Basic mathematical tools

The following pages provide a brief summary of the mathematical material used in the lectures of 442. No rigorous proofs are provided. The reader unfamiliar with the material will find a distillation of the "bare-bones" facts, and may wish to read up on more details in the appropriate mathematics textbook, such as Butkov, *Applied Mathematics*.

A.1 Complex variable analysis

1.1 Complex numbers

The Schrödinger equation is a complex-valued differential equation, and its solutions are complex functions. This behavior of the wavefunction is very important as the basis of quantum mechanical interference effects, and a solid understanding of complex variable analysis is necessary to really appreciate the properties of wavefunctions.

We start with the usual definition $i^2 = -1$, in terms of which a general complex number can be written as $z = x + iy$. A convenient way to think of a complex number is as a vector in a 2-dimensional coordinate system, with two axes: the ‘real’ axis $|R$ along which lies the real part of $z$ (the “$x$”); and an ‘imaginary axis $|I$ along which lies the imaginary part $iy$. This coordinate system is called ‘the complex plane.”

For a product of two complex numbers the standard rules of algebra then yield

$$(x + iy)(c + id) = xc - yd + i(xd + yc)$$

The complex conjugate of $z$: $x + iy$ is defined as

$$z^* = x - iy,$$ with $zz^* = x^2 + y^2$. 

The latter expression appears like the square modulus of a cartesian vector in two dimensions. Thus, we can view complex numbers as 2-D cartesian vectors, with length given by

$$|z| = \sqrt{|z|^2} = \sqrt{zz^*}$$

Addition and multiplication follow the usual associative rules.

1.2 Functions of a complex variable: $g(z) = g(x + iy)$

Functions of a complex variable can be defined in terms of analytic continuation of real functions into the complex plane. In practice, this means that a complex variable can be
substituted into standard expressions instead of a real variable, and manipulated using the algebra outlined in a). For example, the exponential function becomes

\[ e^z = e^{x+iy} = e^x e^{iy}; \]

\[ e^{iy} = 1 + iy - \frac{1}{2} y^2 - \frac{i}{3!} y^3 + \ldots \]

The Taylor expansion in the last line corresponds to \( \exp[iy] = \cos[y] + i \sin[y] \), the famous Euler formula for the exponentiation of an imaginary number. Note that the square modulus of \( \exp[iy] \) is always 1, and \( y \) can be interpreted as an angular phase factor as it varies from 0 to \( 2\pi \). Any complex number can be written in this exponential form:

\[ x + iy = re^{i\theta}; \quad r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1} \frac{y}{x}. \]

In a similar way, any function of complex variables can be expressed in terms of function of real variables. For example,

\[ \ln z = \ln(x + iy) = \ln(\text{re}^{i\theta}) = \ln r + \ln e^{i\theta} = \ln r + i\theta, \]

so the natural logarithm of a complex number automatically yields \( r \) and \( \theta \) separated in its real and imaginary parts. Since the \( \cos[] \) and \( \sin[] \) function in the Euler formula cycle after \( 2\pi \), it is clear that some complex functions will be multivalued, just as the real square-root function is twovalued. Consider the square root function

\[ f = \sqrt{z} = \frac{1}{2} z^{1/2} = \left(\text{re}^{i\theta}\right)^{1/2} = r^{1/2} e^{i\theta/2}. \]

This function does not return to its original value after we traverse a circle in the complex plane and return to the x axis:

\[ \theta \rightarrow 0 \text{ to } 2\pi, \quad f \rightarrow r^{1/2} \text{ to } -r^{1/2} \]

\[ \theta \rightarrow 2\pi \text{ to } 4\pi, \quad f \rightarrow -r^{1/2} \text{ to } +r^{1/2} \]

However, we like to think of them as one function with two Riemann surfaces, as shown in the plot below. The seam where the sheets interleave is the “branch cut.”

Fig. A-2. Double-leafed Riemann sheet for a double-valued complex function such as \( \sqrt{z} \).
1.3 Derivatives of wavefunctions

Wavefunctions are complex-valued functions, but of a real variable – \( x \). This distinguishes them from real complex functions of a complex variable \( z = x + iy \). For this reason, when we introduced \( \Psi \), we treated it as a pair of functions \( \Psi^{(\text{r})}(x) \) and \( \Psi^{(\text{i})}(x) \) (the real and imaginary parts). Thus one can take derivatives of \( \Psi \) as of any other function, using the rules for complex algebra discussed above.

