Introduction to Theoretical Physics

Michal Stanislaw Wojcik

October 23, 2022

Contents

Preface

These are my private notes which help me to understand and remember some topics from theoretical physics. If you find them useful, it is my pleasure to be helpful in your learning or research process. However, don't be disappointed if not everything is to your liking – they are only my personal notes, which just happen to be placed on github.com.

Chapter 1

Introduction to Physics

1.0.1 Planck Units

We will give values of physical constants up to 4 digits in SI units.

$$c = 2.9979 \cdot 10^8 \, [\frac{m}{s}]. \tag{1.0.1.1}$$

$$G = 6.6741 \cdot 10^{-11} \, [\frac{m^3}{kg\, s^2}]. \tag{1.0.1.2}$$

$$\hbar = 1.0546 \cdot 10^{-34} \, [\frac{kg\, m^2}{s}]. \tag{1.0.1.3}$$

$$k_e = 8.9876 \cdot 10^9 \, [\frac{m^3\, kg}{s^2 C}]. \tag{1.0.1.4}$$

$$k_b = 1.3806 \cdot 10^{-23} \, [\frac{kg\, m^2}{s^2 K}]. \tag{1.0.1.5}$$

Planck units will be t_p, l_p, m_p, q_p, T_p which satisfies the below 5 equations.

$$l_p = ct_p. \tag{1.0.1.6}$$

$$F_p = \frac{m_p l_p}{t_p^2} = G\frac{m_p^2}{l_p^2}. \tag{1.0.1.7}$$

$$E_p = F_p l_p = \frac{\hbar}{t_p}. \tag{1.0.1.8}$$

$$F_p = k_e \frac{q_p^2}{l_p^2}. \tag{1.0.1.9}$$

$$E_p = k_b T_p. \tag{1.0.1.10}$$

1. Equation (1.0.1.6) says that the light in a vacuum travels length l_p in time t_p.

2. Equation (1.0.1.7) introduces Planck unit of force F_p which by definition is equal to m_p times acceleration $\frac{l_p}{t_p^2}$. Also we require F_p to be equal to the gravitational force between two physical points with mass m_p being at distance l_p.

3. Equation (1.0.1.8) introduces Planck unit of energy as an energy needed to shift an object at distance l_p with a friction force F_p. Also we require E_p to be equal to an energy of photon with an angular frequency $1/t_p$.

4. Equation (1.0.1.9) States that the force F_p is also set to be equal to the electrostatic force between two physical points with charge q_p being at distance l_p.

5. Equation (1.0.1.10) provides a mapping from this characteristic microscopic energy E_p to the macroscopic temperature T_p.

Corollary 1.0.1.1. *In Planck units $c = G = \hbar = k_e = k_b = 1$.*

Theorem 1.0.1.2. *Equations (1.0.1.6), (1.0.1.7) and (1.0.1.8) are sufficient to uniquely establish l_p, t_p, m_p as*

$$t_p = \sqrt{\frac{G\hbar}{c^5}}, \tag{1.0.1.11}$$

$$l_p = \sqrt{\frac{G\hbar}{c^3}}, \tag{1.0.1.12}$$

$$m_p = \sqrt{\frac{\hbar c}{G}}. \tag{1.0.1.13}$$

Proof. Substituting l_p with ct_p in equation (1.0.1.7) leads to relation

$$m_p = t_p \frac{c^3}{G}. \tag{1.0.1.14}$$

On the other hand substituting l_p with ct_p in equation (1.0.1.8) leads to relation

$$m_p = \frac{\hbar}{c^2 t_p}. \tag{1.0.1.15}$$

From those two we get directly $t_p = \sqrt{\frac{G\hbar}{c^5}}$, which leads to (1.0.1.12) and (1.0.1.13). □

Corollary 1.0.1.3. *Additionally from equations (1.0.1.9) and (1.0.1.10) follows:*

$$q_p = \sqrt{\frac{\hbar c}{k_e}}, \tag{1.0.1.16}$$

$$T_p = \sqrt{\frac{\hbar c^5}{G k_b}}. \tag{1.0.1.17}$$

In equations in Planck units, all mentioned above physical constants are set to 1, like

$$E = \sqrt{p^2 + m^2} \tag{1.0.1.18}$$

or

$$i\frac{d\psi}{dt} = H\psi. \tag{1.0.1.19}$$

Taking the above equations as an example, we will investigate how to reconstruct constants to get equations in SI. In equation (1.0.1.18) we have $E[\frac{m_p l_p^2}{t_p^2}]$ and $p[\frac{m_p l_p}{t_p}]$ and $m[m_p]$. Thus in SI we need 2 constants C_1 and C_2 such that $E = \sqrt{C_1^2 p^2 + C_2^2 m^2}$. We have

$$\frac{m_p l_p^2}{t_p^2} = C_1 \frac{m_p l_p}{t_p} \tag{1.0.1.20}$$

and

$$\frac{m_p l_p^2}{t_p^2} = C_2 m_p. \tag{1.0.1.21}$$

Thus

$$C_1 = \frac{l_p}{t_p} = c. \tag{1.0.1.22}$$

and

$$C_2 = \frac{l_p^2}{t_p^2} = c^2. \tag{1.0.1.23}$$

Therefore equation (1.0.1.18) in SI has a form

$$E = \sqrt{c^2p^2 + c^4m^2}. \tag{1.0.1.24}$$

In equation (1.0.1.19) we have $\frac{d\psi}{dt}[\frac{1}{t_p}]$ and $H[E_p]$, thus for SI we need constant C such that $iC\frac{d\psi}{dt} = H\psi$. We have

$$C\frac{1}{t_p} = \frac{\hbar}{t_p}, \tag{1.0.1.25}$$

Thus

$$C = \hbar. \tag{1.0.1.26}$$

Therefore equation (1.0.1.19) in SI has a form

$$i\hbar\frac{d\psi}{dt} = H\psi. \tag{1.0.1.27}$$

1.0.2 Natural Units

Assume we have a unit of energy U_E. We will express time, length and momentum as powers of U_E.

Definition 1.0.2.1. *The unit of length $1U_E^{-1}$ is equal to a wavelength of a photon with energy of $2\pi U_E$.*

Definition 1.0.2.2. *The unit of time $1U_E^{-1}$ is equal to the period of the wave of a photon with energy of $2\pi U_E$.*

Corollary 1.0.2.3. *In the above units $c = 1$.*

Corollary 1.0.2.4. *In the above units $\hbar = 1$*

Proof. For a photon, we have

$$E = \frac{2\pi\hbar c}{\lambda}, \tag{1.0.2.1}$$

where E is energy of the photon and λ is a wavelength. We already established that $c = 1$, thus

$$\hbar = \frac{E}{2\pi}\lambda. \tag{1.0.2.2}$$

The above equation holds for a photon with energy $2\pi U_E$ but such a photon, by Definition 1.0.2.1 has a wavelength $1U_E^{-1}$. Thus, after substitution to (1.0.2.2) we get $\hbar = 1$. □

Proposition 1.0.2.5. *To covert length of* $1U_E^{-1}$ *to units system* X*, one needs to calculate in the units system* X

$$\lambda = \frac{\hbar c}{U_E}. \tag{1.0.2.3}$$

Proof. Follows from the equation (1.0.2.2). □

Proposition 1.0.2.6. *To covert time of* $1U_E^{-1}$ *to units system* X*, one needs to calculate in the units system* X

$$t = \frac{\hbar}{U_E}. \tag{1.0.2.4}$$

Proof. Follows directly from the Definition (1.0.2.2). □

Definition 1.0.2.7. *The unit of mass* $1U_E$ *is equal to the mass of an object with an rest energy* U_E*.*

Proposition 1.0.2.8. *To covert mass of* $1U_E$ *to units system* X*, one needs to calculate in the units system* X

$$m = \frac{U_E}{c^2}. \tag{1.0.2.5}$$

Proof. Follows from the equation $E = mc^2$. □

Definition 1.0.2.9. *The unit of momentum* $1U_E$ *is equal to the magnitude of momentum of a photon in an energy* U_E*.*

Proposition 1.0.2.10. *To convert momentum of* $1U_E$ *to units system* X*, one needs to calculate in the units system* X

$$p = \frac{U_E}{c}. \tag{1.0.2.6}$$

Definition 1.0.2.11. *The unit of force* $1U_E^2$ *is equal to the force which is equivalent to the change of* $1U_E$ *momentum in* $1U_E^{-1}$ *time.*

Proposition 1.0.2.12. *To convert a force of* $1U_E^2$ *to units of system X, one needs to calculate in the units of system* X

$$F = \frac{U_E^2}{\hbar c}. \tag{1.0.2.7}$$

Proof. By Preposition (1.0.2.10) and Proposition (1.0.2.6) □

Assume that we will take $Q = \sqrt{\varepsilon_0 \hbar c}$ in SI as our unit of electric charge. Let's calculate a force F with which 2 Q charges will repel each other from the distance of $1U_E^{-1}$. Let's do calculations in SI.

$$F = \frac{(\sqrt{\varepsilon_0 \hbar c})^2}{4\pi\varepsilon_0 (\frac{\hbar c}{U_E})^2} = \frac{1}{4\pi}\frac{U_E^2}{\hbar c}. \tag{1.0.2.8}$$

That means that in units U_E, $F = \frac{1}{4\pi}U_E^2$. Therefore that equation for the Coulomb force in units U_E is

$$F = \frac{q^2}{4\pi r^2}, \tag{1.0.2.9}$$

where q is charge dimensionless (or U_E^0) and r is in U_E^{-1}. From that follows that in U_E system $\varepsilon_0 = 1$.

Proposition 1.0.2.13. *To convert 1 unit of electric charge to the units of system X, one needs to calculate in units of system X*

$$Q = \sqrt{\varepsilon_0 \hbar c} \tag{1.0.2.10}$$

For example in SI $Q = 5.290817690 \cdot 10^{-19}C$. Since elementary charge (electron charge) in SI is

$$e = 1.60218 \cdot 10^{-19}C, \tag{1.0.2.11}$$

$e = 0.3028221209$ dimensionless in units U_E.

Note that we established a system where time and length have dimension U_E^{-1}, mass, energy and momentum have dimension U_E, force has dimension U_E^2 and charge is dimensionless. Moreover

$$\boxed{\hbar = c = \varepsilon_0 = 1} \tag{1.0.2.12}$$

and elementary charge is 0.3028221209.

Chapter 2

Classical Mechanics

2.1 Energy

2.1.1 System with potential energy

In the considerations below, we will not distiguish between column and raw vectors. We will consider a system of particles, which positions will be denoted by

$$q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_n \end{bmatrix} \tag{2.1.1.1}$$

where

$$\vec{q}_i = \begin{bmatrix} q_i^1 & q_i^2 & q_i^3 \end{bmatrix}. \tag{2.1.1.2}$$

We consider q as depentent on t. By $\dot{q}$ we denote $\frac{dq}{dt}$.

Let

$$F = \begin{bmatrix} \vec{F}_1 & \vec{F}_2 \dots \vec{F}_n \end{bmatrix} \tag{2.1.1.3}$$

where $\vec{F}_i$ is a force acting on i-th partcile.

Assume that we have some scalar value V which dependents on q and

$$F = -\frac{\partial V}{\partial q}. \tag{2.1.1.4}$$

By this we assume that for our system of partciles forces depend only on position of particles.

Assume now that our system evolve from a state q_0 at $t = t_0$ to a state q_1 at $t = t_1$. Let's try to calculate work of the forces F when system changed from q_0 to q_1 for an arbitrary evolution in time $q(t)$ with this constrain only that $q(t_0) = q_0$ and $q(t_1) = q_1$.

$$W = \int_{t_0}^{t_1} F \cdot \frac{dq}{dt} dt = -\int_{t_0}^{t_1} \frac{\partial V}{\partial q} \cdot \frac{dq}{dt} dt = -\int_{t_0}^{t_1} \frac{dV}{dt} = V(q_0) - V(q_1). \quad (2.1.1.5)$$

$$\boxed{W = V(q_0) - V(q_1)} \quad (2.1.1.6)$$

That's why we call V a potential energy in a state q. The energy that the forces of system needs to use is independent on the path of evolution. You can know this just but substracting respectively potential energy of an start poit and end point.

2.1.2 Lagrangian picture

Assume that the system from Subsection 2.1.1 is described by Lagranian

$$L = T - V. \quad (2.1.2.1)$$

Where T is a kinetic energy of the system

$$T = \sum_{i=1}^{n} \frac{1}{2} m_i |\dot{\vec{q_i}}|^2. \quad (2.1.2.2)$$

And we know about V only that it is dependent only on q. Now, let's check what we can get from Lagrange equation

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0. \quad (2.1.2.3)$$

Firstly, note that it translates into

$$-\frac{\partial V}{\partial q} - \frac{d}{dt}\frac{\partial T}{\partial \dot{q}} = 0, \quad (2.1.2.4)$$

hence

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}} = -\frac{\partial V}{\partial q}. \quad (2.1.2.5)$$

Note that

$$\frac{\partial T}{\partial \dot{\vec{q_i}}} = m_i \dot{\vec{q_i}}, \quad (2.1.2.6)$$

thus

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\vec{q_i}}} = m_i \ddot{\vec{q_i}} = \vec{F_i}. \quad (2.1.2.7)$$

Where $\vec{F}_i$ is just a newtonian dynamical force, which can be measured just by mass times acceleration.

From that we get

$$F = -\frac{\partial V}{\partial q}, \tag{2.1.2.8}$$

which is the equation (2.1.1.4) and thus the rest of the Subsection 2.1.1 applies, so we can call V potential energy.

Because L doesn't depend on t, we get

$$L - \frac{\partial L}{\partial \dot{q}}\dot{q} = \text{const}. \tag{2.1.2.9}$$

The above follows imediatelly from Noether Theorem (Theorem 11.2.2.3) but also can be derived directly from calculating $\frac{d}{dt}(L - \frac{\partial L}{\partial \dot{q}}\dot{q})$ and applying (2.1.2.3) to $\frac{dL}{dt}$.

$$\frac{\partial L}{\partial \dot{q}}\dot{q} = \frac{\partial T}{\partial \dot{q}}\dot{q} = \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{\vec{q}}_i} \cdot \dot{\vec{q}}_i = \sum_{i=1}^{n} m_i |\dot{\vec{q}}_i|^2 = 2T. \tag{2.1.2.10}$$

Thus $L - 2T = \text{const}$, from which by (2.1.2.1) follows imediatelly

$$T + V = \text{const}. \tag{2.1.2.11}$$

Which is energy conservation principle.

2.1.3 Hamiltonian picture

Assume that the system from from Subsection 2.1.1 is described by Hamiltonian

$$H = T + V, \tag{2.1.3.1}$$

where V is defined like in Subsection 2.1.2 and

$$T = \sum_{i=1}^{n} \frac{1}{2m_i} |\vec{p}_i|^2. \tag{2.1.3.2}$$

System satisfies Hamilton's equations:

$$\frac{\partial H}{\partial q} = -\dot{p}, \tag{2.1.3.3}$$

$$\frac{\partial H}{\partial p} = \dot{q}. \tag{2.1.3.4}$$

Let's try to calculate $\ddot{q}$.

$$\ddot{\vec{q_i}} = \frac{d}{dt}\frac{\partial H}{\partial \vec{p_i}} = \frac{\partial^2 H}{\partial \vec{p_i}\partial t} + \frac{\partial^2 H}{\partial \vec{p_i}\partial q}\frac{dq}{dt} + \frac{\partial^2 H}{\partial \vec{p_i}\partial p}\frac{dp}{dt}$$
$$= \frac{\partial^2 H}{\partial \vec{p_i}\partial p}\dot{p} = \frac{\partial^2 H}{\partial \vec{p_i}\partial \vec{p_i}}\dot{\vec{p_i}} = \frac{1}{m_i}\dot{\vec{p_i}}. \quad (2.1.3.5)$$

Thus

$$m_i\ddot{\vec{q_i}} = \dot{\vec{p_i}}. \quad (2.1.3.6)$$

Therefore, from definition $\vec{F_i} = \dot{\vec{p_i}}$ and $F = \dot{p}$. Hence, from (2.1.3.4)

$$F = -\frac{\partial V}{\partial q}, \quad (2.1.3.7)$$

which is the equation (2.1.1.4) and thus the rest of the Subsection 2.1.1 applies, so we can call V potential energy. Note that

$$\frac{d}{dt}H = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} = \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial H}{\partial q} = 0. \quad (2.1.3.8)$$

Thus $H = \text{const}$ and it imediatelly follows from (2.1.3.1) that

$$T + V = \text{const}. \quad (2.1.3.9)$$

Which is energy conservation principle.

2.2 Reduced Mass

Consider Lagrangian

$$L(q, \dot{q}) = \frac{1}{2}m_1(\dot{\vec{q_1}})^2 + \frac{1}{2}m_2(\dot{\vec{q_2}})^2 + V(|\vec{q_1} - \vec{q_2}|). \quad (2.2.0.1)$$

We apply convention similiar as in Subsection 2.1.1, where $q = [\vec{q_1}, \vec{q_2}]$. Recall that we say that, q_0 extremises L, if $\int_{t_0}^{t_1} L(q_0, \dot{q_0})$ is a local extremum for all evolutions q for which $q(t_0) = q_0(t_0)$ and $q(t_1) = q_0(t_1)$ for any moments of time t_0, t_1.

Take any constant velocity $\vec{v}$ and constant ponit c_0. Let

$$\vec{x_i} := \vec{q_i} + t\vec{v} + c_0 \text{ for } i = 1, 2. \quad (2.2.0.2)$$

Following our convention $x = [\vec{x_1}, \vec{x_2}]$. First, we will show that

Fact 2.2.0.1. *x extremises $L(x, \dot{x})$ if and only if q extremises $L(q, \dot{q})$.*

Obviously V part of L stays the same as $\vec{q_1} - \vec{q_2} = \vec{x_1} - \vec{x_2}$. Let's calculate

$$\frac{1}{2}m_1(\dot{\vec{x_1}})^2 + \frac{1}{2}m_2(\dot{\vec{x_2}})^2 = \frac{1}{2}m_1(\dot{\vec{q_1}})^2 + \frac{1}{2}m_2(\dot{\vec{q_2}})^2 + (m_1\dot{\vec{q_1}} + m_2\dot{\vec{q_2}}) \cdot \vec{v} + (m_1 + m_2)(\vec{v})^2. \quad (2.2.0.3)$$

Assume that q maximises L. From Noether Theorem, we know that $\frac{\partial L}{\partial \dot{q}} =$ const, which is a momentum conservation principle, thus $m_1\dot{\vec{q_1}} + m_2\dot{\vec{q_2}} =$ const. Hence,

$$\frac{1}{2}m_1(\dot{\vec{x_1}})^2 + \frac{1}{2}m_2(\dot{\vec{x_2}})^2 = \frac{1}{2}m_1(\dot{\vec{q_1}})^2 + \frac{1}{2}m_2(\dot{\vec{q_2}})^2 + \text{const}. \quad (2.2.0.4)$$

And from that

$$L(q, \dot{q}) = L(x, \dot{x}) + \text{const}. \quad (2.2.0.5)$$

From the above x extremises L. The situation is symetric if we want to show that x extremises L implies q extremises L. This means that for Lagrangian of type (2.2.0.1) any change of choordinates between frames of reference moving with constant velocity relative to each other does not change the Lagrangian of the system.

From the statement above follows that without loss of generality in case of the Lagrangian of type (2.2.0.1), we can always assume that the center of the mass is in the point $(0, 0, 0)$, i.e.

$$m_1\vec{q_1} + m_2\vec{q_2} = 0 \quad (2.2.0.6)$$

Now, if we take

$$\vec{r} := \vec{q_1} - \vec{q_2}, \quad (2.2.0.7)$$

from equations (2.2.0.6) and (2.2.0.7), we have

$$\begin{cases} \vec{q_1} = \dfrac{m_2}{m_1 + m_2}\vec{r}, \\ \vec{q_2} = -\dfrac{m_1}{m_1 + m_2}\vec{r}. \end{cases} \quad (2.2.0.8)$$

Now, after simple calculation

$$L = \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}(\dot{\vec{r}})^2 + V(|\vec{r}|). \quad (2.2.0.9)$$

Hence $\vec{r}$ evolves in the same way as one particle of mass

$$\boxed{m = \frac{m_1 m_2}{m_1 + m_2}} \tag{2.2.0.10}$$

In a central force field $\vec{F} = -\frac{\partial V}{\partial \vec{r}}$. We call m a reduced mass.

2.3 Statistical Mechanics

2.3.1 Flux

Let $\rho(t, x)$ will be a density of a certain abstract continous entity in time t and point $x \in \mathbb{R}^n$. Let $\vec{v}(t, x)$ will be a velocity of an infinitesimal element of the continous entity in time t at point $x \in \mathbb{R}^n$. The current of the continous entity is

$$\vec{j} = \rho \vec{v}. \tag{2.3.1.1}$$

Imagine an infinitesimal $n-1$-dimentional almost hyperplanar surfice elment dS with it's normal unit vector $\vec{n}$. We will try to calculate an amount of the continous entity dm which flew through surface dS in the direction of the arrow of $\vec{n}$ during infenitesimal time Δt.

If we consider an infenitesimal element of the entity, the component of its movement which is in a plane of dS doesn't play any role in its going through surface element dS. Since $\vec{v} \cdot \vec{n}$ is a component of veliocity $\vec{v}$ which is perpendicular to dS, this is the velocity with which the entity passes through dS. Thus

$$\Delta m = \rho dS \Delta t (\vec{v} \cdot \vec{n}). \tag{2.3.1.2}$$

Hence

$$\Delta m = \vec{j} \cdot \vec{n} dS \Delta t. \tag{2.3.1.3}$$

If we take any nice enough connected open $\Omega \subset \mathbb{R}^n$ with compact clousure. We assume that entity is conserved in time, which simply means that it's not disapearing anywhere. Thus, the entity gain inside Ω during time Δt must be equal to the negative total flux of the entity through $\partial\Omega$. Thus

$$\int_\Omega \rho(t + \Delta t, x) dx - \int_\Omega \rho(t, x) dx = -\int_{\partial\Omega} \Delta m = -\Delta t \int_{\partial\Omega} \vec{j} \cdot \vec{n} dS. \tag{2.3.1.4}$$

Hence

$$\int_\Omega \frac{\partial \rho}{\partial t} dx = -\int_{\partial\Omega} \vec{j} \cdot \vec{n} dS. \tag{2.3.1.5}$$

Note that the reasoning above has sense, because by continuity of ρ the change of ρ during time Δt is infinitesimal while at the same time

$$\frac{\rho(t+\Delta t, x) - \rho(t, x)}{\Delta t} = \frac{\partial \rho}{\partial t}(x, t)$$

is a significant value because of differentiability of ρ over t. We need as well assume continuity of $\frac{\partial \rho}{\partial t}(x, t)$ to be able to go with diffrentiation under the intergral.

Now by Theorem 12.1.2.17 (Gauss's Theorem), we have

$$\int_\Omega \frac{\partial \rho}{\partial t} dx = -\int_\Omega \nabla \cdot \vec{j} dx. \tag{2.3.1.6}$$

Since Ω can be arbitrary, we have

$$\boxed{\frac{\partial \rho}{\partial t} = -\nabla \cdot \vec{j}} \tag{2.3.1.7}$$

This is the relation between current and density of any continous entity which is conserved in time.

2.3.2 Liouville's equation

Consider a phase space with hamiltonian H. Let's recall Hamilton equations

$$\frac{\partial H}{\partial x} = -\dot{p}, \tag{2.3.2.1}$$

$$\frac{\partial H}{\partial p} = \dot{x}. \tag{2.3.2.2}$$

For convenience, in phase space we usually use Poisson brackets

$$\{f, g\} := \sum_i \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial x_i}\frac{\partial f}{\partial p_i}. \tag{2.3.2.3}$$

Let $\vec{v} = (\dot{x}, \dot{p}) = (\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x})$ be a veliocity of a point in phase space. Assume that we have certain probability density of position and momentum $f(t, x, p)$. Let $\vec{j} = f(t, x, p)\vec{v}$ be a current of probability. We assume that probability is conserved in time. Which means

$$\frac{\partial f}{\partial t} = -\nabla \cdot \vec{j}. \tag{2.3.2.4}$$

Now,

$$\nabla \cdot \vec{j} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial p}) \cdot (f\frac{\partial H}{\partial p}, -f\frac{\partial H}{\partial x}) =$$
$$\frac{\partial f}{\partial x} \cdot \frac{\partial H}{\partial p} + f\frac{\partial H}{\partial x} \cdot \frac{\partial H}{\partial p} - \frac{\partial f}{\partial p} \cdot \frac{\partial H}{\partial x} - f\frac{\partial H}{\partial p} \cdot \frac{\partial H}{\partial x} =$$
$$\frac{\partial f}{\partial x} \cdot \frac{\partial H}{\partial p} - \frac{\partial f}{\partial p} \cdot \frac{\partial H}{\partial x} = \{f, H\}. \quad (2.3.2.5)$$

Thus the time evolution of probability density is described by Liouville's equation

$$\boxed{\frac{\partial f}{\partial t} = -\{f, H\}} \quad (2.3.2.6)$$

Chapter 3

Classical Electromagnetism

3.1 Introduction

3.1.1 Maxwell's equations

Despite we are in the realm of classical theory we will still use Plank units. We are all the time able to restore constants.

Definition 3.1.1.1. *We say that in some region Ω there exists an electric field $\vec{E} : \Omega \to \mathbb{R}^3$, if in each point of Ω a force $\vec{F} = q\vec{E}$ acts on an test charge q.*

Definition 3.1.1.2. *We say that in some region Ω there exists a magnetic field $\vec{B} : \Omega \to \mathbb{R}^3$, if in each point of Ω a force $\vec{F} = q\vec{v} \times \vec{B}$ acts on a test charge moving with a velocity $\vec{v}$.*

According to Maxwell's theory relations between charge, electric field and magnetic filed are governed by four Maxwell's equations.

$$\nabla \cdot \vec{E} = 4\pi\rho. \tag{3.1.1.1}$$

$$\nabla \cdot \vec{B} = 0. \tag{3.1.1.2}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{3.1.1.3}$$

$$\nabla \times \vec{B} = 4\pi \vec{J} + \frac{\partial \vec{E}}{\partial t}. \tag{3.1.1.4}$$

Where ρ is electric charge density in space and $\vec{J}$ is an electric charge current in space.

3.1.2 Magnetic Moment

Imagine a closed circuit Γ. Let $(\Gamma, \vec{l})$ be an orientation of the circuit consistent with direction of movemet of positive charges. Consider an infinitesimal element ds of Γ. Define a strength of electric current as

$$I = \frac{dq}{dt}. \tag{3.1.2.1}$$

where dq is a portion of charge which passes through element ds in time dt. Assume that the strength of electric current is constant through the whole circuit Γ.
Let's investigate what is an infinitesimal force $d\vec{F}$ acting on an infinitesimal element ds in the presence of magnetic field $\vec{B}$. Let's calculate a velocity $\vec{v}$ of an infinitesimal charge dq in an infinitesimal element ds of the ciruit Γ.

$$\vec{v} = \frac{(ds)\vec{l}}{dt}. \tag{3.1.2.2}$$

Thus

$$d\vec{F} = dq\vec{v} \times \vec{B} = \frac{dq(ds)\vec{l}}{dt} \times \vec{B} = I(ds)\vec{l} \times \vec{B}. \tag{3.1.2.3}$$

$$\boxed{d\vec{F} = I(ds)\vec{l} \times \vec{B}} \tag{3.1.2.4}$$

Assume that Γ is an edge of certain C^1 surface $(S, \vec{n})$ oriented consistently with Γ. We will calculate torque $\vec{M}$ which homogenous magnetic filed $\vec{B}$ exerts on Γ.

$$\vec{M} = \int_{\Gamma} \vec{r} \times d\vec{F}. \tag{3.1.2.5}$$

Thus

$$\vec{M} = I\int_\Gamma \vec{r}\times((ds)\vec{l}\times\vec{B}) = I\int_\Gamma [x_1,x_2,x_3]\times([dx_1,dx_2,dx_3]\times[B_1,B_2,B_3]) = \\ I\int_\Gamma \\ [B_2x_2dx_1 - B_1x_2dx_2 + B_3x_3dx_1 - B_1x_3dx_3, \\ -B_2x_1dx_1 + B_1x_1dx_2 + B_3x_3dx_2 - B_2x_3dx_3, \\ -B_3x_1dx_1 + B_1x_1dx_3 - B_3x_2dx_2 + B_2x_2dx_3] \tag{3.1.2.6}$$

By Stokes Theorem (Theorem 12.1.2.14), we have:

$$\vec{M} = I\int_S \\ [B_3dx_3\wedge dx_1 - B_2dx_1\wedge dx_2, \\ B_1dx_1\wedge dx_2 - B_3dx_2\wedge dx_3, \\ B_2dx_2\wedge dx_3 - B_1dx_3\wedge dx_1] \tag{3.1.2.7}$$

Now, by Example 12.1.2.16, we have

$$\vec{M} = I\int_S \\ [n_2B_3 - n_3B_2, \\ n_3B_1 - n_1B_3, \\ n_1B_2 - n_2B_1]dS \\ = I\int_S \vec{n}\times\vec{B}dS = (I\int_S \vec{n}dS)\times\vec{B}. \tag{3.1.2.8}$$

Thus

$$\vec{M} = (I\int_S \vec{n}dS)\times\vec{B}. \tag{3.1.2.9}$$

Equation (3.1.2.9) tells us two important things. 1. $\vec{M}$ doesn't depend on choice of a central point, and thus total force exerted by homogenous magnetic field $\vec{B}$ on Γ is $\vec{F} = 0$. 2. $I\int_S \vec{n}dS$ doesn't depend on choice of surface S. Then we can define magnetic moment $\vec{\mu}$ of circuit Γ.

$$\boxed{\vec{\mu} := I \int_S \vec{n} dS} \tag{3.1.2.10}$$

$$\boxed{\vec{M} = \vec{\mu} \times \vec{B}} \tag{3.1.2.11}$$

Note that if Γ is contained in plane, $\vec{\mu}$ is perpendicular to the plane and from its tip the positive current is going counter-clockwise, moreover $\|\vec{\mu}\| = IS$ where S is an area cut by circuit Γ.

3.1.3 Magnetic Moment in Inhomogenous Magnetic Field

Imagine a closed circuit Γ in inhomogenous magnetic field $\vec{B}$. Let $(\Gamma, \vec{l})$ be an orientation of the circuit consistent with direction of movemet of positive charges. By equation (3.1.2.4) the force exerted on the circiut is

$$\begin{aligned} \vec{F} = \int_\Gamma I(ds)\vec{l} \times \vec{B} = \\ I \int_\Gamma [B_3 dx_2 - B_2 dx_3, B_1 dx_3 - B_3 dx_1, B_2 dx_1 - B_1 dx_2] \end{aligned} \tag{3.1.3.1}$$

Assume that Γ is an edge of certain C^1 surface $(S, \vec{n})$ oriented consistently with Γ. We will prepare to use Stokes's Theorem (Theorem 12.1.2.14). Let's do a bit of form calculus (see 12.1.2)

$$\begin{aligned} d(B_3 dx_2 - B_2 dx_3) = \\ \frac{\partial B_3}{\partial x_1} dx_1 \wedge dx_2 + \frac{\partial B_3}{\partial x_3} dx_3 \wedge dx_2 \\ - \frac{\partial B_2}{\partial x_1} dx_1 \wedge dx_3 - \frac{\partial B_2}{\partial x_2} dx_2 \wedge dx_3 = \\ \frac{\partial B_3}{\partial x_1} dx_1 \wedge dx_2 - \frac{\partial B_3}{\partial x_3} dx_2 \wedge dx_3 \\ + \frac{\partial B_2}{\partial x_1} dx_3 \wedge dx_1 - \frac{\partial B_2}{\partial x_2} dx_2 \wedge dx_3 \end{aligned} \tag{3.1.3.2}$$

By Maxwell Equation (3.1.1.2) $\nabla \cdot \vec{B} = 0$. Thus

$$d(B_3 dx_2 - B_2 dx_3) = \frac{\partial B_1}{\partial x_1} dx_2 \wedge dx_3 + \frac{\partial B_2}{\partial x_1} dx_3 \wedge dx_1 + \frac{\partial B_3}{\partial x_1} dx_1 \wedge dx_2. \tag{3.1.3.3}$$

Hence

$$F_1 = \int_S \frac{\partial \vec{B}}{\partial x_1} \cdot \vec{n} dS. \tag{3.1.3.4}$$

Making analogous calculations for the rest of coordinates of $\vec{F}$, we get

$$F_i = \int_S \frac{\partial \vec{B}}{\partial x_i} \cdot \vec{n} dS \text{ for } i = 1, 2, 3. \tag{3.1.3.5}$$

Assuming that $\nabla \vec{B}$ is changing insignificanlty within the size of circuit Γ, we can write

$$\boxed{F_i = \vec{\mu} \cdot \frac{\partial \vec{B}}{\partial x_i} \text{ for } i = 1, 2, 3.} \tag{3.1.3.6}$$

3.1.4 Classical Relation Between Angular Momentum and Magnetic Moment

Consider one particle with mass m and charge q moving around in a circle of radious r with velocity v. Angular Momentum of such a particle is

$$\vec{L} = rmv\vec{n} \tag{3.1.4.1}$$

where $\vec{n}$ is a unit vector normal to the plain of rotation from which tip the rotation is counter-clockwise.

The electric current caused by circulation of the charge is $I = \frac{qv}{2\pi r}$. As we proved in the subsection 3.1.2, the magnetic moment will be $\vec{\mu} = \pi r^2 I \vec{n} = \frac{1}{2} rqv\vec{n}$. Thus we have a relation

$$\boxed{\vec{\mu} = \frac{q}{2m}\vec{L}.} \tag{3.1.4.2}$$

Because of additivity the equation above is also true for rotating rigid body with charge and mass homogonously distributed.

3.1.5 Lagrangian and Hamiltonian Formulation of Magnetic Field

In this subsection we will use bold font to indicate space vectors. Let $\boldsymbol{x}$ indicate trajectory of the charged particle, we assume that $\boldsymbol{x}$ is dependent on time t.

Assume that Lagrangian for the charged particle with charge q and mass m is

$$L = \frac{m\dot{\boldsymbol{x}}^2}{2} + q(\boldsymbol{A} \cdot \dot{\boldsymbol{x}}), \tag{3.1.5.1}$$

where $\boldsymbol{A}$ is a vector field (vector dependent on position). In case of analysing trajectory, we can assume that $\boldsymbol{A}$ depends on $\boldsymbol{x}$. Vector $\boldsymbol{A}$ is usually called a vector potencial.

