Math Review

0.1 Complex Numbers

The imaginary number i is defined by:

$$i^2 = -1 \quad \text{or} \quad i = \sqrt{-1} \tag{1}$$

In terms of i, then, a complex number is written:

$$z = x + iy \tag{2}$$

where x and y are both real. It can be thought of as a vector, or phasor, in a two-dimensional plane with the real part, x, is along the abscissa and the imaginary part, y is along the ordinate [see Figure 1].

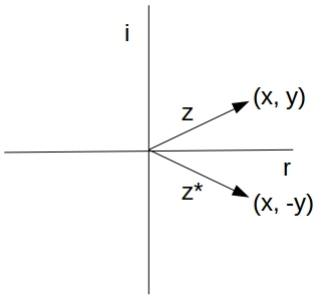

Figure 1: Complex plane representation (phasor) of complex number z. z^* is the complex conjugate of z.

The complex conjugate of a complex number reverses the sign of the imaginary part [see Figure 1].

$$z^* = x - iy \tag{3}$$

The magnitude of an imaginary number is (as should be obvious from Figure 1) the square root of the squared modulus of the imaginary number:

$$|z|^2 = z^* z = (x + iy)(x - iy) = x^2 + y^2 \tag{4}$$

In polar coordinates, where r is the modulus and θ is the phase, $x = r\cos\theta$ and $y = r\sin\theta$. In terms of these quantities,

$$\cos\theta = \frac{x}{r} \qquad \sin\theta = \frac{y}{r} \qquad r^2 = x^2 + y^2 \tag{5}$$

and therefore

$$z = r(\cos\theta + i\sin\theta) \tag{6}$$

and

$$|z|^2 = r^2 \tag{7}$$

An infinitessimal displacement at constant magnitude (r = constant) of a complex quantity, dz, results from an infinitessimal rotation by $d\theta$ in the complex plane:

$$dz = izd\theta \tag{8}$$

A finite displacement at constant magnitude then results from integrating this infinitessimal from θ_{initial} to θ_{final}, $\Theta \equiv \Delta\theta = \theta_{\text{final}}$ - θ_{initial}. Where $r = 1$, all of this leads to the complex exponential,

$$e^{i\Theta} = \cos\Theta + i\sin\Theta \tag{9}$$

Problems

1. Fill in the steps from 8 to 9
2. Because $(e^{i\Theta})^* = e^{-i\Theta}$, show that

$$\cos\Theta = \frac{e^{i\Theta} + e^{-i\Theta}}{2} \tag{10}$$

and

$$\sin\Theta = \frac{e^{i\Theta} - e^{-i\Theta}}{2i} \tag{11}$$

0.2 Dirac Delta "Function"

To model the mass or charge density of a point object, Paul Dirac introduced the distribution heuristicaly defined as:

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0 \end{cases} \tag{12a}$$

$$\int_{-\infty}^{+\infty} \delta(x)dx = 1 \tag{12b}$$

If $f(x)$ is a continuous function, then, since the product $f(x)\delta(x)$ is finite only at $x = 0$ (and zero otherwise),

$$f(x)\delta(x) = f(0)\delta(x) \tag{13}$$

Thus,[1]

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = f(0)\int_{-\infty}^{+\infty} \delta(x)dx = f(0) \tag{14}$$

The distribution can be shifted from zero:

$$\delta(x-a) = \begin{cases} +\infty, & x = a \\ 0, & x \neq a \end{cases} \tag{15a}$$

$$\int_{-\infty}^{+\infty} \delta(x-a)dx = 1 \tag{15b}$$

which generalize Equations 13:

$$f(x)\delta(x-a) = f(a)\delta(x-a) \tag{16}$$

and 14:

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a)\int_{-\infty}^{+\infty} \delta(x-a)dx = f(a) \tag{17}$$

and amounts to a sampling (or sifting) of $f(x)$.

On the other hand, $\delta(kx)$, where k is a non-zero real number, scales $f(x)$. If $y = kx$, then $x = y/k$, so $dx = dy/k$, and[2]

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)\delta(kx)dx &= \pm\int_{-\infty}^{+\infty} f(y/k)\delta(y)\frac{dy}{k} \\ &= \pm\frac{1}{k}f(0) = \frac{1}{|k|}f(0) \end{aligned} \tag{18}$$

In other words,

$$\delta(kx) = \frac{1}{|k|}\delta(x) \tag{19}$$

Each of these is a special case of $\delta(g(x))$, where $g(x)$ is a continuous function of x.

$$\delta(g(x)) = \begin{cases} +\infty, & g(x) = 0 \\ 0, & g(x) \neq 0 \end{cases} \tag{20}$$

That is, at the zeroes of $g(x)$, the distribution $\delta(g(x))$ spikes. Expanding in a Taylor series $g(x)$ around the ith zero,

[1] Any limits either side of zero are sufficient.

