

# More Hermitian : adjoint and self-adjoint

Hermitian means “**self-adjoint**”

What is **adjoint**?                      adjoint of  $\hat{A} \equiv \hat{A}^\dagger$

Adjoint means: take the complex conjugate and **operate backwards**

$$\int f^* \hat{A}^\dagger g \, d\tau = \int g \hat{A}^* f^* \, d\tau$$

The adjoint operator does **not** generally give the same result as the original operator.

“The adjoint of an operator  $\mathbf{A}$  may also be called the **Hermitian adjoint**, **Hermitian conjugate** or **Hermitian transpose**” [https://en.wikipedia.org/wiki/Hermitian\\_adjoint](https://en.wikipedia.org/wiki/Hermitian_adjoint)

Hermitian = self adjoint means  $\hat{A} = \hat{A}^\dagger$

$$\int f^* \hat{A} g \, d\tau = \int f^* \hat{A}^\dagger g \, d\tau = \int g \hat{A}^* f^* \, d\tau = \int g (\hat{A} f)^* \, d\tau$$

In other words, one gets the same number whether using a certain operator or using its adjoint, which leads to the definition used in the previous lecture.

**Other identities:**

$$|n\rangle^\dagger = \langle n|$$

$$\langle m|n\rangle^\dagger = \langle n|m\rangle = \langle m|n\rangle^*$$

$$\langle m|\hat{A}|n\rangle \equiv A_{mn}$$

$$\langle m|\hat{A}^\dagger|n\rangle = \langle n|\hat{A}^*|m\rangle = A_{mn} = A_{nm}^*$$

i.e., the definition of a **Hermitian matrix**

# Math of QM = math of analytic geometry (relation to row and column vectors)

ket = column vector      bra = row vector

$$|a\rangle = \begin{array}{|c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array}$$

$$\langle a| = \begin{array}{|c} a_1^* & a_2^* & a_3^* & a_4^* \end{array}$$

dot product = scalar product = inner product

$$\langle a | a \rangle = \begin{bmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} a_1^* a_1 + a_2^* a_2 + a_3^* a_3 + a_4^* a_4 \end{bmatrix}$$

a positive, real number

= 1 if  $|a\rangle$  is normalized

$$\langle a | b \rangle = \begin{bmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = a_1^* b_1 + a_2^* b_2 + a_3^* b_3 + a_4^* b_4$$

a real or complex number

= 0 if  $|a\rangle$  and  $|b\rangle$  are orthogonal

# Matrices and Matrix Multiplication

row – column rule:

the row subscript is always FIRST e.g.,  $a_{rc}$

Matrix multiplication:  $\mathbf{A} \bullet \mathbf{B} = \mathbf{C}$

$$\begin{array}{|c|} \hline a_{11} & a_{12} & a_{13} \\ \hline a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \\ \hline \end{array} \quad \begin{array}{|c|} \hline b_{11} & b_{12} & b_{13} \\ \hline b_{21} & b_{22} & b_{23} \\ \hline b_{31} & b_{32} & b_{33} \\ \hline \end{array} = \begin{array}{|c|} \hline c_{11} & c_{12} & c_{13} \\ \hline c_{21} & c_{22} & c_{23} \\ \hline c_{31} & c_{32} & c_{33} \\ \hline \end{array}$$

The result  $\mathbf{C}$  of “multiplying” two matrices  $\mathbf{A}$  and  $\mathbf{B}$  is all possible **scalar products** of the **rows of  $\mathbf{A}$**  with the **columns of  $\mathbf{B}$** . In other words:

$$C_{mn} = \sum_k A_{mk} B_{kn}$$

Non-square matrices may be multiplied, **but only if the number of columns of the left-hand matrix = the number of rows in the right-hand matrix**

For example a 4 row by 3 column matrix x a 3 row, 2 column matrix  
(4 by 3 x 3 by 2) yields a 4 row by 2 column matrix

$$\begin{array}{|c|} \hline \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{array} \\ \hline \end{array} \times \begin{array}{|c|} \hline \begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{array} \\ \hline \end{array} = \begin{array}{|c|} \hline \begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{array} \\ \hline \end{array}$$

