n f(z_o)$ so I guess you could call X "the residue of the multiple pole".

Now he is going to do a Laurent series,

The Laurent series for a complex function $f(z)$ about a point c is given by:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z-c)^n$$

where the a_n are constants, defined by a line integral which is a generalization of Cauchy's integral formula:

$$a_n = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)\,dz}{(z-c)^{n+1}}.$$

but I am not sure how he is going to apply this, not sure what f(z) is.

[But the idea of Laurent is that you include some negative powers. Really below we are just going to do a regular Taylor series!]

Derivation of page 716 A. [Skip the raw notes here, ***see summary about 5 below!*** This whole subject is summarized in the Round Three notes, section **"Kato's Bag of Tricks"**. I recommend looking there and skipping here unless there is some particular detail of interest.]

This I think is the key equation of this whole section, but Messiah only gives a tiny hint as to where it comes from. I could not derive it quickly in a half hour of scratch paper, so now have to be more careful. He draws our attention to (56) which I can write like this: (the λ=0 limit of (56), that is to say)

$G_0(z) = P_0/(z-E_0) + \Sigma_{i\neq 0}\, P_i/(z-E_i)$

where these other "i" things are for the other (possibly degenerate) subspaces that make up the full spectrum of H_o when $\lambda=0$. The union of all these spaces is the perp space to $\mathcal{E}_a^0$.

I refer to E_a^0 as E_0 (apologies). Notice that we have already produced the first isolated term in 716A by starting in this manner. So we want now to show this:

$$\Sigma_{i\neq 0}\, P_i/(z-E_i) \;=\; \Sigma_{k=1}^{\infty}\,(-1)^{k-1}\,(z-E_0)^{k-1}\,Q_{0a}^{\,k}$$

First, notice that these sums are not the same. The sum on i goes over the other subspaces of H_o , while the k sum is some kind of Laurent series thing. So we now recall from earlier that

$$Q_{0a} \equiv Q_0(E_0 - H_0)^{-1}Q_0 \;=\; Q_0(E_0 - H_0)^{-1} \;=\; (E_0 - H_0)^{-1}Q_0 \qquad \text{// subject to warning!}$$

[The warning is that the above equalities are only true if you restrict to action on $(\mathcal{E}_a^0)^{\perp}$]

where of course

$$Q_0 = 1 - P_0 \;=\; \Sigma_{i\neq 0}\, P_i \qquad = \text{sum of perp space projectors.}$$

The Laurent power series ingredient is something like this

$$f(z) = \Sigma_{k=0,\infty}\,(z-E_0)^k\, f^{(k)}(E_0)/\,k!$$

but I don't yet know what function f(z) we are going to apply this to. Let's rewrite our *desired* result

$$\Sigma_{i\neq 0}\, P_i/(z-E_i) \;=\; \Sigma_{k=1}^{\infty}\,(-1)^{k-1}\,(z-E_0)^{k-1}\,Q_{0a}^{\,k} \;= -\,1/(z-E_0)\,[\,\Sigma_{k=1}^{\infty}\;(z-E_0)^k\,(-Q_{0a})^k\,]$$

where now we have exposed something that at least "fits" the above power series in some sense.

So, where do we begin? (a common question!)

Note Added 1.31.09: skip over all this scratch paper like Plan A stuff and just read the Review which follows!

Plan A. One approach would be to try to identify f(z). We might write the above sum as

$$[\,\Sigma_{k=1}^{\infty}\;(z-E_0)^k\,(-Q_{0a})^k\,] \;=\; -\,1 \;+\; \Sigma_{k=0}^{\infty}\;(z-E_0)^k\,(-Q_{0a})^k$$

Then we might try to identify

$$f(z) = \Sigma_{k=0,\infty}\,(z-E_0)^k\, f^{(k)}(E_0)/\,k! \;=\; \Sigma_{k=0}^{\infty}\;(z-E_0)^k\,(-Q_{0a})^k$$

which would imply that

$$f^{(k)}(E_0)/\,k! \;=\; (-Q_{0a})^k \;=\; [Q_0(H_0 - E_0)^{-1}Q_0]^k$$

Can we reverse engineer this somehow? Maybe start with k = 1,

$$f^{(1)}(E_0)/\,k! \;=\; (-Q_{0a}) \;=\; [Q_0(H_0 - E_0)^{-1}Q_0] \;= df/dz|^{z=E0}\,/\,k!$$

Suppose $f^{(1)}(z) = [Q_0(H_0 - z)^{-1}Q_0]$. That would agree with the above for k = 1. Then scratch shows we would have

$$f^{(k)}(z) = (k-1)!\; Q_o\,(H_0-z)^{-k}Q_o$$

Now what are our algebra rules for playing with these Q_0 packed operators? For example, consider:

$$[Q_0(H_0 - E_0)^{-1}Q_0]^2 \;= Q_0(H_0 - E_0)^{-1}Q_0\; Q_0(H_0 - E_0)^{-1}Q_0 \;= Q_0(H_0 - E_0)^{-1}Q_0(H_0 - E_0)^{-1}Q_0$$

Now we can use the fact that Q_0 is all of unity except subspace 0, so with crude notation write

$$Q_0 = \Sigma_{k\neq 0}\, |k><k|$$

If we insert this three times on the right above, we find that (this is trivial to show, k are eigenstates of H_0 so you just write it out).

$$[Q_0(H_0 - E_0)^{-1}Q_0]^2 \;=\; \Sigma_{k\neq 0}\, |k>(E_k - E_0)^2<k|$$

Now we wonder if this might the same as $[Q_0(H_0 - E_0)^{-2}Q_0]$. If we do the same thing here, we do in fact get the same result. You can just see this.

What is the general theorem we are showing here? Well, we already know this theorem

$$Q_0\;\; f(H_0)\; Q_0 \;=\; f(H_0)\; Q_0 \;\; = Q_0\;\; f(H_0) \;=\;\; \Sigma_{k\neq 0}\, |k>\, f(E_k)\, <k|$$

because we get $<k|\, f(H_o)|k'> = \delta_{k,k'}\, f(E_k)$. So consider

$$[Q_0\;\; f(H_0)\; Q_0]^k \;=\; Q_0\;\; f(H_0)\; Q_0\; Q_0\;\; f(H_0)\; Q_0\; Q_0\;\; f(H_0)\; Q_0 \ldots .\; Q_0\;\; f(H_0)\; Q_0$$

First, we replace each $Q_0\; Q_0$ by a single Q_0 then when we expand each Q_0, we get the $<k|\, f(H_o)|k'> = \delta_{k,k'}\, f(E_k)$. happening at each point and this we get

$$= \Sigma_{k,k',k''\ldots}\;\; |k><k|\, f(H_0)\, |k'><k'|\, f(H_0)\, |k''><k''|\;\; .....$$

$$= \Sigma_{k,k',k''\ldots}\;\; |k><k|\, f(E_{k'})\, |k'><k'|\, f(E_{k''})\, |k''><k''|\;\; .....$$

$$= \Sigma_{k,k',k''\ldots}\;\; |k>\, \delta_{k,k'}\, f(E_k)\, \delta_{k',k''}\, f(E_{k''})<k''|\;\; .....$$

$$= \Sigma_k\;\; |k>\, f(E_k)^k\, <k|$$

Therefore, we may conclude that $[Q_0\ f(H_0)^k\ Q_0]$ will be exactly this same sum. Thus we find that

$$[Q_0\ f(H_0)\ Q_0]^k\ =\ [Q_0\ f(H_0)^k\ Q_0]$$

and this interestingly says that we can do this:

$$Q_0\ f(H_0)\ Q_0\ Q_0\ f(H_0)\ Q_0\ Q_0\ f(H_0)\ Q_0 \ldots .\ Q_0\ f(H_0)\ Q_0$$

$$= Q_0\ f(H_0)\ Q_0\ f(H_0)\ Q_0\ f(H_0)\ Q_0 \ldots .\ Q_0\ f(H_0)\ Q_0$$

$$= Q_0\ f(H_0)\ 1\ f(H_0)\ 1\ f(H_0)\ 1 \ldots .\ 1\ f(H_0)\ Q_0$$

How do I explain this? Well, we start at the right and say $|\varphi\rangle \equiv Q_o\,|\psi\rangle$ projects out only that portion of ψ that lies in the perp space to $\mathcal{E}_a^0$. Then we know that $H_o\,|\varphi\rangle$ is linear combination of perp space basis functions. That is

$$|\psi\rangle = \Sigma_i\ a_i\ u_i \qquad |\varphi\rangle = \Sigma_{i\neq 0}\ a_i\ u_i \qquad H_o\,|\varphi\rangle = \Sigma_{i\neq 0}\ a_i\ H_o\ u_i\ \ = \Sigma_{i\neq 0}\ a_i\ E_i\ u_i$$

so $H_o\,|\varphi\rangle$ stays in the perp space. Then of course $f(H_0)\,|\varphi\rangle$ is also in the perp space, and that is why all those internal Q_0 are unnecessary. Even the left one is unnecessary as we already know from long ago.

