Kato Round Two PhL 1.31.09

Section 15. ..... **1**
Section 16. ..... **2**
What do the P_n depend on? 3
What does such an expansion of P in powers of λ *mean*? 4
Section 17. ..... **4**
Part I : Assume $g_a = 1$. **4**
Kato on state corrections: 5
Kato on energy corrections. 6
Part II : Assume $g_a > 1$ **7**
The Meaning of the Generalized Eigenvalue Equation. 10

Section 15.

(1) Preliminary Detail. Let's ponder this claimed theorem

$$G_{0a} \equiv P_0\ f(H_0)\ P_0 = f(H_0)\ P_0 = P_0\ f(H_0)$$

$$P_0\ f(H_0) = \Sigma_{r\in\varepsilon a0}\ |r><r|\ f(H_0) = \Sigma_{r\in\varepsilon a0}\ |r><r|\ f(E_r)$$

If I insert $P_0 = \Sigma_{r\in\varepsilon a0}\ |r><r|$ then all three become equal, but what about domain, range and singularities? The triple operator obviously maps any perp space vector to 0, but the rightmost operator would do this:

$$P_0\ f(H_0)\ |k> = P_0\ f(E_k)\ |k> = f(E_k)\ P_0|k> = 0$$

So it seems that these two operators give the same results on the perp space as well as on the subspace. An example is that if we have $f(H_0) = (E_m - H_0)^{-1}$ this does not diverge when we act on |k>.

Notice another fact: since the E_r are the same for all states in the subspace, call that value E_a^0, then we have

$$P_0\ f(H_0) = \Sigma_{r\in\varepsilon a0}\ |r><r|\ f(H_0) = \Sigma_{r\in\varepsilon a0}\ |r><r|\ f(E_r) = f(E_a^0)\ \Sigma_{r\in\varepsilon a0}\ |r><r| = f(E_a^0)\ P_0$$

So we really can make this full claim:

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

2.Preliminary Detail. Now, in this section of Messiah we are no longer dealing with H_0 and the $\mathcal{E}_a^0$ space. We are instead dealing with the full H, and we are assuming that even when this full H is applied, we still have some degeneracy and so P_i applies to one of these degenerate subspaces. (of course it might have dimension 1). So the theorem of interest then becomes this:

$$G_a \equiv P_a\ f(H)\ P_a\ = f(H)\ P_a\ = P_a\ f(H)$$

We prove validity here just as above. Our degenerate space is now $\mathcal{E}_a$ with common energy E_a and we have $P_a\ = \Sigma_{r\in\mathcal{E}a}\ |r><r|$ where |r> are now eigenstates of H in this $\mathcal{E}_a$ space, and the final result is still valid,

$$G_a \equiv P_a\ f(H)\ P_a\ = f(H)\ P_a\ = P_a\ f(H) = \text{ replace H in any of these by } E_a$$

Now suppose $f(H) = (z-H)^{-1}$ where z is a complex variable. Then we will have

$$P_a\ f(H)\ |k> = P_a\ (z-H)^{-1}\ |k>\ = P_a\ (z-E_k)^{-1}\ |k>\ =\ (z-E_k)^{-1}\ P_a\ |k> = 0$$

so we have no problem with singularities. If z happens to equal E_k , we just move away from the point a bit in the usual analytic sense. Here |k> is some full SE solution in the perp space of $\mathcal{E}_a$.

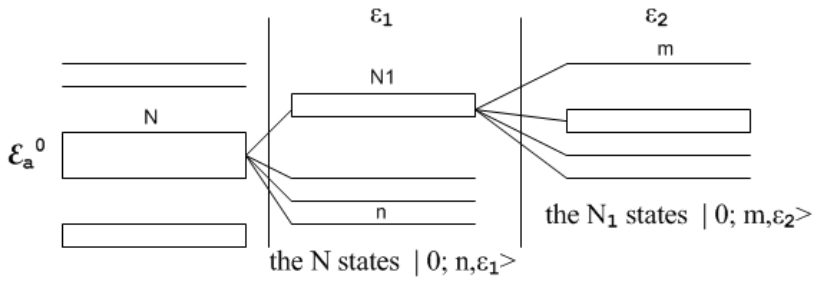
(3) At this point, I am happy with all of Section 15. We have derived contour integral expressions for the operators P_Γ and the product HP_Γ where H is the full Hamiltonian, and where P_Γ is the projector associated with certain poles of G(z) [ie, the eigenvalues of H] , those poles included by your arbitrary contour Γ. There is nothing mysterious about P_Γ . If you apply it to some ket, it projects out the parts of that ket (state) which are eigenstates with the eigenvalues your contour surrounds.

Section 16.

1. We can write $G = \Sigma_{n=0}\ \lambda^n\ G_o(VG_o)^n$ which I have no problem with.

At this point, we have to have a very clear understanding of the meaning of an operator he calls simply P.