1.4 Complex integrals

Integrals in the complex plane follow a contour line \( C(\ x=x(t),\ y=y(t)\ ) \), much like line integrals in the \( x\)-\( y \) plane. However, since \( f(z) \) cannot be a completely arbitrary function of the real and imaginary parts of \( z \), contour integrals have certain special properties. We can split any complex integral into two separate real and imaginary parts:

\[
\int_C f(z)dz = \int_C \left( u + iv \right)(dx + idy)
\]

\[
= \int_C (udx - vdy) + i\int_C vdx + udy = \int_C F \cdot dx + i\int_C G \cdot dx
\]

where \( F = u(x, y)\hat{x} - v(x, y)\hat{y} \) and \( G = u(x, y)\hat{x} + v(x, y)\hat{y} \). Finally, if we choose a specific curve \( C \) with \( x = x(t) \) and \( y = y(t) \), the line integrals become

\[
\int_C f(z)dz = \int_{C(t)} \left( u \frac{dx}{dt} \right)dt - \left( v \frac{dy}{dt} \right)dt + i\int_{C(t)} \left( G \cdot \frac{dx}{dt} \right)dt
\]

This formula is the general recipe for evaluating complex integrals along any curve in the complex plane. Consider the following example integrating the function \( \exp[z] \) along a parabolic path \( y = x^2 \) in the complex plane from \( x=y=0 \) to \( x=y=1 \):

\[
e^z = e^{x+iy} = e^x(\cos y + i \sin y), \quad x = t, \quad y = t^2;
\]

\[
\int_{0,0}^{1,1} e^z dz = \int_0^1 dt \left( e^t \cos t^2 - 2te^t \sin t^2 \right) + i\int_0^1 dt \left( e^t \sin t^2 + 2te^t \cos t^2 \right)
\]
Sometimes, variable substitution helps. In particular, when the path is circular, a common case, the use of polar coordinates is recommended. For example, consider the integral of the function \( f(z) = \frac{1}{z} \) around the origin on a circle of radius \( r \). Then

\[
\int_C dz f(z) = \int_0^{2\pi} re^{i\theta} d\theta = i \int_0^{2\pi} d\theta = 2\pi i.
\]

For wavefunctions, since \( x \) and \( p \) are real variables in the position and momentum representations, we can consider the contour to lie along the real axis, making things simpler. Then the integral just becomes like two ordinary integrals, one for the real part, one for the imaginary part:

\[
\int dx \Psi_1^*(x) \Psi_2(x) = \int dx \left[ \Psi_1^{(r)}(x) \Psi_2^{(r)}(x) + i \Psi_1^{(i)}(x) \Psi_2^{(i)}(x) \right]
\]

A.2 Fourier transforms

2.1 Definition

Fourier transforms play an important role in quantum mechanics as they convert representations from one conjugate variable to another (e.g. \( x \) to \( p \) or \( \omega \) to \( t \)). When the variables are time \( \tau \) and frequency \( \omega \), the purpose of the Fourier transform is to take an arbitrary function \( g(t) \), and to ‘filter’ out of that function the part that oscillates at frequency \( \omega \). For example, the function \( g(t) = 2\cos(3t) + 4\cos(5t) \) would oscillate at frequencies of \( \omega = 3 \) and \( 5 \). Its Fourier transform therefore should have spikes at those two frequencies, with the spike at frequency “5” twice as large as the spike at frequency “3”.