So here is a little way to say this:

$$Q_0A(H_o)\ Q_0\ B(H_o)\ Q_0C(H_o)\ Q_0D(H_o)\ Q_0\ =\ A(H_o)\ B(H_o)\ C(H_o)\ D(H_o)\ Q_0$$

So the main point so far is this:

$$(-Q_{0a})^k\ =\ [Q_0(H_0 - E_0)^{-1}Q_0]^k\ =\ \ Q_0(H_0 - E_0)^{-k}Q_0$$

and so we at least now have some simplification of our "desired result"

$$\Sigma_{i\neq 0}\ P_i/(z\text{-}E_i)\ =\ \Sigma_{k=1}^{\infty}\ (-1)^{k-1}\ (z\text{-}E_0)^{k-1}\ Q_{0a}^k\ = -\ \Sigma_{k=1}^{\infty}\ (z\text{-}E_0)^{k-1}\ (-Q_{0a})^k$$

$$= -\ \Sigma_{n=1}^{\infty}\ (z\text{-}E_0)^{n-1}\ Q_0(H_0 - E_0)^{-n}Q_0$$

Now suppose we just expand both sides. For LHS get

$$\text{LHS} = \Sigma_{k\neq 0}\ P_k/(z\text{-}E_k) = \Sigma_{k\neq 0}\ |k\rangle\ (z\text{-}E_k)^{-1}\langle k|$$

Meanwhile, we know that

$$Q_0(H_0 - E_0)^{-n}Q_0\ =\ \Sigma_{k\neq 0}\ |k\rangle\ (E_k - E_0)^{-n}\ \langle k|$$

Thus, we seem to have

$$\Sigma_{k\neq 0} \mid k\rangle (z-E_k)^{-1}\langle k\mid \;=\; \Sigma_{k\neq 0} \mid k\rangle \{ - \Sigma_{n=1}^{\infty} (z-E_0)^{n-1} (E_k - E_0)^{-n} \} \langle k\mid$$

And this certainly would be true if we could show that

$$(z-E_k)^{-1} = - \Sigma_{n=1}^{\infty} (z-E_0)^{n-1} (E_k - E_0)^{-n} \;=\; -(z-E_0)^{-1} \Sigma_{n=1}^{\infty} (E_k - E_0)^{-n} (z-E_0)^{n}$$

which can be rewritten:

$$(z-E_0)/ (z-E_k) \;=\; -\Sigma_{n=1}^{\infty} (E_k - E_0)^{-n} (z-E_0)^{n}$$

Is this a valid Laurent series? Let

$$f(z) = (z-E_0)/ (z-E_k)$$

$$f^{(1)}(z) = [(z-E_k) - (z-E_0)] / (z-E_k)^2 = (E_o - E_k) (z-E_k)^{-2}$$

$$f^{(2)}(z) = (E_o - E_k) (-2) (z-E_k)^{-3}$$

$$f^{(3)}(z) = (E_o - E_k) (-2)(-3) (z-E_k)^{-4}$$

$$f^{(4)}(z) = (E_o - E_k) (-2)(-3)(-4) (z-E_k)^{-5}$$

$$f^{(n)}(z) = (E_o - E_k) (-1)^{n+1} n! (z-E_k)^{-(n+1)}$$

$$f^{(n)}(E_0) = (E_o - E_k) (-1)^{n+1} n! (E_o-E_k)^{-(n+1)} = (-1)^{n+1} n! (E_o-E_k)^{-n} \qquad n>0$$

$$f^{(0)}(E_0) = f(E_0) = (E_0-E_0)/ (E_o-E_k) = 0 \qquad n=0$$

Then we would expect our Laurent expansion of f(z) to be

$$f(z) = \Sigma_{n=0,\infty} (z-E_0)^{n} f^{(n)}(E_0)/ n!$$

$$(z-E_0)/ (z-E_k) \;=\; \Sigma_{n=1,\infty} (z-E_0)^{n} \;*\; (-1)^{n+1} n! (E_o-E_k)^{-n} \;/n!$$

$$= \Sigma_{n=1,\infty} (-1)^{n+1} \; (z-E_0)^{n} \; (E_o-E_k)^{-n}$$

$$= - \Sigma_{n=1,\infty} (E_k-E_0)^{-n} (z-E_0)^{n}$$

and lo and behold this is in fact our desired Laurent series.

Review of derivation of page 716 A. [Again, skip this and see Kato's Bag in Round Three]

(1) Our first step is to do a Taylor series expansion of this function

$$f(z) = (z-E_0)/(z-E_k)$$

We know the Taylor series, expanding around the point $z-E_o$, is this:

$$f(z) = \Sigma_{n=0,\infty} (z-E_0)^n f^{(n)}(E_0)/ n!$$

By direct computation, we find that

$$f^{(n)}(E_0) = (-1)^{n+1} n! (E_o-E_k)^{-n} \quad \text{for } n = 1,2,3....$$
$$= 0 \quad \text{for } n=0$$

Thus, our Taylor series is this:

$$(z-E_0)/(z-E_k) = -\Sigma_{n=1,\infty} (E_k-E_0)^{-n} (z-E_0)^n$$

which we then rewrite as

$$(z-E_k)^{-1} = -(z-E_0)^{-1} \Sigma_{n=1}^{\infty} (E_k - E_0)^{-n} (z-E_0)^n \qquad (*)$$

We are expanding around the point $z = E_0$. Exactly at this point, the sum on the RHS = 0 and this somehow cancels the pole factor to give the finite factor on the left.

More generally we have merely shown this fact

$$(z-a)/(z-b) = -\Sigma_{n=1,\infty} (b-a)^{-n} (z-a)^n$$

or

$$1/(z-b) = -\Sigma_{n=1,\infty} (b-a)^{-n} (z-a)^{n-1}$$

which is how you would expand the function on the left about the point z = a.

(2) We now take our result (*) derived above

$$(z-E_k)^{-1} = -(z-E_0)^{-1} \Sigma_{n=1}^{\infty} (E_k - E_0)^{-n} (z-E_0)^n \qquad (*)$$

and we use this as a weight function in a projection sum where $k\neq0$ means we are summing over all kets in our Hilbert Space except those in the $\mathcal{E}_a^0$eigenmanifold. Thus:

$$\Sigma_{k\neq0} |k\rangle (z-E_k)^{-1}\langle k| = \Sigma_{k\neq0} |k\rangle \{ -\Sigma_{n=1}^{\infty} (z-E_0)^{n-1} (E_k - E_0)^{-n} \} \langle k|$$

$$= -\Sigma_{n=1}^{\infty} (z-E_0)^{n-1} \Sigma_{k\neq0} |k\rangle (E_k - E_0)^{-n} \langle k|$$

(3) Next, we show these facts to be true (if this is unclear, see notes above in Plan A)

$$(-Q_{0a})^n = [Q_0(H_0 - E_0)^{-1}Q_0]^n = Q_0(H_0 - E_0)^{-n}Q_0 = \Sigma_{k\neq 0} \mid k\rangle (E_k - E_0)^{-n} \langle k\mid$$

Thus, we arrive at this point:

$$\Sigma_{k\neq 0} \mid k\rangle (z-E_k)^{-1}\langle k\mid = -\Sigma_{n=1}^{\infty} (z-E_0)^{n-1}(-Q_{0a})^n = \Sigma_{n=1}^{\infty} (-1)^{n+1}(z-E_0)^{n-1}Q_{0a}{}^n$$

Meanwhile, for the LHS, we know that

$$\Sigma_{k\neq 0} \mid k\rangle (z-E_k)^{-1}\langle k\mid = \Sigma_{k\neq 0} P_k / (z-E_k)$$

but according to (56) page 714 we know that

$$G_0(z) = \Sigma_k P_k / (z-E_k) = P_0 / (z-E_0) + \Sigma_{k\neq 0} P_k / (z-E_k)$$

Therefore we have

$$\Sigma_{k\neq 0} \mid k\rangle (z-E_k)^{-1}\langle k\mid = G_0(z) - P_0 / (z-E_0)$$

and then we have finally

$$G_0(z) - P_0 / (z-E_0) = \Sigma_{n=1}^{\infty} (-1)^{n+1}(z-E_0)^{n-1}Q_{0a}{}^n$$

or

$$G_0(z) = P_0 / (z-E_0) + \Sigma_{n=1}^{\infty} (-1)^{n+1}(z-E_0)^{n-1}Q_{0a}{}^n$$ // which verifies page 716 A

→ This derivation is a pretty big jump for the reader I think, since it took me about **3 hours** to do it.

OK, equations (65) and (66) are obvious given 716 A, so we are now at the pencil line on page 716.

Derive (67) and (68). First we insert our new expansion (66) twice into (64):

$$G_0 = \Sigma_{k=0} (-1)^{k+1}(z-E_0)^{k-1}S^k$$ // our new expansion

Then we have from (68)

$$A^{(n)} = (1/2\pi i) \oint dz \{ \Sigma_{k=0} (-1)^{k+1}(z-E_0)^{k-1}S^k \} V \{\Sigma_{k'=0} (-1)^{k'+1}(z-E_0)^{k'-1}S^{k'}\} \ldots\ldots$$
$$V \{\Sigma_{k"=0} (-1)^{k"+1}(z-E_0)^{k"-1}S^{k"}\}$$

where we have n+1 occurrences of { } and n occurrences of V.

Now, how do we do this contour integral? Recall that it surrounds all the poles for the $\mathcal{E}_a^0$ subspace both before and after they move a little which some λ is applies. But of course when λ=0, the poles are all at the same place, namely, $z = E_0$. This is the case we are now considering.

