Things before QFT for Sloppy Engineers

Sizhe Liu

University of Illinois at Urbana-Champaign

Version 1.0

Contents

Chapter 1

Classic Mechanics

1.1 Formulations

we call the description of a given theory in a particular mathematical arena a formulation of the theory.

Lagrangian

$$L = T - V$$

where T denotes the kinetic energy and V the potential energy.

Euler-Lagrange equation

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

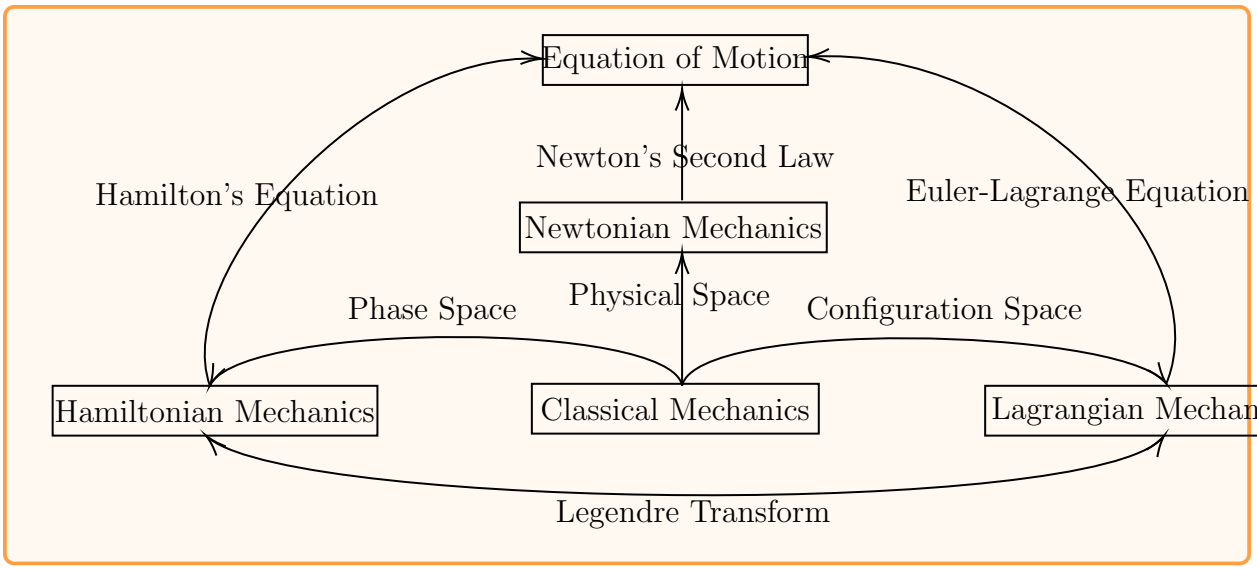
Hamiltonian

$$H = p\dot{q} - L$$

where p denotes the momentum, $\dot{q}$ the velocity and L, as before, the Lagrangian.

Hamiltonian's equation

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q}$$
$$\frac{dq}{dt} = \frac{\partial H}{\partial p}$$



1.2 Basic quantities

Momentum is the velocity of an object times its mass

$$\vec{p}(t) \equiv m\dot{\vec{q}}(t)$$

$$\frac{dp}{dt} = F$$
$$\int_{t_i}^{t_f} dp = \int_{t_i}^{t_f} F dt$$
$$\Delta p = F\Delta t$$

In words, this means that the momentum of an object tells us how long it takes a given force F to stop it. Formulated differently, we need a much bigger force to stop it quickly.

Angular momentum is defined as the cross product of the position vector and the momentum vector

$$\vec{L}(t) = \vec{q}(t) \times \vec{p}(t) = m\vec{q}(t) \times \dot{q}(t)$$

While (linear) momentum tells us how hard it is to stop an object using a specific force, angular momentum tells us how hard it is to stop it spinning.

Momentum and angular momentum are conserved.

1.3 Configuration Space

In a physical space, we use one function $f(x)$ to describe each object in the system separately. In configuration space, we use a vector of functions $\mathbf{r} = [f(x), g(x), h(x)...]$ to describe the system as a whole. **Using the idea of gluing the spaces together, we always only need one vector which lives in an an $\mathbb{R}^N$ -dimensional space. If the objects are allowed to move freely in three dimensions, our vector $\vec{r}$ lives in $\mathbb{R}^{3N}$ since we are gluing N times $\mathbb{R}^3$ together.**

In configuration space, we can imagine the whole system as just one point that moves through this higher-dimensional space. Each point in configuration space corresponds to one specific configuration the system can be in.

To summarize: while individual objects move in the three dimensional physical space, the time evolution of a system as a whole takes place in a higher-dimensional configuration space. A single trajectory in configuration space describes the evolution of a system as a whole.

1.4 Phase Space

Configuration space only keeps track of the locations of the various objects. But in addition to a vector that keeps track of the locations, we need a vector that keeps track of the momenta.

This motivates the construction of the **phase space** which works completely analogously to how we constructed configuration space. However, this time we also act as if **the momenta live in a different space and then glue the momentum spaces to our location spaces**. As a result, we can describe the complete state (not just the configuration) of our system with a single vector.

- Classical mechanics in physical space is what we call the Newtonian formulation.
- Classical mechanics in configuration space is what we call the Lagrangian formulation.
- Classical mechanics in phase space is what we call the Hamiltonian formulation.

1.5 Lagrangian Mechanics

Any system evolves in such a way that the action required is minimal.

A path in configuration space between two fixed points X and Y corresponds to one specific possibility for how our system evolves between two fixed configurations. **We denote the action by $S[q(t)]$ for path the path $q(t)$ from X to Y.** And it is calculated as:

$$\text{action}[q(t)] = S[q(t)] = \int_{t_i}^{t_f} dtL = \int_{t_i}^{t_f} dt(T - V)$$

1.5.1 Action and Lagrangian

Question

why we integrate over L to get the action?

Answer

Kinetic energy is a measure for how much is going on in our system.
Potential energy is a measure for how much could happen, but isn't.
The Lagrangian is therefore a direct measure for the "liveliness" within a system at a specific moment in time. A high kinetic energy implies a large Lagrangian and that our system is extremely lively. A high potential energy implies a small Lagrangian and a less lively system.

The action $S[q(t)]$ is not a function but a functional.

A function eats a number x and spits out a number $f(x)$. Specifically, the action functional $S[q(t)]$ yields a number for each path $q(t)$. We call this number the action of the path.

The Lagrangian is a function of the location q and the velocity $\frac{dq}{dt}$: ' $L = L(q, \dot{q})$. This means that the Lagrangian does not depend on the acceleration $\frac{d^2q}{dt^2}$ or even higher derivatives.

1.5.2 Variational Calculus

For an ordinary function $f(x)$, we can find the minimum by calculating the zeroes of its derivative:

$$\frac{df(x)}{dx} \stackrel{!}{=} 0$$

A minimum is characterized by its neighborhood. If all neighboring points lie higher, the point in question is a minimum.

Let's use this idea to once more find the minimum of the function $f(x) = 3x^2+x$.

We now pick one specific location $x = a$ and start investigating its neighborhood $a \to a+\epsilon$, where ϵ is an infinitesimally small (positive or negative) number. In general, we call ϵ **a variation.**

Putting this into the function yields

$$\begin{aligned} f(a+\epsilon) &= 3(a+\epsilon)^2 + (a+\epsilon) \\ &= 3\left(a^2 + 2a\epsilon + \epsilon^2\right) + a + \epsilon \end{aligned}$$

If the location a is a minimum, we can't get lower by going in any direction ϵ. Mathematically, this implies:

All terms first order in ϵ must vanish. Otherwise, for a negative ϵ the function value $f(a+\epsilon)$ would be smaller than $f(a)$ and therefore, a wouldn't be a minimum.

If we collect all terms linear in ϵ and demand that they vanish,

$$3 \cdot 2a\epsilon + \epsilon \stackrel{!}{=} 0$$
$$\therefore \quad 6a + 1 \stackrel{!}{=} 0$$

we find

$$a = \frac{-1}{6}$$

The variational method of finding minima can also be applied to functionals like the action $S[q(t)]$, not just functions. Take note that for functionals, our goal isn't to find a location like a which is the minimum of a function but instead, to find a function $q(t)$ which is the minimum of a function.

1.5.3 The Euler-Lagrange Equation

Our task now is to find a method which allows us to calculate the path $q_m(t)$ for which the action functional is a minimum.

We start again with a concrete choice $q(t)$ and consider small variations around this specific path

$$q(t) \to q(t) + \epsilon(t)$$

$$\dot{q}(t) \to \dot{q}(t) + \dot{\epsilon}(t)$$

where ϵ is again an infinitesimally small variation.

We consider variations between two fixed configurations $(q_i(t_i) \;\; q_f(t_f))$. Therefore, the variation ϵ has to vanish at t_i and t_f :

$$0 = \epsilon(t_i) = \epsilon(t_f)$$

$$S = \int_{t_i}^{t_f} dt L(q(t) + \epsilon(t), \dot{q}(t) + \dot{\epsilon}(t))$$

According to the Taylor expansion,

$$L(q + \epsilon, \dot{q} + \dot{\epsilon}) = L(q, \dot{q}) + \epsilon \frac{\partial L}{\partial q} + \dot{\epsilon} \frac{\partial L}{\partial \dot{q}} + \ldots$$

$$\begin{aligned} S &= \int_{t_i}^{t_f} dt L(a(t) + \epsilon(t), \dot{a}(t) + \dot{\epsilon}(t)) \\ &= \int_{t_i}^{t_f} dt \left(L(a, \dot{a}) + \epsilon \frac{\partial L}{\partial a} + \dot{\epsilon} \frac{\partial L}{\partial \dot{a}} + \dots \right) \end{aligned}$$

Minima are characterized by vanishing first oerder variations, so again:

$$\int_{t_i}^{t_f} dt \left[\epsilon \frac{\partial L}{\partial q} + \dot{\epsilon} \frac{\partial L}{\partial \dot{q}} \right] \stackrel{!}{=} 0$$

So, first of all, we integrate by parts the second term on the right-hand side:

$$\begin{aligned} \int_{t_i}^{t_f} dt \dot{\epsilon} \frac{\partial L}{\partial \dot{q}} &= \int_{t_i}^{t_f} dt \left(\frac{d}{dt} \epsilon \right) \frac{\partial L}{\partial \dot{q}} \\ &= \epsilon \frac{\partial L}{\partial \dot{q}} \Big|_{t_1}^{t_2} - \int_{t_i}^{t_f} dt \epsilon \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \end{aligned}$$

since the variation $\epsilon(t)$ vanishes for $t = t_i$ and $t = t_f$, we have

$$\begin{aligned} &\int_{t_i}^{t_f} dt \left[\epsilon \frac{\partial L}{\partial q} + \dot{\epsilon} \frac{\partial L}{\partial \dot{q}} \right] \stackrel{!}{=} 0 \\ \therefore &\int_{t_i}^{t_f} dt \left[\epsilon \frac{\partial L}{\partial q} - \epsilon \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \stackrel{!}{=} 0 \\ \therefore &\int_{t_i}^{t_f} dt \epsilon \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \stackrel{!}{=} 0 \end{aligned}$$

Now we recall that if $q(t)$ is indeed the path of least action that we are looking for, the condition must be correct for any possible variation $\epsilon = \epsilon(t)$. This can only be correct if:

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \stackrel{!}{=} 0 \tag{1.5.1}$$

In particular, for a general potential $V = V(q)$, the first term on LHS yields **generalized force**:

$$\frac{\partial L}{\partial q} = \frac{\partial (T(\dot{q}) - V(q))}{\partial q} = -\frac{\partial V(q)}{\partial q} \equiv F$$

In general, the **conjugate momentum** is:

$$p \equiv \frac{\partial L}{\partial \dot{q}} \tag{1.5.2}$$

Therefore, the Euler-Lagrange equation says:

The rate of change of the momentum equals the force.

In eqn1.5.2 we call the momentum "**conjugate momentum**" because it is related to a specific variable q. For example, if we describe our system using angles instead of Cartesian coordinates, the corresponding conjugate momenta would be angular momenta.

1.6 Hamiltonian Mechanics

1.6.1 Hamilton's Equations

To derive the equations, the first key idea is that we act as if q_i and p_i are completely independent variables. We recall the definition of the momentum as:

$$p \equiv \frac{\partial L}{\partial \dot{q}} \tag{1.6.1}$$

According to the Euler-Lagrange equation, we have

$$\frac{dp}{dt} = \frac{\partial L}{\partial q} \tag{1.6.2}$$

Moreover, the definition of the velocity is:

$$\frac{dq}{dt} = \dot{q} \tag{1.6.3}$$

Our goal is to rewrite the equations (1.6.2),(1.6.3) in such a way that they only depend on p and no longer on $\dot{q}$.

First of all we invert eqn(1.6.1) to get $\dot{q} = \dot{q}(q, p)$, from which we get new function:

$$\tilde{L}(q, p) \equiv L(q, \dot{q}(q, p)) \tag{1.6.4}$$

In words, $\tilde{L}(q, p)$ is the function that we get if we use the formula $\dot{q} = \dot{q}(q, p)$ to eliminate $\dot{q}$ from $L(q, \dot{q})$. In particular, take note that, in general,

$$\tilde{L}(q, p) \neq L(q, p) \tag{1.6.5}$$

For example, for the free Lagrangian

$$\begin{aligned}\tilde{L}(q,p) &\overset{(1.6)}{=} L(q,\dot{q}(q,p)) \\ &= \frac{m(\dot{q}(q,p))^2}{2} \\ &= \frac{m\left(\frac{p}{m}\right)^2}{2} \\ &= \frac{p^2}{2m}\end{aligned}$$

Therefore

$$L(q,p) = \frac{mp^2}{2} \neq \frac{p^2}{2m} = \tilde{L}(q,p)$$

Now when we calculate the derivative, we find:

$$\begin{aligned}\frac{\partial \tilde{L}(q,p)}{\partial q} &\overset{(1.6)}{=} \frac{\partial L(q,\dot{q}(q,p))}{\partial q} \\ \frac{\partial \tilde{L}(q,p)}{\partial q} &= \frac{\partial L(q,\dot{q})}{\partial q} + \frac{\partial L(q,\dot{q})}{\partial \dot{q}}\frac{\partial \dot{q}(q,p)}{\partial q} \\ \frac{\partial \tilde{L}(q,p)}{\partial q} &= \frac{\partial L(q,\dot{q})}{\partial q} + p\frac{\partial \dot{q}(q,p)}{\partial q}\end{aligned}$$

Since $\frac{\partial}{\partial q}pq = p\frac{\partial}{\partial q}q$ because $\frac{\partial}{\partial q}p = 0$, we have

$$\begin{aligned}\frac{\partial L(q,\dot{q})}{\partial q} &= \frac{\partial \tilde{L}(q,p)}{\partial q} - p\frac{\partial \dot{q}(q,p)}{\partial q} \\ \frac{\partial L(q,\dot{q})}{\partial q} &= \frac{\partial}{\partial q}(\tilde{L}(q,p) - p\dot{q}(q,p))\end{aligned}$$

From eqn.(1.6.2), we have

$$\begin{aligned}\frac{dp}{dt} &= \frac{\partial L}{\partial q} \\ &= \frac{\partial}{\partial q}(L(q,p) - p\dot{q}(q,p)) \\ &= -\frac{\partial H}{\partial q}\end{aligned} \tag{1.6.6}$$

where we define **the Hamiltonian as**

$$H \equiv p\dot{q}(q,p) - \tilde{L}(q,p) \tag{1.6.7}$$

Similarly, we calculate the derivative of $\tilde{L}(q,p)$ with respect to p:

$$\begin{aligned}
\frac{\partial \tilde{L}(q,p)}{\partial p} &\overset{(1.6)}{=} \frac{\partial L(q,\dot{q}(q,p))}{\partial p} \\
\frac{\partial \tilde{L}(q,p)}{\partial p} &= \frac{\partial L(q,\dot{q})}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial p} \\
\frac{\partial \tilde{L}(q,p)}{\partial p} &= p\frac{\partial \dot{q}}{\partial p} \\
\frac{\partial \tilde{L}(q,p)}{\partial p} &= \frac{\partial}{\partial p}(p\dot{q}) - \dot{q}
\end{aligned}$$

Rearrange terms, we obtain

$$\frac{\partial H}{\partial p} = \dot{q} \tag{1.6.8}$$

In summary, the **Hamilton's equations** are simply:

$$\begin{aligned}
\frac{dp}{dt} &= -\frac{\partial H}{\partial q} \\
\frac{dq}{dt} &= \frac{\partial H}{\partial p}
\end{aligned} \tag{1.6.9}$$

As before, if there are multiple objects in the system moving in three dimensions, we need to take all their locations and momenta into account. Hamilton's equations then read

$$\begin{aligned}
\frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i} \\
\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i}
\end{aligned} \tag{1.6.10}$$

1.6.2 Alternative way to derive Hamilton's equation

From the definition of action and Hamilton, we have:

$$\begin{aligned}
S &= \int_{t_{\mathrm{i}}}^{t} dt L \\
&= \int_{t_{\mathrm{i}}}^{t} dt(p\dot{q} - H)
\end{aligned}$$

Now we can once more use the least action princople to derive the correct equations of motion. **We are searching a path in phase space that minimize the** S:

$$S = \int_{t_i}^{t_f} dt(p\dot{q} - H(q,p)) = \int_{t_i}^{t_f} dt \left(p\frac{d}{dt}q - H(q,p) \right)$$

$$S = \int_{t_i}^{t_f} dt \left(((p+\tilde{\epsilon}))\frac{d}{dt}(q+\epsilon) - H(q+\epsilon, p+\tilde{\epsilon}) \right)$$

$$
\begin{aligned}
&= \int_{t_\mathrm{i}}^{t_f} dt \left((p+\tilde{\epsilon}) \frac{d}{dt}(q+\epsilon) \right. \\
&\qquad \left. -H(q,p) - \epsilon \frac{\partial H(q,p)}{\partial q} - \tilde{\epsilon} \frac{\partial H(q,p)}{\partial p} - \dots \right) \\
&= \int_{t_\mathrm{i}}^{t_\mathrm{f}} dt \left(p\frac{dq}{dt} - H + \tilde{\epsilon} \left(\frac{d}{dt} q - \frac{\partial H}{\partial p} \right) + \dots \right. \\
&\qquad \left. -\epsilon \frac{\partial H}{\partial q} + p\frac{d\epsilon}{dt} + \tilde{\epsilon}\frac{d\epsilon}{dt} \right)
\end{aligned}
$$

In the last line we have a term proportional to $\frac{d\epsilon}{dt}$, but we can turn it into a term proportional to ϵ by integrating by parts:

$$
\begin{aligned}
\int_{t_\mathrm{i}}^{t_f} dt p \frac{d\epsilon}{dt} &= \epsilon p|_{t_\mathrm{i}}^{t_f} - \int_{t_\mathrm{i}}^{t_f} dt \frac{dp}{dt}\epsilon \\
&= -\int_{t_\mathrm{i}}^{t_f} dt \frac{dp}{dt}\epsilon
\end{aligned}
$$

because $\epsilon\left(t_i\right) = \epsilon\left(t_f\right) = 0$. By using this, we can factor out ϵ:

$$
\begin{aligned}
S &= \int_{t_i}^{t_f} dt \left(p\frac{dq}{dt} - H + \tilde{\epsilon} \left(\frac{d}{dt} q - \frac{\partial H}{\partial p} \right) + \dots \right. \\
&\qquad \left. -\epsilon \frac{\partial H}{\partial q} + p\frac{d\epsilon}{dt} + \tilde{\epsilon}\frac{d\epsilon}{dt} \right)
\end{aligned}
$$

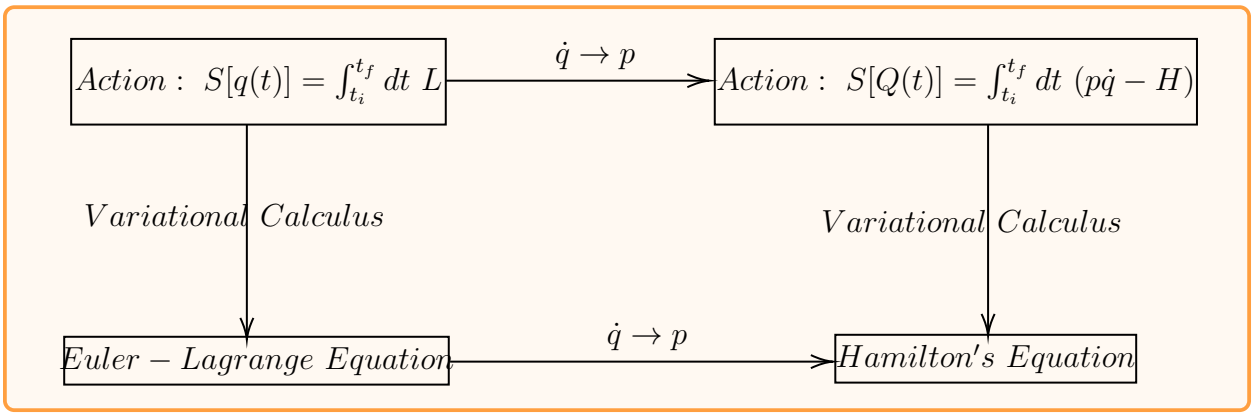
$$
\begin{aligned}
&= \int_{t_\mathrm{i}}^{t} dt \left(p\frac{dq}{dt} - H + \tilde{\epsilon} \left(\frac{d}{dt} q - \frac{\partial H}{\partial p} \right) + \dots \right. \\
&\qquad \left. -\epsilon \frac{\partial H}{\partial q} - \frac{dp}{dt}\epsilon + \tilde{\epsilon}\frac{d\epsilon}{dt} \right) \\
&= \int_{t_\mathrm{i}}^{t} dt \left(p\frac{dq}{dt} - H + \tilde{\epsilon} \left(\frac{d}{dt} q - \frac{\partial H}{\partial p} \right) + \dots \right. \\
&\qquad \left. -\epsilon \left(\frac{\partial H}{\partial q} + \frac{dp}{dt} \right) + \tilde{\epsilon}\frac{d\epsilon}{dt} \right)
\end{aligned}
$$

This means that the terms linear in ϵ and $\tilde{\epsilon}$ only vanish, in general, if the following two conditions are fulfilled:

$$
\begin{aligned}
\frac{\partial H}{\partial q} + \frac{d}{dt} p &\stackrel{!}{=} 0 \\
\frac{\partial H}{\partial p} - \frac{d}{dt} q &\stackrel{!}{=} 0
\end{aligned}
\tag{1.6.11}
$$

The path $(q(t), p(t))$ which fulfills these two conditions is the correct path which minimizes the action and therefore describes the evolution of our system.

The following diagram illustrates the relationship between the derivations discussed above:



1.6.3 Meaning of Hamilton's Equation

First of all, **we need to calculate the momentum explicitly:**

$$\begin{aligned} p &\equiv \frac{\partial L}{\partial \dot{q}} \\ &= \frac{\partial \left(\frac{1}{2}m\dot{q}^2 - V(q)\right)}{\partial \dot{q}} \\ &= m\dot{q} \end{aligned}$$

Using this result, we ca derive the Hamiltonian:

$$\begin{aligned} H &= p\dot{q} - L \\ &= p\dot{q} - \left(\frac{1}{2}m\dot{q}^2 - V(q)\right) \\ &= p\frac{p}{m} - \left(\frac{1}{2}m\left(\frac{p}{m}\right)^2 - V(q)\right) \\ &= \frac{p^2}{2m} + V(q) \end{aligned}$$

This is exactly the total energy of the object.The Hamiltonian represents the total energy.

From the first Hamilton's equation, we have

$$\begin{aligned}\frac{dp}{dt} &= -\frac{\partial H}{\partial q} \\ &= -\frac{\partial\left(\frac{p^2}{2m} + V(q)\right)}{\partial q} \\ &= -\frac{\partial V(q)}{\partial q}\end{aligned}$$

That is, The rate of change of momentum equals the force.

Hamilton's second equation reads:

$$\begin{aligned}\frac{dq}{dt} &= \frac{\partial H}{\partial p} \\ &= \frac{\partial\left(\frac{1}{2}\frac{p^2}{m} + V(q)\right)}{\partial p} \\ &= \frac{p}{m}\end{aligned}$$

Therefore, we can now understand that the purpose of Hamilton's second law is to establish a relationship between the momentum and the rate of change of the position.

1.6.4 Hamilton's General Equation

First of all, we can imagine that sometimes we are not only interested in the locations and momenta of the various objects in the system but other quantities too. For instance, the temperature or how the kinetic energy evolves as time passes can be interesting things to investigate. In phase space, the quantities like the temperature or kinetic energy are functions of q_i and p_i.

Question

But how can we calculate the time evolution of such functions depending on the locations $q_i(t)$ and momenta $p_i(t)$?

For simplicity, let's restrict ourselves to one object moving in one dimension. Then the

total rate of change of a function $F = F(q(t), p(t))$ along a single object's trajectory reads:

$$\frac{d}{dt}F(q,p) = \frac{\partial F(q,p)}{\partial q}\frac{dq}{dt} + \frac{\partial F(q,p)}{\partial p}\frac{dp}{dt}$$

Using Hamilton's equations, we can write:

$$\begin{aligned}\frac{d}{dt}F(q,p) &= \frac{\partial F(q,p)}{\partial q}\frac{dq}{dt} + \frac{\partial F(q,p)}{\partial p}\frac{dp}{dt} \\ &= \frac{\partial F(q,p)}{\partial q}\frac{\partial H(q,p)}{\partial p} - \frac{\partial F(q,p)}{\partial p}\frac{\partial H(q,p)}{\partial q}\end{aligned}$$

since the structure that appears on the right-hand side here is so important, it is conventional to introduce a more compact notation. We therefore introduce the **Poisson bracket $\{,\}$ of two phase space functions $A(q,p), B(q,p)$ by defining**:

$$\{A,B\} \equiv \frac{\partial A}{\partial q}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} \tag{1.6.12}$$

This means that you can't combine two functions in phase space which describe properties of our system arbitrarily and expect to get something that describes another useful property of the system. But **if you calculate the Poisson bracket of the two functions, you'll get something sensible.**

Thus, the equation

$$\frac{d}{dt}F = \{F,H\} \tag{1.6.13}$$

describes the time evolution of a general phase spack function and we call it **Hamilton's equation of motion.**

Notice that

$$\begin{aligned}\frac{d}{dt}q &= \{q,H\} \\ &= \frac{\partial q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial q}{\partial p}\frac{\partial H}{\partial q} \\ &= \frac{\partial H}{\partial p}\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}p &= \{p, H\} \\ &= \frac{\partial p}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial p}{\partial p}\frac{\partial H}{\partial q} \\ &= -\frac{\partial H}{\partial q}\end{aligned}$$

So one way to understand Hamilton's general equation of motion is by imagining that we have a new kind of object $\{, H\}$ (an operator) which eats any function F on phase space ($\{F, H\}$) and spits out the correct time evolution of F.

The Hamiltonian generates time evolution in phase space.

Sometimes we are déaling with a function in phase space which not only depends on q and p, but also explicitly on t. For example, this is necessarily the case if there is a time-dependent potential $V = V(q, t)$. The total rate of change then reads:

$$\frac{d}{dt}F(q, p, t) = \frac{dq}{dt}\frac{\partial F}{\partial q} + \frac{dp}{dt}\frac{\partial F}{\partial p} + \frac{\partial F}{\partial t} \tag{1.6.14}$$

In other words,

$$\frac{d}{dt}F(q, p, t) = \{F, H\} + \frac{\partial F}{\partial t} \tag{1.6.15}$$

1.7 The Newtonian Algorithm

Question

We stand at the top of the Leaning Tower of Pisa and let a ball fall to the ground. How is the ball moving?

$$\vec{F} = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix}$$

Newton's second law tells us

$$\begin{aligned} \frac{d\vec{p}}{dt} &= \vec{F} \\ \frac{d(m\vec{v})}{dt} &= \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix} \\ m\frac{d}{dt}\begin{pmatrix} v_x \\ v_z \end{pmatrix} &= \begin{pmatrix} 0 \\ -mg \end{pmatrix} \\ \frac{d}{dt}\begin{pmatrix} v_x \\ v_z \end{pmatrix} &= \begin{pmatrix} 0 \\ -g \end{pmatrix} \end{aligned}$$

Our task is to solve these three equations of motion. Luckily, we can simply integrate the equations twice since gravity is constant:

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \\ \int_0^t dt' \frac{d}{dt'}\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} &= \int_0^t dt' \begin{pmatrix} 0 \\ -g \end{pmatrix} \\ \begin{pmatrix} v_x(t) \\ v_z(t) \end{pmatrix} - \begin{pmatrix} v_x(0) \\ v_z(0) \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ -gt \end{pmatrix} \end{aligned}$$

$$\int_0^t dt' \begin{pmatrix} v_x(t') \\ v_y(t') \\ v_z(t') \end{pmatrix} - \int_0^t dt' \begin{pmatrix} v_x(0) \\ v_y(0) \\ v_z(0) \end{pmatrix} = \int_0^t dt' \begin{pmatrix} 0 \\ 0 \\ -gt \end{pmatrix}$$

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} - \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} - \begin{pmatrix} v_x(0)t \\ v_y(0)t \\ v_z(0)t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2}gt^2 \end{pmatrix}$$

Next, we need to determine the integration constants

$$v_x(0), v_y(0), v_z(0), x(0), y(0), z(0)$$

Let

$$\begin{pmatrix} v_x(0) \\ v_y(0) \\ v_z(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Finally

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2}gt^2 \end{pmatrix}$$

1.8 Lagrangian Algorithm

Question

We stand at the top of the Leaning Tower of Pisa and let a ball fall to the ground. How is the ball moving?

The Lagrangian reads

$$\begin{aligned} L = T - V &= \frac{1}{2}m\vec{v}^2 - mgz \\ &= \frac{1}{2}m\left(v_x^2 + v_y^2 + v_z^2\right) - mgz \end{aligned}$$

Since $\frac{\partial L}{\partial q} = \frac{d}{dt}\left(\frac{\partial L}{\partial v_q}\right)$,

$$0 = m\frac{d}{dt}v_x$$
$$0 = m\frac{d}{dt}v_y$$
$$-mg = m\frac{d}{dt}v_z$$

Take note that in the derivations above, we didn't have to think about vectors at all. This is one advantage of the Lagrangian formalism. In addition, the Lagrangian formulation of classical mechanics is always a good choice whenever we are dealing with a system which is subject to constraints.

1.8.1 Constraints

In mathematical terms, a constraint is a relationship between coordinates. For example, for a mass attached to a circular loop with radius l, we have the constraint

$$x^2 + y^2 = l^2$$

More generally, a **holonomic constraint** is a formula of the form

$$f\left(q, 1, q_2, \ldots, t\right) = \text{ const.}$$

Non-holonomic constraints

$$f\left(q_1, 1, q_2, \ldots\right) \geq \text{ const.}$$

$$f(q, \dot{q}, t) = \text{ const.}$$

Additionally, take note that it is conventional to call a constraint which does not explicitly depend on t scleronomic (Greek for "rigid") and a constraint with explicit dependence on t rheonomic (Greek for "moving").

The trick which allows us to incorporate constraints in the Lagrangian formalism is known as the method of **Lagrange multipliers** and works as follows.

Question

How to use Lagrange multiplier to add a holonomic constraint?

Answer

First, we rewrite the constraint as:

$$g\left(q_1, 1, q_2, \ldots, t\right) \equiv f\left(q_1, 1, q_2, \ldots, t\right) - \text{ const.}$$

We then take the Lagrangian $L_{\text{free that we would use if the}}$ object could move around freely without constraints and add a new term L_{con} which encodes the constraint

$$\begin{aligned} L_{\text{full}} &= L_{\text{free}} + L_{\text{con}} \\ &= L_{\text{free}} + \lambda g(q,t) \end{aligned}$$

If we treat λ as a new coordinate, the Euler-Lagrange equation for λ tells us

$$\begin{aligned} \frac{\partial L}{\partial \lambda} &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\lambda}}\right) \\ \frac{\partial (L_{\text{free}} + \lambda g(q,t))}{\partial \lambda} &= \frac{d}{dt}\left(\frac{\partial (L_{\text{free}} + \lambda g(q,t))}{\partial \lambda}\right) \\ \frac{\partial (\lambda g(q,t))}{\partial \lambda} &= \frac{d}{dt}\left(\frac{\partial (\lambda g(q,t))}{\partial \dot{\lambda}}\right) \\ \frac{\partial (\lambda g(q,t))}{\partial \lambda} &= 0 \\ g(q,t) &= 0 \end{aligned}$$

In addition, by using the Euler-Lagrange equation for our ordinary coordinates q, we find

$$\begin{aligned} \frac{\partial L}{\partial q} &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) \\ \frac{\partial \left(L_{\text{free}} + \lambda g(q,t)\right)}{\partial q} &= \frac{d}{dt}\left(\frac{\partial \left(L_{\text{free}} + \lambda g(q,t)\right)}{\partial \dot{q}}\right) \\ \frac{\partial \left(L_{\text{free}} + \lambda g(q,t)\right)}{\partial q} &= \frac{d}{dt}\left(\frac{\partial L_{\text{free}}}{\partial \dot{q}}\right) \\ \frac{\partial L_{\text{free}}}{\partial q} + \lambda \frac{\partial g(q,t)}{\partial q} &= \frac{d}{dt}\left(\frac{\partial L_{\text{free}}}{\partial \dot{q}}\right) \end{aligned}$$

we've learned that the term on the right-hand side $\frac{d}{dt}\left(\frac{\partial L_{\text{free}}}{\partial q}\right)$ is analogous to $\frac{dp}{dt}$ in the Newtonian formalism. Moreover, the first term on the left-hand side $\frac{\partial L_{\text{free}}}{\partial q}$ describes the forces. In other words, $\lambda \frac{\partial g(q,t)}{\partial q}$ add new forces to the equation of motion.

If there is more than one constraint

$$L_{\text{full}} = L_{\text{free}} + \lambda_{1g_1}(q,t) + \lambda_2 g_1(q,t) + \dots$$

Moreover, using the Euler-Lagrange equation for the regular coordinate q yields the equation of motion including all constraint forces:

$$\frac{\partial L_{\text{free}}}{\partial q} + \lambda_1 \frac{\partial g_1(q,t)}{\partial q} + \lambda_2 \frac{\partial g_2(q,t)}{\partial q} + \dots = \frac{d}{dt}\left(\frac{\partial L_{\text{free}}}{\partial \dot{q}}\right)$$

1.8.2 Point Transformations and Generalized Coordinates

In general, we call a transformation from one set of coordinates in configuration space $q = (q_1, q_2, \dots)$ to a new set of coordinates $q' = (q'_1, q'_2, \dots)$ a **point transformation.**

The Euler-Lagrange equation is valid for any choice of coordinates.

It is conventional to rephrase this by saying that in the Lagrangian formalism, we are free to use generalized coordinates. Formulated differently, if we find such coordinates for which the constraint term vanishes, we've found the most natural coordinates to describe our system.

Question

Now, why does the Euler-Lagrange equation look exactly the same no matter which coordinates we choose?

Answer

We want to show that no matter which transformation formulas we use

$$q' = q'(q,t)$$

$$\dot{q}' = \dot{q}'(q,\dot{q},t)$$

the Euler-Lagrange equation in terms of the new coordinates

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

where

$$L\left(q',\dot{q}',t\right) = L\left(q\left(q',t\right),\dot{q}\left(q',\dot{q}',t\right),t\right)$$

also

$$\begin{aligned}\dot{q} &= \frac{d}{dt}q\left(q',t\right)\\ &= \frac{\partial q}{\partial q'}\frac{dq'}{dt} + \frac{\partial q}{\partial t}\\ &= \frac{\partial q}{\partial q'}\dot{q}' + \frac{\partial q}{\partial t}\end{aligned}$$

which implies

$$\begin{aligned}\frac{\partial}{\partial \dot{q}'}\dot{q} =& \frac{\partial}{\partial \dot{q}'}\left(\frac{\partial q}{\partial q'}\dot{q}' + \frac{\partial q}{\partial t}\right)\\ &= \frac{\partial q}{\partial q'}\end{aligned}$$

Secondly

$$\begin{aligned}\frac{\partial \tilde{L}}{\partial q'} &= \frac{\partial L\left(q\left(q',t\right),\dot{q}\left(q',\dot{q}',t\right),t\right)}{\partial q'}\\ &= \frac{\partial L}{\partial q}\frac{\partial q}{\partial q'} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q'}\end{aligned}$$

Thirdly

$$\begin{aligned}\frac{\partial \tilde{L}}{\partial \dot{q}'} &= \frac{\partial L\left(q\left(q',t\right),\dot{q}\left(q',\dot{q}',t\right),t\right)}{\partial \dot{q}'}\\ &= \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}'}{\partial \dot{q}'}\\ &= \frac{\partial L}{\partial \dot{q}}\frac{\partial q}{\partial q'}\end{aligned}$$

which implies

$$\begin{aligned}\frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{q}'} &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\frac{\partial q}{\partial q'}\right)\\ &= \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\frac{\partial q}{\partial q'} + \frac{\partial L}{\partial \dot{q}}\left(\frac{d}{dt}\frac{\partial q}{\partial q'}\right)\\ &= \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\frac{\partial q}{\partial q'} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q'}\end{aligned}$$

Answer

With these formulas at hand, let's rewrite the "new" Euler-Lagrange equation

$$
\begin{aligned}
0 &\overset{!}{=} \frac{\partial \tilde{L}}{\partial q'} - \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}'}\right) \\
&= \left(\frac{\partial L}{\partial q}\frac{\partial q}{\partial q'} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q'}\right) - \left(\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\frac{\partial q}{\partial q'} + \frac{\partial L}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial q'}\right) \\
&= \frac{\partial L}{\partial q}\frac{\partial q}{\partial q'} - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\frac{\partial q}{\partial q'} \\
&= \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\frac{\partial q}{\partial q'} = 0
\end{aligned}
$$

However, take note that we've only shown that the Euler-Lagrange equation has the same form no matter which coordinates we use. In technical terms, **we say the Euler-Lagrange equation is only covariant but not invariant under general coordinate transformations.**

1.8.3 Gauge Transformation

We can not only change our configuration space coordinates q, but we can also change the Lagrangian itself and consequently also the action itself without changing the Euler-Lagrange equations.This is possible because **the number which the action functional $S[q(t)]$ assigns to each path $q(t)$ is not really important. Instead, the only thing we really care about is the relative difference between the numbers assigned to different paths.**

In particular, the Euler-Lagrange equation does not change when we add the total time derivative of any (differentiable) function F which only depends on the location q and possibly the time coordinate t to the Lagrangian

$$
L \to L' = L + \frac{dF(q,t)}{dt} \tag{1.8.1}
$$

We call this kind of transformation a **gauge transformation**. Now the "new" action can be calculated as:

$$\begin{aligned} S'[q(t)] &= \int_{t_i}^{t_f} dt L'(q(t), \dot{q}(t), t) \\ &= \int_{t_i}^{t_f} dt (L(q(t), \dot{q}(t), t) + F(q(t), t)) \end{aligned}$$

$$\begin{aligned} &= \left(\int_{t_i}^{t_f} dt L \right) + F\left(q\left(t_f\right), t_f\right) - F\left(q\left(t_i\right), t_i\right) \\ &= S + F\left(q\left(t_f\right), t_f\right) - F\left(q\left(t_i\right), t_i\right) \end{aligned}$$

Where $F\left(q\left(t_f\right), t_f\right) - F\left(q\left(t_i\right), t_i\right)$ is a constant which only depends on $q\left(t_f\right), q\left(t_i\right), t_f$ and t_i.37 The initial and final points are fixed.

Moreover, we cannot only add a constant to the action but also multiply it by a constant (non-zero) factor

$$S \to kS$$

This observation directly implies that we can also scale the Lagrangian by a constant factor

$$L \to L' = kL$$

1.9 Hamiltonian Algorithm

If we want to use the Hamiltonian formalism to describe a given system, our main task is to write down the correct Hamiltonian H. However, usually this requires a few intermediate steps.

- First of all, we write down the Lagrangian $L = T - V$
- Then we calculate the corresponding generalized momenta $p = \frac{\partial L}{\partial q}$
- Afterwards, we solve the formulas we found in the previous step for $\dot{q}$
- This allows us to eliminate $\dot{q}$ from $L = L(q, \dot{q}, t)$ and to use the formula $H(q, p, t) = p\dot{q} - L(q, \dot{q}(q, p, t), t)$
- Finally, use the Hamilton's equations $\frac{dp}{dt} = -\frac{\partial H}{\partial q}, \frac{dq}{dt} = \frac{\partial H}{\partial p}$.

However, the Hamiltonian way to derive them was far more cumbersome. So **the real advantage of the Hamiltonian method is not a practical but mostly a conceptual one**.

1.9.1 Canonical Transformation and Canonical Coordinates

The Hamilton's equation and Hamiltonian formalism remains valid no matter which coordinates we choose.

