Physics Notes

Release 0.1.2

Lei Ma

Mar 24, 2018

Contents

This is a notebook on physics. WIP.

Use the source. Keep the source open.

CHAPTER 1

Whatever

I alway believe that making things open to everyone is one of the most powerful things that drive the world forward. So I DO think open source and open data, even open education are reforming the world.

This was a notebook for myself when I was studying at Fudan University. At that time, I learned how to use LaTeX and I was so excited. So I thought I should start writing something using LaTeX since it's so beautiful. Well, the bad thing is, I just randomly wrote down my notes on some specific topics.

I was so greedy back then. Then I tried to build up my own framework of physics by writing notes here. It never did the work by the way. Then I realized a framework should be somethng organized much better than this one. (I should draw a map of physics.)

Though these notes didn't help me building up my framework of phyiscs, I learned an important lesson. A physicist should build up his/her own style: the way to think, the way to solve problems, the way to check answers, the way to write, etc.

Anyway, I got frastrated and gave up the effort to utilize it as a framework-building thing. However I won't just dump these notes. As I have more and more to add, I think I'll just let it be my notebook, which, of course, is open source and accessible to everyone.

Yes. Use the source. Keep the source open.

CHAPTER 2

Fun

Here are some interesting physics related problems.

2.1 Math4Fun

2.1.1 A funny expression for gradient

$$\odot\nabla\oslash + \oslash\nabla\odot = \nabla(\odot\cdot\oslash)$$

2.2 Quantum

Why Rutherford was wrong about the atom model?

Link to this: *rutherford-atom*

A Rutherford atom model is the combination of protons and electrons to give us a neutral nucleus.

Question 1: What is the energy of the electron if we confine it inside the nucleus? Consider only the kinetic energy due to uncertainty principle is enough to construct a contradiction.

Question 2: From the point of view of nuclear magnetic moment, the electron magnetic moment is way to large. Measurement tells us that nuclei usually have nuclear magnetic moment $-3\mu_N$ to $10\mu_N$ where $\mu_N = e\hbar/2m_p$ and m_p is the mass of proton.

Hint: The magnetic moment is of the magnitude $e\hbar/2m_x$.

2.3 Relativity

Blazers in astrophysics

Blazers has a very large apparant velocity, which is usually much larger than speed of light, eg, 34c.

This problem is very nicely explained on this page: Apparent Superluminal Velocity of Galaxies .

Fig. 2.1: This is taken form the link metioned above. We measure the distance at 1 Jan then at 1 Feb. The apparent velocity would be the travelled distance divided by 1 month. However, the first measurement only measured the light from a place that is $1 + d/c$ further than the second spot due the the fact that light travels at a finite speed. So the distance we measured is larger than the actual distance at 1 Jan and 1 Feb. Thus leading to a apparent larger velocity and this velocity can exceed the limit of light speed.

2.4 Electrodynamics

2.4.1 Electrodynamics in 2+1 Spacetime

Maxwell's equations are mostly experiment determined, except for one term by Maxwell involves the induced current. The only hope to write down a real 2+1 electrodynamcs formalism is to really understand the most fundamental properties of electrodynamics which I don't have at this moment.

So I turned to another approach. First of all we need to reach some basic agreement that which is not changed from our 3+1 theory to a 2+1 theory. As this being said, there could be a bunch of different versions of 2+1 theory.

To make sure we have a consistant theory, the following terms should be applied.

1. Something should be unchanged which will act as a connection between our 2+1 theory and 3+1 theory.

Maxwell's Theory

The equations could be written as

$$\partial_\mu F^{\mu\nu} = 0,$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

However we would like to check the four laws independently since we really need to look into the meaning of the equations.

$$\begin{aligned}
\partial_i E_i &= 4\pi\rho \\
\partial_i B_i &= 0 \\
\epsilon_{ijk}\partial_j E_k &= -\frac{1}{c}\partial_t B_i \\
\epsilon_{ijk}\partial_j B_k &= \frac{4\pi}{c} J_i + \frac{1}{c}\partial_t E_i.
\end{aligned}$$

Here we write down the component form because the cross product doesn't have a clear meaning as we move to different dimensions.

The first naive version

Assumptions

1. The dimension of energy is not changed.
2. The dimension of length and time are kept.

Gauss's Law

Gauss's law shows the source of the electric field, which should be in the form

$$\oint \vec{E} \cdot \vec{dl} = 2\pi \int \rho dS.$$

We have 2π instead of 4π is because we have only a integral of a closed loop not a closed surface.

Applying Stokes Theorem?

At first thought, we need to math the integral on the two sides thus Stokes theorem should be applied.

Surprisingly, we don't really get to the familiar Gauss's law of differential form. Instead, we have

$$\int\int \nabla \times \vec{E} \cdot \vec{dS} = 2\pi \int \rho dS.$$

BUT think about this. Is this really true? We DO NOT have a third dimension! How could we define a curl? Back to the component form,

$$\nabla \times \vec{E} = \hat{e_i} \epsilon_{ijk} \partial_j E_k.$$

As a reminder,

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is } (1,2,3), (2,3,1) \text{ or } (3,1,2), \\ -1 & \text{if } (i,j,k) \text{ is } (3,2,1), (1,3,2) \text{ or } (2,1,3), \\ 0 & \text{if } i=j \text{ or } j=k \text{ or } k=i \end{cases}$$

Now the problem is we have all the elements of this Levi-Civita symbol 0 because only 2 dimensions can be used in this theory.

That means we have no Gaussian theorem or no charge as a naive interpretation if we follow our idea that charge is source of static curl free electric field and followed up by using Stokes theorem.

This argument is WRONG. We need to reconsider the meaning of equations. This is 2D we don't have a third dimension to use Stokes theorem. We need divergence theorem.

Two match up the dimensions we do need to apply the divergence theorem, in 2D.

$$\oint \vec{E} \cdot \vec{dl} = \int\int \partial_i E_i dS,$$

from which we are able to determine the differential form

$$\partial_i E_i = 2\pi \rho,$$

in which we have $i = 1, 2$.

Vectors in 2D

TBD

Faraday's Law

The change of magnetic flux generate electric field,

$$\oint \vec{E} \cdot \vec{dl} = -\frac{1}{c}\frac{d}{dt}\int \vec{B}d\vec{S}.$$

Wave

We still have a wave solution.

Refs

1. Electrodynamics in 2D by Kirk T. McDonald @ Princeton

CHAPTER 3

Vocabulary

3.1 Useful Math Tricks

3.1.1 Functional Derivative

By definition,[1] functional derivative of a functional $G[f]$ with respect to f along the 'direction' of h is

$$\delta G[f][h] = \frac{d}{d\epsilon} G[f + \epsilon h]|_{\epsilon=0}.$$

As an example, the functional derivative of $G[f] = \int dx f^n(x)$ $\delta G[f][h]$ is

$$\begin{aligned}\delta G[f][h] &= \frac{d}{d\epsilon} G[f + \epsilon h]|_{\epsilon=0} \\ &= \int n f^{n-1}(x) h(x) dx.\end{aligned}$$

Now the problem appears. We have an unknown function h which makes sense because we haven't specify a direction of the derivative yet.

For a physicist, the savior of integral is Dirac delta. So we use delta distribution as the direction in the functional derivative of action which is an integral,

$$\frac{\delta G[f]}{\delta f(y)} = \delta G[f][\delta_y].$$

It can be ambiguous to just write down δ_y without an example. Here is the previous example continued,

$$\begin{aligned}\frac{\delta G[f(y)]}{\delta f(y)} &= \int n f^{n-1}(x) h(x) dx|_{h(x)=\delta(x-y)} \\ &= \int n f^{n-1}(x) \delta(x-y) dx \\ &= n f^{n-1}(y).\end{aligned}$$

[1] Chapter 15 of Physical Mathematics

It seems that we can just think of f as a variable then take the ordinary derivative with respect to it. It is NOT true. Consider such a functional $G[f] = \int (f'(x))^2 dx$ where ' means the derivative of $f(x)$.

$$\frac{\delta G[f]}{\delta f} = -\int dx 2f''(x)h(x)|_{h(x)=\delta(x-y)}$$
$$= -2f''(y),$$

which is not that straightforward to understand from function derivatives.

3.1.2 Legendre Transformation

Legendre transformation is NOT just some algebra. Given $f(x)$ as a function of x, which is shown in blue, we could find the distance between a line $y = px_i$ and the function value $f(x_i)$.

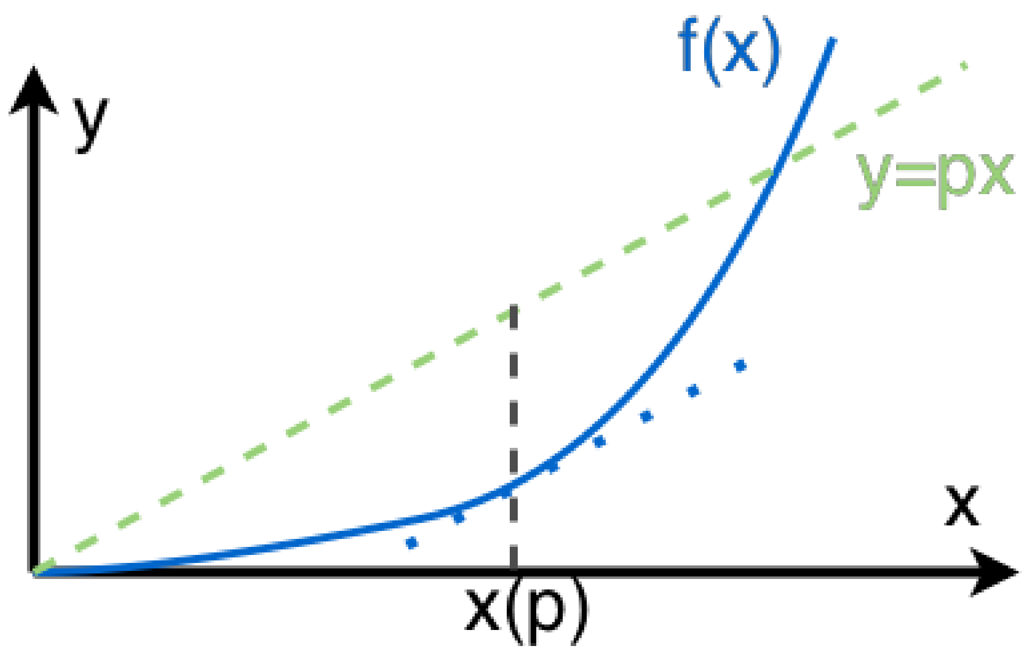


Fig. 3.1: Meaning of Legendre transformation

However, as we didn't fix x, this means that the distance

$$F(p, x) = px - f(x)$$

varies according to x. This is a transformation that maps a function $f(x)$ to some other function $F(p, x)$ which depends on the parameter p. A more pedagogical way of writing this is

$$px = F(x, p) + f(x).$$

To have a Legendre transformation, let's choose a relation between x and p. One choice is to make sure we have a maximum distance given p, which means the x we choose is the point that makes the slope of $f(x)$ the same as the line $y = px$. In the language of math, the condition we require is

$$0 = \frac{\partial F(p, x)}{\partial x} \equiv f'(x) - p,$$

which indeed shows that the slope of function and slope of the straight line match eath other at the specified point. Thus we have a relation between x and p.

Substitute $x(p)$ back into $F(p, x)$, we will get the Legendre transformation $F(p, x(p))$ of $f(x)$.

Back to the math we learned in undergrad study. A Legendre transformation transforms a function of x to another function with variable $\frac{f(x)}{x}$. Using $f(x)$ and its Legendre transformation $F(p = px - f(x(p))$ as an example, we can show that the slope of $F(p)$ is x,

$$\frac{dF(p)}{dp} = x,$$

which is intriging because the slope of $f(x)$ is p in our requirement. We removed the dependence of x in $F(p)$ because we have this extra constrain.

Let's Move to Another Level

We require the function $f(x)$ is convex (second order derivative is not negative). This is required because otherwise we would NOT have a one on one mapping of x and p.

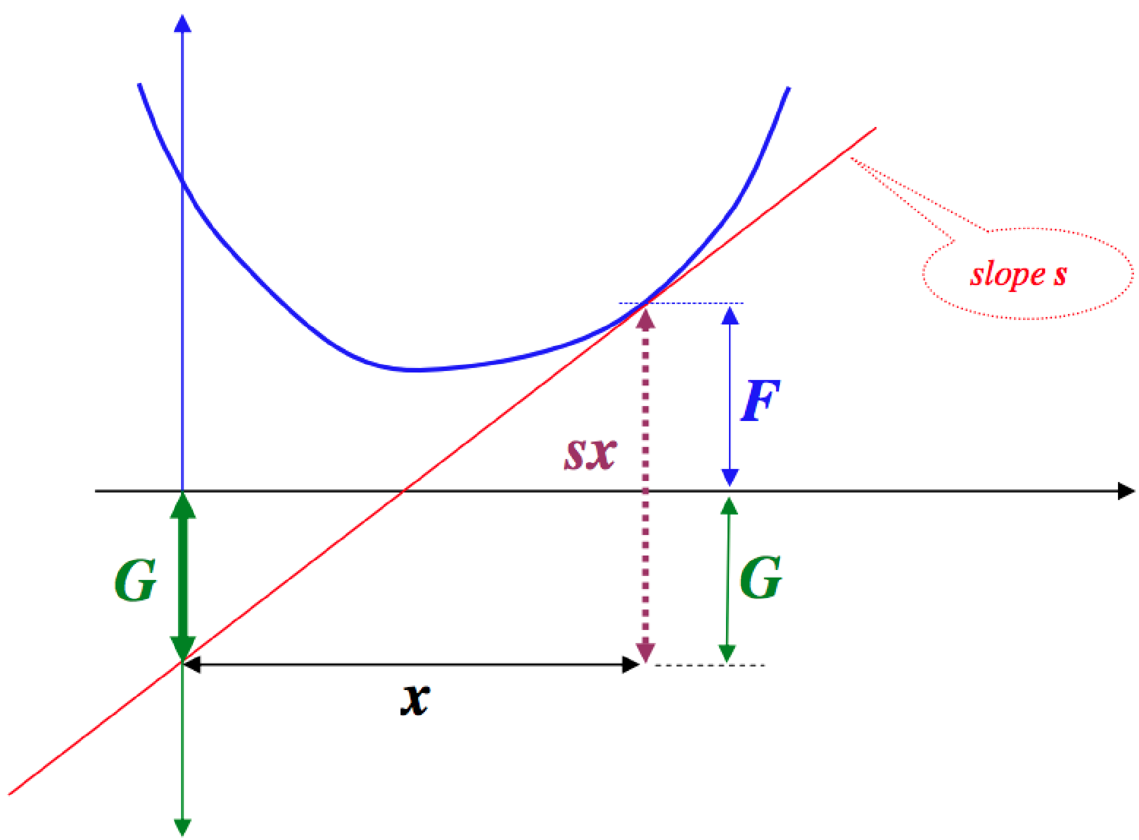


Fig. 3.2: This graph shows the Legendre transformation and triangles in which G is actually the F we used before and F in the graph corresponds to f.

One imediately notices the symmety of Legendre transformation on interchanging of F and f.

This graph is taken from this paper Making Sense of the Legendre Transform .

This is the triangle that represents the Legendre transformation.

If we have a slope that vanishes, which means $f(x)$ is at minimium, then we have the relation

3.1.3 Vector Analysis

The ultimate trick is to use component form.

$$
\begin{aligned}
&\vec{a} \times (\vec{b} \times \vec{c}) \\
=&\hat{e}_i \epsilon_{ijk} a_j (\epsilon_{kmn} b_m c_n) \\
=&\hat{e}_i \epsilon_{kij} \epsilon_{kmn} a_j b_m c_n \\
=&\hat{e}_i (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_j b_m c_n \\
=&\hat{e}_i \delta_{im}\delta_{jn} a_j b_m c_n - \hat{e}_i \delta_{in}\delta_{jm} a_j b_m c_n \\
=&\hat{e}_i a_j b_i c_j - \hat{e}_i a_j b_j c_i \\
=&\vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b}).
\end{aligned}
$$

One should be able to find the component forms of gradient $\vec{\nabla}\cdot$, divergence $\vec{\nabla}\times$, Laplace operator, in **spherical coordinates, cylindrical coordinates and cartisian coordinates**.

3.1.4 Refs & Notes

3.2 Useful Physics Concepts and Tools

3.2.1 Dimension

How to find the relationship between two quantities? For example, what is the dimensional relationship between length and mass.

* Plank constant: $\hbar \sim [\text{Energy}] \cdot [\text{Time}] \sim [\text{Mass}] \cdot [\text{Length}]^2 \cdot [\text{Time}]^{-1}$
* Speed of light in vacuum: $\text{c} \sim [\text{Length}] \cdot [\text{Time}]^{-1}$
* Gravitational constant: $\text{G} \sim [\text{Length}]^3 \cdot [\text{Mass}]^{-1} \cdot [\text{Time}]^{-2}$

Then it is easy to find that a combination of $c/\hbar$ cancels the dimension of mass and leaves the inverse of length. That is

$$[L]^2 = \left[\frac{\hbar G}{c^3}\right]$$

$$[M]^2 = \left[\frac{\hbar c}{G}\right]$$

$$[T]^2 = \left[\frac{\hbar G}{c^5}\right]$$

As we can see, it is possible to use $c = 1, \hbar = 1, G = 1$ because we can always restore the units in a deterministic way. $c, \hbar, G$ are function of mass, length, time, and with $c = \hbar = G = 1$ give us only one solution of mass, length and time: three equations + three variables.

Planck Scales

As we have seen, the three constant can make up a length scale, a mass scale, a time scale. Then what are they?

Planck length:

$$l_P = \sqrt{\frac{\hbar G}{c^3}}$$

Planck mass:

$$m_P = \sqrt{\frac{\hbar c}{G}}$$

Planck time:

$$t_P = \sqrt{\frac{\hbar G}{c^5}}$$

Equations and Dimensions

Before solving equations, it is good to reform them in to dimensionless ones.

To make the equation dimensionless doesn't mean we can just divide arbitary terms on both sides. We need to find out the characteristic quantity of the system. For example, we can divide by $\hbar\omega$ on both sides of Schrodinger equation for Harmonic Oscillators. This is a good step because $\hbar\omega$ is the characteristic energy scale of system. At the same time, we can make the length terms dimensionless using the characteristic length. DO NOT use an arbitary length!

3.2.2 Most Wonderful Equations That Should Never Be Forgotten

Electrodynamics

Maxwell Equations

$$\nabla \times \vec{E} = -\partial_t \vec{B}$$

$$\nabla \times \vec{H} = \vec{J} + \partial_t \vec{D}$$

$$\nabla \cdot \vec{D} = \rho$$

$$\nabla \cdot \vec{B} = 0$$

For linear meterials,

$$\vec{D} = \epsilon \vec{E}$$

$$\vec{B} = \mu \vec{H}$$

$$\vec{J} = \sigma \vec{E}$$

Dynamics

Hamilton conanical equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$

Thermodynamics and Statistical Physics

Liouville's Law

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} \equiv \frac{\partial \rho}{\partial t} + \sum_i \left[\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right] = 0$$

CHAPTER 4

Qualitative Methods

4.1 Dimensional Analysis

Reference Books

1. Dimensional Analysis and Group Theory in Astrophysics, by Rudolf Kurth

4.1.1 Notations

I find the Kurth notation being convinient. Kurth used l^a, t^b, m^c to denote length to the ath power, time to the bth powe and mass to the cth power.

4.1.2 Useful Physical Constants

1. Newton's gravitational constant: $[G] \rightarrow l^3 t^{-2} m^{-1}$. The quick derivation is $-GM/r^2 = F/m = a$.

4.1.3 Nondegenerate Case

1. We find all the physical quantities and relevant physical constants.
2. Check if the combination is unique.

Deflection of Light by Sun

The example given in Kurth's book is the deflection of light by gravitational field. The relevant quantities are

1. deflection angle: θ which is dimensionless, $[\theta] \rightarrow l^0 t^0 m^0$,
2. the mass of the Sun: M which has dimension $[M] \rightarrow l^0 t^0 m^1$,

3. the distance from the center of the Sun at the point of closest contact, $[r] \to l^1 t^0 m^0$.

The physical constant we can think of in the first place is G, which has dimension $[G] \to l^3 t^{-2} m^{-1}$. However, we now combine G and M and r to produce θ, which is identical to solving the equation

$$\theta^a = G^b M^c r^d,$$

or the system of equations

$$\begin{aligned} 0 =& 3b + d \\ &\text{for l} \\ 0 =& -2b \\ &\text{for t} \\ 0 =& (-b) + c \\ &\text{for m.} \end{aligned}$$

The system of equations has no nontrivial solutions. Where to find the other physical constant? We are dealing with light, thus one of the choices is the speed of light, c. Add in $[c^e] \to l^e t^{-e} m^0$, we have the equations

$$\begin{aligned} 0 =& 3b + d + e \\ &\text{for l} \\ 0 =& (-2b) + (-e) \\ &\text{for t} \\ 0 =& (-b) + c \\ &\text{for m,} \end{aligned}$$

which has unique nontrivial set of solutions

$$\begin{aligned} c =& b \\ d =& -b \\ e =& -2b. \end{aligned}$$

Now we find that

$$\theta^a = \left(\frac{GM}{rc^2} \right)^b.$$

In general, since $\frac{GM}{rc^2}$ is dimensionless, the general form is

$$\theta = f(\frac{GM}{rc^2}).$$

This result is already satisfactory.

Is there more we can conclude from here? Kurth took a step further and **used limits**. We expect $\theta \to 0$ for small mass since we do not observe this effect in daily life. $M \to 0$ leads to $\frac{GM}{rc^2} \to 0$. We can even use the simplest form

$$\theta \propto \frac{GM}{rc^2}.$$

This is identical to the fact that the Taylor expansion of function $f(x)$ at $x \to 0$ has a neglectable zeroth order.

To summarize, we used the following techniques.

1. Dimensions.
2. Limits of physical problems compared with observations.

If GR is part of the knowledge pool, we notice that the radius R of the celestial body is not consider in this analysis. When we add in this, we find three dimensionless quantities,

$$\frac{GM}{rc^2}$$
$$\frac{GM}{Rc^2}$$
$$\frac{R}{r}.$$

They are not linearly independent. We choose two of them,

$$\frac{GM}{rc^2}$$
$$\frac{R}{r}.$$

CHAPTER 5

Mathematics

5.1 Kindergarten

Binominal theorem

$$(1+x)^n = \sum_{k=0}^{n} C_n^k x^k.$$

5.2 Special Functions

There are a lot of useful special function in physics. Some of them provides physics understanding of the problem, some of them helps us writing down a solution quickly.

Among them, Gamma functions, Legendre polynomials, Bessel functions, spherical harmonics, modified bessel functions, spherical bessel functions, and elliptical functions are the most used ones.

5.2.1 Gamma Functions

Gamma function satisfies the following relatioin,

$$\Gamma(z+1) = z\Gamma(z).$$

For some cases, it can also be written as

$$\Gamma(n) = \int_0^\infty dt t^{n-1} e^{-t}.$$

One can prove that

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

5.2.2 Legendre Polynomials

Legendre polynomials are solutions to Legendre equation, which is

$$\left(\frac{d}{dx} \left[(1-x^2) \frac{d}{dx} \right] + n(n+1) \right) P_n(x) = 0.$$

Legendre polynomials has many different representations.

Integral

$$P_n(z) = \frac{1}{2\pi i} \oint (1 - 2tz + t^2)^{1/2} t^{-n-1} dt.$$

Rodrigues representation

$$P_n(z) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2-1)^l.$$

It's generation function is

$$\frac{1}{\sqrt{1+\eta^2 - 2\eta x}} = \sum_{k=0}^{\infty} \eta^k P_k(x).$$

Properties

Orthogonality

$$\int_{-1}^{1} P_m(x) P_n(x) dx = \frac{2}{2n+1} \delta_{mn}.$$

They all have value 1 at $z = 1$.

The parity is alternating.

Examples

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1). \end{aligned}$$

Through these, we can solve out

$$\begin{aligned} x &= P_1(x) \\ x^2 &= \frac{1}{3}(P_0(x) + 2P_2(x)). \end{aligned}$$

Notice that they have physics meanings although it's better to understand it together with spherical harmonics.

5.2.3 Associated Legendre Polynomials

The associated Legendre equation is

$$\left(\frac{d}{dx} \left[(1-x^2) \frac{d}{dx} \right] + n(n+1) - \frac{m^2}{1-x^2} \right) P_n(x) = 0.$$

The solution to this equation is Associated Legendre polynomial, which can be represented by

$$P_n^{\nu}(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x).$$

5.2.4 Bessel Functions

Bessel functions are solutions to Bessel equation,

$$\left(x\frac{d}{dx}x\frac{d}{dx}+x^2-\nu^2\right)J_\nu(x)=0.$$

They all satisfy these recurrence relations,

$$Z_{n+1}+Z_{n-1}=\frac{2n}{x}Z_n$$
$$Z_{n-1}-Z_{n+1}=2\frac{d}{dx}Z_n.$$

Bessel Function of the first kind

Use notation $J_n(x)$ for the first kind.

Generating function is

$$e^{\frac{z}{2}\left(t-\frac{1}{t}\right)}=\sum_{n=-\infty} inftyt^n J_n(z).$$

Integral representation

$$J_n(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i(n\tau-x\sin\tau)}d\tau.$$

It also has a **summation representation**,

$$J_\alpha(z)=\sum_{m=0}^{\infty}\frac{(-1)^m}{m!\Gamma(m+\alpha+1)}\left(\frac{x}{2}\right)^{2m+\alpha}.$$

At large $|x|$ limits, we have

$$\lim_{|x|\to\infty}J_l(x)=\frac{\sin(z-l\frac{\pi}{2})}{x}$$
$$\lim_{|x|\to\infty}J_l'(x)=\frac{\cos(z-l\frac{\pi}{2})}{x}.$$

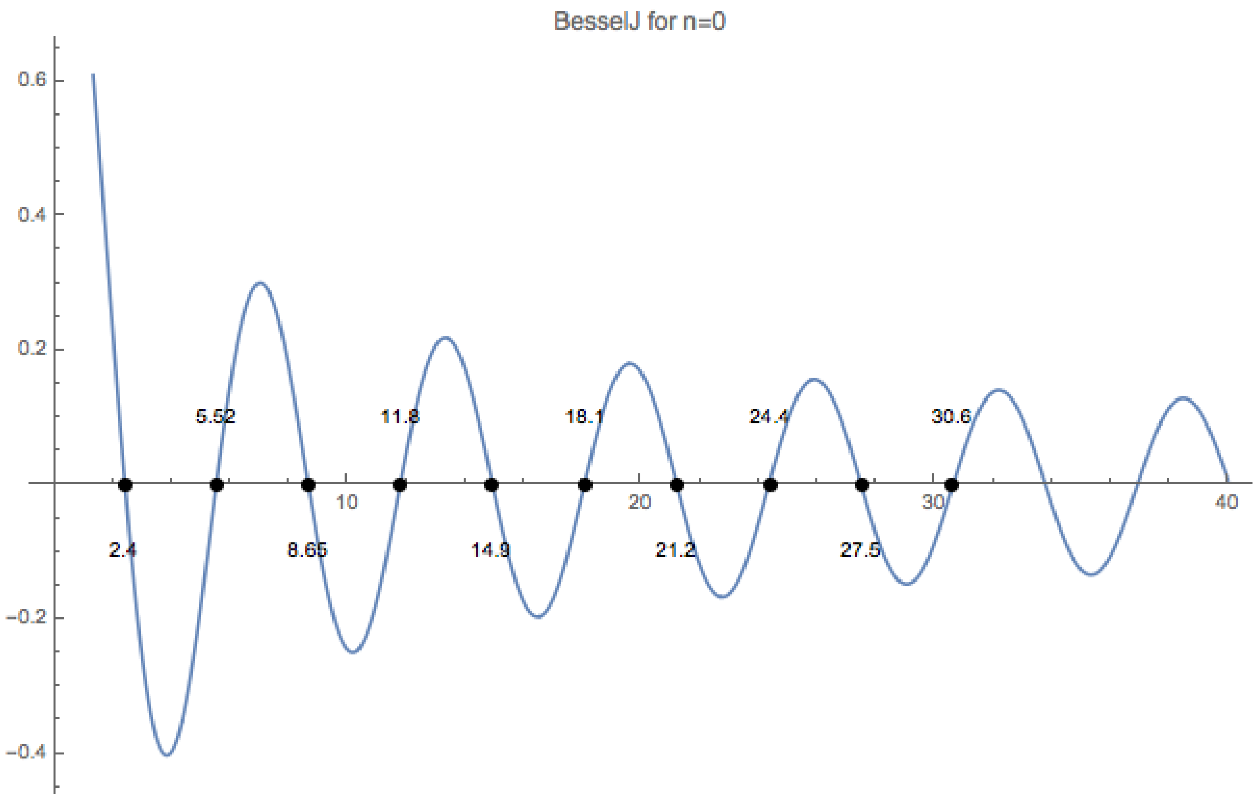
By playing with the recurrence relation,

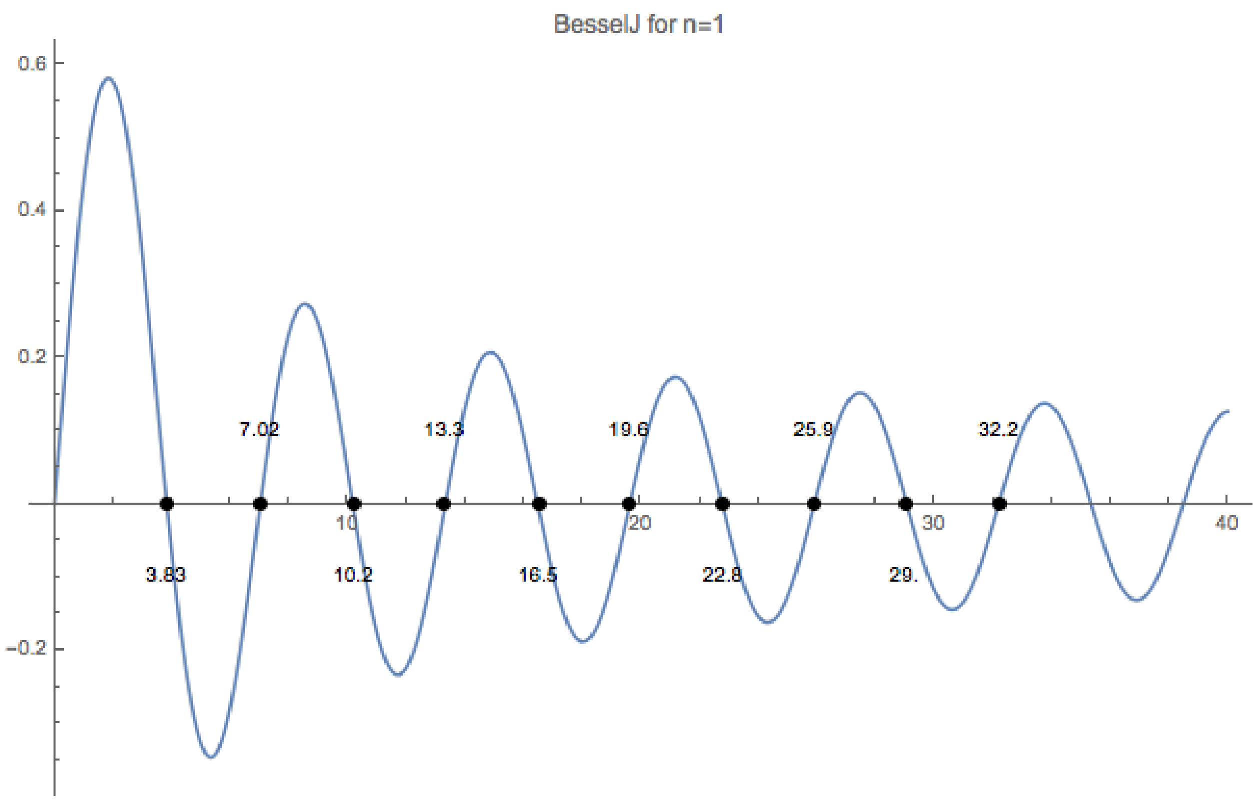
$$2J_n'=J_{n-1}-J_{n+1}$$
$$2nJ_n=J_{n+1}+J_{n-1},$$

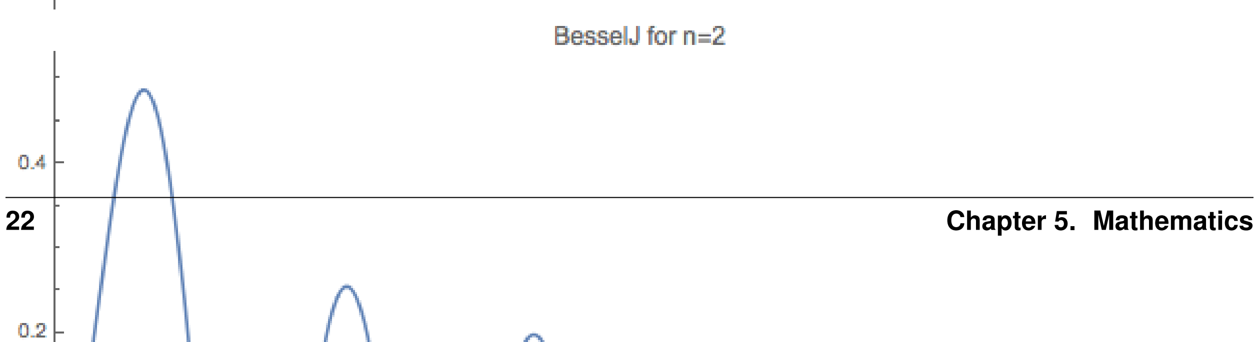
we can get two more useful relations,

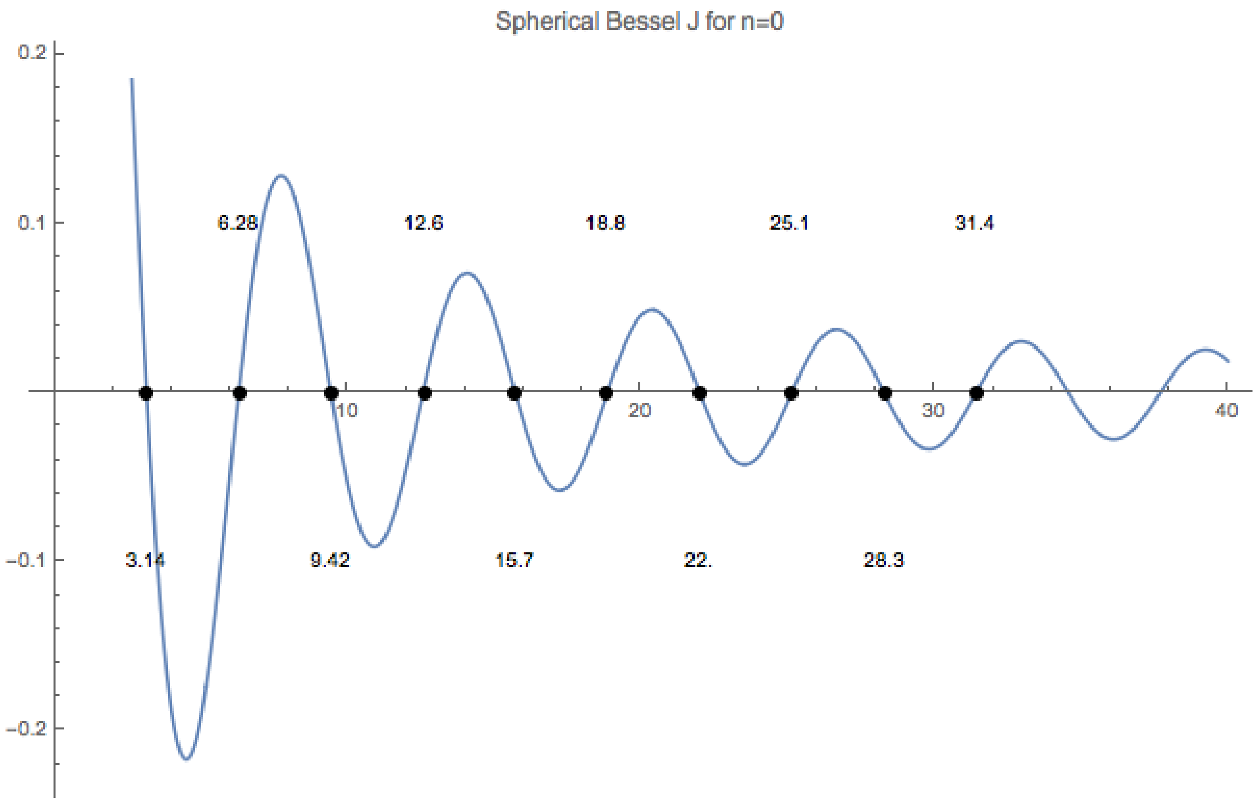
$$\frac{d}{dz}(z^nJ_n)=z^nJ_{n-1}$$
$$\frac{d}{dz}(z^{-n}J_n)=-z^{-n}J_{n+1}.$$

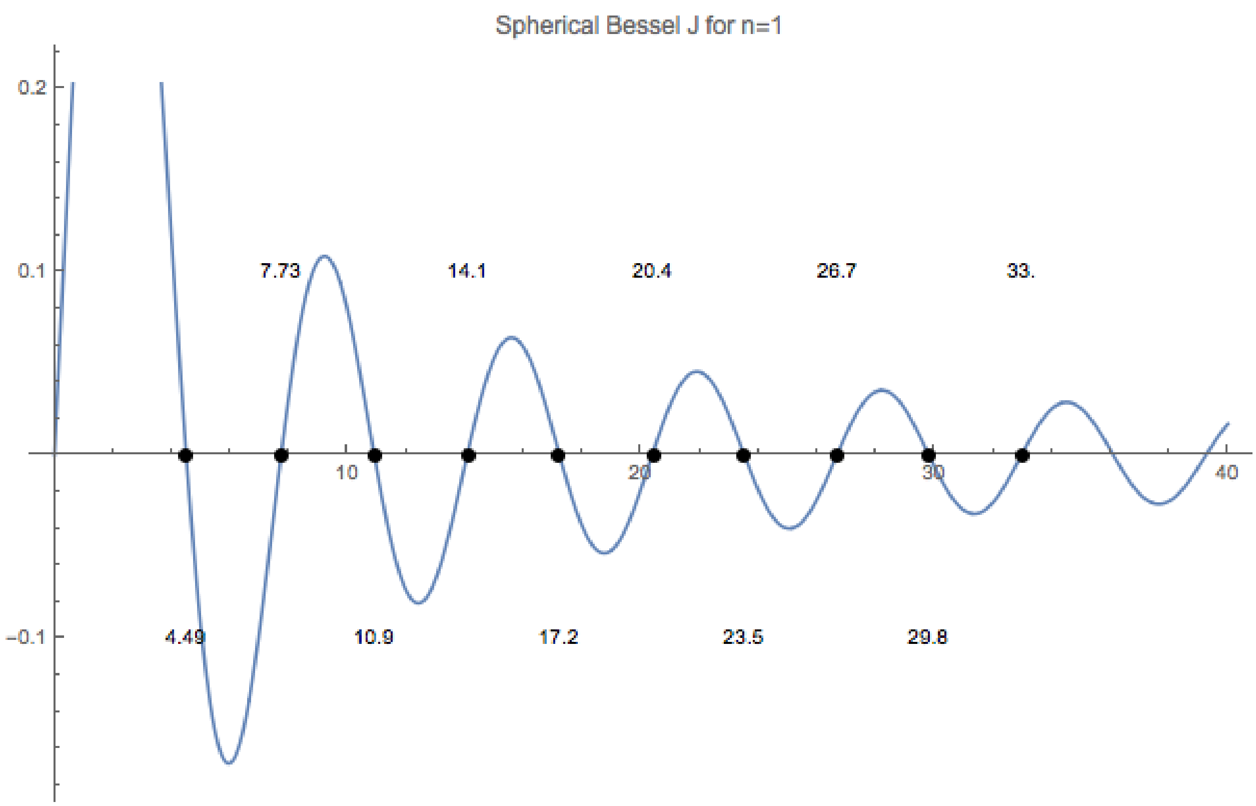
They are very useful when integrating by part.

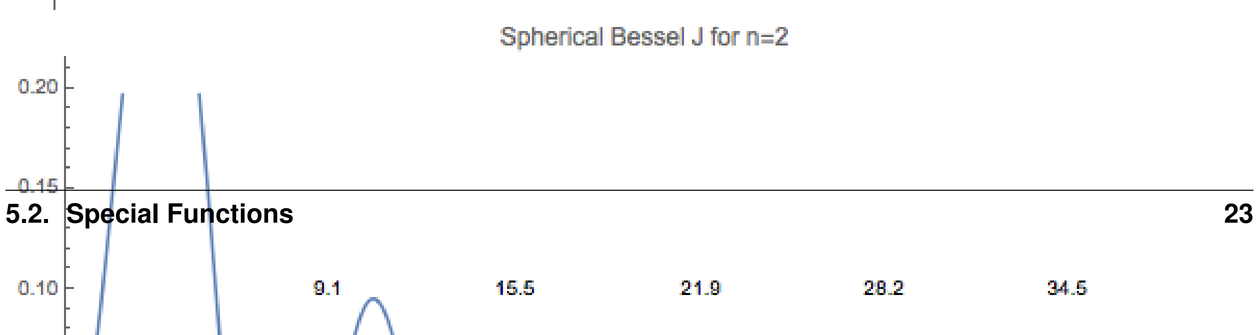

BesselJ for n=0
0.6
0.4
0.2
−0.2
−0.4
10
20
30
40
5.52
11.8
18.1
24.4
30.6
2.4
8.65
14.9
21.2
27.5


BesselJ for n=1
0.6
0.4
0.2
−0.2
10
20
30
40
7.02
13.3
19.6
25.9
32.2
3.83
10.2
16.5
22.8
29.


BesselJ for n=2
0.4
0.2


Spherical Bessel J for n=0
0.2
0.1
−0.1
−0.2
10
20
30
40
6.28
12.6
18.8
25.1
31.4
3.14
9.42
15.7
22.
28.3


Spherical Bessel J for n=1
0.2
0.1
−0.1
10
20
30
40
7.73
14.1
20.4
26.7
33.
4.49
10.9
17.2
23.5
29.8


Spherical Bessel J for n=2
0.20
0.15
0.10
9.1
15.5
21.9
28.2
34.5

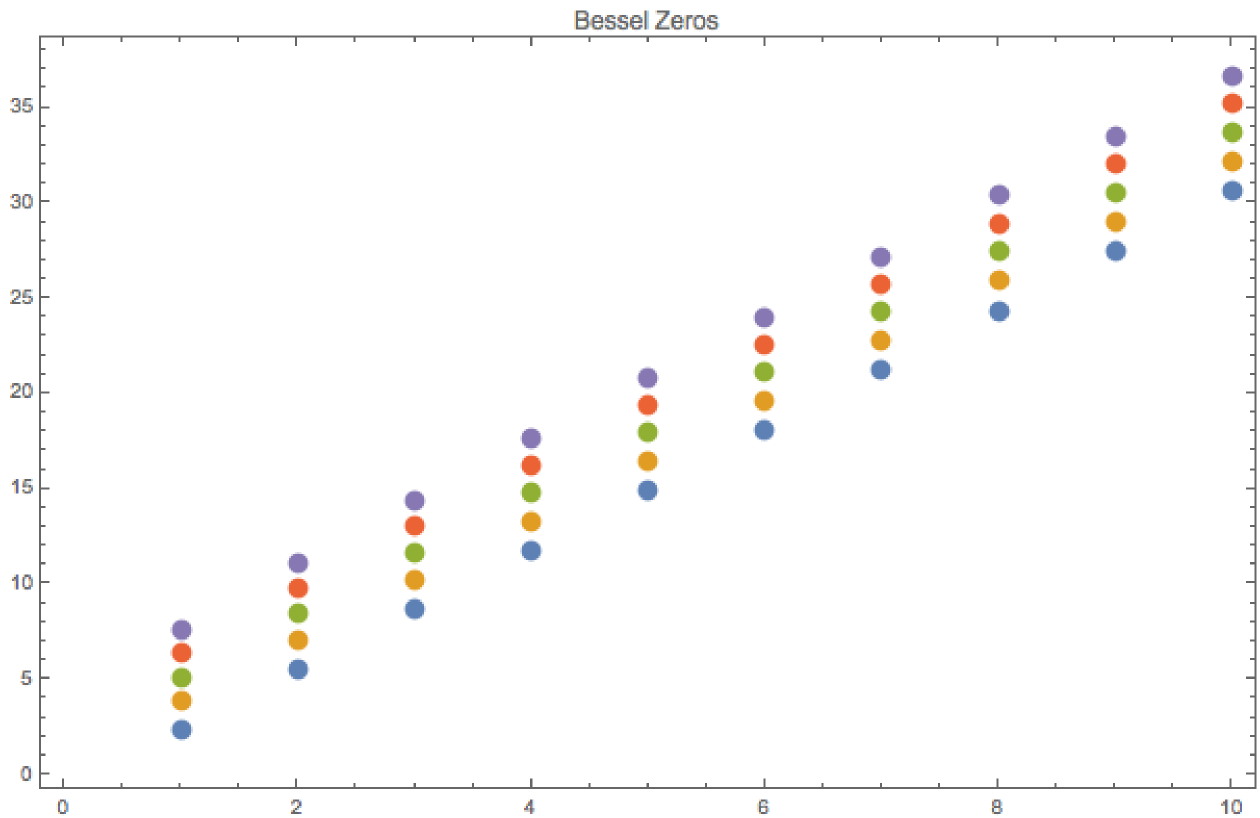


Fig. 5.3: Bessel function zeros in a list plot. Horizontal axis is nth zero point, while vertical axis is the value.