Let's schematize the above equation like this:

$$A^{(n)} = (1/2\pi i) \oint dy \{ a/y + b + cy + dy^2 + ...\} V \{ a/y + b + cy + dy^2 + ...\} V .... V \{ a/y + b + cy + dy^2 + ...\}$$

Each combination of terms for all the sums gives a certain power of y times a constant. How do we list off the terms which have the form C/y? There are *lots* of them! Here are some of them

a/y * b * b * b * b.....b

a/y * cy * a/y * b * b.....b

a/y * dy^2 * a/y * a/y * b.....b

So now that we see the point, let's reorder our integral:

$$A^{(n)} = \Sigma k_{1=0} \Sigma k_{2=0} .... \Sigma k_{n+1=0} (1/2\pi i) \oint dz (-1)^{k1+k2+..kn+1 + (n+1)}$$

$$(z-E_0)^{k1+k2 +...+kn+1 - (n+1)} \{S^{k1}\} V \{S^{k2}\} V .... V S^{kn+1}$$

Now, where is the simple pole? It requires that

$$k_1 + k_2 + ... + k_{n+1} - n - 1 = - 1$$

or

$$k_1 + k_2 + ... + k_{n+1} = n \qquad \text{// **which is result (68)**}$$

For any set of k_i satisfying this rule, the (-1) power is just (-1), so we get

$$A^{(n)} = - \Sigma k_{1=0} \Sigma k_{2=0} .... \Sigma k_{n+1=0} (\text{rule}) (1/2\pi i) \oint dz (z-E_o)^{-1} \{S^{k1}\} V \{S^{k2}\} V .... V S^{kn+1}$$

$$= - \Sigma_{k1,k2...kn+1}(\text{rule}) \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\} \qquad \text{// **which is result (67)**}$$

and this is result (67). He writes $\Sigma_{k1,k2...kn+1}(\text{rule}) = \Sigma_{(k,n)}$ as a compact notation (I modify it)

Now looking back at (63) we have finally achieved our goal of having "an expansion of P in powers of λ", where P is defined above in bold: it is the projector onto all the subspaces into which $\mathcal{E}_a^0$ breaks when we turn on λ. So we have

$$P = P_0 - \Sigma_{n=1} \lambda^n \Sigma_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$$

Next, we want to look at various terms in this sum! The λ=0 term is just P_0. We then have

$$P,\lambda=0 = P_0$$

$$P,\lambda=1 = -\lambda \sum_{(k,1)} \{S^{k1}\} V \{S^{k2}\} = -\lambda [S^0 V S^1 + S^1 V S^0] = +\lambda [P_0 V Q_{0a} + Q_{0a} V P_0]$$

$$P,\lambda=2 = -\lambda^2 \sum_{(k,2)} \{S^{k1}\} V \{S^{k2}\} V \{S^{k2}\}$$

$$= -\lambda^2 [S^0 V S^0 V S^2 + S^0 V S^2 V S^0 + S^2 V S^0 V S^0 + S^0 V S^1 V S^1 + S^1 V S^0 V S^1 + S^1 V S^1 V S^0]$$

$$= -\lambda^2 [P_0 V P_0 V Q_{0a}^2 + P_0 V Q_{0a}^2 V P_0 + Q_{0a}^2 V P_0 V P_0 - P_0 V Q_{0a} V Q_{0a} - Q_{0a} V P_0 V Q_{0a} - Q_{0a} V Q_{0a} V P_0]$$

$$= +\lambda^2 [(P_0 V Q_{0a} V Q_{0a} + Q_{0a} V P_0 V Q_{0a} + Q_{0a} V Q_{0a} V P_0) -(P_0 V P_0 V Q_{0a}^2 + P_0 V Q_{0a}^2 V P_0 + Q_{0a}^2 V P_0 V P_0)]$$

In this way, I see I could write the terms for each order in λ, very clear.

Derive (70) and (71). We are told to "repeat the process we did for P for HP". We are supposed to start with (59), but not clear where to go next. This would give page 715 C with extra factor of z. If we replaced P by HP in (63) we could get (64) with an extra factor of z. Our equation for the $A^{(n)}$ object would then have this extra z in it:

$$A'^{(n)} = (1/2\pi i) \oint dz\, z \{ \Sigma_{k=0} (-1)^{k+1}(z-E_0)^{k-1} S^k \} V \{\Sigma_{k'=0} (-1)^{k'+1}(z-E_0)^{k'-1} S^{k'}\} ..... V \{\Sigma_{k"=0} (-1)^{k"+1}(z-E_0)^{k"-1} S^{k"}\}$$

so now the coefficient of each power of $(z-E_0)$ is no longer a constant. I think we would then pick up both the single poles and the double poles! Here is the regrouping

$$A'^{(n)} = \Sigma k_{1=0} \Sigma k_{2=0} .... \Sigma k_{n+1=0} (1/2\pi i) \oint dz\, z\, (-1)^{k1+k2+..kn+1 + (n+1)} (z-E_0)^{k1+k2 +...+kn+1 - (n+1)} \{S^{k1}\} V \{S^{k2}\} V .... V S^{kn+1}$$

The contribution from the single pole is the same as last time but we have an extra factor of E_0 from the extra z (which matches change in dimensions going from P to HP). So this set of contributions would yield

$$A'^{(n)} = -E_0 \sum_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$$

The other set of contributions will come from double poles due to the z in the numerator. This requires then that

$$(z\text{-}E_0)^{k1+k2\ +\ldots+kn+1\ -\ (n+1)} = (z\text{-}E_0)^{-2}$$

$$\Rightarrow \qquad k_1 + k_2 + ... + k_{n+1} - n - 1 = -2$$

$$\Rightarrow \qquad k_1 + k_2 + ... + k_{n+1} = n - 1$$

In this case the (-1) power is just +1. So then we have

$$A'^{(n)} = \Sigma_{(k,n-1)}\ (1/2\pi i) \oint dz\ z\ (z\text{-}E_0)^{-2} \{S^{k1}\}\ V\ \{S^{k2}\} V\ ....\ V\ \{S^{kn+1}\}$$

The residue rule creates no new factors, and we now get

$$A'^{(n)} = \Sigma_{(k,n-1)}\ \{S^{k1}\}\ V\ \{S^{k2}\} V\ ....\ V\ \{S^{kn+1}\}$$

So here is our complete result at this point:

$$HP = H_0P_0 + \Sigma_{n=1}\ \lambda^n\ A'^{(n)}$$

where $[HP]_0$ is what you have when $\lambda = 0$, so that must be H_0P_0. And then $A'^{(n)}$ has these two terms:

$$A'^{(n)} = -\ E_0\ \Sigma_{(k,n)}\ \{S^{k1}\}\ V\ \{S^{k2}\} V\ ....\ V\{S^{kn+1}\}$$
$$+\ \Sigma_{(k,n-1)}\ \{S^{k1}\}\ V\ \{S^{k2}\} V\ ....\ V\ \{S^{kn+1}\}$$

What about dimensions? We know that Q_{0a} is 1/E but P_0 and Q_0 have no dimensions. Now consider

$$S_k = -P_0 \qquad k = 0 \qquad \dim = 1$$
$$= Q_{oa}{}^k \qquad k > 0 \qquad \dim = E^{-k}$$

So we can say $\dim(S_k) = E^{-k}$ for all k. Then our terms are like this:

$$E^1 * E^{-(\text{sum of k's})} E^n \quad + \ E^{-(\text{sum of k's})} E^n$$

$$= E^1\ E^{-n} E^n + \ E^{-(n-1)} E^n \qquad \text{// } E^n \text{ from the n factors of V}$$

$$= E\ + E\ = E$$

Looks good. Now we have for LHS of (70)

$$(H\text{-}E_o)P\ =\ HP - E_0P\ = \{\ H_0P_0 + \Sigma_{n=1}\ \lambda^n\ A'^{(n)}\} - E_o\{\ P_0 + \Sigma_{n=1}\ \lambda^n\ A^{(n)}\}$$

$$= (H_0 - E_o)\ P_0\ + \Sigma_{n=1}\ \lambda^n\ A'^{(n)} - E_0\ \Sigma_{n=1}\ \lambda^n\ A^{(n)}$$

$= \Sigma_{n=1} \lambda^n [A'^{(n)} - E_o A^{(n)}]$ // since $(H_0 - E_o) P_0 = 0$

$= \Sigma_{n=1} \lambda^n [[- E_0 \Sigma_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$
$+ \Sigma_{(k,n-1)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}]$

$- [- E_0 \Sigma_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}]]$

$= \Sigma_{n=1} \lambda^n \Sigma_{(k,n-1)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$ // first and third terms cancel

$= \Sigma_{n=1} \lambda^n B^{(n)}$ **// verifies (71) and (70)**

I probably could have gotten there faster, but at least we have the results.

Now as before, let's write out the first few terms,

$(H-E_o)P = \Sigma_{n=1} \lambda^n \Sigma_{(k,n-1)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$

n = 1: $= \lambda [\{S^0\} V \{S^0\}] = \lambda P_0 V P_0$

n=2: $= \lambda^2 [\{S^1\} V \{S^0\} V \{S^0\}$ + permutations

$= \lambda^2 [Q_{0a} V P_0 V P_0 + P_0 V Q_{0a} V P_0 + P_0 V P_0 V Q_{0a}]$

and this verifies (72).

Summary: This was a bunbuster little section, what did we actually do here? We expanded the following four operators in powers of λ:

G P HP $(H-E_o)P$

$G = \Sigma_{n=0} G_o (V G_o)^n$

$P = P_0 - \Sigma_{n=1} \lambda^n \Sigma_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$

$HP = H_0 P_0 + \Sigma_{n=1} \lambda^n A'^{(n)}$

where $A'^{(n)} = - E_0 \Sigma_{(k,n)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$
$+ \Sigma_{(k,n-1)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\}$

$(H-E_o)P = \Sigma_{n=1} \lambda^n \Sigma_{(k,n-1)} \{S^{k1}\} V \{S^{k2}\} V .... V \{S^{kn+1}\} = \Sigma_{n=1} \lambda^n B^{(n)}$

where

$S_k = -P_0 \qquad k = 0$

$= Q_{oa}{}^k \qquad k > 0$

where

G(z) is the full resolvent = 1/(z-H).

