Quantum Mechanics

E. F. Faust

with $\LaTeX$

August 25, 2025

Mathematics

Linear Algebra

Introduction

Elementary row operation. Rule as follows.

1. Interchange two rows

2. Multiply (or divide) a row by a (nonzero) constant

3. Add a multiple of one row to another

Rank. Definition as follows.

The number of nonzero rows remaining when a matrix has been row reduced is called the rank of the matrix.

Or:

The order of the largest nonzero determinant is the rank of the matrix.

Consider m equation with n constants. We define matrix M and A, where M has m rows and n column–which corresponds to m equation and n unknown–while A has m rows and $n+1$–which corresponds to unknown plus the constant. There are few possible cases.

1. If (rank M) $<$ (rank A), the equations are inconsistent and there is no solution.

2. If (rank M) $=$ (rank A) $= n$ (number of unknowns), there is one solution.

3. If (rank M) $=$ (rank A) $= R < n$, then R unknowns can be found in terms of the remaining $n - R$ unknowns.

Cramer's Rule. The equations

$$\begin{cases} a_1x + b_1y & = c_1 \\ a_2x + b_2y & = c_2 \end{cases}$$

has the solution:

$$x = \frac{1}{D}\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \qquad \text{and} \qquad y = \frac{1}{D}\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

Dot Product.

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos \theta$$
$$= A_x B_x + A_y B_y + A_z B_z$$

The following applies if vector perpendicular:

$$\mathbf{A} \cdot \mathbf{B} = 0$$

The following applies if vector parallel:

$$\frac{A_x}{B_x} = \frac{A_y}{B_y} = \frac{A_z}{B_z}$$

Cross Product.

$$\mathbf{A} \times \mathbf{B} = |\mathbf{A}||\mathbf{B}| \sin\theta$$

$$= \det \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

The following applies if vector parallel or antiparallel:

$$\mathbf{A} \cdot \mathbf{B} = 0$$

Vector in Braket Notation

Vector space. Linear vector space $\mathbb{V}$ is a collection of vectors $|1\rangle, \ldots, |n\rangle$ for which there exists definitive rule for addition and multiplication. Said rules are as follows.

1. **Closure:** $|V\rangle + |W\rangle \in \mathbb{V}$

2. **Distributive in the vector**: $a(|V\rangle + |W\rangle) = a|V\rangle + a|W\rangle$

3. **Distributive in the scalar**: $(a + b)|V\rangle = a|V\rangle + b|V\rangle$

4. **Associative in the scalar**: $a(b|V\rangle = ab|V\rangle)$

5. **Commutative in the addition**: $|V\rangle + |W\rangle + |W\rangle + |V\rangle$

6. **Associative in the addition**: As follows.

$$|V\rangle + (|W\rangle + |P\rangle) = (|V\rangle + |W\rangle) + |P\rangle$$

7. **Null vector**: $|V\rangle + |0\rangle + |V\rangle$

8. **Inverse under addition**: $|-V\rangle + |V\rangle = |0\rangle$

Vector space has n dimension if it can accommodate n linear independent vectors. We denote $\mathbb{V}^n(R)$ if the field–that is the scalar used to scale the vector–and $\mathbb{V}^n(C)$ if it is complex.

Dual space. Column vectors are concrete manifestations of an abstract vector $|V\rangle$ ket in a basis, while row vector are bra's $\langle V|$. They are adjoint of each other. Thus, there are two vector space: space of KET $|V\rangle$ and dual space of bra $\langle V|$.

Vector expansion in an orthonormal base. Suppose we are to expand vector $|V\rangle$ in an orthonormal base. First, we take the dot product of said vector with an orthonormal base $\langle j|$

$$\langle j|V\rangle = \sum_i v_i \langle j|i\rangle = v_j$$

and obtain the j-th component of the vector. Using this, we then write

$$|V\rangle = \sum_i |i\rangle \langle i|V\rangle$$

Lines and Plane

Suppose we have vector $\mathbf{A} = a\,\hat{\mathbf{x}} + b\,\hat{\mathbf{y}} + c\,\hat{\mathbf{z}}$ and vector $\mathbf{r} - \mathbf{r_0} = (x - x_0)\,\hat{\mathbf{x}} + (y - y_0)\,\hat{\mathbf{y}} + (z - z_0)\,\hat{\mathbf{z}}$, which parallel to $\mathbf{A}$. We can write:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

which is the symmetric equations of a straight line. Note that $\mathbf{r}$ and $\mathbf{r_0}$ is not necessarily parallel with $\mathbf{A}$, but $\mathbf{r} - \mathbf{r_0}$ do. The parameter equation is:

$$\mathbf{r} - \mathbf{r_0} = \mathbf{A}t$$
$$\mathbf{r} = \mathbf{r_0} + \mathbf{A}t$$

The previous equation is obtained by the dot identity of parallel vector

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

also called the equation of plane.

Matrix and Operation

Multiplication. Matrix AB can be multiplied if they are conformable, that is if column A = row B. Matrix multiplication in index notation is:

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}$$

where i denote row and j denote column. For 2×2 matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Commutator. In general, matrix do not commute. We define the commutator of the matrices A and B by

$$[A, B] = AB - BA$$

Two identity involving commutators are

$$[\Omega, \Lambda\theta] = \Lambda[\Omega, \theta] + [\Omega, \Lambda]\theta$$
$$[\Lambda\Omega, \Theta] = \Lambda[\Omega, \theta] + [\Lambda, \theta]\Omega$$

Inverse If a matrix has an inverse we say that it is invertible; if it doesn't have an inverse, it is called singular.

$$M^{-1} = \frac{1}{\det M} C^T$$

where C_{ij} is cofactor of m_{ij} or the checker thing you use on determining determinant. The inverse of a product follows

$$(\Lambda\Omega)^{-1} = \Lambda^{-1}\Omega^{-1}$$

By thin we can obtain the desired result

$$(\Omega\Lambda)(\Omega\Lambda)^{-1} = \Omega\Lambda\Lambda^{-1}\Omega^{-1} = \Omega\Omega^{-1} = I$$

Derivative (with respect to parameter). Consider the function $\theta(\lambda) = e^{\lambda\Omega}$. Its derivative with respect to parameter λ is evaluated in the usual sense

$$\frac{d}{d\lambda}\theta(\lambda) = \Omega e^{\lambda\Omega} = e^{\lambda\Omega}\Omega = \theta(\lambda)\Omega$$

The second and third terms holds true because $[\Omega, e^{\Omega}] = 0$. If we are presented with the differential equation with this form, the solution is given by

$$\theta(\lambda) = ce^{\Omega\lambda}$$

with c as operator constant of integration.

Proof. We can proof the evaluated method by writing the exponential function as an operator then applying the derivative operator

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \frac{d}{d\lambda}\begin{bmatrix} e^{\lambda\omega_1} & & \\ & \ddots & \\ & & e^{\lambda\omega_n} \end{bmatrix} = \begin{bmatrix} \omega_1 e^{\lambda\omega_1} & & \\ & \ddots & \\ & & \omega_n e^{\lambda\omega_n} \end{bmatrix}$$

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \begin{bmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{bmatrix}\begin{bmatrix} \lambda e^{\lambda\omega_1} & & \\ & \ddots & \\ & & \lambda e^{\lambda\omega_n} \end{bmatrix} = \Omega e^{\lambda\Omega}$$

If Ω is not a hermitian however, we can produce the same result by working using the power series representation

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \frac{d}{d\lambda}\sum_{m=0}^{\infty}\frac{\lambda^n\Omega^n}{n!} = \sum_{m=1}^{\infty}\frac{n\lambda^{n-1}\Omega^n}{n!} = \Omega\sum_{m=1}^{\infty}\frac{\lambda^{n-1}\Omega^{n-1}}{(n-1)!}$$

Shifting the index

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \Omega\sum_{m=0}^{\infty}\frac{\lambda^n\Omega^n}{n!} = \Omega e^{\lambda\Omega}$$

Determinant

For 2×2 matrix:

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Here are some determinant rule.

$$\det(kA) = k^2 \det A \qquad\qquad (2\times 2)$$
$$\det(kA) = k^3 \det A \qquad\qquad (3\times 3)$$
$$\det(AB) = \det(BA) = \det(A) \times \det(B)$$

Minor. Minor of element a_{ij} is the determinant of submatrix order $(n-1)$ you get after crossing i-th row and h-th column from order n matrix. Consider the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Each minor is then

$$M_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix},$$

$$M_{21} = \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{22} = \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{23} = \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix},$$

$$M_{31} = \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}, \quad M_{32} = \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}, \quad M_{33} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Cofactor Cofactor is expressed is the minor that includes sign factors

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

As a mnemonic, use the sign pattern to check the sign of the cofactor

$$(-1)^{i+j} \sim \begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Linear Algebra

Orthogonality. Matrix, in context of linear transformation, that preserve the length of vector is said to be orthogonal. Matrix M is orthogonal if

$$M^{-1} = M^T$$

with determinant

$$\det M = \pm 1$$

det M = 1 corresponds geometrically to a rotation, and det M = −1 means that a reflection is involved.

Transformation matrix. Matrix that rotate vector $\vec{r} = (x, y)$ into $\vec{R} = (X, Y)$ (in 2D)is

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and the one that rotate its axis instead

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

in 3D

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

$$R_y = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$R_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Homogeneous equations. The definition is as follows.

> Sets of linear equations when the constants on the right-hand sides are all zero are called homogeneous equations.

Homogeneous equations are never inconsistent; they always have the solution of zero–often called the trivial solution. If the number of independent equations–that is, the rank of the matrix–is the same as the number of unknowns, this is the only solution. If the rank of the matrix is less than the number of unknowns, there are infinitely many solutions.

Consider set of n homogeneous equations in n unknowns. These equations have only the trivial solution unless the rank of the matrix is less than n. This means that at least one row of the row reduced n by n matrix of the coefficients is a zero row. Which mean that the determinant D of the coefficients is zero. This fact will be used in eigenvalue problem.

> A system of n homogeneous equations in n unknowns has solutions other than the trivial solution if and only if the determinant of the coefficients is zero.

Special Matrices and Operator

Theorem.

1. $(ABC)^T = C^T B^T A^T$

2. $(ABC)^{-1} = C^{-1} B^{-1} A^{-1}$

3. $Tr(ABC) = Tr(BCA) = Tr(CAB)$. Trace is the sum of main diagonal. It is a theorem that the trace of a product of matrices is not changed by permuting them in cyclic order.

4. If H is a Hermitian matrix, then $U = e^{iH}$ is a unitary matrix.

Table of special matrices. Consider this.

Definition	Condition
Real	$A = \bar{A}$
Symmetric	$A = A^T$
Antisymmetric	$A = -A^T$
Orthogonal	$A^{-1} = A^T$
Pure Imaginary	$A = -\bar{A}$
Hermitian	$A = A^\dagger$
Antihermitian	$A = -A^\dagger$
Unitary	$A^{-1} = A^\dagger$
Normal	$AA^\dagger = A^\dagger A$

Identity matrix. A matrix, who also acts as an operator, which leave the operated vector unchanged

$$I \left| V \right\rangle = \left| V \right\rangle \quad \left\langle V \right| I = \left\langle V \right|$$

Its element may be written as

$$I_{ij} = \langle i|I|k \rangle = \langle i|j \rangle = \delta_{ij}$$

or using the Kronecker delta. Then $n \times n$, using the Kronecker delta, identity matrix may be written

$$I = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1n} \\ \vdots & \ddots & \vdots \\ \delta_{n1} & \cdots & \delta_{nn} \end{bmatrix}$$

Projection Operator The projection operator is defined as

$$\mathbb{P}_i = \left| i \right\rangle \left\langle i \right|$$

This operator can be used to write the expansion of vector

$$\left| V \right\rangle = \sum_i \left| i \right\rangle \left\langle i | V \right\rangle = \sum_i \mathbb{P}_i \left| V \right\rangle$$

Or the identity matrix, which is called the completeness relation

$$I = \sum_i \left| i \right\rangle \left\langle i \right| = \int_i \left| x \right\rangle \left\langle i \right| \; dx = \sum_i \mathbb{P}_i$$

Its action on bra is all the same

$$\left\langle V \right| \mathbb{P}_i = \left\langle V | i \right\rangle \left\langle i \right| = v_i^* \left\langle i \right|$$

Projection operator obey

$$\mathbb{P}_i \mathbb{P}_j = \delta_{ij} \mathbb{P}_j$$

Hermitian and anti Hermitian. An operator is called Hermitian fit satisfies $\Omega = \Omega^\dagger$, in other hand an operator is called anti-Hermitian if $\Omega = -\Omega^\dagger$. In the world of linear algebra, Hermitian and anti-Hermitian play the role of pure real and pure imaginary number. Just like how we can decompose every number into a sum of pure real and pure imaginary

$$\begin{aligned} \alpha &= \frac{\alpha + \alpha^*}{2} + \frac{\alpha - \alpha^*}{2} \\ &= \frac{(a + ib) + (a - ib)}{2} + \frac{(a + ib) - (a - ib)}{2} \\ \alpha &= a + ib \end{aligned}$$

we can decompose every operator into its Hermitian and anti Hermitian

$$\Omega = \frac{\Omega + \Omega^\dagger}{2} + \frac{\Omega - \Omega^\dagger}{2}$$

Unitary operator. Unitary operator is defined as

$$UU^\dagger = I$$

This definition also implies that U and its Hermitian conjugate are inverse of each. As a comparison with complex number, Unitary operator is a unit modulus $e^{i\theta}$, just as $e^{i\theta}e^{-i\theta} = 1$.

Inner product. This theorem state that unitary operator preserves the inner product between the vectors they act on. This can be proved by considering two vectors

$$|V_1\rangle = U\,|W_1\rangle \qquad |V_2\rangle = U\,|W_2\rangle$$

Then

$$\langle V_2|V_1\rangle = \langle W_2|\,U^\dagger U\,|W_1\rangle = \langle W_2|W_2\rangle$$

Unitary operator is the generalization of rotation operator from $\mathbb{V}^3(R)$ to $\mathbb{V}^n(C)$, for its preserves the inner product of the vector like the rotation matrix.

For $n \times n$ unitary operator, the column, or the row really, can be seen as the component of n vector, just like operator in general. These vector, then, is orthonormal to each other. The reason for this is that the operator preserve the inner product, the transformed set of vector is also orthonormal.

Dirac delta. Under the integral sign, Dirac delta is defined as

$$\int \delta(x - x')f(x')\,dx = f(x)$$

where it samples the value of function $f(x')$ at one point x. This expression also used to define continuous orthogonal basis

$$\langle x|x'\rangle = \delta(x - x')$$

Dirac delta is an even function proved by

$$\delta(x - x') = \langle x|x'\rangle = \langle x'|x\rangle^* = \delta(x' - x)^* = \delta(x' - x)$$

Dirac delta is normalized such that

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

Dirac delta is also defined by the limit of a Gaussian function

$$\delta(x - x') = \lim_{\epsilon \to 0} \frac{1}{\epsilon\sqrt{\pi}} \exp\left[-\frac{(x - x')^2}{\epsilon^2}\right]$$

The derivative is defined as following

$$\delta'(x - x') = \frac{d}{dx}\delta(x - x') = -\frac{d}{dx}\delta(x' - x)$$

The action of this operator is

$$\int \delta'(x - x')f(x')\,dx = \frac{d}{dx}f(x)$$

which can be proved by

$$\int \delta'(x - x')f(x') \, dx = \frac{d}{dx} \int \delta(x - x')f(x') \, dx = \frac{d}{dx}f(x)$$

In general, the action of n-th order of derivative is

$$\delta^{(n)}(x - x') = \delta(x - x')\frac{d^n}{dx^n}$$

The integral representation of Dirac delta is

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

This is obtained from Fourier transformation of given function

$$f(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) \, dx$$

and its inverse

$$f(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx'} f(k) \, dk$$

Substituting the transformation into the inverse

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

The term inside parenthesis is the Dirac delta function.

Differential operator. The action of this operator is described by the following equation

$$\langle x|D|f \rangle = \left\langle x \left| \frac{df}{dx} \right. \right\rangle = \frac{df(x)}{dx}$$

The operator also can act as the differential Dirac delta function under different basis

$$D_{xx'} = \langle x|D|x' \rangle = \delta'(x - x') = \delta(x - x')\frac{d}{dx'}$$

This can be seen by inserting the completion identity

$$\int \langle x|D|x' \rangle \langle x'|f \rangle \, dx' = \int D_{xx'} f(x') dx' = \frac{df(x)}{dx}$$

and comparing it with the action of the derivative of the Dirac delta

$$\int \delta'(x - x')f(x') \, dx = \frac{d}{dx}f(x)$$

The differential operator is not a Hermitian. For D to be a Hermitian, it must satisfy

$$D_{xx'} = D_{x'x}^\dagger \qquad \text{or} \qquad \langle x|D|x' \rangle = (\langle x'|D|x \rangle)^\dagger$$

The proof is as following

$$\langle x|D|x' \rangle = \delta'(x - x')$$

and

$$(\langle x'|D|x \rangle)^\dagger = \delta'(x' - x)^* = \delta'(x' - x) - \delta'(x - x')$$

Wave number operator. On the x basis, K acts as differential operator. This is misleading, actually. The correct term is, if K acts on a state $|\psi\rangle$, and then we project onto the position basis $|x\rangle$, the resulting expression is given by a differential operator acting on the projected state

$$\langle x|K|\psi\rangle = -i\frac{d}{dx}\psi(x)$$

On the other hand, K acts multiplicatively at the k basis

$$K|k\rangle = k|k\rangle$$

since $|k\rangle$ is an eigenstate of K.

Just like the differential operator, the wave number operator maybe expressed in terms of Dirac delta function

$$K_{xx'} = \langle x|K|x'\rangle = -i\delta'(x-x')$$

which is obtained the following expression with the definition of the Dirac delta function

$$\int \langle x|K|x'\rangle\langle x'|f\rangle \ dx' = \int K_{xx'}f(x')dx' = -i\frac{df(x)}{dx}$$

Let $|f\rangle$ and $|g\rangle$ be two kets in function space within $[a, b]$ Suppose we transform $|f\rangle$ by K and project it onto $\langle g|$. In this basis, K is a Hermitian if the surface term vanishes

$$-ig^*(x)f(x)\Big|_a^b = 0$$

The expression is obtained by considering the Hermitian requirement

$$\langle g|K|f\rangle = (\langle f|K|g\rangle)^\dagger$$

For K to be a Hermitian, it must also obey

$$\int_a^b\int_a^b \langle g|x\rangle\langle x|K|x'\rangle\langle x'|f\rangle \ dx \ dx' =$$

$$\left(\int_a^b\int_a^b \langle f|x\rangle\langle x|K|x'\rangle\langle x'|g\rangle \ dx \ dx'\right)^\dagger$$

Rewriting the left-hand side and using integration by part

$$\int_a^b\int_a^b g^*(x)K_{xx'}f(x') \ dx \ dx' = \int_a^b g^*(x)\left[-i\frac{d}{dx}f(x)\right]dx$$

$$= i\int_a^b f(x)\frac{d}{dx}g^*(x) \ dx - ig^*(x)f(x)\Big|_a^b$$

We are then equating it with the right-hand side, which we write as

$$\left[\int_a^b\int_a^b f^*(x)K_{xx'}f(x') \ dx \ dx'\right]^\dagger = \left[\int_a^b\int_a^b f^*(x)\left(-i\frac{d}{dx}g(x)\right) dx\right]^\dagger$$

$$= i\int_a^b f(x)\frac{d}{dx}g^*(x) \ dx$$

Clearly, both side are equal if the surface term is zero.

The basis of k has the form of

$$\psi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$$

where $\psi_k(x) \equiv \langle k|\psi \rangle$ position representation of the abstract momentum eigenbasis $|k\rangle$. This expression can be derived by considering the act of K on the k basis and project it on x basis

$$\langle x|K|k \rangle = k \langle x|k \rangle$$

$$\int \langle x|K|x' \rangle \langle x'|k \rangle \ dx' k = \int K_{xx'} k(x') \ dx' = \psi_k(x)$$

$$i \frac{d}{dx} \psi_k(x) = k \psi_k(x)$$

This differential equation is solved by simple integration

$$\int \frac{1}{\psi_k(x)} \ d\psi_k(x) = \int ik \ dx$$

$$\ln \psi_k(x) = ik + A$$

$$\psi_k(x) = A e^{ik}$$

For continuous function, normalize any function into Dirac delta function. So

$$\int_{-\infty}^{\infty} \psi_k^*(x) \psi_k'(x) \ dx = \delta(k - k')$$

The integral gives

$$\int_{-\infty}^{\infty} A^2 e^{i(k-k')x} \ dx = A^2 \pi \delta(k - k')$$

So the value of A that normalize $\psi_k(x)$ is

$$A = \frac{1}{2\pi} \qquad \text{so that} \qquad \langle x|k \rangle = \psi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$$

On using this basis, the Fourier transforms is just the passage from one basis $|x\rangle$ to another $|k\rangle$

$$f(k) = \langle k|f \rangle = \int_{-\infty}^{\infty} \langle k|x \rangle \langle x|f \rangle \ dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) \ dx$$

$$f(x) = \langle x|f \rangle = \int_{-\infty}^{\infty} \langle x|k \rangle \langle k|f \rangle \ dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} f(k) \ dk$$

Position operator. The action in x basis is described as follows

$$X |x\rangle = x |x\rangle$$

while is the k basis

$$\langle k|X|\psi \rangle = i \frac{d}{dk} \psi(k)$$

Inner Product

Finite dimension. The inner product is defined in such way to obey the following requirement.

1. **Skew-symmetry**: $\langle V|W \rangle = \langle W|V \rangle^*$

2. **Positive semidefinitness**: $\langle V|V \rangle \geq 0$. If $|V \rangle = 0\,|0 \rangle$, then $\langle V|V \rangle = |0 \rangle$

3. **Linearity in ket** $\langle V|(a\,|W \rangle + b\,|P \rangle)) = a\,\langle V|W \rangle + b\,\langle V|P \rangle$

In following this axiom, we arrive at the following definition for inner product

$$\langle V|W \rangle = \sum_{i,j} v_i^* w_i \langle i|j \rangle = \begin{bmatrix} v_1^* & \cdots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

If we think function as element of vector space, two functions can be said to be orthogonal on (a, b) if

$$\int_a^b A(x)B(x) \, dx = 0$$

Infinite dimension. The inner product of two continuous function along some interval is defined as

$$\langle f|g \rangle = \int_a^b f^*(x)g(x) \, dx$$

This definition comes after recalling the completeness relation for infinite space

$$\langle f|g \rangle = \langle f|\, I \,|g \rangle = \int_a^b \langle f|x \rangle \langle x|g \rangle \, dx = \int_a^b f^*(x)g(x) \, dx$$

Norm

Norm of a vector is defined as

$$|V| \equiv \sqrt{\langle V|V \rangle}$$

Two vectors with zero inner product are said to be orthogonal. If said vector has unit norm instead, it is referred as normalized. Set of basis with orthogonal and normalized condition are called orthonormal basis.

Continuous function. The norm of a continuous function is defined as

$$\int_a^b A^*(x)A(x) \, dx = N^2$$

We also said the function $N^{-1}A(x)$ to be normalized and has the norm of one.

Adjoint Operation

The following is a summary on how to perform adjoint in braket notation.

> Reverse the order of all factors and make the substitutions $\Omega \leftrightarrow \Omega^\dagger, |\rangle \leftrightarrow \langle|, a \leftrightarrow a^*$.

From this we also obtain the conjugate expressions

$$
\begin{aligned}
|aV\rangle = a\,|V\rangle &\quad\leftrightarrow\quad \langle aV| = \langle V|\,a^* \\
|\Omega V\rangle = \Omega\,|V\rangle &\quad\leftrightarrow\quad \langle \Omega V| = \langle V|\,\Omega^\dagger
\end{aligned}
$$

Bra and ket. Here is how the method in action. The following vector is an adjoint of each other

$$
|V\rangle = \sum_i v_i\,|i\rangle \qquad \langle V| = \sum_i \langle i|\,v_i^*
$$

We can also write it in terms of projection operation by recalling $v_i = \langle i|V\rangle$ and $v_i^* = \langle V|i\rangle$

$$
|V\rangle = \sum_i |i\rangle\,\langle i|V\rangle \qquad \langle V| = \sum_i \langle V|i\rangle\,\langle i|
$$

Operator. The matrix $\Omega^\dagger$–also called Hermitian adjoint–represent transpose conjugate of Ω

$$
\Omega^\dagger = (\Omega^*)^T = (\Omega^T)^*
$$

Our general rule of adjoint state

$$
(\Omega\Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger
$$

This can however be proved in another method. Consider $\langle \Omega\Lambda V|$. First treat $\Omega\Lambda$ as one operator

$$
\langle (\Omega\Lambda)V| = \langle V|\,(\Omega\Lambda)^\dagger
$$

Then, treat both as separate operator and pull them out of the bra one by one

$$
\langle \Omega\Lambda V| = \langle \Lambda V|\,\Omega^\dagger = \langle V|\,\Lambda^\dagger \Omega^\dagger
$$

Thus, we have

$$
(\Omega\Lambda)^\dagger = \Lambda^\dagger \Omega^\dagger
$$

Equation. Suppose we have the equation involving

$$
a_1\,|V_1\rangle = a_2\,|V_2\rangle + a_3\,|V_3\rangle\,\langle V_4|V_5\rangle + a_4\Omega\Lambda\,|V_6\rangle
$$

and we want to take its adjoint

$$
\langle V_1|\,a_1^* = \langle V_2|\,a_2^* + \langle V_4|V_5\rangle\,\langle V_3|\,a_3^* + \langle V_6|\,\Lambda^\dagger \Omega^\dagger a_4^*
$$

Gram-Schmidt Theorem

The Gram-Schmidt procedure is used to convert linearly independent basis $|I\rangle, \ldots, |N\rangle$ into an orthonormal one $|1\rangle, \ldots, |n\rangle$. We begin with the first basis vector and normalize it

$$|1\rangle = \frac{|I\rangle}{|I|} = \frac{|I\rangle}{\sqrt{\langle I|I\rangle}}$$

For our i-th basis, we create the projection along all $(i-1)$ vector

$$|i'\rangle = |I_i\rangle - \sum_{j=1}^{i-1} |j\rangle \langle j|I_i\rangle$$

and normalize it to obtain the orthonormal basis

$$|i\rangle = \frac{|i'\rangle}{|i'|}$$

Three basis. Let $|I\rangle, |II\rangle, |III\rangle$ be linearly independent basis. The first orthonormal vector is

$$|1\rangle = \frac{|I\rangle}{|I|}$$

For the second basis

$$|2'\rangle = |II\rangle - |1\rangle \langle 1|II\rangle$$

The second term is the projection of $|II\rangle$ along the first orthonormal basis. By subtracting $|II\rangle$ by this, the only thing that remains is the perpendicular partial. Then we normalize the vector

$$|2\rangle = \frac{|2'\rangle}{|2'|}$$

Same goes for the third basis. We construct

$$|3'\rangle = |II\rangle - |1\rangle \langle 1|III\rangle - |2\rangle \langle 2|III\rangle$$

then normalize it

$$|3\rangle = \frac{|3'\rangle}{|3'|}$$

As a sanity check, we see that the projection of $|1\rangle$ along itself is norm

$$\langle 1|1\rangle = \frac{I|I}{|I|^2} = 1$$

and that the $|2'\rangle$ or $|2\rangle$ along the first is orthogonal

$$\langle 1|2'\rangle = \langle 1|II\rangle - \langle 1|1\rangle \langle 1|II\rangle = 0$$

$$\langle 1|2\rangle = \frac{1|2'}{|I||2'|} = 0$$

Schwarz Inequality

Theorem that ensure the magnitude of inner product never exceed the product of vector magnitude

$$|\langle V|W\rangle| \leq |V||W|$$

Another related theorem is the triangle inequality, which state that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides

$$|V + W| \leq |V| + |W|$$

Schwarz inequality proof. First we define

$$|Z\rangle = |V\rangle - \frac{\langle W|V\rangle}{|W|^2}|W\rangle$$

Its inner product is

$$\langle P|P\rangle = \left\langle V - \frac{\langle W|V\rangle}{|W|^2}W \middle| V - \frac{\langle W|V\rangle}{|W|^2}W \right\rangle$$

$$= \langle V|V\rangle - \frac{\langle W|V\rangle\langle V|W\rangle}{|W|^2} - \frac{\langle V|W\rangle\langle W|V\rangle}{|W|^2}$$

$$- \frac{\langle V|W\rangle\langle W|V\rangle\langle W|W\rangle}{|W^4|}$$

$$= \langle V|V\rangle - \frac{\langle W|V\rangle\langle V|W\rangle}{|W|^2}$$

According to semidefinitness axiom

$$\langle V|V\rangle \geq \frac{\langle W|V\rangle\langle V|W\rangle}{|W|^2}$$

Multiply by $|W|^2$ and taking the square root to obtain

$$|V|^2|W|^2 \geq |\langle V|W\rangle|^2$$
$$|V||W| \geq |\langle V|W\rangle|$$

Subspace

Given a vector space $\mathbb{V}$, a subset of its elements that form a vector space among themselves is called a subspace. We will denote a particular subspace i of dimensionality n by $\mathbb{V}_i^n$.