Imagine that when λ=0, we have some degeneracy group of size N. When we turn on λ and think on all orders of λ, we know that the states will split apart in some way, and there may be some degeneracies left, such as in our picture



My former meaning for the middle column above was that it corresponded to first order only. But now let's just think of it as representing all orders. So as we turn on λ, our N degenerate states move apart in energy and in my example we are left with only one degenerate subgroup of size N_1. Corresponding to this group we have some subspace called $\mathcal{E}_{N1}$ spanned by the full eigenfunctions, whatever they are. This subspace $\mathcal{E}_{N1}$ has a projector $P^{(N1)}$. But P is supposed to be *the union of* all such subspaces, and this P

acts as a projector for all states (those I show) in the middle column above. When λ is turned off, this P becomes the P_0 for all of $\mathcal{E}_a^0$. The trick is that our selected contour is supposed to capture the same poles whether λ is on or off, despite the fact that the poles moved a small amount. The claim is that we can find such a fixed contour.

2. So equation C on page 715 is our expression for P as an integral on our special contour now called Γ_a . If we simply insert the expansion (62) into equation C, we get the pair of results (63) and (64). Our goal is to evaluate (64) so that we have a usable expansion (63). The evaluation is extremely non-obvious, but I have done all the details in the raw notes, and I agree with the general result (67) and the specific lowest order terms in (69). The derivation is entirely mechanical.

What do the P_n depend on?

$P = \Sigma_n \lambda^n P_n$

Notice this fact: when we talk about the poles of C which give P in the limit $\lambda = 0$, we are saying

$$P_0 = (2\pi i)^{-1} \int_{\Gamma a} dz\, G_0(z)$$

and contour Γ_a captures all the poles, but all the poles are at the exact same location which is E_a^0. So really there is exactly one pole at $z = E_a^0$ and its residue is P_0. When we turn on λ, the N poles tend to move away from $z = E_a^0$, but our contour still captures them by design.

Now consider:

$$P = (2\pi i)^{-1} \int_{\Gamma a} dz\, G(z) = (2\pi i)^{-1} \int_{\Gamma a} dz\, G(H_0, V, \lambda; z) = h(H_0, V, \lambda, E_i)$$

so here we claim that P is a function of all four variables shown, and E_i are the full SE energy eigenvalues which we don't know. Note that we replaced H dependence with H_0 and V dependence. Now look what happens when we put in the λ expansion for G,

$$P = (2\pi i)^{-1} \int_{\Gamma a} dz\, G(z) = \Sigma_n \lambda^n (2\pi i)^{-1} \int_{\Gamma a} dz\; G_0 (VG_o)^n = \Sigma_n \lambda^n P_n$$

The left integral has poles at E_i as just stated, but the right integral has all its "poles" at E_a^0, so for any value of n, the integrand of the dz integral $G_0 (VG_o)^n$ only has singularities at $z = E_a^0$. We see however that the terms don't just have simple poles, they have poles of degree n+1. We expect the residues of all these poles of various orders to be a function of the singularity location E_a^0. Therefore, we expect to see this dependence :

$$P_n = (2\pi i)^{-1} \int_{\Gamma a} dz\; G_0 (VG_o)^n = P_n(H_0, V, E_a^0)$$

$P_0 = (2\pi i)^{-1} \int_{\Gamma_a} dz\ G_0 \qquad = P_0(H_0, E_a^0) \qquad$ // no V dependence here.

and that is exactly what we see in the terms shown in (69). The E_a^0 dependence is inside the Q_{a0} factors and some is also inside the P_0 factors

What does such an expansion of P in powers of λ *mean*?

All I know to say is this:

(1) P is an operator that depends on λ and, since λ is assumed small, it is reasonable to expand P in a power series in λ, so $P = \Sigma_n \lambda^n P_n$.

(2) Imagine that we start with subspace $\mathcal{E}_a^0$ when λ=0 and it contains N basis vectors. When we turn λ on, the correctly chosen basis vectors fly off into a set of spaces we might call $\mathcal{E}_a^{(i)}$ and we could talk about the total space $\mathcal{E}_a = \mathcal{E}_a^{(1)} \oplus \mathcal{E}_a^{(2)} \oplus \ldots \oplus \mathcal{E}_a^{(n)}$ where there were n *distinct* eigenvalues for the final energy E_i . Each of these spaces has g_i basis vectors (Messiah page 698). The dimensionality of the space $\mathcal{E}_a$ is still $N = g_a = \Sigma_i g_i$, but it is not the same space as $\mathcal{E}_a^0$. The operator P acts on the space of all full solutions of the SE, and it projects out any solutions that are within the space $\mathcal{E}_a$, and it rejects those solutions which are in the space $(\mathcal{E}_a)^{\perp}$. That is what P does! Note that $\mathcal{E}_a$ is a subspace of full $\mathcal{H}$. The *component* projector P_n projects out the portion of a state in $\mathcal{E}_a$ that is of order λ^n.