H is the full Hamiltonian $H = H_0 + \lambda V$

P is the projector that grabs all states in the $\mathcal{E}_a{}^0$ spectrum when $\lambda = 0$ and also grabs them after they move when $\lambda > 0$, where you have to write some $\mathcal{E}_1 + \mathcal{E}_2 + .. \mathcal{E}_n$ as on page 698.

To the lowest few orders we found that:

$$P = P_0 + \lambda [P_0 V Q_{0a} + Q_{0a} V P_0] + \lambda^2 [(P_0 V Q_{0a} V Q_{0a} + Q_{0a} V P_0 V Q_{0a} + Q_{0a} V Q_{0a} V P_0) - (P_0 V P_0 V Q_{0a}{}^2 + P_0 V Q_{0a}{}^2 V P_0 + Q_{0a}{}^2 V P_0 V P_0)]$$

$$(H-E_o)P = 0 + \lambda P_0 V P_0 + \lambda^2 [Q_{0a} V P_0 V P_0 + P_0 V Q_{0a} V P_0 + P_0 V P_0 V Q_{0a}]$$

I suspect that this last item somehow replaces that famous set of equations Schiff 31.4. And this is a pure operator expansion, there are no wavefunctions anywhere.

17. Calculation of Eigenvalues and Eigenvectors (717).

Case $g_a = 1$.

Here our entire $\mathcal{E}_a{}^0$ space has only one state (because that is what $g_a = 1$ says), and it will move somehow with λ. It ends up in space $\mathcal{E}_a$ which has dimension 1.

[Note: All of Messiah's mysterious statements in the opening paragraph of this section are verified and explained in my Kato Meta notes, section 2B and environs.]

Verify the following sentence: " In this case, the eigenvector of H is $P|0\rangle$ "

Proof: We know that $H|\psi_i\rangle = E_i|\psi_i\rangle$ for our entire problem in $\mathcal{H}$. Consider the expansion

$$|\psi_i\rangle = |0; \psi_i\rangle + \lambda |1; \psi_i\rangle + \lambda^2 |2; \psi_i\rangle \ ...$$

where we show the various state corrections at various λ^n levels. Since the $|\psi_i\rangle$ form a complete basis in $\mathcal{H}$ (not just in $\mathcal{E}_a$ or some other subset of $\mathcal{H}$) , we can certainly expand $|0; \psi_i\rangle$ on this basis:

$$|0; \psi_i\rangle = \Sigma_j \, a_j \, |\psi_j\rangle$$

Now suppose some specific $|\psi_0\rangle$ is our non-degenerate eigenstate in question and it has a projector P which we might really call $P^{(0)}$ since it only passes state $|\psi_0\rangle$. Then we know that

$$P|0;\psi_0\rangle = \Sigma_j a_j P|\psi_j\rangle = \Sigma_j a_j \delta_{0,j}|\psi_j\rangle = a_0|\psi_0\rangle$$

Therefore, since $|\psi_0\rangle$ is an eigenstate of H, $P|0;\psi_0\rangle$ must also be an eigenstate of H. The meaning of the notation $|0\rangle$ is this $|0;\psi_0\rangle$ for our state of interest. It is just a confusing coincidence here that we have used the numeral 0 to label our state ψ_0 and then we talk about the level 0 component of this state.

Now if we ignore normalization, we have a very detailed solution to our problem:

$$\begin{aligned}|\psi_0\rangle = P|0\rangle &= P_0|0\rangle + \lambda[P_0 V Q_{0a} + Q_{0a} V P_0]|0\rangle \\ &+ \lambda^2[(P_0 V Q_{0a} V Q_{0a} + Q_{0a} V P_0 V Q_{0a} + Q_{0a} V Q_{0a} V P_0) \\ &\quad -(P_0 V P_0 V Q_{0a}^2 + P_0 V Q_{0a}^2 V P_0 + Q_{0a}^2 V P_0 V P_0)]|0\rangle\end{aligned}$$

We can simplify slightly using these two rules:

$P_0|0\rangle = |0\rangle$ $\qquad$ $Q_0|0\rangle = 0 = Q_{0a}|0\rangle$ $\qquad$ the perp space operator

$$\begin{aligned}|\psi_o\rangle = P|0\rangle &= |0\rangle + \lambda[Q_{0a} V]|0\rangle \\ &+ \lambda^2[(Q_{0a} V Q_{0a} V) \\ &\quad -(P_0 V Q_{0a}^2 V + Q_{0a}^2 V P_0 V)]|0\rangle\end{aligned}$$

$$= |0\rangle + \lambda Q_{0a} V|0\rangle + \lambda^2[(Q_{0a} V Q_{0a} V) - (P_0 V Q_{0a}^2 V + Q_{0a}^2 V P_0 V)]|0\rangle \mid \ldots$$

This implies that, in Messiah notation,

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

$$|2\rangle = [(Q_{0a} V Q_{0a} V) - (P_0 V Q_{0a}^2 V + Q_{0a}^2 V P_0 V)]|0\rangle$$

[But this is the "Kato series" which has the extra renormalization terms like the middle one seen here. See Kato Meta for an explanation of all this stuff.]

Explain the extra middle term. The first equation appears as (14) on page 680. The second appears as (27) on page 694, but I don't see that we agree. What appears in (27) are the first and third terms, if we say $P_0 = |0\rangle\langle 0|$. Is there some reason that the middle term is 0?

$$P_0 V Q_{0a}^2 V|0\rangle \;=?= 0$$

Let's try to compute this thing: Use $1 = |0\rangle\langle 0| + \Sigma_k |k\rangle\langle k|$ where second sum is over the perp space.

$P_0\ V\ Q_{0a}^2\ V|0\rangle = P_0\ V\ Q_{0a}^2\ \{\ |0\rangle\langle 0| + \Sigma_k\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= P_0\ V\ Q_{0a}^2\ \{\ \Sigma_k\ |k\rangle\langle k|\ \}V\ |0\rangle$ // because Q_0 kills all but the perp space

$= P_0\ V\ Q_0\ (E_a^0 - H_0)^{-2}\ \{\ \Sigma_k\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= \Sigma_k\ P_0\ V\ Q_0\ (E_a^0 - E_k)^{-2}\ \{\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ P_0\ V\ \{\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ P_0\ \{\ |0\rangle\langle 0| + \Sigma_{k'}\ |k'\rangle\langle k'|\ \}V\ \{\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ P_0\ \{\ |0\rangle\langle 0|\ \}V\ \{\ |k\rangle\langle k|\ \}V\ |0\rangle$ // because P_0 kills the perp space

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ \{\ |0\rangle\langle 0|\ \}V\ \{\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ \{\ |0\rangle V_{ok} V_{k0}$

$= \Sigma_k\ (E_a^0 - E_k)^{-2}\ |V_{k0}|^2\ |0\rangle$

$=$ (a number) $|0\rangle$

In passing, we note that $\varepsilon_{2,n} = \langle 0; n, \varepsilon_1|\ V\ |1; n, \varepsilon_1\rangle = \Sigma_k\ (E_a^0 - E_k)^{-1}\ |V_{kn}|^2$, so our "number" is not equal to this energy correction.

So OK, I think I see the light here. This extra term that does not appear in (27) just gives a multiple of $|0\rangle$ and we would interpret this term as $\langle 0|2\rangle$ using the above definition of $|2\rangle$. So this is a question of normalization! If you normalize states as in (27), where you have $\langle 0|2\rangle = 0$, you don't get this kind of term. So I conclude that we are NOT now working with the normalization scheme $\langle 0|n\rangle = 0$!

How might we identify terms that just give renormalization like this? Why doesn't the last term do this same thing? Let's compute it:

$Q_{0a}^2\ V\ P_0\ V|0\rangle = Q_{0a}^2\ V\ P_0\ \{\ |0\rangle\langle 0| + \Sigma_k\ |k\rangle\langle k|\ \}V\ |0\rangle$

$= Q_{0a}^2\ V\ P_0\ \{\ |0\rangle\langle 0|\ \}V\ |0\rangle$ // because P_0 kills off the perp space

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

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

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

$= Q_0 (E_a^0 - H_0)^{-2}\ \{\ \Sigma_k\ |k\rangle\langle k|\ \}V\ \{\ |0\rangle\langle 0|\ \}V\ |0\rangle$

$= \Sigma_k Q_0 (E_a^0 - E_k)^{-2} \{ |k><k| \} V \{ |0><0| \} V |0>$

$= \Sigma_k (E_a^0 - E_k)^{-2} \{ |k> V_{k0} V_{00}$

$= \Sigma_k [(E_a^0 - E_k)^{-2} V_{k0} V_{00}] |k> \quad \neq$ (a number) * $|0>$!!!

It seems that if you end up with a perp space sum close to the left end of things, you won't be getting a normalization situation.

Let's go back to the previous case again

$P_0 V Q_{0a}^2 V|0>$

No matter what we get from $V Q_{0a}^2 V|0>$, the left end P_0 will kill everything except $|0>$!

Theorem: So, any term with a P_0 on the left end is going to involve normalization only.

So, let's go back to our expansion:

$$|2> = [(Q_{0a} V Q_{0a} V) - (P_0 V Q_{0a}^2 V + Q_{0a}^2 V P_0 V)] |0>$$

I know by inspection that $P_0 V Q_{0a}^2 V|0>$ = (number) $|0>$, whereas I know that the other two terms can only give perp space results and cannot therefore contain any $|0>$ component. Thus, if I want to impose the condition $<0|2> = 0$, I just drop this middle term.