Given two subspaces $\mathbb{V}_i^{n_i}$ and $\gtrsim_j^{m_j}$, we define their sum $\mathbb{V}_i^{n_i} \oplus \mathbb{V}_j^{m_j}$ as the set containing

1. All vector in the $\mathbb{V}_i^{n_i}$.

2. All vector in the $\mathbb{V}_j^{m_j}$.

3. All linear combination of them.

Vector along three spatial dimension is denoted as $\mathbb{V}_3(R)$, while its component is $\mathbb{V}_x^1$ on the x direction and the same convention on other two. All vector along the xy plane is denoted $\mathbb{V}_{xy}^2$. Adding the x and y subspace will result in the same subspace $\mathbb{V}_x^1 \oplus \mathbb{V}_y^1 = \mathbb{V}_{xy}^2$, which makes sense since the plane is a linear combination of two basis.

Linear Independence

Linear combination of $\mathbf{A}$ and $\mathbf{B}$ means $a\mathbf{A} + b\mathbf{B}$ where a and b are scalars.

Vector. The vector $\mathbf{r}r = x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + z\,\hat{\mathbf{z}}$ with tail at the origin is a linear combination of the unit basis vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ where a is a scalar. Set of vectors are said to independent if the only solution to the following equation is trivial

$$\sum_i a_i \left|i\right\rangle = \left|0\right\rangle$$

If the basis of the vector are independent, we can expand into

$$\left|V\right\rangle = \sum_i v_i \left|i\right\rangle$$

A function of a vector, say $f(r)$, is called linear if

$$f(r_1 + r_2) = f(r_1) + f(r_2) \qquad \text{and} \qquad f(ar) = a\,f(r)$$

Operator. Linear operator obey the following rule

$$\Omega\alpha\left|V\right\rangle = \alpha\Omega\left|V\right\rangle$$
$$\Omega\left[\alpha\left|V_1\right\rangle + \beta\left|V_2\right\rangle\right] = \Omega\alpha\left|V_1\right\rangle + \beta\Omega\left|V_2\right\rangle$$
$$\left\langle V\right|\alpha\Omega = \left\langle V\right|\Omega\alpha$$
$$\left[\left\langle V_1\right|\alpha + \left\langle V_2\right|\beta\right]\Omega = \alpha\left\langle V_1\right|\Omega + \beta\left\langle V_2\right|\Omega$$

Ω is a linear operator if

$$\Omega(r_1 + r_2) = \Omega(r_1) + \Omega(r_2) \qquad \text{and} \qquad \Omega(ar) = a\,\Omega(r)$$

Complete function. A set of function $f_n(x)$ is said to be complete if any other function $f(x)$ can be expressed as linear combination of them

$$f(y) = \sum_{n=1}^{\infty} C_n f(y)$$

Linear dependence f $f_1(x), f_2(x), \cdots, f_n(x)$ have derivatives of order $n-1$, and if the Determinant

$$W = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \cdots & f_n^{n-1}(x) \end{vmatrix} \not\equiv 0$$

then the functions are linearly independent.

Eigenvalue Problem

Consider linear operator Ω transforming non-trivial $|V\rangle$

$$\Omega |V\rangle = |W\rangle$$

Each operator has certain ket on which its transformation is simply recalling

$$\Omega |V\rangle = \omega |V\rangle$$

This is the eigenvalue equation which state that $|V\rangle$ is an eigenket of Ω with eigenvalue ω. We begin to solve the problem by writing the eigenvalue equation as

$$(\Omega - \omega I) |V\rangle = |0\rangle$$

Operating $(\Omega - \omega I)^{-1}$ on both side

$$|V\rangle = (\Omega - \omega I)^{-1} |0\rangle$$

Any operator acting on null vector can only give null vector, not arbitrary vector $|V\rangle$. Therefore, our assumption that operator $(\Omega - \omega I)^{-1}$ exist is false. Recalling the inverse of matrix M

$$M^{-1} = \frac{1}{\det M} \mathbf{C}^T$$

we see that the condition for non-existent inverse is zero determinant. Thus, the condition for non-zero eigenvector is

$$\det(\Omega - \omega I)^{-1} = 0$$

In practice, this equation is enough to determine the eigenvalue and eigenvector. For theoretical purpose however, we can determine what form does the solution take. To do so, we project the equation with basis bra $\langle i|$

$$\langle i| \Omega - \omega I |V\rangle = 0$$

or in summation form

$$\sum_j (\Omega_{ij} - \omega \delta_{ij}) v_j = 0$$

Setting the determinant to zero gives us the characteristic equation

$$\sum_{m=0}^{n} c_m \omega^m = 0$$

Once the eigenvalue ω are found, we can move to the next step of determining the eigenvector. We do this by substituting the value of ω into the equation

$$(\Omega - \omega I) |V\rangle = |0\rangle$$

Then we will obtain set of equations that the equation must obey. It is conventional to normalize the eigenvector also. In the case of degeneracy, we choose the eigenvector such that they are orthogonal to each other.

Propagator

Propagator is an operator that propagate a state forward in space or time. It can be used as a solution to a differential equation. For an equation with the form

$$|\ddot{x}\rangle = \Omega \,|x(t)\rangle$$

have the solution written as

$$|x(t)\rangle = U(t)\,|x(0)\rangle$$

The propagator is constructed as

$$U(t) = \sum_i |i\rangle\,\langle i|\cos\omega_i t$$

with $|i\rangle$ as the i-th eigenvector and ω_i as the i-th eigenvalue.

The steps of solving initial value problem are as follows.

1. Consider the equation $|\ddot{\psi}(t)\rangle = \Omega\,|\psi(0)\rangle$

2. Solve the eigenvalue of operator Ω

3. Find its eigenvector

4. Find the propagator $U(t)$ in terms of the eigenvectors and eigenvalues of Ω

$$U(t) = \sum_i |i\rangle\,\langle i|\cos\omega_i t$$

5. Write the solution

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle$$

Function of Operator

We shall consider the function, such as e^x, where x is an operator and determine its meaning. We restrict ourselves to those function that can be written as power series an operator that are hermitian.

Exponential function. Has the following form

$$e^\Omega = \begin{bmatrix} e_1^\omega & & \\ & \ddots & \\ & & e_n^\omega \end{bmatrix}$$

Proof. By going to the eigenbasis of hermitian operator and raising it to the power of m

$$\Omega = \begin{bmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{bmatrix}, \qquad \Omega^m = \begin{bmatrix} \omega_1^m & & \\ & \ddots & \\ & & \omega_n^m \end{bmatrix}$$

Inserting this into the series representation of exponential function

$$e^{\Omega} = \sum_{m=0}^{\infty} \frac{\Omega^m}{m!}$$

we have

$$e^{\Omega} = \sum_{m=0}^{\infty} \frac{1}{m!} \begin{bmatrix} \omega_1^m & & \\ & \ddots & \\ & & \omega_n^m \end{bmatrix} = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{\omega_1^m}{m!} & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} \frac{\omega_n^m}{m!} \end{bmatrix}$$

$$e^{\Omega} = \begin{bmatrix} e^{\omega_1} & & \\ & \ddots & \\ & & e^{\omega_n} \end{bmatrix}$$

Geometric series. For hermitian Ω

$$\sum_{m=0}^{\infty} \Omega^n = (1 - \Omega)^{-1}$$

Proof. We write the power series as

$$f(\Omega) = \sum_{m=0}^{\infty} \begin{bmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{bmatrix} = \begin{bmatrix} \sum_{m=0}^{\infty} \omega_1^n & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} \omega_n^m \end{bmatrix}$$

$$f(\Omega) = \begin{bmatrix} (1 - \omega_1)^{-1} & & \\ & \ddots & \\ & & (1 - \omega_n)^{-1} \end{bmatrix} = (1 - \Omega)^{-1}$$

Application: Gram-Schmidt Theorem

Convert these linearly independent basis into orthonormal basis

$$|I\rangle = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \quad |II\rangle = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \quad |III\rangle = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$$

First normalize

$$|1\rangle = \frac{|I\rangle}{\sqrt{\langle I|I\rangle}} = \frac{1}{3} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Then we construct

$$|2'\rangle = |II\rangle - |1\rangle \langle 1|II\rangle = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

and normalize

$$|2\rangle = \frac{|2'\rangle}{|2'|} = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Doing the something for the third base

$$|3'\rangle = |III\rangle - |1\rangle \langle 1|III\rangle - |2\rangle \langle 2|III\rangle$$

$$= \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 0 \\ 12/5 \\ 24/5 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/5 \\ 1/5 \end{bmatrix}$$

And normalize it

$$|3\rangle = \frac{|3'\rangle}{|3'|} = \sqrt{5} \begin{bmatrix} 0 \\ -2/5 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

Application: Determining determinant

Consider the matrix

$$A = \begin{bmatrix} 2 & -5 & 2 \\ 7 & 3 & 4 \\ 2 & 1 & 5 \end{bmatrix}$$

We use the elements of third column first

$$\det A = \begin{vmatrix} 2 & -5 & 2 \\ 7 & 3 & 4 \\ 2 & 1 & 5 \end{vmatrix} = 2 \begin{vmatrix} 7 & 3 \\ 2 & 1 \end{vmatrix} - 4 \begin{vmatrix} 2 & -5 \\ 2 & 1 \end{vmatrix} + 5 \begin{vmatrix} 2 & -5 \\ 7 & 3 \end{vmatrix}$$

$$= 2 \cdot 1 - 4 \cdot 11 + 5 \cdot 38 = 148$$

Then, as a check we use the first row's

$$\det A = 1 \begin{vmatrix} 3 & 4 \\ 1 & 5 \end{vmatrix} + 5 \begin{vmatrix} 7 & 4 \\ 2 & 5 \end{vmatrix} + 2 \begin{vmatrix} 7 & 3 \\ 2 & 1 \end{vmatrix}$$

$$= 11 + 135 + 2 =$$

Application: Inverse matrix

We use inverse matrix to solve the following equation

$$\begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & 0 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -10 \end{bmatrix}$$

Notice the equation has the following form

$$\Omega |V\rangle = \omega$$
$$|V\rangle = \Omega^{-1}\omega$$

To find vector $|V\rangle$ that satisfy the equation, we need to determine the inverse of Ω. The minor of each element are

$$M_{11} = 6, \quad M_{12} = -4, \quad M_{13} = 3,$$
$$M_{21} - 3 =, \quad M_{22} = 3, \quad M_{23} = -3,$$
$$M_{31} = 3, \quad M_{32} = -2, \quad M_{33} = 3$$

Thus, the cofactor is

$$C = \begin{bmatrix} 6 & 4 & 3 \\ 3 & 3 & 3 \\ 3 & 2 & 3 \end{bmatrix}$$

Next, using the first row to find the determinant

$$\det \Omega = 1 \cdot 6 + 1 \cdot (6 - 3) = 3$$

Finally the inverse is

$$\Omega^{-1} = \frac{1}{\det \Omega} C^T = \frac{1}{3} \begin{bmatrix} 6 & 3 & 3 \\ 4 & 3 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

Now we can use the inverse to find the value of the vector

$$|V\rangle = \frac{1}{3} \begin{bmatrix} 6 & 3 & 3 \\ 4 & 3 & 2 \\ 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ -10 \end{bmatrix} = \begin{bmatrix} 10 + 1 - 10 \\ \frac{20+3-20}{3} \\ 5 + 1 - 10 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$$

Application: Eigenvalue Problem

We shall find the eigenvalue of the hermitian matrix

$$\Omega = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

First write the eigenvalue equation

$$(\Omega - \omega I)\,|\omega\rangle = \begin{bmatrix} -\omega & 0 & 1 \\ 0 & -\omega & 0 \\ 1 & 0 & -\omega \end{bmatrix} |\omega\rangle = 0$$

The characteristic equation is

$$-\omega^3 + \omega = \omega(\omega^2 + 1) = 0$$

This implies the eigenvalues are

$$\omega = 0, \pm 1$$

Next, we substitute the eigenvalue into the eigenvalue equation. First consider the eigenvalue $\omega = 0$

$$(\Omega - \omega I)\,|\omega_0\rangle = 0 \implies \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_3 = 0 \\ 0 = 0 \\ v_1 = 0 \end{bmatrix}$$

And we get arbitrary v_2, to normalize the eigenvector we choose

$$\langle \omega_0 | = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

Next is the case of $\omega = 1$

$$(\Omega - \omega I)\,|\omega_1\rangle = 0 \implies \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -v_1 + v_3 = 0 \\ -v_2 = 0 \\ v_1 - v_3 = 0 \end{bmatrix}$$

the eigenvector corresponds to the eigenvalue is

$$|\omega_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

Finally the last eigenvalue $\omega = -1$

$$(\Omega - \omega I)\,|\omega_{-1}\rangle = 0 \implies \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + v_3 = 0 \\ v_2 = 0 \\ v_1 + v_3 = 0 \end{bmatrix}$$

with the eigenvector of

$$\langle \omega_{-1} | = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$$

We can also use these eigenvectors to diagonal the operator Ω. From the eigenvector, we construct the unitary matrix

$$U = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

As a check, we can also confirm the unitary identity of unitary matrix

$$U^\dagger U = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

As expected. We can now jump to the diagonalization

$$U^\dagger \Omega U = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$U^\dagger \Omega U = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Which is just the eigenvalue as the diagonal element.

Application: Eigenvalue Problem With Degeneracy

Now consider operator with matrix element

$$\Omega = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

The eigenvalue equation is

$$(\Omega - \omega I)\ket{\omega} = 0 \implies \begin{bmatrix} 1-\omega & 0 & 1 \\ 0 & 2-\omega & 0 \\ 1 & 0 & 1-\omega \end{bmatrix} \ket{\omega} = 0$$

while the characteristic equation

$$(2-\omega)[(1-\omega)^2 - 1] = (2-\omega)[\omega^2 - 2\omega] = (2-\omega)\omega(\omega - 2) = 0$$

which result in eigenvalues of

$$\omega = 0, 2, 2$$

with $\omega = 2$ as degenerate. First we consider the eigenvalue $\omega = 0$

$$(\Omega - \omega I)\ket{\omega_0} = 0 \implies \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 = 0 \\ 2v_2 = 0 \\ v_1 + v_3 = 0 \end{bmatrix}$$

with eigenvector of

$$\bra{\omega_0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$$

Next is the degenerate eigenvalue of $\omega = 2$

$$(\Omega - \omega I)\ket{\omega_2} = 0 \implies \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -v_1 + v_2 = 0 \\ 0 = 0 \\ -v_1 + v_3 = 0 \end{bmatrix}$$

The eigenvector, which are also degenerate, has the arbitrary component v_2. This mean that the two degenerate eigenvector lies on the

same plane. For the first degenerate eigenvector, let us just choose the simplest normalized vector that also satisfies the equation above

$$|\omega_2, \alpha\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

For the second degenerate eigenvector, we choose in a way such that it is orthonormal with the first. The orthogonal condition state that

$$\langle \omega_2, \alpha|_2^\circ, \beta\rangle = 0$$

Since the degenerate vectors lies in the same plane on arbitrary v_2, we shall determine the value of v_2 such that $|\omega_2, \alpha\rangle$ is orthogonal with $|\omega_2, \beta\rangle$

$$\langle \omega_2, \alpha|\omega_2', \beta\rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ \beta \\ 1 \end{bmatrix} = \beta + 2$$

or $\beta = -2$. All that left is normalizing it

$$|\omega_2, \beta\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Application: Solving Second Order Coupled ODE Using Propagator

---

Consider the case of two masses m that are coupled to each other and to the wall by spring force constant k. We denote x_1 and x_2 as the masses' displacement from the equilibrium. Doing Newtonian analysis to the first mass, we have the equation of motion

$$m\ddot{x}_1 = -kx_1 - k\xi_2$$
$$\ddot{x}_1 = -\frac{2k}{m}x_1 + \frac{k}{m}x_2$$

where $\xi_2 = x_1 - x_2$ is the displacement of the second spring. For the second mass,

$$m\ddot{x}_2 = -kx_2 - k\xi_1$$
$$\ddot{x}_2 = -\frac{2k}{m}x_2 + \frac{k}{m}x_1$$

where $\xi_1 = x_2 - x_1$. Writing these equation in matrix form

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} -2k/m & k/m \\ k/m & -2k/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Or

$$|\ddot{x}(t)\rangle = \Omega |x(t)\rangle$$

As per our usual steps, we first solve the eigenvalue problem of matrix Ω. We write the eigenvalue equation as

$$(\Omega + \omega^2 I) |x(t)\rangle = 0$$

The eigenvalue of Ω is written as $-\omega^2$ in anticipation of the fact that ω being real. Setting the determinant to zero

$$\det(\Omega + \omega^2 I) = \det \begin{vmatrix} -2k/m + \omega^2 & k/m \\ k/m & -2k/m + \omega^2 \end{vmatrix}$$

$$0 = \left(\omega^2 - \frac{2k}{m} \right)^2 = \left(\frac{k}{m} \right)^2$$

$$\omega^2 = \frac{2k}{m} \pm \frac{k}{m}$$

$$\omega = \pm \sqrt{\frac{2k}{m} \pm \frac{k}{m}}$$

Since we are taking the square of ω, both positive and negative value of ω produce the same eigenvalue. We're then taking the positive quantity only

$$\omega = \sqrt{\frac{3k}{m}}, \sqrt{\frac{k}{m}}$$

The equation governing eigenvector corresponding to the eigenvalue of $\sqrt{3k/m}$ is

$$(\Omega - \omega^2 I) |\omega_{\sqrt{3k/m}}\rangle = \begin{bmatrix} k/m & k/m \\ k/m & k/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{k}{m}(x_1 + x_2) \\ \frac{k}{m}(x_1 + x_2) \end{bmatrix}$$

The normalized eigenvector is

$$|\omega_{\sqrt{3k/m}}\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The equation for the second eigenvalue

$$(\Omega - \omega^2 I) |\omega_{\sqrt{k/m}}\rangle = \begin{bmatrix} -k/m & k/m \\ k/m & -k/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{k}{m}(-x_1 + x_2) \\ \frac{k}{m}(x_1 - x_2) \end{bmatrix}$$

With the eigenvector of

$$|\omega_{\sqrt{3k/m}}\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We now construct the propagator

$$U(t) = \sum_i |\omega_i\rangle \langle\omega_i| \cos \omega_i t$$

$$= |\omega_{\sqrt{3k/m}}\rangle \langle\omega_{\sqrt{3k/m}}| \cos \sqrt{\frac{3k}{m}} + |\omega_{\sqrt{k/m}}\rangle \langle\omega_{\sqrt{k/m}}| \cos \sqrt{\frac{k}{m}}$$

$$= \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \cos \sqrt{\frac{3k}{m}} t$$

$$= \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \cos \sqrt{\frac{k}{m}} t$$

$$= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \cos \sqrt{\frac{3k}{m}} t + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cos \sqrt{\frac{k}{m}} t$$

$$U(t) = \begin{bmatrix} \cos \sqrt{\frac{3k}{m}} + \cos \sqrt{\frac{3k}{m}} & \cos \sqrt{\frac{3k}{m}} - \cos \sqrt{\frac{3k}{m}} \\ \cos \sqrt{\frac{3k}{m}} - \cos \sqrt{\frac{3k}{m}} & \cos \sqrt{\frac{3k}{m}} + \cos \sqrt{\frac{3k}{m}} \end{bmatrix}$$

For given $|x(0)\rangle$ the solution of said ODE is

$$|x(t)\rangle = \begin{bmatrix} \cos\sqrt{\frac{3k}{m}}t + \cos\sqrt{\frac{3k}{m}}t & \cos\sqrt{\frac{3k}{m}}t - \cos\sqrt{\frac{3k}{m}}t \\ \cos\sqrt{\frac{3k}{m}}t - \cos\sqrt{\frac{3k}{m}}t & \cos\sqrt{\frac{3k}{m}}t + \cos\sqrt{\frac{3k}{m}}t \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$$

Application: Yet Another Propagator Problem

Consider a string of length L clamped at its two ends $x = 0$ and L. The displacement $\psi(x,t)$ obeys the wave equation

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial^2 \psi}{\partial x^2}$$

We recognize the differential operator as $-K^2$ which is Hermitian since $\psi(0) = \psi(L) = 0$.

Our general strategy is still the same as previously, that is we find the solution as a function of time $\psi(t)$. There is one more step, however, to find the solution as a function of both space and time: we project the solution $\psi(t)$ to x basis.

We move to the first step of solving differential equation using propagator by solving the eigenvalue problem of

$$|\psi\rangle = K^2 |\psi\rangle$$

If we project this in the x basis, the action reads

$$-\frac{\partial^2 \psi_k(x)}{\partial x^2} = k^2 \psi_k(x)$$

with $\psi_k(x) = \langle x|\psi_k\rangle$. The general solution to this differential equation is

$$\psi_k(x) = A\cos kx + B\sin kx$$

We have also the boundaries condition of

$$\psi_k(x) = \begin{cases} \psi_k(0) = 0 \\ \psi_k(L) = 0 \end{cases}$$

Applying the first condition return

$$A = 0$$

For the second condition, if we do not consider trivial solution of $B \neq 0$, we obtain the eigenvalue of

$$k = \frac{m\pi}{L}$$

with $k = [1, -\infty]$. We do not consider the zero and the negative value since they do not

Thus we obtain discrete set of eigenvector, or rather eigenfunction, with label m

$$\psi_m(x) = B\sin\left(\frac{m\pi}{L}x\right)$$

For discrete set of continuous, the normalization equation reads

$$\int_0^\lambda \psi_m^*(x)\psi_{m'}(x)\,dx = \delta_{mm'}$$

Substituting the expression for the solution

$$\int_0^L B^2 \sin\left(\frac{m\pi}{L}x\right)\sin\left(\frac{m'\pi}{L}x\right)\,dx = \frac{B^2 L}{m\pi}\int_0^{m\pi}\sin^2 u\,du$$

$$= \frac{B^2 L}{m\pi}\int_0^{m\pi}\frac{1-\cos u}{2}\,du$$

$$= \frac{B^2 L}{m\pi}\left[\frac{u}{2}\Big|_0^{m\pi} - \frac{\cos 2u}{4}\Big|_0^{2m\pi}\right]$$

$$\int_0^L B^2 \sin\left(\frac{m\pi}{L}x\right)\sin\left(\frac{m'\pi}{L}x\right)\,dx = \frac{B^2 L}{2}$$

which give the normalization constant of

$$B = \sqrt{\frac{2}{L}}$$

The normalized eigenfunction now reads

$$\psi_k(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{m\pi}{L}x\right)$$

Let us associate each solution by integer m an abstract $|m\rangle$, which on x basis reads as the said solution

$$\langle x|m\rangle = \sqrt{\frac{2}{L}}\sin\left(\frac{m\pi}{L}x\right)$$

Instead of writing the propagator, we'll just skip to writing the solution for the differential equation since we do not know the expression for the $|m\rangle$. As stated previously, we project the solution $\psi(t)$ into x basis to obtain

$$\psi(x,t) = \langle x|\psi(t)\rangle = \langle x|\,U(t)\,|\psi(0)\rangle$$

$$= \sum_{m=1}^\infty \langle x|m\rangle\,\langle m|\psi(0)\rangle \cos\omega_m t$$

Using the relation $\omega = kv$, we may the express the angular frequency as the spatial frequency. Although feeding $\omega_m = m\pi/L$ directly to the argument of the cosine term does not give a dimensionless quantity, it gives one if one consider that ω_m is equal to k_m times v which happens to be one in this case. Next, we insert the completeness relation

$$\psi(x,t) = \int_0^\lambda \sum_{m=1}^\infty \langle x|m\rangle\,\langle m|x'\rangle\,\langle x'|\psi(0)\rangle \cos\omega_m t$$

Finally, we obtain the solution

$$\psi(x,t) = \sum_{m=1}^\infty \frac{2}{L}\sin\left(\frac{m\pi}{L}x\right)\cos\left(\frac{m\pi}{L}t\right)\int_0^L \sin\left(\frac{m\pi}{L}x'\right)\psi(x',0)\,dx$$

Vector Analysis

Vector Operation

There are four vector operation: Addition, Multiplication by a scalar, Dot product, and Cross Product. (i) Addition of two vectors. Addition is commutative and associative

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$
$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

(ii) Multiplication by a scalar. Scalar multiplication is distributive.

$$a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B}$$

(iii) Dot product of two vectors. The dot product of two vectors is defined

$$\mathbf{A} \cdot \mathbf{B} \equiv AB \cos \theta$$

where θ is the angle they form. Note that dot product is commutative and distributive.

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$$
$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$

(iv) Cross product of two vectors. The cross product of two vectors is defined by

$$\mathbf{A} \times \mathbf{B} \equiv AB \sin \theta \, \hat{\mathbf{n}}$$

where $\hat{\mathbf{n}}$ is a unit vector pointing perpendicular to the plane of $\mathbf{A}$ and $\mathbf{B}$. The cross product is distributive, but not commutative.

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C})$$
$$(\mathbf{B} \times \mathbf{A}) = -(\mathbf{A} \times \mathbf{B})$$

Few rule for manipulating vector. (i): To add vectors, add like components

$$\mathbf{A} + \mathbf{B} = (A_x + B_x)\hat{\mathbf{x}} + (A_y + B_y)\hat{\mathbf{y}} + (A_z + B_z)\hat{\mathbf{z}}$$

(ii): To multiply by a scalar, multiply each component.

$$a\mathbf{A} = (aA_x)\hat{\mathbf{x}} + (aA_y)\hat{\mathbf{y}} + (aA_z)\hat{\mathbf{z}}$$

Rule (iii): To calculate the dot product, multiply like components, and add.

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z$$

Rule (iv): To calculate the cross product, form the determinant whose first row is unit vector, whose second row is A (in component form), and whose third row is B.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

Triple Product

(i) Scalar triple product.

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = B \cdot (\mathbf{C} \times A) = C \cdot (\mathbf{A} \times B)$$

They are cyclic and in component form

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}$$

(ii) Vector triple product.

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}).$$

The product is linear combination of vector in parentheses.

Separation Vector

Separation vector defined as vector from the source point $\bar{r}'$ to the field point $\bar{r}$

$$\boldsymbol{\imath} \equiv \bar{\mathbf{r}} - \mathbf{r}'.$$

Del Operator

Vector operator defined as follows.

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

Operation Involving Del Operator

There are three ways the operator ∇ can act:

1. On a scalar function T : ∇T (the gradient);

2. On a vector function $\mathbf{v}$, via the dot product: $\nabla \cdot \mathbf{v}$ (the divergence);
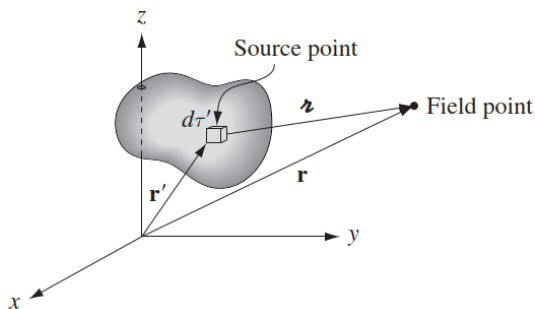

Figure: Separation vector

3. On a vector function $\mathbf{v}$, via the cross product: $\nabla \times \mathbf{v}$ (the curl).

Gradient of scalar function $T(x, y, z)$

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

can be used to define partial derivative of T

$$dT = \left(\hat{\mathbf{x}}\frac{\partial T}{\partial x} + \hat{\mathbf{y}}\frac{\partial T}{\partial y} + \hat{\mathbf{z}}\frac{\partial T}{\partial z} \right) \cdot (dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}})$$

$$= \nabla T \cdot \hat{\mathbf{u}}$$

Note that ∇T is a vector quantity, with three components. The gradient ∇T points in the direction of maximum increase of the function T. Moreover, The magnitude ∇T gives the slope (rate of increase) along this maximal direction.

Divergence of vector function $\mathbf{V}$ is

$$\nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

which is a scalar. Divergence is a measure of how much the vector V spreads out (diverges) from the point in question.

Curl of vector function $\mathbf{V}$ is

$$\nabla \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix}$$

The name curl is also well-chosen, for $\nabla \times \mathbf{v}$ is a measure of how much the vector $\mathbf{v}$ swirls around the point in question.

Product Rule

There are two ways to construct a scalar as the product of two functions

$$fg \quad \text{(product of two scalar functions)}$$
$$\mathbf{A} \cdot \mathbf{B} \quad \text{(dot product of two vector functions)}$$

and two ways to make a vector

$$f\mathbf{A} \quad \text{(scalar times vector)}$$
$$\mathbf{A} \times \mathbf{B} \quad \text{(cross product of two vectors)}$$

Accordingly, there are six product rule, two for gradients

$$\nabla(fg) = f\nabla g + g\nabla f$$

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A}$$

two for divergences

$$\nabla(f\mathbf{A}) = f(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\nabla f)$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$$

and two for curls

$$\nabla \times (f\mathbf{A}) = f(\nabla \times \mathbf{A}) - \mathbf{A} \times (\nabla f)$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A})$$

Second Derivative

(1) Divergence of gradient: $\nabla \cdot (\nabla T)$. Called Laplacian of T. Notice that the Laplacian of a scalar T is a scalar.

