Appendix A Review of Essential Math

Here we review the math that is used in the text.

A.1 Differential Equations

Differential equations for physical systems establish a relationship between differentials. A differential such as $\frac{df(t)}{dt}$ or $\frac{df(x)}{dx}$ represents the change in f for a small change in the independent variable for space (x) or time (t). A second order differential, like say $\frac{d^2f(t)}{dt^2}$, is then the change of $\frac{df(t)}{dt}$ for a small change in t. From simple mechanics, Newton's law has the form $\frac{d^2x(t)}{dt^2} = \frac{F(x)}{m}$ where $F(x)$ is the force, m is the mass, and now $x(t)$ (the position) is the dependent variable and t for time is the independent variable. So $\frac{dx(t)}{dt} \equiv \dot{x}$ is then the change in x for a small change in t and is the velocity. And $\frac{d^2x(t)}{dt^2} \equiv \ddot{x}$ is the change in $\frac{dx(t)}{dt}$ for a small change in t and is the acceleration. Newton's law then says that the acceleration, $\frac{d^2x(t)}{dt^2}$, is $\frac{F(x)}{m}$ where F is the force and m is the mass.

We are interested in *linear* **differential equations**, meaning that terms such as $\frac{df(x)}{dx}, \frac{d^2f(x)}{dx^2}, \frac{df(t)}{dt}, \frac{d^2f(t)}{dt^2}$ may appear, but not terms that involve higher powers of these terms such as $\left(\frac{df}{dx}\right)^2$. The order of the differential equation is given by the number of times a given operator such as $\frac{d}{dx}$ or $\frac{d}{dt}$ operates on the function $f(x)$ or $f(t)$, respectively. So[1]

$$\frac{df(x)}{dx} + f(x) = 0 \tag{A.1}$$

is **first order**, while

$$\frac{d^2f(x)}{dx^2} + a\frac{df(x)}{dx} + cf(x) = 0 \tag{A.2}$$

is **second order**.

When all the operators and the unknown function are on the left-hand side and the right-hand side is zero such as

$$\frac{df(x)}{dx} + f(x) = 0 \tag{A.3}$$

---

[1] Note that in this discussion, there is just one independent variable for the sake of simplicity. That variable is usually either space or time (x or t, respectively). Rather than give each discussion in what follows in terms of say first time and then space, we will just arbitrarily pick x or t, assuming then that the reader can easily substitute the variable interest for their own problem. Language such as the word "gradient" for $\frac{df(x)}{dx}$ or "rate" for $\frac{df(t)}{dt}$ should be changed, as required.

or

$$\frac{d^2 f(x)}{dx^2} + a\frac{df(x)}{dx} + cf(x) = 0 \tag{A.4}$$

then the equations are **homogeneous**. When there is a function on the right such as

$$\frac{df(x)}{dx} + f(x) = g(x) \tag{A.5}$$

or

$$\frac{d^2 f(t)}{dt^2} + a\frac{df(t)}{dt} + cf(t) = \sin \omega t \tag{A.6}$$

the equations are **inhomogeneous**.

First order linear homogeneous differential equations have one solution, second order linear differential equations have two solutions, etc. An inhomogeneous differential equation has the **particular solution**, $f_p(t)$ or $f_p(x)$, associated with the function on the right-hand side, and also with the solution or solutions to the homogeneous equation. So for Eq. A.6, the total solution is $f(t) = A_1 f_1(t) + A_2 f_2(t) + f_p(t)$. In the case of the a time dependent problem, A_1 and A_2 are determined by the initial conditions, two in the case of a second order differential equation.

Examples of Solutions to Homogeneous Differential Equations

1. The solution to a first order homogeneous equation of the form

$$\frac{df(x)}{dx} + bf(x) = 0 \tag{A.7}$$

can be found by rearranging and integrating:

$$\frac{df(x)}{f(x)} = -bdx \tag{A.8}$$

$$\int \frac{df(x)}{f(x)} = -b\int dx \tag{A.9}$$

$$\ln f(x) = -bx + a \tag{A.10}$$

$$f(x) = Ae^{-bx} \tag{A.11}$$

where $A = e^a$ is a constant. to be determined by the boundary conditions.

Note on boundary and initial conditions: In solving a differential equation, since the equation represents a relationship between different derivatives of a function and the function, the solution for the function is going to depend on the starting value of that function when the independent variable (say x or t) is zero. So, in the above example, A is going to be determine by the value of $f(x)$ when $x = 0$. That is called a **boundary condition**. If the problem involved time instead of space, like $f(t)$, you would need to know $f(t = 0)$ which is then an **initial condition**. If a differential equation is nth order, meaning the highest order derivative is $\frac{d^n f(t)}{dt^n}$. or $\frac{d^n f(x)}{dx^n}$, then there must be n initial or boundary conditions. Think of Newton's second law, which is second order in time, then the position $x(t)$ and velocity $v(t) = \frac{dx(t)}{dt}$ must be specified at $t = 0$.

2. The solution to a second order homogeneous equation of the form

$$\frac{d^2 f(t)}{dt^2} + a\frac{df(t)}{dt} + bf(t) = 0 \tag{A.12}$$

can be found by assuming a solution of the form

$$f(t) = e^{\gamma t} \tag{A.13}$$

The same approach used in case 1 above would work, but the above approach is now a little simpler. Substituting this form, taking the derivatives and dividing through by $e^{\gamma x}$ results in a quadratic equation given by

$$\gamma^2 + a\gamma + b = 0 \tag{A.14}$$

With the solution given by the quadratic formula

$$\gamma_\pm = \frac{1}{2}\left(-a \pm \sqrt{a^2 - 4b}\right) \tag{A.15}$$

Note that when $a = 0$ and b is real and > 0, then

$$\gamma_\pm = \pm i\frac{1}{2}\sqrt{b} \tag{A.16}$$

where

$$i = \sqrt{-1} \tag{A.17}$$

The two solutions to the homogeneous equation are

$$f_{h1} = A_+ e^{\gamma_+ t} \tag{A.18}$$

and

$$f_{h2} = A_- e^{\gamma_- t} \tag{A.19}$$

The complete solution is

$$f_h(t) = A_+ e^{\gamma_+ t} + A_- e^{\gamma_- t} \tag{A.20}$$

where the constants are determined by the initial conditions (boundary conditions if the independent variable is space).

In case

$$a^2 - 4b = 0 \tag{A.21}$$

then

$$\gamma_+ = \gamma_- = \frac{\gamma - a}{2} \tag{A.22}$$

and the general solution is then

$$f_h(t) = A_1 e^{\gamma t} + A_2 t e^{\gamma t} \tag{A.23}$$

Solutions to most common and physically relevant differential equations are tabulated, meaning they can be found in published tables.[2]

Examples of Solutions to Inhomogeneous Differential Equations

The solution to an inhomogeneous differential equation is called the *particular solution* as indicated earlier.