(1) We have shown that the P_n are functions of H_0, V, E_a^0, and in practice this comes out meaning that the P_n are functions of operators P_0 , V, and Q_{a0}.

Comment: There is a *huge* amount of underlying math going on in Section 16. I did it all in the raw notes, and I may review it all later in this document, but on my first pass I derived every single thing including a typo found on page 717 in the last line. For the moment, let us completely ignore all this mathematical machinery and just accept the two results:

(1) expansion (69) showing the first terms in the expansion of $P = \Sigma_{n=0} \lambda^n P_n$.

(2) expansion (72) showing the first terms in the expansion of $(H-E_a^0)P = \Sigma_{n=1} \lambda^n B_n$.

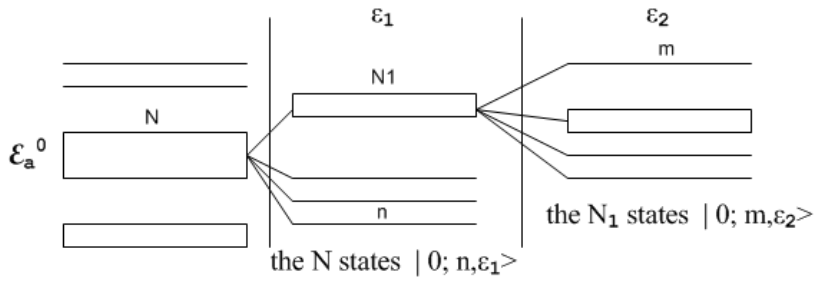
Here, the P_n and B_n are finite sums of products of operators P_0, V and Q_{a0} ! There are no other base operators involved. It is all really quite simple when the dust settles.

Section 17.

Part I : Assume $g_a = 1$.

This means N=1 in my picture below! The dimension of $\mathcal{E}_a^0$ is 1.

So this analysis will apply only to one of the isolated singlet states on the left edge of my picture. This analysis does not apply to singlet states like "n" and "m" !



For now we shall refer to our all-orders singlet state of interest just as $|\psi_0\rangle$ where subscript 0 is unrelated to component 0.

Kato on state corrections:

[see update of this subject in Kato Meta notes Section 2A] We assume that our state has the usual components expansion:

$$|\psi_0\rangle = |0;\psi_0\rangle + \lambda\,|1;\psi_0\rangle + \lambda^2\,|2;\psi_0\rangle \ \ldots$$

Theorem: $P\,|0;\psi_0\rangle = \text{constant} * |\psi_0\rangle$

This is easily shown in the notes when we expand $|0;\psi_0\rangle = \Sigma_i\, a_i|\psi_i\rangle$ and then apply P.

Corollary: If we apply P on the left and our expansion for $P = \Sigma_n\, \lambda^n\, P_n$ on the right, then we have produced a full all-orders expansion for the state $|\psi_0\rangle$ where we now ignore overall normalization on $|\psi_0\rangle$

$$|\psi_0\rangle = P\,|0;\psi_0\rangle = \Sigma_n\, \lambda^n\, P_n\,|0;\psi_0\rangle$$

In other words, this says:

$|n;\psi_0\rangle = P_n\,|0;\psi_0\rangle$ = n^{th} order state correction

where P_n are the operators we computed in the last section. Here are some examples

$$|1;\psi_0\rangle = Q_{0a}\,V\,|0;\psi_0\rangle$$

$$|2;\psi_0\rangle = [Q_{0a}\,V\,Q_{0a}\,V - Q_{0a}^2\,V\,P_0\,V\,]\,|0;\psi_0\rangle$$

You can throw out terms in these formulas which have a P_0 on the left (they give multiples of $|0\rangle$ and only affect normalization), or which have Q_{0a} on the right (these terms are 0 since $Q_0|0\rangle = 0$). I have already thrown out several terms to get the above results. [Again, see Kato Meta notes on this.]

Comment: Prior to Kato, we were able to compute state corrections in the non-degenerate theory only by a rather clumsy iterative method, which I quote:

$\|1\> = Q_{0a}V\|0\>$		1	s = 1
$\varepsilon_1 = \<0\| V \|0\> = V_{nn}$	**// Messiah p 689 (12)**	1	
$\|2\> = Q_{0a}V\|1\> - \varepsilon_1 Q_{0a}\|1\>$		2	s = 2
$\varepsilon_2 = \<0\| V \|1\> = \<0\|V Q_{0a}V\|0\>$	**// Messiah p 694 (26)**	1	
$\|3\> = Q_{0a}V\|2\> - \varepsilon_1 Q_{0a}\|2\> - \varepsilon_2 Q_{0a}\|1\>$		5	s = 3
$\varepsilon_3 = \<0\| V \|2\>$		2	
$\|4\> = Q_{0a}V\|3\> - \varepsilon_1 Q_{0a}\|3\> - \varepsilon_2 Q_{0a}\|2\> - \varepsilon_3 Q_{0a}\|1\>$		13	s = 4
$\varepsilon_4 = \<0\| V \|3\>$		5	
$\|5\> = Q_{0a}V\|4\> - \varepsilon_1 Q_{0a}\|4\> - \varepsilon_2 Q_{0a}\|3\> - \varepsilon_3 Q_{0a}\|2\> - \varepsilon_4 Q_{0a}\|1\>$		34	s = 5
$\varepsilon_5 = \<0\| V \|4\>$		13	

In order to compute |5>, for example, we need to know all these things:

$|4\>, |3\>, |2\>, |1\>, \varepsilon_4, \varepsilon_3, \varepsilon_2, \varepsilon_1$

Kato gives as a direct formula for |5> (it may have 34 operator terms, however). You can see that somehow the Kato method sums all these things up magically.

Thus, at this point we have a Kato formula for all (non-degenerate case) state corrections. Now we want:

Kato on energy corrections.

Theorem: $HP = E_a P$. Here, E_a refers to the final all-orders corrected energy of state $|\psi_0\>$, and P is as we have been using it, it projects only $|\psi_0\>$. We can prove this theorem by applying it to all states in $\mathcal{H}$. For the state $|\psi_0\>$ which has energy E_a (odd notation now), it is obviously true. Then for all other states in $\mathcal{H}$, it says that 0 = 0. I think this is enough to show operator equality.

Theorem: Tr(P) = Tr(P_0) = 1. Whether we do P or P_0, our projector punches out just one state, and if we go to a basis where the P's are diagonal, there is just one diagonal element and it is 1. Then trace is independent of basis, QED.

Now recall from the end of Section 16 that

$$(H-E_a^0)P = \Sigma_{n=1} \lambda^n B_n$$

Take the trace of both sides:

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

But we know $HP = E_a P$ from above, so get

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

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

which tells us how to compute the correction in any order,

$\varepsilon_n = tr(B_n)$

The first few B_n we wrote out in (72) on page 717, and we have a machine that can push out any B_n we want! So now we can compute both |n> and ε_n independently using Kato formulas, and we have no need for that "clumsy iterative sequence".

So even in our non-degenerate theory, Kato has made a contribution.

Theorem: $Tr(P_0 A) = \langle 0|A|0\rangle$

Proof: $Tr(P_0 A) = Tr(|0\rangle\langle 0|A) = \langle i|0\rangle\langle 0|A|i\rangle = \langle 0|A|i\rangle\langle i|0\rangle = \langle 0|A|0\rangle$

Fact: It turns out that all the terms in the sum that is B_n, such as the examples shown in (72) for n=1,2, contain at least one factor of P_0, and my raw notes explain why this is so. A typical term might then have the form

$Tr(B_n \text{ term}) = Tr(MP_0N) = Tr(P_0NM) = \langle 0|NM|0\rangle$ // using the above theorem.

Thus it is that every term in $tr(B_n)$ can be expressed in the form $\langle 0|X|0\rangle$ where X is some operator. The bottom of page (74) shows the lowest three cases, I calculate them in my raw notes. He replaces any residual P_0 by $|0\rangle\langle 0|$ since this simplifies the results!

So now we have the energy corrections all in the form $\langle 0|X|0\rangle$.

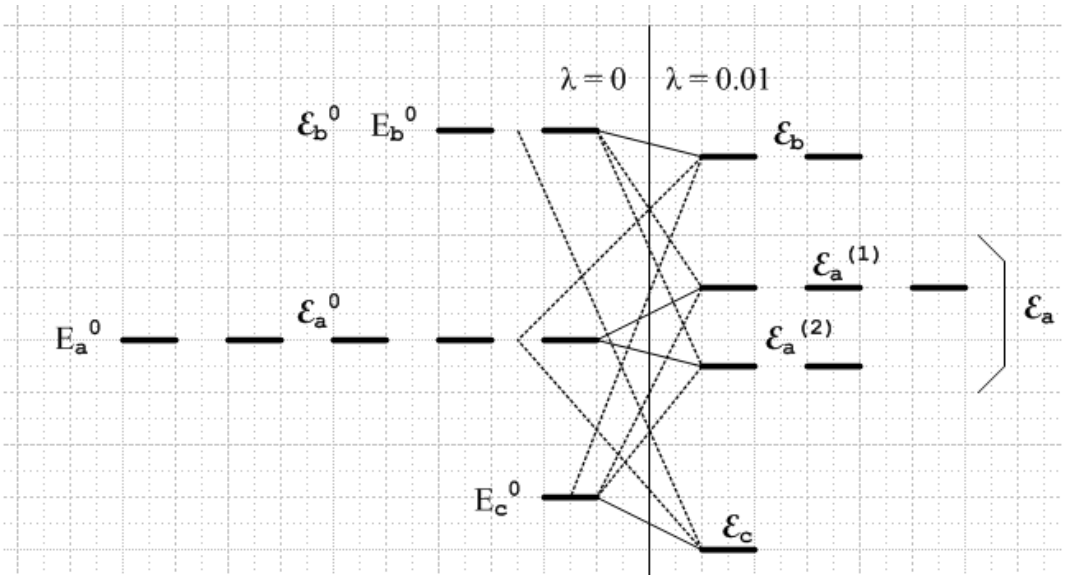
Part II : Assume $g_a > 1$.