[2] If $k > 0$, the limits are as before, but if $k < 0$ they reverse.

$$g(x) = g(x_i) + (x - x_i)g'(x_i) + \frac{1}{2}(x - x_i)^2 g''(x_i) \simeq (x - x_i)g'(x_i) \tag{21}$$

The size of the spike at $x = x_i$ can be found from equations 21, 16, and 19,

$$\delta(g(x)) = \delta(g'(x_i)(x - x_i)) = \frac{1}{|g'(x_i)|}\delta(x - x_i) \qquad (x \simeq x_i) \tag{22}$$

That is, the value of the $\delta(g(x))$ spike at $x = x_i$ is $\delta(x - x_i)$ scaled by $1/|g'(x_i)|$. Therefore,

$$\delta(g(x)) = \sum_{i=1}^{n} \frac{1}{|g'(x_i)|}\delta(x - x_i) \tag{23}$$

Problems

1. Evaluate $\delta(x^2 + x - 6)$
2. Evaluate and sketch $\delta(\cos x)$
3. Evaluate $\int_0^3 ln(x + 1)\delta(\pi - x)dx$
4. Evaluate $\int_0^3 3x^3 - 4x + 1)\delta(x - 2)dx$

The three-dimensional generalization of $\delta(x)$ is

$$\delta^3(\vec{r}) = \delta(x)\delta(y)\delta(z) \tag{24}$$

which obviously is zero except at the origin:

$$\int_{-\infty}^{+\infty} \delta^3(\vec{r})d^3r = \int_{-\infty}^{+\infty} \delta(x)\delta(y)\delta(z)dxdydz = 1 \tag{25}$$

Likewise,

$$\int_{-\infty}^{+\infty} f(\vec{r}\delta^3(\vec{r} - \vec{r}_0)d^3r = f(\vec{r}_0) \tag{26}$$

And so, finally, the charge density (charge per unit volume) of a point charge q at position $\vec{r}_0$ can be modeled as

$$\rho(\vec{r}) = q\delta(\vec{r} - \vec{r}_0) \tag{27}$$

Further, consider the Heaviside step-function,

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases} \tag{28}$$

and notice that the derivate $d\theta/dx$ is zero everywhere but at the origin, so,

$$\int_{-\infty}^{+\infty} \frac{d\theta}{dx}dx = \theta(+\infty) - \theta(-\infty) = 1 - 0 = 1 \tag{29}$$

Thus, the derivative of the Heaviside step-function and the Dirac delta function behave the same.

0.3 Dirac Notation

A *ket* represents a column vector of n complex numbers (n is the dimension of the vector):

$$|a\rangle \equiv \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \tag{30}$$

where the a_i are complex.

The *ket* contains what is known about, or the initial condition or state, of the system. A system in a state Ψ is representated by the state vector $|\Psi\rangle$. For example, $|p\rangle$ says that the system (particle) has momentum $\vec{p}$, and $|x\rangle$ says that the particle is at position $\vec{x}$.

Since a *ket* represents a vector, the sum of two *kets* is equivalent to the component-wise sum of two vectors:

$$|a\rangle + |b\rangle = |a+b\rangle = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \\ \vdots \\ a_n + b_n \end{pmatrix} \tag{31}$$

Note that the dimensions of $\vec{a}$ and $\vec{b}$ must be the same.

It follows from this that the multiplication of a *ket* by a (complex) number results in each term being multiplied, or *scaled*, by that number:

$$w\,|a\rangle = w \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} wa_1 \\ wa_2 \\ wa_3 \\ \vdots \\ wa_n \end{pmatrix} = |wa\rangle \tag{32}$$

A *bra* is the adjoint of a *ket*; it represents a row vector that is the transpose of a column vector with elements that are complex conjugates of the elements of the column vector:

$$\langle a| = (|a\rangle)^\dagger \equiv (a_1^*, a_2^*, a_3^*, \ldots, a_n^*) \tag{33}$$

It represents the final state of the system.

The multiplication of a *ket* by a *bra*, $\langle b|a\rangle$, is equivalent to a dot-product between two vectors–one a row vector, the other a column vector–of the same dimension:

$$\langle b|a\rangle = b_1^* a_1 + b_2^* a_2 + b_3^* a_3 + \cdots + b_n^* a_n \tag{34}$$

which is a (complex) number. Note that $\langle a|b\rangle \neq \langle b|a\rangle$, but $(\langle a|b\rangle)^* = \langle b|a\rangle$.