Since there is a close connection between the p and $q(\dot{q}$ more precisely), there are important restrictions which our coordinate transformations $q \to Q(q, p)$ and $p \to P(q, p)$ must fulfill.**Transformations fulfilled the restrictions are called canonical transformations.** In other words, a canonical transformation is a switch in coordinates which does not "break" our Hamiltonian framework.

We learned that the time-evolution of any phase space function is described by Hamilton's general equation of motion:

$$\frac{d}{dt}F = \{F, H\}$$

since our new coordinates (Q, P) are defined as functions of the old coordinates, we have

$$\begin{aligned} \frac{d}{dt}Q &= \{Q, H\}_{q,p} = \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q} \\ \frac{d}{dt}P &= \{P, H\}_{q,p} = \frac{\partial P}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial P}{\partial p}\frac{\partial H}{\partial q} \end{aligned} \tag{1.9.1}$$

Using the chain rule, we have

$$\begin{aligned}\frac{\partial H(q(Q,P),p(Q,P))}{\partial q} &= \frac{\partial H(Q,P)}{\partial q} \\ &= \frac{\partial \tilde{H}}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial \tilde{H}}{\partial P}\frac{\partial P}{\partial q} \\ \frac{\partial H(q(Q,P),p(Q,P))}{\partial p} &= \frac{\partial H(Q,P)}{\partial p} \\ &= \frac{\partial \bar{H}}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial p}\end{aligned}$$

Substituting these expression into (1.9.1) yields

$$\begin{aligned}\frac{d}{dt}Q &= \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q} \\ &= \frac{\partial Q}{\partial q}\left(\frac{\partial H}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial p}\right) - \frac{\partial Q}{\partial p}\left(\frac{\partial \hat{H}}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial H}{\partial P}\frac{\partial P}{\partial q}\right) \\ &= \frac{\partial Q}{\partial q}\frac{\partial \hat{H}}{\partial P}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial \hat{H}}{\partial P}\frac{\partial P}{\partial q} \\ &= \frac{\partial \hat{H}}{\partial P}\left(\frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q}\right) \\ &= \frac{\partial \bar{H}}{\partial P}\{Q,P\}_{q,p}\end{aligned}$$

Similarly,

$$\frac{d}{dt}P == -\frac{\partial \tilde{H}}{\partial Q}\{Q,P\}_{q,p}$$

In order not to break the Hamilton's formalism, we have

$$\{Q,P\}_{q,p} = 1 \tag{1.9.2}$$

The equation above is the defining condition of canonical coordinates, and is call canonical Poisson bracket relation.

1.9.2 Canonical Point and Gauge Transformations

A point transformation in configuration space reads

$$q \to Q = Q(q) \tag{1.9.3}$$

And the velocities are modified as:

$$\dot{q} \to \dot{Q} = \dot{Q}(q) = \frac{d}{dt}Q(q) = \frac{\partial Q(q)}{\partial q}\frac{dq}{dt} = \frac{\partial Q(q)}{\partial q}\dot{q}$$

This indicates

$$\frac{\partial \dot{Q}}{\partial \dot{q}} = \frac{\partial}{\dot{q}}\left(\frac{\partial Q(q)}{\partial q}\dot{q}\right) = \frac{\partial Q(q)}{\partial q}$$

This, in turn, implies that the momentum p is also directly modified whenever we perform a point transformation

$$\begin{aligned} P &= \frac{\partial \tilde{L}(Q,\dot{Q}))}{\partial \dot{Q}} \\ &= \frac{\partial L(q(Q),\dot{q}(Q,\dot{Q}))}{\partial \dot{Q}} \\ &= \frac{\partial L(q,\dot{q})}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial \dot{Q}} \\ &= p\frac{\partial q}{\partial Q} \end{aligned} \tag{1.9.4}$$

To check that transformation in (1.9.3) and (1.9.4) really correspond to a canonical transformation in phase space, we calculate $\{Q,P\}_{q,p}$:

$$\begin{aligned} \{Q,P\}_{q,p} &= \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \\ &= \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} \\ &= \frac{\partial Q}{\partial q}\frac{\partial}{\partial p}\left(p\frac{\partial q}{\partial Q}\right) \\ &= \frac{\partial Q}{\partial q}\frac{\partial q}{\partial Q} = 1 \end{aligned}$$

Therefore,the point transformations indeed correspond to canonical phase space transformations.

From Hamiltonian-Lagrangian transformation, we also found that:

$$\begin{aligned} H(q,p) &\equiv p\dot{q} - L(q,p) \\ \tilde{H}(Q,P) &= p(Q,P)\dot{q}(Q,\dot{Q},P) - L(q(Q,P),\dot{q}(Q,\dot{Q},P)) \\ &= P\frac{\partial Q}{\partial q}\dot{Q}\frac{\partial q}{\partial Q} - L(q(Q,P),\dot{q}(Q,\dot{Q},P)) \\ &= P\dot{Q} - L(q(Q,P),\dot{q}(Q,\dot{Q},P)) \end{aligned}$$

Now secondly, let's check what gauge transformations look like in the Hamiltonian formalism.

$$\begin{aligned} P &= \frac{\partial L'}{\partial \dot{q}} \\ &= \frac{\partial \left(L + \frac{dF(q,t)}{dt} \right)}{\partial \dot{q}} \\ &= \frac{\partial L}{\partial \dot{q}} + \frac{\partial}{\partial \dot{q}} \frac{dF(q,t)}{dt} \\ &= \frac{\partial L}{\partial \dot{q}} + \frac{\partial F}{\partial q} \\ &= p + \frac{\partial F}{\partial q} \end{aligned} \tag{1.9.5}$$

Moreover, the location coordinates remain completely unaffected by a gauge transformation.This is, after all, one of the defining features of a gauge transformation.

$$q \to Q(q) = q \tag{1.9.6}$$

While a point transformation modifies the location coordinates (plus conjugate momentum coordinates), a gauge transformation only modifies the Lagrangian and, as a result, also the conjugate momenta.

The following diagram summarizes what we've discovered in this section:

	Point Transformation	*Gauge Transformation*
Configuration Space	$q \to Q = Q(q)$ $\dot{q} \to \dot{Q} = \frac{\partial Q}{\partial q}\dot{q}$ $L \to \tilde{L}(Q, \dot{Q}, t)$	$q \to Q = Q(q)$ $\dot{q} \to \dot{Q} = \dot{q}$ $L \to L' = L + \frac{\partial F(q,t)}{\partial t}$
Phase Space	$q \to Q = Q(q)$ $p \to P = \frac{\partial q}{\partial Q}p$ $H \to \tilde{H}(Q, P, t)$	$q \to Q = q$ $p \to P = p + \frac{\partial F(q,t)}{\partial q}$ $L \to H' = H - \frac{\partial F(q,t)}{\partial t}$

1.9.3 Infinitesimal Canonical Transformation

Mathematically, we can write a tiny transformation of our phase space coordinates as:

$$\begin{aligned} q_i \to Q_i &= q_i + \epsilon D_i\left(q_i, p_i\right) \\ p_i \to P_i &= p_i + \epsilon E_i\left(q_i, p_i\right) \end{aligned} \tag{1.9.7}$$

where ϵ is an infinitesimally small number. According to eqn.(1.9.1), we have

$$\begin{aligned} \{Q, P\}_{q,p} &= \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \\ &= 1 + \epsilon\frac{\partial D}{\partial q} + \epsilon\frac{\partial E}{\partial p} + \epsilon^2\frac{\partial D}{\partial q}\frac{\partial E}{\partial p} - \epsilon^2\frac{\partial D}{\partial p}\frac{\partial E}{\partial q} \\ &= 1 + \epsilon\frac{\partial D}{\partial q} + \epsilon\frac{\partial E}{\partial p} \end{aligned}$$

Thus,

$$\frac{\partial D}{\partial q} + \frac{\partial E}{\partial p} \stackrel{!}{=} 0 \tag{1.9.8}$$

This implies that there must be a connection between D and E. Specifically, we can conclude that the equation above is fulfilled if

$$D = \frac{\partial G}{\partial p}, \qquad E = -\frac{\partial G}{\partial q} \tag{1.9.9}$$

where $G = G(q, p)$ is some arbitrary function. Motivated by this observation, it is conventional to say that the new object G generates the transformation. Moreover, G **is called the generator of the canonical transformation.** The transformation in (1.9.7) therefore reads

$$\begin{aligned} q_i \to Q_i &= q_i + \epsilon\frac{\partial G}{\partial p} \\ p_i \to P_i &= p_i - \epsilon\frac{\partial G}{\partial q} \end{aligned} \tag{1.9.10}$$

The Poisson bracket is, in some sense, the natural product in phase space. Therefore, if we consider an abstract phase space transformation:

$$\begin{aligned} q \to Q = T \circ q, \quad p \to P = T \circ p \\ = \{T, q\}, \quad = \{T, p\} \end{aligned}$$

Now **we let our generator G act on the coordinates q, p via the Poisson bracket.** To that end, we define an infinitesimal transformation as a linear combination of the identity transformation $I \circ X = X$ and the generator action:

$$T_{\text{inf}} \circ X \equiv (I + \epsilon G) \circ X = X + \epsilon G \circ X$$

We can then calculate:

$$\begin{aligned} q \to Q &= T_{\text{inf}} \circ q \\ &= q + \epsilon G \circ q \\ &= q + \epsilon \{q, G\} \quad == q + \epsilon \frac{\partial G}{\partial p} \end{aligned}$$

Therefore, **our generator G really acts on the phase space coordinates via the Poisson bracket**. We can transform our phase space coordinates using any phase space function $G(q, p)$ We only have to make sure that G acts on our coordinates via the appropriate phase space product, i.e., the Poisson bracket.

1.9.4 Generating Functions

In this section, we discover a new method to construct transformations between canonical coordinates systematically. We first write down the following transformation using a single function $F = F(q, Q, t)$ through the formulas:

$$\begin{aligned} P &= \frac{\partial F}{\partial Q} \\ p &= -\frac{\partial F}{\partial q} \end{aligned} \tag{1.9.11}$$

since F is a function of q, Q, and t, we know that $\frac{\partial F}{\partial Q}$ will be a function of q, Q, and t too. Therefore, the second line in Eq. (1.9.11) is an equation of the form $p = p(q, Q, t)$. We can then invert this equation to get an expression of the form $Q = Q(q, p, t)$. So in other words, given some function F, the second line tells us what our new location coordinates look like in terms of the original location and momentum coordinates. The first line yields a formula of the form $P = P(q, Q, t)$. But we already know from the second line what Q looks like

in terms of the original coordinates. Therefore, we can use this expression to get a formula that specifies the new momentum coordinates P solely in terms of the original phase space coordinates $P = P(q, p, t)$. Therefore, **Eq.(1.9.11) tells us everything that we need to know about how our old coordinates are related to the new ones.**

As a concrete example, let's calculate the canonical transformation generated by

$$F(q, Q) = qQ \tag{1.9.12}$$

Using Eq.1.9.11, we find

$$\begin{aligned} P &= \frac{\partial F}{\partial Q} = \frac{\partial (qQ)}{\partial Q} = q \\ p &= -\frac{\partial F}{\partial q} = -\frac{\partial (qQ)}{\partial q} = -Q \end{aligned} \tag{1.9.13}$$

This means that the canonical transformation generated by $F(q, Q) = qQ$ flips the roles of the momentum and location coordinates.

Hamilton's equations follow as a condition on the path of least action for the Hamiltonian action

$$S = \int_{t_i}^{t_f} L dt = \int_{t_i}^{t_f} (p\dot{q} - H) dt \tag{1.9.14}$$

where $L \equiv p\dot{q} - H$ is the "Hamiltonian Lagrangian". After the transformation $q, p \to Q, P$, the least action principle with an analogous Hamiltonian action

$$\tilde{S} = \int_{t_i}^{t_f} L' dt = \int_{t_i}^{t_f} dt \left(P\dot{Q} - H' \right) \tag{1.9.15}$$

must still be valid, where $L' \equiv P\dot{Q} - H'$ is the transformed "Hamiltonian Lagrangian".

Because the transformation is based on a generating function, the original Lagrangian L and the transformed Lagrangian differ by the total derivative of the generating function $F = F(q, Q, t)$:

$$L' - L = \frac{dF}{dt} \tag{1.9.16}$$

and

$$\begin{aligned}\frac{dF(q,Q)}{dt} &= \frac{\partial F}{\partial q}\frac{\partial q}{\partial t} + \frac{\partial F}{\partial Q}\frac{\partial Q}{\partial t} + \frac{\partial F}{\partial t} \\ &= \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}Q + \frac{\partial F}{\partial t}\end{aligned}$$

With these two equations, we have

$$\begin{gathered}L' - L = \tfrac{dF}{dt} \\ \left(P\dot{Q} - H'\right) - (p\dot{q} - H) = \tfrac{\partial F}{\partial q}\dot{q} + \tfrac{\partial F}{\partial Q}\dot{Q} + \tfrac{\partial F}{\partial t} \\ \left(P - \tfrac{\partial F}{\partial Q}\right)\dot{Q} - \left(p + \tfrac{\partial F}{\partial q}\right)\dot{q} = H' - H + \tfrac{\partial F}{\partial t}\end{gathered}$$

None of the functions in last line depend explicitly on $\dot{q}$. Therefore, if we calculate the partial derivative with respect to $\dot{q}$, we find:

$$\begin{aligned}\frac{\partial}{\partial \dot{q}}\left(P - \frac{\partial F}{\partial Q}\right)\dot{Q} - \frac{\partial}{\partial \dot{q}}\left(p + \frac{\partial F}{\partial q}\right)\dot{q} = \frac{\partial}{\partial \dot{q}}\left(H' - H + \frac{\partial F}{\partial t}\right) \\ - \left(p + \frac{\partial F}{\partial q}\right) = 0\end{aligned} \tag{1.9.17}$$

Similarly

$$\begin{aligned}\frac{\partial}{\partial \dot{Q}}\left(P - \frac{\partial F}{\partial Q}\right)\dot{Q} - \frac{\partial}{\partial \dot{Q}}\left(p + \frac{\partial F}{\partial q}\right)\dot{q} = \frac{\partial}{\partial \dot{Q}}\left(H' - H + \frac{\partial F}{\partial t}\right) \\ \left(P - \frac{\partial F}{\partial Q}\right) = 0\end{aligned} \tag{1.9.18}$$

Therefore, we conclude:

$$\begin{aligned}P &= \frac{\partial F}{\partial Q} \\ p &= -\frac{\partial F}{\partial q}\end{aligned}$$

We can also calculate how the Hamiltonian gets modified by the canonical transformation generated by F as:

$$H' = H - \frac{\partial F}{\partial t} \tag{1.9.19}$$

The algorithm to produce a canonical transformation goes like this:

- Write down a specific generating function $F = F(q, Q, t)$
- Use Eq.1.9.11 to find the formulas which express the new coordinates (Q, P) in terms of the old ones (q, p)
- Use Eq.1.9.19 to caculate the new Hamiltonian $\tilde{H}(Q, P, t)$

1.10 The Harmonic Oscillator

We call all systems with a potential of the form $V = \frac{1}{2}kx^2$ where k is some constant, a harmonic oscillator.

1.10.1 Solve the equations of motion

We first derive the equation of motion using Lagrangian and then solve it using physical coordinates and canonical coordinates. The Lagrangian of a harmonic oscillator

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \tag{1.10.1}$$

We can then calculate the correct equation of motion by putting this Lagrangian into the Euler-Lagrange equation:

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)$$

$$-kx = m\frac{d^2x}{dt^2} \tag{1.10.2}$$

In the Hamiltonian formalism, **our first task is always to determine the Hamiltonian H.**

$$H = p\dot{x} - L$$

where the generalized momentum $p = \frac{\partial L}{\partial \dot{x}}$. Using Hamilton's equations:

$$\begin{aligned}\frac{dp}{dt} &= -\frac{\partial H}{\partial x} \\ &= -\frac{\partial\left(\frac{1}{2}\frac{p^2}{m}+\frac{1}{2}kx^2\right)}{\partial x} \\ &= -kx\end{aligned}$$

$$\begin{aligned}\frac{dx}{dt} &= \frac{\partial H}{\partial p} \\ &= \frac{\partial\left(\frac{1}{2}\frac{p^2}{m}+\frac{1}{2}kx^2\right)}{\partial p} \\ &= \frac{p}{m}\end{aligned}$$

To solve Eq.1.10.2, we first multiply the equation of motion by the velocity $\dot{x}$:

$$m\dot{x}\frac{d^2x}{dt^2} = -k\dot{x}x$$

and

$$\begin{aligned}\dot{x}\frac{d^2x}{dt^2} = \dot{x}\frac{d\dot{x}}{dt} &= \frac{1}{2}\frac{d}{dt}\dot{x}^2 \\ x\dot{x} = x\frac{dx}{dt} &= \frac{1}{2}\frac{d}{dt}x^2\end{aligned}$$

Thus,

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2\right) &= -\frac{d}{dt}\left(\frac{1}{2}kx^2\right) \\ \frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2+\frac{1}{2}kx^2\right) &= 0\end{aligned}$$

Let $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$, we solve for t as:

$$\begin{aligned}\sqrt{\frac{2E-kx^2}{m}} &= \frac{dx}{dt} \\ dt &= \sqrt{\frac{m}{2E-kx^2}}dx \\ \int dt &= \int\sqrt{\frac{m}{2E-kx^2}}dx \\ t &= \int\sqrt{\frac{m}{2E-kx^2}}dx\end{aligned}$$

To solve the integral above, we make the following substitution:

$$x \equiv \sqrt{\frac{2E}{k}}\sin(\phi)$$

which implies:

$$dx = \sqrt{\frac{2E}{k}}\cos(\phi)d\phi$$

and the integral now becomes:

$$t = \sqrt{\frac{m}{k}}(\phi + \phi_0)$$

$$\phi = \sqrt{\frac{k}{m}}t - \phi_0$$

and

$$\begin{aligned} x &= \sqrt{\frac{2E}{k}}\sin\left(\sqrt{\frac{k}{m}}t - \phi_0\right) \\ &= \sqrt{\frac{2E}{k}}\left(\sin(\sqrt{\frac{k}{m}}t)\cos(\phi_0) - \cos(\sqrt{\frac{k}{m}}t)\sin(\phi_0)\right) \\ &\equiv B\sin(\sqrt{\frac{k}{m}}t) + A\cos(\sqrt{\frac{k}{m}}t) \\ &\equiv B\sin(\omega t) + A\cos(\omega t) \end{aligned}$$

Let's solve the equation of motion using a Canonical transformation. First of all, we use the following generating function:

$$F(q, Q) = -\frac{1}{2}m\omega q^2\cot(Q) \tag{1.10.3}$$

where $\omega \equiv \sqrt{\frac{k}{m}}$ and $\cot(x)$ denotes the cotangent function: $\cot(x) \equiv \frac{1}{\tan(x)} = \frac{\cos(x)}{\sin(x)}$. Using Eq.1.9.11, we have:

$$\begin{aligned} P &= \frac{\partial F}{\partial Q} \\ &= \frac{\partial\left(-\frac{1}{2}m\omega q^2\cot(Q)\right)}{\partial Q} \\ &= \frac{m\omega q^2}{2\sin^2(Q)} \end{aligned} \tag{1.10.4}$$

$$\begin{aligned} p &= -\frac{\partial F}{\partial q} \\ &= -\frac{\partial\left(-\frac{1}{2}m\omega q^2\cot(Q)\right)}{\partial q} \\ &= m\omega q\cot(Q) \end{aligned} \tag{1.10.5}$$

These two formulas implicitly define the new coordinates Q, P in terms of the old ones q, p. **Our next task is to use these formulas to rewrite the Hamiltonian.** From the equation above:

$$\sqrt{\frac{2}{m\omega}} P \sin(Q) = q \tag{1.10.6}$$

$$p = \sqrt{2m\omega P}\cos(Q) \tag{1.10.7}$$

and the Hamiltonian becomes:

$$\begin{aligned} H &= \frac{1}{2m}\left(2m\omega P\cos^2(Q) + 2m\omega P\sin^2(Q)\right) \\ &= P\omega\left(\cos^2(Q) + \sin^2(Q)\right) \\ &= P\omega \equiv H(Q,P) \end{aligned}$$

Using Hamilton's equation with new coordinates, we have:

$$\begin{aligned} \frac{dP}{dt} &= -\frac{\partial \hat{H}}{\partial Q} \\ &= -\frac{\partial(P\omega)}{\partial Q} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{dQ}{dt} &= \frac{\partial \tilde{H}}{\partial P} \\ &= \frac{\partial(P\omega)}{\partial P} \\ &= \omega \end{aligned}$$

Thus, $Q = \omega t + Q_0$.

1.11 The Pendulum

A pendulum consists of a bob of mass m which is suspended from some fixed ceiling by a string of length l. We want to describe how the bob swings back and forth under the influence of the earth's gravitational field. For a pendulum, if we use Cartesian coordinates to describe the position of the bob, we have the following constraint:

$$x^2 + y^2 = l^2$$

In Newtonian formalism, we have the equation of motion as:

$$\frac{d}{dt}\begin{pmatrix} m\dot{x} \\ m\dot{y} \end{pmatrix} = \begin{pmatrix} 0 \\ mg \end{pmatrix} + \begin{pmatrix} -t\sin(\phi) \\ -t\cos(\phi) \end{pmatrix}$$

$$\begin{pmatrix} m\ddot{x} \\ m\ddot{y} \end{pmatrix} = \begin{pmatrix} 0 \\ mg \end{pmatrix} + \begin{pmatrix} -t\sin(\phi) \\ -t\cos(\phi) \end{pmatrix}$$

where t is the magnitude of the tension force which we need to figure out. Let $x = l\sin(\phi)$ and $y = l\cos(\phi)$, we have the following two equations of motion:

$$\begin{aligned} \ddot{\phi} &= -\frac{g}{l}\sin(\phi) \\ t &= ml\dot{\phi}^2 + mg\cos(\phi) \end{aligned} \tag{1.11.1}$$

In Lagrangian formalism, we add the Lagrange multiplier term in the Lagrangian:

$$\begin{aligned} L_{\text{pendulum}} &= L_{\text{free}} + L_{\text{con}} \\ &= \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) + mgy + \frac{1}{2}\lambda\left(x^2 + y^2 - l^2\right) \end{aligned} \tag{1.11.2}$$

For $q = x$, we use the Euler-Lagrange equation and find:

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)$$

$$\lambda x = m\ddot{x}$$

For $q = y$, we find:

$$mg + \lambda y = m\ddot{y}$$

In fact, we've seen in the Newtonian formalism that when we switch from our Cartesian coordinates (x, y) to a description in terms of the angle ϕ, the constraint becomes trivially true. In terms of ϕ, the full pendulum Lagrangian therefore reads:

$$L = \frac{1}{2}ml^2\phi^2 + mgl\cos(\phi)$$

Because Euler-Lagrange equation is always valid for any choice of coordinates, we have:

$$\frac{\partial L}{\partial \phi} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right)$$

$$-mg\sin(\phi) = ml\ddot{\phi}$$

In Hamiltonian formalism, we again first calculate the generalized momentum p as:

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{\phi}} \\ &= \frac{\partial\left(\frac{1}{2}ml^2\dot{\phi}^2 + mgl\cos(\phi)\right)}{\partial\dot{\phi}} \\ &= ml^2\dot{\phi} \end{aligned}$$

and the Hamiltonian:

$$H = p\dot{\phi} - L = \frac{1}{2}\frac{p^2}{ml^2} - mgl\cos(\phi)$$

Then, from Hamilton's equations:

$$\begin{aligned} \frac{dp}{dt} &= -\frac{\partial H}{\partial \phi} \\ &= -\frac{\partial\left(\frac{1}{2}\frac{p^2}{ml^2} - mgl\cos(\phi)\right)}{\partial\phi} \\ &= -mgl\sin(\phi) \end{aligned}$$

$$\begin{aligned} \frac{d\phi}{dt} &= \frac{\partial H}{\partial p} \\ &= \frac{\partial\left(\frac{1}{2}\frac{p^2}{ml^2} - mgl\cos(\phi)\right)}{\partial p} \\ &= \frac{p}{ml^2} \end{aligned}$$

Again, we arrive at the equation of motion:

$$\frac{d^2\phi}{dt^2} = -\frac{g}{l}\sin(\phi)$$

1.12 Noether's Theorem

In a nutshell, **Noether's theorem tells us that for each symmetry of a given system, there must be a conserved quantity**

In particular:

- If the system possesses rotational symmetry, we know immediately that angular momentum is conserved. In other words, this means that if we can rotate our system without changing anything, angular momentum is conserved.
- If the system is symmetric with respect to spatial translations $x \to x+\epsilon$, we know immediately that momentum is conserved. This means that if we can change the position of the whole system and nothing changes, momentum is conserved.
- If the system is symmetric with respect to temporal translations $t \to t+\epsilon$, we know immediately that energy is conserved. Formulated differently, if the system behaved yesterday exactly as it does today, energy is conserved.

1.12.1 In Hamiltonian Formalism

Hamilton's equations are always valid before and after the canonical transformation. **When it comes to Noether's theorem, we only care about a symmetry which does not change the formation of** H:

$$H(Q,P) \overset{!}{=} H(Q,P) \tag{1.12.1}$$

where $\tilde{H}(Q,P) = H(q(Q,P),p(Q,P))$ because only then will **Hamilton's equations not only have the same form but are actually equivalent before and after the transformation.**

Let's consider the following example:

$$H = T + U = \frac{1}{2}m\dot{x}^2 + ax^2$$

If we make the transformation:$x \to X = x - s$, we have

$$\tilde{H} = H(x(X)) = \frac{1}{2}m\dot{X}^2 + a(X+s)^2$$

This is not equal to $H(X) = \frac{1}{2}m\dot{X}^2 + aX^2$ and therefore the shift $x \to X = x - s$ is not a symmetry.

In contrast, if there is no potential:

$$H = T = \frac{1}{2}m\dot{x}^2$$

$$\tilde{H} = H(x(X)) = T = \frac{1}{2}m\dot{X}^2$$

This is equal to $H(X) = \frac{1}{2}mX^2$ and therefore the shift $x \to X = x - s$ is indeed a symmetry.

Next, recall that we describe infinitesimal canonical transformations using so-called generators $G(q, p)$:

$$q \to Q = q + \epsilon\frac{\partial G}{\partial p}$$
$$p \to P = p - \epsilon\frac{\partial G}{\partial q}$$

Now the invariance condition (1.12.1) becomes:

$$H(q(Q,P), p(Q,P)) \stackrel{!}{=} H\left(q + \epsilon\frac{\partial G}{\partial p}, p - \epsilon\frac{\partial G}{\partial q}\right)$$

$$H(q(Q,P), p(Q,P)) \stackrel{!}{=} H(q,p) + \epsilon\frac{\partial H}{\partial q}\frac{\partial G}{\partial p} - \epsilon\frac{\partial H}{\partial p}\frac{\partial G}{\partial q} + \dots$$

$$\begin{aligned} 0 &\stackrel{!}{=} \epsilon\tfrac{\partial H}{\partial q}\tfrac{\partial G}{\partial p} - \epsilon\tfrac{\partial H}{\partial p}\tfrac{\partial G}{\partial q} + \dots \\ 0 &\stackrel{!}{=} \{H, G\} \end{aligned} \tag{1.12.2}$$

This is how we can check if the transformation generated by G is a symmetry.

Now, since the time-evolution of a generator $G(q, p)$ is given by:

$$\frac{d}{dt}G = \{G, H\}$$

According to the feature of Poisson bracket, we also have

$$\{A, B\} = -\{B, A\} \tag{1.12.3}$$

we obtain

$$\frac{d}{dt}G = \{G, H\} = 0 \tag{1.12.4}$$

We therefore learn that if G generates a symmetry, it automatically describes a conserved quantity. This is Noether's theorem.

1.12.2 Noether's extended theorem

We have learned that gauge transformation have no influence on Hamilton's equations and are therefore always sysmmetries. Therefore, it seems reasonable to relax the condition (1.12.1) to:

$$\tilde{H}(Q,P) \stackrel{!}{=} H(Q,P) - \frac{dF}{dt} \tag{1.12.5}$$

and

$$0 \stackrel{!}{=} -\{G,H\} - \frac{dF}{dt} \tag{1.12.6}$$

This means that even if $\{H,G\}$ is non-zero, **the transformation generated by** G **can be a symmetry as long as** $\{H,G\}$ **is equal to the total derivative of an arbitrary function** $F = F(Q,t)$

Now let's assume that we have found a function G so that

$$0 = \{G,H\} + \frac{dF}{dt} \tag{1.12.7}$$

Since $\frac{d}{dt}G = \{G,H\}$ we obtain:

$$\frac{d}{dt}Q = \frac{d}{dt}(G+F) = 0 \tag{1.12.8}$$

This is the extended Noether theorem. This means that the corresponding conserved quantity is described solely by the function F if we are dealing with a pure gauge transformation.

Accordingly, we have **Noether's Converse Theorem**:

For each symmetry, we can find a conserved quantity.

1.12.3 In Lagrangian Formalism

Let a Lagrangian transformed through:

$$\begin{aligned} q \to q' &= q + \epsilon g \\ \dot{q} \to \dot{q}' &= \dot{q} + \epsilon \dot{g} \end{aligned} \tag{1.12.9}$$

where $g = g(q)$ is a configuration space function that describes the infinitesimal transformation in question. So that

$$\begin{aligned} L\left(q\left(q'\right), \dot{q}\left(q', \dot{q}'\right)\right) &= L(q + \epsilon g, \dot{q} + \epsilon \dot{g}) \\ L(q, q) &\overset{!}{=} L(q, q) + \epsilon \tfrac{\partial L}{\partial q} g + \epsilon \tfrac{\partial L}{\partial \dot{q}} g \\ 0 &= \epsilon \tfrac{\partial L}{\partial q} g + \epsilon \tfrac{\partial L}{\partial q} g \\ 0 &= \epsilon \frac{\partial L}{\partial q} g + \epsilon \frac{\partial L}{\partial q} g \\ 0 &= \epsilon \frac{d}{dt}\left(\frac{\partial L}{\partial q}\right) g + \epsilon \frac{\partial L}{\partial q} g \\ 0 &= \frac{d}{dt}\left(\frac{\partial L}{\partial q} g\right) \end{aligned}$$

This tells us that if the transformation is indeed a symmetry, the function

$$Q_0 \equiv \frac{\partial L}{\partial \dot{q}} g \tag{1.12.10}$$

describes a conserved quantity. Q_0 **is usually called as Noether charge.**

Generalized Noether Theorem that if our Lagrangian changes at most by a total derivative of some function $F = F(q, t)$,

$$L \to L' = L - \frac{dF(q,t)}{dt}$$

we get a conserved quantity

$$Q \equiv \frac{\partial L}{\partial \dot{q}} g - \frac{1}{\epsilon} F = Q_0 - \frac{1}{\epsilon} F \tag{1.12.11}$$

In this context it is conventional to call Q the full Noether charge and Q_0 the bare Noether charge.

1.13 Additional Formulations of Classical Mechanics

1.13.1 Hamilton-Jacobi Mechanics

While Hamilton-Jacobi mechanics is a useful tool for some applications (especially chaotic systems and geometrical optics), from a more fundamental perspective we are interested in it **because it provides a bridge to quantum mechanics.**

If we use a canonical transformation generated by some function $F = F(q, Q, t)$, the Hamiltonian becomes:

$$H \to H' = H + \frac{\partial}{\partial t} F \tag{1.13.1}$$

where we simply change the minus sign in Eq. (1.9.19).

But this means that **we can try to find a Hamilton's principal function** $W = W(q, Q, t)$ **for which the Hamiltonian becomes zero:**

$$H \to H' = 0 = H + \frac{\partial}{\partial t} W \tag{1.13.2}$$

and

$$\begin{aligned} \frac{dp}{dt} &= -\frac{\partial H}{\partial q} \to \frac{dP}{dt} = -\frac{\partial H'}{\partial Q} = 0 \\ \frac{dq}{dt} &= \frac{\partial H}{\partial p} \to \frac{dQ}{dt} = \frac{\partial H'}{\partial P} = 0 \end{aligned} \tag{1.13.3}$$

This implies that P and Q are constant.Because these constants will be important, we give them concrete names:

$$Q = \alpha \quad P = \beta \tag{1.13.4}$$

Once we find appropriate $W(q, Q, t)$, we can write down the new coordinates directly by using

$$\begin{aligned} P &= -\frac{\partial W}{\partial Q} \\ p &= \frac{\partial W}{\partial q} \end{aligned} \tag{1.13.5}$$

We use Hamilton-Jacobi equation to find such W:

$$0 = H\left(q, \frac{\partial W}{\partial q}, t\right) + \frac{\partial W}{\partial t} \tag{1.13.6}$$

Taking a total time derivative of $W(q, \alpha, t)$:

$$\begin{aligned} \frac{dW}{dt} &= \frac{\partial W}{\partial q}\frac{dq}{dt} + \frac{\partial W}{\partial t} \\ &= \frac{\partial W}{\partial q}\frac{dq}{dt} - H \\ &= p\frac{dq}{dt} - H \end{aligned} \tag{1.13.7}$$

And therefore

$$\frac{dW}{dt} = L \tag{1.13.8}$$

$$W = \int L dt \tag{1.13.9}$$

Comparing to the definition of action, for Hamilton's principal function W the boundary condition are not fixed.

1.13.2 Statistical Mechanics

Using the concept of probability density, the probability of finding our objects within a region R in the phase space is

$$P(R, t) = \int_R \rho(x, p, t) dV \tag{1.13.10}$$

where $dV = dpdq$.

we have $N_R(t_0)$ microstates within the region R

$$N_R(t_0) = \int_R \tilde{\rho}(q, p, t_0)\, dV \tag{1.13.11}$$

After some short time period,

$$N_R(t_0+\epsilon) = \int_R \tilde{\rho}(q,p,t_0+\epsilon)\, dV \tag{1.13.12}$$

If we imagine that all phase space point are constantly moving around as described by Hamilton's equations. This implies that pur region R is moved around, too

$$R \to \tilde{R}$$

and

$$N_R(t_0+\epsilon) = \int_R \tilde{\rho}(q,p,t_0+\epsilon)\, dV$$

which leads to a key observation:

$$N_R(t_0) = N_R(t_0+\epsilon) \tag{1.13.13}$$

As time passes, all these points get "dragged" around by the Hamiltonian. Analogously, all points outside the region R get dragged around. But no new point will ever enter our moving region. Thus,

The paths of different phase space points never intersect each other.

Following this insight, we know that

$$\frac{d\rho(q,p,t)}{dt} = 0 \tag{1.13.14}$$

which leads to

Liouville equation

$$\frac{\partial \rho}{\partial t} = -\{\rho, H\} \tag{1.13.15}$$

Using Liouville equation, we can derive the Liouville's theorem

Phase space volume remains constant in time.

1.13.3 Koopmann-von Neumann Mechanics

Hilbert Space

Question

Why should we add one axis for each possible state a given object can be in?

Answer

Because of principle of uncertainty, we don't know the position and momentum of a particular object with 100% certainty. Thus, we need to introduce an individual axis for each possible state. We can then describe how likely a given state is by using a particular point on the corresponding axis.

For simplicity, let's assume that our object can only be in four different states $(q_1, p_1), (q_2, p_2)$ $(q_3, p_3)(q_4, p_4)$. Mathematically, we describe these four possible states using four different basis vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{e}_4$ because we now introduce an individual axis for each possible state.

If each of the four states is equally likely, we can then describe the situation by the vector:

$$\Psi = \frac{1}{2}\hat{e}_1 + \frac{1}{2}\hat{e}_2 + \frac{1}{2}\hat{e}_3 + \frac{1}{2}\hat{e}_4$$

The number in front of each basis vector is related to the possibility of finding our object occupying that state as:

The absolute value squared of the number yields the probability.

Koopman-von Neumann Equation

There are two key ideas which allow us to describe classical mechanics in Hilbert space:

- The probability density $\rho(x,p,t)$ is related to the state vectors Ψ $\mathbb{V}$ by

$$\rho(x,p,t) = |\Psi(x,p,t)|^2$$

If the integral form of state vector is used

$$\Psi(x,p,t) = \int c(x,p,t)\hat{e}_{x,p}dxdp$$

we have

$$\rho(x,p,t) = |\Psi(x,p,t)|^2 = |c(x,p,t)|^2$$

The integral vanishes because basis vectors are orthogonal.

- The time-evolution of our state vectors is described by the equation

$$i\frac{\partial}{\partial t}\Psi(x,p,t) = \hat{L}\Psi(x,p,t) \tag{1.13.16}$$

where $\hat{L}$ is the so-called Liouville operator:

$$L \equiv i\left(-\frac{\partial H}{\partial p}\frac{\partial}{\partial x} + \frac{\partial H}{\partial x}\frac{\partial}{\partial p}\right)$$

We call Eq.(1.13.16) the Koopman-von Neumann equation.

We now derive Liouville equation from Koopman-von Neumann equation:

We first multiply Eq.(1.13.16) by the conjugate state vector

$$\Psi^\star\frac{\partial}{\partial t}\Psi = -\Psi^\star\frac{\partial H}{\partial p}\frac{\partial \Psi}{\partial x} + \Psi^\star\frac{\partial H}{\partial x}\frac{\partial \Psi}{\partial p} \tag{1.13.17}$$

Secondly, we write similar equation for conjugate state vector:

$$\Psi\frac{\partial}{\partial t}\Psi^\star = -\Psi\frac{\partial H}{\partial p}\frac{\partial \Psi^\star}{\partial x} + \Psi\frac{\partial H}{\partial x}\frac{\partial \Psi^\star}{\partial p} \tag{1.13.18}$$

The sum of the equations above gives

$$\Psi^* \frac{\partial}{\partial t}\Psi + \Psi \frac{\partial}{\partial t}\Psi^* = -\Psi^* \frac{\partial H}{\partial p}\frac{\partial \Psi}{\partial x} + \Psi^* \frac{\partial H}{\partial x}\frac{\partial \Psi}{\partial p} - \Psi \frac{\partial H}{\partial p}\frac{\partial \Psi^*}{\partial x} + \Psi \frac{\partial H}{\partial x}\frac{\partial \Psi^*}{\partial p}$$
$$\frac{\partial}{\partial t}\left(\Psi^*\Psi\right) = -\frac{\partial H}{\partial p}\frac{\partial(\Psi^*\Psi)}{\partial x} + \frac{\partial H}{\partial x}\frac{\partial(\Psi^*\Psi)}{\partial p}$$
$$\frac{\partial}{\partial t}(\rho) = -\frac{\partial H}{\partial p}\frac{\partial(\rho)}{\partial x} + \frac{\partial H}{\partial x}\frac{\partial(\rho)}{\partial p}$$
$$\frac{\partial}{\partial t}\rho = \{H, \rho\}$$
$$\frac{\partial}{\partial t}\rho = -\{\rho, H\}$$

1.14 A Few more Words on Path Integral

And Feynmann says:

Every path between a fixed initial and final point in configuration space is possible in principle.

However, the probability of each path is a different one. Specifically, we can calculate a probability amplitude for each path $q(t)$

$$\Psi = \mathrm{e}^{iS[q(t)]/\hbar} \tag{1.14.1}$$

where $S[q(t)]$ is the action associated with the path, and $\hbar$ is a fundamental constant known as Planck's constant.

To calculate the total probability that our system evolves from a specific initial point A to a specific final point B we sum over the probability amplitudes for all possible paths:

$$\psi_{A\to B} = \sum_{i=1}^{N} \mathrm{e}^{iS[q_i(t)]/\hbar}$$

and

$$P(A \to B) = \left|\psi_{A\to B}\right|^2$$

Usually there are infinitely many paths possible, so we write the sum as a **path integral**:

$$\psi_{A\to B} = \int Dq(t) e^{iS[q(t)]/\hbar} \tag{1.14.2}$$

In general, since the action for a specific path ($S[q(t)]$) is an ordinary number, $e^{iS}[q(t)]$ is a complex number with absolute value 1. In the complex plane, these numbers lie as vectors on the unit circle. The contribution of each path to the total path integral is therefore simply a unit complex number. To calculate the path integral, we have to add the little arrows for each path like we would add vectors from complex plane. The total value of the path integral is then the resulting arrow.

The main point to take away, however, is that **each path between a fixed initial and final point is possible. Each path contributes exactly one little arrow (one probability amplitude) to the total path integral.**

Chapter 2

Quantum Mechaaaaanics

This is a cheat sheet for QM. The focus here is not giving comprehensive explanations but summarizing concise "quantum" bullet points.

2.1 State vectors and Their Conjugates

1. **Quantum operators**: quantities we can measure.

2. A state vector represents an eigenstate of a particular operator **if measuring the associated quantity always yields the same result**:

3. operator $\times$ eigenstate = eigenvalue $\times$ eigenstate

4. **state vector** is usually a linear combination of eigenstates:

$$|\Psi\rangle = \sum_i a_i |o_i\rangle = a_1 |o_1\rangle + a_2 |o_2\rangle + \dots$$

and probability to measure $o_i = |a_i|^2$.