Let's recall Euler-Lagrange equation, which is satisfied along trajectory $\boldsymbol{x}$:

$$\frac{\partial L}{\partial \boldsymbol{x}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{x}}} = 0. \tag{3.1.5.2}$$

Since $\boldsymbol{A}$ depends on $\boldsymbol{x}$ note that

$$\frac{\partial L}{\partial \boldsymbol{x}} = q \sum_{k=1}^{3} \frac{\partial A_k}{\partial \boldsymbol{x}} \dot{x}_k, \tag{3.1.5.3}$$

Also, note that

$$\frac{\partial L}{\partial \dot{\boldsymbol{x}}} = m\dot{\boldsymbol{x}} + q\boldsymbol{A}, \tag{3.1.5.4}$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{x}}} = m\ddot{\boldsymbol{x}} + q\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}}. \tag{3.1.5.5}$$

Mind that in the above equation $\dfrac{\partial \boldsymbol{A}}{\partial \boldsymbol{x}}$ is simply a Jacobian matrix acting on vector $\dot{\boldsymbol{x}}$.

From Euler-Lagrange equation, we have then

$$m\ddot{\boldsymbol{x}} = q \sum_{k=1}^{3} \frac{\partial A_k}{\partial \boldsymbol{x}} \dot{x}_k - q\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}}. \tag{3.1.5.6}$$

Let's see the above equation in the index notation:

$$m\ddot{x}_i = q \sum_{k=1}^{3} \left(\frac{\partial A_k}{\partial x_i}\dot{x}_k - \frac{\partial A_i}{\partial x_k}\dot{x}_k\right) = q \sum_{k=1}^{3} \dot{x}_k \left(\frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}\right). \tag{3.1.5.7}$$

Note that for $i = k$, we have $\dfrac{\partial A_k}{\partial x_i} - \dfrac{\partial A_i}{\partial x_k} = 0$. Thus,

$$m\ddot{x}_i = q \sum_{k \in \{1,2,3\} \setminus \{i\}} \dot{x}_k \left(\frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}\right). \tag{3.1.5.8}$$

Let's define a new vector field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, which means simply that

$$B_1 = \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}, \tag{3.1.5.9}$$

$$B_2 = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}, \tag{3.1.5.10}$$

$$B_3 = \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}. \tag{3.1.5.11}$$

[Compare the above with (13.5.2.20)].

Then from equation 3.1.5.6, we have:

$$m\ddot{x}_1 = q\dot{x}_2\big(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\big) + q\dot{x}_3\big(\frac{\partial A_3}{\partial x_1} - \frac{\partial A_1}{\partial x_3}\big) = q(\dot{x}_2 B_3 - \dot{x}_3 B_2), \tag{3.1.5.12}$$

$$m\ddot{x}_2 = q\dot{x}_1\big(\frac{\partial A_1}{\partial x_2} - q\frac{\partial A_2}{\partial x_1}\big) + \dot{x}_3\big(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}\big) = q(-\dot{x}_1 B_3 + \dot{x}_3 B_1), \tag{3.1.5.13}$$

$$m\ddot{x}_3 = q\dot{x}_1\big(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\big) + q\dot{x}_2\big(\frac{\partial A_2}{\partial x_3} - \frac{\partial A_3}{\partial x_2}\big) = q(\dot{x}_1 B_2 - \dot{x}_2 B_1). \tag{3.1.5.14}$$

Now, we can go back to vector notation to get:

$$m\ddot{\boldsymbol{x}} = q(\dot{\boldsymbol{x}} \times \boldsymbol{B}), \tag{3.1.5.15}$$

which is a well known equation of force exerted by magnetic filed $\boldsymbol{B}$ on particle with charge q and velocity $\dot{\boldsymbol{x}}$.

Chapter 4

Special Relativity

4.1 Minkowski space-time

Definition 4.1.0.1. *We call a tensor $g_{\mu\nu}$ Minkowski metric tensor iff $g_{00} = 1$, $g_{ii} = -1$ for $i = 1, 2, 3$ and $g_{\mu\nu} = 0$ for $\mu \neq \nu$.*

Note that if we treat g as matrix, then

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{4.1.0.1}$$

Definition 4.1.0.2. *We will say that a tensor $\Lambda^{\mu}{}_{\nu}$ preserves Minkowski metric tensor iff*

$$g_{\mu\nu} = g_{\alpha\beta}\Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu} \tag{4.1.0.2}$$

Definition 4.1.0.3. *Two contravariant vectors A^μ and B^μ are said to be orthogonal iff*

$$g_{\nu\mu}A^{\mu}B^{\mu} = 0, \tag{4.1.0.3}$$

i.e.

$$A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 = 0. \tag{4.1.0.4}$$

Definition 4.1.0.4.

$$\mathcal{Q}(A^\mu) = g_{\mu\nu}A^{\mu}A^{\nu}. \tag{4.1.0.5}$$

Lemma 4.1.0.5. *If $\hat{x}^\mu$ and x^μ are two set of axes for which $\hat{x}^\mu = \Lambda^{\mu}{}_{\alpha}x^{\alpha}$, then*

$$\frac{\partial \hat{x}^\mu}{\partial x^\nu} = \Lambda^{\mu}{}_{\nu}. \tag{4.1.0.6}$$

It is good to visualize $\Lambda^\mu{}_\nu$ as matrix.

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{bmatrix}. \tag{4.1.0.7}$$

Proposition 4.1.0.6. *If $\hat{x}^\mu$ and x^μ are two set of axes for which $\hat{x}^\mu = \Lambda^\mu{}_\alpha x^\alpha$, the following conditions are equivalent:*

1. *$\Lambda^\mu{}_\nu$ preserves Minkowski metric tensor.*

2. *$g_{\mu\nu}x^\mu x^\nu = g_{\mu\nu}\hat{x}^\mu\hat{x}^\nu$ in each point.*

3. *Contravariants vectors $\Lambda^\mu{}_0, \Lambda^\mu{}_1, \Lambda^\mu{}_2, \Lambda^\mu{}_3$ are pairwise orthogonal and $\mathcal{Q}(\Lambda^\mu{}_0) = 1$ and $\mathcal{Q}(\Lambda^\mu{}_1) = \mathcal{Q}(\Lambda^\mu{}_2) = \mathcal{Q}(\Lambda^\mu{}_3) = -1$.*

Proposition 4.1.0.7. *If $\Lambda^\mu{}_\nu$ preserves Minkowski metric tensor, then*

$$\begin{bmatrix} \Lambda^0{}_0 & -\Lambda^1{}_0 & -\Lambda^2{}_0 & -\Lambda^3{}_0 \\ -\Lambda^0{}_1 & \Lambda^1{}_1 & \Lambda^2{}_1 & \Lambda^3{}_1 \\ -\Lambda^0{}_2 & \Lambda^1{}_2 & \Lambda^2{}_2 & \Lambda^3{}_2 \\ -\Lambda^0{}_3 & \Lambda^1{}_3 & \Lambda^2{}_3 & \Lambda^3{}_3 \end{bmatrix} \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{bmatrix} = I. \tag{4.1.0.8}$$

Definition 4.1.0.8.

$$\Lambda_\nu{}^\mu = \begin{bmatrix} \Lambda^0{}_0 & -\Lambda^1{}_0 & -\Lambda^2{}_0 & -\Lambda^3{}_0 \\ -\Lambda^0{}_1 & \Lambda^1{}_1 & \Lambda^2{}_1 & \Lambda^3{}_1 \\ -\Lambda^0{}_2 & \Lambda^1{}_2 & \Lambda^2{}_2 & \Lambda^3{}_2 \\ -\Lambda^0{}_3 & \Lambda^1{}_3 & \Lambda^2{}_3 & \Lambda^3{}_3 \end{bmatrix}.$$

Proposition 4.1.0.9. *If $\Lambda^\mu{}_\nu$ preserves Minkowski metric tensor and $\hat{x}^\mu$, x^μ are two set of axes for which $\hat{x}^\mu = \Lambda^\mu{}_\alpha x^\alpha$, then the following are true:*

1. *$\Lambda^\alpha{}_\mu \Lambda_\beta{}^\mu = \Lambda_\beta{}^\mu \Lambda^\alpha{}_\mu = \delta^\alpha_\beta$.*

2. *$x^\mu = \Lambda_\nu{}^\mu \hat{x}^\nu$.*

3. *$\Lambda_\nu{}^\mu$ preserves Minkowski metric tensor.*

Assume that $\hat{x}^\mu = \Lambda^\mu{}_\alpha x^\alpha$ such that $\hat{x}^2 = x^2$ and $\hat{x}^3 = x^3$. Then

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & 0 & 0 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.1.0.9}$$

Note that

$$\begin{cases} \hat{x}^0 = \Lambda^0{}_0 x^0 + \Lambda^0{}_1 x^1, \\ \hat{x}^1 = \Lambda^1{}_0 x^0 + \Lambda^1{}_1 x^1. \end{cases} \tag{4.1.0.10}$$

$$\begin{cases} x^0 = \Lambda^0{}_0 \hat{x}^0 - \Lambda^1{}_0 \hat{x}^1, \\ x^1 = -\Lambda^0{}_1 \hat{x}^0 + \Lambda^1{}_1 \hat{x}^1. \end{cases} \tag{4.1.0.11}$$

By Preposition 4.1.0.6 (c) applied to $\Lambda^\mu{}_\nu$, we have $(\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 = 1$, thus

$$\Lambda^1{}_0 = \pm((\Lambda^0{}_0)^2 - 1)^{1/2}. \tag{4.1.0.12}$$

It follows that $(\Lambda^0{}_0)^2 \geq 1$.

By Preposition 4.1.0.6 (c) applied to $\Lambda_\mu{}^\nu$, we have $(\Lambda^1{}_0)^2 - (\Lambda^1{}_1)^2 = -1$. Thus $(\Lambda^0{}_0)^2 - (\Lambda^1{}_1)^2 = 0$ and

$$\Lambda^1{}_1 = \pm\Lambda^0{}_0. \tag{4.1.0.13}$$

By Preposition 4.1.0.6 (c) column vectors from $\Lambda^\mu{}_\nu$ are orthogonal, thus

$$\Lambda^0{}_0 \Lambda^0{}_1 = \Lambda^1{}_0 \Lambda^1{}_1. \tag{4.1.0.14}$$

So, if $\Lambda^1{}_1 = \Lambda^0{}_0$, then

$$\Lambda^1{}_0 = \pm((\Lambda^0{}_0)^2 - 1)^{1/2}, \tag{4.1.0.15}$$

$$\Lambda^0{}_1 = \pm((\Lambda^0{}_0)^2 - 1)^{1/2}. \tag{4.1.0.16}$$

If $\Lambda^1{}_1 = -\Lambda^0{}_0$, then

$$\Lambda^1{}_0 = \pm((\Lambda^0{}_0)^2 - 1)^{1/2}, \tag{4.1.0.17}$$

$$\Lambda^0{}_1 = \mp((\Lambda^0{}_0)^2 - 1)^{1/2}. \tag{4.1.0.18}$$

Thus we have 2 possibilities:

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \Lambda^0{}_0 & \pm((\Lambda^0{}_0)^2 - 1)^{1/2} & 0 & 0 \\ \pm((\Lambda^0{}_0)^2 - 1)^{1/2} & \Lambda^0{}_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.1.0.19}$$

or

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \Lambda^0{}_0 & \mp((\Lambda^0{}_0)^2 - 1)^{1/2} & 0 & 0 \\ \pm((\Lambda^0{}_0)^2 - 1)^{1/2} & -\Lambda^0{}_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.1.0.20}$$

Note that in case (4.1.0.19), $\det(\Lambda^\mu{}_\nu) = 1$, while in case (4.1.0.20) $\det(\Lambda^\mu{}_\nu) = -1$.

Note that coordinates $(t, 0, 0, 0)$ in x^μ translates into $(\Lambda^0{}_0 t, \Lambda^1{}_0 t, 0, 0)$ in $\hat{x}^\mu$. On the other hand $(t, 0, 0, 0)$ in $\hat{x}^\mu$ translates into $(\Lambda^0{}_0 t, -\Lambda^0{}_1 t, 0, 0)$ in x^μ.

If we want to interpret transformation $\Lambda^\mu{}_\nu$ as a change of axes in a physical experiment, we need to impose additional conditions. We need to require:

$$\boxed{\Lambda^0{}_0 \geq 1}, \tag{4.1.0.21}$$

and Λ^1_0 needs to have the same sign as $\Lambda^0{}_1$, which is equivalent to

$$\boxed{\det(\Lambda^\mu{}_\nu) = 1}. \tag{4.1.0.22}$$

Let

$$\boxed{\beta = -\frac{\Lambda^0{}_1}{\Lambda^0{}_0} \text{ be a velocity of a particle } (\Lambda^0{}_0 t, -\Lambda^0{}_1 t, 0, 0) \text{ in } x^\mu},$$

which corresponds to the particle $(t, 0, 0, 0)$ in $\hat{x}^\mu$. Then

$$\beta^2 = \frac{(\Lambda^0{}_0)^2 - 1}{(\Lambda^0{}_0)^2}, \tag{4.1.0.23}$$

Thus

$$\Lambda^0{}_0 = (1 - \beta^2)^{-1/2}. \tag{4.1.0.24}$$

We will usually denote

$$\boxed{\gamma = (1 - \beta^2)^{-1/2}}.$$

We can finally write the $\Lambda^\mu{}_\nu$ in a form:

$$\boxed{\Lambda^\mu{}_\nu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}} \tag{4.1.0.25}$$

Thus once again we will write transformations explicitly:

$$\boxed{\begin{cases} \hat{x}^0 = \gamma x^0 - \beta\gamma x^1, \\ \hat{x}^1 = -\beta\gamma x^0 + \gamma x^1. \end{cases}} \tag{4.1.0.26}$$

$$\boxed{\begin{cases} x^0 = \gamma\hat{x}^0 + \beta\gamma\hat{x}^1, \\ x^1 = \beta\gamma\hat{x}^0 + \gamma\hat{x}^1. \end{cases}} \tag{4.1.0.27}$$

Let's rewrite (4.1.0.26) using t for time and x for one space dimension.

$$\begin{cases} \hat{t} = \gamma t - \beta\gamma x, \\ \hat{x} = x\gamma - \beta\gamma t. \end{cases} \tag{4.1.0.28}$$

Assume that $\beta \ll 1$, such that β^2 vanishes. It is easy to show that with this assumption $1 - \gamma$ also vanishes and thus we can write

$$\begin{cases} \hat{t} = t - \beta x, \\ \hat{x} = x - \beta t. \end{cases} \tag{4.1.0.29}$$

It is not yet clear how this relates to Galilean transformation, because of potentially non-vanishing shift in time.

We will show that any „slowly" moving particle will be described in the frame reference $\hat{x}^\mu$ according to Galilean transformation with an approximation to the first order.

Let's assume we have a „slowly" moving particle with equation of motion $t \to (t, \beta_0 t, 0, 0)$ in the frame x^μ where β_0^2 vanishes. Let's observe the trajectory of this particle in the frame $\hat{x}^\mu$:

$$\begin{cases} \hat{t} = t - \beta\beta_0 t, \\ \hat{x} = \beta_0 t - \beta t. \end{cases} \tag{4.1.0.30}$$

But because $\beta\beta_0$ vanishes we get

$$\begin{cases} \hat{t} = t, \\ \hat{x} = (\beta_0 - \beta)t. \end{cases} \tag{4.1.0.31}$$

Fact 4.1.0.10. *$g_{\mu\nu}x^\mu x^\nu < 0$ (x is space-like) if and only if there exists a restricted Lorentz transformation $u \mapsto \hat{u}$ such that $\hat{x}^0 = 0$.*

Proof. We can take Lorentz transformation, where $x \mapsto (x^0, z^1, 0, 0)$. Now, compose it with another Lorentz transformation

$$\begin{cases} \hat{x}^0 = \gamma x^0 - \beta\gamma z^1, \\ \hat{x}^1 = \gamma z^1 - \beta\gamma x^0. \end{cases} \tag{4.1.0.32}$$

Because we want $\hat{x}^0 = 0$, there must be $\beta = \dfrac{x^0}{z^1}$. Since $\beta \in (-1, 1)$, this is possible if and only if $(x^0)^2 - (z^1)^2 < 0$. But $(x^0)^2 - (z^1)^2 = g_{\mu\nu}x^\mu x^\nu$, because of Lorentz invariance. $\square$

Corollary 4.1.0.11. *$x-y$ is space-like if and only if there exists a restricted Lorentz transformation such that $\hat{x}^0 = \hat{y}^0$.*

The physical interpretation of the above is that for two events x and y which are space-like separated ($x-y$ is space-like), there exists and observer for whom these events occurred at the same time.

Fact 4.1.0.12. *$g_{\mu\nu}x^\mu x^\nu < 0$ (x is space-like) if and only if there exists a restricted Lorentz transformation $u \mapsto \hat{u}$ such that $\hat{x} = -x$.*

Proof. We can take Lorentz transformation, where $x \mapsto (x^0, z^1, 0, 0)$. Now, compose it with another Lorentz transformation

$$\begin{cases} \hat{x}^0 = \gamma x^0 - \beta\gamma z^1, \\ \hat{x}^1 = \gamma z^1 - \beta\gamma x^0. \end{cases} \tag{4.1.0.33}$$

To get $-x^0 = \gamma x^0 - \beta\gamma z^1$, we want

$$\frac{x^0}{z^1} = \frac{\beta\gamma}{1+\gamma}. \tag{4.1.0.34}$$

Consider function

$$f(\beta) = \frac{\beta\gamma}{1+\gamma} = \frac{\beta}{1+(1-\beta^2)^{1/2}} \tag{4.1.0.35}$$

Note that f is continuous and strictly increasing on $(-1,1)$ and $\lim\limits_{\beta\to-1} f(\beta) = -1$ and $\lim\limits_{\beta\to1} f(\beta) = 1$. Then there exists $\beta \in (-1,1)$ such that

$$f(\beta) = \frac{x^0}{z^1}, \tag{4.1.0.36}$$

if and only if $g_{\mu\nu}x^\mu x^\nu = (x^0)^2 - (z^1)^2 < 0$.

We can assume now that 4.1.0.34 holds, then

$$\hat{x}^1 = \gamma z^1 - \frac{\beta^2\gamma^2}{1+\gamma}z^1 = \frac{\gamma+\gamma^2-\beta^2\gamma^2}{1+\gamma}z^1 = \frac{\gamma+1}{1+\gamma}z^1 = z^1. \tag{4.1.0.37}$$

Thus, we have

$$\begin{cases} \hat{x}^0 = -x^0, \\ \hat{x}^1 = z^1. \end{cases} \tag{4.1.0.38}$$

Now, we compose two Lorentz transformations (which are just space rotations):

$$(-x^0, z^1, 0, 0) \mapsto (-x^0, -z^1, 0, 0) \mapsto -x. \tag{4.1.0.39}$$

□

Lemma 4.1.0.13. *$a_1, a_2 \in (-1,1)$ or $a_1, a_2 \notin (-1,1)$, if and only if*

$$\frac{a_1 + a_2}{1 + a_1 a_2} \in (-1,1). \tag{4.1.0.40}$$

Proof. Note that

$$\frac{a_1 + a_2}{1 + a_1 a_2} \in (-1,1). \tag{4.1.0.41}$$

if and only if

$$\left(\frac{a_1 + a_2}{1 + a_1 a_2}\right)^2 < 1. \tag{4.1.0.42}$$

Indeed, observe the equivalent inequalities:

$$\begin{gathered}
\frac{a_1^2 + 2a_1 a_2 + a_2^2}{1 + 2a_1 a_2 + a_1^2 a_2^2} < 1, \\
a_1^2 + \cancel{2a_1 a_2} + a_2^2 < 1 + \cancel{2a_1 a_2} + a_1^2 a_2^2, \\
0 < 1 - a_1^2 - a_2^2 + a_1^2 a_2^2, \\
0 < (1 - a_1^2)(1 - a_2^2).
\end{gathered}$$

It is apparent that the last inequality holds either if and only if $a_1, a_2 \in (-1,1)$ or $a_1, a_2 \notin (-1,1)$. $\square$

Lemma 4.1.0.14. *Let $\beta \in \mathbb{R}$ and $\gamma = (1-\beta^2)^{-1/2}$, $\alpha = \dfrac{x^1}{x^0}$ and $\hat{\alpha} = \dfrac{\hat{x}^1}{\hat{x}^0}$, where*

$$\begin{cases} \hat{x}^0 = \gamma x^0 - \beta\gamma x^1, \\ \hat{x}^1 = \gamma x^1 - \beta\gamma x^0, \end{cases} \tag{4.1.0.43}$$

then

$$\beta = \frac{\hat{\alpha} - \alpha}{\alpha\hat{\alpha} - 1}. \tag{4.1.0.44}$$

Lemma 4.1.0.15. *If $(x^0)^2 - (x^1)^2 = (\hat{x}^0)^2 - (\hat{x}^1)^2 > 0$ and $x^0 > 0$ and $\hat{x}^0 > 0$, then for*

$$\beta = \frac{\hat{\alpha} - \alpha}{\alpha\hat{\alpha} - 1}, \tag{4.1.0.45}$$

where $\alpha = \dfrac{x^1}{x^0}$ and $\hat{\alpha} = \dfrac{\hat{x}^1}{\hat{x}^0}$, we have $\beta \in (-1,1)$ and there exists a restricted Lorentz transformation

$$\begin{cases} \hat{x}^0 = \gamma x^0 - \beta\gamma x^1, \\ \hat{x}^1 = \gamma x^1 - \beta\gamma x^0, \end{cases} \tag{4.1.0.46}$$

where $\gamma = (1-\beta^2)^{-1/2}$.

Proof. Note that

$$(x^0)^2(1-\alpha^2) = (\hat{x}^0)^2(1-\hat{\alpha}^2).$$

Thus

$$\left(\frac{\hat{x}^0}{x^0}\right)^2 = \frac{1-\alpha^2}{1-\hat{\alpha}^2} > 0. \tag{4.1.0.47}$$

Therefore either $\alpha, \hat{\alpha} \in (-1,1)$ or $\alpha, \hat{\alpha} \notin (-1,1)$, thus by Lemma 4.1.0.13 $\beta \in (-1,1)$.

Let's calculate

$$\begin{aligned} 1-\beta^2 &= \frac{\alpha^2\hat{\alpha}^2 - 2\alpha\hat{\alpha} + 1 - \alpha^2 - \hat{\alpha}^2 + 2\alpha\hat{\alpha}}{\alpha^2\hat{\alpha}^2 - 2\alpha\hat{\alpha} + 1} \\ &= \frac{(1-\alpha^2)(1-\hat{\alpha}^2)}{(\alpha\hat{\alpha}-1)^2}. \end{aligned} \tag{4.1.0.48}$$

Since $(x^0)^2 - (x^1)^2 = (\hat{x}^0)^2 - (\hat{x}^1)^2 > 0$, then $\alpha, \hat{\alpha} \in (-1,1)$, and so

$$\gamma = \frac{1-\alpha\hat{\alpha}}{\left((1-\alpha^2)(1-\hat{\alpha}^2)\right)^{1/2}}. \tag{4.1.0.49}$$

Let's verify the first equation of 4.1.0.46.

$$\begin{aligned} &\gamma x^0 - \beta\gamma x^1 = \gamma x^0(1-\beta\alpha) = \\ &= x^0\gamma\frac{\alpha\hat{\alpha} - 1 - \alpha\hat{\alpha} + \alpha^2}{\alpha\hat{\alpha}-1} = x^0\gamma\frac{\alpha^2-1}{\alpha\hat{\alpha}-1} \\ &= x^0\left(\frac{1-\alpha^2}{1-\hat{\alpha}^2}\right)^{1/2} = x^0\frac{\hat{x}^0}{x^0} = \hat{x}^0. \end{aligned} \tag{4.1.0.50}$$

Again, the second from the end equality holds since $x^0 > 0$ and $\hat{x}^0 > 0$. Let's verify the second equation of 4.1.0.46.

$$\begin{aligned}\gamma x^1 - \beta\gamma x^0 &= x^0\gamma(\alpha-\beta) = x^0\gamma\frac{\alpha^2\hat{\alpha} - \alpha - \hat{\alpha} + \alpha}{\alpha\hat{\alpha} - 1} \\ &= x^0\hat{\alpha}\gamma\frac{\alpha^2 - 1}{\alpha\hat{\alpha} - 1} = x^0\hat{\alpha}\left(\frac{1-\alpha^2}{1-\hat{\alpha}^2}\right)^{1/2} = x^0\frac{\hat{x}^1}{\hat{x}^0}\frac{\hat{x}^0}{x^0} = \hat{x}^1.\end{aligned} \tag{4.1.0.51}$$

The second from the end equality holds since $x^0 > 0$ and $\hat{x}^0 > 0$. □

Lemma 4.1.0.16. *If $(x^0)^2 - (x^1)^2 = (\hat{x}^0)^2 - (\hat{x}^1)^2 < 0$ and $x^0, x^1 > 0$ and $\hat{x}^0, \hat{x}^1 > 0$, then for*

$$\beta = \frac{\hat{\alpha} - \alpha}{\alpha\hat{\alpha} - 1}, \tag{4.1.0.52}$$

where $\alpha = \frac{x^1}{x^0}$ and $\hat{\alpha} = \frac{\hat{x}^1}{\hat{x}^0}$, we have $\beta \in (-1,1)$ and there exists a restricted Lorentz transformation

$$\begin{cases}\hat{x}^0 = \gamma x^0 - \beta\gamma x^1, \\ \hat{x}^1 = \gamma x^1 - \beta\gamma x^0,\end{cases} \tag{4.1.0.53}$$

where $\gamma = (1-\beta^2)^{-1/2}$.

Proof. Note that (4.1.0.47) and (4.1.0.48) holds. By that obviously $\beta \in (-1,1)$.

Since $(x^0)^2 - (x^1)^2 = (\hat{x}^0)^2 - (\hat{x}^1)^2 < 0$, then $\alpha, \hat{\alpha} \notin (-1,1)$. But since $x^0, x^1 > 0$ and $\hat{x}^0, \hat{x}^1 > 0$, we have $\alpha, \hat{\alpha} > 0$, then $\alpha, \hat{\alpha} > 1$. Therefore, as (4.1.0.48) stays the same

$$\gamma = \frac{\alpha\hat{\alpha} - 1}{\big((\alpha^2-1)(\hat{\alpha}^2-1)\big)^{1/2}}. \tag{4.1.0.54}$$

One can check that for the above γ (4.1.0.50) and (4.1.0.51) also holds and this proves (4.1.0.53). □

Fact 4.1.0.17. *If $g_{\mu\nu}y^\mu y^\nu = g_{\mu\nu}x^\mu x^\nu > 0$ (x, y are the same length and time-like) such that $x^0 > 0$ and $y^0 > 0$, then there exists a restricted Lorentz transformation $u \mapsto \hat{u}$ such that $\hat{x} = y$.*

Proof. Let's start with rotations which are restricted Lorentz transformations

$$(x^0, x^1, x^2, x^3) \mapsto (x^0, z^1, 0, 0). \tag{4.1.0.55}$$

$$(y^0, y^1, y^2, y^3) \mapsto (y^0, v^1, 0, 0). \tag{4.1.0.56}$$

Since they are Lorentz transformations $(x^0)^2 - (z^1)^2 = (y^0)^2 - (v^1)^2 > 0$, by Lemma 4.1.0.15 there exists a restricted Lorentz transformation

$$(x^0, z^1, 0, 0) \mapsto (y^0, v^1, 0, 0). \tag{4.1.0.57}$$

□

Fact 4.1.0.18. *If $g_{\mu\nu}y^\mu y^\nu = g_{\mu\nu}x^\mu x^\nu < 0$ (x, y are space-like) such that, then there exists a restricted Lorentz transformation $u \mapsto \hat{u}$ such that $\hat{x} = y$.*

Proof. By combining Fact 4.1.0.12 and the fact that each space rotation is a restricted Lorentz transformation, we can have a restricted Lorentz transformations:

$$(x^0, x^1, x^2, x^3) \mapsto (|x^0|, |z^1|, 0, 0). \tag{4.1.0.58}$$

$$(y^0, y^1, y^2, y^3) \mapsto (|y^0|, |v^1|, 0, 0). \tag{4.1.0.59}$$

Since they are Lorentz transformations, we have $(x^0)^2 - (z^1)^2 = (y^0)^2 - (v^1)^2 < 0$ and thus by Lemma 4.1.0.16, there exists a restricted Lorentz transformation

$$(|x^0|, |z^1|, 0, 0) \mapsto (|y^0|, |v^1|, 0, 0). \tag{4.1.0.60}$$

□

4.2 Lorentz group

Let us remind that $x_\mu = g_{\mu\nu}x^\nu$, hence in Special Relativity we have

$$x_0 = x^0, x_1 = -x^1, x_2 = -x^2, x_3 = -x^3. \tag{4.2.0.1}$$

Note that with $g^{00} = 1$ and $g^{ii} = -1$ for $i = 1, 2, 3$ and $g^{\mu\nu} = 0$ for $\mu \neq \nu$, we have $x^\mu = g^{\mu\nu}x_\nu$.

4.2.1 Rotation in ij-plane

Consider an infinitesimal rotation in direction from an arrow of x^i to an arrow of x^j.

$$\begin{cases} \hat{x}^i = x^i - \varepsilon x^j, \\ \hat{x}^j = x^j + \varepsilon x^i, \\ \hat{x}^k = x^k \text{ where } k \notin \{i, j\}. \end{cases} \tag{4.2.1.1}$$

where ε^2 vanishes.

Let's show that the above transformation preserves Minkowski metric tensor.

$$\begin{aligned} g_{\mu\nu}\hat{x}^\mu\hat{x}^\nu - g_{\mu\nu}x^\mu x^\nu &= -(x^i - \varepsilon x^j)^2 - (x^j + \varepsilon x^i)^2 + (x^i)^2 + (x^j)^2 \\ &= -\varepsilon^2((x^i)^2 + (x^j)^2) \end{aligned}$$

The last term vanishes, so the transformation above preserves Minkowski metric tensor to the first order. Let's define

$$\Delta R_{ij}(x^\mu) = \hat{x}^\mu. \tag{4.2.1.2}$$

$$(G_{ij})^k_l := \begin{cases} -1 \text{ for } k = i, l = j, \\ 1 \text{ for } k = j, l = i, \\ 0 \text{ otherwise.} \end{cases} \tag{4.2.1.3}$$

We may picture G_{ij} as

$$G_{ij} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{4.2.1.4}$$

If we restrict it to i-th and j-th rows and i-th and j-th columns and remember that there is 0s everywhere else.

Then $\Delta R_{ij} = I + \varepsilon G_{ij}$. Now we can describe rotation of angle θ in direction from an arrow of x^i to an arrow of x^j as

$$R_{ij}(\theta) = \lim_{n\to\infty} \left(I + \frac{\theta G_{ij}}{n} \right)^n = \exp(\theta G_{ij}). \tag{4.2.1.5}$$

It can be also show that

$$R_{ij}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4.2.1.6}$$

If we restrict it to i-th and j-th rows and i-th and j-th columns and remember that there are 1s on the remaining part of diagonal and 0s everywhere else.

It is easy to notice that

$$\left.\frac{\partial R_{ij}(\theta)}{\partial\theta}\right|_{\theta=0} = G_{ij}. \tag{4.2.1.7}$$

which nicely confirms that G_{ij} is a generator of a subgroup $R_{ij}(\theta)$.

Consider a transformation $f \mapsto f \circ R_{ij}(-\theta)$. Note that this transformation rotates the graph of function against its domain in the same direction as $R_{ij}(\theta)$. Its generator is then $f \mapsto -(\nabla f) \cdot G_{ij}$, which we can write in a differential form as

$$x^j \frac{\partial}{\partial x^i} - x^i \frac{\partial}{\partial x^j}. \tag{4.2.1.8}$$

By convention we introduce a generator in an quantum-mechanical form (i.e. multiplying the above by i):

$$\mathcal{J}_{nm} = i(x_n \frac{\partial}{\partial x^m} - x_m \frac{\partial}{\partial x^n}) \tag{4.2.1.9}$$

Note that we lowered indices and this changed sign.

4.2.2 Boost in $0j$-plane

Consider an infinitesimal boost in direction from an arrow of x^0 to an arrow of x^j.

$$\begin{cases} \hat{x}^0 = x^0 - \varepsilon x^j, \\ \hat{x}^j = x^j - \varepsilon x^0, \\ \hat{x}^k = x^k \text{ where } k \notin \{0, j\}. \end{cases} \tag{4.2.2.1}$$

where ε^2 vanishes. Note this is approximated Lorentz transformation for a frame of reference moving with a small velocity ε towards and arrow of x^j as in (4.1.0.29).

Let's show that the above transformation preserves Minkowski metric tensor.

$$\begin{aligned} g_{\mu\nu}\hat{x}^\mu\hat{x}^\nu - g_{\mu\nu}x^\mu x^\nu &= (x^0 - \varepsilon x^j)^2 - (x^j - \varepsilon x^0)^2 - (x^0)^2 + (x^j)^2 \\ &= \varepsilon^2((x^0)^2 - (x^j)^2) \end{aligned}$$

The last term vanishes, so the transformation above preserves Minkowski metric tensor to the first order. Let's define

$$\Delta B_j(x^\mu) = \hat{x}^\mu. \tag{4.2.2.2}$$

$$(G_{ij})^k_l := \begin{cases} -1 \text{ for } k = 0, l = j, \\ -1 \text{ for } k = j, l = 0, \\ 0 \text{ otherwise.} \end{cases} \tag{4.2.2.3}$$

We may picture G_j as

$$G_j = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \tag{4.2.2.4}$$

If we restrict it to 0-th and j-th rows and 0-th and j-th columns and remember that there is 0s everywhere else.

Then $\Delta B_j = I + \varepsilon G_j$. Now we can describe rotation of angle θ in direction from an arrow of x^i to an arrow of x^j as

$$B_j(\omega) = \lim_{n\to\infty} \left(I + \frac{\omega G_{ij}}{n} \right)^n = \exp(\omega G_{ij}). \tag{4.2.2.5}$$

It can be also show that

$$B_j(\omega) = \begin{bmatrix} \cosh\omega & -\sinh\omega \\ -\sinh\omega & \cosh\omega \end{bmatrix} \tag{4.2.2.6}$$

If we restrict it to 0-th and j-th rows and 0-th and j-th columns and remember that there are 1s on the remaining part of diagonal and 0s everywhere else.

Let's put $\beta = \tanh\omega$. Let $\gamma = (1-\beta^2)^{-1/2}$. It is easy to calculate then $\gamma = \cosh\omega$. Then

$$B_j(\omega) = \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix}, \tag{4.2.2.7}$$

which is exactly Lorentz transformation for a frame moving with a velocity β along j-th ax towards an arrow as in (4.1.0.25).

It is easy to notice that

$$\left.\frac{\partial B_j(\omega)}{\partial\omega}\right|_{\omega=0} = G_j. \tag{4.2.2.8}$$

which nicely confirms that G_j is a generator of a subgroup $B_j(\theta)$.