Recall that the dot- or inner-product of two perpendicular (orthogonal) vectors is zero, and the inner-product of two (normalized) parallel vectors is 1. In Dirac notation, this is expressed:

$$\langle b|a\rangle = \delta_{ab} \tag{35}$$

where δ_{ab} is the Dirac Delta function, which equals 1 when $a = b$ and 0 otherwise.

With normalized vectors, $\langle b|a\rangle$ is the *probability amplitude* and $|\langle b|a\rangle|^2$ is the *probability* that a state $|a\rangle$ transforms into a state $|b\rangle$, or, alternatively, that an object in state $|a\rangle$ is measured in an experiment to be in state $|b\rangle$. Both quantities are zero if $\vec{a}$ and $\vec{b}$ are orthogonal.

Problems

1. Let

$$|\alpha\rangle = \begin{pmatrix} i \\ 2 \\ 3 \end{pmatrix} \qquad |\beta\rangle = \begin{pmatrix} 1 \\ 1 \\ i \end{pmatrix}$$

   Compute

   (a) $\langle\alpha|\alpha\rangle$
   (b) $\langle\beta|\beta\rangle$
   (c) $\langle\alpha|\beta\rangle$
   (d) $\langle\beta|\alpha\rangle$

2. Show $\langle b|a\rangle = (\langle a|b\rangle)^*$

0.4 Matrices

An *operator* is a matrix which maps one vector onto another vector.

Consider a matrix $\mathcal{M}$, with m rows and n columns. Its elements are identified by m_{ij}, given by $\langle i|\,\mathcal{M}\,|j\rangle$ The order of a matrix is $m \times n$. Its rank is $\leq n$.

The complex conjugate of $\mathcal{M}$, $\mathcal{M}^*$, has elements m^*_{ij}, and the transpose of $\mathcal{M}$, $\mathcal{M}^\mathcal{T}$, has elements m_{ji} (that is, the columns and rows are interchanged).

The adjoint or conjugate transpose of $\mathcal{M}$, $\mathcal{M}^\dagger$ (sometimes referred to as "$\mathcal{M}$ dagger"), transposes rows and columns and conjugates all elments, so that the elements are m^*_{ji}. A matrix satisfying $\mathcal{M} = \mathcal{M}^\dagger$ is called *Hermitian*, and the adjoint is called the Hermitian conjugate.

The identity matrix, $\mathcal{I}$, consists of ones along the diagonal and zeroes everywhere else, $m_{ii} = 1$, $m_{ij} = 0$ $(i \neq j)$. A matrix which, when multiplied by its adjoint (conjugate transpose) results in the identity matrix, is called a *unitary* matrix, $\mathcal{U}$. Note that in this case the the matrix and its adjoint are inverse matrices.

$$\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = \mathcal{U}\mathcal{U}^{-1} = \mathcal{I} \tag{36}$$

Note that a unitary matrix must be square, that is, its *order* is $n \times n$, and its *rank* is n.

The importance of matrices to physics is that they transform vectors, that is, they transform a system from one state to another. For square matrices, a scaling factor or magnitude of a transformation is given by the determinant of the matrix, $\det\{\mathcal{M}\}$. In two dimensions, this is

$$\det\{\mathcal{M}\} = \begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} = m_{11}m_{22} - m_{12}m_{21} \tag{37}$$

In three dimensions, this is

$$\begin{aligned}
\det\{\mathcal{M}\} &= \begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} \\
&= m_{11}\begin{vmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{vmatrix} - m_{12}\begin{vmatrix} m_{21} & m_{23} \\ m_{31} & m_{33} \end{vmatrix} + m_{13}\begin{vmatrix} m_{21} & m_{22} \\ m_{31} & m_{32} \end{vmatrix} \\
&= m_{11}m_{22}m_{33} + m_{12}m_{23}m_{31} + m_{13}m_{21}m_{32} \\
&\quad -m_{11}m_{23}m_{32} - m_{12}m_{21}m_{33} - m_{13}m_{22}m_{31}
\end{aligned} \tag{38}$$

And, in general,

$$\det\{\mathcal{M}\} = \sum_{k=1}^{n} m_{ik}C_{ik} = \sum_{k=1}^{n} m_{kj}C_{kj} \tag{39}$$

where C_{ij} is the i, j cofactor of $\mathcal{M}$:

$$C_{ij} = (-1)^{i+j}M_{ij} \tag{40}$$

where M_{ij} is the minor of $\mathcal{M}$, the determinant of the $(n-1)\times(n-1)$ matrix that results from deleting the i-th row and j-th column of $\mathcal{M}$.