$|2'> = |2> + c |0>$ is what our result above says, $|2'>$ be our resolvent section result

OK, let this ride for now. [this issue is reviewed and resolved in Kato Meta Notes section 2A]

The point is that we have in principle an exact expression for the state $|\psi_o>$ *to any order we want*, and this comes from being able to do the operator expansion for P.

$$|\psi_0> = k'|0> + \lambda Q_{0a} V |0> + \lambda^2 [Q_{0a} V Q_{0a} V - Q_{0a}^2 V P_0 V] |0> | + \lambda^3 P_3 |0> + ...$$

Next mystery: " eigenenergy comes from equation $HP = E_a P$ "

This is a pure operator equation. I am not use to a "pure operator eigenvector problem" since there are no vectors here. If we apply this to state $|0>$ we get $H |\psi_o> = E_a |\psi_o>$ so it is at least consistent. So I will allow this one tentatively.

Note added 1.31.09: Well, having reread earlier parts, I think this idea is common. If you apply an operator equation to the state $|0>$ and that produces an EV problem, as in the last paragraph, then you can make the claim quoted above.

Next mystery: $Tr(P) = Tr(P_0) = 1$

Why on earth would this be true, and what size matrices are these? Well, let's say our problem has some dimension N. If we go into the perturbed basis $|\psi_i>$, call it, then for $|\psi_0>$ the matrix P is NxN with a single 1 on the diagonal at the 0,0 entry, and all other elements are 0 (in this basis). So in this basis, Tr(P) = 1. But that is then true in any basis and is independent of basis. A similar argument for P_o in the unperturbed basis says $Tr(P_0) = 1$, so that is true as well in any basis. Fine. This would work fine even for $N=\infty$ because there is only one non-zero diagonal element.

Now apply this to our equation above which says $HP = E_aP$, then $tr(HP) = E_a$!

Next, apply this to (7) to get

$$tr(HP) - E_a^0 = \Sigma_{n=1}\lambda^n tr(B^{(n)}) = E_a - E_a^0$$

Wow, we suddenly have a full expansion for the energy correction, and we learn that

ε_n = the correction at the λ^n level $= tr(B^{(n)})$

Next mystery: $Tr(P_0A) = <0|A|0>$

Write out the trace like so, in the unperturbed basis,

$$Tr(P_0A) = (P_0A)_{ii} = (P_0)_{ij}A_{ji}$$

But $P_0 = |0><0|$ so $(P_0)_{ij} = <i|0><0|j>$

Then

$$(P_0)_{ij}A_{ji} = <i|0><0|j>A_{ji} = <0|j>A_{ji}<i|0> = <0|j><j|A|i><i|0> = <0|A|0>$$ QED

Maybe this is faster

$$Tr(P_0A) = Tr(|0><0|A) = <i|0><0|A|i> = <0|A|i><i|0> = <0|A|0>$$

Now lets look at

$$E_a - E_a^0 = \Sigma_{n=1}\lambda^n tr(B^{(n)}) = \Sigma_{n=1}\lambda^n tr[(\Sigma_{(k,n-1)}) \{S^{k1}\} V \{S^{k2}\}V .... V \{S^{kn+1}\}]$$

$$0 + \lambda tr(P_0 V P_0) + \lambda^2 tr[Q_{0a} V P_0 V P_0 + P_0 V Q_{0a} V P_0 + P_0 V P_0 V Q_{0a}] + \lambda^3 ...$$

The lowest order has

$$tr(P_0 V P_0) = tr(P_0^2 V) = tr(P_0 V) = <0|V|0> = \varepsilon_1$$ // our very familiar first order result

For second order we can write:

$\varepsilon_2 = tr\,[Q_{0a}\,V\,P_0\,V\,P_0 + P_0\,V\,Q_{0a}\,V\,P_0 + P_0\,V\,P_0\,V\,Q_{0a}\,]$

$tr\,[P_0\,Q_{0a}\,V\,P_0\,V + P_0\,V\,Q_{0a}\,V\,P_0 + P_0\,V\,P_0\,V\,Q_{0a}\,]$

$= tr\,[\,P_o\,\{\,Q_{0a}\,V\,P_0\,V + V\,Q_{0a}\,V\,P_0 + V\,P_0\,V\,Q_{0a}\,\}\,]$

$= \langle 0|\,Q_{0a}\,V\,P_0\,V + V\,Q_{0a}\,V\,P_0 + V\,P_0\,V\,Q_{0a}|\,0\rangle$

$= \langle 0|\,V\,Q_{0a}\,V|\,0\rangle \;\; = \Sigma_{k\neq 0}\,\langle 0|\,V\,|\,k\rangle\,(E_k - E_0)^{-1}\,\langle k|\;V|\,0\rangle$ // since Q_0 kills off $|0\rangle$

Maybe I can write down the general result! We always need this

$tr(B^{(n)}) = tr\,[\;(\Sigma_{(k,n-1)})\,\{S^{k1}\}\,V\;\{S^{k2}\}V\;\ldots.\,V\,\{S^{kn+1}\}\,]$

What terms are going to survive our process above? He claims that each term in the above contains P_0 somewhere. We have n+1 exponents and they have to add up to n-1. If they were all ≥ 1, the sum would end up being $\geq$ n+1 . So at least some of them have to have the value $k_i = P_0$, I agree. So P_0 appears somewhere in every contributing term. We then think of this term as

$tr(P_0AB) = tr(ABP_0) = tr(BP_0A) = \langle\varphi|AB|\varphi\rangle$

So imagine processing each term in this way. I call this step above "**knocking out** a P_0". Then if we have a Q_{0a} power on either end of a term after doing the knockout, the term is 0 !

So OK, let's now try n = 3. The exponents have to add up to n-1 = 2.

$tr(B^{(3)}),\,0002 = tr\,[\;\{S^0\}V\,\{S^0\}V\{S^0\}\,V\,\{S^2\}\;\;+\text{permutations}]$

Notice that this form does not have cyclic symmetry! We cannot say for example 0002 = 2000. The reason is that 0002 means 0v0v0v2 and if you rotate this, you get 20v0v0v $\neq$ 2v0v0v0. On the other hand, we do have reverse-ordering symmetry, so that 0002 = 2000 since 0v0v0v2 = 2v0v0v0.

Which terms contribute?

0002	no	$S^2 = (Q_{0a})^2$ and kills against $\|0\rangle$
0020	yes	
0200	yes	// and same as 0020 by reversal
2000	no	$S^2 = (Q_{0a})^2$ and kills against $\langle 0\|$

So only terms with S^2 on the interior can contribute! So we get

$tr(B^{(3)}),\,0002\text{ class} = tr\,[\;\{S^0\}V\,\{S^2\}V\{S^0\}\,V\,\{S^0\}\;+\;\{S^0\}V\,\{S^0\}V\{S^2\}\,V\,\{S^0\}\;]$

$= 2\,tr\,[\;\{S^0\}V\,\{S^2\}V\{S^0\}\,V\,\{S^0\}$ // reverse order rule

$= -2\ \mathrm{tr}\ [\ \{P_0\} V\ \{ Q_{0a}^2 \} V \{ P_0 \}\ V\ \{ P_0 \}$

$= -2 \langle 0 |\ V\ \{ Q_{0a}^2 \} V \{ P_0 \}\ V\ |0\rangle\ =\ -2 \langle 0 |\ V\ Q_{0a}^2 V\ P_0 V\ |0\rangle$

$= -2 \langle 0 |\ V\ P_0 V Q_{0a}^2 V\ |0\rangle$ // if we reverse the other term

We can replace any residual interior P_0 with $|0\rangle\langle 0|$ to simplify the matrix element into a product of two simpler ones, so the above becomes

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

This agrees with the second term in (74) except for the factor of 2. But see below!

Next, consider

$\mathrm{tr}(B^{(3)}),\ 0011\ = \mathrm{tr}\ [\ \{S^0\} V\ \{S^0\} V \{S^1\}\ V\ \{S^1\}\ +\ \text{permutations}\ ]$

0011	no
0101	no
0110	yes
1001	**yes, see below**
1010	no
1100	no

First, here is the 0110 term:

$\mathrm{tr}\ [\ \{S^0\} V\ \{S^1\} V \{S^1\}\ V\ \{S^0\}\ ] = \mathrm{tr}(P_0 V Q_{oa} V Q_{oa} V\ P_0)\ =\ \langle 0|\ V Q_{oa} V Q_{oa} V\ P_0\ |0\rangle$

$= \langle 0|\ V Q_{oa} V Q_{oa} V\ |0\rangle$

and this is the first term in (74). Now consider the 1001 term

$\mathrm{tr}\ [\ \{S^1\} V\ \{S^0\} V \{S^0\}\ V\ \{S^1\}\ ] = \mathrm{tr}(Q_{oa} V P_0 V P_0 V\ Q_{oa})\ =\ \langle 0|\ V P_0 V\ Q_{oa}\ Q_{oa} V |0\rangle$

$= \langle 0|\ V P_0 V Q_{oa}^2 V\ |0\rangle\ = \langle 0|\ V\ Q_{oa}^2 V\ P_0 V\ |0\rangle$

where again the two forms shown are equivalent, which is obvious if you insert $P_0 = |\varphi\rangle\langle\varphi|$.