Consider a transformation $f \mapsto f \circ (B_j(-\theta))$. Its generator is then $f \mapsto -(\nabla f) \cdot G_j$, which we can write in a differential form as

$$x^0 \frac{\partial}{\partial x^j} + x^j \frac{\partial}{\partial x^0}. \quad (4.2.2.9)$$

By convention we introduce a generator in an quantum-mechanical form (i.e. multiplying the above by i):

$$\mathcal{J}_{0j} = i(x_0 \frac{\partial}{\partial x^j} - x_j \frac{\partial}{\partial x^0}) \quad (4.2.2.10)$$

Note that we lowered indices and this changed the sign in front of x_j.

4.2.3 Lorentz Group Representation

It is well known fact that each Lorentz transformation Λ can be represented as

$$\Lambda = B_1(\omega_1)B_2(\omega_2)B_3(\omega_3)R_{23}(\theta_1)R_{31}(\theta_2)R_{12}(\theta_3). \quad (4.2.3.1)$$

4.3 Dynamics

Consider particle of mass m described in the frame reference x^μ by equation of motion

$$t \mapsto (x^0 = t, x^1(t), x^2(t), x^3(t)). \quad (4.3.0.1)$$

Let u^μ be a 4-velocity of the particle along the path $t \mapsto x^\mu(t)$. $p^\mu = mu^\mu$ is 4-momentum. Let

$$v^i := \frac{dx^i}{dt}, \quad (4.3.0.2)$$

$$\vec{v} := [v^1, v^2, v^3], \quad (4.3.0.3)$$

$$\vec{u} := [u^1, u^2, u^3], \quad (4.3.0.4)$$

$$\vec{p} = [p^1, p^2, p^3]. \quad (4.3.0.5)$$

Note that $\vec{p} := m\vec{u}$ is a Newtonian momentum measured in the frame of reference x^μ. In this section $\cdot$ will be standard inner product in $\mathbb{R}^3$. Let

$$\vec{F} := \frac{d\vec{p}}{dt}. \quad (4.3.0.6)$$

Note that $\vec{F}$ is a Newtonian force acting on the particle measured in the frame of reference x^μ. Let

$$\beta := (-g_{nm}v^n v^m)^{\frac{1}{2}} = ((v^1)^2 + (v^2)^2 + (v^3)^2)^{\frac{1}{2}} = |\vec{v}|, \tag{4.3.0.7}$$

and

$$\alpha := (1 - \beta^2)^{\frac{1}{2}}. \tag{4.3.0.8}$$

Note that

$$u^\mu = \frac{dx^\mu}{d\tau}, \tag{4.3.0.9}$$

where

$$d\tau = (g_{\mu\nu}dx^\mu dx^\nu)^{\frac{1}{2}}. \tag{4.3.0.10}$$

Note that

$$(u_0)^2 - \vec{u} \cdot \vec{u} = g_{\mu\nu}u^\mu u^\nu = 1. \tag{4.3.0.11}$$

Now

$$\frac{d\tau}{dt} = \left(g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\right)^{\frac{1}{2}} = (1 - \beta^2)^{\frac{1}{2}} = \alpha. \tag{4.3.0.12}$$

$$\boxed{d\tau = \alpha dt.} \tag{4.3.0.13}$$

Note that

$$u^0 = \alpha^{-1}, u^i = \alpha^{-1}v^i. \tag{4.3.0.14}$$

Let's find the acceleration $\vec{a}$ of the particle in the frame of reference x^μ. Obviously

$$\vec{a} = \left[\frac{d^2x^1}{dt^2}, \frac{d^2x^2}{dt^2}, \frac{d^3x^i}{dt^2}\right] = \left[\frac{dv^1}{dt}, \frac{dv^2}{dt}, \frac{dv^3}{dt}\right] \tag{4.3.0.15}$$

$$\frac{d^2x^i}{dt^2} = \frac{d}{dt}\left(\frac{dx^i}{dt}\right) = \frac{d}{dt}\left(\frac{dx^i}{d\tau}\frac{d\tau}{dt}\right) = \frac{d}{dt}(u^i\alpha) = \frac{du^i}{dt}\alpha + u^i\frac{d\alpha}{dt}. \tag{4.3.0.16}$$

By (4.3.0.14) and (4.3.0.11),

$$\frac{d\alpha}{dt} = \frac{d}{dt}(u^0)^{-1} = \frac{d}{dt}(1 + \vec{u} \cdot \vec{u})^{-\frac{1}{2}} = -\frac{1}{2}(2\vec{u} \cdot \frac{d\vec{u}}{dt})(1 + \vec{u} \cdot \vec{u})^{-\frac{3}{2}} = -\alpha^3\vec{u} \cdot \frac{d\vec{u}}{dt}. \tag{4.3.0.17}$$

Thus

$$\frac{d\alpha}{dt} = -\alpha^2\vec{v} \cdot \frac{d\vec{u}}{dt} = -\frac{\alpha^2}{m}\vec{v} \cdot \frac{d\vec{p}}{dt} = -\frac{\alpha^2}{m}\vec{v} \cdot \vec{F}. \tag{4.3.0.18}$$

By (4.3.0.16) we have

$$\frac{d^2x^i}{dt^2} = \frac{\alpha}{m}\frac{dp^i}{dt} - u^i\frac{\alpha^2}{m}\vec{v}\cdot\vec{F} = \frac{\alpha}{m}\Big(F^i - (\vec{v}\cdot\vec{F})v^i\Big). \tag{4.3.0.19}$$

Thus

$$\boxed{\vec{a} = \frac{\alpha}{m}\Big(\vec{F} - (\vec{v}\cdot\vec{F})\vec{v}\Big)} \tag{4.3.0.20}$$

It might be useful to see how $\vec{a}$ depends on 4-momentum and 4-force.

$$\vec{a} = \frac{\alpha^2}{m}\frac{d\vec{p}}{d\tau} - \frac{\alpha^4}{m^3}(\vec{p}\cdot\frac{d\vec{p}}{d\tau})\vec{p}. \tag{4.3.0.21}$$

and on 4-velocity and 4-acceleration

$$\vec{a} = \alpha^2\frac{d\vec{u}}{d\tau} - \alpha^4(\vec{u}\cdot\frac{d\vec{u}}{d\tau})\vec{u}. \tag{4.3.0.22}$$

It is also interesting to see $\frac{d\alpha}{dt}$ from a bit different perspective.

$$\frac{d\alpha}{dt} = \frac{d}{dt}(1-\beta^2)^{\frac{1}{2}} = -\alpha^{-1}\beta\frac{d\beta}{dt}. \tag{4.3.0.23}$$

Note that

$$\frac{d\beta}{dt} = \beta^{-1}\vec{v}\cdot\frac{d\vec{v}}{dt}. \tag{4.3.0.24}$$

Thus

$$\frac{d\alpha}{dt} = -\alpha^{-1}\vec{v}\cdot\frac{d\vec{v}}{dt}. \tag{4.3.0.25}$$

4.4 Symmetries - Noether's Theorem for Relativistic Fields

The approach in this chapter is inspired by [**?**][8.7]

Consider a generic continuous symmetry, which is very close to identity mapping.

$$x^\mu \mapsto x^\mu + \delta x^\mu. \tag{4.4.0.1}$$

We can describe δx^μ e.g. using generators.

$$\delta x^\mu = \delta\varepsilon G^\mu_\nu(x^\nu + a^\nu). \tag{4.4.0.2}$$

Where $\delta\varepsilon$ is an infinitesimal variation around 0, G^μ_ν is rotation/boost generator and a^μ is a constant shift vector.

For a given field ϕ^α we will usually consider an infinitesimal variation of the field:

$$(\delta\phi^\alpha)(x) \leftarrowtail \phi'^\alpha(x) - \phi^\alpha(x). \tag{4.4.0.3}$$

Note that for this kind of variation we have $\partial_\mu\delta = \delta\partial_\mu$.

Assume that we have a functional $\mathcal{L} = \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)$ which depends on all dimensions of field ϕ^α, all their derivatives and on position in space-time.

For an infinitesimal variation $\delta\phi^\alpha$ we define

$$\delta\mathcal{L} = \mathcal{L}(\phi^\alpha + \delta\phi^\alpha, \partial_\mu(\phi^\alpha + \delta\phi^\alpha), x^\mu) - \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu). \tag{4.4.0.4}$$

Let's introduce a symbol

$$\Pi^\mu_\alpha := \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}. \tag{4.4.0.5}$$

Lemma 4.4.0.1.

$$\delta\mathcal{L} = \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} - \partial_\mu\Pi^\mu_\alpha\right)\delta\phi^\alpha + \partial_\mu(\Pi^\mu_\alpha\delta\phi^\alpha). \tag{4.4.0.6}$$

Proof.

$$\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\delta(\partial_\mu\phi^\alpha)\\
&= \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha + \Pi^\mu_\alpha\delta(\partial_\mu\phi^\alpha)\\
&= \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha + \Pi^\mu_\alpha\partial_\mu\delta\phi^\alpha = \frac{\partial\mathcal{L}}{\partial\phi^\alpha}\delta\phi^\alpha + \partial_\mu(\Pi^\mu_\alpha\delta\phi^\alpha) - \partial_\mu(\Pi^\mu_\alpha)\delta\phi^\alpha\\
&= \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} - \partial_\mu\Pi^\mu_\alpha\right)\delta\phi^\alpha + \partial_\mu(\Pi^\mu_\alpha\delta\phi^\alpha).
\end{aligned}$$

$\square$

Let's define action functional which will act on a given field

$$S[\phi^\alpha, M] = \int_M \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)d^4x. \tag{4.4.0.7}$$

We will concentrate on finding ϕ_0^α which locally extremises $S[\phi^\alpha, M]$ (i.e. $\delta S[\phi_0^\alpha, M] = 0$) for a neighbourhood of fields $\phi^\alpha = \phi_0^\alpha + \delta\phi^\alpha$ where $\delta\phi^\alpha$ and $\partial_\mu \delta\phi^\alpha$ vanishes on ∂M. We will sometimes call such a field ϕ_0^α a stationary point of $S[\cdot, M]$.

Based on [see **?** , Appendix B] we can formulate the following:

Theorem 4.4.0.2. *(Divergence Theorem) Let M be a n-dimentional compact manifold with some metric tensor $g_{\mu\nu}$. Integrals below are over volume element induced by g. Let ∂M be a $(n-1)$-dimentional manifold, which is a boundary of M. Let n^μ be a continous vector field of unit normal vectors which are "pointing outward" if n^μ is spacelike ($g_{\mu\nu}n^\mu n^\nu < 0$) and "pointing inward" if n^μ is timelike ($g_{\mu\nu}n^\mu n^\nu > 0$). If v^μ is C^1 vector field on M, then*

$$\int_M \nabla_\mu v^\mu = \int_{\partial M} n_\mu v^\mu. \tag{4.4.0.8}$$

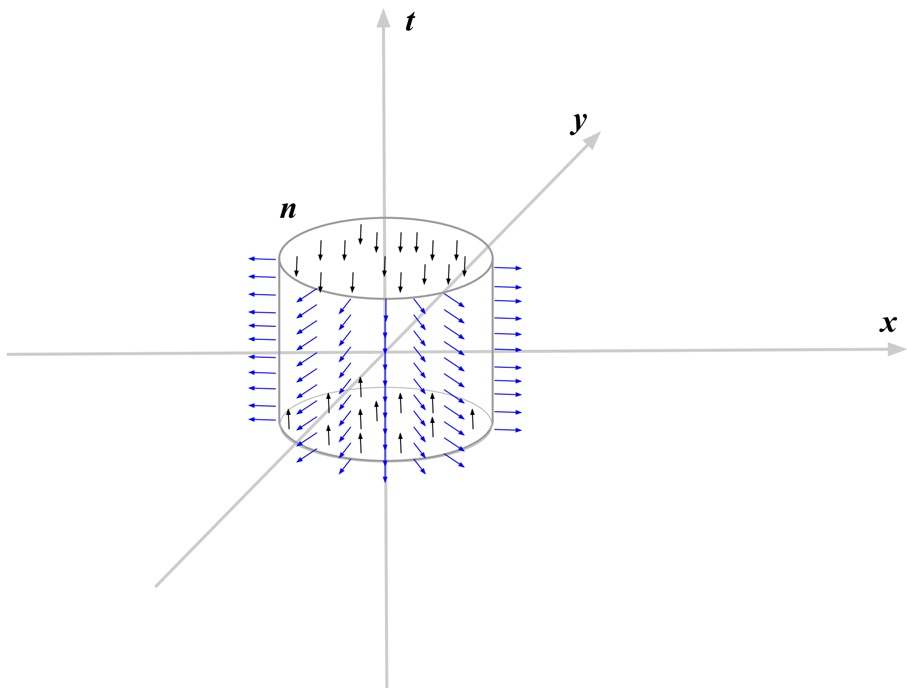

Figure 4.1: Normal vectors to the oriented surface in spacetime. Blue arrows indicate spacelike vectors. Black arrows indicate timelike vectors.

In the context of Special Theory of Relativity the equation 4.4.0.8 becomes

$$\int_M \partial_\mu v^\mu d^4x = \int_{\partial M} v^\mu dn_\mu. \tag{4.4.0.9}$$

Theorem 4.4.0.3. *(Relativistic Field Euler-Lagrange equations) If*

$$\delta(S[\phi_0^\alpha, M]) = 0$$

for variation δ with all partial derivatives $\partial_\mu \delta\phi^\alpha$ and $\partial_\nu \partial_\mu \delta\phi^\alpha$ infinitesimal of the first order and $\delta\phi^\alpha$ vanishes on ∂M, then

$$\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \partial_\mu \Pi^\mu_\alpha = 0 \tag{4.4.0.10}$$

for ϕ_0^α on the interior of M.

Proof. Let $S = S[\phi_0^\alpha, M]$, then $\delta S = \int_M \delta\mathcal{L} d^4x$. From (4.4.0.1) it follows that

$$\delta S = \int_M \left(\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \partial_\mu \Pi^\mu_\alpha\right) \delta\phi^\alpha d^4x + \int_M \partial_\mu (\Pi^\mu_\alpha \delta\phi^\alpha) d^4x. \tag{4.4.0.11}$$

By Divergence Theorem (4.4.0.9), we have

$$\int_M \partial_\mu (\Pi^\mu_\alpha \delta\phi^\alpha) d^4x = \int_{\partial M} \Pi^\mu_\alpha \delta\phi^\alpha dn_\mu = 0, \tag{4.4.0.12}$$

because $\delta\phi^\alpha$ vanishes on ∂M. Then

$$\delta S = \int_M \left(\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \partial_\mu \Pi^\mu_\alpha\right) \delta\phi^\alpha d^4x. \tag{4.4.0.13}$$

Since $\delta S = 0$ and $\delta\phi^\alpha$ is arbitrary enough, we have

$$\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \partial_\mu \Pi^\mu_\alpha = 0. \tag{4.4.0.14}$$

□

Consider an infinitesimal continuous symmetry $x \mapsto x + \hat{\delta}x$. From now on in this section, we will use $\hat{\delta}$ to denote an infinitesimal continuous symmetry variation. Note that in this section, we don't apply $\hat{\delta}$ as transformation of coordinates, we will use it only to shift the field relative to the frame of reference. We work all the time in an arbitrarily fixed frame of reference x^μ and we don't change it through out the whole chapter. This is why we don't need to transform Lagrangian density $\mathcal{L}$ and we use it all the time as exactly the same expression.

Let $\hat{\phi}^\alpha$ denote a field which was shifted in the direction $-\hat{\delta}x$ namely $\hat{\phi}^\alpha(x) = \phi^\alpha(x + \hat{\delta}x)$. Note that up to first order

$$\hat{\phi}^\alpha = \phi^\alpha + \partial_\mu \phi^\alpha \hat{\delta}x^\mu, \tag{4.4.0.15}$$

where $\hat{\delta}x^\mu = \hat{\delta}x^\mu(x)$ is a function of x as in example (4.4.0.3). We will use variation of ϕ^α defined as $\hat{\delta}\phi^\alpha = \partial_\mu \phi^\alpha \hat{\delta}x^\mu$. Note that this variation moves the field in direction $-\hat{\delta}x$.

By $\hat{A}$ we will denote a set shifted in the direction $-\hat{\delta}x$.

$$\hat{A} = \{x - \hat{\delta}x : x \in A\}. \tag{4.4.0.16}$$

Lemma 4.4.0.4. *If F is an arbitrary continuous function on space-time then*

$$\int (\mathbb{1}_{\hat{A}} - \mathbb{1}_A) F d^4x = \int_{\partial A} F \hat{\delta}x^\mu dn_\mu. \tag{4.4.0.17}$$

n_μ is understood as in (4.4.0.9).

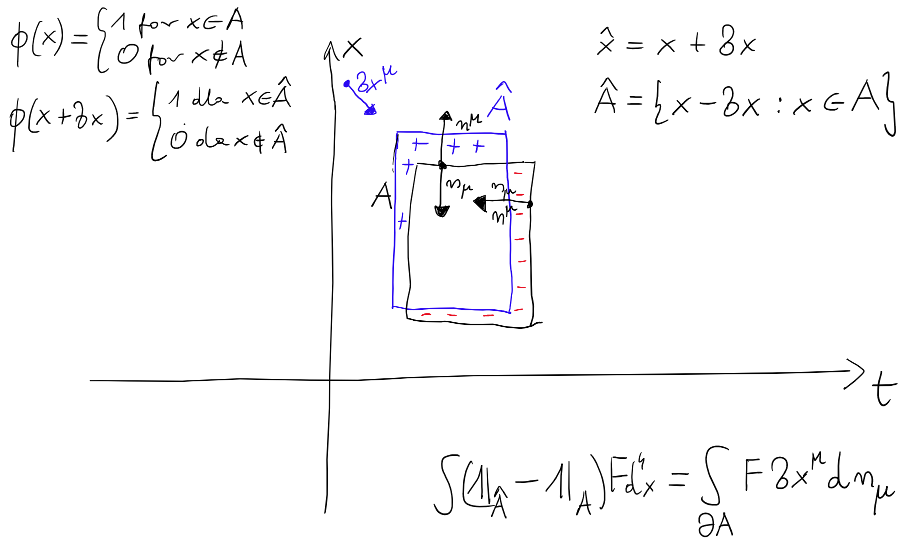

Figure 4.2: Illustration to visualise a proof of Lemma 4.4.0.4

Note that

$$\hat{\delta}S[\phi^\alpha, M] = S[\hat{\phi}^\alpha, \hat{M}] - S[\phi^\alpha, M]. \tag{4.4.0.18}$$

$\hat{\delta}S[\phi^\alpha, M]$ denotes difference in action when we move the field ϕ^α by an infinitesimal continuous symmetry $-\hat{\delta}x$.

Lemma 4.4.0.5. *If a field ϕ^α is a stationary point of $S[\cdot, M]$, then*

$$\hat{\delta}S[\phi^\alpha, M] = \int\limits_M \partial_\mu \left(\Pi^\mu_\alpha \hat{\delta}\phi^\alpha + \hat{\delta}x^\mu \mathcal{L} \right) d^4x. \tag{4.4.0.19}$$

Proof. To make expressions shorter we will omit (x) from $\phi^\alpha(x)$ and $\hat{\phi}^\alpha(x)$ when variable x is obvious and we will write just $\mathcal{L}$ when applied to $(\phi^\alpha(x), \partial_\mu\phi^\alpha(x), x^\mu)$.

$$\begin{aligned}
\hat{\delta}S[\phi^\alpha, M] &= \int\limits_{\hat{M}} \mathcal{L}(\hat{\phi}^\alpha(x), \partial_\mu\hat{\phi}^\alpha(x), x^\mu) d^4x - \int\limits_M \mathcal{L}(\phi^\alpha(x), \partial_\mu\phi^\alpha(x), x^\mu) d^4x \\
&= \int (\mathbb{1}_M + \mathbb{1}_{\hat{M}} - \mathbb{1}_M)\mathcal{L}(\hat{\phi}^\alpha, \partial_\mu\hat{\phi}^\alpha, x^\mu) d^4x - \int\limits_M \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu) d^4x \\
&= \int\limits_M \mathcal{L}(\hat{\phi}^\alpha(x), \partial_\mu\hat{\phi}^\alpha(x), x^\mu) d^4x - \int\limits_M \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu) d^4x + \int (\mathbb{1}_{\hat{M}} - \mathbb{1}_M)\mathcal{L}(\hat{\phi}^\alpha, \partial_\mu\hat{\phi}^\alpha, x^\mu) d^4x \\
&= \int\limits_M \mathcal{L}(\hat{\phi}^\alpha(x), \partial_\mu\hat{\phi}^\alpha(x), x^\mu) - \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu) d^4x + \int (\mathbb{1}_{\hat{M}} - \mathbb{1}_M)(\mathcal{L} + \hat{\delta}\mathcal{L}) d^4x \\
&= \int\limits_M \left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha} - \partial_\mu\Pi^\mu_\alpha \right) \hat{\delta}\phi^\alpha + \partial_\mu(\Pi^\mu_\alpha\hat{\delta}\phi^\alpha) d^4x + \int\limits_{\partial M} (\mathcal{L} + \hat{\delta}\mathcal{L})\hat{\delta}x^\mu dn_\mu \\
&= \int\limits_M \partial_\mu(\Pi^\mu_\alpha\hat{\delta}\phi^\alpha) d^4x + \int\limits_{\partial M} \mathcal{L}\hat{\delta}x^\mu dn_\mu = \int\limits_M \partial_\mu(\Pi^\mu_\alpha\hat{\delta}\phi^\alpha) d^4x + \int\limits_{\partial M} \partial_\mu(\mathcal{L}\hat{\delta}x^\mu) d^4x \\
&= \int\limits_M \partial_\mu \left(\Pi^\mu_\alpha \hat{\delta}\phi^\alpha + \hat{\delta}x^\mu \mathcal{L} \right) d^4x.
\end{aligned}$$

In transition from line 4 to 5 we used Lemma 4.4.0.1 and Lemma 4.4.0.4 correspondingly. Calculations are done up to first order.

□

For $\mathcal{L} = \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)$ to have physical meaning one might require that if a field ϕ^α is a stationary point of $S[\cdot, M]$, then the same field but shifted in space-time by an infinitesimal continuous symmetry $-\hat{\delta}x$ denoted as $\hat{\phi}^\alpha$ should be also a stationary point of $S[\cdot, \hat{M}]$. In other words, orientation of the field against arbitrary fixed frame of reference x^μ shouldn't matter. This is an assumption of isotropy and homogeneity of space.

In the next Lemma we will show that if $\hat{\delta}\mathcal{L} = \hat{\delta}\varepsilon\partial_\mu f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x)$ then $\mathcal{L}$ has these symmetry invariant stationary points in the above sense. In this context ∂_μ is a full derivative with respect to x^μ. Note that a particular case of this is $\hat{\delta}\mathcal{L} = 0$.

Lemma 4.4.0.6. *If $\hat{\delta}\mathcal{L} = \hat{\delta}\varepsilon\partial_\mu f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)$ and a field ϕ_0^α is a stationary point of $S[\cdot, M]$, then a field $\hat{\phi}_0^\alpha$ is a stationary point of $S[\cdot, \hat{M}]$.*

Proof. Let's take an arbitrary field ϕ^α not necessarily stationary. We will show that

$$S[\hat{\phi}^\alpha, \hat{M}] = S[\phi^\alpha, M] + \int\limits_{\partial M} \hat{\delta}\varepsilon f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x) + \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)\hat{\delta}x^\mu dn_\mu. \tag{4.4.0.20}$$

Indeed,

$$\begin{aligned}
S[\hat{\phi}^\alpha, \hat{M}] &= \int\limits_{\hat{M}} \mathcal{L}(\hat{\phi}^\alpha, \partial_\mu\hat{\phi}^\alpha, x^\mu)d^4x = \int\limits_{\hat{M}} \mathcal{L} + \hat{\delta}\mathcal{L}d^4x \\
&= \int\limits_{M} \mathcal{L} + \hat{\delta}\mathcal{L}d^4 + \int (\mathbb{1}_{\hat{M}} - \mathbb{1}_M)(\mathcal{L} + \hat{\delta}\mathcal{L})d^4x \\
&= S[\phi^\alpha, M] + \int\limits_{M} \hat{\delta}\varepsilon\partial_\mu f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)d^4x + \int\limits_{\partial M} (\mathcal{L} + \hat{\delta}\mathcal{L})\hat{\delta}x^\mu dn_\mu \\
&= S[\phi^\alpha, M] + \int\limits_{\partial M} (\hat{\delta}\varepsilon f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu) + \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)\hat{\delta}x^\mu)dn_\mu.
\end{aligned}$$

Take any variation of the field $\delta\phi^\alpha$ for which $\delta\phi^\alpha = \partial_\mu\delta\phi^\alpha = 0$ on ∂M. Note that because $\delta\hat{\phi}^\alpha(x) = \delta\hat{\phi}^\alpha(x + \hat{\delta}x)$, we have $\delta\hat{\phi}^\alpha = \partial_\mu\delta\hat{\phi}^\alpha = 0$ on $\partial\hat{M}$.

Since (4.4.0.20) was for an arbitrary field we have

$$\begin{aligned}
&S[\hat{\phi}_0^\alpha + \delta\hat{\phi}^\alpha, \hat{M}] = S[\phi_0^\alpha + \delta\phi^\alpha, M] \\
&+ \int\limits_{\partial M} \left(\hat{\delta}\varepsilon f^\mu(\phi_0^\alpha + \delta\phi^\alpha, \partial_\mu\phi_0^\alpha + \partial_\mu\delta\phi^\alpha, x^\mu) + \mathcal{L}(\phi_0^\alpha + \delta\phi^\alpha, \partial_\mu\phi_0^\alpha + \partial_\mu\delta\phi^\alpha, x^\mu)\hat{\delta}x^\mu\right)dn_\mu.
\end{aligned}$$

But since $\delta\phi^\alpha = \partial_\mu \delta\phi^\alpha = 0$ on ∂M, we have

$$S[\hat{\phi}_0^\alpha + \delta\hat{\phi}^\alpha, \hat{M}] = S[\phi_0^\alpha + \delta\phi^\alpha, M] + \int\limits_{\partial M} \Big(\hat{\delta}\varepsilon f^\mu(\phi_0^\alpha, \partial_\mu\phi_0^\alpha, x^\mu) + \mathcal{L}(\phi_0^\alpha, \partial_\mu\phi_0^\alpha, x^\mu)\hat{\delta}x^\mu \Big) dn_\mu.$$

Hence,

$$S[\hat{\phi}_0^\alpha + \delta\hat{\phi}^\alpha, \hat{M}] = S[\phi_0^\alpha + \delta\phi^\alpha, M] + C(\phi_0^\alpha). \tag{4.4.0.21}$$

and from that it follows that a field $\hat{\phi}_0^\alpha$ is a stationary point of $S[\cdot, \hat{M}]$. □

Now, we are ready to prove a celebrated Noether's theorem, in which one derives a locally conserved Noether's current J^μ from the invariance of stationary-point field under a chosen infinitesimal continuous symmetry of time-space for a given action functional.

Theorem 4.4.0.7. *(Noether's Theorem) If $\hat{\delta}\mathcal{L} = \hat{\delta}\varepsilon\partial_\mu f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x)$ and a field ϕ_0^α is a stationary point of $S[\cdot, M]$ for an arbitrary M, then for*

$$J^\mu = \Pi_\alpha^\mu \partial_\nu \phi_0^\alpha \hat{\delta}x^\nu - \hat{\delta}\varepsilon f^\mu, \tag{4.4.0.22}$$

we have

$$\partial_\mu J^\mu = 0. \tag{4.4.0.23}$$

Proof. We will make all calculations for fixed $\phi^\alpha = \phi_0^\alpha$. From Lemma 4.4.0.5, we have

$$\hat{\delta}S[\phi^\alpha, M] = \int\limits_M \partial_\mu \Big(\Pi_\alpha^\mu \hat{\delta}\phi^\alpha + \hat{\delta}x^\mu \mathcal{L} \Big) d^4x. \tag{4.4.0.24}$$

On the other hand from (4.4.0.20) we have

$$\hat{\delta}S[\phi^\alpha, M] = \int\limits_M \partial_\mu \Big(\hat{\delta}\varepsilon f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x) + \mathcal{L}(\phi^\alpha, \partial_\mu\phi^\alpha, x^\mu)\hat{\delta}x^\mu \Big) d^4x. \tag{4.4.0.25}$$

Thus

$$0 = \int\limits_M \partial_\mu \Big(\Pi_\alpha^\mu \hat{\delta}\phi^\alpha - \hat{\delta}\varepsilon f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x) \Big) d^4x. \tag{4.4.0.26}$$

and since M is arbitrary

$$\partial_\mu \Big(\Pi_\alpha^\mu \hat{\delta}\phi^\alpha - \hat{\delta}\varepsilon f^\mu(\phi^\alpha, \partial_\mu\phi^\alpha, x) \Big) = 0. \tag{4.4.0.27}$$

From the above and because $\hat{\delta}\phi^\alpha = \partial_\mu\phi^\alpha\hat{\delta}x^\mu$, for

$$J^\mu = \Pi_\alpha^\mu \partial_\nu \phi^\alpha \hat{\delta}x^\nu - \hat{\delta}\varepsilon f^\mu, \tag{4.4.0.28}$$

we have $\partial_\mu J^\mu = 0$. □

Momentum preservation Let ϕ^α be a field which is a stationary point of an action functional for certain Lagrangian density $\mathcal{L}$. Consider infinitesimal translation $\hat{\phi}^\alpha(x) = \phi^\alpha(x + \hat{\delta}\varepsilon a)$ where a is a constant 4-vector.

Note that

$$\begin{aligned}\hat{\delta}\mathcal{L} &= \mathcal{L}(\phi^\alpha(x + \hat{\delta}\varepsilon a), \partial_\mu \phi^\alpha(x + \hat{\delta}\varepsilon a), x^\mu) - \mathcal{L}(\phi^\alpha, \partial_\mu \phi^\alpha, x^\mu) \\ &= \hat{\delta}\varepsilon \partial_\mu \mathcal{L} a^\mu = \hat{\delta}\varepsilon \partial_\mu (\mathcal{L}a)^\mu .\end{aligned}$$

Mind that ∂_μ is a full derivative with respect to x^μ. The locally conserved Noether's current for infinitesimal translation is then:

$$\begin{aligned}J^\mu &= \delta\varepsilon(a^\nu \Pi^\mu_\alpha \partial_\nu \phi^\alpha - a^\mu \mathcal{L}) \\ &= \delta\varepsilon a^\nu (\Pi^\mu_\alpha \partial_\nu \phi^\alpha - \delta^\mu_\nu \mathcal{L}).\end{aligned}$$

If we consider translation along x^0 axis, let's assume $a = (1, 0, 0, 0)$, then

$$J^0 = \delta\varepsilon(\Pi^0_\alpha \partial_0 \phi^\alpha - \mathcal{L}), \tag{4.4.0.29}$$

which is related to energy conservation and

$$J^k = \delta\varepsilon \Pi^k_\alpha \partial_k \phi^\alpha, \tag{4.4.0.30}$$

which is related to momentum conservation along axis x^k.

Chapter 5

General Relativity

5.1 Basic properties

5.1.1 Preliminaries

$$\eta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad \delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

5.1.2 Tensors

$(\hat{T})^{i_1 \dots i_n}_{i_{n+1} \dots i_m} = \frac{\partial \hat{x}^{i_1}}{\partial x^{j_1}} \cdots \frac{\partial \hat{x}^{i_n}}{\partial x^{j_n}} \frac{\partial x^{j_{n+1}}}{\partial \hat{x}^{i_{n+1}}} \cdots \frac{\partial x^{j_m}}{\partial \hat{x}^{i_m}} T^{j_1 \dots j_n}_{j_{n+1} \dots j_m}$

$$(\hat{A})^{\mu} = \frac{\partial \hat{x}^{\mu}}{\partial x^{\nu}} A^{\nu}$$

5.1.3 Metric Tensor

$$\boxed{d\tau^2 = g_{\mu\nu} dx^{\mu} dx^{\nu}}$$

Theorem 5.1.3.1. *$g_{\mu\nu}$ is a tensor.*

Proof. Notice that $dx^{\mu} = \frac{\partial x^{\mu}}{\partial \hat{x}^{\alpha}} d\hat{x}^{\alpha}$.
Hence $d\tau^2 = g_{\mu\nu} dx^{\mu} dx^{\nu} = (g_{\mu\nu} \frac{\partial x^{\mu}}{\partial \hat{x}^{\alpha}} \frac{\partial x^{\nu}}{\partial \hat{x}^{\beta}}) d\hat{x}^{\alpha} d\hat{x}^{\beta}$. □

Theorem 5.1.3.2. *For each point ω there exists a frame of reference*

$$(x^0, x^1, x^2, x^3)$$

such that $g_{\mu\nu} = \eta_{\mu\nu}$ at ω.

$$\boxed{g_{\mu\lambda} g^{\lambda\nu} = g^{\nu}_{\mu} = \delta^{\nu}_{\mu}}$$

5.1.4 Christoffel symbols

$$\boxed{\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} \left(\frac{\partial g_{\rho\mu}}{\partial x^{\nu}} + \frac{\partial g_{\rho\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}} \right)}$$

5.1.5 Geodecis

$$\boxed{\frac{d^2\phi^{\mu}}{d\tau^2} + \Gamma^{\mu}_{\alpha\beta} \frac{d\phi^{\alpha}}{d\tau} \frac{d\phi^{\beta}}{d\tau} = 0}$$

5.1.6 Motionless particle

Assume that we have a frame of reference $(t = x_0, x_1, x_2, x_3)$. Let's consider a motionless particle

$$(\phi^0(t) = t, \phi^1(t), \phi^2(t), \phi^3(t)).$$

As the particle is motionless, we have $\dfrac{d\phi^n}{dt} = 0$ for $n = 1, 2, 3$ at the moment $t = 0$. The particle is motionless in our space frame of reference (x, y, z) at the moment $t = 0$, but we assume that it's still following geodesic in the space-time.

$$0 = \frac{d^2\phi^{\mu}}{d\tau^2} + \Gamma^{\mu}_{\alpha\beta} \frac{d\phi^{\alpha}}{d\tau} \frac{d\phi^{\beta}}{d\tau} = \frac{d^2\phi^{\mu}}{d\tau^2} + \Gamma^{\mu}_{\alpha\beta} \frac{d\phi^{\alpha}}{dt} \frac{d\phi^{\beta}}{dt} \Big(\frac{dt}{d\tau} \Big)^2 = \frac{d^2\phi^{\mu}}{d\tau^2} + \Gamma^{\mu}_{00} \Big(\frac{dt}{d\tau} \Big)^2.$$

$$\frac{d^2\phi^{\mu}}{d\tau^2} = -\Gamma^{\mu}_{00} \Big(\frac{dt}{d\tau} \Big)^2.$$

The above holds in the point $(0, \phi^1(0), \phi^2(0), \phi^3(0))$.

$$\frac{d^2\phi^n}{d\tau^2} = \frac{d}{d\tau} \Big(\frac{d\phi^n}{dt} \frac{dt}{d\tau} \Big) = \frac{d}{d\tau} \Big(\frac{d\phi^n}{dt} \Big) \frac{dt}{d\tau} + \frac{d\phi^n}{dt} \frac{d^2 t}{d\tau^2} = \frac{d}{dt} \Big(\frac{d\phi^n}{dt} \Big) \Big(\frac{dt}{d\tau} \Big)^2 = \frac{d^2\phi^n}{dt^2} \Big(\frac{dt}{d\tau} \Big)^2$$

for $n = 1, 2, 3$. Thus,

$$\frac{d^2\phi^n}{dt^2} = -\Gamma^n_{00} \text{ for } n = 1, 2, 3.$$

Therefore,

$$\frac{d^2\phi^n}{dt^2} = -\Gamma^n_{00} = -\frac{1}{2} g^{n\rho} \left(\frac{\partial g_{\rho 0}}{\partial x^0} + \frac{\partial g_{\rho 0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^{\rho}} \right).$$

Now we will assume that the curvature is constant in time.