1. To solve an inhomogeneous first order differential equation of the form

$$\frac{df(x)}{dx} + bf(x) = g(x) \tag{A.24}$$

we introduce an **integrating factor:**

$$e^{bx} \tag{A.25}$$

Then we can recover the above inhomogeneous equation by noting that

$$\frac{d}{dx}\left(e^{bx}f(x)\right) = e^{bx}\left(\frac{df(x)}{dx} + bf(x)\right) = e^{bx}g(x) \tag{A.26}$$

We can integrate this equation:

$$\int dx \frac{d}{dx}\left(e^{bx}f(x)\right) = \int dx\, e^{bx}g(x) \tag{A.27}$$

$$e^{bx}f(x) = \int dx' e^{bx'} g(x') \tag{A.28}$$

where x' has been substituted on the right to distinguish between the x that is in the integral and the x on the left that is not in the integral, so that

$$f_p(x) = \int dx' e^{-b(x-x')} g(x') \tag{A.29}$$

The subscript p denotes the particular solution to the inhomogeneous equation.

As indicated earlier, the complete solution to an inhomogeneous first order differential equation is then the sum of the homogeneous solution and the inhomogeneous solution:

$$f(x) = f_h(x) + f_p(x) = Ae^{-bx} + e^{-bx}\int dx' e^{bx'} g(x') \tag{A.30}$$

and A is determined for the boundary condition for $f(x)$ (not just for the homogeneous part).

---

[2] For this and many more helpful mathematical relationships, see for example Murray R. Spiegel, Seymour Lipschutz and John Liu, *Schaum's Outline of Mathematical Handbook of Formulas and Tables*, 4th Edition, McGraw Hill (2013). Also, Milton Abramowitz and Irene A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards Applied Mathematics Series 55, US Government Printing Office (1964).

2. To solve a second order inhomogeneous differential equation, the complete solution is of the form

$$f(x) = f_h(x) + f_p(x) \tag{A.31}$$

where the two constants in the solution $f_h(x)$ are again determined for the boundary conditions for $f(x)$. The steps for generally finding the **particular solution**, $f_p(x)$ are beyond the current discussion, but this is often done with a Green's function approach.[3]

A.2 Partial Differential Equations (PDE)—Method of Separation of Variables

Much of the analysis and models in this text are limited to one dimension in order to keep the math simple and because in today's technology, a one-dimensional system has technological advantages.

A one-dimensional differential equation of the form:

$$\frac{d^2 f(x)}{dx^2} + bf(x) = 0 \tag{A.32}$$

may be generalized to two or three dimensions when appropriate. For example, the equation for electromagnetic waves including light and radio is often a three-dimensional problem. Like mechanical vibrations of a three-dimensional object, quantum systems like atoms, and thermal transport for heat management, many devices require a solution to the equivalent three-dimensional equations. We can deal with this by replacing the differential operator with the appropriate ∇ operator where, in three dimensions in Cartesian coordinates, for example:

$$\nabla^2 f(x, y, z) + bf(x, y, z) = 0 \tag{A.33}$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{A.34}$$

This equation is solved by the **method of separation of variables** and, since the three coordinates are independent of each other, it must be that this equation holds only when

$$f(x, y, z) = f_x(x)f_y(y)f_z(z) \tag{A.35}$$

If we substitute this into the partial differential equation and then divide by $f(x, y, z)$ we get

$$\frac{1}{f_x(x)}\frac{\partial^2 f_x(x)}{\partial x^2} + \frac{1}{f_y(y)}\frac{\partial^2 f_y(y)}{\partial y^2} + \frac{1}{f_z(z)}\frac{\partial^2 f_z(z)}{\partial z^2} + b = 0 \tag{A.36}$$

Since this has to hold for all values of $x, y,$ and z, it means that

$$\frac{1}{f_x(x)}\frac{\partial^2 f_x(x)}{\partial x^2} = a_x \tag{A.37}$$

[3] George B Arfken, Hans J. Weber and Frank E. Harris Mathematical Methods for Physicists, 7th ed., Elsevier, Amsterdam (2013).

$$\frac{1}{f_y(y)} \frac{\partial^2 f_y(y)}{\partial y^2} = a_y \tag{A.38}$$

$$\frac{1}{f_z(z)} \frac{\partial^2 f_z(z)}{\partial z^2} = a_z \tag{A.39}$$

where

$$a_x + a_y + a_z = -b \tag{A.40}$$

A.3 **Eigenvalue Problems**

In general, if $\hat{A}$ is an operator like $\frac{d^2}{dx^2}$ or ∇^2, then the equation

$$\hat{A}u_a(x) = au_a(x) \tag{A.41}$$

where a is a constant, is called an **eigenvalue equation**. Then $u_a(x)$ is the eigenfunction and a is the **eigenvalue**. So, if $\hat{A} = \frac{d^2}{dx^2}$, then

$$\frac{d^2}{dx^2}u_a(x) = -au_a(x) \tag{A.42}$$

The solution is

$$u_a(x) = e^{\pm i\sqrt{a}x} \tag{A.43}$$

If there are no constraints like boundary conditions, then a can take on any positive value., and $\frac{d^2}{dx^2}$ is said to have a *continuous* **spectrum of eigenvalues.** Sometimes, in this case, the eigenfunction is written as

$$u_a(x) \equiv u(a, x) \tag{A.44}$$

If there are constraints, such as that the function must satisfy **periodic boundary conditions**, e.g.,

$$e^{\pm i\sqrt{a}x} = e^{\pm i\sqrt{a}(x+L)} \tag{A.45}$$

then this is only satisfied for specific values of a. Namely,

$$e^{\pm i\sqrt{a}L} = 1$$

Limiting the values of a to

$$\sqrt{a}L = 2m\pi \quad \text{where } m \text{ is a positive integer} \tag{A.46}$$

or

$$a = \left(\frac{2m\pi}{L}\right)^2 \tag{A.47}$$

In this case, the spectrum of eigenvalues is discrete and we write $u_a(x)$ as

$$u_a(x) \equiv u_m(x) \tag{A.48}$$

For either a discrete or continuous spectrum of eigenvalues and assuming that the operators are **Hermitian**, meaning an operator $\hat{P}$ is Hermitian if $\int_0^\infty dx\, g^*(x)\hat{P}f(x) = \int_0^\infty dx\big(\hat{P}g(x)\big)^* f(x)$ for an arbitrary and well behaved $g(x)$ and $f(x)$, there are an infinite number of eigenfunctions and *the set of all eigenfunctions for each case forms a complete set*. If the boundary conditions allow $-\infty < x < \infty$, then any $f(x)$ can be expanded in terms of the eigenfunctions. In the case that the spectrum of eigenvalues is discrete, it means that $f(x)$ can be written as:

$$f(x) = \sum_{m=0}^{\infty} c_m u_m(x) \tag{A.49}$$

where c_m is called an expansion coefficient. This is a countably infinite series. This is very similar to the ideas learned in the study of the Fourier series. If the spectrum of eigenvalues is continuous, we could then expand $f(x)$ in terms of an uncountably infinite series,

$$f(x) = \int_0^\infty da\, c(a)u(a,x) \tag{A.50}$$

where again $c(a)$ is an expansion coefficient.