So, a scalar product of two vectors:

$$\langle a | b \rangle = \begin{bmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} a_1^* b_1 + a_2^* b_2 + a_3^* b_3 + a_4^* b_4 \end{bmatrix}$$

a number

Is: a 1 by 4 matrix multiplying a 4 by 1 matrix

= ?

a 1 by 1 matrix = **a number**

# Outer Products

Outer Product =  $|b\rangle\langle a|$  = Operator

let  $\langle a|$  and  $\langle b|$  be normalized and orthogonal:

$$|b\rangle\langle a| \text{ operating on } |a\rangle = |b\rangle\langle a|a\rangle = |b\rangle$$

$$|a\rangle\langle b| \text{ operating on } |a\rangle = |a\rangle\langle b|a\rangle = 0$$

$|b\rangle\langle a|$  = column vector x row vector

$|b\rangle\langle a|$  = column vector x row vector

$$|b\rangle\langle a| = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \begin{bmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{bmatrix} = \begin{matrix} \text{a } 4 \text{ by } 1 \text{ matrix} \\ \text{x a } 1 \text{ by } 4 \text{ matrix} \\ \\ \text{= a } 4 \text{ by } 4 \text{ "matrix"'} \end{matrix}$$

$$|b\rangle\langle a|a\rangle = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \begin{bmatrix} a_1^* & a_2^* & a_3^* & a_4^* \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = |b\rangle$$

$$|a\rangle\langle b|a\rangle = \begin{array}{|c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array} \begin{array}{|c} b_1^* \quad b_2^* \quad b_3^* \quad b_4^* \end{array} \begin{array}{|c} a_1 \\ a_2 \\ a_3 \\ a_4 \end{array} = 0$$

## Relation between wavefunctions and vectors:

In the previous lecture we noted that:

$$\int \Psi_m^*(x) \Psi_n(x) dx \equiv \langle \Psi_m | \Psi_n \rangle \equiv \langle m | n \rangle \equiv A_{mn}$$

which clearly means that the common “overlap” integral on the left must be an *inner product* of two vectors.

In what sense is can we think of the integral as the scalar product of two vectors when we are used to think of the wavefunctions in the integral as common functions of  $x$ ?

Before giving a formal answer in terms of the Dirac delta function, one can simply recall the definition of acceptable quantum mechanical state functions, for which at every point in space there exists a single complex number, the absolute square of which is proportional to probability density for finding the particle. This is **Postulate 1** of quantum mechanics in most text books.

Imagine, now, *infinitely* long column and row vectors in which are stored all the numbers which are the values of  $\Psi(x)$  and  $\Psi^*(x)$ , respectively at all possible values of  $x$ .

Thus, all integrals in this notation are actually just scalar products of two vectors. They are formally infinite because in calculus the result is only exactly correct in the limit of  $dx \rightarrow 0$ . In all actual calculations of numbers, on computers,  $dx$  is a very small—but finite— $\Delta x$ . This truly turns the function into a vector of numbers, as large as we care to make by using smaller values of  $\Delta x$  and larger range of  $x$ . This is a fine, because the human mind cannot grasp infinitesimal and infinity.

This idea is true for any number of coordinates, of course.

## Wavefunctions are “representations of bra and ket vectors”:

Expanding arbitrary state functions as linear combinations of the complete set of eigenfunctions of an observable, dynamical variable operator **A**:

An arbitrary state vector  $|f\rangle$  may be expressed as a superposition (linear combination) of members of a complete set:  $|f\rangle = \sum_k c_{kf} |k\rangle$

where  $\hat{A}|k\rangle = a_k|k\rangle$ , where  $a_k$  are the eigenvalues of  $\hat{A}$ , and  $c_{kf} = \langle k|f\rangle$

The numbers  $c_{kf}$  are said to be: “ **$|f\rangle$  in the **A** representation”.**

One may of course use any other complete set to “represent”  $|f\rangle$ . For example, the  $c_{kf}$  from the complete set using the eigenfunctions of the Hermitian operator,  $d^2/dx^2$ , i.e.,  $\sin(kx)$  and  $\cos(kx)$ , is the Fourier representation, better known as the **Fourier Transform**.

The set of numbers  $\langle k'|B|k\rangle$  is similarly said to be the operator **B** in the **A** representation.