Assumption 5.1.6.1. $\frac{\partial g_{\mu\nu}}{\partial x^0} = 0.$

Now,

$$\frac{d^2\phi^n}{dt^2} = \frac{1}{2} g^{nm} \frac{\partial g_{00}}{\partial x^m}.$$

5.1.7 The Ricci tensor

$$\boxed{R_{\mu\nu} = \frac{\partial \Gamma^{\alpha}_{\mu\alpha}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}}$$

Einstein assumption for empty space:

$$\boxed{R_{\mu\nu} = 0}$$

5.1.8 A slow particle in weakly curved empty space-time: The Newtonian aproximation

Let's consider a slow particle

$$(\phi^0(t) = t, \phi^1(t), \phi^2(t), \phi^3(t)).$$

Assume that $\phi^1(0) = \phi^2(0) = \phi^3(0) = 0$ and that for our frame of refrence $g_{\mu\nu} = \eta_{\mu\nu}$ in $(0, 0, 0, 0)$. The particle is slow so we assume:

Assumption 5.1.8.1. $\frac{d\phi^n}{dt}$ *is an infinitesimal of the first order for* $n = 1, 2, 3$.

We will asume that curvature is constant in time.

Assumption 5.1.8.2. $\frac{\partial g_{\mu\nu}}{\partial x^0} = 0.$

We will assume also that the curvature is weak.

Assumption 5.1.8.3. $\frac{\partial g_{\mu\nu}}{\partial x^n}$ *is an infinitesimal of the first order for* $n = 1, 2, 3$.

Proposition 5.1.8.4. $g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon_{\mu\nu}$*, where* $\varepsilon_{\mu\nu}$ *is an infinitesimal of the second order in reasonable range that we care about.*

Proof.

$$g_{\mu\nu}(d\phi) = \eta_{\mu\nu} + \frac{\partial g_{\mu\nu}}{\partial x^n} d\phi^n$$

□

The particle is following the geodecis in the spacetime, so with neglecting second-order infinitesimals we have:

$$0 = \frac{d^2\phi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{d\phi^\alpha}{d\tau}\frac{d\phi^\beta}{d\tau} = \frac{d^2\phi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{d\phi^\alpha}{dt}\frac{d\phi^\beta}{dt}\Big(\frac{dt}{d\tau}\Big)^2 = \frac{d^2\phi^\mu}{d\tau^2} + \Gamma^\mu_{00}\Big(\frac{dt}{d\tau}\Big)^2.$$

Thus

$$\frac{d^2\phi^\mu}{d\tau^2} = -\Gamma^\mu_{00}\Big(\frac{dt}{d\tau}\Big)^2. \tag{5.1.8.1}$$

Putting $\mu = 0$ we may conclude that $\frac{d^2t}{d\tau^2}$ is a first order infinitesimal.

$$\frac{d^2\phi^n}{d\tau^2} = \frac{d}{d\tau}\Big(\frac{d\phi^n}{dt}\frac{dt}{d\tau}\Big) = \frac{d}{d\tau}\Big(\frac{d\phi^n}{dt}\Big)\frac{dt}{d\tau}+\frac{d\phi^n}{dt}\frac{d^2t}{d\tau^2} = \frac{d}{dt}\Big(\frac{d\phi^n}{dt}\Big)\Big(\frac{dt}{d\tau}\Big)^2 = \frac{d^2\phi^n}{dt^2}\Big(\frac{dt}{d\tau}\Big)^2$$

for $n = 1, 2, 3$, neglecting second-order infinitesimals. Therefore applying the above to the (5.1.8.1) we get

$$\frac{d^2\phi^n}{dt^2} = -\Gamma^n_{00} \text{ for } n = 1, 2, 3.$$

so

$$\frac{d^2\phi^n}{dt^2} = -\Gamma^n_{00} = -\frac{1}{2}g^{n\rho}\left(\frac{\partial g_{\rho 0}}{\partial x^0} + \frac{\partial g_{\rho 0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^\rho}\right).$$

and by Assumption 5.1.8.2, neglecting second-order infinitesimals:

$$\frac{d^2\phi^n}{dt^2} = \frac{1}{2}\frac{\partial g_{00}}{\partial x^n}.$$

Since $R_{\mu\nu} = 0$, neglecting second-order infinitesimals we have:

$$\frac{\partial\Gamma^\alpha_{\mu\alpha}}{\partial x^\nu} - \frac{\partial\Gamma^\alpha_{\mu\nu}}{\partial x^\alpha} = 0$$

Note that with neglecting of second-order infinitesimals,

$$\frac{\partial\Gamma^\alpha_{\mu\alpha}}{\partial x^\nu} = \frac{1}{2}g^{\alpha\rho}\Big(\frac{\partial^2 g_{\rho\mu}}{\partial x^\alpha\partial x^\nu} + \frac{\partial^2 g_{\rho\alpha}}{\partial x^\mu\partial x^\nu} - \frac{\partial^2 g_{\mu\alpha}}{\partial x^\alpha\partial x^\nu}\Big),$$

$$\frac{\partial\Gamma^\alpha_{\mu\nu}}{\partial x^\nu} = \frac{1}{2}g^{\alpha\rho}\Big(\frac{\partial^2 g_{\rho\mu}}{\partial x^\nu\partial x^\alpha} + \frac{\partial^2 g_{\rho\nu}}{\partial x^\mu\partial x^\alpha} - \frac{\partial^2 g_{\mu\nu}}{\partial x^\rho\partial x^\alpha}\Big).$$

Hence,

$$\frac{1}{2}g^{\alpha\rho}\Big(\frac{\partial^2 g_{\rho\nu}}{\partial x^\mu \partial x^\alpha} - \frac{\partial^2 g_{\mu\alpha}}{\partial x^\alpha \partial x^\nu} - \frac{\partial^2 g_{\rho\nu}}{\partial x^\mu \partial x^\alpha} + \frac{\partial^2 g_{\mu\nu}}{\partial x^\rho \partial x^\alpha}\Big) = 0.$$

Choose $\mu = \nu = 0$. Then by Assumption 5.1.8.2

$$g^{mn}\frac{\partial^2 g_{00}}{\partial x^m \partial x^n} = 0 \text{ for } n, m = 1, 2, 3.$$

By Proposition 5.1.8.4 with neglecting second-order infinitesimals we have

$$\eta_{nn}\frac{\partial^2 g_{00}}{\partial x^n \partial x^n} = 0 \text{ for } n = 1, 2, 3.$$

Which is a Laplace equation in $\mathbb{R}^3$.

5.1.9 Four-velocity

Assume that we describe a particle in a frame of reference x^μ as $(\phi^0(t) = t, \phi^1(t), \phi^2(t), \phi^3(t))$. We define a four-velocity vector as:

$$\boxed{v^\mu = \frac{d\phi^\mu}{d\tau}}$$

where $d\tau$ is a infinitesimal length element of curve ϕ.

Note that

$$d\tau = (g_{\mu\nu} d\phi^\mu d\phi^\nu)^{\frac{1}{2}}. \tag{5.1.9.1}$$

Hence

$$\frac{d\tau}{dt} = (g_{\mu\nu}\frac{d\phi^\mu}{dt}\frac{d\phi^\nu}{dt})^{\frac{1}{2}}.$$

Notive that

$$v^0\frac{d\tau}{dt} = \frac{d\phi^0}{d\tau}\frac{d\tau}{dt} = \frac{d\phi^0}{dt} = 1.$$

So

$$\boxed{v^0 = (g_{\mu\nu}\frac{d\phi^\mu}{dt}\frac{d\phi^\nu}{dt})^{-\frac{1}{2}}}$$

Note that v^0 is what we usually denote in literature as γ. On the other hand:

$$\frac{d\tau}{dt} = (v^0)^{-1}$$

And thus:

$$\frac{d\phi^\mu}{dt} = \frac{d\phi^\mu}{d\tau}\frac{d\tau}{dt} = v^1(v^0)^{-1}.$$

$$\boxed{\frac{d\phi^\mu}{dt} = v^1(v^0)^{-1}}$$

Which means that once you know the four-velocity in given frame of reference, you know space velocities as well.

There is one more implication from (5.1.9.1):

$$\boxed{g_{\mu\nu}v^\mu v^\nu = 1} \tag{5.1.9.2}$$

We will show that v^μ is a tensor. Assume that y^μ is a new frame of reference. Note that

$$d\hat{\phi}^\mu = \frac{\partial y^\mu}{\partial x^\nu}d\phi^\nu.$$

Since $d\tau$ is invariant $\hat{v}^\mu d\tau = \dfrac{\partial y^\mu}{\partial x^\nu}v^\nu d\tau$, hence

$$\boxed{\hat{v}^\mu = \frac{\partial y^\mu}{\partial x^\nu}v^\nu}$$

Let $\omega^\mu = \dfrac{d\phi^\mu}{dt}$. ($\omega^\nu$ is not a tensor). Then $\omega^\mu = v^\mu(v^0)^{-1}$.
Note that usually encountered in literature $\beta = (\omega^1 + \omega^2 + \omega^3)^{\frac{1}{2}}$. Note

$$\boxed{v^0 = (1-\beta^2)^{-\frac{1}{2}}} \tag{5.1.9.3}$$

To simplify calculations, we can assume that $g_{\mu\nu} = \eta_{\mu\nu}$ at $(0,0,0,0)$. $\phi^\mu(0) = 0$ and that particle at the moment $t = 0$ is moving along x^1, i.e. $\omega^2 = \omega^3 = 0$. Thus $v^2 = v^3 = 0$. Let

$$\begin{cases} y^0 = v^0x^0 - v^1x^1 \\ y^1 = v^0x^1 - v^1x^0 \\ y^2 = x^2 \\ y^3 = x^3 \end{cases} \tag{5.1.9.4}$$

It's easy to show that $\hat{v}^0 = 1$ and $\hat{v}^1 = \hat{v}^2 = \hat{v}^3 = 0$, which simply means that for $t = 0$ the particle is in rest in the frame of reference y^μ. Obviously

$$\begin{cases} x^0 = v^0y^0 + v^1y^1 \\ x^1 = v^0y^1 + v^1y^0 \\ x^2 = y^2 \\ x^3 = y^3 \end{cases} \tag{5.1.9.5}$$

$$\hat{g}_{00} = g_{\mu\nu}\frac{\partial x^\mu}{\partial y^0}\frac{\partial x^\nu}{\partial y^0} = (v^0)^2 - (v^1)^2 = 1,$$
$$\hat{g}_{01} = g_{\mu\nu}\frac{\partial x^\mu}{\partial y^0}\frac{\partial x^\nu}{\partial y^1} = v^0v^1 - v^1v^0 = 0,$$
$$\hat{g}_{11} = g_{\mu\nu}\frac{\partial x^\mu}{\partial y^1}\frac{\partial x^\nu}{\partial y^1} = (v^1)^2 - (v^0)^2 = -1.$$

Thus also $\hat{g}_{\mu\nu} = \eta_{\mu\nu}$ for in $(0,0,0,0)$.

Consider now particle ϕ in some very short (in fact as short as we need) time after $t = 0$. We can assume that in the frame of reference x we can describe it as:

$$t \to (t, \omega^1 t, 0, 0)$$

.

Let's see this particle in the frame y. Assume that time interval that we consider is so short that we all the time can use (5.1.9.4):

$$t \to (v^0t - v^1w^1t, v^0w^1t - v^1t, 0, 0) = (v^0(1-(w^1)^2)t, 0, 0, 0) = ((v^0)^{-1}t, 0, 0, 0)$$

It follows that if for the observer in x interval Δt passed, for observer in y interval $(v^0)^{-1}\Delta t$ passed, which is of course time dilatation. Note that t in frame of reference y is merely a parameter. To summarise, this is the law of time dilatation

$$\boxed{\Delta t' = (v^0)^{-1}\Delta t} \tag{5.1.9.6}$$

Where is Δt is time passed measured by an observer who is in rest relative to frame of reference x^μ and $\Delta t'$ is time passed measured by an observer who is in rest relative to a particle frame of reference y^μ.

Now consider second particle that is moving in exactly the same direction as first one but proceeding it with a distance $\Delta x'$ in frame of reference x. We can describe it as:

$$t \to (t, w^1t + \Delta x', 0, 0).$$

Now look at this in frame of reference y.

$$t \to ((v^0)^{-1}t - v^1\Delta x', v^0\Delta x', 0, 0).$$

So the distance from particles in frame of reference y is $v^0\Delta x'$. To summarise, this is the law of length contraction

$$\boxed{\Delta x' = (v^0)^{-1}\Delta x} \tag{5.1.9.7}$$

Where Δx is a rest distance between particles and $\Delta x'$ is a distance between particles measured by an observer who is in rest in the frame of refernce x^μ.

Theorem 5.1.9.1. *If u^μ, v^μ are 4-velocities of two observers at the point of intersection of their geodesics, then*

$$g_{\mu\nu}u^\mu v^\nu = (1-\omega)^{-\frac{1}{2}} \tag{5.1.9.8}$$

where ω is a value of the relative velocity measured by observers,

$$g_{\mu\nu}x^\mu v^\nu = -(1-\omega)^{-\frac{1}{2}}\omega_x. \tag{5.1.9.9}$$

where x_μ is a vector of length -1, orthogonal to u_μ, and ω_x is a velocity of v_ν measured by an observer u^μ along the vector x_μ.

Proof. Let p be the point of intersection. As the value of $g_{\mu\nu}u^\mu v^\nu$ doesn't depend on frame of reference, we are free to choose any frame of reference we need. Let's choose the frame of reference where $u = (1,0,0,0)$ and $v = (v^0, v^1, v^2, v^3)$ and $g_{\mu\nu} = \eta_{\mu\nu}$ at point p. Such frame of reference will be just Riemannian coordinates with $0-axis$ along vector u^μ. Thus $g_{\mu\nu}u^\mu v^\nu = \eta_{00}v_0 = v_0$, which proves equation 5.1.9.8. We may also require that in the frame of reference where $u = (1,0,0,0)$, $x = (0,1,0,0)$. Thus $g_{\mu\nu}x^\mu v^\nu = -v_1 = -v_0\omega_1$. which proves equation 5.1.9.9. □

Fact 5.1.9.2. *If u^μ, v^ν are 4-velocities of two observers at the point of intersection of their geodesics and λ is an arbitrary density of some quantity measured locally by an observer v^ν, then $\lambda g_{\mu\nu}u^\mu v^\nu$ is a density of this quantity measured locally by an observer u^μ.*

Fact 5.1.9.3. *If p^ν is a 4-momentum of a particle and u^μ is a 4-velocity of an observer at the point of interection with the geodesic of the particle, then*

$$E = g_{\mu\nu}u^\mu p^\nu, \tag{5.1.9.10}$$

$$p_x = g_{\mu\nu}x^\mu p^\nu. \tag{5.1.9.11}$$

where E is an energy of the particle measured by an observer, x_μ is a vector of length -1, orthogonal to u_μ, and p_x is a momentum of a particle measured by an observer u^μ along the vector x_μ.

5.1.10 Static spacetime

Definition 5.1.10.1. *We will say that system of coordinates x is static if $\dfrac{\partial g_{\mu\nu}}{\partial x^0} = 0$ and $g_{m0} = 0$ for $m = 1,2,3$.*

Fact 5.1.10.2. *If the system of coordinates x is static, then*

$$g^{00} = (g_{00})^{-1}, \tag{5.1.10.1}$$

$$g^{n0} = 0, \tag{5.1.10.2}$$

$$g_{na}g^{am} = \delta_n^m, \tag{5.1.10.3}$$

$$\Gamma_{00}^a = -\frac{1}{2}g^{ab}\frac{\partial g_{00}}{\partial x^b}, \tag{5.1.10.4}$$

$$\Gamma_{nm}^a = \frac{1}{2}g^{ab}\left(\frac{\partial g_{bm}}{\partial x^n} + \frac{\partial g_{bn}}{\partial x^m} - \frac{\partial g_{mn}}{\partial x^b}\right), \tag{5.1.10.5}$$

$$\Gamma_{n0}^a = \Gamma_{00}^0 = 0. \tag{5.1.10.6}$$

Proof.

$$\Gamma_{00}^a = \frac{1}{2}g^{ab}\left(\frac{\partial g_{b0}}{\partial x^0} + \frac{\partial g_{b0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^b}\right) = -\frac{1}{2}g^{ab}\frac{\partial g_{00}}{\partial x^b}.$$

$$\Gamma_{nm}^a = \frac{1}{2}g^{ab}\left(\frac{\partial g_{bm}}{\partial x^n} + \frac{\partial g_{bn}}{\partial x^m} - \frac{\partial g_{mn}}{\partial x^b}\right) + \frac{1}{2}g^{a0}(\cdots) = \frac{1}{2}g^{ab}\left(\frac{\partial g_{bm}}{\partial x^n} + \frac{\partial g_{bn}}{\partial x^m} - \frac{\partial g_{mn}}{\partial x^b}\right).$$

$$\Gamma_{n0}^a = \frac{1}{2}g^{ab}\left(\frac{\partial g_{b0}}{\partial x^n} + \frac{\partial g_{bn}}{\partial x^0} - \frac{\partial g_{0n}}{\partial x^b}\right) + \frac{1}{2}g^{a0}(\cdots) = 0.$$

□

Note that condition $g_{n0} = 0$ provides that $(dt, 0, 0, 0)$ and $(0, dx^1, dx^2, dx^3)$ are orthogonal at each point. We may say that all static observers agree locally on what is their space. Motivated by this we can define space as $t = const$.

Theorem 5.1.10.3. *If x is a static coordinate system and $t \to (t, \phi_1(t), \phi_2(t), \phi_3(t))$ is an equation of free particle (geodesic in coordinates x) and $\omega_\nu = \frac{d\phi^\nu}{dt}$, then*

$$\omega^a \nabla_a \omega^b = -\Gamma_{00}^a, \tag{5.1.10.7}$$

$$i.e.\ \frac{d^2\phi^a}{dt^2} + \Gamma_{nm}^a \frac{d\phi^n}{dt}\frac{d\phi^m}{dt} = -\Gamma_{00}^a. \tag{5.1.10.8}$$

where ∇ is a covariant derivative in 3 dimensional manifold with metric g_{nm}.

Theorem 5.1.10.4. *If x is a static coordinate system and $t \to (t, \phi_1(t), \phi_2(t), \phi_3(t))$ is a null geodesic in coordinates x, then for $t \to (\phi_1(t), \phi_2(t), \phi_3(t))$, we have $\delta \int dt = 0$.*

Proof. Let's define a new metric $H_{\mu\nu} = g_{\mu\nu}(g_{00})^{-1}$. H is conformally related to g. Thus they have the same null geodesics. Note that x is a static coordinate system in H. Then by Theorem 5.1.10.3

$$\frac{d^2\phi^a}{dt^2} + \overset{H}{\Gamma}{}^a_{nm}\frac{d\phi^n}{dt}\frac{d\phi^m}{dt} = -\overset{H}{\Gamma}{}^a_{00}. \tag{5.1.10.9}$$

But $\overset{H}{\Gamma}{}^a_{00} = 0$. Then $t \to (\phi_1(t), \phi_2(t), \phi_3(t))$ is a geodesic equation in a 3 dimensional manifold with metric H_{mn}. So $\delta \int ds_H = 0$. Note that

$$ds_H^2 = -H_{mn}dx^m dx^n = -(g_{00})^{-1}g_{mn}dx^m dx^n = (g_{00})^{-1}ds^2.$$

Hence

$$\delta \int g_{00}^{-\frac{1}{2}} ds = 0. \tag{5.1.10.10}$$

Because we are on the null geodesic in x with metric g, we have

$$0 = g_{00}dt^2 + g_{nm}dx^n dx^m = g_{00}dt^2 - ds^2.$$

So

$$dt = (g_{00})^{-\frac{1}{2}} ds.$$

□

And 5.1.10.10 becomes

$$\delta \int dt = 0.$$

5.1.11 Stress-energy tensor

We will construct some illustrative example of the stress-energy tensor. Assume that we have distribution of matter whose velocity varies continuously from one point to a neighboring one. Let ρ be a scalar field of rest density measured in the rest frame of reference of an infinitesimal part of the matter distribution. Let v^μ be a vector field of 4-velocities of an infinitesimal part of the matter distribution at the point. Let define

$$T^{\mu\nu} = \rho v^\mu v^\nu. \tag{5.1.11.1}$$

Let ω be a relative velocity between observer u^μ and an infenitesimal matter element v^μ at each point. Note that rest density of the matter distribution is just some quantity distributed in space. Forget for a moment that this is a rest density, let's treat this as some kind of abstract quantity. Then

$$\underbrace{\rho(1-\omega)^{-\frac{1}{2}}}_{\text{rest density}*\frac{1}{\text{length contraction}}} \tag{5.1.11.2}$$

is a density of this quantity measured by the observer u^μ, i. e. density of rest density measured by the observer u^μ.

Fact 5.1.11.1. *$T_{\mu\nu}u^\mu u^\nu$ is a energy density measured locally by an observer u^ν.*

Proof. We will use notation from Theorem 5.1.9.1

$$T_{\mu\nu}u^\mu u^\nu = \rho v^\alpha v^\beta g_{\mu\alpha}g_{\mu\beta}u^\mu u^\nu = \overbrace{\underbrace{\rho(1-\omega)^{-\frac{1}{2}}}_{\text{rest density}*\frac{1}{\text{length contraction}}} \cdot(1-\omega)^{-\frac{1}{2}}}^{\text{relative energy density}}. \tag{5.1.11.3}$$

□

Fact 5.1.11.2. *$T_{\mu\nu}x^\mu u^\nu$ is a density of momentum along the vector x^μ, measured locally by an observer u^ν – where x^μ is an orthogonal to u^μ vector of length -1.*

Proof.

$$T_{\mu\nu}x^\mu u^\nu = \rho v^\alpha v^\beta g_{\mu\alpha}g_{\mu\beta}x^\mu u^\nu = \overbrace{\underbrace{\rho(1-\omega)^{-\frac{1}{2}}}_{\text{rest density}*\frac{1}{\text{length contraction}}} \cdot(1-\omega)^{-\frac{1}{2}}\omega_x}^{\text{relative momentum density}}. \tag{5.1.11.4}$$

□

Fact 5.1.11.3. *$T_{\mu\nu}u^\mu x^\nu$ is a energy flux across the surface with normal vector x^μ, measured locally by an observer u^ν – where x^μ is an orthogonal to u^μ vector of length -1.*

Proof.

$$T_{\mu\nu}u^\mu x^\nu = \rho v^\alpha v^\beta g_{\mu\alpha}g_{\mu\beta}u^\mu x^\nu = \overbrace{\underbrace{\rho(1-\omega)^{-\frac{1}{2}}}_{\text{rest density}*\frac{1}{\text{length contraction}}} \cdot(1-\omega)^{-\frac{1}{2}}}^{\text{relative energy density}}\omega_x. \tag{5.1.11.5}$$

□

Of course $T_{\mu\nu}x^\mu u^\nu = T_{\mu\nu}u^\mu x^\nu$.

Fact 5.1.11.4. *$T_{\mu\nu}x^\mu y^\nu$ is a flux of momentum along x^μ across the surface with normal vector y^μ, measured locally by an observer u^ν – where x^μ, y^ν are orthogonal to u^μ vectors of length -1.*

Proof.

$$T_{\mu\nu}x^\mu y^\nu = \rho v^\alpha v^\beta g_{\mu\alpha} g_{\mu\beta} u^\mu x^\nu = \overbrace{\underbrace{\rho(1-\omega)^{-\frac{1}{2}}}_{\text{rest density}*\frac{1}{\text{length contraction}}} \cdot (1-\omega)^{-\frac{1}{2}}\omega_x}^{\text{relative momentum density}} \omega_y. \tag{5.1.11.6}$$

□

Chapter 6

Quantum Mechanics

6.1 Preliminaries

6.1.1 Free particle (first overview)

Assume, we want to find a function $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ which satisfies the following equation:

$$i\frac{\partial\psi(t,x)}{\partial t} = -\frac{1}{2m} \cdot \frac{\partial^2\psi(t,x)}{\partial x^2}. \tag{6.1.1.1}$$

This happen to be a Schrödinger equation of a free particle in Quantum Mechanics. But we will deal with it purely mathematically and in this subsection we won't need any of Quantum Mechanics. Very often when the context is clear, we will write $\psi_t(x) := \psi(t,x)$. We need of course to assume that $t \to \psi(t,x)$ is differentiable for all $x \in \mathbb{R}$ and that $x \to \psi(x,t)$ is two times differentiable for all $t \in \mathbb{R}$. If we assume that $\psi_t \in L^1(\mathbb{R})$, we can define:

$$\hat{\psi}(t,p) = \mathcal{F}(\psi_t)(p). \tag{6.1.1.2}$$

Then by Theorem 11.2.1.8, the equation (6.1.1.1) implies

$$\frac{\partial\hat{\psi}(t,p)}{\partial t} = -i\frac{p^2}{2m} \cdot \hat{\psi}(t,p). \tag{6.1.1.3}$$

The solution for the above equation is given by

$$\hat{\psi}(t,p) = A(p)\exp\left(-it\frac{p^2}{2m}\right). \tag{6.1.1.4}$$

If we define $\hat{\psi}(p) := \hat{\psi}(0,p)$ then

$$\hat{\psi}(t,p) = \hat{\psi}(p)\exp\left(-it\frac{p^2}{2m}\right). \tag{6.1.1.5}$$

If we assume that $\hat{\psi}(p)$ is $L^1(\mathbb{R})$ we have a solution of (6.1.1.1) $\psi(t,x) = \mathcal{F}^{-1}(\hat{\psi}_t)(x)$ which expands to

$$\psi(t,x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\hat{\psi}(p)\exp\left(ip(x-\frac{tp}{2m})\right)dp. \tag{6.1.1.6}$$

This is a wave packet. Note that if $|\hat{\psi}(p)|^2$ is a probability density, then $|\hat{\psi}_t|^2 = |\hat{\psi}(p)|^2$ is a probability density for each $t \in \mathbb{R}$. Then also Theorem 11.2.1.10 (The Parserval Formula) implies that $|\psi_t|^2$ is a probability density for all $t \in \mathbb{R}$. We will proceed further under the assumption that $|\hat{\psi}(p)|^2$ is a probability density. To analyze how the wave packet moves in time it will be useful to introduce some abstract random processes X_t and P_t. Let X_t be an arbitrary random process where probability density of X_t is $|\psi_t|^2$ and let P_t be an arbitrary process where probability density of P_t is $|\hat{\psi}_t|^2$. Given that $|\hat{\psi}_t|^2 = |\hat{\psi}(p)|^2$ we will assume $P = P_t$ with probability density $|\hat{\psi}(p)|^2$. We don't assume a joint distribution for X_t and P_t, so they can't be read as momentum and position random processes, because in general as Nelson Theorem states: "The observables R_1 and R_2 have a joint probability distribution in all states if and only if they commute, that is if and only if $[R_1, R_2] = 0$." ([**?**], [**?**]). We have

$$E(X_t) = \int_{-\infty}^{\infty} x\psi(x,t)\overline{\psi(x,t)}dx, \tag{6.1.1.7}$$

and

$$E(P) = \int_{-\infty}^{\infty} p\hat{\psi}(p,t)\overline{\hat{\psi}(p,t)}dp = \int_{-\infty}^{\infty} p\hat{\psi}(p)\overline{\hat{\psi}(p)}dp. \tag{6.1.1.8}$$

By Theorem 11.2.1.14

$$E(X_t) = \int_{-\infty}^{\infty} i\frac{\partial\hat{\psi}}{\partial p}(p,t)\cdot\overline{\hat{\psi}(p,t)}dp. \tag{6.1.1.9}$$

Note that

$$\frac{\partial\hat{\psi}}{\partial p}(p,t) = \left(\frac{d\hat{\psi}}{dp}(p) - \frac{itp}{m}\hat{\psi}(p)\right)\exp\left(-it\frac{p^2}{2m}\right). \tag{6.1.1.10}$$

Thus

$$E(X_t) = \int_{-\infty}^{\infty} i\Big(\frac{d\hat{\psi}}{dp}(p) - \frac{itp}{m}\hat{\psi}(p)\Big)\overline{\hat{\psi}(p)}dp. \tag{6.1.1.11}$$

And from the above, by (6.1.1.9) and (6.1.1.8) we are getting a very nice equation

$$E(X_t) = E(X_0) + \frac{t}{m}E(P). \tag{6.1.1.12}$$

Which says that if we consider t as time, the mean value of a random variable with density $|\psi_t|^2$ travels along x axis with a constant velocity $v = \frac{\langle P \rangle}{m}$. Let's now analyze $\mathrm{Var}(X_t)$ to observe ψ_t wave dispersion. Again by Theorem 11.2.1.14, we have

$$E(X_t^2) = \int_{-\infty}^{\infty} -\frac{\partial^2\hat{\psi}}{\partial p^2}(p,t) \cdot \overline{\hat{\psi}(p,t)}dp. \tag{6.1.1.13}$$

Note that

$$\frac{\partial^2\hat{\psi}}{\partial p^2}(p,t) = \big(\frac{d^2\hat{\psi}}{dp^2}(p) - 2\frac{itp}{m}\cdot\frac{d\hat{\psi}}{dp}(p) - \frac{it}{m}\hat{\psi}(p) - \frac{t^2p^2}{m^2}\hat{\psi}(p)\big)e^{-it\frac{p^2}{2m}}. \tag{6.1.1.14}$$

Thus

$$E(X_t^2) = \int_{-\infty}^{\infty} \Big(-\frac{d^2\hat{\psi}}{dp^2}(p) + \frac{t^2p^2}{m^2}\hat{\psi}(p) + \frac{it}{m}\big(2p\frac{d\hat{\psi}}{dp}(p) - \hat{\psi}(p)\big)\Big)\overline{\hat{\psi}(p)}dp. \tag{6.1.1.15}$$

Hence

$$E(X_t^2) = E(X_0^2) + \frac{t^2}{m^2}E(P^2) + \int_{-\infty}^{\infty} \frac{it}{m}\Big(2p\frac{d\hat{\psi}}{dp}(p) - \hat{\psi}(p)\Big)\overline{\hat{\psi}(p)}dp, \tag{6.1.1.16}$$

and

$$\mathrm{Var}(X_t) = \mathrm{Var}(X_0) + \frac{t^2}{m^2}\mathrm{Var}(P) + \frac{t}{m}\Big(\int_{-\infty}^{\infty} i\big(2p\frac{d\hat{\psi}}{dp}(p) - \hat{\psi}(p)\big)\overline{\hat{\psi}(p)}dp - 2E(X_0)E(P)\Big). \tag{6.1.1.17}$$

It is easy to notice that for $\hat{\psi}(p) \in \mathbb{R}$ for all $p \in \mathbb{R}$ (it might be helpful to note first that in such case $E(X_0) = 0$.) holds:

$$\mathrm{Var}(X_t) = \mathrm{Var}(X_0) + \frac{t^2}{m^2}\mathrm{Var}(P). \tag{6.1.1.18}$$

Note that the reasoning up to this point is purely mathematical, we didn't use any postulates of quantum theory or we didn't use consciously momentum or position operators. However obviously they appeared in calculations. Similar but more detailed approach you may find in [**?**].

6.2 Quantum theory

6.2.1 Naive Momentum and Position Operators

In this section, unless not stated otherwise, we assume that all $L^p(X)$, $C^n(X)$, $C_0(X)$ etc. are sets of complex valued functions. We will use symbols P i Q to denote momentum and position operators respectively.

$$P\psi = -i\frac{d}{dx}\psi. \tag{6.2.1.1}$$

$$(Q\psi)(x) = x\psi(x). \tag{6.2.1.2}$$

For any two linear operators A, B understood at the moment as very abstract linear operations with no explicit domain, we define commutator as $[A, B] = AB - BA$.

Theorem 6.2.1.1. *Let $\psi : \mathbb{R} \to \mathbb{C}$ be differntiable function, then*

$$[Q, P]\psi = i\psi. \tag{6.2.1.3}$$

Proof.

$$QP\psi = Q(-i\frac{d\psi}{dx}) = -ix\frac{d\psi}{dx}. \tag{6.2.1.4}$$

$$PQ\psi = P(x\psi) = -i\psi + -ix\frac{d\psi}{dx}. \tag{6.2.1.5}$$

If we subtract equation (6.2.1.5) from equation (6.2.1.4), we got (6.2.1.3). □

6.2.2 Momentum operator in n-dimensional space

Below we will develop in more details what is mentioned in [?] (Quantum Probability 2.1) and [?] (Perspectives on the Spectral Theorem 6.6).

In this subsection $\mathcal{F}$ will be a Fourier transform defiend by Definition 11.4.0.11. The most usual case will be $n = 3$, but we also quite often consider particles moving only along the stright line when $n = 1$. Mathematically speaking though, we can think even about $n > 3$.

We will define momentum operator, from the perspective of spectral measure. This aproach has following advantages. It is clear from the begining that momentum operator is self-adjoint and there is no problem what is its domain, we don't need to define its domain, we mearly need to discover it. The below definition will be fully understood in the context of Theorem 11.3.2.1

Definition 6.2.2.1. *Let $H = L^2(\mathbb{R}^n)$. Let E be a spectral measure defined as follows*

$$E_i(\omega)(\psi) = \mathcal{F}^{-1}(1_\omega(x_i)\mathcal{F}(\psi)), \tag{6.2.2.1}$$

where ψ is treated as a function of $x = (x_1, \dots, x_n)$. The momentum operator along the coordinate i is

$$P_i = E_i(id). \tag{6.2.2.2}$$

Recall that $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a unitary mapping. Let's see how much we can derive from the above definition by Theorem 11.3.2.1. First we know that P_i is sef-adjoin and that the domain $\mathcal{D}(P_i)$ is dense in $L^2(\mathbb{R}^n)$. Next, that

$$\mathcal{D}(P_i) = \{\phi \in L^2(\mathbb{R}^n) : x_i\mathcal{F}(\psi) \in L^2(\mathbb{R}^n)\}. \tag{6.2.2.3}$$

and

$$P_i\psi = \mathcal{F}^{-1}(x_i\mathcal{F}(\psi)). \tag{6.2.2.4}$$

Which by Theorem 11.4.0.19 translates imediatelly to

$$\mathcal{D}(P_i) = \{\psi \in L^2(\mathbb{R}^n) : \mathsf{D}^i\psi \in L^2(\mathbb{R}^n)\}. \tag{6.2.2.5}$$

and

$$P_i\psi = -i\mathsf{D}^i\psi. \tag{6.2.2.6}$$

Which for any differentiable $\psi \in L^2(\mathbb{R}^n)$ means simply

$$P_i\psi = -i\frac{\partial\psi}{\partial x_i}. \tag{6.2.2.7}$$

Also we see immediately that $\mathcal{F}(\psi)$ is momentum representation of a wave function ψ with

$$\langle\phi, P_i\psi\rangle = \int x_i\mathcal{F}(\psi)\overline{\mathcal{F}(\phi)}dx^n. \tag{6.2.2.8}$$

Example 6.2.2.2. *Let $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be the extension of Fourier transform. We will define a spectral measure*

$$E(\omega)\psi = \mathcal{F}^{-1}(1_\omega \cdot \mathcal{F}(\psi)). \tag{6.2.2.9}$$

Note that for $\psi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ and $\mathcal{F}(\psi) \in L^1(\mathbb{R})$.

$$(E(\omega)\psi)(x) = \frac{1}{\sqrt{2\pi}}\int_\omega e^{ixk}\mathcal{F}(\psi)(k)dk. \tag{6.2.2.10}$$

$P\psi = E(id)$ is obviously a 1-dimensional momentum operator.