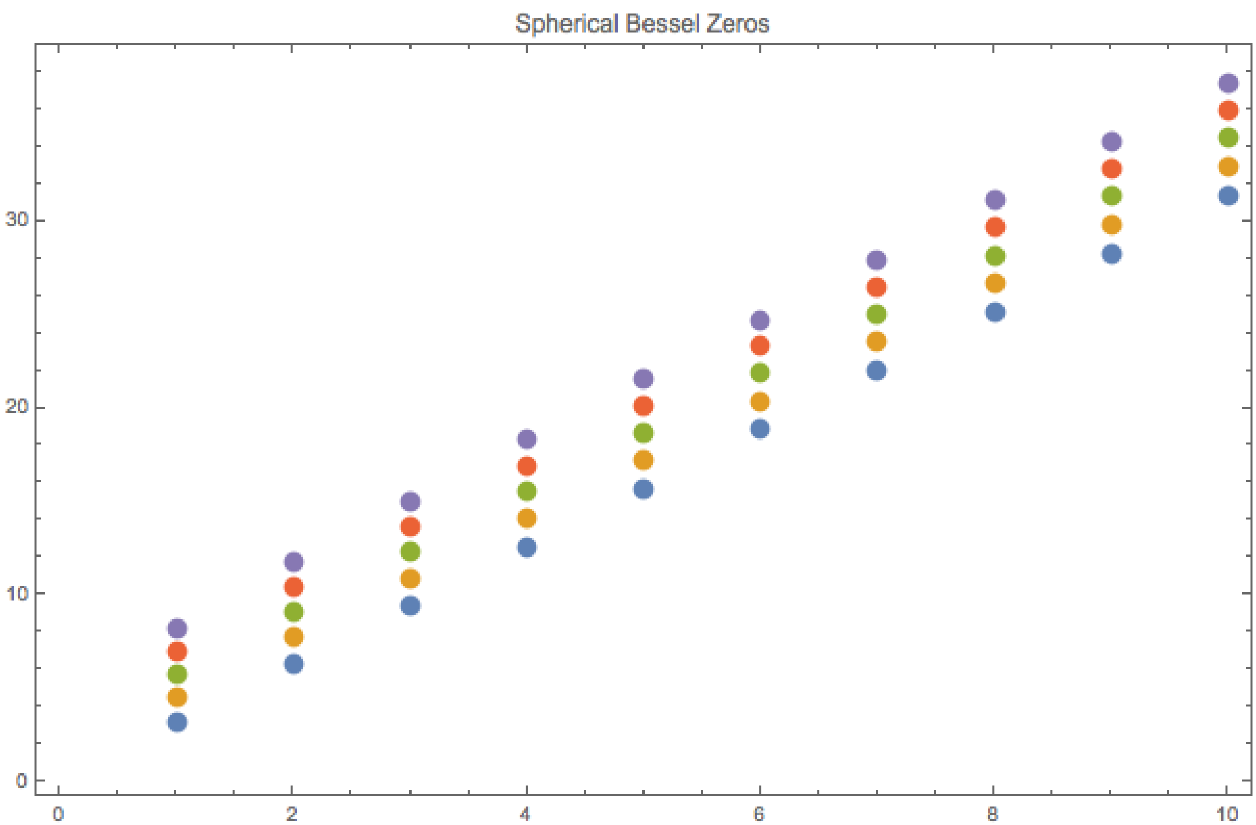


Fig. 5.4: Spherical Bessel function zeros.

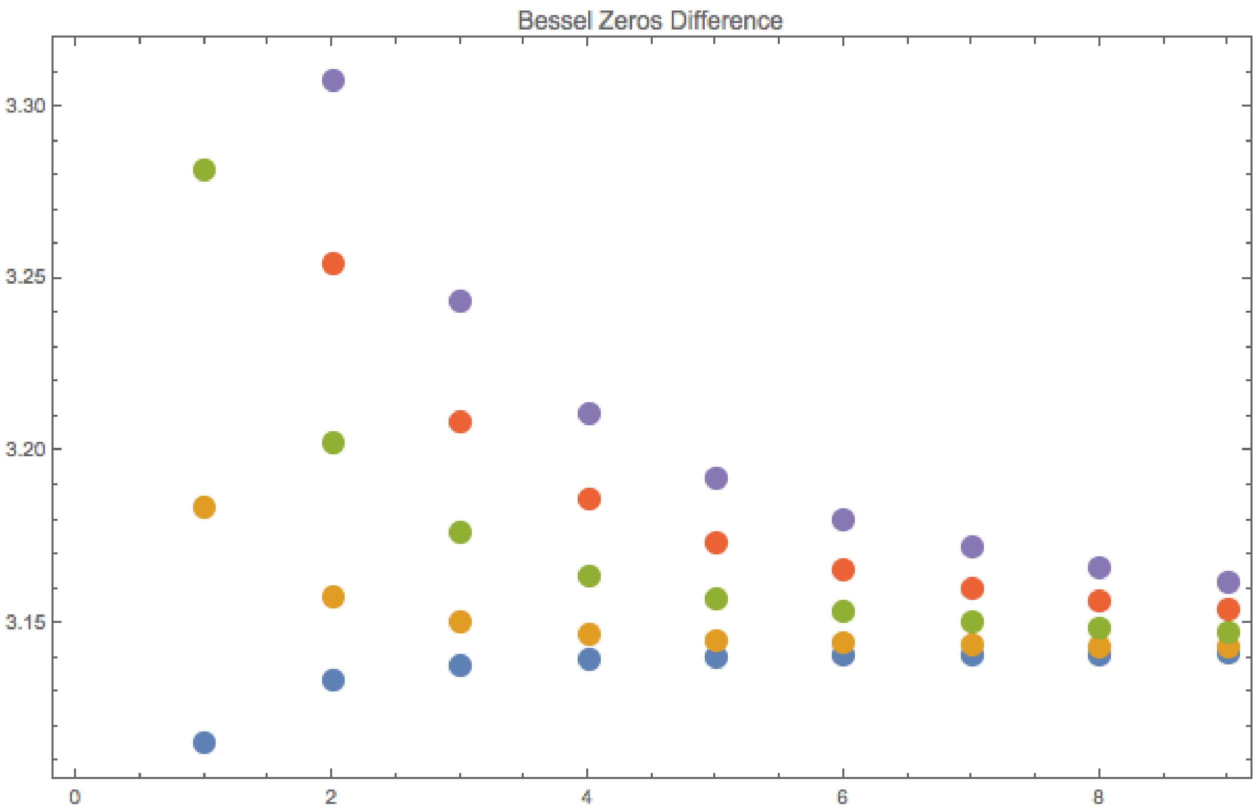


Fig. 5.5: The difference between zeros of Bessel functions. They are almost the same, which a around Pi.

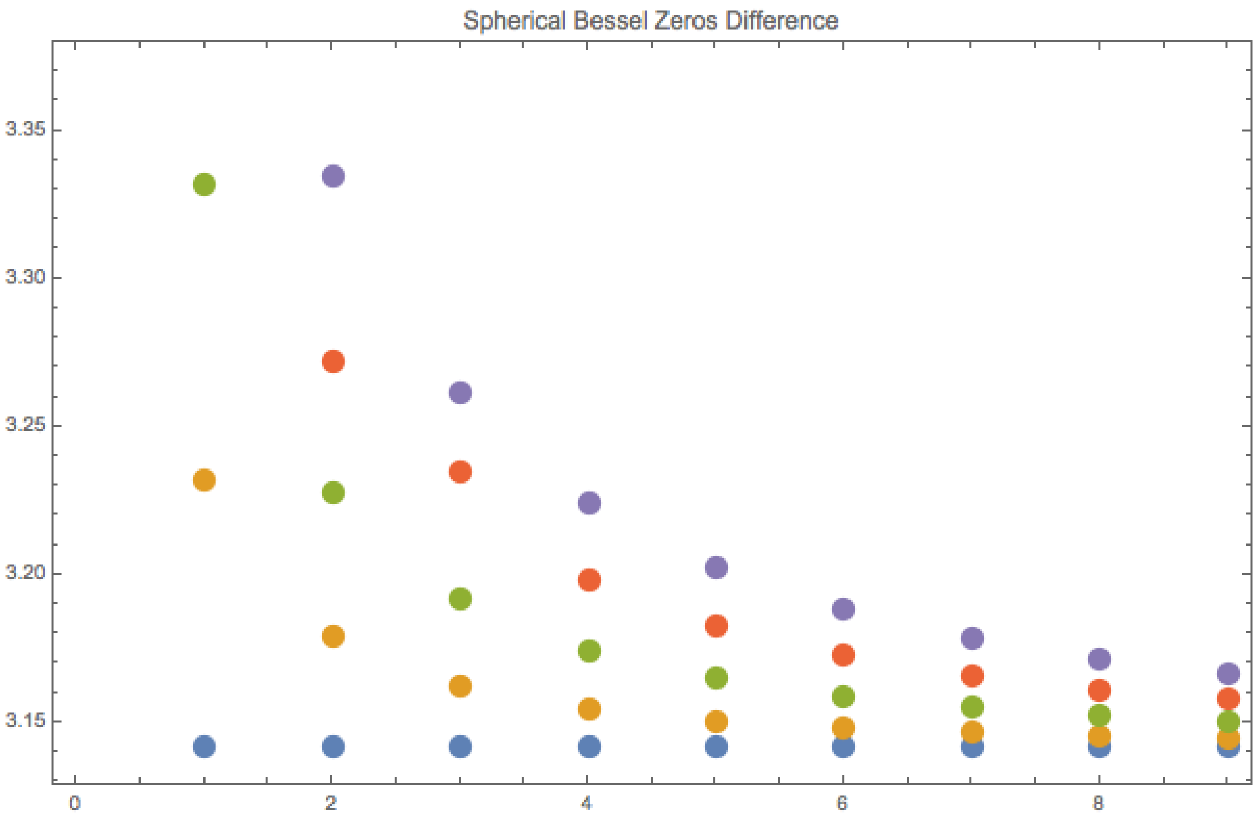


Fig. 5.6: Spherical Bessel function zeros differences.

Graphics and Properties

5.2.5 Refs & Notes

5.3 Equation Solving

More about Equation Solving

For more about equation solving please refer to another notebook of mine: Intelligence.

There are so many methods and techniques to solve an equation. Here we will review only some of them.

5.3.1 Ordinary Differential Equations

There are many important equations in physics.

The are many methods to solve an ODE,

1. Green's function.
2. Series solution
3. Laplace transform
4. Fourier transform

Green's Function

Definition of Green's Function

The idea of Green/s function is very simple. To solve a general solution of equation

$$\frac{d^2}{dx^2}y(x) + y(x) = f(x), \tag{5.1}$$

where $f(x)$ is the source and some given boundary conditions. To save ink we define

$$\hat{L}_x = \frac{d^2}{dx^2} + 1,$$

which takes a function $y(x)$ to $f(x)$, i.e.,

$$\hat{L}_x y(x) = f(x). \tag{5.2}$$

Now we define the Green's function to be the solution of equation (5.2) but replacing the source with delta function $\delta(x-z)$

$$\hat{L}_x G(x,z) = \delta(z-x).$$

Why do we define this function? The solution to equation (5.1) is given by

$$y(x) = \int G(x,z) f(z) dz.$$

Equation	Regular singularities	Essential singularities
Hypergeometric $z(1-z)y'' + [c-(a+b+1)z]y' - aby = 0$	$0, 1, \infty$	—
Legendre $(1-z^2)y'' - 2zy' + \ell(\ell+1)y = 0$	$-1, 1, \infty$	—
Associated Legendre $(1-z^2)y'' - 2zy' + \left[\ell(\ell+1) - \frac{m^2}{1-z^2}\right] y = 0$	$-1, 1, \infty$	—
Chebyshev $(1-z^2)y'' - zy' + \nu^2 y = 0$	$-1, 1, \infty$	—
Confluent hypergeometric $zy'' + (c-z)y' - ay = 0$	0	∞
Bessel $z^2y'' + zy' + (z^2 - \nu^2)y = 0$	0	∞
Laguerre $zy'' + (1-z)y' + \nu y = 0$	0	∞
Associated Laguerre $zy'' + (m+1-z)y' + (\nu - m)y = 0$	0	∞
Hermite $y'' - 2zy' + 2\nu y = 0$	—	∞
Simple harmonic oscillator $y'' + \omega^2 y = 0$	—	∞

Table 16.1 Important second-order linear ODEs in the physical sciences and engineering.

Fig. 5.7: Taken from Riley's book.

To verify this conclusion we plug it into the LHS of equation (5.1)

$$\begin{aligned}
&\left(\frac{d^2}{dx^2}+1\right)\int G(x,z)f(z)dz\\
=&\int\left[\left(\frac{d^2}{dx^2}+1\right)G(x,z)\right]f(z)dz\\
=&\int\delta(z-x)f(z)dz\\
=&f(x),
\end{aligned}$$

in which we used one of the properties of Dirac delta distribution

$$\int f(z)\delta(z-x)dz = f(x).$$

Also note that delta function is even, i.e., $\delta(-x)=\delta(x)$.

So all we need to do to find the solution to a standard second differential equation

$$\left(\frac{d^2}{dx^2}+p(x)\frac{d}{dx}+q(x)\right)y(x)=f(x)$$

is do the following.

1. Find the general form of Green's function (GF) for operator for operator $\hat{L}_x$.
2. Apply boundary condition (BC) to GF. This might be the most tricky part of this method. Any ways, for a BC of the form $y(a)=0=y(b)$, we can just choose it to vanish at a and b. Otherwise we can move this step to the end when no intuition is coming to our mind.
3. Continuity at $n-2$ order of derivatives at point $x=z$, that is

   $$G^{(n-2)}(x,z)|_{x<z}=G^{(n-2)}|_{x>z}, \qquad \text{at } x=z.$$

4. Discontinuity of the first order derivative at $x=z$, i.e.,

   $$G^{(n-1)}(x,z)|_{x>z}-G^{(n-1)}(x,z)|_{x<z}=1, \qquad \text{at } x=z.$$

   This condition comes from the fact that the integral of Dirac delta distribution is Heaviside step function.
5. Solve the coefficients to get the GF.
6. The solution to an inhomogeneous ODE $y(x)=f(x)$ is given immediately by

   $$y(x)=\int G(x,z)f(z)dz.$$

   If we haven't done step 2 we know would have some unkown coefficients which can be determined by the BC.

How to Find Green's Function

So we are bound to find Green's function. Solving a nonhonogeneous equation with delta as source is as easy as solving homogeneous equations.

We do this by demonstrating an example differential equation. The problem we are going to solve is

$$\left(\frac{d^2}{dx^2}+\frac{1}{4}\right)y(x)=f(x),$$

with boundary condition

$$y(0) = y(\pi) = 0. \tag{5.3}$$

For simplicity we define

$$\hat{L}_x = \frac{d^2}{dx^2} + \frac{1}{4}.$$

First of all we find the GF associated with

$$\hat{L}_x G(x,z) = \delta(z-x).$$

We just follow the steps.

- The general solution to

  $$\hat{L}_x G(x,z) = 0$$

  is given by

  $$G(x,z) = \begin{cases} A_1 \cos(x/2) + B_1 \sin(x/2), & x \leq z, \\ A_2 \cos(x/2) + B_2 \sin(x/2), & x \geq z. \end{cases}$$

- Continuity at $x = z$ for the 0th order derivatives,

  $$G(z_-, z) = G(z_+, z),$$

  which is exactly

  $$A_1 \cos(z/2) + B_1 \sin(z/2) = A_2 \cos(z/2) + B_2 \sin(z/2). \tag{5.4}$$

- Discontinuity condition at 1st order derivatives,

  $$\frac{d}{dx} G(x,z)\bigg|_{x=z_+} - \frac{d}{dx} G(x,z)\bigg|_{x=z_-} = 1,$$

  which is

  $$-\frac{A_2}{2} \sin\frac{z}{2} + \frac{B_2}{2} \cos\frac{z}{2} - \left(-\frac{A_1}{2} \sin\frac{z}{2} + \frac{B_1}{2} \cos\frac{z}{2}\right) = 1 \tag{5.5}$$

  Now we combine ((5.4)) and ((5.5)) to eliminate two degrees of freedom. For example, we can solve out A_1 and B_1 as a function of all other coefficients. Here we have

  $$B_1 = \frac{-2/\sin(z/2)}{\tan(z/2) + \cot(z/2)} + B_2,$$

  $$A_1 = A_2 + B_2(\tan(z/2) - 1) + \frac{2}{\sin(z/2) + \cot(z/2)\cos(z/2)}.$$

- Write down the form solution using $y(x) = \int G(x,z) f(z) dz$. Then we still have two unknown free coefficients A_2 and B_2, which in fact is to be determined by the BC equation (5.3).

Series Solution

A second order ODE,

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0$$

Wronskian of this is

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix},$$

where y_1 and y_2 are linearly independent solutions, i.e., $c_1 y_1 + c_2 y_2 = 0$ is only satisfied when $c_1 = c_2 = 0$. **Wronskian is NOT zero if they are linearly independent.**

Singularities of an ODE is are defined when $p(x)$ or $q(x)$ or both of them have singular points. For example, Legendre equation

$$(1 - z^2)y'' - 2zy' + l(l+1)y = 0$$

has three singular points which are $z = \pm 1, \infty$ while $z = 0$ is an ordinary point.

Solution at Ordinary Points

Series expansion of the solution can be as simple as

$$y(z) = \sum_{n=0}^{\infty} a_n z^n,$$

which converges in a radius R where R is the distance from $z = 0$ to the nearest singular point of our ODE.

Solution at Regular Singular Points

Frobenius series of the solution

$$y(z) = z^\sigma \sum_{n=0}^{\infty} a_n z^n.$$

The next task is to find the indicial equation.

If the roots are not differing by an integer, we just plug the two solutions to σ in and find two solutions independently.

If the roots differ by an integer, on the other side, we can only plug in the **larger** root and find one solution. As for the second solution, we need some other techniques, such as Wronskian method and derivative method.

Wronskian method requires two expression of Wronskian, which are

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix},$$

and

$$W(z) = Ce^{-\int^z p(u)\mathrm{d}u}.$$

From the first expression, we have

$$y_2(z) = y_1(z) \int^z \frac{W(u)}{y_1(u)^2} \mathrm{d}u.$$

However, we don't know $W(z)$ at this point. We should apply the second expression of Wronskian,

$$y_2(z) = y_1(z) \int^z \frac{Ce^{-\int^z p(u)\mathrm{d}u}}{y_1(u)^2}\mathrm{d}u,$$

where the constant C can be set to 1 as one wish.

TO DO

The **derivative method** is on my to do list.

Comparing With A General Form

For equation that take the following form,

$$y'' + \frac{1-2a}{x}y' + \left((bcx^{c-1})^2 + \frac{a^2 - p^2c^2}{x^2}\right) y = 0,$$

where $y \equiv y(x)$, we can write down the solutions immediately,

$$y(x) = x^a \mathscr{Z}_p(bx^c),$$

in which $\mathscr{Z}_p$ is the solution to Bessel equation, i.e., is one kind of Bessel function with index p.

A Pendulum With A Uniformly Chaning String Length

As an example, let's consider the case of length changing pendulum,

$$\frac{d}{dt}\left(ml^2\dot{\theta}\right) = -mgl\sin\theta \approx= -mgl\theta.$$

Notice that l is a function of time and

$$l = l_0 + vt.$$

Then the equation can be rewritten as

$$\frac{d^2}{dl^2}\theta + \frac{2}{l}\frac{d}{dl}\theta + \frac{g/v^2}{l}\theta = 0.$$

Comparing with the general form, we have one of the possible solutions

$$\begin{aligned} a &= -1/2, \\ pc &= 1/2, \\ c &= 1/2, \\ p &= 1, \\ b &= 2\sqrt{g}/v. \end{aligned}$$

This solution should be

$$\begin{aligned} \theta &= l^a \mathscr{Z}_p(bl^c) \\ &= \frac{1}{\sqrt{l}} J_1(\frac{2\sqrt{g}}{v}\sqrt{l}). \end{aligned}$$

Airy Equatioin

Time-independent Schrödinger equation with a simple potential,

$$\ddot{\Psi} + \alpha x \Psi = 0.$$

Comparing it with general form, we should set

$$\begin{aligned} a &= 1/2, \\ |pc| &= 1/2, \\ c &= 3/2, \\ b^2 c^2 &= \alpha^2. \end{aligned}$$

So the two possible solutions are

$$\begin{aligned} \Psi_1(x) &= \sqrt{x} \mathscr{Z}_{1/3}(2/3 \alpha x^{3/2}), \\ \Psi_2(x) &= \sqrt{x} \mathscr{Z}_{-1/3}(2/3 \alpha x^{3/2}). \end{aligned}$$

The general solution is

$$\Psi(x) = a\Psi_1(x) + b\Psi_2(x).$$

Second Order Differential Equations and Gauss' Equation

Gauss' equation has the form

$$z(z-1)\frac{d^2}{dz^2}u(z) + \left[(a+b+1)z - c\right]\frac{d}{dz}u(z) + abu(z) = 0,$$

which has a solution of the hypergeometric function form

$$u(z) = {}_2F_1(a, b; c; z).$$

The interesting part about this equation is that its Papperitz symbol is

$$\begin{Bmatrix} 0 & 1 & \infty & \\ 0 & 0 & a & z \\ 1-c & c-a-b & b & \end{Bmatrix},$$

in which the first three columns are the singularities at points $0, 1, \infty$ while the last column just points out that the argument of this equation is z.

This means, in some sense, the solution to any equation with three singularities can be directly written down by comparing the equation with Gauss' equation. If you care, the actual steps are changing variables, rewriting the equation into Gauss' equation form, writing down the solutions.

5.3.2 Integral Equations

Neumann Series AKA WKB

For differential equation, whenever the highest derivative is multiplied by a small parameter, try this. But generally, the formalism is the following.

First of all, we use Hilbert space $\mathscr{L}[a, b; w]$ which means the space is defined on $[a, b]$ with a weight w, i.e.,

$$\langle f \mid g \rangle = \int_a^b dx w(x) f(x) g(x).$$

Quantum Mechanics Books

Notice that this is very different from the notation we used in most QM books.

What is the catch? Try to write down $\langle x \mid u \rangle$. It's not that different because one can alway go back to the QM notation anyway.

With the help of Hilbert space, one can alway write down the vector form of some operators. Suppose we have an equation

$$\hat{L} u(x) = f(x),$$

where $\hat{L} = \hat{I} + \hat{M}$. So the solution is simply

$$\begin{aligned} u(x) &= \hat{L}^{-1} f(x) \\ &= (\hat{I} + \hat{M})^{-1} f(x). \end{aligned}$$

However, it's not a solution until we find the inverse. A most general approach is the Neumann series method. We require that

$$\|\hat{M} u\| < \gamma \|u\|,$$

where $\gamma \in (0, 1)$ and should be independent of x.

As long as this is satisfied, the equation can be solved using Neumann series, which is an iteration method with

$$\begin{aligned} u(x) &= u_0(x) + \delta u_1(x) + \delta^2 u_2(x) + \cdots \\ u_0(x) &= f(x). \end{aligned}$$

As an example, we can solve this equation

$$(\hat{I} + |t\rangle \langle \lambda|) u(t) = f(t).$$

We define $\hat{M} = |t\rangle \langle \lambda|$ and check the convergence condition for λ.

Step one is always checking condition of convergence.

Step two is to write down the series and zeroth order. Then we reach the key point. The iteration relation is

$$u_n(t) + \int_0^1 ds s u_{n-1}(s) = 0.$$

One can write down u_1 imediately

$$u_1(t) = -\int_0^1 ds s u_0(s).$$

Keep on going.

Using Dyads in Vector Space

For the same example,

$$\hat{L}u(x) = f(x),$$

where $\hat{L} = \hat{I} + \hat{M}$, we can solve it using vector space if operator is linear.

Suppose we have a $\hat{M} = |a\rangle \langle b|$, the equation, in some Hilbert space, is

$$|u\rangle + |a\rangle \langle b \mid u\rangle = |f\rangle .$$

Multiplying through by $\langle b|$, we have

$$\langle b \mid u\rangle + \langle b \mid a\rangle\langle b \mid u\rangle = \langle b \mid f\rangle,$$

which reduces to a linear equation. We only need to solve out $\langle b \mid u\rangle$ then plug it back into the original equation.

5.4 Complex Analysis

Some useful concepts:[1]

- Representation of a complex number and its conjugate
- Complex functions
- curves, closed curves, simple curves
- Ininity point
- Analytic functions: depends only on z not its complex conjugate
- Entire function: single-valued analytic all over C
- Liouville theorem
- Pole
- Singularity, Essential Singularity
- Meromorphic function

For multi-valued functions,

- A branch of a function
- Analyticity of multi-valued function
- Branch point
- Cut

Operations

- Contour integral of a continuous function arround some simple curve
- Cauchy's Integral Theorem

[1] A handout note by Finly

5.4.1 Cauchy-Riemann Equation

A function $f(z) = u(z) + iv(z)$ is a function of a complex variable $z = x + iy$.

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

5.4.2 Singularities

There are 3 common singularities,

1. Pole
2. Branch point
3. Essential singularity

Pole is very useful since it's related to the Residue Theorem. Thus one of the task in physics is to calculate the residue of a function.

The residue at a simple pole is given by

$$\text{Residue}(f(z_0)) = \lim_{z\to z_0}\left((z - z_0)f(z)\right).$$

Meanwhile, the residue at a pole of nth order is

$$\text{Residue}(f(z_0)) = \frac{1}{(n-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{n-1}}{\mathrm{d}z^{n-1}}\left((z - z_0)^n f(z)\right).$$

Branch points are points when we go around it in circles the values of our function would change. Examples of such points are $z = 0$ for $f(z) = ln(z)$ and $z = 1$ for $f(z) = (z - 1)^{1/2}$.

5.4.3 Refs & Notes

5.5 Calculus

5.5.1 Differential of Functions

5.5.2 Integrals

Sometimes a integral on Real plane can be very hard, one of the techniques is to work on Complex plane and use contour integral.

1. Contours: use Ghost Contours so that we don't need to calculate these complicated integrals.
2. Branch Cut: cuts are needed if we have got branch points on the complex plane.
3. Residue Theorem: we can write down the integral by calculating the residue of the integrand,

   $$\int_C f(z)\mathrm{d}z = 2\pi i\sum_j \text{Residue}(f(z_j)),$$

   where z_j are the poles.

Calculus on One Page

Differential	Integral
Fundamental Theorem of Calculus	
$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t = f(x)$	$\int_a^x \frac{\mathrm{d}}{\mathrm{d}t} f(t)\mathrm{d}t = f(x) - f(a)$
Rules	
$[c_1 u(x) + c_2 v(x)]' = c_1 u'(x) + c_2 v'(x)$	$\int [c_1 f(x) + c_2 g(x)]\,\mathrm{d}x = c_1 \int f(x)\mathrm{d}x + c_2 \int g(x)\mathrm{d}x$
$(uv)' = u'v + uv'$	$\int u'v\mathrm{d}x = uv - \int uv'\mathrm{d}x$
$\frac{\mathrm{d}f(u(x))}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}$	$\int \frac{\mathrm{d}f}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x}\mathrm{d}x = f(u(x)) + c$
Mean Value Theorem	
$\frac{F(b)-F(a)}{b-a} = F'(\xi)$	$\int_a^b f(x)\mathrm{d}x = (b-a)f(\xi)$
Useful Equations	
$(x^n)' = nx^{n-1}$	$\int x^n \mathrm{d}x = \frac{x^{n+1}}{n+1} + c$
$(\sin x)' = \cos x$	$\int \cos x \mathrm{d}x = \sin x + c$
$(\ln x)' = \frac{1}{x}$	$\int \frac{1}{x}\mathrm{d}x = \ln\|x\| + c$
$(e^x)' = e^x$	$\int e^x \mathrm{d}x = e^x + c$
Multivariable Calculus (Fundamental Theorem)	
$\int_{\partial\Omega} \boldsymbol{\omega} = \int_\Omega \mathbf{d}\boldsymbol{\omega}$	

Notes:

- One can derive other useful equation using just these in this table.
- The three equations of fundamental theorem for multivariable calculus can be derived using the exterior derivative.

Fig. 5.8: LaTeX source of this image is here .

5.6 Linear Algebra

5.6.1 Basic Concepts

Trace

Trace should be calculated using the metric. An example is the trace of Ricci tensor,

$$R = g^{ab} R_{ab}$$

Einstein equation is

$$R_{ab} - \frac{1}{2} g_{ab} R = 8\pi G T_{ab}$$

The trace is

$$\begin{aligned} g^{ab} R_{ab} - \frac{1}{2} g^{ab} g_{ab} R &= 8\pi G g^{ab} T_{ab} \\ \Rightarrow R - \frac{1}{2} 4R &= 8\pi G T \\ \Rightarrow -R &= 8\pi G T \end{aligned}$$

Determinant

Some useful properties of determinant.

1. Interchange rows (colomns) once will generate a negative sign.
2. Determinant can be calculated recursively when implemented numerically.
3. Determinant for block matrix can be expressed using the blocks.

Here is an example of the determinant of block matrix. Suppose our block matrix is

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where each block is a square matrix. We calculate the determinant through

$$\mathrm{Det}(A) = \mathrm{Det}(BE - CD).$$

This is useful when we have a block diagonalized matrix.

5.6.2 Technique

Inverse of a matrix

Many methods to get the inverse of a matrix. Check wikipedia for Invertible matrix.

Adjugate matrix method for example is here.

$$A^{-1} = \frac{A^*}{|A|}$$

in which, A^* is the adjugate matrix of A.

Eigenvalues of $A^\dagger A$

One can prove that the eigenvalues of any matrix B that can be written as $A^\dagger A$ are positive semidefinite.

Proof

Suppose the eigenvectors are V_i with corresponding eigenvalues λ_i, i.e.,

$$BV_i = \lambda_i V_i.$$

We now construct a number

$$V_i^\dagger B V_i.$$

On one hand, we have

$$V_i^\dagger B V_i = V_i^\dagger \lambda_i V_i = \lambda_i V_i^\dagger V_i,$$

where $V_i^\dagger V_i \geq 0$.

On the other hand,

$$V_i^\dagger B V_i = V_i^\dagger A \dagger A V_i = (AV_i)^\dagger A V_i \geq 0.$$

As long as $V_i^\dagger V_i \neq 0$, we have

$$\lambda_i = (AV_i)^\dagger A V_i / V_i^\dagger V_i \geq 0.$$

5.6.3 Tensor Product Space

$|\phi\rangle_1$ and $|\phi\rangle_2$ are elements of Hilbert space H_1 and H_2. **Tensor Product** of $|\phi\rangle_1$ and $|\phi\rangle_2$ is denoted as $|\phi\rangle_1 \otimes |\phi\rangle_2$. This operation is linear and distributive.

Tensor product space $H_1 \otimes H_2$ is composed of all the linear combinations of all possible tensor products of elements in H_1 and H_2.

Inner Product

Inner product of two tensor products

$$(\langle\phi|_1 \otimes \langle\phi|_2)(|\psi\rangle_1 \otimes |\psi\rangle_2) = ({}_1\langle\phi \mid \psi\rangle_1)({}_2\langle\phi \mid \psi\rangle_2)$$

Operators Applied to Tensor Product

Two operators $\hat{O}_1$ and $\hat{O}_2$ works on H_1 and H_2 respectively applied to tensor product

$$(\hat{O}_1 \otimes \hat{O}_2)(|\phi\rangle_1 \otimes |\phi\rangle_2) = (\hat{O}_1 |\phi\rangle_1) \otimes (\hat{O}_2 |\phi\rangle_2)$$

5.6.4 Solving Linear Equations

First of all, write down the augmented matrix for the equation set.

Elementary row operations are allowed on the augmented matrix. Operate on the matrix until one can read out the solutions.

5.7 Differential Geometry

5.7.1 Metric

Definitions

Denote the basis in use as $\hat{e}_\mu$, then the metric can be written as

$$g_{\mu\nu} = \hat{e}_\mu \hat{\cdot} e_\nu$$

if the basis satisfies

Inversed metric

$$g_{\mu\lambda} g^{\lambda\nu} = \delta_\mu^\nu = g_\mu^\nu$$

How to calculate the metric

Let's check the definition of metric again.

If we choose a basis $\hat{e}_\mu$, then a vector (at one certain point) in this coordinate system is

$$x^a = x^\mu \hat{e}_\mu$$

Then we can construct the expression of metric of this point under this coordinate system,

$$g_{\mu\nu} = \hat{e}_\mu \cdot \hat{e}_\nu$$

For example, in spherical coordinate system,

$$\vec{x} = r \sin\theta \cos\phi \hat{e}_x + r \sin\theta \sin\phi \hat{e}_y + r \cos\theta \hat{e}_z \tag{5.6}$$

Now we have to find the basis under spherical coordinate system. Assume the basis is $\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi$. Choose some scale factors $h_r = 1, h_\theta = r, h_\phi = r\sin\theta$. Then the basis is

$$\hat{e}_r = \frac{\partial \vec{x}}{h_r \partial r} = \hat{e}_x \sin\theta\cos\phi + \hat{e}_y \sin\theta \sin\phi + \hat{e}_z \cos\theta,$$

etc. Then collect the terms in formula (5.6) is we get $\vec{x} = r\hat{e}_r$, this is incomplete. So we check the derivative.

$$\begin{aligned}
\mathrm{d}\vec{x} = & \hat{e}_x(\mathrm{d}r \sin\theta\cos\phi + r\cos\theta\cos\phi\mathrm{d}\theta - r\sin\theta\sin\phi\mathrm{d}\phi) \\
& \hat{e}_y(\mathrm{d}r \sin\theta\sin\phi + r\cos\theta\sin\phi\mathrm{d}\theta + r\sin\theta\cos\phi\mathrm{d}\phi) \\
& \hat{e}_z(\mathrm{d}r\cos\theta - r\sin\theta\mathrm{d}\theta) \\
= & \mathrm{d}r(\hat{e}_x \sin\theta\cos\phi + \hat{e}_y\sin\theta\sin\phi - \hat{e}_z\cos\theta) \\
& \mathrm{d}\theta(\hat{e}_x\cos\theta\cos\phi + \hat{e}_y\cos\theta\sin\phi - \hat{e}_z\sin\theta)r \\
& \mathrm{d}\phi(-\hat{e}_x\sin\phi + \hat{e}_y\cos\phi)r\sin\theta \\
= & \hat{e}_r\mathrm{d}r + \hat{e}_\theta r\mathrm{d}\theta + \hat{e}_\phi r\sin\theta\mathrm{d}\phi
\end{aligned}$$

Once we reach here, the component (e_r, e_θ, e_ϕ) of the point under the spherical coordinates system basis $(\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi)$ at this point are clear, i.e.,

$$\begin{aligned}
\mathrm{d}\vec{x} &= \hat{e}_r\mathrm{d}r + \hat{e}_\theta r\mathrm{d}\theta + \hat{e}_\phi r\sin\theta\mathrm{d}\phi \\
&= e_r\mathrm{d}r + e_\theta\mathrm{d}\theta + e_\phi\mathrm{d}\phi
\end{aligned}$$

In this way, the metric tensor for spherical coordinates is

$$g_{\mu\nu} = (e_\mu \cdot e_\nu) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

5.7.2 Connection

First class connection can be calculated

$$\Gamma^{\mu}{}_{\nu\lambda} = \hat{e}^{\mu} \cdot \hat{e}_{\mu,\lambda}$$

Second class connection isfootnote{Kevin E. Cahill}

$$[\mu\nu, \iota] = g_{\iota\mu}\Gamma^{\mu}{}_{\nu\lambda}$$

5.7.3 Gradient, Curl, Divergence, etc

Gradient

$$T^{b}{}_{c;a} = \nabla_a T^{b}{}_{c} = T^{b}{}_{c,a} + \Gamma^{b}_{ad} T^{d}{}_{c} - \Gamma^{d}_{ac} T^{b}{}_{d}$$

Curl

For an anti-symmetric tensor, $a_{\mu\nu} = -a_{\nu\mu}$

$$\begin{aligned}\mathrm{Curl}_{\mu\nu\tau}(a_{\mu\nu}) &\equiv a_{\mu\nu;\tau} + a_{\nu\tau;\mu} + a_{\tau\mu;\nu} \\ &= a_{\mu\nu,\tau} + a_{\nu\tau,\mu} + a_{\tau\mu,\nu}\end{aligned}$$

Divergence

$$\begin{aligned}\mathrm{div}_{\nu}(a^{\mu\nu}) &\equiv a^{\mu\nu}{}_{;\nu} \\ &= \frac{\partial a^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\nu\tau} a^{\tau\nu} + \Gamma^{\nu}_{\nu\tau} a^{\mu\tau} \\ &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\nu}} (\sqrt{-g} a^{\mu\nu}) + \Gamma^{\mu}_{\nu\lambda} a^{\nu\lambda}\end{aligned}$$

For an anti-symmetric tensor

$$\mathrm{div}(a^{\mu\nu}) = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\nu}} (\sqrt{-g} a^{\mu\nu})$$

Annotation Using the relation $g = g_{\mu\nu} A_{\mu\nu}$, $A_{\mu\nu}$ is the algebraic complement, we can prove the following two equalities.

$$\Gamma^{\mu}_{\mu\nu} = \partial_{\nu} \ln \sqrt{-g}$$

$$V^{\mu}{}_{;\mu} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\mu}} (\sqrt{-g} V^{\mu})$$

In some simple case, all the three kind of operation can be demonstrated by different applications of the del operator, which $\nabla \equiv \hat{x}\partial_x + \hat{y}\partial_y + \hat{z}\partial_z$.

- Gradient, ∇f, in which f is a scalar.
- Divergence, $\nabla \cdot \vec{v}$
- Curl, $\nabla \times \vec{v}$
- Laplacian, $\Delta \equiv \nabla \cdot \nabla \equiv \nabla^2$

5.8 Statistics

5.8.1 Famous Distributions

- Binomial distribution
- Poisson Distribution
- Chi-squared Distribution

CHAPTER 6

Classical Mechanics

6.1 Lagrangian and Equation of Motion

Euler-Lagrangian equation is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) - \frac{\partial L}{\partial \boldsymbol{q}} = 0. \tag{6.1}$$

The component form is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{6.2}$$

Conserved Quantities

A quantity is conserved through time if $\frac{d}{dt}Q = 0$.

We notice that the second term in (6.2) vanishes if the lagragian doesn't depend on q_i. That is

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

for Lagragian that doesn't depend on q_i.

We immediately spot that the quantity

$$\frac{\partial L}{\partial \dot{q}_i}$$

is a conserved quantity.

6.2 Regid Body

Center of mass $\vec{R}$ is defined as

$$\int \rho(\vec{r})\vec{r}d^3\vec{r} = \int \rho(\vec{r})\vec{R}d^3\vec{r}.$$

Equivalently,

$$\int \rho(\vec{r})\left(\vec{r} - \vec{R}\right)d^3\vec{r} = 0.$$

6.3 Central Force Fields

Central force fields are widely used in physics and they have simple yet important properties.

In general, central force is described using

$$\vec{F}(\vec{r}) = f(r)\hat{r}.$$

The Lagrangian for an object of mass m in a central force field is

$$\begin{aligned} L &= \frac{1}{2}m\dot{\mathbf{r}}^2 - V(r) \\ &= \frac{1}{2}m(\dot{\mathbf{r}}^2 + r^2\theta^2) - V(r). \end{aligned}$$

The interesting thing for such a system is that there is always a conserved quantity since the Lagrangian has no explicit θ dependence. It is obvious that

$$\frac{\partial L}{\partial \theta} = 0.$$

Now we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = 0,$$

which leads to the conservation of angular momentum as the first equation of motion,

$$\dot{l} \equiv \dot{p}_\theta = \frac{d}{dt}\left(mr^2\dot{\theta}\right) = 0$$

The second equation of motion is given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0,$$

which simplifies to

$$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V(r)}{\partial r} = 0.$$

Applying the conserved quantity, we find an effective potential

$$V_{eff}(r) = V(r) + \frac{1}{2}\frac{l^2}{mr^2}.$$

6.4 Oscillators

In general, the Lagragian for a system with n general coordinates can be

$$L = \frac{1}{2} m_{jk} \dot{q}_j \dot{q}_k - V(q_1, \cdots, q_n)$$

To write down equation of motion, we need the following terms,

$$\frac{\partial L}{\partial \dot{q}_j} = m_{jk} \dot{q}_k \frac{\partial L}{\partial q_j} = \frac{1}{2} \frac{\partial m_{kl}}{\partial q_j} \dot{q}_k \dot{q}_l - \frac{\partial V}{\partial q_j}$$

Then equation of motion is

$$m_{jk} \ddot{q}_k + \frac{\partial m_{jk}}{\partial q_l} \dot{q}_k \dot{q}_l - \frac{1}{2} \frac{\partial m_{kl}}{\partial q_j} \dot{q}_k \dot{q}_l = -\frac{\partial V}{\partial q_j}$$

Generally, we can't solve this system. But there is an interesting limit. The system may have equilibrium points. We can study systems oscillating around equilibrium points.

At equilibrium, the system can stay steady, i.e., $\dot{q}_j^0 = 0$. This gives us

$$\frac{\partial V}{\partial q_j} = 0,$$

for all j.

Now for small deviations, we can expand the system around equilibrium points.

$$q_j = q_j^0 + \eta_j$$

Then

$$T = \frac{1}{2} m_{jk}|_0 \dot{\eta}_j \dot{\eta}_k \equiv \frac{1}{2} T_{jk} \dot{\eta}_j \dot{\eta}_k$$

$$V = V|_0 + \frac{\partial V}{\partial q_j}|_0 \eta_j + \frac{1}{2} \frac{\partial^2 V}{\partial q_j \partial q_k}|_0 \eta_j \eta_k + \cdots \equiv \frac{1}{2} V_{jk} \eta_j \eta_k$$

So we have the Lagrangian for small oscillations,

$$L = \frac{1}{2} T_{jk} \dot{\eta}_j \dot{\eta}_k - \frac{1}{2} V_{jk} \eta_j \eta_k$$

Typing indices using LaTeX is so annoying. So we'll use matrix notations and Lagragian becomes

$$L = \frac{1}{2} \dot{\tilde{\eta}} T \dot{\eta} - \frac{1}{2} \tilde{\eta} V \eta,$$

in which T and V matrices are n by n real and symmetric.

(We need to diagonalize T and V. First question comes to us is:

Is is possible to diagonalize both T and V at the same time?

We can have a look at the surface $\tilde{p} T p = C$, which is a elliptical surface with coordinates p.)

Use the following transformation

$$\xi = T^{1/2} \eta$$

Then transpose

$$\tilde{\xi} = \tilde{\eta} T^{1/2}$$

$$\dot{\tilde{\xi}}\dot{\xi} = \dot{\tilde{\eta}} T \dot{\eta}$$

So we have the new Lagragian

$$L = \frac{1}{2}\dot{\tilde{\xi}}\dot{\xi} - \frac{1}{2}\tilde{\xi} T^{-1/2} V T^{-1/2} \xi$$

Define $T^{-1/2} V T^{-1/2} \equiv V'$.

Next we need to diagonalize V' by using its eigen vectors.

$$V' b = \lambda b$$

is equivalent to

$$V a = \lambda T a$$

with $b = T^{1/2} a$. So we have

$$\det(V' - \lambda \mathbf{I}) = 0$$

is same as

$$\det(V - \lambda T) = 0$$

in which λ is the eigen value of this function.

6.4.1 Simplest Harmonic Oscillators

Harmonic oscillators are described by

$$-kx = m\ddot{x},$$

which has solution

$$x = x(t=0) e^{i\omega x},$$

where $\omega = \pm\sqrt{\frac{k}{m}}$ and the final solution is determined by the second initial condition, i.e., the first order derivative of displacement.

6.5 Refs & Notes

6.6 Hamiltonian Dynamics

Hamiltonian equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$
$$\dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

Some constant of motion can be read out from the equations by recogonizing the fact that the time derivative of a constant of motion, q_i or p_i, is zero. For example, if the Hamiltonian doesn't explicitly depend on p_k, we have $\frac{\partial H}{\partial p_k} = 0 = \dot{q}_k$, which means that q_k is a constant of motion.

The evolution of the system in phase space obeys the Liouville's theorem, which describes the motion of phase space density $\rho(\{q_i\}, \{p_i\}, t)$,

$$\frac{d\rho}{dt} = 0.$$

Phase Space Density

The probability that the system will be found in a phase space interval $d^n p d^n q$ is given by $\rho(\{q_i\}, \{p_i\}, t) d^n p d^n q$.