If we now add up all terms, we get

$\mathrm{tr}(B^{(3)}),\ 2000\ \text{class} = -2 \langle 0 |\ V\ \{ Q_{0a}^2 \} V \{ P_0 \}\ V\ |0\rangle\ =\ -2 \langle 0 |\ V\ Q_{0a}^2 V\ P_0 V\ |0\rangle$

$\mathrm{tr}(B^{(3)}),\ 1100\ \text{class} = \langle 0|\ V Q_{oa} V Q_{oa} V\ |0\rangle\ +\ \langle 0|\ V\ Q_{oa}^2 V\ P_0 V\ |0\rangle$

We now see that one of the 1100 class terms cancels one of the 2000 terms and our result is then

$$\mathrm{tr}(B^{(3)}) = \langle 0| V Q_{oa} V Q_{oa} V |0\rangle - \langle 0| V Q_{oa}^2 V P_0 V |0\rangle$$

and this agrees with Messiah page 717 (74).

Now you might ask: how do you compute something like $\langle 0 | V Q_{0a}^2 V |0\rangle$? I suppose you just do it this way:

$$\langle 0 | V Q_{0a}^2 V |0\rangle = \Sigma_{k',k''} \langle 0|V|k'\rangle\langle k'| Q_{0a}^2 |k''\rangle\langle k''|V |0\rangle$$

But then

$$Q_{0a}^2 = Q_0(H_0 - E_0)^{-2} Q_0 = \Sigma_{k\neq 0} | k\rangle (E_k - E_0)^{-2} \langle k|$$

so then have

$$\langle 0 | V Q_{0a}^2 V |0\rangle = \Sigma_{k',k''} \langle 0 | V|k'\rangle \langle k'| [\Sigma_{k\neq 0} | k\rangle (E_k - E_0)^{-2} \langle k|] |k''\rangle\langle k''|V |0\rangle$$

$$= \Sigma_{k',k''} \Sigma_{k\neq 0} \langle 0 | V|k'\rangle \delta_{k',k} (E_k - E_0)^{-2} \delta_{k,k''} \langle k''|V |0\rangle$$

$$= \Sigma_{k\neq 0} \langle 0 |V|k\rangle (E_k - E_0)^{-2} \langle k|V |0\rangle$$

Easier to start over:

$$\langle 0 | V Q_{0a}^2 V |0\rangle = \Sigma_{k\neq 0} \langle 0 | V | k\rangle (E_k - E_0)^{-2} \langle k| V |0\rangle$$

So here would be the full 3rd order energy correction:

$$\varepsilon_3 = \langle 0| V Q_{oa} V Q_{oa} V |0\rangle - 2 \langle 0 | V Q_{0a}^2 V |0\rangle\langle 0| V |0\rangle$$

$$= \Sigma_{k\neq 0} \Sigma_{k'\neq 0} \langle 0| V | k\rangle (E_k - E_0)^{-1} \langle k|V | k'\rangle (E_{k'} - E_0)^{-1} \langle k'|V |0\rangle$$

$$- 2 \langle 0| V |0\rangle \Sigma_{k\neq 0} \langle 0 | V | k\rangle (E_k - E_0)^{-2} \langle k| V |0\rangle$$

where the first term involves a sort of triple communication channel from our state back to itself.

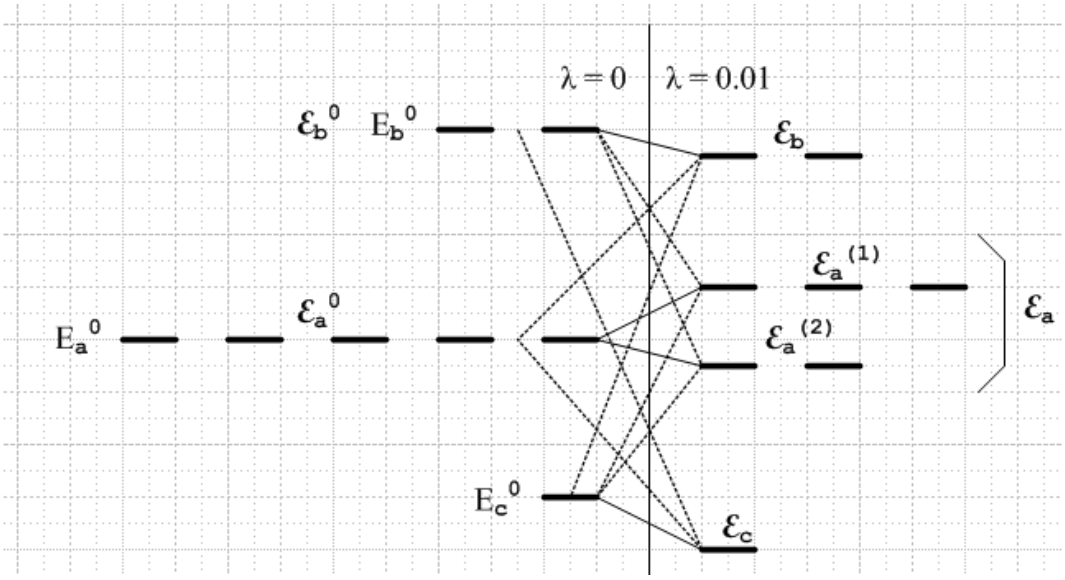
So OK, I think everything is well-defined here, albeit not easy to compute in practice.

Now we are ready for the Real McCoy:

Case $g_a > 1$.

Preliminaries.

Question: What does Messiah mean by the notations $\mathcal{E}_a$ and $\mathcal{E}_a^0$? He never defines $\mathcal{E}_a$ as far as I can see, anywhere in this chapter. Here is my interpretation, where I show a "perturbation situation"



I know from my general theory that

$$\psi_{mi} = u_{mi} + \Sigma_{j=1,N}\, B_{ij}\, u_{mj} + \Sigma_{k\neq m}\, B_{ik}\, u_k$$

that a perturbed state picks up pieces from *all* eigenstates, not just from eigenstates within its initial degenerate group. But for very small λ, we know that the contributions of the local degenerate group are going to be larger, and I show these contributions by solid lines. The dashed lines represent small contributions. The above picture shows my interpretations of $\mathcal{E}_a$ and $\mathcal{E}_a^0$. On page 718 Messiah says that $\mathcal{E}_a$ and $\mathcal{E}_a^0$ "have the same number of dimensions". In my picture above that number would be 5. But we cannot claim that one of the 5 states in $\mathcal{E}_a$ is a linear combination *only* of states from $\mathcal{E}_a^0$ due to the dotted line contributions. Regardless, I think we can regard the 5 states of $\mathcal{E}_a$ as being a Hilbert Space spanned by those 5 vectors (heavy dash = state = vector). The states on the right side are eigenstates of the full H, while those on the left are eigenstates of H_0.

Question: What is the meaning of the various "projection operators" like P = P_a and P_0 = P_{a0} with respect to our picture? Here are some examples:

P_a | any of the states in $\mathcal{E}_b$> = 0
P_a | any of the states in $\mathcal{E}_a$> = | exactly that state in $\mathcal{E}_a$>

P_{a0} | any of the states in $\mathcal{E}_b^0$> = 0
P_{a0} | any of the states in $\mathcal{E}_a^0$> = | exactly that state in $\mathcal{E}_a^0$>

Now he makes a statement which I have a lot of trouble with, so I will quote it:

"Let us suppose that any vector in $\mathcal{E}_a$ may be considered as the *projection* in $\mathcal{E}_a$ of a certain well defined vector of $\mathcal{E}_a^0$."

Well, I need maybe to build up to this with some claims of my own. In the above picture, if I pick one of the $\mathcal{E}_a$ states and apply P_{0a} to it, I know I will get some linear combination of the 5 states in $\mathcal{E}_a^0$. I could also apply P_{0b} and get the smaller contribution from states in $\mathcal{E}_b^0$. Think

$$\psi_{mi} = u_{mi} + \Sigma_{j=1,N}\, B_{ij}\, u_{mj} + \Sigma_{k\neq m}\, B_{ik}\, u_k$$

So, certainly if you apply P_{0a} to a specific state in $\mathcal{E}_a$, you will get a unique and specific state in $\mathcal{E}_a^0$.

Question: what are the characteristics of an operator like P_{0a}? We know that $P_{0a}^2 = 1$. It's domain is all of the Hilbert Space H_0 or H. But its range is the space $\mathcal{E}_a^0$. I am claiming that the spaces H and H_0 are the same space in effect, and we just have different base states comparing $\lambda = 0$ to $\lambda = \lambda_1$. So we can say this:

$P_{a0}: H \rightarrow \mathcal{E}_a^0$ or $P_{a0}: H_0 \rightarrow \mathcal{E}_a^0$

You can see that in general P_{a0} has a huge nullspace which agrees with the partial range. So P_{a0} is certainly not 1-to-1 on H_0, and thus P_{a0}^{-1} does not exist.

But then we could consider the following *restriction* of P_{a0}:

$P_{a0}: \mathcal{E}_a \rightarrow \mathcal{E}_a^0$ $\qquad\qquad$ $P_{a0}^{-1}: \mathcal{E}_a^0 \rightarrow \mathcal{E}_a$

If we define this operator to be so restricted, it might be 1-to-1 . I know that going in the direction shown, for each state in $\mathcal{E}_a$ you get exactly one unique state in $\mathcal{E}_a^0$, so it is at least a function. It has no nullspace as so defined, so I think the inverse mapping exists. I might then write that as on the right above. I now conjecture that $P_{a0}^{-1} = P_a$. So let's first talk about P_a a bit. I copy from above and edit. I think there is no question but that any state in $\mathcal{E}_a^0$ is a mixture of ALL states on the right side, not just those in $\mathcal{E}_a$.