The desired mathematical property of filtering out the frequencies and turning \( g(t) \) into \( G(\omega) \) or \( g(x) \) into \( G(k) \) (\( k = p \) in units of Planck’s constant) can be achieved by

\[
\text{F.T.}(g(x)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} g(x) dx = G(k) \quad \text{(or } 1 \cdot \int dt \ldots \text{ or } \int dt e^{-i\omega t} \ldots) \quad \text{A-27a}
\]

Many definitions are in use; in math, a symmetrical one with a \( 1/(2\pi) \) factor for the F.T. and another for the inverse F.T. is used. In chemistry or physics text (parentheses), the factor \( 1/(2\pi) \) is often built into just one of the functions:

\[
\text{F.T.}^{-1}(G(k)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{-ikx} G(k) = g(x) \quad \text{(or } \frac{1}{2\pi} \cdot \int d\omega e^{-i\omega t} \ldots) \quad \text{A-27b}
\]
If the function contains just one oscillating frequency component at $\omega_0$, its Fourier transform will be a spike at that frequency, called a “delta function” $\delta(\omega - \omega_0)$, or more accurately a "delta distribution":

$$\text{ex: } F.T.(\sin \omega_0 t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \sin \omega_0 t \, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{2i} (e^{i\omega_0 t} - e^{-i\omega_0 t}) \, dt$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{2i} \int_{-\infty}^{\infty} e^{i(\omega - \omega_0)t} - e^{i(\omega + \omega_0)t} \, dt$$

$$= \frac{1}{2i} \frac{1}{\sqrt{2\pi}} 2\pi \left( \delta(\omega - \omega_0) - \delta(\omega + \omega_0) \right)$$

The example shows that the function $\sin\omega t$ only has frequency contributions at $\omega = \pm |\omega|$. The frequency spectrum is "infinitely" narrow and given by delta-distributions. In general, the function should be absolutely integrable (i.e. $\int |g|dt$ is finite) in order to have a well-defined Fourier transform. Note that for position $x$ as opposed to time, the F.T. and its inverse are often defined with the sign switched, to be compatible with wave expressions of the type $\exp[\text{i}kx - \text{i}\omega t]$. The $\pm$ definition is arbitrary, but of course needs to be used consistently throughout. The figure on the next page shows how functions that are “wide” in the time domain can be mapped into functions “narrow” in the frequency domain, and vice versa. For such Fourier-transform related pairs of functions, the time-frequency product is a constant. The Heisenberg uncertainty principle is just a special case of this, so is Bragg’s law.

2.2 Some useful properties of Fourier transforms

The following equations summarize a series of properties of Fourier transforms:

-Shift yields frequency phase factor:

$$F.T. (g(x - x_0)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixk} g(x - x_0) \, dx = \frac{1}{\sqrt{2\pi}} e^{ikx_0} \int_{-\infty}^{\infty} e^{ixk'} g(x') \, dx' = e^{ikx_0} G(k)$$

-F.T. of derivative turns into product factor (useful for solving differential equations, and important in quantum mechanics to relate $x$ and $p$). This is the relationship that proves that $p \sim \partial/\partial x$ and $x \sim \partial/\partial p$:
F.T. \left( g'(x) \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx g'(x) e^{ikx} = \frac{1}{\sqrt{2\pi}} \left[ g(x) e^{ikx} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} ik e^{ikx} g(x) dx \\
= -ik F \left( g(x) \right); \quad \text{In general } F.T. \left( g^{(n)}(x) \right) = (-ik)^n F.T. \left( g(x) \right)

-F.T. of a convolution is the product of the F.T (here, let’s use the \exp[-i\omega t] definition of the Fourier transform, usually used for time variables):

\[ g \ast h = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' g(t') f(t-t') \quad \Rightarrow \quad F.T. \left( g \ast h \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{-i\omega t} \int_{-\infty}^{\infty} dt' g(t') f(t-t') \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' g(t') \int_{-\infty}^{\infty} dt e^{-i\omega t(t-t')} \quad \text{(let } i'' = t-t' \text{ and } dt = dt'') \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{-i\omega t} g(t') \frac{2\pi}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{-i\omega t''} f(t'') = \sqrt{2\pi} G(\omega) F(\omega) \]

The convolution function occurs in many applications, for example the increase in observed linewidths due to convolution of the natural lineshape with a "broadening factor" due to laser resolution, molecular motion (Doppler effect), etc. In a laser autocorrelation function, the output intensity is obtained by convolving the input electric fields with respect to one another:

\[ I(t_0) = \int_{-\infty}^{\infty} dt A(t) A(t_0 - t) \]

\[ \Rightarrow \quad \int I(t_0) dt = \int |A(\omega)|^2 d\omega \]

The formula at the bottom is a form the Wiener-Khintchine theorem, a fundamental theorem of statistical mechanics relating the integral of power over time to the power in the frequency domain. These F.T. formulas are practical because of a computer algorithm known as the FFT developed by Cooley and Tukey in the early 1960s, which allows rapid evaluation of Fourier transforms.

