Mathematics recap
– closing chapter of linear algebra –

Some notes about (☺) on page 41:

⊖ Why is the $\underline{\underline{\lambda}}^{diag}$ form interesting?

First note, that both matrix $\underline{\underline{A}}$ and matrix $\underline{\underline{\lambda}}^{diag}$ are representations of the same operator $\hat{A}$, just on a different basis.

{ See the proof of this in SZO, chapter 1.1.5, page 13. }

In case we wanted to compute functions of operator $\hat{A}$, e.g. $\hat{A}^2$ or $\hat{A}^{1/2}$, then it is easy with the diagonal matrix representation, but can be hard with $\underline{\underline{A}}$ in general.

E.g. $$\left(\underline{\underline{\lambda}}^{diag}\right)^2 = \begin{pmatrix} \lambda_1^2 & & \\ & \lambda_2^2 & \\ & & \ddots \end{pmatrix}$$

$$\left(\underline{\underline{\lambda}}^{diag}\right)^{1/2} = \begin{pmatrix} \lambda_1^{1/2} & & \\ & \lambda_2^{1/2} & \\ & & \ddots \end{pmatrix}$$

{HW:} Check that the square of the above $\left(\underline{\underline{\lambda}}^{diag}\right)^{1/2}$ is indeed $\underline{\underline{\lambda}}^{diag}$.

We are usually interested in operator $\hat{A}$, and do not mind particularly which basis we use for representing it as a matrix. However, the basis that represents $\hat{A}$ as a diagonal matrix is very convenient. The importance of (☺) on page 41 is that it provides a recipe for obtaining a diagonal matrix representation of $\hat{A}$.

⊖ When $\hat{A}$ is self-adjoint or Hermitian $\underline{\underline{\lambda}}^{diag}$ and $\underline{\underline{V}}$ have special properties.

T→ Eigenvalues of Hermitian operators are real.

P→ Start from the eigenvalue equation

$$\hat{A}\,|v_i\rangle = \lambda_i\,|v_i\rangle \qquad (\lrcorner)$$

(New notation: $|v_i\rangle$ instead of $\underline{v}_i$.)

Multiply by $|v_i\rangle^{+}$ and integrate.

(New notation further: $|v_i\rangle^{+} = \langle v_i|$
Whenever a "bra" vector $\langle v_i|$ meets a "ket" vector $|v_j\rangle$, i.e. $\langle v_i||v_j\rangle$ it is denoted by $\langle v_i|v_j\rangle$ and a scalar product is implied. So-called bra-ket notation.)

Hence we get

$$\underbrace{\langle v_i | \hat{A} | v_i \rangle} = \langle v_i | \lambda_i | v_i \rangle = \lambda_i \underbrace{\langle v_i | v_i \rangle} \quad (6)$$

Better write this as $\langle v_i | \hat{A} v_i \rangle$ for now, to indicate that $\hat{A}$ acts to the right.

This can be assumed 1, c.f. homogeneous equation (5)

Now take the complex conjugate of the above

$$\langle v_i | \hat{A} v_i \rangle^* = \lambda_i^*$$

And use the properties of the scalar product

$$\langle \hat{A} v_i | v_i \rangle = \lambda_i^*$$

Now we move operator $\hat{A}$ from the "bra" of the scalar product to the "ket", which can be done at the expense of adjungation, c.f. page 37 of these notes

$$\langle v_i | \hat{A}^+ v_i \rangle = \lambda_i^* \quad (7)$$

Since the left hand side of (6) and (7) are the same we can read:

$$\lambda_i = \lambda_i^*$$

$\boxed{\text{QED}}$

I → Eigenvectors of Hermitian operators corresponding to $\lambda_i \neq \lambda_j$ are orthogonal.

P → Start from the eigenvalue equation for $|v_i\rangle$ and $|v_j\rangle$:

$$\hat{A}\,|v_i\rangle = \lambda_i\,|v_i\rangle \qquad \binom{\wedge}{\wedge}$$

$$\hat{A}\,|v_j\rangle = \lambda_j\,|v_j\rangle$$

Take the adjoint of the latter :

$$\langle v_j|\,\hat{A}^{+} = \langle v_j|\,\lambda_j^{*} \qquad \binom{\vee}{\vee}$$

(Note: you can take the adjoint in two steps, first $|\hat{A}\,v_j\rangle^{+} = \langle \hat{A}\,v_j| = \langle v_j|\hat{A}^{+}$ hence taking $\hat{A}$ from the "bra" costs writing its adjoint in parallel with complex conjugation of numbers when taken out from the "bra".)