CHAPTER 7

Quantum Mechanics

7.1 Preliminary

7.1.1 Quantum Vocabulary

Vocabulary of physics, the fountain of research ideas.

0. Fine Structure Constant

$\alpha = \frac{k_e e^2}{\hbar c} = \frac{1}{(4\pi\varepsilon_0)}\frac{e^2}{\hbar c} = \frac{e^2 c\mu_0}{2h}$

In electrostatic cgs units, $\alpha = \frac{e^2}{\hbar c}$.

In natural units, $\alpha = \frac{e^2}{4\pi}$.

1. Hydrogen Atom

Potential $V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}$.

Energy levels: $E_n = -\left(\frac{Z^2\mu e^4}{32\pi^2\epsilon_0^2\hbar^2}\right)\frac{1}{n^2} = -\left(\frac{Z^2\hbar^2}{2\mu a_\mu^2}\right)\frac{1}{n^2} = \frac{\mu c^2 Z^2\alpha^2}{2n^2}$.

Ground state of hydrogen atom $\psi_{100}(r) = \frac{1}{\sqrt{\pi}}\frac{1}{a^{3/2}}e^{-Zr/a}$.

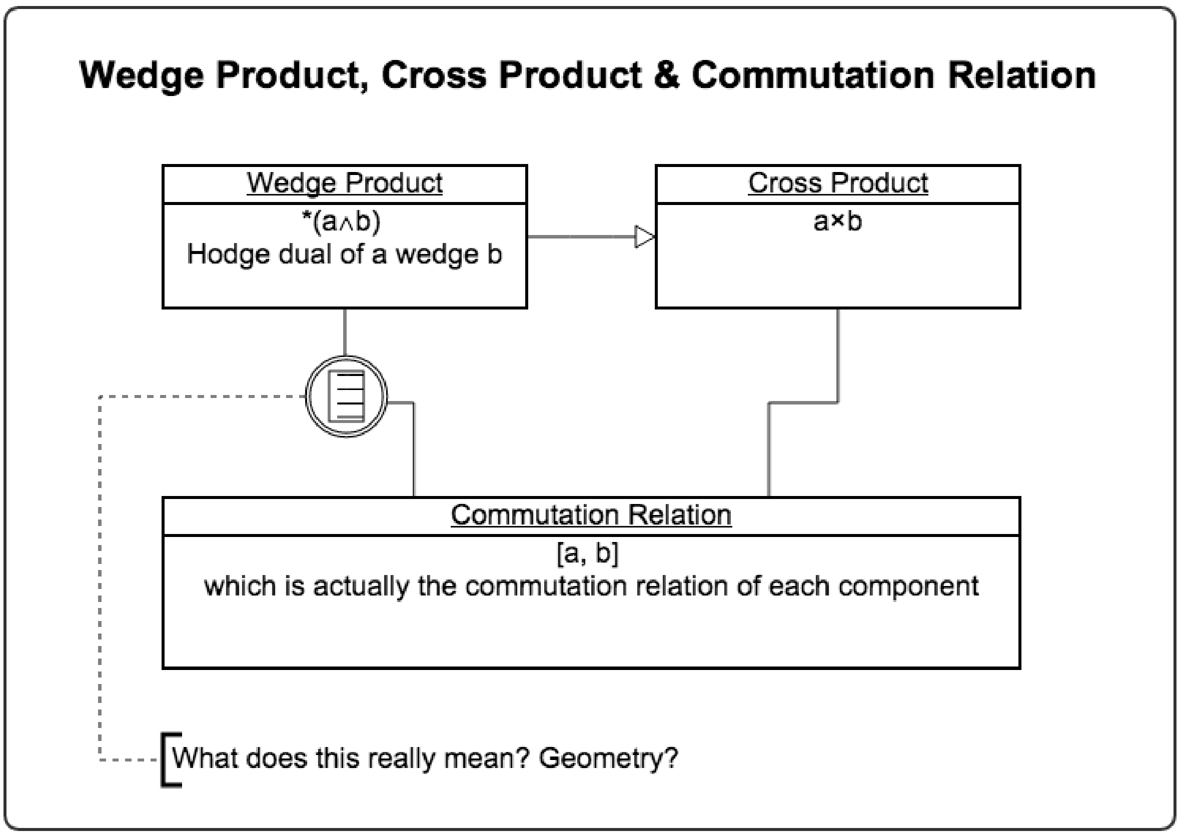


Fig. 7.1: Geometry language here?

7.1.2 Quantum Questions

Wedge Product, Cross Prodcut & Commutation relation

7.1.3 Tensors and Groups in Quantum

A rank-k tensor $\hat{T}_k^q$ is defined as

$$\begin{aligned} \left[\hat{J}_z, \hat{T}_k^1\right] &= q\hbar \hat{T}_k^q \\ \left[\hat{J}_\pm, \hat{T}_k^q\right] &= \sqrt{(k \mp q)(k \pm q + 1)}\hbar \hat{T}_k^{q\pm 1}. \end{aligned}$$

Wigner-Eckart Theorem

Wigner-Eckart theorem is

$$\langle n'j'm'| \hat{T}_k^q |njm\rangle = \langle n'j'| |\hat{T}_k| |nj\rangle \langle j'm'; kj \mid kq; jm\rangle,$$

where j, j' are the angular momentum quantum numbers and n, n' are quantum numbers which are not related to angular momentum.

It seems that tensor $\hat{T}_k^q$ is a source of angular momentum. The maximum angular momentum it can provide is k.

7.2 Quantum Mechanics Beginners

7.2.1 Fundamental Concepts

What're the most important tricks in QM calculations?

- **Remember what basis we are working in**
- **Identity**

First Three Postulates

- Physical state is described by kets in a Hilbert space. We need to specify a complete basis $\{|i\rangle\}$ to do calculations.

$$|\psi\rangle = \sum_i |i\rangle \langle i| |\psi\rangle = \sum_i C_i |i\rangle$$

- Operators are given by Hermitian operators; A measurement of the variable $\hat{\Omega}$ will yield one of the eigenvalues ω with the probability

$$|\langle \omega \mid \psi\rangle|^2 .$$

  And the state of the system will change to $|\omega\rangle$.

- The state vector obeys the Schrödinger equation,

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle ,$$

where $\hat{H}$ is the Hamiltonian operator.

Comments

The logic here is that we first find the way to describe a system, then think about how to find out the information we need from the state vector and also find the evolution of the state vector. Then we need the operator and Schrodinger equation. Finally, we would like to relate the theory to experiments, and it comes the measurement postulate.

Later we will need the relation between position and momentum, which becomes the fourth postulate.

- How to solve the evolution of a system? We just define a magical operator, propagator

$$\hat{U} |\psi(t_0)\rangle = |\psi(t)\rangle .$$

This operator just gives us the evolution of state vector! Wait, can we write down the explicit expression of it?

Let's find out. The only thing we know about the evolution of a state vector is the third postulate up there.

$$\begin{aligned} i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle &= \hat{H} |\psi(t)\rangle \\ i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \hat{U} |\psi(t_0)\rangle &= \hat{H}\hat{U} |\psi(t_0)\rangle \\ i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \hat{U} &= \hat{H}\hat{U} \end{aligned}$$

Looks familiar? This just gives us a exponential result, **if the Hamiltonian is time independent**.

$$\hat{U} = e^{-i\hat{H}(t-t_0)/\hbar}$$

We can prove that this operator is Unitary because $\hat{H}$ is Hermitian.

This is just the abstract representation, we work in some basis, and the most convenient basis is the eigenstates of Hamiltonian, { $|\epsilon_i\rangle$ },

$$\begin{aligned} \hat{U} |\phi\rangle &= e^{-i\hat{H}(t-t_0)/\hbar} |\psi\rangle \\ \hat{U} |\phi\rangle &= \sum_i e^{-i\hat{H}(t-t_0)/\hbar} |\epsilon_i\rangle \langle\epsilon_i| |\psi\rangle \\ \hat{U} |\phi\rangle &= \sum_i e^{-i\epsilon_i(t-t_0)/\hbar} |\epsilon_i\rangle \langle\epsilon_i| |\psi\rangle \end{aligned}$$

And we are going to use

$$\hat{U} = \sum_i e^{-i\epsilon_i(t-t_0)/\hbar} |\epsilon_i\rangle \langle\epsilon_i|$$

from now on. (Well, only on discrete eigenvalues ones.)

(See that? Identity does the work again.)

Position and Momentum Space

Summary

- **Position**
    1. Define $\{|x\rangle\}$ basis.
    2. Define $\hat{x}$ operator.
    3. Find wave function in this basis.
    4. Find measurement.
- **Evolution**
    1. Need propagator $\hat{U}$.
    2. Propagator needs the solution of Hamiltonian eigensystem.
    3. (Free particles) Hamiltonian needs the solution of momentum eigensystem.
- **Momentum**
    1. Before we define some arbitrary momentum space, we should check the relation between momentum and position. And it turns out to be related by a commutator.(Postulate IV)
    2. Use the postulate to momentum operator.
    3. Find eigenstates.
    4. (Calculate the propagator.)

Position Space

1. Define $|x\rangle$ basis.

   Orthonormal condition is

   $$\langle x \mid x' \rangle = \delta(x - x').$$

   Completeness condition is

   $$\int \langle x' \mid x' \rangle \mathrm{d}x' = \mathbb{I}$$

2. Define position operator.

   The position operator is defined as

   $$\hat{x} \,|x\rangle = x \,|x\rangle$$

   And in $\{|x\rangle\}$ basis, this operator becomes a function, which is

   $$\begin{aligned}&\langle x| \,\hat{x}\, |x'\rangle \\ =& \left(\langle x| \,\hat{x}\right) |x'\rangle \\ =& x\langle x \mid x' \rangle \\ =& x\delta(x - x')\end{aligned}$$

3. Find state vector in $\{|x\rangle\}$ basis.

$$\psi(t,x) = \langle x \mid \psi(t)\rangle$$

Normalized

$$\int |\psi(t,x)|^2 \,\mathrm{d}x = 1.$$

And we are interpreting $|\psi(t,x)|^2$ as probability density.

4. Calculate probability of a measurement. Taking $\hat{x}$ as an example.

$$\begin{aligned}
&\langle\psi|\,\hat{x}\,|\psi\rangle \\
&= \iint \langle\psi \mid x\rangle \langle x|\,\hat{x}\,|x'\rangle \langle x' \mid \psi\rangle \mathrm{d}x\mathrm{d}x' \\
&= \iint \psi^*(t,x) x \delta(x-x')\psi(t,x')\mathrm{d}x\mathrm{d}x' \\
&= \int |\psi(t,x)|^2 \, x\mathrm{d}x
\end{aligned}$$

Momentum Space

To find the momentum operator, we need to check the relation between momentum and position before we just randomly define one. Truth is, we have a fourth postulate states the relation between them.

Postulate IV

The commutator of $\hat{x}$, $\hat{p}$ is

$$[\hat{x},\hat{p}] = i\hbar$$

Two comments:

- Why i ? Eigenvalue of Anti-Hermitian operator.
- Why $\hbar$? Because people define the dimensions of position and momentum differently before they know this commutator. We would like to assign them the same dimension if we already know this relation.

Momentum Space

1. Find momentum operator in position basis $\{|x\rangle\}$.

$$\langle x|\,[\hat{x},\hat{p}]\,|x'\rangle = i\hbar\delta(x-x')$$

And write out the commutator and use the relation of delta function $x\delta'(x) = -\delta(x)$, we find out the momentum operator in $\{|x\rangle\}$ basis,

$$\langle x|\,\hat{p}\,|x'\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\delta(x-x')$$

Let's talk physics. What does that operator mean? We need to see what the result is when momentum operator is applied to a state. And remember we would work in $\{|x\rangle\}$ basis.

$$
\begin{aligned}
&\langle x|\hat{p}|\psi\rangle \\
&= \iint \langle x \mid x'\rangle \langle x'|\hat{p}|x''\rangle \langle x'' \mid \psi\rangle \mathrm{d}x'\mathrm{d}x'' \\
&= \int \langle x|\hat{p}|x''\rangle \psi(t,x'')\mathrm{d}x'' \\
&= \int \left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x')\psi(t,x')\right)\mathrm{d}x' \\
&= \int \left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x'}\delta(x'-x)\psi(t,x')\right)\mathrm{d}x'
\end{aligned}
$$

Integrate by parts, we will find the expression. (I am having a problem finding the right answer.)

$$\langle x|\hat{p}|\psi\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x).$$

2. Eigenfunction for momentum.

$$\hat{p}|p\rangle = p|p\rangle .$$

Again, we are going to project it on the $\{|x\rangle\}$ basis,

$$\langle x|\hat{p}|p\rangle = \langle x|p|p\rangle ,$$

where $\langle x \mid p\rangle$ is the eigenstates in $\{|x\rangle\}$ basis, we call it $\phi_p(x)$.

$$
\begin{aligned}
\langle x|\hat{p}|p\rangle &= p\phi_p(x) \\
\int \langle x|\hat{p}|x'\rangle \langle x' \mid p\rangle \mathrm{d}x' &= p\phi_p(x) \\
-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\phi_p(x) &= p\phi_p(x)
\end{aligned}
$$

The solution is

$$\phi_p(x) = \mathrm{C}e^{ipx/\hbar}$$

This constant C is found by the normalization condition,

$$\langle p \mid p'\rangle = \int \phi_p^*(x)\phi_{p'}(x)\mathrm{d}x = \delta(p-p')$$

The final results should be

$$\phi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(ipx/\hbar\right)$$

3. Find the dynamics of free particles in quantum mechanics. **Find the propagator and everything solves.** The hamiltonian for a free particle is

$$\hat{H} = \frac{\hat{p}^2}{2m}.$$

We argue here that the eigenvectors of momentum are also the eigenvectors of this hamiltonian. And we can easily guess the eigenvalues are $p^2/2m$. So the propagator is

$$\hat{U} = \int e^{-ip^2t/2m\hbar}|p\rangle\langle p|\,\mathrm{d}p$$

But that is too abstract to use, we can find the expression in $\{|x\rangle\}$ basis.

$$\begin{aligned}\langle x|\hat{U}|x\rangle &= \int e^{-ip^2t/2m\hbar}\langle x \mid p\rangle\langle p \mid x\rangle \mathrm{d}p \\ &= \int e^{-ip^2t/2m\hbar}|\phi_p|^2 \,\mathrm{d}p\end{aligned}$$

7.2.2 Quantum in 1D

General

Always start with the propagator for time independent Hamiltonian.

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

For cases that Hamiltonian with discrete eigenvalues,

$$|\psi(t)\rangle = \sum_n e^{-i\epsilon_n t/\hbar}\,|n\rangle\,\langle n \mid \psi(0)\rangle$$

If the initial state is just one of the eigenstates of Hamiltonian, say the mth one (normalized),

$$|\psi(t)\rangle = e^{-i\epsilon_m t/\hbar}\,|m\rangle$$

Well, that phase factor doesn't have any effect for the topic we discuss. So our time evolution will stay on the same state forever.

The same thing happens for continuous cases.

So our task is simplified to solve the eigensystem of Hamiltonian, which is

$$\hat{H}\,|\epsilon\rangle = \epsilon\,|\epsilon\rangle$$

Infinite Barriers

Math

Setup

- Potential in a box

$$\begin{aligned}V(x) =& 0, 0 < x < L \\ & \infty, \text{Other}\end{aligned}$$

Solve the Problem

- Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$$

- Dynamic equation

  $$\hat{H} |\psi(t)\rangle = \epsilon |\psi(t)\rangle$$

  We are happy to work in $\{|x\rangle)\}$ basis,

  $$\langle x| \hat{H} |\psi(t)\rangle = \langle x| \epsilon |\psi(t)\rangle .$$

  Put the Hamiltonian in, and remember that in position basis

  $$\langle x| \hat{p} |\psi\rangle = -i\hbar \frac{\mathrm{d}}{\mathrm{d}x}\psi,$$

  the equation of motion becomes

  $$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x,t) + V(x)\psi(x,t) = \epsilon\psi(x,t)$$

- Boundary conditions

  $$\psi_I(0,t) = \psi_{II}(0,t)$$
  $$\psi_{II}(L,t) = \psi_{III}(L,t)$$

- Guess the Solutions

  $$\psi_{II} = \psi = C\sin(kx) + D\cos(kx)$$

- Find the wavenumber k, by putting the assumed solutions into equation of motion

  $$k = \pm\sqrt{\frac{2m\epsilon}{\hbar^2}}$$

  Since we can always merge the negative into the constants, it is fine to use

  $$k = \sqrt{\frac{2m\epsilon}{\hbar^2}}$$

- Use Boundary Condition 1. At x=0,

  $$\psi(0,t) = 0.$$

  This gives us $D = 0$.

  2. At $x = L$,

     $$\psi(L,t) = 0.$$

     This leads to

     $$kL = n\pi.$$

     Since $n = 0$ gives us a 0 wave function, we would just drop $n = 0$. For the same reason why we drop the negative values of k, we would drop all the negative values of n. This BC gives us the possible values of energy because wavenumber k is related to energy,

     $$\epsilon = \frac{\hbar^2}{2mL^2}(n\pi)^2,$$

     with

     $$n = 1, 2, 3, \cdots$$

- Normalization as the last constraint for the last undetermined parameter,

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

Physics

1. Estimation
   - Find the expression for energy using dimensional analysis.
   - Using uncertainty relation to estimate the expression for energy.
2. Comments
   - Why is the solution quantized? 1. Too many constraints. BCs + normalization.
   - **Why do the n in the solution goes into the expression for energy?**
     (a) Have a look at the kinetic energy term, the derivative does it.
   - **What's so weird?**
     (a) For $n = 2$, no particles found at $x = L/2$. And so on.

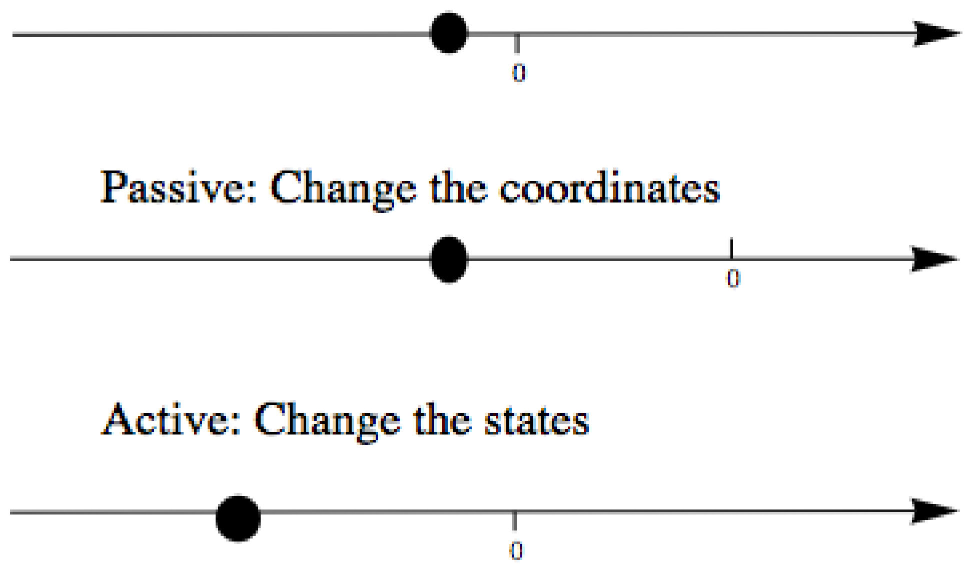
Some General Properties

1. 1D bound states have no degeneracy. Prove it by assume that there is a degeneracy state.
2. 1D bound states' wave function can be chosen to be real. (if potential V is real.)

7.2.3 Parity

Passive and Active Transformations

Generally, there are two ways of interpreting a transformation.



Here in QM, passive means transform the operator $\hat{\Omega}$, while active means change the state $|\psi\rangle$. Suppose we have a system $|\psi\rangle$, an operator $\hat{\Omega}$, a transformation $\hat{U}$.

Transformation $\hat{U}|\psi\rangle$ is identical to $\hat{U}^\dagger\hat{\Omega}\hat{U}$ because they give the same observation results. The first one is called active, while the second one is called passive.

Parity

Definition

$$\hat{\Pi}|x\rangle = |-x\rangle$$

Properties

1. Act on momentum eigenvectors,

$$\hat{\Pi}|p\rangle = |-p\rangle .$$

- Physics: Parity changes the coordinate, so the direction of momentum is also changed.

- Math:

$$\hat{\Pi}\left|p\right\rangle = \int \hat{\Pi}\left|x\right\rangle \left\langle x \mid p\right\rangle \mathrm{d}x = \int \left|-x\right\rangle \left\langle x \mid p\right\rangle \mathrm{d}x$$

  Change coordinate from x to -x,

$$\hat{\Pi}\left|p\right\rangle = \int \left|x\right\rangle \left\langle -x \mid p\right\rangle \mathrm{d}x = \int \left|x\right\rangle \left\langle x \mid -p\right\rangle \mathrm{d}x = \left|-p\right\rangle$$

2. Hermitian,

$$\left\langle x\right| \hat{\Pi} \left|x'\right\rangle = \delta(x+x')(\left\langle x'\right| \hat{\Pi} \left|x\right\rangle)^{\dagger} = \left\langle x\right| \hat{\Pi}^{\dagger} \left|x'\right\rangle = \delta(x+x')$$

3. Unitary

$$\left\langle x\right| \hat{\Pi}^{\dagger}\hat{\Pi} \left|x'\right\rangle = \left\langle -x \mid -x'\right\rangle = \delta(-x+x') = \delta(x-x') = \left\langle x \mid x'\right\rangle$$

4. Inverse of parity

$$\hat{\Pi}\hat{\Pi} = \hat{\Pi}\hat{\Pi}^{\dagger} = \hat{I}$$

5. Eigensystem of parity.

$$\hat{\Pi}\left|\pi\right\rangle = \pi\left|\pi\right\rangle$$

   Apply another operator

$$\hat{\Pi}^2\left|\pi\right\rangle = \pi^2\left|\pi\right\rangle$$

   So, * Eigenvalues: 1, -1; * Eigenvactors: Even function, Odd function

6. Parity applied to operators a. Apply to position operator,

$$\hat{\Pi}^{\dagger}\hat{X}\hat{\Pi} = -\hat{X}$$

   Proof:

$$\left\langle x\right| \hat{\Pi}^{\dagger}\hat{X}\hat{\Pi} \left|x'\right\rangle = \left\langle -x\right| \hat{X} \left|-x'\right\rangle = -x'\delta(x-x') = \left\langle x\right| (-\hat{X}) \left|x'\right\rangle$$

   (a) Apply to momentum operator,

$$\hat{\Pi}^{\dagger}\hat{p}\hat{\Pi} = -\hat{p}$$

   Proof: Similar to the previous one, just change x basis to momentum basis.

2. Symmetry related to Hamiltonian.

$$\left[\hat{\Pi}, \hat{H}\right] = 0$$

   When this happens, parity of Hamiltonian won't change the wave function. Or the wave function should have an specific parity for 1D problem.

7.2.4 Classical Limit of QM

Ehrenfest's Theorem

Schrödinger equation and its adjoint

$$
\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\psi(t)\right\rangle &= \hat{H}\left|\psi(t)\right\rangle \\
-i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi(t)\right| &= \left\langle\psi(t)\right|\hat{H}
\end{aligned}
$$

For any observable $\hat{\Omega}$,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\hat{\Omega}\right\rangle &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\psi(t)\right|\right)\hat{\Omega}\left|\psi(t)\right\rangle + \left\langle\psi(t)\right|\dot{\hat{\Omega}}\left|\psi(t)\right\rangle + \left\langle\psi(t)\right|\hat{\Omega}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left|\psi(t)\right\rangle\right) \\
&= \frac{1}{i\hbar}\left(-\left\langle\psi(t)\right|\hat{H}\hat{\Omega}\left|\psi(t)\right\rangle + \left\langle\psi(t)\right|\hat{\Omega}\hat{H}\left|\psi(t)\right\rangle\right) + \left\langle\psi(t)\right|\dot{\hat{\Omega}}\left|\psi(t)\right\rangle \\
&= \frac{1}{i\hbar}\left\langle\psi(t)\right|\left[\hat{\Omega},\hat{H}\right]\left|\psi(t)\right\rangle + \left\langle\psi(t)\right|\dot{\hat{\Omega}}\left|\psi(t)\right\rangle
\end{aligned}
$$

This is called Ehrenfest's Theorem.

Simple Example of Ehrenfest's Theorem

Suppose we have a system with Hamiltonian

$$
\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})
$$

We need to figure some commutators first.

$$
2m\left[\hat{x},\hat{H}\right] = \left[\hat{x},\hat{p}^2\right] = \hat{x}\hat{p}\hat{p} - \hat{p}\hat{p}\hat{x} = \hat{x}\hat{p}\hat{p} - \hat{p}\hat{x}\hat{p} + \hat{p}\hat{x}\hat{p} - \hat{p}\hat{p}\hat{x} = \left[\hat{x},\hat{p}\right]\hat{p} + \hat{p}\left[\hat{x},\hat{p}\right] = 2i\hbar\hat{p}
$$

$$
\left[\hat{p},\hat{H}\right] = \left[\hat{p},V(\hat{x})\right] = \left[\hat{p},\sum_0^\infty \frac{V^{(n)}}{n!}\hat{x}^n\right] = \cdots = -i\hbar V'(\hat{x})
$$

1. Position average

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\hat{x}\right\rangle &= \frac{1}{i\hbar}\left\langle\psi(t)\right|\left[\hat{x},\hat{H}\right]\left|\psi(t)\right\rangle \\
&= \frac{\left\langle\hat{p}\right\rangle}{m}
\end{aligned}
$$

   We are familiar with this in classical mechanics.

2. Momentum average

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\hat{p}\right\rangle &= \frac{1}{i\hbar}\left\langle\psi(t)\right|\left[\hat{p},\hat{H}\right]\left|\psi(t)\right\rangle \\
&= \frac{1}{i\hbar}\left\langle\psi(t)\right|\left(-i\hbar V'(\hat{x})\right)\left|\psi(t)\right\rangle \\
&= -\left\langle V'(\hat{x})\right\rangle
\end{aligned}
$$

   In classical mechanics, the derivative of potential is force. And the result is just like Newton's 2n Law except the right hand side is not exactly like a force which should be $-\frac{\mathrm{d}}{\mathrm{d}x}\left\langle V(\hat{x})\right\rangle$.

What does $-\langle V'(\hat{x})\rangle$ mean

Suppose the potential area is fairly small and distributed around some coordinate $x_0 = \langle \hat{x}\rangle$, we can do Taylor expansion around x_0.

$$\begin{aligned}< V(\hat{x}) > &= V(x_0) + V'(x_0) < (x - x_0) > + V''(x_0) < (x - x_0)^2 > /2 + \cdots \\ &= V(x_0) + 0 + V''(x_0)(\Delta x)^2 + \cdots\end{aligned}$$

If the uncertainty is small enough, every term except the first one becomes small. So to the lowest order, average of potential is approximately the potential at x_0.

Similarly, the average of first derivative of potential $< V'(\hat{x}) >$ is approximately $V'(x_0)$.

These gives us a hint for the previous result we got for the time evolution of average momentum. The result reduces to classical mechanics one as long as we keep the lowest order of Taylor expansion. Those higher order terms show the quantum effect.

Picture

We can see deeper into Ehrenfest's Theorem through Heisenberg Picture of quantum mechanics.

Schrödinger & Heisenberg Pictures

Pictures are the ways we look at the evolution of systems.

Schrödinger Picture

In Schrödinger picture the states are evolving with time.

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi\rangle_S = \hat{H} |\psi\rangle_S$$

And for time independent Hamiltonian,

$$|\psi\rangle_S = U^\dagger |\psi_0\rangle_S$$

Heisenberg Picture

In Heisenberg Picture, the states do not change with time.

$$|\psi\rangle_H = |\psi_0\rangle_H ,$$

and of course the initial is the same with Schrödinger Picture,

$$|\psi_0\rangle_H = |\psi_0\rangle_S .$$

How do we relate to Heisenberg Picture to Schrödinger Picture? Through investigation of observables. We should have the same observation results in both Pictures.

$$\begin{aligned}{}_H\langle\psi| \hat{\Omega}_H |\psi\rangle_H &= {}_S\langle\psi| \hat{\Omega}_S |\psi\rangle_S \\ {}_H\langle\psi| \hat{\Omega}_H |\psi\rangle_H &= {}_S\langle\psi_0| \hat{U}^\dagger \hat{\Omega}_S \hat{U} |\psi_0\rangle_S \\ \hat{\Omega}_H &= \hat{U}^\dagger \hat{\Omega}_S \hat{U}\end{aligned}$$

So the operators change with time in Heisenberg Picture.

Ehrenfest's Theorem in Heisenberg Picture

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\Omega}_H = \frac{1}{i\hbar}\left[\hat{\Omega}_H, \hat{H}\right] + \hat{U}^\dagger \frac{\partial}{\partial t}\Omega_H \hat{U}$$

This can be easily proved by throwing every definition need in to it. We also need the following equations.

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{U} = \frac{\mathrm{d}}{\mathrm{d}t}e^{-i\hat{H}t/\hbar} = \frac{\hat{H}}{i\hbar}\hat{U}$$

And REMEMBER that propagator commute with time independent Hamiltonian, so

$$\hat{H} = \hat{U}^\dagger \hat{U}\hat{H} = \hat{U}^\dagger \hat{U}\hat{U} \equiv \hat{H}_H$$

So this Ehrenfest's Theorem can also be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\Omega}_H = \frac{1}{i\hbar}\left[\hat{\Omega}_H, \hat{H}_H\right] + \hat{U}^\dagger \frac{\partial}{\partial t}\Omega_H \hat{U}$$

We can **define**

$$\frac{\partial}{\partial t}\hat{\Omega}_H \equiv \hat{U}^\dagger \frac{\partial}{\partial t}\hat{\Omega}_S \hat{U},$$

which is the time derivative of operator in Heisenberg Picture.

Reminder: The time derivative of an observable (average) depends not only the time derivative of itself, but also the commutator of the observable and Hamiltonian.

Example of Ehrenfest's Theorem in Heisenberg Picture

We will show why it is better to work in Heisenberg Picture to show the meanings of Ehrenfest's Theorem.

Suppose we have a Hamiltonian in Heisenberg Picture,

$$\hat{H}_H = \frac{\hat{p}_H^2}{2m} + V(\hat{x}_H).$$

Time derivative of position operator

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_H = \frac{1}{i\hbar}\left[\hat{x}_H, \hat{H}_H\right] = \frac{\hat{p}_H}{m}$$

Time derivative of momentum operator

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{p}_H = \frac{1}{i\hbar}\left[\hat{p}_H, \hat{H}\right] = -V'(\hat{x}_H)$$

So the operator in Heisenberg Picture just have a sense of the physical quantities in classical mechanics. That's why we like it.

Comparison of Picutres

Conservation

We say a observable is conserved if the corresponding operator commutes with Hamiltonian,

$$\left[\hat{\Omega}, \hat{H}\right] = 0$$

Aspects	Schrödinger Picture	Heisenberg Picture	Dirac Picture
Hamiltonian	$\hat{H}$	$\hat{H}_H$	$\hat{H}_0 + \hat{W}(T)$
State	$\vert\psi\rangle_S = \vert\psi(t)\rangle$	$\vert\psi\rangle_H = \hat{U}(t,t_0)^{-1}\vert\psi\rangle_S$ $= \vert\psi(t_0)\rangle$	$\vert\psi\rangle_I = \hat{U}_0^{-1}\vert\psi\rangle_S$ $= e^{i\hat{H}_0(t-t_0)/\hbar}\vert\psi\rangle_S$
Operators	$\hat{\Omega}_S = \hat{\Omega}(t)$	$\hat{\Omega}_H = \hat{\Omega}(t)$ $= \hat{U}^{-1}\hat{\Omega}_S\hat{U}$	$\hat{\Omega}_I = \hat{U}_0^{-1}\Omega_S\hat{U}_0$
EoM	$i\hbar\frac{d}{dt}\hat{U} = \hat{H}\hat{U}$	$\frac{d}{dt}\hat{\Omega}_H = \frac{1}{i\hbar}[\hat{\Omega}_H, \hat{H}_H] + \frac{\partial}{\partial t}\hat{\Omega}_H$	$i\hbar\frac{d}{dt}\vert\psi\rangle_I = \hat{W}_I\vert\psi\rangle_I$ $\frac{d}{dt}\hat{\Omega}_I = \frac{1}{i\hbar}\left[\hat{\Omega}_I, \hat{H}_0\right] + \frac{\partial}{\partial t}\hat{\Omega}_I$
Observables			
Propagator	$\hat{U}$	$\hat{U}$	$\hat{U}_I$
Equation of Propagator	$i\hbar\frac{d}{dt}\hat{U} = \hat{H}\hat{U}$	-	$i\hbar\frac{d}{dt}\hat{U}_I = \hat{W}_I\hat{U}_I$

Fig. 7.2: Comparison of different pictures. Notice that in Dirac picture, $\hat{W}_I = \hat{U}_0^{-1}\hat{W}\hat{U}_0$, $\vert\psi(t)\rangle_I = \hat{U}_I\vert\psi(0)\rangle_I$. A markdown file that is used to make this table can be downloaded here .

1. Energy Hamiltonian always commutes with itself.

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\epsilon\rangle = \langle\psi|\left(\frac{\partial}{\partial t}\hat{H}\right)|\psi\rangle$$

If Hamiltonian is time independent, then energy is conserved. (If Hamiltonian is tide dependent, energy is not conserved. This is kind of obvious in classical mechanics.)

What is the nature of time dependence

We can see this by looking at a simple example.

Assume we have a system with energy eigenstates $|\epsilon_n\rangle$, and initially,

$$|\psi_0\rangle = \sum_n C_n |\epsilon_n\rangle .$$

So

$$|\psi(t)\rangle = \sum_n C_n e^{-i\epsilon_n t/\hbar} |\epsilon_n\rangle .$$

We can calculate the expectation value of some operator $\hat{\Omega}$,

$$\langle\omega(t)\rangle = \sum_{n,m}\left(C_n^* e^{i\epsilon_n t/\hbar}\langle\epsilon_n|\right)\hat{\Omega}\left(C_m e^{-i\epsilon_m t/\hbar}|\epsilon_m\rangle\right) = \sum_{n,m} C_n^* C_m e^{-i(\epsilon_m-\epsilon_n)t/\hbar}\langle\epsilon_n|\hat{\Omega}|\epsilon_m\rangle$$

If $|\epsilon_n\rangle$ are also the eigenvectors of $\hat{\Omega}$, then

$$\langle\epsilon_n|\hat{\Omega}|\epsilon_m\rangle = \omega_m\delta_{n,m}$$

And the expectation value

$$\langle\omega(t)\rangle = \sum_n C_n^* C_n \omega_n$$

The important thing is that the time dependence of this expectation value actually arise from this term

$$e^{-i(\epsilon_m-\epsilon_n)t/\hbar}.$$

As it is so important, we call

$$(\epsilon_m - \epsilon_n)/\hbar$$

Bohr frequency.

7.2.5 Harmonic Oscillators

Why Harmonic Oscillators

Many systems can reduce to it. Use Taylor expansion for the potential and redefine parameters we will find harmonic oscillators in the potential.

Hamiltonian for 1D is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}k\hat{x}^2$$

Standard Solution

We can use polynomial expansion for part of the solution.

Dimension Schrodinger Equation

First step is always finding out the characteristic length scale and characteristic energy scale. Assume we have an characteristic length η and characteristic energy scale ϵ_0. Through uncertainty principle we know only for dimensional analysis

$$[\hat{p}] = \frac{\hbar}{\eta}$$

Kinetic energy and potential energy have the same dimension

$$\frac{\hbar^2}{\eta^2 m} = k\eta^2,$$

so we have

$$\eta = \sqrt{\frac{\hbar}{m\omega}}$$

with $\omega^2 = k/m$. A dimensional analysis shows that $\epsilon_0 = \hbar\omega$.

Now we can define dimensionless variables,

$$z = x/\eta, e = \epsilon/\epsilon_0$$

The time independent Schrodinger equation in position basis is

$$-\hbar^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi''/m + kx^2 = 2\epsilon\psi.$$

Using those characteristic scales, we can rewrite this equation into a dimensionless one, which is

$$\psi'' + (2e - z^2)\psi = 0$$

in which $\psi' = \frac{\mathrm{d}}{\mathrm{d}z}\psi$.

Take Limits

We need to look at the behavior of the solutions before we can guess a proper general solution.

$z \to \infty$, we have $\psi'' - z^2\psi = 0$. Solution to this equation is $\psi(z)\ e^{-z^2/2}$.

The solution of the the equation should be in the form

$$\psi(z) = u(z)e^{-z^2/2}.$$

Insert this to time independent Schrodinger equation, we can get the equation of $u(z)$.

$$u'' - 2zu' + (2e - 1)u = 0$$

Polynomial Method

The simplest form of $u(z)$ is polynomial,

$$u(z) = \sum_{n=0}^{\infty} u_n z^n.$$

Put this back to equation of u, we can get the recursion relation,

$$(n+2)(n+1)u_{n+2} = [2n - (2e-1)]\, u_n.$$

If u_0 and u_1 are given, we can get the whole polynomial.

Notice that we have definite parity here. So u_1 branch vanish because they are even.

u_0 is set by the normalization condition.

Terminate The Series

The series blow up if it doesn't terminate. So we need to terminate the series using the following relation,

$$2e - 1 = 2n.$$

Then we have the energy levels, which is $e = n + 1/2$.

Complete Series

By picking proper normalization factor, we can write down the energy levels and corresponding wave functions. In fact, this polynomial can be found in mathematical phyisics books.

$$H_{n+1} = 2zH_n - nH_{n-1}$$

Tricky Solution

Find out the characteristic length and energy

$$\eta = \sqrt{\frac{\hbar}{m\omega}}$$
$$\epsilon = \hbar\omega$$
$$\omega = \sqrt{\frac{k}{m}}$$

One way to get the intrinsic length without writing down the dimensions of each quantity is to use the following relation

$$[E] = \left[m\omega^2 \hat{x}^2\right]$$
$$\hbar\omega = m\omega^2\eta^2$$
$$\eta = \sqrt{\frac{\hbar}{m\omega}}$$

Or if we are given the Hamiltonian in terms of k,

$$\left[\frac{\hat{p}^2}{2m}\right] = \left[k\hat{x}^2\right]$$
$$\frac{\hbar^2/\eta^2}{m} = k\eta^2$$
$$\eta = \sqrt{\hbar}\sqrt{mk} = \sqrt{\hbar m\omega}$$

Rewrite the Hamiltonian

$$\begin{aligned}
\hat{H} &= \frac{1}{2m}\left[\left(\frac{\hat{p}}{\hbar/\eta}\right)^2\left(\frac{\hbar}{\eta}\right)^2 + \frac{1}{2}m\omega^2\left(\frac{\hat{x}}{\eta}\right)^2\right] \\
&= \frac{1}{2}\hbar\omega\left[\left(\frac{\hat{p}}{\hbar/\eta}\right)^2 + \left(\frac{\hat{x}}{\eta}\right)^2\right] \\
&= \frac{1}{2}\hbar\omega\left(\frac{\hat{x}}{\eta} - i\frac{\hat{p}}{\hbar/\eta}\right)\left(\frac{\hat{x}}{\eta} + i\frac{\hat{p}}{\hbar/\eta}\right) - \frac{i}{\hbar}\left[\hat{x},\hat{p}\right] \\
&= \frac{1}{2}\hbar\omega(\sqrt{2}\hat{a}^\dagger\sqrt{2}\hat{a} + 1) \\
&= \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)
\end{aligned}$$

Now we can define $\hat{a}^\dagger\hat{a} = \hat{N}$, which is just like an operator for (energy) quanta numbers.

An impoertan relation is

$$\left[\hat{a},\hat{a}^\dagger\right] = 1$$
$$\left[\hat{a},\hat{N}\right] = \hat{a}$$

The eigen equation for this weird energy quanta number operator is

$$\hat{N}\left|n\right\rangle = n\left|n\right\rangle$$

To find out the eigen state of $\hat{a}$ and $\hat{a}^\dagger$, we try this,

$$\hat{N}(\hat{a}\left|n\right\rangle) = (n-1)(\hat{a}\left|n\right\rangle)$$
$$\hat{N}(\hat{a}^\dagger\left|n\right\rangle) = (n+1)(\hat{a}^\dagger\left|n\right\rangle)$$

This means $\hat{a}\left|n\right\rangle$ and $\hat{a}^\dagger\left|n\right\rangle$ are also eigen states of $\hat{N}$.

The next step is very crucial. Since $\hat{a}\left|n\right\rangle$ and $\hat{a}^\dagger\left|n\right\rangle$ are eigen states of $\hat{N}$, we know that

$$\hat{a}\left|n\right\rangle = C1\left|n\right\rangle$$
$$\hat{a}^\dagger\left|n\right\rangle = C2\left|n\right\rangle$$

Then our next step is to find out what are $C1$ and $C2$ exactly.

They way of finding them is to use invariant quantities, such as the inner product. Here we use average of $\hat{N}$ operator.

$$\hat{a}\left|n\right\rangle = \sqrt{n}\left|n-1\right\rangle$$
$$\hat{a}^\dagger\left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle$$

Final step is to constrain on n, which should be integrals. This is true because we need a cut off for the eigen equation of $\hat{N}$, whose avarage is n and it should be non negative.

$$\left\langle n\right|\hat{N}\left|n\right\rangle \geq 0$$

leads to $n \geq 0$. To get this proper cut off, n should be integer because if it's not, according to

$$\hat{a} \left| n \right\rangle = \sqrt{n} \left| n-1 \right\rangle$$

n can go to negative numbers. If n is positive integer,

$$\hat{a} \left| 1 \right\rangle = \left| 0 \right\rangle$$
$$\hat{a} \left| 0 \right\rangle = 0 \left| 0 \right\rangle$$

show an cut off at 0.

We can even find out the wave functions of these $\left| n \right\rangle$ by finding the ground state first and apply $\hat{a}^\dagger$ to the ground state.

Ground state in $\left| x \right\rangle$ basis can be found by solving the differential equation,

$$\left\langle x \right| \hat{a} \left| 0 \right\rangle = 0$$

Very important:

- The Hermitian conjugate of $\hat{a} \left| n \right\rangle$ is $\left\langle n \right| \hat{a}^\dagger$.
- Hermitian conjugate of $\hat{a}\hat{a}^\dagger$ is $\hat{a}\hat{a}^dagger$. This can be a trap. Hermitian conjugate is the complex conjugate AND TRANSPOSE!

Semiclassical

Classical

In phase space, the trajectory of phase space points ($\{x/\eta$ and $p/(\hbar/\eta)\}$) is on a circle of radius x_{max}/η.

Quantum semiclassical

Key points:

1. What is the trajectory of $\langle \hat{x}/\eta \rangle$ and $\langle \hat{p}/(\hbar/\eta) \rangle$
2. Can we make the trajectory just like the classical case by choosing some special conditions?
3. What do these special cases mean?

- Expectation value of creation and annihilation operators

Apply Ehrenfest theorem to annihilation operator,

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \hat{a}(t) \rangle = \left\langle \psi \right| \left[\hat{a}(t), \hat{H} \right] \left| \psi \right\rangle = \hbar\omega \langle \hat{a}(t) \rangle$$

Excellent. Now we can solve out $\langle \hat{a}(t) \rangle$, which is

$$\langle \hat{a}(t) \rangle = \alpha_0 \exp(-i\omega t)$$

Take the hermitian conjugate,

$$\left\langle \hat{a}^\dagger(t) \right\rangle = \alpha_0^* \exp(i\omega t)$$

- Expectation value of position and momentum

With these two operators, we can find out the average of $\hat{x}$ and $\hat{p}$ because

$$\hat{x} = \eta \frac{1}{\sqrt{2}} \left(\hat{a}^\dagger + \hat{a}\right)$$
$$\hat{p} = \frac{\hbar}{\eta} i \frac{1}{\sqrt{2}} \left(\hat{a}^\dagger - \hat{a}\right),$$

we have

$$\langle \hat{x}(t) \rangle = \eta \frac{1}{\sqrt{2}} \left(\langle \hat{a}^\dagger(t) \rangle + \langle \hat{a}(t) \rangle \right)$$
$$\langle \hat{p}(t) \rangle = \frac{\hbar}{\eta} i \frac{1}{\sqrt{2}} \left(\langle \hat{a}^\dagger(t) - \langle \hat{a}(t) \rangle \rangle \right)$$

We can have a look at these two averages,

$$\frac{\langle \hat{x}(t) \rangle}{\eta} = \frac{1}{\sqrt{2}} \left[(\alpha_0 + \alpha_0^*) \cos(\omega t) + i(\alpha_0^* - \alpha_0) \sin(\omega t) \right]$$
$$\frac{\langle \hat{p}(t) \rangle}{\hbar/\eta} = \frac{1}{\sqrt{2}} \left[(\alpha_0 + \alpha_0^*) \sin(\omega t) + i(\alpha_0 - \alpha_0^*) \cos(\omega t) \right]$$

It is obvious that the average reduces to classical case if $\alpha_0 = \alpha_0^*$. **But this is too strong for a semiclassical limit.**

- Coherent state

Coherent state is the eigenstate of creation operator. Its wave package has the smallest spread allowed by quantum mechanics.