The *trace* of a square matrix is the sum of the main diagonal's elements:

$$\mathrm{Tr}\{\mathcal{M}\} = \sum_{k=1}^{n} m_{kk} \tag{41}$$

As a *ket* represents a vector, applying a matrix to a *ket* is the same as left-multiplication of a vector by a matrix (the number of matrix columns must be the same as the dimension of the vector):

$$\mathcal{M}\left|a\right\rangle = \begin{pmatrix} m_{11}a_1 + m_{12}a_2 + m_{13}a_3 + \cdots + m_{1n}a_n \\ m_{21}a_1 + m_{22}a_2 + m_{23}a_3 + \cdots + m_{2n}a_n \\ m_{31}a_1 + m_{32}a_2 + m_{33}a_3 + \cdots + m_{3n}a_n \\ \vdots \\ m_{m1}a_1 + m_{m2}a_2 + m_{m3}a_3 + \cdots + m_{mn}a_n \end{pmatrix} = \left|b\right\rangle \tag{42}$$

That is, left-multiplying a *ket* by a matrix (an operator) results in another *ket*. The original *ket* is transformed into the resulting *ket*, and the matrix is called a *transformation matrix*. Note that in general it is not necessary that the number of matrix rows equals the vector dimension. In quantum mechanics, they invariably are equal.

The *projection operator* is a square matrix formed from the outer product of a normalized vector with itself. In Dirac notation:

$$\left|i\right\rangle\left\langle i\right| \tag{43}$$

The *bra* forms an inner product with a state, giving the probability amplitude for a transition from the original state into state $\vec{i}$. This amplitude becomes the numerical coefficient–a scaling factor–of the state $\vec{i}$.

The sum of projection operators, over a complete set of possible transitions, is the identity matrix:

$$\sum_i |i\rangle \langle i| = I \tag{44}$$

so any state can be decomposed into a linear sum of all possible transitions:

$$|a\rangle = \sum_i |i\rangle \langle i|a\rangle \tag{45}$$

If the matrix $\mathcal{M}$ is Hermitian (refer to it then as $\mathcal{H}$), and the *ket* $|a\rangle$ is an *eigenvector* (refer to it then as $|v\rangle$ of $\mathcal{H}$, then left-multiplying $|v\rangle$ by $\mathcal{H}$ returns $|v\rangle$ scaled by a (real) number called an *eigenvalue*, λ_v:

$$\mathcal{H}\,|v\rangle = \lambda_v\,|v\rangle \tag{46}$$

An Hermitian matrix may have multiple eigenvectors with corresponding (usually, but not always, different) eigenvalues.

In quantum mechanics, measurements of a system are represented by Hermitian matrices (operators), the eigenvectors represent the possible states of the system, and the eigenvalues represent the outcomes or results of the measurements.

Problems

Prove:

1. $\mathrm{Tr}\{\mathcal{A}+\mathcal{B}\} = \mathrm{Tr}\{\mathcal{A}\} + \mathrm{Tr}\{\mathcal{B}\}$
2. $\mathrm{Tr}\{\alpha\mathcal{A}\} = \alpha\,\mathrm{Tr}\{\mathcal{A}\}$
3. $\mathrm{Tr}\{\mathcal{A}\mathcal{B}\} = \mathrm{Tr}\{\mathcal{B}\mathcal{A}\}$

0.5 Lagrangian and Hamiltonian Formulations

0.5.1 Lagrange Equation

The Principle of Least Action, also known as Hamilton's Principle, is considerd the most general formulation leading to the laws of motion. It states that every (classical) system can be characterized by a function, called the Lagrangian, of so-called generalized coordinates, $q_i \quad (i = 1, \ldots, q_n)$, which are position-like, and $\dot{q}_i \quad (i = 1, \ldots, q_n)$, which are velocity-like.

$$L(q, \dot{q}, t) = T(q, \dot{q}, t) - U(q, \dot{q}, t) \tag{47}$$

where T is the "kinetic energy" and U is the "potential energy." The coordinates, chosen for convenience without reference to a coordinate system, characterize the dynamical behavior of a system, with a time-like variable as the independent variable.