A.3 Linear algebra, function spaces, Dirac notation

Quantum mechanics uses complex-valued wavefunctions to describe the evolution of a system, operators that act on these wavefunctions to transform them into other wavefunctions, and expectation values of operators, which correspond to observable quantities. This is directly analogous to vectors, matrices that act on them to yield other vectors, and scalar products of vectors with such transformed vectors:

\[ |\Psi > \leftrightarrow v \]

\[ \hat{A} |\Psi > |\Psi > = |\Psi > \leftrightarrow Av = v' \]

A-32

\[ < \Psi | \hat{A} |\Psi > = <\Psi |\Psi > \leftrightarrow v \bullet (Av) = v \bullet v' \]

The analogy is twofold: the function space of wavefunctions is directly analogous to the n-dimensional vector spaces of linear algebra; and by using basis set representations of wavefunctions and operators,
the quantum mechanical problem in effect becomes a linear algebra problem. We briefly review these issues below.

### 3.1 Hilbert space

Among the many notations attempted for quantum mechanics, the most powerful has proven to be in terms of a "function space", in which the functions can be thought of as vectors, much like n-dimensional vectors in ordinary vector space. To be useful for quantum mechanics, such a function space \( h \) must have a number of properties:

- The functions \( \Psi \) must have a norm \( |\Psi| \). In practice \( \int dx \Psi_k^* \Psi_k = 1 \) is the square of this norm (\( \int dx \Psi_k^* \Psi_{k'} = \delta(k-k') \) for free particles).

- There exists a scalar product operation \( < | > \) which maps any two functions \( \Psi, \Psi' \in h \) onto the complex numbers with properties analogous to ordinary scalar products. In practice, this is done by an operation \( \int dx \Psi^*(x) \_ \_ \int \). The complete set of operations \( \int dx \Psi^*(x) \_ \_ \) dual to \( \Psi \), insert \( \Psi' \) in the placeholder \( \_ \_ \). The results is the scalar product \( \langle \Psi | \Psi' \rangle \), a complex number. In particular, \( \sqrt{\langle \Psi | \Psi \rangle} \) is the positive-definite norm of \( \Psi \). Also, \( \langle \Psi' | \Psi \rangle = \langle \Psi | \Psi' \rangle^* \).

The first of these properties is sufficient for a "Banach space" \( h \). An entity \( \mathcal{H} = (h; < | >) \) which consists of a Banach space on which a scalar product is defined, is known as a "Hilbert space". These concepts are entirely analogous to the ones from ordinary linear algebra. For example, an ordinary vector \( v \) is element of a vector space \( V \). The norm of \( v \) is the length of that vector. The dual space of \( V \) has elements \( v' \) which are operations acting on the \( v \in V \), and the scalar product between \( v \) and \( v' \) is defined as \( v \cdot v' \). The norm is \( \sqrt{v \cdot v} \).

For a Hilbert space, we can define a complete orthonormal basis \( \{ \Psi_i \} \). Orthonormal implies \( \langle \Psi_i | \Psi_j \rangle = \delta_{ij} \). Completeness implies that any function \( \Psi \in \mathcal{H} \) can be expressed as a linear combination of these basis functions. (If not, we would add the part of \( \Psi \) that's left over as another basis function to obtain completeness!) This implies that \( \Psi \) can be fully characterized by its scalar products \( \langle \Psi | \Psi' \rangle \) with the basis functions \( \Psi_i \). These scalar products can be assembled into a vector \( \psi \) with coefficients \( \langle \Psi_i | \Psi \rangle \) and the same dimension as the Hilbert space. This allows one to reduce function spaces to ordinary vector spaces, as discussed below.