It can be shown that, for the operators of interest in these notes and that are associated with an observable (i.e., the operators are Hermitian), the eigenfunctions are orthogonal and can be normalized (i.e., they are orthonormal), mathematically meaning that, for eigenfunctions with discrete eigenvalues,

$$\int_{-\infty}^{\infty} dx\, u_n^*(x)u_m(x) \equiv (u_n(x)|u_m(x)) = \delta_{nm} \equiv \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases} \tag{A.51}$$

where δ_{nm} is a Kronecker delta. A shorthand notation has been introduced to simplify the writing and calculations and looks similar to Dirac notation though it is definitely not Dirac notation, which describes eigenvectors in a Hilbert space. For eigenfunctions with a continuous spectrum of eigenvalues, the orthonormality is given by

$$\int_{-\infty}^{\infty} dx\, u^*(a',x)\,u(a,x) \equiv (u(a',x)\,|u(a,x)) = \delta(a'-a) \tag{A.52}$$

where $\delta(a'-a)$ is the **Dirac delta-function** with the property that $\int_{-\infty}^{\infty} dx f(x)\delta(x-x_0) = f(x_0)$, discussed below.

We can use these results to find the expansion coefficient above. For the case of a discrete spectrum of eigenvalues,

$$f(x) = \sum_{m=0}^{\infty} c_m u_m(x) \tag{A.53}$$

We then multiply both sides by $u_n^*(x)$ and integrate over x:

$$\int_{-\infty}^{\infty} dx\, u_n^*(x) f(x) = \int_{-\infty}^{\infty} dx \sum_{m=0}^{\infty} c_m u_n^*(x) u_m(x) = \sum_{m=0}^{\infty} c_m \int_{-\infty}^{\infty} dx\, u_n^*(x) u_m(x)$$

$$= \sum_{m=0}^{\infty} c_m \delta_{nm} = c_n \tag{A.54}$$

Therefore, with n reverting now to the symbol m:

$$c_m = \int_{-\infty}^{\infty} dx\, u_m^*(x) f(x) \tag{A.55}$$

For the case of a continuous spectrum of eigenvalues,

$$f(x) = \int_0^{\infty} da\, c(a) u(a, x) \tag{A.56}$$

We then multiply both sides by $u^*(a', x)$ and again integrate over x:

$$\int_{-\infty}^{\infty} dx\, u^*(a', x) f(x) = \int_{-\infty}^{\infty} dx \int_0^{\infty} da\, c(a) u^*(a', x)\, u(a, x)$$

$$= \int_0^{\infty} da\, c(a) \int_{-\infty}^{\infty} dx\, u^*(a', x)\, u(a, x) \int_0^{\infty} da\, c(a)\, \delta(a' - a) = c(a') \tag{A.57}$$

Therefore, after a' reverting back to the symbol a,

$$c(a) = \int_{-\infty}^{\infty} dx\, u^*(a, x) f(x) \tag{A.58}$$

A.4 Complex Numbers and Euler's Theorem

For real numbers x, and y, a complex number z is written as

$$z = x + iy \tag{A.59}$$

where

$$i = \sqrt{-1} \tag{A.60}$$

Euler's theorem says that

$$z = x + iy = R \exp(i\theta) = R \cos\theta + iR \sin\theta \tag{A.61}$$

where

$$R = \sqrt{x^2 + y^2} \text{ and } \tan\theta = \frac{y}{x} \tag{A.62}$$

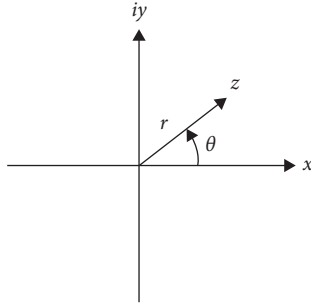Complex numbers are represented in the complex plane as shown in Fig. A.1

Fig. A.1 Complex numbers are often represented in the complex plane.

A.5 **Dirac Delta-Function**

The Dirac delta-function has physical meaning only under an integral and defined by the property that

$$\int_{-\infty}^{\infty} dx \, f(x)\delta(x) = f(0) \tag{A.63}$$

Or, by changing variables, it is easy to show that

$$\int_{-\infty}^{\infty} dx \, f(x)\delta(x - x_0) = f(x_0) \tag{A.64}$$

Likewise,

$$\int_{-\infty}^{\infty} dx \, f(x)\delta(ax) = \frac{1}{|a|}f(0); |a| > 0 \tag{A.65}$$

From the Fourier theorem it can be shown that

$$\int_{-\infty}^{\infty} dk \, e^{-ik(x-x')} = 2\pi\delta(x - x') \tag{A.66}$$

Proof: It follows from Fourier theory that if

$$\mathcal{F}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, e^{ikx} f(x) \tag{A.67}$$

and

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \, e^{-ikx} \, \mathcal{F}(k) \tag{A.68}$$

Then the Fourier integral theorem says that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \, e^{-ikx} \int_{-\infty}^{\infty} dx' \, e^{ikx'} f(x') \tag{A.69}$$

Reorganizing, we get

$$f(x) = \int_{-\infty}^{\infty} dx' \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-ik(x-x')} \right) f(x') \tag{A.70}$$

This is only true if

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-ik(x-x')} = \delta(x - x') \tag{A.71}$$

Another form for the Dirac delta-function that is convenient to use is

$$\delta(k - k_0) = \lim_{a \to \infty} \frac{1}{a\sqrt{\pi}} e^{-\left(\frac{k-k_0}{a}\right)^2} \tag{A.72}$$

A.6 Gaussian Integral

The functional form $e^{-\alpha x^2}$ or $e^{-\alpha x^2 + bx}$ is called a Gaussian. The second form is equivalent to displaced Gaussian since the exponent can be rewritten as $\alpha x^2 - bx = a\left(x^2 - \frac{b}{a} + \frac{b^2}{4a^2}\right) - \frac{b^2}{4a} = a\left(x - \frac{b}{2a}\right)^2 - \frac{b^2}{4a}$

$$\int_{-\infty}^{\infty} dx\, e^{-\alpha x^2} = \sqrt{\frac{\pi}{a}}; \quad \mathrm{Re}\ a > 0 \tag{A.73}$$

$$\int_{-\infty}^{\infty} dx\, e^{-\alpha x^2 + bx} = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}}; \quad \mathrm{Re}\ a > 0 \tag{A.74}$$

Note that more complex integrals involve a Gaussian form, such as $\int_{-\infty}^{\infty} dx\, xe^{-\alpha x^2 + bx}$ or $\int_{-\infty}^{\infty} dx\, x^2 e^{-\alpha x^2 + bx}$, which can be evaluated by **differentiating under the integral**. For example: $\int_{-\infty}^{\infty} dx\, xe^{-\alpha x^2 + bx} = \frac{d}{db} \int_{-\infty}^{\infty} dx\, e^{-\alpha x^2 + bx} = \frac{b}{2a} \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}}$ and $\int_{-\infty}^{\infty} dx\, x^2 e^{-\alpha x^2} = -\frac{d}{da} \int_{-\infty}^{\infty} dx\, e^{-\alpha x^2 + bx} = \sqrt{\frac{\pi}{a^3}}$.