5. A **conjugated state vectors ("bra")** is a Hermitian conjugated ket:

$$\langle\Psi| = |\Psi\rangle^\dagger = (|\Psi\rangle^\star)^T$$

6. The **expectation value of an observable is then**

$$\text{expectation value } = \langle\Psi|\hat{O}|\Psi\rangle \tag{2.1.1}$$

7. **Wave function:** function $\psi(x)$ that we get by expanding a state vector in terms of position eigenstates:

$$|\Psi\rangle = \int dx\psi(x)|x\rangle \tag{2.1.2}$$

8. **Schrödinger equation:**

$$i\hbar\partial_t|\Psi\rangle = -\frac{\hbar^2\partial_i^2}{2m}|\Psi\rangle + V(\hat{x})|\Psi\rangle \tag{2.1.3}$$

9. The frequency of the wave associated with a given particle is directly related to its energy as:

$$v = \frac{E}{h} \tag{2.1.4}$$

10. the wavelength is directly related to its momentum

$$\lambda = \frac{h}{p} \tag{2.1.5}$$

2.2 Quantum Framework

2.2.1 Wave Function

Use **location projector operators**, we formulate the **wave function** as:

$$\begin{aligned} |\Psi\rangle &= \int dx|x\rangle\langle x|\Psi\rangle \\ &\equiv \int dx\Psi(x)|x\rangle \end{aligned} \tag{2.2.1}$$

where $\Psi(x) \equiv \langle x|\Psi\rangle$

The physical interpretation of $\Psi(x)$ is again as a probability amplitude.

Using momentum basis states, we obtain the **momentum representation** of the wave function:

$$\begin{aligned} |\Psi\rangle &= \int dp|p\rangle\langle p|\Psi\rangle \\ &\equiv \int dp\Psi(p)|p\rangle \end{aligned} \tag{2.2.2}$$

The operators $\int dx|x\rangle\langle x|$, $\int dp|p\rangle\langle p|$ are the **identity operator**, i.e., operators that do not change anything.

An important idea is that by using a given wave function $\Psi(x)$ we can immediately calculate important quantities like the expectation value:

$$\begin{aligned} \langle\Psi|\hat{O}|\Psi\rangle &= \left\langle \Psi \left| \hat{O} \int dx\Psi(x) \right| x \right\rangle \\ &= \int dx' \left\langle x' \left| \Psi^{+}\left(x'\right) O \int dx\Psi(x) \right| x \right\rangle \\ &= \int dx' \int dx \left\langle x' \left| \Psi^{\dagger}\left(x'\right) \hat{O}\Psi(x) \right| x \right\rangle |x\rangle \\ &= \int dx' \int dx\Psi^{\dagger}\left(x'\right) O\Psi(x) \underbrace{\langle x'|x\rangle}_{=\delta(x-x')} \\ &= \int dx\Psi^{+}(x)O\Psi(x) \end{aligned}$$

and

$$\begin{aligned} \langle\Phi|\Psi\rangle &= \int dx' \left\langle x' \left| \Phi\left(x'\right) \int dx\Psi(x) \right| x \right\rangle \\ &= \int dx' \int dx \left\langle x' \left| \Phi^{\dagger}\left(x'\right) \Psi(x) \right| x \right\rangle \\ &= \int dx' \int dx\Phi^{\dagger}\left(x'\right) \Psi(x) \underbrace{\langle x'|x\rangle}_{=\delta(x-x')} \\ &= \int dx\Phi^{\dagger}(x)\Psi(x) \end{aligned}$$

2.2.2 Quantum Operators

We now talk more about the operators from a perspective of **symmetries**. The part of mathematics which deals with symmetries is called **group theory**. A group is a set of

transformations which fulfill special rules plus an operation that tells us how to combine the transformations. Also, **we only need one special part of group theory, namely the part that deals with continuous symmetries.**

There is one property that makes continuous symmetries especially nice to deal with:

they have elements which are arbitrarily close to the identity transformation.

For example, think about the symmetries of a circle. Any rotation about the origin is a symmetry of a circle. Therefore, a rotation extremely close to the identity transformation, say a rotation by $0.000001°$, is a symmetry of the circle.

Mathematically, we write an element g close to the identity I as:

$$g(\epsilon) = I + \epsilon G \tag{2.2.3}$$

where ϵ is a really, really small number and G is an object, called a generator. In the smallest possible case, such transformations are called **infinitesimal transformations**.

Let's return to our discussion about rotations. Many small rotations in one direction are equivalent to one big rotation in the same direction. Mathematically, we can write the idea of repeating a small transformation many times as follows:

$$h(\theta) = (I + \epsilon G)(I + \epsilon G)(I + \epsilon G)\ldots = (I + \epsilon G)^k \tag{2.2.4}$$

If θ denotes some finite transformation parameter, e.q.,50^o or so, and N is some huge number that makes sure we are close to the identity. We write g as:

$$g(\theta) = I + \frac{\theta}{N} G$$

At the limit of $N \to \infty$, we have

$$h(\theta) = \lim_{N\to\infty} \left(I + \frac{\theta}{N} G \right)^N = e^{\theta G}$$

The bottom line is that the object G generates the finite transformation h. This is why we call objects like this **generators**.

Let's consider a function $f(x,t)$ and assume that our goal is to generate a spatial translation such that $Tf(x,t) = f(x+a,t)$ The following generator do the job:

$$G_{\text{xtrans}} = \partial_x \tag{2.2.5}$$

and

$$\begin{aligned} e^{aG_{\text{xtrans}}} f(x,t) &= \left(1 + aG_{\text{xtrans}} + \frac{a^2}{2} G_{\text{xtrans}}^2 + \dots\right) f(x,t) \\ &= \left(1 + a\partial_x + \frac{a^2}{2}\partial_x^2 + \dots\right) f(x,t) \\ &= f(x+a,t) \end{aligned} \tag{2.2.6}$$

$G_{\text{xtrans}} = \partial_x$ **generates spatial translations.**

Analogously, $G_{\textbf{ttrans}} = \partial_t$ **generates temporal translation**.

The core of each continuous symmetry is the corresponding generator.

quantum operator $\leftrightarrow$ ***generator of symmetry***

1. momentum $\hat{p}_i \leftrightarrow$ generator of spatial translations $(-i\hbar\partial_i)$
2. energy $\hat{E} \leftrightarrow$ generator of temporal translations $(i\hbar\partial_t)$
3. Since there is no symmetry connected to the conservation of position, the position operator stays as $\hat{x}$

Canonical commutation relation reads:

$$[\hat{p}_i, \hat{x}_j] = -i\hbar\delta_{ij} \tag{2.2.7}$$

In physical terms, we cannot measure the position and the momentum of a particle at the same time with arbitrary precision. We can now also understand how this comes about in our framework. We identified the momentum operator as the generator of spatial translations.Thus, each time we measure the momentum, we perform a tiny spatial translation.

Hamiltonian operator

$$\hat{H} \equiv \left(-\frac{\hbar^2 \partial_i^2}{2m} + V(\hat{x}) \right) \tag{2.2.8}$$

Which leads to more general form of the Schrödinger equation:

$$i\hbar \partial_t |\Psi\rangle = \hat{H} |\Psi\rangle \tag{2.2.9}$$

which is valid even for relativistic systems and quantum field theories.

A convenient alternative way to describe the time-evolution of quantum systems is with the so-called **time evolution operator** $U(t)$:

$$|\Psi(x,t)\rangle = U(t)|\Psi(x,0)\rangle \tag{2.2.10}$$

Since

$$i\hbar \partial_t |\Psi(x,t)\rangle = H|\Psi(x,t)\rangle$$
$$i\hbar \partial_t U(t)|\Psi(x,0)\rangle = HU(t)|\Psi(x,0)\rangle$$

This equation holds for any $|\Psi(x,0)\rangle$ and we can therefore write it without it:

$$\therefore \quad i\hbar \partial_t U(t) = HU(t)$$
$$\therefore \quad i\hbar \frac{\partial_t U(t)}{U(t)} = H$$

Hence

$$U(t) = \mathrm{e}^{-\frac{i}{\hbar}\int_0^t dt' H(t')} \tag{2.2.11}$$

Question

Why Quantum Mechanics is About Waves?

Answer

Let's consider an 1D case:

$$
\begin{aligned}
i\hbar\partial_t|\Psi\rangle &= -\frac{\hbar^2\partial_x^2}{2m}|\Psi\rangle \\
i\hbar\partial_t \int dx\psi(x,t)|x\rangle &= -\frac{\hbar^2\partial_x^2}{2m}\int dx\psi(x,t)|x\rangle \\
\therefore \quad i\hbar\partial_t\psi(x,t) &= -\frac{\hbar^2\partial_x^2}{2m}\psi(x,t)
\end{aligned}
$$

we integrated over. One solution to this equation is

$$\psi(x,t) = \mathrm{e}^{-i(Et-px)/\hbar} \tag{2.2.12}$$

A function of the form above is known as a **plane wave**.

An important observation: **the Schrödinger equation is linear in** ψ. This means that we can use the **superposition principle**.

$$\psi_{\text{sup}} = a\psi_1 + b\psi_2 + \ldots \tag{2.2.13}$$

This observation allow us to construct **wave packets** through suitable linear combinations of plane waves. A wave packet is what we use to describe a particle which is localized within some region.

2.2.3 Angular Momentum

Using the explicit form of the angular momentum operators, we have:

$$
\begin{aligned}
\hat{L}_i &= \epsilon_{ijk}\hat{x}_j\hat{p}_k \\
&= \epsilon_{ijk}\hat{x}_j\left(-i\hbar\partial_k\right) \\
&= -i\hbar\epsilon_{ijk}\hat{x}_j\partial_k
\end{aligned}
$$

Angular momentum commutation relation

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\epsilon_{ijk}\hat{L}_k \tag{2.2.14}$$

The measurements of the angular momentum along different axis affect each other. In particular, a measurement of the angular momentum along the x-axis affects the angular momentum along the y- and z-axes.

2.2.4 Spin

Sometimes, we need to describe our quantum systems with objects that have more than one component:

$$\Psi(x,t) = \begin{pmatrix} \Psi_1(x,t) \\ \Psi_2(x,t) \end{pmatrix}$$

Acting on the object with symmetry generators will result in two things:

- Change of the spatial and temporal coordinates $x \to 1 + Gx$, where G denotes a generator. $\Psi(x) \to \Psi(x+\epsilon)$
- mixing the components

$$\begin{pmatrix} \Psi_1(x,t) \\ \Psi_2(x,t) \end{pmatrix} \to \begin{pmatrix} \Psi_2(x,t) \\ \Psi_1(x,t) \end{pmatrix}$$

Question

Now we have a generator composed of two parts, how can we construct a general quantum operator that describes angular momentum?

Answer

The generator that causes the rotation of the argument of a function is a **differential operator**. The naive operator $\vec{L} = \vec{x} \times (-i\hbar\vec{\partial})$ correctly describe the first change. To avoid confusion, this type of angular momentum is referred to as **orbital angular momentum**.

The generator that causes the mixing of the components is a **matrix**. And the operator identified with this generator describes **spin** or **internal angular momentum**. Specifically, the correct generators of spin for two-component states are:

$$\hat{S}_i = \frac{\hbar}{2}\sigma_i \tag{2.2.15}$$

where σ_i are (2×2) matrices known as **Pauli matrices**:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.2.16}$$

Here we use the labels 1, 2, and 3 instead of x, y, and z.

In contrast to the classical mechanics regime, under which angular momentum can assume any value, quantum mechanical spin can only take on one of two values. Either we measure the value $\hbar/2$, or we measure the value $-\hbar/2$.

Spin is quantized.

Using the explicit matrices for the spin operators, we can calculate the commutation relations as:

$$\left[\hat{S}_i, \hat{S}_j\right] = i\hbar\epsilon_{ijk}\hat{S}_k \tag{2.2.17}$$

2.2.5 Quantum Numbers

We specify the angular momentum of a given quantum system using the z-component of the angular momentum $\hat{L}_z$, and the total angular momentum, $\hat{L}^2$.The quantum operator defined for total angular momentum is

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{2.2.18}$$

and

$$\left[\hat{L}^2, \hat{L}_i\right] = 0 \tag{2.2.19}$$

It is conventional to use the label m for the angular momentum in the z-direction:

$$\hat{L}_z|m\rangle = \hbar m|m\rangle \tag{2.2.20}$$

and the label l for the total angular momentum:

$$\hat{L}^2|l\rangle = \hbar^2 l(l+1)|l\rangle \tag{2.2.21}$$

In addition, the energy operator $\hat{H}$ often commutes with $\hat{L}_z$ and L^2, so we can use it as a third label. The conventional label for the energy eigenvalues is n.

So in summary: We often label our states using the three labels m, l and n : $|n, m, l\rangle$. Labels like this are known as **quantum number**.

2.3 The Classical Limit

To connect QM to classical mechanics, we notice that **we can extend QM results to the classical mechanics through calculating the expectation values.**

The following statement for the momentum expectation value is exactly Newton's second law of classical mechanics:

Enrenfest's theorem

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi|\hat{p}|\Psi\rangle = -\left\langle\Psi\left|\partial_x V(\hat{x})\right|\Psi\right\rangle \tag{2.3.1}$$

Question

How to derive Enrenfest's theorem from QM principles?

Answer

We start with the expectation value for a general operator

$$\langle\Psi|\hat{O}|\Psi\rangle = \int d^3x \Psi^\star \hat{O}\Psi$$

and Schrödinger equation

$$\frac{d}{dt}\Psi = \frac{1}{i\hbar}H\Psi$$

and its conjugate

$$\begin{aligned}\frac{d}{dt}\Psi^+ &= -\frac{1}{i\hbar}\Psi^+ \overbrace{H^\dagger}^{H=H^+} \\ &= -\frac{1}{i\hbar}\Psi^\dagger H\end{aligned}$$

Taking the time derivative of the expectation value then yields

$$\begin{aligned}\frac{d}{dt}\langle\hat{O}\rangle &= \frac{d}{dt}\int d^3x \Psi^+\hat{O}\Psi \\ &= \int d^3x \left(\left(\frac{d}{dt}\Psi^+\right)\hat{O}\Psi + \Psi^+\left(\frac{d}{dt}\hat{O}\right)\Psi + \Psi^+\hat{O}\left(\frac{d}{dt}\Psi\right)\right)\end{aligned}$$

Next, we use $\frac{d}{dt}\hat{O} = 0$, which is correct for many operators. For example, for $\hat{O} = \hat{\vec{p}} = -i\hbar\vec{\nabla} \neq \hat{O}(t)$. This yields

$$\begin{aligned}\frac{d}{dt}\langle\hat{O}\rangle &= \int d^3x \left(\left(\frac{d}{dt}\Psi^+\right)\hat{O}\Psi + \underbrace{\Psi^+\left(\frac{d}{dt}\hat{O}\right)\Psi}_{=0} + \Psi^+\hat{O}\left(\frac{d}{dt}\Psi\right)\right) \\ &= \int d^3x \left(\left(-\frac{1}{i\hbar}\Psi^+ H\right)\hat{O}\Psi + \Psi^+\hat{O}\left(\frac{1}{i\hbar}H\Psi\right)\right) \\ &= \frac{1}{i\hbar}\int d^3x \left(-\Psi^\dagger H\hat{O}\Psi + \Psi^\dagger\hat{O}H\Psi\right) \\ &= \frac{1}{i\hbar}\int d^3x \Psi^\dagger[\hat{O}, H]\Psi \\ &= \frac{1}{i\hbar}\langle[\hat{O}, H]\rangle\end{aligned}$$

Answer

Now we can evaluate this equation specifically for the momentum operator and, in addition, use the explicit form of the Hamiltonian operator:

$$\begin{aligned}\frac{d}{dt}\langle\hat{p}\rangle &= \frac{1}{i\hbar}\langle[\hat{p},H]\rangle \\ &= \frac{1}{i\hbar}\left\langle\left[\hat{p},\frac{\hat{p}^2}{2m}+V\right]\right\rangle \\ &= \frac{1}{i\hbar}\langle\underbrace{\left[\hat{p},\frac{\hat{p}^2}{2m}\right]}_{=0}+[\hat{p},V]\rangle\end{aligned}$$

$$\begin{aligned}&= \frac{1}{i\hbar}\langle[\hat{p},V]\rangle \\ &= \frac{1}{i\hbar}\int d^3x\Psi^\dagger[\hat{p},V]\Psi \\ &= \frac{1}{i\hbar}\int d^3x\Psi^\dagger\hat{p}V\Psi - \frac{1}{i\hbar}\int d^3x\Psi^\dagger V\hat{p}\Psi \\ &= \frac{1}{i\hbar}\int d^3x\Psi^\dagger(-i\hbar\nabla)V\Psi - \frac{1}{i\hbar}\int d^3x\Psi^\dagger V(-i\hbar\nabla)\Psi \\ &= -\int d^3x\Psi^\dagger(\nabla V)\Psi - \int d^3x\Psi^\dagger V\nabla\Psi + \int d^3x\Psi^\dagger V\nabla\Psi \\ &= -\int d^3x\Psi^\dagger(\nabla V)\Psi \\ &= \langle-\nabla V\rangle = \langle F\rangle\end{aligned}$$

2.3.1 Classification of Solutions to QM systems

1. If the energy of the state is smaller than the potential in the system at infinity ($E < V(\infty)$ and $E < V(-\infty)$), we are dealing with a **bound state.** We label bound states using a discrete index n.

2. If the energy of the state is larger than the potential at infinity: $E > V(\infty)$ or $E > V(-\infty)$, we are dealing with a **scattering state**. We label scattering states using a continuous index k. A general solution is an integral $\int dk$ over such solutions.

2.4 Harmonic Quantum Mechanics

The potential of the harmonic oscillator is usually written as

$$V(x) = \frac{1}{2}m\omega^2 x^2$$

where $\omega = \sqrt{k/m}$ denotes the classical oscillation frequency and m the mass at the end of the spring. The stationary Schrödinger equation therefore reads

$$-\hbar^2 \frac{\partial_x^2}{2m}\psi + \frac{1}{2}m\omega^2 x^2 \psi = E\psi \tag{2.4.1}$$

The solutions to the Eq. (2.4.1) look like this:

$$\psi_n(x) = \left(\frac{1}{2}\right)^{n/2} H_n(\sqrt{\frac{m\omega}{\hbar}}x)e^{-\frac{m\omega}{2\hbar}x^2} \tag{2.4.2}$$

where H_n denotes the so-called Hermite polynomials

$$H_n(u) = (-1)^n e^{u^2/2}\frac{d}{du}e^{-u^2/2}$$

The corresponding energy eigenvalues are

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{2.4.3}$$

2.4.1 Lowering and Raising Operator

To solve Eq.(2.4.1) efficiently, we introduce the lowering and raising operators as:

$$a \equiv \sqrt{\frac{m\omega}{2\hbar}}x + i\frac{1}{\sqrt{2m\omega\hbar}}p$$
$$a^\dagger \equiv \sqrt{\frac{m\omega}{2\hbar}}x - i\frac{1}{\sqrt{2m\omega\hbar}}p \tag{2.4.4}$$

and

$$x = \sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right)$$
$$p = -i\sqrt{\frac{\hbar m\omega}{2}}\left(a - a^\dagger\right)$$

The commutator is simply 1:

$$\left[a, a^\dagger\right] = aa^\dagger - a^\dagger a = 1 \tag{2.4.5}$$

The Eq.(2.4.1) is then rewritten as:

$$E\Psi = \frac{1}{2m}\left(i\sqrt{\frac{\hbar m\omega}{2}}\left(a^\dagger - a\right)\right)^2\Psi + \frac{m\omega^2}{2}\left(\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right)\right)^2\Psi = \frac{\hbar\omega}{2}\left(a^\dagger a + aa^\dagger\right)\Psi \tag{2.4.6}$$

Using the commutator relation, we simply obtain that

$$\begin{aligned} E\Psi &= \frac{\hbar\omega}{2}\left(a^\dagger a + aa^\dagger\right)\Psi \\ &= \frac{\hbar\omega}{2}\left(a^\dagger a + aa^\dagger - a^\dagger a + a^\dagger a\right)\Psi \\ &= \frac{\hbar\omega}{2}\left(2a^\dagger a + \left[a, a^\dagger\right]\right)\Psi \\ &= \frac{\hbar\omega}{2}\left(2a^\dagger a + 1\right)\Psi \\ &= \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)\Psi \end{aligned}$$

Question

what a and $a^\dagger$ do when they act on our system $a\left|E_1\right\rangle =?$

Answer

First we calculate the commutator $[H, a]$:

$$\begin{aligned}
[H, a] &= Ha - aH \\
&= \left(\hbar\omega\left(a^\dagger a + \frac{1}{2}\right)\right) a - a\left(\hbar\omega\left(a^\dagger a + \frac{1}{2}\right)\right) \\
&= -\hbar\omega\left[a, a^\dagger\right] a \\
&= -\hbar\omega a
\end{aligned}$$

Similarly,

$$\left[H, a^\dagger\right] = \hbar\omega a^\dagger$$

With this information at hand, we are finally ready to calculate the energy of our new state $a\,|E_1\rangle$:

$$\begin{aligned}
\hat{H}\left(a\,|E_1\rangle\right) &= \left(Ha - a\hat{H} + a\hat{H}\right)|E_1\rangle \\
&= a\hat{H}\,|E_1\rangle + [\hat{H}, a]\,|E_1\rangle \\
&= aE_1\,|E_1\rangle + [\hat{H}, a]\Big) \\
&= (aE_1 - \hbar\omega a)\,|E_1\rangle \\
&= (E_1 - \hbar\omega)\,(a\,|E_1\rangle)
\end{aligned}$$

Analogously for $a^\dagger$ we find

$$\hat{H}a^\dagger\,|E_1\rangle = (E_1 + \hbar\omega)\,a^\dagger\,|E_1\rangle$$

a and $a^\dagger$ are known as **ladder operators**. They allow us to move between the energy eigenstates. Using $a^\dagger$ we can jump to the next higher eigenstate. Using a we can jump to the next lower eigenstate.

Question

How to normalize the states after ladder operators?

Answer

In general, we have something of the form:

$$a^\dagger|n\rangle = C|n+1\rangle$$

Now we solve the following expression using commutation relation:

$$\begin{aligned}
\langle n\,|aa^\dagger|\,n\rangle &= \langle n\,|\big(aa^\dagger - a^\dagger a + a^\dagger a\big)|\,n\rangle \\
&= \langle n\,|\big([a,a^\dagger] + a^\dagger a\big)|\,n\rangle \\
&= \langle n\,|\big(1 + a^\dagger a\big)|\,n\rangle \\
&= \langle n|(1+N)|n\rangle \\
&= \langle n|(N|n\rangle + 1|n\rangle) \qquad\qquad = \langle n|(n|n\rangle + 1|n\rangle) \\
&= \langle n|(n+1)|n\rangle \\
&= (n+1)\langle n|n\rangle \\
&= n+1
\end{aligned}$$

Also, we can use the conjugate states to solve the same expression as:

$$\begin{aligned}
\big(a^\dagger|n\rangle\big)^\dagger &= (C|n+1\rangle)^\dagger \\
\langle n\,|a &= \langle n+1|C^\dagger
\end{aligned}$$

$$\begin{aligned}
\langle n\,|aa^\dagger|\,n\rangle &= \langle n+1\,|C^\dagger C|\,n+1\rangle \\
&= C^\dagger C\langle n+1|n+1\rangle
\end{aligned}$$

Thus,

$$\begin{aligned}
C^\dagger C &= n+1 \\
C &= \sqrt{n+1}
\end{aligned}$$

In summary,

$$a|n\rangle = \sqrt{n}|n-1\rangle \tag{2.4.7}$$

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{2.4.8}$$

2.5 Quantum Systems with Spin

2.5.1 Spin Measurements

The eigenstates of the S_z operator

$$S_z = \frac{\hbar}{2}\sigma_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are

$$|\hbar/2\rangle_z \triangleq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-\hbar/2\rangle_z \triangleq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.5.1}$$

To find the probability to get the result $-\hbar/2$ for a spin measurement along the z -axis of a system described by a general ket $|X\rangle = a|\hbar/2\rangle_z + b|-\hbar/2\rangle_z$, we have to calculate

$${}_z\langle -\hbar/2|X\rangle = a\underbrace{\langle -\hbar/2|\hbar/2\rangle}_{=0} + \underbrace{b_z\langle -\hbar/2|-\hbar/2\rangle_z}_{=1} = b \tag{2.5.2}$$

If we want to measure the spin along another axis, say the x-axis, we first need to expand our state in terms of the eigenstates of $\hat{S}_x$:

$$S_x = \begin{pmatrix} 0 & \hbar/2 \\ \hbar/2 & 0 \end{pmatrix}$$

The corrsponding normalized eigenvectors are

$$|\hbar/2\rangle_x \triangleq \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |-\hbar/2\rangle_x \hat{=} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

So if we want to calculate the probability to measure $-\hbar/2$ for the spin in the x -direction, we first need to rewrite $|\hbar/2\rangle_z$ and $|-\hbar/2\rangle_z$ in terms of $|\hbar/2\rangle_x$ and $|-\hbar/2\rangle_x$:

$$|\hbar/2\rangle_z = \frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x + |-\hbar/2\rangle_x\right)$$

and

$$|-\hbar/2\rangle_z = \frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x - |-\hbar/2\rangle_x\right)$$

Our general state reads in this new basis

$$\begin{aligned}|X\rangle &= a|\hbar/2\rangle_z + b|-\hbar/2\rangle_z \\ &= a\frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x + |-\hbar/2\rangle_x\right) + b\frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x - |-\hbar/2\rangle_x\right)\end{aligned}$$

The probability amplitude to measure $-\hbar/2$ for the spin along the x -axis is therefore

$$\begin{aligned}{}_x\langle -\hbar/2|X\rangle =& {}_x\left\langle -\hbar/2\right|\Bigg(a\frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x + |-\hbar/2\rangle_x\right) \\ &+b\frac{1}{\sqrt{2}}\left(|\hbar/2\rangle_x - |-\hbar/2\rangle_x\right)\Bigg) \\ =&\frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}}\end{aligned}$$

The probability is

$$P_{x=-h/2} = \left|\frac{a}{\sqrt{2}} - \frac{b}{\sqrt{2}}\right|^2$$

2.5.2 Spin Addition

For concreteness, let's consider two particles with spin 1/2. For such spin 1/2 particles there are always only two possible spin alignments: spin up $|\uparrow\rangle$ and $|\downarrow\rangle$.**The total spin operator acting on this combined system to yield:**

$$\hat{S}\psi_1\psi_2 = \sqrt{s(s+1)}\hbar\psi_1\psi_2 \tag{2.5.3}$$

where s is the total spin number, which is $1/2 + 1/2 = 1$ in this case.

Similarly,

$$\begin{aligned}S_z\psi_1\psi_2 &= \left(S_z^{(1)} + S_z^{(2)}\right)\psi_1\psi_2 = \left(S_z^{(1)}\psi_1\right)\psi_2 + \psi_1\left(S_z^{(2)}\psi_2\right) \\ &= \left(\hbar m_1\psi_1\right)\psi_2 + \psi_1\left(\hbar m_2\psi_2\right) = \hbar\left(m_1 + m_2\right)\psi_1\psi_2\end{aligned}$$

The possible values for the overall spin in the z-direction $m = (m_1 + m_2)$ are therefore

$\uparrow\uparrow$: $m = 1$

$\uparrow\downarrow$: $m = 0$

$\downarrow\uparrow$: $m = 0$

$\downarrow\downarrow$: $m = -1$

The physical interpretation of this observation is that the two spin 1/2 particles can form together a system with a total spin of 1 or a total spin of 0. For the total spin 1 case, we can measure for the z-components the values $-1, 0$ or 1. We say the state is in a **triplet state** and denote it by:

$$\begin{aligned} |11\rangle &= \uparrow\uparrow \\ |10\rangle &= \frac{1}{\sqrt{2}} \uparrow\downarrow + \frac{1}{\sqrt{2}} \downarrow\uparrow \\ |1-1\rangle &= \downarrow\downarrow \end{aligned}$$

The other possibility is that the complete system has a total spin of zero. In this case, we say the system is in a **singlet state**. In terms of the individual spins this state reads:

$$|00\rangle = \frac{1}{\sqrt{2}} \uparrow\downarrow - \frac{1}{\sqrt{2}} \downarrow\uparrow$$

So the only difference to the $|10\rangle$ state is a minus sign, which however is crucial if we determine the energy levels, for example, in the Hydrogen atom. These numbers $\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}$ etc. are known as **Clebsch-Gordan coefficients**. They tell us how a system looks like in terms of the individual angular momentums that it consists of.

2.6 Perturbation Theory

In general, the situation we are now interested in looks like this

$$H = H_0 + \lambda V$$

where H is the Hamiltonian for the full system, H_0 the Hamiltonian of a system that we have already solved and V is the difference between the two. Here, λ is a parameter that we introduce to keep track of the order of perturbation theory.

Our goal is to calculate the energy levels E_n of this system which correspond to the eigenvalues of H :

$$\begin{aligned} H|n\rangle &= E_n|n\rangle \\ \therefore \quad (H_0 + \lambda V)\,|n\rangle &= E_n|n\rangle \end{aligned}$$

We make the following ansatz for the states:

$$|n\rangle = |n\rangle_0 + \lambda|n\rangle_1 + \lambda^2|n\rangle_2 + \dots \tag{2.6.1}$$

Next, we simply put these ansätze into our Schrödinger equation:

$$(H_0 + \lambda V)\left(|n\rangle_0 + \lambda|n\rangle_1 + \dots\right) = E_n\left(|n\rangle_0 + \lambda|n\rangle_1 + \dots\right)$$
$$(H_0 + \lambda V)\left(|n\rangle_0 + \lambda|n\rangle_1 + \dots\right) = \left(E_n^{(0)} + \lambda E_n^{(1)} + \dots\right)\left(|n\rangle_0 + \lambda|n\rangle_1 + \dots\right)$$

$$H_0|n\rangle_0 + \lambda V|n\rangle_0 + \lambda H_0|n\rangle_1 + \lambda^2 V|n\rangle_1 = E_n^{(0)}|n\rangle_0 + \lambda E_n^{(1)}|n\rangle_0 + \lambda E_n^{(0)}|n\rangle_1 + \lambda^2 E_n^{(1)}|n\rangle_1 + \dots$$

Then we collect all terms with λ in front of them:

$$V|n\rangle_0 + H_0|n\rangle_1 = E_n^{(1)}|n\rangle_0 + E_n^{(0)}|n\rangle_1 \tag{2.6.2}$$

Multiply the equation above by ${}_0\langle n|$:

$${}_0\langle n|V|n\rangle_0 + {}_0\langle n|H_0|n\rangle_1 = {}_0\left\langle n\left|E_n^{(1)}\right|n\right\rangle_0 + {}_0\left\langle n\left|E_n^{(0)}\right|n\right\rangle_1$$
$${}_0\langle n|V|n\rangle_0 + {}_0\left\langle n\left|E_n^{(0)}\right|n\right\rangle_1 = {}_0\left\langle n\left|E_n^{(1)}\right|n\right\rangle_0 + {}_0\left\langle n\left|E_n^{(0)}\right|n\right\rangle_1$$

and

$${}_0\langle n|V|n\rangle_0 = E_n^{(1)} \tag{2.6.3}$$

Following analogous steps

$$E_n^{(2)} = \sum_{m\neq n} \frac{|{}_0\langle m|V|n\rangle_0|^2}{E_n^{(0)} - E_m^{(0)}} \tag{2.6.4}$$

Following similar steps we can also calculate how the kets $|n\rangle_0$ change in the presence of the perturbations [12], i.e., the corrections $|n\rangle_{16}|n\rangle_{\jmath\ell}$ etc.

Again,we collect all terms with λ in front of them:

$$V|n\rangle_0 + H_0|n\rangle_1 = E_n^{(1)}|n\rangle_0 + E_n^{(0)}|n\rangle_1$$
$$\left(V - E_n^{(1)}\right)|n\rangle_0 = -\left(H_0 - E_n^{(0)}\right)|n\rangle_1$$

Using the following ansatz:

$$|n\rangle_1 = \sum_m c_{nm}|m\rangle_0$$

we have

$$\left(V - E_n^{(1)}\right)|n\rangle_0 = -\sum_m c_{nm}\left(E_m^{(0)}|m\rangle_0 - E_n^{(0)}|m\rangle_0\right)$$

To isolate the coefficient c_{nm} we multiply this equation by ${}_0\langle l|$ and get"

$$\frac{{}_0\left\langle l\left|\left(V - E_n^{(1)}\right)\right|n\right\rangle_0}{\left(E_n^{(0)} - E_l^{(0)}\right)} = c_{nl} \tag{2.6.5}$$

2.7 Beyond Wave Equation

2.7.1 Bohmian mechanics

The starting point is the following ansatz for a general wave function:

$$\Psi(\mathbf{r},t) = \sqrt{\rho(\mathbf{r},t)}e^{iS(\mathbf{r},t)/\hbar} \tag{2.7.1}$$

where $\rho = \Psi\Psi^*$ is the usual probability density and $S(\mathrm{r},t)$ the phase of the wave function.

putting this ansatz into the Schrödinger equation yields

$$\begin{aligned} i\hbar\partial_t\Psi &= \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\Psi \\ \frac{\partial\rho}{\partial t} + \nabla\left(\rho\frac{\nabla S}{m}\right) &= 0 \\ \frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + Q &= 0 \end{aligned}$$

where

$$\begin{aligned} Q &= -\frac{\hbar^2}{8m}\left[2\left(\frac{\nabla^2\rho}{\rho}\right) - \left(\frac{\nabla\rho}{\rho}\right)^2\right] \\ &= -\frac{\hbar^2}{2m}\left\{\mathrm{Re}\left(\frac{\nabla^2\Psi}{\Psi}\right) + \left[\mathrm{Im}\left(\frac{\nabla\Psi}{\Psi}\right)\right]^2\right\} \end{aligned} \tag{2.7.2}$$

is known as the **quantum potential**.Now we can do the same thing as in classical mechanics but additionally take the quantum potential into account. We can then calculate the trajectories of particles using Newton's usual second law $m\vec{a} = \vec{F} = -\vec{\nabla}(V + Q)$.

However, we never know $r(t_0)$ for any particle with absolute accuracy. The quantum potential is extremely sensitive to small changes in the initial conditions. Therefore, the best we can do is average over ensembles of particles. This way we again end up with probabilistic predictions just as in the wave function formulation. The motivation for the name pilot wave formulation comes from the observation that we have a wave-like potential Q that guides our particles as they move through space.

2.7.2 Again, Path Integral

What we are interested in is the probability that a particle that starts at a point A ends up after some time T at another point B. Using the standard quantum framework, we have:

$$\langle B|\Psi(A,T)\rangle$$

Using the time evolution operator:

$$\begin{aligned}\langle B|\Psi(A,T)\rangle &= \langle B|U(T)|A\rangle \\ &= \left\langle B\left|\mathrm{e}^{-iHT}\right|A\right\rangle\end{aligned}$$

Let's consider one specific path where the particle travels from A via some intermediate point q_1 to B. The corresponding probability amplitude is

$$\left\langle B\left|\mathrm{e}^{-iH(T-t_1)}\right|q_1\right\rangle\left\langle q_1\left|\mathrm{e}^{-iHt_1}\right|A\right\rangle$$

Of course we need to take into account the probability amplitudes that after t_1 seconds the particle is at any possible locations q. And mathematically this means that

$$\int dq_1 \left\langle B\left|\mathrm{e}^{-iH(T-t_1)/\hbar}\right|q_1\right\rangle\left\langle q_1\left|\mathrm{e}^{-iHt_1}\right|A\right\rangle \tag{2.7.3}$$

However, **to consider all possible paths we have to do the same thing for all points in time between 0 and T.** For this purpose, we slice the interval [0, T] into N equally sized pieces: $\delta = T/N$. The time evolution operator between two points in time is then $U(\delta) = \mathrm{e}^{-iH\delta/\hbar}$ and we have to sum after each time evolution step over all possible locations.

Mathematically, we have completely analogous to Eq. (2.7.3) for the amplitude $\psi_{A\to B}$ that we want to calculate

$$\begin{aligned} = \psi_{A\to B} = \int dq_1 \cdots dq_{N-1} \left\langle B \left| e^{-iH\delta} \right| q_{N-1} \right\rangle \left\langle q_{N-1} \left| e^{-iH\delta} \right| q_{N-2} \right\rangle \cdots \\ \cdots \left\langle q_1 \left| e^{-iH\delta} \right| A \right\rangle \end{aligned} \tag{2.7.4}$$

Our task is therefore to calculate the products of the form

$$\left\langle q_{N-1} \left| e^{-iH\delta} \right| q_j \right\rangle \equiv K_{q_{j+1},q_j} \tag{2.7.5}$$

which is usually called the **propagator.** We expand the exponential function in a series since δ is tiny:

$$\begin{aligned} K_{q_{j+1},q_j} &= \left\langle q_{j+1} \left| \left(1 - iH\delta - \frac{1}{2}H^2\delta^2 + \cdots \right) \right| q_j \right\rangle \\ &= \langle q_{j+1}|q_j\rangle - i\delta \langle q_{j+1}|H|q_j\rangle + \ldots \end{aligned} \tag{2.7.6}$$

It has been shown that

$$K_{q_{j+1},q_j} = \int \frac{dp_j}{2\pi} e^{ip_j(q_{j+1}-q_j)} \exp\left(-i\delta H\left(p_j, q_{j+1}\right)\right) \tag{2.7.7}$$

the amplitude $\psi_{A\to B}$ now becomes

$$\psi_{A\to B} = \int \prod_{j=1}^{N-1} dq_j \int \prod_{j=0}^{N-1} \frac{dp_j}{2\pi} \exp\left(i\delta \sum_{j=0}^{N-1} \left(p_j \frac{(q_{j+1}-q_j)}{\delta} - H\left(p_j, \bar{q}_j\right) \right) \right) \tag{2.7.8}$$

Now, in the limit $N \to \infty$ our interval δ becomes infinitesimal Therefore, in this limit the term $\frac{(q_{j+1}-q_j)}{\delta}$ becomes the velocity $\dot{q}$[19] So the term in the exponent reads $p\dot{q} - H = L$. And the amplitude can be rewritten as:

$$\psi_{A\to B} = \left(\frac{m}{2\pi i\delta}\right)^{N/2} \int \prod_{j=1}^{N-1} dq_j \exp\left(i\delta \sum_{j=0}^{N-1} \left(L\left(q_j\right)\right)\right) \tag{2.7.9}$$

In a compact form:

$$\psi_{A\to B} = \int \mathcal{D}q(t) e^{iS[q(t)]/\hbar} \tag{2.7.10}$$

where $S[q(t)]$ is the action that we always use in the Lagrangian formalism and $\mathcal{D}q(t)$ the so-called path integral measure.

2.7.3 Heisenberg Formulation

Let's say we want to calculate the expectation value of some operator $\hat{O}$ for a system in the state $|\Psi(x,t)\rangle$:

$$\langle\Psi(x,t)|\hat{O}|\Psi(x,t)\rangle = \left\langle \Psi(x,0) \left| U^{+}(t)\hat{O}U(t) \right| \Psi(x,0) \right\rangle$$

So far, we have always assumed that our states change over time. In particular, this means that our operator $U(t)$ acts on the ket $|\Psi(x,0)\rangle$. However, without changing any result, we can equally say that U **acts on the operator** $\hat{O}$ **instead and the kets remain unchanged.** The time evolution of an operator is then given by

$$\mathcal{O}(t) = U^{\dagger}(t)OU(t) \tag{2.7.11}$$

This change of perspective is known as the **Heisenberg picture.**

In the Heisenberg picture, the Schrödinger equation (which describes the time evolution of states) gets replaced with the so-called **Heisenberg equation**.

$$\frac{d}{dt}\mathcal{O} = \frac{i}{\hbar}[\hat{H},\mathcal{O}] + (\partial_t\mathcal{O}) \tag{2.7.12}$$

Chapter 3

Electromagnetism

3.1 Fundamental Concepts

3.1.1 Charge, Current, Flux

Electric charge is conserved no matter how small our system is, even for elementary particle processes

The total charge inside any volume V is then given by the integral over the charge density:

$$Q = \int_V d^3x \rho(\vec{x}, t) \tag{3.1.1}$$

Of course, **we can also use the charge density if there is only one charged object in our system.** In this case, the charge density is zero everywhere except at one specific point. Any integral over a region which contains the location of the object, yields simply the charge of the object.