This means N=1 in my picture below! The dimension of $\mathcal{E}_a^0$ is 1.

In my Round One, I had a lot of trouble with the first half of page 718, so here is another attack:

Setting up the 1:to:1 mappings between $\mathcal{E}_a^0$ and $\mathcal{E}_a$.

We start with this picture:



On either side of the vertical line, we have a complete set of states that span $\mathcal{H} = \mathcal{H}^{(0)}$.

Theorem 1: The states |n> and |0;n> are always connected by a non-vanishing matrix element.

Let's assume that the five states in $\mathcal{E}_a^0$ shown on the left are the λ=0 limits of the 5 perturbed states on the right. We know we can find linear combinations of any starting set of 5 basis functions on the left and find the right ones to make this true. Call these states on the left |0;n> for n = 1,2,3,4,5. Call the states on the right just |n> for n=1,2,3,4,5. We can write |n> = |0;n> + λ |1;n> + etc. In this notation, <n|0:n> = 1 when λ = 0, but becomes something else when λ = 0, but it never becomes 0 ! That is my claim. The states |n> and |0;n> are always connected by a non-vanishing matrix element. This is I guess an appeal to continuity near the location λ = 0. If this connecting matrix element is lost, then I don't think we can claim to be doing "perturbation theory". We are just perturbing a state slightly, we are not jerking it completely out of its box so it has no connection with the initial state. So I think this Theorem 1 is OK.

Claim A1: P $|\mathcal{E}_a^0\alpha\rangle = |\mathcal{E}_a\beta\rangle$. For any state α in $\mathcal{E}_a^0$, Pα gives us some state β in $\mathcal{E}_a$. Details:

any state in $\mathcal{E}_a^0$ = linear combination of states in $\mathcal{E}_a$ + linear comb of other states on the right
due to the fact that all states span $\mathcal{H}$.

P(any state in $\mathcal{E}_a^0$) = P(linear combination of states in $\mathcal{E}_a$) = a particular state in $\mathcal{E}_a$

Let's pick a specific state "α" in $\mathcal{E}_a^0$ and call it $|E_a^0\alpha\rangle$. Let's say the last line above projects it into state $|\mathcal{E}_a\beta\rangle$ which is some specific linear combination of states that make up $\mathcal{E}_a$. So we have

P $|\mathcal{E}_a^0\ \alpha\rangle = |\mathcal{E}_a\beta\rangle$

We certainly have not made any profound claim here. We have just said that P acting on an arbitrary state in $\mathcal{E}_a^0$ results in some state in $\mathcal{E}_a$.

Theorem A1: The mapping shown above treated as P: $\mathcal{E}_a^0 \rightarrow \mathcal{E}_a$ is in fact 1:to:1 and is invertible.

Proof: Let's consider this mapping as P: $\mathcal{E}_a^0 \rightarrow \mathcal{E}_a$, a restriction of P: $\mathcal{H} \rightarrow \mathcal{E}_a$. Does this first mapping have a nullspace within $\mathcal{E}_a^0$? Could there be a state in $\mathcal{E}_a^0$ that is entirely a linear combination of states in $(\mathcal{E}_a)^{\perp}$? If so, then we do have a nullspace. But in Theorem 1 above we claim that this can never happen, and any state in $\mathcal{E}_a^0$ will have an overlap with some state in $\mathcal{E}_a$. This is true for the 5 tuned states we show on the left and described above, and is true then for any linear combination of them. Thus, there are no elements in $\mathcal{E}_a^0$ such that $P\psi_o = 0$, there is no nullspace, and so P is 1:to:1 and is invertible.

Normally a projection operator is non-invertible, but when described in this way, P is invertible. P of course is not invertible when described as P: $\mathcal{H} \rightarrow \mathcal{E}_a$.

Now we can repeat everything above in the other direction and say

Claim A2: $P_0 |\mathcal{E}_a\delta\rangle = |\mathcal{E}_a^0\gamma\rangle$. For any state δ in $\mathcal{E}_a$, $P_0\delta$ gives us some state γ in $\mathcal{E}_a$. Details:

Theorem A2: The mapping shown above treated as P_0: $\mathcal{E}_a \rightarrow \mathcal{E}_a^0$ is in fact 1:to:1 and is invertible.