### 3.2 Dirac notation and operators

Functions such as \( \Psi \) can be expressed in terms of many different kinds of variables such as \( x \) or \( p \), depending on what is most convenient. Very often, one only wants to discuss the functions themselves, irrespective of which variables they are represented in terms of. It is then convenient to use the bracket notation of Dirac and replace

\[
\Psi \in \mathcal{H} \text{ by the ket } |\Psi\rangle \\
\int dx \Psi^*(x) \_ \_ \in \mathcal{H}^* \text{ by the bra } <\Psi|.
\]

\[ A-33 \]

Note that while \( |\Psi\rangle \) is indeed the function \( \Psi \) as an abstract member of the function space \( \mathcal{H} \), \( <\Psi| \) is not \( \Psi^*(x) \) or any function at all, but rather an overlap operation in the dual space \( \mathcal{H}^* \). Lack of realization of this fact leads to endless confusion and misuse of the Dirac notation.

We now have a compact representation for functions and their overlap integrals (or scalar products) with other functions. The final ingredient is the addition of linear operators, which can turn functions into new functions. Again, the best analogy is conventional vector space: \( v' = av \) where \( a \) is a scalar.
The new vector \( \mathbf{v}' \) points in the same direction as \( \mathbf{v} \), but has different normalization. Similarly \( \Psi' = \Psi \). To change the orientation of a vector, we need to apply a matrix \( M: \mathbf{v}' = M \mathbf{v} \). \( M \) may induce a translation, rotation, or shearing on the vector. Similarly we can have \( \Psi' = M \Psi \), where \( M \) is a linear operator (i.e. it acts on \( \Psi \) to the first power as shown). A wide class of linear operators is allowed, such as differentiation, multiplication by another function, etc. Consider for example the angular-variable function \( \Psi(\phi) \) acted on by operator \( \exp[\alpha d/d\phi] \):

\[
e^{\alpha d/d\phi} \Psi(\phi) = (1 + \alpha \frac{d}{d\phi} \Psi + \frac{1}{2} \alpha^2 \frac{d^2}{d\phi^2} \Psi + \cdots) = \Psi(\phi + \alpha)
\]

Clearly, \( \exp[\alpha d/d\phi] \) is the wavefunction equivalent of a rotation matrix for a vector, rotating the wavefunction by an angle \( \alpha \).

The analogy with linear algebra vector spaces is even more direct. Let us express wavefunctions \( \Psi \) and \( \Psi' \) in terms of their coefficients \( c_i \) and \( c_i' \) in a basis \( \{\Psi_i\} \) as vectors \( c \) and \( c' \). Then \( c_i = \int dx \Psi_i^* \Psi \) and \( c_i' = \int dx \Psi_i^* \Psi' \). We can compute "matrix elements" of any operator \( M \) using the dual basis: \( M_{ij} \equiv \langle \Psi_i | M | \Psi_j \rangle \). If \( \Psi' = M \Psi \), then

\[
\int dx \Psi'_j \Psi' = \int dx \Psi'_j M \Psi
\]

\[
\int dx \Psi'_i \sum_j c_j' \Psi_j = \int dx \Psi'_i M \sum_j c_j \Psi_j = \sum_i < \Psi'_i | M | \Psi_j > c_i \text{ or}
\]

\[
c_i' = \sum_j < \Psi'_i | M | \Psi_j > c_i
\]

The first step simply involves application of the "overlap operator"; in the second step, \( \Psi \) and \( \Psi' \) are expressed in terms of the basis; the third step switches to Dirac notation for the overlap integrals, AND assumes the basis functions are orthonormal. Steps 4 and 5 are the same as step 3, in pure algebra or pure Dirac notation.

Thus, acting with an operator on a wavefunction can be replaced by a linear algebra problem of matrix multiplication if we write the wavefunction as a vector of basis set coefficients, and the operator as a matrix of overlap integrals. Dirac notation is an intermediate form of notation which allows us to write equations in terms of the abstract wavefunction, its "dual" bra, and abstract operators without reference to a basis, while maintaining the simplicity of pure matrix notation.

In the above, the dual space overlap operation \( \int dx \Psi_i(x) \) is referred to as an operation, not an operator. This is because its application to a wavefunction yields a scalar (the overlap coefficient \( c_i \)) rather than another function, as an operator would. We can write an operator which is the direct analog of the dual space operators as follows:

\[
P_i = \Psi_i \int dx \Psi_i^* = |\Psi_i > < \Psi_i|
\]

When applied to a function \( \Psi \), \( P_i \) computes its overlap with \( \Psi_i \) and returns the function \( \Psi_1 \) times that overlap. It effectively projects the part of \( \Psi \) that looks like \( \Psi_i \) out and returns it. Hence \( P_i \) is referred to as a projection operator. Generally, operators can be constructed by combining functions \( \in \mathbb{H} \) and duals \( \in \mathbb{H}^* \) in "reverse" order from matrix elements.