Specifically, the charge density of a system with just one charged object located at $\vec{x}_0$ is

$$\rho(\vec{x}) = q\delta(\vec{x} - \vec{x}_0) \tag{3.1.2}$$

where q is the charge of the object. Any integral over a region V_0 which contains $\vec{x}_0$ yields exactly q, as it should be

$$\begin{aligned}\int_{V_0} d^3x\rho(\vec{x}, t) &= \int_{V_0} d^3xq\delta(\vec{x} - \vec{x}_0) \\ &= q\int_{V_0} d^3x\delta(\vec{x} - \vec{x}_0) \\ &= q\end{aligned}$$

In electrodynamics, the the electric current is defined as:

$$I = \frac{dQ}{dt} \tag{3.1.3}$$

The correct tool to describe the flow of charges in three dimensions is called **current density**. A current density yields a vector at each point in space(i.e., vector space). The direction of the vector at a given point describes the direction of the flow. The length of the vector describes how much is flowing.

By introducing a vector $\vec{n}$ of unit length normal to a frame with area of A, we can calculate the current passing through the frame as:

$$I \equiv \frac{\Delta Q}{\Delta t} = \rho_0 A\vec{v}_c \cdot \vec{n} \tag{3.1.4}$$

If we define **the electric current density** as

$$\vec{J} \equiv \rho_0\vec{v}_c \tag{3.1.5}$$

we have

$$\vec{J} \cdot \vec{n}A = \rho_0\vec{v}_c \cdot \vec{n}A = I \tag{3.1.6}$$

In general, **the magnitude of the current density $|\vec{J}|$ at one specific point de-**

scribes the amount of electric charge which passes per unit time through an infinitesimal surface element which is at right angles to the direction of the flow. The direction of the charge density vector $\vec{J}$ encodes where the charges are flowing.

If we want to know how much electric charge is flowing through a more complicated surface S, we have to calculate the **surface integral**:

$$I = \int_S \vec{J} \cdot \overrightarrow{dS} \tag{3.1.7}$$

The total amount of charge passing through the surface S during some time interval Δt is then given by

$$\Delta Q = \Delta t \int_S \vec{J} \cdot \overrightarrow{dS} \tag{3.1.8}$$

3.1.2 Electromagnetic Field

An important feature of the electromagnetic field is that even a vector field is not sufficient and we need instead a tensor field. **The electromagnetic field tensor is an antisymmetric** (4×4) **matrix**. This tensor field, in some sense, assigns exactly two vectors to each point in space $\vec{x}$ at each point in time t.

It is conventional to introduce the electric vector field $\vec{E}(t, \vec{x})$ and the magnetic vector field $\vec{B}(t, \vec{x})$ and work with them, instead of with the electromagnetic tensor field $F_{\mu\nu}(t, \vec{x})$.

$$F_{\mu\nu}(t, \vec{x}) \equiv \begin{pmatrix} 0 & -E_1(t,\vec{x})/c & -E_2(t,\vec{x})/c & -E_3(t,\vec{x})/c \\ E_1(t,\vec{x})/c & 0 & -B_3(t,\vec{x}) & B_2(t,\vec{x}) \\ E_2(t,\vec{x})/c & B_3(t,\vec{x}) & 0 & -B_1(t,\vec{x}) \\ E_3(t,\vec{x})/c & -B_2(t,\vec{x}) & B_1(t,\vec{x}) & 0 \end{pmatrix} \tag{3.1.9}$$

Unfortunately, the interpretation of the electric field $\vec{E}(t, \vec{x})$ and the magnetic field $\vec{B}(t, \vec{x})$ is not so simple. Instead, the little arrows encode information about the somewhat abstract "physical field", which we can understand as follows: The direction of the vector at a given location encodes in which direction a test charge would move if it were placed here. The

magnitude of the vector encodes how fast the test charge would accelerate as a result of, for example, the electric field. In this sense, these more abstract fields encode how something would flow if it were there.

3.1.3 Electromagnetic potential

For the moment, we only note that the electromagnetic potential is characterized by 4 numbers at each point in space and time $A_\mu(t,\vec{x}) = (A_0(t,\vec{x}), A_1(t,\vec{x}), A_2(t,\vec{x}), A_3(t,\vec{x}))^T$ and that the electric and magnetic fields can be calculated immediately once the electromagnetic potential is specified:

$$\begin{aligned} E_i &= \left(-\partial_i A_0 - \partial_0 A_i\right)/c \\ B_i &= \epsilon_{ijk}\partial_j A_k \end{aligned} \tag{3.1.10}$$

where $i,j,k \in \{1,2,3\}$. We can also write this as a vector equations

$$\begin{aligned} \vec{E} &= \left(-\nabla A_0 - \partial_t \vec{A}\right)/c \\ \vec{B} &= \nabla \times \vec{A} \end{aligned} \tag{3.1.11}$$

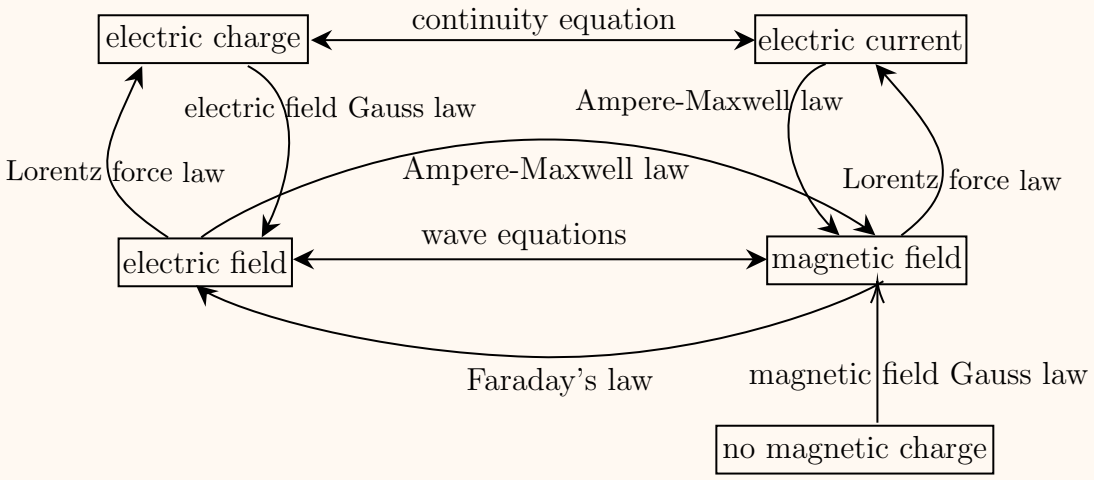
We can also express the tensor field itself using the electromagnetic potential

$$F_{\mu v} = \partial_\mu A_v - \partial_v A_\mu \tag{3.1.12}$$

3.2 Fundamental Equations

The Lorentz force law describes how charged object react to the presence of the electric and magnetic field. Maxwell's equations describe how non-zero electric and magnetic field strengths are generated by charges and currents. Moreover, they also describe how the electric and magnetic fields influence each other. In addition, the continuity equation encodes

the fact that electric charge is conserved. All this is summarized by the following diagram:



3.2.1 Continuity equation

$$\frac{\partial}{\partial t}\int_V \rho dV = \oint_S \vec{J}\cdot\overrightarrow{dS} \tag{3.2.1}$$

The continuity equation really describes the local conservation of electric charge. In words, the continuity equation tells us: the at which the charge density changes in a given volume is exactly equal to the net amount of electric charge flowing through the surface of the volume per unit time.

Using Gauss's theorem, we have

$$\oint_S \vec{J}\cdot\overrightarrow{dS} = \int_V \nabla\cdot\vec{J}dV \tag{3.2.2}$$

and we get the **differential form of the continuity equation**

$$\frac{\partial}{\partial t}\rho - \nabla\cdot\vec{J} = 0 \tag{3.2.3}$$

3.2.2 Lorentz force law

$$\vec{F}_E = q\vec{E} \tag{3.2.4}$$

$$\vec{F}_B = q\vec{v} \times \vec{B} \tag{3.2.5}$$

The Lorentz force law consists of two parts. The first part, tells us that the force resulting from the electric field $\vec{E}$ is directly proportional to $\vec{E}$. The second part take into account that only moving charged objects react to the presence of the magnetic field. When the velocity $\vec{v}$ is completely perpendicular to the magnetic field $\vec{B}$, the trajectory is simply a circle.

The general Lorentz force law therefore reads

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{3.2.6}$$

3.2.3 Gauss's law for the electric field

Each electric charge generates a particular structure in the electromagnetic field and this, in turn, has a direct impact on other charges.

We first introduce the **flux of the electric field through some surface** S:

$$\phi \equiv \int_S \vec{E} \cdot \overrightarrow{dS} \tag{3.2.7}$$

Our goal is to find an equation which describes that **a nonzero charge within some volume has a direct impact on the structure of the electric field**. The charge within some volume V is given by $\int_V \rho dV$ and we just argued that **a reasonable object represent the electric field in such an equation is the flux of the electric field**:

$$\oint_S \vec{E} \cdot \overrightarrow{dS} = \frac{1}{\epsilon_0} \int_V \rho dV \tag{3.2.8}$$

where ϵ_0 is a constant known as **electric permittivity** which *describes how strongly the electric field reacts to the presence of charges.* This is the integral form of Gauss's law for electric fields. In words, it tells us:

> The flux of the electric field passing through a closed surface is directly proportional to the amount of electric charge contained inside the surface.

Note that "flux" here is simply a statement about the **length of the vectors at the boundary of our volume.** Moreover, if the sign of the flux is positive, the vectors point outwards and *vice versa*.

- $\vec{J}$ and $\vec{E}$ assign a vector to each point in space.
- For $\vec{J}$ these vectors represent the real movement of our electric charges. For the more abstract electric field $\vec{E}$, the arrows only represent how a charge would move if it were there.

Using Gauss's theorem, we also have **the differential form of Gauss's law for the electric field**:

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \tag{3.2.9}$$

In general, the divergence of a vector field gives us information about the tendency of the field to flow towards or away from a specific point. With this in mind, we can say that the differential form of Gauss's law for the electric field tells us:

> The structure of the electric field generated by electric charges is such that it converges upon negative charges and diverges from positive charges.

The integral form is especially useful for macroscopic problems which posses a high degree of symmetry, e.g. when there is a rotationally symmetric charge distribution. The differential form is useful, for example, when we want to talk about electromagnetic waves.

3.2.4 Gauss's law for the magnetic field

So far no non-zero magnetic charge density has ever been observed, and **the integral form of Gauss's law for the magnetic field** reads

$$\oint_S \vec{B} \cdot \overrightarrow{dS} = 0 \tag{3.2.10}$$

Thus, this equation tells us

The flux of the magnetic field passing through any closed surface is zero.

Again, **the differential form is**

$$\nabla \cdot \vec{B} = 0 \tag{3.2.11}$$

Since the right-hand side of our equation is zero, we can conclude that **the tendency of the magnetic field to flow towards a point P is always exactly equal to its tendency to flow away from P. Any "inflow" of the magnetic field is always accompanied by an "outflow" of exactly the same magnitude.**

If the divergence is non-zero, the structure was created by a fundamental charge. If the divergence is zero, the structure was created by a different field (or a current).

The non-zero magnetic field strengths generated by conventional magnets are generated by the billions of individual electrons which exist inside any macroscopic object. **Each electron orbits around a nucleus and possesses a non-zero internal angular momentum**. The special thing about magnets is that in these objects **the majority of electrons rotate around approximately the same axis**.

Note that **we can also achieve a non-zero electric field with zero divergence. This is described by Faraday's law. The structure of the electric field generated by a changing magnetic field circulates back on itself**.

3.2.5 Faraday's law

We want a description in which a non-zero electric field strength shows up when the magnetic field is changing. The correct mathematical notion this task are time-derivatives. If the time-derivative of the magnetic field is non-zero, this means that the magnetic field is changing over time. Hence, our first puzzle piece is

$$\cdots = \int_S \frac{\partial}{\partial t} \vec{B} \cdot \overrightarrow{dS}$$

The second puzzle piece that we need is something which describes the structure of the electric field that results from such a changing magnetic field.

A first hint is that the left-hand side must depend on our choice of the surface S too. However, a volume integral over $\vec{E}$ is not a good idea since then S would represent the surface of the volume. As a result S would be a closed surface and, as mentioned above, Gauss's law for the magnetic field tells us that the integral over the magnetic field for any closed surface vanishes. Thus we use the boundary of a surface:

$$\oint_C \vec{E} \cdot \overrightarrow{dl} = -\int_S \frac{\partial}{\partial t} \vec{B} \cdot \overrightarrow{dS} \tag{3.2.12}$$

The contour integral of LHS describes the circulation of the electric field. The minus sign encodes in which direction the resulting electric field circulates.

To summarize, Faraday's law tells us

A changing magnetic field flux generates a circulating structure in the electric field.

Question

What's the physical meaning of a contour integral like $\oint_C \vec{E} \cdot \overrightarrow{dl}$?

Answer

From Lorentz force laws, we have

$$\begin{aligned}\oint_C \vec{E}\cdot\overrightarrow{dl} &= \oint_C \frac{\vec{F}}{q}\cdot\overrightarrow{dl}\\ &= \frac{1}{q}\oint_C \vec{F}\cdot\overrightarrow{dl}\\ &\equiv \frac{W}{q}\end{aligned}$$

where$W$ denotes the work done by the electric field if we move a test charge q along the path C.

Therefore, we can conclude that the circulation of the electric field represents the energy for each unit of charge moving along the contour C.

Here, we use Stokes' theorem to derive the differential form of Faraday's theorem.

$$\oint_C \vec{E}\cdot\overrightarrow{dl} = \int_S \nabla\times\vec{E}\cdot\overrightarrow{dS} \tag{3.2.13}$$

$$\nabla\times\vec{E} = -\frac{\partial}{\partial t}\vec{B} \tag{3.2.14}$$

3.2.6 Ampere-Maxwell law

The Ampere-Maxwell law describes how a changing electric field generates a non-zero magnetic field strength.**A non-zero magnetic field strength can be generated through a changing electric field. We can also generate a non-zero magnetic field strength using moving electric charge.** Thus the integral form of the Ampere-Maxwell law reads

$$\oint_C \vec{B}\cdot\overrightarrow{dl} = \mu_0\left(I_{\text{enc}} + \epsilon_0\int_S \frac{\partial}{\partial t}\vec{E}\cdot\overrightarrow{dS}\right) \tag{3.2.15}$$

μ_0 and ϵ_0 are proportionality constants known as vacuum permeability and vaccum permittivity. Moreover, I_{enc} denotes the electric current enclosed by the contour C.

By using Stokes' theorem, we find

$$\oint_C \vec{B} \cdot \overrightarrow{dl} = \mu_0 \left(I_{\text{enc}} + \epsilon_0 \int_S \frac{\partial}{\partial t} \vec{E} \cdot \overrightarrow{dS} \right)$$
$$\int_S \nabla \times \vec{B} \cdot \overrightarrow{dS} = \mu_0 \left(I_{\text{enc}} + \epsilon_0 \int_S \frac{\partial}{\partial t} \vec{E} \cdot \overrightarrow{dS} \right)$$
$$\int_S \nabla \times \vec{B} \cdot \overrightarrow{dS} = \mu_0 \left(\int_S \overrightarrow{J}\, \overrightarrow{dS} + \epsilon_0 \int_S \frac{\partial}{\partial t} \vec{E} \cdot \overrightarrow{dS} \right)$$
$$\int_S \nabla \times \vec{B} \cdot \overrightarrow{dS} = \mu_0 \int_S \left(\vec{J} + \epsilon_0 \frac{\partial}{\partial t} \vec{E} \right) \cdot \overrightarrow{dS}$$

And **the differential form of the Ampere-Maxwell law** becomes

$$\nabla \times \vec{B} = \mu_0 \left(\vec{J} + \epsilon_0 \frac{\partial}{\partial t} \vec{E} \right) \tag{3.2.16}$$

3.2.7 The wave equations

To derive the equation of motion for the electric field $\overrightarrow{E}$, we start by taking the curl on both sides of Faraday's law:

$$\nabla \times \vec{E} = -\frac{\partial}{\partial t} \vec{B}$$
$$\nabla \times \nabla \times \vec{E} = -\nabla \times \frac{\partial}{\partial t} \vec{B}$$
$$\nabla \times \nabla \times \vec{E} = -\frac{\partial}{\partial t} \nabla \times \vec{B}$$

Using the follwoing vector identity

$$\nabla \times \nabla \times \vec{E} = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} \tag{3.2.17}$$

We have:

$$\nabla \times \nabla \times \vec{E} = -\frac{\partial}{\partial t} \nabla \times \vec{B}$$
$$\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{\partial}{\partial t} \nabla \times \vec{B}$$
$$\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{\partial}{\partial t} \left(\mu_0 \left(\vec{J} + \epsilon_0 \frac{\partial}{\partial t} \vec{E} \right) \right)$$

Gauss's law for the electric field reads $\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}$, we have

$$\nabla\left(\frac{\rho}{\epsilon_0}\right) - \nabla^2\vec{E} = -\frac{\partial}{\partial t}\left(\mu_0\left(\vec{J} + \epsilon_0\frac{\partial}{\partial t}\vec{E}\right)\right)$$

And finally, we obtain **the wave equation for the electric field**

$$\nabla^2\vec{E} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\vec{E} \tag{3.2.18}$$

By following completely analogues steps, we have **the wave equation for the magnetic field** as

$$\nabla^2\vec{B} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\vec{B} \tag{3.2.19}$$

So in summary, static charged object $\rightarrow$ static radial electric field

moving charged object $\rightarrow$ static circulating magnetic field

changing magnetic field $\rightarrow$ changing circulating electric field

changing electric field $\rightarrow$ changing circulating magnetic field.

3.3 Eletrostatics and Magnetostatics

If the charge density is static $(\partial_t\rho = 0)$, the resulting electric field configuration is time-independent. Analogously, if our current is steady $\left(\partial_t\vec{J} = 0\right)$, the resulting magnetic field configuration is time-independent.

In such systems we can treat the electric field and the magnetic field independently because only if the electric field is changing does it have an effect on the magnetic field and vice versa.

For the electric field, we have

$$\begin{aligned} \nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \times \vec{E} &= 0 \end{aligned} \tag{3.3.1}$$

The general solution to the electrostatic equations in (3.3.1) is known as Coulomb's law

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r})\,(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} d^3\vec{r}' \tag{3.3.2}$$

If we writing our charge distribution in terms of individual charges

$$\rho(\vec{r}) = \sum_i q_i \delta\left(\vec{r}-\vec{r}_i\right)$$

we have

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r})\,(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} d^3\vec{r}' = \sum_i \frac{q_i}{4\pi\epsilon_0} \frac{\vec{r}-\vec{r}_i}{|\vec{r}-\vec{r}_i|^3}$$

For the magnetic field

$$\begin{aligned} \nabla \cdot \vec{B} &= 0 \\ \nabla \times \vec{B} &= \mu_0 \vec{J} \end{aligned} \tag{3.3.3}$$

The solution has form of

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{J}(\vec{r}') \times (\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} d^3r' \tag{3.3.4}$$

This is known as **Biot-Savart law**. It tells us that the magnetic field configuration generated by a general current density is simply the sum over the field configurations generated by the individual moving charges:

$$\vec{J}(\vec{r}) = \sum_i q_i \delta\left(\vec{r}-\vec{r}_i\right) \vec{v}_i$$

Putting this into 3.3.4

$$\vec{B}(\vec{r}) = \sum_i \frac{\mu_0 q_i}{4\pi} \vec{v}_i \times \frac{\vec{r}-\vec{r}_i}{|\vec{r}-\vec{r}_i|^3}$$

3.3.1 Electric field of a single point charge

Single point charge $\rho(\vec{r}) = q\delta(\vec{r})$, using Gauss's law

$$\oint_S \vec{E} \cdot \overrightarrow{dS} = \frac{1}{\epsilon_0} \int_V \rho dV$$

$$\int_S \vec{E} \cdot \overrightarrow{dS} = \frac{q}{\epsilon_0} \underbrace{\int_V \delta(\vec{r}) dV}_{=1}$$

$$\int_S \vec{E} \cdot \overrightarrow{dS} = \frac{q}{\epsilon_0}$$

$$\int_0^{2\pi} \int_0^{\pi} E \underbrace{\vec{e}_r \cdot \vec{e}_r} r^2 \sin(\theta) d\theta d\phi = \frac{q}{\epsilon_0}$$

Thus,

$$E = \frac{q}{r^2 4\pi\epsilon_0}$$

And the **Coulomb potential** is

$$\phi = \frac{Qq}{r 4\pi\epsilon_0} \tag{3.3.5}$$

3.3.2 Electric field of a sphere

charge distribution reads

$$\rho(\vec{r}) = \begin{cases} \rho_0, & \text{on the sphere: } |\vec{r}| = R \\ 0, & \text{otherwise} \end{cases}$$

Thus,

$$\vec{E}(\vec{r}) = 0 \quad \text{for } |\vec{r}| < R$$

and

$$\vec{E}(\vec{r}) = \frac{\rho_0}{4\pi\epsilon_0} \int \frac{\delta\left(|\vec{r}'| - R\right)(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} d^3r' = \frac{\rho_0}{4\pi\epsilon_0} \int_0^{2\pi} \int_0^{\pi} \int_0^{\infty} \frac{\delta\left(|\vec{r}'| - R\right)(\vec{r} - \vec{r}')}{|\vec{r} - \vec{r}'|^3} r'^2 dr' \sin(\theta) d\phi d\theta$$

Thus, outside of the sphere the electric field is described by

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0 r^2} \vec{e}_r$$

3.3.3 Electric field of an electric dipole

For simplicity, we assume that we are dealing with two point charges and choose our coordinate system such that one charge sits at the origin. The charge density then reads

$$\rho(\vec{x}) = q\delta(\vec{x}) - q\delta(\vec{x} - \vec{d})$$

where $\vec{d}$ is a vector pointing from the charge at the origin to the second charge. We can now immediately write down the resulting electric field configuration because we know that we simply have to **use a superposition of the individual configurations**.

$$\vec{E}(r) = \vec{E}_q(r) + \vec{E}_{-q}(r) = \frac{q\vec{r}}{4\pi\epsilon_0|\vec{r}|^3} - \frac{q(\vec{r} - \vec{d})}{4\pi\epsilon_0|\vec{r} - \vec{d}|^3}\vec{e}_r \tag{3.3.6}$$

3.3.4 Electric field of more complicated charge distributions

The general formula for the electric potential is

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho\,(\vec{r}')}{|\vec{r} - \vec{r}'|}d^3r' \tag{3.3.7}$$

We assume that the charge distribution $\rho\,(\vec{r}')$ is localized within some region V. In mathematical terms, being far away from our charge distribution then means $|\vec{r}| \gg |\vec{r}'|$ for all $\vec{r}'$ in V

The main idea is that **we can then use Taylor expansion to simplify the general formula (3.3.7)**. In particular, we need the expansion

$$\begin{aligned}\frac{1}{|\vec{r} - \vec{r}'|} &= \frac{1}{r} - \vec{r}' \cdot \nabla\frac{1}{r} + \dots \\ &= \frac{1}{r} + \vec{r}' \cdot \frac{\vec{r}}{r^3} + \dots\end{aligned} \tag{3.3.8}$$

and

$$\begin{aligned}\phi(\vec{r}) =& \frac{1}{4\pi\epsilon_0}\int \rho\left(\vec{r}'\right)\frac{1}{|\vec{r}-\vec{r}'|}d^3r' \\ =& \frac{1}{4\pi\epsilon_0}\int \rho\left(\vec{r}'\right)\left(\frac{1}{r}+\vec{r}'\cdot\frac{\vec{r}}{r^3}+\ldots\right)d^3r' \\ =& \frac{1}{4\pi\epsilon_0 r}\int \rho\left(\vec{r}'\right)d^3r' \\ &+\frac{1}{4\pi\epsilon_0 r^3}\int \rho\left(\vec{r}'\right)\vec{r}'\cdot\vec{r}d^3r'+\ldots\end{aligned} \tag{3.3.9}$$

If we are sufficiently far away, we can approximate our general formula as

$$\phi(\vec{r}) \approx \frac{q}{4\pi\epsilon_0 r}$$

If we look closely enough, we should describe the potential using

$$\phi(\vec{r}) \approx \frac{q}{4\pi\epsilon_0 r}+\frac{1}{4\pi\epsilon_0 r^3}\int \rho\left(\vec{r}'\right)\vec{r}'\cdot\vec{r}d^3r'$$

where we can define

the **dipole moment**

$$\vec{p} \equiv \int \rho\left(\vec{r}'\right)\vec{r}'d^3r' \tag{3.3.10}$$

And the potential then reads

$$\phi(\vec{r}) \approx \frac{q}{4\pi\epsilon_0 r}+\frac{1}{4\pi\epsilon_0 r^3}\vec{p}\cdot\vec{r}$$

Analogously it's possible to introduce additional quantities which characterize the higher order terms in the Taylor expansion. For example,

$$\phi(\vec{r}) \approx \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r}+\frac{\vec{p}\cdot\vec{r}}{r^3}+\sum_{ij}\frac{Q_{ij}r_ir_j}{2r^5}\right)$$

where we have

the quadrupole moment defined as

$$Q_{ij} \equiv \int_V \rho\left(\vec{r}'\right)\left(3r_i'r_j'-\delta_{ij}r'^2\right)d^3r' \tag{3.3.11}$$

This method of collecting step-by-step information about complicated charge distributions is known as **multipole expansion**.

3.3.5 Boundary Value Problem

Finding a description of electrostatic systems where the charge distribution is not fixed and known is usually called a boundary value problem. The main task in all these problems is to **solve the integral in Coulomb's law**.

In the context of boundary value problems it is often more convenient to use the electric potential instead of the electric field. The correct electrostatic equation for the electric potential can be derived using Maxwell's equations and the relation between the potential ϕ and the electric field $\vec{E}$:

$$\begin{aligned} \nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0} \\ -\nabla^2 \phi &= \frac{\rho}{\epsilon_0} \end{aligned} \tag{3.3.12}$$

There are clever methods for solving the Poisson equation. The main idea behind the most famous one is that **we can calculate the solution for a general charge distribution by using the known solutions for individual charges**. Hence, the basic building block for a general solution is the solution for a single point charge. The charge distribution of a single point charge at $\vec{r}'$ is $\rho(\vec{r}) \sim \delta(\vec{r} - \vec{r}')$

Therefore, our main task in constructing a general solution is to solve the equation:

$$\nabla^2 G(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}') \tag{3.3.13}$$

The solution $G(\vec{r}, \vec{r}')$ is known as the Green's function for the Laplacian operator $\nabla^2 \equiv \Delta$. The Green's function for the Laplacian operator reads

$$G(\vec{r}, \vec{r}') = \frac{1}{4\pi} \frac{1}{|\vec{r} - \vec{r}'|} \tag{3.3.14}$$

The main idea is that as soon as we have the solution $G(\vec{r}, \vec{r}')$ for a single point charge, **we can calculate directly the solution of the Poisson equation for a general charge distribution $\rho(\vec{r}')$ as follows:**

$$\phi_{\text{sol}}(\vec{r}) = \frac{1}{\epsilon_0} \int \rho\left(\vec{r}'\right) G\left(\vec{r}, \vec{r}'\right) d^3r' \tag{3.3.15}$$

3.3.6 Magnetic field of a wire

The Ampere-Maxwell law tells us:

$$\oint_C \vec{B} \cdot \overrightarrow{dl} = \mu_0 \int_S \vec{J} \cdot \overrightarrow{dS} \tag{3.3.16}$$

Since

$$\begin{aligned} \oint_C \vec{B} \cdot \overrightarrow{dl} &= \int_0^{2\pi} |\vec{B}| \vec{e}_\varphi \cdot \vec{e}_{\varphi^r} d\varphi \\ &= |\vec{B}| r \int_0^{2\pi} d\varphi \\ &= |\vec{B}| r 2\pi \end{aligned}$$

and

$$\mu_0 \int_S \vec{J} \overrightarrow{dS} = \mu_0 I_{\text{wire}}$$

we obtain

$$|\vec{B}| = \frac{\mu_0 I_{\text{wire}}}{r 2\pi}$$

Therefore

$$\vec{B}(\vec{r}) = \frac{\mu_0 q |\vec{v}|}{2\pi r} \vec{e}_\varphi \tag{3.3.17}$$

where $\vec{e}_\varphi$ is a unit vector always pointing tangential to the contour.

The direction in which it circles can be determined by the so-called "right hand rule". If your thumb points in the direction of the current, the remaining fingers curl in the direction of the magnetic field.

There is also a Poisson equation for the magnetic potential:

$$\nabla^2 \vec{A} = -\mu_0 \vec{J} \tag{3.3.18}$$

The Poisson equation for the vector potential can then be derived by rewriting the magnetic field in terms of the potential: $\vec{B} = \nabla \times \vec{A}$. This yields $\mu_0 \vec{J} = \nabla \times \vec{B} = \nabla \times \nabla \times \vec{A}$. Then, we can rewrite this using the identity

$$\nabla \times \nabla \vec{A} = \nabla(\nabla \vec{A}) - \nabla^2 \vec{A}$$

To simplify this further, we can use the observation that potentials cannot be directly measured and only potential differences can. Hence, we can always add a constant to our potentials, and this so-called gauge freedom can be used to achieve $\nabla \cdot \vec{A} = 0$.If we use the gauge freedom like this, we get the poisson equation above. The general solution can be derived again using the Green's function method and reads

$$\vec{A}_{\text{sol}}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{J}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r' \tag{3.3.19}$$

3.4 Electrodynamics

3.4.1 General Solution of the Wave Function

The most basic solutions of the electric wave equations look like

Plane wave solution

$$\vec{E} = \vec{E}_0 \cos(\vec{k} \cdot \vec{r} \pm \omega t + \delta) \tag{3.4.1}$$

where $\vec{k}$ is a vector that determines in which direction the wave is traveling and δ encodes possible phase shifts.
Moreover, $\vec{E}_0$ is a vector whose magnitude $\left|\vec{E}_0\right|$ describes the amplitude of the wave and its direction determines the direction in which the electric field oscillates. The choice of the sign $\pm$ in the cosine function determines whether the wave travels in the positive direction or negative direction on the axis defined by $\vec{k}$.

we can use sums of the basic plane wave solutions to create more complicated solutions. A simple example is a **standing wave**:

$$\vec{E} = \overrightarrow{E_0}(\cos(\vec{k}\cdot\vec{r} - \omega t) + \cos(\vec{k}\cdot\vec{r} + \omega t)) = \vec{E}_0(\cos(\vec{k}\cdot\vec{r})\cos(\omega t)) \tag{3.4.2}$$

The crucial point is that the spatial dependence $\vec{k}\cdot\vec{r}$ of the wave is completely separated from the time dependence ωt. This means that the spatial shape of the wave is fixed and does not change as time moves on. However, since we are still multiplying this fixed form of the wave by $cos(\omega t)$, **the amplitude of the wave at each point still varies (but not the wave form).**

In particular, the most general solution of the wave equation is linear combination of all possible plane waves:

$$\vec{E}(\vec{r}, t) = \int_{-\infty}^{\infty} \frac{d^3k}{2\pi^3}\vec{E}(\vec{k})\cos(\vec{k}\cdot\vec{r} - \omega t) \tag{3.4.3}$$

3.4.2 Basic properties of electromagnetic waves

1. The argument $\varphi \equiv (\vec{k}\cdot\vec{r} \pm \omega t)$ of our periodic function $\cos\varphi$ is called the **phase** of the wave.

2. The vector $\vec{k}$ is usually called the wave vector. The direction of $\vec{k}$ tells us in which direction the wave is traveling. **The length of the wave vector $|\vec{k}|$ describes how many oscillations there are per meter.** To understand this, imagine that we could stop the time, i.e. keep t fixed and then move through space. As we move along the axis defined by $\vec{k}$ we count how many full wave shapes we encounter per meter. This number is the wave number. One full oscillation is over as soon as the phase of the wave has increased by 2π. Hence, we can say a bit more precisely that $|\vec{k}|$ measures how many 2π cycles there are per meter. For this reason, $|\vec{k}|$ is known as spatial angular frequency or **wave number**.

For example, if we move 2 meters and observe that the phase changes by 20π, we know that the wave number is 10π radians per meter. The wave number is directly related to the

wavelength:

$$\lambda = \frac{2\pi}{|\vec{k}|} \tag{3.4.4}$$

3. The constant ω is known as temporal angular frequency or simply angular frequency. The angular frequency describes how many oscillations there are per second.

4. The direction in which the amplitude vector $\vec{E}_0$ points, **specifies the geometrical orientation of the oscillation.** Usually, the direction of $\vec{E}_0$ is called the **polarization** of the wave. The length of the amplitude vector $\left|\vec{E}_0\right|$ encodes the peak magnitude of the oscillation.An extremely important observation is that the amplitude vector cannot point in any arbitrary direction for electromagnetic waves.According to Gauss's law, **The electromagnetic waves cannot be polarized in the direction they are traveling.**

5. The sign between the two terms in the cosine function determines whether our wave moves up or down on the axis defined by $\vec{k}$.

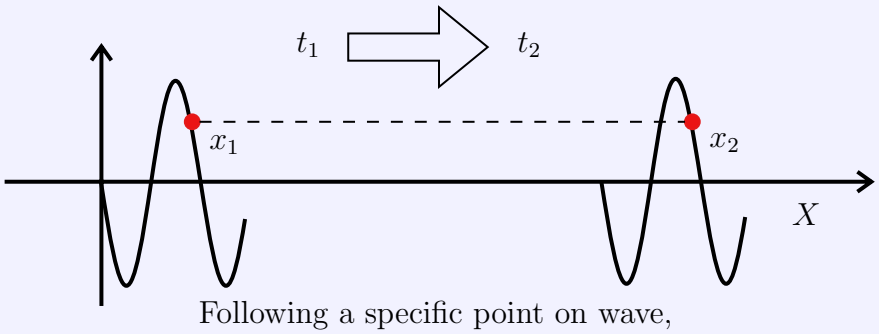
6.The absolute phase δ encodes the phase of the wave at $\vec{r} = 0$ and $t = 0$.

3.4.3 Advanced Properties of electromagnetic waves

Question

1. How fast are electromagnetic waves traveling?

Answer



Following a specific point on wave,

the phase of the point remains the same:

$kx_1 - wt_1 = kx_2 - wt2 \Rightarrow \boxed{v = \frac{x_2 - x_1}{t_2 - t_1} = \frac{w}{k}}$

In words, the velocity of each point in our wave form is given by the ratio of the angular frequency ω and the wave number k. From a different perspective, a velocity has units meter per second. The only combination of our basic wave quantities gives such unit is

$$v = \frac{\lambda}{\tau}$$

In words, this equation tells us that a wave travels one wavelength λ per period τ. We can rewrite the velocity of the wave in terms of the angular frequency ω and wave number $k \equiv |\vec{k}|$ as follows:

$$v = \frac{\lambda}{\tau} = \frac{2\pi/k}{2\pi/\omega} = \frac{\omega}{k} \tag{3.4.5}$$

2. If we put the general form of the solution in (3.2.18), we obtain:

the dispersion relation

$$k^2 = \mu_0 \epsilon_0 \omega^2 \tag{3.4.6}$$

And we have

$$\frac{\omega}{k} = \sqrt{\frac{1}{\mu_0 \epsilon_0}} \tag{3.4.7}$$

Putting the experimental values, we find the **speed of light**:

$$c \equiv 2.9979 \times 10^8 \text{m/s} \tag{3.4.8}$$

3. Let's now derive **Poynting theorem**, which is analogous to the continuity equation. And Poynting theorem proves that **electromagnetic wave carries energy.**

Using the Lorentz force law, we can calculate the work done by the electromagnetic field on a single charge q is:

$$\vec{F} \cdot d\vec{l} = q(\vec{E} + \vec{v} \times \vec{B}) \cdot \vec{v} dt = q\vec{E} \cdot \vec{v} dt$$

To calculate the work done on all charges, we use $q = \rho dV$ and $\vec{J} = \rho \vec{v}$. The rate at which work is done on all charges is therefore

$$\frac{dW}{dt} = \int_V (\vec{E} \cdot \vec{J}) dV \tag{3.4.9}$$

So, $\vec{E} \cdot \vec{J}$ is **the power per unit volume.** If we use Maxwell's equation to rewrite this power, we have:

$$\vec{E} \cdot \vec{J} = \frac{1}{\mu_0} \vec{E} \cdot (\vec{\nabla} \times \vec{B}) - \epsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t}$$

Since

$$\nabla \cdot (\vec{E} \times \vec{B}) = \vec{B} \cdot (\vec{\nabla} \times \vec{E}) - \vec{E} \cdot (\vec{\nabla} \times \vec{B}) \tag{3.4.10}$$

and

$$\frac{\partial}{\partial t} \vec{B}^2 = \left(\frac{\partial}{\partial t} \vec{B} \right) \vec{B} + \vec{B} \left(\frac{\partial}{\partial t} \vec{B} \right) = 2\vec{B} \frac{\partial}{\partial t} \vec{B} \tag{3.4.11}$$

we have

$$\begin{aligned}\vec{E}\cdot\vec{J} &= \frac{1}{\mu_0}(\vec{B}\cdot(\vec{\nabla}\times\vec{E}) - \nabla\cdot(\vec{E}\times\vec{B})) - \epsilon_0\vec{E}\cdot\frac{\partial\vec{E}}{\partial t} \\ &= \frac{1}{\mu_0}\vec{B}\cdot\left(-\frac{\partial}{\partial t}\vec{B}\right) - \frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B}) - \epsilon_0\vec{E}\cdot\frac{\partial\vec{E}}{\partial t} \\ &= -\frac{1}{2\mu_0}\frac{\partial}{\partial t}\vec{B}^2 - \frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B}) - \epsilon_0\frac{1}{2}\frac{\partial}{\partial t}\vec{E}^2\end{aligned} \tag{3.4.12}$$

Putting this into Eq.3.4.9 yields

$$\begin{aligned}\frac{dW}{dt} &= \int_V(\vec{E}\cdot\vec{J})dV \\ &= \int_V\left(-\frac{1}{2\mu_0}\frac{\partial}{\partial t}\vec{B}^2 - \frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B}) - \epsilon_0\frac{1}{2}\frac{\partial}{\partial t}\vec{E}^2\right)dV \\ &= -\frac{1}{2}\frac{\partial}{\partial t}\int_V\left(\frac{1}{\mu_0}\vec{B}^2 + \epsilon_0\vec{E}^2\right)\Big)dV - \int_V\frac{1}{\mu_0}\nabla\cdot(\vec{E}\times\vec{B})dV \\ &= -\frac{1}{2}\frac{\partial}{\partial t}\int_V\left(\frac{1}{\mu_0}\vec{B}^2 + \epsilon_0\vec{E}^2\right)dV - \oint_S\frac{1}{\mu_0}(\vec{E}\times\vec{B})\cdot d\vec{a} \\ &\equiv -\frac{dU_{\text{em}}}{dt} - \oint_S\vec{S}\cdot d\vec{a}\end{aligned} \tag{3.4.13}$$

Thus, the Poynting theorem reads:

$$\frac{dW}{dt} = -\frac{dU_{\text{em}}}{dt} - \oint_S\vec{S}\cdot d\vec{a} \tag{3.4.14}$$

Where

$$\vec{S} \equiv \frac{1}{\mu_0}(\vec{E}\times\vec{B}) \tag{3.4.15}$$

is usually called the **Poynting vector**. Moreover, U_{em} is the total energy stored in the electromagnetic field configuration.

The Poynting theorem tells us that the work done on charges by the Lorentz force is equal to the decrease in energy stored in the field $\frac{dU_{\text{em}}}{dt}$ minus the energy which has flowed through the surface S. In other words, **if the energy stored in the electromagnetic field gets smaller, it must have either been used to move charges around** $\left(\frac{dW}{dt}\right)$ **or has flown out of the volume we are considering** $\left(\oint_S\vec{S}\cdot d\vec{a}\right)$.