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

Occasionally, we shall speak of the Laplacian of a vector, $\nabla^2 \mathbf{v}$. By this we mean a vector quantity whose x-component is the Laplacian of V_x, and so on:

$$\nabla^2 \mathbf{v} \equiv (\nabla^2 V_x)\hat{\mathbf{x}} + (\nabla^2 V_y)\hat{y} + (\nabla^2 V_z)\hat{\mathbf{z}}$$

(2) The curl of a gradient: $\nabla \times (\nabla T)$. Always zero.

$$\nabla \cdot (\nabla T) = 0$$

(3) Gradient of divergence: $\nabla(\nabla \cdot \mathbf{v})$. $\nabla(\nabla \cdot \mathbf{v})$ is not the same as the Laplacian of a vector.

$$\nabla(\nabla \cdot \mathbf{v}) \neq \nabla^2 \mathbf{v} = (\nabla \cdot \nabla)\mathbf{v}$$

(4) The divergence of a curl: $\nabla \cdot (\nabla \times \mathbf{v})$. Always zero.

$$\nabla \cdot (\nabla \times \mathbf{v}) = 0$$

(5) Curl of curl:$\nabla \times (\nabla \times \mathbf{v})$. From the definition of ∇,

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v}$$

Fundamental Theorem of Calculus

The fundamental theorem of calculus says the integral of a derivative over some region is given by the value of the function at the end points (boundaries).

$$\int_a^b \frac{df}{dx}\, dx = f(b) - f(a)$$

Gradient. The fundamental theorem for gradients; like the "ordinary" fundamental theorem, it says that the integral (line integral) of a derivative (gradient) is given by the value of the function at the boundaries (a and b).

$$\int_a^b (\nabla T) \cdot d\mathbf{l} = T(\mathbf{b}) - T(\mathbf{a})$$

Corollary 1: $\int_a^b (\nabla T) \cdot d\mathbf{l}$ is independent of the path.
Corollary 2: $\oint (\nabla T) \cdot d\mathbf{l} = 0$ since the beginning and end points are identical.

Divergences . Like the other "fundamental theorems," it says that the integral of a derivative (divergence) over a region (volume V) is equal to the value of the function at the boundary (surface S).

$$\int_V (\nabla \cdot \mathbf{v}) \, d\tau = \oint_S \mathbf{v} \cdot d\mathbf{a}$$

If v represents the flow of an incompressible fluid, then the flux of **v** is the total amount of fluid passing out through the surface, per unit time. There are two ways we could determine how much is being produced: (a) we could count up all the faucets, recording how much each puts out, or (b) we could go around the boundary, measuring the flow at each point, and add it all up. Alternatively,

$$\int (\text{faucets within the volume}) = \oint (\text{flow out through the surface})$$

Curl. As always, the integral of a derivative (curl) over a region (patch of surface, S) is equal to the value of the function at the boundary (perimeter of the patch, P). Now, the integral of the curl over some surface (flux of the curl) represents the "total amount of swirl," and we can determine that just as well by going around the edge and finding how much the flow is following the boundary.

$$\int_S (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = \oint_P \mathbf{v} \cdot d\mathbf{l}$$

Corollary 1. $\int (\nabla \times \mathbf{v}) \cdot d\mathbf{a}$ depends only on the boundary line. It doesn't matter which way you go as long as you are consistent. For a closed surface (divergence theorem), $d\mathbf{a}$ points in the direction of the outward normal; but for an open surface is given by the right-hand rule: if your fingers point in the direction of the line integral, then your thumb fixes the direction of $d\mathbf{a}$.

Corollary 2. $\oint (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = 0$ for any closed surface, since the boundary line, like the mouth of a balloon, shrinks down to a point.

Integration by Parts

It applies to the situation in which you are called upon to integrate the product of one function (f) and the derivative of another (g); it says you can transfer the derivative from g to f, at the cost of a minus sign and a boundary term.

$$\int_a^b f \left(\frac{dg}{dx} \right) dx = - \int_a^b g \left(\frac{df}{dx} \right) dx + fg \Big|_a^b$$

Curvilinear Coordinates

I shall use arbitrary (orthogonal) curvilinear coordinates (u, v, w), developing formulas for the gradient, divergence, curl, and Laplacian in any such system. Infinitesimal displacement vector can be written

$$d\mathbf{l} = f \, du \, \hat{\mathbf{u}} + g \, dv \, \hat{\mathbf{v}} + h \, dw \, \hat{\mathbf{w}}$$

System	u	v	w	f	g	h
Cartesian	x	y	z	1	1	1
Spherical	r	θ	ϕ	1	r	$r\sin\theta$
Cylindrical	s	ϕ	z	1	s	1

where f, g, and h are functions of position characteristic of the particular coordinate system. While infinitesimal volume is

$$d\tau = fgh\, du\, dv\, dw$$

Use table 1.1 for references.

Gradient. The gradient of t is

$$\nabla \mathbf{t} \equiv \frac{1}{f}\frac{\partial t}{\partial u}\hat{\mathbf{u}} + \frac{1}{g}\frac{\partial t}{\partial v}\hat{\mathbf{v}} + \frac{1}{h}\frac{\partial t}{\partial w}\hat{\mathbf{w}}$$

Divergence. The divergence of $\mathbf{A}$ in curvilinear coordinates:

$$\nabla \cdot \mathbf{A} \equiv \frac{1}{fgh}\left[\frac{\partial}{\partial u}(ghA_u) + \frac{\partial}{\partial v}(fhA_v) + \frac{\partial}{\partial w}(fgA_w)\right]$$

Curl.

$$\nabla \times \mathbf{A} \equiv \frac{1}{gh}\left[\frac{\partial}{\partial v}(hA_w) - \frac{\partial}{\partial w}(gA_v)\right]\hat{\mathbf{u}} + \frac{1}{fh}\left[\frac{\partial}{\partial w}(fA_u) - \frac{\partial}{\partial u}(hA_w)\right]\hat{\mathbf{v}}$$
$$+ \frac{1}{fg}\left[\frac{\partial}{\partial u}(gA_v) - \frac{\partial}{\partial v}(fA_u)\right]\hat{\mathbf{w}}$$

Laplacian.

$$\nabla^2 t \equiv \frac{1}{fgh}\left[\frac{\partial}{\partial u}\left(\frac{gh}{f}\frac{\partial t}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{fh}{g}\frac{\partial t}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{fg}{h}\frac{\partial t}{\partial w}\right)\right]$$
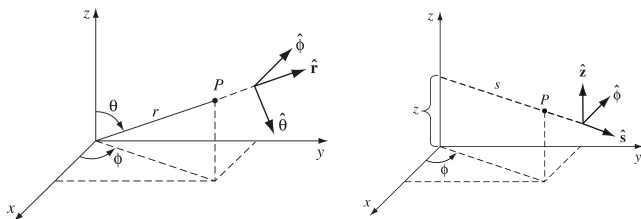
Spherical.

$$\begin{cases} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{cases} \qquad \begin{cases} \hat{\mathbf{x}} = \sin\theta\cos\phi\,\hat{\mathbf{r}} + \cos\theta\cos\phi\,\hat{\boldsymbol{\theta}} - \sin\phi\,\hat{\boldsymbol{\phi}} \\ \hat{\mathbf{y}} = \sin\theta\sin\phi\,\hat{\mathbf{r}} + \cos\theta\sin\phi\,\hat{\boldsymbol{\theta}} + \cos\phi\,\hat{\boldsymbol{\phi}} \\ \hat{\mathbf{z}} = \cos\theta\,\hat{\mathbf{r}} - \sin\theta\,\hat{\boldsymbol{\theta}} \end{cases}$$

$$\begin{cases} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arctan\sqrt{x^2 + y^2}/z \\ \phi &= \arctan y/z \end{cases} \qquad \begin{cases} \hat{\mathbf{r}} = \sin\theta\cos\phi\,\hat{\mathbf{x}} + \sin\theta\sin\phi\,\hat{\mathbf{y}} + \cos\theta\,\hat{\mathbf{z}} \\ \hat{\boldsymbol{\theta}} = \cos\theta\cos\phi\,\hat{\mathbf{x}} + \cos\theta\sin\phi\,\hat{\mathbf{y}} - \sin\theta\,\hat{\mathbf{z}} \\ \hat{\boldsymbol{\phi}} = -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}} \end{cases}$$

Cylindrical.

$$\begin{cases} x &= s\cos\phi \\ y &= s\sin\phi \\ z &= z \end{cases} \qquad \begin{cases} \hat{\mathbf{x}} = \cos\phi\,\hat{\mathbf{s}} - \sin\phi\,\hat{\boldsymbol{\phi}} \\ \hat{\mathbf{y}} = \sin\phi\,\hat{\mathbf{s}} + \cos\phi\,\hat{\boldsymbol{\phi}} \\ \hat{\mathbf{z}} = \hat{\mathbf{z}} \end{cases}$$

Spherical Coordinates and Cylindrical Coordinates

$$
\begin{cases}
s & = \sqrt{x^2 + y^2} \\
\phi & = \arctan y/z \\
z & = z
\end{cases}
\qquad
\begin{cases}
\hat{\mathbf{s}} = \cos\phi\,\hat{\mathbf{x}} + \sin\phi\,\hat{\mathbf{y}} \\
\hat{\boldsymbol{\phi}} = -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}} \\
\hat{\mathbf{z}} = \hat{\mathbf{z}}
\end{cases}
$$

Dirac Delta

The one-dimensional Dirac delta function, $\delta(x)$, can be pictured as an infinitely high, infinitesimally narrow "spike," with area 1. That is to say

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

with

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

It follows that

$$
f(x)\delta(x - a) = f(a)\delta(x - a)
$$

Since the product is zero anyway except at $x = a$, we may as well replace $f(x)$ by the value it assumes at the origin. In particular

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

It's best to think of the delta function as something that is always intended for use under an integral sign. In particular, two expressions involving delta functions (say, $D_1(x)$ and $D_2(x)$) are considered equal if

$$
\int_{-\infty}^{\infty} f(x)D_1(x)\, dx = \int_{-\infty}^{\infty} f(x)D_2(x)\, dx
$$

It is easy to generalize the delta function to three dimensions

$$
\delta^3(\mathbf{r}) = \delta(x)\delta(y)\delta(z)
$$

with $\mathbf{r} \equiv x\hat{x} + y\hat{y} + z\hat{z}$, and it's integral

$$
\int_{\text{all space}} \delta^3(\mathbf{r}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x)\delta(y)\delta(z)\, dxdydz = 1
$$

Generalizing Delta function, we get

$$
\int_{\text{all space}} f(\mathbf{r})\delta^3(\mathbf{r}\text{-}\mathbf{a})\, d\tau = f(\mathbf{a})
$$

Few Dirac delta function

$$\nabla \cdot \left(\frac{\hat{\imath}}{\imath^2} \right) = 4\pi\delta^3(\imath)$$

$$\nabla \left(\frac{1}{\imath} \right) = -\frac{\hat{\imath}}{\imath^2}$$

$$\nabla^2 \frac{1}{\imath} = -4\pi\delta^3(\imath)$$

Fourier Transform of a δ function. Using the definition of a Fourier transform, we write

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

and its inverse transform

$$\delta(x - a) = \int_{-\infty}^{\infty} g(\alpha)e^{i\alpha x} \, d\alpha = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\alpha(x-a)} \, d\alpha$$

The integral however does not converge. If we replace the limits by $-n, n$, we obtain a set of functions which are increasingly peaked around $x = a$ as n increases, but all have area 1.

Derivative of a δ function. Using repeated integrations by parts gives

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

Few formulas involving δ function. For step function

$$u(x - a) = \begin{cases} 1, x > a \\ 0, x < a \end{cases}$$

$$u'(x - a) = \delta(x - a)$$

It is easy to see how the derivative of step function is equal to delta function.

Helmholtz Theorem

Suppose we are told that the divergence of a vector function F(r) is a specified scalar function D(r):

$$\nabla \cdot \mathbf{F} = D$$

and the curl of F(r) is a specified vector function C(r):

$$\nabla \times \mathbf{F} = \mathbf{C}$$

For consistency, C must be divergenceless $\nabla \cdot \mathbf{C} = 0$. Helmholtz theorem state if the divergence $D(r)$ and the curl $C(r)$ of a vector function F(r) are specified, and if they both go to zero faster than $1/r^2$ as $r \to \infty$ and if F(r) goes to zero as $r \to \infty$, then F is given uniquely by

$$\mathbf{F} = -\nabla U + \nabla \times \mathbf{W}$$

Potential Theorem

Curl-less (or "irrotational") fields. The following conditions are equivalent (that is, F satisfies one if and only if it satisfies all the others):

- $\nabla \times \mathbf{F} = 0$ everywhere.

- $\int_a^b \mathbf{F} \cdot d\mathbf{l}$ is independent of path, for any given end points.

- $\oint \mathbf{F} \cdot d\mathbf{l} = 0$ for any closed loop.

- $\mathbf{F}$ is the gradient of some scalar function: $F = -\nabla V$.

Divergence-less (or "solenoidal") fields. The following conditions are equivalent:

- $\nabla \cdot \mathbf{F} = 0$ everywhere.

- $\int \mathbf{F} \cdot d\mathbf{a}$ is independent of surface, for any given boundary line.

- $\oint \mathbf{F} \cdot d\mathbf{a} = 0$ for any closed surface.

- $\mathbf{F}$ is the curl of some scalar function: $\mathbf{F} = -\nabla \mathbf{A}$.

State Function and Conservative Field

State function, such as internal energy U and entropy S, can be thought as conservative field. The condition that must be satisfied by conservative field $\mathbf{V}$ is

$$\nabla \times \mathbf{V} = 0$$

Suppose we actually evaluate the curl of vector function $\mathbf{V}(x, y, z)$, we get

$$\nabla \times \mathbf{V} = \begin{pmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ V_x & V_y & V_z \end{pmatrix}$$

$$\nabla \times \mathbf{V} = \hat{\mathbf{i}}\left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right) + \hat{\mathbf{j}}\left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right) + \hat{\mathbf{k}}\left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)$$

Since $\mathbf{V}$, as a conservative field, has curl of zero, those term inside parenthesis can be evaluated into

$$\frac{\partial V_z}{\partial y} = \frac{\partial V_y}{\partial z}, \quad \frac{\partial V_x}{\partial z} = \frac{\partial V_z}{\partial x}, \quad \frac{\partial V_y}{\partial x} = \frac{\partial V_x}{\partial y}$$

For state function $U(S, V, N)$, the equation reads

$$\frac{\partial U_N}{\partial V} = \frac{\partial U_V}{\partial N}, \quad \frac{\partial U_S}{\partial N} = \frac{\partial U_N}{\partial S}, \quad \frac{\partial U_V}{\partial S} = \frac{\partial U_S}{\partial V}$$

Of course you can't evaluate the curl of state function, but hear me out. What we consider is not the function U itself, but rather, the differential dU. Its total differential may be written as

$$dU(S, V, N) = \left.\frac{\partial U}{\partial S}\right|_{V,N} dS + \left.\frac{\partial U}{\partial V}\right|_{S,N} dV + \left.\frac{\partial U}{\partial N}\right|_{S,V} dN$$

Here, the differentials (dS, dT, dN) act like unit vector, thus we can pretend that dU is a vector field with components of

$$U_S = \left.\frac{\partial U}{\partial S}\right|_{V,N}, \quad U_V = \left.\frac{\partial U}{\partial V}\right|_{S,N}, \quad U_N = \left.\frac{\partial U}{\partial N}\right|_{S,V}$$

Therefore

$$\frac{\partial}{\partial V}\frac{\partial U}{\partial N} = \frac{\partial}{\partial N}\frac{\partial U}{\partial V}, \quad \frac{\partial}{\partial N}\frac{\partial U}{\partial S} = \frac{\partial}{\partial S}\frac{\partial U}{\partial N}, \quad \frac{\partial}{\partial S}\frac{\partial U}{\partial V} = \frac{\partial}{\partial V}\frac{\partial U}{\partial S}$$

This is what it means to be an exact differential.

Partial Derivative

Total Diferential

For a function $f = f(x, y, z, \dots)$, its total derivative is defined as

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz$$

Identity Involving Partial Derivative

The Jacobian of $[u(x, y), v(x, y)]$ with respect to (x, y) is defined by

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2mm] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$$

Here are some identity relating the Jacobian with partial derivative.

Unity. Unity as in one

$$\frac{\partial(u, v)}{\partial(x, y)} = 1$$

Proof. Trivial

$$\frac{\partial(x, y)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial x}{\partial x} & \dfrac{\partial x}{\partial y} \\[2mm] \dfrac{\partial y}{\partial x} & \dfrac{\partial y}{\partial y} \end{vmatrix} = \frac{\partial x}{\partial x}\frac{\partial y}{\partial y} - \frac{\partial x}{\partial y}\frac{\partial y}{\partial x} = 1 \quad \blacksquare$$

Change of order. It can be proved that change of order cost the minus sign

$$\frac{\partial(u, v)}{\partial(x, y)} = -\frac{\partial(v, u)}{\partial(x, y)} = -\frac{\partial(u, v)}{\partial(y, x)}$$

Proof. Those three terms literally have the same value when evaluated

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2mm] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}$$

$$-\frac{\partial(v, u)}{\partial(x, y)} = -\begin{vmatrix} \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \\[2mm] \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \end{vmatrix} = \frac{\partial v}{\partial y}\frac{\partial u}{\partial x} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}$$

$$\frac{\partial(u,v)}{\partial(y,x)} = - \begin{vmatrix} \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial x} \\[2mm] \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial x} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}$$

See? ∎

Jacobian. In terms of Jacobian, partial derivative of u with respect to x can be written as

$$\left.\frac{\partial u}{\partial x}\right|_y = \frac{\partial(u,y)}{\partial(x,y)}$$

Proof. Just evaluate the Jacobian

$$\frac{\partial(u,y)}{\partial(x,y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2mm] \dfrac{\partial y}{\partial x} & \dfrac{\partial y}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial y}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial u}{\partial x} \quad ∎$$

Chain rule for partial derivative. The expression is

$$\frac{\partial(u,y)}{\partial(x,y)} = \frac{\partial(u,y)}{\partial(w,z)}\frac{\partial(w,z)}{\partial(x,y)}$$

Proof. The total differential of u and v as function w and z read

$$du = \frac{\partial u}{\partial w}\,dw + \frac{\partial u}{\partial v}\,dz \quad \wedge \quad dv = \frac{\partial v}{\partial w}\,dw + \frac{\partial v}{\partial z}\,dz$$

We can therefore evaluate the Jacobian

$$\frac{\partial(u,y)}{\partial(x,y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2mm] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial u}{\partial w}\dfrac{\partial w}{\partial x} + \dfrac{\partial u}{\partial z}\dfrac{\partial z}{\partial x} & \dfrac{\partial u}{\partial w}\dfrac{\partial w}{\partial y} + \dfrac{\partial u}{\partial z}\dfrac{\partial z}{\partial y} \\[2mm] \dfrac{\partial v}{\partial w}\dfrac{\partial w}{\partial x} + \dfrac{\partial v}{\partial z}\dfrac{\partial z}{\partial x} & \dfrac{\partial v}{\partial w}\dfrac{\partial w}{\partial y} + \dfrac{\partial v}{\partial z}\dfrac{\partial z}{\partial y} \end{vmatrix}$$

$$= \left| \begin{pmatrix} \dfrac{\partial u}{\partial w} & \dfrac{\partial u}{\partial z} \\[2mm] \dfrac{\partial v}{\partial w} & \dfrac{\partial v}{\partial z} \end{pmatrix} \begin{pmatrix} \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} \\[2mm] \dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y} \end{pmatrix} \right| = \begin{vmatrix} \dfrac{\partial u}{\partial w} & \dfrac{\partial u}{\partial z} \\[2mm] \dfrac{\partial v}{\partial w} & \dfrac{\partial v}{\partial z} \end{vmatrix} \begin{vmatrix} \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} \\[2mm] \dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y} \end{vmatrix}$$

$$\frac{\partial(u,y)}{\partial(x,y)} = \frac{\partial(u,y)}{\partial(w,z)}\frac{\partial(w,z)}{\partial(x,y)} \quad ∎$$

The real chain rule. We have

$$\left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial z}{\partial x}\right|_y = 1$$

Proof. Trivial

$$1 = \frac{\partial(x,y)}{\partial(x,y)} = \frac{\partial(x,y)}{\partial(z,y)}\frac{\partial(z,y)}{\partial(x,y)} = \left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial z}{\partial x}\right|_y \quad ∎$$

Yet another chain rule... Even more chain rule...

$$\left.\frac{\partial x}{\partial y}\right|_w = \left.\frac{\partial x}{\partial z}\right|_w \left.\frac{\partial z}{\partial y}\right|_w$$

Proof. Trivial

$$\left.\frac{\partial x}{\partial y}\right|_w = \frac{\partial(x,w)}{\partial(y,w)} = \frac{\partial(x,w)}{\partial(z,w)}\frac{\partial(z,w)}{\partial(y,w)} = \left.\frac{\partial x}{\partial z}\right|_w \left.\frac{\partial z}{\partial y}\right|_w$$

Cyclic rule. This is chain rule all over again...

$$\left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial z}{\partial y}\right|_x \left.\frac{\partial y}{\partial x}\right|_z = -1$$

Proof. Trivial

$$1 = \frac{\partial(x,y)}{\partial(x,y)} = \frac{\partial(x,y)}{\partial(z,y)}\frac{\partial(z,y)}{\partial(z,x)}\frac{\partial(z,x)}{\partial(x,y)} = -\frac{\partial(x,y)}{\partial(z,y)}\frac{\partial(y,z)}{\partial(x,z)}\frac{\partial(z,x)}{\partial(y,x)}$$

$$= -\left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial y}{\partial x}\right|_z \left.\frac{\partial z}{\partial y}\right|_x \quad \blacksquare$$

Application in Thermodynamics

Here we will derive some useful intensive parameter used in thermodynamics. We assumed entropy function S has the form of

$$S = S(U, V, N_{i|r})$$

where N is number of chemical potential and $N_{i|r} \equiv N_1, \ldots N_r$. Therefore, its total differential is

$$dS = \left.\frac{\partial S}{\partial U}\right|_{V,N_{i|r}} dU + \left.\frac{\partial S}{\partial V}\right|_{U,N_{i|r}} dV + \sum_{j=1}^{r} \left.\frac{\partial S}{\partial N_j}\right|_{U,V,N_{i\neq r}} dN_j$$

We also assume the following quantities

$$T = \left.\frac{\partial U}{\partial S}\right|_{V,N_i} \quad ; P = -\left.\frac{\partial U}{\partial V}\right|_{S,N_i} \quad ; \mu_j = \left.\frac{\partial U}{\partial N}\right|_{S,V,N_{i\neq j}}$$

First identity. As follows.

$$\left.\frac{\partial S}{\partial U}\right|_{V,N_i} = \frac{1}{T}$$

Proof. We use chain rule with $x \to U, y \to V, z \to S$; while keeping all the N_i constant

$$\left.\frac{\partial U}{\partial S}\right|_{V,N_i} \left.\frac{\partial S}{\partial U}\right|_{V,N_i} = 1 \implies \left.\frac{\partial S}{\partial U}\right|_{V,N_i} = \left(\left.\frac{\partial U}{\partial S}\right|_{V,N_i}\right)^{-1}$$

Then, from the definition of temperature

$$\left.\frac{\partial S}{\partial U}\right|_{V,N_i} = \frac{1}{T} \quad \blacksquare$$

Second identity. The identity written as

$$\left.\frac{\partial S}{\partial V}\right|_{U,N_i} = \frac{P}{T}$$

Proof. We invoke cyclic rule with $x \to U, y \to V, z \to S$; while keeping all the N_i constant

$$1 = -\left.\frac{\partial U}{\partial S}\right|_{V,N_i} \left.\frac{\partial S}{\partial V}\right|_{U,N_i} \left.\frac{\partial V}{\partial U}\right|_{U,N_i}$$

Then, from the first identity and the definition of pressure

$$1 = T\left.\frac{\partial S}{\partial V}\right|_{U,N_i} \frac{1}{P} \implies \left.\frac{\partial S}{\partial V}\right|_{U,N_i} = \frac{P}{T} \quad \blacksquare$$

Third Identity. Expressed as

$$\left.\frac{\partial S}{\partial N_j}\right|_{U,N_{i\neq j}} = -\frac{P}{T}$$

Proof. We again invoke cyclic with $x \to U, y \to Nj, z \to S$; while keeping V and all N except N_i constant

$$1 = -\left.\frac{\partial U}{\partial S}\right|_{V,N_i} \left.\frac{\partial S}{\partial N_j}\right|_{U,N_{i\neq j}} \left.\frac{\partial N_j}{\partial U}\right|_{U,N_{i\neq j}}$$

Then, from the definition of temperature and chemical potential

$$1 = -T\left.\frac{\partial S}{\partial N_j}\right|_{U,N_{i\neq j}} \frac{1}{\mu_j} \implies \left.\frac{\partial S}{\partial N_j}\right|_{U,N_{i\neq j}} = -\frac{\mu_j}{T} \quad \blacksquare$$

Lagrange Multipliers

Let $f(x, y, z)$ be our function that we want to optimize and $\phi(x, y, z) =$ const be our constraint. We then set the total differential of $f(x, y, z)$ and $\phi(x, y, z)$ equal to zero

$$\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial y}dz = 0$$

$$\frac{\partial \phi}{\partial x}dx + \frac{\partial \phi}{\partial y}dy + \frac{\partial \phi}{\partial y}dz = 0$$

Next, we construct the function

$$F(x, y, z) = f(x, y, z) + \lambda\phi(x, y, z)$$

and set its total derivative to zero

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z}\right)dz = 0$$

It follows that, for any value of dx, dy, dz, we choose λ such that

$$\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 0, \quad \frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 0, \quad \frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 0$$

Putting it all together, to optimize $f(x, y, z)$ with constraint $\phi(x, y, z)$, we need to optimize $F(x, y, z)$, which obtained by solving three partial derivative equations and constraint equation $\phi(x, y, z) =$ const. The equations in question are

$$\frac{\partial F}{\partial x} = 0, \quad \frac{\partial F}{\partial y} = 0,$$

$$\frac{\partial F}{\partial z} = 0, \quad \phi = \text{const.}$$

Multiple constraint. If there are multiple constraints, say ϕ_1 and ϕ_2, we function F we construct instead is

$$F(x, y, z) = f(x, y, z) + \lambda_1 \phi_1(x, y, z) + \lambda_2 \phi_2(x, y, z)$$

As aside, the function that we want to optimize need not to a function of three variable x, y, z. The previous derivation can be justified for any number of variable. Of course, with more variable there are more variables.

Leibniz' rule for Integral

Differentiation under integral sign stated by Leibniz' rule

$$\frac{d}{dx} \int_{u(x)}^{v(x)} f(x, t) \, dt = \int_u^v \frac{\partial f}{\partial x} \, dt + f(x, v) \frac{dv}{dx} - f(x, u) \frac{du}{dx}$$

Proof. Suppose we want dI/dx where

$$I = \int_u^v f(t) \, dt$$

By the fundamental theorem of calculus

$$I = F(v) - F(u) = \mathcal{F}(v, u)$$

or I is a function of v and u. Finding dI/dx is then a partial differentiation problem. We can write

$$\frac{dI}{dx} = \frac{\partial I}{\partial v} \frac{dv}{dx} + \frac{\partial I}{\partial u} \frac{du}{dx}$$

By the fundamental theorem of calculus, we have

$$\frac{d}{dv} \int_a^v f(x) \, dt = \frac{d}{dv} \big[F(v) - F(a) \big] = f(v)$$

$$\frac{d}{dv} \int_u^b f(x) \, dt = \frac{d}{dv} \big[F(b) - F(u) \big] = -f(u)$$

where u and v are a function of x, while a and b are a constant. This is the case when we consider $\partial I / \partial v$ or $\partial I / \partial v$; the other variable is constant. Then

$$\frac{d}{dx} \int_u^v f(t) \, dt = f(v) \frac{dv}{dx} - f(u) \frac{du}{dx}$$

Under not too restrictive conditions,

$$\frac{d}{dx} \int_a^b f(x, t) \, dt = \int_a^b \frac{\partial f(x, t)}{\partial x} \, dt$$

where, as before, a and b are constant. In other words, we can differentiate under the integral sign. It is convenient to collect these formulas into one formula known as Leibniz' rule:

$$\frac{d}{dx} \int_{u(x)}^{v(x)} f(x,t) \, dt = \int_{u}^{v} \frac{\partial f}{\partial x} \, dt + f(x,v)\frac{dv}{dx} - f(x,u)\frac{du}{dx} \qquad \blacksquare$$

Appendix: Lagrange Multipliers

Single constraint. Consider this example.

> Determine the largest volume of parallelepiped–that is, a three-dimensional figure formed by six parallelograms– whose edges parallel with the x, y, z axis inside ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

The ellipsoid function above acts as constraints, that left is to determine the function that we want to optimize. This requires some clever thinking. We begin by defining point (x, y, z) be the corner of our parallelepiped. Now, this point is located in the first octant of our parallelepiped. The volume of this octant is

$$v = xyz$$

Since the parallelepiped's sides are parallel the axis, its total volume is

$$V = 8v$$

Hence, the volume of our parallelepiped is

$$V = 8xyz$$

This is the function that we want to maximize. We then construct the function

$$F(x, y, z) = 8xyz + \lambda \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right)$$

The partial derivatives of F read as

$$\frac{\partial F}{\partial x} = 8yz + \frac{2\lambda}{a^2}x, \quad \frac{\partial F}{\partial y} = 8xz + \frac{2\lambda}{b^2}y, \quad \frac{\partial F}{\partial z} = 8xy + \frac{2\lambda}{c^2}z$$

To find the maximum of F, we then must solve the partial derivative equations and constraint equation

$$8yz + \frac{2\lambda}{a^2}x = 0$$

$$8xz + \frac{2\lambda}{b^2}y = 0$$

$$8xy + \frac{2\lambda}{c^2}z = 0$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Multiplying the first equation by x, the second by y, the third by z and adding them all together, we get

$$24xyz + 2\lambda \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right) = 24xyz + 2\lambda = 0$$

Hence

$$\lambda = -12xyz$$

Substituting this into the partial derivative equation to obtain

$$8yz - \frac{24yz}{a^2}x^2 = 0 \implies x = \frac{\sqrt{3}}{3}a$$

$$8xz - \frac{24xz}{b^2}y^2 = 0 \implies y = \frac{\sqrt{3}}{3}b$$

$$8xy - \frac{24xy}{c^2}z^2 = 0 \implies z = \frac{\sqrt{3}}{3}c$$

Therefore, the maximum volume of said parallelepiped is

$$V = \frac{24\sqrt{3}}{27}abc$$

Two constraints. Here's an example.