A.7 Linear Algebra: Matrices, Determinants, Permanents, and the Eigenvector

Multiplication

Matrices are a two-dimensional array of numbers: $n - rows \times m - column$. A square matrix, such as frequently encountered in quantum problems, is one where the number of columns and rows are the same. An $n \times n$ matrix $\hat{A}$ is written as:

$$\hat{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \tag{A.75}$$

The subscripts refer to the rows and columns, respectively. A square $n \times n$ matrix has **order** $n \times n$ which is sometimes just n. Two matrices, $\hat{A}$ and $\hat{B}$ of order $j \times k$ and $m \times n$ can be multiplied together if $k = m$ to give a resulting matrix that is order $j \times n$. **Matrix multiplication** is given by

$$\hat{A} * \hat{B} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & a_{jm} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} a_{1i}b_{i1} & \cdots & \sum_{i=1}^{m} a_{1i}b_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} a_{ji}b_{i1} & \cdots & \sum_{i=1}^{m} a_{ji}b_{in} \end{bmatrix} \tag{A.76}$$

Three examples:

$$\begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1*1+4*2 & 1*6+4*3 \\ 3*1+5*2 & 3*6+5*3 \end{bmatrix} = \begin{bmatrix} 9 & 18 \\ 13 & 33 \end{bmatrix} \tag{A.77a}$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1*1+2*2 & 1*6+2*3 \end{bmatrix} = \begin{bmatrix} 5 & 12 \end{bmatrix} \tag{A.77b}$$

$$\begin{bmatrix} 1 & 6 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 1*4+6*5 \\ 2*4+3*5 \end{bmatrix} = \begin{bmatrix} 34 \\ 23 \end{bmatrix} \tag{A.77c}$$

Note that when a scalar multiplies a matrix, it multiples every term in that matrix:

$$3 * \begin{bmatrix} 1 & 6 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 18 \\ 6 & 9 \end{bmatrix} \tag{A.77d}$$

Determinants

The **determinant** of a square matrix is a scalar and is given by selecting any row or column in the matrix, and for each element in that row or column, a_{ij}, multiply it by $(-1)^{i+j}$ and the determinant of the matrix formed by removing the row i and column j. As an example, we evaluate the determinant of the square matrix $\hat{A}$ by removing the top row:

$$\det \hat{A} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} - a_{21} \begin{vmatrix} a_{11} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots \tag{A.78}$$

where the determinant associated with the multiplying factor out front, a_{j1}, is missing the first column and j^{th} row of the original determinant. The process is continued with each resulting matrix until the final result is a scalar. This can be written more succinctly as:

$$\det \hat{A} = \sum_{i \text{ or } j}^{n} (-1)^{i+j} a_{ij} M_{ij} \tag{A.79}$$

M_{ij} is the **first minor** of the a_{ij} element and is computed by forming the sub-matrix of $\hat{A}$ by removing row i and column j and calculating the determinant. Then $(-1)^{i+j} M_{ij}$ is called the **cofactor** of the a_{ij} element.

For the matrix:

$$\hat{A} = \begin{bmatrix} 3 & 1 & 2 \\ 4 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix} \tag{A.80}$$

the minor of $\hat{A}_{12}$ $\left(\text{note that } \hat{A}_{12} = a_{12} = 1, \text{ the second entry from the upper left}\right)$ is

$$\hat{M}_{12} = \det \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \tag{A.81}$$

So for the determinant of $\hat{A}$:

$$\det \begin{vmatrix} 3 & 1 & 2 \\ 4 & 2 & 3 \\ 2 & 5 & 1 \end{vmatrix} = 3 \begin{vmatrix} 2 & 3 \\ 5 & 1 \end{vmatrix} - 4 \begin{vmatrix} 1 & 2 \\ 5 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}$$

$$= 3(2 - 15) - 4(1 - 10) + 2(3 - 4) = -5. \tag{A.82}$$

The **rank** of a square matrix is the order of the largest sub-matrix with linearly independent rows (columns). This corresponds to the largest **sub-matrix** with a non-zero determinant. In the case above with the matrix of order three, corresponding to the determinant with value -45, since the determinant is non-zero, the rank and order are both three.

The determinant is used to preserve exchange symmetry for fermions.

Permanents

The **permanent** of a square matrix is identical to the determinant except that there is no $(-1)^{i+j}$ factor. So in the case of Eq. A.82, switch the minus sign in front of the 4 in the first equality to a plus sign: i.e., $-4 \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} \rightarrow +4 \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix}$.

The permanent is used to preserve the exchange symmetry for bosons.

Adjoint, Hermiticity, and Unitarity

The **adjoint** of a matrix $\hat{A}$ is the **complex-transpose** of the original matrix. With the matrix $\hat{A}$ above then the corresponding adjoint, $\hat{A}^\dagger$, is given by

$$\hat{A}^\dagger = \begin{bmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{bmatrix} \tag{A.83}$$

If $\hat{A} = \hat{A}^\dagger$, then the matrix is **Hermitian**. If $\hat{A}^\dagger = \hat{A}^{-1}$, then the matrix is **unitary**.

Vectors: Inner and Outer Products and Dirac Notation

A matrix $\hat{V}$ with a single column is an $n \times 1$ **column vector**:

$$\hat{V} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \tag{A.84}$$

and the complex transpose is a $1 \times n$ **row vector**:

$$\hat{V}^\dagger = \begin{bmatrix} v_1^* & \cdots & v_n^* \end{bmatrix} \tag{A.85}$$

In **Dirac notation**, the **ket** can be represented as a column vector. So for a three-dimensional **Hilbert space**,

$$|V\rangle = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{A.86}$$

The bra is the corresponding complex transpose (row vector) of the **ket**. The **inner product** is similar to a dot product and produces a **scalar** (a matrix of order unity) that can be complex.

$$\langle W|V \rangle = \begin{bmatrix} w_1^* & w_2^* & w_3^* \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = w_1^* v_1 + w_2^* v_2 + w_3^* v_3 \tag{A.87}$$

The **outer product** of two vectors of the same size is a square matrix:

$$|V\rangle\langle W| = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1^* & w_2^* & w_3^* \end{bmatrix} = \begin{bmatrix} v_1 w_1^* & v_1 w_2^* & v_1 w_3^* \\ v_2 w_1^* & v_2 w_2^* & v_2 w_3^* \\ v_3 w_1^* & v_3 w_2^* & v_3 w_3^* \end{bmatrix} \tag{A.88}$$

An arbitrary matrix can then be written out as

$$\hat{A} = \sum_{ij} \langle i|\hat{A}|j \rangle \, |i\rangle \langle j| \tag{A.89}$$

where $|i\rangle$ ($\langle j|$) are unit vectors with 0's in all the positions except position i (j), where there is a 1.

$$\langle i|\hat{A}|j \rangle \equiv a_{ij} = \begin{bmatrix} 0_1 & \cdots & 1_i & \cdots & 0_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 0_1 \\ \vdots \\ 1_j \\ \vdots \\ 0_n \end{bmatrix} \tag{A.90}$$

Example for an operator of order two:

$$\langle 1|\hat{A}|2 \rangle = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = a_{12} \tag{A.91}$$

For consistency in notation, we note that

$$\hat{A}_{ij} = \langle i|\hat{A}|j \rangle = a_{ij} \tag{A.92}$$

The Identity Matrix and the Inverse Matrix

The **identity matrix** has zero's in all the positions except along the diagonal where it has ones. For example for a matrix of arbitrary order,

$$\hat{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{bmatrix} \tag{A.93}$$

Therefore, for any matrix $\hat{A}$ of order n,

$$\hat{A} * \hat{I} = \hat{I} * \hat{A} = \hat{A} \tag{A.94}$$

where $\hat{I}$ is the identity matrix of order n.