We can make the analogy between Poynting theorem and continuity equation more con-

crete by introduce the mechanical energy density:

$$\frac{dW}{dt} \equiv \frac{d}{dt}\int_V u_{\text{mech}} dV$$

where

$$u_{\text{em}} \equiv \frac{1}{2}\left(\epsilon_0 \vec{E}^2 + \frac{1}{\mu_0}\vec{B}^2\right)$$

Thus,

$$\frac{dW}{dt} = -\frac{dU_{\text{em}}}{dt} - \oint_S \vec{S}\cdot d\vec{a}$$
$$\frac{d}{dt}\int_V u_{\text{mech}} dV = -\frac{d}{dt}\int_V u_{\text{em}} dV - \oint_S \vec{S}\cdot d\vec{a}$$

and **the differential form of Poynting theorem is**

$$\frac{d}{dt}\left(u_{\text{mech}} + u_{\text{em}}\right) = -\nabla\cdot\vec{S}$$

Question

4. How can we produce changing field configuration?

Answer

Through accelerating charges

3.5 Gauge Symmetry

Gauge symmetry refers to the observation that we can't measure potentials like the electric potential A_0 or the magnetic potential $\vec{A}$ directly. **The electric and magnetic field strengths remain completely unchanged if we shift our potentials by a constant amount**:

$$A_0 \to A_0 + \eta$$
$$A_i \to A_i + \xi_i$$

Thus,

$$\begin{aligned} E_i = -\partial_i A_0 - \partial_0 A_i \to \tilde{E}_i &= -\partial_i \left(A_0 + \eta\right) - \partial_0 \left(A_i + \xi_i\right) \\ &= -\partial_i A_0 - \underbrace{\partial_i \eta}_{=0} - \partial_0 A_i - \underbrace{\partial_0 \xi}_{=0} \\ &= -\partial_i A_0 - \partial_0 A_i \\ &= E_i \end{aligned} \tag{3.5.1}$$

and

$$\begin{aligned} B_i = \epsilon_{ijk}\partial_j A_k \to B_i &= \epsilon_{ijk}\partial_j \left(A_k + \xi_k\right) \\ &= \epsilon_{ijk}\partial_j A_k + \epsilon_{ijk}\partial_j \xi_k \\ &= \epsilon_{ijk}\partial_j A_k \\ &= B_i \end{aligned} \tag{3.5.2}$$

More importantly, We can not only add constants to the potentials but also derivatives of an arbitrary scalar function $\eta(t, \vec{x})$:

Gauge Freedom/gauge symmetry

$$\begin{aligned} A_0 &\to A_0 + \partial_0 \eta(t, \vec{x}) \\ A_i &\to A_i + \partial_i \eta(t, \vec{x}) \end{aligned} \tag{3.5.3}$$

We can check this as:

$$\begin{aligned} E_i = -\partial_i A_0 - \partial_0 A_i \to \tilde{E}_i &= -\partial_i \left(A_0 + \partial_0 \eta(t, \vec{x})\right) - \partial_0 \left(A_i + \partial_i \eta(t, \vec{x})\right) \\ &= -\partial_i A_0 - \partial_i \partial_0 \eta(t, \vec{x}) - \partial_0 A_i - \partial_0 \partial_i \eta(t, \vec{x}) \\ &= -\partial_i A_0 - \partial_0 \partial_i \eta(t, \vec{x}) - \partial_0 A_i - \partial_0 \partial_i \eta(t, \vec{x}) \\ &= -\partial_i A_0 - \partial_0 A_i = E_i \end{aligned}$$

and

$$\begin{aligned} \tilde{B}_i &= \epsilon_{ijk}\partial_j \left(A_k + \partial_k \eta(t, \vec{x})\right) \\ &= \epsilon_{ijk}\partial_j A_k + \underbrace{\epsilon_{ijk}\partial_j \partial_k \eta(t, \vec{x})}_{=0} \\ &= \epsilon_{ijk}\partial_j A_k = B_i \end{aligned}$$

where we used that ϵ_{ijk} is antisymmetric but $\partial_j\partial_k$ is symmetric under the switching of the indices $j \leftrightarrow k$.

Question

How to choose a gauge so that $\nabla \cdot \vec{A} = 0$ (Coulomb Gauge)?

Answer

The modified potential reads:

$$\vec{A'} = \vec{A} + \nabla\eta(t, \vec{x})$$

The divergence of this reads

$$\begin{aligned}\nabla \cdot \vec{A'}(t, \vec{x}) &= \nabla \cdot \vec{A}(t, \vec{x}) + \nabla \cdot \nabla\eta(t, \vec{x}) \\ &= f(t, \vec{x}) + \nabla \cdot \nabla\eta(t, \vec{x})\end{aligned}$$

By solving the following equation:

$$\nabla \cdot \nabla\eta_{\text{sol}}(t, \vec{x}) = -\nabla \cdot f(t, \vec{x})$$

we can find our Coulomb gauge. In practice, we don't have to find the function explicitly, but instead can simply use the knowledge that a function with the required properties exists and set $\nabla \cdot \vec{A} = 0$.

3.5.1 Electrodynamics as a gauge theory

Financial toy system

Here we use a simple financial toy system to show the definition of gauge redundancy and symmetry. And the difference that active/passive transformation made to the toy system.

we now introduce independent local currencies in, say Germany, which we call Deutsche Mark (DM). Moreover, we introduce Francs (F) in France, in England Pounds (P) and Lira

(L) in Italy. The exchange rates are

$$DM/P = 1$$
$$P/F = 2$$
$$F/L = 10$$
$$DM/L = 20$$

When a local currency changes, the bookkeepers simply adjust their exchange rates accordingly and the situation remains the same. Thanks to these bookkeepers, we can see that such passive transformations really make no difference as it should be.Our bookkeeper aren't dynamical actors which start actions on their own. They do not have any physical influence on the dynamics of the system.

The change in the local currency is an example of invariance under passive transformations. It is called as a **redundancy**. An invariance under active transformations is called a **symmetry**. The redundancy that exists in our description of the financial toy example after the introduction of the bookkeepers is an example of a **local gauge symmetry.**

Gauge dynamics

So far, our bookkeepers are purely mathematical ingredients which we introduced to make our description invariant under local passive transformations. Let's now imagine the exchange rates set by the bank as:

$$DM/P = 1$$
$$P/F = 2$$
$$F/L = 10$$
$$DM/L = 10$$

Now our trader is able to earn money simply by exchanging money. If he starts with 1DM, he can trade it for 1P, then use the pound to trade it for 2F, then trade these for 20L and finally trade these for 2DM. In the financial world this is known as an **arbitrage opportunity**

If there is a change in the local money coordinate system, for examole $L \to \tilde{L} = L/10$,

we have to adjust the exchange rates accordingly as:

$$F/\tilde{L} = F/(L/10) = 1$$
$$DM/\tilde{L} = DM/(L/10) = 1$$

However, the amount of money our trader earns is unchanged by such a re-scaling. If he starts again with 1DM, he can still trade it for 1P, which he can trade for 2F, then trade these for 2L and finally trade these for 2DM. The final result is the same as before.

Here we show a crucial aspect of gauge theory: an important task is to find quantities which do not depend on local conventions.

There is another way to make risk-free money which involves interest rates. Let's assume that the interest rate is 2% per year for the Deutsche Mark. Moreover, we assume that the exchange rate changes over time and is now

$$DM/L = 10$$

and in one year

$$DM/L = 5$$

Our trader can borrow 1DM and after 1 year he has to pay 1.02DM back. Moreover, he can now exchange his 1DM for 10L. After one year he can then exchange his 10L for 2DM. Therefore, after paying back the 1.02DM he has made 0.98DM in profit.

Promoting bookkeepers to dynamical objects which follow their own rules is a crucial step to make our toy model more realistic.

Mathematical description of the toy model

Mathematically, we imagine that some countries live on a lattice. Each point on the lattice is labelled by d -numbers: $\vec{n} = (n_1, n_2, \cdots, n_d)$. In other words, each country can be identified by a vector $\vec{n}$ which points to its location. We can move from one country to a neighboring country by using a basis vector $\vec{e}_i$, where i denotes the direction we are moving. For example, $\vec{e}_2 = (0, 1, 0 \cdots, 0)$

We denote the **currency exchange rates** between the country labelled by the vector $\vec{n}$ and its neighbor in the i -direction by $R_{\vec{n},i}$. For example, if the country at the location labelled by $\vec{n}$ uses Deutsche Marks and its neighbor in the 2 -direction uses Francs, $R_{\vec{n},2}$ tells us how many Francs we get for each Deutsche Mark.

In physics, we usually introduce the corresponding logarithms

$$R_{\vec{n},i} \equiv e^{A_i(\vec{n})}$$

where $A_i(\vec{n}) \equiv \ln(R_{\vec{n},i})$.

We use the notation $f(\vec{n})$ to denote a change of the currency in the country at $\vec{n}$ by a factor of $f(\vec{n})$. In addition, we again introduce the corresponding logarithm

$$f(\vec{n}) \equiv e^{\varepsilon(\vec{n})}$$

In general, when we perform such a gauge transformation in the country labelled by $\vec{n}$ and also in the neighboring country in the i -direction, the corresponding exchange rate changes as follows:

$$\begin{aligned} R_{\vec{n},i} = e^{A_i(\vec{n})} \rightarrow & \frac{f_{\vec{n}} + \vec{e}_i}{f_{\vec{n}}} R_{\vec{n},i} \\ & = \frac{e^{\epsilon(\vec{n}+\vec{e}_i)}}{e^{\epsilon(\vec{n})}} e^{A_i(\vec{n})} \\ & = e^{A_i(\vec{n})+\epsilon(\vec{n}+\vec{e}_i)-\epsilon(\vec{n})} \end{aligned}$$

Thus, we conclude

$$A_i(\vec{n}) \rightarrow A_i(\vec{n}) + \epsilon(\vec{n} + \vec{e}_i) - \epsilon(\vec{n})$$

An arbitrage opportunity exists when we can trade currencies in such a way that we end up with more money than we started with. But we can only make such a statement when the starting currency and the final currency are the same. **The total gain we can earn by following a specific loop can be quantified by the gain factor**

$$G = R_{\vec{n},i} R_{\vec{n}+\vec{e}_i,j} \frac{1}{R_{\vec{n}+\vec{e}_j,i}} \frac{1}{R_{\vec{n},j}}$$

When this gain factor is larger than one, we can earn money by trading money following the loop. Once more we introduce the corresponding logarithm

$$G \equiv e^{F_{ij}(\vec{n})}$$

and

$$F_{ij}(\vec{n}) = A_j\left(\vec{n}+\vec{e}_i\right) - A_j(\vec{n}) - \left[A_i\left(\vec{n}+\vec{e}_j\right) - A_i(\vec{n})\right]$$

A crucial consistency check is that G and F_{ij} are unchanged by gauge transformations. Quantities like this are usually called **gauge invariant**. Moreover, an important technical observation is that $F_{ij}(\vec{n})$ is antisymmetric: $F_{ij}(\vec{n}) = -F_{ji}(\vec{n})$, which follows directly from thdefinition.

Now let's introduce temporal exchange rate by specifying a point on the lattice using $d+1$ coordinates $\vec{n} = (n_0, n_1, ...)$, where the zeroth component indicates the point in time. Thus,

$$F_{\mu\nu}(\vec{n}) = A_\mu\left(\vec{n}+\vec{e}_\mu\right) - A_v(\vec{n}) - \left[A_\mu\left(\vec{n}+\vec{e}_v\right) - A_\mu(\vec{n})\right] \tag{3.5.4}$$

In the continuum limit,

$$\begin{aligned} A_\mu\left(x_\mu\right) &\to A_\mu\left(x_\mu\right) + \frac{\partial\epsilon}{\partial x^\mu} \\ F_{\mu\nu}\left(x_\mu\right) &\equiv \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} \end{aligned} \tag{3.5.5}$$

We can not only earn money by trading money itself, but also by trading goods like, for example, copper. Depending on the local prices it can be lucrative to buy copper in one country and sell it in another country. The gain factor for such process is given by:

$$g = \frac{p\left(\vec{n}+\vec{e}_i\right)}{p(\vec{n})R_{\vec{n},i}} \tag{3.5.6}$$

where $p(\vec{n})$ is the price of copper at $\vec{n}$. Again, we introduce the corresponding logarithm

$$g \equiv e^{J_i(\vec{n})} \tag{3.5.7}$$

and

$$\begin{aligned} e^{J_i(\vec{n})} &= \frac{e^{\varphi(\vec{n}+\vec{e}_i)}}{e^{\varphi(\vec{n})}e^{A_i(\vec{n})}} \\ J_i(\vec{n}) &= \varphi\left(\vec{n}+\vec{e}_i\right) - \varphi(\vec{n}) - A_i(\vec{n}) \end{aligned} \tag{3.5.8}$$

The amount of money we earn depends on the amount of copper q we carry around. Thus,

$$J_\mu(\vec{n}) = q\left(\varphi\left(\vec{n}+\vec{e}_\mu\right) - \varphi(\vec{n}) - A_\mu(\vec{n})\right) \tag{3.5.9}$$

In the trade described by $J_i(\vec{n}), (i = 1, 2, 3)$ copper is transported from the country at $\vec{n}$ to the neighboring country at $\vec{n} + \vec{e}_i$. Hence, $J_i(\vec{n})$ is a measure of the amount of copper that flows between the two countries. $J_0(\vec{n})$ tells us how much money we can earn by buying copper and selling it at a later point in time in the **same country**.

Gauge rules

we start with the crucial assumption that copper is conserved (i.e. no copper is destroyed or produced). And it must be true that:change of amount of copper in country $\vec{n}$ = total net flow:

$$J_0\left(\vec{n}+\vec{e}_0\right) - J_0(\vec{n}) = -\left(\sum_{i=1}^{d} J_i(\vec{n}) - \sum_{i=1}^{d} J_i\left(\vec{n}-\vec{e}_i\right)\right) \tag{3.5.10}$$

In the continuum limit, we get continuity equation:

$$\sum_{\mu=0}^{3} \partial_\mu J_\mu = 0 \tag{3.5.11}$$

Moreover, we already know that the good quantities to describe our system are J_μ and $F_{\mu\nu}$ since these do not depend on local conventions. The quantities J_μ contain information about the positions and flow of copper, while $F_{\mu\nu}$ represent arbitrage opportunities. $F_{\mu\nu}$ is anti-symmetric and $\partial_\mu\partial_\nu = \partial_\nu\partial_\mu$,so

$$\sum_{\mu=0}^{3}\sum_{v=0}^{3} \partial_v\partial_\mu F_{\mu v} = 0 \tag{3.5.12}$$

From here, we obtain the famous **inhomogeneous Maxwell equation** as:

$$\sum_{v=0}^{3} \partial_v F_{\mu v} = \mu_0 J_\mu \tag{3.5.13}$$

and

$$\sum_{v=0}^{3}\sum_{v=0}^{3} \partial_\mu\partial_v F_{\mu v} = \mu_0 \sum_{v=0}^{3} \partial_\mu J_\mu \tag{3.5.14}$$

Gauss's law can be written as:

$$\sum_{i=0}^{3} \partial_i F_{0i} = \mu_0 J_0 \tag{3.5.15}$$

To summarize: quantum mechanics and electrodynamics are invariant under active global gauge transformations. Hence, the global gauge symmetry is a real symmetry. However, only our description is invariant under passive local transformations. Therefore, local gauge symmetry is a redundancy.

In particular, this means that whenever we perform a local passive transformation, we have to accompany:

$$\Psi \to \mathrm{e}^{i\varepsilon(x)}\Psi \tag{3.5.16}$$

always with

$$A_\mu \to A_\mu + \partial_\mu \epsilon(x) \tag{3.5.17}$$

When we perform a passive transformation these two transformations always go hand in hand. In contrast, there is no need to change A_μ if we only have global gauge transformation:

$$A_\mu(x) \to A_\mu(x) + a_\mu$$

in electrodynamics, and

$$\Psi(x) \to \mathrm{e}^{i\varepsilon}\Psi(x)$$

in quantum mechanics.

Chapter 4

Tensor Calculus

4.1 Total and Partial Derivatives

To understand the different kinds of derivatives, let's say we have a function $\rho(t, x(t), p(t))$ which, in general, depends on the location $x(t)$ and momentum $p(t)$ plus the time t. A key observation is that the location $x(t)$ and momentum $p(t)$ are functions of t too. Therefore, we need to be extremely careful what we mean when we calculate the derivative with respect to the time t.

$$\frac{d\rho}{dt} = \lim_{\Delta t \to 0} \frac{\rho(t + \Delta t, x(t + \Delta t), p(t + \Delta t)) - \rho(t, x(t), p(t))}{\Delta t}$$

The result is the total rate of change of ρ.

$$\frac{\partial\rho}{\partial t} = \lim_{\Delta t \to 0} \frac{\rho(t + \Delta t, x(t), p(t)) - \rho(t, x(t), p(t))}{\Delta t}$$

The key difference is that we only vary t if it appears explicitly in ρ but not if it only appears implicitly because $x(t)$ and $p(t)$ also depend on t. Thus

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial x}\frac{dx}{dt} + \frac{\partial\rho}{\partial p}\frac{dp}{dt} + \frac{\partial\rho}{\partial t}$$

4.2 Taylor Expansion

In general, we want to estimate the value of some function $f(x)$ at some value of x by using our knowledge of the function's value at some fixed point a. The Taylor series then reads

$$\begin{aligned} f(x) =& \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!} \\ =& \frac{f^{(0)}(a)(x-a)^0}{0!} + \frac{f^{(1)}(a)(x-a)^1}{1!} + \frac{f^{(2)}(a)(x-a)^2}{2!} \\ &+ \frac{f^{(3)}(a)(x-a)^3}{3!} + \dots \end{aligned} \tag{4.2.1}$$

or

$$f(x+a) = f(x) + (a\cdot\partial)f(x) + \frac{1}{2}(a\cdot\partial)^2 f(x) + \cdots \tag{4.2.2}$$

Taylor expansion of a scalar field (function f that maps $\mathbb{R}^n$ to $\mathbb{R}$).Now, identify $\partial f/\partial t$ as $\hat{\boldsymbol{n}}\cdot\nabla f$. In addition, see that $t\hat{\boldsymbol{n}} = \boldsymbol{x} - \boldsymbol{x}_0$. Some clever recombining of terms gives

$$f(x) = f\left(x_0\right) + \left(x - x_0\right)\cdot\nabla f|_{x_0} + \frac{1}{2}\left(\left[x - x_0\right]\cdot\nabla\right)^2 f\Big|_{x_0} + \dots \tag{4.2.3}$$

and

$$\hat{\boldsymbol{n}}\cdot\nabla = \partial_t \tag{4.2.4}$$

4.3 Vector Identities

$$\begin{aligned} \vec{\nabla}\cdot(\vec{\nabla}\times\vec{A}) &\equiv \operatorname{div}(\operatorname{rot}\vec{A}) = (\vec{\nabla}\times\vec{\nabla})\cdot\vec{A} \equiv 0 \\ \vec{\nabla}\times(\vec{\nabla}\varphi) &\equiv \operatorname{rot}\operatorname{grad}\varphi = (\vec{\nabla}\times\vec{\nabla})\varphi \equiv 0 \end{aligned} \tag{4.3.1}$$

$$\vec{\nabla}\cdot(\vec{A}\varphi) = \varphi\vec{\nabla}\cdot\vec{A} + \vec{A}\cdot\vec{\nabla}\varphi \quad\Longleftrightarrow\quad \operatorname{div}(\vec{A}\varphi) = \varphi\operatorname{div}\vec{A} + \vec{A}\cdot\operatorname{grad}\varphi$$

$$\vec{\nabla}\times(\vec{A}\varphi) = \varphi\vec{\nabla}\times\vec{A} - \vec{A}\times\vec{\nabla}\varphi \quad\Longleftrightarrow\quad \operatorname{rot}(\vec{A}\varphi) = \varphi\operatorname{rot}\vec{A} - \vec{A}\times\operatorname{grad}\varphi$$

$$\vec{\nabla}\cdot(\vec{A}\times\vec{B}) = \vec{B}\cdot(\vec{\nabla}\times\vec{A}) - \vec{A}\cdot(\vec{\nabla}\times\vec{B}) \quad\Longleftrightarrow\quad \operatorname{div}(\vec{A}\times\vec{B}) = \vec{B}\cdot\operatorname{rot}\vec{A} - \vec{A}\cdot\operatorname{rot}\vec{B} \tag{4.3.2}$$

$$\begin{aligned} \vec{\nabla}\times(\vec{A}\times\vec{B}) &= (\vec{B}\cdot\vec{\nabla})\vec{A} - (\vec{A}\cdot\vec{\nabla})\vec{B} + \vec{A}(\vec{\nabla}\cdot\vec{B}) - \vec{B}(\vec{\nabla}\cdot\vec{A}) \\ &\Longleftrightarrow \operatorname{rot}(\vec{A}\times\vec{B}) = (\vec{B}\operatorname{grad})\vec{A} - (\vec{A}\operatorname{grad})\vec{B} + \vec{A}(\operatorname{div}\vec{B}) - \vec{B}(\operatorname{div}\vec{A}) \end{aligned} \tag{4.3.3}$$

$$\begin{aligned}\vec{\nabla}(\vec{A}\cdot\vec{B}) =&(\vec{B}\cdot\vec{\nabla})\vec{A}+(\vec{A}\cdot\vec{\nabla})\vec{B}+\vec{A}\times(\vec{\nabla}\times\vec{B})+\vec{B}\times(\vec{\nabla}\times\vec{A})\\ &\Longleftrightarrow \operatorname{grad}(\vec{A}\cdot\vec{B})=(\vec{B}\cdot\operatorname{grad})\vec{A}+(\vec{A}\cdot\operatorname{grad})\vec{B}+\vec{A}\times\operatorname{rot}\vec{B}+\vec{B}\times\operatorname{rot}\vec{A}\end{aligned} \tag{4.3.4}$$

$$\begin{aligned}&\vec{\nabla}\cdot(\vec{\nabla}\varphi)\equiv\operatorname{div}(\operatorname{grad}\varphi)\equiv\Delta\varphi=\frac{\partial^2\varphi}{\partial x^2}+\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2},\quad \Delta=\text{ Laplace Operator}\\ &\vec{\nabla}\times(\vec{\nabla}\times\vec{A})\equiv\operatorname{rot}(\operatorname{rot}\vec{A})=\vec{\nabla}(\vec{\nabla}\cdot\vec{A})-(\vec{\nabla}\cdot\vec{\nabla})\vec{A}\equiv\operatorname{grad}\operatorname{div}\vec{A}-\Delta\vec{A}\end{aligned} \tag{4.3.5}$$

4.4 Linear Spaces, Vectors, and Tensors

The linear space, say L, consists of the elements (vectors) that **permit linear operations with the properties described below**:

- Summing up the vectors
- Multiplication by a number

Consider the set of vectors, elements of the linear space L

$$\{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\}=\{\mathbf{a}_i|i=1,\ldots,n\}$$

and the set of real numbers $k_1,k_2,\ldots k_n$. Vector

$$\mathbf{k}=\sum_{i=1}^{n}k^i\mathbf{a}_i=\mathbf{k}^i\mathbf{a}_i=0\Longrightarrow\sum_{i=1}^{n}\left(k^i\right)^2=0$$

then vectors $\{\mathbf{a_i},i=1,\ldots,n\}$ are called **linearly independent**. The last condition means that no one of the coefficients k_i can be different from zero.

For example, in 2D, two vectors are linearly dependent if and only if they are parallel; in 3D, three vectors are linearly dependent if and only if they belong to the same plane, etc.

Δ Definition

The maximal number of linearly independent elements of the linear space L is called its dimension. It proves useful to denote L_D a linear space of dimension D.

Δ Theorem

Consider a D-dimensional linear space L_D and the set of linearly independent vectors $\mathbf{e}_i = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_D)$. Then, for any vector a one can write

$$\mathbf{a} = \sum_{i=1}^{D} a^i \mathbf{e}_i = a^i \mathbf{e}_i \tag{4.4.1}$$

where the coefficients a^i are defined in a unique way.

The coefficients a^i are called components or **contravariant components of the vector a**.The word "contravariant" here means that the components a^i have upper indices.

4.5 Direct product of the two linear spaces

Let's consider the example of phase space in the classic mechanics. The coordinate system in the linear space L_D consists of the initial point O and the basis $\mathbf{e}_i$. The position of a point P can be characterized by **its position vector or radius vector** $\mathbf{r} = \overrightarrow{OP} = x^i \mathbf{e}_i$.

If the phase space is describing one particle, the space is composed of the radius vectors $\mathbf{r}$ and velocities $\mathbf{v}$ of the particle:

$$\mathbf{r} = x^i \mathbf{e}_i, \quad \mathbf{v} = v^j \mathbf{f}_j$$

The two set of bases can be related and independent. This example is a particular case of the linear space which is called a direct product of the two linear spaces. The element of the direct product of the two linear spaces L_{D_1} and L_{D_2} (**they can have different dimensions**) is the ordered set of the elements of each of the two spaces L_{D_1} and L_{D_2}.

The notation for the basis in the case of configuration space is $\mathbf{e}_i \otimes \mathbf{f}_j$. Hence, the state of the point-like particle in the phase space is characterized by the element of this linear space, which can be presented as:

$$x^i v^j \mathbf{e}_i \otimes \mathbf{f}_j$$

In general, one can define the space that is a direct product of several linear spaces with different individual basis sets $\mathbf{e}_i$ each. In this case, we will have $\mathbf{e}_i^{(1)}, \mathbf{e}_i^{(2)}, \ldots \mathbf{e}_i^{(N)}$. The basis in the direct product space will be

$$\mathbf{e}_{i_1}^{(1)} \otimes \mathbf{e}_{i_2}^{(2)} \otimes \ldots \otimes \mathbf{e}_{i_N}^{(N)}$$

and the element becomes

$$T^{i_1 i_2 \ldots i_N} \mathbf{e}_{i_1}^{(1)} \otimes \mathbf{e}_{i_2}^{(2)} \otimes \ldots \otimes \mathbf{e}_{i_N}^{(N)}$$

4.6 Vector basis and its transformation

Let us start from the components of the vector, which were defined in (4.4.1). Consider, along with the original basis $\mathbf{e}_i$, another basis $\mathbf{e}'_i$. since each vector of the new basis belongs to the same space, it can be expanded using the original basis as

$$\mathbf{e}'_i = \wedge^j_{i'} \mathbf{e}_j \tag{4.6.1}$$

and

$$\mathbf{a} = a^i \mathbf{e}_i = a^{j'} \mathbf{e}_{j'} = a^{j'} \wedge^i_{j'} \mathbf{e}_i$$

$$a^i = a^{j'} \wedge^i_{j'} \tag{4.6.2}$$

Similarly, we can make inverse transformation:

$$a^{k'} = \left(\wedge^{-1}\right)^{k'}_l a^l$$

and

$$\left(\wedge^{-1}\right)_{l}^{k'} \cdot \wedge_{i'}^{l} = \delta_{i'}^{k'} \quad \text{and} \quad \wedge_{i'}^{l} \cdot \left(\wedge^{-1}\right)_{k}^{i'} = \delta_{k}^{l}$$

Taking the partial derivatives, we arrive at the relations

$$\wedge_{j'}^{i} = \frac{\partial x^{i}}{\partial x^{j'}} \quad \text{and} \quad \left(\wedge^{-1}\right)_{l}^{k'} = \frac{\partial x^{k'}}{\partial x^{l}}$$

and

$$\left(\wedge^{-1}\right)_{l}^{k'} \cdot \wedge_{k'}^{i} = \frac{\partial x^{k'}}{\partial x^{l}} \frac{\partial x^{i}}{\partial x^{k'}} = \frac{\partial x^{i}}{\partial x^{l}} = \delta_{l}^{i}$$

is nothing but the chain rule for partial derivatives.

4.7 Scalar, vector, and tensor fields

Δ Definition

Function $\varphi(x)$ is called scalar field or simply scalar if it does not transform under the change of coordinates

$$\varphi(x) = \varphi'(x') \tag{4.7.1}$$

Let us give a clarifying example in 1D. Consider a function

$$y = x^2$$

Now, let us change the variables

$$x' = x + 1$$

The function $y = (x')^2$, obviously, represents another parabola. In order to preserve the plot intact, we need to modify the form of the function, that is, to go from φ to φ'. The new function $y' = (x' - 1)^2$ will represent the original parabola,because the change of the variable is completely compensated by the change of the form of the function.

Question

Discuss whether the three numbers temperature $T(x)$, pressure $p(x)$ and density $\rho(x)$ form a contravariant vector

Answer

We can only form contravariant vector if the numbers transform by following the rule (4.6.2). Since these numbers are scalar fields, it transforms like $T(\mathbf{r}) = T'(\mathbf{r}')$. Thus, these three parameters can not form a contravariant vector.

· **Example**

From the definition above we know $\varphi'(x) \neq \varphi(x)$ and $\varphi(x') \neq \varphi(x)$. Let us calculate these quantities explicitly for the special case of infinitesimal transformation $x'^i = x^i + \xi^i$ where ξ are constant coefficients. Now we have

$$\varphi(x') = \varphi(x+\xi) \overset{Taylor}{=} \varphi(x) + \frac{\partial\varphi}{\partial x^i}\xi^i$$

and,

$$\varphi'\left(x^i\right) = \varphi'\left(x^i - \xi^i\right) = \varphi'(x') - \frac{\partial\varphi'}{\partial x^{i'}} \cdot \xi^i$$

rewrite the two equations above, we find:

$$\frac{\partial\varphi'}{\partial x'^i} = \frac{\varphi'(x'^i) - \varphi'(x^i - \xi^i)}{\xi^i}$$

$$\frac{\partial\varphi}{\partial x^i} = \frac{\varphi(x'^i) - \varphi(x^i)}{\xi^i}$$

At the limit of $\xi \to 0$, we have

$$\frac{\partial\varphi'}{\partial x'^i} = \frac{\partial\varphi(x)}{\partial x^i} + \mathcal{O}(\xi)$$

As mentioned above,

$$\varphi'\left(x^i\right) = \varphi'(x') - \frac{\partial\varphi'}{\partial x'^i} \cdot \xi^i = \varphi'(x') - \frac{\partial\varphi(x)}{\partial x^i} \cdot \xi^i$$

Since $\varphi(x) = \varphi'(x')$, we have

$$\varphi'(x) = \varphi(x) - \xi^i\partial_i\varphi$$

where we have introduced a useful notation $\partial_i = \partial/\partial x^i$.

For the vector field, we have the following rule for the coordinate transformation.

The set of three functions $\{a^i(x)\} = \{a^1(x), a^2(x), a^3(x)\}$ forms contravariant vector field (or simply vector field) if they transform, under the change of coordinates $\{x^i\} \to \{x'j\}$, as

$$a^{j'}(x') = \frac{\partial x^{j'}}{\partial x^i} \cdot a^i(x) \tag{4.7.2}$$

The components of the vector in a given geometrical point of space modify under the coordinate transformation, while the scalar field does not.The scalar and vector fields can be considered as examples of the more general objects called tensors. Tensors are also defined through their transformation rules.

The set of 3^n functions $\{a^{i_1\cdots i_n}(x)\}$ is called a contravariant tensor of rank n, if these functions transform, under $x^i \to x^i$, as

$$a^{i'_1\cdots i'_n}(x') = \frac{\partial x^{i'_1}}{\partial x^{j_1}} \cdots \frac{\partial x^{i'_n}}{\partial x^{j_n}} a^{j_1\cdots j_n}(x) \tag{4.7.3}$$

4.8 Orthonormal Basis and Cartesian Coordiantes

The scalar product of two vectors **a** and **b** in 3D is defined in a ususal way,

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \cos\theta \tag{4.8.1}$$

Special orthonormal basis $\{\hat{\mathbf{n}}_a\}$ is the one with

$$(\hat{\mathbf{n}}_a, \hat{\mathbf{n}}_b) = \delta_{ab} = \begin{cases} 1 \text{ if } a = b \\ 0 \text{ if } a \neq b \end{cases} \tag{4.8.2}$$

· **Example**

Making transformations of the basis vectors, verify that the change of coordinates

$$x' = \frac{x+y}{\sqrt{2}} + 3, \quad y' = \frac{x-y}{\sqrt{2}} - 5$$

does not modify the type of coordinates x', y', which remains Cartesian. **Solution**:

$$x' = \frac{x+y}{\sqrt{2}} + 3, \quad y' = \frac{x-y}{\sqrt{2}} - 5$$

The change of initial point $(0,0) \to (3,-5)$ does not have relation to the change of basis. Then

$$\begin{aligned} x'\hat{i}' + y'\hat{j}' &= \left(\frac{x+y}{\sqrt{2}}\right)\hat{i}' + \left(\frac{x-y}{\sqrt{2}}\right)\hat{j}' \\ &= \frac{x}{\sqrt{2}}\left(\hat{i}' + \hat{j}'\right) + \frac{y}{\sqrt{2}}\left(\hat{i}' - \hat{j}'\right) = x\hat{i} + y\hat{j} \end{aligned}$$

Thus, $\hat{i} = \frac{1}{\sqrt{2}}\left(\hat{i}' + \hat{j}'\right), \hat{j} = \frac{1}{\sqrt{2}}\left(\hat{i}' - \hat{j}'\right)$. Obviously, $\hat{i}^2 = \hat{j}^2 = 1$ and $\hat{i} \cdot \hat{j} = 0$

Now we can introduce a conjugated covariant basis.

Consider basis $\{\mathbf{e}_i\}$. The conjugated basis is defined as a set of vectors $\{\mathbf{e}^j\}$ which satisfy the relations

$$\mathbf{e}_i \cdot \mathbf{e}^j = \delta_i^j \tag{4.8.3}$$

The special property of the orthonormal basis is that $\hat{\mathbf{n}}^a = \hat{\mathbf{n}}_a$.

Any vector a can be expanded using the conjugated basis $\mathbf{a} = a_i \mathbf{e}^i$. The coefficients a_i are called covariant components of the vector $\mathbf{a}$.

In the case of covariant vector components, the transformation is done by means of the matrix inverse to the one for the contravariant components.

Δ Theorem

If we change the basis of the coordinate system from $\mathbf{e}_i$ to $\mathbf{e}'_i$, then the covariant components of the vector a transform as

$$a'_i = \frac{\partial x^j}{\partial x^i} a_j \tag{4.8.4}$$

The set of three functions $\{A_i(x)\}$ forms a covariant vector field, if they transform from one coordinate system to another one as

$$A'_i(x') = \frac{\partial x^j}{\partial x^i} A_j(x) \tag{4.8.5}$$

The set of 3^n functions $\{A_{i_1 i_2 \ldots i_n}(x)\}$ form a covariant tensor of rank n if they transform from one coordinate system to another as

$$A'_{i_1 i_2 \ldots i_n}(x') = \frac{\partial x^{j_1}}{\partial x^{i_1}} \frac{\partial x^{j_2}}{\partial x^{i_2}} \cdots \frac{\partial x^{j_n}}{\partial x^{n_n}} A_{j_1 j_2 \ldots j_n}(x) \tag{4.8.6}$$

In general

The set of 3^{n+m} functions $\{B_{i_1 \ldots i_n} j_{1 \ldots j_n}(x)\}$ forms the tensor of the type (m, n), if these functions transform, under the change of coordinate basis, as

$$B'^{j'_1 \ldots j'_m}_{i'_1 \ldots i'_n}(x') = \frac{\partial x^{j'_1}}{\partial x^{l_1}} \cdots \frac{\partial x^{j'_m}}{\partial x^{l_m}} \frac{\partial x^{k_1}}{\partial x'^{i_1}} \cdots \frac{\partial x^{k_n}}{\partial x'^{i_n}} B^{l_1 \ldots l_m}_{k_1 \ldots k_n}(x) \tag{4.8.7}$$

Other possible names are the mixed tensor of covariant rank n and contravariant rank m, or simply (m, n) -tensor.

Tensors are important due to the fact that they offer the coordinate-independent description of geometrical and physical laws. The following example shows this observation:

· **Example**

For an arbitrary $\mathbf{a}(x) = a^i(x)\mathbf{e}_i$ we have

$$a^{j'}(x') = \frac{\partial x^{j'}}{\partial x^i} a^i(x), \quad \mathbf{e}_{j'}(x') = \mathbf{e}_k(x) \frac{\partial x^k}{\partial x^{j'}}$$

Then

$$a^{j'}(x')\, \mathbf{e}_{j'}(x') = \frac{\partial x^{j'}}{\partial x^i} a^i(x) \mathbf{e}_k(x) \frac{\partial x^k}{\partial x^{j'}} = \delta^k_l a^i(x) \mathbf{e}_k(x) = a^i(x) \mathbf{e}_i(x)$$

If the Kronecker symbol transforms as a mixed $(1, 1)$ tensor,

$$\delta^{i'}_{j'} = \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^l}{\partial x^{j'}} \delta^k_l = \frac{\partial x^{i'}}{\partial x^{j'}} \tag{4.8.8}$$

then in any coordinates x^i it has the same form

$$\delta_j^i = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

This property is very important, as it enables us to use the Kronecker symbol in any coordinates

· **Example**

Show that the product $A^i(x)B_j(x)$ of covariant and contravariant vectors transforms as a $(1,1)$ -type mixed tensor.

$$A^{i'}(x')\,A_{j'}(x') = \frac{\partial x^{i'}}{\partial x^k}A^k(x)\frac{\partial x^l}{\partial x^{j'}}B_l(x) = \frac{\partial x^{i'}}{\partial x^k}\frac{\partial x^l}{\partial x^{j'}}A^k(x)B_l(x)$$

4.9 Orthogonal transformation

The rotation transformation around $\hat{\mathbf{z}}-$axis is given by the following relation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \hat{\Lambda}_z \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \text{where} \quad \hat{\Lambda}_z = \hat{\Lambda}_z(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.9.1}$$

the matrix above has the following property:

$$\hat{\Lambda}_z^T = \hat{\Lambda}_z^{-1} \tag{4.9.2}$$

Δ **Definition**

The matrix $\hat{\Lambda}_z$ which satisfies $\hat{\Lambda}_z^{-1} = \hat{\Lambda}_z^T$ and the corresponding coordinate transformation is called **orthogonal**.

Similarly, we can write the rotation matrix around other axis:

$$\hat{\Lambda}_x(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix} \tag{4.9.3}$$

$$\hat{\Lambda}_y(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \tag{4.9.4}$$

In 3D space, any rotation of the rigid body may be represented as a combination of the rotation around the axes $\hat{\mathbf{z}}, \hat{\mathbf{y}}$, and $\hat{\mathbf{x}}$ to the angles α, β, and γ:

$$\hat{\Lambda} = \hat{\Lambda}_z(\alpha)\hat{\Lambda}_y(\beta)\hat{\Lambda}_x(\gamma)$$

Since $(A \cdot B)^T = B^T \cdot A^T$ and $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$, we can easily obtain that the general 3D rotation matrix satisfies the orthogonal relation.

For the orthogonal matrix, one can take the determinant and arrive at $det\hat{\Lambda} = det\hat{\Lambda}^{-1}$. Therefore,

$$det\hat{\Lambda} = \pm 1$$

As far as any rotation matrix has a determinant equal to one, there must be some other orthogonal matrices with the determinant equal to -1.

In case where the matrix elements are allowed to be complex, the matrix that satisfies the property $U^\dagger = U^{-1}$ is called unitary. The operation $U^\dagger$ is called **Hermitian conjugation** and consists of complex conjugation plus transposition $U^\dagger = (U^*)^T$.

· **Example**

Consider the transformation from one orthonormal basis $\hat{\mathbf{n}}_i$ to another such basis $\hat{\mathbf{n}}'_k$, namely, $\hat{\mathbf{n}}'_k = R^i_k \hat{\mathbf{n}}_i$. Prove that the matrix $\|R^i_k\|$ is orthogonal.

$\hat{\mathbf{n}}'_k = R^i_k \hat{\mathbf{n}}_i$ and also $\hat{\mathbf{n}}'' = (R^{-1})^l_j \hat{\mathbf{n}}^j$. since $\hat{\mathbf{n}}'_k = \hat{\mathbf{n}}^{k'}$ and $\hat{\mathbf{n}}_i = \hat{\mathbf{n}}^i$ also have $(R^{-1})^l_j = \left(R^T\right)^l_j$

4.10 Operations over Tensors, Metric Tensor

Multiplication of a tensor by a number produces a tensor of the same type. This operation is equivalent to the multiplication of all tensor components to the same number α, namely:

$$(\alpha A)_{i_1 \ldots i_n} = \alpha \cdot A^{j_1 \ldots j_m}_{i_1 \ldots i_n} \tag{4.10.1}$$

Multiplication of two tensors is defined for a couple of tensors of any type. The product of a (m, n)-tensor and a (t, s)-tensor results in the (m + t, n + s)- tensor, e.g.,

$$A_{i_1 \ldots i_n}{}^{j_1 \cdots j_m} \cdot C_{l_1 \ldots l_s}{}^{k_1 \ldots k_i} = D_{i_1 \ldots i_n}{}^{j_1 \cdots j_m}{}_{l_1 \ldots l_s}{}^{k_1 \ldots k_l} \tag{4.10.2}$$

The order of indices is important here, because a_{ij} may be different from a_{ji}.