The most special part about coherent state is that the system stays on coherent state if it start with coherent state.

$$\hat{a} \left|\alpha(t)\right\rangle = \alpha(t) \left|\alpha(t)\right\rangle$$

Take the hermitian conjugate,

$$\left\langle \alpha(t) \right| \hat{a}^\dagger = \left\langle \alpha(t) \right| \alpha(t)^*$$

At $t = 0$, we have

$$\left\langle \psi(0) \right| N \left| \psi(0) \right\rangle = |\alpha_0|^2$$

That is to say, energy should be

$$\left\langle \psi(0) \right| \hat{H} \left| \psi(0) \right\rangle = \hbar\omega \left(|\alpha_0|^2 + \frac{1}{2} \right)$$

Initially, we also have

$$\left\langle \psi(0) \right| (\hat{a} - \alpha_0)^\dagger (\hat{a} - \alpha_0) \left| \psi(0) \right\rangle = 0$$

This means

$$\hat{a} \left|\psi(0)\right\rangle = \alpha_0 \left|\psi(0)\right\rangle$$

- Coherent state expanded using energy eigenstates

(This result)

(To Be Finished…)

7.3 Quantum Mechanics Intermediates

7.3.1 Tensor Product Space

This part has been moved to *Tensor Product Space*

7.3.2 Density Matrix

7.3.3 Angular Momentum

Angular Momentum

For an new operator, we would like to know

1. Commutation relation: with their own components, with other operators;
2. Eigenvalues and their properties;
3. Eigenstates and their properties;
4. Expectation and classical limit.

Definition of Angular Momentum

In classical mechanics, angular momentum is defined as

$$\vec{L} = \vec{X} \times \vec{P}.$$

One way of defining operator is to change position and momentum into operators and check if the operator is working properly in QM. So we just define

$$\hat{\vec{L}} = \hat{\vec{X}} \times \hat{\vec{P}}.$$

It is Hermitian. So it can be an operator. We also find

$$\hat{\vec{L}} \times \hat{\vec{L}} = i\hbar\hat{\vec{L}}$$

$$\left[\hat{L}_i, \hat{L}_j\right] = \sum_k i\epsilon_{ijk}\hat{L}_k.$$

More generally, we can define angular momentum as

$$\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\sum_k \epsilon_{ijk}\hat{J}_k$$

We can prove that

$$\left[\hat{J}^2, \hat{J}_z\right] = 0.$$

So they can have the same eigenstates

$$\hat{J}_z \left|\lambda m\right\rangle = m\hbar \left|\lambda m\right\rangle$$

$$\hat{J}^2 \left|\lambda m\right\rangle = \lambda^2\hbar^2 \left|\lambda m\right\rangle$$

To find the constraints on these eigenvalues, we can use positive definite condition of certain inner porducts, such as,

$$\langle\psi| \hat{J}_+\hat{J}_- |\psi\rangle \geq 0$$

$$\langle\psi| \hat{J}_-\hat{J}_+ |\psi\rangle \geq 0$$

where

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$$

and we have

$$\left[\hat{J}_+, \hat{J}_-\right] = 2\hbar\hat{J}_z$$

$$\left[\hat{J}_z, \hat{J}_\pm\right] = \pm\hbar\hat{J}_\pm.$$

It's easy to find out that

$$\hat{J}_z(\hat{J}_\pm |\lambda m\rangle) = (m \pm 1)\hbar(\hat{J}_\pm |\lambda m\rangle)$$

i.e., $\hat{J}_\pm |\lambda m\rangle$ is eigenstate of $\hat{J}_z$.

Follow the plan of finding out the bounds through these positive inner products, we can prove that

$$\hat{J}^2 |jm\rangle = j(j+1)\hbar^2 |jm\rangle$$

$$\hat{J}_\pm |jm\rangle = \sqrt{j(j+1) - m(m\pm 1)}\hbar |j, m \pm 1\rangle$$

Eigenstates of Angular Momentum

As we have proposed, the eigenstates of both $\hat{J}_z$ and $\hat{\vec{J}}^2$ are $|j, m\rangle$, where $j = 0, 1, 2, \cdots$ and $m = -j, -j+1, \cdots, j-1, j$.

We can also find out the wave function in $|\theta, \phi\rangle$ basis. Before we do that, the definition of this basis should be made clear. This basis spans the surface of a 3D sphere in Euclidean space and satisfies the following orthonormal and complete condition.

$$\int \mathrm{d}\Omega\langle\theta', \phi' \mid \theta, \phi\rangle = \delta(\cos\theta' - \cos\theta, \phi' - \phi) \int \mathrm{d}\Omega \, |\theta', \phi'\rangle \, \langle\theta, \phi| = 1$$

Now we have an arbitary state $|\psi\rangle$,

$$\begin{aligned} |\psi\rangle &= \sum_{l,m} \psi_{lm} \, |l, m\rangle \\ &= \sum_{l,m} \int \mathrm{d}\Omega \, |\theta', \phi'\rangle \, \langle\theta, \phi| \, \psi_{lm} \, |l, m\rangle \\ &= \sum_{l,m} \int \mathrm{d}\Omega \, |\theta', \phi'\rangle \, (\langle\theta, \phi \mid l, m\rangle)\psi_{lm} \end{aligned}$$

Then we define

$$\langle\theta, \phi \mid l, m\rangle = Y_l^m(\theta, \phi)$$

which is the spherical harmonic function.

Then

$$|\psi\rangle = \sum_{l,m} \psi_{lm} \int \mathrm{d}\Omega Y_l^m(\theta, \phi) \, |\theta', \phi'\rangle$$

So as long as we find out what ψ_{lm} is, any problem is done.

7.4 Quantum Mechanics Advanced

7.4.1 Symmetries in QM

Time and Space Translation

First of all I want to know what is not changed or what is the invariant quantity in a transformation.

There are three kind of common transformations.

1. Time translation: move the system in time. In this sense time translation is just the time evolution operator or propagator.
2. Space translation: move the system in space.
3. Gauge transformation

The invariance of them corresponds to:

1. Time translation invariance (T.T.I.) means the evolution of the system is not changing under time translations. **Hamiltonian is invariant.**
2. Space translation invariance (S.T.I.) means that the

Time Translation Symmetry

Time translation Gliffy Source

Definition of Time Translation

Move the system in time.

Generator of Time Translation

T.T.I. is generated by Hamiltonian which can be easily understood by looking into Shcrödinger equation.

Hint: Starting from Schrödinger equation,

$$i\hbar\frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t} = H(t)\psi(t)$$

Then we get the state after a evolution of time Δt,

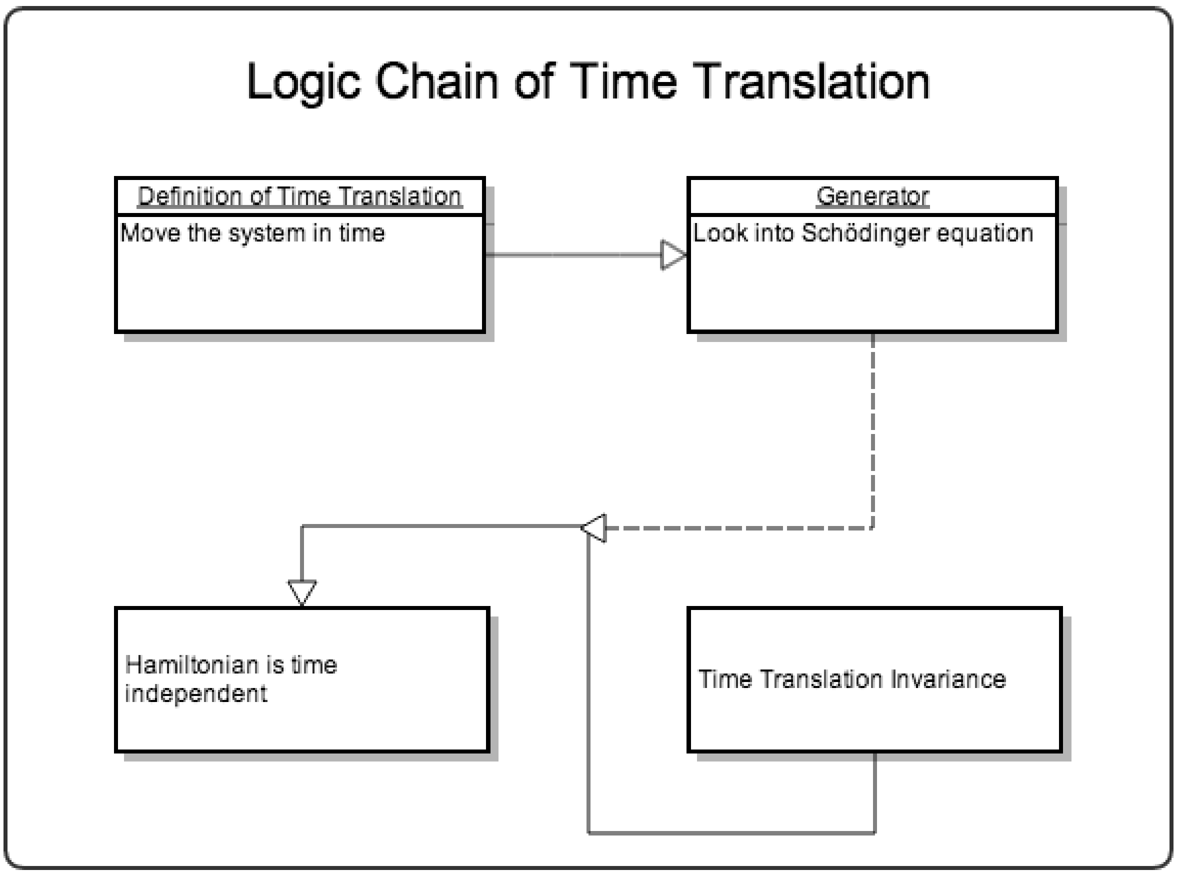
$$|\psi(t+\Delta t)\rangle = \left(\hat{I} - i\frac{\Delta t\hat{H}(t)}{\hbar}\right)|\psi(t)\rangle$$

Time translation symmetry means the state evolution in the same time interval Δt no matter when to start the evolution. Mathematically,

$$|\psi(t_1+\Delta t)\rangle = \left(\hat{I} - i\frac{\Delta t\hat{H}(t_1)}{\hbar}\right)|\psi(t_1)\rangle$$

should get the same final state if we start from some other time t_2,

$$|\psi(t_2+\Delta t)\rangle = \left(\hat{I} - i\frac{\Delta t\hat{H}(t_2)}{\hbar}\right)|\psi(t_2)\rangle$$


Logic Chain of Time Translation
Definition of Time Translation
Move the system in time
Generator
Look into Schödinger equation
Hamiltonian is time independent
Time Translation Invariance

That means the two Hamiltonian should be the same. Now we reach the conclusion that Hamiltonian is time independent.

The logic is to prove that Hamiltonian is time independent by using infinitesimal time translation approach. Given that Hamiltonian is time independent, we imediately know that time translation operator is just the propagator with the form

$$\hat{T}_{\Delta t} \equiv \hat{U}(\Delta t) = e^{-i\hat{H}\Delta t/\hbar}$$

All other conclusions come from the fact that Hamiltonian is a constant of motion.

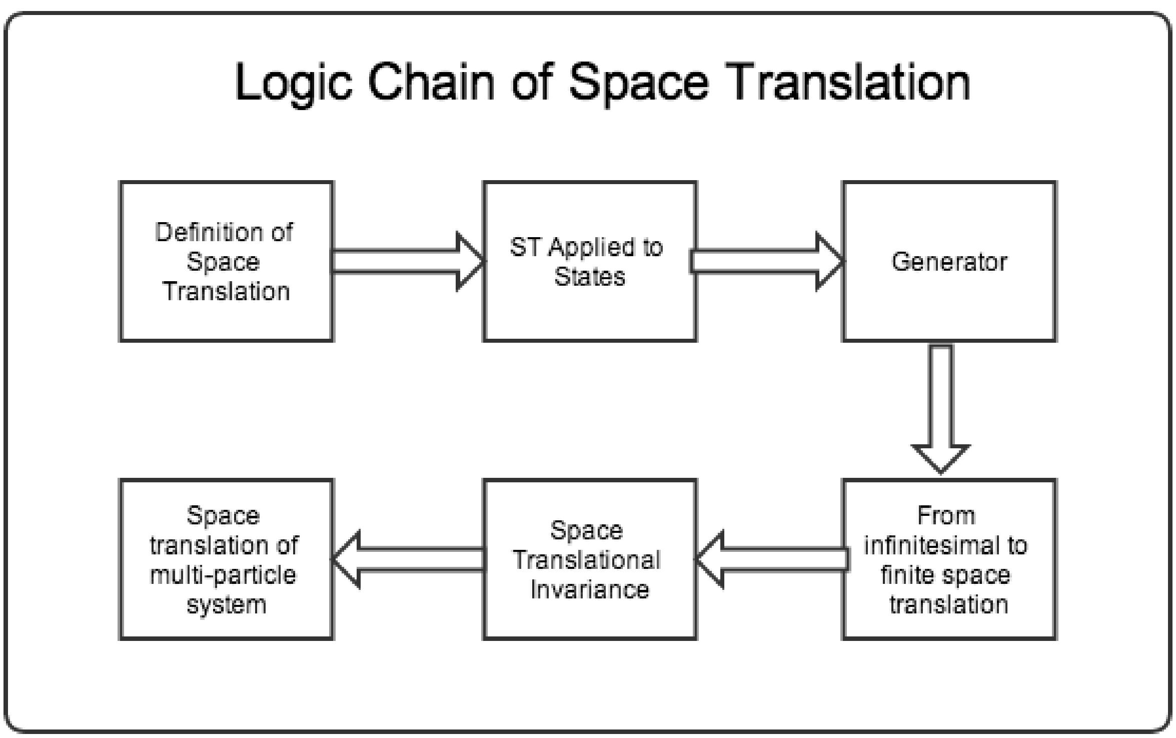
Hint: Ehrenfest theorem tells us that time independent Hamiltonian is a constant of motion.

$$\frac{d}{dt}\langle H\rangle = \frac{1}{i\hbar}\left\langle [\hat{H},\hat{H}]\right\rangle + \left\langle \frac{\partial}{\partial t} H\right\rangle = 0$$

Important: For an isolated system, T.T.I. should always be satisfied because there is nothing more else to change the system but to leave the system with energy conserved.

My concern is if we don't have an Hamiltonian for $T\mathrm{d}S$, we can't actually says this because of what the second law of thermodynamics tells us.

Space Translation Symmetry



Space Translation Gliffy Source

S.T.I. is generated by canonical momentum. This is not so obvious as time translation. To prove this we need to understand what space translation really means.

Definition of Space Translation

Space translation means we change the position of the system by some spatial distance a. In math this means a transformation from $|x\rangle$ to $|x+a\rangle$ where the plus sign is by definition. We invent this space translation operator,

$$\hat{T}_a |x\rangle = |x+a\rangle .$$

Space Translation Applied to States

Next we can obtain the result of space translation operator applied to state in position basis

$$\langle x| \hat{T}_a |\psi\rangle = (\langle x| \hat{T}_a^\dagger) |\psi\rangle = \langle x-a \mid \psi\rangle = \psi(x-a)$$

where we used the relation

$$(\langle x| \hat{T}_a) = (\hat{T}_a^\dagger |x\rangle)^\dagger = (\hat{T}_{-a} |x\rangle)^\dagger = (|x-a\rangle)^\dagger = \langle x-a|$$

which of course is because the normalization of coordinate basis tells us that space translation operator is unitary,

$$\langle x+a \mid x+a\rangle = \langle x| \hat{T}_a^\dagger \hat{T}_a |x\rangle$$

Generator of Space Translation

Similarly to time translation, we can find out the generator out of this definition. For infinitesimal translation,

$$-i\hbar \frac{|\psi(x)\rangle - |\psi(x-\Delta)\rangle}{\Delta} = \hat{p} |\psi(x)\rangle$$

i.e.,

$$|\psi(x-\Delta)\rangle = |\psi(x)\rangle - \frac{i\Delta}{\hbar} |\psi(x)\rangle$$

which shows that the generator of space translation is momentum operator.

From Infinitesimal to Finite Space Translation

$$\hat{T}_a = \lim_{N\to\infty} \hat{T}_{a/N}^N = \lim_{N\to\infty} \left(1 - \frac{i\hat{p}}{\hbar}\frac{a}{N}\right)^N = exp\left(-\frac{i\hat{p}a}{\hbar}\right)$$

Now we have the explicity expression for space translation operators.

Space Translation on Operators

1. Use the invariant scalar – inner product.
2. Passive vs Active

Space Translational Invariance

Space translational invariance of arbitary operator is

$$\hat{\Omega} = \hat{T}_a^\dagger \hat{\Omega} \hat{T}_a$$

is equivalent to

$$\hat{T}_a \hat{\Omega} = \hat{\Omega} \hat{T}_a \Rightarrow [\hat{T}_a, \hat{O}mega] = 0$$

We say some system has space translational invariance we mean the Hamiltonian is space translational invariant,

$$[\hat{H}, \hat{T}_a] = 0.$$

Such a system has space translational invariance.

Hint: I once thought Hamiltonian is space/time translational invariant is not enough for the statement that the whole system is invariant under space or time translation for all observables. Of course I was wrong. Once the Hamiltonian and initial condition is given the whole system can be determined completely in principle.

Gauge Symmetry

Global Gauge Transformation

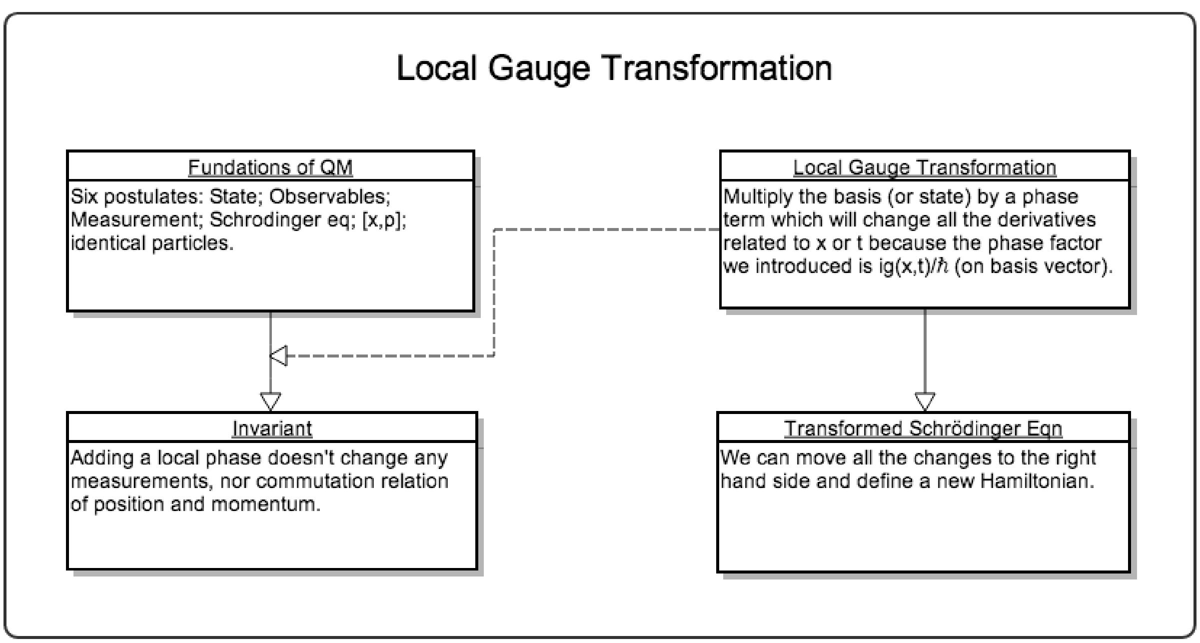
$$|\psi\rangle \to e^{ig\hat{I}} |\psi\rangle$$

All quantum states are invariant under such transformation. This is not a nonsene transformation because the two states are different in some sense if we put them in a phase space where the phase factor assigns a position for the state vector in the phase space and we can see the difference directly in this image.

The invariant thing is the probability density which is obvious.

Hint: This is global because the phase factor doesn't depend on position and time.

Local Gauge Transformation



Local Gauge Transformation Gliffy Source

What if we have a local phase factor: $g(x, t)$?

One way of implenment this phase factor is to transform the basis, for example:

$$|x\rangle \rightarrow e^{ig(x,t)/\hbar} |x\rangle$$

By changing the basis, we can transform anything on position basis. Since the first principle of QM is Schrödinger equation, we would like to check what happens to that.

It turns out that both space derivative and time derivative of the wave function changed. For both of them,

$$\frac{d}{dw}(\exp(-ig/\hbar)\phi) = \exp(-ig/\hbar)\frac{d}{dw}\phi - i/\hbar\left(\frac{d}{dw}g\right)\phi$$

equivalently, we can just change all the derivatives to

$$\frac{d}{dw} \rightarrow \exp(-ig/\hbar)\frac{d}{dw} - i/\hbar\frac{d}{dw}g$$

where w can be x or t.

Parity

Logic

The only thing we need is the definition:

$$\hat{\Pi}|\vec{x}\rangle = |-\vec{x}\rangle$$

Starting from that, we can derive properties.

1. Hermition? **The way to find out something is Hermition or not is to take the Hermitian conjugate of the inner product sandwiched by the operator.**

   We know

   $$\langle x| \hat{\Pi} |x\rangle = \delta(x + x')$$

   Take the Hermitian conjugate of the whole expression,

   $$(\langle x'| \hat{\Pi} |x\rangle)^\dagger = \delta(x + x')$$

   We know the LHS is $\langle x| \hat{\Pi}^\dagger |x'\rangle$. So we have

   $$\langle x| \hat{\Pi}^\dagger |x'\rangle = \langle x| \hat{\Pi} |x'\rangle$$

   Then we get that parity operator is Hermitian.

2. Inversion? Parity operator is Unitary.

   $$\hat{\Pi}\hat{\Pi}|\pi\rangle = \hat{\Pi}\pi|\pi\rangle = \pi^2|\pi\rangle$$

   By physics we know that parity twice gets back to the original state. So $\pi^2 = 1$. Then we can find inverse parity operator. What's important is that it's unitary.

3. Acts on states? From definition, we need to go to position basis.

   $$\langle x|\hat{\Pi}|\psi\rangle = \langle -x \mid \psi\rangle.$$

   We can also find the results on momentum eigenbasis, which is

   $$\langle x|\hat{\Pi}|p\rangle = \langle -x \mid p\rangle.$$

   We already know mometum eigen state in postion is some kind of plane wave and it's easily proved that $\langle -x \mid p\rangle = \langle x \mid -p\rangle$.

4. Commutators with any observables? Just sandwich $\hat{\Pi}^\dagger\hat{\Omega}\hat{\Pi}$ then act on arbitary state and put it into position basis.

   As an example, find commutation relation with position operator.

   $$\langle x|\hat{\Pi}^\dagger\hat{X}\hat{\Pi}|\psi\rangle = \langle -x|\hat{X}\hat{\Pi}|x\rangle = -x\langle -x|\hat{\Pi}|\psi\rangle = -x\langle x \mid \psi\rangle$$

   which is $\langle x|(-\hat{X})|\psi\rangle$. This proves the following equation.

   $$\hat{\Pi}^\dagger\hat{X}\hat{\Pi} = -\hat{X}$$

   which can also be interpreted as passive transformation.

   Another example is the commutation relation with (canonical) momentum.

   $$\langle x|\hat{\Pi}^\dagger\hat{P}\hat{\Pi}|\psi\rangle = \langle -x|\hat{P}\hat{\Pi}|\psi\rangle = \int \langle -x|\hat{P}|x'\rangle\langle x'|\hat{\Pi}|\psi\rangle\, dx'.$$

   By carefully applying parity on position basis, we have

   $$\int \langle -x|\hat{P}|x'\rangle\langle -x' \mid \psi\rangle dx' = \int \langle -x|\hat{P}|-x'\rangle\langle x' \mid \psi\rangle dx'$$

   Because commutation relation tells us

   $$\langle x'|[\hat{X},\hat{P}]|x\rangle = \langle x'|\hat{X}\hat{P}|x\rangle - \langle x'|\hat{P}\hat{X}|x\rangle = (x'-x)\langle x'|\hat{P}|x\rangle = i\hbar\delta(x'-x)$$

   Here comes the keypoint. Recall that

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

   we know that

   $$(x-x')\langle x|\hat{P}|x'\rangle = i\hbar\delta(x'-x)$$

   gives us the expression of momentum in position basis,

   $$\langle x'|\hat{P}|x\rangle = -i\hbar\partial_x\delta(x'-x)$$

   So to continue our calculation of parity applied to momentum,

   $$\int \langle -x|\hat{P}|-x'\rangle\langle x' \mid \psi\rangle dx' = \int \langle x|(-\hat{P})|x'\rangle\langle x' \mid \psi\rangle dx'$$

   So we can prove that mometum actually invserses when parity is applied to it.

7.4.2 Quantum Approximation Methods

Variational Method

Trial functions

Some of the calculable trial functions:

1. $\psi(x) = \cos\alpha x$, for $|\alpha x| < \pi/2$, otherwise 0.
2. $\psi(x) = \alpha^2 - x^2$, for $|x| < \alpha$, otherwise 0.
3. $\psi(x) = C\exp(-\alpha x^2/2)$.
4. $\psi(x) = C(\alpha - |x|)$, for $|x| < \alpha$, otherwise 0.
5. $\psi(x) = C\sin\alpha x$, for $|\alpha x| < \pi$, otherwise 0.

Procedure

1. Pick a trial function.

Note: How to pick a trial function? For ground state energy, we should pick a function that has the same property as the real ground state. This requires some understanding of the problem we are dealing with.

Things to consider:

(a) The new problem is just a modification of a known solved problem. Then we can easily find out what really is different and interprete the new problem in terms of the old one.

(b) If the Hamiltonian have definite parity, the ground state wave function should pick up some parity which is usually even to make it the lowest energy.

(c) Continious function? A C^∞ Hamiltonian can only have continious functions as solutions for a finite system.

(d) Nodes deteremines the kinetic energy so check the nodes for ground state wave function.

(e) Check the behivor of the wave function at different limits. In most cases, the Shrödinger equation can be reduced to something solvable at some limits.

(f) **One more thing, the trial function should make the problem calculable.**

Why Not General Viriational Method

Why don't we just use a most general variational method to find out the ground state? Because we will eventually come back to the time-independent Shrödinger equation.

Suppose we have a functional form

$$E(\psi^*, \psi, \lambda) = \int dx \psi^* H \psi - \lambda \left(\int dx \psi^* \psi - 1 \right)$$

The reason we have this Lagrange multiplier method is that the wave function should be normalized and this multiplier provides the degree of freedom. We would only get a wrong result if we don't include this DoF.

Variation of ψ^*,

$$\delta E = \int dx \delta\psi^* H\psi - \int dx \delta\psi^* \psi = 0$$

Now what?

$$H\psi - \lambda\psi = 0$$

Not helpful.

Variational Method and Virial Theorem

For a potential $V(x) = bx^n$, we can prove that virial theorem is valid for ground state if we use Gaussian trial function $e^{-\alpha x^2/2}$.

A MMA proof is here.

Virial theorem is pretty interesting. It shares the same math with equipartition theorem.

WKB

This is a semi-classical method. It is semi classical because we will use the classical momentum

$$\hbar k(x) = \sqrt{2m(E - V(x))}$$

The following points are important for this method.

0. WKB start from a classical estimation of wave number at a certain energy E which is later quantified by the Bohr-Sommerfeld quantization rule.
1. Conservation law:

   $$\frac{\partial}{\partial t}\rho + \nabla \cdot \vec{j} = 0$$

   where $\rho = \psi^*\psi$, $\vec{j} = -\frac{\hbar}{2mi}(\psi\nabla\psi^* - \psi^*\nabla\psi)$. This can be derived from Shrödinger equation easily.
2. Phase: Wave function is generally $A(x)\exp(\phi(x))$. However, $\phi(x)$ should be the area of the phase function starting from some initial point. For example in WKB, $k(x) = \phi'(x)$ and $\phi(x) = \int \phi'(x')dx' = \int k(x')dx'$.

   Using this general wave function and conservation law we find out that $A(x) \frac{1}{\sqrt{k(x)}}$. Then we can apply the two boundary conditions. However we will find two different wave functions given by two boundary conditions. Now we should connect them because $\psi(a) = \psi(b)$ exactly. By comparing the two wave functions we can find something like Bohr-Sommerfeld quantization rule.
3. Correction at bouldary: However, this method requires that the potential varies slowly or equivalently the wave number varies slowly. Basicly we are just using the following approximation:

   $$A'(x) = 0, k'(x) = 0$$

   For example when taking the derivative of wave function,

   $$\psi'(x) = A'(x)e^{i\int\cdots} + A(x)k(x)e^{i\int\cdots} \approx A(x)k(x)e^{i\int\cdots}$$

   where we drop the term with $A'(x)$. That is to say

   $$|A'| \ll |Ak| \Rightarrow |k'| \ll k^2$$

But at boundary where $E = V$, this is obviously not valid because $k = 0$. So we need to fix this problem.

The solution is to use first order of the potential in a Taylor expansion. Then solve the problem exactly. Finally we connect regions that is far out from the boundary, need the boundary and between the boundary.

If we can have a good boundary condition, then the energy spectrum given by WKB can be very good. Even we don't have a good boundary condtion, the excited states given by this method are always close to the exact ones.

How does it work

7.4.3 Super-Symmetric Quantum Mechanics

Here is a note on this.

The idea of supersymmetric quantum mechanics is to introduce a hamiltonian related to supercharge, which is defined through

$$[H, Q_i] = 0,$$

for all charges Q_i and

$$Q_i, Q_j = \delta_{ij} H.$$

In the 2 charge case, I can define two charges,

$$Q_1 = \frac{1}{2}(\sigma_1 p + \sigma_2 W(x))$$
$$Q_2 = \frac{1}{2}(\sigma_2 p - \sigma_1 W(x)).$$

Harmonic Oscillators

The harmonic oscillators can be solved using ladder operators,

$$a = \sqrt{i/2\hbar}(\hat{p}/\sqrt{m\omega} - i\sqrt{m\omega}\hat{x})$$
$$a^\dagger = \sqrt{-i/2\hbar}(\hat{p}/\sqrt{m\omega} + i\sqrt{m\omega}\hat{x}).$$

This is a hint why we define the charges in that way.

With these charges, we can solve the state that is annihlated by Q_1.

$$Q\psi_0 = 0,$$

which is the ground state.

The result is

$$\psi_0(x) = exp(\int_0^x dy W(y)\sigma_3/\hbar)\psi_0(0).$$

CHAPTER 8

Statistical Physics

This part has been moved to http://statisticalphysics.openmetric.org .

CHAPTER 9

Electrodynamics

Note for electrodynamics course.

Coulomb Potential Energy for a point charge Q with the appearance of a test charge q at distance r

$$V(r) = k\frac{qQ}{r}.$$

The ability to keep storage of charge is called capacitance, which is straight forward to have such a definition as

$$C = \frac{q}{U},$$

where U is the electric potential (not the potential energy).

Maxwell's equations are

$$\begin{aligned}
\mathbf{E} \cdot \mathrm{d}\mathbf{S} &= \frac{1}{\varepsilon_0} \iiint_\Omega \rho \,\mathrm{d}V \\
\mathbf{B} \cdot \mathrm{d}\mathbf{S} &= 0 \\
\oint_{\partial\Sigma} \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} &= -\frac{d}{dt} \iint_\Sigma \mathbf{B} \cdot \mathrm{d}\mathbf{S} \\
\oint_{\partial\Sigma} \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} &= \mu_0 \iint_\Sigma \mathbf{J} \cdot \mathrm{d}\mathbf{S} + \mu_0\varepsilon_0 \frac{d}{dt} \iint_\Sigma \mathbf{E} \cdot \mathrm{d}\mathbf{S}
\end{aligned}$$

or

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{B} &= \mu_0 \left(\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right)
\end{aligned}$$

9.1 Electrodynamics Vocabulary and Program

9.1.1 Vocabulary

Units

Gaussian Units

Gaussian unit is very useful in electrodynamics. Back to two equations used in SI unit system to define some standard units,

$$F_e = k_e \frac{Qq}{r^2}$$
$$F_m = k_m \frac{2 I_1 I_2 L_2}{\rho},$$

where we have the two constants defined as $k_e = \frac{1}{4\pi\epsilon_0}$ and $k_m = \frac{\mu_0}{4\pi}$.

The idea of Gaussian unit is as simple as setting $k_e = 1$ and $k_m = 1/c^2$. The consequences are, however,

$$\epsilon_0 = \frac{1}{4\pi}$$
$$\mu_0 = \frac{4\pi}{c^2}.$$

These two equalities are the most useful ones to help us switch between SI and Gaussian.

Math

Vector Analysis

A lot of vector analysis equations will be deployed in this subject. The way to quickly prove a vector or tensor relation is to write down the component form, mess with the orders and use some specific relations.

One of the most useful symbol involved is Levi Civita symbol, which has a relation with the Kronecker delta,

$$\epsilon_{ijk}\epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}.$$

As an example, we consider the case $i = l$, the determinant reduces to

$$\epsilon_{ijk}\epsilon_{imn} = \begin{vmatrix} \delta_{jm} & \delta_{jn} \\ \delta_{km} & \delta_{kn} \end{vmatrix} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}.$$

This relation is useful in many situations.

$$\begin{aligned} \vec{a} \times (\vec{b} \times \vec{c}) &= \hat{e}_i \epsilon_{ijk} a_j (\epsilon_{kmn} b_m c_n) \\ &= \hat{e}_i \epsilon_{kij} \epsilon_{kmn} a_j b_m c_n \\ &= \hat{e}_i (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}) a_j b_m c_n \\ &= \hat{e}_i \delta_{im}\delta_{jn} a_j b_m c_n - \hat{e}_i \delta_{in}\delta_{jm} a_j b_m c_n \\ &= \hat{e}_i a_j b_i c_j - \hat{e}_i a_j b_j c_i \\ &= \vec{b}(\vec{a}\cdot\vec{c}) - \vec{c}(\vec{a}\cdot\vec{b}). \end{aligned}$$

Integrals

Gaussian integral is

$$\int_{-\infty}^{\infty} e^{-x^2/a^2} dx = a\sqrt{\pi}.$$

To calculate higher orders we can use parity and derivitives.

x^{2n+1} are odd function thus

$$\int_{-\infty}^{\infty} x^{2n+1} e^{-x^2/a^2} dx = 0.$$

For those x^{2n} we can use this derivative trick,

$$\int_{-\infty}^{\infty} e^{-x^2/a^2} dx = \left(\frac{-\partial}{\partial 1/a^2}\right)^2 \left(a\sqrt{\pi}\right).$$

E & M

1. Static E field and B field,
2. Scalar potential and vector potential,
3. Multipole expansion,
4. Force of objects in E field and B field, both for arbitary field and each multipoles,
5. Torque of objects in E field and B field, both for arbitary field and each multipoles.

Radiation Pressure

There are many ways to understand the pressure produced by light. In classical electrodynamics, we have the momentum current density, electromagnetic stress tensor, surface current density and quantum as the tools.

These are three different levels of the phenomenon.

Momentum Current Density

Momentum current density is

$$\vec{g} = \frac{1}{8\pi c}\mathrm{Re}(\vec{E}^* \times \vec{H}).$$

Pressure is force per unit area or momentum change per unit time per unit area. Momentum change, meanwhile, is momentum current density times volume.

To carry out in the language of math, the volume in time Δt and on area a is given by $c\delta ta$, where c is the speed of light. Here we used c because we are basically considering the process in vacuum.

Pressure is given by

$$\begin{aligned} P &= \frac{\vec{F}\cdot\hat{n}}{a} \\ &= \frac{\Delta\vec{p}\hat{n}}{\Delta t a} \\ &= \frac{\hat{n}\cdot(\Delta\langle\vec{g}\rangle)\Delta t c a}{\Delta t a} \\ &= \hat{n}\cdot(\Delta\langle\vec{g}\rangle). \end{aligned}$$

The next step is to plug in the momentu current density and calculate the difference. **Here we calculate an example.**

Suppose we have our incident wave normal to the surface of perfect reflection. The wave has

$$\langle \vec{g}_1 \rangle = -\langle \vec{g}_1' \rangle = -\frac{|\vec{E}_0|^2 \hat{n}}{8\pi c}.$$

Finally the radiation pressure becomes

$$P = \frac{|\vec{E}_0|^2}{4\pi}.$$

Interaction Between Magnetic Field and Surface Current as Radiation Pressure

On the surface of metal, electromagnetic waves could induce surface current which in return interacts with the magnetic component in the electromagnetic wave thus producing radiation pressure.

Surface current induced is calculated using

$$\hat{n} \times (\vec{B}_1 - \vec{B}_2) = \frac{4\pi}{c} \vec{K},$$

in which $\vec{B}_2 = 0$ since it is the magnetic field inside the good conductor.

The force is given by

$$\frac{a}{2} \frac{1}{2} \mathrm{Re} \left(\vec{K}^* \times \frac{\vec{B}_1 + \vec{B}_2}{2} \right).$$

The average of magnetic field is the key point here. The reason behind this 1/2 factor is that infact only half of the magnetic field outside of the conductor is the original part while the other half is induced by the surface current density however the $\vec{B}_1$ includes all the magnetic field outside.

The pressure is

$$P = \frac{-\hat{n} \cdot \Delta \vec{F}}{a} = \frac{|\vec{E}_0|^2}{4\pi}.$$

The Balance of Mechanic Pressure and Radiation Pressure

Due to conservation law, we have

$$\frac{d}{dt} (\vec{P}_{mech} + \vec{P}_{EM}) = \oint_S d\vec{a} \cdot \mathbf{T},$$

which says the energy stress tensor integrated over a close surface on the interface of metal is the change in the momentum of mechanical part and electromagnetic part in total.

By using the monochromatic wave expression we can find the pressure.

$$P = \frac{|\vec{E}_0|^2}{4\pi}$$

Refs & Note

9.1.2 Models

Drude Model

Drude models models the electromagnetic properties of conductors. In this simple model, electromagnetic interaction with the electromagnetic waves comes from the free charges governed by this equation of motion.

$$m\frac{d\vec{v}}{dt} = q\vec{E} - \frac{m\vec{v}}{\tau},$$

where τ is the damping term like the one in Brownian motion. In fact this is a Brownian motion like equation since $\vec{E}$ is periodic and averages to 0 in a simple case **except that this external electric force is not random**.

Math

Now consider the solution to this equation which is simple using Green's function,

$$\begin{aligned}\vec{v}(t) &= \vec{v}(0)e^{-t/\tau} + \frac{q\vec{E}_0}{m}\int_0^t dt' e^{-(t-t')/\tau} e^{-i\omega t'} \\ &= \vec{v}(0)e^{-t/\tau} + \frac{q\vec{E}_0}{m} e^{-t/\tau} \frac{i\tau(-1 + e^{t(1/2-i\omega)})}{\tau\omega + i}.\end{aligned}$$

The first term is the damping contribution to the velocity of the charges and the second term is the contribution of electric field.

As a practice, we can show that the ratio of two contributions is

$$\begin{aligned}M &= \frac{|v_2|}{|v_1|} \\ &= \frac{1}{v(0)}\frac{qE_0}{m}\frac{i\tau(-1 + e^{t(1/2-i\omega)})}{\tau\omega + i} \\ &= \frac{qE_0}{mv(0)}\frac{i\tau(e^{t/\tau(1-i\omega\tau)} - 1)}{i + \omega\tau}.\end{aligned}$$

It is obvious that in the limit $\omega\tau \gg 1$, and after a long time this ratio becomes 0 which actually comes from the fact that ω is very large.

This means that in a long run, the damping always takes over the system if we have a very large frequency. Now the next question is what is the energy contribution of electric field after a long time?

$$\vec{v}_2(t) = \frac{q\vec{E}_0}{m} e^{-t/\tau} \frac{i\tau(-1 + e^{t(1/2-i\omega)})}{\tau\omega + i}$$

becomes 0 as ω becomes large enough.

So in a high frequency limit, the system take no energy from electric field.

Note that this result is complex. Explain why is a velocity is complex.

In this model, what to calculate is the electromagnetic response of the material like reflection ration etc.

Maxwell Equations

Maxwell Equations

$$\begin{aligned}\nabla\cdot\vec{D} &= 4\pi\rho,\\ \nabla\cdot\vec{B} &= 0,\\ \nabla\times\vec{E} &= -\frac{1}{c}\partial_t\vec{B},\\ \nabla\times\vec{H} &= \frac{1}{c}\partial_t\vec{D}+\frac{4\pi}{c}\vec{j}.\end{aligned}$$

In Drude model, for each mode $\vec{E},\vec{B}\propto e^{i(\vec{k}\cdot\vec{})}$, I can replace ∂_t with $-i\omega$ and $\vec{\nabla}$ with $i\vec{k}$. With the help of this, the Maxwell's equation becomes

$$\begin{aligned}\vec{k}\cdot\vec{D} &= 4\pi\rho,\\ \vec{k}\cdot\vec{B} &= 0,\\ \vec{k}\times\vec{E} &= -\frac{1}{c}(-i\omega)\vec{B},\\ \vec{k}\times\vec{H} &= \frac{1}{c}(-i\omega)\vec{D}+\frac{4\pi}{c}\vec{j}.\end{aligned}$$

Current Density

The current density is

$$\vec{j} = qn_q\langle\vec{v}\rangle.$$

Meanwhile, the Drude model tells us the velocity is governed by

$$m\frac{d\vec{v}}{dt} = q\vec{E}-\frac{m\vec{v}}{\tau},$$

where $-\frac{m\vec{v}}{\tau}$ is a damping term.

Damping Term

This damping term can also be interpreted as the mean free time or some kind of probability.

Plugin the current density we could find the equation for current density.

$$\begin{aligned}\vec{j} &= qn_q\langle\vec{v}\rangle\\ &= \vec{E}\left(\frac{q^2n_q\tau}{m(1-i\omega\tau)}\right)\end{aligned}$$

Ohm's law tells us that

$$\vec{j} = \sigma\vec{E}.$$

Now we can have a phenomelogical conductivity from Drude model,

$$\sigma(\omega) = \frac{q^2n_q\tau}{m(1-i\omega\tau)}.$$

Maxwell's Equation in Neutral Matter

Pluging in the current density we calculated previously, the Maxwell's equation becomes

$$\begin{aligned}\vec{k}\cdot\vec{D} &= 4\pi\rho,\\ \vec{k}\cdot\vec{B} &= 0,\\ \vec{k}\times\vec{E} &= -\frac{1}{c}(-i\omega)\vec{B},\\ \vec{k}\times\vec{H} &= \frac{1}{c}(-i\omega)\epsilon\vec{E} + \frac{4\pi}{c}\sigma\vec{E}.\end{aligned}$$

Comparing with the equations in matter without free charge, where the transverse wave satisfies

$$\begin{aligned}\vec{k}\cdot\vec{E} &= 0,\\ \vec{k}\cdot\vec{B} &= 0,\\ \vec{k}\times\vec{E} &= \frac{\omega}{c}\vec{B},\\ \vec{k}\times\vec{H} &= -\frac{\omega}{c}\vec{D}\end{aligned}$$

we can find the expression for permitivity,

$$\epsilon = 1 + \frac{i\omega_p^2\tau}{\omega(1-i\omega\tau)}.$$

since $\mu = 1$. In the result I defined

$$\omega^2 = e^2 n_e \tau/m.$$

The next quantity is to calculate the refractive index with $\mu = 1$.

$$n = \sqrt{\epsilon\mu} = \sqrt{1 + \frac{i\omega_p^2\tau}{\omega(1-i\omega\tau)}}.$$

In the limit of $\omega\tau \gg 1$, refractive index becomes

$$n^2 = 1 - \frac{\omega_p^2}{\omega^2}.$$

Dispersion Relation

$$\begin{aligned}\vec{k}\times\vec{E} &= \frac{\omega}{c}\vec{B},\\ \vec{k}\times\vec{H} &= -\frac{\omega}{c}\vec{D}\end{aligned}$$

gives us the dispersion relation. However, we need to make a choice that the field need to be broken into parts that is perpendicular and parallel to wave vector. For the transverse wave, we could write down

$$k^2 = \frac{\omega}{c}\epsilon.$$

Lorentz Model

Drude model only considers the damping part of conducting charges. Lorentz model, considers the actually polarization inside medium, using a simple but efficient model.

In models about matter response to electromagnetic waves, we have to get the permitivity out of it and furthure calculate the refractive index.

Suppose we already know how to write down polarization,

$$\vec{P} = n\vec{p},$$

which means the polarization is caused by a lot of small dipoles. At this point we are not binded to the calculation of the detailed expression of these small dipoles. Instead we are going to calculate the permitivity first then come back to have a look at the details.