If $\hat{A}$ is a matrix with a non-singular determinant, then there exists an **inverse matrix** $\hat{A}^{-1}$ such that

$$\hat{A} * \hat{A}^{-1} = \hat{A}^{-1} * \hat{A} = \hat{I} \tag{A.95}$$

To find the inverse of a matrix, we use **Cramer's rule**:

$$\hat{A}^{-1} = \frac{1}{|\hat{A}|}\hat{C}^T \tag{A.96}$$

where

$$|\hat{A}| \equiv \det \hat{A} \tag{A.97}$$

and $\hat{C}^T$ is the *transpose* of the matrix of **cofactors** (see A.79 and discussion) where the cofactor of the a_{ij} element is again $(-1)^{i+j}M_{ij}$, as discussed above. So for say a 4th order matrix:

$$\hat{C}_{24} \equiv (-1)^{2+4}\hat{M}_{24} = (-1)^{2+4}\det\begin{bmatrix} 1 & 4 & 7 & 5 \\ 3 & 0 & 5 & 3 \\ -1 & 9 & 2 & 7 \\ 2 & -4 & 8 & 1 \end{bmatrix} = (-1)^{2+4}\begin{vmatrix} 1 & 4 & 7 \\ -1 & 9 & 2 \\ 2 & -4 & 8 \end{vmatrix}$$

$$= 1\begin{bmatrix} 9 & 2 \\ -4 & 8 \end{bmatrix} - 4\begin{bmatrix} -1 & 2 \\ 2 & 8 \end{bmatrix} + 7\begin{bmatrix} -1 & 9 \\ 2 & -4 \end{bmatrix} = 80 + 48 - 98 = 30 \tag{A.98}$$

Finally, with $\hat{C}$ being given by

$$\hat{C} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix} \tag{A.99}$$

$$\hat{A}^{-1} = \frac{1}{|\hat{A}|}\hat{C}^T = \begin{bmatrix} \frac{c_{11}}{|\hat{A}|} & \cdots & \frac{c_{n1}}{|\hat{A}|} \\ \vdots & \ddots & \vdots \\ \frac{c_{1n}}{|\hat{A}|} & \cdots & \frac{c_{nn}}{|\hat{A}|} \end{bmatrix} \tag{A.100}$$

The factor $\frac{1}{|\hat{A}|}$ was taken inside to emphasize that when a number multiplies a matrix, it multiplies every element of that matrix (A.77d).

Cramer's Rule for Solving n Linear Inhomogeneous Algebraic Equations in n Unknowns

Suppose we have n linear equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{A.101}$$

These can be written in matrix form by putting the coefficients in a matrix and converting the unknowns, x_i and the right-hand side into vectors. Specifically,

$$
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \vdots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ \vdots \\ x_n
\end{bmatrix}
=
\begin{bmatrix}
b_1 \\ b_2 \\ \vdots \\ b_n
\end{bmatrix}
\tag{A.102}
$$

Or written more compactly:

$$
\hat{A}\hat{X} = \hat{B} \tag{A.103}
$$

If one or more of the $b_i's \neq 0$, the set of algebraic equations is **inhomogeneous**. In this case, one can derive **Cramer's rule**, which says that the value of ith entry of $\hat{X}$ (a column vector) is:

$$
x_i = \frac{\det \hat{A}_i}{\det \hat{A}}; i = 1, \ldots, n \tag{A.104}
$$

where the matrix $\hat{A}_i$ is formed by taking the matrix $\hat{A}$ above and replacing the ith column with the vector $\hat{B}$.

For example

$$
\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
=
\begin{bmatrix} -1 \\ 3 \end{bmatrix}
\tag{A.105}
$$

$$
\det A = \det \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} = 5 \tag{A.106}
$$

$$
x_1 = \frac{\det \hat{A}_1}{\det \hat{A}} = \frac{\det \begin{bmatrix} -1 & 3 \\ 3 & 4 \end{bmatrix}}{\det \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}} = \frac{-13}{5} \tag{A.107}
$$

$$
x_2 = \frac{\det \hat{A}_2}{\det \hat{A}} = \frac{\det \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}}{\det \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}} = \frac{7}{5} \tag{A.108}
$$

The Eigenvector Problem

In the above case, a special case is had if

$$
\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
= \lambda
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\tag{A.109}
$$

In this case,

$$
\hat{A}\hat{X} = \lambda \hat{X} = \lambda \hat{I}\hat{X} \tag{A.110}
$$

This is called an **eigenvector** or **eigenvalue** equation. Then $\hat{X}$ is the eigenvector and λ is the eigenvalue.

Then

$$(\hat{A} - \lambda\hat{I})\,\hat{X} = \begin{bmatrix} 0_1 \\ 0_2 \\ \vdots \\ 0_n \end{bmatrix} = 0 \tag{A.111}$$

We rewrite this as

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0 \tag{A.112}$$

This is again a set of n equations in n unknowns but now the right-hand side is 0 and so the equations are a set of **homogeneous** algebraic equations. *In order for a solution to exist, the determinant of coefficients must be 0:*

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{A.113}$$

Solving this equation will yield n different values for λ and requires solving an nth order polynomial. We consider the case for two unknowns:

$$\hat{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{A.114}$$

$$\hat{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{A.115}$$

and

$$\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \tag{A.116}$$

We require that

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0 \tag{A.117}$$

or

$$(a - \lambda)(d - \lambda) - cb = 0 \tag{A.118}$$

and after rearranging

$$\lambda^2 - (a + d)\lambda + ad - cb = 0 \tag{A.119}$$

Giving as expected two different eigenvalues,

$$\lambda_\pm = \frac{1}{2}\left((a+d) \pm \sqrt{(a+d)^2 - 4(ad-cb)}\right) \tag{A.120}$$

The complete solution only gives one unknown, say x_2, in terms of x_1 for each of two eigenvalues. So, using the equation from above:

$$\begin{bmatrix} a-\lambda_\pm & b \\ c & d-\lambda_\pm \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \tag{A.121}$$

$$(a-\lambda_\pm)x_1 + bx_2 = 0 \tag{A.122}$$

or

$$x_2 = x_1\frac{(\lambda_\pm - a)}{b} \tag{A.123}$$

$$\widehat{X}_{\lambda_\pm} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ \frac{(\lambda_\pm - a)}{b} \end{bmatrix} \tag{A.124}$$

Assuming that the eigenvalues are not degenerate (i.e., $\lambda_+ \neq \lambda_-$), each eigenvector is different. You can also show that they are orthogonal; i.e., $\widehat{X}_{\lambda_+}^* \cdot \widehat{X}_{\lambda_-} = 0$. In quantum systems using Dirac notation, it would be

$$\widehat{X}_{\lambda_\pm} \rightarrow |\lambda_\pm\rangle = x_1\begin{bmatrix} 1 \\ \frac{(\lambda_\pm - a)}{b} \end{bmatrix} \tag{A.125}$$

and x_1 would be determined then by normalization:

$$x_1 = \frac{1}{\sqrt{1 + \left(\frac{(\lambda_\pm - a)}{b}\right)^2}} \tag{A.126}$$

$$\langle \lambda_i | \lambda_j \rangle = \delta_{ij}, i, j = +/- \tag{A.127}$$

We note now that we have found that

$$\hat{A}\widehat{X}_{\lambda_\pm} = \lambda_\pm \widehat{X}_{\lambda_\pm} \tag{A.128}$$

If we converted

$$\hat{A} \rightarrow \hat{A}' = \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} \tag{A.129}$$

then it is easy to show that

$$\begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix}\widehat{X}'_{\lambda_\pm} = \lambda_\pm \widehat{X}'_{\lambda_\pm} \tag{A.130}$$

where the prime represents the eigenstates for this form of the matrix; the corresponding eigenvectors are

$$\widehat{X}'_{\lambda_+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \widehat{X}'_{\lambda_-} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{A.131}$$

So, we say the matrix $\hat{A}$ has been **diagonalized**.