6.3 Harmonic oscillator (first attempt)

Let's write Schrödinger equation of harmonic oscillator in the following form

$$i\frac{\partial \psi}{\partial t} = \frac{1}{2m}P^2\psi + \frac{1}{2}m\omega^2 Q^2\psi. \tag{6.3.0.1}$$

The Hamiltonian in the equation is $H = \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 Q^2$. We want to find all possible states of energy $E \in \mathbb{R}$, which are H eigenvalues

$$H\psi = E\psi. \tag{6.3.0.2}$$

for any differentiable $\psi \in L^2(\mathbb{R})$. We will put aside domain considerations. All transformation should be true in a Schwartz space S_1 (Definition 11.4.0.1) which is dense in $L^2(\mathbb{R})$. This statement is not necesarly removing all difficulties, we will perhaps solve this more rigoriously later. At some point it will be good to find proper theorems to deal rigorously with solutions of equations, as for having P self-adjoin we extended domain to weakly differentiable functions – so now all theorems that say about uniquess of solutions must work for weak deriviatives.

Let's define

$$P_0 := \frac{1}{\sqrt{2m\omega}}P, \tag{6.3.0.3}$$

$$Q_0 := \sqrt{\frac{m\omega}{2}}Q. \tag{6.3.0.4}$$

Now $H = \omega(P_0^2 + Q_0^2)$. Note that

$$[Q_0, P_0] = \frac{1}{2}i. \tag{6.3.0.5}$$

Thus

$$(Q_0 - iP_0)(Q_0 + iP_0) = Q_0^2 + P_0^2 - \frac{1}{2}I. \tag{6.3.0.6}$$

Then if we define

$$A^\dagger = Q_0 - iP_0, \tag{6.3.0.7}$$

and

$$A = Q_0 + iP_0. \tag{6.3.0.8}$$

Now, we can write H in the following form

$$H = \omega(A^\dagger A + \frac{1}{2}). \tag{6.3.0.9}$$

Let $N = A^\dagger A$. We will establish first some facts about $A^\dagger$ and A.

We will write now $A^\dagger$ and A explicitly.

$$A^\dagger\psi = \sqrt{\frac{m\omega}{2}}x\psi - \frac{1}{\sqrt{2m\omega}}\frac{\partial\psi}{\partial x}. \tag{6.3.0.10}$$

$$A\psi = \sqrt{\frac{m\omega}{2}}x\psi + \frac{1}{\sqrt{2m\omega}}\frac{\partial\psi}{\partial x}. \tag{6.3.0.11}$$

Fact 6.3.0.1. *If $\psi \in L^2(\mathbb{R})$, ψ is differentiable and satisfies the equation*

$$A\psi = 0, \tag{6.3.0.12}$$

then there exists $C \in \mathbb{C}$ such that

$$\psi(x) = C\exp\left(-\frac{m\omega x^2}{2}\right). \tag{6.3.0.13}$$

Proof. Equation (6.3.0.12) is equivalent to

$$\frac{\partial\psi}{\partial x} = -m\omega x\psi. \tag{6.3.0.14}$$

It can be proven that all solutions of the above have a form $\psi(x) = C\exp\left(-\frac{m\omega x^2}{2}\right)$. Since $\psi \in L^2(\mathbb{R})$ it is a valid solution of (6.3.0.12). □

Fact 6.3.0.2. *If $\psi \in L(\mathbb{R})^2$, ψ is differentiable $A^\dagger\psi = 0$, then $\psi = 0$.*

Proof. Here $A\psi = 0$ is equivalent to

$$\frac{\partial\psi}{\partial x} = m\omega x\psi. \tag{6.3.0.15}$$

So the solution has a form $\psi(x) = C\exp\left(\frac{m\omega x^2}{2}\right)$, where $C \in \mathbb{C}$. But we must have $\psi \in L(\mathbb{R})^2$, thus the only possible solution is $\psi = 0$. □

Let's now check commutator $[A, A^\dagger]$.

$$A^\dagger A = Q_0^2 + P_0^2 - \frac{1}{2}. \tag{6.3.0.16}$$

$$AA^\dagger = Q_0^2 + P_0^2 + \frac{1}{2}. \tag{6.3.0.17}$$

Thus

$$[A, A^\dagger] = 1. \tag{6.3.0.18}$$

Fact 6.3.0.3. *If $\psi, \phi \in S_1$, then*

$$\langle A\psi, \phi\rangle = \langle \psi, A^\dagger \phi\rangle \tag{6.3.0.19}$$

and

$$\langle N\psi, \phi\rangle = \langle \psi, N \phi\rangle . \tag{6.3.0.20}$$

Fact 6.3.0.4. *If γ is an eigenvalue of N with the corresponding eigenvector $\psi \in S_1$, then $\gamma \geq 0$ and $\gamma + 1$ is an eigenvalue of N with the corresponding eigenvector $A^\dagger\psi$.*

Proof. By (6.3.0.20), we have $\gamma \geq 0$. By (6.3.0.18) we can calculate the following

$$NA^\dagger\psi = A^\dagger AA^\dagger\psi = A^\dagger(1 + A^\dagger A)\psi = A^\dagger\psi + A^\dagger N\psi = A^\dagger\psi + \gamma A^\dagger\psi = (\gamma + 1)A^\dagger\psi. \tag{6.3.0.21}$$

By Fact 6.3.0.2, we know that $A^\dagger\psi \neq 0$, thus $A^\dagger\psi$ is an eigenvector. □

Fact 6.3.0.5. *If $\gamma > 0$ is an eigenvalue of N with the corresponding eigenvector $\psi \in S_1$, then $\gamma - 1$ is an eigenvalue of N with the corresponding eigenvector $A\psi$.*

Proof. By (6.3.0.18), we can calculate the following

$$NA\psi = A^\dagger AA\psi = (AA^\dagger - 1)A\psi = AN\psi - A\psi = A\gamma\psi - A\psi = (\gamma - 1)A\psi. \tag{6.3.0.22}$$

Now we need to show that $A\psi \neq 0$. Indeed, if $A\psi = 0$, then $N\psi = 0$, but since ψ is an eigenvector with eigenvalue $\gamma > 0$, this is a contradiction. □

6.4 Dirac Formulation of Quantum Mechanics

Bra-ket notation is described in first chapters of [**?**].

$$\left(x^\dagger\right)^\dagger = x\,. \tag{6.4.0.1}$$

$$\big(c_1|\psi_1\rangle + c_2|\psi_2\rangle\big)^\dagger = c_1^*\langle\psi_1| + c_2^*\langle\psi_2|\,. \tag{6.4.0.2}$$

$$|\phi\rangle = A|\psi\rangle \quad \text{if and only if} \quad \langle\phi| = \langle\psi|A^\dagger\,. \tag{6.4.0.3}$$

$$\langle\phi|\big(c_1|\psi_1\rangle + c_2|\psi_2\rangle\big) = c_1\langle\phi|\psi_1\rangle + c_2\langle\phi|\psi_2\rangle\,. \tag{6.4.0.4}$$

6.4.1 Momentum and Position operators

Let us consider states of one non-relativistic particle for which we assume that can be fully described as superpositions of locations in R^3. By $|x\rangle$ we will denote a quantum state which describe a particle being "around" point $x \in \mathbb{R}^3$. We understand "around" in Dirac delta sense

$$\langle x_1 | x_2 \rangle = \delta(x_2 - x_1). \tag{6.4.1.1}$$

Dirac delta provides nice normalisation where

$$\int |x\rangle \langle x|z\rangle \, dx = \int |x\rangle \, \delta(z - x) dz = |z\rangle \,, \tag{6.4.1.2}$$

as expected.

Similarity by $|p\rangle$ we denote a quantum state, which describes a quantum state of particle which has a momentum "around" $p \in \mathbb{R}^3$ (as you can notice usage of letter p is important for this convention, because only by usage of the letter we distinguish this from position representation).

Now, it seems to be a law of the nature that

$$\boxed{\langle x|p\rangle = (2\pi\hbar)^{-\frac{3}{2}} \exp\left(-i\frac{p \cdot x}{\hbar}\right)} \tag{6.4.1.3}$$

Where $(2\pi\hbar)^{-\frac{3}{2}}$ is a normalisation factor needed to have

$$\langle p_1 | p_2 \rangle = \delta(p_2 - p_1). \tag{6.4.1.4}$$

Indeed, first note that (as proved for for (13.15.1.2)):

$$\mathcal{F}(1)(x) = (2\pi)^{\frac{n}{2}} \delta(x). \tag{6.4.1.5}$$

in the context of distributions.

Now let's do calculations integrating in the sense of distributions:

$$\langle p_1|p_2\rangle = \int \langle p_1|x\rangle \langle x|p_2\rangle \, dx =$$

$$(2\pi\hbar)^{-3}\int \exp\left(-i\frac{(p_2-p_1)x}{\hbar}\right)dx \underset{\hbar z\to x}{=} (2\pi\hbar)^{-3}\int \exp(-i(p_2-p_1)z)|\det\frac{\partial x}{\partial z}|dz$$
$$= (2\pi)^{-\frac{3}{2}}\hbar^{-3}\hbar^3\int (2\pi)^{-\frac{3}{2}}\exp(-i(p_2-p_1)z)dz = (2\pi)^{-\frac{3}{2}}\mathcal{F}(1)(p_2-p_1)$$
$$= (2\pi)^{-\frac{3}{2}}(2\pi)^{\frac{3}{2}}\delta(p_2-p_1) = \delta(p_2-p_1).$$

6.4.2 Some properties of Dirac δ

Fact 6.4.2.1. *If δ is Dirac delta and $x \in \mathbb{R}^n$, then*

$$\delta(f(x)) = \sum_i \frac{1}{|\det(\nabla f)(x_i)|}\delta(x-x_i), \tag{6.4.2.1}$$

where x_i are all discrete zeros of a real function f.

Proof. Let y depends on z in a following way $y = f(z)$.

$$\delta(f(x)) = \int \delta(f(z))\delta(z-x)dz = \sum_i \int\limits_{O_\varepsilon(x_i)} \delta(f(z))\delta(z-x)dz$$
$$= \sum_i \int\limits_{f(O_\varepsilon(x_i))} \delta(y)\delta(z-x)|\det\frac{\partial z}{\partial y}|dy = \sum_i \int\limits_{f(O_\varepsilon(x_i))} \delta(y)\delta(z-x)|\frac{1}{\det\frac{\partial y}{\partial z}}|dy$$
$$= \sum_i \frac{1}{|\det(\nabla f)(x_i)|}\delta(x_i-x),$$

where the last equality holds because on a $f(O_\varepsilon(x_i))$ variable z is a function of y and for $y = 0$ we have $z = x_i$. $\square$

Corollary 6.4.2.2. *If δ is Dirac delta and $x \in \mathbb{R}$, then*

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x - x_i), \tag{6.4.2.2}$$

where x_i are all discrete zeros of a real function f.

6.4.3 Representations

Short and incomplete digest from [**?**][III. Representations]

We will use all the time symbol δ but it depends on measure if it is Dirac delta or Kronecker delta. Let $|\alpha\rangle$ be states of orthogonal basis normed in the following way $\langle\alpha|\beta\rangle = 0$ for $\alpha \neq \beta$ and $\int \langle\alpha|\beta\rangle \, d\alpha = 1$.

Definition 6.4.3.1. *Let $A_1, \ldots, A_k$ be set of commuting observables for which all basis states are eigenstates and f is a real valued function. Then we define observable*

$$f(A_1, \ldots, A_k) |\alpha\rangle = f(\lambda_1, \ldots, \lambda_n) |\alpha\rangle, \tag{6.4.3.1}$$

where $A_i |\alpha\rangle = \lambda_i |\alpha\rangle$.

Definition 6.4.3.2. *Commuting observables $A_1, \ldots, A_k$ with the same eigenstates basis are called independent, if and only if for any observable A_i there is no real function f such that $A_i = f(A_1, \cdots, A_{i-1}, A_{i+1}, \cdots A_k)$.*

Definition 6.4.3.3. *Commuting observables $A_1, \ldots, A_k$ with the same eigenstates basis are called complete, if for any observable B with the same eigenstates basis*

$$B = f(A_1, \ldots, A_k). \tag{6.4.3.2}$$

Fact 6.4.3.4. *If $A_1, \ldots, A_n$ are commuting and complete observables with the same eigenstates basis then any basis states can be uniquely represented as*

$$|\alpha\rangle = |\lambda_1, \ldots, \lambda_n\rangle, \tag{6.4.3.3}$$

where $A_i |\alpha\rangle = \lambda_i |\alpha\rangle$.

Proof. Assume to the contrary that we have the whole subspace of states $|\beta\rangle$ such that $A_i |\beta\rangle = \lambda_i' |\beta\rangle$ for fixed λ_i'. Then the basis state $|\alpha\rangle$ will be represented as

$$|\alpha\rangle = |\lambda_1, \ldots, \lambda_n, \theta_1, \ldots, \theta_k\rangle. \tag{6.4.3.4}$$

Let's define $L_i |\alpha\rangle = \theta_i |\alpha\rangle$. But because $A_1, \ldots, A_n$ are complete $L_i = f_i(A_1, \ldots, A_n)$ and thus $\theta_i = f_i(\lambda_1, \cdots, \lambda_n)$. Therefore unique set of $\lambda_1, \ldots, \lambda_n$ defines unique basis vector $|\alpha\rangle$. □

The above representation can be normalised in the following way

$$\langle \lambda_1', \ldots, \lambda_n' | \lambda_1, \ldots, \lambda_n \rangle = \delta(\lambda_1' - \lambda_1) \ldots \delta(\lambda_n' - \lambda_n). \tag{6.4.3.5}$$

6.4.4 Unitary transformations

Assume $|\alpha\rangle$ is a normalised basis where α is a vector of eigenvalues for the system of complete commuting observables. Two different basis states $|\alpha\rangle$ and $|\beta\rangle$ must be orthogonal, because at least one eigenvalue needs to be different. We will additionally assume they are normalised $\langle \alpha|\beta\rangle = \delta(\alpha - \beta)$.

Assume we have an isometry u on a space of these vectors. Then

Consider transformation of quantum states

$$U|\alpha\rangle = |u(\alpha)\rangle. \tag{6.4.4.1}$$

Because U transforms basis states to basis states it can be extended to linear operator on the whole states space.

We will show that U is unitary (i.e. $U^\dagger U = U^\dagger U = I$.)

$$\langle \beta| U^\dagger U |\alpha\rangle = \langle u(\beta)|u(\alpha)\rangle = \delta(u(\beta) - u(\alpha)) = \delta(\beta - \alpha) = \langle \beta|\alpha\rangle.$$

The equality $\delta(u(\alpha) - u(\beta)) = \delta(\alpha - \beta)$ holds because u is an isometry (i.e. $|\det(\nabla u)| = 1$) and by Fact 6.4.2.1.

In general case where u is just an arbitrary $1-1$ and "onto" map from the representation space of all eigenvalues vectors for the system of complete observables into the other representation space of all eigenvalues of vectors for the system of complete observables (we don't even require the kets will be from the same states space) the transformation given by

$$U|\alpha\rangle = |\det(\nabla u(\alpha))|^{1/2} |u(\alpha)\rangle \tag{6.4.4.2}$$

is unitary provided both $|\alpha\rangle$ and $|u(\alpha)\rangle$ are normalised.

The other fact which is worth to note is that if we transform space of states with an arbitrary unitary transformation U as

$$|\alpha\rangle \mapsto U|\alpha\rangle, \tag{6.4.4.3}$$

then any linear operator A transforms in a way

$$A \mapsto UAU^\dagger. \tag{6.4.4.4}$$

Indeed, it is enough to notice

$$\langle \beta| U^\dagger (UAU^\dagger) U |\alpha\rangle = \langle \beta| A |\alpha\rangle. \tag{6.4.4.5}$$

6.4.5 Wave Packet

Let us consider a quantum state

$$|\phi\rangle = \int_{\mathbb{R}^3} \frac{A}{(\Delta p)^3} e^{-\frac{1}{2(\Delta p)^2}(p-p_0)^2} e^{\frac{i}{\hbar}x_0 p} |p\rangle \, dp, \tag{6.4.5.1}$$

where factor A is to be established later.

Let's calculate amplitude

$$\langle x|\phi\rangle = \int_{\mathbb{R}^3} A \frac{(2\pi\hbar)^{-\frac{3}{2}}}{(\Delta p)^3} e^{-\frac{1}{2(\Delta p)^2}(p-p_0)^2} e^{-\frac{i}{\hbar}(x-x_0)p} dp. \tag{6.4.5.2}$$

To calculate apply to this equation (11.4.1.2). Recall:

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{izy} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} dz = \exp\left(i\mu y - \frac{\sigma^2 y^2}{2}\right). \tag{6.4.5.3}$$

Convert this to $\mathbb{R}^3$ case, using Fubini's Theorem and property of multiplying values when adding arguments for exp.

$$\int_{\mathbb{R}^3} \frac{(2\pi)^{-\frac{3}{2}}}{\sigma^3} e^{izy} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} dz = \exp\left(i\mu y - \frac{\sigma^2 y^2}{2}\right). \tag{6.4.5.4}$$

Let's do substitutions

$$\begin{aligned} p &\to z, \\ p_0 &\to \mu, \\ \Delta p &\to \sigma, \\ -\frac{x-x_0}{\hbar} &\to y. \end{aligned}$$

Thus

$$\langle x|\phi\rangle = A\hbar^{-\frac{3}{2}} \exp\left(-i\frac{p_0(x-x_0)}{\hbar} - \frac{(\Delta p)^2(x-x_0)^2}{2\hbar^2}\right). \tag{6.4.5.5}$$

Then

$$|\langle x|\phi\rangle|^2 = A^2\hbar^{-3} \exp\left(-\frac{(\Delta p)^2(x-x_0)^2}{\hbar^2}\right). \tag{6.4.5.6}$$

To get $\int |\langle x|\phi\rangle|^2 dx = 1$, we need to set $A = (\Delta p)^{\frac{3}{2}}\pi^{-\frac{3}{4}}$. You can also check that with this choice of A, we have $|\langle\phi|\phi\rangle|^2 = 1$.

Let's summarize

$$|\phi\rangle = \int_{\mathbb{R}^3} \frac{\pi^{-\frac{3}{4}}}{(\Delta p)^{3/2}} e^{-\frac{1}{2(\Delta p)^2}(p-p_0)^2} e^{\frac{i}{\hbar}x_0 p} |p\rangle \, dp, \tag{6.4.5.7}$$

$$\langle x|\phi\rangle = \pi^{-\frac{3}{4}}(\frac{\Delta p}{\hbar})^{\frac{3}{2}} \exp\left(-i\frac{p_0(x-x_0)}{\hbar}\right) \exp\left(-\frac{(\Delta p)^2(x-x_0)^2}{2\hbar^2}\right). \tag{6.4.5.8}$$

6.5 Schrödinger and Heisenber equations of motion

6.5.1 Time independent Hamiltonian

Suppose we have equation of quantum state evolution

$$i\hbar\frac{\partial}{\partial t}|\phi\rangle = H|\phi\rangle. \tag{6.5.1.1}$$

Let's consider it in units where $\hbar = 1$. Note that here $|\phi\rangle$ depends on time. Assume that H does not depend on time. Then we have

$$|\phi(t)\rangle = e^{-itH}|\phi(0)\rangle. \tag{6.5.1.2}$$

Assume we have an observable A_S which doesn't depend on time.

We may now propose a process of measurement, which goes as follows. We start with state $|\phi(0\rangle$ let it evolve with Hamiltonian H for time t and then measure observable A_S. One may claim that this is a process of measuring some quantity for the state $|\phi(0)\rangle$, thus we should have observable A_H which acts on $|\phi(0)\rangle$ related to this measurement process. Assume that in described process of measurement, we got value λ. That means that when we measured A_S against $e^{-itH}|\phi(0)\rangle$ we got state $|\psi\rangle$ and $A_S|\psi\rangle = \lambda|\psi\rangle$. Now, we want an observable A_H which will be giving us the same results of state $|\phi(0)\rangle$.

The natural candidate is

$$A_H = e^{itH}A_S e^{-itH}. \tag{6.5.1.3}$$

When we go back in time of time t and take $|\psi_0\rangle = e^{itH}|\psi\rangle$, we got $A_H|\psi_0\rangle = \lambda|\psi_0\rangle$, which shows that A_H acting on $|\phi(0)\rangle$ has exactly the same spectrum as A_S acting on $|\phi(t)\rangle$.

Let's see an evolution of A_H in time.

$$\frac{\partial A_H}{\partial t} = i(He^{itH}A_S e^{-itH} - e^{itH}A_S He^{-itH})$$

$$= i(He^{itH}A_S e^{-itH} - e^{itH}A_S e^{-itH}H) = i[H, A_H].$$

Thus

$$\boxed{\frac{\partial A_H}{\partial t} = i[H, A_H]} \tag{6.5.1.4}$$

Which is a Heisenberg equation of observable evolution. With reinstated $\hbar$ it is as follows:

$$\boxed{\frac{\partial A_H}{\partial t} = \frac{i}{\hbar}[H, A_H]} \tag{6.5.1.5}$$

6.5.2 Time dependent Hamiltonian

Suppose we have equation of quantum state evolution

$$i\hbar\frac{\partial}{\partial t}|\phi\rangle = H|\phi\rangle\,. \tag{6.5.2.1}$$

Let's consider it in units where $\hbar = 1$. Note that here $|\phi\rangle$ depends on time and also H might depend on time.

Let

$$U(t_1, t_2)\,|\phi_0\rangle := |\phi(t_2)\rangle\,. \tag{6.5.2.2}$$

where ϕ is a solution of differential equation (6.5.2.1) for a boundary condition $|\phi(t_1)\rangle = |\phi_0\rangle$. Linearity of $U(t_1, t_2)$ is obvious as (6.5.2.1) is a linear equation. It is also straightforward to see that

$$U(t_1, t_3) = U(t_2, t_3)U(t_1, t_2) \tag{6.5.2.3}$$

and

$$U(t, t) = I. \tag{6.5.2.4}$$

We will show that U is unitary. Assume that ϕ and ψ are solutions of (6.5.2.1). Consider

$$\frac{\partial}{\partial t}\langle\phi(t)|\psi(t)\rangle = \{\frac{\partial}{\partial t}\langle\phi(t)|\}\,|\psi(t)\rangle + \langle\phi(t)|\,\{\frac{\partial}{\partial t}|\psi(t)\rangle\}$$

$$= \langle\phi(t)|\,(-iH)^\dagger\,|\psi(t)\rangle + \langle\phi(t)| - iH\,|\psi(t)\rangle$$
$$= \langle\phi(t)|\,iH\,|\psi(t)\rangle + \langle\phi(t)| - iH\,|\psi(t)\rangle = 0.$$

Thus

$$\langle\phi_0|\psi_0\rangle = \langle\phi_0|\,U^\dagger(t_1,t_2)U(t_1,t_2)\,|\psi_0\rangle\,. \qquad (6.5.2.5)$$

Hence $U^\dagger(t_1,t_2)U(t_1,t_2) = I$. From that, it is easy to show that $U^\dagger(t_1,t_2) = U(t_2,t_1)$, thus also $U(t_1,t_2)U^\dagger(t_1,t_2) = I$.

Note that from (6.5.2.1) we have

$$\frac{\partial}{\partial t}U(t_0,t) = -iH(t)U(t_0,t). \qquad (6.5.2.6)$$

and

$$\frac{\partial}{\partial t}U^\dagger(t_0,t) = iU^\dagger(t_0,t)H(t). \qquad (6.5.2.7)$$

Assume we have an observable A_S acting on $|\phi(t)\rangle$. Now if we want to replace it by observable A_H giving the same measurements but acting on $|\phi(0)\rangle$ by similar argument than for (6.5.1.3) we get

$$A_H = U^\dagger(0,t)A_SU(0,t). \qquad (6.5.2.8)$$

and similarly

$$H_H(t) = U^\dagger(0,t)H(t)U(0,t). \qquad (6.5.2.9)$$

Observe now, how A_H evolves in time

$$\begin{aligned}\frac{\partial A_H}{\partial t} &= iU^\dagger(t_0,t)H(t)A_S U(0,t) - iU^\dagger(t_0,t)A_H H(t)U(t_0,t)\\ &= iU^\dagger(t_0,t)H(t)U(0,t)U^\dagger(0,t)A_S U(0,t)\\ &- iU^\dagger(t_0,t)A_S U(0,t)U^\dagger(0,t)H(t)U(t_0,t)\\ &= i(H_H(t)A_H(t) - A_H H_H(t)) = i[H_H(t), A_H].\end{aligned}$$

Thus

$$\boxed{\frac{\partial A_H}{\partial t} = i[H_H(t), A_H]} \tag{6.5.2.10}$$

With reinstated $\hbar$:

$$\boxed{\frac{\partial A_H}{\partial t} = \frac{i}{\hbar}[H_H(t), A_H]} \tag{6.5.2.11}$$

Note that in case when $H = \text{const}$, since $U(0,t) = e^{-itH}$ and since it commutes with H, from (6.5.2.9), we have $H_H = H$. Thus equation (6.5.1.4) is just a particular case of (6.5.2.10).

6.6 Uncertainty principle

Lemma 6.6.0.1. *For any two observables A and B we have*

$$\langle\phi|A^2|\phi\rangle^{\frac{1}{2}}\ \langle\phi|B^2|\phi\rangle^{\frac{1}{2}} \geq \frac{1}{2}|\,\langle\phi|[A,B]|\phi\rangle\,|. \tag{6.6.0.1}$$

Morover, the equality holds iff there exist $\lambda_1, \lambda_2 \in \mathbb{R}$, not both zero, such that

$$i\lambda_1 A\phi + \lambda_2 B\phi = 0. \tag{6.6.0.2}$$

Proof. By Schwarz inequality we have

$$\langle\phi|A^2|\phi\rangle^{\frac{1}{2}}\ \langle\phi|B^2|\phi\rangle^{\frac{1}{2}} \geq |\,\langle\phi|AB|\phi\rangle\,|, \tag{6.6.0.3}$$

on the other hand

$$\langle\phi|A^2|\phi\rangle^{\frac{1}{2}}\ \langle\phi|B^2|\phi\rangle^{\frac{1}{2}} \geq |\,\langle\phi|BA|\phi\rangle\,|. \tag{6.6.0.4}$$

Thus

$$\langle\phi|A^2|\phi\rangle^{\frac{1}{2}}\ \langle\phi|B^2|\phi\rangle^{\frac{1}{2}} \geq \frac{1}{2}\Big(|\langle\phi|AB|\phi\rangle| + |\langle\phi|BA|\phi\rangle|\Big) \geq$$
$$\frac{1}{2}|\langle\phi|AB|\phi\rangle - \langle\phi|BA|\phi\rangle| = \frac{1}{2}|\langle\phi|[A,B]|\phi\rangle|. \quad (6.6.0.5)$$

We have proven inequality (6.6.0.1). We will prove moreover part. Assume that we have equality in (6.6.0.1). If we have equality in (6.6.0.1), then we must also have equality in (6.6.0.3) and (6.6.0.3). From the properties of Schwarz inequality this means that Ax and Bx are lineary dependant. If $A\phi = 0$ then obviously $A\phi$ and $B\phi$ are lineary dependant and the equation (6.6.0.2) holds for $\lambda_1 = 1$ and $\lambda_2 = 0$. Assume then that $A\phi \neq 0$. Thus, we have $\lambda \in \mathbb{C}$ such that $B\phi = \lambda A\phi$. Since we have equality in (6.6.0.1), we must have also

$$|\langle\phi|AB|\phi\rangle| + |\langle\phi|BA|\phi\rangle| = |\langle\phi|AB|\phi\rangle - \langle\phi|BA|\phi\rangle|. \quad (6.6.0.6)$$

Which implies the following.

$$|\langle\phi|A^2\lambda|\phi\rangle| + |\langle\phi|\overline{\lambda}A^2|\phi\rangle| = |\langle\phi|A^2\lambda|\phi\rangle - \langle\phi|\overline{\lambda}A^2|\phi\rangle|. \quad (6.6.0.7)$$

Hence $2|\lambda| = |\lambda - \overline{\lambda}|$, which is equivalent $|\lambda| = |\operatorname{Im}\lambda|$. Thus $\operatorname{Re}\lambda = 0$ and the morover part is proved. □

Definition 6.6.0.2. *Let A be an observable and ϕ be a state of the system.*

$$\langle A\rangle_\phi := \langle\phi|A|\phi\rangle\,. \quad (6.6.0.8)$$

We can interpret physically $\langle A\rangle_\phi$ in the following way. When in some way we create copies of the state ϕ and measure the value of observable A as average value of the measurments in limit we get $\langle A\rangle_\phi$.

Definition 6.6.0.3. *Let A be an observable and ϕ be a state of the system.*

$$\sigma_\phi(A) := \langle\phi|(A - \langle A\rangle_\phi)^2|\phi\rangle^{\frac{1}{2}}\,. \quad (6.6.0.9)$$

The above is the standard deviation of observable A if we measure it in copies of state ϕ.

Theorem 6.6.0.4. *For any two observables A and B we have and state ϕ, we have*

$$\sigma_\phi(A)\sigma_\phi(B) \geq \frac{1}{2}|\langle\phi|[A,B]|\phi\rangle|. \quad (6.6.0.10)$$

Corollary 6.6.0.5. *Let Q be a position operator and P be a momentum operator and ϕ be a state, then*

$$\sigma_\phi(Q)\sigma_\phi(P) \geq \frac{\hbar}{2}. \tag{6.6.0.11}$$

6.7 Hydrogen Atom

In this section we will restore SI units, as it will be more convenient to make references to measurable quantities.

6.7.1 Angular Momentum

Recall that for "nice enough" $\psi : \mathbb{R}^3 \to \mathbb{C}$, we define momentum operators along each axis as

$$P_i\psi = -i\hbar\frac{\partial}{\partial x^i}\psi \text{ for } i = 1, 2, 3, \tag{6.7.1.1}$$

and position operator as

$$Q_i\psi = x^i\psi \text{ for } i = 1, 2, 3. \tag{6.7.1.2}$$

Let's define the angular momentum operator

$$\begin{aligned} L_1 &= Q_2P_3 - Q_3P_2, \\ L_2 &= Q_3P_1 - Q_1P_3, \\ L_3 &= Q_1P_2 - Q_2P_2. \end{aligned} \tag{6.7.1.3}$$

We will use interchangeably x^1, x^2, x^3 convention with x, y, z.

$$\begin{aligned} L_x &= -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\ L_y &= -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\ L_z &= -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \end{aligned} \tag{6.7.1.4}$$

Let
$$L^2 := L_1^2 + L_2^2 + L_3^2. \tag{6.7.1.5}$$
If we denote symbolically $\vec{L} := [L_1, L_2, L_3]$ (it is not really a vector – just a convention). And if we treat composition of operators as multiplication in a standard definition of vector cross product, we can nicely formulate commutations rules as
$$\vec{L} \times \vec{L} = i\hbar\vec{L}. \tag{6.7.1.6}$$
The are proved e.g in [see **?** , 4.8 Angular-Momentum Operators]. For convenience assume spherical coordinates:

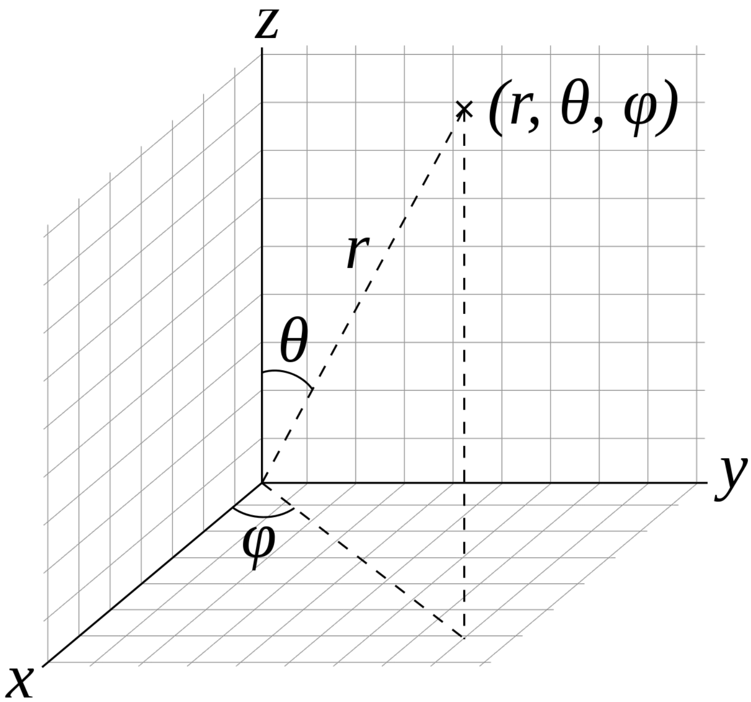

Figure 6.1: Graphical demonstration of spherical coordinates.

$$\begin{cases} x = r\sin\theta\cos\phi, \\ y = r\sin\theta\sin\phi, \\ z = r\cos\theta. \end{cases} \tag{6.7.1.7}$$
It can be shown that
$$\boxed{L_z = -i\hbar\frac{\partial}{\partial\phi}.} \tag{6.7.1.8}$$

Indeed, $\frac{\partial}{\partial\phi} = \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} + 0.$

Let's define

$$\Delta_{\phi,\theta} := \frac{1}{\sin^2\theta}\cdot\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\cdot\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial}{\partial\theta}\right). \tag{6.7.1.9}$$

It can be also shown that

$$\boxed{L^2 = -\hbar\Delta_{\phi,\theta}.} \tag{6.7.1.10}$$

Interestingly, it is well known that:

$$\Delta = \frac{1}{r^2}\left(\Delta_{\phi,\theta} + \frac{\partial}{\partial r}\left(r^2\cdot\frac{\partial}{\partial r}\right)\right). \tag{6.7.1.11}$$

From (6.7.1.8) and (6.7.1.10) it is apparent that $[L^2, L_z] = 0$. Thus L_z and L^2 can be measured simultaneously.