Multiply $\binom{\wedge}{\wedge}$ by $\langle v_j|$ and $\binom{\vee}{\vee}$ by $|v_i\rangle$ and subtract the two :

$$\underbrace{\langle v_j|\hat{A}\,v_i\rangle - \langle v_j|\hat{A}^{+}\,v_i\rangle}_{\text{since } \hat{A}=\hat{A}^{+} \text{ this side is zero.}} = \underbrace{\lambda_i\,\langle v_j|v_i\rangle - \lambda_j^{*}\,\langle v_j|v_i\rangle}_{\underbrace{(\lambda_i - \lambda_j)}_{\text{This is nonzero.}}\,\langle v_j|v_i\rangle}$$

Then this must be zero.

I → linear combination of degenerate eigenvectors is also an eigenvector.

Terminology: $|v_i\rangle$ and $|v_j\rangle$ are degenerate eigenvectors when $\lambda_i = \lambda_j$

P → {HW:} Take $|w\rangle = \gamma_i |v_i\rangle + \gamma_j |v_j\rangle$ substitute into $\hat{A}|w\rangle$, utilize (1) on page 43 and see what you get.

Remark: {home study:} G. Example 14.4 on page 221
This is a nice example of linear combination, of degenerate eigenvectors. The operator under consideration there is

$$\hat{A} = -\frac{1}{2}\underbrace{\langle \hat{\underline{\nabla}} | \hat{\underline{\nabla}} \rangle}_{\substack{\text{This is denoted} \\ \text{by } \hat{\Delta} \text{ often,} \\ \text{and called the} \\ \text{Laplacian operator}}} - \frac{1}{r}$$

The degenerate eigenvalue is:

$$\lambda = -\frac{1}{8}$$

The eigenfunctions of $\hat{A}$ are regarded as depending on the so-called spherical polar coordinates of a point in 3D space, that are: r, ϑ, φ instead of Cartesian coordinates: x_1, x_2, x_3.

In case you are not familiar with spherical polar coordinates, the relation to Cartesians reads:

$$x_1 = r \cdot \sin\vartheta \cos\varphi$$

$$x_2 = r \cdot \sin\vartheta \sin\varphi$$

$$x_3 = r \cdot \cos\vartheta$$

c. f. chapter G 3.1., page 33.

This example was a detour, but at least it already shows that the things we are now learning will be useful for speaking about electrons in atoms and molecules.

For those of you who know these things from previous studies:

- $\hat{A}$ is the Hamiltonian of the hydrogen atom in the so called Born-Oppenheimer approximation
- λ is the eigenvalue, of the triply degenerate p-orbitals
- superposition is allowed since G (14.9) is a homogeneous DE

Let us get back to pure mathematics for a while yet!

A consequence of Theorem on page 46 is that degenerate eigenvectors can be chosen as orthonormal. Together with Theorem on top of page $\boxed{45}$ this means that matrix $\underline{\underline{V}}$ introduced on page $\boxed{40}$ is unitary. (Of course we suppose that $|v_i\rangle$ are normalized.)

Let us check this!

Going back to page $\boxed{40}$ we see, that the eigenvector index was introduced as column index. Hence we can write

$$\begin{array}{cc} & \begin{array}{cccc} |v_1\rangle & |v_2\rangle & \cdots & |v_N\rangle \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \underline{\underline{V}} & \left(\begin{array}{cccc} & & & \\ \\ \\ \end{array} \right) \end{array}$$

$$\begin{array}{cc} \underline{\underline{V}}^+ & \left(\begin{array}{c} \\ \\ \\ \\ \end{array} \right) \end{array} \begin{array}{l} \leftarrow \langle v_1| \\ \leftarrow \langle v_2| \\ \quad\vdots \\ \leftarrow \langle v_N| \end{array}$$

That is, the eigenvectors are filled into columns of $\underline{\underline{V}}$ with their components, and their adjoints are filled into rows $\underline{\underline{V}}^+$ with their components.

Then the product gives

$(\underline{\underline{V}}^+ \underline{\underline{V}})_{ij} = \langle v_i | v_j \rangle = \delta_{ij}$, proving unitary nature of $\underline{\underline{V}}$.