So, certainly if you apply projector P_a to a specific state in $\mathcal{E}_a^0$, you will get a unique and specific state in $\mathcal{E}_a$. Question: what are the characteristics of an operator like P_a? We know that $P_a^2 = 1$. It's domain is all of the Hilbert Space H_0 or H. But its range is the space $\mathcal{E}_a$. So we can say this:

$P_a: H \rightarrow \mathcal{E}_a$ or $P_a: H_0 \rightarrow \mathcal{E}_a$

You can see that in general P_a has a huge nullspace which agrees with the partial range. So P_a is certainly not 1-to-1 on H. We could consider the following *restriction* of P_a:

$P_a: \mathcal{E}_a^0 \rightarrow \mathcal{E}_a$ $\qquad\qquad$ $P_a^{-1}: \mathcal{E}_a \rightarrow \mathcal{E}_a^0$

Now I will let my conjecture that $P_{a0}^{-1} = P_a$ for the *restricted* operators ride, I think it is exactly correct. Now let's go back to Messiah's sentence above which we still need to get happy with:

"Let us suppose that any vector in $\mathcal{E}_a$ may be considered as the *projection* in $\mathcal{E}_a$ of a certain well defined vector of $\mathcal{E}_a^0$."

This then would go with my restricted P_a : $\mathcal{E}_a^0 \rightarrow \mathcal{E}_a$. OK, and we could also make the reverse statement, and it would involve P_{a0} : $\mathcal{E}_a \rightarrow \mathcal{E}_a^0$. I am now happy with the entire large paragraph at the start of page 718.

Next Messiah sentence is this:

" Thus any eigenvector of H in $\mathcal{E}_a$ can be put into the form $P_a \mid E_a^0\alpha\rangle$ "

Well, in my picture all the states on the right are meant to be eigenstates of H. And I just said that any state on the right in $\mathcal{E}_a$ is the image via P_a of some unique state in $\mathcal{E}_a^0$. To find this unique state, we could just write out this eigenstate of H in the H_o basis and look at just the contributions from $\mathcal{E}_a^0$ states, the first terms in my usual ψ expansion, and that combination of states would be that unique state which Messiah is going to call $\mid E_a^0\alpha\rangle$. I guess we just regard α as the label on this unique state. *Note that* this state $\mid E_a^0\alpha\rangle$ is not itself an eigenstate of H, though it is certainly an eigenstate of H_0 since it is a mixture of degenerate states all with the same energy E_a^0. So I agree that, for any eigenstate $|\varphi\rangle$ in $\mathcal{E}_a$, we could find the unique state $\mid E_a^0\alpha\rangle$ in $\mathcal{E}_a^0$ such that $|\varphi\rangle = P_a \mid E_a^0\alpha\rangle$. Happy again!

Comment: I just wonder what the typical reader of this section made of it!

Then he says $H \{ P|E_a^0\alpha\rangle \} = E_\alpha \{ P|E_a^0\alpha\rangle \}$ where H is the full H and E_α is the fully corrected energy, etc. This is just fine, since we just said $|\varphi\rangle = P|E_a^0\alpha\rangle$ is an eigenstate of H. Both sides of this equation are the same ket in $\mathcal{E}_a$, easier perhaps to access this on the RHS. So here goes the processing:

$H\, P|E_a^0\alpha\rangle = E_\alpha\, P|E_a^0\alpha\rangle$ // $|E_a^0\alpha\rangle$ is the back image of an $\mathcal{E}_a$ eigenstate

$P_0 H\, P|E_a^0\alpha\rangle = E_\alpha\, P_0 P|E_a^0\alpha\rangle$

In this last step, we take our H eigenstate state $|\varphi\rangle = P_a \mid E_a^0\alpha\rangle$ and apply P_0 to it. I would have thought this would give you $|E_a^0\alpha\rangle$ with my idea above that $P_{a0}^{-1} = P_a$ *for the restricted operators*. In any event, applying it to both sides of an equality certainly gives you another equality. Now one more process

$P_0 H\, P\, P_0|E_a^0\alpha\rangle = E_\alpha\, P_0 P\, P_0|E_a^0\alpha\rangle$

$H_a\, |E_a^0\alpha\rangle = E_\alpha\, K_a\, |E_a^0\alpha\rangle$ $\qquad H_a = P_0 H\, P\, P_0$ $\qquad K_a = P_0 P P_0$

where we extracted P_0 from the $\mathcal{E}_a^0$ state on both sides. No problem, and the new operators are as defined by comparison.

Messiah now wants to talk about these operators acting on $\mathcal{E}_a^0$. We can see that if you act on a ket in H^0 outside this subspace, you will get 0 due to the right-side P_0 , and the same is true if you act on a bra outside $\mathcal{E}_a^0$. So truly these operators really only act in $\mathcal{E}_a^0$.

Theorem 1: Any projection operator is a "part of unity" and you can think of it as having a partial diagonal of 1's in some region. It is therefore symmetric and Hermitian in that basis. The Hermiticity will survive any unitary transformation of basis. So projection operators are Hermitian.

Theorem 2: Projection operators have positive matrix elements.

Proof: Let $\psi = \Sigma_i a_i u_i$ $\quad < \psi \mid P_j \mid \psi > = \Sigma_i a_i^* u_i^\dagger a_j u_j = |a_j|^2$

This is probability of being in state j. So a projector is positive definite. This would be true also if we had a projector which projected out a set of base states. I assume orthonormal basis.

Theorem 3: The operator $K_a = P_0 P P_0$ is **positive definite**.

Lemma 3: If you can write $A = BB^\dagger$, then A is positive definite.

Proof of lemma: $A\psi = \lambda\psi$ then $\lambda \|\psi\|^2 = <\psi|A\psi> = <\psi| BB^\dagger|\psi> = <B^\dagger\psi|B^\dagger\psi> = \|B^\dagger\psi\|^2$,
so $\lambda = \|B^\dagger\psi\|^2 / \|\psi\|^2 > 0$. Could we have $B^\dagger\psi = 0$? Yes, I suppose that is possible, but somehow we I think that does not happen in our application below. This is a loose bolt.

Proof of theorem: $K_a = P_0 P P_0 = P_0 P P P_0 = (P_0 P)(P_0 P)^\dagger = BB^\dagger$, QED.
This would be true for any K = CPC where C was Hermitian and P a projection operator

Now back to

$$H_a |E_a^0 \alpha> = E_\alpha K_a |E_a^0 \alpha> \qquad H_a = P_0 H P P_0 \qquad K_a = P_0 P P_0$$

Write this as

$$(H_a - E_\alpha K_a) |E_a^0 \alpha> = 0$$

We expect this might have non-zero solutions provided that

$$\det(H_a - E_\alpha K_a) = 0$$

Now our amazing claim: THIS secular equation determines the eigenvalues E_α for our states in $\mathcal{E}_a$, to all orders! If you can solve this "**generalized eigenvalue problem**" also for the eigenvectors that go with these eigenvalues, then you have found eigenvectors $|E_a^0 \alpha>$ of this equation, and THEN you know that $P |E_a^0 \alpha>$ are the eigenvectors of H that go with these same E_α eigenvalues. This seems amazingly convoluted.

Comments on the Generalized Eigenvalue Problem.

(1) Here is a Wolfram claim:

"The determinant of a positive definite matrix is always positive, so a positive definite matrix is always nonsingular."

Well in particular, if a matrix is Hermitian we know we can diagonalize it with a unitary similarity, and the diagonal elements (eigenvalues) will all be positive, so the determinant will be positive. And the determinant is invariant under similarity, so certainly a positive definite Hermitian matrix is non-singular because the determinant cannot be 0. It can therefore be inverted.

(2) We know from above that K is positive definite, so we may then conclude that K^{-1} exists. At first this seems odd since $K = P_0PP_0$ which is a product of three non-invertible projection operators, but the triplet shown must be invertible.

(3) Then our *generalized* eigenvalue problem can be reduced to a *regular* eigenvalue problem

$H\psi = \lambda K\psi \quad \Rightarrow \quad (K^{-1}H)\,\psi = \lambda\psi$

(4) The generalized eigenvalue problem is in itself a whole field of study. From

http://www.cs.umu.se/~dacke/foweb00/new/NGEP/index.html

we have this little introduction, including the new term "**matrix pencil**".

Introduction

In the generalized eigenvalue problem $Ax = \lambda Bx$, the scalar λ is called the eigenvalue and x is the associated eigenvector. The matrix $A - \lambda B$ is called a matrix pencil and is often denoted by the matrix pair (A, B). The eigenvalue λ is a root of the characteristic equation $\det(A - \lambda B) = 0$.

If $A, B \in \mathcal{R}^{n\times n}$ and $\det(A - \lambda B)$ is not identically zero, the pencil is called regular. For a nonsingular B matrix the eigenvalues of a regular matrix pair are in theory the same as those of $B^{-1}A$ (or $AB^{-1} = B(B^{-1}A)B^{-1}$), since, from $Ax = \lambda Bx$ we have $B^{-1}Ax = \lambda x$. Notice, for $B = I$ the generalized problem reduces to the standard eigenvalue problem $Ax = \lambda x$. If B is singular the generalized eigenvalue problem has $r < n$ finite eigenvalues. The remaining n-r eigenvalues are, by convention, set to be infinity. Matrix pencils where $\det(A - \lambda B) = 0$ for any λ is called singular and any number λ can be an eigenvalue of (A, B).

we get this little introduction. The page goes on to show how you would solve the problem numerically.