Let's recall associated Legendre polynomials:

$$P_{l,m}(x) := \frac{(-1)^m}{2^l l!}\left(1-x^2\right)^{m/2}\frac{\partial^{l+m}\left(x^2-1\right)^l}{\partial x^{l+m}} \tag{6.7.1.12}$$

for $l = 0, 1, \ldots$ and $m = -l, \ldots, l$. Now we can define

$$Y_{l,m}(\theta,\phi) := \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\exp(im\phi)P_{l,m}(\cos(\theta)). \tag{6.7.1.13}$$

They are eigenvectors of both L^2 and L_z with the eigenvalues as in equations below.

$$\boxed{\begin{cases} L_z Y_{l,m} = m\hbar Y_{l,m}, \\ L^2 Y_{l,m} = l(l+1)\hbar^2 Y_{l,m} \end{cases}} \tag{6.7.1.14}$$

for $l = 0, 1, \ldots$ and $m = -l, \ldots, l$.

6.7.2 Motivation for potential operator

We will assume that potential observable V measures potential $V(x)$ (we overload V symbol here) in a space point x as an eigenvalue of a Dirac delta at point x (i.e. $\delta_x(z) := \delta(x-z)$).

$$V\delta_x = V(x)\delta_x. \tag{6.7.2.1}$$

Let's calculate $V(\psi)$.

$$(V\psi)(x) = \int \delta_x(z)(V\psi)(z)dz = \int (V\delta_x)(z)\psi(z)dz = \int V(z)\cdot\delta_x(z)\psi(z)dz = V(x)\cdot\psi. \tag{6.7.2.2}$$

Second transformation above holds because V is self-adjoint.

6.7.3 Proton-electron system

Let's consider a system of two particles, one with negative elementary charge $-e$ and mass m_e and the other with positive elementary e and mass m_p.

To investigate probability amplitude of the relative position of the two particles, according to considerations in Subsection 13.10.1, we need to get the following hamiltonian:

$$H\psi = \sum_{i=1}^{3} \frac{P_i^2}{2\mu}\psi + V \cdot \psi, \tag{6.7.3.1}$$

where μ is a reduced mass, namely $\mu = \dfrac{m_e m_p}{m_e + m_p}$ and V is an potential energy of the system of two charges governed by Coulomb force:

$$V = -\frac{e^2}{4\pi\varepsilon_0 r}. \tag{6.7.3.2}$$

For completeness let's check if $-\dfrac{\partial V}{\partial x}$ gives us Coulomb force. Let's assume $x = [x_1, x_2, x_3]$ is a radial vector of electron and that proton is in the center of our frame of reference. Obvioulsy $r := (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}}$. Now

$$F_i = -\frac{\partial V}{\partial x_i} = -\big(-\frac{1}{2}\cdot 2x_i \cdot (-\frac{e^2}{4\pi\varepsilon_0}(x_1^2+x_2^2+x_3^2)^{-\frac{3}{2}})\big) = -x_i\frac{e^2}{4\pi\varepsilon_0 r^3}. \tag{6.7.3.3}$$

Thus

$$\vec{F} = \frac{e^2}{4\pi\varepsilon_0 r^3}x, \tag{6.7.3.4}$$

which is exactly Coulomb force. Let's then put our Hamiltonian in an abreviated form:

$$H = -\frac{\hbar^2}{2\mu}\Delta - \frac{e^2}{4\pi\varepsilon_0 r}. \tag{6.7.3.5}$$

There exist eigen vectors $\psi_{n,l,m}$ of H such that

$$\begin{cases} H\psi_{n,l,m} = -\dfrac{e^4\mu}{32\pi^2 n^2 \varepsilon_0^2 \hbar^2}\psi_{n,l,m}, \\ L^2\psi_{n,l,m} = l(l+1)\hbar^2\psi_{n,l,m} \\ L_z\psi_{n,l,m} = m\hbar\psi_{n,l,m}, \end{cases} \tag{6.7.3.6}$$

for $n = 1, 2, \ldots,$ $l = 0, 1, 2, \ldots$ and $m = -l, \ldots, l$. The value

$$\boxed{E_n = -\frac{e^4\mu}{32\pi^2 n^2 \varepsilon_0^2 \hbar^2}} \tag{6.7.3.7}$$

is an energy level of hydrogen for $n = 1, 2, \ldots$.

6.8 Perturbation Theory

6.8.1 Stationary Perturbation

Let's assume our hamiltonian has a form

$$H = H_0 + \varepsilon W, \tag{6.8.1.1}$$

where complete set of eignestates and eigenvalues of unperturbed hamiltonian H_0 is known

$$H_0 \psi_n^0 = E_n^0 \psi_n^0. \tag{6.8.1.2}$$

We assume also that energy levels are not degenerated, i.e. $E_i \neq E_j$ for $i \neq j$. The goal of perturbation theory is to determine solutions of

$$H\psi = E\psi, \tag{6.8.1.3}$$

as

$$E = E_k = \sum_n E_k^{(n)} \varepsilon^n \tag{6.8.1.4}$$

$$\psi = \psi_k = \sum_m a_{m,k} \psi_m^0, \tag{6.8.1.5}$$

where

$$a_{m,k} = \sum_{n=0}^{\infty} a_{m,k}^{(n)} \varepsilon^n. \tag{6.8.1.6}$$

It can be proven (e.g. [**?**][11.1]) that

$$\begin{cases} E_k^{(0)} = E_k^0, \\ E_k^{(1)} = \langle \psi_k^0 | W | \psi_k^0 \rangle, \\ E_k^{(2)} = \sum_{n \neq k} \dfrac{| \langle \psi_n^0 | W | \psi_k^0 \rangle |^2}{E_n^0 - E_k^0}, \\ \ldots \end{cases} \tag{6.8.1.7}$$

and

$$\begin{cases} a_{m,k}^{(0)} = \delta_{m,k}, \\ \ldots \end{cases} \tag{6.8.1.8}$$

6.8.2 Degeneracy

Assume that for a given level of energy E_n^0 in unperturbed hamiltonian H_0, we have a series of eigenfunctions $\psi_{n,\beta}^0$ where $\beta = 1, 2, ..., f_n$. Such level is called f_n-fold degenerate.

Solutions $E_{n,\beta}^{(1)}$ of the below equation (6.8.2.1) with unknown variable $E_n^{(1)}$ determin the split of energy level E_n^0 into levels $E_{n,\beta} \approx E_n^0 + \varepsilon E_{n,\beta}^{(1)}$ in perturbed Hamiltonian $H_0 + \varepsilon W$.

$$\det \begin{bmatrix} \langle \psi_{n,1}^0 | W | \psi_{n,1}^0 \rangle - E_n^{(1)} & \langle \psi_{n,1}^0 | W | \psi_{n,2}^0 \rangle & \dots & \langle \psi_{n,1}^0 | W | \psi_{n,f_n}^0 \rangle \\ \langle \psi_{n,2}^0 | W | \psi_{n,1}^0 \rangle & \langle \psi_{n,2}^0 | W | \psi_{n,2}^0 \rangle - E_n^{(1)} & \dots & \langle \psi_{n,2}^0 | W | \psi_{n,f_n}^0 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \psi_{n,f_n}^0 | W | \psi_{n,1}^0 \rangle & \langle \psi_{n,f_n}^0 | W | \psi_{n,2}^0 \rangle & \dots & \langle \psi_{n,f_n}^0 | W | \psi_{n,f_n}^0 \rangle - E_n^{(1)} \end{bmatrix} = 0. \tag{6.8.2.1}$$

The above is proven in (e.g. [**?**][11.2]). To get above equation from matrix equation there, one needs to notice that $E^0 - E = -\varepsilon E^{(1)}$ and then remove factor ε from the matrix. The equation in form similiar to (6.8.2.1) is also in [**?**][16.4 Degenerate States].

6.8.3 Time-Dependent Perturbation

We assume that the hamiltonian is given by

$$H = H_0 + W(t) \tag{6.8.3.1}$$

We assume that $W(t)$ is a small perturbation and acts only in time interval $[0, T]$. Let ψ_n^0 be stationary solutions of

$$H_0 \psi_n^0 = E_n \psi_n^0. \tag{6.8.3.2}$$

Time evolution is given by

$$\psi_n^0(t) = e^{-\frac{i}{\hbar} E_n t} \psi_n^0. \tag{6.8.3.3}$$

We will be interested in an evolution

$$i\hbar \frac{\partial \psi_n(t)}{\partial t} = H \psi_n(t), \tag{6.8.3.4}$$

where

$$\psi_n(0) = \psi_n^0. \tag{6.8.3.5}$$

Next we will be intrested in the probability for the transition from state $\psi_n(t)$ to ψ_m^0 in the period $[0,T]$. We assume that we start in the state $\psi_n(0) = \psi_n^0$, then the state evolve under the perturbed hamiltonian H according to equation 6.8.3.4. At the time $t = T$ perturbation W is "turned off" and we measure the energy of the state. We want to know the probability that our final state will be ψ_m^0, which is expressed by

$$|\left\langle \psi_m^0 \middle| \psi_n(t) \right\rangle|^2. \tag{6.8.3.6}$$

Let's express $\psi_n(t)$ in basis $\psi_k^0(t)$:

$$\psi_n(t) = \sum_k a_{n,k}(t)\psi_k^0(t). \tag{6.8.3.7}$$

Note that

$$|\left\langle \psi_m^0 \middle| \psi_n(t) \right\rangle|^2 = |a_{n,m}(t)|^2. \tag{6.8.3.8}$$

Coefficients $a_{n,m}$ are given by the following system of differential equations

$$i\hbar \frac{da_{n,m}(t)}{dt} = \sum_k a_{n,k}(t) \left\langle \psi_m^0 \middle| W(t) \middle| \psi_k^0 \right\rangle e^{i\omega_{mk}t}, \tag{6.8.3.9}$$

where

$$\omega_{mk} = \frac{E_m - E_k}{\hbar}. \tag{6.8.3.10}$$

The first order aproximation of $a_{n,m}(t)$ is

$$a_{n,m}(t) \approx a_{n,m}^{(1)} = \delta_{nm} - \frac{i}{\hbar}\int_0^t \left\langle \psi_m^0 \middle| W(\tau) \middle| \psi_n^0 \right\rangle e^{i\omega_{mn}\tau} d\tau. \tag{6.8.3.11}$$

The above is proven in (e.g. [**?**][11.4]).

Chapter 7

Quantum Field Theory

7.1 Introduction

7.1.1 Preliminaries

We will have convention that for a space time vector x in a chosen frame of reference $\boldsymbol{x}$ will mean space part of the vector x. $x \cdot y$ will mean Minkowsky inner product:

$$x \cdot y = x^0 y^0 - \boldsymbol{x} \cdot \boldsymbol{y}, \tag{7.1.1.1}$$

where $\boldsymbol{x} \cdot \boldsymbol{y}$ is euclidean inner product in $\mathbb{R}^3$. We will also write $x^2 = x \cdot x$ and $\boldsymbol{x}^2 = \boldsymbol{x} \cdot \boldsymbol{x}$.

When a particle of mass m has a 4-momentum p, we will denote it's energy by $E_{\boldsymbol{p}} = p^0$. Note that

$$E_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2}. \tag{7.1.1.2}$$

Fact 7.1.1.1. *If p, q are 4-momentums of particles of mass m, then the expression*

$$E_{\boldsymbol{q}} \delta^{(3)}(\boldsymbol{p} - \boldsymbol{q}) \tag{7.1.1.3}$$

is Lorentz invariant.

Proof. Take Lorentz transformation

$$\begin{cases} E_{\boldsymbol{p}'} = p'^0 = \gamma p^0 - \beta \gamma p^1, \\ p'^1 = \gamma p^1 - \beta \gamma E_{\boldsymbol{p}}. \end{cases} \tag{7.1.1.4}$$

Let's calculate

$$\delta^{(3)}(\boldsymbol{p'} - \boldsymbol{q'}) = \delta((p'^1 - q'^1, p'^2 - q'^2, p'^3 - q'^3))$$
$$= \delta((\gamma p^1 - \beta\gamma E_{\boldsymbol{p}} - \gamma q^1 + \beta\gamma E_{\boldsymbol{q}}, p^2 - q^2, p^3 - q^3)).$$

Note that

$$\frac{\partial}{\partial p^k} E_{\boldsymbol{p}} = \frac{\partial}{\partial p^k}\sqrt{\boldsymbol{p}^2 + m^2} = \frac{p^k}{E_{\boldsymbol{p}}}. \tag{7.1.1.5}$$

Then

$$\frac{\partial}{\partial p^1}(\gamma p^1 - \beta\gamma E_{\boldsymbol{p}} - \gamma q^1 + \beta\gamma E_{\boldsymbol{q}})$$

$$= \gamma - \frac{\beta\gamma p^1}{E_{\boldsymbol{q}}} = \frac{E_{\boldsymbol{p}}\gamma - \beta\gamma p^1}{E_{\boldsymbol{p}}} = \frac{E_{\boldsymbol{p'}}}{E_{\boldsymbol{p}}}.$$

Let's write Jacobian

$$\left|\frac{\partial(\boldsymbol{p'} - \boldsymbol{q'})}{\partial \boldsymbol{p}}\right| = \begin{vmatrix} \frac{E_{\boldsymbol{p'}}}{E_{\boldsymbol{p}}} & \frac{\beta\gamma p^2}{E_{\boldsymbol{p}}} & \frac{\beta\gamma p^3}{E_{\boldsymbol{p}}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{E_{\boldsymbol{p'}}}{E_{\boldsymbol{p}}}. \tag{7.1.1.6}$$

By Theorem 6.4.2.1, we have

$$\delta(\boldsymbol{p}' - \boldsymbol{q}') = \left(\left|\frac{\partial(\boldsymbol{p'} - \boldsymbol{q'})}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=\boldsymbol{q}}\right)^{-1} \delta(\boldsymbol{p} - \boldsymbol{q}) = \frac{E_{\boldsymbol{q}}}{E_{\boldsymbol{q}'}}\delta(\boldsymbol{p} - \boldsymbol{q}). \tag{7.1.1.7}$$

Thus

$$E_{\boldsymbol{q}'}\delta^{(3)}(\boldsymbol{p'} - \boldsymbol{q'}) = E_{\boldsymbol{q}}\delta^{(3)}(\boldsymbol{p} - \boldsymbol{q}). \tag{7.1.1.8}$$

□

Klein-Gordon equation For a relativistic particle $E^2 = \boldsymbol{p}^2 + m^2$, there is an idea for relativistic Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\phi(t)\rangle = \sqrt{\boldsymbol{P}^2 + m^2} \, |\phi(t)\rangle \, . \tag{7.1.1.9}$$

which is a motivation for a "squared" version (we will set $\hbar = 1$):

$$-\frac{\partial^2}{\partial t^2} = \boldsymbol{P}^2 + m^2, \tag{7.1.1.10}$$

which is

$$-\frac{\partial^2}{\partial t^2} = -\nabla^2 + m^2, \tag{7.1.1.11}$$

or

$$\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2 = 0. \tag{7.1.1.12}$$

Can be noted using raised index of partial derivative and Einstein summation convention:

$$\partial^\mu \partial_\mu + m^2 = 0. \tag{7.1.1.13}$$

From this immediately follows Lorentz invariance, but we will give a direct proof in language of tensor transformations Figure 7.1.1.

$$\frac{\partial}{\partial x_\mu}\frac{\partial}{x^\mu}\phi(\hat{x}) = g^{\sigma\nu}\frac{\partial}{\partial x^\sigma}\frac{\partial}{\partial x^\nu}\phi(\hat{x}) = (*)$$

$$\boxed{\frac{\partial}{\partial x^\nu}\phi(\hat{x}) = \frac{\partial}{\partial x^\nu}\phi\left(x^\xi \frac{\partial \hat{x}^\eta}{\partial x^\xi}\right) = \partial_\eta \phi(\hat{x})\frac{\partial \hat{x}^\eta}{\partial x^\nu}}$$

$$(*) = g^{\sigma\nu}\left(\partial_{\eta_1}\partial_{\eta_2}\phi\right)(\hat{x})\frac{\partial \hat{x}^{\eta_1}}{\partial x^\nu}\frac{\partial \hat{x}^{\eta_2}}{\partial x^\sigma} = g^{\sigma\nu}\Lambda^{\hat{\eta}_1}_{\nu}\Lambda^{\hat{\eta}_2}_{\sigma}\left(\partial_{\eta_1}\partial_{\eta_2}\phi\right)(\hat{x}) =$$

$$= \hat{g}^{\eta_1\eta_2}\left(\partial_{\eta_1}\partial_{\eta_2}\phi\right)(\hat{x}) = \left(\partial^\mu\partial_\mu\phi\right)(\hat{x})$$

Figure 7.1: Proof of Lorentz invariance of Klein-Gordon equation

Chapter 8

Experiments

8.1 Secondary Cosmic Rays Detection with Two Geiger-Müller Tubes

8.1.1 Introduction

The idea to detect cosmic rays by registering simultanous discharges in Geiger-Müller tubes is almost as old as GM tubes themselves. Walther Bothe and Werner Kolhörster published the results from a coincidence experiment with discharges in GM counters in 1929 ("Das Wesen der Höhenstrahlung", Zeitschrift für Physik 1929, 56: 751), while GM tube was invented in 1928. They registered deflections of fiber electrometers on moving film to detect symyultanous discharges [**?**]. The first modern, analogous to AND gate, system for detection of symultanous discharges was proposed by Bruno Rossi in 1930 [**?**].

In XIX ceuntry, due to very low prices of electrical components, the experiments which were conducted on the state of art devices in 30's, are possible to repeat in schools' and hobbyists' laboratories. Recently, many experiments detecting muons by the means of two Geiger-Müller tubes were reported on various physical blogs on the Internet and were performed by many students in school laboratories.

The aim of this artcile is to report in possibly rigorous way one of such experiments conducted in hobbiest's laboratory and compare empirical count rate of symultanous discharges in GM tubes with theoretically predicted count rate of muons passing exactly through both GM tubes. We hope that the results could help studends to project their own experiments and to make necessary calculations.

8.1.2 Description of the Experimental Setup

In the experiment, two electric boards with Geiger-Muller (GM) tubes were fixed to a retord stand in such a way that the tubes were in horisontal position, parallel to each other, one exatly above the other in such a way that the vertical line could be drawn through its centers. The intention of this setup was to register the rate of "simultaneous" discharges caused by particles traveling down from the upper parts of the Earth atmosfere. The experiment was performed for (?) various distances of the tubes (...).

Equimpent used:

1. Assembled DIY Geiger-Muller Counter Kit with GM Tube SBM20 with date mark 8903
2. Assembled DIY Geiger-Muller Counter Kit with GM Tube SBM20 with date mark 8904
3. Microcontroler Atmel ATMega328P embedded in the board Elegoo Uno R3 with 16MHz clock.

Both Geiger-Muller Counter boards were built from the same specification. 3 pins from Geiger-Muller Counter boards were of our particular interest: +5V, GND and INT. On INT pin the special signal is generated whenever there is a discharge in the GM tube. This signal can trigger ATMega328P interruption if INT output from Geiger-Muller Counter board is connected to the interrupt pin (PIN2 or PIN3) of ATMega328P.

Both Geiger-Muller Counter boards were calibarted before experiment to $400 \pm 0.8V$ anode-cathode voltage on GM tube. The readings were done by $1M\Omega$ multimeter and the value read was $6.56 \pm 0.01V$ with multiplying factor 61.

In the experiment, Elegoo Uno R3 was powered by USB cable from PC. Both Geiger-Muller Counter boards were connected through their +5V pin pararelly to the source of +5V from Elegoo Uno R3 board and grounded by their GND pins to Elegoo Uno R3 board GND pin. INT pin from one Geiger-Muller Counter board (which was placed down) was connected to ATMega328P interrupt pin PIN2 and INT pin from the other Geiger-Muller Counter board (which was placed up) was connected to ATMega328P interrupt pin PIN3.

The software on ATMega328P registered signals on PIN2 and PIN3 caused by discharges in GM tubes and fire up interuptions respectively. It sent a signal through USB to PC if both interruptions fired up in less than

$$\Delta t_{th} = 100 \pm 4\mu\text{s} \tag{8.1.2.1}$$

time interval. Time of arrival of the signal was logged on PC together with time interval between interuptions.

It might have been temppting to use simple AND gate to register symultanous discharges in GM tubes. But even by AND gate will not answer if impulses were really absolutelly symultanous. They always come in certain, however small, intervals and the impulses also last for some time. While the time interval between impulses for AND gate is certainly of order of magnitute shorter than the time interval measured while registering by microcontroller, we had no practical means of quantitive analisys of the former. Thus we decided on the latter, accepting potentially greater, but measurable error.

8.1.3 Background readings

In order to determine background readings, both GM Counter boards were monitored for 55 908 s commenced at 2019-01-10 23:00:01.

	SBM20 (8903)	SBM20 (8904)
discharges:	20531	21522
average waiting time:	2.7232s	2.5978s
std of waiting time:	2.7537s	2.6158s
est. expect. waiting time (β):	$2.72 \pm 0.06s$	2.60 ± 0.05s
est. expect. rate (λ)	$0.367 \pm 0.008\text{s}^{-1}$	$0.385 \pm 0.008\text{s}^{-1}$

8.1.4 Differentiating discharges in tubes caused by secondary cosmic ray from random nearly simultaneous discharges

To detect a charged particle flying from the upper parts of Earth atmosfere by registering “simultaneous” discharges in both tubes, we need to define what “simultaneous” means in our laboratory conditions and prove that we can distiguish them from random nearly simultaneous discharges in tubes. As first test we put one Geiger-Muller Counter board very close (2.7 ± 0.1 cm) above the other and log all events in which discharges in both tubes were registered by microcontroler in an interval Δt where $\Delta t < \Delta t_{th} = 100 \pm 4\mu s$. The microcontroler with the Arduino libraries we used could measure time in microseconds with a resolution of 4μs.

We observed tubes for 36 627 s starting at 2019-01-12 12:39:29. During this time we registered 801 such events. In the table below we sumarise the distribution of measured time interval Δt between discharges in tubes.

$\Delta t[\mu\text{s}]$	number of events
12	385
16	128
20	114
24	87
28	45
32	28
36	12
40	1
44	1
$\Delta t > 44$	0

First thing to notice is that the number of events decreases when the length of interval increases, which should have been exactly in an opposit way if nearly simultaneous discharges had happend only by chance. The mean value is $m(\Delta t) = 17.12\mu\text{s}$ and the standard diviation is $\sigma(\Delta t) = 6.38\mu s$. A good candidate for a threshold is $m(\Delta t) + 3\sigma(\Delta t) = 36.3\mu\text{s}$. Keeping in mind that the resolution of the microcontroler is $4\mu\text{s}$, the natural choice is $36\mu\text{s}$.

Therefore, our definition of "simultaneous" will be, if two discharges appear in both tubes in the time interval less or equal

$$\Delta t_s = 36\mu\text{s}. \tag{8.1.4.1}$$

Considering the error of our microcontroller, we need to remeber that real $\Delta t_s = 36 \pm 4\mu\text{s}$

We need to keep in mind that based on empirical data in the above table this definition will introduce additional systematic error $\approx 0.25\%$ for measured count rate of simulatnous discharges.

Now, we need to estimate what are the chances of random discharges in both tubes in interval less than $\Delta t_s = 36\mu \pm 4\text{s}$. To calculate this, we will assume that discharges in tubes are completly random and independent and as such they follow the Poisson process with $\lambda_A = 0.367 \pm 0.008\text{s}^{-1}$ for the tube with date mark 8903 (called further tube A) and $\lambda_B = 0.385 \pm 0.008\text{s}^{-1}$ for tube with date mark 8904 (called further tube B) as established in Subsection 8.1.3. Let T_A and T_B be random variables which denote the waiting times for discharge respectively for tubes A and B. Now, consider two types of events:

1. Discharge in tube A followed by discharge in a tube B in an interval equal or less Δt_s.

2. Discharge in tube B followed by discharge in a tube A in an interval equal or less Δt_s.

They constitute two independent Poisson processes with rates respectively.

1. $$\lambda^A_{\Delta t_s} = \lambda_A P(T_B \leq \Delta t_s) = \lambda_A (1 - e^{-\lambda_B \Delta t_s}). \tag{8.1.4.2}$$

2. $$\lambda^B_{\Delta t_s} = \lambda_B P(T_A \leq \Delta t_s) = \lambda_B (1 - e^{-\lambda_A \Delta t_s}). \tag{8.1.4.3}$$

Now if we conceptually merge events of the above two types in one Poisson process, it will have the following rate

$$\lambda_{\Delta t_s} = \lambda_A P(T_B \leq \Delta t_s) = \lambda_A (1 - e^{-\lambda_B \Delta t_s}) + \lambda_B (1 - e^{-\lambda_A \Delta t_s}). \tag{8.1.4.4}$$

For our current values

$$\lambda_{\Delta t_s} = (1.02 \pm 0.16) \cdot 10^{-5} \mathrm{s}^{-1}. \tag{8.1.4.5}$$

Thus, expected waiting time for a spontanous "simultaneous" discharge (i.e. in interval equal or less $\Delta t_s = 36 \pm 4\mu$s) is $(1.00 \pm 0.16) \cdot 10^5$s which is more than 23 hours and a half. This rate of spontanous "simultaneous" discharges is quite satisfactory, however we need to keep in mind that $\lambda_{\Delta t_s}$ is our lower bound of an error for any count rate of particles that we will be able to detect in our experiment. Thus we will assume that any conclusions from the experimnent can be drawn only for the rates which are of at least 2 orders of magnitude greater than $\lambda_{\Delta t_s}$.

8.1.5 Theoretical prediction of mouns count rate in our experimental setup

We attemped to predict theoretically the count rate of muons passing exactly through both of our GM tubes (more precisely, through the active volumes of the tubes). We used the value of vertical intensity of muon flux

$$I_V = (8.4 \pm 0.4) \cdot 10^{-3} \mathrm{cm}^2 \mathrm{s}^{-1} \mathrm{sr}^{-1} \tag{8.1.5.1}$$

as determined in [**?**] and some of our calculation were inspiried by explanations given in [**?**].

The count rate $d\lambda$ of muons coming from an infinitesimal solid angle element $d\Omega$ of celestial sphere at a zenith angle θ and passing through an infinitesimal element of surfuce with a vector area $d\vec{S}$ is assumed to be equal to

$$d\lambda = I_V \cos^n \theta (\vec{\eta} \cdot d\vec{S}) d\Omega, \tag{8.1.5.2}$$

where $\vec{\eta}$ is a unit vector directed from the infinitesimal element of surfuce $d\vec{S}$ to the infinitesimal solid angle element of celestial sphere $d\Omega$, as ilustarted on Figure 8.1. After [?], we assumed $n = 1.85 \pm 0.10$.

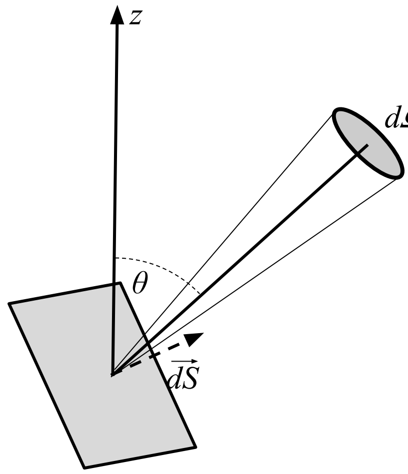

Figure 8.1: Muon flux through the infinitesimal element of surfuce with a vector area $d\vec{S}$ from an infinitesimal element of celestial sphere with a solid angle $d\Omega$ at a zenith angle θ.

In case of two cylindrical tubes of the same radious R and of the same length L placed horizontally one above the other in a distance H, the total count rate λ of muons going exactly through both of the tubes will be given by

$$\lambda = I_V \int_{\cup} \int_{\cap} \cos^n \theta (\vec{\eta} \cdot d\vec{S}) d\Omega. \tag{8.1.5.3}$$

Where an infinitesimal vector area element $d\vec{S}$ runs over the upper part of the lower tube's surface and an infinitesimal solid angle element $d\Omega$ runs over the part of celestial sphere as visible from the point $d\vec{S}$ through the lower half of the upper tube's surface.

Let's prepare to express an integral 8.1.5.3 in variables x, α and θ, ϕ where x, α describe surface coordinates of an element $d\vec{S}$ on the lower tube's surface as in Figure 8.2 and θ, ϕ are respectivelly zenith and azimuth angles of an element $d\Omega$ looking from an element $d\vec{S}$.

Note that

$$d\Omega = \sin\theta d\theta d\phi. \tag{8.1.5.4}$$

We will introduce now help variables x, z, y and x', z', y' in such a way that (x, z, y) are cartesian coordinates of an infinitesimal vector area element $d\vec{S}$ and (x', z', y') are cartesian coordinates of the point where an infinitesimal solid angle element $d\Omega$ crosses the lower half of the upper tube's surface

Note that

$$d\vec{S} = d\alpha dx R(\frac{1}{R}[0, y, z]) = d\alpha dx[0, y, z]. \tag{8.1.5.5}$$

Now, we can express the integral 8.1.5.3 as dependent on α, x, θ, ϕ.

$$\lambda = I_V \int_{\cup} \int_{\cap} \cos^n\theta \sin\theta(\vec{\eta} \cdot [0, y, z]) d\theta d\phi dx d\alpha. \tag{8.1.5.6}$$

Next, we prepare to integrate by substitution

$$x, \alpha, \theta, \phi \mapsto x, \alpha, x', \alpha' \tag{8.1.5.7}$$

where x', α' describe surface coordinates of the point where an infinitesimal solid angle element $d\Omega$ crosses the lower half of the upper tube's surface as in Figure 8.2.

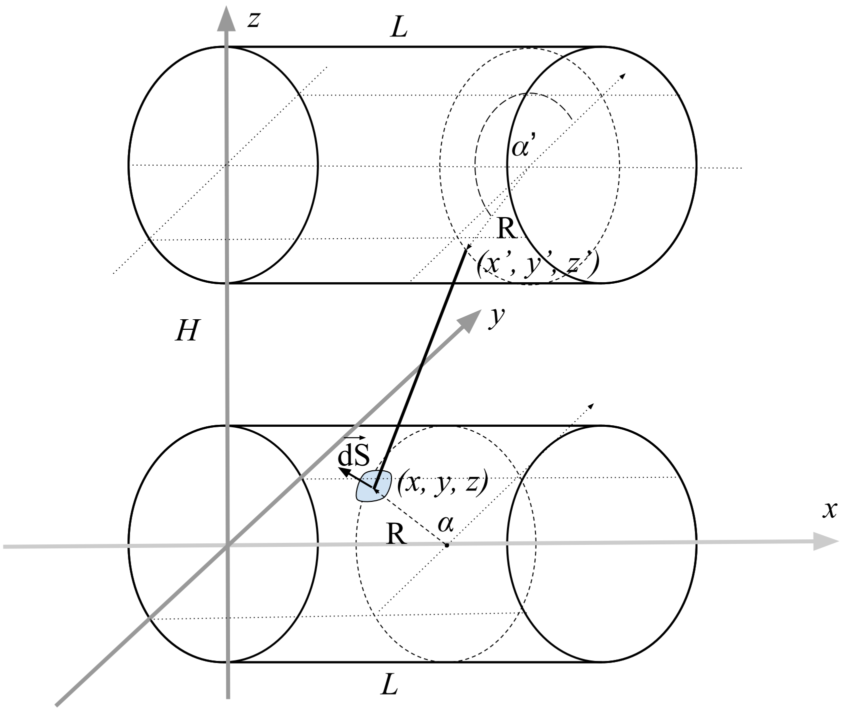

Figure 8.2: Muon flux through two GM tubes.

Let's calculate how $\theta, \phi, \vec{\eta}$ and y, z depends on x, α, x', α'. Note that

$$\begin{cases} y = R\cos\alpha, \\ z = R\sin\alpha \\ y' = R\cos\alpha', \\ z' = R\sin\alpha' + H. \end{cases} \tag{8.1.5.8}$$

Now, we can calculate $\theta, \phi, \vec{\eta}$.

$$\begin{cases} \theta = \arccos\left(\frac{z'-z}{\sqrt{(x'-x)^2+(y'-y)^2+(z'-z)^2}}\right), \\ \phi = \mathrm{arc}((x'-x)+i(y'-y)), \\ \vec{\eta} = \frac{1}{\sqrt{(x'-x)^2+(y'-y)^2+(z'-z)^2}}[x'-x, y'-y, z'-z], \end{cases} \tag{8.1.5.9}$$

It is quite obvious that all above variables depend only on x, α, x', α'.

To do substitution of variables (8.1.5.7) in the integral 8.1.5.6 we need to introduce Jacobian

$$J = \begin{bmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial x'} & \frac{\partial x}{\partial \alpha'} \\ \frac{\partial \alpha}{\partial x} & \frac{\partial \alpha}{\partial \alpha} & \frac{\partial \alpha}{\partial x'} & \frac{\partial \alpha}{\partial \alpha'} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial \alpha} & \frac{\partial \theta}{\partial x'} & \frac{\partial \theta}{\partial \alpha'} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial \alpha} & \frac{\partial \phi}{\partial x'} & \frac{\partial \phi}{\partial \alpha'} \end{bmatrix} \tag{8.1.5.10}$$

After the substitution we have

$$\lambda = I_V \int_0^L \int_0^\pi \int_0^L \int_\pi^{2\pi} \cos^n \theta \sin \theta (\vec{\eta} \cdot [0, y, z]) |\det(J)| dx d\alpha dx' d\alpha'. \tag{8.1.5.11}$$

We didn't solve the integral (8.1.5.11) manually, instead we used software Mathematica 11 [**?**] to calculate J symbolically and then applied Monte Carlo method (with $5 \cdot 10^7$ iterations) for fixed R, L, H to calculate theoretically expected muon count rate λ for particular experimental setups.