In statics we know,

$$\vec{D} = \vec{E} + 4\pi\vec{P} \equiv \vec{E} + 4\pi\chi\vec{E} = \epsilon\vec{E}.$$

To find ϵ we need to establish the relation between $\vec{P}$ and $\vec{E}$ which is equivalently setting up the relation between $n\vec{p}$ and $\vec{E}$.

Here we introduce Lorentz model. In this context, we consider the case that equation of motion for the charges are governed by

$$m\ddot{\vec{x}} = -e\vec{E} - m\omega_0^2\vec{x} - \gamma m\dot{\vec{x}}.$$

Solve the equation of motion we have the relation between $\vec{x}$ and $\vec{E}$ thus we can write down

$$\begin{aligned}\vec{P} &= n\vec{p} \\ &= -en\vec{x} \\ &= -en\frac{-e\vec{E}/m}{\omega_0^2 - \omega^2 - i\gamma\omega} \\ &= \frac{e^2n/m}{\omega_0^2 - \omega^2 - i\gamma\omega}\vec{E}.\end{aligned}$$

Imediately, we have the permitivity

$$\begin{aligned}\epsilon &= \epsilon_0 + 4\pi\chi \\ &= 1 + 4\pi\frac{e^2n/m}{\omega_0^2 - \omega^2 - i\gamma\omega} \\ &= 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega},\end{aligned}$$

where we used the definition of plasma frequency

$$\omega_p^2 = \frac{4\pi n e^2}{m}.$$

Limits

We have got three important parameters or arguments in Lorentz model, ω_0, ω, γ and on overall ω_p. One should notice that in normal matter we would see $\gamma \ll \omega_0 \ll \omega_p$.

Three limits can be considered,

1. low frequency, ω is very small like $\omega_0 - \omega \gg \gamma$;
2. critical, $\omega = \omega_0$ where we have only $-i\gamma\omega$ appears in denominator;
3. intermediate, $\omega_0 \ll \omega \ll \omega_p$;
4. very high frequency, $\omega \gg \omega_p$.

The interesting thing is that in situation 3, we get back to Drude model.

9.1.3 Program

Most problems in stat mech have similiar procedures. This page is for the programs of solving problems.

9.1.4 Causality

Field theory shows a lot of casality conditions. Here is a collection of them.

Radiation

Maxwell's equations in vacuum are

$$\begin{aligned}
\nabla \cdot \vec{E} &= 4\pi\rho, \\
\nabla \cdot \vec{B} &= 0, \\
\nabla \times \vec{E} &= -\frac{1}{c}\partial_t \vec{B}, \\
\nabla \times \vec{B} &= \frac{1}{c}\partial_t E + \frac{4\pi}{c}\vec{j}.
\end{aligned}$$

To write down the wave equation, we could switch to the scalar potential ϕ and vector potential $\vec{A}$.

Divergence free means that we can always have

$$\vec{B} = \nabla \times \vec{A}.$$

By using the above relation, I could rewrite this to

$$\begin{aligned}
\nabla \times \vec{E} &= -\frac{1}{c}\partial_t(\nabla \times \vec{A}), \\
\nabla \times (\vec{E} + \frac{1}{c}\partial_t \vec{A}) &= 0.
\end{aligned}$$

This means I can write all inside divergance written as a gradient of a scalar function or a constant time the gradient of some scalar function.

$$\vec{E} = -\nabla\phi - \frac{1}{c}\partial_t \vec{A}.$$

With the definition of scalar and vector potentials, we could plug them in and find the wave equations. However, since the values of these potential are gauge dependent, I should choose a convinient gauge. Hereby, I use Lorenz gauge.

$$\frac{1}{c}\partial_t\phi + \nabla \cdot \vec{A} = 0.$$

The importance of this gauge is that it is Lorentz invariant.

Using this gauge the two other Maxwell's equations, I have the wave equations,

$$(\frac{1}{c^2}\partial_t^2 - \nabla^2)\phi = 4\pi\rho,$$
$$(\frac{1}{c^2}\partial_t^2 - \nabla^2)\vec{A} = \frac{4\pi}{c}\vec{j}.$$

Solving these Helmholtz equations, I get the solution as a function of retarted time $t_{ret} = t - \frac{R}{c}$, where $R = |\vec{x} - \vec{x}'|$.

$$\phi(t,\vec{x}) = \int d^3x' \frac{\rho(t_{ret},\vec{x}')}{R},$$
$$\vec{A}(t,\vec{x}) = \int d^3x' \frac{\vec{j}(t_{ret},\vec{x}')}{R}.$$

Here it clearly shows that the observation depends on the history R/c ago. This is the signal propagation time.

Response of Matter

$$\vec{P} = \int \chi(t,t')\vec{E}(t')dt'.$$

9.2 Electricity and Magnetism

9.2.1 Circuits

Capacitor

Capacitance is defined as

$$C = \frac{dQ}{dV}.$$

Since current is defined as $I = \frac{dQ}{dt}$, we derive the current and potential relation for capacitor

$$CdV = dQ$$
$$\Rightarrow C\frac{dV}{dt} = \frac{dQ}{dt}$$
$$\Rightarrow C\frac{dV}{dt} = I.$$

Inductor

Inductor is defined as

$$L = \frac{d\Phi}{dI},$$

where Φ is the magnetic flux of the inductor and I is the current going through the inductor.

Refs & Notes

9.2.2 Electrostatics and Magnetostatics

Electric Field

The electric potential energy of a point charge in an electric field is given by

$$U = q\phi.$$

In the case of fixed point charge Q field, the amount electric energy of test charge q is

$$U_q = q\frac{Q}{r}.$$

Suppose we have N point charges places randomly in the space, the total electric energy of the system can be calculated through the a superposition of the electric potential.

$$\begin{aligned} U &= q_1 \sum_{i=1}^{N} \frac{q_i}{|\vec{x_i} - \vec{x}|} + \cdots \\ &= \frac{1}{2}\left(\sum_{i\neq j} \frac{q_i q_j}{|\vec{x_i} - \vec{x_j}|} \right), \end{aligned}$$

in which the half is due to the double counting of the interactions.

Energy Density

The next question we are going to consider is the electric potential energy of a bulk charged object.

$$U = \frac{1}{2}\int \rho(\vec{x})\phi(\vec{x}) d^3 x,$$

where

$$\phi(\vec{x}) = \int \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} d^3 x'.$$

Derivation of Electric Energy of Charged Object

Construct the system by adding small amount of charge to it from zero charge distribution.

Charge distribution $\rho(\vec{x})$ can be changed linearly $\lambda\rho(\vec{x}')$ where λ is a scalar number.

The energy change due to adding charge to the system to change the charge distribution linearly $\delta\lambda$. Energy change is

$$\delta U = \int \delta\lambda \phi(\vec{x}) d^3 x.$$

where

$$\phi(\vec{x}) = \int d^3 x' \frac{\lambda\rho(\vec{x}')}{|\vec{x} - \vec{x}'|}.$$

Integrate from 0 to 1 will give us the total energy for an object with charge distribution $\rho(\vec{x})$.

It is obvious that the $1/2$ comes from the linear dependence of $\lambda\rho(\vec{x})$.

Now we have the expression for energy of electric field, it is straightforward to find the energy density.

1. To calculate that, we need to put together two parts of energy, i.e., E-field outside of the conductor and the surface electric potential energy of the conductor.
2. Then apple the divergence theorem, counting the infinite surface and also all the surfaces of the conductor.
3. Notice the surface integral of the potential energy on the conductor surface is exactly the negative of the surface energy density we put in to the expression in the first step.
4. The surface integral at infinity surface is zero.
5. The only term left is $-\frac{1}{8\pi}\int d^3x E_i \partial_i \phi = \frac{1}{8\pi}\int d^3x E_i E_i$.

Electric Forces

Alright, move on to the concept of force in E&M. Coloumb force is

$$F_i = \int \rho(\vec{x}')E_i(\vec{x}')d^3x'.$$

Maxwell's equations tells that

$$\rho(\vec{x}') = \frac{1}{4\pi}\nabla' \cdot \vec{E}(\vec{x}').$$

So force can be rewritten as

$$\begin{aligned} F_i &= \int d^3x' \frac{1}{4\pi}(\partial_j E_j(\vec{x}'))E_i(\vec{x}') \\ &= \frac{1}{4\pi}\int d^3x'(\partial_j(E_jE_i) - E_j\partial_j E_i) \end{aligned}$$

Now notice that $\partial_i(E_jE_j) = 2E_j\partial_i E_j$ and $\partial_i E_j = \partial_j E_i$.

Proof

$$\begin{aligned} \nabla \times \vec{E} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial_i & \partial_j & \partial_k \\ E_i & E_j & E_k \end{vmatrix} \\ &= \hat{i}\epsilon_{ijk}\partial_j E_k. \end{aligned}$$

So $\nabla \times \vec{E} = 0$ means $\partial_i E_j - \partial_J E_i = 0$.

Force becomes

$$\begin{aligned} F_i &= \frac{1}{4\pi}\int d^3x'(\partial_j(E_jE_i) - E_j\partial_j E_i) \\ &= \frac{1}{4\pi}\int d^3x'(\partial_j(E_jE_i) - \frac{1}{2}\partial_i(E_jE_j)) \\ &= \frac{1}{4\pi}\int d^3x'\partial_j((E_jE_i) - \frac{1}{2}E^2) \\ &= \frac{1}{4\pi}\int d^3x'\partial_j((E_iE_j) - \frac{1}{2}\delta_{ij}E_iE_j) \\ &= \frac{1}{4\pi}\int d^3x'\partial_j((1 - \frac{1}{2}\delta_{ij})E_iE_j) \\ &= \frac{1}{4\pi}\int dS_j(1 - \frac{1}{2}\delta_{ij})E_iE_j \end{aligned}$$

Recall that force is momentum per unit time. What is inside the integral means some force density or momentum flow density per unit time, for some reason we use the definition

$$T_{ij} = (1 - \frac{1}{2}\delta_{ij}) E_i E_j,$$

which is symmetric under the exchange of i and j.

Stress Tensor in General Relativity or Fluid Dynamics

Electromagnetic energy momentum tensor using Gaussian units and +2 signature is

$$T^{\mu\nu} = c^2 \left(F^{\mu\alpha} F^{\nu}{}_{\alpha} - \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right),$$

where $\mathbf{F} = \mathbf{dA} \Rightarrow F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

$$F^{\mu\nu}|_{\text{matrix form}} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}.$$

To change the index to $F^{\mu}{}_{\nu}$ we just use the Minkowski metric which just put plus and minus signs on the components,

$$F^{\mu}{}_{\nu}|_{\text{matrix form}} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}.$$

Why is this the case? This is because of the Lorentz group requires it. Generators of $L_+^\uparrow$ which means the orthochronous patch combine together with the infinitesimal change to form the tiny change in Lie algebra, that is to form ω in $L = I + \omega$

$$\omega = \vec{\theta} \cdot \vec{R} + \vec{\lambda} \cdot \vec{B},$$

where the six matrices are basically the generators to construct the Lorentz group.

Energy momentum tensor can be decomposed using the generators,

$$F^{\mu}{}_{\nu}|_{\text{matrix form}} = -E_i B_i + B_i R_i,$$

where we are using the Einstein summation rule.

For more information about this please read arXiv:physics/0005084 .

Force on Capacitor

Suppose we have a capacitor with two parallel round plates of radius R and separation d. The top plate has charge Q and bottom plate has charge $-Q$. Now ask the question, what is the force on the top plate?

Is its magnitude $F = QE$ because of the fact that the charge Q is in a electric field $\vec{E}$? NO.

We can calculate the force using several methods like

1. principle of virtual work,
2. stress tensor.

However the result shows that the magnitude of force is only $F = \frac{1}{2}QE$, half of the expectation we had at first.

The intuition is that electric field only exists on one side of the plate while $F = qE$ describes the force of a charged particle emerged in the electric field. Once one knows this, the half comes from the fact that stress only applies on one side of the top plate.

Multipole Expansion

For a charge object with charge distribution $\rho(\vec{x})$, the electic potential is

$$\phi(\vec{x}) = \int \rho(\vec{x}') \frac{1}{|\vec{x} - \vec{x}'|} d^3 x'.$$

For any general distribution, we can always have $\frac{1}{|\vec{x}-\vec{x}'|}$ expanded and define differential multipoles.

Monopole and Dipole

Before we go into this expansion, a review of the idea of monopole and dipole is very useful.

Monopole is the case of spherical symmetric distribution of electric potential.

$$\rho_0(\vec{x}) = \frac{Q}{r} = \frac{1}{r} \int d^3 x' \rho(\vec{x}') = \int d^3 x' \frac{\rho(\vec{x}')}{|\vec{x}|}.$$

Dipole, using the simplest model, is a system of two charged particles with $\pm Q$ and directed from negative charge to positive charge.

Fig. 9.1: Electric dipole from wikipedia .

In the language of math, $\vec{q} = Q\vec{d}$. The electric potential is the superposition of the two charges.

$$\begin{aligned}\phi(\vec{x}) &= \frac{-Q}{|\vec{x} - (-\vec{d}/2)|} + \frac{Q}{|\vec{x} - \vec{d}/2|} \\ &= \int d^3 x' \frac{-Q\delta(\vec{x}' + \vec{d}/2) + Q\delta(\vec{x}' - \vec{d}/2)}{|\vec{x} - \vec{x}'|} ??\end{aligned}$$

I need some time to make clear of different methods.

Dipole comes from the expansion at $\vec{x} \gg \vec{x}'$,

$$\frac{1}{|\vec{x} - \vec{x}'|} = \frac{1}{r} + \vec{x}' \cdot \vec{x} \frac{1}{r^3} \cdots$$

thus the dipole part is actually

$$\phi_1(\vec{x}) = \frac{1}{r^3} \vec{x} \cdot d^3 x' \vec{x}' \rho(\vec{x}').$$

Multipole

A nth multipole has the dimnsion of $\frac{[Q]}{[r]} \left(\frac{d}{r}\right)^2$.

In most of the cases, we have a very small dipole compared to the distance at where the field is measured. So we could **take the limit that the dipole is point like**. The actually meaning of a point dipole is that the dipole is a constant viewed from very far away.

$$\begin{aligned}\rho(\vec{x}) &= -q\delta(\vec{x} + \vec{d}/2) + q\delta(\vec{x} - \vec{d}/2) \\ &\approx -q\nabla \cdot \delta(\vec{x}).\end{aligned}$$

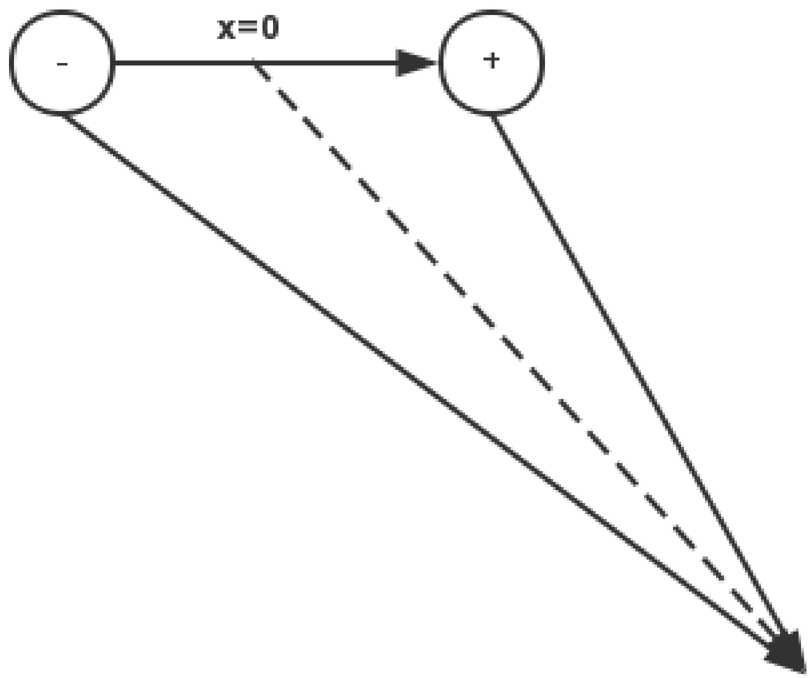


Fig. 9.2: The dashed line is the position vector $\vec{x}$.

The next question to be answered is the electric field generated by this point dipole.

$$\begin{aligned}\vec E &= -\nabla \frac{\vec p\cdot \vec x}{r^3}\\ &= \frac{3(\vec p\cdot\vec x)\vec x - \vec p r^2}{r^5}\\ &= \frac{3(\vec p\cdot\vec x)\vec x}{r^5} - \frac{\vec p}{r^3},\end{aligned}$$

which is **only true for** $\vec x \neq 0$.

Force feels by a dipole is

$$\vec F = \nabla(\vec p(\vec x)\cdot \vec E(\vec x)).$$

Torque can also be calculated

$$\vec \tau = \vec p \times \vec E + \vec x \times \vec F.$$

Force and Torque of Dipole

Force can be calculated using principle of virtual work. $-\vec p \cdot \vec E$ is the energy which can be explained using simple two charge model.

Torque has two parts, one with a precession-like term which is called spin torque while another one is a orbital torque since it tries to align the dipole with the field lines.

Dielectric Material

In the presence of dielectric material, a new set of quantities will be introduced.

Polarization

Why do we introduce this polarization $\vec P$? To make the picture and the math easier to understand, in some sense.

Imagine dielectric material in electric field. Since the material is dielectric, the external electric field induces additional electric field.

We use subscript P for quantities that is induced.

For a neutral object, one conservation law is

$$\int \rho_P(\vec x') d^3 x' + \oint \sigma_P \hat n \cdot d\vec S = 0,$$

where ρ_P is the charge volume density distribution inside the object while σ_P is the surface charge density. We need the surface charge because charge will be induced on the surface.

Maxwell's equations are very beautiful, so we definitely want to preserve Gauss' law. Thus we define

$$\begin{aligned}\rho_P(\vec x) &= -\nabla\cdot \vec P_P(\vec x)\\ \sigma_P(\vec x) &= \vec P(\vec x)\cdot \hat n(\vec x).\end{aligned}$$

Polarization and Electric Field

So what is the relation of $\vec{P}$ and $\vec{E}$? This is not trivial. The physics here is the electric field (vector), space distribution, polarization (vector).

$$P_i = ?E_j$$

How many ranks do we need to describe a spatially distribution in Euler space? Two maximum. So

$$P_i = \chi_{ij} E_j.$$

However, the tensor χ_{ij} can be a function of E_j and even P_i.

For **isotropic** material, we would have no direction dependence of the spatial distribution

$$P_i = \chi E_i.$$

This is because no shear stress is possible thus :math:'P_i' doesn't depend on E field on other directions.

χ can still be a function of E_i though. Now we require that this material is **linear**, which means χ is independent of E_i.

With the requirement of **isotropic and linear**, we have a simple relation,

$$P_i = \chi_e E_i,$$

where, for some reason, χ_e is called **electric suscptibility**.

Why is this isotropic and linear approximation useful?

Think about the microscopic structure of material.

Electric Potential

The potential should be related dipole since this is dielectric material. (<- This is a joke.) Anyway, it is related to dipole.

The electric potential is composed of two parts, one produced by the volume charge density and another produced by the surface charge density.

$$\begin{aligned}
\phi_P(\vec{x}) &= \int d^3x' \frac{\rho(\vec{x}')}{|\vec{x} - vecx'|} + \oint d\vec{S}' \cdot \frac{\sigma_P \hat{n}'}{|\vec{x} - \vec{x}'|} \\
&= \int d^3x' \frac{-\nabla' \cdot \vec{P}(\vec{x}')}{|\vec{x} - vecx'|} + \oint d\vec{S}' \cdot \frac{\vec{P}(\vec{x}')}{|\vec{x} - vecx'|} \\
&= \int d^3x' \frac{-\nabla' \cdot \vec{P}(\vec{x}')}{|\vec{x} - vecx'|} + \int d^3x' \nabla' \cdot \frac{\vec{P}(\vec{x}')}{|\vec{x} - vecx'|} \\
&= \int d^3x' \left(\vec{P}(\vec{x}') \cdot \nabla' \cdot \frac{1}{|\vec{x} - vecx'|} \right) \\
&= \int d^3x' \frac{\vec{P}(\vec{x}') \cdot (\vec{x} - \vec{x}')}{|\vec{x} - vecx'|^3}.
\end{aligned}$$

Recall that the electric potential by a dipole is

$$\phi_d = \frac{\vec{p} \cdot \vec{x}}{r^3}.$$

The electrical potential generated by induced charge distribution in dielectric material is a integral of a lot dipole electric potential fields. This can be seen by the following approximation in x direction,

$$(x - x') = x(1 - \frac{x'}{x}) \approx x,$$

then

$$\phi_P(\vec{x}) \approx \int d^3x' \frac{\vec{P}(\vec{x}') \cdot \vec{x}}{\vec{x}^3}.$$

At this point we only considered the induced field.

Total Field

Gauss' law tells us the field about the total charge, both the original and the induced, and the total field, whatever generated it.

$$\nabla \cdot \vec{E} = 4\pi(\rho + \rho_P),$$

where ρ means the free charge.

As we have introduced

$$\rho_P = -\nabla \cdot \vec{P},$$

$$\nabla \cdot \vec{E} = 4\pi(\rho - \nabla \cdot \vec{P}).$$

Combine terms

$$\nabla \cdot (\vec{E} + 4\pi\vec{P}) = 4\pi\rho.$$

The point is we don't want to mess with $\vec{P}$, so we define a **displancement vector**

$$\vec{D} = \vec{E} + 4\pi\vec{P},$$

so that the Gauss' law only involve with free charge ρ and the displacement vector,

$$\nabla \cdot \vec{D} = 4\pi\rho.$$

For **isotropic and linear** material, we already know $\vec{P} = \chi_e\vec{E}$,

$$\vec{D} = \vec{E} + 4\pi\chi\vec{E} = (1 + 4\pi\chi_e)\vec{E},$$

then we define

$$\epsilon = 1 + 4\pi\chi_e,$$

which is the **permittivity** by definition.

Energy and Stress

Just as a reference,

$$\delta U = \frac{1}{4\pi}\int \vec{E} \cdot \delta\vec{D} d^3x' = \frac{1}{4\pi}\int \frac{1}{2\epsilon}\delta E^2 d^3x',$$

which indicates that the energy density is

$$u = \frac{1}{8\pi} \vec{E} \cdot \vec{D}.$$

Stress tensor T_{ij} which appears in the formula of force $\vec{F} = \int d\vec{S} \cdot \overleftrightarrow{T}$ is

$$T_{ij} = \frac{1}{4\pi} \left(D_i E_j - \frac{1}{2} D_k E_k \delta_{ij} \right).$$

Magnetic Field

Lorentz force is

$$\vec{F} = q\vec{E} + q\frac{\vec{v}}{c} \times \vec{B}.$$

Ampere's Law

Force between two charged wire is given by

$$\vec{F} = \frac{1}{c^2} \int \int \frac{I_1 d\vec{x} \times (I_2 d\vec{x}' \times (\vec{x} - \vec{x}'))}{|\vec{x} - \vec{x}'|^3}$$

Force on A Charged in Electric Field

Force on a charge in a electric field generated by $\rho(\vec{x})$ is

$$\vec{F} = q\vec{E} = q \int d^3 x' \frac{\rho(\vec{x}')(\vec{x} - \vec{x}')}{|\vec{x} - \vec{x}'|^3}.$$

Charge Conservation in Magnetostatistics

Conservation of charge means

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0,$$

we require the source of current density 0 which means

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \vec{j} = 0.$$

The question is how can the electric current feels each other? By intermediate field we say.

Biot-Savart Law

The magnetic field generated by electric current is given by Biot-Savart law.

$$\vec{B} = \frac{1}{c} \int d^3 x' \frac{vecj(\vec{x}') \times (\vec{x} - \vec{x}')}{|\vec{x} - \vec{x}'|^3}$$

Starting from magnetic field we could find the force,

$$\vec{F} = \frac{1}{c} \int d^3x' \vec{j}(\vec{x}) \times \vec{B}(\vec{x}).$$

Force of Charge Distribution in Electric Field

vec F = int d^3x' frac{rho(vec x')} vec E(vec x') .

Divergence and Curl of Magnetic Field

Using the formula of magnetic field generated by current,

$$\begin{aligned}\nabla \cdot B &= \frac{1}{c}\nabla \cdot \int d^3x' \vec{j}(\vec{x}') \times \frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3} \\ &= \frac{1}{c}\partial_i \int \epsilon_{ikl} j_k(\vec{x}') \frac{x_l - x_l'}{|\vec{x}-\vec{x}'|^3} d^3x' \\ &= \frac{1}{c}\int j_k \epsilon_{kli} \partial_i \partial_l \frac{1}{|\vec{x}-\vec{x}'|} d^3x' \\ &= 0\end{aligned}$$

The last line is true since

$$\begin{aligned}& j_k \epsilon_{kli} \partial_i \partial_l \frac{1}{|\vec{x}-\vec{x}'|} \\ &= \vec{j} \cdot (\nabla \times \nabla \frac{1}{|\vec{x}-\vec{x}'|}) \\ &= 0\end{aligned}$$

The curl on the other hand is not zero.

$$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{j}.$$

Apply a loop integral,

$$\oint d\vec{x} \cdot \vec{B} = \frac{4\pi}{c} I.$$

Vector Potential

For a electric field, we have a scalar potential,

$$\vec{E} = -\nabla \phi,$$

since electrostatics is curl free,

$$\nabla \times \vec{E} = -\nabla \times \nabla \phi = 0.$$

Magnetostatics, on the other hand, is divergence free. Thus we expect

$$\vec{B} = \nabla \times \vec{A}.$$

Not surprisingly, we have

$$\vec{A}' = \vec{A} + \nabla\chi,$$

is also a valid vector potential for $\vec{B}$.

Gauges

Coulomb gauge is

$$\nabla \cdot \vec{A} = 0.$$

Example of Gauge Freedom

To have a magnetic field $\vec{B} = B_0 \hat{e_z}$, we could use different vector potentials,

$$\vec{A}_1 = B_0 x \hat{e_y},$$
$$\vec{A}_2 = -B_0 y \hat{e_x},$$
$$\vec{A}_3 = \frac{\vec{A}_1 + \vec{A}_2}{2}.$$

This is the gauge freedom of the magnetic field.

Refs & Notes

9.2.3 Comparison of E & M

Useful Tricks

$$\partial_i (x_k j_i) = j_k + x_k \partial_i j_i$$

This is useful because we have

$$\nabla \cdot \vec{j} = \partial_i j_i = \frac{d}{dt}\rho = 0,$$

and the LHS can be turned into a surface integral as one wish and disappears.

A similar one is

$$\partial_i (x_k x_l j_i).$$

Statics in Vacuum

Source of Fields

1. Source of Electric Field in Electrostatics

Source of electric field is charge

$$\nabla \cdot \vec{E} = 4\pi\rho$$

2. Source of Maganetic field in Magnetostatics

Current is the source of maganetic field

$$\nabla \times \vec{B} = \frac{4\pi}{c}\vec{j}.$$

Potentials

1. Electric Potential

Electric potential is given by

$$\vec{E} = -\nabla\phi.$$

Imediately, we have the curl of electrical field being 0, i.e.,

$$\nabla \times \vec{E} = 0.$$

By implementing Gauss's law, the equation for potential becomes

$$\nabla^2\phi = -4\pi\rho.$$

The solution, apply the Green's function to Laplace equation,

$$\phi(\vec{x}) = \int d^3x' \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|}$$

2. Magnetic Potential

Magnetic potential is given by

$$\vec{B} = \nabla \times \vec{A}.$$

Applying curl of magnetic field and solving the equation,

$$\vec{A} = \frac{1}{c}\int \frac{\vec{j}}{|\vec{x} - \vec{x}'|} d^3x'.$$

Gauge of fields

By definition, electric potential and maganetic potential are, repectively,

$$\vec{E} = -\nabla\phi,$$

$$\vec{B} = -\nabla \times \vec{A}.$$

Electric field is invariant under a transform

$$\phi' = \phi + \phi_0,$$

where $\nabla\phi_0 = 0$.

Similarly, the potential for magnetic field is

$$\vec{A}' = \vec{A} + \nabla\psi,$$

in which ψ can be any scalar fields.

Gauge

As expected, these definitions of fields do not determine the potential completely. This is gauge freedom.

It might seem strange to talk about such a freedom. As we would ask why we have such freedom for potentials?

In class electrodynamics, potentials are merely mathematical tools. So the notion that potental has gauge freedom comes only from the mathematical definition of potentials.

However, we do expect such a freedom is part of nature as we step into quantum ralm. In quantum world, Aharonov-Bohm effect proves that potentials are actually real existance. In such cases, the gauge freedom do have a very important impact on our theory. Gauge freedom is part of the internal structure of fields and goes deep into group theory, topology and differential geometry.

Multipole Expansion

Requirement

One should be able to derive these multipole expansions of fields without refering to any material.

In the expression for potentials,

$$\phi(\vec{x}) = \int d^3x' \frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|},$$

and

$$\vec{A} = \frac{1}{c}\int \frac{\vec{j}}{|\vec{x}-\vec{x}'|} d^3x',$$

the term

$$\frac{\vec{j}}{|\vec{x}-\vec{x}'|}$$

can be Taylor expanded when $\vec{x}' \ll \vec{x}$,

$$\begin{aligned} &\frac{1}{|\vec{x}-\vec{x}'|} \\ =&\frac{1}{r} - x_i'\partial_i\frac{1}{r} + \frac{1}{2}x_i'x_j'\partial_i\partial_j\frac{1}{r} + cdots \end{aligned}$$

where $1/r$ is $1/|x|$.

Apply this expansion, we can find the dipole and quadrapole of a charge distribution, which are

$$\vec{p} = \int \rho(\vec{x}')\vec{x}' d^3x',$$

$$Q_{ij} = \frac{1}{2}\int \rho(\vec{x}')(3x_i'x_j' - r'^2\delta_{ij})d^3x'.$$

The corresponding potentials are

$$\vec{\phi}_d = \frac{\vec{p}\cdot\vec{x}}{r^3},$$

and

$$\vec{\phi}_q = \frac{x_i Q_{ij} x_j}{r^5}.$$

The electric field can be calculated using $\vec{E} = -\nabla\phi$.

For magnetic field, a dipole expansion shows that

$$\vec{A} = \frac{\mu \times \vec{x}}{r^3},$$

where

$$\vec{\mu} = \frac{1}{2c}\int \vec{x}' \times \vec{j}(\vec{x}') d^3x'.$$

Force and Torque

Requirement

1. Write the most general expression for force and torque.
2. Derive the expression with dipole and quadrapole approximations.

Among the tricks, virtual work principle could be a nice one.

Force and torque can be calculated using virtual work principle. However, for dipoles, they can be calculated directly.

9.3 Response of Matter

Maxwell's Equations and Transfer of Momentum and Energy

$$\begin{aligned}
\nabla \cdot \vec{D} &= 4\pi\rho \\
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{E} &= -\frac{1}{c}\partial_t \vec{B} \\
\nabla \times \vec{H} &= \frac{1}{c}\partial_t \vec{D} + \frac{4\pi}{c}\vec{j}.
\end{aligned}$$

9.3.1 Energy and Momentum

Energy transfer can be described through Poynting vector which in the simplist plane wave case is basically $\langle \vec{S} \rangle = u_{EM}\vec{v}$ where u_{EM} is the energy density. Obviously Poynting vector is the energy transfer rate.

Momentum density is $\vec{g} = \frac{u_{EM}}{c}$ in the same plane wave.

Let's start from conservation of charge. Taking the divergence of the following equation,

I get the conservation law,

$$\partial_t \rho + \nabla \cdot \vec{j} = 0.$$

On the next step we find energy density of the wave. The starting point is the interaction of charge and electric field, e.g., $\vec{j}\cdot\vec{E}$, $q\vec{E}\cdot d$.

$$u_{EM} = \frac{1}{8\pi}(\vec{D}\cdot\vec{E} + \vec{B}\cdot\vec{H}).$$

Upon this we can find $\vec{S}$ which is

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{H}.$$

As for momentum, we have

$$\vec{F}_{EM} = \oint d\vec{a}\cdot\mathbf{T} - \frac{d}{dt}\int \vec{g}(\vec{x})d^3x,$$

where

$$\vec{g} = \frac{\vec{S}}{c^2} = \frac{\vec{E}\times\vec{B}}{4\pi c}.$$

9.3.2 Response of Matter

The equation for monochromatic wave,

$$\begin{aligned}\vec{k}\cdot\vec{D} &= 4\pi\rho\\ \vec{k}\cdot\vec{B} &= 0\\ \vec{k}\times\vec{E} &= i\omega\vec{B}\\ \vec{k}\times\vec{H} &= -\frac{i\omega}{c}\vec{D} + \frac{4\pi}{c}\vec{j}.\end{aligned}$$

In the case of $\rho = 0$ (no free charge) and $\vec{j} = 0$ no free current,

$$\begin{aligned}\vec{k}\cdot\vec{E} &= 0\\ \vec{k}\cdot\vec{B} &= 0\\ \vec{k}\times\vec{E} &= i\omega\vec{B}\\ \vec{k}\times\vec{B} &= -\frac{i\omega}{c}\epsilon\mu\vec{E}.\end{aligned}$$

The important relation is dispersion relation which is the relation between wave number and frequency. In this context,

$$\vec{k}\cdot\vec{k} = \frac{\omega^2}{c^2}\mu\epsilon,$$

which reduces to

$$k = \frac{\omega}{c/n}$$

as now we can define a new velocity which stands for the phase velocity by

$$v_{ph} = c/n.$$

Impedance

Since $\vec{E}$ corresponds to the magnetic part $\vec{H}$, we would like to rewrite one of the equation to

$$Z\vec{H} = \vec{k} \times \vec{E}.$$

where

$$Z = \sqrt{\frac{\mu}{\epsilon}}.$$

If $\mu = 1$, we have

$$Z = \frac{1}{n}.$$

Matching Condition

Matching conditions are not just something written down on the textbook. They all have meanings.

$$\begin{aligned}
\hat{n} \cdot (\vec{D}_1 + \vec{D}_1' - \vec{D}_2) &= 0, \\
\hat{n} \cdot (\vec{B}_1 + \vec{B}_1' - \vec{B}_2) &= 0, \\
\hat{n} \times (\vec{E}_1 + \vec{E}_1' - \vec{E}_2) &= 0, \\
\hat{n} \times (\vec{H}_1 + \vec{H}_1' - \vec{H}_2) &= 0.
\end{aligned}$$

In order to satisfy these matching conditions, the phase of all fields should be the same. Thus leading to the Snell's law

$$\begin{aligned}
\theta_1 &= \theta_1' \\
n_1 \sin\theta_1 &= n_2 \sin\theta_2 \qquad \text{perpendicular part.}
\end{aligned}$$

The parallel parts of the fields give us the other relation,

$$\begin{aligned}
E_1 \cos\theta_1 - E_1' \cos\theta_1 - E_2 \cos\theta_2 &= 0 \\
H_1 + H_1' &= H_2.
\end{aligned}$$

Reflected and Transmitted Wave

Recall that impedance is $Z = \sqrt{\frac{\mu}{\epsilon}}$ which leads to or corresponds to

$$|E| = Z|H|.$$

Impedance shows the scaling between the magnetic field and the electric field. In fact for monochromatic TEM,

$$\vec{k} \times \vec{E} = Z\vec{H}.$$

Then we derive the ratios,

$$\begin{aligned}
\frac{E_1'}{E_1} &= \frac{Z_1 \cos\theta_1 - Z_2 \cos\theta_2}{Z_1 \cos\theta_1 + Z_2 \cos\theta_2}, \\
\frac{E_2}{E_1} &= \frac{2Z_2 \cos\theta_1}{Z_1 \cos\theta_1 + Z_2 \cos\theta_2}.
\end{aligned}$$

By definition, impedance becomes $Z = 1/n$ when $\mu = 1$. In this limit,

$$\frac{E_1'}{E_1} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2},$$
$$\frac{E_2}{E_1} = \frac{2n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2}.$$

Now what I need to be careful is that in this derivation, I used the geometry to project the fields on to the surface where the polarization of the wave matters. For the result above, they are only valid for waves with polarization parallel to the surface.

Using similar tricks, I can write down the result for waves with polarization perpendicular to surface. The matching conditions are

$$E_1 + E_1' = E_2$$
$$H_1 \cos\theta_1 - H_1' \cos\theta_1 = H_2 \cos\theta_2.$$

This can be abtained by just draw a graph of the incident wave, reflected wave and refracted wave.

Solving the equations

$$\frac{E_1'}{E_1} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2},$$
$$\frac{E_2}{E_1} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2}.$$

Reflection, Refraction, Transparent, Dissipative

Reflection coefficient and transmission coefficient find the energy reflected and transmitted.

$$\begin{aligned} R &= \frac{\langle \vec{S}_1' \rangle \cdot \hat{n}}{\langle \vec{S}_1 \rangle \cdot \hat{n}} \\ &= \frac{\frac{c}{4\pi}\frac{1}{2} Re(\vec{E}_1'^* \times \vec{H}_1')}{\frac{c}{4\pi}\frac{1}{2} Re(\vec{E}_1^* \times \vec{H}_1)} \\ &= \cdots \\ T &= \frac{\langle \vec{S}_2 \rangle \cdot \hat{n}}{\langle \vec{S}_1 \rangle \cdot \hat{n}} \\ &= \cdots \end{aligned}$$

For normal incident, these becomes easier to calculate because all thetas becomes 0. The result is

$$R = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2}$$
$$T = \frac{4n_1 n_2}{(n_1 + n_2)^2}.$$

Evanescent wave is the case when the wave vector becomes imaginary and the wave tenuates to 0 quickly. In the situation of total reflection, the transmitted wave can be calculated. To find out the evanescent wave one need to calculate the condition for total reflection then plug in the condition for a assumed wave in medium 2.

Evanescent wave doesn't mean energy lose in reflection, it only proves that wave can not go deep into the material and all waves are reflected. The material does NOT nessarily obsorbe all the energy of the wave. One can show that in total reflection, energy flowing in all flows out, i.e.,

$$R = 1.$$

The question is, what is dissipative material? They are those with complex wave vectors such that wave dissipates as they passing through the bulk material.

Ohmic Matter

A term of current

$$\frac{4\pi}{c}\vec{j}$$

is added to the equations so that it describes matter which supports current.

Ohm's law shows the relations between current and field,

$$\vec{j} = \sigma\vec{E}.$$

By plug in this to the Maxwell's equations, we can have a new permitivity. For monochromatic wave,

$$\epsilon = \epsilon_R + i\frac{4\pi\sigma}{\omega}.$$

Using dispersion relation derived, we can define the complex refractive index.

Given the dispersion relation we could also find the waves and calculate Poynting vector

$$\langle\vec{S}\rangle \propto e^{2(\omega n_i/c)\hat{k}\cdot x}.$$

This is our anticipation since the wave vector has a length of $k = \frac{\omega}{c/(n_R+in_I)}$ due to dispersion relation $\vec{k} = \frac{\omega}{c/n}\hat{k}$. Plug this into the expression for plane wave, we have the spatial part proportional to

$$e^{\vec{k}\cdot\vec{x}} = e^{i(n_R+in_I)\omega/c\hat{k}\cdot\vec{x}}$$

.

We can directly see the damping part. This is how one finds the skin depth which means the Poynting vector drops to $1/e$ of that of the incident wave,

$$\delta = \frac{c}{2\omega n_i}.$$

If the imaginary part becomes very large, i.e.,

$$\frac{4\pi\sigma}{\omega} \gg 1,$$

the atenuation becomes significant or both real part and imaginary part of refractive index becomes much larger than 1. This is equivalent to good conductor. We can see the skin depth

$$\delta \ll \frac{c}{\omega}.$$

Refrective index also becomes 1.

Remember surface current is defined as

$$\vec{K} = \int \vec{j} dz.$$

We could use this to find a effective surface current. Current density is

$$\begin{aligned}\vec{j} &= \sigma\vec{E}\\ &= \sigma\vec{E}_0 e^{i(\vec{k}\cdot\vec{x}-\omega t)}.\end{aligned}$$

Using dispersion relation, and suppose we have wave vector in z direction. We can write down the incident wave

$$\vec{E}_1 = \vec{E}_0 e^{i\omega(nz/c-t)}.$$

Fresnel's relation tells us the refracted wave, which is

$$\begin{aligned}\vec{E}_2 &= \frac{2\vec{E}_1}{n+1} \\ &\approx \frac{2\vec{E}_1}{n} \\ &= \frac{2\vec{E}_1}{(1+i)n_R}.\end{aligned}$$

The last step is true for good conductor. The current density is therefore clear so is the surface current.

In summary, good conductor has a

1. small skin depth;
2. large n for both real and imaginary part;
3. almost 1 as the reflection coefficient.

Dispersive Media

Dispersive media can be modeled using Drude model, Lorentz model and many other. Read *Drude Model* and *Lorentz Model* in vocabulary part.

By definition, group velocity is the result of dispersion relation,

$$v_g = \frac{d\omega}{dk}|_{k=k_0},$$

while phase velocity is always $v_{phase} = nc$ where n can be larger than 1.

Read chapte 5 of Kevin Cahill's book for knowledge of complex dispersion relation and more. Notice that both group velocity and phase velocity can be larger than the vacuum speed of light.

9.4 Radiation

Usually, radiation is something that can be propagated to infinity, which means, those wave that drops according to $\frac{1}{r^3}$ or even faster won't get propagated to very far away. But we still might be interested in those near fields.

Understand retarded time is something always good for radiation study.

$$t_{ret} = t - \frac{|\vec{x} - \vec{x}'|}{c},$$

which means retarded time is the current time minus the travel time of the radiation.

The mathematical reason for this retardation is that the Green's function for electromagnetic wave,

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) A = 0.$$

This is for the fields in vacuum. For radiation with source, the wave equations gains a source term. To solve the equation, I need to write down the Green's function, which, in this case, has a retarded term in it. In fact all such waves have a retarded time in it even just by looking at the math. Retarded time has a great impact on the solution since it delays the effect of source.

This is also causality.

9.4.1 Radiation In General

In general the wave equation with a source is

$$\left(\nabla^2 - \frac{1}{c^2}\partial_t^2\right) A^\mu = -s^\mu.$$

By solving that, the retarded solution comes into effect,

$$A^\mu = \int d^3x' \frac{s^\mu(t_{ret}, \vec{x}')}{|\vec{x} - \vec{x}'|}.$$

Here the vector potential is

$$\vec{A} = \int \frac{\vec{j}(t_{ret}, \vec{x}')/c}{|\vec{x} - \vec{x}'|},$$

while the scalar potential is

$$\phi = \int \frac{\rho(t_{ret}, \vec{x}')}{|\vec{x} - \vec{x}'|}.$$

Static Fields

The static field has similar structure except we have no retardation. Here are the expressions.

The important thing in radiation is the angle depenence of radiation power or total radiation power. To find that, there are many procedures.

One of them is to use the fact that electric field in radiation is always transverse which means

$$\vec{E} = -\hat{r} \times \vec{B}.$$

So we only need to find out the magnetic field thus the first thing is to calculate the vector potential.

Why B field first?

We can also find out electric field first. But in dynamics,

$$\vec{E} = -\nabla\phi - \partial_t \vec{A},$$

which means we need to find both ϕ and $\vec{A}$.

While in the procedure stated previously, we only need $\vec{A}$.

To summarize, here is the procedure.

1. Calculate $\vec{A}$.
2. Find magnetic field using $\vec{B} = \nabla \times \vec{A}$.
3. Find electric field is needed using $\vec{E} = -\hat{r} \times \vec{B}$.
4. Find Poynting vector $\vec{S} = \frac{c}{4\pi}\vec{E} \times \vec{B}$.
5. Find radiation power $\frac{dP}{d\Omega} = r^2 \langle \vec{S} \rangle \cdot \hat{r}$, which is angle dependent in general.

Zangwill's Method

There is a radiation vector method in Zangwill's *Modern Electrodynamics* book.

Dipole Radiation

Dipole Radiatioin Can Be Calculated Exactly.

Approximations

Since we are talking about radiation which is radiated away. Looking at a far zone radiation is good.

Define $r = |\vec{x}|$. Expanding the vector potential, we have the multipole expansion of the vector potential in the limit that $|\vec{x}'| \ll |\vec{x}|$ is all true for the whole integral,

$$\begin{aligned}\vec{A} &\approx \frac{1}{cr}\int d^3x'\vec{j}(t-r/c+\hat{r}\cdot\vec{x}'/c,\vec{x}') \\ &\approx \frac{1}{cr}\int d^3x'\left(\vec{j}(t-r/c,\vec{x}')+\frac{\hat{r}\cdot\vec{x}'}{c}\partial_t\vec{j}(t-r/c,\vec{x}')\right)\end{aligned}$$

So the vector potential under this degree of approximation can be splited into two terms at this point.

$$\vec{A}_{E1} = \frac{1}{cr}\int d^3x'\vec{j}(t-r/c,\vec{x}'),$$

is the electric dipole.