The idea at this point is then to insert our perturbation expansions for HP and P to get H_a and K_a to some definite order, then you can find E_α to that order. Recall from earlier that we found

$$P = P_0 + \lambda [P_0 V Q_{0a} + Q_{0a} V P_0] + \lambda^2 [(P_0 V Q_{0a} V Q_{0a} + Q_{0a} V P_0 V Q_{0a} + Q_{0a} V Q_{0a} V P_0) - (P_0 V P_0 V Q_{0a}^2 + P_0 V Q_{0a}^2 V P_0 + Q_{0a}^2 V P_0 V P_0)]$$

$$(H-E_o)P = 0 + \lambda P_0 V P_0 + \lambda^2 [Q_{0a} V P_0 V P_0 + P_0 V Q_{0a} V P_0 + P_0 V P_0 V Q_{0a}]$$

Now when we compute $H_a = P_0 \; HP \; P_0$ or $K_a = P_0 \; P \; P_0$, we can throw out terms where we will get $Q_{oa}P_o$ adjacent anywhere, so that really means we only keep terms that begin and end with P_0 and then we don't have to replicate these leading and trailing values. So here we go:

$$K_a = P_0 \; P \; P_0 = P_0 - \lambda^2 P_0 V Q_{0a}^2 V P_0$$

$$H_a = P_0 \; HP \; P_0 = P_0 \{ (H-E_o)P \} P_0 + E_o P_0 \; P \; P_0 = P_0 \{ (H-E_o)P \} P_0 + E_o K_a$$

$$= E_o K_a + \lambda P_0 V P_0 + \lambda^2 P_0 \; V Q_{0a} V P_0$$

and so we have verified (77) and (78).

So let's write our secular equation through order λ^2:

$$\det(H_a - E_\alpha K_a) = 0$$

$$\det[\; E_o K_a + \lambda P_0 V P_0 + \lambda^2 P_0 \; V Q_{0a} V P_0 - E_\alpha K_a] = 0$$

$$\det [\lambda P_0 V P_0 + \lambda^2 P_0 \; V Q_{0a} V P_0 - (E_\alpha - E_o)K_a \;] = 0$$

$$\det [\lambda P_0 V P_0 + \lambda^2 P_0 \; V Q_{0a} V P_0 - (E_\alpha - E_o)\{ P_0 - \lambda^2 P_0 V Q_{0a}^2 V P_0\} \;] = 0$$

I would guess that since we are in our $\mathcal{E}_a^0$ space, we can now just throw out all the P_0 things:

$$\det [\lambda V + \lambda^2 V Q_{0a} V - (E_\alpha - E_o)\{ 1 - \lambda^2 V Q_{0a}^2 V\} \;] = 0$$

$$\det [\lambda V + \lambda^2 V Q_{0a} V - (E_\alpha - E_o)\{ 1 - \lambda^2 V Q_{0a}^2 V\} \;] = 0$$

If we keep only λ^0 terms, we get $(E_\alpha - E_o)\det(1) = 0$ so that $E_\alpha = E_o$.

If we keep only λ^0 and λ^1 terms, this says

$$\det [\lambda V - (E_\alpha - E_o) \, 1] = 0$$

Of course the size of this determinant is the size of the space $\mathcal{E}_a^0$. I could write this out as

$$\det = \varepsilon_{ijk} \ldots. \ [\lambda V_{1i} - (E_\alpha - E_o)\,\delta_{1i}]\,[\lambda V_{2j} - (E_\alpha - E_o)\,\delta_{2j}]\,[\lambda V_{3k} - (E_\alpha - E_o)\,\delta_{3k}] \ \ldots$$

Let's do the 2x2 case please:

$$\det = \ [\lambda V_{11} - (E_\alpha - E_o)\ ]\,[\lambda V_{22} - (E_\alpha - E_o)\ ] \ - \lambda V_{12}\,\lambda V_{21} \ = 0$$

Absorb λ into V and set $\varepsilon_\alpha = \ (E_\alpha - E_o)$ to get

$$\det = \ [V_{11} - \varepsilon_\alpha\ ]\,[V_{22} - \varepsilon_\alpha\ ] \ - V_{12}\,V_{21} \ = 0$$
$$\det = \ [\varepsilon_\alpha - V_{11}]\,[\varepsilon_\alpha - V_{22}] \ - V_{12}\,V_{21} \ = 0$$

This is a quadratic equation for ε_α and

$$\varepsilon_\alpha^2 - (V_{11} + V_{22})\varepsilon_\alpha \ + (V_{11}\,V_{22} - V_{12}\,V_{21}) = 0$$

which looks pretty much like **Saxon page 205**. Notice that

$$B^2 - 4AC \ = (V_{11} + V_{22})^2 - 4 \ (V_{11}\,V_{22} - V_{12}\,V_{21}) \ = (V_{11} - V_{22})^2 + 4\,V_{12}\,V_{21}$$

So the answer is this

$$\varepsilon_\alpha \ = \ (V_{11} + V_{22})/2 \ \pm \ [(V_{11} - V_{22})^2/4 + V_{12}\,V_{21}]^{\,1/2}$$

If you happen to have a problem where $V_{11} = V_{22} = 0$, then the result is this

$$\varepsilon_\alpha \ = \ \pm \ \lambda\,[V_{12}\,V_{21}]^{\,1/2}$$

and this is the famous result that the two states are pushed apart, equally up and down. They "repel" each other.

Now it could happen that all four elements V_{ij} in our 2x2 subspace are 0. In this case and this case only we get $B^2 - 4AC = 0$ so there is no splitting at all, and no shift either. This is where we would have to go to higher order and that will involve the effect of distant matrix elements.

So go back to

$$\det\,[V + V\,Q_{0a}\,V \ - \varepsilon_\alpha\,\{\,1 - V\,Q_{0a}^2\,V\}\ \ ] = 0$$

where again I have embedded the λ into the B . Recall now that

$$Q_{0a}^n = Q_0(H_0 - E_0)^{-n}Q_0 \ = \ \Sigma_{k\neq 0}\,|\,k\rangle\,(E_k - E_0)^{-n}\,\langle k|$$

so to solve for the energies at this level, you have to go off and first compute

$$A_{ij} \ = \Sigma_{k\neq 0}\,\langle i|\,V\,|\,k\rangle\,(E_k - E_0)^{-1}\,\langle k|V|j\rangle$$

$$B_{ij} = \Sigma_{k \neq 0} \langle i| V | k\rangle (E_k - E_0)^{-2} \langle k|V|j\rangle$$

Then we have

$$\det [V + A - \varepsilon_\alpha \{ 1 - B\}] = 0$$

and this is another messy quadratic in ε_α . In our λ^1 situation, we have A = B = 0 and we recover our previous equation which was

$$\det [V - \varepsilon_\alpha 1] = 0$$

Conclusion: Suppose we have a degenerate subspace $\mathcal{E}_a^0$ of dimension N. When we turn on the perturbation, we can compute the corrected energies through second order as follows:

$$\det [V + V Q_{0a} V - \varepsilon_\alpha \{ 1 - V Q_{0a}^2 V\}] = 0 \qquad \varepsilon_\alpha = (E_\alpha - E_o)$$

[I think we can discard the $V Q_{0a}^2 V$ factor on the grounds that it is down one power of λ.]

and again the λ are absorbed into the V. This is an NxN determinant. What is involved in the computation is the matrix elements V_{ij} within $\mathcal{E}_a^0$ and also V_{ik} where k is outside the subspace, and we also need to know the E_k outside the subspace. This solves the problem completely for any N !!! Only when you actually compute things will you know if there is still some degeneracy left when λ is turned on. I should have said: The above det = 0 equation is a polynomial of degree N in ε_α and so there will be N roots. Degeneracy will still exist if there is some algebraic multiplicity > 1 for a root! Very good!

Now back for a moment to this EV problem

$$(H_a - E_\alpha K_a) |E_a^0 \alpha\rangle = 0$$

Suppose we have found the eigenvalues E_α to some order, how do we then solve this thing?

$$(H - \lambda_\alpha K)v_\alpha = 0 \qquad \text{within } \mathcal{E}_a^0$$

Normally if there were an inverse for K, you would change this to

$$(K^{-1}H - \lambda_\alpha)v_\alpha = 0 \qquad \text{within } \mathcal{E}_a^0$$

and then it is a standard problem. For a given λ_α, there are N unknowns to find. But we have N equations: I guess this is the same with the original equation:

$$(H - \lambda_\alpha K)v_\alpha = 0 \qquad \text{within } \mathcal{E}_a^0$$

$$\Sigma_j (H - \lambda_\alpha K)_{ij} (v_\alpha)_j = 0 \qquad i = 1,2....N \qquad \text{so N equations each in N unknowns.}$$

How do we know there are solutions? Well, this problem really imposes that fact somehow. But more generally, if both are Hermitian, you can probably show it.

All of the above is the Kato method of 1949, reference is given p 712. Messiah concludes this section with a slight variation due to **Bloch 1958** which brings in yet another operator $\mathcal{U}$ and the claim is that the expansions are a bit simpler. But not enough detail is given here to follow the work. For example, it is not obvious why you can "define" an operator $\mathcal{U}$ as shown in (79). Bloch must talk about this. Here is some information on the reference, etats lies means bound states. [I see that Messiah was a friend of this Bloch at Saclay and write Bloch's obit in 1972] Notice his mention of "diagrams".

Sur la théorie des perturbations des états liés

Claude Bloch[a]

[a]Centre d'études nucléaires de Saclay, Gif-sur-Yvette, France (S. & O.)

Reçu le 16 décembre 1967. Available online 11 September 2003.

Purchase the full-text article

PDF and HTML
All references
All images
All tables

Abstract

Two methods leading to general prescriptions for the construction of the time independent perturbation expansions for bound states to all orders are discussed. The two methods differ by the choice of the zeroth-order wave functions. In both cases, the expansions are simplified by the use of non-orthonormalized zeroth-order wave functions. The terms of the expansions are described by means of diagrams. The accuracy of the expansions is tested by a calculation of two eigen-values of the Mathieu equation.