8.1.6 Experimental symultanous discharge count rate v. theoretical muon count rate

In the table below we will compare theoretical results calculated by integral (8.1.5.11) with experimental count rate of symultanous discharges. For used GM tubes SBM20, we assumed after official specification for active length and active radious

$$\begin{cases} L = 9.1 \pm 0.1 \text{cm}, \\ R = 1.0 \pm 0.1 \text{cm}. \end{cases} \tag{8.1.6.1}$$

GM tubes distance [cm]	predicted muon count rate [s^{-1}]	observed symult. discharges count rate [s^{-1}]	observation time [s]	observed symult. discharges
2.7 ± 0.1	0.031 ± 0.009	0.022 ± 0.002	36 627	801
4.2 ± 0.1	0.0185 ± 0.0055	0.014 ± 0.002	43 557	611
7.9 ± 0.1	$(7.9 \pm 2.1) \cdot 10^{-3}$	$(5.2 \pm 0.7) \cdot 10^{-3}$	106 971	545
11.2 ± 0.1	$(4.7 \pm 1.3) \cdot 10^{-3}$	$(3.1 \pm 0.4) \cdot 10^{-3}$	141 383	437

Chapter 9

Examples

9.1 Quantum Probability

Example 9.1.0.1. *In the lowest energy state ψ_0 of harmonic oscillator, there exists a joint distribution for P_{ψ_0} (momentum) and Q_{ψ_0} (position) random varibles ([**?**]).*

Chapter 10

Ideas

10.1 Quantum Physics

10.1.1 Mathematical formalism

There are many problems in Quantum Mechanics which relates to imprecise use of mathematical apparatus, which cause at least propaedeutical problems. Some of them are described in an excellent paper [?]. On the other hand it is well known fact that quantum mechanics done on finite dimensional Hilbert spaces is more adoptable mathematically to comprehend. Isn't any way to make Quantum Mechanics in some kind of limit space of finite dimentinal Hilber spaces?

10.1.2 Conservation of wave function modulus

In finite dimension Hilbert space it is relatively easy to prove, using a finite spectral theorem, that if observable A and H commute than for any quantum state which evolves according to Schrödinger equation

$$i\frac{d\psi}{dt} = H\psi. \tag{10.1.2.1}$$

$$\langle A\psi(t), e_i\rangle = \text{const}, \tag{10.1.2.2}$$

for any eigenvector e_i of A. How this translates to the infinite dimensional case to have

$$|\hat{\psi}(t)|^2 = \text{const} \tag{10.1.2.3}$$

when momentum operator P commutes with H. Does it take for P and H to have the same spectral basis? In case of free particle it holds (see equation 6.1.1.5)

Chapter 11

Mathematics

11.1 Vector Analysis

Definition 11.1.0.1. *(**Cross product**) Let $x, y \in \mathbb{R}^3$ such that*

$$x = (x_1, x_2, x_3) \text{ and } y = (y_1, y_2, y_3). \tag{11.1.0.1}$$

We define

$$x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1). \tag{11.1.0.2}$$

Theorem 11.1.0.2. *Let $a, b, c \in \mathbb{R}^3$. The following conditions holds:*

1. $a \times a = 0$
2. $a \times b = -b \times a$.
3. $(ta) \times b = t(a \times b)$ *for any* $t \in \mathbb{R}$.
4. $(a + b) \times c = a \times c + b \times c$.
5. $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$.

By $\cdot$ we will denote inner product.

Theorem 11.1.0.3. *Let $a, b, c \in \mathbb{R}^3$. The following conditions holds:*

1. $a \cdot (a \times b) = 0$.
2. $a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b)$.
3. $\|a \times b\|^2 + (a \cdot b)^2 = \|a\|^2\|b\|^2$.

11.1.1 Vector Spaces

In this chapter when ranges of summation are evident, we will also use Einstein summation convention. We will also use symbol δ for Kronecker delta.

$$\delta_{\mu\nu} = \delta^{\mu}_{\nu} = \delta^{\mu\nu} = \begin{cases} 1 \text{ for } \mu = \nu, \\ 0 \text{ for } \mu \neq \nu. \end{cases} \tag{11.1.1.1}$$

Definition 11.1.1.1. *Let X be a vector space over field $\mathbb{K}$. By $X^\star$ we will denote a space of all linear functions $X \to \mathbb{K}$.*

If $u_1, \ldots, u_k$ is a set of vectors in vector space X, for any $y \in X^\star$ we might write a row vector $[y(u_1), \ldots, y(u_k)]$.

Lemma 11.1.1.2. *Let X be a vector space. For any sequence of vectors $u_1, \ldots, u_k \in X$, for any sequence of functionals $y^1, \ldots, y^k \in X^\star$ and for any sequence of scalars $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$, we have*

$$\lambda_\mu y^\mu = 0 \tag{11.1.1.2}$$

if and only if

$$\lambda_\mu [y^\mu(u_1), \ldots, y^\mu(u_k)] = 0. \tag{11.1.1.3}$$

Proof. Implication from (11.1.1.2) to (11.1.1.3). Now, assume that (11.1.1.3) holds. Take any $x \in X$, assume that $x = x^\nu u_\nu$.

$$\lambda_\mu y^\mu(x) = \lambda_\mu y^\mu(x^\nu u_\nu) = \lambda_\mu x^\nu y^\mu(u_\nu) = x^\nu \lambda_\mu y^\mu(u_\nu) = 0. \tag{11.1.1.4}$$

□

As a simple conclusion, we will formulate the following.

Theorem 11.1.1.3. *Let X be a vector space and let $u_1, \ldots, u_n$ be a linear basis. Then there exist a basis of $X^\star$ $e^1, \ldots, e^n$ such that*

$$e^\mu(u_\nu) = \begin{cases} 1 \textit{ for } \mu = \nu, \\ 0 \textit{ for } \mu \neq \nu. \end{cases} \tag{11.1.1.5}$$

Proof. We can easily define functionals by $e^\mu(x^\nu u_\nu) := x^\mu$. Linearity is easy to show. By Lemma 11.1.1.2 $u^1, \ldots, e^n$ are linearly independent. Now we will show that all $y \in X^\star$ can be represented linearly by $e^1, \ldots, e^n$. Take any $x = x^\nu u_\nu$ and calculate

$$y(x) = y(x^\nu u_\nu) = x^\nu y(u_\nu) = y(u_\nu)e^\nu(x) = (y(u_\nu)e^\nu)(x). \tag{11.1.1.6}$$

□

Corollary 11.1.1.4. *Let X be a vector space with $\dim X < \infty$, then $\dim X = \dim X^{\star}$.*

Definition 11.1.1.5. *Let X be a vector space and $u_1, \ldots, u_n$ be an linear basis. We define a dual basis in $X^{\star}$ as $u^1, \ldots, u^n$ where*

$$u^{\mu}(u_{\nu}) = \delta^{\mu}_{\nu}. \tag{11.1.1.7}$$

The above tells us that for any chosen basis of X, for functionals $y \in X^{\star}$ there exists a natural parametrisation

$$y \mapsto [y(u_1), \ldots, y(u_n)] \in \mathbb{R}^n. \tag{11.1.1.8}$$

Note that this parametrisation doesn't establish identity between elements of X and $X^{\star}$. For this we will need a metric tensor.

If not stated other wise we will use the following convention. If $\{\hat{u}_{\mu}\}$ is also a basis of a vector space X, then we will denote transformation matrices using a capital letter - in this case U like this

$$\hat{u}_{\mu} = U^{\nu}_{\hat{\mu}} u_{\nu} \tag{11.1.1.9}$$

and

$$u_{\mu} = U^{\hat{\nu}}_{\mu} \hat{u}_{\nu}. \tag{11.1.1.10}$$

Where $U^{\hat{}}$ and $U_{\hat{}}$ are treated as a different variable which denote matrices.

Theorem 11.1.1.6. *Let X be a real vector space where $\dim X = n$ and let $u_1, \ldots, u_n$ and $\hat{u}_1, \ldots, \hat{u}_n$ be two bases of X, then*

$$\hat{u}^{\mu} = U^{\hat{\mu}}_{\nu} u^{\nu} \tag{11.1.1.11}$$

and

$$u^{\mu} = U^{\mu}_{\hat{\nu}} u^{\nu} \tag{11.1.1.12}$$

Proof. Note that $U^{\hat{\mu}}_{\nu} U^{\nu}_{\hat{\sigma}} = \delta^{\mu}_{\sigma}$. Assume that

$$\hat{u}^{\mu} = X^{\mu}_{\nu} u^{\nu}. \tag{11.1.1.13}$$

Then

$$\delta^{\mu}_{\sigma} = X^{\mu}_{\nu} u^{\nu}(\hat{u}_{\sigma}) = X^{\mu}_{\nu} u^{\nu}(U^{\rho}_{\hat{\sigma}} u_{\rho}) = X^{\mu}_{\nu} U^{\rho}_{\hat{\sigma}} \delta^{\nu}_{\rho} = X^{\mu}_{\nu} U^{\nu}_{\hat{\sigma}}.$$

Thus $X^{\mu}_{\nu} = U^{\hat{\mu}}_{\nu}$. We prove (11.1.1.12) analogously. □

11.1.2 Metric Tensor

It is called otherwise inner product.

Definition 11.1.2.1. *Let X be a real vector space. A functional $g : X^2 \to \mathbb{R}$ is called a metric tensor iff*

1. $\forall_{x,y \in X}\, g(x, y) = g(y, x)$,
2. $\forall_{x,y,z \in X} \forall_{a,b \in K}\, g(z, ax + by) = ag(z, x) + bg(z, y)$,
3. $\forall_{x \neq 0} \exists_y g(x, y) \neq 0$.

The property in the point 3 is called *nondegeneracy* of metric tensor. Throughout this chapter if not stated otherwise g will denote a metric tensor.

Lemma 11.1.2.2. *Let X be a real vector space with a metric tensor g. If S is a subspace of X such that $g(x, x) = 0$ for all $x \in S$, then $g(x, y) = 0$ for all $x, y \in S$.*

Proof. Take any $x, y \in S$

$$0 = g(x + y, x + y) = g(x, x) + 2g(x, y) + g(y, y) = 2g(x, y). \tag{11.1.2.1}$$

□

Lemma 11.1.2.3. *Let X be a real vector space with a metric tensor g. If S is a proper subspace of X, then there exists $u \notin S$ such that $g(u, u) \neq 0$.*

Proof. Assume by contradiction that $g(v, v) = 0$ for all $v \notin S$. Let V be a vector subspace such that $X = S + V$ and $S \cap V = \{0\}$. Then we have $g(v, v) = 0$ for all $v \in V$ and by Lemma 11.1.2.2, we have $g(p, q) = 0$ for all $p, q \in V$. Take a non-zero vector $x \in V$. By definition of metric tensor (Definition 11.1.2.1), we have $y \in X$ such that $g(x, y) \neq 0$. But $y = s + v$ where $s \in S$ and $v \in V$.

$$g(x, y) = g(x, s) + g(x, v) = g(x, s). \tag{11.1.2.2}$$

Thus $g(x, s) \neq 0$. Let λ be an arbitrary scalar. Consider

$$g(x + \lambda s, x + \lambda s) = 2\lambda g(x, s) + \lambda^2 g(s, s). \tag{11.1.2.3}$$

Since $g(x, s) \neq 0$, we can choose such a λ that $g(x + \lambda s, x + \lambda s) \neq 0$. Let $u = x + \lambda s$. We have $g(u, u) \neq 0$ and since $x \notin S$, $u \notin S$. □

Corollary 11.1.2.4. *Let X be a real vector space with a metric tensor g where* $\dim X < +\infty$. *There exists a basis $u_1, \ldots, u_n$ such that $g(u_i, u_i) \neq 0$ for any $i = 1, \ldots, k$.*

Definition 11.1.2.5. *Let X be a real vector space and g be a metric tensor. We say that vectors $u_1, \ldots, u_k$ are orthogonal with respect to g iff*

$$g(u_i, u_i) \neq 0 \text{ for } i = 1, \ldots, k, \tag{11.1.2.4}$$

and

$$g(u_i, u_j) = 0 \text{ for } i \neq j. \tag{11.1.2.5}$$

Theorem 11.1.2.6. *Let X be a real vector space with a metric tensor g. If vectors $u_1, \ldots, u_k$ are orthogonal with respect to g, then vectors $u_1, \ldots, u_k$ are linearly independent.*

Proof. We will prove this by induction. For $k = 1$, thesis is trivially true. By induction, assume that thesis holds for $k - 1$. Thus vectors $u_1, \ldots, u_{k-1}$ are linearly independent. Assume to the contrary that vectors $u_1, \ldots, u_{k-1}, u_k$ are linearly dependent. Therefore we have parameters for which $u_k = \sum_{i=1}^{k-1} \lambda_i u_i$. But then

$$g(u_k, u_k) = g(u_k, \sum_{i=1}^{k-1} \lambda_i u_i) = \sum_{i=1}^{k-1} g(u_k, u_i)\lambda_i = 0. \tag{11.1.2.6}$$

Hence, contradiction with $g(u_k, u_k) \neq 0$, which proves that $u_1, \ldots, u_{k-1}, u_k$ are linearly independent which completes proof by induction. □

Lemma 11.1.2.7. *Let X be a real vector space with* $\dim X = n$. *If $u_1, \ldots, u_k$ are orthogonal and $k < n$ then there exists such $u \in X$ that $u_1, \ldots, u_k, u$ are orthogonal.*

Proof. Let $S = \operatorname{span}\{u_1, \ldots, u_k\}$. Choose $v_i, \ldots, v_{n-k} \in X$ such that $u_1, \ldots, u_k, v_1, \ldots, v_{n-k}$ form a linear basis of X. Let's define

$$\hat{v}_j = v_j - \sum_{i=1}^{k} \frac{g(u_i, v_j)}{g(u_i, u_i)} u_i \tag{11.1.2.7}$$

for $j = 1, \ldots, n - k$. Note that

$$g(u_i, \hat{v}_j) = g(u_i, v_j) - \frac{g(u_i, v_j)}{g(u_i, u_i)} g(u_i, u_i) = 0 \tag{11.1.2.8}$$

for $i = 1, \ldots, k$ and $j = 1, \ldots, n - k$. Note that $u_1, \ldots, u_k, \hat{v}_1, \ldots, \hat{v}_{n-k}$ is an independent linear basis of X. Let $V = \operatorname{span}\{\hat{v}_1, \ldots, \hat{v}_{n-k}\}$. It is enough to show that we have a $v \in V$ such that $g(v,v) \neq 0$. Assume to the contrary that for any $v \in V$, we have $g(v,v) = 0$. From Lemma 11.1.2.2, we have $g(v_0, v) = 0$ for all $v_0, v \in V$. Take any $z \in X$, since $S + V = X$, we have $z = v + s$ where $v \in V$ and $s \in S$. Let's calculate

$$g(v_0, z) = g(v_0, v+s) = g(v_0, v) + g(v_0, s) = g(v_0, s) = 0. \tag{11.1.2.9}$$

Where the last equality is because of (11.1.2.8). But this is in contradiction with nondegeneracy of g. □

As a conclusion we can formulate the following:

Theorem 11.1.2.8. *Let X be a real vector space with $\dim X = n$. There exists an orthogonal linear basis of X.*

Definition 11.1.2.9. *Let X be a real vector space with a tensor metric g. We say that vectors $e_1, \ldots, e_k$ are orthonormal with respect to g iff*

$$g(e_i, e_i) = \pm 1 \text{ for } i = 1, \ldots, k, \tag{11.1.2.10}$$

and

$$g(e_i, e_j) = 0 \text{ for } i \neq j. \tag{11.1.2.11}$$

If we have orthogonal vectors $u_1, \ldots, u_k$, we can transform them easily into orthonormal by

$$e_i = |g(u_i, u_i)|^{-\frac{1}{2}} u_i. \tag{11.1.2.12}$$

Corollary 11.1.2.10. *Let X be a real vector space with $\dim X = n$. There exists an orthonormal linear basis of X.*

Theorem 11.1.2.11. *Let X be a real vector space with $\dim X = n$ and let g be a metric tensor. If $e_1, \ldots e_n$ and $f_1, \ldots f_n$ are orthonormal bases with respect to tensor g then*

$$|\{i : g(e_i, e_i) = 1\}| = |\{i : g(f_i, f_i) = 1\}|. \tag{11.1.2.13}$$

Proof. Assume to the contrary that thesis doesn't hold. Let $g(e_i, e_i) = 1$ for $i = 1, \ldots, k_e$ and $g(e_i, e_i) = -1$ for $i = k_e + 1, \ldots, n$ and $g(f_i, f_i) = 1$ for $i = 1, \ldots, k_f$ and $g(f_i, f_i) = -1$ for $i = k_f + 1, \ldots, n$. By symmetry assume that $k_f < k_e$.

Assume that A_i^j is a transition matrix and $e_i = \sum_{j=1}^n A_i^j f_j$. Let $V^+ = \mathrm{span}\{e_1, \ldots, e_{k_e}\}$. Take any $v = \lambda^1 e_1 + \cdots + \lambda^{k_e} e_{k_e} \in V^+$. Note that

$$v = \sum_{i=1}^{k_e} \sum_{j=1}^{n} \lambda^i A_i^j f_j = \sum_{j=1}^{n} (\sum_{i=1}^{k_e} \lambda^i A_i^j) f_j. \tag{11.1.2.14}$$

Let's consider a linear mapping L

$$\mathbb{R}^{k_e} \ni (\lambda^1, \dots, \lambda^{k_e}) \mapsto (\sum_{i=1}^{k_e} \lambda^i A_i^1, \dots, \sum_{i=1}^{k_e} \lambda^i A_i^{k_f}) \in \mathbb{R}^{k_f}. \tag{11.1.2.15}$$

Since $\dim(\ker L) + \dim(\mathrm{im} L) = k_e$, we have $k_e - k_f \leq \dim(\ker L)$. Thus we have a non-zero

$$v_0 = \lambda_0^1 e_1 + \dots + \lambda_0^{k_e} e_{k_e} \in V^+ \tag{11.1.2.16}$$

such that $(\sum_{i=1}^{k_e} \lambda_0^i A_i^1, \dots, \sum_{i=1}^{k_e} \lambda_0^i A_i^{k_f}) = 0$. Therefore, by (11.1.2.14) we have

$$v_0 = (\sum_{i=1}^{k_e} \lambda_0^i A_i^{k_f+1}) f_{k_f+1} + \dots + (\sum_{i=1}^{k_e} \lambda_0^i A_i^n) f_n. \tag{11.1.2.17}$$

By (11.1.2.16), we have

$$g(v_0, v_0) = \sum_{i=1}^{k_e} (\lambda_0^i)^2 > 0. \tag{11.1.2.18}$$

But by (11.1.2.17), we have

$$g(v_0, v_0) = \sum_{j=k_f+1}^{n} (\sum_{i=1}^{k_e} \lambda_0^i A_i^j))^2 (-1) < 0, \tag{11.1.2.19}$$

which concludes proof by contradiction. □

Lemma 11.1.2.12. *Let X be a real vector space and $e_1, \dots, e_n$ is an orthonormal basis. Then for any x such that $g(x, e_i) = 0$ for all $i = 1, \dots, n$, we have $x = 0$.*

Proof. It follows from non-degeneracy of g. □

Corollary 11.1.2.13. *Let X be a real vector space and $e_1, \dots, e_n$ is an orthonormal basis. Then for any x, y such that $g(x, e_i) = g(y, e_i)$ for all $i = 1, \dots, n$, we have $x = y$.*

Theorem 11.1.2.14. *Let X be a real vector space. If $e_1, \dots, e_n$ is an orthonormal basis, then for any $x \in X$, we have*

$$x = \sum_{i=1}^{n} \frac{g(x, e_i)}{g(e_i, e_i)} e_i. \tag{11.1.2.20}$$

Proof. Note that for all $k = 1, \ldots, n$

$$g(\sum_{i=1}^{n} \frac{g(x, e_i)}{g(e_i, e_i)} e_i, e_k) = g(x, e_k). \tag{11.1.2.21}$$

Thus by Lemma 11.1.2.13 we have thesis. □

Corollary 11.1.2.15. *Let X be a real vector space. If $e_1, \ldots, e_n$ is an orthonormal basis, then for any $x, y \in X$*

$$g(x, y) = \sum_{i=1}^{n} g(x, e_i) g(y, e_i) g(e_i, e_i). \tag{11.1.2.22}$$

From now on when it is obvious we will use Einstein summation convention.

Theorem 11.1.2.16. *Let X be a real vector space and $u_1, \ldots, u_n$ is any linear basis of X. If $x = x^\mu u_\mu$ and $y = y^\mu u_\mu$, then*

$$g(x, y) = g_{\mu\nu} x^\mu y^\nu, \tag{11.1.2.23}$$

where

$$g_{\mu\nu} = g(u_\mu, u_\nu). \tag{11.1.2.24}$$

Proof. Follows directly from bilinearity of g. □

We will call $g_{\mu\nu}$ a representation of g in basis $\{u_\mu\}$. It also follows that if $g_{\mu\nu}$ is a representation of g in orthonormal basis $g_{\mu\nu} = 0$ for $\mu \neq \nu$ and $g_{\mu\mu} = \pm 1$.

In this convention if $g_{\mu\nu}$ is a representation of a metric tensor g in basis u_μ, then by $\hat{g}_{\mu\nu}$ we will denote a representation of a metric tensor g in basis $\{\hat{u}_\mu\}$ (i.e. $\hat{g}_{\mu\nu} = g(\hat{u}_\mu, \hat{u}_\nu)$).

Theorem 11.1.2.17. *Let X be a real vector space where $\dim X = n$ with a metric tensor g. If $\{u_\mu\}$ and $\{\hat{u}_\mu\}$ are two linear bases of X, then*

$$\hat{g}_{\mu\nu} = g_{\rho\sigma} U^\rho_{\hat{\mu}} U^\sigma_{\hat{\nu}}. \tag{11.1.2.25}$$

Proof.

$$g(\hat{u}_\mu, \hat{u}_\nu) = g(U^\rho_{\hat{\mu}} u_\rho, U^\sigma_{\hat{\nu}} u_\sigma) = g(u_\rho, u_\sigma) U^\rho_{\hat{\mu}} U^\sigma_{\hat{\nu}}. \tag{11.1.2.26}$$

□

As a particular case of the above theorem we will formulate the following:

Theorem 11.1.2.18. *Let X be a real vector space with a metric tensor g and let $e_1, \ldots, e_n$ be an orthonormal basis of X and let $g_{\mu\nu} = g(e_\mu, e_\nu)$. If $f_1, \ldots f_n$ is an orthonormal basis of X and*

$$F^\mu_\nu = \frac{g(f_\nu, e_\mu)}{g(e_\mu, e_\mu)}, \tag{11.1.2.27}$$

(i.e. $f_\mu = F^\mu_\nu e^\nu$), then

$$g_{\mu\nu} F^\mu_\sigma F^\nu_\rho = \eta_{\sigma\rho}, \tag{11.1.2.28}$$

where $\eta_{\sigma\rho} = g(f_\sigma, f_\rho)$.

Next we will show how metric tensor g induces a natural linear isomorphism between X and $X^\star$.

Theorem 11.1.2.19. *Let X be a real vector space where $\dim X < +\infty$ with a tensor metric g. If $L : X \to X^\star$ is defined as*

$$(Ly)(x) = g(y, x) \text{ for all } x \in X. \tag{11.1.2.29}$$

then L is a linear isomorphism $L : X \underset{onto}{\overset{1-1}{\to}} X^\star$. Moreover

$$Ly = g(y, u_\mu) u^\mu \tag{11.1.2.30}$$

where $u_1, \ldots, u_n$ is any basis of X.

Proof. Since $\dim X = \dim X^\star$, to show that L is an isomorphism it is enough to show that $\ker L = \{0\}$. But this follows from nondegeneracy of g. Let $u_1, \ldots, u_n$ be any basis of X. Take any $y \in X$ and any $x \in X$. Assume that $x = x^\mu u_\mu$. Let's calculate

$$(Ly)(x) = g(y, x) = g(y, x^\mu u_\mu) = x^\mu g(y, u_\mu) = x^\nu \delta^\mu_\nu g(y, u_\mu)$$

$$= x^\nu u^\mu(u_\nu) g(y, u_\mu) = g(y, u_\mu) u^\mu(x^\nu u_\nu) = g(y, u_\mu) u^\mu(x).$$

□

Lemma 11.1.2.20. *Let X be a real vector space where $\dim X < +\infty$ with a tensor metric g and let $u_1, \ldots, u_n$ be a basis of X. If $g_{\mu\nu} = g(u_\mu, u_\nu)$ then $g_{\mu\nu}$ understood as matrix is invertible.*

Proof. Because L from Theorem 11.1.2.19 is linear isomorphism, for any $\sigma = 1, \dots, n$ we have such $x^{\sigma\nu}$ that

$$u^{\sigma} = g(x^{\sigma\nu} u_{\nu}, u_{\mu}) u^{\mu}. \tag{11.1.2.31}$$

Thus

$$\delta^{\sigma}_{\mu} = g(x^{\sigma\nu} u_{\nu}, u_{\mu}) = x^{\sigma\nu} g(u_{\nu}, u_{\mu}) = x^{\sigma\nu} g(u_{\nu}, u_{\mu}) = x^{\sigma\nu} g_{\mu\nu}. \tag{11.1.2.32}$$

Hence thesis. □

Note that for a real vector space X where $\dim X < \infty$ with metric tensor g, g induces an isometric metric tensor $g^{\star}$ on $X^{\star}$ in a following way:

$$g^{\star}(p, q) = g(L^{-1}p, L^{-1}q) \text{ for } p, q \in X^{\star}. \tag{11.1.2.33}$$

where L is from Theorem 11.1.2.19.

Theorem 11.1.2.21. *Let X be a real vector space where $\dim X < +\infty$ with a tensor metric g and let $u_1, \dots, u_n$ be a basis of X and $u^1, \dots, u^n$ be dual basis of $X^{\star}$. If $g_{\mu\nu} = g(u_{\mu}, u_{\nu})$ and $g^{\mu\nu} = g^{\star}(u^{\mu}, u^{\nu})$ then*

$$g_{\rho\nu} g^{\mu\nu} = \delta^{\mu}_{\rho}. \tag{11.1.2.34}$$

Proof. For u_{σ} and u_{ρ} we have

$$g^{\star}(Lu_{\sigma}, Lu_{\rho}) = g(u_{\sigma}, u_{\rho}). \tag{11.1.2.35}$$

Thus

$$g^{\star}(g(u_{\sigma}, u_{\mu}) u^{\mu}, g(u_{\rho}, u_{\nu}) u^{\nu}) = g_{\sigma\rho}. \tag{11.1.2.36}$$

Hence

$$g_{\sigma\mu} g_{\rho\nu} g^{\mu\nu} = g_{\sigma\rho} = g_{\sigma\mu} \delta^{\mu}_{\rho}. \tag{11.1.2.37}$$

But by Lemma 11.1.2.20, we have

$$g_{\rho\nu} g^{\mu\nu} = \delta^{\mu}_{\rho}. \tag{11.1.2.38}$$

□

This establishes an interesting fact that if $g_{\mu\nu}$ is representation of g in certain basis, then its inverse denoted as $g^{\mu\nu}$ is representation of $g^{\star}$ in a dual basis.

Assume we talk about real vector space X with $\dim X < +\infty$ with a metric tensor g. We know already that metric tensor g induces metric tensor $g^{\star}$ on $X^{\star}$ through establishing a natural linear isomorphism L between X

and $X^\star$. But then $g^\star$ establishes a natural linear isomorphism $L^\star$ between $X^\star$ and $X^{\star\star}$. Because $X^{\star\star}$ has a natural isomorphism with X defined by $x(y) := y(x)$, $L^\star$ will really establish a linear isomorphism between $X^\star$ and X. In that sense we will consider it as $L^\star : X^\star \to X$ such that for any $y, z \in X^\star$

$$(L^\star y)(z) = z(L^\star y) = g^\star(y, z). \tag{11.1.2.39}$$

We will show now that $L^\star = L^{-1}$, which means that this is exactly the same isomorphism as L, thus $g^{\star\star} = g$. First we will show that for a chosen basis $u_1, \ldots, u_n$

$$L^\star y = g^\star(y, u^\mu)u_\mu \text{ for any } y \in X^\star. \tag{11.1.2.40}$$

Let's calculate

$$\begin{aligned}(L^\star y)(z) = g^\star(y, z) = g^\star(y, z_\mu u^\mu)\\ = g^\star(y, u^\mu)z(u_\mu) = z(L^\star y) = z(g^\star(y, u^\mu)u_\mu)).\end{aligned} \tag{11.1.2.41}$$

To show that $L^{-1} = L^\star$ it is enough to show that $L^\star L u_\sigma = u_\sigma$.

$$\begin{aligned}L^\star L u_\sigma = g^\star(g(u_\sigma, u_\nu)u^\nu, u^\mu)u_\mu = g(u_\sigma, u_\nu)g^\star(u^\nu, u^\mu)u_\mu\\ = g_{\sigma\nu}g^{\nu\mu}u_\mu = \delta^\mu_\sigma u_\mu = u_\sigma.\end{aligned}$$

It is trivial now to show that $g^{\star\star} = g$.
Take any $x, y \in X$

$$g^{\star\star}(x, y) = g^\star((L^\star)^{-1}x, (L^\star)^{-1}x) = g^\star(Lx, Ly) = g(x, y). \tag{11.1.2.42}$$

All the above can be summarised shortly that metric tensor g establishes an identity between elements of X and X^* by transformation $x \mapsto g(x, \cdot)$. By doing so we obtain a metric tensor $g^\star$ on $X^\star$ which assumes the same values as metric tensor g on corresponding elements of X. Next when you repeat this operation for $g^\star$ with respect to $X^\star$ and $X^{\star\star} = X$, we go back to X and to the same metric tensor g (i.e. $g^{\star\star} = g$).

11.1.3 Contravariant and covariant coordinates

Let X be a vector space with a metric tensor g where $\dim X = n < +\infty$. Let $u_1, \ldots, u_n$ be a basis of X and let $u^1, \ldots, u^n$ be a dual basis of $X^\star$.

When we consider any vector $x \in X$ we might write this as

$$x = x^\mu u_\mu. \tag{11.1.3.1}$$

We call $x^1, \ldots, x^n$ contravariant coordinates of x. Recall that the metric tensor g establishes a linear isomorphism between X and $X^\star$. In this isomorphism a functional $x^\star(z) = g(x, z)$ corresponds to x. We might write $x^\star$ as

$$x^\star = x_\mu u^\mu. \tag{11.1.3.2}$$

We call $x_1, \ldots, x_n$ covariant coordinates of x (sic: we describe x in 2 ways because under the isomorphism we treat x and $x^\star$ as if there were the same object).

Theorem 11.1.2.19 establishes a relation between x^μ and x_μ. Indeed

$$\begin{aligned} x_\mu u^\mu = x^\star = Lx = g(x, u_\mu)u^\mu &= g(x^\nu u_\nu, u_\mu)u^\mu \\ = x^\nu g(u_\nu, u_\mu)u^\mu &= x^\nu g_{\nu\mu} u^\mu. \end{aligned}$$

Thus

$$\boxed{x_\mu = g_{\nu\mu} x^\nu} \tag{11.1.3.3}$$

The operation above is so called "index lowering". Because situation is symmetrical if we talk about linear isomorphism from $X^\star$ to X, we have:

$$\boxed{x^\mu = g^{\nu\mu} x_\mu} \tag{11.1.3.4}$$

The operation above is so called "index raising".

Let's now investigate how x^μ and x_μ transform under change of coordinates. Assume that we have other basis of X - $\hat{u}_1, \ldots, \hat{u}_n$ and $\hat{u}^1, \ldots, \hat{u}^1$ is a dual basis in $X^\star$.

Recall that by convention

$$\hat{u}_\mu = U^\nu_{\hat{\mu}} u_\nu, \tag{11.1.3.5}$$

and

$$u_\mu = U^{\hat{\nu}}_\mu \hat{u}_\nu. \tag{11.1.3.6}$$

By Theorem 11.1.1.6 we have

$$\hat{u}^{\mu} = U^{\hat{\mu}}_{\nu} u^{\nu}, \tag{11.1.3.7}$$

and

$$u^{\mu} = U^{\mu}_{\hat{\nu}} u^{\nu}. \tag{11.1.3.8}$$

Note the following

$$\hat{x}^{\nu} \hat{u}_{\nu} = x^{\mu} u_{\mu} = x^{\mu} U^{\hat{\nu}}_{\mu} \hat{u}_{\nu}. \tag{11.1.3.9}$$

Thus

$$\boxed{\hat{x}^{\nu} = U^{\hat{\nu}}_{\mu} x^{\mu}} \tag{11.1.3.10}$$

Note that that coordinates x^{μ} transform with the inverse of transformation matrix for vectors of basis - thus the name "contravariant".

On the other hand note following

$$\hat{x}_{\mu} \hat{u}^{\mu} = x_{\nu} u^{\nu} = x_{\nu} U^{\mu}_{\hat{\nu}} u^{\nu}. \tag{11.1.3.11}$$

Thus

$$\boxed{\hat{x}_{\mu} = U^{\mu}_{\hat{\nu}} x_{\nu}} \tag{11.1.3.12}$$

Note that the coordinates x_{μ} transform with the same transformation matrix as vectors of basis - thus the name "covariant".

11.2 Mathematical Analysis

Definition 11.2.0.1. *Let $U \subset \mathbb{R}^n$. By $C^n(U, \mathbb{R})$ we denote a set of all functions $f : U \to R$ for which a mapping*

$$U \ni (x^1, \ldots, x^n) \to \frac{\partial^n f}{\partial x^{k_1} \ldots \partial x^{k_m}} (x^1, \ldots, x^n) \tag{11.2.0.1}$$

is continuous for any ordered $(k_1, \ldots, k_m)$ where $k_i \in \{1, \ldots, m\}$ and $m \leq n$.

The above definition in equivalent to the one you can find in [**?**].

Theorem 11.2.0.2. *Let (X, d) be a metric space, let (Y, ρ) be complete metric space. If $BC(X, Y)$ is a space of all continuous and bounded functions from X to Y with a metric $\delta(f, g) = \sup_{x \in X} \rho(f(x), g(x))$, then $BC(X, Y)$ is a complete metric space.*

Proof. You may find a proof in [**?**] (V.4). □

Definition 11.2.0.3. *Let $(X, d), (Y, \rho)$ be metric spaces. We say that $f_n : X \to Y$ converges uniformly to $f_0 : X \to Y$, if and only if*

$$\lim_{n\to\infty} \sup_{x\in X} \rho(f_n(x), f_0(x)) = 0. \tag{11.2.0.2}$$

We will also denote uniform convergence as $f_n \rightrightarrows f_0$.

Definition 11.2.0.4. *Let $(X, d), (Y, \rho)$ be metric spaces. We say that $f_n : X \to Y$ converges almost uniformly to $f_0 : X \to Y$, if and only if $f|_K \rightrightarrows f_0|_K$ for each compact $K \subset X$.*

Theorem 11.2.0.5. *Let U be an open and connected subset of $\mathbb{R}$. If $g_n \in C^1(U)$, $\frac{dg_n}{dt} \to g$ almost uniformly and there exists at least one $x_0 \in U$ such that $g_n(x_0)$ converges to a certain real value, then $g_n \to g_0$ almost uniformly, where $\frac{dg_0}{dt} = g$.*

Proof. You will find a proof in [?] (V.4). □

Theorem 11.2.0.6. *Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of differentiable functions. If $\frac{df_n}{dt} \rightrightarrows g$ and there exists $t_0 \in [a, b]$ such that $\lim_{n\to\infty} f_n(t_0) \in \mathbb{R}$, then $f_n \rightrightarrows f$ and $\frac{df}{dt} = g$.*

Proof. You may find a proof in [?] (Uniform Convergence and Differentiation). □

Theorem 11.2.0.7. ***(Lebesgue's Dominated Convergence Theorem)*** *Let X be a measurable space with a positive measure μ. Let f_n be a sequence of a complex measurable functions,*

$$f(x) = \lim_{n\to\infty} f_n(x)$$

for each $x \in X$ and $|f_n| \le g$ where $g \in L^1(X)$, then $f \in L^1(X)$ and

$$\int_X f(x)\mu(dx) = \lim_{n\to\infty} \int_X f_n \mu(dx). \tag{11.2.0.3}$$

Proof. You may find a proof in [?] (Abstract Integration). □

.

