How to do First Order Perturbation Theory PhL 1.19.05

After being tripped up 4 times on the same thing in Jim's paper on 4 different occasions, the time has now come to write down how this works. I will use Schiff page 245 as a basis.

Let's start with this problem, where the ϕ_i are orthonormal $(\phi_m, \phi_n) = \delta_{m,n}$

$$T\,\phi_i = \lambda_i\,\phi_i \qquad (1)$$

where λ_i are eigenvalues of T and ϕ_i are the eigenvectors. Now make this perturbation to T

$$T \rightarrow T + \varepsilon T' \qquad (2)$$

where ε is a smallness parameter. Assume the solutions now look like this

$$(T + \varepsilon T')\,\psi_i = \Lambda_i\,\psi_i \qquad (3)$$

Now do the following expansions

$$\psi_i = \phi_i + \varepsilon\,\Sigma_j a_{ij}\phi_j \qquad (4)$$ // note that the sum includes some ϕ_i

where we are now assuming that $\psi_i \approx \phi_i$ At the same time, let us assume that

$$\Lambda_i = \lambda_i + \varepsilon\,L_i \qquad (5)$$ // how the eigenvalue changes

where again we are assuming that $\Lambda_i \approx \lambda_i$. Our equation (3) becomes

$$(T + \varepsilon T')\,[\,\phi_i + \varepsilon\,\Sigma_j\,a_{ij}\phi_j\,] = (\,\lambda_i + \varepsilon\,L_i\,)\;[\,\phi_i + \varepsilon\,\Sigma_j\,a_{ij}\phi_j\,] \qquad (6)$$

We can now group terms by powers of ε:

$$T\phi_i = \lambda_i\,\phi_i \qquad \varepsilon^0 \qquad (7)$$

$$T'\,\phi_i + \Sigma_j\,a_{ij}\,T\phi_j = L_i\,\phi_i + \lambda_i\,\Sigma_j\,a_{ij}\phi_j \qquad \varepsilon^1 \qquad (8)$$

But $T\phi_j = \lambda_j\,\phi_j$ in the second term, so we then have

$$T'\,\phi_i + \Sigma_j\,a_{ij}\,\lambda_j\,\phi_j = L_i\,\phi_i + \lambda_i\,\Sigma_j\,a_{ij}\,\phi_j \qquad \varepsilon^1 \qquad i = 1,2,3... \qquad (9)$$

which we can rewrite as

$$T'\,\phi_i = L_i\,\phi_i + \lambda_i\,\Sigma_j\,(\lambda_i - \lambda_j)\,a_{ij}\,\phi_j \qquad (10)$$

Suppose we now close with $(\phi_k$, then we have [not specifying Hilbert space here!]

$$(\phi_{k,} T' \phi_i) = L_i \delta_{k,i} + (\lambda_i - \lambda_k) a_{ik} \qquad (11)$$

Now separate this into the two cases:

$$(\phi_{i,} T' \phi_i) = L_i \qquad k = i \qquad (11a)$$
$$(\phi_{k,} T' \phi_i) = (\lambda_i - \lambda_k) a_{ik} \qquad k \neq i \qquad (11b)$$

The first result here is "the classic result" [see Schiff page 246 eq 31.8] for the shift in an energy level, say. The second equation looks like Schiff (31.10).

Now one extra fact. Suppose we go from ϕ_i to some other $f_i = k_i\phi_i$ which are *not* normalized. Then clearly we would replace our "classic result" with

$$L_i = (\phi_{i,} T' \phi_i) = (1/k_i)^2 \ (f_{i,} T' f_i) = (f_{i,} T' f_i) / (f_{i,} f_i) \qquad (12)$$

We can summarize the "classic result" in this way. If we have $Tf_i = \lambda_i f_i$ then

$$\delta\lambda_i = (f_{i,} \delta T f_i) / (f_{i,} f_i) \qquad \text{// eigenvalue shift due to } \delta T \text{ perturbation} \qquad (13)$$

Alternative Approach. Suppose we start with this eigenvalue equation where ϕ_i need not be normalized, and where T is real symmetric,

$$T \phi_i = \lambda_i \phi_i$$

Apply variation δ to get

$$(\delta T) \phi_i + T(\delta\phi_i) = (\delta\lambda_i)\phi_i + \lambda_i (\delta\phi_i) \qquad (13)$$

or

$$(\delta\lambda_i)\phi_i = (\delta T) \phi_i + [\ T - \lambda_i 1] (\delta\phi_i) \qquad (14)$$

Now close this in a scalar product with $(\phi_i$, to get

$$(\delta\lambda_i)(\phi_i , \phi_i) = (\phi_i , (\delta T) \phi_i) + (\phi_i , [\ T - \lambda_i 1] (\delta\phi_i))$$

The last scalar product can be written as $(\ [\ T - \lambda_i 1]\phi_i , \ (\delta\phi_i))$ as long as T is Hermitian. But then we know that $[\ T - \lambda_i 1]\phi_i = 0$, so we find that $(\phi_i , [\ T - \lambda_i 1] (\delta\phi_i)) = 0$ and we get our desired result

$$(\delta\lambda_i)(\phi_i , \phi_i) = (\phi_i , (\delta T) \phi_i) \Rightarrow (\delta\lambda_i) = (\phi_i , (\delta T) \phi_i) / (\phi_i , \phi_i) \qquad \text{QED.}$$

So we do NOT claim that $(\delta\phi_i) \sim \varepsilon^2$ or anything like that! We need both the $(\delta\phi_i)$ terms and we only get these terms to "go away" when we do the scalar product.

We are done, but just for fun here is another way to get this same result:

From above we have

$$(\delta\phi_i) = \varepsilon \Sigma_j a_{ij}\phi_j \qquad (15)$$

so that

$$[T - \lambda_i 1] (\delta\phi_i) = [T - \lambda_i 1] \varepsilon \Sigma_j a_{ij}\phi_j = \varepsilon \Sigma_j a_{ij} [T - \lambda_i 1]\phi_j = \varepsilon \Sigma_j a_{ij} (\lambda_j - \lambda_i) \phi_j = \varepsilon \Sigma_{j\neq i} a_{ij} (\lambda_j - \lambda_i) \phi_j \qquad (16)$$

Now watch what happens when we close with (ϕ_i,

$$(\phi_i, [T - \lambda_i 1] (\delta\phi_i)) = (\phi_i, \varepsilon \Sigma_{j\neq i} a_{ij} (\lambda_j - \lambda_i) \phi_j) = \varepsilon \Sigma_{j\neq i} a_{ij} (\lambda_j - \lambda_i) (\phi_i,\phi_j) = 0$$

In other words, the quantity [T - λ_i 1] ($\delta\phi_i$) is orthogonal to ϕ_i to this order of ε. Then when we do this same closure onto (14) we get

$$(\delta\lambda_i) = (\phi_i, (\delta T) \phi_i)$$

which agrees with the "classic result".