· **Example**

Prove, by checking the transformation law, that the product of the contravariant vector a^i and mixed tensor b^k_i is a mixed $(2,1)$ -type tensor.

$$a^{i'}(x')\, b^{k'}_{j'}(x') = \frac{\partial x^{i'}}{\partial x^m} a^m(x) \frac{\partial x^{k'}}{\partial x^l} \frac{\partial x^n}{\partial x^{j'}} b^l_n(x)$$

Contraction reduces the (n,m)-tensor to the (n-1,m-1)-tensor through the summation over two (always upper and lower, of course) indices. For example,

$$A^{ln}_{ijk} \longrightarrow A^{lk}_{ijk} = \sum_{k=1}^{3} A^{lk}_{ijk} \tag{4.10.3}$$

The internal product of the two tensors consists in their multiplication with the consequent contraction over some couple of indices. **Internal product of (m, n) and (r, s)-type tensors results in the (m+r-1, n+s-1)-type tensor.**

$$A_{ijk} \cdot B^{lj} = \sum_{j=1}^{3} A_{ijk} B^{lj} \tag{4.10.4}$$

· **Example**

Prove that the internal product $a_i \cdot b^i$ is a scalar if $a_i(x)$ and $b^i(x)$ are co- and contravariant vectors.

$$b^{i'}(x') = \frac{\partial x^{i'}}{\partial x^l} b^l(x), \quad a_{i'}(x') = \frac{\partial x^k}{\partial x^{i'}} a_k(x)$$

Then

$$a_{i'}(x')\, b^{i'}(x') = \frac{\partial x^{i'}}{\partial x^l} \frac{\partial x^k}{\partial x^{i'}} b^l(x) a_k(x) = \delta_l^k b^l(x) a_k(x) = b^k(x) a_k(x)$$

Δ **Definition**

Consider a basis $\{\mathbf{e}_i\}$. The **scalar product** of the two basis vectors,

$$g_{ij} = (\mathbf{e}_i, \mathbf{e}_j) \tag{4.10.5}$$

is called **metric**. Here **the scalar product is not always inner product.**

Properties of metric:

1. Symmetry of the metric $g_{ij} = g_{ji}$ follows from the symmetry of a scalar product $(\mathbf{e}_i, \mathbf{e}_j) = (\mathbf{e}_j, \mathbf{e}_i)$
2. For the orthonormal basis $\hat{\mathbf{n}}_a$, the metric is nothing but the Kronecker symbol $g_{ab} = (\hat{\mathbf{n}}_a, \hat{\mathbf{n}}_b) = \delta_{ab}$
3. Metric is a (0, 2) - tensor, as

$$g_{i'j'} = \left(\mathbf{e}'_i, \mathbf{e}'_j\right) = \left(\frac{\partial x^l}{\partial x^{i'}} \mathbf{e}_l, \frac{\partial x^k}{\partial x'^j} \mathbf{e}_k\right)$$
$$= \frac{\partial x^l}{\partial x'^i} \frac{\partial x^k}{\partial x'^j} (\mathbf{e}_l, \mathbf{e}_k) = \frac{\partial x^l}{\partial x'^i} \frac{\partial x^k}{\partial x'^j} \cdot g_{kl}$$

4. The distance between two points: $M_1(x^i)$ and $M_2(y^i)$ is defined by the inner product:

$$S_{12}^2 = g_{ij} \left(x^i - y^i\right) \left(x^j - y^j\right)$$

Since g_{ij} is $(0, 2)$ - tensor and $(x^i - y^i)$ is $(1, 0)$ -tensor (contravariant vector) $-S_{12}^2$ is a scalar.

Therefore, S^2_{12} is the same in any coordinate system.

The conjugated metric is defined as

$$g^{ij} = \left(\mathbf{e}^i, \mathbf{e}^j\right) \quad \text{where} \quad \left(\mathbf{e}^i, \mathbf{e}_k\right) = \delta^i_k \tag{4.10.6}$$

$$g^{ik} g_{kj} = \delta^i_j \tag{4.10.7}$$

We can **Raise and lower indices of a tensor** by taking an appropriate internal product of a given tensor and the corresponding metric tensor.

$$\begin{aligned} &\text{Lowering the index, } A_i(x) = g_{ij} A^j(x), \quad B_{ik}(x) = g_{ij} B^j_k(x) \\ &\text{Raising the index, } \quad C^l(x) = g^{lj} C_j(x), \quad D^{ik}(x) = g^{ij} D^k_j(x) \end{aligned} \tag{4.10.8}$$

$$\mathbf{e}_i g^{ij} = \mathbf{e}^j, \quad \mathbf{e}^k g_{kl} = \mathbf{e}_l \tag{4.10.9}$$

Let the metric g_{ij} correspond to the basis ek and to the coordinates x^k. The determinants of the metric tensors and of the matrices of the transformations to Cartesian coordinates X^a satisfy the following relations:

$$g = \det\left(g_{ij}\right) = \det\left(\frac{\partial X^a}{\partial x^k}\right)^2, \quad g^{-1} = \det\left(g^{kl}\right) = \det\left(\frac{\partial x^l}{\partial X^b}\right)^2 \tag{4.10.10}$$

The possibility of lowering and raising the indices may help us in contracting two contravariant or two covariant indices of a tensor.

· **Example**

Suppose we need to contract the two first indices of the $(3,0)$ -tensor B^{ijk}. After we lower the second index, we arrive at the tensor

$$B^i{}_l{}^k = B^{ijk} g_{jl}$$

And now we can contract the indices i, j. But, if we forget to indicate the order of indices, we obtain:B^{ik}_l, and it is not immediately clear which index was lowered and in which couple of indices one has to perform the contraction.

4.11 Symmetric, Skew(Anti) Symmetric Tensors, and Determinants

4.11.1 Definitions and general considerations

Tensor $A^{ijk\cdots}$ is called **symmetric** in the indices i and j, if

$$A^{ijk\cdots} = A^{jik\cdots} \tag{4.11.1}$$

Tensor $A^{i_1 i_2 \cdots i_n}$ is called **completely (or absolutely) symmetric** in the indices $(i_1, i_2, \ldots, i_n)$, if it is symmetric in any couple of these indices.

Tensor A^{ij} is called skew-symmetric or antisymmetric, if

$$A^{ij} = -A^{ji} \tag{4.11.2}$$

The advantage of tensors is that their (anti)symmetry holds under the transformation from one basis to another.

· **Example**

Consider the basis in 2D,

$$\mathbf{e}_1 = \hat{\mathbf{i}} + \hat{\mathbf{j}}, \quad \mathbf{e}_2 = \hat{\mathbf{i}} - \hat{\mathbf{j}}$$

(i) Derive all components of the absolutely antisymmetric tensor ε^{ij} in the new basis, taking into account that $\varepsilon^{ab} = \epsilon^{ab}$, where $\epsilon^{12} = 1$ in the orthonormal basis $\hat{\mathbf{n}}_1 = \hat{\mathbf{i}}, \quad \hat{\mathbf{n}}_2 = \hat{\mathbf{j}}$. The calculation should be performed directly and also by using the formula for antisymmetric tensor. Explain the difference between the two results.

(ii) Repeat the calculation for ε_{ij}. Calculate the metric components and verify that $\varepsilon_{ij} = g_{ik} g_{jl} \varepsilon^{kl}$ and that $\varepsilon^{ij} = g^{ik} g^{jl} \varepsilon_{kl}$

Solution: From the basis relation, we have:

$$x' = x + y$$

$$y' = x - y$$

Thus,

$$\frac{\partial x^{j'}}{\partial x^i} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and $\varepsilon^{12} = \frac{\partial x'^1}{\partial x^k}\frac{\partial x'^2}{\partial x^l}\epsilon^{kl} = \epsilon^{21} - \epsilon^{12} = -2$. Since **the metric tensor of the orthonormal basis is**

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

we have $\epsilon_{ab} = \epsilon^{ab}$. Thus, $\varepsilon_{12} = \frac{\partial x^k}{\partial x'^1}\frac{\partial x^l}{\partial x'^2}\epsilon_{kl} = -1/2$. Also, the metric tensor for our new basis (**w.r.t. the orthonormalized vector basis**) is

$$g_{ij} = \frac{\partial x^a}{\partial x^i}\frac{\partial x^b}{\partial x^j}g_{ab} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

Using $g^{ik}g_{kj} = \delta^i_j$, we also have $g^{ij} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Obviously, we have:

$$\varepsilon_{ij} = g_{ik}g_{jl}\varepsilon^{kl}$$
$$\varepsilon^{ij} = g^{ik}g^{jl}\varepsilon_{kl}$$

4.11.2 Completely Antisymmetric Tensors

$(n, 0)$ -tensor $A^{i_1\dots i_n}$ is called completely (or absolutely) antisymmetric, if it is antisymmetric in any couple of its indices

$$\forall(k, l), \quad A^{i_1\dots i_l\dots i_k\dots i_n} = -A^{i_1\dots i_k\dots i_l\dots i_n}$$

In the case of an absolutely antisymmetric tensor, the sign changes when we perform permutation of any two indices.

- In a D-dimensional space, all completely antisymmetric tensors of rank $n > D$ are equal to zero.
- Theorem 2 An absolutely antisymmetric tensor $A^{i_1 \cdots i_D}$ in a D-dimensional space has only one independent component.
- Prove that for symmetric a^{ij} and antisymmetric b_{ij} tensors the scalar internal product is zero $a^{ij} \cdot b_{ij} = 0$

For the last point, we can use the following argument to prove it:

$$a^{ji} b_{ji} = \left(+a^{ij}\right)\left(-b_{ij}\right) \overset{i \to j, j \to i}{=} - a^{ij} b_{ij}$$

The quantity that equals itself with an opposite sign is zero.

· **Example**

(i) Prove that if $b^{ij}(x)$ is a tensor and $a_l(x)$ and $c_k(x)$ are covariant vector fields, then $f(x) = b^{ij}(x) \cdot a_i(x) \cdot c_j(x)$ is a scalar field.

(ii) In case $a_i(x) \equiv c_i(x)$, formulate the sufficient condition for $b^{ij}(x)$, such that the product $f(x) \equiv 0$

Solution:

(i) $b^{ij}(x) \cdot a_i(x) \cdot c_j(x) = d^{ij}{}_{ij}(x)$, which is a scalar.

(ii) when b is anti-symmetric: $b^{ij} = b^{[ij]}$

Determinants

Consider first $2D$ and special coordinates X^1 and X^2 corresponding to the orthonormal basis $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$. We first define a special maximal absolutely antisymmetric object in 2D using the relations:

$$E_{ab} = -E_{ba} \quad \text{and} \quad E_{12} = 1$$

For and matrix $\|C_b^a\|$ we can write its determinant as

$$\det \|C_b^a\| = E_{ab} C_1^a C_2^b \tag{4.11.3}$$

and

$$\det (C_b^a) = \frac{1}{2} E_{ab} E^{de} C_d^a C_e^b \tag{4.11.4}$$

The difference between these two equations is that the latter admits two expressions $C_1^a C_2^b$ and $C_2^a C_1^b$. Thus, we have the $1/2$ factor in the second equation.

Because E_{ab} is maximal absolutely anti-symmetric, we also have:

$$E_{ab} C_e^a C_d^b = -E_{ab} C_e^b C_d^a \tag{4.11.5}$$

Proof

$$E_{ab} C_e^a C_d^b = E_{ba} C_e^b C_d^a = -E_{ab} C_d^a C_e^b$$

Here in the first equality, we have exchanged the names of the umbral(dummy) indices $a \leftrightarrow b$ and in the second one used antisymmetry $E_{ab} = -E_{ba}$. Now we consider an arbitrary dimension of space D.

Consider $a, b = 1, 2, 3, \ldots, D$. $\forall$ matrix $\|C_b^a\|$, the determinant is

$$\det \|C_b^a\| = E_{a_1 a_2 \ldots a_D} \cdot C_1^{a_1} C_2^{a_2} \cdots C_D^{a_D} \tag{4.11.6}$$

where $E_{a_1 a_2 \ldots a_D}$ is absolutely antisymmetric and $E_{123\ldots D} = 1$. The determinant in the equation above is a sum of D! terms, each of which is a product of elements from different lines and columns, taken with the positive sign for even and with the negative sign for odd parity of permutations.

In general, the repression

$$\mathcal{A}_{a_1 \ldots a_D} = E_{b_1 \ldots b_D} A_{a_1}^{b_1} A_{a_2}^{b_2} \ldots A_{a_D}^{b_D} \tag{4.11.7}$$

is absolutely antisymmetric in the indices $\{a_1, \ldots, a_D\}$ and therefore is proportional to $E_{a_1 \ldots a_D}$.

In an arbitrary dimension D

$$\det \|C_b^a\| = \frac{1}{D!} E_{a_1 a_2 \ldots a_D} E^{b_1 b_2 \ldots b_D} \cdot C_{b_1}^{a_1} C_{b_2}^{a_2} \ldots C_{b_D}^{a_D} \tag{4.11.8}$$

and

$$E^{a_1 a_2 \ldots a_D} \cdot E_{b_1 b_2 \ldots b_D} = \begin{vmatrix} \delta_{b_1}^{a_1} & \delta_{b_2}^{a_1} & \ldots & \delta_{b_D}^{a_1} \\ \delta_{b_1}^{a_2} & \delta_{b_2}^{a_2} & \ldots & \delta_{b_D}^{a_2} \\ \ldots & \cdots & \cdots & \ldots \\ \delta_{b_1}^{a_D} & \ldots & \cdots & \delta_{b_D}^{a_D} \end{vmatrix} \tag{4.11.9}$$

Following the theorem above, we have for the special dimension 3D:

$$E^{abc} E_{def} = \begin{vmatrix} \delta_d^a & \delta_e^a & \delta_f^a \\ \delta_d^b & \delta_e^b & \delta_f^b \\ \delta_d^c & \delta_e^c & \delta_f^c \end{vmatrix}$$

Now, let's contract the indices c, f:

$$E^{abc} E_{dec} = \begin{vmatrix} \delta_d^a & \delta_e^a & \delta_c^a \\ \delta_d^b & \delta_e^b & \delta_c^b \\ \delta_d^c & \delta_e^c & 3 \end{vmatrix} = \delta_d^a \delta_e^b - \delta_e^a \delta_d^b \tag{4.11.10}$$

The last formula is sometimes called magic, as it is an extremely useful tool for the calculations in analytic geometry and vector calculus. Let us proceed and contract indices b and e. We obtain

$$E^{abc} E_{dbc} = 3\delta_d^a - \delta_d^a = 2\delta_d^a$$

Finally, contracting the last remaining couple of indices, we arrive at

$$E^{abc} E_{abc} = 6$$

In general

$$E^{a_1a_2\ldots a_D}E_{a_1a_2\ldots a_D} = D! \tag{4.11.11}$$

Because for the first index, we have D choices, for the second $D-1$ choices, for the third $D-2$ choices, etc.

· **Example**

Using definition of the maximal antisymmetric symbol $E^{a_1a_2\ldots a_D}$ prove the rule for the product of matrix determinants:

$$\det(A\cdot B) = \det A\cdot\det B$$

Here both $A = \|a^i_k\|$ and $B = \|b^i_k\|$ are $n\times n$ matrices.

Solution:

$$\det(A\cdot B) = E_{i_1i_2\ldots i_D}(A\cdot B)^{i_1}_1(A\cdot B)^{i_2}_2\ldots(A\cdot B)^{i_D}_D$$

expand, we have

$$\det(A\cdot B) = E_{i_1i_2\ldots i_D}a^{i_1}_{k_1}a^{i_2}_{k_2}\ldots a^{i_D}_{k_D}b^{k_1}_1b^{k_2}_2\ldots b^{k_D}_D$$

According to (5.11.8), we have

$$E_{i_1i_2\ldots i_D}a^{i_1}_{k_1}a^{i_2}_{k_2}\ldots a^{i_D}_{k_D} = E_{k_1k_2\ldots k_D}\cdot\det A$$

Therefore,

$$\det(A\cdot B) = \det AE_{k_1k_2\ldots k_D}b^{k_1}_1b^{k_2}_2\ldots b^{k_D}_D$$

Finally, let's consider two more statements

· **Lemma**

For any nondegenerate $D\times D$ matrix M^a_b with the determinant M, the elements of the inverse matrix $(M^{-1})^b_c$ are given by the expressions

$$\left(M^{-1}\right)^b_a = \frac{1}{M(D-1)!}E_{aa_2\ldots a_D}E^{bb_2\ldots b_D}M^{a_2}_{b_2}M^{a_3}_{b_3}\ldots M^{a_D}_{b_D} \tag{4.11.12}$$

Consider the nondegenerate $D\times D$ matrix $\|M^a_b\|$ with the elements $M^a_b = M^a_b(\kappa)$ being functions of some parameter κ.In general, the determinant of this matrix $M = \det\|M^a_b\|$

also depends on κ. Suppose all functions $M^a_b(\kappa)$ are differentiable. Then the derivative of the determinant equals

$$\dot{M} = \frac{dM}{d\kappa} = M\left(M^{-1}\right)^b_a \frac{dM^a_b}{d\kappa} \tag{4.11.13}$$

where $(M^{-1})^b_a$ are the elements of the matrix imverse to (M^a_b) and the dot over a function indicates its derivative with respect to κ.

4.11.3 Applications to Vector Algebra

Let's come back to the following definition for a **special object**:

$$E^{abc}E_{dec} = \begin{vmatrix} \delta^a_d & \delta^a_e & \delta^a_c \\ \delta^b_d & \delta^b_e & \delta^b_c \\ \delta^c_d & \delta^c_e & 3 \end{vmatrix} = \delta^a_d\delta^b_e - \delta^a_e\delta^b_d$$

We first working on a special orthonormal basis $\hat{\mathbf{n}}_a = (\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$, corresponding to the Cartesian coordinates X^a. In a vector product notations, we have

$$[\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2] = \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2 = \hat{\mathbf{n}}_3$$

$$[\hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3] = \hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_3 = \hat{\mathbf{n}}_1$$

$$[\hat{\mathbf{n}}_3, \hat{\mathbf{n}}_1] = \hat{\mathbf{n}}_3 \times \hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2$$

Using the previously defined E^{abc}, we can write this in a more compact way:

$$[\hat{\mathbf{n}}_a, \hat{\mathbf{n}}_b] = E_{abc} \cdot \hat{\mathbf{n}}^c \tag{4.11.14}$$

For orthonormal basis, $\hat{\mathbf{n}}_a = \hat{\mathbf{n}}^a$ and $E^{abc} = E_{abc}$. In general, if we consider two vectors $\mathbf{V} = V^a\hat{\mathbf{n}}_a$ and $\mathbf{W} = W^a\hat{\mathbf{n}}_a$, then

$$[\mathbf{V}, \mathbf{W}] = E_{abc} \cdot V^a \cdot W^b \cdot \hat{\mathbf{n}}^c$$

Using this object, we can solve many problems of vector algebra:

· **Example**

Consider the mixed product of the three vectors

$$(\mathbf{U},\mathbf{V},\mathbf{W}) = (\mathbf{U},[\mathbf{V},\mathbf{W}]) = U^a \cdot [\mathbf{V},\mathbf{W}]_a$$

$$= U^a \cdot E_{abc}V^bW^c = E_{abc}U^aV^bW^c = \begin{vmatrix} U^1 & U^2 & U^3 \\ V^1 & V^2 & V^3 \\ W^1 & W^2 & W^3 \end{vmatrix}$$

The similar arguments can be made to prove the following properties of the **mixed product**:

(i) Cyclic identity: $(\mathbf{U},\mathbf{V},\mathbf{W}) = (\mathbf{W},\mathbf{U},\mathbf{V}) = (\mathbf{V},\mathbf{W},\mathbf{U})$

(ii) Antisymmetry: $(\mathbf{U},\mathbf{V},\mathbf{W}) = -(\mathbf{V},\mathbf{U},\mathbf{W}) = -(\mathbf{U},\mathbf{W},\mathbf{V})$

· **Example**

Prove $[\mathbf{U},\mathbf{V}] \times [\mathbf{W},\mathbf{Y}] = \mathbf{V}\cdot(\mathbf{U},\mathbf{W},\mathbf{Y}) - \mathbf{U}\cdot(\mathbf{V},\mathbf{W},\mathbf{Y})$

$$([\mathbf{U},\mathbf{V}] \times [\mathbf{W},\mathbf{Y}])_a = E_{abc}[\mathbf{U},\mathbf{V}]^b[\mathbf{W},\mathbf{Y}]^c = E_{abc}E^{bde}U_dV_eE^{cfg}W_fY_g$$

$$= -E_{bac}E^{bde}E^{cfg}U_dV_eW_fY_g = -\left(\delta^d_a\delta^e_c - \delta^e_a\delta^d_c\right)E^{cfg}U_dV_eW_fY_g$$

$$= E^{cfg}\left(U_cV_aW_fY_g - U_aV_cW_fY_g\right) = V_a(\mathbf{U},\mathbf{W},\mathbf{Y}) - U_a(\mathbf{V},\mathbf{W},\mathbf{Y})$$

Some more identities

$$[[\mathbf{U},\mathbf{V}],[\mathbf{W},\mathbf{Y}]] = \mathbf{W}\cdot(\mathbf{Y},\mathbf{U},\mathbf{V}) - \mathbf{Y}\cdot(\mathbf{W},\mathbf{U},\mathbf{V})$$

$$\mathbf{V}(\mathbf{W},\mathbf{Y},\mathbf{U}) - \mathbf{U}(\mathbf{W},\mathbf{Y},\mathbf{V}) = \mathbf{W}(\mathbf{U},\mathbf{V},\mathbf{Y}) - \mathbf{Y}(\mathbf{U},\mathbf{V},\mathbf{W})$$

$$[\mathbf{U}\times\mathbf{V}]\cdot[\mathbf{W}\times\mathbf{Y}] = (\mathbf{U}\cdot\mathbf{W})(\mathbf{V}\cdot\mathbf{Y}) - (\mathbf{U}\cdot\mathbf{Y})(\mathbf{V}\cdot\mathbf{W})$$

$$[\mathbf{A},[\mathbf{B},\mathbf{C}]] = \mathbf{B}(\mathbf{A},\mathbf{C}) - \mathbf{C}(\mathbf{A},\mathbf{B})$$

$$[\mathbf{U},[\mathbf{V},\mathbf{W}]] + [\mathbf{W},[\mathbf{U},\mathbf{V}]] + [\mathbf{V},[\mathbf{W},\mathbf{U}]] = 0$$

Now, let's construct a tensor version of the object E_{abc}. To do so, one has to use the **transformation rule for the covariant tensor of the third rank,starting from the special Cartesian coordinates** X^a:

$$\varepsilon_{ijk} = \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial X^c}{\partial x^k}E_{abc} \tag{4.11.15}$$

The component ε_{123} is a square root of the metric determinant,

$$\varepsilon_{123} = g^{1/2}, \qquad \text{where } g = \det \|g_{ij}\| \tag{4.11.16}$$

or

$$\varepsilon_{123} = \det \left\| \frac{\partial X^a}{\partial x^i} \right\| \tag{4.11.17}$$

while we can use (4.11.17) without much thought, we need to be careful using (4.11.16). The reason is that the coordinate transformations that break parity can change the sign in the r.h.s. in Eq.(4.11.16)

Despite E_{ijk} is not being a tensor, we can easily control its transformation from one coordinate system to another:

$$E_{ijk} = \frac{1}{\sqrt{g}}\varepsilon_{ijk}$$

E_{ijk} is a particular example of objects which are called **tensor densities**.

Δ Definition

The quantity $A_{i_1\ldots i_n}{}^{j_1\ldots j_n}$ is a tensor density of the (m,n) -type with the weight r, if the quantity

$$g^{-r/2} \cdot A_{i_1\ldots i_n}{}^{j_1\ldots j_m}$$

is a tensor.

In general,

$$\varepsilon_{i_1 i_2\cdots i_D} = \frac{\partial x^{a_1}}{\partial x^{i_1}}\frac{\partial x^{a_2}}{\partial x^{i_2}}\cdots\frac{\partial x^{a_D}}{\partial x^{i_D}}E_{a_1 a_2\cdots a_D} \tag{4.11.18}$$

Since

$$g_{ij} = \frac{\partial x^a}{\partial x^i}\frac{\partial x^b}{\partial x^j}\delta^a_b \tag{4.11.19}$$

assuming that the orientation of axes is such that $\det\left(\frac{\partial x^a}{\partial x^i}\right) > 0$, we get

$$\det\left\|\frac{\partial x^a}{\partial x^i}\right\| = \sqrt{g} \tag{4.11.20}$$

Then

$$\varepsilon_{123\ldots D} = \frac{\partial x^{a_1}}{\partial x^1}\frac{\partial x^{a_2}}{\partial x^2}\cdots\frac{\partial x^{a_D}}{\partial x^D}E_{a_1a_2\cdots a_D} = \det\left(\frac{\partial x^a}{\partial x^i}\right) = \sqrt{g} \tag{4.11.21}$$

Similarly, $\varepsilon^{12\cdots D} = \frac{1}{\sqrt{g}}$.

4.12 Curvilinear Coordinates, Local Coordinate Transformations

So far, we have learned the following transformation rule for scalar field, vector, and tensor:

$$\varphi'(x') = \varphi(x)$$

$$a'^i(x') = \frac{\partial x^i}{\partial x^j}a^j(x)$$

$$b'_l(x') = \frac{\partial x^k}{\partial x'^l}b_k(x)$$

$$A^{i'}_{j'}(x') = \frac{\partial x^{i'}}{\partial x^k}\frac{\partial x^l}{\partial x'^j}A^k_l(x)$$

and for metric tensor:

$$g_{ij}(x) = \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}g_{ab} \tag{4.12.1}$$

$$g^{ij} = \frac{\partial x^i}{\partial X^a}\frac{\partial x^j}{\partial X^b}\delta^{ab} \tag{4.12.2}$$

where $g_{ab} = \delta_{ab}$ is a metric in Cartesian coordinates.

It is so easy to generalize the notion of tensor and algebraic operations over tensors, **because all these operations are defined in the same point of the space.** Thus, the main difference between general coordinate transformation $x'^\alpha = x'^\alpha(x)$ and the special

one $x'^{\alpha} = \wedge^{\alpha'}_{\beta} x^{\beta} + B^{\alpha'}$ with $\wedge^{\alpha'}_{\beta} = \text{const}$ and $B^{\alpha'} = \text{const}$ is that **in the general case, the transition coefficients $\partial x^i / \partial x'^j$ are not necessary constants.**

One of important consequences is that the metric tensor g_{ij} also depends on the point. Additionally, the antisymmetric tensor ε^{ijk} also depends on the coordinates.

4.12.1 Polar coordinates on the plane

Our purpose in this section is to learn how to transform an arbitrary tensor to polar coordinates. From (x, y) to (r, φ), we have

$$\mathbf{e}_r = \frac{\partial x}{\partial r}\hat{\mathbf{n}}_x + \frac{\partial y}{\partial r}\hat{\mathbf{n}}_y = \hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi$$
$$\mathbf{e}_\varphi = \frac{\partial x}{\partial \varphi}\hat{\mathbf{n}}_x + \frac{\partial y}{\partial \varphi}\hat{\mathbf{n}}_y = r(-\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi)$$

Because these basis vectors are orthogonal, the metric in polar coordinates is diagonal

$$g_{ij} = \begin{pmatrix} g_{rr} & g_{r\varphi} \\ g_{\varphi r} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_r \cdot \mathbf{e}_r & \mathbf{e}_r \cdot \mathbf{e}_\varphi \\ \mathbf{e}_\varphi \cdot \mathbf{e}_r & \mathbf{e}_\varphi \cdot \mathbf{e}_\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \tag{4.12.3}$$

The first observation is that the basic vectors $\mathbf{e}_r$ and $\mathbf{e}_\varphi$ are orthogonal $\mathbf{e}_r \cdot \mathbf{e}_\varphi = 0$. As a result, the metric in polar coordinates is diagonal. Again, this metric tensor is w.r.t. the coordinate basis, not w.r.t. orthonormalized vector basis. To get metrix tensor w.r.t. orthonormalized vector, we can simply use Eq.(4.12.1).

In what follows we shall mark the components of the vector in this basis by tilde. The simpler notations without tilde are always reserved for the components of the vector in the normalized basis. Generally, using the normalized basis means that all components of the vector have the same dimension.

· **Example**

For the polar coordinates on the 2D plane, find the metric by performing transformation of the metric in Cartesian coordinates

Solution:

$$g_{ab} = \delta_{ab}, \quad \text{that is} \quad g_{xx} = 1 = g_{yy}, \quad g_{xy} = g_{yx} = 0$$

$$g_{\varphi\varphi} = \frac{\partial x}{\partial \varphi}\frac{\partial x}{\partial \varphi}g_{xx} + \frac{\partial x}{\partial \varphi}\frac{\partial y}{\partial \varphi}g_{xy} + \frac{\partial y}{\partial \varphi}\frac{\partial x}{\partial \varphi}g_{yx} + \frac{\partial y}{\partial \varphi}\frac{\partial y}{\partial \varphi}g_{yy} =$$
$$= \left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2 = r^2\sin^2\varphi + r^2\cos^2\varphi = r^2$$

and

$$g_{rr} = \frac{\partial x}{\partial r}\frac{\partial x}{\partial r}g_{xx} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial r}g_{yy} = \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 = \cos^2\varphi + \sin^2\varphi = 1$$

Thus, the metric is

$$\begin{pmatrix} g_{rr}g_{r\varphi} \\ g_{\varphi r}g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

This derivation of the metric has a simple geometric interpretation. Consider two points that have infinitesimally close φ and r. The distance between a these two points ds is defined by the relation

$$ds^2 = dx^2 + dy^2 = g_{ab}dX^a dX^b = g_{ij}dx^i dx^j = dr^2 + r^2 d\varphi^2$$
$$= g_{rr}drdr + g_{\varphi\varphi}d\varphi d\varphi + 2g_{\varphi r}d\varphi dr$$

The tensor form of the transformation of the metric corresponds to the coordinate-independent distance between two infinitesimally close points

4.12.2 Cylindrical and Spherical Coordinates

Cylindrical coordinates in 3D are defined by the relations:

$$x = r\cos\varphi, y = r\sin\varphi, \quad z = z$$
$$\text{where} \quad 0 \le r < \infty, 0 \le \varphi < 2\pi \quad \text{and} \quad -\infty < z < \infty$$

The basic vectors are

$$\mathbf{e}_r = \hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi, \quad \mathbf{e}_\varphi = -\hat{\mathbf{i}}r\sin\varphi + \hat{\mathbf{j}}r\cos\varphi, \quad \mathbf{e}_z = \hat{\mathbf{k}}$$

and the metric tensor is

$$g_{ij} = \begin{pmatrix} g_{rr}g_{r\varphi}g_{rz} \\ g_{\varphi r}g_{\varphi\varphi}g_{\varphi z} \\ g_{zr}g_{z\varphi}g_{zz} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.12.4}$$

For spherical coordinates, we have the following relation with Cartesian coordiantes

$$x = r\cos\varphi\sin\chi, y = r\sin\varphi\sin\chi, \quad z = r\cos\chi$$

$$\text{where} \quad 0 \le r < \infty, 0 \le \varphi < 2\pi \quad \text{and} \quad 0 \le \chi \le \pi$$

The basic vector is then

$$\mathbf{e}_r = \hat{\mathbf{i}}\cos\varphi\sin\chi + \hat{\mathbf{j}}\sin\varphi\sin\chi + \hat{\mathbf{k}}\cos\chi$$
$$\mathbf{e}_x = \hat{\mathbf{i}}r\cos\varphi\cos\chi + \hat{\mathbf{j}}r\sin\varphi\cos\chi - \hat{\mathbf{k}}r\sin\chi$$
$$\mathbf{e}_\varphi = -\hat{\mathbf{i}}r\sin\varphi\sin\chi + \hat{\mathbf{j}}r\cos\varphi\sin\chi$$

with the metric tensor as

$$g_{ij} = \begin{pmatrix} g_{rr} & g_{rx} & g_{r\varphi} \\ g_{\chi r} & g_{\chi x} & g_{x\varphi} \\ g_{\varphi r} & g_{\varphi x} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2 x \end{pmatrix}$$

And the normalized basis is

$$\begin{aligned} \hat{\mathbf{n}}_r &= \hat{\mathbf{i}}\cos\varphi\sin\chi + \hat{\mathbf{j}}\sin\varphi\sin\chi + \hat{\mathbf{k}}\cos\chi \\ \hat{\mathbf{n}}_\varphi &= -\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi \\ \hat{\mathbf{n}}_\chi &= \hat{\mathbf{i}}\cos\varphi\cos\chi + \hat{\mathbf{j}}\sin\varphi\cos\chi - \hat{\mathbf{k}}\sin\chi \end{aligned} \tag{4.12.5}$$

Now let us derive the expression for the velocity and acceleration in 3D for the case of the spherical coordinates. The starting point is Eq.(4.12.5). A simple calculus gives the following result for the first derivatives:

$$\begin{aligned} \dot{\hat{\mathbf{n}}}_r &= \dot{\varphi}\sin\chi(-\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi) \\ &\quad + \dot{\chi}(\hat{\mathbf{i}}\cos\varphi\cos\chi + \hat{\mathbf{j}}\sin\varphi\cos\chi - \hat{\mathbf{k}}\sin\chi) \\ \dot{\hat{\mathbf{n}}}_\varphi &= -\dot{\varphi}(\hat{\mathbf{i}}\cos\varphi + \hat{\mathbf{j}}\sin\varphi) \\ \dot{\hat{\mathbf{n}}}_\chi &= \dot{\varphi}\cos\chi(-\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi) \\ &\quad - \dot{\chi}(\hat{\mathbf{i}}\sin\chi\cos\varphi + \hat{\mathbf{j}}\sin\chi\sin\varphi + \hat{\mathbf{k}}\cos\chi) \end{aligned}$$

Next, we need to perform an inverse transformation from the Cartesian basis to the new basis:

$$\begin{aligned} \hat{\mathbf{i}} &= \frac{\partial r}{\partial x}\mathbf{e}_r + \frac{\partial\varphi}{\partial x}\mathbf{e}_\varphi + \frac{\partial\chi}{\partial x}\mathbf{e}_x \\ \hat{\mathbf{j}} &= \frac{\partial r}{\partial y}\mathbf{e}_r + \frac{\partial\varphi}{\partial y}\mathbf{e}_\varphi + \frac{\partial\chi}{\partial y}\mathbf{e}_\chi \\ \mathbf{k} &= \frac{\partial r}{\partial z}\mathbf{e}_r + \frac{\partial\varphi}{\partial z}\mathbf{e}_\varphi + \frac{\partial\chi}{\partial z}\mathbf{e}_\chi \end{aligned}$$

Using these relations, one can derive the first derivatives of the vector $\hat{\mathbf{n}}_r, \hat{\mathbf{n}}_\varphi, \hat{\mathbf{n}}_\chi$ in the final form:

$$\dot{\mathbf{n}}_r = \dot{\varphi}\sin\chi\hat{\mathbf{n}}_\varphi + \dot{\chi}\hat{\mathbf{n}}_\chi$$
$$\dot{\mathbf{n}}_\varphi = -\dot{\varphi}\left(\sin\chi\hat{\mathbf{n}}_r + \cos\chi\hat{\mathbf{n}}_\chi\right)$$
$$\dot{\mathbf{n}}_\chi = \dot{\varphi}\cos\chi\hat{\mathbf{n}}_\varphi - \dot{\chi}\hat{\mathbf{n}}_r$$

Now the particle's velocity and acceleration are

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{d}{dt}\left(r\hat{\mathbf{n}}_r\right) = \dot{r}\hat{\mathbf{n}}_r + r\dot{\chi}\hat{\mathbf{n}}_\chi + r\dot{\varphi}\sin\chi\hat{\mathbf{n}}_\varphi \tag{4.12.6}$$

$$\begin{aligned}\mathbf{a} = \ddot{\mathbf{r}} = &\left(\ddot{r} - r\dot{\chi}^2 - r\dot{\varphi}^2\sin^2\chi\right)\hat{\mathbf{n}}_r + \left(2r\dot{\varphi}\dot{\chi}\cos\chi + 2\dot{\varphi}\dot{r}\sin\chi + r\ddot{\varphi}\sin\chi\right)\hat{\mathbf{n}}_\varphi \\ &+ \left(2\dot{r}\dot{\chi} + r\ddot{\chi} - r\dot{\varphi}^2\sin\chi\cos\chi\right)\hat{\mathbf{n}}_\chi\end{aligned} \tag{4.12.7}$$

One can also find the components of Levi-Civita tensor ϵ_{123} and ϵ^{123} for cylindric and spherical coordinates using the following relations:

$$g = \det\|g_{ij}\| \quad \text{and} \quad \det\left\|g^{ij}\right\| = \frac{1}{g} \tag{4.12.8}$$

$$\varepsilon^{123} = \frac{1}{\sqrt{g}}$$

$$\varepsilon_{123} = g^{1/2}$$

For the hyperbolic coordinates, we have

$$x = r\cdot\cosh\varphi, \quad y = r\cdot\sinh\varphi$$

Since

$$\cosh^2\varphi - \sinh^2\varphi = 1; \quad \cosh'\varphi = \sinh\varphi; \quad \sinh'\varphi = \cosh\varphi$$

we find

$$g_{rr} = \cosh 2\varphi, \quad g_{\varphi\varphi} = r^2\cosh 2\varphi, \quad g_{r\varphi} = r\sinh 2\varphi$$

in this case the basis is not orthogonal.

4.13 Deivatives of Tensors,Covariant Derivatives

Let us start from a scalar field φ. Consider its partial derivative

$$\partial_i \varphi = \varphi_{,i} = \frac{\partial \varphi}{\partial x^i} \tag{4.13.1}$$

In the transformed coordinates $x^i = x^{ii}\left(x^j\right)$ we obtain, using the chain rule and $\varphi'\left(x'\right) = \varphi(x)$

$$\partial_{i'} \varphi' = \frac{\partial \varphi'}{\partial x^{i'}} = \frac{\partial x^j}{\partial x'^i} \frac{\partial \varphi}{\partial x^j} = \frac{\partial x^j}{\partial x'^i} \partial_j \varphi$$

The last formula shows that the partial derivative of a scalar field $\varphi_{,i} = \partial_i \varphi$ transforms as a covariant vector field. Certainly, there is no need to modify a partial derivative in this case.

Now we introduce vector differential operator ∇.

When acting on a scalar, ∇ produces a covariant vector which is called gradient.

The next step is to consider a partial derivative of a covariant vector $b_i(x)$.Let us make a corresponding transformation

$$\partial_{j'} b_{i'} = \frac{\partial}{\partial x'^j} b_{i'} = \frac{\partial}{\partial x'^j}\left(\frac{\partial x^k}{\partial x'^i} b_k\right) = \frac{\partial x^k}{\partial x^{i'}} \frac{\partial x^l}{\partial x'^j} \partial_l b_k + \frac{\partial^2 x^k}{\partial x'^j \partial x'^i} b_k \tag{4.13.2}$$

The parenthesis contains a product of the two expressions that are functions of different coordinates x'^j and x^j. When the partial derivative $\partial/\partial x'^j$ acts on the function of the coordinates x^i, it must be applied following the chain rule

$$\frac{\partial}{\partial x'^j} = \frac{\partial x^l}{\partial x'^j} \frac{\partial}{\partial x^l} \tag{4.13.3}$$

The last term in Eq.(4.13.2) can be zero when $\partial x^k / \partial x'^i$ is constant. Then its derivatives are zeros and the gradient of a covariant vector is a tensor. But, **for a general case of curvilinear coordinates (e.g., polar in 2D or spherical in 3D), the formula (4.13.2) shows the non-tensor nature of the transformation.**

For contravariant vector $a^i(x)$ we have similar relations to (4.13.2):

$$\partial_{k'} a^{i'} = \frac{\partial}{\partial x^{k'}} \left(\frac{\partial x^{i'}}{\partial x^j} a^j \right) = \frac{\partial x^{i'}}{\partial x^j} \frac{\partial x^l}{\partial x^{k'}} \partial_l a^j + \frac{\partial^2 x^{i'}}{\partial x^j \partial x^l} \frac{\partial x^l}{\partial x^{k'}} a^j \tag{4.13.4}$$

Remember, $\frac{\partial x^{i'}}{\partial x^j}$ is a function of x^j and chain rule (4.13.3) must apply here.

Similarly, for mixed tensor $T_i^j(x)$

$$\begin{aligned} \partial_{k'} T_{j'}^{i'} &= \frac{\partial}{\partial x^{k'}} \left(\frac{\partial x^{i'}}{\partial x^l} \frac{\partial x^m}{\partial x^{j'}} T_m^l \right) \\ &= \frac{\partial x^n}{\partial x^{k'}} \frac{\partial x^{i'}}{\partial x^l} \frac{\partial x^m}{\partial x^{j'}} \partial_n T_m^l + \frac{\partial^2 x^{i'}}{\partial x^n \partial x^l} \frac{\partial x^n}{\partial x^{k'}} \frac{\partial x^m}{\partial x^{j'}} T_m^l + \frac{\partial x^{i'}}{\partial x^l} \frac{\partial^2 x^m}{\partial x^{k'} \partial x^{j'}} T_m^l \end{aligned} \tag{4.13.5}$$

We now construct a covariant derivative of a tensor.