> Given two equation $z^2 = x^2 + y^2$ and $x + 2z + 3 = 0$, find the shortest and longest distance from the origin and the intersection of those two equations.

Here we want to minimize $f = x^2 + y^2 + z^2$ as usual. We construct auxiliary function

$$F = x^2 + y^2 + z^2 + \lambda_1(z^2 - x^2 - y^2) + \lambda_2(x + 2z)$$

The partial differentials of F read

$$\frac{\partial F}{\partial x} = 2x - 2\lambda_1 x + \lambda_2,$$

$$\frac{\partial F}{\partial y} = 2y - 2\lambda_1 y,$$

$$\frac{\partial F}{\partial z} = 2z + 2\lambda_1 z + 2\lambda_2$$

Putting it all together, we have these equations

$$2x - 2\lambda_1 x + \lambda_2 = 0 \tag{1.1}$$
$$2y - 2\lambda_1 y = 0 \tag{1.2}$$
$$2z + 2\lambda_1 z + 2\lambda_2 = 0 \tag{1.3}$$
$$z^2 - x^2 - y^2 = 0 \tag{1.4}$$
$$x + 2z + 3 = 0 \tag{1.5}$$

By equation 1.2, we have two possible cases

$$2y - 2\lambda_1 y = y(1 - \lambda_1) = 0 \implies y = 0 \lor \lambda_1 = 1$$

First we consider $y = 0$. Equation 1.4 reads

$$z^2 = x^2 \implies z = \pm x$$

Then in the subcase $y = 0$, $z = x$; equation 1.5 evaluates into

$$3x + 3 = 0 \implies x = 3$$

In other hand, for subcase $y = 0$, $z = -x$; the same equation evaluates into

$$x = 3$$

Now we consider the case when $\lambda_1 = 1$. Equation 1.1 reduces into

$$\lambda_2 = 0$$

which means equation 1.5 turns into

$$4z = 0 \implies z = 0$$

and equation 1.5

$$x = -3$$

Using this result, equation 1.4 reads

$$y^2 = -9$$

which is impossible unless we are willing to take a complex value. Suppose we are willing, we have the $y = 3i$. Hence, we have three possibilities that the optimized points might take

$$\{\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}\} = \{(-1, 0, -1), (3, 0, -3), (-3, 3i, 0)\}$$

The distance from origin then evaluated by

$$d_1 = \sqrt{\mathbf{P_1} \cdot \mathbf{P_2}} = \sqrt{2}$$
$$d_2 = \sqrt{\mathbf{P_2} \cdot \mathbf{P_2}} = \sqrt{18}$$
$$d_3 = \sqrt{\mathbf{P_3} \cdot \overline{\mathbf{P_3}}} = \sqrt{18}$$

Hence the shortest distance is $d = \sqrt{2}$ and the longest is $d = \sqrt{18}$.

Variation Calculus

The Euler Equation

Any problem in the calculus of variations is solved by setting up the integral which is to be stationary, writing what the function F is, substituting it into the Euler equation

$$\frac{d}{dx}\frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0$$

and solving the resulting differential equation. When the function $F = F(r, \theta, \theta')$, the Euler's equation read

$$\frac{d}{dr}\frac{\partial F}{\partial \theta'} - \frac{\partial F}{\partial \theta} = 0$$

If $F = F(t, x, \dot{x})$

$$\frac{d}{dt}\frac{\partial F}{\partial x'} - \frac{\partial F}{\partial x} = 0$$

Notice that the first derivative in the Euler equation is with respect to the integration variable in the integral. The partial derivatives are with respect to the other variable and its derivative.

Proof. We will try to find the y which will make stationary the integral

$$I = \int_{x_1}^{x_2} F(x, y, y')\, dx$$

where F is a given function. Let $\eta(x)$ represent a function of x which is zero at x_1 and x_2, and has a continuous second derivative in the interval x_1 to x_2, but is otherwise completely arbitrary. We define the function $Y(x)$ by the equation

$$Y(x) = y(x) + \epsilon\eta(x)$$

where $y(x)$ is the desired extremal and ϵ is a parameter. Differentiating with respect to x, we get

$$Y(x) = y(x)' + \epsilon\eta'(x)$$

Then we have

$$I(\epsilon) = \int_{x_1}^{x_2} F(x, Y, Y')\, dx$$

Now I is a function of the parameter ϵ; when $\epsilon = 0$, $Y = y(x)$, the desired extremal. Our problem then is to make $I(\epsilon)$ take its minimum value when $\epsilon = 0$. In other words, we want

$$\left.\frac{dI}{d\epsilon}\right|_{\epsilon=0} = 0$$

Remembering that Y and Y' are functions of ϵ, and differentiating under the integral sign with respect to ϵ

$$\frac{dI}{d\epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial Y}\frac{dY}{d\epsilon} + \frac{\partial F}{\partial Y'}\frac{dY'}{d\epsilon}\right) dx$$

$$= \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial Y} \eta(x) + \frac{\partial F}{\partial Y'} \eta'(x) \right] dx$$

We want $dI/d\epsilon = 0$ at $\epsilon = 0$

$$\frac{dI}{d\epsilon}\bigg|_{\epsilon=0} = \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} \eta(x) + \frac{\partial F}{\partial y'} \eta'(x) \right] dx$$

Assuming that y'' is continuous, we can integrate the second term by parts

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y'} \eta'(x) \, dx = - \int_{x_1}^{x_2} \frac{d}{dx} \frac{\partial F}{\partial y'} \eta(x) \, dx + \frac{\partial F}{\partial y'} \eta(x) \bigg|_{x_1}^{x_2}$$

The first term is zero as before because $\eta(x)$ is zero at x_1 and x_2. Then we have

$$\frac{dI}{d\epsilon}\bigg|_{\epsilon=0} = \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right] \eta(x) \, dx$$

Since $\eta(x)$ is arbitrary, we must have

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = 0 \qquad \blacksquare$$

Notice carefully here that we are not saying that when an integral is zero, the integrand is also zero; this is not true. What we are saying is that the only way $\int f(x)\eta(x) \, dx$ can always be zero for every $\eta(x)$ is for $f(x)$ to be zero.

Several Variables

If there are n dependent variables in the original integral, there are n Euler-Lagrange equations. For instance, an integral of the form

$$S = \int_{u_1}^{u_2} f[x(u), y(u), x'(u), y'(u), u] \, du$$

with two dependent variables $[x(u)$ and $y(u)]$, is stationary with respect to variations of $x(u)$ and $y(u)$ if and only if these two functions satisfy the two equations

$$\frac{\partial f}{\partial x} = \frac{d}{du} \frac{\partial f}{\partial x'} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = \frac{d}{du} \frac{\partial f}{\partial y'}$$

Application: Shortest Between two points

Arbitrary path is given by

$$L = \int_1^2 \sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2} \sqrt{1 + y'^2}\, dx$$

We factor dx from the integrand in order to make the function we are optimizing not dependent on the y variable and make the evaluation using Euler-Lagrange equation easier

$$f(y, y', x) = \sqrt{1 + y'^2}$$

Then the Euler-Lagrange equation takes the form of

$$\frac{\partial f}{\partial y} = \frac{d}{dt}\frac{\partial f}{\partial y'}$$

$\partial f/\partial y = 0$ implies simply that $\partial f/\partial y'$ is a constant. Accordingly,

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}} = C$$
$$y'^2 = C^2(1 + y^2)$$
$$y'^2(1 - C^2) = C^2$$
$$y = \int \frac{C}{\sqrt{1 - C^2}}\, dx = \mathcal{C}x$$

which is the equation for straight line.

Application: Brachistochrone Given two points 1 and 2, with 1 higher above the ground, in what shape should we build a frictionless roller coaster track so that a car released from point 1 will reach point 2 in the shortest possible time?

The speed at which the coaster descend can be determined by the conservation energy principle

$$mgy = \frac{1}{2}mv^2 v \qquad\qquad = \sqrt{2gy}$$

Thus the time to travel between points

$$t = \int_{t_1}^{t_2} \frac{ds}{v} = \int_{t_1}^{t_2} \sqrt{\frac{dx^2 + dy^2}{2gh}}$$

Since v gives a function of y, we take it as independent variable for the same reason as previously

$$t = \frac{1}{\sqrt{2g}} \int_{t_1}^{t_2} \sqrt{\frac{1 + x'^2}{y}}\, dy$$

Ignoring the constant, the function we want to optimize is
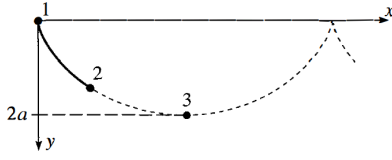
$$f(x, x', y) = \sqrt{\frac{1 + x'^2}{y}}$$

Figure: Brachistochrone problem

Then the Euler-Lagrange equation takes the form of

$$\frac{\partial f}{\partial x} = \frac{d}{dy}\frac{\partial f}{\partial x'}$$

$\partial f/\partial x = 0$ implies simply that $\partial f/\partial x'$ is a constant. Accordingly,

$$\frac{\partial f}{\partial x'} = \frac{x'}{\sqrt{y(1+x'^2)}} = C$$

Here we take the constant as $\sqrt{1/2a}$

$$x'^2 = \frac{y(1+x'^2)}{2a}$$

$$x'^2\left(1 - \frac{y}{2a}\right) = \frac{y}{2a}$$

$$x' = \sqrt{\frac{y}{2a}\frac{1}{2 - y/2a}}$$

$$x' = \sqrt{\frac{y}{2a - y}}$$

$$x = \int \sqrt{\frac{y}{2a - y}}\, dy$$

To solve this integral, we substitute $y = a(1-\cos\alpha)$ and $dy = a\sin\alpha\ d\theta$

$$x = \int \left[\frac{a(1 - \cos\theta)}{a(1 + \cos\theta)}\right]^{1/2} a\sin\theta\ s\theta$$

$$= a \int \left[\frac{1 - \cos\theta}{1 + \cos\theta}\right]^{1/2} [(1 + \cos\theta)(1 - \cos\theta)]^{1/2}\ d\theta$$

$$= a \int (1 - \cos\theta)\ d\theta$$

$$x = a(\theta - \sin\theta) + c$$

Therefore the path of the coaster is given by the following parametric equation

$$\begin{cases} x = a(\theta - \sin\theta) + c \\ y = a(1 - \cos\theta) \end{cases}$$

Classical Mechanics

Unsorted Classical Mechanics Topics

Newton's Law

First law. In the absence of an external force, when viewed from an inertial frame, an object at rest remains at rest and an object in uniform motion in a straight line maintains that motion.

Second law. Simply put

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

Third law. States that if two objects interact, the force exerted by object 1 on object 2 is equal in magnitude and opposite in direction to the force exerted by object 2 on object 1.

Particle Under Constant Acceleration

Here's some kinematics equation for position

$$x(t) = x_i + \frac{1}{2}(v_i + v_f)t$$

$$x(t) = x_i + v_i t + \frac{1}{2}at^2$$

and for velocity

$$v(t) = v_i + at$$

$$v(t)^2 = v_i^2 + 2a(x_f - x_i)$$

Particle in Uniform Circular Motion

If a particle moves in a circular path of radius r with a constant speed v, the magnitude of its centripetal acceleration is given by

$$a_r = \frac{v^2}{r}$$

while its period and angular velocity is

$$T = \frac{2\pi r}{v}, \quad \omega = \frac{2\pi}{T}$$

Applying Newton's second law

$$\sum F = ma_r = m\frac{v^2}{r}$$

Rigid Object Under Constant Angular Acceleration

Analogous to those for translational motion of a particle under constant acceleration

$$\omega(t) = \omega_i + \alpha t$$

$$\omega(t)^2 = \omega_i^2 + 2\alpha(\theta_t - \theta_i)$$

$$\theta(t) = \theta_i + \omega t + \frac{1}{2}\alpha t^2$$

$$\theta(t) = \theta_i + \frac{1}{2}(\omega_i + \omega_f)t$$

Relation of Linear and Rotational Motion

The following equations show the relation of linear and rotational motion

$$s = r\theta, \quad v = r\omega, \quad a_t = r\alpha$$

Torque

The torque associated with a force $\mathbf{F}$ acting on an object

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F} = I\alpha = \frac{d\mathbf{L}}{dt}$$

Moment of Inertia

The moment of inertia of a rigid object is

$$I = \sum mr^2 = \int r^2 dm$$

Parallel Axis Theorem. To calculate the moment inertia from any axis, we use parallel axis theorem

$$I = I_{\text{CM}} + Md^2$$

Terminal velocity

$\boldsymbol{r} \propto \boldsymbol{v}$. The velocity as a function of time is

$$v = \frac{mg}{b}\left[1 - \exp\left(-\frac{bt}{m}\right)\right] = v_T\left[1 - \exp\left(-\frac{bt}{m}\right)\right]$$

where b is a resistive constant whose value depends on the properties of the medium.

$r \propto v^2$. Given by

$$v_T = \sqrt{\frac{2mg}{D\rho A}}$$

where D is a dimensionless empirical quantity called the drag coefficient, ρ is the density of air, and A is the cross-sectional area of the moving object.

Escape velocity. The speed required by an object to escape from any planet orbit is

$$v_{\text{esc}} = \sqrt{\frac{2GM}{R}}$$

Work Energy Theorem

It states that if work is done on a system by external forces and the only change in the system is in its speed,

$$W = \Delta T$$

Kinetic Energy

For an object in linear motion, the kinetic energy of said object is

$$T = \frac{1}{2}mv^2$$

whereas for rotational motion

$$T = \frac{1}{2}I\omega^2$$

Hence the total kinetic energy of a rigid object rolling on a rough surface without slipping

$$T = \frac{1}{2}mv_{\text{CM}}^2 + \frac{1}{2}I\omega_{\text{CM}}^2$$

Potential Energy Function

For conservative energy $\mathbf{F}$, applies

$$V_f - V_i = -\int_{\mathbf{r_i}}^{\mathbf{r_f}} \mathbf{F} \cdot d\mathbf{r}$$

For particle-Earth system, the gravitational potential energy is

$$V = mgy$$

and elastic potential stored in spring

$$V = \frac{1}{2}kx^2$$

Effective potential

Effective potential energy $U_{\text{eff}}(r)$ is the sum of the actual potential energy $U(r)$ and the centrifugal $U_{\text{cf}}(r)$:

$$U_{\text{eff}}(r) = U(r) + \frac{l^2}{2\mu r^2}$$

where l is the angular momentum and μ is the reduced mass

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

Momentum Impulse

The linear momentum and impulse are defined as

$$\mathbf{p} = m\mathbf{v}, \quad \mathbf{I} = \int_{t_i}^{t_f} \sum \mathbf{F} \, dt$$

Angular Momentum The angular momentum about an axis through the origin of a particle having linear momentum

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

The z component of angular momentum of a rigid object rotating about a fixed z axis is

$$L_z = I\omega$$

Center of Mass and Velocity

The position vector of the center of mass of a system of particles is defined as

$$\mathbf{r}_{\text{CM}} = \frac{1}{M} \sum m\mathbf{r} = \frac{1}{M} \int \mathbf{r} \, dm$$

where M is the total mass. The velocity of the center of mass for a system of particles is

$$\mathbf{v}_{\text{CM}} = \frac{1}{M} \sum m\mathbf{v} =$$

Collision

Inelastic collision. One for which the total kinetic energy of the system of colliding particles is not conserved.

Elastic collision. One in which the kinetic energy of the system is conserved.

Perfectly inelastic. A collision which the colliding particles stick together after the collision.

Rocket Propulsion The expression for rocket propulsion is

$$v_f - v_i = v_e \ln \frac{M_i}{M_f}$$

Power

The rate at which work is done by an external force, called power, is

$$P = \frac{dE}{dt} = Fv = \tau\omega$$

Newton's Law on Gravity

$$\mathbf{F} = G\frac{m_1 m_2}{r^2}\hat{\mathbf{r}}$$

For an object at a distance h above the Earth's, the gravitational acceleration is

$$g = \frac{GM_E}{r^2} = \frac{GM_E}{(R_E + h)^2}$$

In general, the gravitational field experienced by mass m is

$$\mathbf{g} = \frac{\mathbf{F}}{m}$$

Kepler's Law

First Law. All planets move in elliptical orbits with the Sun at one focus.

Second Law The radius vector drawn from the Sun to a planet sweeps out equal areas in equal time intervals.

Third Law Simply put

$$T^2 = \frac{4\pi^2 a^3}{GM_S}$$

where a is semimajor axis and M_S is the mass of the sun.

Energy of Gravitational system

Potential energy. The gravitational potential energy associated with a system of two particles is
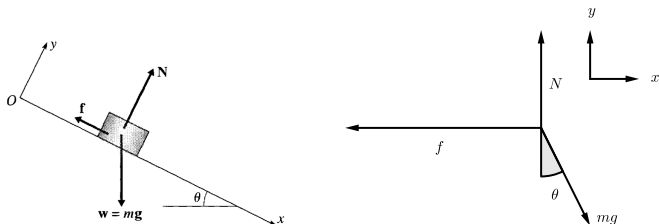
$$V = -\frac{Gm_1 m_2}{r}$$

Total energy. The total energy of the system is the sum of the kinetic and potential energies

$$E = \frac{1}{2}mv^2 - G\frac{Mm}{r} = -\frac{GMm}{2r}$$

Incline Problem

Consider:

A block of mass m is observed accelerating from rest down an incline that has coefficient of friction μ and is at angle θ from the horizontal. How far will it travel in time t?

Force approach. First we define the direction of displacement as positive x axis and the normal force as positive y axis. The resultant force in y axis written as

$$\sum F_y = N - mg\cos\theta = 0 \implies N = mg\cos\theta$$

and x axis

$$\sum F_x = mg\sin\theta - f = mg\sin\theta - \mu mg\cos\theta = m\ddot{x}$$

we then solve for x by

$$\ddot{x} = g\left(\sin\theta - \mu\cos\theta\right)$$
$$\dot{x} = g\left(\sin\theta - \mu\cos\theta\right)t$$
$$x(t) = \frac{1}{2}g\left(\sin\theta - \mu\cos\theta\right)t^2$$

Since the block stared from rest, its constant of integration is zero.

Energy approach. Here define the zero potential energy at the bottom of the incline. Using the work energy theorem for non-conservative force

$$\Delta T + \Delta U = W_{\text{fric}}$$
$$\frac{1}{2}mv^2 - mgh = -fd$$
$$\frac{1}{2}mv^2 - mgd\sin\theta = -\mu mg\cos\theta d$$
$$v = \sqrt{2gd(\sin\theta - \mu\cos\theta)}$$

Using the kinematics relation $v(t)^2 = v_i^2 + 2a\Delta x$

$$v^2 = 2gd(\sin\theta - \mu\cos\theta) = 2ad \implies a = g(\sin\theta - \mu\cos\theta)$$

and the relation $v(t) = v_i + at$, we are able to rewrite it as the previous form

$$v = g(\sin\theta - \mu\cos\theta)t$$

Central Force Problem

Consider also:

> A "half-pipe" at a skateboard park consists of a concrete trough with a semicircular cross section of radius $R = 5$m. I hold a frictionless skateboard on the side of the trough pointing down toward the bottom and release it. Find the equation of motion for this system.
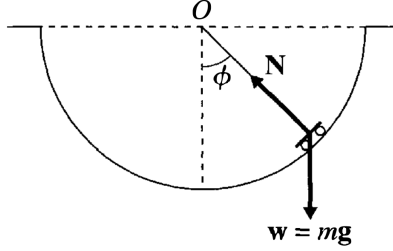
In this case r is held constant, thus the expression for resultant force in polar coordinate reads

$$\mathbf{F} = -m\dot{\phi}^2 R\,\hat{\mathbf{r}} + mR\ddot{\phi}\,\hat{\boldsymbol{\phi}}$$

We also know that the acting force in this system are the normal and the skateboard weigh. Applying this force into equation above

$$(mg\cos\phi - N)\,\hat{\mathbf{r}} - mg\sin\phi\,\hat{\boldsymbol{\phi}} = -m\dot{\phi}^2 R\,\hat{\mathbf{r}} + mR\ddot{\phi}\,\hat{\boldsymbol{\phi}}$$

We can't do anything with the radial component, we only use the angular component

$$mR\ddot{\phi} = mg\sin\phi$$

$$\ddot{\phi} = \frac{g}{R}\sin\phi$$

This differential equation is solved by

$$\phi(t) = A\sin\sqrt{\frac{g}{R}}t + B\cos\sqrt{\frac{g}{R}}t$$

Since this is released from rest, we have the initial condition of $\phi(0) = \phi_0$ and $\dot{\phi}(0) = 0$. Apllying the first condition

$$\phi_0 = B$$

and the second

$$\dot{\phi}(t) = A\sqrt{\frac{g}{R}}\cos\sqrt{\frac{g}{R}}t - \phi_0\sqrt{\frac{g}{R}}\sin\sqrt{\frac{g}{R}}t$$

$$\dot{\phi}(0) = 0 = A\sqrt{\frac{g}{R}}$$

Hence the equation of motion reads

$$\phi(t) = \phi_0\cos\sqrt{\frac{g}{R}}t$$

Central Force

Newton's Second Law in Polar Coordinate

Acceleration in polar coordinate expressed as

$$\ddot{\mathbf{r}} = \left(\ddot{r} - r\dot{\phi}^2\right)\,\hat{\mathbf{r}} + \left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right)\,\hat{\boldsymbol{\phi}}$$
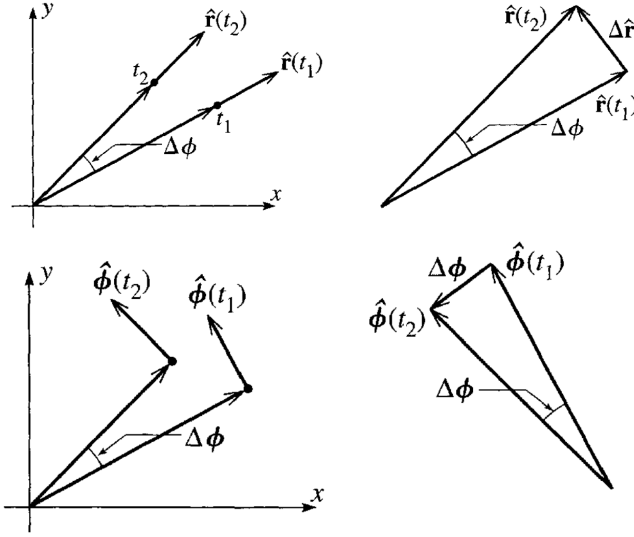
and velocity as

$$\mathbf{v} = \dot{r}\,\hat{\mathbf{r}} + r\dot{\phi}\,\hat{\boldsymbol{\phi}}$$

Hence Newton's law transform into

$$\mathbf{F} = m\mathbf{a} = \begin{cases} F_r & = m\left(\ddot{r} - r\dot{\phi}^2\right) \\ F_\phi & = m\left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right) \end{cases}$$

Derivation

The value of $d\hat{\mathbf{r}}$ and $d\hat{\boldsymbol{\phi}}$.

From the figure, we have

$$d\hat{\mathbf{r}} = d\phi\,\hat{\boldsymbol{\phi}}, \quad d\hat{\boldsymbol{\phi}} = -d\phi\,\hat{\mathbf{r}}$$

or equivalently

$$\frac{d\hat{\mathbf{r}}}{dt} = \dot{\phi}\,\hat{\boldsymbol{\phi}}, \quad \frac{d\hat{\boldsymbol{\phi}}}{dt} = -\dot{\phi}\,\hat{\mathbf{r}}$$

Using these we can now proceed to derive the Newton's law in polar coordinate. In cartesian coordinate, position vector can be writen as

$$\mathbf{r} = x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}}$$

converting it into polar

$$\mathbf{r} = r\,\hat{\mathbf{r}}$$

Next, we determine the velocity as

$$\dot{\mathbf{r}} = \frac{d}{dt}r\,\hat{\mathbf{r}} = \dot{r}\,\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt} = \dot{r}\,\hat{\mathbf{r}} + r\dot{\phi}\,\hat{\boldsymbol{\phi}}$$

and acceleration as

$$\ddot{\mathbf{r}} = \frac{d}{dt}\left(\dot{r}\,\hat{\mathbf{r}} + r\dot{\phi}\,\hat{\boldsymbol{\phi}}\right) = \ddot{r}\,\hat{\mathbf{r}} + \dot{r}\dot{\phi}\,\hat{\boldsymbol{\phi}} + r\frac{d}{dt}\left(\dot{\phi}\,\hat{\boldsymbol{\phi}}\right)$$

$$= \ddot{r}\hat{\mathbf{r}} + 2\dot{r}\dot{\phi}\,\hat{\boldsymbol{\phi}} + r\left(\ddot{\phi}\,\hat{\boldsymbol{\phi}} - \dot{\phi}\hat{\mathbf{r}}\right)$$

$$= \left(\ddot{r} - r\dot{\phi}^2\right)\hat{\mathbf{r}} + \left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right)\hat{\boldsymbol{\phi}}$$

Finally

$$F = F_r\,\hat{\mathbf{r}} + F_\phi\,\hat{\boldsymbol{\phi}} \begin{cases} F_r & = m\left(\ddot{r} - r\dot{\phi}^2\right) \\ F_\phi & = m\left(r\ddot{\phi} + 2\dot{r}\dot{\phi}\right) \end{cases}$$

Energy

Potential Energy

Gravitational. The potential energy is defined

$$U = mgh$$

where the reference point is chosen to be the ground. This makes it so that

$$F = -\nabla U = -\frac{d}{dr} mgr \, \hat{\mathbf{r}} = -mg \, \hat{\mathbf{r}}$$

the gravitational force is negative, or point downward. Now if we define the downward as positive displacement, the potential energy reads

$$U = -mgh$$

and the gravitational force

$$F = -\nabla U = -\frac{d}{dr} \left(-mgr \, \hat{\mathbf{r}} \right) = mg \, \hat{\mathbf{r}}$$

is positive, or point downward all the same.

Lagrangian Mechanics

The Lagrangian is defined as

$$\mathcal{L} = T - U$$

which mean that the Lagrangian is a function of position and velocity. The path of particle is determined by Hamilton's principle

$$S = \int_{t_1}^{t_2} \mathcal{L}\, dt$$

that is, the particle's path is such that the action integral S is stationary.

We can express the coordinate in other generalized coordinate

$$\mathbf{r} = \mathbf{r}(q_1, q_2, q_2)$$

And the same for velocity

$$\mathbf{v} = \mathbf{v}(\dot{q}_1, \dot{q}_2, \dot{q}_3)$$

Now the action integral reads

$$S = \int_{t_1}^{t_2} \mathcal{L}(q_1, q_2, q_3, \dot{q}_1, \dot{q}_2, \dot{q}_3, t)\, dt$$

Therefore we have three Euler-Lagrange equation that must be satisfied by the particle

$$\frac{\partial \mathcal{L}}{\partial q_1} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_1}, \quad \frac{\partial \mathcal{L}}{\partial q_2} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_2}, \quad \frac{\partial \mathcal{L}}{\partial q_3} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_3}$$

This is the case of one unconstrained particle. For N unconstrained particle, then, we shall have $3N$ Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \qquad i = 1, \ldots, 3N$$

The Lagrangian equation

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

takes the form

Generalized force = Rate of change of Generalized momentum

where

$$F_i = \frac{\partial \mathcal{L}}{\partial q_i} = i\text{-th component of the Generalized force}$$

and

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = i\text{-th component of the Generalized momentum}$$

Set of generalized coordinates that minimize the number of Euler-Lagrangian equation and be able to uniquely describe said system is said to be natural. This implies that natural coordinate also have minimum degree of freedom, that is the number of coordinate that can be varied independently. In natural coordinate, the generalized coordinate has no time dependence with its Cartesian relative.

Consider the Lagrangian

$$\mathcal{L} = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$$

If the Lagrangian does not depend on the coordinate q_i, the coordinate is said to be cyclic. Correspondingly, the canonical momentum $p_i = \partial\mathcal{L}/\partial q_i$ is conserved. The condition for q_i to be cyclic can be written as

$$\frac{\partial\mathcal{L}}{\partial q_i} = 0$$

System with n degree of freedom that can be described by n generalized coordinate is called holonomic. In general

$$\text{DoF} = \text{No. of Coordinate} - \text{No. Constraint}$$

For example, double pendulum have 4 coordinates and two constraint, thus having two degree of freedom.

The steps to solve problem using Lagrangian formalism are as follows.