### The Identity operator

Given that the  $c_{kf}$  are numbers  $= \langle k|f\rangle$

$$|f\rangle = \sum_k c_{kf} |k\rangle = \sum_k |k\rangle c_{kf} = \sum_k |k\rangle \langle k|f\rangle$$

$$|f\rangle = \sum_k |k\rangle \langle k|f\rangle$$

Thus,  $\sum_k |k\rangle \langle k| = \hat{I}$ , the very widely used "Identity" operator

Next we introduce the special complete set known as the “position eigenfunctions.

# Position Eigenfunctions

The position operator  $x_{op}$  is simple multiplication by the position  $x$ . The eigenfunctions are the ingenious and widely used Dirac delta functions, with the symbol  $\delta(x-a)$ , where  $a$  is a real number, and the properties:

$$\delta(x - a) = 0 \quad \text{if } x \neq a$$

$$\delta(x - a) = \infty \quad \text{if } x = a$$

$$\int_{-\infty}^{\infty} \delta(x - a) dx = 1$$

$$\hat{x}\delta(x - a) = a\delta(x - a)$$

To rationalize the normalization, one may think of the function near  $x=a$  to be very large but finite in a very narrow, but finite, region such that area under the curve = 1.

Thus expanding  $|f\rangle$  in the eigenfunctions of  $x$ , i.e., the Dirac delta function, changes  $|f\rangle$  into the wavefunction  $\Psi_f(x) = |f\rangle$  **in the position representation**, in analogy with the previous page.

# “The” Variation Principle and Variation Methods

One cannot overemphasize the immense importance of “the” variation principle to quantum chemistry, but it is equally important in numerous other fields, many of them in non-quantum engineering applications. It turns out that eigenvalues and eigenvectors always emerge when a differential equation is subjected to boundary conditions. Schrödinger himself invoked a variation principle to support his formulation of the Schrödinger Equation.

Quoting from [https://en.wikipedia.org/wiki/Variational\\_principle](https://en.wikipedia.org/wiki/Variational_principle): A **variational principle** is a scientific principle used within the [calculus of variations](#), which develops general methods for finding functions which extremize the value of quantities that depend upon those functions. “For example...the shape of a chain suspended at both ends... is found by minimizing the gravitational potential energy”

## Examples

- [Lord Rayleigh's variational principle](#)
- [Ekeland's variational principle](#)
- [Fermat's principle](#) in [geometrical optics](#)
- The [principle of least action](#) in [mechanics](#), [electromagnetic theory](#), and [quantum mechanics](#)
- [Maupertuis' principle](#) in [classical mechanics](#)
- The [Einstein equation](#) also involves a variational principle, the [Einstein–Hilbert action](#)
- [Gauss's principle of least constraint](#)
- [Hertz's principle of least curvature](#)
- [Palatini variation](#)
- The [variational method](#) in quantum mechanics
- The [finite element method](#)

# The Variation Principle of Quantum Systems

$$\frac{\int \phi_{trial}^* H \phi_{trial} d\tau}{\int \phi_{trial}^* \phi_{trial} d\tau} = \langle E \rangle \geq E_0$$

The power of the variation principle is that **any** approximate wavefunction for the ground state of a quantum system is guaranteed to give an energy expectation value,  $\langle E_{approx} \rangle$ , that will be higher than the true ground state energy,  $E_0$ . This allows one to vary the “shape” of the trial wavefunction by any means available until the derivative of  $\langle E_{approx} \rangle$  with respect to all variables being varied = 0, and be satisfied that this will be the most accurate wave function of the type being used as  $\phi_{trial}$ .

There is no danger of finding an energy that is too low.

In addition to the above “wavefunction” scheme, Kohn and Sham shared the Nobel Prize recently for proving and helping implement the so called density functional theory (DFT). With DFT, one directly varies the electron density  $\phi_{trial}^* \phi_{trial}$  to reach the energy minimum, instead of varying  $\phi_{trial}$ , which must always be squared to get the energy anyway.

There are different variation methods which can be classified into two broad classes: (1) linear variation method, and (2) non-linear variation methods. We will begin with **the linear variation method**, which pervades most of the computational methods we will encounter in chemistry.