Δ Definition

The covariant derivative ∇_i satisfies the following two conditions:

- Tensor transformation rule when the new derivative is applied to a tensor
- In the cartesian coordiantes, $\{X^a\}$ the covariant derivative coincides with the usual partial derivative

$$\nabla_a = \partial_a = \frac{\partial}{\partial X^a}$$

Let's now consider an arbitrary tensor, say mixed (1,1)-type one W_i^j. In order to define its covariant derivative, we perform the following steps: 1. Transform it to Cartesian coordinates

$$W_b^a = \frac{\partial X^a}{\partial x^i} \frac{\partial x^j}{\partial X^b} W_j^i$$

2. Take the partial derivative w.r.t. the Cartesian coordinates X^c

$$\partial_c W_b^a = \frac{\partial}{\partial X^c} \left(\frac{\partial X^a}{\partial x^i} \frac{\partial x^j}{\partial X^b} W_j^i \right)$$

3. Transform this derivative back to the original coordinates x^i

$$\nabla_k W_m^l = \frac{\partial X^c}{\partial x^k} \frac{\partial x^l}{\partial X^a} \frac{\partial X^b}{\partial x^m} \partial_c W_b^a$$

By construction, **the covariant derivative follows the Leibnitz rule for the product of two tensor A and B**:

$$\nabla_i(A \cdot B) = \nabla_i A \cdot B + A \cdot \nabla_i B \tag{4.13.6}$$

· **Example**

Let us make an explicit calculation for the covariant vector T_i:

$$\begin{aligned}\nabla_i T_j &= \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}(\partial_a T_b) = \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\cdot\frac{\partial}{\partial X^a}\left(\frac{\partial x^k}{\partial X^b}T_k\right)\\ &= \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial x^k}{\partial X^b}\frac{\partial x^l}{\partial X^a}\frac{\partial T_k}{\partial x^l} + \frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}T_k\frac{\partial^2 x^k}{\partial X^a \partial X^b}\\ &= \partial_i T_j - \Gamma^k_{ji}T_k\end{aligned}$$

The last equivalence holds because we have the following relation for Kronecker symbol

$$\frac{\partial X^a}{\partial x^i}\frac{\partial x^l}{\partial X^a} = \delta^l_i$$

$$\Gamma^k_{ji} = -\frac{\partial X^a}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial^2 x^k}{\partial X^a \partial X^b} \tag{4.13.7}$$

Δ **Theorem**

Suppose the elements of the matrix $\wedge$ depend on the parameter κ and $\wedge(\kappa)$ is a differentiable and invertible matrix for $\kappa \in (a, b)$. Within the region (a, b) the inverse matrix $\wedge^{-1}(\kappa)$ is also differentiable and its derivative is

$$\frac{d\wedge^{-1}}{d\kappa} = -\wedge^{-1}\frac{\partial\wedge}{\partial\kappa}\wedge^{-1} \tag{4.13.8}$$

Proof: from $\wedge \cdot \wedge^{-1} = I$ we obtain

$$\frac{d\wedge}{d\kappa}\wedge^{-1} + \wedge\frac{\partial\wedge^{-1}}{\partial\kappa} = 0$$

After multiplying this equation by $\wedge^{-1}$ from the left, we arrive at the theorem above.

Consider now that

$$\wedge^b_k = \frac{\partial X^b}{\partial x^k}$$

and its inverse matrix is

$$\left(\wedge^{-1}\right)^k_a = \frac{\partial x^k}{\partial X^a}$$

Using (4.13.8) with x^i playing the role of parameter κ, we arrive at

$$\begin{aligned}\frac{\partial^2 X^b}{\partial x^i \partial x^k} &= \frac{\partial}{\partial x^i}\wedge^b_k = -\wedge^b_l \frac{\partial \left(\wedge^{-1}\right)^l_a}{\partial x^i}\wedge^a_k \\ &= -\frac{\partial X^b}{\partial x^l}\left(\frac{\partial X^c}{\partial x^i}\frac{\partial}{\partial X^c}\frac{\partial x^l}{\partial X^a}\right)\frac{\partial X^a}{\partial x^k} = -\frac{\partial X^b}{\partial x^l}\frac{\partial X^a}{\partial x^k}\frac{\partial X^c}{\partial x^i}\frac{\partial^2 x^l}{\partial X^c \partial X^a}\end{aligned} \tag{4.13.9}$$

Use the equation above, we find another expression for Γ^j_{ki}:

$$\Gamma^j_{ki} = \frac{\partial x^j}{\partial X^b}\frac{\partial^2 X^b}{\partial x^i \partial x^k} = -\frac{\partial^2 x^j}{\partial X^b \partial X^a}\frac{\partial X^a}{\partial x^k}\frac{\partial X^b}{\partial x^i} \tag{4.13.10}$$

By making direct calculation and using the equation above, we have:

$$\nabla_i S^j = \partial_i S^j + \Gamma^j_{ki} S^k \tag{4.13.11}$$

$$\nabla_i W^j_k = \partial_i W^j_k + \Gamma^j_{li} W^l_k - \Gamma^l_{ki} W^j_l \tag{4.13.12}$$

Using the definition of Γ^k_{ij}, we can **formulate the general rule for constructing a covariant derivative of an arbitrary tensor,**

$$\begin{aligned}\nabla_i T^{j_1 j_2 \cdots}{}_{k_1 k_2 \ldots} = \partial_i T^{j_1 j_2 \cdots}{}_{k_1 k_2 \ldots} + \Gamma^{j_1}_{li} T^{l j_2 \cdots}{}_{k_1 k_2 \ldots} + \Gamma^{j_2}_{li} T^{j_1 l \ldots}{}_{k_1 k_2 \ldots} \\ - \Gamma^l_{k_1 i} T^{j_1 j_2 \cdots}{}_{l k_1 \ldots} - \Gamma^l_{k_2 i} T^{j_1 j_2 \cdots}{}_{l k_2 \ldots} - \ldots\end{aligned} \tag{4.13.13}$$

According to the definition of affine connection (4.13.10), we can derive the transformation rule for Γ^i_{kj} as:

$$\begin{aligned}\Gamma^{l'}_{m'n'} &= \frac{\partial x^{l'}}{\partial X^a}\frac{\partial^2 X^a}{\partial x^{m'}\partial x^{n'}} = \frac{\partial x^{l'}}{\partial x^i}\frac{\partial x^i}{\partial X^a}\frac{\partial}{\partial x^{m'}}\left[\frac{\partial X^a}{\partial x^j}\frac{\partial x^j}{\partial x^{n'}}\right] \\ &= \Gamma^i_{kj}\frac{\partial x^{l'}}{\partial x^i}\frac{\partial x^k}{\partial x^{m'}}\frac{\partial x^j}{\partial x^{n'}} + \frac{\partial x^{l'}}{\partial x^i}\frac{\partial^2 x^i}{\partial x^{m'}\partial x^{n'}}\end{aligned}$$

Thus,

$$\Gamma^{i'}_{j'k'} = \frac{\partial x^{i'}}{\partial x^l}\frac{\partial x^m}{\partial x^{j'}}\frac{\partial x^n}{\partial x^{k'}}\Gamma^l_{mn} + \frac{\partial x^{i'}}{\partial x^r}\frac{\partial^2 x^r}{\partial x^{j'}\partial x^{k'}} \tag{4.13.14}$$

The most useful form of the affine connection is expressed via the metric tensor. In any coordinate system

$$\nabla_i g_{jk} = \frac{\partial X^c}{\partial x^i}\frac{\partial X^b}{\partial x^j}\frac{\partial X^a}{\partial x^k}\partial_c g_{ab} = 0 \tag{4.13.15}$$

where $g_{ab} \equiv \delta_{ab}$, and hence $\partial_c g_{ab} \equiv \nabla_c g_{ab} \equiv 0$. If we apply to the above equation, the explicit form of the covariant derivative, we arrive at the equation:

$$\nabla_i g_{jk} = \partial_i g_{jk} - \Gamma^1_{ji} g_{lk} - \Gamma^l_{ki} g_{lj} = 0$$

Making permutations of indices, we get:

$$\begin{aligned}
\partial_i g_{jk} &= \Gamma^l_{ji} g_{lk} + \Gamma^l_{ki} g_{lj} (i) \\
\partial_j g_{ik} &= \Gamma^l_{ij} g_{lk} + \Gamma^l_{kj} g_{il} (ii) \\
\partial_k g_{ij} &= \Gamma^l_{ik} g_{lj} + \Gamma^l_{jk} g_{li} (iii)
\end{aligned}$$

Taking $(i) + (ii) - (iii)$, we arrive at the relationship:

$$2\Gamma^l_{ij} g_{lk} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}$$

Contracting both parts with g^{km} (remember that $g^{km} g_{ml} = \delta^k_l$), we arrive at

$$\Gamma^m_{ij} = \frac{1}{2} g^{km} \left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right) \tag{4.13.16}$$

As an application of these formulas, we can consider the derivation of the Laplace operator acting on scalar and vector fields. For any kind of field, the Laplace operator can be defined as

$$\Delta = g^{ij} \nabla_i \nabla_j \tag{4.13.17}$$

In the case of a scalar field Ψ, we have

$$\Delta \Psi = g^{ij} \nabla_i \nabla_j \Psi = g^{ij} \nabla_i \partial_j \Psi \tag{4.13.18}$$

The second covariant derivative acts on the vector $\partial_i \Psi$, hence

$$\Delta\Psi = g^{ij}\left(\partial_i\partial_j - \Gamma^k_{ij}\partial_k\right)\Psi = \left(g^{ij}\partial_i\partial_j - g^{ij}\Gamma^k_{ij}\partial_k\right)\Psi \tag{4.13.19}$$

Similarly, for the vector field $A^i(x)$, we obtain

$$\begin{aligned}
\Delta A^i &= g^{jk}\nabla_j\nabla_k A^i = g^{jk}\left[\partial_j\left(\nabla_k A^i\right) - \Gamma^l_{kj}\nabla_l A^i + \Gamma^i_{lj}\nabla_k A^l\right] \\
&= g^{jk}[\partial_j\left(\partial_k A^i + \Gamma^i_{ik}A^l\right) - \Gamma^l_{kj}\left(\partial_l A^i + \Gamma^i_{ml}A^m\right) + \Gamma^i_{lj}\left(\partial_k A^l + \Gamma^l_{mk}A^m\right)] \\
&= g^{jk}(\partial_j\partial_k A^i + \Gamma^i_{ik}\partial_j A^l + A^l\partial_j\Gamma^i_{lk} - \Gamma^l_{kj}\partial_l A^i + \Gamma^i_{lj}\partial_k A^l \\
&- \Gamma^l_{kj}\Gamma^i_{ml}A^m + \Gamma^i_{lj}\Gamma^l_{mk}A^m)
\end{aligned} \tag{4.13.20}$$

The formulas in the equations above for scalar and vector fields hold for any dimension D and for an arbitrary choice of coordinates.

· **Example**

Verify that for the scalar field Ψ, the following relation takes place:

$$\Delta\Psi = div(grad\Psi)$$

Solution:

$$\operatorname{div}\operatorname{grad}\Psi = \nabla_i(\operatorname{grad}\Psi)^i = \nabla_i\left(g^{ij}\nabla_j\Psi\right) = \Delta\Psi$$

The second equivalence is the result of raising subscript. And we also have $\nabla_i g^{ij}$.

· **Example**

Prove the following relation between the contraction of the affine connection and the derivative of the metric determinant $g = det||g_{\mu\nu}||$:

$$\Gamma^j_{ij} = \frac{1}{2g}\frac{\partial g}{\partial x^i} = \partial_i \ln\sqrt{g} \tag{4.13.21}$$

Solution:

$$\Gamma^j_{ij} = \frac{1}{2}\delta^j_k\Gamma^k_{ij} = \frac{1}{2}\delta^j_k g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right) = \frac{1}{2}g^{jl}\partial_i g_{jl} = \frac{1}{2g}\partial_i g$$

The last equivalence holds because the derivative of the determinant g equals:

$$\dot{g} = \partial_i g = g g^{lj}\partial_i g_{jl} = g g^{jl}\partial_i g_{jl}$$

4.14 Grad,div and rot

Here, we consider some important operations over vector and scalar fields. All the consideration here will be restricted to the Cartesian coordinates, but the results can be easily generalized to any other coordinates by using the transformation rule. When transforming the relations into curvilinear coordinates, one has to take care to use only the Levi-Civita tensor ϵ_{ijk} and not the maximal antisymmetric symbol E_{ijk}. The two objects are related as $\epsilon_{ijk} = \sqrt{g} Eijk$ and $E_{123} = 1$.

We already know

$$\begin{aligned} \operatorname{div} \mathbf{V} = \partial_{\mathrm{a}} \quad & \mathrm{V}^{\mathrm{a}} = \nabla \mathbf{V} \\ \operatorname{grad} \Psi = \hat{\mathbf{n}}^a \partial_a \Psi \quad & = \nabla \Psi \\ \Delta \quad & = g^{ab} \partial_a \partial_b \end{aligned}$$

$$\nabla = \mathbf{e}^i \nabla_i = g^{ij} \mathbf{e}_j \nabla_i$$

$$\operatorname{rot} \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ V_x & V_y & V_z \end{vmatrix} = \begin{vmatrix} \hat{\mathbf{n}}_1 & \hat{\mathbf{n}}_2 & \hat{\mathbf{n}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ V_1 & V_2 & V_3 \end{vmatrix} = E^{abc} \hat{\mathbf{n}}_a \partial_b V_c$$

One can easily prove the following important relations using antisymmetry of E^{abc}:

$$\operatorname{rot} \operatorname{grad} \Psi \equiv 0 \tag{4.14.1}$$

$$\operatorname{div}(\operatorname{rot} \mathbf{V}) \equiv 0 \tag{4.14.2}$$

4.14.1 Some important definitions

Δ Definition

Consider a differentiable vector $\mathbf{C}(\mathbf{r})$. It is called **potential vector field**, if it can be presented as a gradient of some scalar field $\Psi(\mathbf{r})$ (called potential)

$$\mathbf{C}(\mathbf{r}) = \operatorname{grad} \Psi(\mathbf{r}) \tag{4.14.3}$$

Δ Definition

A differentiable vector field $\mathbf{B}(\mathbf{r})$ is called **solenoidal**, if it can be presented as a rotor of some vector field $\mathbf{A}(\mathbf{r})$

$$\mathbf{B}(\mathbf{r}) = \operatorname{rot} \mathbf{A}(\mathbf{r}) \tag{4.14.4}$$

The most known physical example of the solenoidal vector field is the magnetic field B, which is derived from the vector potential A exactly through the equation above.

Δ Theorem

Suppose $\mathbf{V}(\mathbf{r})$ is a smooth vector field, defined in the whole $3D$ space, which falls sufficiently fast at infinity. Then $\mathbf{V}(\mathbf{r})$ has unique (up to a gauge transformation) representation as a sum

$$\mathbf{V} = \mathbf{C} + \mathbf{B}$$

where C and B are potential and solenoidal fields correspondingly.

4.15 Grad, div, rot in other coordinates

In cylindrical coordinates, we have a scalar field Ψ and vector field $\mathbf{V}$:

$$(\operatorname{grad}\Psi)_r = \frac{\partial\Psi}{\partial r}, \quad (\operatorname{grad}\Psi)_\varphi = \frac{1}{r}\frac{\partial\Psi}{\partial\varphi}, \quad (\operatorname{grad}\Psi)_z = \frac{\partial\Psi}{\partial z}$$

$$\operatorname{div}\mathbf{V} = \frac{1}{r}\frac{\partial}{\partial r}(rV^r) + \frac{1}{r}\frac{\partial V^\varphi}{\partial\varphi} + \frac{\partial V^z}{\partial z}$$

$$(\operatorname{rot}\mathbf{V})^z = \frac{1}{r}\frac{\partial(rV^\varphi)}{\partial r} - \frac{1}{r}\frac{\partial V^r}{\partial_\varphi}, \quad (\operatorname{rot}\mathbf{V})^r = \frac{1}{r}\frac{\partial V^z}{\partial\varphi} - \frac{\partial V^\varphi}{\partial z}$$

$$(\operatorname{rot}\mathbf{V})^\varphi = \frac{\partial V^r}{\partial z} - \frac{\partial V^z}{\partial r}$$

$$\Delta\Psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\varphi^2} + \frac{\partial^2\Psi}{\partial z^2}$$

In spherical coordinates

$$(\operatorname{grad}\Psi)_r = \frac{\partial\Psi}{\partial r}, \quad (\operatorname{grad}\Psi)_\varphi = \frac{1}{r\sin\theta}\frac{\partial\Psi}{\partial\varphi}, \quad (\operatorname{grad}\Psi)_\theta = \frac{1}{r}\frac{\partial\Psi}{\partial\theta}$$

$$\operatorname{div}\mathbf{V}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2V^r\right)+\frac{1}{r\sin\theta}\frac{\partial V^\varphi}{\partial\varphi}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(V^\theta\sin\theta\right)$$

$$\Delta\Psi=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\Psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2}$$

$$(\operatorname{rot}\mathbf{V})^r=\frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}\left(V^\varphi\cdot\sin\theta\right)-\frac{\partial V^\theta}{\partial\varphi}\right]$$

$$(\operatorname{rot}\mathbf{V})^\varphi=\frac{1}{r}\frac{\partial}{\partial r}\left(rV^\theta\right)-\frac{1}{r}\frac{\partial V^r}{\partial\theta},\quad(\operatorname{rot}\mathbf{V})^\theta=\frac{1}{r\sin\theta}\frac{\partial V^r}{\partial\varphi}-\frac{1}{r}\frac{\partial}{\partial r}\left(rV^\varphi\right)$$

4.16 D-dimensional Integrals

The first integral we will formulate is the D -dimensional volume integral in curvilinear coordinates $x^i=x^i\left(X^a\right)$. For such a formulation, we shall use the metric tensor and Levi-Civita tensor in the curvilinear coordinates. The square of the distance between the two points X^a and Y^a in Cartesian and arbitrary global coordinates is given by

$$s_{xy}^2=\sum_{a=1}^{D}\left(X^a-Y^a\right)^2=g_{ij}\left(x^i-y^i\right)\left(x^j-y^j\right)$$

the similar formula holds for the infinitesimal distances

$$dl^2=g_{ij}dx^idx^j \tag{4.16.1}$$

Therefore, if we have a curve in D -dimensional space $x^i=x^i(\tau)$ (here τ is an arbitrary monotonic parameter along the curve), then the length of the curve between the points A with the coordinates $x^i(a)$ and B with the coordinates $x^i(b)$ is

$$l_{AB}=\int_{(AB)}dl=\int_a^b d\tau\sqrt{g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau}} \tag{4.16.2}$$

The direct geometric sense of the metric is that it defines a distance between two infinitesimally close points and also the length of the finite curve in the D-dimensional space, in arbitrary curvilinear coordinates.

For the infinitesimal volume in D-dimensional space, we have

$$dV = \frac{1}{D!}\left|\varepsilon_{i_1 i_2 .. i_p}\right| dx^{i_1} \dots dx^{i_D} = \sqrt{g} \cdot dx^1 dx^2 \dots dx^D \tag{4.16.3}$$

The quantity

$$J = \sqrt{g} = \sqrt{\det \|g_{ij}\|} = \left|\det\left(\frac{\partial X^a}{\partial x^i}\right)\right| = \left|\frac{D\left(X^a\right)}{D\left(x^i\right)}\right| \tag{4.16.4}$$

is nothing but the well-known **Jacobian of the coordinate transformation.**

4.16.1 Curvilinear Integrals

We now start to consider integrals over curves and curved surfaces.

Δ Definition

curvilinear integral of the first type

Consider the continuous function $f(\mathbf{r}) = f\left(x^i\right)$, where x^i are the coordinates in D -dimensional space, $x^i = \left\{x^1, x^2, \dots, x^D\right\}$ defined along the curve $L = \int_{(L)} dl = \int_a^b \sqrt{g_{ij}\dot{x}^i\dot{x}^j}dt$. One can define the curvilinear integral of the first type

$$I_1 = \int_{(L)} f(\mathbf{r})dl \tag{4.16.5}$$

It is easy to see that this sum is also an integral sum for the Riemann integral.

$$\int_a^b \sqrt{g_{ij}\dot{x}^i\dot{x}^j} f\left(x^i(t)\right) dt \tag{4.16.6}$$

Δ Definition

curvilinear integral of the second type

Consider a vector field $\mathbf{A}(\mathbf{r})$ defined along the curve (L). Let us consider the $3D$ case, one can construct the following infinitesimal scalar:

$$\mathbf{A} \cdot d\mathbf{r} = A_x d_x + A_y d_y + A_z d_z = A_i dx^i \tag{4.16.7}$$

This expression can be integrated along the curve (L) to give

$$\int_{(L)} \mathbf{A}\cdot d\mathbf{r} = \int_{(L)} A_x dx + A_y dy + A_z dz = \int_{(L)} A_i dx^i \tag{4.16.8}$$

If the curve is parametrized by the continuous monotonic parameter t, the last expression can be presented as

$$A_i dx^i = \mathbf{A}\cdot d\mathbf{r} = \left(A_x\dot{x} + A_y\dot{y} + A_z\dot{z}\right) dt = A_i\dot{x}^i dt \tag{4.16.9}$$

It is easy to establish the following relation between the curvilinear integral of the second type, the curvilinear integral of the first type, and the Riemann integral:

$$\int_{(L)} A_x dx + A_y dy + A_z dz = \int_{(L)} A_i \frac{dx^i}{dl} dl = \int_a^b A_i \frac{dx^i}{dt} dt \tag{4.16.10}$$

where l is a natural parameter along the curve

$$l = \int_a^t \sqrt{g_{ij}\dot{x}^i\dot{x}^j}\, dt, \quad 0 \le l \le L$$

In Cartesian coordinates, we have

$$\begin{aligned} \int_{(L)} \mathbf{A}\cdot d\mathbf{r} &= \int_0^L \left(A_x\cos\alpha + A_y\cos\beta + A_z\cos\gamma\right) dl \\ &= \int_a^b \left(A_x\cos\alpha + A_y\cos\beta + A_z\cos\gamma\right) \sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}\, dt \end{aligned} \tag{4.16.11}$$

where we used $dl^2 = g_{ab}\dot{X}^a\dot{X}^b dt^2 = \left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) dt^2$. The main properties of the curvilinear integral of the first type are

$$\text{additivity } \int_{AB} + \int_{BC} = \int_{AC} \text{ and symmetry } \int_{AB} = \int_{BA}$$

The main properties of the curvilinear integral of the second type are

$$\text{additivity } \int_{AB} + \int_{BC} = \int_{AC} \text{ and antisymmetry } \int_{AB} = -\int_{BA}$$

4.16.2 2D surfaces Integrals in a 3D Space

Consider the integrals over the $2D$ surface, (S) in the $3D$ space $(S) \subset \mathbb{R}^3$. Suppose the surface is defined by the three smooth functions

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)$$

where u and v are independent variables which are called internal coordinates on the surface. One can regard the relations above as a mapping of a figure (G) on the plane with coordinates u and v into the $3D$ space with the coordinates x, y, z.

Let us start by considering the infinitesimal line element dl linking two points of the surface. We can suppose that these two infinitesimally close points belong to the same smooth curve $u = u(t), v = v(t)$, situated on the surface. Then the line element is

$$dl^2 = dx^2 + dy^2 + dz^2$$

where

$$\begin{aligned} dx &= \left(\frac{\partial x}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial x}{\partial v}\frac{\partial v}{\partial t} \right) dt = x'_u du + x'_v dv \\ dy &= \left(\frac{\partial y}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial t} \right) dt = y'_u du + y'_v dv \\ dz &= \left(\frac{\partial z}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial t} \right) dt = z'_u du + z'_v dv \end{aligned}$$

Thus,

$$dl^2 = g_{uu}du^2 + 2g_{uv}dudv + g_{vd}dv^2 \tag{4.16.12}$$

where

$$\begin{aligned} g_{uu} &= x'^2_u + y'^2_u + z'^2_u \\ g_{uv} &= x'_u x'_v + y'_u y'_v + z'_u z'_v \\ g_{vv} &= x'^2_v + y'^2_v + z'^2_v \end{aligned} \tag{4.16.13}$$

are the elements of the matrix which is called **induced metric on the surface**.

We now calculate the infinitesimal element of **the area of the surface**. Here, we introduce two basis vectors on the surface in such a way that the scalar products of these vectors are equal to the corresponding metric components:

$$\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u} \quad \text{and} \quad \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v} \tag{4.16.14}$$

Below we suppose that at any point of the surface $\mathbf{r}_u \times \mathbf{r}_v \neq 0$. This means that the lines of constant coordinates u and v are not parallel or, equivalently, that the internal coordinates are not degenerate.

$$g_{uu} = \mathbf{r}_u \cdot \mathbf{r}_u, \quad g_{uv} = \mathbf{r}_u \cdot \mathbf{r}_v, \quad g_{vv} = \mathbf{r}_v \cdot \mathbf{r}_v \tag{4.16.15}$$

After we map the infinitesimal surface element $dudv$ on (G) to a curve surface (S), we have a parallelogram spanned on the vectors $\mathbf{r}_u du$ and $\mathbf{r}_v dv$.The area of this parallelogram equals to the absolute value of the vector product of the two vectors,

$$\begin{aligned} dS = (d\mathbf{A})^2 &= |\mathbf{r}_u \times \mathbf{r}_v|^2 du^2 dv^2 = r_u^2 r_v^2 \sin^2\alpha \cdot du^2 dv^2 \\ &= du^2 dv^2 \left\{ r_u^2 r_v^2 \left(1 - \cos^2\alpha\right) \right\} = du^2 dv^2 \left\{ \mathbf{r}^2 \cdot \mathbf{r}^2 - (\mathbf{r}_u \cdot \mathbf{r}_v)^2 \right\} \\ &= \left(g_{uu} \cdot g_w - g_{uv}^2\right) du^2 dv^2 = g \cdot du^2 dv^2 \end{aligned} \tag{4.16.16}$$

where g is the determinant of the induced metric. Finally, the area S of the whole surface (S) may be defined as an integral

$$S = \iint_{(G)} \sqrt{g} du dv \tag{4.16.17}$$

Δ Definition

Suppose we have a surface (S) defined as a mapping of the closed finite figure (G) in a uv-plane, and a function $f(u, v)$ on this surface. The surface integral of the first type is defined as:

$$\mathcal{I}_1 = \iint_{(S)} f(u,v) dS = \iint_{(G)} f(u,v) \sqrt{g} du dv \tag{4.16.18}$$

This surface integral is a scalar, if the function $f(u, v)$ is a scalar. when we change the internal coordinates on the surface to (u', v'), the following aspects must change:

- The form of the area $(G) \to (G')$
- the surface element $\sqrt{g} \to \sqrt{g'}$
- the form of the integrand $f'(u', v') = f\left(u\left(u', v'\right), v\left(u', v'\right)\right)$

Below we define **the surface integral of the second type on smooth two-sided oriented surfaces**, for non-oriented (one-sided) surface, like the Möbius band does not have the property of

$$|\mathbf{r}_u \times \mathbf{r}_v| \neq 0$$

where $\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u}$ and $\mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$. The area of the corresponding parallelogram on (S) is $dS = |\mathbf{r}_u \times \mathbf{r}_v|\, dudv$, and $d\mathbf{S} = \vec{\mathbf{n}} \cdot dS$. For calculational purposes, it is useful to introduce the components of $\hat{\mathbf{n}}$; using the same cosines, we already considered for the curvilinear second-type integral,

$$\cos\alpha = \hat{\mathbf{n}} \cdot \hat{\mathbf{i}}, \quad \cos\beta = \hat{\mathbf{n}} \cdot \hat{\mathbf{j}}, \quad \cos\gamma = \hat{\mathbf{n}} \cdot \hat{\mathbf{k}} \tag{4.16.19}$$

Δ Definition

Consider a continuous vector field $\mathbf{A}(\mathbf{r})$ defined on the surface (S). The surface integral of the second type is

$$\iint_{(S)} \mathbf{A} \cdot d\mathbf{S} = \iint_{S} \mathbf{A} \cdot \mathbf{n} dS \tag{4.16.20}$$

By construction, the surface integral of the second type equals to the following double integral over the figure (G) in the uv -plane:

$$\iint_{(S)} \mathbf{A} \cdot d\mathbf{S} = \iint_{(G)} \left(A_x \cos\alpha + A_y \cos\beta + A_z \cos\gamma\right) \sqrt{g} dudv \tag{4.16.21}$$

In a more general situation, we have an $n-$dimensional surface embedded into $D-$dimensional space R^D, with $D > n$. Introducing internal coordinates u^i on the surface, we obtain its parametric equation of the surface in the form

$$x^\mu = x^\mu\left(u^i\right) \tag{4.16.22}$$

where

$$\begin{cases} i = 1, \ldots, n \\ \mu = 1, \ldots, D \end{cases}$$

The expression for the distance between the two infinitesimally close points is then

$$ds^2 = g_{\mu v} dx^\mu dx^v = \frac{\partial x^\mu}{\partial u^i} \frac{\partial x^v}{\partial u^j} g_{\mu v} du^i du^j \tag{4.16.23}$$

Therefore, in this general case, we meet the same relation as for the usual coordinate transformation

$$g_{ij}(u) = \frac{\partial x^\mu}{\partial u^i}\frac{\partial x^v}{\partial u^j} g_{\mu v}(x) \tag{4.16.24}$$

The main difference between this relation and the usual tensor law of transforming metric is that the matrix $\frac{\partial x^\mu}{\partial u^i}$ has dimension $n \times D$ and hence it cannot be inverted. The volume element of the area (surface element) in the n-dimensional case also looks as

$$dS = J du^1 du^2 \dots du^n, \quad \text{where} \quad J = \sqrt{g} \tag{4.16.25}$$

4.17 Theorems of Green, Stokes, and Gauss

Δ Theorem

For the surface (S) dividable to a finite number of pieces (S_i) of the form:

$$\begin{cases} x = a \\ x = b \end{cases} \quad \text{and} \quad \begin{cases} y = \alpha(x) \\ y = \beta(x) \end{cases}$$

and for the functions $P(x,y)$, $Q(x,y)$ that are continuous on (S) together with their first derivatives P'_x, P'_y, Q'_x, Q'_y, the following relation holds:

$$\iint_{(S)} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \oint_{(\partial S+)} P dx + Q dy \tag{4.17.1}$$

One can use the Green's formula to calculate the area of the surface. To this end, one can take $Q = x, P = 0$ or $Q = 0, P = -y$, or $Q = x/2, P = -y/2$ etc. Then

$$S = \oint_{(\partial S^+)} x dy = -\oint_{(\partial S^+)} y dx = \frac{1}{2}\oint_{(\partial S^+)} x dy - y dx \tag{4.17.2}$$

Δ Theorem

Stoke's Theorem

For any continuous vector function **F** (r) with continuous partial derivatives $\partial_i F_j$,

the following relation holds:

$$\iint_{(S)} \operatorname{rot}\mathbf{F}\cdot d\mathbf{S} = \oint_{(\partial S+)} \mathbf{F}\cdot d\mathbf{r} \tag{4.17.3}$$

The last integral is called circulation of the vector field.

At any point of the space

$$\mathbf{F} = \mathbf{F}(u,v,w) = F^u\hat{\mathbf{n}}_u + F^v\hat{\mathbf{n}}_v + F^w\hat{\mathbf{n}}_w$$

The components of the vectors $\mathbf{F}$, rot $\mathbf{F}, d\mathbf{S}$, and $d\mathbf{r}$ transform in a standard way, such that the products

$$\operatorname{rot}\mathbf{F}\cdot d\mathbf{S} = \varepsilon^{ijk}\,(dS_i)\,\partial_j F_k \quad \text{and} \quad \mathbf{F}\cdot d\mathbf{r} = F_i dx^i$$

$$\oint_{(\partial S+)} \mathbf{F}\cdot d\mathbf{r} = \oint_{(\partial G^+)} \mathbf{F}\cdot d\mathbf{r} = \oint_{(\partial G^+)} F_u du + F_v dv$$

because the contour (∂G^+) lies in the (u,v) plane. Further, since the figure (G) belongs to the (u,v) plane, the oriented area vector $d\mathbf{G}$ is parallel to the $\hat{\mathbf{n}}_w$ axis. Therefore,

$$(d\mathbf{G})_u = (d\mathbf{G})_v = 0$$

Then

$$\begin{aligned}\iint_{(S)} \operatorname{rot}\mathbf{F}\cdot d\mathbf{S} &= \iint_{(G)} \operatorname{rot}\mathbf{F}\cdot d\mathbf{G} = \int_{(G)} (\operatorname{rot}\mathbf{F})_w dG^w \\ &= \iint_{(G)} \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) du dv\end{aligned} \tag{4.17.4}$$

and

$$\oint_{(\partial G^+)} F_u du + F_v dv = \iint_{(G)} \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) du dv \tag{4.17.5}$$

Δ Theorem

For the $\mathbf{F}(\mathbf{r}) = \operatorname{grad} U(\mathbf{r})$, where $U(\mathbf{r})$ is a smooth function of coordinates, the curvilinear integral between two points A and B doesn't depend on the path (AB) and is equal to the difference

$$\int_{(AB)} \mathbf{F}\cdot d\mathbf{r} = U(B) - U(A) \tag{4.17.6}$$

A useful criterion of $\mathbf{F}$ being $\operatorname{grad} U$ is

$$\frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x} \tag{4.17.7}$$

and the same for any couple of partial derivatives.

Δ Theorem

Gauss-Ostrogradsky Theorem

Consider a $3D$ figure $(V) \subset R^3$ and also define (∂V^+) to be the externally oriented boundary of (V). Consider a vector field $\mathbf{E}(\mathbf{r})$ defined on (V) and on its boundary (∂V^+) and suppose that the components of this vector field are continuous, as are their partial derivatives

$$\frac{\partial E^x}{\partial x}, \quad \frac{\partial E^y}{\partial y}, \quad \frac{\partial E^z}{\partial z}$$

Then these components satisfy the following integral relation:

$$\oiint_{(\partial V+)} \mathbf{E} \cdot d\mathbf{S} = \iiint_{(V)} \operatorname{div} \mathbf{E} dV \tag{4.17.8}$$

Chapter 5

Special Relativity

5.1 Maxwell Equations and Lorentz Transformations

Our purpose in this section is to use tensor methods to clarify the relation between electrodynamics and relativity. Let's first write Lorentz law and Maxwell equations in the Gaussian system of units, where electric and magnetic fields have the same dimension. The quantity c is the speed of light:

$$\mathbf{F} = q\mathbf{E} + \frac{q}{c}\mathbf{v} \times \mathbf{H} \tag{5.1.1}$$

$$\begin{aligned} \operatorname{div} \mathbf{E} &= 4\pi\rho, & \operatorname{rot} \mathbf{H} &= \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j} \\ \operatorname{div} \mathbf{H} &= 0, & \operatorname{rot} \mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \end{aligned} \tag{5.1.2}$$

where $\mathbf{j} = \mathbf{v} \cdot \rho(t, \mathbf{r})$ is the current density.

The equations above contradict the laws of Newtonian mechanics. To see this, consider a pair of point-like charges q_1,q_2 that interact by means of Coulomb force. Now if the same two charges in another inertial frame that moves with the constant speed $-\mathbf{v}$ in the direction orthogonal to the line between the two charges.

According to the standard logic of classical mechanics, the electric Coulomb force should not change because of the choice of inertial frame. At the same time, in the new frame, each of the charges will move with a constant speed $+\mathbf{v}$. But now the two charges become currents and hence each of them creates a magnetic field. This magnetic field will act on another charge according to Lorentz's law; therefore, there will be an increase in the force acting on each of the charges. On the other hand, according to the Galilean principle, the force cannot change when we switch from one inertial frame to another such frame—hence this example shows that there is a contradiction. Thus, **this sytem of equations is not invariant under the Galilean transformations between inertial references frames**:

$$\mathbf{r}' = \mathbf{r} + \mathbf{v}t, \quad \mathbf{t}' = \mathbf{t} \tag{5.1.3}$$

The successful solution the conflict between mechanics and electromagnetism lies in the generalization of the transformation(5.1.3). The new transformation should be performed in such a way that the tranformed laws are compatible with the Maxwell equations. **Being compatible here means that changing from one inertial frame to another one does not change the form of these equations, while the quantities such as $\mathbf{H}, \mathbf{E}, \rho, \mathbf{j}$, and $\mathbf{v}$ do transform in a controllable tensorical way.**

Consider Maxwell equations without sources:

$$\begin{aligned} \operatorname{div}\mathbf{E} &= 0, & \operatorname{rot}\mathbf{H} &= \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \\ \operatorname{div}\mathbf{H} &= 0, & \operatorname{rot}\mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \end{aligned} \tag{5.1.4}$$

In order to obtain the wave equation, we remember the formula:

$$\operatorname{rot}(\operatorname{rot}\mathbf{E}) = \operatorname{grad}(\operatorname{div}\mathbf{E}) - \Delta\mathbf{E}$$

Apply the *rot* operator to both side of the second equation in (5.1.4). Using the fact of $div\mathbf{E} = 0$, we arrive at the equation:

$$-\Delta\mathbf{E} = -\frac{1}{c}\operatorname{rot}\frac{\partial \mathbf{H}}{\partial t} = -\frac{1}{c}\frac{\partial \operatorname{rot}\mathbf{H}}{\partial t} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{5.1.5}$$

or

$$\Box\mathbf{E} = 0, \quad \Box = \frac{1}{c}\frac{\partial^2}{\partial t^2} - \Delta \tag{5.1.6}$$

where we introduce the d'Alembert operator $\Box$. Similarly, we have $\Box\mathbf{H} = 0$.

Another important observation is that the lines of the magnetic field $\mathbf{H}$ are always closed and do not have their ends on some sources. This feature can be provided by assuming that there is a **vector potential field A** such that

$$\mathbf{H} = \operatorname{rot} \mathbf{A} \tag{5.1.7}$$

Contrary to this, in the presence of sources the lines of electric field $\mathbf{E}$ always have their ends on the point-like sources or at the infinity. This feature corresponds to the definition

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \varphi \tag{5.1.8}$$

where φ is an additional scalar potential.

5.1.1 Invariant Interval and Minkowski Space

Let us start by defining the Minkowski space, which has the coordinates

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z \tag{5.1.9}$$

The standard abbreviations for Minkowski space are $M_{1,3}$ or M_4. In what follows we shall use Greek indices for the four-dimensional notations in Minkowski space, e.g., $\alpha, \beta, \ldots, \mu$ $0, 1, 2, 3$ and Latin indices for the space variables, e.g., $a, b, \ldots, ij, \cdots = 1, 2, 3$ One can write, for instance, $x^\mu = (x^0, x^i)$ or $x^\mu = (x^0, \mathbf{r})$.

The point in Minkowski space is frequently called the **event**. The event is defined by time and space coordinates. Correspondingly, the interval between two infinitesimally separated events has the form

$$ds^2 = \left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2 = c^2dt^2 - dl^2 \tag{5.1.10}$$

where $dl^2 = dx^2 + dy^2 + dz^2$ is the space interval between the two close events. The last formula can be also written in the form

$$ds^2 = \eta_{\mu v} dx^\mu dx^v, \quad \text{where} \quad \eta_{\mu v} = \operatorname{diag}(1, -1, -1, -1) \tag{5.1.11}$$

is the metric of the *p*seudo-Euclidean Minkowski space, which is also called space- time.

The nonlinear changes of coordinates that involve both space and time variables can be consistently considered in general relativity, but better avoided in special relativity. The reason is that this type of the change in the space–time coordinates means that we go to the non-inertial reference frame. Under the change of inertial frame, we can use the definition of ds^2 above arrive at the **invariance of the interval**:

$$ds'^2 = ds^2 \tag{5.1.12}$$

There is a clear classification of the intervals between the two events into three distinct groups: The events of the first type have a positive square of the interval, $s_{12}^2 > 0$, and are called **time-like**. Themotion from one event to another one in this case can be done with the constant velocity $v < c$. In particular, the two events that occur at the same space point are separated by a time-like interval. The second type of separation is the **light-like interval**, $s_{12}^2 = 0$. The motion with constant velocity between the two intervals can be performed only with the speed of light $v = c$. The third type of separation is by the **space-like interval**, $s_{12}^2 = 0$. This case only happens when $v > c$.

5.1.2 Lorentz Transformations

Consider the two inertial reference frames, the static one K and another K', which are moving with the constant velocity $\mathbf{v}$ with respect to K. Since the space is isotropic, without loss of generality it is possible to assume that the direction of the velocity is along the axis X.