The Principle of Least Action states further that the system proceeds between the boundary conditions $q(t_1) \equiv q_1$ and $q(t_2) \equiv q_2$ in such a way as to minimize the action,

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt \tag{48}$$

That is, as the system evolves $q \to q + \delta q$, and $\dot{q} \to \dot{q} + \delta\dot{q}$ in such a way that the action is minimized and the boundary conditions are met, $\Delta S = 0$, $\delta q_1 = 0$, and $\delta q_2 = 0$. Thus,

$$\Delta S = \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta\dot{q}, t)dt - \int_{t_1}^{t_2} L(q, \dot{q}, t)dt = 0 \tag{49}$$

Expanding the first term,

$$\int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta\dot{q}, t)dt = \int_{t_1}^{t_2} \left(L(q, \dot{q}, t) + \frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} \right) dt + \cdots \tag{50}$$

Then, to first-order,

$$\Delta S = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} \right) dt = 0 \tag{51}$$

Since $\delta\dot{q} = \frac{d\delta q}{dt}$, Equation 51 can be integrated by parts:

$$\Delta S = \left[\frac{\partial L}{\partial \dot{q}}\delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) \delta q \; dt = 0 \tag{52}$$

The first term is zero, due to the boundary conditions, and the only way for the integral to be always zero is for the integrand to be zero:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0 \tag{53}$$

This is the single-particle Lagrangian, which can be generalized to a system with more particles (n degrees of freedom).

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \tag{54}$$

for $i = 0, 1, \ldots, n$. These are a set of n second-order differential equations whose general solution must include $2n$ constants, determined from inital "coordinates" and "velocities."

Lagrangians are additive, $L = L_A + L_B + \cdots$, and defined (unique) only up to the addition of a total-time derivative of a function that depends on coordinates and time (not velocity).

Because free space is isotropic and also homogenous, as is time, the free-particle Lagrangian is independent of coordinate, time or direction of motion. Therefore, the free-particle Lagrangian can depend only on the magnitude of the velocity, or, equivalently, $\dot{q}^2$, $L \sim \dot{q}^2$, which we can write, classically:

$$L = \frac{1}{2}m\dot{q}^2 \tag{55}$$

or, for a system of non-interacting free particles,

$$L = \sum \frac{1}{2}m_i\dot{q}_i^2 \tag{56}$$

The proportionality constant, m, must have dimensions of mass; $\frac{1}{2}$ is convention.

For motion in an external field, the Lagrangian takes the form of

$$L = \frac{1}{2}m\dot{q}^2 - U(q,t) \tag{57}$$

where U is a potential field.

Problem

A particle of mass m moves under the influence of a conservative force, that is, a force derived from the gradient of a time-independent scalar potential,

$$\mathbf{F} = -\nabla V(r)$$

Use the Lagrange methods to show that this system satisfies Newton's 2nd Law.

0.5.2 The Hamiltonian and Hamilton's Equations

The Hamiltonian is the total energy of the system:

$$H = T + U = 2T - L \tag{58}$$

It's written in terms of generalized coordinates, q, and generalized momenta, p. The equations of motion are determined by partial derivatives of the Hamiltonian with respect to these, the so-called Hamilton's or canonical equations:

$$\frac{\partial H}{\partial p} = \dot{q} \tag{59}$$

$$\frac{\partial H}{\partial q} = -\dot{p} \tag{60}$$

Note that if the Hamiltonian of a system has no explict time dependence, then the energy of the system is conserved.

Problem

A particle of mass m moves under the influence of a conservative force, that is, a force derived from the gradient of a time-independent scalar potential,

$$\mathbf{F} = -\nabla V(r)$$

Use Hamilton's methods to show that this system satisfies Newton's 2nd Law.

0.6 Symmetries and Groups

A symmetry is an operation that can (at least in principle) be performed on a system yet leaves it invariant: the system's final configuration is indistinguishable from its original configuration.

As we now understand physics, symmetries govern interactions. Emily Noether's 1917 theorem that symmetries of nature yield conservation laws, and, conversely, that conservation laws reflect underlying

symmetries, made clear the dynamical implications of symmetry. The inverse implies that when a symmetry is spoiled (broken), the quantity associated with it may not be conserved. Consider 1, which lists well-known, global relationships.

Table 1: Associated symmetries and conservation laws.

Symmetry	**Conservation Law**
space translation	momentum
time displacement	energy
rotation	angular momentum
gauge transformation	charge

The set of all symmetry operations on a given system forms a group, and therefore:

1. the set is closed under "multiplication": if two elements of the set, A_i and A_j act successively on the system, [3] A_iA_j, there exists another element of the set, A_k, whose action on the system is the same as the product's, $A_iA_j = A_k$.

2. the set contains an identity element, I: $IA_i = A_iI = A_i$ for all i elements in the set.

3. the set contains inverse elements, A_i^{-1} for every element: $A_iA_i^{-1} = A_i^{-1}A_i = I$ for all i elements in the set.