### 3.3 Hermitian operators

Since observable quantities must be real (as far as we know) and are interpreted as expectation values in quantum mechanics, operators with real expectation values are of particular interest. For a matrix to have this property, it must be Hermitian, i.e. it must equal its complex conjugate-transpose: \( M = M^\dagger \). An analogous property must therefore hold for Hermitian operators. For any operator,

\[
\langle \Psi | F | \Phi \rangle = \langle \Psi | F' \Phi \rangle = \langle F \Phi | \Psi \rangle^* = \langle \Phi | F^\dagger | \Psi \rangle^*
\]
The fourth step can be taken as the definition of the hermitian conjugate $F^\dagger$ of $F$. In the special case of hermitian operators $\langle \Psi | F | \Phi \rangle = \langle \Phi | F^\dagger | \Psi \rangle$ must hold for any matrix elements in any basis representation, or the expectation values will not be real. This immediately implies that $F = F^\dagger$ must hold for a hermitian operator, i.e. in its matrix representation it must equal its own complex conjugate-transpose. This can easily be shown in matrix notation:

Let $\langle \Psi | F | \Phi \rangle = \sum_i c_i^* F_{ij} d_j$

or

$$\left(\begin{array}{c} f_{11} \\ f_{12} \\ \vdots \end{array}\right) \left(\begin{array}{c} d_1 \\ d_2 \\ \vdots \end{array}\right) = c_1^* f_{11} d_1 + c_1^* f_{12} d_2 + \ldots$$

We claim that the complex conjugate transpose matrix element of operator $f^{\dagger}$, namely

$$\left(\begin{array}{c} f_{11}^* \\ f_{12}^* \\ \vdots \end{array}\right) \left(\begin{array}{c} c_1 \\ c_2 \\ \vdots \end{array}\right) = d_1^* f_{11}^* c_1 + d_1^* f_{12}^* c_2 + \ldots$$

should be the same. As shown by the one example of $f_{12}$ by comparison of the two explicit equations, this is indeed the case, if we make use of the hermitian property of the operator's matrix representation.

### 3.4 Other useful matrix properties

- Inverse, transpose, and complex-conjugate transpose of products:

  These can be shown by direct evaluation:

  $$\left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array}\right] \left[\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array}\right]^* = \left[\begin{array}{cc} a_{11}^* b_{11} + a_{12}^* b_{21} + a_{21}^* b_{12} + a_{22}^* b_{22} + \ldots \\ a_{21}^* b_{11} + a_{22}^* b_{21} + a_{11}^* b_{12} + a_{12}^* b_{22} + \ldots \end{array}\right]$$

- Antiterminal property of commutators:

  Let $A$ and $B$ be hermitian; then

  $$[A, B]^\dagger = (AB)^\dagger - (BA)^\dagger = B^* A^\dagger - A^* B^\dagger = BA - AB = -[A, B]$$

  This implies that $[A, B]$ is anti-Hermitian: its Hermitian conjugate equals minus itself. $[A, B]$ can therefore be written as $iC$, where $C$ is a Hermitian operator.

- Diagonalization of Hermitian matrices by unitary matrices:
Let $M^{-1}HM = \Lambda$. $H$ is hermitian $\Rightarrow$

$$(M^{-1}HM)^+ = \Lambda^+ \Rightarrow M^+H(M^{-1})^+ = \Lambda \text{ or } M^{-1} = M^+$$

$M^*M = I$

The last property defines $M$ to be a unitary matrix.