Now we resolve Eq.(5.1.12) in the two-dimensional case:

$$c^2dt'^2 - dx^2 = c^2dt^2 - dx^2$$

The solution can be easily found by analogy with the Euclidean rotations. We know that the continuous transformation that preserves the length. One can make an analytic continuation

of time $ct = ix^4$ and remember that

$$\sin ix = -\frac{\sinh x}{i} = (-1)\cdot\frac{\sinh x}{i} = \left(i^2\right)\cdot\frac{\sinh x}{i} = i\sinh x$$

$$\cos(iz) = \cosh z$$

we have the hyperbolic rotation:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{5.1.13}$$

Since $dx = 0$, we have

$$v = c\beta = \frac{dx'}{dt'} = -c\tanh\psi$$

Thus,

$$\cosh\psi = \frac{1}{\sqrt{1-\beta^2}}, \quad \sinh\psi = -\frac{\beta}{\sqrt{1-\beta^2}}, \quad \text{where} \quad \beta = \frac{v}{c}$$

Now we are in a position to write the Lorentz transformations in terms of v or β,

$$ct' = \frac{ct-\beta x}{\sqrt{1-\beta^2}}, \quad x' = \frac{x-\beta ct}{\sqrt{1-\beta^2}}, \quad y' = y, \quad z' = z \tag{5.1.14}$$

The inverse transformation is:

$$ct = \frac{ct'+\beta x'}{\sqrt{1-\beta^2}}, \quad x = \frac{x'+\beta ct'}{\sqrt{1-\beta^2}}, \quad y = y', \quad z = z' \tag{5.1.15}$$

Using Lorentz transformation, we can also show that the apparent length of the object is contracting in the direction of motion according to the formula:

$$l = l_0\sqrt{1-\beta^2} \tag{5.1.16}$$

Similarly, the space volume contracts too:

$$V = V_0\sqrt{1-\beta^2} \tag{5.1.17}$$

5.2 Laws of Relativistic Mechanics

The Lorentz transformations (also called boosts) and orthogonal space transformations (rotations and inversion of coordinates) **do not change the form of the metric.** An important relation concerning ds for a massive particle that moves with the speed $v = \beta c \leq c$ is as follows:

$$ds = \sqrt{c^2dt^2 - dl^2} = \sqrt{c^2dt^2 - v^2dt^2} = c\sqrt{1-\beta^2}dt \tag{5.2.1}$$

Along with the interval, one can introduce a **proper time** of the particle, which is defined as $d\tau = ds/c = \sqrt{1-\beta^2}dt$.

The evolution of a particle in the Minkowski space M_4 is described by the world line $x^\mu(\lambda)$ where λ is a parameter along the curve. The tangent vector along the curve is defined as $k^\mu = dx^\mu/d\lambda$. One can choose λ in different ways, for example, take $d\lambda = dx^0$. However, in many cases, it is most useful to choose $d\lambda = ds$. Indeed, this choice works well only for the time-like curves, where the tangent vector k^μ to the world line of a particle is time-like, $k^\mu k_\mu \geq 0$. Certainly, **the interval is not the best choice in the case of a light-like tangent vector**, because in this case $ds^2 = 0$. One of the possibilities for this case is to use the time variable, t or x^0.

Consider the world line of a massive particle, which (as we shall see in what follows) has a time-like tangent vector. Taking $d\lambda = ds$, we arrive at the four-velocity of the particle,

$$u^\mu = \frac{dx^\mu}{ds} \tag{5.2.2}$$

and

$$u_\mu u^\mu = \eta_{\mu\nu}u^\mu u^\nu = \frac{ds^2}{ds^2} \equiv 1 \tag{5.2.3}$$

It is interesting to find the relation between four-velocity and usual 3D velocity of the particle, $\mathbf{v} = d\mathbf{r}/dt$:

$$u^i = \frac{dx^i}{ds} = \frac{dx^i}{cd\tau} = \frac{1}{c\sqrt{1-\beta^2}}\frac{dx^i}{dt} = \frac{v^i}{c\sqrt{1-\beta^2}} \tag{5.2.4}$$

and

$$u^0 = \frac{dx^0}{ds} = \frac{1}{\sqrt{1-\beta^2}} \tag{5.2.5}$$

The next relevant quantity is four-acceleration,

$$\omega^\mu = \frac{du^\mu}{ds} = \frac{d^2x^\mu}{ds^2} \tag{5.2.6}$$

Since $u_\mu u^\mu = \eta_{\mu\nu}u^\mu u^\nu = \frac{ds^2}{ds^2} \equiv 1$, we have

$$\frac{d(u_\mu u^\mu)}{ds} = 2\eta_{\mu\nu}u^\mu \frac{du^\nu}{ds} = 0$$

Thus,

$$u_\mu \omega^\mu \equiv 0 \tag{5.2.7}$$

The last thing about kinematics is to obtain the formula for **summing velocities**.Let us suppose that the inertial reference frame K' moves with velocity $\mathbf{V}$ with respect to the static inertial frame K. Consider the particle that moves with the constant velocity $\mathbf{v}'$ with respect to K'. What is its velocity with respect to K? In classical mechanics, we have

$$\mathbf{v} = \mathbf{V} + \mathbf{v}'$$

In the frame K', we have the relations:

$$dx' = v'_x dt', \quad dy' = v'_y dt', \quad dz' = v'_z dt' \tag{5.2.8}$$

Without loss of generality one can assume that the velocity V has only X component, $\mathbf{V} = V\hat{\mathbf{i}}$. Then, according to the differential version of (5.1.15),

$$cdt = \frac{cdt' + \beta dx'}{\sqrt{1-\beta^2}}, \quad dx = \frac{dx' + \beta cdt'}{\sqrt{1-\beta^2}}, \quad dy = dy', \quad dz = dz' \tag{5.2.9}$$

where $\beta = V/c$. Combining (5.2.8) and (5.2.9), we arrive at the result:

$$v_x = \frac{v'_x + V}{1 + Vv'_x/c^2}, \quad v_y = \frac{\sqrt{1-\beta^2}v'_y}{1 + Vv'_x/c^2}, \quad v_z = \frac{\sqrt{1-\beta^2}v'_z}{1 + Vv'_x/c^2} \tag{5.2.10}$$

· **Example**

Consider the pure space transformations (e.g., rotations and parity transformation, that is, the simultaneous inversion of all space coordinates) of $F_{\mu v} = -F_{v\mu}$. Show that with respect to these particular transformations, the components of $F_{\mu\nu}$ transform as the components of usual V_i and axial A_i vectors, such that

$$F_{\mu v} = \begin{pmatrix} 0 & V_1 & V_2 & V_3 \\ -V_1 & 0 & A_3 & -A_2 \\ -V_2 & -A_3 & 0 & A_1 \\ -V_3 & A_2 & -A_1 & 0 \end{pmatrix} \tag{5.2.11}$$

or

$$F_{0i} = V_i \text{ and } F_{ij} = \epsilon_{ijk} A_k \tag{5.2.12}$$

Solution: Consider the transformation of the desired type, with $x^0 = x^0$. Then

$$F'_{0i}(x') = \frac{\partial x^\alpha}{\partial x^0} \frac{\partial x^\beta}{\partial x^{i'}} F_{\alpha\beta}(x)$$

The first factor is nonzero only when $\alpha = 0$, and the second only when $\beta = j$ therefore

$$F'_{0i} = \frac{\partial x^j}{\partial x^{i'}} F_{0j}$$

The last formula means that the components F_{0i} form a vector under continuous transformations, such as rotations. Under the parity transformation, the matrix of derivatives is

$$P = \frac{\partial x^j}{\partial x^{i'}} = \text{diag}(-1, -1, -1)$$

and it is easy to see that F_{0i} behaves like a normal vector $V'_i = -V_i$. The second part of the exercise can be most easily solved by noting that $A_k = \frac{1}{2}\epsilon_{ijk} F^{ij}$. As we know from the first part of the book, in this case, the components of A_k do not change under parity, and hence they form an axial vector. Sometimes the notation $F_{\mu v} = (\mathbf{V}, \mathbf{A})$ is used.

5.2.1 Relativistic dynamics of a free particle

Consider the particle moving between the instants of time t_1 and t_2. In order to have Lorentz-covariant equations of motion, the action of the particle should be a constant scalar with respect to the Lorentz rotations (boosts), and also with respect to orthogonal transformations of space coordinates. The simplest scalar expression that satisfies this condition has the form

$$S = -\alpha \int_{t_1}^{t_2} ds \tag{5.2.13}$$

where α is an unknown constant and one can find it by requiring that the nonrelativistic limit of the expression coincides with the action of a free particle in classical mechanics

$$S_{cl} = \int_{t_1}^{t_2} dt L_{cl}, \quad L_{cl} = T_{cl} = \frac{mv^2}{2} \tag{5.2.14}$$

Let us first assume that $\beta \ll 1$ and make a series expansion:

$$ds = c\sqrt{1-\beta^2}dt = c\left(1 - \frac{\beta^2}{2} - \frac{\beta^4}{8} + \dots\right)dt \tag{5.2.15}$$

Replacing this expression into (5.2.13) and using $\beta = v/c$, we arrive at the nonrelativistic limit,

$$S_{nr} = \int_{t_1}^{t_2} \left(-\alpha c + \frac{\alpha v^2}{2c} + \frac{\alpha v^4}{8c^3} \cdot + \dots\right) dt \tag{5.2.16}$$

The first term in the parentheses is an irrelevant constant. The second term becomes exactly the desired expression (5.2.14) if we assume that

$$\alpha = mc$$

The later terms become small $\mathcal{O}(\beta^4)$ corrections. Since the value of α is now fixed, we can write down the final form of the action (5.2.13):

$$S = -mc \int_{t_1}^{t_2} ds = \int_{t_1}^{t_2} dt L \tag{5.2.17}$$

where

$$L = -mc^2\sqrt{1-\beta^2} \tag{5.2.18}$$

The next steps will be to derive the expressions for canonical momenta and energy, and construct the Lagrange equations. In 4D formalism, remember that $ds = \sqrt{\eta_{\mu\nu}dx^\mu dx^\nu}$,$ds' = \sqrt{\eta_{\mu\nu}d(x^\mu + \delta x^\mu)d(x^\nu + \delta x^\nu)}$ and consider an arbitrary variation of four-coordinate $\delta x^\mu(t)$, such that

$$\delta x^\mu\left(t_1\right) = \delta x^\mu\left(t_2\right) = 0 \tag{5.2.19}$$

Notice that $ds'^2 - ds^2 = \delta s^2$,$\delta s = \frac{d(\delta s^2)}{2ds}$, the variation of the action is

$$\begin{aligned}\delta S = -mc\int_{t1}^{t2}\delta s = -mc\int_{t_1}^{t_2}\frac{1}{2ds}\cdot\left(2\eta_{\mu v}dx^\mu\delta dx^v\right) = -mc\int_{t_1}^{t_2}ds\eta_{\mu v}\frac{dx^\mu}{ds}\frac{\delta dx^v}{ds} \\ = -mc\eta_{\mu v}u^\mu\delta x^v\Big|_{t_1}^{t_2} + mc\int_{t_1}^{t_2}ds\omega_\mu\delta x^\mu\end{aligned} \tag{5.2.20}$$

In the last expression, we integrated by parts, used the standard relation $\delta dx^v = d\delta x^v$. The first term in the last expression vanishes. As far as δS should be zero for an arbitrary δx^μ, the equations of motion for a free relativistic particle have the form

$$\omega^\mu = \frac{du^\mu}{ds} = \frac{d^2x^\mu}{ds^2} = 0 \tag{5.2.21}$$

Now we can compare this result with the usual 3D treatment. Starting from the Lagrange function(5.2.18), one can take the derivative w.r.t. velocity,

$$p_i = \frac{\partial L}{\partial v^i} = \frac{mv^i}{\sqrt{1-\beta^2}} \Longrightarrow \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} \tag{5.2.22}$$

which is the relativistic expression for the momentum of the particle. According to classical mechanics, the energy of the particle is

$$\varepsilon = \mathbf{p}\cdot\mathbf{v} - L = \frac{mc^2}{\sqrt{1-\beta^2}} \tag{5.2.23}$$

The Lagrange equation $\frac{\partial L}{\partial q_i}(t, \boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i}(t, \boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)) = 0$ for the particle and energy conservation provide us with the relations

$$\frac{d\mathbf{p}}{dt} = 0, \quad \frac{d\varepsilon}{dt} = 0 \tag{5.2.24}$$

Both expression (5.2.22) and (5.2.23) are very interesting and tell us a lot about relativity. First of all, the nonrelativistic ($\beta \ll 1$) limits for momentum and energy are compatible with

classical mechanics, plus small corrections,

$$\mathbf{p} = m\mathbf{v} + \frac{1}{2}m\beta^2\mathbf{v} + \mathcal{O}\left(\beta^4\right), \quad \varepsilon = mc^2 + \frac{mv^2}{2} + \mathcal{O}\left(\beta^4\right) \tag{5.2.25}$$

where

$$\varepsilon_0 = mc^2$$

is called the rest energy of the particle. This formula is indeed very important, because in relativistic quantum theory one can consistently describe how to transform the rest energy of the composite or even elementary particle into other forms of energy.

Finally, if we compare the Eq.(5.2.5) and (5.2.4) with (5.2.22) and (5.2.23), we arrive at the following identification for the four-momentum:

$$p^\mu = mcu^\mu = \left(\frac{\varepsilon}{c}, \mathbf{p}\right) \tag{5.2.26}$$

Using the identity $u^\mu u_\mu = 1$ for the four-velocity, we also have the following **dispersion relation**:

$$p^2 = p^\mu p^v \eta_{\mu v} = \frac{\varepsilon^2}{c^2} - \mathbf{p}^2 = m^2c^2 \tag{5.2.27}$$

An interesting feature of (5.2.27) is that it applies not only to normal massive particles but also to massless particles and even to tachyons (assuming their masses are complex).

5.2.2 Charged Particle in Electromagnetic Field

It proves useful to introduce a four-vector of electromagnetic potential $A^\mu(t, \mathbf{r}) = (\varphi, \mathbf{A})$. And define:

$$\mathbf{H} = \operatorname{rot} \mathbf{A}$$
$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} \varphi$$

E and **H** have six degrees of freedom (that means the number of independent functions), while $A^\mu(t, \mathbf{r})$ has only four of them. This reduction is one manifestation of the fact that the

electric and magnetic fields are not independent, but form the unique electromagnetic field.

The fact that scalar φ and vector A potentials form the four-vector A^μ is important, as it enables one to construct a covariant form of the interacting term for the charged particle in the form

$$S_{\text{int}} = -\frac{e}{c}\int_{t_1}^{t_2} A_\mu dx^\mu \tag{5.2.28}$$

Let us note that $A_\mu = (\varphi, -\mathbf{A})$ and hence $A_\mu dx^\mu = -\mathbf{A}\cdot d\mathbf{r} + \varphi c dt$. Therefore, in the $3D$ notations the total action of the particle together with the interaction term can be written as follows:

$$\begin{aligned} S &= -\int_{s_1}^{s_2}\left\{mc + \frac{e}{c}A_\mu\frac{dx^\mu}{ds}\right\}ds = \int_{t_1}^{t_2} dtL \\ L &= -mc^2\sqrt{1-\beta^2} + \frac{e}{c}\mathbf{A}\cdot\mathbf{v} - e\varphi \end{aligned} \tag{5.2.29}$$

The canonically conjugated momentum is then

$$\mathbf{P} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} + \frac{e}{c}\mathbf{A} = \mathbf{p} + \frac{e}{c}\mathbf{A} \tag{5.2.30}$$

the energy of the particle is

$$\varepsilon = \mathbf{P}\cdot\mathbf{v} - L = \frac{mc^2}{\sqrt{1-\beta^2}} + e\varphi \tag{5.2.31}$$

If we make the Legendre transformation and express ε in terms of momentum as:

$$H = \sqrt{m^2c^4 + c^2\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2} + e\varphi \tag{5.2.32}$$

Using Hamilton's equation $\mathbf{P} \to -\frac{\partial S}{\partial \mathbf{r}}$, we arrive at the Hamilton-Jacobi equation in the form

$$\left(\nabla S - \frac{e}{c}\mathbf{A}\right)^2 - \frac{1}{c^2}\left(\frac{\partial S}{\partial t} + e\varphi\right)^2 + m^2c^2 = 0 \tag{5.2.33}$$

Let us now use the variational principle to obtain the equations of motion. To this end consider an arbitrary variation of the coordinates of the particle $\delta x^\mu(t)$, which is frozen at

the ends of the interval of integration, $\delta x^{\mu}(t_1) = \delta x^{\mu}(t_2) = 0$. Then, as before,

$$\delta s = u_{\mu} \frac{d\delta x^{\mu}}{ds} \tag{5.2.34}$$

hence the first variation of the action is:

$$\delta S = -\int_{s_1}^{s_2} ds \left\{ mcu_{\mu} \frac{d\delta x^{\mu}}{ds} + \frac{e}{c} u^{\mu} \left(\partial_{\alpha} A_{\mu}\right) \delta x^{\alpha} + \frac{e}{c} A_{\mu} \frac{d\delta x^{\mu}}{ds} ds \right\} \tag{5.2.35}$$

Integrating by parts and changing the names of indices, we arrive at

$$\delta S = \int_{s_1}^{s_2} ds \left\{ mc \frac{du_{\mu}}{ds} + \frac{e}{c} u^{\alpha} \left(\partial_{\alpha} A_{\mu}\right) - \frac{e}{c} \left(\partial_{\mu} A_{\alpha}\right) u^{\alpha} \right\} \delta x^{\mu} \tag{5.2.36}$$

which provides the following equation of motion:

$$mc \frac{du_{\mu}}{ds} - \frac{e}{c} u^{\alpha} F_{\mu\alpha} = 0 \tag{5.2.37}$$

In the last formula, we introduced a very important object

$$F_{\mu v} = \partial_{\mu} A_{v} - \partial_{v} A_{\mu} \tag{5.2.38}$$

which is called as **tensor of electromagnetic field**.

The explicit calculation shows that

$$F_{\mu v} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & H_z & -H_y \\ -E_y - H_z & 0 & H_x & \\ -E_z & H_y & -H_x & 0 \end{pmatrix} = (\mathrm{E}, \mathrm{H}) \tag{5.2.39}$$

$H_{1,2,3} = -H^{1,2,3}$, and $E_{1,2,3} = -E^{1,2,3}$ are regarded as components of $4D$ vector in Minkowski space. The first index μ in the $l.h.s.$ of (5.2.39) indicates the number of lines, while v numbers the columns. It is clear from (5.2.39) that the tensor under discussion is antisymmetric $F_{\mu v} = -F_{v\mu}$. For the space components F_{ik}, there is the following useful relation:

$$F_{ik} = \sum_{l=1,2,3} \varepsilon_{ikl} H_l, \qquad \text{where} \qquad H_l = H_x, H_y, H_z \tag{5.2.40}$$

The example in the previous section shows that **H** is axial vector while **E** is the usual vector, which changes the sign of its components under parity transformation. By raising indices, we have

$$F^{\mu v} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & H_z & -H_y \\ E_y & -H_z & 0 & H_x \\ E_z & H_y & -H_x & 0 \end{pmatrix} = (-\mathbf{E}, \mathbf{H}) \tag{5.2.41}$$

The dispersion relation between energy and momentum of a particle interacting with electromagnetic potential becomes

$$\left(\frac{\varepsilon - e\varphi}{c}\right)^2 = m^2c^2 + \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2 \tag{5.2.42}$$

Let's now introduce an absolutely antisymmetric Levi-Civita symbol $\varepsilon_{\alpha\beta\mu\nu}$ in Minkowski space. It is customary to define $\varepsilon^{0123} = 1$. With $\varepsilon_{\alpha\beta\mu\nu}$, we can prove that

$$\tilde{F}^{\alpha\beta} = \frac{1}{2}\varepsilon^{\alpha\beta\mu\nu}F_{\mu\nu} \tag{5.2.43}$$

where

$$\tilde{F}_{\alpha\beta} = \begin{pmatrix} 0 & H_1 & H_2 & H_3 \\ -H_1 & 0 & E_3 & -E_2 \\ -H_2 & -E_3 & 0 & E_1 \\ -H_3 & E_2 & -E_1 & 0 \end{pmatrix} = (-\mathbf{H}, \mathbf{E}) \tag{5.2.44}$$

Gauge Invariance of the Particle Action

The most important feature of Eq.(5.2.37) is that it possesses a symmetry called gauge invariance. To see this, we note that the electromagnetic potential A_μ enters these equations only through the relation (5.2.38). And this combination does not change under gauge transformation:

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu f, \quad f = f(x) = f(\mathbf{r}, t) \tag{5.2.45}$$

since

$$F'_{\mu v} = \partial_\mu A'_v - \partial_v A'_\mu = \partial_\mu A_v + \partial_\mu \partial_v f - \partial_v A_\mu - \partial_v \partial_\mu f$$
$$= \partial_\mu A_v - \partial_v A_\mu = F_{\mu v}$$

Hence, the motion of a charged particle cannot distinguish between the potentials A'_μ and A_μ. In other words, at least one out of the four degree of freedom of A_μ is not physical, as it does not affect the motion of a test particle.

The gauge transformation does not represent an essential change of the action of the particle, either. Obviously, after the transformation we only need to worry about the interaction term since the free action off the partickes does not depend on A_μ:

$$S_{\text{int}} \to S'_{\text{int}} = -\frac{e}{c}\int_{s_1}^{s_2} A'_\mu dx^\mu = -\frac{e}{c}\int_{s_1}^{s_2} (A_\mu + \partial_\mu f)\, dx^\mu$$
$$= S_{\text{int}} - \frac{e}{c}\int_{s_1}^{s_2} ds \frac{\partial f}{\partial x^\mu}\frac{dx^\mu}{ds} = S_{\text{int}} - \frac{e}{c}\int_{s_1}^{s_2} ds \frac{df}{ds}$$

The last term is the integral of the total derivative and hence does not represent an essential change of the action.

Continuous description and electric current

Let's now introduce the notion of electric current. To this end, we consider many point-like charges e_a. We assume that these particles create electric **E** and magnetic H fields, also there may be external electric and magnetic fields. Then each of the particles interacts with the field created by all the particles and by an external field, if the last is present. The interaction part of the action has the form

$$S_{\text{int}} = -\sum_a \frac{e_a}{c}\int A_\mu (x_a)\, dx_a^\mu \tag{5.2.46}$$

Let us now use **continuous description** and introduce the notation $\rho(t, \mathbf{r})$ for the density of charge. The average density in the volume V is defined as $\rho = Q/V$ where $Q = \sum_b e_b$ is the total charge since the last sum is taken over all the charges inside the volume V. It is important to note that $\rho(t, \mathbf{r})$ is not a scalar. In order to understand this, consider the same volume at rest, V_0, and in the inertial reference frame that moves with the constant speed

$\mathbf{v}$. The volume in this frame is $V = V_0\sqrt{1-\beta^2}$. Thus, the charge density transforms as

$$\rho = \frac{\rho_0}{\sqrt{1-\beta^2}} \tag{5.2.47}$$

which is not the scalar rule. As a result, one can define the four-current

$$j^\mu = \rho\frac{dx^\mu}{dt} = \rho\frac{ds}{dt}\frac{dx^\mu}{ds} = c\rho\sqrt{1-\beta^2}\frac{dx^\mu}{ds} \tag{5.2.48}$$

The time and space components of j^μ form scalar and $3D$ vector with respect to the purely spacial change of coordinates,

$$j^\mu = (c\rho, \mathbf{j}), \quad \mathbf{j}(t,\mathbf{r}) = \rho(t,\mathbf{r})\mathbf{v}$$

With the definition of j^μ, we can rewrite S_{int} in the continuous notation as:

$$S_{\text{int}} = -\int_{t_1}^{t_2} d^3x dt A_\mu \frac{dx^\mu}{cdt}\rho(t,\mathbf{r}) = \frac{1}{c^2}\int d^4x A_\mu j^\mu \tag{5.2.49}$$

and the total action is now

$$S = -\sum\int mcds - \frac{1}{c^2}\int d^4x A_\mu j^\mu \tag{5.2.50}$$

Now we are in a position to understand the fundamental role of the gauge symmetry. Let us perform the gauge transformation in the action above. The variation of the action is

$$\delta S = -\frac{1}{c^2}\int d^4x j^\mu \partial_\mu f \tag{5.2.51}$$

Integrating by parts, **we can ignore the surface term just assuming that the current j^μ vanishes on the boundary of the volume where the charges are moving**. Then the requirement of vanishing δS for any choice of the gauge parameter f leads to the identity:

$$\partial_\mu j^\mu = 0 \quad \Longrightarrow \quad \frac{\partial\rho}{\partial t} + \operatorname{div}\mathbf{j} = 0 \tag{5.2.52}$$

And this relation is the continuity equation for the charge.

5.3 Maxwell Equations in Relativistic Form

5.3.1 Minimal action principle for fields

Consider the field $\varphi(x)$, where $x = x^\mu = (x^0, x^1, x^2, x^3)$. This field $\varphi(x)$ can multicomponent, $\varphi = \varphi_i$, it can be scalar, vector, tensor, or have another nature. since these details are not relevant in what follows, for the sake of simplicity we use the condensed notation $\varphi(x)$. The action of the field $\varphi = \varphi(x)$ has the form of the $4D$ integral

$$S = \int_{\mathcal{M}} d^4x \mathcal{L}\left(\varphi, \partial_\mu \varphi\right) \tag{5.3.1}$$

Let's take an arbitrary variation of the field $\delta\varphi(x)$ in the bulk $\mathcal{M}$, but require that it vanish on the boundary,

$$\delta\varphi(x)|_{\partial\mathcal{M}} \equiv 0 \tag{5.3.2}$$

The variation $\delta\varphi(x)$ is not related to the variation of independent coordinates x hence

$$\delta\left(\frac{\partial\varphi}{\partial x^\mu}\right) = \delta\partial_\mu\varphi(x) = \partial_\mu\delta\varphi(x) = \frac{\partial(\delta\varphi)}{\partial x^\mu} \tag{5.3.3}$$

As usual in the variational principle, we require that the first variation of S vanishes, hence

$$\delta S = \int_{\mathcal{M}} d^4x \left\{\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial_\mu\varphi}\delta\left(\partial_\mu\varphi\right)\right\} \tag{5.3.4}$$

Starting from the above equation, the requirement of $\delta S = 0$ for an arbitrary variation $\delta\varphi(x)$ provides the Lagrange equation for the field,

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \frac{\partial}{\partial x^\mu}\left[\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\varphi\right)}\right] \tag{5.3.5}$$

5.3.2 Maxwell Equations in Relativistic Form

Now we shall see whether it is possible to achieve the same invariance for the action of the electromagnetic field itself. As far aswe demand Lorentz invariance, the action should be

written in terms of the 4D vector $A^\mu = (\mathbf{A}, \varphi)$. Since our goal is to reproduce the following equations:

$$\operatorname{div}\mathbf{E} = 4\pi\rho, \quad \operatorname{rot}\mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j}$$
$$\operatorname{div}\mathbf{H} = 0, \quad \operatorname{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}$$

And those are of the second order in derivatives of the potentials A and φ. Furthermore, the desired equations are linear in **E** and **H** and therefore linear in A and φ. Thus, we can expect that the action of A^μ should be **quadratic (bilinear) in the potential**. The standard definition leads to the action

$$S_{\text{em}} = -\frac{1}{16\pi c}\int d^4x F_{\mu\nu}F^{\mu\nu} \tag{5.3.6}$$

In quantum field theory, it is more common to use the coefficient $-1/4$ instead of $-1/(16\pi c)$ and the units with $c = 1$

Before using the least action principle to derive the equations of motion, let us explore the definition of the electromagnetic field tensor $F_{\mu\nu}$. It is easy to check that this antisymmetric tensor satisfies the identity

$$\partial_\alpha F_{\mu\nu} + \partial_\nu F_{\alpha\mu} + \partial_\mu F_{\nu\alpha} = 0 \tag{5.3.7}$$

and

$$\varepsilon^{\alpha\beta\mu\nu}\left(\partial_\alpha F_{\mu v} + \partial_v F_{\alpha\mu} + \partial_\mu F_{v\alpha}\right) = 0 \quad \Longrightarrow \quad \partial_\alpha \tilde{F}^{\alpha\beta} = 0 \tag{5.3.8}$$

Let us consider the $3D$ representation of Eq. (5.3.8) Consider the case $\beta = 0$ Taking into account that the terms with $\alpha = \beta$ vanish, we arrive at the equation:

$$\partial_k \tilde{F}^{k0} = 0 \Longrightarrow \operatorname{div}\mathbf{H} = 0 \tag{5.3.9}$$

In the case $\beta = i$, the equation that results from (5.3.8) can be elaborated as:

$$\partial_o \tilde{F}^{ko} + \partial_k \tilde{F}^{ki} = 0 \Longrightarrow \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = -\operatorname{rot}\mathbf{E} \tag{5.3.10}$$

In order to arrive at the last couple of Maxwell equations, consider the action of electromagnetic field (5.3.6) together with the term that corresponds to the interaction with current from (5.2.49). The relevant part of the total action is

$$S_{\rm t} = \int d^4x \left\{ -\frac{1}{16\pi c} F_{\mu v} F^{\mu v} - \frac{1}{c^2} A_\mu j^\mu \right\} \tag{5.3.11}$$

To obtain the dynamical equations we use variational principle. Let us perform an arbitrary variation $\delta A_\mu(x)$ of the potential and require that $\delta A_\mu(x)|_{(\partial M)} = 0$, where (∂M) is the boundary of the volume (M), where the dynamics of the fields occurs.

The variation of the action can be easily elaborated by using antisymmetry of $F_{\mu v}$ and integrating by parts,

$$\begin{aligned} \delta S_{\rm t} &= \int d^4x \left\{ \frac{1}{4\pi c} F^{\mu v} \cdot \partial_v \delta A_\mu - \frac{1}{c^2} j^\mu \cdot \delta A_\mu \right\} \\ &= \int d^4x \delta A_\mu \left\{ -\frac{1}{4\pi c} \partial_v F^{\mu v} - \frac{1}{c^2} j^\mu \right\} \end{aligned} \tag{5.3.12}$$

Thus, using antisymmetry of $F_{\mu\nu}$, we arrive at **the nonhomogeneous equations**

$$\partial_\mu F^{\mu v} = \frac{4\pi}{c} j^v \tag{5.3.13}$$

For $\nu = 0$:

$$\partial_k F^{k0} = \operatorname{div} \mathbf{E} = \frac{4\pi}{c} j^0 = 4\pi\rho \tag{5.3.14}$$

And for $\nu = i$,

$$-\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \operatorname{rot} \mathbf{H} = \frac{4\pi}{c} \mathbf{j} \tag{5.3.15}$$

5.3.3 Lorentz Transformation of EM Field

Assuming that the reference frame K' moves with the velocity $v = \beta c$ in the direction OX with respect to the static frame K, we get

$$\begin{aligned} ct' = x^{0'} = \frac{x^0 - \beta x^1}{\sqrt{1-\beta^2}}, \quad x' = x^{1'} = \frac{x^1 - \beta x^0}{\sqrt{1-\beta^2}} \\ y' = x^{2'} = x^2, \quad z' = x^{3'} = x^3 \end{aligned} \tag{5.3.16}$$

and the inverse transformation

$$x^0 = \frac{x^\sigma + \beta x^{1'}}{\sqrt{1-\beta^2}}, \quad x^1 = \frac{x^{1'} + \beta x^0}{\sqrt{1-\beta^2}}, \quad x^2 = x^2, \quad x^3 = x^{3'} \tag{5.3.17}$$

Let us show in detail the calculation of only the two coefficients:

$$E'_x = F'_{01} = \frac{\partial x^\alpha}{\partial x^0}\frac{\partial x^\beta}{\partial x^{1'}}F_{\alpha\beta} = \frac{\partial x^0}{\partial x^0}\frac{\partial x^1}{\partial x^{1'}}F_{01} + \frac{\partial x^1}{\partial x^{0'}}\frac{\partial x^0}{\partial x^{1'}}F_{10} = \tag{5.3.18}$$

Using the Lorentz transformation listed above, we get

$$E'_x = F'_{01} = \frac{1}{\sqrt{1-\beta^2}}\frac{1}{\sqrt{1-\beta^2}}E_x + \frac{\beta}{\sqrt{1-\beta^2}}\frac{\beta}{\sqrt{1-\beta^2}}(-E_x) = E_x \tag{5.3.19}$$

Similarly

$$\begin{aligned} E'_y = F'_{02} &= \frac{\partial x^\alpha}{\partial x^0}\frac{\partial x^\beta}{\partial x^2}F_{\alpha\beta} = \frac{\partial x^0}{\partial x^0}\frac{\partial x^2}{\partial x^{2\prime}}F_{02} + \frac{\partial x^1}{\partial x^0}\frac{\partial x^2}{\partial x^{2\prime}}F_{12} \\ &= \frac{E_y - \beta H_z}{\sqrt{1-\beta^2}} \end{aligned} \tag{5.3.20}$$

$$E'_z = \frac{E_z + \beta H_y}{\sqrt{1-\beta^2}}, \quad H'_x = H_x, \quad H'_y = \frac{H_y + \beta E_z}{\sqrt{1-\beta^2}}, \quad H'_z = \frac{H_z - \beta E_y}{\sqrt{1-\beta^2}} \tag{5.3.21}$$

5.3.4 Energy-Momentum Tensor

If we consider the dynamics of the closed system, which includes both particles and field, we need to formulate the notion of energy and momentum for the field itself and then understand how the exchange of energy and momentum occurs between particles and fields.

Indeed, the approach of the modern physics assumes that charged particles do not interact directly between themselves. **Each of them exchanges energy and momentum with the field, and thus field plays the role of intermediator in the interaction between particles.** From this perspective, the formulation of energy and momentum of the field has fundamental importance.

Let us take the two of the Maxwell equations and combine them as

$$\frac{1}{c}\left(\mathbf{E}\cdot\frac{\partial \mathbf{E}}{\partial t} + \mathbf{H}\cdot\frac{\partial \mathbf{H}}{\partial t}\right) = -\frac{4\pi}{c}\mathbf{j}\cdot\mathbf{E} - (\mathbf{H}\cdot\operatorname{rot}\mathbf{E} - \mathbf{E}\cdot\operatorname{rot}\mathbf{H}) \tag{5.3.22}$$

where

$$\mathbf{H}\cdot\operatorname{rot}\mathbf{E} - \mathbf{E}\cdot\operatorname{rot}\mathbf{H} = \operatorname{div}[\mathbf{E},\mathbf{H}]$$

Now, we have

$$\frac{1}{2c}\frac{\partial}{\partial t}\left(E^2+H^2\right)=-\frac{4\pi}{c}\mathbf{j}\cdot\mathbf{E}-\operatorname{div}[\mathbf{E},\mathbf{H}]$$

or

$$\frac{\partial}{\partial t}\left(\frac{E^2+H^2}{8\pi}\right)=-\mathbf{j}\cdot\mathbf{E}-\operatorname{div}\mathbf{S} \tag{5.3.23}$$

where

$$\mathbf{S}=\frac{c}{4\pi}\mathbf{E}\times\mathbf{H} \tag{5.3.24}$$

is called **Poynting vector**.

Consider a simple space volume (V) with the boundary (∂V) and write down the integrated form of (5.3.23), we arrive at

$$\frac{d}{dt}\left\{\iiint_{(V)}\frac{E^2+H^2}{8\pi}+\sum\varepsilon_{\text{kin}}\right\}=\oiint_{(\partial V)}\mathbf{S}\cdot d\Sigma \tag{5.3.25}$$

where

$$-\mathbf{j}\cdot\mathbf{E}=-\sum e\mathbf{v}\cdot\mathbf{E}=-\frac{d}{dt}\sum\varepsilon_{\text{kin}}$$

and

$$\rho_{\text{r}}=\frac{E^2+H^2}{8\pi}$$

is the energy density of the EM field.The physical interpretation of the Poynting vector is that it is **the space density of the vector flux of energy**. One can guess that this vector is related to the density of momentum of the field.

5.3.5 Conservation of the Energy-Momentum Tensor

Consider a closed system of fields φ, without the action of external forces. This means that the density of the Lagrange function cannot depend on x^μ explicitly, but only through the fields and their partial derivatives, $\mathcal{L}=\mathcal{L}(\varphi,\partial\varphi)$. Then

$$\frac{\partial\mathcal{L}}{\partial x^\mu}=\frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial x^\mu}+\frac{\partial\mathcal{L}}{\partial\left(\partial_\alpha\varphi\right)}\frac{\partial\left(\partial_\alpha\varphi\right)}{\partial x^\mu} \tag{5.3.26}$$

Using $\frac{\partial \mathcal{L}}{\partial \varphi} = \frac{\partial}{\partial x^\mu}\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\right]$, we have

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x^\mu} &= \frac{\partial}{\partial x^\alpha}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \varphi\right)}\right)\frac{\partial \varphi}{\partial x^\mu} + \frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \varphi\right)}\frac{\partial}{\partial x^\alpha}\left(\frac{\partial \varphi}{\partial x^\mu}\right) \\ &= \frac{\partial}{\partial x^\alpha}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \varphi\right)}\frac{\partial \varphi}{\partial x^\mu}\right)\end{aligned}$$

Using Kronecker delta symbol, we cast this relation into the form

$$\partial_\alpha \theta^\alpha{}_\mu = 0 \tag{5.3.27}$$

where

$$\theta^\alpha{}_\mu = \frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \varphi\right)}\partial_\mu \varphi - \delta^\alpha_\mu \mathcal{L} \tag{5.3.28}$$

is called **energy momentum tensor of the field** φ. The relation (5.3.27) is called conservation law.

Now we further write

$$\theta^{\alpha\beta} = \theta^\alpha{}_\mu \eta^{\mu\beta} \tag{5.3.29}$$

In general, tensor $\theta^{\alpha\beta}$ may not be symmetric, but one can modify it by adding a new term

$$\theta^{\alpha\beta} \longrightarrow T^{\alpha\beta} = \theta^{\alpha\beta} + \partial_\gamma \Psi^{\alpha\beta\gamma}, \qquad \text{where} \qquad \Psi^{\alpha\beta\gamma} = -\Psi^{\gamma\beta\alpha} \tag{5.3.30}$$

It is easy to see that $T^{\alpha\beta}$ satisfies the same conservation law, because $\partial_\alpha \partial_\gamma \Psi^{\alpha\beta\gamma} = 0$. Thus, we arrive at the new version of the energy-momentum tensor, which is symmetric $T^{\beta\alpha} = T^{\alpha\beta}$ and satisfies the conservation law

$$\partial_\alpha T^{\alpha\beta} = 0 \tag{5.3.31}$$

which reflects the same fundamental symmetry, that is, homogeneity of the space–time. An important example of the energy–momentum tensor for the of electromagnetic field with the action (5.3.6) and Lagrangian

$$\mathcal{L} = -\frac{1}{16\pi}F^2_{\mu v} = -\frac{1}{8\pi}\partial_\mu A_v\left(\partial^\mu A^v - \partial^v A^\mu\right)$$

In this formula, we used condensed notations $F^2_{\mu\nu} = F_{\mu\nu}F^{\mu\nu}$ and $\partial^\mu = \eta^{\mu\lambda}\partial_\lambda$. Taking the

required derivatives, we get

$$\frac{\partial \mathcal{L}}{\partial\left(\partial_\rho A_\sigma\right)}=-\frac{1}{4\pi}\delta_\mu^\rho\delta_v^\sigma\left(\partial^\mu A^v-\partial^v A^\mu\right)=-\frac{1}{4\pi}F^{\rho\sigma}$$
$$\text{and}\quad \frac{\partial \mathcal{L}}{\partial A_\rho}=0$$

then

$$\theta^\rho{}_\mu=\frac{\partial \mathcal{L}}{\partial\left(\partial_\rho A_\sigma\right)}\cdot\partial_\mu A_\sigma-\delta_\mu^\rho\mathcal{L}=\frac{1}{4\pi}\left(-F^{\rho\sigma}\partial_\mu A_\sigma+\frac{1}{4}\delta_\mu^\rho F_{\alpha\beta}^2\right) \tag{5.3.32}$$

It is easy to note that $\theta^{\rho\sigma}$ is not symmetric, and the symmetric version of the tensor for EM field is

$$T_{\alpha\beta}=\frac{1}{4\pi}\left(-F_\alpha^\rho\cdot F_{\beta\rho}+\frac{1}{4}g_{\alpha\beta}F_{\rho\sigma}^2\right) \tag{5.3.33}$$

where

$$\begin{aligned}
T_{00}&=\frac{E^2+H^2}{8\pi}\\
T_{0k}&=-\frac{1}{c}S^k\\
T_{ik}=\sigma_{ik}&=\frac{1}{4\pi}\left\{-E_iE_k-H_iH_k+\frac{1}{2}\delta_{ik}\left(E^2+H^2\right)\right\}
\end{aligned} \tag{5.3.34}$$

5.4 Summary: Symmetries of Maxwell Equations

Maxwell equations are Lorentz invariant. This invariance can be seen in that the action of both the electromagnetic field (5.3.6) and the massive charged particles (5.2.51) is scalars under Lorentz transformations, space rotations, and inversion of coordinates (parity transformation). Taken together, Lorentz symmetry, gauge invariance, absence of higher derivatives, and the linearity of dynamical equations (superposition principle) enable one to establish the action of the electromagnetic field in a unique way. Thus, the two fundamental symmetries form the basis of electromagnetic theory, including the possibility of the Lorentz mixing of electric and magnetic fields.

Another important sysmmetry of the actions (5.3.6) and (5.2.51) is the gauge symmetry, **which implies the conservation of electric charge.** Another symmetry of the action of a free electromagnetic field (5.3.6) is the **global rescaling of coordinates and potentials**

$$x^\mu\longrightarrow x''=x^\mu e^{-\lambda},\quad A^\mu\longrightarrow A^\mu=A^\mu e^{\lambda} \tag{5.4.1}$$

where λ is an arbitrary constant. It is easy to see that this transformation leaves the action invariant. As we shall see, this symmetry is related to the local (or global) conformal symmetry that takes place in curved space–time, typical for general relativity.