Fact: We are *not claiming* that $P_0^{-1} = P$. I suspect this is not true, but whether or not true, we do not need this fact.

Corollary C1 . Any element β of $\mathcal{E}_a$ may be uniquely represented as P acting on some state α in $\mathcal{E}_a^0$.

$$|\mathcal{E}_a\beta\rangle = P\,|\mathcal{E}_a^0\alpha\rangle$$

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

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

Corollary C2 . Any element γ of $\mathcal{E}_a^0$ may be uniquely represented as P_0 acting on some state δ in $\mathcal{E}_a$.

$$|\mathcal{E}_a^0\ \gamma\rangle = P_0\,|\,\mathcal{E}_a\delta\rangle =$$

Now finally we can march ahead on page 718. Let's now change our notation to match Messiah. We will rewrite corollary C1 to say

$$|\mathcal{E}_a\alpha\rangle = P\,|\mathcal{E}_a^0\alpha\rangle$$

rather than use α and β. We are certainly not implying that $|\mathcal{E}_a\alpha\rangle = |\mathcal{E}_a^0\alpha\rangle$. We are just using the same label α on both sides to show that they are related in this way by P.

Now let $|\mathcal{E}_a\alpha\rangle$ be an eigenket of H. Then we have

$$H\,|\,\mathcal{E}_a\alpha\rangle = E_\alpha\,|\,\mathcal{E}_a\alpha\rangle$$

and we can thus write

$$HP\,|\mathcal{E}_a{}^0\alpha\rangle = E_\alpha\,P\,|\mathcal{E}_a{}^0\alpha\rangle \qquad \text{// which is page 718 A}$$

We are free to apply P_0 to both sides to get

$$P_0HP\,|\mathcal{E}_a{}^0\alpha\rangle = E_\alpha\,P_0P\,|\mathcal{E}_a{}^0\alpha\rangle \qquad \text{// which is page 718 B}$$

We are free again to express $|\mathcal{E}_a{}^0\alpha\rangle = P_0\,|\mathcal{E}_a{}^0\alpha\rangle$ on both sides to get

$$P_0HPP_0|\mathcal{E}_a{}^0\alpha\rangle = E_\alpha\,P_0PP_0\,|\mathcal{E}_a{}^0\alpha\rangle$$

and now we define

$$H_a \equiv P_0HPP_0 \qquad K_a \equiv P_0PP_0 \qquad \text{// which is page 718 (75)}$$

and we then have

$$H_a\;|\mathcal{E}_a{}^0\alpha\rangle = E_\alpha\,K_a\;|\mathcal{E}_a{}^0\alpha\rangle \qquad \text{// which is p 718 (76)}$$

and this is our famous "generalized eigenvalue problem". At this point in the raw notes I show the following facts:

The Meaning of the Generalized Eigenvalue Equation. What exactly does the equation

$$P_0HPP_0|\mathcal{E}_a{}^0\alpha\rangle = E_\alpha P_0PP_0\,|\mathcal{E}_a{}^0\alpha\rangle$$

tell us? The state $|\mathcal{E}_a{}^0\alpha\rangle$ is not "just any state". It is this *specific* state:

$$|\mathcal{E}_a{}^0\alpha\rangle = P^{-1}\,|\,\mathcal{E}_a\alpha\rangle \qquad |\,\mathcal{E}_a\alpha\rangle = P\,|\mathcal{E}_a{}^0\alpha\rangle$$

It is the "back-projection" into $\mathcal{E}_a{}^0$ of a state $|\,\mathcal{E}_a\alpha\rangle$ in $\mathcal{E}_a$ which solves the full SE $H\psi = E_\alpha\psi$. There are N such states in our notation, so think of $\alpha = 1,2...N$. So there are N of these back-projections. Only these certain states satisfy the above equation. If we could exactly solve the above equation, we would exactly now the eigenvalues E_α and we would exactly know the back-projections $|\mathcal{E}_a{}^0\alpha\rangle$.

Now consider a more generic problem in some N dimensional space

$$A|i\rangle = k_iB|i\rangle$$

which we want to solve for k_i and the $|i\rangle$. Suppose A and B are series in a smallness parameter λ. Suppose we were to write

$$A(\text{thru } \lambda^2)\ |i,\lambda> = k_i(\text{thru } \lambda^2)B(\text{thru } \lambda^2)\ |i, \text{thru } \lambda^2>$$