1. Write down the Lagrangian $\mathcal{L} = T - U$.

2. Choose generalized n coordinate q_n and $\dot{q}_n$.

3. Rewrite $\mathcal{L}$ in terms of q_n and $\dot{q}_n$.

4. Write n Lagrange equation.

Proof of Lagrange Equation with Constraint

Suppose a particle has two degree of freedom with two kinds of forces act on it: constraint force $\mathbf{F}_{\text{cstr}}$, say interatomic forces that bind rigid body atom together; and the is non constraint conservative forces $\mathbf{F}$, which at minimum must be able to be derived from potential energy $F = -\nabla U(\mathbf{r}, t)$, say gravitational force. The total energy is then

$$\mathbf{F}_{\text{tot}} = \mathbf{F}_{\text{cstr}} + \mathbf{F}$$

and the Lagrangian

$$\mathcal{L} = T - U$$

where U is the potential energy which can be derived into non constraint conservative force.

The path of the particle can be denoted as

$$\mathbf{R}(t) = \mathbf{r}(t) + \boldsymbol{\epsilon}(t)$$

with $\mathbf{r}(t)$ as the correct path and $\boldsymbol{\epsilon}(t)$ as infinitesimal vector pointing away from the correct path. Then we have two Lagrangians

$$\mathcal{L} = \mathcal{L}(\mathbf{R}, \dot{\mathbf{R}}, t), \qquad \mathcal{L}_0 = \mathcal{L}_0(\mathbf{r}, \dot{\mathbf{r}}, t)$$

and two action integral

$$S = \int_{t_1}^{t_2} \mathcal{L} \, dt, \qquad S_0 = \int_{t_1}^{t_2} \mathcal{L}_0 \, dt$$

It can be proven that the difference in action integral $\delta S = S - S_0$ is zero, to the first order.

We write the difference in Lagrangian as

$$\delta \mathcal{L} = \frac{1}{2} m \dot{\mathbf{R}}^2 - U(\mathbf{R}, t) - \frac{1}{2} m \dot{\mathbf{r}}^2 + U(\mathbf{r}, t)$$

$$= \frac{1}{2} m \left[(\dot{\mathbf{r}}^2 + \dot{\boldsymbol{\epsilon}}^2) - \dot{\mathbf{r}}^2 \right] - [U(\mathbf{r} + \boldsymbol{\epsilon}, t) - U(\mathbf{r}, t)]$$

$$= \frac{1}{2} m \left[\dot{\mathbf{r}}^2 + \dot{\boldsymbol{\epsilon}}^2 + 2 \dot{\mathbf{r}} \cdot \dot{\boldsymbol{\epsilon}}^2 - \dot{\mathbf{r}}^2 \right] - dU$$

$$\delta \mathcal{L} = \frac{1}{2} m \dot{\boldsymbol{\epsilon}}^2 + m \dot{\mathbf{r}} \cdot \dot{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon} \cdot \nabla U$$

The difference in action integral, in the first order, is then

$$S = \int_{t_1}^{t_2} \mathcal{L}_0 \, dt = \int_{t_1}^{t_2} \left(m \dot{\mathbf{r}} \cdot \dot{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon} \cdot \nabla U \right) \, dt$$

Using integration by parts

$$\int_a^b f \left(\frac{dg}{dx} \right) dx = - \int_a^b g \left(\frac{df}{dx} \right) dx + fg \Big|_a^b$$

on the first term

$$\delta S = - \int_{t_1}^{t_2} \boldsymbol{\epsilon} \cdot [m \ddot{\mathbf{r}} + \nabla U] \, dt + m \dot{\mathbf{r}} \boldsymbol{\epsilon} \Big|_{t_1}^{t^2}$$

The difference of $\boldsymbol{\epsilon}$ is zero between two end point, so

$$\delta S = - \int_{t_1}^{t_2} \boldsymbol{\epsilon} \cdot [m \ddot{\mathbf{r}} + \nabla U] \, dt$$

The path $\mathbf{r}(t)$ satisfies Newton second law, thus $m\ddot{\mathbf{r}} = \mathbf{F}_{\text{tot}}$. Meanwhile, the gradient of potential energy is the negative of non constraint force

$$\delta S = - \int_{t_1}^{t_2} \boldsymbol{\epsilon} \cdot [\mathbf{F}_{\text{tot}} - \mathbf{F}] \, dt = i \int_{t_1}^{t_2} \boldsymbol{\epsilon} \cdot \mathbf{F}_{\text{cstr}} \, dt$$

Note that the constraint force is normal to the particle path and the $\boldsymbol{\epsilon}$, which lies to the same surface of particle path. Therefore, their dot product is zero

$$\delta S = 0$$

and the action integral is stationary.

This justifies the Lagrange equation for system with two degree of freedom where its constraint lie in the same surface as the particle path. In other words, it only applies to particle, or particles in that case, constrained to move in two dimension. Accordingly, the action integral in this case is written as

$$S = \int_{t_1}^{t_2} \mathcal{L}(q_1, q_2, \dot{q}_1, \dot{q}_2, t) \, dt$$

which will result in two Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial q_1} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_1}, \qquad \frac{\partial \mathcal{L}}{\partial q_2} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_2}$$

Newton Law in 2D Cartesian

The Lagrangian in this case is

$$\mathcal{L}(x, y, \dot{x}, \dot{y}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - U(x, y)$$

Here we have two Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}, \qquad \frac{\partial \mathcal{L}}{\partial y} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}$$

The derivative with respect to position is force

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial U}{\partial x} = F_x, \qquad \frac{\partial \mathcal{L}}{\partial y} = -\frac{\partial U}{\partial y} = F_y$$

while the derivative with respect to velocity is momentum

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial T}{\partial \dot{x}} = m\dot{x}, \qquad \frac{\partial \mathcal{L}}{\partial \dot{y}} = \frac{\partial T}{\partial \dot{y}} = m\dot{y}$$

Substituting this result in the two Euler-Lagrange equation, we have

$$F_x = m\ddot{x}, \qquad F_y = m\ddot{y}$$

which are the two component of Newton second law $\mathbf{F} = m\ddot{\mathbf{r}}$.

Newton Law in Polar Coordinate

The Lagrangian is

$$\mathcal{L}(r, \phi, \dot{r}, \dot{\phi}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - U(r, \phi)$$

which result into two Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial r} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{r}} \qquad \frac{\partial \mathcal{L}}{\partial \phi} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\phi}}$$

Before moving into evaluating the derivative with respect to r and ϕ, recall the gradient of potential energy in polar coordinate

$$\nabla U = \frac{\partial U}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial U}{\partial \phi}\hat{\phi} - F_r\hat{\mathbf{r}} - F_\phi\hat{\phi}$$

Evaluating the radial derivative

$$mr\dot{\phi} - \frac{\partial U}{\partial r} = m\ddot{r}$$
$$F_r = m(\ddot{r} + r\dot{\phi})$$

which is simply the radial component of the Newton's second law in polar coordinate. Evaluating the angular derivative

$$-\frac{\partial U}{\partial \phi} = \frac{d}{dt}mr^2\dot{\phi} = m(2r\dot{r}\dot{\phi} + r^2\ddot{\phi})$$

$$-\frac{1}{r}\frac{\partial U}{\partial \phi} = m(2\dot{\mathbf{r}}\dot{\phi} + r^2\ddot{\phi})$$

$$F_\phi = m(2\dot{\mathbf{r}}\dot{\phi} + r\ddot{\phi})$$

which is, as previously mentioned, the angular component of Newton's second law. It should be noted that the quantity $mr^2\dot{\phi}$ can be recognized as angular momentum L, and the rate of change of it is torque Γ

$$\Gamma = F_\phi r = \frac{dL}{dt} = \frac{d}{dt}mr^2\dot{\phi}$$

Lagrangian with Explicit Constraint Forces Using Lagrange Multipliers

The modified Euler-Lagrangian that include constraining force can be obtained by Lagrange Multipliers. For two dimension Lagrangian $\mathcal{L}(x, \dot{x}, y, \dot{y}, t)$ with one constraint $f(x, y) = C$, we must solve two modified Lagrange equation plus one for the constraint

$$\frac{\partial \mathcal{L}}{\partial x} + \lambda\frac{\partial f}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}$$

$$\frac{\partial \mathcal{L}}{\partial y} + \lambda\frac{\partial f}{\partial y} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}$$

$$f(x, y) = \text{Constant}$$

Lagrange multiplier is not simply mathematical technique, in fact given partial derivatives of the constraint function $f(x, y)$, the Lagrange multiplier $f(x, y)$ gives the corresponding components of the constraint force

$$\lambda\frac{\partial f}{\partial q_i} = F_i^{\text{cstr}}$$

Derivation. We consider the case of two dimension Lagrange. Then deviate the correct path $x(t)$ and $y(t)$ into

$$x(t) \rightarrow x(t) + \delta x$$

$$y(t) \rightarrow y(t) + \delta y$$

With the Lagrangian of $\mathcal{L}(x, \dot{x}, y, \dot{y})$, the action integral takes the form

$$S = \int_{t_1}^{t_2} \mathcal{L}\, dt$$

Assuming the constraint is consisted with the displacement, then, as proved before, the integral is unchanged and $\delta S = 0$. In other words

$$\int \left(\frac{\partial \mathcal{L}}{\partial x}\delta x + \frac{\partial \mathcal{L}}{\partial \dot{x}}\delta\dot{x} + \frac{\partial \mathcal{L}}{\partial y}\delta y + \frac{\partial \mathcal{L}}{\partial \dot{y}}\delta\dot{y}\right)\, dt = 0$$

Using integration by part on second and fourth term to move the derivative sign

$$\int \left(\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right)\delta x\, dt + \int \left(\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\right)\delta y\, dt = 0$$

For displacement δx and δy, the terms inside parenthesis must be zero for the integral to be zero. Now, this is just the proof of Euler-Lagrange equation. However, we want to explicitly including the constraint, which can be achieved by multiplying the deviation of the constraint equation δf with Lagrange multipliers

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y$$

$$0 = \lambda\frac{\partial f}{\partial x}\delta x + \lambda\frac{\partial f}{\partial y}\delta y$$

This step is justified because the value is zero anyway. Then we can add it into the integral of δS without changing the integral itself

$$\int \left(\frac{\partial \mathcal{L}}{\partial x} + \lambda\frac{\partial f}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) \delta x \, dt + \int \left(\frac{\partial \mathcal{L}}{\partial y} + \lambda\frac{\partial f}{\partial \hat{y}} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\right) \delta y \, dt = 0$$

Since the multipliers is arbitrary, we define Lagrange multipliers such that the terms inside parenthesis is zero, thus resulting in two modified Euler-Lagrange equation written previously

$$\frac{\partial \mathcal{L}}{\partial x} + \lambda\frac{\partial f}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}$$

$$\frac{\partial \mathcal{L}}{\partial y} + \lambda\frac{\partial f}{\partial y} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}$$

Now we add the constraint equation since we need three equation to find three equation in three unknown.

Lagrange multipliers physical meaning. Given that the kinetic energy of the system does not depend on the position and the potential energy does not depend on velocity, the modified Euler-Lagrange equation reads

$$-\frac{\partial U}{\partial q_i} + \lambda\frac{\partial d}{\partial q_i} = m\ddot{q}_i$$

$$\lambda\frac{\partial f}{\partial q_i} = m\ddot{q}_i + \frac{\partial U}{\partial q_i}$$

Recall that the negative gradient of potential energy is the non constraint force, while the product of mass and generalized acceleration is total force. The total force is the sum of non constraint force and constraint force, so

$$\lambda\frac{\partial f}{\partial q_i} = F_i^{\text{cstr}}$$

Comparison with the usual function optimization using Lagrange multipliers. Let us compare the optimization method of Lagrange multipliers when the function we are concerned is $f(x, y)$ instead of the previous case of action integral S. To optimize $f(x, y)$, set its derivative to zero

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0$$

while to optimize S, we use the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0, \qquad \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} = 0$$

If we include constraint $\phi(x, y)$, we construct function $F = f + \lambda\phi$ such that the multipliers is defined

$$\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 0$$

while in the action integral case

$$\frac{\partial \mathcal{L}}{\partial x} + \lambda\frac{\partial \phi}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0, \qquad \frac{\partial \mathcal{L}}{\partial y} + \lambda\frac{\partial \phi}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} = 0$$

In any case, we need to solve those two resulting function alongside the constraint equation.

Lagrangian for a Charge in an Electromagnetic Field

The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\dot{\mathbf{r}}^2 - q(V - \dot{\mathbf{r}}\cdot\mathbf{A})$$

This Lagrangian is defined is such way to produce the Lorentz force law

$$\mathbf{F} = q\left(\mathbf{E} + \dot{\mathbf{r}}\times\mathbf{B}\right)$$

Derivation. The derivation of Lagrangian for particle in electromagnetic field assumes you know field and gauge theory. Since you didn't, we shall only prove that the Lagrangian does give the correct equation of motion.

First we write the Euler-Lagrangian equation in three dimension as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} = \frac{\partial \mathcal{L}}{\partial \mathbf{r}}$$

where both scalar and vector potential are a function of space and time. Feeding the given Lagrangian

$$\frac{d}{dt}\left(m\dot{\mathbf{r}} + q\mathbf{A}\right) = -q\nabla\left(V - \dot{\mathbf{r}}\cdot\mathbf{A}\right)$$

then rewriting it as such

$$\frac{d}{dt}\left(m\dot{\mathbf{r}}\right) = -q\nabla V - q\frac{d\mathbf{A}}{dt} + q\nabla\left(\dot{\mathbf{r}}\cdot\mathbf{A}\right)$$

The total derivative of $\mathbf{A}(\mathbf{r}, t)$ with respect to time has two part: spatial variation and explicit dependence on time

$$\frac{d\mathrm{A}}{dt} = \sum_i \frac{\partial\mathbf{A}}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial\mathbf{A}}{\partial t} = \sum_i \dot{q}_i\frac{\partial}{\partial q_i}\mathbf{A} + \frac{\partial\mathbf{A}}{\partial t} = \left(\dot{\mathbf{r}}\cdot\nabla\right)\mathbf{A} + \frac{\partial\mathbf{A}}{\partial t}$$

Now we have

$$\mathbf{F} = -q\left(\nabla V + \frac{\partial\mathbf{A}}{\partial t}\right) + q\left[\nabla\left(\mathbf{r}\cdot\dot{\mathbf{A}}\right) - \left(\dot{\mathbf{r}}\cdot\nabla\right)\mathbf{A}\right]$$

Recalling the vector triple product

$$\dot{\mathbf{r}} \times (\nabla \times \mathbf{A}) = \nabla (\dot{\mathbf{r}} \cdot \mathbf{A}) - (\dot{\mathbf{r}} \cdot \nabla) \mathbf{A}$$

to write it as

$$\mathbf{F} = -q \left(\nabla V + \frac{\partial \mathbf{A}}{\partial t} \right) + q\dot{\mathbf{r}} \times (\nabla \times \mathbf{A})$$

In terms of potential, both electromagnetic field maybe expressed as

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \qquad \text{and} \qquad \mathbf{B} = \nabla \times \mathbf{A}$$

Hence, we obtain our desired result

$$\mathbf{F} = q \left(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B} \right)$$

Application: Atwood's Machine

Atwood's machine consist of masses m_1 and m_2 are suspended by an in extensible string (length l) which passes over a massless pulley with frictionless bearings and radius R. The length of the string acts as constraint

$$x + y + 2\pi R = l$$

This implies $y = -x + C$ and $\dot{y} = -\dot{x}$. Thus, the kinetic energy of both mass

$$T = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\dot{y}^2 = \frac{1}{2}(m_1 + m_2)\dot{x}^2$$

We define the downward displacement of m_1 and upward displacement of m_2 as positive displacement in our generalized coordinate x. This is the same case of upward acceleration of m_2 being the same as downward acceleration of m_1. In any case, the potential energy is

$$U = -m_1gx - m_2gy = -(m_1 - m_2)gx + C_2$$

We can now write the Lagrangian as

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + (m_1 - m_2)gx$$

With only one generalized coordinate, we only have one Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}$$

which on substituting the Lagrangian yield

$$(m_1 - m_2)g = (m_1 + m_2)\ddot{x}$$
$$\ddot{x} = \frac{m_1 - m_2}{m_1 + m_2}g$$

Now let's compare it with Newtonian approach, we should obtain the same result. Considering the acceleration direction for both masses, the net forces on both m_1 and m_2 respectively are

$$m_1g - F_t = m_1\ddot{x}$$
$$F_t - m_2g = m_2\ddot{x}$$

Adding both equation

$$(m_1 - m_2)g = (m_1 + m_2)\ddot{x}$$
$$\ddot{x} = \frac{m_1 - m_2}{m_1 + m_2}g$$

Now suppose we explicitly include the constraint using Lagrange multipliers. Here, the constraint is the wire and the constraining force is the tension, which takes the form

$$f(x, y) = x + y = \text{Constant}$$

Since we do not need to reduce the number of coordinate, we express the Lagrangian as

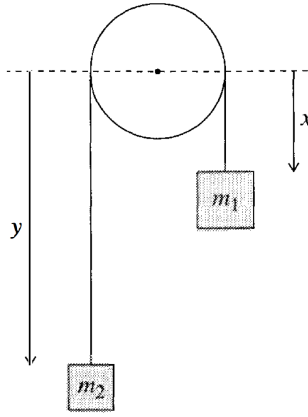$$\mathcal{L} = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\dot{y}^2 + m_1gx + m_2gy$$

Figure: Atwood's machine configuration

Then the equations that wee need to solve are as follows

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}$$
$$\frac{\partial \mathcal{L}}{\partial y} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}$$

$$x + y = \text{Constant}$$

The first two equation yield

$$m_1 g + \lambda = m_1 \ddot{x}, \qquad m_2 g + \lambda = m_2 \ddot{y}$$

Performing second derivative to the third equation with respect to time results

$$\ddot{x} = -\ddot{y}$$

Using this and subtracting the first equation by the second

$$(m_1 + m_2)\ddot{x} = (m_1 - m_2)g$$
$$\ddot{x} = \frac{m_1 - m_2}{m_1 + m_2}g$$

which is the same result. It can be seen that, by comparing to the Newtonian result, that the Lagrange multipliers gives

$$\lambda = -F_t$$

Application: Particle Constrained on a Cylinder

Consider a particle of mass m constrained to move on a frictionless cylinder of radius R. Besides the force of constraint, the only force on the mass is a force $\mathbf{F} = -k\mathbf{r}$ directed toward the origin. With $\mathbf{r}$ as the position vector of the particle, this force is the three dimension version of Hooke's law.

We shall use cylindrical coordinate to solve this problem. It is known that the radius component is fixed $\rho = R$, so we use (ϕ, z) as our generalized coordinate. The kinetic energy is

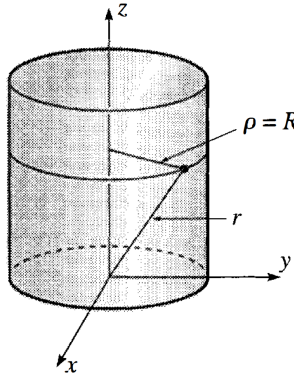$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(R^2\dot{\phi}^2 + \dot{z}^2)$$

Figure: Particle constrained to move on a cylinder

Recall $\mathbf{F} = -\nabla U$, hence the potential energy

$$-\frac{dU}{dr}\,\hat{\mathbf{r}} = -kr\,\hat{\mathbf{r}}$$

$$U = \frac{1}{2}kr^2$$

The distance of particle from origin is given by $r^2 = R^2 + z^2$, so

$$U = \frac{1}{2}k(z^2 + R^2)$$

Therefore, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}m(R^2\dot{\phi}^2 + \dot{z}^2) - \frac{1}{2}k(R^2 + z^2)$$

We use two generalized coordinates, thus we have two Euler-Lagrangian equation

$$\frac{\partial\mathcal{L}}{\partial z} = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{z}}, \qquad \frac{\partial\mathcal{L}}{\partial\phi} = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}}$$

The z equation is

$$-kz = m\ddot{z} \implies z = A\cos(\omega t - \delta)$$

This mean that the mass perform simple harmonic motion in the z direction. Now, the ϕ equation

$$0 = \frac{d}{dt}mR^2\dot{\phi}$$

This mean that the angular momentum $L = mR^2\dot{\phi}$ is conserved and the particle rotate in constant velocity $\dot{\phi}$.

Application: Sliding Block on a Frictionless Wedge

The block (mass m) is free to slide on the wedge, and the wedge (mass M) can slide on the horizontal table, both with negligible friction. The block is released from the top of the wedge, with both initially at rest.

The system has two coordinates and no constraint whatsoever, so it has two degree of freedom. We choose q_1 and q_2 as our generalized coordinate which denote the distance form the block from the top of the wedge and the distance of the wedge from convenient fixed point on the table. We also define positive x displacement to the right and downward as positive y displacement.

The kinetic energy is

$$T = \frac{1}{2}mv_m^2 + \frac{1}{2}Mv_M^2$$

The wedge velocity is simply

$$v_M = \dot{q}_2$$

Meanwhile, the block velocity have two component, which are

$$\mathbf{v}_m = (\dot{q}_2 + \dot{q}_1 \cos \alpha)\ \hat{\mathbf{x}} + \dot{q}_2 \sin \alpha\ \hat{\mathbf{y}}$$

In terms of generalized coordinate, the kinetic energy reads

$$T = \frac{1}{2}m\left[(\dot{q}_2 + \dot{q}_1 \cos \alpha)^2 + \dot{q}_2 \sin \alpha^2\right] + \frac{1}{2}m\dot{q}_2^2$$
$$= \frac{1}{2}m\left[\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_1\dot{q}_2 \cos \alpha+\right] + \frac{1}{2}M\dot{q}_2^2$$
$$T = \frac{1}{2}(m+M)\dot{q}_2^2 + \frac{1}{2}m\left(\dot{q}_1^2 + 2\dot{q}_1\dot{q}_2 \cos \alpha\right)$$

In other hand, we defined downward as positive displacement, so the potential energy reads

$$U = -mgy = -mgq_1 \sin \alpha$$

The Lagrangian can be evaluated as

$$\mathcal{L} = \frac{1}{2}(m+M)\dot{q}_2^2 + \frac{1}{2}m\left(\dot{q}_1^2 + 2\dot{q}_1\dot{q}_2 \cos \alpha\right) + mgq_1 \sin \alpha$$

With the generalized coordinates we used, we have two Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial q_1} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_2} \qquad \frac{\partial \mathcal{L}}{\partial q_2} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_2}$$

The q_1 equation yield

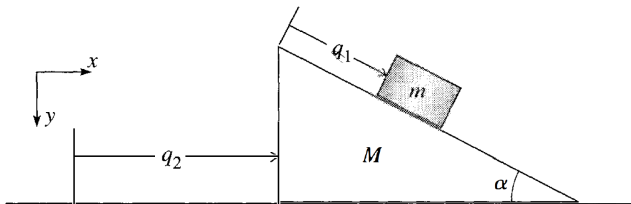$$mg \sin \alpha = \frac{d}{dt}\left[m(\dot{q}_1 + \dot{q}_2 \cos \alpha)\right]$$

Figure: Block slides on a wedge which is free to move without friction

$$g \sin \alpha = \ddot{q}_1 + \ddot{q}_2 \cos \alpha$$

while the q_2 equation yield

$$0 = \frac{d}{dt} \left[(M + m)\dot{q}_2 + m(\dot{q}_2 + \dot{q}_1 \cos \alpha) \right]$$
$$= (M + m)\ddot{q}_2 + m(\ddot{q}_2 + \ddot{q}_1 \cos \alpha)$$
$$\ddot{q}_2 = -\frac{m}{M + m}\ddot{q}_1 \cos \alpha$$

which is just conservation of momentum in the x direction

$$m\dot{q}_2 + m\left(\dot{q}_2 + \dot{q}_1 \cos \alpha\right) = \text{Constant}$$

Now, combining the q_1 and q_2 result in

$$\ddot{q}_1 = g \sin \alpha + \frac{m}{M + m}\ddot{q}_1 \cos^2 \alpha$$
$$\ddot{q}_1 = \frac{g \sin \alpha}{1 - \frac{m}{M+m} \cos^2 \alpha}$$

Suppose we want to determine the time it took for the block to reach the bottom of wedge, we can use the kinematic relation $x(t) = x_i + v_i t + at^2/2$ or $t = \sqrt{2l/a}$, with l as the length of the slope, to obtain

$$t = \left(2l \frac{1 - \frac{m}{M+m} \cos^2 \alpha}{g \sin \alpha} \right)^{1/2}$$

As a sanity check, consider the case for $\alpha = 90°$. The acceleration $\ddot{q}_1 = g$, which is correct. Another is the case for $M \to \infty$. The acceleration $\ddot{q}_1 = g \sin \alpha$, which is the acceleration for a block on a fixed incline.

Application: Simple Pendulum

A bob of mass m is fixed to a massless rod length $l = \sqrt{x^2 + y^2}$, which is pivoted at O and free to swing without friction in the xy plane. One way to integrate the constraint into the Lagrangian is by expressing both coordinate in terms single generalized coordinate ϕ or by writing one of them in terms other variable, say $y = \sqrt{l^2 - x^2}$.

In terms of ϕ, the kinetic energy is

$$T = \frac{1}{2}ml^2\dot{\phi}^2$$

and the potential energy

$$U = mg(l - l \cos \phi) = mgl(1 - \cos \phi)$$

So, the Lagrangian reads

$$\mathcal{L} = \frac{1}{2}ml^2\dot{\phi}^2 - mgl(1 - \cos \phi)$$

In this case, we only have one Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$$
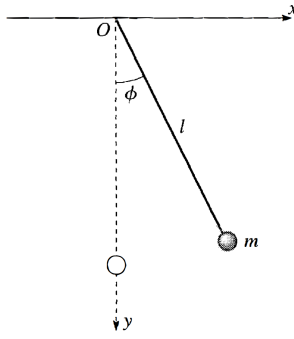
Figure: A simple pendulum

Substituting the Lagrangian

$$-mgl \sin \phi = ml^2 \ddot{\phi}$$

$$\sin \phi = -\frac{l}{g} \ddot{\phi}$$

This is the differential equation describing simple pendulum motion, which, on assuming small angle ϕ, has the solution $\phi = A \cos(\omega t + \delta)$. Also, recall $\ddot{\phi}$ is the angular acceleration. This mean that the Euler-Lagrange equation reproduce the formula for torque $\Gamma = I\alpha = ml^2 \ddot{\phi}$ or $\Gamma = Fr = -mgl \sin \phi$.

Application: Bead Spinning on a Wire Hoop

A bead of mass m is threaded on a frictionless circular wire hoop of radius R. The hoop lies in a vertical plane, which is forced to rotate about the hoop's vertical diameter with constant angular velocity $\dot{\phi} = \omega$. The bead's position on the hoop is specified by the angle θ measured up from the rotation axis. We shall use θ as our only generalized coordinate.

The kinetic energy is

$$T = \frac{1}{2}m \left(v_\theta^2 + v_\phi^2 \right)$$

with v_θ denote the bead tangential velocity with respect to non-rotating hoop, while v_ϕ denote the rotation velocity of the hoop. The tangential velocity is simply $v_\theta = R\dot{\theta}$,

$$v_\theta = R\dot{\theta}$$

with R as the distance of the bead with the axis of rotation–the hoop radius in other words. With the same principle, the hoop velocity is $v_\phi = \rho\dot{\phi}$. From the figure and known quantity, $\rho = R \cos \theta$ and $\dot{\phi} = \omega$, thus

$$v_\phi = R \sin \theta \omega$$

In terms of generalized coordinate, the kinetic energy reads

$$T = \frac{1}{2}mR^2 \left(\dot{\theta}^2 + \omega^2 \sin^2 \theta \right)$$

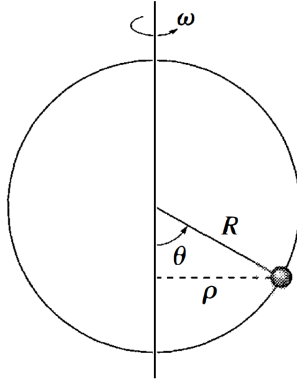Figure: Bead constrained into moving within wire hoop

The potential energy is

$$U = mg(R - R\cos\theta) = mgR(1 - \cos\theta)$$

Therefore, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}mR^2\left(\dot{\theta}^2 + \omega^2\sin^2\theta\right) - mgR(1 - \cos\theta)$$

which yield one Euler-Lagrange equation

$$\frac{\partial\mathcal{L}}{\partial\theta} = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\theta}}$$

Substituting the Lagrangian

$$mR^2\omega^2\sin\theta\cos\theta - mgR\sin\theta = mR^2\ddot{\theta}$$

dividing by mR^2

$$\ddot{\theta} = \omega^2\sin\theta\cos\theta - \frac{g}{R}\sin\theta$$

$$\ddot{\theta} = \left(\omega^2\cos\theta - \frac{g}{R}\right)\sin\theta$$

This equation can be used to determine the equilibrium point–that is, the point where the position of the system does not change–and the stable point–that is, the position at which the system returns after slightly disturbed–of the system. The requirement of equilibrium point is $\dot{\theta} = 0$, but we can obtain the same result by setting both $\dot{\theta}$ and $\ddot{\theta}$ to zero

$$\left(\omega^2\cos\theta - \frac{g}{R}\right)\sin\theta = 0$$

If we set $\sin\theta$ to zero, we obtain the following equilibrium points

$$\theta = 0, \pi$$

If we set the term inside parenthesis to zero

$$\cos\theta = \frac{g}{\omega^2 R}$$

or, due to cosine being even function

$$\cos(-\theta) = \frac{g}{\omega^2 R}$$

hence we have the following equilibrium

$$\theta = \pm \arccos \frac{g}{\omega^2 R}$$

Arccos function only has real value at $\theta \in [-1, 1]$, so when

$$\left| \frac{g}{\omega^2 R} \right| > 1 \quad \text{or} \quad \omega^2 < \frac{g}{R}$$

This equation undefined and the stable point disappear. These equilibrium points only appear when $\omega^2 > g/R$ and located on either side of bottom $\theta = 0$.