4. the set's "multiplication" rule is associative: $A_i(A_jA_k) = (A_iA_j)A_k$

The set's multiplication rule may be communitive (Abelian), $A_iA_j = A_jA_i$, for all elements of the set, but generally is not (non-Abelian).

Groups can be infinite or finite, continuous or discrete. The elements of continuous groups typically depend on one or more continuous parameters. For reference, if that dependence takes the form of an analytical function, the group is called a Lie group.

Most physically relevant groups can be represented as matrices, and, of these, the most relevant are unitary $n \times n$ matrices [see Table 2].

Table 2: Most relevant groups and their $n \times n$ matrix representations.

Group	Matrix Representation
$U(n)$	unitary ($U^\dagger U = 1$)
$SU(n)$	unitary, determinant 1
$O(n)$	orthogonal ($O^T O = 1$)
$SO(n)$	orthogonal, determinant 1

Rotations (of either the reference frame, keeping the system fixed, or of the system, keeping the axes fixed) about the origin of three-dimensional Euclidean space form a symmetry group, R.

[3] A_j is performed first, and then A_i is performed on the result.

This group is a Lie group: every finite rotation can be expressed as the composition (product) of successive infinitesimal rotations. If ϵ is an infinitesimal angle, then, to first order, an infinitesimal rotation about the α-axis, $\alpha = 1,\ 2,\ 3$ ($x,\ y,\ z$), can be written

$$U(\epsilon) = 1 - i\epsilon J_\alpha \tag{61}$$

U is the rotation operator, and J_α are called the generators of rotations about the α axis. In this notation, a rotation is written $\psi' = U\psi$.

A finite rotation, again, would be constructed by successive (say, n) of such infinitesimal rotations:

$$\begin{aligned} U(\theta) = (U(\epsilon))^n &= \left(1 - i\frac{\theta}{n}J_\alpha\right)^n \\ &\underset{n\to\infty}{=} e^{-i\theta J_\alpha} \end{aligned} \tag{62}$$

A rotation of this sort, that is, $|\psi\rangle \to |\psi'\rangle = U|\psi\rangle$, cannot change the probability that $|\psi\rangle$ will be found in state $|\phi\rangle$: $|\langle\phi|\psi\rangle|^2 = |\langle\phi'|\psi'\rangle|^2 = |\langle\phi|U^\dagger U|\psi\rangle|^2$. This implies that U is unitary. The operations $U(R_1),\ U(R_2), \ldots$ form a unitary representation of the rotation group composed of $R_1,\ R_2, \ldots$.

A rotation of this sort will also not change the total energy of the system, so the Hamiltonian is unchanged: $\langle\phi'|H|\psi'\rangle = \langle\phi|U^\dagger HU|\psi\rangle = \langle\phi|H|\psi\rangle$, so $H = U^\dagger HU$, which is to say

$$UH - HU \equiv [U,\ H] = 0 \tag{63}$$

In other words, U and H commute. $[U,\ H]$ is called the commutator of U and H.

Note that the rotation operator has no explicit time dependence and leaves the equations of motion (Schrödinger's equation) unchanged:

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \quad \text{and} \quad H\langle\psi(t)| = -i\frac{d}{dt}\langle\psi(t)| \tag{64}$$

Therefore, the expectation value of U is a constant of the motion:

$$i\frac{d}{dt}\langle\psi(t)|U|\psi(t)\rangle = \langle\psi(t)|UH - HU|\psi(t)\rangle = 0 \tag{65}$$

The unitarity of U says something important about J_α: $1 = U^\dagger U = (1 + i\epsilon J_\alpha^\dagger)(1 - i\epsilon J_\alpha) = 1 + i\epsilon(J_\alpha^\dagger - J_\alpha) + \mathcal{O}(\epsilon^2)$. Because $J_\alpha^\dagger - J_\alpha = 0$, J_α is hermitian and therefore an observable.

To see which observable J_α is, we take the point of view that the system is rotated around the origin of a fixed reference frame,

$$\psi'(\mathbf{r}) = \psi(R^{-1}\mathbf{r}) \tag{66}$$

the wave function of the system rotated to position $\mathbf{r}$, $\psi'(\mathbf{r})$, equals the original wave function at its position before the rotation $R^{-1}\mathbf{r}$, $\psi(R^{-1}\mathbf{r})$. Then, an infinitesimal rotation about, say, the 3- or z-axis,

$$\begin{aligned} U\psi(\mathbf{r}) &= \psi(R^{-1}\mathbf{r} = \psi(x + \epsilon y, y - \epsilon x, z) \\ &\approx \psi(x,\ y,\ z) + \epsilon\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right) \\ &= [1 - i\epsilon(xp_y - yp_x)]\psi \\ &= (1 - i\epsilon L_z)\psi \end{aligned} \tag{67}$$

shows that J_α is the α-component of the angular momentum operator. Rotational symmetry (measurements of the system in different orientations give the same results) leads, since the expectation value of U is a constant of the motion, to conservation of angular momentum, as listed in Table 1.