If we could solve this equation exactly, we would interpret $k_i(\text{thru } \lambda^2)$ and $|i, \text{thru } \lambda^2>$ as being approximations to k_i and $|i>$. We would expect as we add more and more terms to A and B, $k_i(\text{thru } \lambda^2)$ and $|i,\text{thru } \lambda^2>$ would converge to the values k_i and $|i>$. Write out the above,

$$[\ A_0 + \lambda A_1 + \lambda^2 A_2]\ |i, \text{thru } \lambda^2> \ = k_i(\text{thru } \lambda^2)\ [\ B_0 + \lambda B_1 + \lambda^2 B_2]\ |i, \text{thru } \lambda^2>$$

Main Point: once you pick an order (here order = 2), you have a well-defined problem which has eigenvalues and eigenkets. I don't think it makes any sense to try to represent this as a sum of problems in each lower order! For example, I think the following does not make sense, where we treat $k_{i,1}$ for example as the solution of some "order λ= 1 only" problem,

$$k_i(\text{thru } \lambda^2) = k_i(\text{at order } \lambda^0) + \lambda\ k_i(\text{at order } \lambda^1) + \lambda^2 k_i(\text{at order } \lambda^2)$$

$$= k_{i,0} + \lambda\ k_{i,1} + \lambda^2\ k_{i,2}$$

and similarly an attempted expansion of $|i, \text{thru } \lambda^2>$ in this manner does not make sense. Suppose we were to solve the "thru λ^0" equation and find the values $K_{i,0}$. And then we go off and solve the "thru λ^1" equation and find $k_i(\text{thru } \lambda^1)$. We can certainly Taylor expand all solutions in an infinite power series in smallness parameter λ (assuming things are nice of course),

$$k_i(\text{thru } \lambda^0) = K_{i,0}$$
$$k_i(\text{thru } \lambda^1) = K'_{i,0} + K'_{i,1}\lambda + K'_{i,2}\lambda^2 + ...$$
$$k_i(\text{thru } \lambda^2) = K''_{i,0} + K''_{i,1}\lambda + K''_{i,2}\lambda^2 + ...$$

I think almost surely you would find that $K_{i,0} \neq K'_{i,0} \neq K''_{i,0}$ and so on. We see this same issue when we treat the set t^n of polynomials as a spanning set in L_2. It does not form an orthogonal basis, so the coefficients of lower terms don't "stay put". You need to do a lot of work (maybe Bloch does this work) to make things stay put.

So here is the MO for handling this kind of problem. You start probably with order λ^1 and solve the eigenvalue problem. If you want to know k_i to more accuracy, or if you want to try to split any remaining degeneracy in the eigenkets, then you start over and solve the λ^2 problem.

Theorem 1: Any projection operator is Hermitian.

Theorem 2: Projection operators have non-negative diagonal matrix elements so $<\psi|R|\psi> \geq 0$

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

Theorem 3: The operator $K_a = P_0PP_0$ is **positive definite**. This is so because you can write the operator as $K_a = P_0PP_0 = P_0PPP_0 = (P_0P)\ (P_0P)^\dagger = BB^\dagger$. I don't have a proof that K_a does not have any 0 eigenvalues, but I think it does not. Messiah says positive definite and that means no zeros.

We can then write our equation above as

$$(K_a^{-1} H_a)\ |\mathcal{E}_a^0\alpha\rangle = E_\alpha\ |\mathcal{E}_a^0\alpha\rangle$$

and this is then a standard EV problem in an NxN space.

But now I am skeptical of this last inversion on K_a, so let's leave our problem as this:

$$H_a\ |\mathcal{E}_a^0\alpha\rangle = E_\alpha K_a\ |\mathcal{E}_a^0\alpha\rangle$$

[But footnote on page 719 states outright that K_a has an inverse! And he comments as I have often done in what sense this inverse exists.]

Since we have

$$H_a \equiv P_0(HP)P_0 \qquad K_a \equiv P_0(P)P_0 \qquad \text{// which is page 718 (75)}$$

and since we have explicit expansions in λ for (P) and for (HP), we can put these in to get a problem to a given order. Let's use (77) and (78) and expand in orders

$$K_a^{(0)} = P_0$$

$$K_a^{(1)} = 0$$

$$K_a^{(2)} = -\lambda^2 P_0 V Q_{a0}^2 V P_0$$

$$H_a^{(0)} = E_a^0 P_0$$

$$H_a^{(1)} = \lambda P_0 V P_0$$

$$H_a^{(2)} = E_a^0 K_a^{(2)} + \lambda^2 P_0 V Q_{a0} V P_0 = -\lambda^2 E_a^0\ P_0 V Q_{a0}^2 V P_0 + \lambda^2 P_0 V Q_{a0} V P_0$$

Also think of

$$E_\alpha = E_a^0 + \lambda\varepsilon_1 + \lambda^2 \varepsilon_2 + \ldots$$

Now let's write a generalized EV problem in each order:

λ^0: $E_a^0 P_0\ |\mathcal{E}_a^0\alpha\rangle = E_a^0\ P_0\ |\mathcal{E}_a^0\alpha\rangle$ just says $E_\alpha \approx E_a^0$

λ^1: $\lambda P_0 V P_0 |\mathcal{E}_a^0\alpha\rangle = \lambda\varepsilon_1 P_0\ |\mathcal{E}_a^0\alpha\rangle$ $P_0 V P_0 |\mathcal{E}_a^0\alpha\rangle = \varepsilon_1 |\mathcal{E}_a^0\alpha\rangle$

I do recognize this as our level 1 EV problem.