Out of these four equilibrium points, not all of them are stable.

1. Top $\theta = \pi$ point is unstable point due to not having restorative force, both gravitational and centrifugal push the bead away.

2. Bottom $\theta = 0$ point depends on ω^2. It is stable if $\omega^2 < g/R$, but become unstable if $\omega^2 > g/R$. This can be proven by approximating small θ displacement

$$\ddot{\theta} = \left(\omega^2 - g/R \right) \theta$$

If $\omega^2 < g/R$, then

$$\ddot{\theta} = -\Omega^2 \theta$$

with

$$\Omega = \sqrt{\frac{g}{r} - \omega^2}$$

which mean the bead perform simple harmonic motion about the stable point. If $\omega^2 > g/R$, then

$$\ddot{\theta} = \Omega^2 \theta$$

which has the solution $\theta = Ae^{\Omega t} + Be^{-\Omega t}$, so it moves in exponential way and the point is unstable.

3&4 These two point that comes after speeding up the rotation such that $\omega^2 > g/R$ are stable. To proof this, we expand the equation

$$\ddot{\theta} = \left(\omega^2 \cos \theta - \frac{g}{R} \right) \sin \theta$$

around stable point θ_0

$$\theta \equiv \theta + \epsilon$$

using Taylor expansion

$$\cos(\theta_0 + \epsilon) \approx \cos(\theta_0) - \epsilon \sin \theta_0, \quad \sin(\theta_0 + \epsilon) \approx \sin \theta_0 + \epsilon \cos \theta_0$$

This result in

$$\ddot{\theta} = \left[\omega^2 \cos \theta_0 - \omega^2 \epsilon \sin \theta_0 - \frac{g}{R} \right] [\sin \theta_0 + \epsilon \cos \theta_0]$$

$$\ddot{\theta} = -\omega^2 \epsilon \sin^2 \theta_0 - \omega^2 \epsilon^2 \sin \theta_0 \cos \theta_0$$

Since ϵ is small, we can ignore the second order

$$\ddot{\theta} = -\epsilon \omega^2 \sin^2 \theta_0$$

Since $\ddot{\theta}$ is the same as $\ddot{\epsilon}$

$$\ddot{\epsilon} = -\omega^2 \sin^2 \theta = -\Omega'^2 \epsilon$$

with

$$\Omega = \omega \sin \theta = \sqrt{\omega - \omega \cos^2 \theta_0} = \sqrt{\omega^2 - \left(\frac{g}{\omega R}\right)^2}$$

This mean ϵ oscillates about zero, and the bead itself oscillates about the equilibrium position θ_0 with frequency Ω'.

Hamiltonian Mechanics

The Hamiltonian is defined as

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}$$

The Hamilton equation derived from this is

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i}$$

The conservation of Hamiltonian is stated as follows

$$\frac{\partial \mathcal{L}}{\partial t} = 0 \implies \frac{dH}{dt} = 0$$

In the case of natural coordinate, that is the relation between the generalized coordinates and Cartesian is time-independent, the Hamiltonian takes the simple form $\mathcal{H} = T + U$

$$q_i = q_i(\mathbf{q}) \implies \mathcal{H} = T + U$$

Consider the Hamiltonian

$$\mathcal{H} = \mathcal{H}\left(\mathbf{q}, \mathbf{p}, t\right)$$

If the Hamiltonian does not depend on the coordinate q_i, the coordinate is said to be cyclic. Correspondingly, the canonical momentum $p_i = \partial \mathcal{L}/\partial q_i$ is conserved. The condition for q_i to be cyclic can be written as

$$\frac{\partial \mathcal{H}}{\partial q_i} = 0$$

The steps to solve problem using Hamilton formalism are as follows.

1. Choose the generalized coordinate $\mathbf{q}$.

2. Write the T and U in terms of $(\mathbf{q}, \dot{\mathbf{q}})$.

3. Find the generalized momenta $p_i = \partial \mathbf{L}/\partial \dot{q}_i$.

4. Solve for $\dot{\mathbf{q}}$ in terms $(\mathbf{q}, \mathbf{p})$.

5. Write $\mathcal{H}$ in terms $(\mathbf{q}, \mathbf{p})$.

6. Write the Hamilton equation for motion.

Notation

To avoid clutter, we define the following notation. The n-dimension system with n-generalized coordinate is represented by

$$\mathbf{q} \equiv (q_1, \ldots, q_n)$$

while n-generalized velocity

$$\dot{\mathbf{q}} \equiv (\dot{q}_1, \ldots, \dot{q}_n)$$

and generalized momentum

$$\mathbf{p} \equiv (p_1, \ldots, p_n)$$

In this case, $\mathbf{p}$ and $\mathbf{p}$ are n-dimensional vectors in the space of generalized position and generalized momentum.

Definition

The definition Hamiltonian comes from Lagrangian. To see this, consider the change of Lagrangian $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ as the time increase

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial q_1}\frac{\partial q_1}{\partial t} \cdots + \frac{\partial \mathcal{L}}{\partial q_n}\frac{\partial q_n}{\partial t} + \frac{\partial \mathcal{L}}{\partial \dot{q}_1}\frac{\partial \dot{q}_2}{\partial t} + \cdots + \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\frac{\partial \dot{q}_n}{\partial t} + \frac{\partial \mathcal{L}}{\partial t}\frac{\partial t}{\partial t}$$

$$= \sum_{i=1}^{n}\left(\frac{\partial \mathcal{L}}{\partial q_i}\dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\ddot{q}_i\right) + \frac{\partial \mathcal{L}}{\partial t}$$

$$= \sum_{i=1}^{n}\left(\dot{p}_i\dot{q}_i + p_i\ddot{q}_i\right) + \frac{\partial \mathcal{L}}{\partial t}$$

$$\frac{d\mathcal{L}}{dt} = \frac{d}{dt}\sum_{i=1}^{n}\left(p_i\dot{q}_i\right) + \frac{\partial \mathcal{L}}{\partial t}$$

If we assume the Lagrangian does not depend explicitly on time, say in the case of $T = T(\mathbf{q}, \dot{\mathbf{q}})$ and $U = U(\mathbf{q})$, the $\partial \mathcal{L}/\partial t$ term is zero. Thus, the quantity

$$\frac{d}{dt}\left(\sum_{i=1}^{n}p_i\dot{q}_i - \mathcal{L}\right) = 0$$

which is defined as Hamiltonian, is conserved. In other words,

$$\frac{\partial \mathcal{L}}{\partial t} = 0 \implies \frac{dH}{dt} = 0$$

Special Case of Hamiltonian

The special case we are referring is $\mathcal{H} = T + U$, which occur when the coordinates are natural

$$\mathbf{r}_\alpha = \mathbf{r}_\alpha(\mathbf{q})$$

where the α subscript used to denote the coordinate as α-th particle's coordinate. We first begin by expressing the generalized velocity in terms of generalized coordinates, which can be obtained by performing partial derivative with respect to the generalized coordinate

$$\frac{\partial \mathbf{r}_\alpha}{\partial t} = \frac{\partial \mathbf{r}_\alpha}{\partial q_1}\frac{\partial q_1}{\partial t} \cdots + \frac{\partial \mathbf{r}_\alpha}{\partial q_n}\frac{\partial q_n}{\partial t} = \sum_{i=1}^{n}\frac{\partial \mathbf{r}_\alpha}{\partial q}\dot{q}_i$$

Then its square is just the dot product with itself

$$\left(\frac{\partial \mathbf{r}_\alpha}{\partial t}\right)^2 = \sum_{j=1}^{n}\frac{\partial \mathbf{r}_\alpha}{\partial q_i}\dot{q}_i \cdot \sum_{k=1}^{n}\frac{\partial \mathbf{r}_\alpha}{\partial q_j}\dot{q}_j$$

The kinetic energy is then product of triple sum

$$T = \frac{1}{2}\sum_{\alpha}m_\alpha\mathbf{r}_\alpha = \frac{1}{2}\sum_{\alpha}m_\alpha\sum_{j}\frac{\partial \mathbf{r}_\alpha}{\partial q_j}\dot{q}_j \cdot \sum_{k}\frac{\partial \mathbf{r}_\alpha}{\partial q_k}\dot{q}_k$$

If we define

$$A_{jk} = \sum_{\alpha}m_\alpha\frac{\partial \mathbf{r}_\alpha}{\partial q_j} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_k}$$

The expression for kinetic energy simplifies into

$$T = \frac{1}{2} \sum_{j,k} A_{jk} \dot{q}_j \dot{q}_k$$

If we assume kinetic energy $T = T(\mathbf{q}, \dot{\mathbf{q}})$ and potential energy $U = U(\mathbf{q})$, which is what natural coordinates imply anyway, then

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} = p_i$$

To evaluate partial derivative, we need the following relation

$$\frac{d}{dv_i} \sum_{j,k} A_{jk} v_j v_k = 2 \sum_{j,k} A_{ij} v_j \quad \text{if } A_{jk} = A_{kj}$$

Applying the identity

$$p_i = \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2} \sum_{jk} A_{jk} \dot{q}_j \dot{q}_k \right) = \sum_j A_{ij} \dot{q}_j$$

Therefore, the Hamiltonian now reads

$$\mathcal{H} = \sum_i p_i \dot{q}_i - \mathcal{L} = \sum_i \sum_j A_{ij} \dot{q}_i \dot{q}_j - (T - U)$$

$$\mathcal{H} = 2T - T + U = T + U$$

Hamilton Equation Derivation

The assumption that must be satisfied first are the constraint must be holonomic–the number of degree of freedom matches the number of generalized coordinates–and that the non constraint force must be derivable from potential energy $\mathbf{F} = -\nabla U$. In principle, to obtain Hamilton equation of motion, we differentiate the Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p})$ with both variable q_i and p_i. To evaluate the derivative, we need to express the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_1 - \mathcal{L}$$

in terms of $\mathbf{q}$ and $\mathbf{p}$

First note that the generalized momentum is a function of generalized velocity and perhaps time, which can be seen from

$$p_i = \frac{\partial}{\partial \dot{q}_i} \mathcal{L}(\mathbf{q}, \dot{q}_i, t) \implies p_i = p_i(\dot{q}, t)$$

Therefore, in principle, we can solve p_i to express $\dot{q}_i$ in terms of variables $\mathbf{p}$ and $\mathbf{q}$, also time perhaps

$$\dot{q}_i = \dot{q}_i(\mathbf{q}, \mathbf{p}_i, t)$$

Then the next step is to actually perform the differentiation. First we differentiate with respect to q_i

$$\frac{\partial \mathcal{H}}{\partial q_i} = \frac{\partial}{\partial q_i} \sum_{j=1}^{n} p_j \dot{q}_j - \frac{\partial \mathcal{L}}{\partial q_i}$$

The sum differentiation can be evaluated as follows

$$\frac{\partial}{\partial q_i} \sum_j p_j \dot{q}_j = \sum_j \left(\dot{q}_j \frac{\partial p_j}{\partial q_i} + p_j \frac{\partial \dot{q}_j}{\partial q_i} \right) = \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_i}$$

While the Lagrangian term

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\partial \mathcal{L}}{\partial q_1} \frac{\partial q_1}{\partial q_i} + \cdots \frac{\partial \mathcal{L}}{\partial q_n} \frac{\partial q_n}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial \dot{q}_1} \frac{\partial \dot{q}_1}{\partial q_i} + \cdots + \frac{\partial \mathcal{L}}{\partial \dot{q}_n} \frac{\partial \dot{q}_n}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial t} \frac{\partial t}{\partial q_i}$$

$$= \sum_j \left(\frac{\partial \mathcal{L}}{\partial q_j} \frac{\partial q_j}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial q_i} \right)$$

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\partial \mathcal{L}}{\partial q_i} + \sum_j p_i \frac{\partial \dot{q}_j}{\partial q_i}$$

Substituting into the Hamiltonian

$$\frac{\partial \mathcal{H}}{\partial q_i} = \sum_j p_i \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial \mathcal{L}}{\partial q_i} - \frac{\partial \dot{q}_j}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i} = -\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = -\dot{p}_i$$

Next is the derivative with respect to p_i

$$\frac{\partial \mathcal{H}}{\partial p_i} = \frac{\partial}{\partial p_i} \sum_{j=1}^n p_i \dot{q}_1 - \frac{\partial \mathcal{L}}{\partial p_i}$$

The sum differentiation can be evaluated as follows

$$\frac{\partial}{\partial p_i} \sum_j p_j \dot{q}_j = \sum_j \left(\dot{q}_j \frac{\partial p_j}{\partial p_i} + p_j \frac{\partial \dot{q}_j}{\partial p_i} \right) = \dot{q}_i + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_j}$$

While the Lagrangian term

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{\partial \mathcal{L}}{\partial q_1} \frac{\partial q_1}{\partial p_i} + \cdots \frac{\partial \mathcal{L}}{\partial q_n} \frac{\partial q_n}{\partial p_i} + \frac{\partial \mathcal{L}}{\partial \dot{q}_1} \frac{\partial \dot{q}_1}{\partial p_i} + \cdots + \frac{\partial \mathcal{L}}{\partial \dot{q}_n} \frac{\partial \dot{q}_n}{\partial p_i} + \frac{\partial \mathcal{L}}{\partial t} \frac{\partial t}{\partial p_i}$$

$$= \sum_j \left(\frac{\partial \mathcal{L}}{\partial q_j} \frac{\partial q_j}{\partial p_i} + \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial p_i} \right)$$

$$\frac{\partial \mathcal{L}}{\partial p_i} = \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i}$$

Substituting into the Hamiltonian

$$\frac{\partial \mathcal{H}}{\partial q_i} = \dot{q}_i + \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_j} - \sum_j p_j \frac{\partial \dot{q}_j}{\partial p_i} = \dot{q}_i$$

Hamiltonian Dependence on Time

Consider the total derivative of the Hamiltonian with respect to time

$$\frac{d\mathcal{H}}{\partial t} = \sum_i \left(\frac{\partial \mathcal{H}}{\partial q_i} \dot{q} + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i \right) + \frac{\partial \mathcal{H}}{\partial t}$$

Then using the Hamilton equation of motion

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t}$$

The total derivative $d\mathcal{H}/dt$ denote the actual changes as t moves forward, which will changes both variables $(\mathbf{q}, \mathbf{p})$ also. In other hand, the partial derivative $\partial\mathcal{H}/\partial t$ denote the change of the Hamilton as t moves forward with $(\mathbf{q}, \mathbf{p})$ being held constant. This will be zero if $\mathcal{H}$ does not explicitly depend on t. Thus, if $\mathcal{H}$ does not explicitly depend on t, then $\mathcal{H}$ is conserved.

Newton's Law in 2D

Since we are using 2D Cartesian, the Hamiltonian is simply the system energy

$$\mathcal{H} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + U(x,y) = \frac{1}{2m}\left(p_x^2 + p_y^2\right) + U(x,y)$$

The x-th component reads

$$\dot{x} = \frac{\partial\mathcal{H}}{\partial p_x} = \frac{p_x}{m} \qquad \dot{p}_x = -\frac{\partial\mathcal{H}}{\partial x} = -\frac{\partial U}{\partial x}$$

while the y-th component

$$\dot{y} = \frac{\partial\mathcal{H}}{\partial p_y} = \frac{p_y}{m} \qquad \dot{p}_y = -\frac{\partial\mathcal{H}}{\partial y} = -\frac{\partial U}{\partial y}$$

Combining both, we obtain the Newton's definition of momentum and force

$$\mathbf{p} = m\mathbf{r} \qquad \mathbf{F} = \frac{d\mathbf{p}}{dt}$$

Newton's Law in Polar Coordinate As previously, the Hamiltonian is the system energy

$$\mathcal{H}(r, \phi, \dot{r}, \dot{\phi}) = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - U(r)$$

Next, we evaluate the momenta

$$p_r = \frac{\partial\mathcal{L}}{\partial\dot{r}} = m\dot{r}, \qquad p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = mr^2\dot{\phi}$$

and solve for generalized velocity

$$\dot{r} = \frac{p_r}{m} \qquad \dot{\phi} = \frac{p_\phi}{mr^2}$$

Then rewrite the Hamiltonian in terms of coordinate and momenta

$$\mathcal{H} = \frac{1}{2}m\left(\frac{p_r^2}{m^2} + \frac{p_\phi^2}{m^2r^2}\right) + U(r) = \frac{1}{2}\left(p_r^2 + \frac{p_\phi^2}{r^2}\right) + U(r)$$

The radial components are

$$\dot{r} = \frac{\partial\mathcal{H}}{\partial p_r} = \frac{p_r}{m} \qquad \dot{p}_r = -\frac{\partial\mathcal{H}}{\partial r} = \frac{p_\phi^2}{mr^3} - \frac{\partial U}{\partial r}$$

Combining both, we obtain

$$m\ddot{\mathbf{r}} = \frac{p_\phi^2}{mr^3} - \frac{\partial U}{\partial r}$$

which is the definition of total force $m\ddot{\mathbf{r}}$ as the sum of radial force $-\partial U/\partial r$ and centrifugal force p_ϕ^2/mr^3. Then the angular component

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial p_\phi} = \frac{p_\phi}{mr^2} \qquad \dot{p}_\phi = -\frac{\partial \mathcal{H}}{\partial \phi} = 0$$

which reproduce the definition of angular momentum and conservation of angular momentum.

Hamiltonian of a Charge in an Electromagnetic Field

The Hamiltonian is

$$\mathcal{H} = \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}\right)^2 + qV$$

From this, we obtain the two equation of motion

$$\dot{\mathbf{r}} = \frac{1}{m}\left(\mathbf{p} - q\mathbf{A}\right) \qquad \text{and} \qquad \dot{\mathbf{p}} = q\left(\dot{\mathbf{r}} \cdot \nabla\right)\mathbf{A} - q\nabla V$$

The momentum evolution is consistent with the Lorentz force law

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)$$

Derivation. By definition, Hamiltonian in three dimension is

$$\mathcal{H} = \sum_{i=1}^{3} q_i p_i - \mathcal{L} = \mathbf{p} \cdot \dot{\mathbf{r}} - \mathcal{L}$$

With the known Lagrangian of a particle in electromagnetic field

$$\mathcal{L} = \frac{1}{2}m\dot{\mathbf{r}}^2 - q(V - \dot{\mathbf{r}} \cdot \mathbf{A})$$

and known canonical momentum

$$\mathbf{p} = m\dot{\mathbf{r}} + q\mathbf{A}$$

the Hamiltonian reads

$$\mathcal{H} = m\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + q\mathbf{A} \cdot \dot{\mathbf{r}} - \frac{1}{2}m\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + q\left(V - \dot{\mathbf{r}} \cdot \mathbf{A}\right)$$
$$\mathcal{H} = \frac{1}{2}m\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + qV$$

Expressing the change of coordinate as momentum

$$\mathcal{H} = \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}\right)^2 + qV$$

For the first equation of motion, consider

$$\dot{\mathbf{r}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \frac{1}{2m}2\left(\mathbf{p} - q\mathbf{A}\right) = \frac{1}{m}\left(\mathbf{p} - q\mathbf{A}\right)$$

On multiplying both side with m, we obtain the kinetic momentum $\pi = m\dot{\mathbf{r}}$, as opposed to canonical momentum $\pi = m\dot{\mathbf{r}} + q\mathbf{A}$.

Next we move to its conjugate. Consider

$$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}} = \frac{1}{2m} 2 \left(\mathbf{p} - q\mathbf{A} \right) q \cdot \frac{\partial \mathbf{A}}{\partial \mathbf{r}} - 1\frac{\partial V}{\partial \mathbf{r}}$$

$$\dot{\mathbf{p}} = \frac{q}{m} \left(\mathbf{p} - q\mathbf{A} \right) \cdot \nabla \mathbf{A} - q\nabla V$$

Substituting the value of kinetic momentum obtained previously

$$\dot{\mathbf{p}} = \frac{q}{m} m\dot{\mathbf{r}} \cdot \nabla \mathbf{A} - q\nabla V = q\dot{\mathbf{r}} \cdot \mathbf{A} - q\nabla \mathbf{A}$$

Writing the first equation of motion as

$$m\dot{\mathbf{r}} = \mathbf{p} - q\mathbf{A}$$

$$m\ddot{\mathbf{r}} = \dot{\mathbf{p}} - q\frac{d\mathbf{A}}{dt}$$

and substituting it back

$$m\ddot{\mathbf{r}} + q\frac{\partial \mathbf{A}}{\partial t} = q\dot{\mathbf{r}} \cdot \nabla \mathbf{A} - q\nabla \mathbf{A}$$

$$m\ddot{\mathbf{r}} = q\dot{\mathbf{r}} \cdot \nabla \mathbf{A} - q\nabla V - q\frac{d\mathbf{A}}{dt}$$

The total derivative of $\mathbf{A}(\mathbf{r}, t)$ with respect to time has two part: spatial variation and explicit dependence on time

$$\frac{d\mathbf{A}}{dt} = \sum_i \frac{\partial \mathbf{A}}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial \mathbf{A}}{\partial t} = \sum_i \dot{q}_i \frac{\partial}{\partial q_i} \mathbf{A} + \frac{\partial \mathbf{A}}{\partial t} = \left(\dot{\mathbf{r}} \cdot \nabla \right) \mathbf{A} + \frac{\partial \mathbf{A}}{\partial t}$$

Hence we write

$$\mathbf{F} = -q \left(\nabla V + \frac{\partial \mathbf{A}}{\partial t} \right) + q \left[\nabla \left(\mathbf{r} \cdot \dot{\ } \mathbf{A} \right) - \left(\dot{\mathbf{r}} \cdot \nabla \right) \mathbf{A} \right]$$

Recalling the vector triple product

$$\dot{\mathbf{r}} \times \left(\nabla \times \mathbf{A} \right) = \nabla \left(\dot{\mathbf{r}} \cdot \mathbf{A} \right) - \left(\dot{\mathbf{r}} \cdot \nabla \right) \mathbf{A}$$

to write it as

$$\mathbf{F} = -q \left(\nabla V + \frac{\partial \mathbf{A}}{\partial t} \right) + q\dot{\mathbf{r}} \times \left(\nabla \times \mathbf{A} \right)$$

In terms of potential, both electromagnetic field maybe expressed as

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \qquad \text{and} \qquad \mathbf{B} = \nabla \times \mathbf{A}$$

Hence, we obtain our desired result

$$\mathbf{F} = q \left(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B} \right)$$

Poisson Bracket

Poisson bracket between two variable $\omega(q, p)$ and $\lambda(q, p)$ is defined as

$$\{\omega, \lambda\} \equiv \sum_i \left(\frac{\partial \omega}{\partial q_i}\frac{\partial \lambda}{\partial p_i} - \frac{\partial \omega}{\partial p_i}\frac{\partial \lambda}{\partial q_i} \right)$$

Key properties of Poisson bracket are as follows.

1. **Bilinearity.** $\{af + bg, \} = a\{f, h\} + b\{g, h\}$

2. **Antisymmetry.** $\{f, g\} = -\{g, f\}$

3. **Jacobi identity.** $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$

4. **Leibniz rule.** $\{fg, h\} = f\{g, h\} + \{f, h\}g$

Time evolution. In Hamiltonian mechanics, any variable that does not depend on explicitly with time $\omega(\mathbf{q}, \mathbf{p})$ whose Poisson brackets with $\mathcal{H}$ vanish

$$\{\omega, \mathcal{H}\} = 0$$

is conserved. Its time evolution is given by the same Poisson brackets, for we can write

$$\frac{d\omega}{dt} = \sum_i \left(\frac{\partial \omega}{\partial q_i} \dot{q}_q + \frac{\partial \omega}{\partial p_i} \dot{p}_i \right) = \sum_i \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right) = \{\omega, \mathcal{H}\}$$

For variable that depend on explicitly with time $\omega(\mathbf{q}, \mathbf{p}, t)$, the time evolution is given by

$$\frac{d\omega}{dt} = \{\omega, \mathcal{H}\} + \frac{\partial \omega}{\partial t}$$

Canonically conjugate. Phase space coordinates that satisfies the following equations are called canonically conjugate.

$$\{q_i, q_j\} = 0 \qquad \{p_i, p_j\} = 0 \qquad \{q_i, p_j\} = \delta_{ij}$$

To see why that is the case, consider the expanded form of Poisson bracket

$$\{q_i, q_j\} = \sum_i \left(\frac{\partial q_i}{\partial q_i} \frac{\partial q_j}{\partial p_i} - \frac{\partial q_i}{\partial p_i} \frac{\partial q_j}{\partial q_i} \right) = 0$$

$$\{p_i, p_j\} = \sum_i \left(\frac{\partial p_i}{\partial q_i} \frac{\partial p_j}{\partial p_i} - \frac{\partial p_i}{\partial p_i} \frac{\partial p_j}{\partial q_i} \right) = 0$$

$$\{q_i, p_j\} = \sum_i \left(\frac{\partial q_i}{\partial q_i} \frac{\partial p_j}{\partial p_i} - \frac{\partial q_i}{\partial p_i} \frac{\partial p_j}{\partial q_i} \right) = \delta_{ij}$$

Note that $(\mathbf{q}, \mathbf{p})$ are independent variables.

Canonical Transformation

In short, canonical transformation from (q, p) to $(\bar{q}, \bar{p})$ is a change of coordinates that preserves the canonical structure of Hamiltonian mechanics. By this requirement, they must satisfy the Hamiltonian equation

$$\dot{\bar{q}}_i = \frac{\partial \mathcal{H}}{\partial \bar{p}_i} \qquad \dot{\bar{p}}_i = -\frac{\partial \mathcal{H}}{\partial \bar{q}_i}$$

where $\mathcal{H} = \mathcal{H}(\bar{q}, \bar{p})$. Since the original coordinate is canonically conjugate, the new coordinate must also be canonically conjugate, which the conditions are given by

$$\{\bar{q}_i, \bar{q}_j\} = 0 \qquad \{\bar{p}_i, \bar{p}_j\} = 0 \qquad \{\bar{q}_i, \bar{p}_j\} = \delta_{ij}$$

Proof. We shall proof that the new transformed coordinate still obeys the Hamiltonian equation.

First note by performing the transformation, the Hamiltonian in Hamilton equation is expressed as a function of the new coordinate $\mathcal{H} = \mathcal{H}(\bar{\mathbf{q}}, \bar{\mathbf{p}})$. Its partial derivative with respect to both variable, then, may be expressed as

$$\frac{\partial \mathcal{H}}{\partial p_i} = \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial p_i} \right)$$

and

$$\frac{\partial \mathcal{H}}{\partial q_i} = \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial q_i} \right)$$

Next we begin the proof for the $\bar{q}_i$. Using Poisson bracket, we express its time evolution as

$$\dot{\bar{q}}_j = \{\bar{q}_j, \mathcal{H}\} = \sum_i \left(\frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right)$$

Inserting the expression for the partial derivative of the Hamiltonian

$$\dot{\bar{q}}_j = \sum_i \frac{\partial \bar{q}_j}{\partial q_i} \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial p_i} \right)$$

$$- \sum_i \frac{\partial \bar{q}_j}{\partial p_i} \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial q_i} \right)$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \bar{q}_k}{\partial p_i} + \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \bar{p}_k}{\partial p_i}$$

$$- \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \bar{q}_k}{\partial q_i} - \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \bar{p}_k}{\partial q_i}$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \left(\frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \bar{q}_k}{\partial p_i} - \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \bar{q}_k}{\partial q_i} \right)$$

$$+ \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \left(\frac{\partial \bar{q}_j}{\partial q_i} \frac{\partial \bar{p}_k}{\partial p_i} - \frac{\partial \bar{q}_j}{\partial p_i} \frac{\partial \bar{p}_k}{\partial q_i} \right)$$

$$= \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \{\bar{q}_j, \bar{q}_k\} + \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \{\bar{q}_j, \bar{p}_k\} \right)$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \delta_{jk}$$

$$\dot{\bar{q}}_j = \frac{\partial \mathcal{H}}{\partial \bar{p}_j}$$

Moving to the proof for the $\bar{p}_i$, which on expressing using Poisson bracket reads

$$\dot{\bar{p}}_j = \{\bar{p}_j, \mathcal{H}\} = \sum_i \left(\frac{\partial \bar{p}_j}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \bar{p}_j}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right)$$

With the same step as the $\bar{q}_i$, we write

$$\dot{\bar{p}}_j = \sum_i \frac{\partial \bar{p}_j}{\partial q_i} \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial p_i} \right)$$

$$- \sum_i \frac{\partial \bar{p}_j}{\partial p_i} \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \frac{\partial \bar{q}_k}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \frac{\partial \bar{p}_k}{\partial q_i} \right)$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \frac{\partial \bar{p}_j}{\partial q_i} \frac{\partial \bar{q}_k}{\partial p_i} + \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \frac{\partial \bar{p}_j}{\partial q_i} \frac{\partial \bar{p}_k}{\partial p_i}$$

$$- \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \frac{\partial \bar{p}_j}{\partial p_i} \frac{\partial \bar{q}_k}{\partial q_i} - \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \frac{\partial \bar{p}_j}{\partial p_i} \frac{\partial \bar{p}_k}{\partial q_i}$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{q}_k} \sum_i \left(\frac{\partial \bar{p}_j}{\partial q_i} \frac{\partial \bar{q}_k}{\partial p_i} - \frac{\partial \bar{p}_j}{\partial p_i} \frac{\partial \bar{q}_k}{\partial q_i} \right)$$

$$+ \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \sum_i \left(\frac{\partial \bar{p}_j}{\partial q_i} \frac{\partial \bar{p}_k}{\partial p_i} - \frac{\partial \bar{p}_j}{\partial p_i} \frac{\partial \bar{p}_k}{\partial q_i} \right)$$

$$= \sum_k \left(\frac{\partial \mathcal{H}}{\partial \bar{q}_k} \{ \bar{p}_j, \bar{q}_k \} + \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \{ \bar{p}_j, \bar{p}_k \} \right)$$

$$= \sum_k \frac{\partial \mathcal{H}}{\partial \bar{p}_k} \left(-\delta_{kj} \right)$$

$$\dot{\bar{p}}_j = - \frac{\partial \mathcal{H}}{\partial \bar{p}_j}$$

Application: Mass on a Cone

Consider a mass m which is constrained to move on the frictionless surface of a vertical cone $\rho = cz$.