Note: it is relation

$$\underline{\underline{V}}^{-1} = \underline{\underline{V}}^{+}$$

that allows to call the transformation on the left hand side of $(\odot)$ on page 41 a change of basis. (Note, that G&O uses the adjoint operator in chapter 1.1.5, not the inverse)

Remark: eigenvectors of Hermitian operators provide an ON basis. This is particularly used in L_2 space. Orthogonal polynomials are in fact generated as eigenfunctions of Hermitian operators. Spherical harmonic functions mentioned on page 20 of these notes are eigenfunctions of the so-called angular momentum operator. You will see more on this in the advanced physical chemistry course.

home study: G. Example 19.8 on page 308

This shows you that Hermite polynomials are obtained from the eigenvalue equation corresponding to the quantum mechanical description of the harmonic oscillator. A famous problem.

We have two more important things to address before stepping to quantum mechanics. One is on simultaneous diagonalizability of matrices.

Terminology: Operators $\hat{A}$ and $\hat{B}$ are simultaneously diagonalizable when their eigenvectors match. Note, that in this case they are represented by a diagonal matrix in the same basis.

$\underline{T} \rightarrow$ Hermitian operators $\hat{A}$ and $\hat{B}$ are simultaneously diagonalizable if and only if $[\hat{A}, \hat{B}] = \emptyset$.

$\underline{P} \rightarrow$ Assume first that $\hat{A}$ and $\hat{B}$ are simultaneously diagonalizable. In this case we can write

$$\underline{\underline{A}} = \underline{\underline{V}}\, \underline{\underline{a}}^{diag}\, \underline{\underline{V}}^{+}$$

$$\underline{\underline{B}} = \underline{\underline{V}}\, \underline{\underline{b}}^{diag}\, \underline{\underline{V}}^{+}$$

based on ☺ on page (24)

Start now from $\underline{\underline{A}}\,\underline{\underline{B}}$ and head for $\underline{\underline{B}}\,\underline{\underline{A}}$:

$$\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{V}}\, \underline{\underline{a}}^{diag} \underbrace{\underline{\underline{V}}^{+} \underline{\underline{V}}}_{\underline{\underline{E}}}\, \underline{\underline{b}}^{diag}\, \underline{\underline{V}}^{+} \underset{\substack{\text{diagonal} \\ \text{matrices} \\ \text{commute}}}{=} \underline{\underline{V}}\, \underline{\underline{b}}^{diag} \underset{\substack{\text{insert} \\ \underline{\underline{V}}^{+}\underline{\underline{V}}}}{\uparrow} \underline{\underline{a}}^{diag}\, \underline{\underline{V}}^{+} =$$

$\xrightarrow[\text{cont'd}]{P}$ $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{V}}\,\underline{\underline{b}}^{diag}\,\underline{\underline{V}}^{+}\underline{\underline{V}}\,\underline{\underline{a}}^{diag}\,\underline{\underline{V}}^{+} = \underline{\underline{B}}\,\underline{\underline{A}}$ ✓ one direction

$\boxed{\text{HW:}}$ Prove, that diagonal matrices commute.

Assume next, that $[\hat{A}, \hat{B}] = \emptyset$ and also assume, that $\hat{A}|v_i\rangle = a_i|v_i\rangle$ holds.

Now represent

$$\hat{A}\hat{B} - \hat{B}\hat{A} = 0$$

on the basis of $\{|v_i\rangle\}_{i=1}^{N}$ to get

$$\sum_{j=1}^{N} A_{ij}B_{jk} - B_{ij}A_{jk} = 0$$

and utilize, that $A_{ij} = a_i\,\delta_{ij}$ to get

$$a_i\,B_{ik} - B_{ik}\,a_k = 0$$

$$B_{ik}\,(a_i - a_k) = 0$$

Now we have to consider two cases. Either $a_i \neq a_k$ which means that $B_{ik} = \emptyset$ ✓

Or $a_i = a_k$ in which case B_{ik} can be nonzero.

Since $\underline{\underline{B}}$ is Hermitian, we can transform this nonzero block of $\underline{\underline{B}}$ into diagonal form

$\xrightarrow{P}$ cont'd

and this does not spoil the eigenvalue equation of $\hat{A}$ since $|v_i\rangle$ and $|v_k\rangle$ correspond to degenerate eigenvalues, and can be combined without making harm, c.f. page ~~46~~.

Important message: when $\hat{A}$ and $\hat{B}$ commute, the matrix representation of $\hat{B}$ on the eigenvectors of $\hat{A}$ can be nondiagonal only in the block where the corresponding eigenvectors have degenerate eigenvalues with $\hat{A}$.