λ^2: $H_a^{(2)}\ |\mathcal{E}_a^0\alpha\rangle = [\ E_\alpha^{(2)} K_a^{(0)} + E_\alpha^{(0)} K_a^{(2)}\ ]\ |\mathcal{E}_a^0\alpha\rangle$

$$[-\lambda^2 E_a^0 \; P_0 V Q_{a0}^2 V P_0 + \lambda^2 P_0 V Q_{a0} V P_0] \quad |\mathcal{E}_a^0 \alpha>$$

$$= \{ \varepsilon_2 \lambda^2 P_0 - E_a^0 \lambda^2 P_0 V Q_{a0}^2 V P_0 \} \; |\mathcal{E}_a^0 \alpha>$$

$$[-E_a^0 \; P_0 V Q_{a0}^2 V P_0 + P_0 V Q_{a0} V P_0] \; |\mathcal{E}_a^0 \alpha> \quad = \{ \varepsilon_2 P_0 - E_a^0 P_0 V Q_{a0}^2 V P_0 \} \; |\mathcal{E}_a^0 \alpha>$$

Two large terms cancel and we get

$$P_0 V Q_{a0} V P_0 \; |\mathcal{E}_a^0 \alpha> = \varepsilon_2 P_0 \; |\mathcal{E}_a^0 \alpha>$$

$$P_0 (V Q_{a0} V) \; |\mathcal{E}_a^0 \alpha> = \varepsilon_2 \; |\mathcal{E}_a^0 \alpha>$$

and I do recognize this as our level 2 EV problem.

λ^3: $\quad H_a^{(3)} \; |\mathcal{E}_a^0 \alpha> = [E_\alpha^{(3)} K_a^{(0)} + E_\alpha^{(1)} K_a^{(2)}] \; |\mathcal{E}_a^0 \alpha>$

The RHS I can write as:

$$= [\lambda^3 \; \varepsilon_3 P_0 - \varepsilon_1 \lambda \lambda^2 P_0 V Q_{a0}^2 V P_0] \; |\mathcal{E}_a^0 \alpha>$$

Rewrite this as

$$\{ H_a^{(3)} + \varepsilon_1 P_0 V Q_{a0}^2 V P_0 \} \; |\mathcal{E}_a^0 \alpha> = \varepsilon_3 \; |\mathcal{E}_a^0 \alpha>$$

But I have to do more work to compute $H_a^{(3)}$ because I need $(HP)^{(3)}$ and we only did this on page 717 to order 2. I would have to do it like this from (70)

$$(HP)^{(3)} = E_a^0 P^{(3)} + B^{(3)}$$

so there are really two new things I would have to compute. **But I could do it**, and I could then find out what that ε_3 equation looked like. [I finally did this in the Kato Meta notes]

This is an important point. In my non-resolvent section notes, I had no way to "thread myself" to the ε_3 equation. But here Kato gives us the exact prescription, we just have to turn the crank! Of course actually solving the level 3 EV equation would probably be a nasty numerical problem like the level 2 problem because all the perp space states are involved.

Now, above we were looking for the EV equations in each order and we saw exactly how you go about that problem.

But suppose we only care about the energies! Well, part of the same problem. We have to do this:

level 1:

$P_0 V P_0 |\mathcal{E}_a^0 \alpha\rangle = \varepsilon_1 |\mathcal{E}_a^0 \alpha\rangle \quad \Rightarrow \quad \det(V - \varepsilon_1) = 0$

level 2:

$P_0 (V Q_{a0} V) |\mathcal{E}_a^0 \alpha\rangle = \varepsilon_2 |\mathcal{E}_a^0 \alpha\rangle \quad \Rightarrow \quad \det(V Q_{a0} V - \varepsilon_2) = 0$

level 3:

$$\{ H_a^{(3)} + \varepsilon_1 P_0 V Q_{a0}^2 V P_0 \} \; |\mathcal{E}_a^0 \alpha\rangle = \varepsilon_3 |\mathcal{E}_a^0 \alpha\rangle$$

$$\Rightarrow \quad \det(H_a^{(3)} + \varepsilon_1 P_0 V Q_{a0}^2 V - \varepsilon_3) = 0$$

and so on. Notice that these determinants are always NxN if we started with $\mathcal{E}_a^0$ of that dimension.

The determinant should be basis independent, so we can compute them in a raw starting basis for the space $\mathcal{E}_a^0$.

This subject is revisited in later Rounds.