Here we use cylindrical coordinate. The kinetic energy then is

$$T = \frac{1}{2}m\left(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2\right) = \frac{1}{2}m\left[(cz\dot{\phi})^2 + (c^2+1)\dot{z}^2\right]$$

while the potential energy is simply due to gravity

$$U = mgz$$

Thus the Hamiltonian read

$$\mathcal{H} = \frac{1}{2}m\left[(cz\dot{\phi})^2 + (c^2+1)\dot{z}^2\right] + mgz$$

Next we determine the generalized momenta

$$p_z = \frac{\partial \mathcal{L}}{\partial \dot{z}} = (c^2+1)m\dot{z} \qquad p_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = (c^2 z^2)m\dot{\phi}$$

Then we express the Hamiltonian as following

$$\mathcal{H} = \frac{1}{2}\left[\frac{(c^2+1)}{(c^2+1)^2 m^2}p_z^2 + +\frac{c^2 z^2}{m^2(c^2 z^2)^2}p_\phi^2\right] + mgz$$

$$\mathcal{H} = \frac{1}{2m}\left[\frac{p_z^2}{c^2+1} + \frac{p_\phi^2}{c^2 z^2}\right] + mgz$$

Finally, we have the equations for z component

$$\dot{z} = \frac{\partial \mathcal{H}}{\partial p_z} = \frac{p_z}{c^2+1} \qquad \dot{p}_z = -\frac{\partial \mathcal{H}}{\partial z} = \frac{p_\phi^2}{mc^2 z^3} - mg$$

and ϕ component

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial p_\phi} = \frac{p_\phi}{mc^2 z^2} \qquad \dot{p}_\phi = -\frac{\partial \mathcal{H}}{\partial \phi} = 0$$

Simple 1D Harmonic Oscillator

In this system, the kinetic energy is given by

$$T = \frac{1}{2}m\dot{x}^2 = \frac{p_x^2}{2m}$$

while the potential energy given by

$$U = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2 x^2$$

Thus the Hamiltonian is

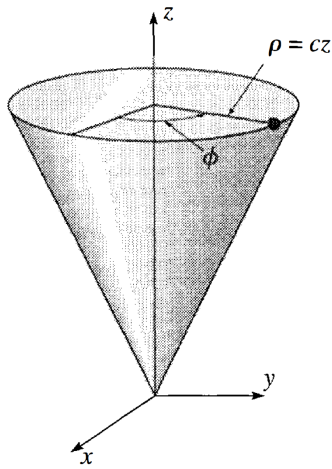$$\mathcal{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

Figure: Particle constrained to move within a cone

This gives the equation

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p_x} = \frac{p_x}{m} \qquad \dot{p}_x = \frac{\partial \mathcal{H}}{\partial x} = -m\omega^2 x$$

Solving the first equation for $\dot{p}_x$ then equating it with the second

$$\ddot{x} = -\omega^2 x$$

This is the equation for simple harmonic oscillator with the solution

$$x = A\cos(\omega t - \delta)$$

Modern Physics

Quantum Numbers

The energy eigenstates of the hydrogen atom are states of definite energy, magnitude of angular momentum, and z component of angular momentum. The values of those quantities are labeled by the quantum numbers n, l, and m_l, respectively. These three quantum numbers must be integers, and the spin quantum number m_s must be $-1/2$ (spin-up) or $-1/2$ (spin-down). Other restrictions are given below.

Letter	Name	Range	Physical interpretation
n	Principal quantum number	$1 \leq n < \infty$	$E_n = -(1/n^2)Ry$, with ($1\ Ry \approx 13.6$ eV)
l	Angular momentum quantum number	$0 \leq l \leq n-1$	$|\vec{L}| = \sqrt{l(l+1)}\hbar$
m_l	Magnetic quantum number	$-l \leq m_l \leq l$	$L_z = m_l\hbar$
m_s	Spin quantum number	$m_s = \pm 1/2$	$S_z = m_s\hbar$

Principal Quantum Number n. The term "principal quantum number" refers to the fact that n determines energy: each eigenstate ψ_{nlm_l} has energy $E_n = -(1/n2)$ Ry, where 1 Ry $\approx$ 13.6 eV is a unit of energy called a "rydberg." So the ground state of hydrogen has energy -1 Ry $\approx$ -13.6 eV, the first excited state has energy -(1/4) Ry $\approx$ -3.4 eV, and so on.

Angular Momentum Quantum Number l. l controls the magnitude of the total angular momentum of the electron about the nucleus is $L = \sqrt{l(l+1)}\hbar$.

Magnetic Quantum Number m_l. While l determines the magnitude of the angular momentum vector, m_l determine the z component of the angular momentum $L_z = m_l\hbar$. The name comes from the fact that the simplest way to measure a particular component of the angular momentum of a charged particle is to measure the magnetic field it creates.

Spin m_s. In 1924 Pauli suggested that each electron in an atom has a fourth quantum number that can only take on two possible values. He gave no physical explanation for this new electron property, which he called a "two-valuedness not describable classically." Consider a ball spinning about its own axis, except it is not a ball and it is not spinning.

For every electron, the magnitude of spin is $\sqrt{s(s+1)}\hbar = \sqrt{3/4}\hbar$, with values of $-1/2$ or $+1/2$. The z component of the spin is therefore $L_z = m_s\hbar = \pm(1/2)\hbar$.

Any particle with half-integer spin is a fermion, for example electrons, protons, and neutrons. Any particle with integer spin is a boson, for example photons. Every particle can be classified as either a "fermion," meaning it obeys the Pauli exclusion principle, or a "boson," meaning it does not.

The energy eigenvalues of a hydrogen atom depend only on n (to a very good approximation). The following formulas use m_e and e for
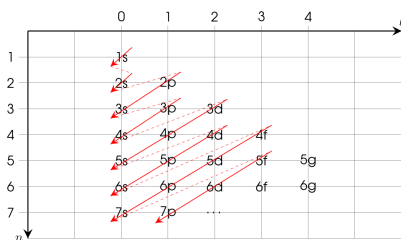
Figure: Perhabs the more familiar name for $n + l$ rule is Aufbau rule

the electron mass and charge, respectively

$$E_n = -\frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2} = -\frac{1}{2n^2} m_e c^2 \alpha^2 = -\frac{1}{n^2}\text{Ry}$$

$$= -\frac{1}{n^2} 13.6\text{eV} = \frac{1}{n^2} 2.18 \times 10^{-18}\text{J}$$

where α is fine structure constant

$$\alpha \equiv \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} \approx \frac{1}{137}$$

For a hydrogen-like atom with Z protons and one electron, the energies are

$$E_n - \frac{Z^2}{n^2}\text{Ry}$$

Energies of Different Quantum State

Pauli exclusion principle. Pauli exclusion principle forbid two or more of the same type of fermion in the same quantum state as each other. That means that, in the ground state of a multielectron atom, the Z electrons (fermion) fill up the Z lowest-energy states.

$n + l$ rule. Given two states ψ_{nl}, the subshell with the lower sum n + l will fill up first. In the case of a tie, the lower-n state fills first.

Hund's Rule. Because each electron can have spin up or spin down, two electrons can be in each combination of n, l, and m_l. That in turn means each subshell consists of $2(2l+1)$ degenerate states, with $2n^2$ degenerate state for each shell. Within a particular subshell, the electrons tend to end up in the states that minimize the energy level. Two nearby electrons whose spins are identical will be, on average, farther apart from each other than if they have opposite spins. The result is that electrons with identical spins shield each other from the nucleus less than electrons with different spins, so same-spin is a lower energy state than opposite-spin.

Energy Eigenstates

The energy eigenstates of the hydrogen atom are of the form

$$\psi_{nlm_l}(r, \theta, \phi) = R_{nl}(r) Y_l^{m_l}(\theta, \phi).$$

A full specification of the quantum state of a hydrogen atom also includes the spin state (up or down) of the electron.

Radial Wavefunction

The radial wavefunction is given by the following:

$$R_{nl}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} \left(\frac{2r}{na_0}\right)^l \exp\left(-\frac{r}{na_0}\right) L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right)$$

where L is called an associated Laguerre polynomial and is defined as

$$L_n^k(x) = (-1)^k \left(\frac{d}{dx}\right)^k \left[e^x \frac{d^{p+k}}{dx^{p+k}} \left(e^{-x} x^{p+k}\right)\right]$$

As you know, $k = 0$ then the above formula calls for the "zeroth derivative," or the function itself. $R_{nl}(r)$ is a polynomial in r times a decaying exponential $e^{-r/(a_0 n)}$, where a_0 is a constant called the "Bohr radius"

$$a_0 \equiv \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} \approx 5 \times 10^{-11} \text{ m}$$

Here's first few Radial Wavefunctions: for $n = 1$

$$R_{1,0} = 2a_0^{-3/2} \exp\left(-\frac{r}{a_0}\right)$$

for $n = 2$

$$R_{2,0} = \frac{1}{\sqrt{2}} a_0^{-3/2} \left(1 - \frac{r}{2a_0}\right) \exp\left(-\frac{r}{2a_0}\right)$$

$$R_{2,1} = \frac{1}{\sqrt{24}} a_0^{-3/2} \frac{r}{a_0} \exp\left(-\frac{r}{2a_0}\right)$$

for $n = 3$

$$R_{3,0} = \frac{2}{\sqrt{27}} a_0^{-3/2} \left[1 - \frac{2r}{3a_0} + \frac{2}{27}\left(\frac{r}{a_0}\right)^2\right] \exp\left(-\frac{r}{3a_0}\right)$$

$$R_{3,1} = \frac{8}{27\sqrt{6}} a_0^{-3/2} \left(1 - \frac{r}{6a_0}\right) \frac{r}{a_0} \exp\left(-\frac{r}{3a_0}\right)$$

$$R_{3,2} = \frac{4}{81\sqrt{30}} a_0^{-3/2} \left(\frac{r}{a_0}\right)^2 \exp\left(-\frac{r}{3a_0}\right)$$

Spherical Harmonics

Spherical harmonics are products of complex exponentials in ϕ and polynomials in $\cos\theta$:

$$Y_l^{m_l}(\theta, \psi) = \pm\sqrt{\frac{2l+1}{4\pi} \frac{(l-|m_l|)!}{(l+|m_l|)!}} P_l^{m_l}(\cos\theta) \exp(im_l\phi)$$

where l and m_l are integers and $-l \leq m_l \leq l$. The plus minus comes from $+1$ for all negative m_l and all even values of m_l, and -1 for odd positive values. So the only differences between $Y_l^{m_l}$ and and Y_l^{-ml} are that they have different signs if m_l is odd, and the exponent in

$e^{im_l \phi}$ changes sign between them. Also, P is an "associated Legendre polynomial"

$$P_l^{m_l}(x) \equiv (1 - x^2)^{|m_l|/2} \left(\frac{d}{dx}\right)^{|m_l|} \left[\frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l (x^2 - 1)^l\right]$$

Not to be confused by Legendre polynomials (defined by Rodrigues formula)

$$P_l(x) \equiv \frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l (x^2 - 1)^l$$

$Y_l^{m_l}$ is a polynomial in $\sin\theta$ and $\cos\theta$ multiplied by $e^{im_l \phi}$. Here's first few Spherical Harmonics: for $l = 0$

$$Y_0^0 = \sqrt{\frac{1}{4\pi}}$$

for $l = 1$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos\theta$$

$$Y_0^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\phi}$$

for $l = 2$

$$Y_2^0 = \sqrt{\frac{5}{16\pi}} (3\cos^2\theta - 1)$$

$$Y_2^0 = \mp\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta e^{\pm 2i\phi}$$

$$Y_2^0 = \sqrt{\frac{15}{31\pi}} \sin^2\theta e^{\pm 2i\phi}$$

Probability Distributions

For any particle with wavefunction Ψ in 3D, the probability per unit volume of finding the particle in a given region is $|\Psi|2$. For an electron in a hydrogen atom energy eigenstate ψ_{nlm_l} , probability of finding the electron between radii r_1 and r_2

$$P(r_1 \le r \le r_2) = \int_{r_1}^{r_2} r^2 R^2 \, dr$$

while the probability of finding the electron in a specified range of angles is

$$\int_{\theta_1}^{\theta_2} \int_{\phi_1}^{\phi_2} \sin\theta \left|Y_l^{m_l}\right|^2 \, d\psi d\theta$$

These integrals are normalized so that

$$\int_0^\infty r^2 R^2 \, dr = \int_0^\pi \int_0^{2\pi} \sin\theta \left|Y_l^{m_l}\right|^2 \, d\psi d\theta = 1$$

Zeeman Effect

The Zeeman effect refers to the changes in atomic spectra caused by an external magnetic field. In classical electrodynamics, magnets tend to align with magnetic fields, or magnet pointing along an external magnetic field has less energy than one pointing opposite to the external field. Because an electron is negatively charged, its magnetic moment points opposite to its angular momentum. So a positive m_l means its magnetic moment is pointing opposite external field, if we define positive z axis as the direction of external field. A negative m_l means the electron's magnetic field is pointing with the external field, which is the preferred (low-energy) orientation.

Therefore, in an external magnetic field pointing in the positive z direction, higher m_l leads to higher energy. For weak magnetic fields (neglecting spin), the Zeeman contribution to the energy of a state is

$$\Delta E = m_l \mu_B B$$

where μ_B is Bohr magneton

$$\mu_B = \frac{e\hbar}{2m_e}$$

If an atom transitions between states in which all its electron spins cancel, then you can ignore spin. In that case the splitting caused by a magnetic field is called the "normal Zeeman effect." in the presence of an external magnetic field, the normal Zeeman effect splits the $n = 2$ to $n = 1$ transition into three separate transitions based on Δm_l. The selection rule limits which electron transitions generally occur

$$\Delta l = \pm 1$$
$$\Delta m_l = -1, 0, 1$$

Quantum Mechanics

Unasorted Quantum Mechanics Topics

Braket notation

Ket. $|\psi\rangle$ represents quantum state. Written in matrix form as

$$|\psi\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_n \end{pmatrix}$$

Bra. $\langle\psi|$ is the Hermitian conjugate (complex conjugate transpose) of the ket $|\psi\rangle$

$$\langle\psi| = \begin{pmatrix} \psi_0 & \psi_1 & \cdots & \psi_n \end{pmatrix}$$

Inner Product. Written

$$\langle\phi|\psi\rangle = \begin{cases} 0, & \text{if orthogonal} \\ 1, & \text{if orthonormal} \end{cases}$$

Operator

Position Operator. Represents the position of a particle.

$$\hat{x} = x$$

Momentum Operator.

$$\hat{p} = -i\hbar\nabla$$

Energy Operator.

$$\hat{E} = i\hbar\frac{\partial}{\partial t}$$

Its action on the energy eigenstates is given by:

$$\langle\psi|\hat{E}|\psi\rangle = E_n$$

Hamiltonian Operator.

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(x)$$

The Hamiltonian can be written in terms of ladder operators as:

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)$$

Its action on the energy eigenstates $|n\rangle$ is given by:

$$H|n\rangle = E_n|n\rangle$$

where the energy eigenvalues are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

Creation operator. Increases the system's energy, thus often said to be raising operator. Defined as

$$a^\dagger = \frac{1}{\sqrt{2\hbar m\omega}} (m\omega x - ip)$$

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \dots & \dots \\ 0 & 0 & 0 & \dots & \sqrt{n} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Its action on the energy eigenstates $|n\rangle$ is given by:

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle$$

Annihilation operator. Decrease the system's energy, thus often said to be lowering operator. Defined as

$$a = \frac{1}{\sqrt{2\hbar m\omega}} (m\omega x + ip)$$

in matrix representation

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots & 0 & \dots \\ 0 & 0 & 0 & 0 & \ddots & \vdots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \sqrt{n} & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Its action on the energy eigenstates $|n\rangle$ is given by:

$$a |n\rangle = \sqrt{n} |n-1\rangle$$

Commutator

Commutator measures how much two physical quantities fail to be simultaneously measurable or well-defined. It is defined as

$$[A, B] = AB - BA$$

If $[A, B] = 0$, then A and B commute and can be simultaneously measured with arbitrary precision. If not, their measurement outcomes interfere with each other.

Expectation value

Braket notation.

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$$

Matrix notation.

$$\langle \hat{A} \rangle = \psi^\dagger \hat{A} \psi$$

Integral notation. If $\psi(x)$ is the wavefunction in the position representation

$$\langle \hat{A} \rangle = \int_{-\infty}^{\infty} \psi^*(x) \hat{A} \psi(x) \, dx$$

Normalization

Braket notation.

$$\langle \psi | \psi \rangle = 1$$

Integral notation. If $\psi(x)$ is the wavefunction in the position representation

$$\int_{-\infty}^{\infty} \psi^*(x) \hat{A} \psi(x) \, dx = 1$$

Normalization Problem

Ex 1. Find the value of A such that the following wavefunction particle inside potential well is normalized.

$$\psi = \frac{1}{\sqrt{10a}} \sin\left(\frac{\pi x}{a}\right) + A\frac{2}{a} \sin\left(\frac{2\pi x}{a}\right) + \frac{3}{\sqrt{5a}} \sin\left(\frac{3\pi x}{a}\right)$$

The wavefunction of said particle is written in the form

$$\psi_n = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{a}\right)$$

Hence we write the wavefunction as

$$\psi = \sqrt{\frac{1}{20}}\psi_1 + A\psi_2 + \sqrt{\frac{9}{10}}\psi_3$$

We then normalized the wavefunction by

$$\langle \psi | \psi \rangle = \left\langle \sqrt{\frac{1}{20}}\psi_1 + A\psi_2 + \sqrt{\frac{9}{10}}\psi_3 \middle| \sqrt{\frac{1}{20}}\psi_1 + A\psi_2 + \sqrt{\frac{9}{10}}\psi_3 \right\rangle$$

Since the wavefunction is orthonormal to itself and orthogonal to another, therefore

$$\langle \psi | \psi \rangle = \frac{1}{20} \langle \psi_1 | \psi_1 \rangle + A^2 \langle \psi_2 | \psi_2 \rangle + \frac{9}{10} \langle \psi_3 | \psi_3 \rangle = A^2 + \frac{19}{20}$$

Thus

$$A = \sqrt{\frac{1}{20}}$$

Expectation Value Problem

Ex 1. From the first normalization problem, find the expectation value of the energy. We have the normalized wavefunction

$$\psi = \sqrt{\frac{1}{20}} \sin\left(\frac{\pi x}{a}\right) + \sqrt{\frac{1}{20}} \sin\left(\frac{2\pi x}{a}\right) + \sqrt{\frac{9}{10}} \sin\left(\frac{3\pi x}{a}\right)$$

or simply

$$\psi = \sqrt{\frac{1}{20}}\psi_1 + \sqrt{\frac{1}{20}}\psi_2 + \sqrt{\frac{9}{10}}\psi_2$$

All that left is to do the algebra

$$\langle \psi | \hat{E} | \psi \rangle = \left\langle \sqrt{\frac{1}{20}}\psi_1 \middle| \hat{E} \middle| \sqrt{\frac{1}{20}}\psi_1 \right\rangle + \left\langle \sqrt{\frac{1}{20}}\psi_2 \middle| \hat{E} \middle| \sqrt{\frac{1}{20}}\psi_2 \right\rangle$$

$$+ \left\langle \sqrt{\frac{9}{10}}\psi_3 \middle| \hat{E} \middle| \sqrt{\frac{9}{10}}\psi_3 \right\rangle$$

Factoring the constant

$$\langle \psi | \hat{E} | \psi \rangle = \frac{1}{20} \left\langle \psi_1 \middle| \hat{E} \middle| \psi_1 \right\rangle + \frac{1}{20} \left\langle \psi_2 \middle| \hat{E} \middle| \psi_2 \right\rangle + \frac{9}{10} \left\langle \psi_3 \middle| \hat{E} \middle| \psi_3 \right\rangle$$

$$= \frac{1}{20} E_1 + \frac{1}{20} E_2 + \frac{9}{10} E_3 = \frac{1}{20}\frac{h^2}{8ma^2} + \frac{1}{20}\frac{4h^2}{8ma^2} + \frac{9}{10}\frac{9h^2}{8ma^2}$$

Quantum Mechanics Postulate

First Postulate

As the following

> The state of the particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space.

Compare it with the classical mechanics postulate via Hamiltonian formalism

> The state of a particle at any given time is specified by the two variables $x(t)$ and $p(t)$, i.e., as a point in a two-dimensional phase space.

Discussion: general discussion. The first three tell us how the system is depicted at a given time, and the fourth specifies how this picture changes with time. The first postulate states that a particle is described by a ket $|\psi\rangle$ in a Hilbert space which, you will recall, contains proper vectors normalizable to unity as well as improper vectors, normalizable only to the Dirac delta functions.

Now, a ket in a Hilbert space has in general an infinite number of components in a given basis. A classical particle has, at any given time, a definite position. One simply has to give this value of x in specifying the state. A quantum particle, on the other hand, can take on any value of x upon measurement and one must give the relative probabilities for all possible outcomes.

Second Postulate

As the following

> The independent variables x and p of classical mechanics are represented by Hermitian operators X and P with the following matrix elements in the eigenbasis of X

$$\langle x|X|x'\rangle = x\delta(x - x')$$
$$\langle x|P|x'\rangle = -i\hbar\delta'(x - x')$$

> The operators corresponding to dependent variables $\omega = \omega(x, p)$ are given Hermitian operators

$$\Omega(X, P) = \omega\left(x \to X, p \to P\right)$$

> where it is understood that Ω is a function of X and P just like ω is a function of x and p.

Compare it with the classical mechanics postulate via Hamiltonian formalism

> Every dynamical variable ω is a function of x and p, that is $\omega = \omega(x, p)$.

Discussion: ambiguity of operator. The definition $\Omega(X, P) = \omega(x \to X, p \to P)$ is ambiguous. Suppose $\omega = xp$. We do not know $\Omega = PX$ or $\Omega = XP$ since $px = xp$ classically. The rule is to use symmetric sum

$$\Omega = \frac{XP + PX}{2}$$

which also render Ω Hermitian.

Discussion: continuous spectrum of operator. In the case of continuous eigenvalues of Ω, ket $|\psi\rangle$ expands as

$$|\psi\rangle = \int |\omega\rangle \langle\omega|\psi\rangle \ d\omega = \int |\omega\rangle \psi(\omega) \ d\omega$$

Since ω varies continuously, so will $\langle\omega|\psi\rangle = \psi(\omega)$, which imply that $\psi(\omega)$ is a smooth function.

$P(\omega)$ in this case refer to the probability density, in particular $P(\omega) \ d\omega$ is the probability of obtaining ω between ω and $\omega + d\omega$. This definition must meet the unity definition

$$\int P(\omega) \ d\omega = \int \langle\psi|\omega\rangle \langle\omega|\psi\rangle \ d\omega = \langle\psi|I\psi\rangle = 1$$

Third Postulate

As the following

> If the particle is in a state $|\psi\rangle$, measurement of the variable corresponding to Ω will yield one of the eigenvalues ω with probability $P(\omega) \propto |\langle\omega|\psi\rangle|^2$. The state of the system will change from $|\psi\rangle$ to $|\omega\rangle$ as a result of the measurement.

Compare it with the classical mechanics postulate via Hamiltonian formalism

> If the particle is in a state given by x and p, the measurement of the variable ω will yield a value $\omega(x, p)$. The state will remain unaffected.

Discussion: probability. The theory makes only probabilistic predictions for the result of a measurement of Ω, which only have the possible values of its eigenvalues, which also all real by the postulate. The postulate also state the relative probabilities, to get the absolute probability, we divide by the sum of all relative probabilities

$$P(\omega_i) = \frac{|\langle\omega_i|\psi\rangle|^2}{\sum_j \langle\omega_j|\psi\rangle^2} = \frac{|\langle\omega_i|\psi\rangle|^2}{\langle\psi|\psi\rangle}$$

As a side note, the probability interpretation breaks down if $|\psi\rangle$ is an improper vector.

Suppose there exist state described by

$$|\psi\rangle = \frac{\alpha|\omega_1\rangle + |\omega_2\rangle}{(|\alpha|^2 + |\beta|^2)^{1/2}}$$

Measurement can either yield ω_1 or ω_2 with the probabilities $|\alpha^2|/(\alpha^2 + \beta^2)$ and $|\beta^2|/(\alpha^2 + \beta^2)$ respectively.

Discussion: probability for degenerate operator. Say that we have degenerate operator Ω with degenerate eigenvalues $\omega_1 = \omega_2 = \omega$ with orthonormal basis $|\omega, 1\rangle$ and $|\omega, 2\rangle$. Then the probability is

$$P(\omega) = |\langle \omega, 1|\psi \rangle|^2 + |\langle \omega, 2|\psi \rangle|^2$$

In terms of projection operator in this eigenspace

$$\mathbb{P}_\omega = |\omega, 1\rangle \langle \omega, 1| + |\omega, 1\rangle \langle \omega, 2|$$

we have

$$P(\omega) = \langle \psi|\mathbb{P}_\omega|\psi \rangle = \langle \mathbb{P}_\omega \psi|\mathbb{P}_\omega \psi \rangle$$

Discussion: change of basis. Suppose our interest switch to measuring Λ from the measurement of Ω–note that this does not mean a successive measurement. Recall that to obtain the probability of ω, we need to $|\psi\rangle$ in the Ω basis. To obtain the probability of λ, then, we can do the same; but there is no need to expand $|\psi\rangle$ in Λ given we already expanded $|\psi\rangle$ in Ω basis. Working in Ω basis, we have

$$|\psi\rangle = \sum_i |\omega_i\rangle \langle \omega_i|\psi \rangle \qquad \text{and} \quad P(\omega_i) = |\langle \omega_i|\psi \rangle|^2$$

If we wanted the probability of λ, we can simply project $|\psi\rangle$ in the eigenvectors $|\lambda\rangle$

$$\langle \lambda_i|\psi \rangle = \sum_j \langle \lambda_i|\omega_j \rangle \langle \omega_j|\psi \rangle$$

Fourth Postulate

As the following

> The state vector $|\psi\rangle$ obeys the Schrödinger equation
>
> $$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle$$
>
> where $H(X, P) = \mathcal{H}(x \to X, p \to P)$ is the quantum Hamiltonian operator and $\mathcal{H}$ is the Hamiltonian for the corresponding classical problem

Compare it with the classical mechanics postulate via Hamiltonian formalism

> The state variables change with time according to Hamilton's equations
>
> $$\dot{x} = \frac{\partial \mathcal{H}}{\partial p} \qquad \dot{p} = -\frac{\partial \mathcal{H}}{\partial x}$$

Measurement

Given state (x, p), we can say that dynamical value has the value $\omega(x, p)$ in classical mechanics. In quantum mechanics, when a particle is in a state $|\psi\rangle$ the particle can have the value of ω for the quantum operator Ω with probability $P(\omega) \propto |\langle \omega|\psi \rangle|^2$. This value is obtained by the following method.

1. Construct the corresponding quantum operator $\Omega = \omega(x \to, p \to P)$ where X and P are the operators defined in postulate II.

2. Find the orthonormal eigenvectors $|\omega\rangle$ and eigenvalues ω_i of Ω.

3. Expand $|\psi\rangle$ in this basis

$$|\psi\rangle = \sum_i |\omega_i\rangle \langle\omega_i|\psi\rangle$$

4. The probability $P(\omega_i)$ that the result ω_i will be obtained is $P(\omega) \propto |\langle\omega|\psi\rangle|^2$, or in terms of projection vector

$$P(\omega) \propto \langle\psi|\omega\rangle \langle\omega|\psi\rangle = \langle\psi|\mathbb{P}_\omega|\psi\rangle = \langle\psi|\mathbb{P}_\omega\mathbb{P}_\omega|\psi\rangle = \langle\mathbb{P}_\omega\psi|\mathbb{P}_\omega\psi\rangle$$

Collapse. Measurement by the operator Ω changes the state vector $|\psi\rangle$, which on this basis expands as

$$|\psi\rangle = \sum_\omega |\omega\rangle \langle\omega|\psi\rangle$$

into the eigenstate $|\omega\rangle$ corresponding to the eigenvalue of ω obtained in measurement.

The effect of the measurement, then, can be represented as

$$|\psi\rangle \xrightarrow{\Omega \text{ measured, } \omega \text{ obtained}} \frac{\mathbb{P}_\omega |\omega\rangle}{\langle\mathbb{P}_\omega\psi|\mathbb{P}_\omega\psi\rangle^{1/2}}$$

with $\mathbb{P}_\omega$ as the projection operator to $|\omega\rangle$. If the eigenvalues ω is degenerate, the $\mathbb{P}_\omega$ is the projection operator for the eigenspace $\mathbb{V}_\omega$.