### 3.5 Behavior of the determinant of matrices

i) the determinant of $A$ is the sum over all possible products of matrix elements including only one from each row and column, with a sign alternation for odd and even permutations:

$$
\begin{vmatrix}
  a_{11} & a_{12} & a_{13} & L \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33} \\
  M & O
\end{vmatrix}
= a_{11} a_{22} a_{33} - a_{12} a_{23} a_{31} + a_{31} - L
$$

ex:

Note: $\det(A)$ unchanged if we add rows to one another or exchange them.

ii) $\det(AB) = \det(A)\det(B)$

iii) $\det(A^{-1}) = \frac{1}{\det(A)}$ since $\det(AA^{-1}) = \det(I) = 1 = \det(A)\det(A^{-1})$

iv) $\Rightarrow$ a similarity transformation leaves the determinant unchanged:

v) $\Rightarrow$ the determinant of a matrix is just the product of its eigenvalues, since diagonalization is just a similarity transformation:

vi) For any matrix, $\det(\tilde{M}) = \det(M)$; $\det(M^*) = \det(M)^*$

vii) For unitary matrices, $\det(M^{-1}M) = \det(M^{-1})\det(M) = 1 \Rightarrow \det(M)\det(M^*) = 1$
or $\det(M) = e^{i\theta}$
all eigenvalues of a unitary matrix have unit modulus.

3.6 Matrix diagonalization

Many problems in quantum mechanics can be posed as eigenvalue problems; the most famous is the solution of the time-independent Schrödinger equation for energy eigenvalues. The eigenvalue problem is the following: which functions in \( \mathcal{H} \) are unchanged by application of a certain operator \( G \), other than a multiplicative factor \( \lambda \), referred to as the eigenvalue? If we convert wavefunctions to vectors \( \textbf{m} \), the question becomes: which vectors are unchanged by application of a matrix \( G \), except for a multiplicative factor \( \lambda \)?

For hermitian and other well-conditioned matrices, the number of such vectors generally equals the dimension of the vector space. (This is of course infinite for a Hilbert space with infinitely many basis functions.) We can group all these vectors \( \textbf{m} \) into a matrix \( M \), and all the eigenvalues into a diagonal matrix \( \Lambda \), in which case the problem is posed as

\[
GM = M\Lambda \quad \text{or} \quad M^{-1}GM = \Lambda
\]

where the columns of \( M \) are the vectors \( \textbf{m} \), and the diagonal elements of \( \Lambda \) are the eigenvalues. Note that the matrix \( M \) appears on the LEFT-hand side of \( \Lambda \). For a single vector \( \textbf{m} \), we can write this explicitly in terms of a set of linear equations to be solved:

\[
G\textbf{m} = \lambda \textbf{m}, \quad \text{or} \\
(G-\lambda I)\textbf{m} = 0; \quad \text{explicitly,} \\
i = 1\cdots n: \quad \sum_{j=1}^{n} (G_{ij} - \delta_{ij}\lambda)x_j = 0
\]

It is not generally possible for \( n \) sums over \( n \) \( x_i \) variables to be simultaneously zero, unless all \( x_i \) are identically zero:

\[
A-52
\]

However, if one of the equations is redundant (can be expressed in terms of the others), then a solution is possible, where all \( x_i \) can be determined as multiples of one of them:

\[
A-53
\]

A condition for such redundancy is that the determinant of the coefficient matrix be zero, or \( \det(G-I\lambda) = 0 \) to make the set of equations singular. The \( \det(G-I\lambda) \) is a polynomial in \( \lambda \) of order \( n \), where \( n \) is the size of the matrix. These eigenvalues \( \lambda \) can be found by solving the polynomial, and can then be inserted into the linear equations above to yield the components of the eigenvectors.

As an example of diagonalization, consider the following simple and very common case:
This matrix corresponds to two “degenerate” energy levels, coupled by a coupling term $V$. This could arise for two bending states of CO$_2$ (its two bending modes have the same frequency), or two electronic states coupled by “non-adiabatic” terms. We subtract $I\lambda$ and compute the determinant.

Now we insert the eigenvalues back and solve the linear equation A51-b. For $E+V$ for example,

and similarly the other eigenvector equals

Thus the matrix $M$ equals

and indeed, one can easily check that $M^\dagger HM$ is diagonal with elements $E\pm V$. Note that the eigenvectors are not uniquely defined. For example,

Are just as good a pair as the ones above. However, the eigenvectors do need to be normalized so that the matrix $M$ is unitary (but not for them to be eigenvectors satisfying eq. A-51b; hence the factors of $1/\sqrt{2}$. 