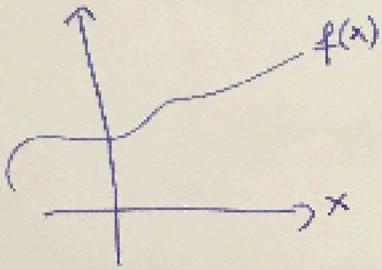
E.g.

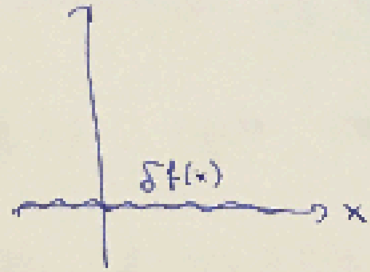
$$\underline{\underline{A}} \begin{pmatrix} 1 & & & \emptyset \\ & 2 & & \\ & & 3 & \\ \emptyset & & & 3 \end{pmatrix} \qquad \underline{\underline{B}} \begin{pmatrix} 4 & & & \emptyset \\ & 5 & & \\ & & 6 & 7 \\ \emptyset & & 8 & 9 \end{pmatrix}$$

One last thing from maths. Eigenvalue equations can be formulated as so-called variational problems.

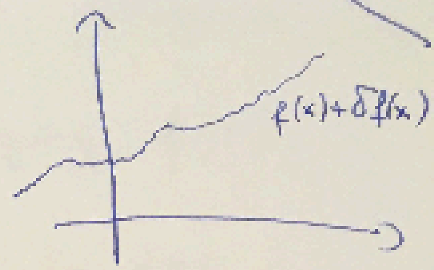
What is a variation?

It is an arbitrary small alteration of a function. E.g.





this is called "variation f" as a function



this is the varied function; altered with respect to $f(x)$ only slightly

One is the interested in the change of the value of expressions, upon changing $f(x)$ for $f(x)+\delta f(x)$.

E.g. what is the variation of $\langle f|f\rangle$? Let us see:

$$\langle f+\delta f|\, f+\delta f\rangle = \langle f|f\rangle + \underbrace{\langle \delta f|f\rangle + \langle f|\delta f\rangle}_{\text{this is the so-called first variation of the scalar product; notation: } \delta\langle f|f\rangle} + \underbrace{\langle \delta f|\delta f\rangle}_{\text{this is smaller by an order of magnitude, than the first variation, gets neglected}}$$

Hence we have:

$$\delta\langle f|f\rangle = \langle \delta f|f\rangle + \langle f|\delta f\rangle$$ just like evaluating derivatives

This can be used as a rule of ~~thumb~~ thumb.

HF: Compute the variation of the integral

$$\int_{-1}^{1} \left(f^2(x) + g(x)\right)\left(f(x) - g(x)\right) dx$$ upon varying $f(x)$.

Hint: substitute $f(x) + \delta f(x)$ into $f(x)$ above and read the terms proportional to δf!

The variational approach to the eigenvalue equation can be formulated as follows.

We seek that function f, for which

$$\delta \underbrace{\frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle}}_{} = 0 \qquad (\oplus)$$

Note: this expression is the eigenvalue, if $|f\rangle$ is an eigenfunction, c. f. ($\oplus$) on page 53

Let us evaluate the first variation on the left hand side of ($\oplus$)

$$\delta \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} = \frac{\left[\langle \delta f | \hat{A} f \rangle + \langle f | \hat{A} \delta f \rangle\right] \langle f | f \rangle - \left[\langle \delta f | f \rangle + \langle f | \delta f \rangle\right] \langle f | \hat{A} f \rangle}{\langle f | f \rangle^2}$$

$$= \frac{1}{\langle f | f \rangle} \left\{ \langle \delta f | \hat{A} f \rangle - \langle \delta f | f \rangle \cdot \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} + \text{c.c.} \right\}$$

The symbol c.c. refers to "complex conjugate" of the term preceding it in the { } bracket.
From the bottom of page 54 one can deduce

$$\langle \delta f | \hat{A} f \rangle - \langle \delta f | f \rangle \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} = 0$$

(and the same for its c.c. which we do not write separately)

Rewriting the above we have

$$\langle \delta f | \hat{A} - \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} | f \rangle = 0$$

Since $\langle \delta f |$ is arbitrary we can deduce

$$\hat{A} - \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} | f \rangle = \emptyset$$

$$\hat{A} | f \rangle = \underbrace{\frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle}}_{\text{this is the eigenvalue}, \ \alpha} | f \rangle \qquad (*)$$

Hence we arrived at the eigenvalue equation.