Problem

Recall that (with $\hbar = 1$) the quantum mechanical momentum operator is $p_\alpha = -i\frac{\partial}{\partial\alpha}$, where $m = x,\ y,\ z$. Referring to Equation 67 (and remember the product rule of derivatives), show that $[J_1,\ J_2] = iJ_3$.

This is a particular instance of the general case

$$[J_j,\ J_k] = i\varepsilon_{jkl}J_l \tag{68}$$

known as the commutation algebra of the rotation group generators. The so-called structure constants of the group, $\varepsilon_{jkl} = +1$ (-1) if jkl are cyclic (anti-cyclic) permutations of 1 2 3 and zero otherwise, define the properties of the rotation group: J_α form a Lie algebra. Note that Equation 68 implies that no two J_αs with different αs can be measured simultaneously; only one component of angular momentum can be measured independently.

Nonlinear functions of group generators, so-called invariants or Casimir operators, may commute with the generators, and thus independently and simultaneously measurable. In the case of the rotation group, there is only one such operator:

$$J^2 = J_1^2 + J_2^2 + J_3^2 \tag{69}$$

and its eigenstate can be measured simultaneously with one of the angular momentum generators, typically J_3, $|jm\rangle$ to produce

$$J^2\,|jm\rangle = j(j+1)\,|jm\rangle \tag{70a}$$
$$J_3\,|jm\rangle = m\,|jm\rangle \tag{70b}$$

where, in quantum mechanics, $j = 0, \frac{1}{2}, 1, \frac{2}{3}, 2, \ldots$ and $m = -j, -j+1, \ldots, j-1, j$, so that there are $2j+1$ different states with the same j. Such states can be the basis of a $(2j+1)$-dimensional irriducible representation of the rotation group

Stepping through the various ms for any given j is accomplished with so-called step-up and step-down (ladder) operators

$$J_\pm = J_1 \pm iJ_2 \tag{71a}$$
$$J_\pm\,|jm\rangle = [j(j+1) - m(m\pm1)]^{1/2}\,|j,\ m\pm1\rangle \tag{71b}$$

The combined operator of a composite system

$$\mathbf{J} = \mathbf{J}_A + \mathbf{J}_B + \cdots \tag{72}$$

also satisfies the Lie algebra, Equation 68. Its basis can be formed thus:

$$|J_A J_B \ldots m_A m_B \ldots\rangle = |J_A m_A\rangle\,|J_B m_B\rangle \ldots \tag{73}$$

The conserved quantities will be J^2 and J_3 with eigenvalues $J(J+1)$ and M, respectively.

For the combination of just two systems, the product of the irreducible $2J_A+1$ and $2J_B+1$ dimension representations can be reduced to the sum of irreducible $2J+1$ dimension representations with bases $|J_A J_B JM\rangle$, where $J = |J_A - J_B|, |J_A - J_B|+1, \ldots, J_A + J_B - 1, J_A + J_B$, and $M = m_A + m_B$.

The two sorts of bases are related in terms of Clebsch-Gordon coefficients, $C(m_A m_B;\ JM)$:

$$|J_A J_B JM\rangle = \sum_{m_A,\ m_B} C(m_A m_B;\ JM)\, |J_A J_B m_A m_B\rangle \tag{74}$$

If, for example, two spin-$\frac{1}{2}$ particles are combined, the total spin could be $J = 0, 1$, leading to the decomposition, using the size of the multiplet to label the irreducible representations:

$$2 \otimes 2 = 3 \oplus 1 \tag{75}$$

Adding a third spin-$\frac{1}{2}$ particle yields a spin-$\frac{3}{2}$ quartet and two spin-$\frac{1}{2}$ doublets:

$$\begin{aligned} (2 \otimes 2) \otimes 2 &= (3 \otimes 2) \oplus (1 \otimes 2) \\ &= 4 \oplus 2 \oplus 2 \end{aligned} \tag{76}$$

0.6.1 SU(2)

While every number is a trivial represention of the $j = 0$ rotation group, the $j = \frac{1}{2}$ group is the smallest not trivial representation. Its generators are usually written in terms of the Pauli matrices, σ_α,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{77}$$

$$J_\alpha = \frac{1}{2}\sigma_\alpha \quad \alpha = 1, 2, 3 \tag{78}$$

Because only one of the Pauli matrices is diagonal–identified with the rank of the group–the generators do not commute. In fact, they anti-commute.