Matrix Diagonalization

In the above example of solving the eigenvalue problem we found two eigenvectors of the original matrix and then saw how replacing $\hat{A}$ with the diagonal matrix, $\hat{A}'$, the new eigenvectors were the unit vectors *in that basis*, meaning in terms of the eigenvectors of $\hat{A}'$. It is important to know how to mathematically convert from $\hat{A}$ to $\hat{A}'$. For an arbitrary non-singular matrix of order n, $\hat{A}$, we assume that we have found the eigenvalues λ_i and the corresponding normalized eigenvectors $\widehat{X}_{\lambda_i}$ such that

$$\hat{A}\widehat{X}_{\lambda_i} = \lambda_i \widehat{X}_{\lambda_i} \tag{A.132}$$

Consider a square matrix where the columns are the eigenvectors,

$$\hat{S} = \begin{bmatrix} \widehat{X}_{\lambda_1} & \widehat{X}_{\lambda_2} & \cdots & \widehat{X}_{\lambda_n} \end{bmatrix} \tag{A.133}$$

Then

$$\hat{A}\hat{S} = \begin{bmatrix} \lambda_1\widehat{X}_{\lambda_1} & \lambda_2\widehat{X}_{\lambda_2} & \cdots & \lambda_n\widehat{X}_{\lambda_n} \end{bmatrix} \tag{A.134}$$

We now form the adjoint of $\hat{S}$, which you recall is the **complex transpose** of $\hat{S}$:

$$\hat{S}^{\dagger} = \hat{S}^{*T} = \begin{bmatrix} \widehat{X}_{\lambda_1}^{*T} \\ \widehat{X}_{\lambda_2}^{*T} \\ \vdots \\ \widehat{X}_{\lambda_n}^{*T} \end{bmatrix} \tag{A.135}$$

Then

$$\hat{S}^{*T}\hat{A}\hat{S} = \hat{S}^{\dagger}\hat{A}\hat{S} = \begin{bmatrix} \widehat{X}_{\lambda_1}^{*T} \\ \widehat{X}_{\lambda_2}^{*T} \\ \vdots \\ \widehat{X}_{\lambda_n}^{*T} \end{bmatrix} \begin{bmatrix} \lambda_1\widehat{X}_{\lambda_1} & \lambda_2\widehat{X}_{\lambda_2} & \cdots & \lambda_n\widehat{X}_{\lambda_n} \end{bmatrix}$$

$$= \begin{bmatrix} \widehat{X}_{\lambda_1}^{*T}\lambda_1\widehat{X}_{\lambda_1} & \widehat{X}_{\lambda_1}^{*T}\lambda_2\widehat{X}_{\lambda_2} & \cdots & \widehat{X}_{\lambda_1}^{*}\lambda_n\widehat{X}_{\lambda_n} \\ \widehat{X}_{\lambda_2}^{*T}\lambda_1\widehat{X}_{\lambda_1} & \widehat{X}_{\lambda_2}^{*T}\lambda_2\widehat{X}_{\lambda_2} & \cdots & \widehat{X}_{\lambda_2}^{*T}\lambda_n\widehat{X}_{\lambda_n} \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{X}_{\lambda_n}^{*T}\lambda_1\widehat{X}_{\lambda_1} & \widehat{X}_{\lambda_n}^{*T}\lambda_2\widehat{X}_{\lambda_2} & \cdots & \widehat{X}_{\lambda_n}^{*T}\lambda_n\widehat{X}_{\lambda_n} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{A.136}$$

Where, by orthonormality

$$\widehat{X}_{\lambda_i}^{*T} \lambda_k \widehat{X}_{\lambda_k} = 0 \text{ for } i \neq k \tag{A.137}$$

and

$$\widehat{X}_{\lambda_j}^{*T} \lambda_j \widehat{X}_{\lambda_j} = \lambda_j \tag{A.138}$$

Unitary Transformation

Consider an operator (e.g., a matrix), $\hat{A}$. Then let $\widehat{U}$ be a unitary operator (A.83 and discussion, meaning $\widehat{U}^\dagger = \widehat{U}^{-1}$) of the same order as $\hat{A}$ in the same Hilbert space. Then a new operator, $\hat{B} = \widehat{U}\hat{A}\widehat{U}^\dagger$, is the result of a **unitary transformation** of $\hat{A}$. For example, let $\hat{A}\widehat{X} = \lambda\widehat{X}$. Since $\widehat{U}^\dagger\widehat{U} = \widehat{U}\widehat{U}^\dagger = \hat{I}$, then inserting $\widehat{U}^\dagger\widehat{U}$ between $\hat{A}$ and $\widehat{X}$ and then multiplying both sides by $\widehat{U}$, we get

$$\widehat{U}\hat{A}\widehat{U}^\dagger\widehat{U}\widehat{X} = \lambda\widehat{U}\widehat{X} \tag{A.139}$$

Hence, since $\widehat{X}$ is an eigenvector of $\hat{A}$ with eigenvalue λ, then $\widehat{U}\widehat{X}$ is an eigenvector of $\hat{B} = \widehat{U}\hat{A}\widehat{U}^\dagger$ with the same eigenvalue. The same is also true for

$$\widehat{U}^\dagger\hat{A}\widehat{U}\widehat{U}^\dagger\widehat{X} = \lambda\widehat{U}^\dagger\widehat{X} \tag{A.140}$$

The transformation above from a matrix $\hat{A}$ to the matrix $\hat{S}^\dagger\hat{A}\hat{S}$, which is now diagonal, is also a unitary transformation. It is easy to show that the magnitude of both eigenvectors is the same.

Vocabulary (page) and Important Concepts

- first order differential equations 329
- second order differential equations 329
- Homogeneous differential equations 330
- Inhomogeneous differential equations 330
- Method of separation of variables 333
- Eigenvalue problems 334
- Euler's Theorem 336
- Dirac delta-function 337
- Gaussian integral 338
- Matrix multiplication 338
- First minor 339
- Determinant 339

Appendix B Power Series for Important Functions

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots \quad -1 < x < 1$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8}$$

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad -1 < x < 1$$

Taylor series expansion of $f(x)$ about a:

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots$$

Appendix C Properties and Representations of the Dirac Delta-Function

Definition:

$$\int_{-\infty}^{\infty} dx\, f(x)\delta\,(x - x_0) = f(x_0)$$

Identities:

$$\delta(ax) = \frac{1}{|a|}\delta(x)$$

Representations:

$$\delta\,(k - k_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx\, e^{i(k - k_0)x}$$

$$\delta\,(k - k_0) = \frac{1}{\sqrt{\pi}} \lim_{a \to +0} \frac{1}{a} e^{-\left(\frac{k-k_0}{a}\right)^2}$$

$$\delta\,(k - k_0) = \frac{1}{\pi} \int_{0}^{\infty} dx \cos x\,(k - k_0)$$

$$\delta\,(k - k_0) = \frac{1}{\pi} \lim_{a \to \infty} \frac{\sin a\,(k - k_0)}{(k - k_0)}$$

$$\delta\,(k - k_0) = \frac{1}{\pi} \lim_{a \to 0} \frac{a}{(k - k_0)^2 + a^2}$$

Appendix D Vector Calculus and Vector Identities

Cartesian coordinates:

$$\nabla f = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z}$$

$$\nabla \cdot \boldsymbol{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

Cylindrical coordinates:

$$\nabla f = \frac{\partial f}{\partial \rho}\hat{\rho} + \frac{1}{\rho}\frac{\partial f}{\partial \varphi}\hat{\varphi} + \frac{\partial f}{\partial z}\hat{z}$$

$$\nabla \cdot \boldsymbol{A} = \frac{1}{\rho}\frac{\partial \rho A_\rho}{\partial \rho} + \frac{1}{\rho}\frac{\partial A_\varphi}{\partial \varphi} + \frac{\partial A_z}{\partial z}$$

$$\nabla^2 f = \frac{1}{\rho}\frac{\partial}{\partial \rho}\rho\frac{\partial f}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2}$$

Spherical coordinates:

$$\nabla f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \varphi}\hat{\varphi}$$

$$\nabla \cdot \boldsymbol{A} = \frac{1}{r^2}\frac{\partial r^2 A_r}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial (\sin\theta A_\theta)}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial A_\varphi}{\partial \varphi}$$

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial f}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\sin\theta\frac{\partial f}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2}$$

$$= \frac{1}{r}\frac{\partial^2}{\partial r^2}rf + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\sin\theta\frac{\partial f}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2}$$

Divergence theorem:

$$\int_{Volume} dv\, \nabla \cdot \boldsymbol{A}(\boldsymbol{x}) = \int_{Surface} ds\, \hat{\boldsymbol{n}} \cdot \boldsymbol{A}(\boldsymbol{x})$$

Stokes' theorem:

$$\int_{Volume} d\boldsymbol{\ell} \cdot \boldsymbol{A}(\boldsymbol{x}) = \int_{Surface} ds\, \hat{\boldsymbol{n}} \cdot \nabla \times \boldsymbol{A}(\boldsymbol{x})$$

Appendix F Maxwell's Equations in Media, the Wave Equation, and Coupling of a Field to a Two-Level System

In general, Maxwell's equations describe the behavior of the electromagnetic field. Of interest here is to show how the observables that were calculated in the text relate directly to Maxwell's equations. For this discussion, we assume that the field generated is classical though the source is an ensemble of quantum systems. The charge and the electric dipole related to the displacement of the charge are central to Maxwell's equations. The electron spin is another quantum system that also couples directly to Maxwell's equations through the magnetic field. The discussion is nearly parallel.

We start by writing down Maxwell's equations in their most general form:

$$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0 \tag{F.1}$$

$$\nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{J} \tag{F.2}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{F.3}$$

$$\nabla \cdot \boldsymbol{D} = \rho \tag{F.4}$$

In the above, $\boldsymbol{J}$ is the current density and ρ is the charge density. Charge is conserved by the continuity relation given by:

$$\nabla \cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t} = 0 \tag{F.5}$$

The constitutive relationship between $\boldsymbol{D}$ and $\boldsymbol{E}$ is

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} \tag{F.6}$$

and between $\boldsymbol{B}$ and $\boldsymbol{H}$ is

$$\boldsymbol{B} = \mu_0 \boldsymbol{H} + \mu_0 \boldsymbol{M} \tag{F.7}$$

where $\boldsymbol{P}$ is polarization per unit volume (corresponding to charge displacement), $\boldsymbol{M}$ is the magnetization per unit volume (corresponding to the magnetic field resulting from extrinsic and intrinsic angular momentum), $\varepsilon_0 \cong 8.85 \times 10^{-12}$ farads/meter the permittivity of free space, and $\mu_0 \cong 1.2566 \times 10^{-6}$ henries/meter is the permeability of free space.

For a single charge, the source terms become

$$\rho(\boldsymbol{r}, t) = q \left| \psi(\boldsymbol{r}, t) \right|^2 \tag{F.8}$$

The quantum current is

$$J = -\frac{i\hbar}{2m}\left(\psi^*\left(r,t\right)\nabla\psi\left(r,t\right) - \psi\left(r,t\right)\nabla\psi^*\left(r,t\right)\right) \tag{F.9}$$

In bulk media, however, with N being the number of quantum systems per unit volume, the polarization is

$$P(t) = qN\langle\psi\left(r,t\right)|r|\psi(r,t)\rangle \tag{F.10}$$

and the magnetization is

$$M = \frac{q}{2m}N\langle\psi\left(r,t\right)|\hat{J}|\psi(r,t)\rangle \tag{F.11}$$

The angular momentum, $\hat{J}$ in Eq. F.11, is intended to be general here, meaning that if the magnetic moment is due to intrinsic spin, it would be S and there would be a corresponding g-factor).

Here we are interested in the electron described by a two-level Hamiltonian. Hence, we focus just on $P(t) = qN\langle\psi\left(r,t\right)|r|\psi\left(r,t\right)\rangle$ and set the other source terms to 0. Maxwell's equations then become

$$\nabla\times E + \frac{\partial B}{\partial t} = 0 \tag{F.12}$$

$$\nabla\times B - \frac{1}{c^2}\frac{\partial E}{\partial t} = \mu_0\frac{\partial P}{\partial t} \tag{F.13}$$

$$\nabla\cdot B = 0 \tag{F.14}$$

$$\nabla\cdot E = 0 \tag{F.15}$$

Combining with the curl equation, we get

$$\nabla\times\nabla\times E + \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = -\mu_0\frac{\partial^2 P}{\partial t^2} \tag{F.16}$$

Assuming that the vector components are in the Cartesian coordinate system, then $\nabla\times\nabla\times E = \nabla\nabla\cdot E - \nabla^2 E = -\nabla^2 E$ since $\nabla\cdot E$. Substituting this result, we get a wave equation,

$$\nabla^2 E - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = \mu_0\frac{\partial^2 P}{\partial t^2} \tag{F.17}$$

with

$$P = N\langle\mu\rangle = \varepsilon_0\chi E \tag{F.19}$$

and

$$\langle\mu\rangle = Tr\mu\hat{\rho} = \left(\mu_{12}\hat{\rho}_{21} + \mu_{21}\hat{\rho}_{12}\right) \tag{F.20}$$

where $\hat{\rho}$ is the density matrix operator in Chapter 18 and χ is the electric susceptibility.