Terminology: $\frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle}$ Rayleigh-quotient

The above derivation tells, that setting the variation of the Rayleigh-quotient ~~to~~ zero

(~~to~~ in other terms: setting it stationary)
is <u>equivalent</u> to solving the eigenvalue
equation.

This is called the <u>variation principle</u>.

The equation ( ) on page 54 is therefore
just an alternative way of writing the
eigenvalue equation ($\oplus$) on page 55.

Based on the variation principle one can
devise a way of solving the eigenvalue
problem of $\hat{A}$. That is assuming a
linear combination with the help of a
set of basis functions $\{|u_i\rangle\}_{i=1}^{N}$

$$|f\rangle = \sum_{i=1}^{N} f_i \, |u_i\rangle \qquad ( )$$

seek the linear combination
coefficients based on

$$\delta \frac{\langle f|\hat{A} f\rangle}{\langle f|f\rangle} = \emptyset$$

that is better written now as

$$\frac{\partial}{\partial f_i} \frac{\langle f|\hat{A} f\rangle}{\langle f|f\rangle} = 0$$

since these are only the
numbers f_i that we can alter

Let us then evaluate the derivative at the bottom of page 56; using notation

$\langle u_i | \hat{A} u_j \rangle = A_{ij}$ and assuming $\langle u_i | u_j \rangle = \delta_{ij}$.

First we write the Rayleigh-quotient:

$$\frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} = \frac{\sum_{ij} f_i^* A_{ij} f_j}{\sum_i f_i^* f_i}$$

$$\frac{\partial}{\partial f_k} \frac{\langle f | \hat{A} f \rangle}{\langle f | f \rangle} = \frac{\left(\sum_j A_{kj} f_j + \sum_i f_i^* A_{ik} \right) \sum_i f_i^* f_i}{\left(\sum_i f_i^* f_i \right)^2} -$$

$$- \frac{(f_k + f_k^*)}{\sum_i f_i^* f_i} \cdot \underbrace{\frac{\sum_{ij} f_i^* A_{ij} f_j}{\sum_i f_i^* f_i}}_{\alpha}$$

In the same manner as before, the condition

$$\frac{\partial}{\partial f_k} \frac{\langle f | A f \rangle}{\langle f | f \rangle} = 0$$

now gives us

$$\sum_j A_{kj} f_j - \alpha f_k = \phi \qquad (☺)$$

$$\sum_j A_{kj} f_j = \alpha f_k$$

This is the matrix eigenvalue equation, written with indices.

We consequently obtained a matrix eigenvalue equation by ~~the~~ assuming form (*) of the eigenfunction on page [56]. This procedure is called the linear variational method, or Ritz method.

{HW:} Take the functions of § 19.53 on page 308 as basis functions, and devise the problem to solve when we aim at the eigenvalue equation of the operator

$$\hat{A} = \frac{d^2}{d\xi^2} - \xi^2 + \lambda \xi^4$$

where λ is an arbitrary number. Note that the above $\hat{A}$ differs from the differential operator of Example 19.8 only in the last term.

Hint: you do not need to solve this problem, but let us suppose you had to implement it as a computer program. What would you do then? What would the computer have to evaluate?

Remark: The matrix eigenvalue equation obtained in (●) on page 57 is just an approximate solution in practice, since the basis we can use in practice is finite dimensional, while the ~~operator has~~ true eigenfunction resides in an infinite dimensional Hilbert-space.

There is however an important theorem on the nature of the approximate eigenvalue associated with the normalized function

$$|f\rangle = \sum_{i=1}^{N} f_i \, |u_i\rangle$$

For this end rewrite $|f\rangle$ on the basis of the exact normalized eigenfunctions of $\hat{A}$

$$|f\rangle = \sum_{i=1}^{N} h_i \, |v_i\rangle$$

And compute α as

$$\alpha = \sum_{ij} h_i^* \underbrace{\langle v_i | \underbrace{\hat{A} v_j}_{\lambda_j |v_j\rangle} \rangle}_{\delta_{ij}\lambda_i} h_j = \sum_i |h_i|^2 \lambda_i \geq \lambda_1$$

For this reason

$$\alpha \geq \lambda_1 \underbrace{\sum_i |h_i|^2}_{1}$$

$$\alpha \geq \lambda_1$$

This is an extremely important result, telling that the Rayleigh-quotient approximates the bottommost eigenvalue from above when evaluated with a trial function. This is called the variational theorem, not to be mixed with the variational principle.