Electric Dipole Radiation

The reason is the conservation of charge. Consider this relation,

$$\nabla\cdot\vec{j}+\partial_t\rho = 0.$$

Combine this with the following trick,

$$\partial_i(j_i x_k) = x_k\partial_i j_i + j_k,$$

and the fact that

$$\oint j_i x_k \mathrm{d}\Sigma_i = 0,$$

for any surface that is large enough.

We have the following relation

$$\int d^3x(j_k - x_k\partial_t\rho) = 0.$$

Then we know that

$$\int \vec{j}d^3x = \dot{\vec{p}}.$$

$$\begin{aligned}\vec{A}_2 &= \frac{1}{cr}\int d^3x'\frac{\hat{r}\cdot\vec{x}'}{c}\partial_t\vec{j}(t-r/c,\vec{x}') \\ &= \frac{1}{c^2r}\int d^3x'\hat{r}\cdot\vec{x}'\partial_t\vec{j}(t-r/c,\vec{x}').\end{aligned}$$

A trick can be applied to this expression,

$$\frac{1}{2}\hat{r}\times(\vec{x}'\times\vec{j}) = \frac{1}{2}(\hat{r}\cdot\vec{x}')\vec{j} - \frac{1}{2}(\hat{r}\cdot\vec{j})\vec{x}'.$$

Recall that magnetic dipole is by definition

$$\vec{\mu} = \frac{1}{2c}\int d^3x'\vec{x}'\times\vec{j}.$$

We would like to have a term similar to the electric dipole, but we have a term like $\hat{r}\times 2c\vec{\mu}$ in $\vec{A}_2$. So we take one step furthur.

$$\begin{aligned}\frac{1}{2}\hat{r}\times(\vec{x}'\times\vec{j}) &= (\hat{r}\cdot\vec{x}')\vec{j} - \frac{1}{2}(\hat{r}\cdot\vec{x}')\vec{j} - \frac{1}{2}(\hat{r}\cdot\vec{j})\vec{x}' \\ &= (\hat{r}\cdot\vec{x}')\vec{j} - \frac{1}{2}\left((\hat{r}\cdot\vec{x}')\vec{j} - (\hat{r}\cdot\vec{j})\vec{x}'\right).\end{aligned}$$

So the term $\vec{A}_2$ becomes

$$\begin{aligned}\vec{A}_2 &= \frac{1}{c^2 r}\int d^3x'\left(\frac{1}{2}\hat{r}\times(\vec{x}'\times\vec{j}) + \frac{1}{2}\left((\hat{r}\cdot\vec{x}')\vec{j} + (\hat{r}\cdot\vec{j})\vec{x}'\right)\right) \\ &= \frac{1}{c^2 r}\partial_t\left(c\hat{r}\times\vec{\mu}_{ret}\right) + \frac{1}{c^2 r}\partial_t\int d^3x'\left(\frac{1}{2}\left((\hat{r}\cdot\vec{x}')\vec{j} + (\hat{r}\cdot\vec{j})\vec{x}'\right)\right) \\ &= \vec{A}_{M1} + \vec{A}_{E2}.\end{aligned}$$

Back to the procedure to find radiation power, we can find the radiation power for a specific case.

9.4.2 Lamor's Formula

Charge Moving in Gravity Feilds

A charge q is moving in a constant gravity field.

Vector potential for a charge moving in gravitational field is

$$\begin{aligned}\vec{A} &\approx \frac{1}{cr}\int d^3x'\vec{j}_{ret} \\ &\approx \frac{1}{cr}qg(t-\frac{r}{c})\hat{e}_z.\end{aligned}$$

The magnetic field is

$$\begin{aligned}\vec{B} &= \nabla\times\vec{A} \\ &\approx \nabla t_{ret}\times\hat{e}_z\frac{qg}{cr} \\ &\approx \frac{qg}{c^2 r}(-\hat{e}_r\times\hat{e}_z)\end{aligned}$$

Go on with the calculation,

$$\begin{aligned}\vec{S}\cdot\hat{e}_r &= \frac{c}{4\pi}\vec{E}\times\vec{B} \\ &= \frac{q^2 g^2}{4\pi c^3 r^2}\sin^2\theta\end{aligned}$$

Then we have the radiation power

$$\frac{dP}{d\Omega} = \frac{q^2 g^2}{4\pi c^3}\sin^2\theta.$$

9.5 Scattering

A total monochromatic wave $\vec{\epsilon} E_0 e^{i\vec{k}\cdot\vec{x}-i\omega t}$ shined on an object is described using

$$\vec{E}_{tot} = \vec{E}_{in} + \frac{E_0 e^{ikr-i\omega t}}{r}\vec{f}(\vec{k}).$$

Here the second term is the spherical scattered wave, while $\vec{f}(\vec{k})$ is for shape of scattered wave.

The nature of this kind of scattered wave is that the incident wave induced the object to emit some radiation. We only consider the radiation part not the close field region.

The differential cross section is defined to be the probility of light being scattered per. For the case of scattering of electromagnetic wave is

$$\frac{d\sigma}{d\Omega} = \frac{r^2 \hat{r}\cdot\langle\vec{S}_{sc}\rangle}{|\langle\vec{S}_{in}\rangle|}.$$

For transverse wave,

$$\hat{r}\cdot\langle\vec{S}_{sc}\rangle = \frac{c}{8\pi}\mathrm{Re}(\vec{E}^*_{sc}\times\vec{B}_{sc}),$$

and the final result of differential cross section becomes

$$\frac{d\sigma}{d\Omega} = |\vec{f}|^2.$$

CHAPTER 10

Relativity

10.1 Math in Relativity

10.1.1 Forms

Forms are used in many contexts of relativity. It might be difficult to visualize a general n-form, 1-form, on the other hand, carries a simple geometrical meaning even to physicists.

1-form can be viewed as the dual space of vectors. In many textbooks, vectors are named as contravariant vectors. In any case, **vectors are visualized using arrows**.

By definition, contraction of 1-form $\tilde{\omega}$ and a vector v^a should result in a number. In the field of relativity, we talk about real fields, so

$$\tilde{\omega} v^a \in \mathscr{R}.$$

A 1-form maps a vector to a real number. From this point of view, 1-form is a set of contour lines. Given this set of contour lines, it maps an arrow to a number.

Refs & Notes

10.2 Special Relativity

Notations

We use abstract index notation for most cases. A tensor with latin indices is an abstract tensor which tells us the rank of it but has no intention to indicate the components or basis.

We talk about transformations all the time in physics. For space and time, we also have the transformations from one reference frame to another, which plays the key role of our theories. Newtonian mechanics transform each space and

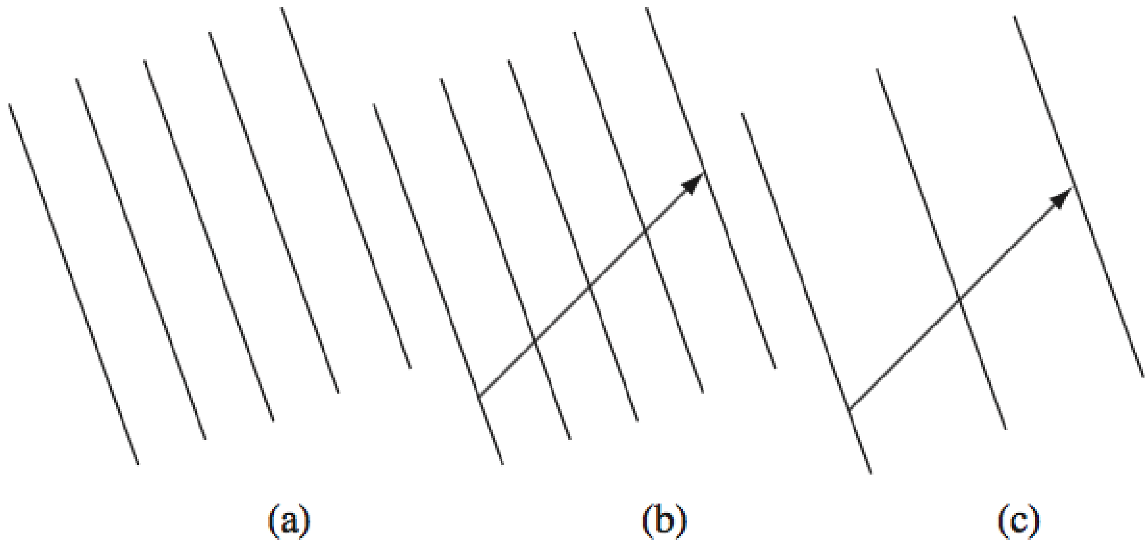


Fig. 10.1: 1-forms as contour lines. Figure (a) shows a 1-form using contour lines in a neighbourhood of a point. Figure (b) shows how a 1-form (contour lines) maps a vector (arrow) to a real number. In this case, we could assign the result real number as the number of contour lines that the arrow crossed. Different 1-forms (contour lines) take the same vector (arrow) to different real number, 4 and 2 using our definition for (b) and (c). Taken from *[Schutz]*.

time coordinates independently in translations. This idea was simply a extrapolation of our daily life experience that translations only change space coordinates accordingly, i.e.,

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}$$

for the two reference systems described in Fig. 10.2. This transformation matrix is neither symmetric nor Hermitian. It is ugly and unexpected. Why is time special and is not related to other coordinates?

As we think of this description, we would expect a most general transformation of coordinates for translation that involves all the coordinates, even for time.

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & L_{13} & L_{14} \\ L_{21} & L_{22} & L_{23} & L_{24} \\ L_{31} & L_{32} & L_{33} & L_{34} \\ L_{41} & L_{42} & L_{43} & L_{44} \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}.$$

How to determine this transformation matrix $\mathbf{L}$? Geometrically, we should conserve length since it's a scalar, which is

$$\eta^a_b x^a x^b, \tag{10.1}$$

where η_{ab} is the metric. More specifically,

$$\eta^a_b (L^a_c x^c)(L^b_d x^d) = \eta^a_b x^a x^b.$$

In order to derive special relativity, we have to deterimin this metric η_{ab}.

Physically or historically, we should preserve Maxwell's equation, since it has been proved to be true in different reference frames.

What people found is that the geometry is hyperbolic geometry, hence the metric is Minkowskian.

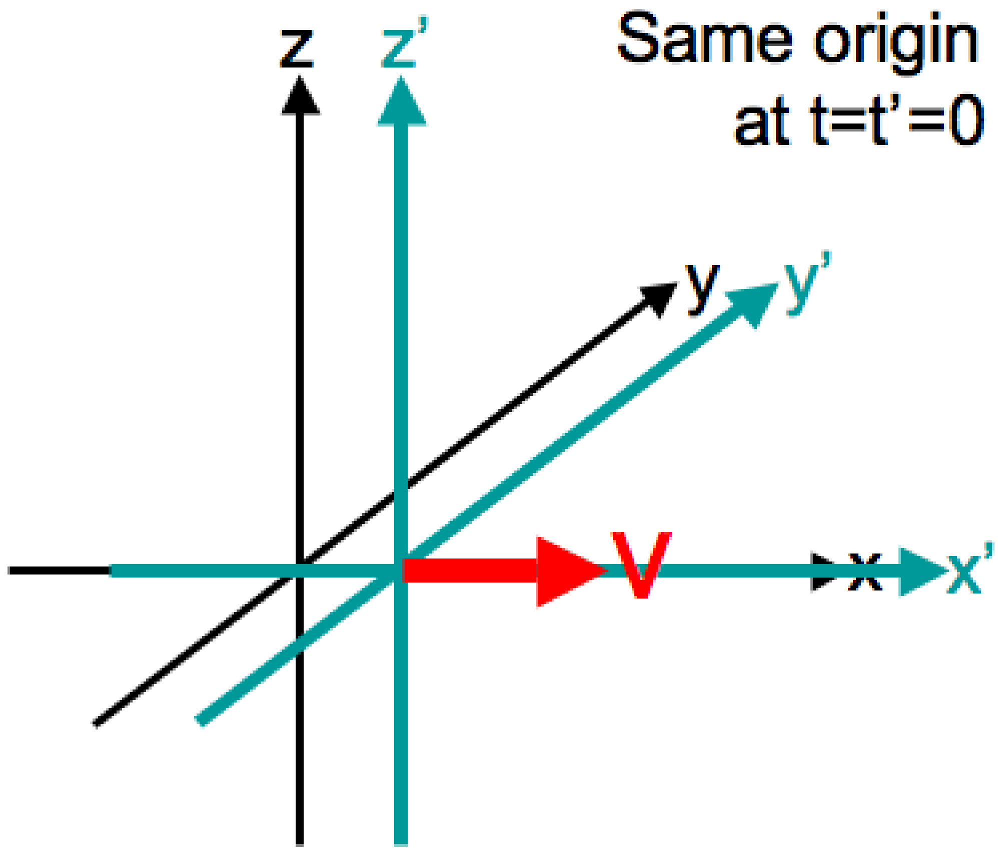


Fig. 10.2: Galilean transformation. Source: Wikipedia

10.2.1 Hyperbolic Geometry

Physics students are in fact quite familiar with the properties of hyperbolic geometry, even though some of the terms are not usually used in physics.

Visualizations of Hyperbolic Space

A lot of models are used to describe hyperbolic space

Numerious models are developed to describe the hyperbolic space.

1. Klein model;
2. Poincare model;
3. Gans model;
4. Weierstrass model or hyperboloid mode.

One of the useful visualizations of hyperbolic space is the the hyerboloid model, a.k.a. Weierstrass model. As the name indicates, hyperbolic space is embeded in Euclid space as a hyperboloid.

Hyperboloid on Two Sheets

A unit hyerboloid is described by the equation

$$x^2 + y^2 - z^2 = -1,$$

where x, y, z are the coordinates.

Lines on a hyperboloid is defined by the intersection of a plane with the hyperboloid.

We would also imagine that the so called light cone is basically

$$x^2 + y^2 - z^2 = 0.$$

For simplicity, we consider two dimensions, space x and time t. To build a theory of special relativity, we have to first specify the distance on hyperbola geometry. It's straightforward as it seems to be, we just extract distance from the definition of hyperbola since it is a conserved quantity,

$$\eta^\mu_\nu x^\mu x^\nu = x^2 - t^2.$$

For a standard hyperbola $x^2 - t^2 = 1$, we can parameterize the coordinates using a single parameter,

$$\begin{aligned} t &= \sinh \beta \\ x &= \cosh \beta, \end{aligned}$$

since

$$\cosh^2 \beta - \sinh^2 \beta = \Delta s^2.$$

Hyperbolic Trig Identities

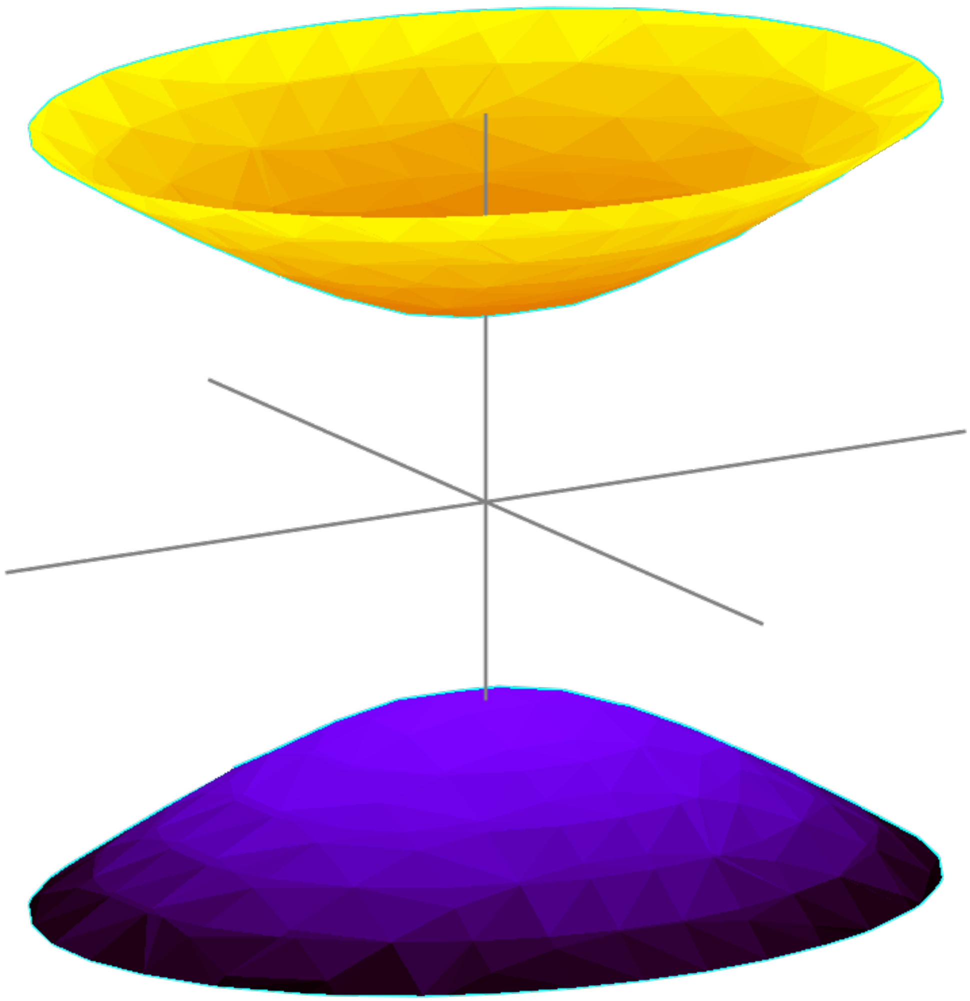


Fig. 10.3: Hyperboloid on two sheets. Source: Wikipedia.

As a review, the hyperbolic functions are defined as

$$\begin{aligned}
\sinh\beta =& \frac{e^x - e^{-x}}{2} \\
\cosh\beta =& \frac{e^x + e^{-x}}{2} \\
\tanh\beta =& \frac{\sinh\beta}{\cosh\beta} \\
\coth\beta =& \frac{1}{\tanh\beta} \\
\operatorname{csch}\beta =& \frac{1}{\sinh\beta} \\
\operatorname{sech}\beta =& \frac{1}{\cosh\beta}.
\end{aligned}$$

Hyperbolic trig functions have several identities that could help us with the understanding of the geometry.

$$\begin{aligned}
\cosh^2\beta - \sinh^2\beta =& 1 \\
\tanh^2\beta + \operatorname{sech}^2\beta =& 1 \\
\coth^2\beta - \operatorname{csch}^2\beta =& 1.
\end{aligned}$$

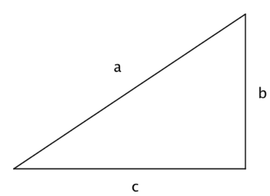


Fig. 10.4: Triangle of hyperbolic trig functions.

To visualize such relations, we draw an triangle, Fig. 10.4. We use this triangle to illustrate $\sin\beta = b/a = 3/4$

We choose the edge $b = 3$ and $a = 4$. Given the condition that

$$\begin{aligned}
\cosh\beta =& c/a \\
=& \sqrt{1 + \sinh^2\beta} \\
=& \sqrt{1 + 3^2/4^2} = 5/4,
\end{aligned}$$

we have to set

$$c = 5,$$

which is not so intuitive from Fig. 10.4.

10.2.2 Spacetime Diagram

How to find the time and space axes of an arbitary intertial reference frame?

The best approach I have ever read is in Schutz's book.

Suppose we have a frame O which we are sitting in. Another frame O' is moving with velocity v.

The time axis is the $x = 0$ points. We simply find out the line of object moving with velocity v.

The space axis is an equal time line. The invariant motion that is assumed in special relative is the motion of light. So we use light to find out this axis. We know light is always travelling with speed 1, which means it is always travelling with 45 degrees of angle in spacetime diagram, no matter what frame we are in. We fine equal time distance on t axis of frame O and t' axis of frame O', light emitted from $t' = -t'_0$ reflected on the $t' = 0$ point will be back to $x' = 0$ point but at time $t' = t'_0$. So we draw 45 degree lines from $(-t'_0, 0)$ and $(t'_0, 0)$ and let the two light beams intersect. The intersection point is a point on the space axis.

10.2.3 Basics of Special Relativity

The Postulates, Spacetime Diagram, and Metric

Special relativity was developed out of two postulates *[Schutz2009]*

1. Princple of relativity (Galileo),
2. Universality of speed of light (Einstein).

Using these two postulates, where the first key definition is interval of events squared

$$\Delta s^2,$$

we can derive basically all the relations we need. Some other intuitions will also be applied to the derivations.

Using a spacetime diagram, we can prove that this is invariant under transformation of frames *[Schutz2009]*.

Hyperbolic Space

If anyone realizes that spacetime is in fact hyperbolic space by looking at the expression of intervals Δs^2, the transformation is determined.

As we know the invariant quantity of the physical laws, the transformation of vectors can be found out of it, which is basically a rotation in hyperbolic space.

Metric Conventions

The metric in Eq. (10.1) is 'derived' from the interval.

To write it down, there are different convention. We choose the signature $+2$ metric in special relativity

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In most cases, we use natural unit $c = 1$.

---

d'Alembert operator

d'Alembert operator, or wave operator, is the Lapace operator in Minkowski space.[1]

$$\Box \equiv \partial_\mu \partial^\nu = \eta_{\mu\nu} \partial^\mu \partial^\nu$$

[1] Actually, there are more general definations for Lapacian, which includes this d'Alembertian of course.

In the usual {t,x,y,z} natural orthonormal basis,

$$\begin{aligned}\Box &= -\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2 \\ &= -\partial_t^2 + \Delta^2 \\ &= -\partial_t^2 + \nabla\end{aligned}$$

On wiki[2] , they give some applications to it.

- klein-Gordon equation $(\Box + m^2)\phi = 0$
- wave equation for electromagnetic field in vacuum: For the electromagnetic four-potential $\Box A^\mu = 0$footnote{Gauge}
- wave equation for small vibrations $\Box_c u(t,x) = 0 \rightarrow u_{tt} - c^2 u_{xx} = 0$

Hyperbolic Geometric Description

A Coincidence?

Let's start from this coincidence.

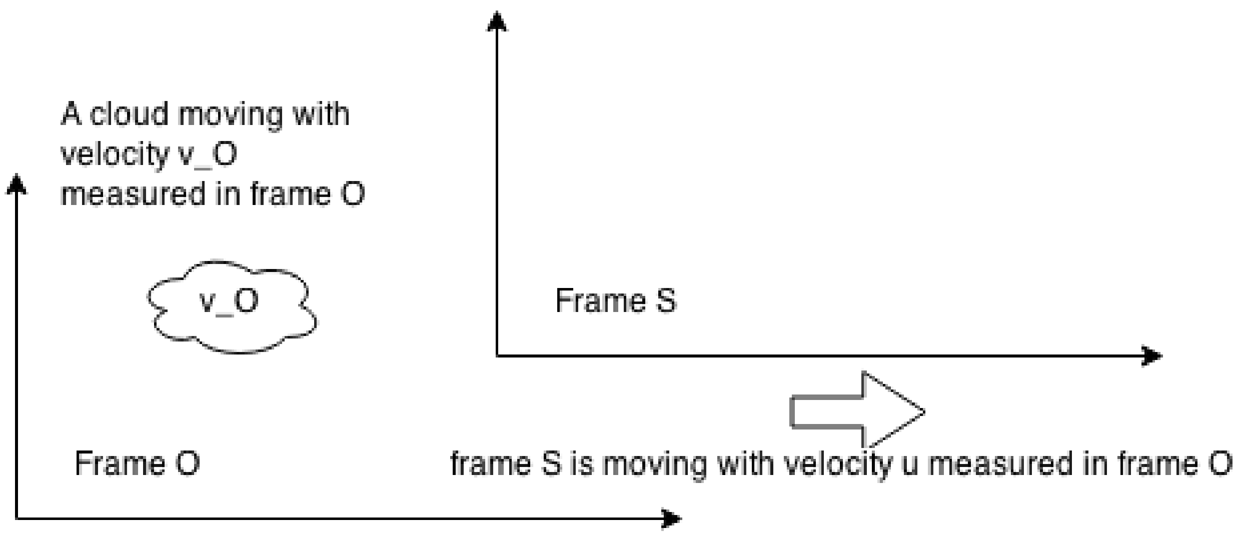


Fig. 10.5: Addition of velocities

Recall that in special relativity, velocity addition is

$$v_S = \frac{u + v_O}{1 + \beta v/c}, \tag{10.2}$$

where v_S is the velocity measured in moving frame S, v_O is the velocity measured in frame O. This β is the factor u/c where u is the velocity of the moving frame measure in frame O.

At the same time, we have the following hyper trig relation.

$$\tanh(\alpha + \beta) = \frac{\tanh\alpha + \tanh\beta}{1 + \tanh\alpha\tanh\beta}.$$

[2] wiki:D'Alembert_operator

Isn't this addition of angles the same as the velocity addition?

The algebra of relativity is mostly based on invariance of a new distance under a new rotation. Here we are not going to repeat the derivation of these transformations from the beginning, instead we would like to have a look at the really amazing part of this mathematical theory.

As shown in Fig. 10.5, we define quantities in two different frames, the frame O and frame S. The velocity of frame S measured in frame O is u. Out of this velocity we define a quantity

$$\tanh\alpha_u = \frac{u}{c},$$

In fact, any velocity divided by speed of light should be a hyperbolic tangent,

$$\tanh\alpha_{v_x} = \frac{v_x}{c}.$$

With this definition of hyperbolic tangent, we notice that

$$\gamma = \frac{1}{\sqrt{1-u^2/c^2}} = \cosh\alpha_u.$$

Suppose we have an object moving with velocity v_S in frame S. The velocity measured in frame O is the addition of the velocity of frame S itself and the velocity v_S, except the addition rule is not the usual plus but the rule stated in Eq. ((10.2)). We apply the definitions of the hyperbolic trig function,

$$\frac{v_S}{c} = \tanh(\alpha_u + \alpha_{v_O}) = \frac{\tanh\alpha_u + \tanh\alpha_{v_0}}{1+\tanh\alpha_U \tanh\alpha_{v_O}} = \frac{u/c + v_O/c}{1+\frac{u}{c}\frac{v_O}{c}}.$$

We could imagine the algebra of velocities would be simply summations if we define 'velocity' as $\arctan\frac{v_x}{c}$.

Addition of velocities is not that fundamental. What's more important is the transformation of coordinate, as we have always been talking about. In the old school language, the coordinate transformation is

$$\begin{pmatrix} t_O \\ x_O \end{pmatrix} = \gamma \begin{pmatrix} 1 & u/c^2 \\ u & 1 \end{pmatrix} \begin{pmatrix} t_S \\ x_S \end{pmatrix},$$

where

$$\gamma = \frac{1}{\sqrt{1-u^2/c^2}} = \cosh\alpha_u.$$

If we use the language of hyperbolic trig functions, this transformation becomes

$$\begin{aligned} \begin{pmatrix} t_O \\ x_O \end{pmatrix} &= \cosh\alpha_u \begin{pmatrix} 1 & (\tanh\alpha_u)/c \\ c(\tanh\alpha_u) & 1 \end{pmatrix} \begin{pmatrix} t_S \\ x_S \end{pmatrix} \\ &= \begin{pmatrix} \cosh\alpha_u & (\sin\alpha_u)/c \\ c(\sin\alpha_u) & \cosh\alpha_u \end{pmatrix} \begin{pmatrix} t_S \\ x_S \end{pmatrix}. \end{aligned}$$

To make the transformation symmetric, we consider

$$\begin{pmatrix} ct_O \\ x_O \end{pmatrix} = \begin{pmatrix} \cosh\alpha_u & \sin\alpha_u \\ \sin\alpha_u & \cosh\alpha_u \end{pmatrix} \begin{pmatrix} ct_S \\ x_S \end{pmatrix}.$$

Natural Unit

Look at these tedious steps. Why not just use natural units and set $c = 1$. We should.

This is basically the rotation matrix in hyperbolic spacetime.

Rotation in Euclidean Space

The rotations in Euclidean space is described as

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

It is quite different from the rotations in Euclidean space.

Since we are talking about geometry, space-time diagram will be extremely important. The length contraction, time dilation, and even doppler shift can be explained and calculated using the hyperbolic trig functions. Triangles on the space-time diagram are described in *Visualizations of Hyperbolic Space*.

Time Dilation

Use a spacetime diagram.

Length Contraction

Use a spacetime diagram.

Footnotes

1. *The Geometry of Special Relativity* by Tevian Dray.

10.2.4 Doppler Effect

Doppler shift in special relativity is always confusing. I'll demonstrate doppler shift in four different ways.

Conservation of Four Momentum

The special relativistic doppler shift can be derived using the fact that 4-momentum is a vector thus it transforms under Lorentz transformation.

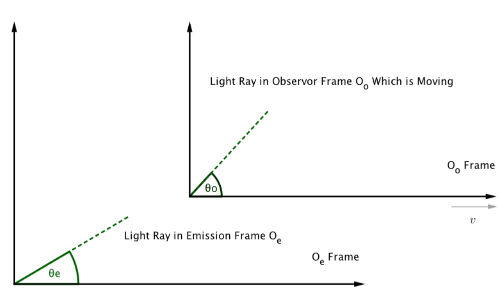


Fig. 10.6: The observer frame is moving in x direction only.

The observer is fixed in observer frame and source of emission is in emission frame.

The observer detects an angle of light ray θ_o. However, the emission angle θ_e is different from this angle, i.e., $\theta_e \neq \theta_0$.

Which Angle to Use

Which angle to use then? In theory, we can derive the Doppler shift in terms of either angles. We present both of the analysis. The redshift expressions looks different because they are measuring different events.

Components of four momentum in observer frame is

$$p_o^\mu = \begin{pmatrix} E_o & E_o \cos\theta_o & E_o \sin\theta_o & 0 \end{pmatrix}.$$

The components in the emission frame is

$$p_e^\mu = \begin{pmatrix} E_e & E_e \cos\theta_e & E_e \sin\theta_e & 0 \end{pmatrix}.$$

Redshift

Redshift is define as

$$\begin{aligned} z &= \frac{\nu_e - \nu_o}{\nu_o} \\ &= \frac{\omega_e - \omega_o}{\omega_o}. \end{aligned}$$

Non-relativistic Doppler Shift

To understand the effect of relativity, we would first recall the non-relativistic doppler shift.

$$\omega_{o,non-rel} = \omega_{e,non-rel}(1 - v/c\cos\theta). \tag{10.3}$$

where no γ is relavent. It's obvious that we have only two kinds of shift, redshift due to the source is closing, or blueshift due to the fact that the source is moving away.

Redshift Relation to Angle in Emission Frame

Since momentum is a vector, we have the Lorentz transformation which transfrom it in to O_o frame,

$$\frac{E_o}{c} = \gamma\left(\frac{E_e}{c} - \beta p_e^1\right),$$

where we also have

$$\begin{aligned} p_e^1 &= p_e^0 \cos\theta_e, \\ p_e^0 &= E_e/c. \end{aligned}$$

Combining these equations, the energy of the photons in O_e frame is

$$E_o = E_e\gamma(1 - \beta\cos\theta_e).$$

In quantum mechanics, energy is related to angular frequency,

$$E_{o/e} = \hbar\omega_{o/e}.$$

The angular frequency in O' frame is

$$\omega_o = \omega_e\gamma(1 - \beta\cos\theta_e). \tag{10.4}$$

In fact, $\beta = v$ if we choose $c = 1$.

Redshift Relation to Angle in Observor Frame

The problem is, we know that $\theta_e \neq \theta_o$. In many cases, it's more convinient to obtain the angle in observation frame θ_o.

In this case, we Lorentz transform the representation of four momentum in observer frame (p_o^μ) to emission frame (p_e^μ).

$$\frac{E_e}{c} = \gamma(\frac{E_o}{c} + \beta p_o^1),$$

where

$$p_o^1 = p_o^0 \cos\theta_o$$
$$p_o^0 = E_o/c.$$

We solve the angular frequency for photons in observer's frame

$$\begin{aligned}\omega_o &= \frac{\omega_e}{\gamma(1+\beta\cos\theta_o)} \\ &= \omega_e \frac{\sqrt{1-\beta^2}}{(1+\beta\cos\theta_o)},\end{aligned} \tag{10.5}$$

where in the last step we applied

$$\gamma = 1/\sqrt{1-\beta^2}.$$

Eq. (10.5) seems to be very different from (10.4). The reason is that we are measuring different events, due to the difference between θ_e and θ_o.

Eq. (10.5) is what usually used in discussion of relativistic Doppler effect.

Line-of-sight Direction Same as Relative Velocity Direction

$\theta_o = \theta_e = 0$ gives us the most used Doppler shift

$$\omega_o = \omega_e \sqrt{\frac{1-\beta}{1+\beta}}. \tag{10.6}$$

$\beta = v$ if we choose $c = 1$.

Relativistic Effect

Nonrelativistic Doppler shift (10.3) contains only the effect of line-of-sight relocity.

Relativistic Doppler shift (10.5) we have new contributions from relative velocity, which is the transverse redshift due to the γ factor or the contraction of time.

If the motion is in line-of-sight, Eq. (10.6) is reduced to the nonrelativistic Doppler shift for slow velocity v as we take only first order of its Tayler series.

For motion that is not along the line-of-sight, angle difference becomes important, since we have to choose the equal time surface.

An gif from wikipedia shows this explicitly,

The change in wavelength is given by

$$\frac{\lambda_{obs}}{\lambda_{src}} = \sqrt{\frac{1-\beta}{1+\beta}}.$$

Fig. 10.7: Image Source: File:XYCoordinates.gif

Four Vector Language

Doppler shift can be solve using abstract four vectors without going into a coordinate system *[CBLiang]*.

First of all we associate the emission frame and observation frame with its own four velocities, V^a and U^a.

From the knowledge of special relativity, we know that

$$E_o/c = -p^a U_a|_{obs},$$

where U^a is the four velocity of observer, subscript $_{obs}$ indicates this is the measurement at observation point.

The photon energy in emission frame is

$$E_e/c = -p^a V_a|_{em}, \tag{10.7}$$

which is calculated at the point of emission. V^a is the four velocity of the emission frame.

Since inner product is independent of the physical point in special relativity, we can calculate the both energy at the same physical point.

We also know that

$$\gamma = -V^a U_a,$$

which associates the four velcoities

$$U_a = \gamma V_a + \gamma u_a, \tag{10.8}$$

where γu_a is the three velocity component viewed by V_a.

The Instataneous Frame Decomposition

We always have to define our equal time surface first. Here we can see that γu_a is indeed a decomposition onto our spatial surface.

Multiply on both sides of Eq. (10.8) by V^a,

$$V^a u_a = 0.$$

We also know from Eq. (10.7)

$$p^a = \omega V^a + k^a.$$

To collect our thoughts, we have obtained from the four momentum and our velocities u^a and k^a, which are the quantities we would like to work on in 3D space.

Finally, we calculate the frequency in observer's frame,

$$\begin{aligned}\omega_o &= -p^a U_a \\ &= -(\omega V^a + k^a)(\gamma V_a + \gamma u_a) \\ &= \gamma(\omega - k^a u_a),\end{aligned}$$

where

$$k^a u_a = v\omega \cos\theta_e.$$

Then we obtain the Doppler shift equation

$$\omega_o = \omega_e \gamma (1 - v\cos\theta_e).$$

We can work out this using another projection of spatial dimensions, which is give us the frequency relation in terms of observation angle θ_o.

Spacetime Diagram

Needless to say, it can be explained using spacetime diagram. The only caveat is to pay attention to the equal time surface.

I am just too lazy to make a spacetime diagram with six axes here. You get the idea.

10.2.5 Relativistic Aberration of Light

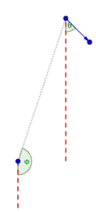


Fig. 10.8: The relativistic aberration of light.

In astrophyics, object moving with a significant fraction of the speed of light v_S with angle θ shown in Fig. 10.8, is measure from the observer in a direction

$$\cos\theta_O = \frac{\cos\theta_O - v_S}{1 - v_S\cos\theta_S}.$$

If the object is moving towards us, we observe $\theta_O = 0$.

Meanwhile the apparent transverse velocity is measured to be

$$v_{\perp,O} = \frac{v}{1 - \beta\cos(\pi - \phi + \theta)}, \tag{10.9}$$

where $\pi - \phi + \theta$ is the angle between line of sight and the velocity, measured in the object's frame. One of the astonishing fact about Eq. ((10.9)) is that it's maximun value can be larger than 1, which means we could observe superluminal objects.

10.3 General Relativity

General relativity is a theory of gravity. The idea is to find a set of "proper" coordinate system to describe physics on a curved space and make connection between these "proper" coordinate systems.

10.3.1 Geometrized Unit

In general relativity, it's quite useful to use geometrized unit, where everything has unit of kilometers. *[Schutz]*

The principle of geometrized unit is to convert everything to length using $c = G = 1$. A precalculated value is

$$1 = G/c^2 = 7.425 \times 10^{-28}\mathrm{mkg}^{-1}.$$

In the spirit of this conversion, we have the mass of the Sun $M_\odot = 2.0 \times 10^{30}\mathrm{kg} = 1.5 \times 10^3\mathrm{m}$ and mass of earth $M_\oplus = 6.0 \times 10^{24}\mathrm{kg} = 4.4 \times 10^{-3}\mathrm{m}$.

References and Notes

10.3.2 Basic Principles of General Relativity

Tidal Force and Equivalence Principle

We can imagine from the Einstein's elevator thought experiment that we could not tell whether we are in a inertial frame or free falling frame by measuring forces. This is generalized to the weak equilvalence principle.

Weak Equivalence Principle

Uniform gravitational field are equivalent to frames that are accelerating uniformly.

On the other hand, we know that tidal force in the frame work of Newtonian gravity can be derived by finding the second order derivative of the displacement difference between two nearby objects. A free falling in non-uniform gravitation is distinguishable from inertial frame because we can measure the tidal force.

However, the free falling frame is no different than inertial frame **if the two object are close enough** since the comoving equipment we are using to measure the tidal effect could not tell the tidal effect.

The weak equivalence principle seems to work in limited circumstances. A stronger version is called the Einstein's equivalence principle which states that all physics are the same in a local spacetime. The word "same" means the equations have the same form thus requires tensor equations.

Examples of Equivalence Principle

The first example that can be easily worked out is the redshift of photons in gravitational field, or the Pound-Rebka-Snider experiment.

Using the equivalence principle, we expect that the photon doesn't change when we measure everything in a freely falling frame.

Suppose the source emits a photon when our free-fall starts. It takes the photon $\Delta t = h$ to climb up to a height h. When we measure the photon at the top, our frame (measurement) is done with a relative velocity gh compared to the beginning of the experiment. Thus we experience Doppler shift (first order) of the photon,

$$\frac{\nu'_h}{\nu_h} = 1 + gh,$$

to the first order, where ν'_h is the frequency measured in free-falling frame and ν_h is the frequency of photon in the lab frame.

By arguing using equivalence principle we know that ν'_h is the same as the emission frequency $\nu'_e = \nu_e$.

Effectively, gh is the potential energy the photon loses during the climbing if we measure in lab frame.

References and Notes

10.3.3 Mathematics in General Relativity

Vectors and Tensors

Two Systems of Notations

Come back to this when the index fearing syndrome becomes too strong.

There are many different systems of notations for vectors.

One of them is to use a tilde on top of the letter to denote a co-vector, i.e.,

$$\tilde{v}, \qquad \text{dual vector, one-form, co-vector}$$

The other notation that is widely used is abstract index notation where we use latin superscript to denote vector and latin subscript to denote co-vector, i.e.,

$$\begin{aligned} v^a, &\qquad \text{vector} \\ v_a, &\qquad \text{co-vector.} \end{aligned}$$

That is to say, v_a is basically $\tilde{v}$.

The question is, obviously, how the components of vectors is denoted. In the first notation, we use subscript (components of vector) and superscript (components of co-vector) for the components,

$$v_\mu \equiv \tilde{v}(e_\mu), \qquad \mu \text{ component of dual vector,}$$

where e_μ is the μ basis.

However, the abstract index notation is using greek superscript for co-vector component and greek subscript for vector component.

$$\begin{aligned} v^a_\mu, &\qquad \mu \text{ component of vector} \\ v^\mu_a, &\qquad \mu \text{ component of co-vector.} \end{aligned}$$

For basis vectors, we usually denote them as $\{e_\mu\}$. The dual space, which are basis of dual vectors are denoted as $\{\tilde{e}^\mu\}$, with

$$\tilde{e}^\mu e_\nu = \delta^\mu_\nu.$$

Whenever a vector is mentioned, it is composed of its components and basis, which should be written as

$$\alpha^a = \alpha^\mu {e^a}_\mu,$$

where $\{\alpha^\mu\}$ are the components and $\{{e^a}_\mu\}$ are the basis.

Then we define its dual w_b

$$w_a \alpha^a = \mathscr{C},$$

where the left hand side can be expanded

$$w_a \alpha^a = w_\lambda {f_a}^\lambda \alpha^\mu {e^a}_\mu = w_\lambda \alpha^\mu {f_a}^\lambda {e^a}_\mu.$$

For orthonormal basis, we have

$$f_a{}^{\lambda} e^a{}_{\mu} = \delta^{\lambda}_{\mu},$$

which gives us the contraction of w_a and α^a

$$w_a \alpha^a = w_\lambda \alpha^\mu \delta^\lambda_\mu = w_\mu \alpha^\mu.$$

Usually the vector basis could be $\{dx^\mu\}$ and we could derive the basis for the dual vector $\{\frac{d}{dx^\mu}\}$.

Metric

We define a scalar product in this way

$$\beta^a \cdot w^a = \beta^a g_{ab} \cdot w^b,$$

where g_{ab} is the metric and can be related to the basis.

The expression will be simplified using basis formalism,

$$\beta^a \cdot w^a = \beta^a g_{ab} \cdot w^b = \beta^\mu e^a{}_\mu g_{ab} w^\lambda f^b{}_\lambda = \beta^\mu w^\lambda g_{\mu\lambda}.$$

Description of Space-time Manifold

How to describe space-time manifold?

- Metric (with a set of local coordinates), connection (Christoffel symbols).
- Metric (in the form of tetrads), connection (Ricci rotation coefficients).
- 1+3 covariantly defined variables.

Description of Space-time Manifold - Coordinates

Description of Space-time Manifold - Tetrads

Description of Space-time Manifold - 1+3 Covariant Description

Physics in description is easier to understand.

Definations

Definations of some physical quantities and operators are listed below.

Here we have

1. **geometrical variables**: Volume
2. **Kinematical variables**: Velocity, Expansion rate, Shear rate
3. **Thermaldynanmical variables**: Energy density, Momentum density, Pressure, Equation of state

Volume

To calculate volume, the volume element should be defined first in order to integrate. Before that, orientation on manifolds is to be figured out.

On an oriented manifold with metric, the defined volume element (a n-form) should be compatible with the orientation and also determined by the metric.[1]

Introducing those requirements, a compatible volume element is

$$\epsilon_{a_1\cdots a_n} = \pm\sqrt{|g|}(e^1)_{a_1} \wedge \cdots \wedge (e^n)_{a_n}$$

Alternatively, this can be expressed in the way Ellis used in arXiv:gr-qc/9812046v5.

$$\eta_{abcd} = \eta_{[abcd]}, \quad \text{with} \eta_{0123} = \sqrt{|\text{det} g_{ab}|}$$

Induced volume element $\hat{\epsilon}_{a_1\cdots a_{n-1}}$ is defined use the normal vector u^a of the hypersurface,

$$\hat{\epsilon}_{a_1\cdots a_{n-1}} = u^b \epsilon_{ba_1\cdots a_{n-1}}$$

4-velocity

4-velocity of observed matter is

$$u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}$$

with $u^\alpha u_\alpha = -1$, τ is the proper time along the worldlines of investaged matter.

Projection Tensors

We can use 4-velocity to project variables to parts that is parallel to u^α and parts that is orthogonal to u^α.

$$\begin{aligned} U^a{}_b &= -u^a u_b \\ h_{ab} &= g_{ab} + u_a u_b, \qquad \text{induced metric from } g_{ab} \end{aligned}$$

Some properties of the two projections.

$$\begin{aligned} &U^a{}_b U^b{}_c = U^a{}_c, U^a{}_a = 1, U_{ab} = g_{ac}U^c{}_b, U_{ab}u^b = -g_{ac}u^c u_b u^b = u_a \\ &h^a{}_b = g^{ac}h_{cb} = \delta^a{}_b + u^a u_b = \delta^a{}_b - U^a{}_b \\ &h^a{}_c h^c{}_b = (\delta^a{}_c - U^a{}_c)(\delta^c{}_b - U^c{}_b) = \delta^a{}_b - U^a{}_b = h^a{}_b \\ &h^a{}_a = 4 - 1 = 3, h_{ab}u^b = 0 \end{aligned}$$

Covariant time derivative (˙)

This is the derivative along the fundamental worldlines (projection on the worldlines),

$$\dot{T}^{ab}{}_{cd} = u^e \nabla_e T^{ab}{}_{cd}$$

[1] For more information, check out Canbin Liang's book. Volume 1, page 115.