Testing quantum theory. Unlike classical mechanics, quantum mechanics does not make deterministic predictions but statistical predictions. To test the predictions of the quantum mechanics, we must be able to create well-defined state $|\psi\rangle$ and check the probabilistic predictions at any time. The procedure are as follows.

1. Begin with particle in an arbitrary state $|\psi\rangle$ and measure a variable Ω.

2. If we get nondegenerate value ω, we have the state $|\omega\rangle$.

3. We immediately thereafter measure variable Λ next so that the state could not change form $|\omega\rangle$. If say the theory state

$$|\omega\rangle = \alpha |\lambda_1\rangle + \beta |\lambda_2\rangle$$

then the theory predict that λ_1 and λ_2 will be obtained with probability α^2 and β^2 respectively.

If the measurement give any other eigenvalue λ_i or other noneigenvalue value, that is the end of the theory. Suppose the resulting measurement give the correct eigenvalue, we can repeat the measurement with quantum ensemble of N particles all in this same state $|\omega\rangle$. The measurement should match with the predictions N/α^2 particle in $|\lambda_1\rangle$ and N/β^2 in $|\lambda_2\rangle$.

Information Inside Abstract State

Probabilities. The probability for obtaining ω for variable Ω is

$$P(\omega) = |\langle \omega|\psi\rangle|^2$$

or continuous eigenvalues, the probability ω between a and b.

$$P(a,b) = \int_a^b |\langle \omega|\psi\rangle|^2 \, d\omega$$

Expectation value. The expectation value is the mean value defined in statistics

$$\langle \Omega \rangle = \sum_i P(\omega_i)\omega_i = \langle \psi|\Omega|\psi\rangle$$

If the particle is an eigenstate of Ω, that is $\Omega|\psi\rangle = \omega|\psi\rangle$, the $\langle \Omega \rangle = \omega$.

To see how the third term comes to be, we can expand the statistical definition as follows.

$$\langle \Omega \rangle = \sum_i \langle \psi|\omega_i\rangle \langle \omega_i|\psi\rangle \, \omega_i = \sum_i \langle \psi|\, \omega_i \, |\omega_i\rangle \langle \omega_i|\psi\rangle = \sum_i \langle \psi|\Omega|\omega_i\rangle \langle \omega_i|\psi\rangle$$

$$\langle \Omega \rangle = \langle \psi|\Omega|\psi\rangle$$

where we have used the completion relation $\sum_i |\omega_i\rangle \langle \omega_i|$.

For continuous variable, we write

$$\langle \Omega \rangle = \int P(\omega)\omega \, d\omega = \int \langle \psi|\omega\rangle \, \omega \, \langle \omega|\psi\rangle \, d\omega$$

Uncertainty. The uncertainty is the standard deviation defined in statistics, which measure the average fluctuation around the mean

$$\Delta \Omega = \left(\sum_i P(\omega_i) \, (\omega_i - \langle \Omega \rangle)^2 \right)^{1/2} = \left\langle (\Omega - \langle \Omega \rangle)^2 \right\rangle^{1/2}$$

Another expression for the uncertainty is

$$\Delta \Omega = \left(\langle \Omega^2 \rangle - \langle \Omega \rangle^2 \right)^{1/2}$$

This expression is related to the definition of standard deviation by expanding the square term

$$\Delta \Omega = \left\langle (\Omega - \langle \Omega \rangle)^2 \right\rangle^{1/2} = \left\langle \Omega^2 + \langle \Omega \rangle^2 - 2\Omega \langle \Omega \rangle \right\rangle^{1/2}$$

$$= \left(\langle \Omega^2 \rangle + \langle \Omega \rangle^2 - 2 \langle \Omega \rangle^2 \right)^{1/2}$$

$$\Delta \Omega = \left(\langle \Omega^2 \rangle - \langle \Omega \rangle^2 \right)^{1/2}$$

As it the case with the expectation value, the uncertainty for continuous variable is

$$\Delta \Omega = \left(\int P(\omega) \left(\omega - \langle \Omega \rangle^2 \right)^2 \, d\omega \right)^{1/2}$$

Appendix: Measurement Example

State collapse example. Suppose there exist state described as

$$|\psi\rangle = \frac{1}{2}|\omega, 1\rangle + \frac{1}{2}|\psi\rangle + \sum_{i=3} c_i |\omega_i\rangle$$

It is also known that the orthonormal basis $|\omega, 1\rangle$ and $|\omega, 2\rangle$ is degenerate and so do their eigenvalues $\omega_1 = \omega_2 = \omega$. If measurement reveal the result ω, the collapsed state is written as

$$|\psi\rangle \xrightarrow{\omega \text{ obtained}} \frac{\mathbb{P}_\omega |\omega\rangle}{\langle \mathbb{P}_\omega \psi | \mathbb{P}_\omega \psi \rangle}^{1/2}$$

where the projection operator $\mathbb{P}_\omega$ for the eigenspace is written as

$$\mathbb{P}\omega = |\omega, 1\rangle \langle \omega, 1| + |\omega, 2\rangle \langle \omega, 2|$$

As such, its action on the state vector is

$$\mathbb{P}\omega |\psi\rangle = |\omega, 1\rangle \langle \omega, 1|\psi\rangle + |\omega, 2\rangle \langle \omega, 2|\psi\rangle = \frac{1}{2}|\omega, 1\rangle + \frac{1}{2}|\omega, 2\rangle$$

Therefore

$$|\psi\rangle = \frac{1}{(1/4 + 1/4)^{1/2}} \left(\frac{1}{2}|\omega, 1\rangle + \frac{1}{2}|\omega, 2\rangle \right) = \frac{1}{\sqrt{2}}|\omega, 1\rangle + \frac{1}{\sqrt{2}}|\omega, 2\rangle$$

Appendix: Exercise 4.2.1 from Shankar's Book

Consider the following operators on a Hilbert space $\mathbb{V}^3(C)$

$$L_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad L_y = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad L_z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

First question. What are the possible values one can obtain if L_z is measured?

Second question. Take the state in which $L_z = 1$. In this state what are $\langle L_x \rangle$, $\langle L_x^2 \rangle$ and $\langle \Delta L_x^2 \rangle$?

Third question. Find the normalized eigenstates and the eigenvalues of L_x in L_z basis.

Fourth question. If the particle is in the state with $L_z = -1$, and L_x is measured, what are the possible outcomes and their probabilities?

Fifth question. Consider the state, in bra to save space

$$\langle \psi | = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

in the L_z basis. If L_z^2 is measured in this state and a result $+1$ is obtained, what is the state after the measurement? How probable was this result? If L_z is measured, what are the outcomes and respective probabilities?

Sixth question. A particle is in a state for which the probabilities are $P(L_z = 1) = 1/4$, $P(L_z = O) = 1/2$, and $P(L_z = -1) = 1/4$. Convince yourself that the most general, normalized state with this property is

$$|\psi\rangle = \frac{e^{i\delta_1}}{2} |L_z = 1\rangle + \frac{e^{i\delta_2}}{\sqrt{2}} |L_z = 0\rangle + \frac{e^{i\delta_3}}{2} |L_z = -1\rangle$$

It was stated earlier on that if $|\psi\rangle$ is a normalized state then the state $e^{i\delta} |y\rangle$ is a physically equivalent normalized state. Does this mean that the factors $e^{i\delta}$ multiplying the L_z eigenstates are irrelevant?

Nuclear Physics

Introduction

Atomic Species

Characterized by the number of neutron N, number of proton Z, and mass number $A = N + Z$

$$(A, Z) = {}^A_Z X = {}^A_Z X_N$$

Nucleon

Defined as bound state of atomic nuclei. The two type are positively charged proton and neutral neutron. Nucleon constitutes three bound fermions called quark: up with charge $(2/3)$ and down with charge $(-1/3)$

$$\text{proton} = \text{uud}$$
$$\text{neutron} = \text{udd}$$

Both of them are fermion with mass

$$m_e = 939.56 \text{ MeV}/c^2$$
$$m_p = 938.27 \text{ MeV}/c^2$$
$$m_n - m_e = 1.29 \text{ MeV}/c^2$$

The magnetic moment projected by both are

$$\mu_p = 2.792847386 \ \mu_N \quad \mu_n = -1.91304275 \ \mu_N$$

where μ_N denote nuclear magneton

$$\mu_N = \frac{e\hbar}{2m_p} = 3.15245166 \ 10^{-14} \text{ MeV/T}$$

Here are the difference in unit used to describe nucleus compared to atom

Properties	Atom	Nucleus
Radius	Angstrom (10^{-10} m)	Femto (10^{-15} m)
Energy	eV	MeV

Radii. In terms of their mass number A, their radius may be approximated as
$$R = r_0 A^{1/3} \quad \text{with} \quad r_0 = 1.2 \text{ fm}$$
This approximation comes from assuming the radius is proportional to the volume which is also assumed to be spherical. Then $\mathcal{V} = 4\pi R^3/3 \approx A$.

Binding energy. Defined as the difference of the sum of nuclei mass and the nuclear mass

$$B(A, Z) = N m_n c^2 + Z m_p c^2 - m(A, Z) c^2$$

Mass. Three unit most common are atomic mass unit (u), the kilogram (kg), and the electron-volt (eV). The atomic mass unit is defined as the mass of ^{12}C atom divided by 12

$$1 \text{ u} = \frac{m(^{12}C)}{12}$$

electron volt is defined as the kinetic energy of an electron after being accelerated from rest through a potential difference of 1 V.

Nuclear Relative

Isotope. Same number of charge Z, but different number of neutron N. Isotope has identical chemical properties, since they have the same electron, but different nuclear properties. Example are

$$^{238}_{92}U \quad \text{and} \quad ^{235}_{92}U$$
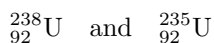
Isobar. Same mass A. Frequently have the same nuclear properties due to the same number of nucleon. Example are

$$^{3}He \quad \text{and} \quad ^{3}H$$
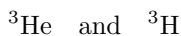
Isotone. Same number of neutron N, but different number of proton Z. Example are

$$^{14}C_6 \quad \text{and} \quad ^{16}O_8$$

Radioactivity

Nuclear Decay

Alpha decay. Occur when parent nucleon decay distribute among daughter nuclei

$$^{A}_{Z}\text{P} \rightarrow ^{A-4}_{Z-2}\text{D} + ^{4}_{2}\text{He}$$
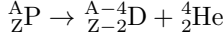
The Q value, which is defined as the total released in a given nuclear decay, of alpha decay is

$$Q = \left[m\left(^{A}_{Z}\text{P}\right) - m(^{A-4}_{Z-2}\text{D}) - m\left(^{4}_{2}\text{He}\right)\right] c^2$$

$$Q = K_D + K_\alpha = \frac{A}{A-4} K_\alpha$$

Q can be determined by applying the energy conservation into given nuclear reaction.

**Beta decay.** Two example are electron capture and neutrino capture

$$e^- + p \leftrightarrow \nu_e + n$$
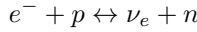
Other interaction can be found by moving particle to different side and changing them to their anti particle, such as
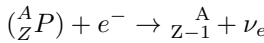
$$\bar{\nu}_e + p \leftrightarrow e^+ + n$$

where the anti particle of electron e^- and electron neutrino ν_e are positron e^+ and electron antineutrino $\bar{\nu}_e$. Another example is beta negative decay

$$\left(^{A}_{Z}P\right) \rightarrow ^{A}_{Z+1}\text{D} + e^- + \bar{\nu}_e$$
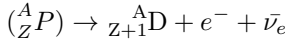
which is the neutron decay inside isotope, and beta positive decay

$$\left(^{A}_{Z}P\right) \rightarrow ^{A}_{Z-1}\text{D} + e^- + \nu_e$$
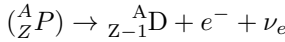
which is the proton decay inside isotope. Electron capture inside isotope takes the form of

$$\left(^{A}_{Z}P\right) + e^- \rightarrow ^{A}_{Z-1} + \nu_e$$

The Q value of beta negative is

$$Q = \left[m\left(\left(^{A}_{Z}P\right)\right) - m(^{A}_{Z-1}\text{D}) + m_e + m_{\nu_e}\right]$$

and beta negative decay

$$Q = \left[m\left(\left(^{A}_{Z}P\right)\right) - m(^{A}_{Z-1}\text{D}) + m_e + m_{\bar{\nu}_e}\right]$$

Radiation

Radioactive decay. The number of decay events over a time inteval is proportional to the number of particle

$$-\frac{dN}{dt} = \lambda N$$

This first order ODE is solved by integration

$$-\int_{N_0}^{N} \frac{1}{N} \, dN = \int_0^t \lambda t \, dt$$

$$\ln \frac{N}{N_0} = -\lambda t$$

$$N(t) = N_0 e^{-\lambda t}$$

Now consider the case of chain radiation, that is the particle undergoes decay $N_A \to N_B \to N_C$. For the first decay of particles N_A into N_B

$$\frac{dN_A}{dt} = -\lambda_A N_A \implies N_A = N_{A0} e^{-\lambda_A t}$$

Now the total rate of creation for the second particle N_B is the sum of the decay of said particle and the decay of the first particle into the second particle.

$$\frac{dN_B}{dt} = -\lambda_B N_B + \lambda_A N_A$$

$$\frac{dN_B}{dt} + \lambda_B N_B = \lambda_A N_{A0} e^{-\lambda_A t}$$

This ODE is solved by integrating factor method

$$I = \int \lambda_B \, dt = \lambda_B t$$

Then

$$N_B(t) = e^{-\lambda_B t} \int \lambda_A N_{A0} e^{-\lambda_A t} \, dt + C e^{-\lambda_B t}$$

$$N_B(t) = \frac{\lambda_A N_{A0}}{\lambda_B - \lambda_1} e^{-\lambda_A t} + C e^{-\lambda_B t}$$

If we assume at $t = 0$ we have zero second particle, then

$$N_B(0) = \frac{\lambda_A N_{A0}}{\lambda_B - \lambda_1} + C = 0 \implies C = -\frac{\lambda_A N_{A0}}{\lambda_B - \lambda_1}$$

Thus, the complete solution is

$$N_B(t) = \frac{\lambda_A N_{A0}}{\lambda_B - \lambda_1} \left(e^{\lambda_A t} - e^{\lambda_B t} \right)$$

In practice, often we are not given the number of particle, but rather the mol n or mass m (in gram) of said particle. Those quantities are related by

$$n = \frac{m}{M_r} \quad \text{or} \quad n = \frac{N}{N_A}$$

where $M_r \approx A$ is the molecular mass (gram/mol) and $N_A = 6.02 \cdot 10^{23}$ (mol^{-1}) is the Avogadro constant.

Halflife. Defined as the time required for particles to reduce to half of its initial value. The value of halflife can be determined by considering $N(t)$ at $t_{1/2}$, which by definition

$$\frac{1}{2}N_0 = N_0 e^{-\lambda t_{1/2}}$$

$$t_{1/2} = \frac{\ln 2}{\lambda}$$

This equation can also be used to determine the decay constant.

Activity. Defined as the number of radioactive transformations per second

$$A \equiv -\frac{dN}{dt} = \lambda N$$

Using the solution for N, we can write

$$A = \lambda N_0 e^{-\lambda t} = A_0 e^{-\lambda t}$$

Specific Activity. Quantity related to activity; specific activity is the activity per unit mass

$$a \equiv \frac{A}{m}$$

Using the relation of mass with mol and halflife relation

$$a = \frac{\lambda N}{\frac{N}{N_A} M_r} = \frac{N_A \lambda}{M_r} = \frac{\ln 2 N_A}{t_{1/2} M_r}$$

On evaluating the numerator constant

$$a = \frac{1.32 \cdot 10^{16}}{t_{1/2} M}$$

Nuclear Stability

Valley of stability. Consist of long-lived isotope that do not simultaneously decay.

Below the valley of stability. Consist of isotopes with more N than those of the valley of stability, thus they decay either by beta negative, more likely, or neutron decay, less likely.

Above the valley of stability. Consist of isotope with more Z than those of the valley of stability, thus it decays either by beta positive or electron capture.

Beyond the valley of stability. Consist of heavy isotope with $Z > 83$, $N > 126$, and $A > 209$. They decay with alpha radiation.
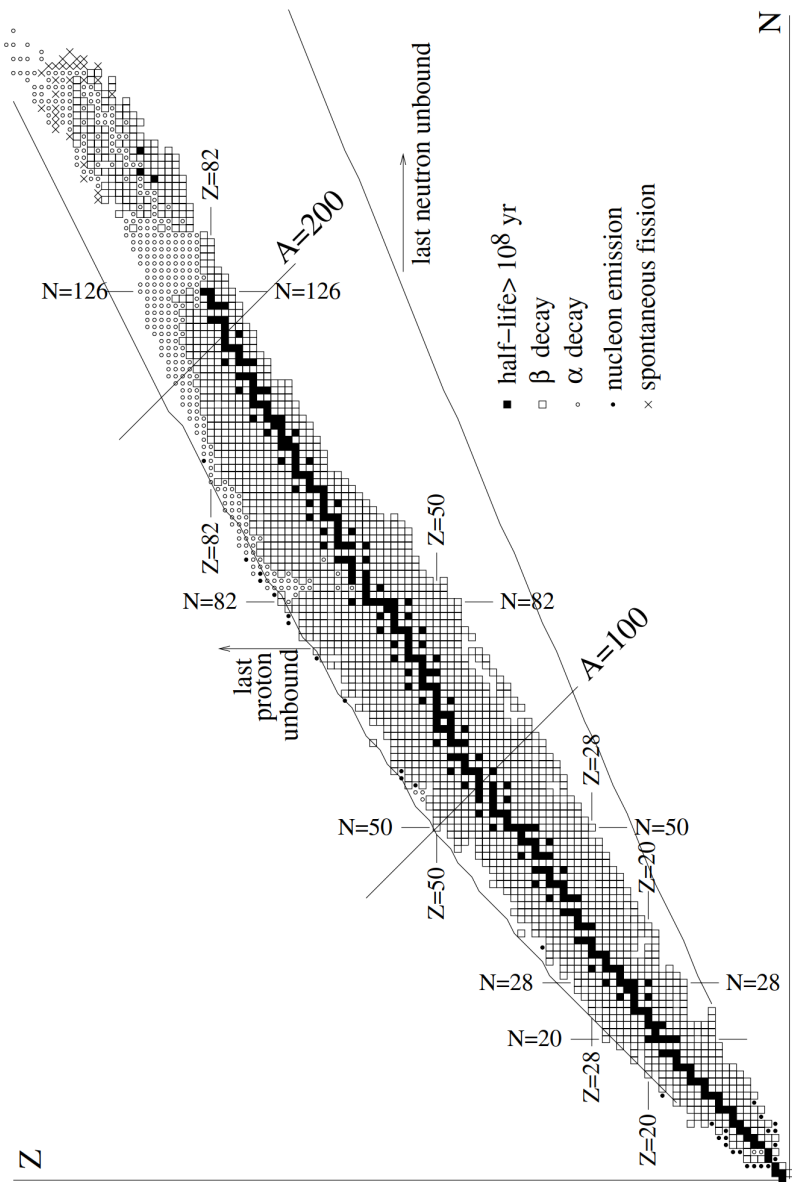
Figure: Valley of stability in ZN graph.

Particle Physics

Elementary Particle

Particle can be categorized by their spin, mass, and the type of interaction.

Boson. A pseudo-particle with interger spin that mediate the interaction. Two types of boson are gauge, or force carrier, and scalar, whose function is to give particle mass. Also, $W^\pm$ and Z^0 boson both belong to the same isospin triplet with $I = 1$

Table: Boson properties

Gauge Boson	
Family	Function/Force
Photon γ	Electromagnetic
Gluon g	Strong: binds the quark
W Boson $W^\pm$	Weak:radioactive decay
Z Boson Z^0	Weak: same as above

Scalar Boson	
Family	Function/Force
Higgs Boson H^0	Gives particle mass

Fermion. Consist of quark and lepton. All has half interger spin. Bound state of quark is called hadron and quark is able to experience all four fundamental forces.

Lepton experiences the weak force, however for charged lepton is also able to experience electromagnetic forces. This is why neutral lepton, that is neutrino, is hard to detect.

Hadron. Two types of hadron are baryon, which consist of three quarks, and meson, which consist of two quarks and one antiquark. Since baryon is made up of three spin-1/2, its spin is always half interger, opposite with meson; thus, both are also able to experience all four fundamental forces. Anti baryon consist of antiquark with the same configuration, for example is anti proton $\bar{p} = \bar{u}\bar{u}\bar{d}$.

One difference is that, unlike baryon, meson do not follow Pauli exclusion principle, since its total spin is interger. Most mesons are not their own antiparticle: charged mesons like K^+ with K^-; and neutral meson like K^0 with $\bar{K}^0$.

Quantum Number. Summarized as follows. First we have the fundamental particles.

Table: fundamental particle properties

Boson						
Name	q	L	B	S	I	I_z
γ	0	0	0	0	0	0
g	0	0	0	0	0	0
W^+	+1	0	0	0	1	+1
Z^0	0	0	0	0	1	0
W^-	-1	0	0	0	1	-1
H^0	0	0	0	0	0	0

Quark						
Name	q	L	B	S	I	I_z
u	+2/3	0	+1/3	0	1/2	+1/2
d	-1/3	0	+1/3	0	1/2	-1/2
s	-1/3	0	+1/3	-1	0	0
c	+2/3	0	+1/3	0	0	0
b	-1/3	0	+1/3	0	0	0
t	+2/3	0	+1/3	0	0	0

Lepton						
Name	q	L	B	S	I	I_z
e	-1	$L_e = 1$	0	0	1/2	+1/2
μ	-1	$L_\mu = 1$	0	0	1/2	-1/2
τ	-1	$L_\tau = 1$	0	0	0	0
ν_e	0	$L_e = 1$	0	0	0	0
ν_μ	0	$L_\mu = 1$	0	0	0	0
ν_τ	0	$L_\tau = 1$	0	0	0	0

Then the Hadron.

Table: Hadron properties

Baryon							
Name	Content	q	L	B	S	I	I_z
p	uud	+1	0	+1	0	1/2	+1/2
n	udd	0	0	+1	0	1/2	-1/2
Λ^0	uds	0	0	+1	-1	I	0
Σ^+	uus	+1	0	+1	-1	1	+1
Σ^0	uds	0	0	+1	-1	1	0
Σ^-	dds	-1	0	+1	-1	1	-1
Ξ^0	uss	0	0	+1	-2	1/2	+1/2
Ξ^-	dss	-1	0	+1	-2	1/2	-1/2
Ω^-	sss	-1	0	+1	-3	0	0

Meson							
Name	Content	q	L	B	S	I	I_z
π^+	$u\bar{d}$	+1	0	0		1	+1
π^0	$u\bar{u}, d\bar{d}$	0	0	0	0	1	0
π^-	$d\bar{u}$	-1	0	0	0	1	-1
K^+	$d\bar{s}$	+1	0	0	+1	1/2	+1/2
K^0	$d\bar{s}$	0	0	0	+1	1/2	-1/2

$\bar{K}^0$	$s\bar{d}$	0	0	0	-1	1/2	+1/2
K^-	$s\bar{u}$	-1	0	0	-1	1/2	-1/2
η	$u\bar{u}, d\bar{d}, s\bar{s}$	0	0	0	0	0	0
η'	$u\bar{u}, d\bar{d}, s\bar{s}$	0	0	0	0	0	0

Conservation Laws

Energy. The total energy $E = K + mc^2$ must be conserved in all types of nuclear reaction.

Momentum. Same as energy conservation law.

Mass number. Same as energy conservation law.

Charge. Same as energy conservation law.

Lepton number. We assign lepton a lepton number $L = +1$, anti lepton $L = -1$, and non lepton $L = 0$. Each family of lepton–such as electron e, muon μ, tau τ and their neutrino sibling–has separate conserved neutrino number–L_e, L_μ, L_τ repectively.
This quantum number almost conserved in all reaction, exception exist in neutrino oscillations; however only family lepton number is violated, while the total lepton number is conserved.

Baryon number. Like before, we assign baryon a baryon number $L = +1$, anti baryon $L = -1$, and non baryon $L = 0$. As an aside, baryon is the defined as the bound state of three quarke and that strong force works on all of them.

Strangeness number. Defined as the negative of the number of strange quarks in it, in particular strange quark s has $S = -1$ and antistrange quark $\bar{s}$ has $S = +1$.

Isospin (Isotropic Spin). We define isospin I such that $2I + 1$ is equal to the multiplet type of the baryon. Recall that the strong force does not differentiate between proton and nucleon, hence both particle can be defined as the different state of the same particle and thus can be categorized as doublet. To differentiate them, then, we define proton with $I_z = 1/2$ and neutron with $I_z = -1/2$. In general,

$$I_z = I, I - 1, \ldots, -I$$

Isospin is conserved in strong force interaction, but may not in other interaction.

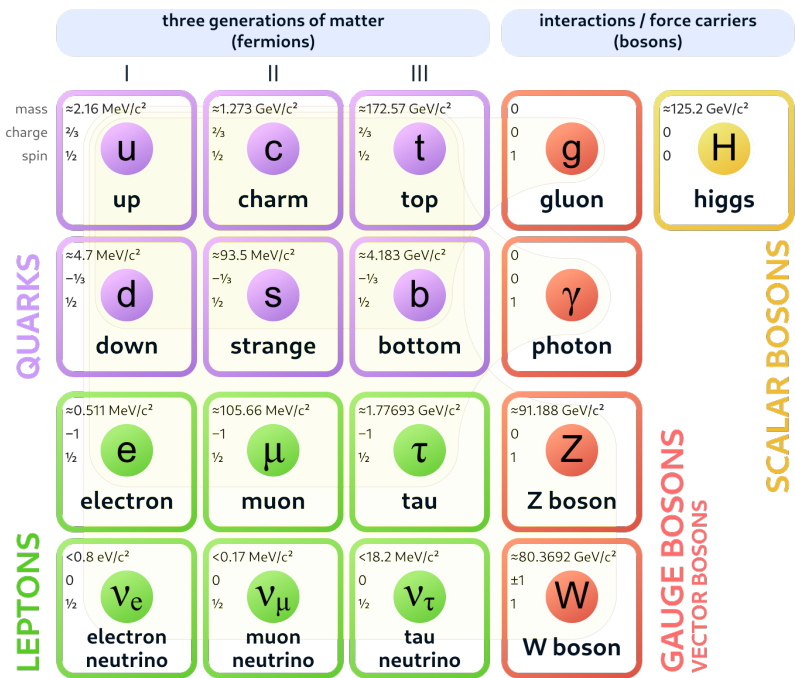# Standard Model of Elementary Particles

Figure: Standard model.

Nuclear Model

Liquid Drop Model

The binding energy by this model is given by

$$B(A, Z) = a_V A - a_s A^{2/3} - a_C \frac{Z^2}{A^{1/3}} - a_a \frac{(N - Z)^2}{A} + \delta(A)$$

where

$$a_V = 15.753$$
$$a_s = 17.804$$
$$a_C = 0.7103$$
$$a_a = 23.69$$

$$\delta(A) = \begin{cases} 33.6A^{-3/4} & \text{if } N \text{ and } Z \text{ are even} \\ -33.6A^{-3/4} & \text{if } N \text{ and } Z \text{ are odd} \\ 0 & \text{if } A = N + Z \text{ is odd} \end{cases}$$

Volume term. Recall that the volume of nucleon is proportional to A; on using this we have obtained $r = r_0 A^{1/3}$, which means that nuclei have constant density, like a drop of water.

Experiment shows that the binding energy per nucleon is rougly constant, $B/A \approx 8$ MeV. The overshoot of the volem term, then, require corection that will lower the value of the binding energy.

Surface term. Like water molecule, internal nucleon experience isotropic force, while surface nucleon only from inside. This resulting the force, and consequentl the energy, to be proportional to area $4\pi r^2 \approx a_s A^{1/3}$, with $r = r_0 A^{1/3}$.

Columb term. The binding energy due to charged particle is proportional to $Q^2/R \approx Z^2/A^{1/3}$.

Asymmetry term. By The Pauli exclusion principle, the configuration with different N and Z will have more energy than that with the same due to the different nucleon will simply occupy the higher energy state. This is the basis of $N - Z$ term.

Quantum pairing term. This term captures the effect of spin coupling. Odd-odd nuclei tend to undergo beta decay to an adjacent even-even nucleus by changing an N to a Z or vice versa.

Stable nuclei. By setting the derivative of B with respect to Z to zero, we have the maximum binding energy, that is the most state nuclei. If we ignore the quantum pairing term, or we are considering the case of odd A,we have

$$Z(A) = \frac{A}{2 + a_c A^{2/3}/2a_a} \approx \frac{A/2}{1 + 0.0075 \, A^{2/3}}$$

Bibliography

[1] Jean-Louis Basdevant, James Rich, and Michael Spiro. *Fundamentals in nuclear physics: from nuclear structure to cosmology*. eng. Softcover version of original hardcover edition 2005. New York, NY: Springer, 2010. ISBN: 978-1-4419-1849-9.

[2] Mary L. Boas. *Mathematical Methods in the Physical Sciences*. Washington: Wiley, 2005.

[3] Ramamurti Shankar. *Principles of quantum mechanics*. eng. Second edition, corrected at 19th printing. New York, NY: Springer, 2014. ISBN: 978-0-306-44790-7.