To see the implication of the field-induced polarization and ignoring the vector nature of the field as well as the transient response, we have from Chapter 18 that $\hat{\rho}_{21} \sim E(z)e^{ikz-i\omega t}$ and $\hat{\rho}_{12} \sim E^*(z)e^{-ikz+i\omega t}$. If we set $P_+ = N\mu_{12}\hat{\rho}_{21}$ and $P_- = N\mu_{21}\hat{\rho}_{12}$, then we can write P_+ as

$$P_+ = \varepsilon_0 \left(\chi_R + i\chi_I\right) E(z)e^{ikz-i\omega t} \tag{F.21}$$

If the field amplitude, $E(z)$, varies slowly on the scale length of a wavelength, then we can ignore terms in the wave equation that go like $\frac{\partial^2}{\partial z^2}E(z)$ (there is no dependence in the field amplitude on x or y for a transverse plane wave). We can then write the wave equation (Eq. F.17) in the form

$$\left(-k^2 + \frac{\omega^2}{c^2}\right) E(z) + ik\frac{\partial}{\partial z}E(z) = -\frac{\omega^2}{c^2}\left(\chi_R + i\chi_I\right) E(z) \tag{F.22}$$

Setting imaginary and real parts equal to each other, we get two equations:

$$\left(-k^2 + \frac{\omega^2}{c^2}\right) = -\frac{\omega^2}{c^2}\chi_R \tag{F.23}$$

and

$$k\frac{\partial}{\partial z}E(z) = -\frac{\omega^2}{c^2}\chi_I E(z) \tag{F.24}$$

where the first equation represents the linear dispersion relation with

$$k^2 = \frac{\omega^2}{c^2}\left(1 + \chi_R\right) \tag{F.25}$$

$$k = n\frac{\omega}{c} \tag{F.26}$$

where $n^2 = (1 + \chi_R)$ is the index of refraction (the ratio of the speed of light in vacuum to the speed of light in the medium)

The second equation can be solved as

$$E(z) = E(z = 0)\, e^{-\frac{\alpha}{2}z} \tag{F.27}$$

where the absorption coefficient (gain if it is negative) is

$$\alpha = \frac{2\omega^2}{kc^2}\chi_I = \frac{2\omega}{n\,c}\chi_I \tag{F.28}$$

To see how to relate the **absorption cross section** to fundamental parameters in the density matrix, we start with Eq. F.20 and use the density matrix from Chapter 18 using first order perturbation theory. From the density matrix,

$$i\dot{\varrho}_{21} = \frac{1}{\hbar}[H,\rho]_{21} - i\left(\frac{d\rho_{21}}{d1}\right)_{decoherence} = (\omega_0 - i\gamma)\rho_{21} - \frac{\mu_{21}\breve{E}e^{-i\omega t}}{2\hbar}(\rho_{11} - \rho_{22}) \tag{F.29}$$

Take for ρ_{21} in the field interaction picture:

$$\rho_{21} = \tilde{\rho}_{21} e^{-i\omega t} \tag{F.30}$$

Then

$$\omega \tilde{\rho}_{21} = \omega_0 \tilde{\rho}_{21} - i\gamma \tilde{\rho}_{21} - \frac{\mu_{21} \tilde{E}}{2\hbar} (\rho_{11} - \rho_{22}) \tag{F.31}$$

$$\tilde{\rho}_{21} = -\frac{\mu_{21} \tilde{E}}{2\hbar \left[(\omega - \omega_0) + i\gamma \right]} (\rho_{11} - \rho_{22}) \tag{F.32}$$

At low power, $\rho_{11} - \rho_{22} = 1$.

$$\tilde{\rho}_{21} = -\frac{\mu_{21} \tilde{E}}{2\hbar \left[(\omega - \omega_0) + i\gamma \right]} \tag{F.33}$$

and so (to first order in E, linear response)

$$\tilde{P}_+ = -N \frac{\mu_{12} \mu_{21} \tilde{E}}{2\hbar \left[(\omega - \omega_0) + i\gamma \right]} \tag{F.34}$$

From this,

$$\alpha = \frac{2\omega^2}{c^2 k} \chi_I = \frac{2\omega}{c} \chi_I \equiv N\sigma \tag{F.35}$$

where σ is the cross section for absorption and is defined by this relationship. Since this is linear theory, we use the phasor $e^{-i\omega t}$ as used in the Maxwell equation work and the quantum work, and we have for the prefactor to the phasor,

$$\tilde{P}_+ = \epsilon_0 (\chi_R + i\chi_I) N\tilde{E} = \mu_{12} \tilde{\rho}_{21} = -\frac{\mu_{12} \mu_{21} \tilde{E}}{2\hbar \left[(\omega - \omega_0) + i\gamma \right]} N = -\frac{\mu_{12} \mu_{21} \tilde{E} \left[(\omega - \omega_0) - i\gamma \right]}{2\hbar \left[(\omega - \omega_0)^2 + \gamma^2 \right]} N \tag{F.36}$$

$$\chi_I = \frac{\mu_{12} \mu_{21} \gamma}{2\hbar\epsilon_0 \left[(\omega - \omega_0)^2 + \gamma^2 \right]} N = \frac{\mu_{12} \mu_{21} \gamma}{2\hbar\epsilon_0 \left[(\omega - \omega_0)^2 + \gamma^2 \right]} N = \frac{\mu_{12} \mu_{21}}{2\hbar\epsilon_0 \gamma} N\mathcal{L}(\Delta) \tag{F.37}$$

where $\mathcal{L}(\Delta) = \frac{\gamma^2}{\left[(\omega - \omega_0)^2 + \gamma^2 \right]}$. Finally, from the form for α above,

$$\sigma = 2\frac{\omega}{c} \frac{\chi_I}{N} = 2\frac{\mu_{12} \mu_{21} \omega}{2\hbar\epsilon_0 c\gamma} \mathcal{L}(\Delta) = \alpha_{FS} \frac{4\pi r_{12} r_{21} \omega}{\gamma} \mathcal{L}(\Delta) \tag{F.38}$$

where the fine structure constant is

$$\alpha_{FS} = \frac{e^2}{4\pi\epsilon_0 \hbar c} \tag{F.39}$$

Note that there is a correction factor of order unity in Eq. F.38 associated with the radial matrix elements in the last expression discussed in Appendix G. The fine structure constant is dimensionless and is approximately $\alpha_{FS} \sim \frac{1}{137}$. The dependence on the fine structure constant is frequently cited

as a hall-mark of optical interactions and it reflects the intrinsically weak interaction between light and charged particles. However, at resonance,

$$\sigma_0 = \alpha_{FS}\frac{4\pi r_{12}r_{21}\omega}{\gamma} \tag{F.40}$$

In the absence of pure dephasing,

$$\gamma = \frac{1}{2}\Gamma_{sp.em.} = \frac{1}{2}A \tag{F.41}$$

where A is the Einstein A-coefficient, or the inverse radiative lifetime.

$$A = \frac{4}{3}\frac{e^2 r_{12}r_{21}\omega^3}{4\pi\epsilon_0\hbar c^3} = \frac{4}{3}\frac{\alpha_{FS}r_{12}r_{21}\omega^3}{c^2} \tag{F.42}$$

Substituting into the cross section at resonance,

$$\sigma_0 = \frac{6\pi c^2}{\omega^2} = \frac{6\pi}{k^2} = \frac{3}{2\pi}\lambda^2 \tag{F.43}$$

There are many important results here for real device studies; however, an important piece of fundamental physics that impacts technology is the following.

In quantum electrodynamics *a famous result is that electromagnetic radiation interacts with charged particles only weakly, resulting from the fine structure constant.* This is misleading when the transition is a resonance and the transition is lifetime broadened. Just for comparison, for the optical wavelength, $\sigma_0 \sim 10^{-12}$ square meters, but for scattering from a free electron (no resonance, Thompson scattering) $\sigma \sim 10^{-28}$ square meters.