Fully orthogonally projected covariant derivative ($\tilde{\nabla}$)

This derivative is the project orghogonal to the normal vector of the hyperspace or orthogonal to the observer's 4-velocity or along the tagent of the hyperspace.

$$\tilde{\nabla}_e T^{ab}{}_{cd} = h^a_f h^b_g h^p_c h^q_d h^r_e \nabla_r T^{fg}{}_{pq}$$

Orthogonal projections of vectors

Orthogonal projection of vectors

$$v^{<a>} = h^a{}_b v^b$$

And the orthogonally projected symmetric trace-free part of tensors

$$T^{<ab>} = [h^{(a}{}_c h^{b)}{}_d - \frac{1}{3} h^{ab} h_{cd}] T^{cd}$$

Othogonal projected covariant time derivatives along u^a

$$\dot{v}^{<a>} = h^a{}_b \dot{v}^b$$
$$\dot{T}^{<ab>} = [h^{(a}{}_b h^{b)}{}_d - \frac{1}{3} h^{ab} h_{cd}] \dot{T}^{cd}$$

Properties

- Projected time and space derivatives of U_{ab}, h_{ab} and η_{abc} vanish.

10.3.4 Curved Spacetime

Christoffel Symbol

By definition, Christoffel symbol is defined through

$$\frac{\partial}{\partial x^\beta} \mathbf{e}_\alpha = \Gamma^\mu_{\alpha\beta} \mathbf{e}_\mu .$$

So what it means geometrically, is the small change in the basis vector $\mathbf{e}_\alpha$ when we change the coordinate x^β, then project it on to the basis vector $\mathbf{e}_\mu$.

In polar coordinate system, the basis change when we move from one point to another. At point A, the basis vectors are shown as red while the basis vectors at point B are shown as black. The two sets of basis vectors are different when we look at it. The change of the vectors are described by $\frac{\partial}{\partial x^\beta}\mathbf{e}_\alpha$, which are shown as dotted vectors $\Delta \mathbf{e}_\theta$ and $\Delta \mathbf{e}_r$. These are calculatable fairly easily. Then we project these vectors onto the basis to get the components of Christoffel symbol.

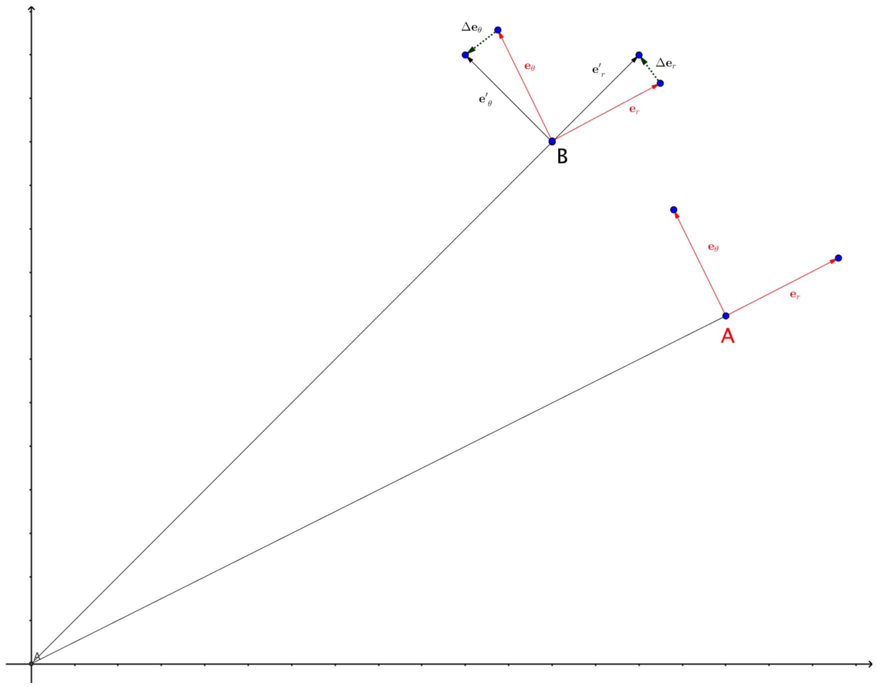


Fig. 10.9: Christoffel symbol tells us the change in the basis when we move around a little bit. This is an example of polar coordinate system.

Curvature

1. Usually curvature is about calculating distances locally. If the distances are the same as flat space, it dosn't have **intrinsic curvature**, for example, a cylinder. However, a cylinder has **extrinsic curvature** because it is curved by our common sense.
2. We can use a plane that is tangent to a point on a surface to determine wether it has intrinsic curvature. Move the plane parallelly up and down from this initial point. If the interceptions are straight lines, we do not expect intrinsic curvature. If the interceptions are conic sections, we expect it to has intrinsic curvature.
3. This indicates that we need to compare two different directions to really know wether the curvature is intrinsic.
4. In fact Gaussian curvature can be calculated using two principal curvatures in different directions.
5. Gauss map is the way that Gauss defined to calculate the Gaussian curvature. It maps an area from the surface onto a unit sphere.

Parallel Transport

1. On a sphere, transporting a vector leads to strange results.
2. All the tangent vectors at different parameters of the curve specifies the curve itself.
3. Parallel transport generally require the vector to be parallel locally with each infinitesimal move.
4. Parallel transport can also be explained as the components along of the vector that we are transporting doesn't change.
5. Mathematically, $\frac{d}{d\lambda} V^\alpha = 0$ at a locally inertial frame.
6. It can be written as covariant derivative thus genralized to all frames. $\frac{d}{d\lambda} V^\alpha = U^\beta V^\alpha_{,\beta} = U^\beta V^\alpha_{;\beta} = 0$.

7. Another notation is $\nabla_{\mathbf{V}}\mathbf{V} = 0$.

Geodesics

1. Euclid's straight lines indicates that the direction is the key to define straight lines.
2. They are lines that is formed by parallel transporting their own tagent vectors.
3. Generalize to curved spacetime. $\nabla_{\mathbf{V}}\mathbf{V} = 0$. It is dubbed as the **geodesic**.
4. We can find all the coordinates of the geodesic by using the definition of tagent vectors.

$$\frac{d}{d\lambda}\left(\frac{dx^\alpha}{d\lambda}\right) + \Gamma^\alpha_{\mu\alpha}\frac{dx^\mu}{d\lambda}\frac{dx^\alpha}{d\lambda} = 0.$$

4. Second order DE: given initial point and the inital tagent vector we can solve it.
5. λ is the **affine parameter**.
6. A linear transformation of the affine parameter is usually still a affine parameter.
7. Solve the equation for inertial frame and draw the lines. We'll see it really describes a straight line.

Geodesics are the lines that describe the extemal length between two points. To prove that we need to write down the length between two lines. From $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$,

$$l = \int_{\lambda_1}^{\lambda^2} \sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} d\lambda.$$

Then we do the variation of length. Before we actually do it, it's good to think about what the plan is. We will derive a equation, aka, the geodesic equation, for all coordinates, which indicates that we need to derive a Euler-Lagrange like equation. So we will finally write down the variation of length in a form

$$\delta l \sim \int_{\lambda_1}^{\lambda_2} (\cdots)_\alpha \delta x^\alpha d\lambda,$$

which should be 0 if we require it to be the extemal length.

Mathematics to Prove Geodesic is the Extemal Line

In principle, we could define a Lagrangian and use Euler-Lagrange equation. But here I will demonstrate it using the variation principle.

$$\begin{aligned}\delta l &= \frac{1}{2}\int_{\lambda_1}^{\lambda_2}\left(g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right)^{-1/2}\left((\delta g_{\alpha\beta})\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} + g_{\alpha\beta}\left(\delta\frac{dx^\alpha}{d\lambda}\right)\frac{dx^\beta}{d\lambda} + g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\left(\delta\frac{dx^\beta}{d\lambda}\right)\right)d\lambda \\ &= \frac{1}{2}\int_{\lambda_1}^{\lambda_2}\left(g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right)^{-1/2}\left((\delta g_{\alpha\beta})\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} + 2g_{\alpha\beta}\left(\delta\frac{dx^\alpha}{d\lambda}\right)\frac{dx^\beta}{d\lambda}\right)d\lambda,\end{aligned}$$

where we used the symmetry in :math:"{}_{alphabeta} in the last step.

So we need to calculate $\delta g_{\alpha\beta}$

$$\begin{aligned}\delta g_{\alpha\beta} &= \frac{g_{\alpha\beta}(x^\mu + \delta x^\mu) - g_{\alpha\beta}(x^\mu)}{\delta x^\mu}\delta x^\mu = g_{\alpha\beta,\mu}\delta x^\mu \\ \delta\frac{dx^\alpha}{d\lambda} &= \frac{d\delta x^\alpha}{d\lambda}.\end{aligned}$$

Plug this in and sort out the total derivatives then we have an expression

$$\delta l = \frac{1}{2}\int_{\lambda_1}^{\lambda_2} S\left(g_{\alpha\beta,\mu}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} - 2\frac{d}{d\lambda}\left(g_{\alpha\mu}\frac{dx^\alpha}{d\lambda}\right)\right),$$

where we defined

$$S = \left(g_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} \right)^{-1/2}$$

which is a constant since it simply measures the scaling of parameters and length.

Then we use the symmetries in metric and get the Euler-Lagrangian equation which is basically the geodesic equation.

---

1. The longest length between two points is the time-like geodesic.
2. The time-like geodesic is not necessarily the longest line between two points.
3. We can not find shortest time-like lines between two points.

References and Notes

1. Bernard F. Schutz, A first course in general relativity. Chapter 6.

10.3.5 Energy Momentum Tensor

Energy momentum tensor is an important concept when dealing with continuum media.

In general, what we would like to define is a tensor that contains the energy density.

First of all, energy density obviously is not a conserved quantity. As an example, we consider a number of particles with number density n and each with mass m. In its comoving frame, we would define the energy density as $\rho = nm$ since every single particle is stationary. When we transform to another frame, say $\bar{O}$ frame, $\bar{\rho} = \gamma^2 nm$, which indicates that this quantity is not a scalar.

So to achieve this goal of an invariant quantity, we need a tensor. Suppose its components are denoted as $T^{\alpha\beta}$, we need to find a definition that carries the following meanings.

1. T^{00} is energy density.
2. T^{0i} is energy flux.
3. T^{i0} is momentum density.
4. T^{ij} is momentum flux. In this sense Tii has the meaning of pressure.

For perfect fluid, the definition that satisfies the requirements is

$$T^{ab} = (\rho + p) U^a U^b + p g^{ab}.$$

10.3.6 Gravitational Waves

In the weak field regime of sourceless Einstein's equation ($T^{\mu\nu} = 0$), the equation for metric with perturbations is reduced to a wave equation,

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2 \right) \bar{h}^{\alpha\beta} = 0,$$

where $\bar{h}^{\alpha\beta}$ is the trace-reversed perturbation of the metric on top of Minkowski metric background, i.e.,

$$\bar{h}^{\alpha\beta} = h^{\alpha\beta} - \frac{1}{2} \eta^{\alpha\beta} h,$$

where $h^{\alpha\beta} = g^{\alpha\beta} - \eta^{\alpha\beta}$ and h is the trace of metric perturbation $h^{\alpha\beta}$.

Trace Reverse

The tensor $\bar{h}^{\alpha\beta}$ is called trace reverse of $h^{\alpha\beta}$ for its trace is $-h$.

Gauge

To solve the equation we introduce a solution of the form $\hat{h}^{\alpha\beta} = A^{\alpha\beta} e^{ik_\mu x^\mu}$, which simiplifies the equation

$$\eta^{\mu\nu} k_\mu k_\nu \bar{h}^{\alpha\beta} = 0.$$

To solve the amplitude A^α we need constraints on it. We can derive that gravitational waves are always null, that is $k^\mu k_\mu = 0$.

Some of the conditions requires a gauge transformation. In any case, we have the second gauge condition as

$$A_{\alpha\beta} U^\beta = 0,$$

which specifies that $A_{\alpha\beta}$ is orthogonal to the vector we chose U^β. A practical choice of U^β is a four velocity. This removes another **four degrees of freedom**. For illustration purpose, we choose $U^\beta \to (1, 0, 0, 0)$ since it's a null vector. The degrees of freedom removed can be visualized as the first rwo and column.

The second one we can think of is a transverse condition,

$$A_{\alpha\beta} k^\beta = 0,$$

which removes **another three degrees of freedom**. This specifies that the wave is transverse, i.e., $A_{\alpha\beta}$ can not have elements that is in the direction of four wavevector. We specify a wavevector $k^\beta \to (\omega, 0, 0, \omega)$, which leads to the removal of the remaining elements of the fourth row and column.

The matrix we have now becomes

$$A_{\alpha\beta} \to \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{xx} & A_{xy} & 0 \\ 0 & A_{yx} & A_{yy} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The last gauge condition is traceless condition $A^\alpha_\alpha = 0$ which also requires the gauge transformation. This condition fixes the phase relations between different spatial directions, that is $A_{xx} = e^{i\pi} A_{yy} = -A_{yy}$. This conditions insists that the two directions of distance oscillations should be quadrupole-like, i.e., contracts in one direction (say x) while extend in the other direction (say y).

Slicing

The first two conditions are basically specifying slicings of spacetime.

Physical Significance of Transverse-traceless Gauge

Transverse-traceless gauge is the very gauge that determines a coordinate system that a test particle is stationary in terms of coordinates.

To show this we assume that we have a test particle being stationary initially, i.e., $U^\alpha|_{\tau=0} \to (1, 0, 0, 0)^{\mathrm{T}}$.

The particle should travel on geodesics,

$$\frac{d}{d\tau}U^\alpha + \Gamma^\alpha_{\mu\nu}U^\mu U^\nu = 0,$$

which leads to

$$\frac{d}{d\tau}U^\alpha|_{\tau=0} = -\Gamma^\alpha_{00} = 0.$$

The four acceleration is 0 for the test particle. No motion would be detected within the coordinate system.

The same is true for a particle moving in z direction. However, the conclusion doesn't hold for other motions. p

10.3.7 General Relativity Revisited

This section lists the experiments which are used to test gravity theories carried out on the earth.

The test of gravity theories can be viewed as test of the fundations of gravity theories and the the theories themselves, say test of equivalent principle and general relativity or f(R) gravity theory. Thus we should break down general relativity theory into several stages. Here, we use the following table to do so.

- **Physical Fundations: Hyperthesis**:

Theory	Mach	WEP	EEP	SEP	GC	Notes
GR	Partial	Y	Y	Y	Y	

- **Mathematical Description**:

Theory	Topoplogy	Manifold	Connection	Metric
GR			No torsion	Non-metricity tensor vanishes

- **Theoretical Implifications**:

Theory	Gravitational Waves	Newtonian Limit	GR Limit	Notes
GR	Y	Y		

Most items in mathematics are the same in different theories.

Hyperthesis

- **WEP**: weak equivalence principle
- **EEP**: Einstein equivalence principle
- **SEP**: strong equivalence principle
- **GC**, General Covariance
- **Mach Principle**: gravity coupled to matter

Derivation of Field Equation

From postulations

1. General covirance
2. Linear approximation should be compatible with Newton's thoery/Weak field and slow motion limit is Newton's thoery of gravity
3. In theory regarding the metric, no higher than second derivative is envolved and the terms of second derivative is linear.

The first point is for the invariance of frames/coordinates. The second point is for the success of Newtonian's theory on our earth.

Why do we believe the third point? The answer is that we don't have to. Here we propose it is because the simplicity of such quasilinear equations, i.e.,

$$F(\phi, \partial\phi)\partial^2\phi + G(\phi, \partial\phi) = 0$$

We have a bunch of theorems on this system, including its existance of solutions, Couchy problem, wave propagation etc.

We can use both 1&2 and 1&3 to derive Einstein's equation. That is 2 and 3 are identical when 1 is considered.

From Action

This is an application of stationary principal and Hilbert action or Hilbert action plus a Λ.

Lovelock's Theorem

The only possible second-order Euler-Lagrange expression obtainable in a four dimensional space from a scalar density of the form $L = L(g_{\mu\nu})$ is

$$E^{\mu\nu} = \alpha\sqrt{-g}[R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R] + \lambda\sqrt{-g}g^{\mu\nu}$$

Thus modification could be

- Metric tensor not a fundamental tensor
- Higher than second order derivatives of the metric in the field equations
- Not a four dimension space
- Not rank (2,0) tensor field equations, non-symmetry of field equations under exchange of indices, or divergence field equations
- non-locality

Birkhoff's Theorem

All spherically symmetric solutions of Einstein's equations in vacuum must be static and asymptotically flat, without Λ.

Actually, this can be extended to a Λ space only keeping the static result.

No-hair Theorems

The generic final state of gravitational collapse is a Kerr-Newman black hole, fully specified by its mass, angular momentum and charge

Also, "in the context of General Relativity with a cosmological constant all expanding universe solutions should evolve towards de Sitter space."[1] This is only valid in some situation.

Vacuum Solutions

The vacuum Einstein equation

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = 0,$$

which indicates that all constant metrics are solutions to vacuum Einstein equation.

Physically this doesn't make any sense, unless we impose that our universe is Minkowski like. From this point of view, vacuum Einstein equation is more general than our universe.

Perturbation Theory of General Relativity

Gauge freedom is the freedom of choosing a coordinate system. Fixing a gauge means choosing a particular coordinate system.

Gauge tranformation is Lie derivative along some arbitary vector here.

Line element

$$\tilde{g}_{00} =$$
$$-a^2(1+2AY)$$
$$\tilde{g}_{0j} =$$
$$-a^2 BY_j$$
$$\tilde{g}_{ij} =$$
$$a^2(\gamma_{ij} + 2H_L Y\gamma_{ij} + 2H_T Y_{ij})$$

Energy momentum tensor is

$$\tilde{T}^0{}_0 = -\rho(1+\delta Y)$$
$$\tilde{T}^0{}_j = (\rho+p)(v-B)Y$$
$$\tilde{T}^j{}_0 = -(\rho+p)vY^j$$

For a infinitesimal gauge transformation along some vector ($X = T\partial_t + L^i\partial_i$), gauge variables are

Symbol	Physics	Gauge Transformation	Note
$\tilde{A}$			

Through that we can find out gauge invariant variables.

[1] **R. M. Wald**. Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant. Phys. Rev. D, 28(8):2118–2120, Oct 1983.

What Frame Are We In

Synge once said, use space and time, and define them.

This post is aimed to make clear what frame are we in.

In general relativity, we often transform coordinates. Here is an example.

The general form of metric with spherical space component is

$$\mathrm{d}s^2 = -\gamma(r,t)c^2\mathrm{d}t^2 + \beta(r,t)c\mathrm{d}r\mathrm{d}t + \alpha(r,t)[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)] \tag{10.10}$$

With a transformation $\alpha(r,t)r^2 = r'^2$,

$$\mathrm{d}s^2 = -\gamma'(r',t)c^2\mathrm{d}t^2 + \beta'(r',t)c\mathrm{d}r\mathrm{d}t + \alpha(r,t)[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)]$$

Then compose the integral multiplier

$$c\mathrm{d}t' = \eta(r',t)[-\gamma'(r',t)c\mathrm{d}t + \frac{1}{2}\beta'(r',t)\mathrm{d}r']$$

And finally,

$$\mathrm{d}s^2 = -\eta^{-2}(r',t)\gamma'^{-1}(r',t)c^2\mathrm{d}t'^2 + [\alpha'(r',t) + \frac{\beta'^2(r',t)}{4r'}]\mathrm{d}r'^2 + r'^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)$$

In general

$$\mathrm{d}s^2 = -b(r,t)c^2\mathrm{d}t^2 + a(r,t)\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) \tag{10.11}$$

Then what? The two forms of metric demonstrate different properties. Take Birkhoff theorem as an example. The results could be very different startting from the form (10.10) and (10.11).

It is obviously very important to show what the coordinate transformation means and what frame are the observers in indicated by the coordinates.

Experiments

Eotvos Torsion Balance

How

- Inertial mass m_I
- Gravitational mass m_G

In Newtonian system, the acceleration of an object will be

$$\vec{a} \propto \frac{\vec{F}}{m_I}.$$

In a static and uniform gravitation field, the gravity force is

$$G = -g m_G \hat{r}$$

Thus the acceleration in this case should be

$$\vec{a} \propto -\hat{r} g \frac{m_G}{m_I}$$

When m_G/m_I is constant, the falling accerelation are the same for different objects with same mass. However, if m_G/m_I is not a constant, say $m_G \neq m_I$, different objects would fall at different acceleration.

Now if we put two ball with different mass on the Eotvos torsion balance, the balance would rotate and we can measure it.

Results

Detection of $R^k_{0l0} = (1/c^2)\partial^2\Phi/\partial x^k \partial x^l \sim 10^{-32}\mathrm{cm}^{-2}$.

Hughes-Drevershiy Experiment, etc

Anisotropy of gravitation/electromagnetism is not proved in our galaxy.

Radio Signal

Similar to Eddington and Dyson's bending light observation, radio signals serve as a more precise experiment to test Einstein's theory. And these experiments are against scalar tensor theories because scalar tensor theories give a smaller bending angle (1.66 second of arc less than the observations).

Summary Table

Tables constructed according to arXiv:1106.2476v3.

Test of fundamental principles

1. WEP 1. Eotvos torsion balance $\eta = (0.3 \pm 1.8) \times 10^{-13}$, More precise in space exp.[1a]_ *[1b]* *[1c]* 2. Gravitational redshift of light[2]
2. EEP: 1. Hughes-Drever Experiment: $n \leq 10^{-27}$, references *[3a]* *[3b]*

Test of GR:

1. Null geodesics test: 1. photon trajectory, spatial deflection: $\theta = (0.99992 \pm 0.00023) \times 1.75''$, where 1.75 is the theoretical value; Achieved through observing star position, etc[4] 2. Shapiro time-delay effect: $\Delta t = (1.00001 \pm 0.00001)\Delta t_{GR}$, references *[5a]* *[5b]*
2. Time like geodesics: 1. Anomalous perihelion precession: Just use the PPN formalism *[6a]* *[6b]* *[6c]* 2. Nordtvedt effect: $\eta = (-1.0 \pm 1.4) \times 10^{-*13}$, references *[7a]* *[7b]* 3. Spinning objects obiting *[8a]* *[8b]*
3. Small-range: 1. Potential probing *[9a]* *[9b]*
4. Radiation 1. Speed of gravitational waves 2. Polarity of gravitational radiation 3. Dynamics of source objects

Footnote

10.3.8 Spherical Solutions to Stars

Static

Static spacetime is defined as *[Schutz]*

1. all metric components are independent of time;
2. geometry is unchanged by time reversal.

[2] To be added

[4] **S. S. Shapiro, J. L. Davis, D. E. Lebach, and J. S. Gregory**. Measurement of the Solar Gravitational Deflection of Radio Waves using Geodetic Very-Long-Baseline Interferometry Data, 1979 1999. Physical Review Letters, 92(12):121101, Mar. 2004.

Equation of Motion

For spherical configuration, we have the line element of the form

$$ds^2 = -e^{2\Phi} dt^2 + e^{2\Lambda} dr^2 + r^2 d\Omega^2.$$

For future reference, we write down some metric elements.

$$g_{00} = -e^{2\Phi}$$
$$g_{rr} = e^{2\Lambda}.$$

Exterior Solutions

Interior Solutions

By defining

$$m(r) = \frac{1}{2} r(1 - e^{-\Lambda(r)}).$$

the 00 component of the Einstein equation is basically the mass function

$$m(r) = \int_0^r dr' 4\pi r'^2 \rho(r').$$

The other components leads to the famous TOV equation

$$\frac{dp}{dr} = -\frac{(\rho + p(r))(m(r) + 4\pi r^3 p(r))}{r(r - 2m(r))}, \tag{10.12}$$

where $m(r)$ is the mass function.

Newtonian Limit

Newtonian limit of the same equation is

$$\frac{dp}{dr} = -\frac{\rho(r) m(r)}{r^2}, \tag{10.13}$$

which can be derived by looking at a shell of mass at radius r.

To derive this we need to construct a shell of mass at radius r. At this shell gravity is balanced by pressue since we assumed a static star.

Interpretation of TOV equation

We can make sense of it and cast it into a more Newtonian form. We know that proper distance

$$dl = \sqrt{g_{rr}} dr = e^{-\Lambda} dr = (1 - 2m/r)^{-1} dr.$$

Then we rewrite (10.12) as

$$\frac{dp}{dl} = -\frac{(\rho + p(r))(m(r) + 4\pi r^3 p(r))}{r^2}. \tag{10.14}$$

The term $(\rho + p(r))$ corresponds to the contribution from the shell of mass $\rho(r)$ in Newtonian theory (10.13). The reason that we obtained a new contribution from pressure is that pressure is also the source of gravity. The second term is the total contribution of mass inside the shell. For a similar reason we pick up a pressure term in relativity.

The TOV equation with proper length (10.14) has some very interesting implications. Suppose we are at a coordinate radius r of a star. We measure the pressure gradient $\frac{dp}{dr}|_r$. Then we restrict ourselves on this shell at radius r. Then what we experience is basically a Newtonian-like $1/r^2$ law.

However

However, the second term in $(m(r) + 4\pi r^3 p(r))$ is kind of strange. It seems that we need no knowledge of pressure inside the shell to get the pressure gradient. It is not simply the total mass contribution from inside.

Something is not right.

References and Notes

10.3.9 Black Holes

Glossaries

1. ZAMO: zero angular-momentum observer

Observations of Black Holes

LIGO!

There are several ways of detecting black holes, directly or indirectly.

1. Gravitational effects: such as lensing, binary systems, gravitational waves.
2. Matter around it: orbit of stars around black holes, emission of accretion discs, final stage of Hawking radition.

Several astronomical tips:

1. As for accretions, pulsars are steady because of the magnetic field they are holding. Black holes do not hold magnetic field like that so it can not be a steady pulsation.
2. There is a star orbiting the center of our galaxy at an orbit of 120AU, which helps us determing the mass of the center object.
3. First generatioin of stars (population III stars) might form supermassive black holes, and might still be around since Hawking radiation of massive black holes is low.
4. Black hole collisions are studied using computers numerically. Such simulations have intrinsic instability due to the fact that GR has coordinate freedom and singularities.
5. Simulations found kicks of black holes. Black hole mergers starting with asymmetries will develop asymmetric emission of gravitational waves thus kicking the system in the opposite direction.

Schwarzschild Metric

The line element for Schwarzschild metric is

$$\begin{aligned} ds^2 &= -(1-\frac{2M}{r})dt^2 + (1-\frac{2M}{r})^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \\ &= g_{tt}dt^2 + g_{rr}dr^2 + g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2. \end{aligned}$$

The inverse of metric elements $g^{\alpha\beta}$ are easily obtained since the metric is diagonal in this basis.

Kerr Metric

To begin with, Kerr spacetime around a Kerr black hole of mass M, spin angular momentum J, is described as the line element

$$ds^2 = -\frac{\Delta - a^2 \sin^2\theta}{\rho^2} dt^2 - 2a\frac{2Mr\sin^2\theta}{\rho^2} dt d\phi + \frac{(r^2+a^2)^2 - a^2\Delta\sin^2\theta}{\rho^2}\sin^2\theta d\phi^2 + \frac{\rho^2}{\Delta^2} dr^2 + \rho^2 d\theta^2,$$

where

$$\begin{aligned} a &= J/M \\ \Delta &= r^2 - 2Mr + a^2 \\ \rho^2 &= r^2 + a^2\cos^2\theta. \end{aligned}$$

The Kerr metric has very nice symmetries.

1. Reflection symmetry with respect to $\theta = \pi/2$;
2. Axial symmetry around the z axis so that the angular momentum $L_\phi = p_\phi$ is conserved;
3. Stationary metric so that energy of test particles is conserved, $E = -p_0$ is conserved.

Frame Dragging

We would be very interested in how the rotation of black holes drag the spacetime. Due to the symmetries, the equatorial plane is easier to think about. So we set $\theta = \pi/2$. For frame dragging, we have a guy riding a rocket so that $p_r = 0$. The frame dragging is best described by a observor that is staying stationary with frame, that is a zero angular momentum observor (ZAMO), that $p_\phi = 0$. Frame dragging angular velocity is

$$\omega = \frac{d\phi}{dt} = \frac{d\phi/d\tau}{dt/d\tau}.$$

It can be derived that the frame dragging angular velocity at radius r is

$$\omega = -\frac{g_{t\phi}}{g_{\phi\phi}} = \frac{2Mra}{(r^2+a^2)^2 - a^2\Delta\sin^2\theta}.$$

Mathematica (11) Code for the Plot

```
In[1]:= omega[r_,a_:0.2,mass_:1,theta_:0]:=Module[{deltaM},
deltaM=r^2-2mass r+a^2;
2mass r/( (r^2+a^2)^2-a^2 deltaM Sin[theta]^2 )
]
In[2]:= Solve[D[omega[r,a],r]==0,r]
Out[2]= {{r->-(a/Sqrt[3])},{r->a/Sqrt[3]}}
In[3]:= Manipulate[
Plot[omega[r,a],{r,0,5},
Frame->True,FrameLabel->{"r","\[Omega]"},
ImageSize->Large,PlotRange->Full,PlotStyle->Black,
GridLines->{{ {a/Sqrt[3],
Directive[Gray,Thick]},{1+Sqrt[1-a^2],Directive[Red,Thick]}},{1}},
Epilog->{Inset[Style["Horizon",13,Red],{(1+Sqrt[1-a^2]),0},{0,0}],
```

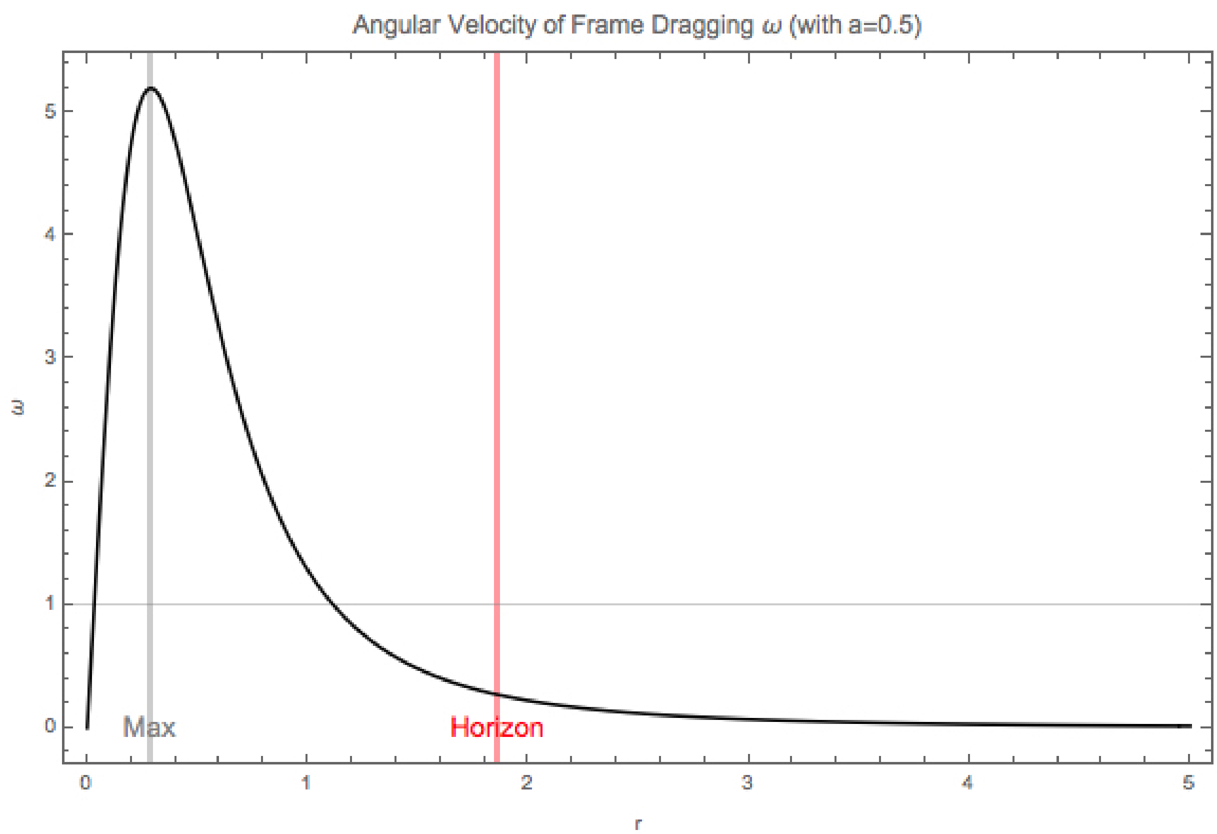


Fig. 10.10: Angular velocity of the frame dragging for a Kerr black hole. Everything is scaled by the mass of black hole. It drops as $1/r^3$ for large r.

```
Inset[Style["Max",13,Gray],{a/Sqrt[3],0},{0,0}]},
PlotLabel->"Angular Velocity of Frame Dragging \[Omega] (with a="<>ToString@a<>")"],
{{a,0.5,"Spin Angular Momentum of Black Hole"},0,1}
]
```

Ergospheres, Horizons

In Schwarzschild black holes, the surface that $g_{tt} = 0$ and $g_{rr} \to \infty$ are the same surface that is defined as the horizon. However, $g_{tt} = 0$ gives us the ergospheres and $g_{rr} \to \infty$ gives us the horizons.

Why?

The ergospheres are the regions that even light can not travel in the counter-rotation direction. That being said, p^ϕ can only be positive for light. To prove this, we set $ds^2 = 0$ and neglect p_r,

$$g_{tt}dt^2 + g_{\phi\phi}d\phi^2 + 2g_{t\phi}dtd\phi = 0,$$

which shows that,

$$p^\phi = \frac{-g_{t\phi} \pm \sqrt{g_{t\phi} - g_{tt}g_{\phi\phi}}}{g_{\phi\phi}}.$$

We can show that p^ϕ can only be negative if $g_{tt} > 0$, while it can be negative or positive if $g_{tt} < 0$. What we found is that for $g_{tt} > 0$, we are entering a region where even light can not travel against the direction of the rotation. This is why we define the condition $g_{tt} = 0$ to be the surface of the ergospheres.

As for $g_{rr} \to \infty$, it is the condition for p^r being always negative, which means we are always travelling inward. It proven by similar techniques.

We have two ergospheres and horizons! Solving $g_{tt} = 0$ gives us

$$r_{e,\pm} = M \pm \sqrt{M^2 - a^2\cos^2\theta},$$

which defines the two surfaces of ergospheres.

Meanwhile, $g_{rr} \to \infty$ indicates that $\Delta = 0$, which proves to be

$$r_{h,\pm} = M \pm \sqrt{M^2 - a^2},$$

which shows us the two horizons.

Mathematica (11) Code

```
gtt[r_,a_,mass_:1,theta_:Pi/2]:=Module[{deltaM,rhosquareM},
deltaM=r^2-2mass r+a^2;
rhosquareM=r^2+a^2Cos[theta]^2;
-(deltaM-a^2Sin[theta]^2)/rhosquareM
]
grr[r_,a_,mass_:1,theta_:Pi/2]:=Module[{deltaM,rhosquareM},
deltaM=r^2-2mass r+a^2;
rhosquareM=r^2+a^2Cos[theta]^2;
rhosquareM/deltaM
]
```

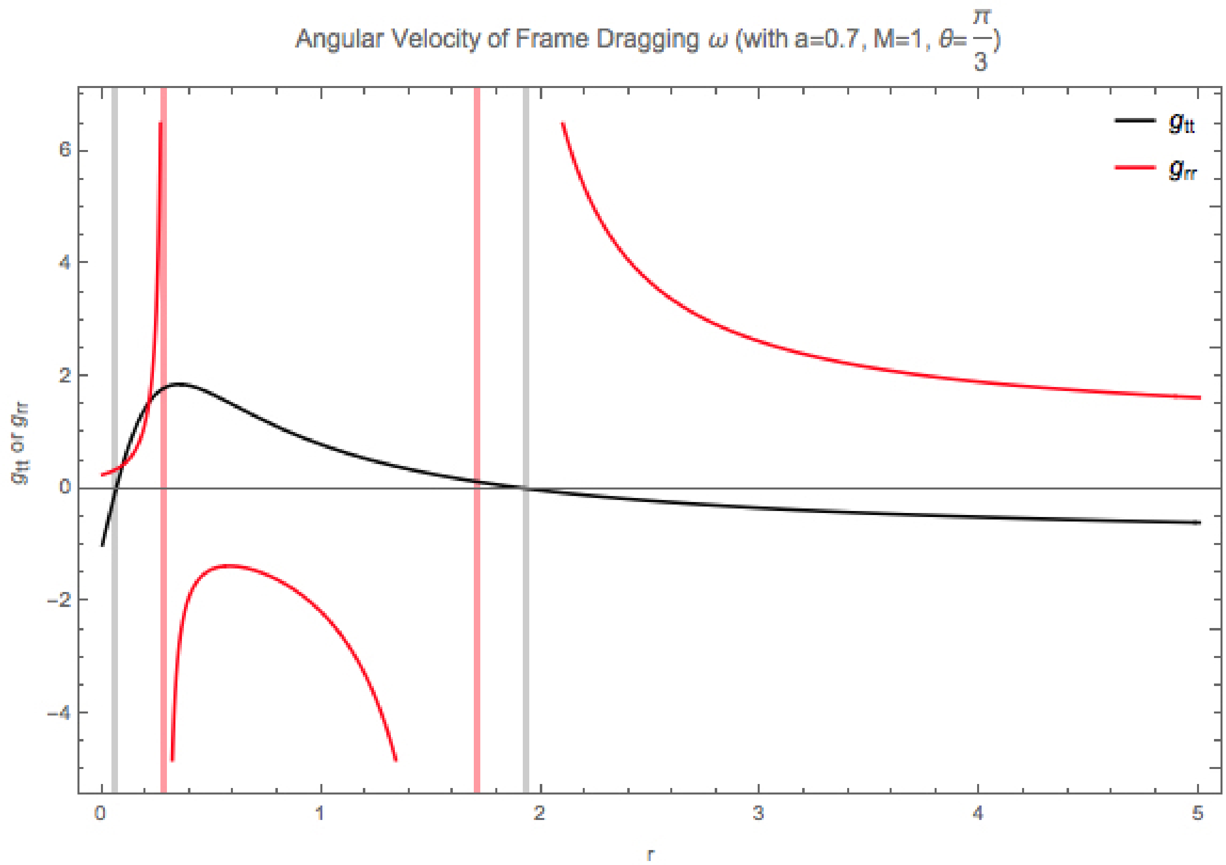


Fig. 10.11: g_{tt} and g_{rr} as function of coordinate r.

```
Manipulate[
Plot[{gtt[r,a,mass,theta],grr[r,a,mass,theta]},{r,0,5},Frame->True,FrameLabel->{"r",
↪"Subscript[g, tt] or Subscript[g, rr]"},ImageSize->Large,PlotRange->Automatic,
↪PlotStyle->{Black,Red},GridLines->{{ {mass+Sqrt[mass^2-a^2],Directive[Red,Thick]},
↪{mass+Sqrt[mass^2-a^2Cos[theta]^2],Directive[Gray,Thick]},{mass-Sqrt[mass^2-a^2],
↪Directive[Red,Thick]},{mass-Sqrt[mass^2-a^2Cos[theta]^2],Directive[Gray,Thick]}},
↪None},PlotLabel->"Angular Velocity of Frame Dragging \[Omega] (with a="<>ToString@a
↪<>", M="<>ToString@mass<>", \[Theta]="<>ToString@TraditionalForm@theta<>")",
↪PlotLegends->Placed[{"Subscript[g, tt]","Subscript[g, rr]"},{Right,Top}]],
{{a,0.7,"Spin Angular Momentum of Black Hole"},0,1},{{mass,1,"Mass of Black Hole"},0.
↪1,10},{{theta,Pi/3,"\[Theta]"},0,Pi}
]
```

In fact we can prove that

1. Within region $r > r_{e,+}$, $g_{tt} < 0$, $g_{rr} > 0$;
2. Within region $r_{h,+} < r < r_{e,+}$, $g_{tt} > 0$, $g_{rr} > 0$;
3. Within region $r_{h,-} < r < r_{h,+}$, $g_{tt} < 0$, $g_{rr} < 0$;
4. Within region $r_{e,-} < r < r_{h,-}$, $g_{tt} < 0$, $g_{rr} > 0$;
5. Within region $r < r_{e,-}$, $g_{tt} < 0$, $g_{rr} > 0$.

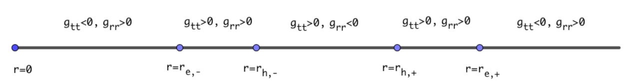


Fig. 10.12: Regions of Kerr black holes. $r_{e,\pm}$ are the two surfaces of ergospheres, $r_{h,\pm}$ are the two horizons, as calculated previously.

What are the significances of the surfaces?

For the outer horizon and outer ergosphere, their properties are discussed. What are the properties of the inner surfaces?

Photons Travelling on Equatorial Plane

The elements of the metric $g_{\alpha\beta}$ as well as $g^{\alpha\beta}$ are frequently used. It is essential to find them, which involves some matrix inversion.

Inverse of Block Diagonal Matrix

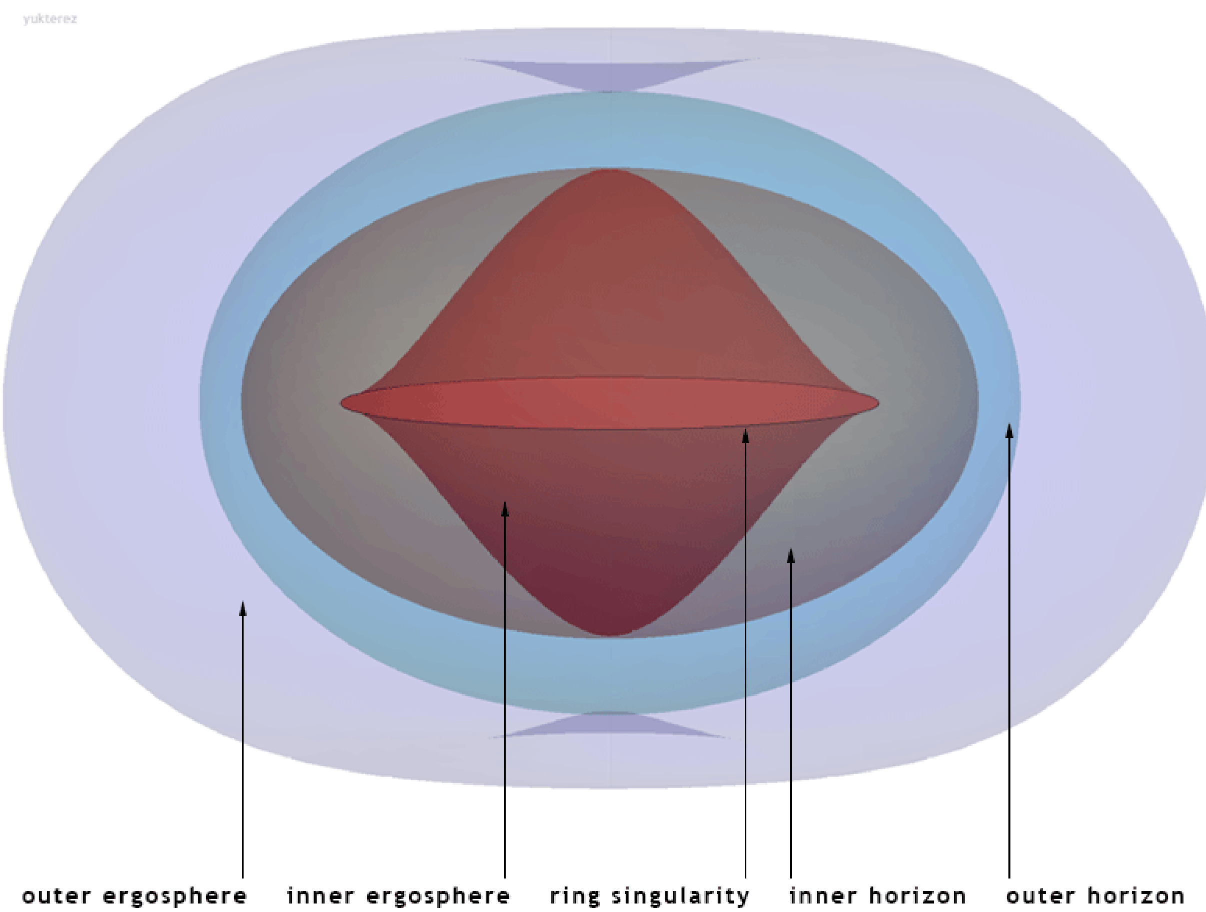


Fig. 10.13: A Kerr black hole is nicely visualized by Simon Tyran, whose work is licensed with CC BY-SA.

For a given matrix

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

the inverse of it A^{-1} is

$$A^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{pmatrix}.$$

This result works for arbitrary dimensions.

Inverse of 2 by 2 Matrix

For a 2 by 2 matrix

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

the inverse is

$$B^{-1} = \frac{1}{D} \begin{pmatrix} B_{22} & -B_{21} \\ -B_{12} & B_{11} \end{pmatrix}.$$

Penrose Process

Suppose we have a particle falling inside a black hole, starting with 0 energy at infty. It falls through the ergosphere, and decays into two particles, A and B. Particle A obtains a negative energy, meanwhile having negative angular momentum, so that it stays in ergosphere or falls through the horizon. The other particle obtains positive energy, and managed to escape. Energy conservation tells us that the escaped particle will have energy at infty that is larger than the initial energy 0.