$$\sigma_\alpha \sigma_\beta = -\sigma_\beta \sigma_\alpha \tag{79a}$$

$$[\sigma_\alpha,\ \sigma_\beta] = 2i\varepsilon_{\alpha\beta\gamma}\sigma_\gamma \tag{79b}$$

$$\{\sigma_\alpha,\ \sigma_\beta\} = 2\delta_{\sigma\beta} I \tag{79c}$$

where $\{a,\ b\} = ab + ba$ (anti-commutation algebra), and $\delta_{ab} = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$ (Kronecker delta).

Usually, the basis for this representation is constructed from the eigenvectors of σ_3

$$|m = +1 \text{ or } \uparrow\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |m = -1 \text{ or } \downarrow\rangle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{80}$$

referring to the spin-up ($m = +1$ or $\uparrow$) and spin-down ($m = -1$ or $\downarrow$) projections along the z-axis of a spin-$\frac{1}{2}$ particle. These states are related through step (ladder) operator:

$$|\uparrow\rangle \Leftrightarrow |\downarrow\rangle : \quad \frac{1}{2}(\sigma_1 \pm i\sigma_2) \tag{81}$$

Clearly, as necessary, the transformation operators

$$U(\theta) = e^{-i\theta\sigma_\alpha/2} \tag{82}$$

are unitary, and the generators, σ_α are hermitian.

The set of all 2×2 unitary matrices forms the group $U(2)$, but $U(\theta_\alpha)$ is a subset of these. Notice that the Pauli matrices all have trace 0. Now,

$$\det\{U(\theta\} = e^{-i\theta\,\mathrm{Tr}\{\sigma_\alpha\}/2} = 1 \tag{83}$$

for each α, and the unit determinant is preserved when any of these matrices is multiplied against another. This special case of traceless, unitary matrices forms a subgroup of $U(2)$, referred to as $SU(2)$, and the Pauli miatrices are identified as the fundamental representation of $SU(2)$.

Problem

By expanding the exponential in a Taylor series, and noting that $\sigma_\alpha^2 = I$, show that rotating a spin-$\frac{1}{2}$ system (finite) angle θ about any axis is the unitary transformation

$$e^{-i\theta\sigma_\alpha/2} = \cos\frac{\theta}{2} - i\sigma_\alpha \sin\frac{\theta}{2}$$

0.6.2 SU(3)

SU(3) is the group of unitary, determinant 1 matrices, with $3^2 - 1 = 8$ linearly independent, traceless 3×3 matrices, λ_α, $\alpha = 1,\ 2, \ldots,\ 8$, as generators. In the Gell-Mann representation,

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda_8 = \sqrt{\frac{1}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{aligned} \tag{84}$$

In this convention, the top left quandrants of λ_1, λ_2, and λ_3 are the Pauli matrices, and so these three matrices form a $SU(2)$ subgroup of $SU(3)$.

No more than two of the eight traceless 3×3 matrices may be diagonal. The rank then is 2, indicating that the maximum number of mutually commuting matrices among the eight is 2, and that the number of Casimir operators is also 2. When the eight generators are the Gell-Mann matrices, the two diagonal matrices are conventionally labeled λ_3 and λ_8, which have simultaneous eigenvectors

$$|R\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad |G\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad |B\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{85}$$

These states are also related through step (ladder) operators:

$$|G\rangle \Leftrightarrow |R\rangle : \quad \frac{1}{2}(\lambda_1 \pm i\lambda_2) \quad |R\rangle \Leftrightarrow |B\rangle : \quad \frac{1}{2}(\lambda_4 \pm i\lambda_5) \quad |B\rangle \Leftrightarrow |G\rangle : \quad \frac{1}{2}(\lambda_6 \pm i\lambda_7) \tag{86}$$

The Lie algebra for $SU(3)$

$$\left[\frac{\lambda_\alpha}{2}, \ \frac{\lambda_\beta}{2}\right] = i \sum_\gamma f_{\alpha\beta\gamma} \frac{\lambda_\gamma}{2} \tag{87}$$

contains the three-index structure constants $f_{\alpha\beta\gamma}$, which are completely antisymmetric under interchange of any pair of indices. The only non-vanishing triplets contain one or three of the indices 2, 5, and 7, the imaginary matrices:

$$f_{123} = 1, \ f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = \frac{1}{2}, \ f_{458} = f_{678} = \frac{\sqrt{3}}{2} \tag{88}$$