This though experiment relies on the effective potential $V(r)$ of the ergosphere. For positive angular momentum, we always fall through the

10.3.10 Fields and Particles

Energy-Momentum Tensor for Particles

$$S_p \equiv -mc \int\int \mathrm{d}s \mathrm{d}\tau \sqrt{-\dot{x}^\mu g_{\mu\nu} \dot{x}^\nu} \delta^4(x^\mu - x^\mu(s)),$$

in which $x^\mu(s)$ is the trajectory of the particle. Then the energy density ρ corresponds to $m\delta^4(x^\mu - x^\mu(s))$.

The Largrange density

$$\mathcal{L} = -\int \mathrm{d}s mc \sqrt{-\dot{x}^\mu g_{\mu\nu} \dot{x}^\nu} \delta^4(x^\mu - x^\mu(s))$$

Energy-momentum density is $\mathcal{T}^{\mu\nu} = \sqrt{-g} T^{\mu\nu}$ is

$$\mathcal{T}^{\mu\nu} = -2 \frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}$$

Finally,

$$\begin{aligned} \mathcal{T}^{\mu\nu} &= \int \mathrm{d}s \frac{mc\dot{x}^\mu \dot{x}^\nu}{\sqrt{-\dot{x}^\mu g_{\mu\nu} \dot{x}^\nu}} \delta(t - t(s)) \delta^3(\vec{x} - \vec{x}(t)) \\ &= m\dot{x}^\mu \dot{x}^\nu \frac{\mathrm{d}s}{\mathrm{d}t} \delta^3(\vec{x} - \vec{x}(s(t))) \end{aligned}$$

10.3.11 Theorems

Killing Vector Related

ξ^a is Killing vector field, T^a is the tangent vector of geodesic line. Then $T^a \nabla_a (T^b \xi_b) = 0$, that is $T^b \xi_b$ is a constant on geodesics.

10.3.12 Specific Topics

Redshift

In geometrical optics limit, the angular frequency ω of a photon with a 4-vector K^a, measured by a observer with a 4-velocity Z^a, is $\omega = -K_a Z^a$.

Stationary vs Static

Stationay

"A stationary spacetime admits a timelike Killing vector field. That a stationary spacetime is one in which you can find a family of observers who observe no changes in the gravitational field (or sources such as matter or electromagnetic fields) over time."

When we say a field is stationary, we only mean the field is time-independent.

Static

"A static spacetime is a stationary spacetime in which the timelike Killing vector field has vanishing vorticity, or equivalently (by the Frobenius theorem) is hypersurface orthogonal. A static spacetime is one which admits a slicing into spacelike hypersurfaces which are everywhere orthogonal to the world lines of our 'bored observers'"

When we say a field is static, the field is both time-independent and symmetric in a time reversal process.

CHAPTER 11

Astrophysics

11.1 Astrophysics Basics

11.1.1 Numbers

One of the most useful concepts in astrophysics is that things should be scaled using a scale that we can understand.

Physical Facts

Ionization energy for atoms

1. H: 13.6eV
2. He: 24.7eV
3. He^{++}: 54.4eV

Solar System

Sun

Some numbers:

1. Mass of Sun: $M_\odot = 1.99 \times 10^{30}\mathrm{kg}$;
2. Radius of Sun: $6.9 \times 10^5\mathrm{km}$;
3. Average Density: $1.4\mathrm{g/cm}^3$;
4. Surface Gravity: $27g_E$;
5. Core Density: $150\mathrm{g/cm}^3$;
6. Core Temperature: $1.5 \times 10^7\mathrm{K}$ which corresponds to energy of $1\mathrm{keV}$.

The sun is composed of

1. X = H mass fraction of 0.70; 1. Y = He mass fraction of 0.26; 1. X = Li mass fraction of 0.04.

11.1.2 Orbital Dynamics

Kepler's law can be easily derived from Newton's dynamics.

The first law is about the

11.1.3 Wave Bands

A table of wavebands in astronomy is something like this.

11.1.4 Doppler Shift

Refer to doppler-shift.

11.1.5 Photon Production

There are two important processes that produces photons in astrophysical environment, namely thermal Bremsstrahlung and synchrotron radiation.

Bremsstrahlung Radiation

11.1.6 Compton Scattering

The quantities we are interested in are cross section and radiated power, both of which are Lorentz invariant.

Thomson Scattering

Total cross section of Thomson scattering can be obtained using only classical electrodynamics. The physics behind it is that the electric field excerts force on an electron then the electron emits photons to all possible direction as it oscillates.

The incoming power per unit area is

$$P_{inc} = c u_{rad},$$

as $u_{rad} = \frac{E^2}{4\pi}$ is the energy density of EM field.

The outgoing or scattered wave power (total) is

$$P_{scatt} = \frac{2}{3}\frac{e^4}{m_e^2 c^3} E^2.$$

The total cross section is ratio of the two quantity, i.e.,

$$\sigma_T \equiv \frac{P_{scatt}}{P_{inc}} = \frac{8\pi e^4}{3 m_e^2 c^4} = \frac{8\pi(\alpha\hbar c)^2}{3 m_e^2 c^4},$$

in which the fine structure constant is defined as $\alpha = \frac{e^2}{\hbar c}$.

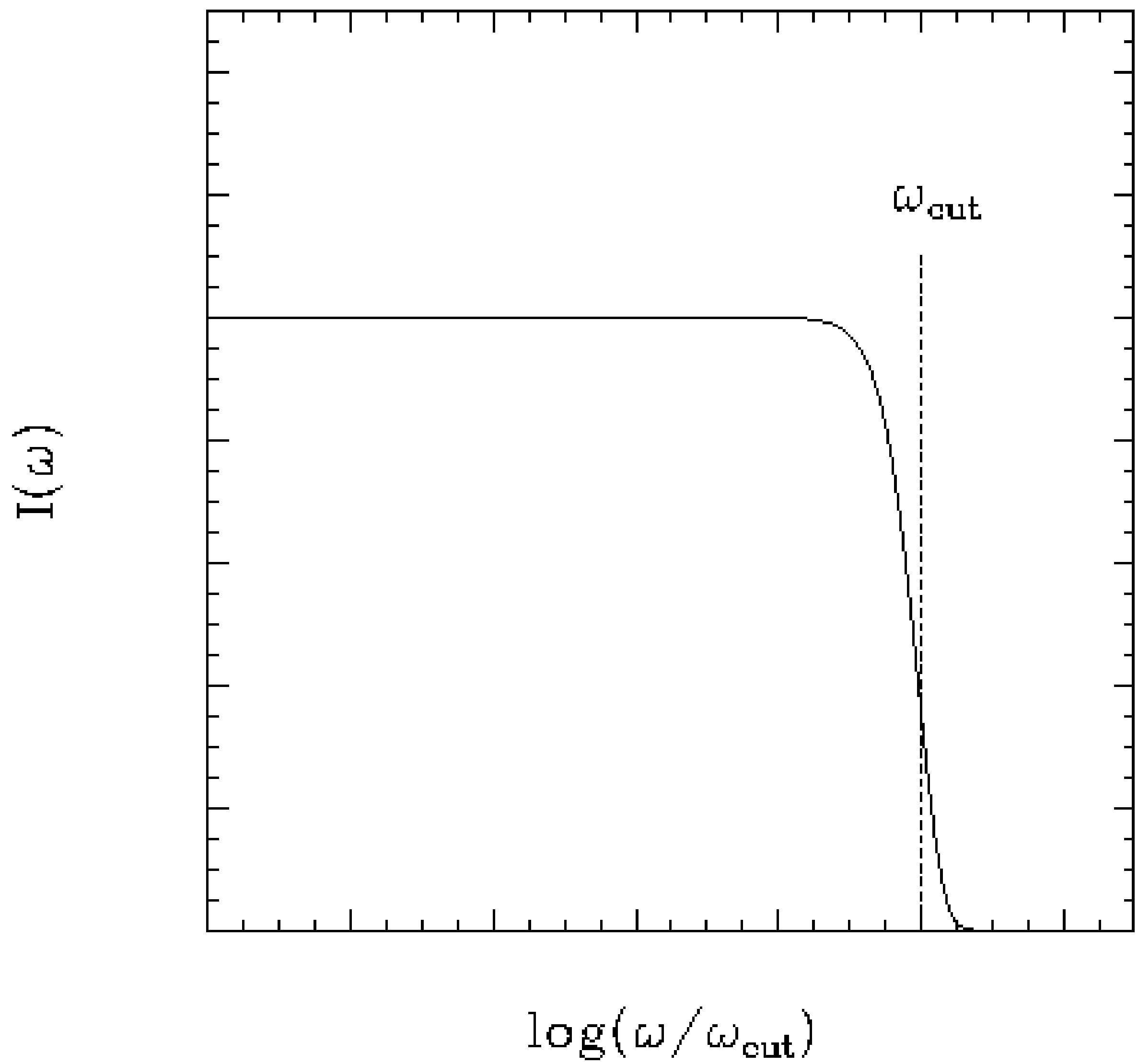


Fig. 11.1: This is the spectrum of frequency-dependent emissivity of the process which happens when a flux of non-relativistic regime with thermal distributon of temperature T is shot into a plasma of ions or protons.

Compton Scattering

The full quantum electrodynamics result is called Klein-Nishima formula, which describes the total cross section of coliding photon and electron,

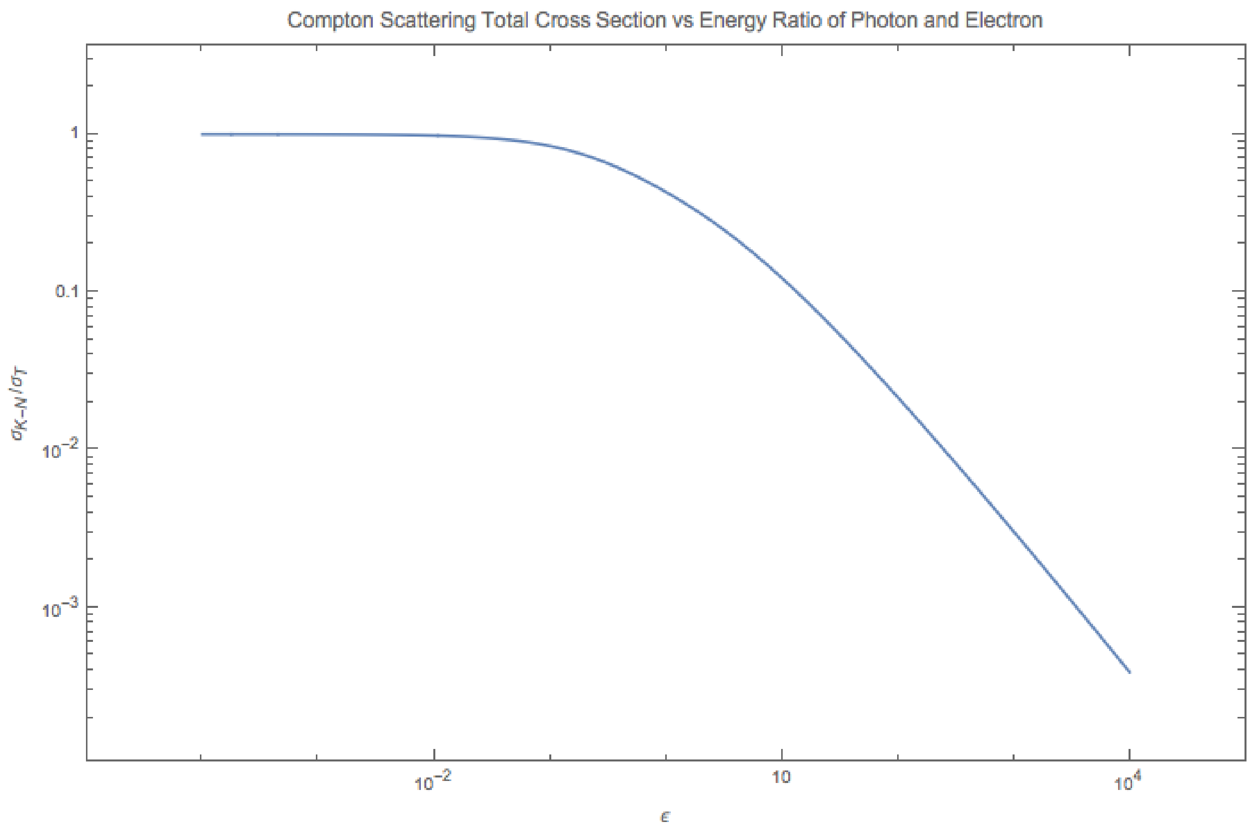
$$\sigma_{K-N} = \frac{\pi e^4}{m_e^2 c^4}\frac{1}{\epsilon}\left[\left(1-\frac{2(\epsilon+1)}{\epsilon^2}\right)\ln(2\epsilon+1)+\frac{1}{2}+\frac{4}{\epsilon}-\frac{1}{2(2\epsilon+1)^2}\right],$$

where $\epsilon = \frac{E}{m_e c^2}$.

In the limit that energy of electron is much larger than photons, we have ϵ is much smaller than 1, we would come back to the Thomson limit, which is true for our equation,

$$\begin{aligned}\sigma_{K-N} &= \sigma_T \frac{3\left(\left(1-\frac{2(x+1)}{x^2}\right)\log(2x+1)+\frac{4}{x}-\frac{1}{2(2x+1)^2}+\frac{1}{2}\right)}{8x} \\ &= \sigma_T(1-2\epsilon+O(x^2)).\end{aligned}$$

To have more understanding on this formula, I ploted σ_{K-N} in terms of σ_T as the energy scale ϵ changes.



11.1.7 Asteroseismology

The stars do shake, from inside out.

Long period variable such as Cepheids pulsate in the luminosity. This is because of the radial oscillation mode with a approximate period of

$$P_{dynamical} \approx \left(\frac{R^3}{GM}\right)^{1/2} \approx (G\bar{\rho})^{-1/2},$$

in which we have the radius of the star as R, mass of the star as M and mean density $\bar{\rho}$. **The good thing of this oscillation immediately shows us the mean density of the star, even without any furthure inspectation.**

There are double mode Cepheids, whose modes provides information about mass and radius.

Our sun, up to now we have identified thousands of individual modes. And more modes as many as 10^6 modes can be determined accurately.[heliooosc]_

Papers, Researches and More

1. Double mode Cepheids, J. Otzen Petersen, 1973, 1974, 1978.
2. An introduction of seismology applied to stars. http://ap.smu.ca/~guenther/seismology/seismology.html

11.1.8 Relativistic Beaming Effect

11.1.9 Refs & Notes

11.2 Stars

Question

What Are The Typical Masses of Stars?

A star is typically formed from a cloud of gas which is contracted and reaches a stable state. Suppose we can find the relation between the temperature and mass of the star, we could use our knowledge of nuclear physics to determine the temperature scale thus the mass of the star could be calculated.

Typical Pressure

A star starts from a gas cloud which is divided into inner part and outer shell in order to estimate the typical pressure caused by gravitational force,

$$\begin{aligned}P_G &\sim \frac{F_G}{A} \\ &\sim \frac{G\frac{M}{2}\frac{M}{2}}{R^2 4\pi r^2} \\ &\sim \frac{2^{2/3}GM^2}{16\pi R^4},\end{aligned}$$

where r is the radius of the inner radius which has a mass of $\frac{M}{2}$, R is the radius of the whole gas cloud which has total mass M, and the outer shell of the gas also has mass $M - \frac{M}{2} = \frac{M}{2}$, A is the area on the contact of the inner shpere and the outer shell.

The cloud collapses with these gravitational potential energy goes into Fermi energy and kinetic energy of the molecules. It is convinient to calculate the gravitatioinal potential energy per proton (from Hydrogen),

$$\epsilon_G = \frac{Gm_p^2 N^2}{RN},$$

in which m_p is the mass of a proton while N is the number of protons. Define number density

$$n = \frac{N}{\frac{4}{3}\pi R^3},$$

which helps rewrite the radius

$$R \sim \left(\frac{N}{n}\right)^{1/3}$$

so that the gravitational potential energy becomes

$$\epsilon_G \sim n^{1/3}.$$

Switching to a microscopic view of the particle we could identify the total potential energy of the particle, which is composed of thermal energy kT and Fermi energy $\frac{\hbar^2}{2m_e}(3\pi^2 n)^{2/3}$. The conservation of energy is, roughly,

$$kT + \frac{\hbar^2}{m_e} n^{2/3} \sim G m_p^2 N^{2/3} n^{1/3}.$$

Solve the thermal energy

$$kT \sim G m_p^2 N^{2/3} n^{1/3} - \frac{\hbar^2}{m_e} n^{2/3},$$

which has a maximum value located at n_*,

$$\frac{\hbar^2}{m_e} n_*^{-1/3} - G m_p^2 N^{2/3} n_*^{-2/3} \sim 0.$$

We find the critical value of number density n_*

$$\begin{aligned} n_* &\sim \left(\frac{G m_p^2 N^{2/3} m_e}{\hbar^2}\right)^3 \\ &\sim \frac{\alpha_G}{\lambda_e} N^{2/3}, \end{aligned}$$

where α_G is a dimension quantity related to the gravity between two protons while λ_e is the Compton wavelength of electron,

$$\begin{aligned} \alpha_G &= \frac{G m_p^2}{\hbar c} \\ \lambda_e &= \frac{\hbar}{m_e c}. \end{aligned}$$

Thus the maximum temperatur is

$$kT_m \sim \alpha_G N^{4/3} m_e c^2.$$

As long as we could find a constraint on temperature, the mass of the star could possibily be determined with this relateion. The constraint on temperature comes from the nuclear reaction since the temperature should be able to sustain some kind of nuclear fusion.

The simplest nuclear fusion is Hydrogen to Deuterium which requires the wave packet of two proton to overlap. This overlap distance is related to a Couloumb potential energy.

The overlap distance is de Broglie wavelength of the protons,

$$\lambda_b \sim \frac{\hbar}{m_p c},$$

in relativistic limit if the temprature is comparable or higher than the mass of a proton.

The coresponding Coulomb potential using Gaussian unit is

$$\epsilon_n \sim \frac{e^2}{\lambda_b} \sim \alpha^2 m_p c^2,$$

where α is fine structure constant. In Gaussian unit

$$\alpha = \frac{e^2}{\hbar c}.$$

More Accuracy

In fact we could put in a factor $\eta \sim 0.1$ for the estimation of energy

$$\epsilon_n = \eta \alpha^2 m_p c^2.$$

This step is not require for an estimation.

For a star to exist, we require the temperature is higher than this nuclear ignition energy.

$$kT > \epsilon_n.$$

To estimate the mass of the star we find the critical number of protons in this problem by using the condition that $kT_m \sim \epsilon_n$, which results in

$$N_* \sim \left(\frac{m_p}{m_e}\right)\left(\frac{\alpha}{\alpha_G}\right)^{3/2}.$$

This is about the value 10^{56} and the corresponding mass is about $10^{29}\mathrm{kg}$ or $0.1 M_\odot$. **We have obtained the lower limit of a star.**

Question

How bright is the Sun?

The Sun has layers callled

1. Corona which goes into the interplanetary medium;
2. Chromosphere;
3. Photosphere which we see in visible light and contains absorption line (Fraunhof lines).

Brightness in a sense can be represented using power. Here we are going to calculate the power of the Sun using the knowledge we know at Earth.

In the Earth atmosphere, the light is partially obsorbed, which requires knowledge of obsorption which is described by

$$I(\theta, z) = I_0 e^{-kz/cos\theta},$$

according to which observation on the surface is enough to find out the actually energy arrived at Earth per unit area assuming perpendicular incident.

The result is $S \equiv I(0,0) = I_0 = 1360\mathrm{W/m^2}$. The fact is only 40 percent of the light reaches the surface and due to the O zone almost no light with wavelength under 1 is coming through.

The total power of the Sun is

$$L_\odot = 4\pi R_\odot^2 S = 3.8 \times 10^{26}\mathrm{W}.$$

Assuming the Sun emmits black body radiation, the corresponding temperature is $T_\odot = 5760\mathrm{K}$.

Stefan-Boltzmann Law

Energy radiated per unit area per unit time of black body radiation is

$$j = \sigma_{SB} T^4,$$

where $\sigma_{SB} = \frac{2\pi^5 k^4}{15 c^2 h^3} = 5.67 \times 10^{-8} \mathrm{W m^{-2} K^{-4}}$.

As a reference, 100K black body radiation has an energy flux of $5.67\mathrm{W/m^2}$.

Question

What is the composition of the Sun?

Absorption line from photosphere is the tool to identify the elements and abundance in the sun.

1. Identify the absorption lines;
2. Physical conditions related to the source of these lines, such as electron and ion temperature, pressure and density;
3. Atomic physics that is responsible for the strenght of the lines.

We can find out the density for a state of the element which is called "column" density, *N h* where *N* is the number density of the state and *h* is the depth the light is going through.

Question

What is the age of the solar system?

Before humanbeings actually realized the radioactivity, people did estimations of the Sun by assuming the energy source is GRAVITY which is so wrong. The way they did it is to calculate the total potential energy contained in a process that a cloud collapses to the size of the Sun. These energy goes to light so we know the upper limit of the age of our Sun since we know the power of it.

A spherical ball has potential energy

$$V = \frac{kGM^2}{R},$$

where $k = \frac{3}{5}$ for a uniform ball and $k = 1.74$ for our sun.

Assuming that the Sun is emmitting light constantly, which is $S = 1360\mathrm{W/m^2}$ per unit area at surface, we have

$$4\pi R_\odot^2 S = -\frac{dV}{dt}$$

we find $t < 55\mathrm{Myr}$.

Nowadays we find out the age of the solar system by looking into the isotop abundance.

We have a lot of decays that can be used for dating.

The idea behind this method is that since some of the isotopes decay the abundance will decrease and the abundance of the daughter will increase.

Decays is

$$\frac{dN(t)}{dt} = -\lambda N(t),$$

which has a solution

$$N(t) = N(0)e^{-\lambda t}.$$

We know the half life τ which determines λ through $\tau = \ln 2/\lambda$.

We don't know $N(0)$ so we need to use the abunddance of daughter,

$$D^*(t) = N(0) - N(t) = N_0(1 - e^{-\lambda t}),$$

then we have

$$\frac{D^*(t)}{N(t)} = e^{\lambda t} - 1.$$

As long as we know the initial daughter abundance we can find out the age.

11.2.1 Refs and Notes

11.3 Supernova

To-do

- Observational facts;
- Theory of explosion;
- Roles in cosmoligical scale;
- Roles in galatical scale;
- Roles in planetary system;
- Effects on life.

11.3.1 Refs & Notes

CHAPTER 12

Cosmology

12.1 Two Parameters

Why is Cosmology Dedicated to Finding Two Parameters Before 90's

Basically, the cosmology before the 90's have only two tasks. The first one is to find out the Hubble constant, while the second one is looking for the deceleration parameter.

We don't rush to define what Hubble constant and deceleration parameter are, but have a look at what observations do at that time.

12.1.1 Observations

Astronomers are really good at measuring distances. They have infinite tricky ways to find out some distance.

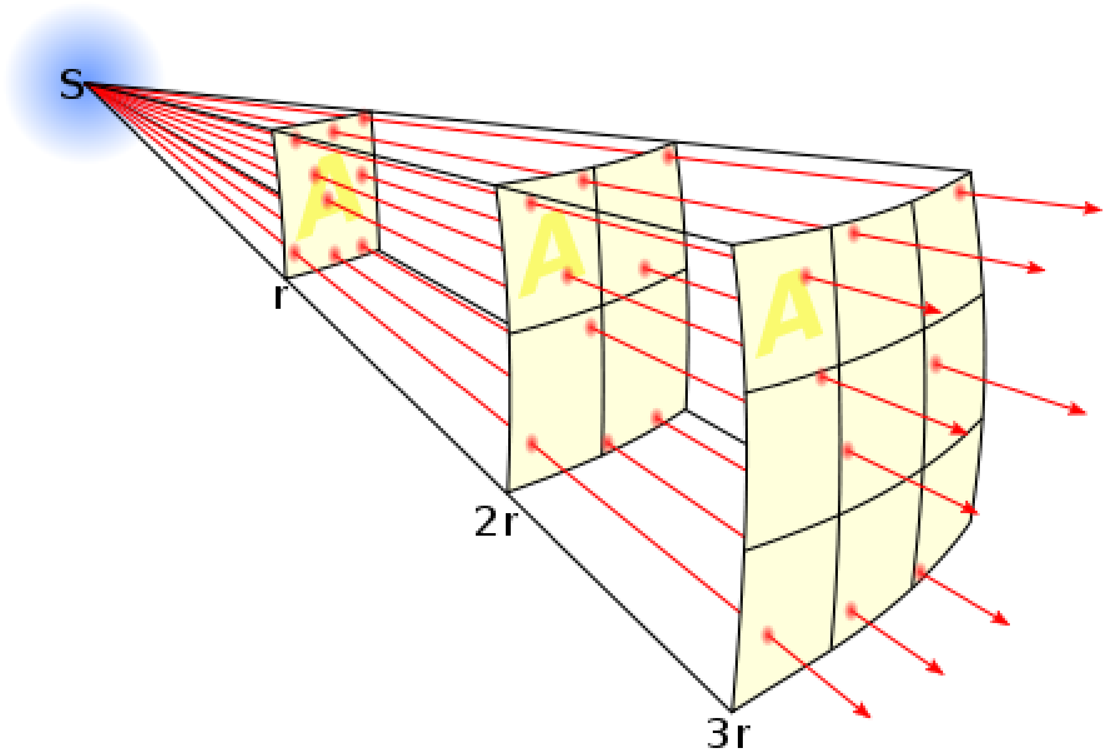
Luminosity Distance

Luminosity Distance from Observation

We can find out how bright a star is by observation. One way to represent the brightness is to use the energy crossed per unit area per unit time at the observer, because this is what our eyes do.

This quantity is related to how much energy was emitted at the star, how far we are from the star. The more energy the star emitted, the brighter it look like. The nearer the star is, the brighter it is. Just like what we feel like with a candle.

This schematic picture shows that energy spread out on a surface because the total energy is conserved. Isotropic energy flux through the same solid angle at different radius must be the same.



Through a very simple calculation, it is as simple as

$$L_0 = \frac{L}{4\pi r^2}.$$

We are dealing with Cosmology now. The space-time manifold should be a great concern. The luminosity turns out to be

$$L_0 = \frac{L_{\text{abs}}}{4\pi d^2} \frac{1}{1+z} \frac{1}{1+z}.$$

Here d is the physical distance between the star and the observer. L is the absolute luminosity of the star, which stands for the power of the star. z is the redshift of the star.

The first $\frac{1}{1+z}$ term comes from the fact that the energy of each photon decrease due to expansion of the universe, while the second is the result that the rate of photons arrived at the observer is less.

We are happy to define

$$d_L = d(1+z),$$

then the luminosity becomes simpler,

$$L_0 = \frac{L_{\text{abs}}}{4\pi d_L^2}.$$

Now we come back to have a look at this luminosity.

- We can measure how much energy is passing through a unit area at a unit time, which means **we can determine this luminosity directly from observations**.
- We can **predict the absolute luminosity** from a star evolution model.

- The $d_L = d(1+z)$ is only valid for a flat universe, with curvature term $K = 0$ in Friedmann equation.

Then we can find out this so called luminosity distance

$$d_L = \frac{L_{\text{abs}}}{4\pi L_0}$$

from some data.

Luminosity Distance from Theory

We don't just do the observation for the luminosity distance itself. We observe to test theories.

What is this distance in theory?

$$d_L = d(1+z)$$

Wait, didn't we just mention that this is only valid for a flat universe? So we just do some extension.

$$d_L = R(d)(1+z)$$

where R(d) is a function of d and can be determined through geometry,

- Spherical: $4\pi \sin^2 d$,
- Flat: $4\pi d$,
- Hyperbolic: $4\pi \sinh^2 d$.

Nearby Objects

For nearby objects, we can always use flat geometry and use Taylor expansion at current time for a(t).

Luminosity distance is

$$d_L = d(1+z) = ra(t)(1+z),$$

where r is the comoving distance and a(t) is the scale factor at time t.

We know

$$r = \int_t^{t_0} \frac{1}{a(t')} \mathrm{d}t'.$$

So we are happy to use Taylor expansion around t_0 for $a(t)$, and keep only up to the second order of time. And do some substitution with

$$H_0 = \dot{a}(t_0)/a(t_0)$$
$$q_0 = \ddot{a}(t_0)/a(t_0)$$

We then do the same thing on redshift

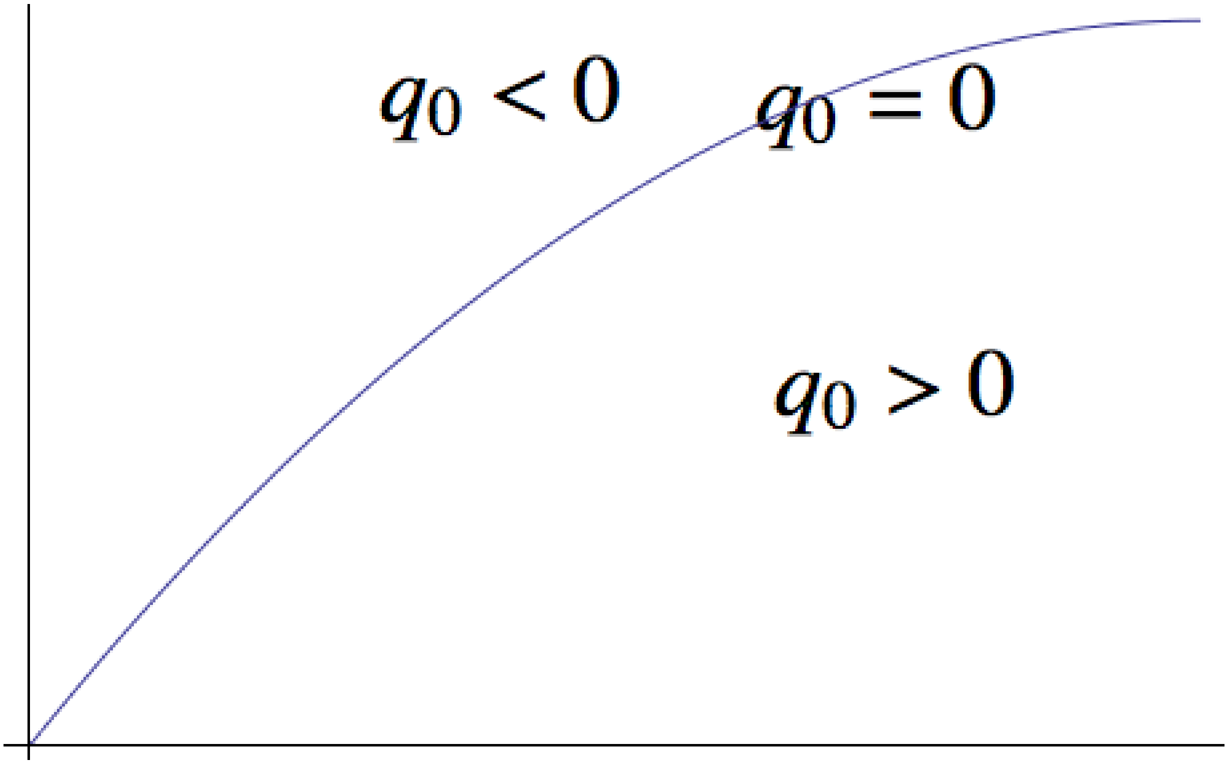
$$z = a(t_0)/a(t) - 1.$$

Finally, we can find out the relation $r(z)$, which leads us to the result we need, $d_L(z) = H_0^{-1}(z - \frac{1}{2}(1+q_0)z^2)$.

- For very near objects (not as near as our sun of course),

$$d_L = H_0^{-1} z.$$

This is a model independent observation and derivation. We can draw a line to represent the case when deceleration parameter is zero, lines higher than this stands for a accelerating universe while lower region show a decelerating universe.



We can show that for a vacuum energy dominated universe, the line would go up and for a matter dominated universe, it would below the zero deceleration line.

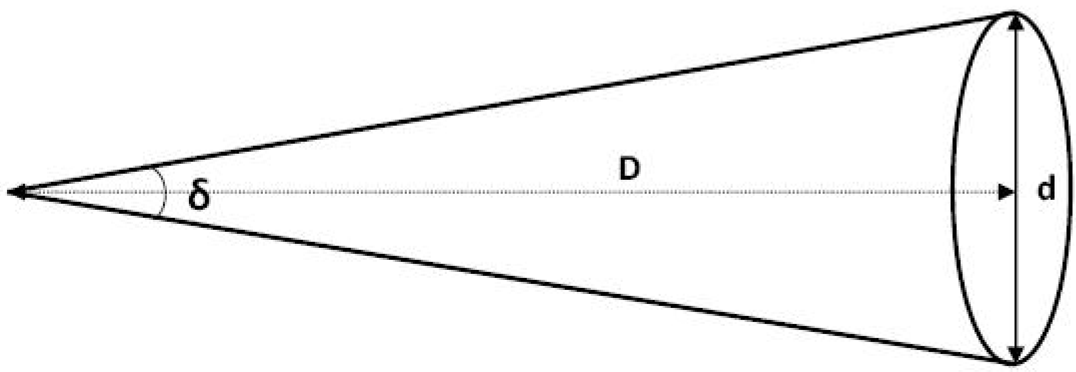
Comment

In this model independent method, the only two parameters occur are Hubble constant H_0 and deceleration parameter q_0 .

Angular Diameter Distance

Observation

Angular diameter distance is really useful if we have some standard ruler. Now assume we have a ruler d, we can find out the angle between the two ends of the ruler, by some kind of measurement.



At the same time, we can use magic of math

$$\theta = d/D.$$

Now as we already find out what θ is by a measurement, and we said about the d is a standard ruler, which means we know the length of it very well. Then we can find out the distance D, which is the distance between us and the standard ruler.

Theory

We can find out this kind of distance, which we will denote it as d_A from now on. What is it for?

A angular diameter distance is the physical distance between us and the standard ruler,

$$d_A = a(t)r.$$

We can use the same trick we used in luminosity distance calculations, and it is easy to find that

$$d_A = H_0^{-1}(z - \frac{1}{2}(3 + q_0)z^2).$$

Again, the observation is related to only two parameters, Hubble constant H_0 and deceleration parameter q_0.

Standard Rulers

It is hard to imagine that we really have some standard rulers. In fact, we do. They are

- Baryon Acoustic Oscillation
- Sound Horizon at Recombination

Galaxy Number Count

Now we can see anything that is only (simply) related to physical or comoving distance can be determined by this trick. The result is that only two cosmological parameters would come in our equation as long as we keep only upper to order two of redshift.

Here another example is the galaxy number count.

$$\frac{\mathrm{d}N_g}{\mathrm{d}z\mathrm{d}\Omega} = z^2\frac{n_0}{H_0^3}(1 - 2(1 + q_0)z).$$

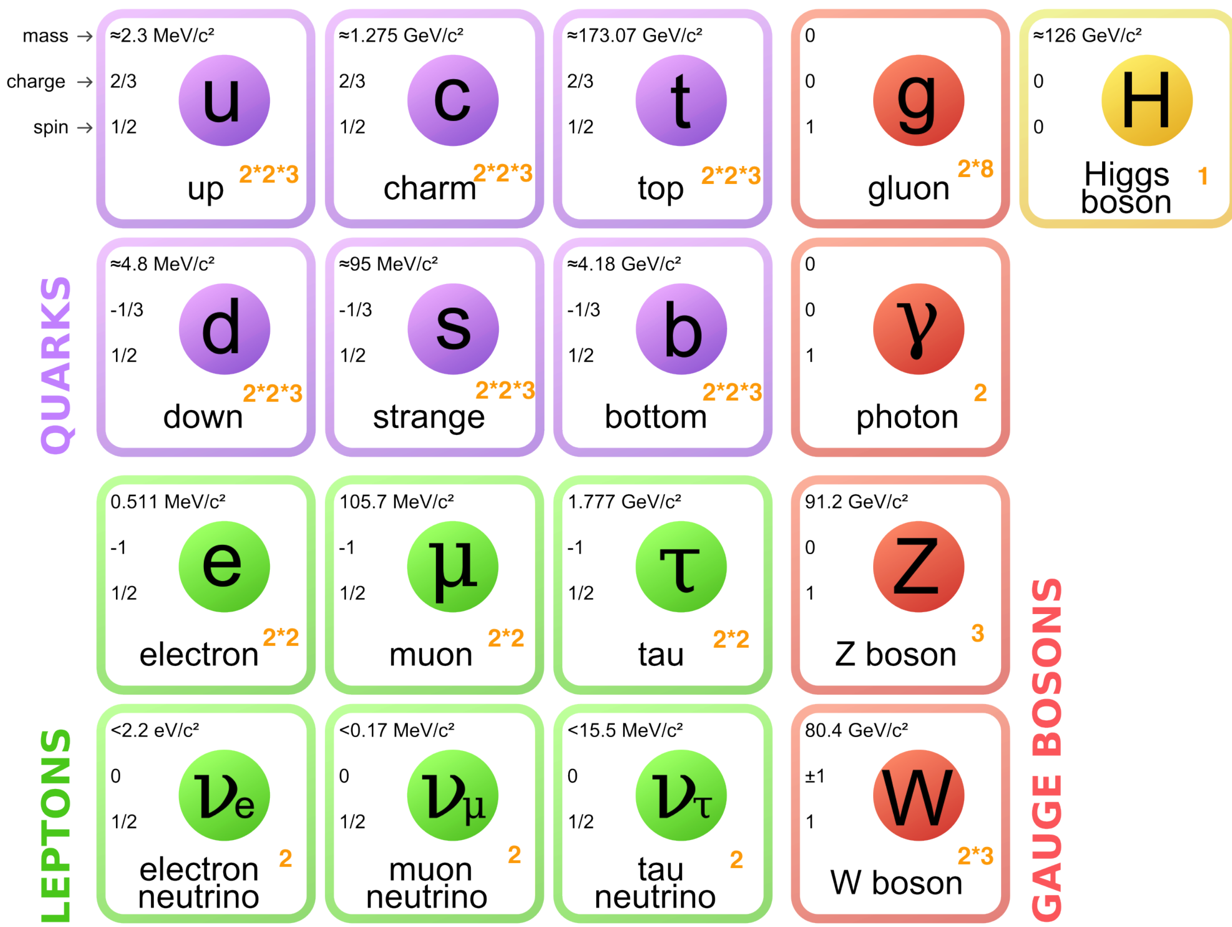
12.2 Thermal History of The Universe

12.2.1 Review of Standard Model for Particle Physics

SM of particle physics

1. describes elementary particles and their interactions.
2. is well test with experiments.

Degree of Freedom of Elementary Particles



IMG Source: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg

The orange numbers at the right bottom of each particle is the degrees of freedom it has. Here are some comments.

1. Photons have only two DoF because it is mass 0. Same reason can apply to gluon. But according to symmetry, there are 8 kinds of gluons.
2. W bosons carry charges. This is where the 2 come from.
3. Electrons and quarks have antiparticles. So there DoF will be doubled after counting the spin.
4. Each quark have 3 different colors and this gives us the 3 when calculating there DoF.

Finally, we can make this table.

Partilces	Higgs	Messengers	Quarks	Leptons
DoF	1	27	72	18

12.2.2 Expansion and Temperature

We can see that the heaviest particle is top quark with a mass of $m_t = 170\mathrm{GeV}$.

Temperature Greater Than Mass of Top Quark

If temperature of the universe $T \gg m_t$, all particles should be in relativistic regime and the decay (annihilation) and inverse decay (inverse annihilation) are in equilibrium so all particles contribute to the thermal quantities in a relativistic way.

$$g_B = 28$$
$$g_F = 90$$

Then

$$g_* = g_B + \frac{7}{8} g_F = 106.75$$

For convinience, define the following reduced Planck mass

$$8\pi G = \frac{1}{M_p^2}$$

And it's good to know its value, which is $2.4 \times 10^{18}\mathrm{GeV}$.

We would like to know the relation between expansion and temperature. We already know that the energy density is

$$\rho = g_* \frac{\pi^2}{30} T^4$$

So the expansion is

$$H^2 = \frac{8\pi G}{3}\rho = 106.75 \times \frac{\pi^2}{30}\frac{T^4}{3M_p^2}$$

So Hubble function is

$$H \approx 3\frac{T^2}{M_p}$$

Temperature Down to Mass of A Particle

As temperature drops down, particle dacay (annihilation) will be greater than its inverse which is suppressed by Boltzmann factor $\exp(-m/T)$. The decay rate is so quick that the particle will almost dispear before the universe expand a lot.

So when the temperature drops below the mass of a particle, it won't contribute to the energy density. Their DoF will just dispear.

For example, if T MeV, Higgs and W and Z will decay and quarks are combined with gluons. So we only have **photons, electrons, neutrinos** as elementary particles, that is $g_* = 10.75$.

The Hubble function,

$$H \approx \frac{T^2}{M_p^2}$$

12.2.3 Decay Rate VS Expansion Rate

We can generally prove that decay rate is much faster than the expansion rate. . . . . To be added.

12.3 FLRW Universe

Why Does Pressure Has Similar Contribution as Density

One of the mysteries is that pressure contributes to the evolution of the universe in a way similar to density. A even weirder, negative pressure would expand the universe. Can you imagine this?

CHAPTER 13

DOI

CHAPTER 14

License

Bibliography

[Schutz] Schutz, A First Course in General Relativity(Second Edition).

[Schutz2009] *A First Course in General Relativity*

[CBLiang]

[Schutz] A First Course in General Relativity (Second Edition)

[Schutz] A First Course in General Relativity (Second Edition)

[1a] arXiv:0712.0607

[1b] **Eotvos experiment**: using torsion balance to test the equality of gravitational mass and inertial mass. Wikipedia has a photo of how this works.

[1c] $\eta = 2\frac{ABS(a1-a2)}{ABS(a1+a2)}$. $a1$ and $a2$ are the accelerations of the two bodies in Eotvos torsion balance. Thus η is the accleration difference of the two objects.

[3a] References: **R. W. P. Drever**. A search for anisotropy of inertial mass using a free precession technique. Philosophical Magazin, 6:683-687, May 1961. ; **V. W. Hughes, H. G. Robinson, and V. Beltran-Lopez**. Upper Limit for the Anisotropy of Inertial Mass from Nuclear Resonance Experiments. Physical Review Letters, 4:342-344, Apr. 1960. ; **S. K. Lamoreaux, J. P. Jacobs, B. R. Heckel, F. J. Raab, and E. N. Fortson**. New limits on spatial anisotropy from optically-pumped 201 Hg and 199 Hg. Physical Review Letters, 57:3125–3128, Dec. 1986. ; **T. E. Chupp, R. J. Hoare, R. A. Loveman, E. R. Oteiza, J. M. Richardson, M. E. Wagshul, and A. K. Thompson**. Results of a new test of local Lorentz invariance: A search for mass anisotropy in 21 Ne. Physical Review Letters, 63:1541–1545, Oct. 1989.

[3b] **Hughes-Drever Experiment**: test the isotropy of mass and space through the NMR spectrum, or the monometric spacetime.

[3c] **n**: four momentum of the test particle is $p_\mu = \frac{m g_{\mu\nu} u^\nu}{\sqrt{-g_{\alpha\beta} u^\alpha u^\beta}} + \frac{n h_{\mu\nu} u^\nu}{-h_{\alpha\beta} u^\alpha u^\beta}$. Thus n is the effect of another metric.

[5a] References, **I. I. Shapiro**. Fourth Test of General Relativity. Physical Review Letters, 13:789–791, Dec. 1964. ; **B. Bertotti, L. Iess, and P. Tortora**. A test of general relativity using radio links with the Cassini spacecraft. Nature, 425:374–376, Sept. 2003.

[5b] **Shapiro time-delay effect**: time delay when light travels through a massive object.

[6a] Observational data for the value of perihelion precession of Mercury are summarized in **E. V. Pitjeva**. Modern Numerical Ephemerides of Planets and the Importance of Ranging Observations for Their Creation. Celestial Mechanics and Dynamical Astronomy, 80:249–271, July 2001.

[6b] PPN formalism is the lowest order of GR.

[6c] **Anomalous precession**:

[7a] **K. Nordtvedt**. Equivalence Principle for Massive Bodies. I. Phenomenology. Physical Review, 169:1014–1016, May 1968. ; **J. G. Williams, S. G. Turyshev, and D. H. Boggs**. Progress in Lunar Laser Ranging Tests of Relativistic Gravity. Physical Review Letters, 93(26):261101, Dec. 2004, arXiv:gr-qc/0411113.

[7b] **Nordtvedt effect**: massive objects in Eotvos torsion balance experiments. We can use the whole Earth-Moon system to test this effect.

[8a] To be added

[8b] There is a Lense Thirring effect here. GPB has done this.

[9a] GR can be reduced to Newtonian potential at small range.

[9b] Currently, most of the modification has a Yukawa potential form.

[Schutz] A First Course in General Relativity (Second Edition)

[helioosc] Jørgen Christensen-Dalsgaard's Lecture Notes

Index