Theorem 11.2.0.8. *Let X be a measurable space with a positive measure μ, U be an open subset of $\mathbb{R}$ and let $f : U \times X \to \mathbb{R}$. If*

1. *the mapping* $X \ni x \to f(t,x)$ *is a* $L^1(X)$ *function for each* $t \in U$,

2. *the mapping* $U \ni t \to \dfrac{\partial f}{\partial t}(t,x)$ *is continuous for each* $x \in X$,

3. $\left|\dfrac{\partial f}{\partial t}(U,x)\right| \leq g(x)$ *for each* $x \in X$, *where* $g \in L^1(X)$,

then the mapping $X \ni x \to \dfrac{\partial f}{\partial t}(t,x)$ *is a* $L^1(X)$ *function for each* $t \in U$ *and*

$$\frac{\partial}{\partial t}\int_X f(t,x)\mu(dx) = \int_X \frac{\partial f}{\partial t}(x,t)\mu(dx) \tag{11.2.0.4}$$

for each $t \in U$.

Proof. Take any $t_0 \in U$. Let $U \supset K(t_0, \varepsilon) \ni \Delta t \to 0$. Since the mapping $U \ni t \to \dfrac{\partial f}{\partial t}(t,x)$ is differentiable on U (thus on $(t_0 - \Delta t, t_0 + \Delta t)$) for each $x \in X$, by Lagrange's Theorem we have

$$\frac{f(t_0+\Delta t,x) - f(t_0,x)}{\Delta t} = \frac{\partial f}{\partial t}(t(t_0,\Delta t,x),x) \tag{11.2.0.5}$$

where $|t_0 - t(t_0,\Delta t,x)| < \Delta t$ for each $x \in X$. Note that since the mapping $U \ni t \to \dfrac{\partial f}{\partial t}(t,x)$ is continuous and $t(t_0,\Delta t,x) \to t_0$ for each $x \in X$, we have

$$\frac{f(t_0+\Delta t,x) - f(t_0,x)}{\Delta t} = \frac{\partial f}{\partial t}(t(t_0,\Delta t,x),x) \to \frac{\partial f}{\partial t}(t_0,x) \tag{11.2.0.6}$$

for each $x \in X$. Since by assumptions, we have $|\dfrac{\partial f}{\partial t}(t(t_0,\Delta t,x),x)| \leq g(x)$ for each $x \in X$ and $g \in L^1(X)$, we can apply Theorem 11.2.0.7 (Lebesgue's Dominated Convergence Theorem). Thus

$$\int_X \frac{f(t_0+\Delta t,x) - f(t_0,x)}{\Delta t}\mu(dx) \to \int_X \frac{\partial f}{\partial t}(t_0,x)\mu(dx), \tag{11.2.0.7}$$

which is the same as

$$\frac{\int_X f(t_0+\Delta t,x)\mu(dx) - \int_X f(t_0,x)\mu(dx)}{\Delta t} \to \int_X \frac{\partial f}{\partial t}(t_0,x)\mu(dx). \tag{11.2.0.8}$$

Hence, we have proved our thesis. $\square$

Theorem 11.2.0.9. *Let X be a compact space with a Borel additive measure μ where $\mu(X) < +\infty$, U be an open subset of $\mathbb{R}$ and $f : U \times X \to \mathbb{R}$. If*

1. *the mapping $X \ni x \to f(t, x)$ is a $L^1(X)$ function for each $t \in U$,*
2. *the mapping $U \times X \ni (t, x) \to \frac{\partial f}{\partial t}(x, t)$ is continuous,*

then the mapping $X \ni x \to \frac{\partial f}{\partial t}(t, x)$ is a $L^1(X)$ function for each $t \in U$ and

$$\frac{\partial}{\partial t}\int_X f(t, x)\mu(dx) = \int_X \frac{\partial f}{\partial t}(x, t)\mu(dx) \tag{11.2.0.9}$$

for each $t \in U$.

Proof. Take any $t_0 \in U$. We have an open neighbourhood $V \subset U$ of the point t_0 such that $Clo(V)$ is compact. Since $\frac{\partial f}{\partial t}$ is continous on $Clo(V) \times X$, we have $M > 0$ such that $\left|\frac{\partial f}{\partial t}(V, x)\right| \leq M$ for each $x \in X$. Now, since $\int_X M\mu(dx) = M\mu(X) < +\infty$, the assumptions of Theorem 11.2.0.8 are satisfied for the open set V. Thus

$$\frac{\partial}{\partial t}\int_X f(t_0, x)\mu(dx) = \int_X \frac{\partial f}{\partial t}(x, t_0)\mu(dx). \tag{11.2.0.10}$$

□

Theorem 11.2.0.10. ***(Fubini theorem)*** *Let $(X, \mathcal{M}_X, \mu)$, $(Y, \mathcal{M}_Y, \nu)$ be measurable spaces and μ and ν are σ-finite. If f is $\mathcal{M}_X \times \mathcal{M}_Y$ measurable and*

$$\int_X\int_Y |f(x, y)|d\nu(y)d\mu(x) < +\infty, \tag{11.2.0.11}$$

then

$$\int_X\int_Y f(x, y)d\nu(y)d\mu(x) = \int_Y\int_X f(x, y)d\mu(x)d\nu(y) < +\infty. \tag{11.2.0.12}$$

Proof. [See **?** , Integration on Product Spaces. The Fubini theorem] □

Definition 11.2.0.11. ***(Fréchet derivative)*** *Let X, Y be Banach spaces and $T : X \to Y$. We will say that T is differentiable at a point $x_0 \in X$, if and only if there exist a mapping $A_{x_0} \in L(X, Y)$ such that*

$$T(x_0 + h) - T(x_0) = A_{x_0}h + r(x_0, h), \tag{11.2.0.13}$$

where

$$\lim_{h\to 0}\frac{\|r(x_0,h)\|}{\|h\|}=0. \tag{11.2.0.14}$$

Let $\nabla T(x_0) := A_{x_0}$.

You will find basic properties of ∇ proven in [**?**] (VII).

Definition 11.2.0.12. *Let* X, Y *be Banach spaces and* U *be an open subset of* X. *Let's define by induction.*

1. $L_1 := L(X, Y)$,
2. $T \in C^1(U, Y)$ *iff.* $\nabla T \in C(U, L_1)$, *i.e.* $U \ni x \to (\nabla T(x)) \in L_1$.
3. $L_{n+1} := L(X, L_n)$,
4. $\nabla^{n+1}T(x_0) := \nabla(\nabla^n(T))(x_0)$,
5. $T \in C^{n+1}(U, X)$ *iff.* $T \in C^n(U, Y)$ *and* $\nabla(\nabla^n T) \in C(U, L_{n+1})$, *i.e.* $U \ni x \to \nabla(\nabla^n T)(x) \in L_{n+1}$ *is continuous.*

Because of the isometry

$$L(X_1\times\cdots\times X_n; Y) \cong L(X_1, L(X_2, \dots, L(X_{n-1}, L(X_n, Y))\dots)) \tag{11.2.0.15}$$

proven in [**?**] (VII.7), we have

$$L_n \cong L(\underbrace{X, \dots, X}_{\text{n times}}; Y). \tag{11.2.0.16}$$

and we can exchangeably use

$$\nabla^n T(x_0)(h_1, h_2, ..., h_n) := (\dots(\nabla^n T(x_0)h_1)h_2\dots)h_n. \tag{11.2.0.17}$$

The below propositions will be written in Einstein summation convention.

Proposition 11.2.0.13. *Let* U *be an open subset of* R^n. *If* $f \in C^1(U, \mathbb{R}^m)$, *then*

$$\nabla f(x)(\Delta x) = \frac{\partial f^i}{\partial x_j}(x)\Delta x^j. \tag{11.2.0.18}$$

Proposition 11.2.0.14.

$$\nabla^n f(x)(\Delta x_1, \dots, \Delta x_n) = \frac{\partial^n f^i}{\partial x_{j_1}\dots\partial x_{j_1}}(x)\Delta x_1^{j_1}\dots\Delta x_n^{j_n}. \tag{11.2.0.19}$$

In the theorem below $h^{(k)}$ is just an abbreviation of $(\underbrace{h, \ldots, h}_{\text{k times}})$.

Theorem 11.2.0.15. ***(The Taylor Formula)*** *Let X, Y be Banach spaces, U be an open subset of X and $T \in C^{n+1}(U, Y)$. If the interval $[x, x+h] \subset U$ then*

$$T(x+h) = T(x) + \sum_{k=1}^{n} \frac{1}{k!}(\nabla^k T(x))h^{(k)} + R_{n+1}(x)h^{(n+1)}, \qquad (11.2.0.20)$$

where

$$R_{n+1}(x) = \int_0^1 \frac{(1-s)^n}{n!}(\nabla^{n+1}T)(x+sh)ds. \qquad (11.2.0.21)$$

Proof. You may find a proof in [**?**] (VII.9). □

Theorem 11.2.0.16. *Let V, Y be normed vector spaces and let X be a topological space. If $A : X \to L(V, Y)$ is continuous at x_0 and $u : X \to V$ is continuous at x_0, then the mapping $X \ni x \to A(x)u(x) \in Y$ is continous at x_0.*

Theorem 11.2.0.17. *Let E be a compact Hausdorff space, let X, Y be Banach spaces. We consider $C(E, X)$ and $C(E, Y)$ as Banach spaces with supremum norm $\|\|_\infty$. If $A \in C(E, L(X, Y))$ and $\hat{A} : C(E, X) \to C(E, Y)$ is defined as*

$$\hat{A}(u)(t) = A(t)u(t) \qquad (11.2.0.22)$$

for each $t \in E$, then $\hat{A} \in L(C(E, X), C(E, Y))$.

Theorem 11.2.0.18. *Let E be a compact Hausdorff space, let X, Y be Banach spaces and let $\mathcal{L} \in C^2(X, Y)$. We consider $C(E, X)$ and $C(E, Y)$ as Banach spaces with supremum norm $\|\|_\infty$. Let $\hat{\mathcal{L}} : C(E, X) \to C(E, Y)$ be defined as*

$$\hat{\mathcal{L}}(u)(t) = \mathcal{L}(u(t)), \qquad (11.2.0.23)$$

then $\nabla \hat{\mathcal{L}}$ exists for all $u \in C(E, X)$ and

$$(\nabla \hat{\mathcal{L}}(u)(\Delta u))(t) = \nabla \mathcal{L}(u(t))(\Delta u(t)). \qquad (11.2.0.24)$$

11.2.1 Fourier Transforms and Related Theorems

In this subsection, if not stated otherwise, we assume that $L^p(X)$ is a space of complex valued functions $f : X \to \mathbb{C}$ for which $\int |f|^p < +\infty$.

Theorem 11.2.1.1. *Let X be a measurable space with a positive measure μ. If $f \in L^1(X)$, then*

$$\left| \int f\mu \right| \leq \int |f| d\mu. \tag{11.2.1.1}$$

Proof. [see **?** , Abstract Integration] □

Theorem 11.2.1.2. *(**Hölder's inequality**) Let X be a measurable space with a positive measure μ. Let $p, q \in (1, +\infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. If $f, g : X \to [0, \infty]$ are measurable, then*

$$\int fg d\mu \leq \left(\int f^p d\mu \right)^{\frac{1}{p}} \left(\int g^p d\mu \right)^{\frac{1}{p}}. \tag{11.2.1.2}$$

Proof. [see **?** , L^p-Spaces] □

Definition 11.2.1.3. *Let $f \in L^1(\mathbb{R})$.*

$$\mathcal{F}(f)(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixt} dx. \tag{11.2.1.3}$$

In many equations with Fourier transform we will be using u as identity function $u(t) := t$ but we will be omitting t for a purpose, making a silent assumption that everything is a function of t. In this convention i.e. $uf(u)$ is the function $t \to u(t)f(u(t)) = tf(t)$.

Theorem 11.2.1.4. *If $f \in L^1(\mathbb{R})$, then $\mathcal{F}(f)(t) \in \mathbb{R}$ is well defined for each $t \in \mathbb{R}$.*

Proof. You may find a proof in [**?**] (Fourier Transforms. Formal Properties.) □

Definition 11.2.1.5. *Let $f \in L^1(\mathbb{R})$.*

$$\mathcal{F}^{-1}(f)(t) := \mathcal{F}(f)(-t). \tag{11.2.1.4}$$

Theorem 11.2.1.6. *If $f \in L^1(\mathbb{R})$ and $\mathcal{F}(f) \in L^1(\mathbb{R})$, then*

$$f \underset{a.e.}{=} \mathcal{F}^{-1}(\mathcal{F}(f)) \in C_0(\mathbb{R}). \tag{11.2.1.5}$$

Proof. You may find a proof in [**?**] (Fourier Transforms. The Inversion Theorem.) □

Theorem 11.2.1.7. *If $f \in L^1(\mathbb{R})$, then*

$$\mathcal{F}^{-1}(f) = -\mathcal{F}(f(-u)). \tag{11.2.1.6}$$

Proof. By definition 11.2.1.5 we have

$$\mathcal{F}^{-1}(f)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ixt} dx. \tag{11.2.1.7}$$

Make the substitution $z = -x$, so $dz = -dx$. Thus

$$\mathcal{F}^{-1}(f)(t) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(-z) e^{-izt} dz. \tag{11.2.1.8}$$

□

Theorem 11.2.1.8. *Let $f \in L^1(\mathbb{R})$ and let $\alpha, \gamma \in \mathbb{R}$. The following conditions holds:*

1. $\mathcal{F}(f(u)e^{i\alpha u})(t) = \mathcal{F}(f)(t - \alpha)$.
2. $\mathcal{F}(f(u - \alpha))(t) = e^{-i\alpha t}\mathcal{F}(f)(t)$.
3. $\mathcal{F}(\,\overline{f(-u)}\,) = \overline{\mathcal{F}(f)}$.
4. *If* $\gamma > 0$ *then* $\mathcal{F}(f(\frac{u}{\gamma}))(t) = \gamma\mathcal{F}(f)(\gamma t)$.
5. *If* $-iuf(u) \in L^1(\mathbb{R})$, *then* $\mathcal{F}(f)$ *is differentiable and*

$$\frac{d}{dt}\mathcal{F}(f) = \mathcal{F}(-iuf(u)). \tag{11.2.1.9}$$

Proof. You may find a proof in [**?**] (Fourier Transforms. Formal Properties.) □

Theorem 11.2.1.9. *If $iuf(u) \in L^1(\mathbb{R})$, then*

$$\frac{d}{dt}\mathcal{F}^{-1}(f) = \mathcal{F}^{-1}(iuf(u)). \tag{11.2.1.10}$$

Proof.

$$\begin{aligned}\frac{d}{dt}\mathcal{F}^{-1}(f) &= -\frac{d}{dt}\mathcal{F}(f(-u)) = -\mathcal{F}(-iuf(-u)) = \mathcal{F}(iuf(-u)) \\ &= -\mathcal{F}^{-1}(i(-u)f(u)) = \mathcal{F}^{-1}(iuf(u)).\end{aligned} \tag{11.2.1.11}$$

□

Theorem 11.2.1.10. *(**The Parseval Formula**) If $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} f\overline{g}\, dx = \int_{-\infty}^{\infty} \mathcal{F}(f)\overline{\mathcal{F}(g)}\, dx. \tag{11.2.1.12}$$

Proof. You may find a proof in [?] (Fourier Transforms. The Plancherel Theorem.) □

Corollary 11.2.1.11. *If $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} f\overline{g}\,dx = \int_{-\infty}^{\infty} \mathcal{F}^{-1}(f)\overline{\mathcal{F}^{-1}(g)}\,dx. \tag{11.2.1.13}$$

Corollary 11.2.1.12. *If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then $\mathcal{F}(f) \in L^2(\mathbb{R})$.*

Theorem 11.2.1.13. *If $\psi_1, \psi_2 \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and the function $x \to x\mathcal{F}(\psi_1)(x)$ is in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} -i\frac{d\psi_1}{dx}(x) \cdot \overline{\psi_2(x)}\,dx = \int_{-\infty}^{\infty} x\mathcal{F}(\psi_1)(x) \cdot \overline{\mathcal{F}(\psi_2)(x)}\,dx. \tag{11.2.1.14}$$

Proof. Let

$$P = \int_{-\infty}^{\infty} \frac{d\psi_1}{dx}(x) \cdot \overline{\psi_2(x)}\,dx = \int_{-\infty}^{\infty} \frac{d}{dx}\mathcal{F}^{-1}(\mathcal{F}(\psi_1)) \cdot \overline{\mathcal{F}^{-1}(\mathcal{F}(\psi_2))}dx. \tag{11.2.1.15}$$

By Theorem 11.2.1.9 we have

$$P = \int_{-\infty}^{\infty} \mathcal{F}^{-1}(ix\mathcal{F}(\psi_1)(x)) \cdot \overline{\mathcal{F}^{-1}(\mathcal{F}(\psi_2))}\,dx. \tag{11.2.1.16}$$

Now, by Corollary 11.2.1.11 we have

$$P = \int_{-\infty}^{\infty} ix\mathcal{F}(\psi_1)(x) \cdot \overline{\mathcal{F}(\psi_2)(x)}\,dx, \tag{11.2.1.17}$$

which completes the proof. □

Theorem 11.2.1.14. *If $\psi_1, \psi_2 \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and the function $x \to x\psi_1(x)$ is in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} x\psi_1(x) \cdot \overline{\psi_2(x)}\,dx = \int_{-\infty}^{\infty} i\frac{d}{dx}\mathcal{F}(\psi)(x) \cdot \overline{\mathcal{F}(\psi)(x)}\,dx. \tag{11.2.1.18}$$

Proof. Let

$$Q = \int_{-\infty}^{\infty} x\psi_1(x) \cdot \overline{\psi_2(x)}\,dx. \tag{11.2.1.19}$$

By Theorem 11.2.1.10 we have

$$Q = \int_{-\infty}^{\infty} \mathcal{F}(x\psi_1(x)) \cdot \overline{\mathcal{F}(\psi_2(x))}\,dx = \int_{-\infty}^{\infty} i\mathcal{F}(-ix\psi_1(x)) \cdot \overline{\mathcal{F}(\psi_2(x))}\,dx. \tag{11.2.1.20}$$

Now, by Theorem 11.2.1.8

$$Q = \int_{-\infty}^{\infty} i \frac{d}{dx} \mathcal{F}(\psi_1)(x) \cdot \overline{\mathcal{F}(\psi_2)(x)}\, dx, \tag{11.2.1.21}$$

which completes the proof. □

Lemma 11.2.1.15. *Let $p_1, p_2 \in [1, +\infty)$ and $f : \mathbb{R} \to \mathbb{R}$. If $f \in L^{p_1}(\mathbb{R})$ and $\dfrac{df}{dx} \in L^{p_2}(\mathbb{R})$, then*

$$\lim_{x \to -\infty} f(x) = 0, \tag{11.2.1.22}$$

and

$$\lim_{x \to +\infty} f(x) = 0. \tag{11.2.1.23}$$

Proof. Without lost of generality we will show only $\lim_{x\to+\infty} f(x) = 0$. By proof by contracdiction, assume that $\limsup\limits_{x\to\infty} |f(x)| > 0$. Without lost of generality we may assume that we have such $\varepsilon > 0$ and a sequence $b_n \to \infty$ such that $f(b_n) > \varepsilon$. Fix some $\Delta s > 0$. There must be

$$\inf_{x \in (c_n - \Delta s, c_n)} f(x) < \frac{1}{2}\varepsilon \tag{11.2.1.24}$$

for almost all $n \in \mathbb{N}$. Otherwise $f \notin L^{p_1}(\mathbb{R})$. Hence, we have a sequence $a_n \to \infty$ such that $a_n \in (b_n - \Delta s, b_n)$ and $f(a_n) < \frac{1}{2}\varepsilon$ for almost all $n \in \mathbb{N}$. Since $f(b_n) > \varepsilon$ and $f(a_n) < \frac{1}{2}\varepsilon$ we have

$$\int_{a_n}^{b_n} \frac{df}{dx} = f(b_n) - f(a_n) > \frac{1}{2}\varepsilon \tag{11.2.1.25}$$

for almost all $n \in \mathbb{N}$. This imediatelly contradicts $\frac{df}{dx} \in L^{p_2}(\mathbb{R})$ for $p_2 = 1$. Assume now that $p_2 > 1$. By Theorem 11.2.1.1 and Theorem 11.2.1.2 (Hölder's inequality), for almost all $n \in \mathbb{N}$ we are getting

$$\frac{1}{2}\varepsilon < \int_{a_n}^{b_n} \left| \frac{df}{dx} \right| \cdot 1 dx \le \left(\int_{a_n}^{b_n} \left| \frac{df}{dx} \right|^{p_2} dx \right)^{\frac{1}{p_2}} (b_n - a_n)^{\frac{1}{q_2}}, \tag{11.2.1.26}$$

where q_2 is such that $\frac{1}{p_2} + \frac{1}{q_2} = 1$. But $b_n - a_n < \Delta s$. Thus

$$\left(\frac{\varepsilon}{(\Delta s)^{\frac{1}{q_2}}} \right)^{p_2} < \left(\frac{\varepsilon}{(b_n - a_n)^{\frac{1}{q_2}}} \right)^{p_2} < \int_{a_n}^{b_n} \left| \frac{df}{dx} \right|^{p_2} dx. \tag{11.2.1.27}$$

Since $b_n \to \infty$ the above contradicts $\dfrac{df}{dx} \in L^{p_2}(\mathbb{R})$. □

The assumption that $\frac{df}{dx} \in L^{p_2}(\mathbb{R})$ is important, for counterexample [see **?** , Example 2].

Theorem 11.2.1.16. *Let $p_1, p_2 \in [1, +\infty)$ and $f : \mathbb{R} \to \mathbb{C}$. If $f \in L^{p_1}(\mathbb{R})$ and $\frac{df}{dx} \in L^{p_2}(\mathbb{R})$, then*

$$\lim_{x \to -\infty} f(x) = 0, \tag{11.2.1.28}$$

and

$$\lim_{x \to +\infty} f(x) = 0. \tag{11.2.1.29}$$

Proof. It is enough to notice $f(x) = \mathrm{Re}(f(x)) + i\mathrm{Im}(f(x))$. From $\max\{|a|, |b|\} \leq |a+ib|$ follows that $\mathrm{Re}(f), \mathrm{Im}(f) \in L^{p_1}(\mathbb{R})$ and $\frac{d\mathrm{Re}f}{dx}, \frac{d\mathrm{Im}f}{dx} \in L^{p_2}(\mathbb{R})$. Thus, applying the above Lemma to both $\mathrm{Re}f$ and $\mathrm{Im}f$ completes the proof. □

Theorem 11.2.1.17. *Let $\mathcal{D} = \{f : \mathbb{R} \to \mathbb{C} | f \in L^2(\mathbb{R})$ and $\frac{df}{dx} \in L^2(\mathbb{R})\}$. If $f, g \in \mathcal{D}$, then*

$$\int_{-\infty}^{\infty} i\frac{df}{dx} \cdot \overline{g}\, dx = \int_{-\infty}^{\infty} f \cdot \overline{i\frac{dg}{dx}}\, dx. \tag{11.2.1.30}$$

Proof. Note that by Theorem 11.2.1.16 we have

$$\int_{-\infty}^{\infty} i\frac{d}{dx}(f\overline{g}) = 0. \tag{11.2.1.31}$$

From the above, we get the equation (11.2.1.30). □

11.2.2 Euler-Lagrange Equations

Let $I = [a, b]$. In this subsection we will consider a function $\mathcal{L} \in C^2(I \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$. We will also consider a space $C^1(I, \mathbb{R}^n)$ equipped with the norm

$$\|u\|_{C^2} := \sum_{i=1}^{n} \sup |u_i(I)| + \sum_{i=1}^{n} \sup |\frac{du_i}{dt}(I)|. \tag{11.2.2.1}$$

By Theorem 11.2.0.2 and Theorem 11.2.0.6 one can easily prove that $(C^1(I, \mathbb{R}^n), \|\cdot\|_{C^2})$ is a Banach space. Define a function J on $C^1(I, \mathbb{R}^n)$.

$$J(u) := \int_{b}^{a} \mathcal{L}(t, u(t), \frac{du}{dt}(t))dt. \tag{11.2.2.2}$$

Definition 11.2.2.1. *We will say that $F_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous symmetry group iff.*

$$F_\varepsilon(x) = v(\varepsilon) + A_\varepsilon x, \tag{11.2.2.3}$$

where $\frac{d}{d\varepsilon} v|_{\varepsilon=0} \in \mathbb{R}$ and A_ε is a continuous group of linear mappings, for which there exists a linear mapping Q_A such that

$$\lim_{\varepsilon\to 0} \left\| \frac{1}{\varepsilon}(A_\varepsilon - I) - Q_A \right\| = 0. \tag{11.2.2.4}$$

We will say that $Q_F := \frac{d}{d\varepsilon} v|_{\varepsilon=0} + Q_A$ is an infinitesimal generator of F_ε.

Lemma 11.2.2.2. *Let F_ε be a continuous symmetry group, let $u \in C^1(I, \mathbb{R}^n)$. Let*

$$\Phi(t, \varepsilon) = F_\varepsilon(u(t)). \tag{11.2.2.5}$$

Then

$$\frac{d}{dt}\frac{d}{d\varepsilon}\Phi(t,\varepsilon)\bigg|_{\varepsilon=0} = \frac{d}{d\varepsilon}\frac{d}{dt}\Phi(t,\varepsilon)\bigg|_{\varepsilon=0} = Q_F \dot{u}(t). \tag{11.2.2.6}$$

Theorem 11.2.2.3. *(**Noether's Theorem**) Let $\mathcal{L} \in C^2(I \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$, let $u \in C^1(\mathbb{R}, \mathbb{R}^n)$ and*

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}(t, u, \dot{u}) - \frac{\partial \mathcal{L}}{\partial x}(t, u, \dot{u}) = 0. \tag{11.2.2.7}$$

Let F_ε be a continuous symmetry group with an infinitesimal generator Q_F and let $Q_T \in \mathbb{R}$. Let $T_\varepsilon(t) = t + Q_T\varepsilon$ and $u_\varepsilon(t) = F_\varepsilon(u(t))$. If

$$\mathcal{L}(T_\varepsilon, u_\varepsilon, \dot{u}_\varepsilon) = \mathcal{L}(t, u, \dot{u}) \tag{11.2.2.8}$$

for all $\varepsilon \in \mathbb{R}$, then

$$(\mathcal{L} - \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{u})Q_T + \frac{\partial \mathcal{L}}{\partial \dot{x}} Q_F u = const. \tag{11.2.2.9}$$

Proof. Let

$$\psi_t(\varepsilon) := (t + Q_T\varepsilon, F_\varepsilon(u(t)), \frac{d}{dt}F_\varepsilon(u(t))). \tag{11.2.2.10}$$

For fixed $t \in \mathbb{R}$, ψ_t is differentiable in $\varepsilon = 0$ and by Lemma 11.2.2.2, we have

$$\frac{d}{d\varepsilon}\psi_t(\varepsilon)\bigg|_{\varepsilon=0} = (Q_T, Q_F u(t), Q_F \dot{u}(t)). \tag{11.2.2.11}$$

By equation (11.2.2.8) we have

$$\mathcal{L}(\psi_t(\varepsilon)) - \mathcal{L}(\psi_t(0)) = 0 \tag{11.2.2.12}$$

for all $\varepsilon \in \mathbb{R}$. Thus

$$\frac{d}{d\varepsilon}\mathcal{L}(\psi_t(\varepsilon))\Big|_{\varepsilon=0} = 0. \tag{11.2.2.13}$$

By the law of derivatives composition ([**?**] VII.4), we have

$$\nabla\mathcal{L}(\psi_t(0))\frac{d}{d\varepsilon}\psi_t(\varepsilon)\Big|_{\varepsilon=0} = 0, \tag{11.2.2.14}$$

Since $\psi_t(0) = (t, u(t), \dot{u}(t))$, the above expands to

$$\frac{\partial\mathcal{L}}{\partial t}(t,u,\dot{u})Q_T + \frac{\partial\mathcal{L}}{\partial x}(t,u,\dot{u})Q_F u + \frac{\partial\mathcal{L}}{\partial \dot{x}}(t,u,\dot{u})Q_F \dot{u} = 0. \tag{11.2.2.15}$$

By (11.2.2.7)

$$\frac{\partial\mathcal{L}}{\partial t}(t,u,\dot{u})Q_T + \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial \dot{x}}(t,u,\dot{u})Q_F u + \frac{\partial\mathcal{L}}{\partial \dot{x}}(t,u,\dot{u})Q_F \dot{u} = 0. \tag{11.2.2.16}$$

From that point we will omit $(t, u, \dot{u})$ but we will consider $\mathcal{L}$ and its all derivatives at point $(t, u(t), \dot{u}(t))$. From the above we get immediately

$$\frac{\partial\mathcal{L}}{\partial t}Q_T + \frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial \dot{x}}Q_F u) = 0. \tag{11.2.2.17}$$

Note that we have

$$\frac{d}{dt}\mathcal{L} = \frac{\partial\mathcal{L}}{\partial t} + \frac{\partial\mathcal{L}}{\partial x}\dot{u} + \frac{\partial\mathcal{L}}{\partial \dot{x}}\ddot{u}. \tag{11.2.2.18}$$

But again from (11.2.2.7) we get

$$\frac{d}{dt}\mathcal{L} = \frac{\partial\mathcal{L}}{\partial t} + \frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial \dot{x}}\dot{u}). \tag{11.2.2.19}$$

Thus using (11.2.2.17) we get

$$\frac{d}{dt}\big((\mathcal{L} - \frac{\partial\mathcal{L}}{\partial \dot{x}}\dot{u})Q_T + \frac{\partial\mathcal{L}}{\partial \dot{x}}Q_F u\big) = 0, \tag{11.2.2.20}$$

which gives (11.2.2.9). □

11.3 Spectral Theory

11.3.1 Spectral Measure

For any vector space X, by $\mathcal{B}(X)$ we will mean a space of all bounded linear operators.

Definition 11.3.1.1. *Let $\mathfrak{M}$ be a σ-algebra in a set Ω and let H be a Hilbert space. The mapping*

$$E : \mathfrak{M} \to \mathcal{B}(H) \tag{11.3.1.1}$$

is a spectral measure iff

1. $E(\Omega) = I$.
2. $E(\omega)$ *is a selfadjoint projection for any* $\omega \in \mathfrak{M}$.
3. $E(\omega_1 \cup \omega_2) = E(\omega_1) + E(\omega_2)$ *where* $\omega_1 \cap \omega_2 = \emptyset$ *for any* $\omega_1, \omega_2 \in \mathfrak{M}$.
4. $E(\bigcup_{i=1}^{\infty} \omega_i)\psi = \sum_{i=1}^{\infty} E(\omega_i)\psi$ *for any pairwise disjoint sequence of sets* $\omega_i \in \mathfrak{M}$ *and any* $\psi \in H$.

Theorem 11.3.1.2. *Let $\mathfrak{M}$ be a σ-algebra in a set Ω and let H be a Hilbert space. If $E : \mathfrak{M} \to \mathcal{B}(H)$ is a spectral measure, then the following holds:*

1. $E(\Omega \setminus \omega) = I - E(\omega)$ *for any* $\omega \in \mathfrak{M}$, *in particular* $E(\emptyset) = 0$.
2. $E(\omega_1 \cap \omega_2) = E(\omega_1)E(\omega_2)$ *for any* $\omega_1, \omega_2 \in \mathfrak{M}$.
3. $\mu^E_{\psi,\phi}(\omega) := \langle E(\omega)\psi, \phi\rangle$ *is a complex valued measure on* $\mathfrak{M}$.

Proof. [see **?** , The Spectral Theorem] □

We will denote

$$\mu^E_{\psi}(\omega) := \mu^E_{\psi,\psi}(\omega). \tag{11.3.1.2}$$

Note that because of $E(\omega)$ is a projection $\mu^E_{\psi}(\omega) \geq 0$.

Definition 11.3.1.3. *Let H be a Hilber space. We say that E is a spectral measure on real line iff $E : \mathfrak{M} \to \mathcal{B}(H)$ where $\mathfrak{M}$ is σ-algebra of Lebesgue measurable sets on $\mathbb{R}$.*

Definition 11.3.1.4. *Let H be a Hilbert space, let $f : \mathbb{R} \to \mathbb{C}$ be a measurable function and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure.*

$$\mathcal{D}^E_f := \{\psi \in H : \int |f|^2 d\mu^E_{\psi} < +\infty\}. \tag{11.3.1.3}$$

Theorem 11.3.1.5. *Let $\mathcal{D}_f^E$ be like in Definition 11.3.1.4, then $\mathcal{D}_f^E$ is a dense subspace of H.*

Proof. [see **?** , Unbounded Operators on a Hilbert Space, Resolution of Identity] □

Theorem 11.3.1.6. *Let H be a Hilbert space, let $f : \mathbb{R} \to \mathbb{C}$ be a measurable function and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure. Then*

$$\int |f|d|\mu^E_{\psi,\phi}| \leq ||\phi|| \left(\int |f|^2 d\mu^E_\psi \right)^{\frac{1}{2}} \text{ for } \psi, \phi \in H. \tag{11.3.1.4}$$

Proof. [see **?** , Unbounded Operators on a Hilbert Space, Resolution of Identity] □

Theorem 11.3.1.7. *Let H be a Hilbert space, let $f : \mathbb{R} \to \mathbb{C}$ be a measurable function and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure. There exists a densly defined normal operator $\Psi^E(f)$ with a domain $\mathcal{D}(\Psi^E(f)) = \mathcal{D}_f^E$ such that*

$$\langle \Psi^E(f)\psi, \phi \rangle = \int f d\mu^E_{\psi,\phi} \text{ for all } \psi \in \mathcal{D}_f^E \text{ and } \phi \in H. \tag{11.3.1.5}$$

Moreover

1. $\left\| \Psi^E(f)\psi \right\| = \int |f|^2 d\mu^E_\psi$ *for all* $\psi \in \mathcal{D}_f^E$.

2. $\Psi^E(f)^* = \Psi^E(\overline{f})$.

3. *If* $f, g : \mathbb{R} \to \mathbb{C}$ *are measurable, then* $\Psi^E(f)\Psi^E(g) \subset \Psi^E(fg)$ *and* $\mathcal{D}(\Psi^E(f)\Psi^E(g)) = \mathcal{D}_g^E \cap \mathcal{D}_{fg}^E$.

4. $\Psi^E(f)\Psi^E(g) = \Psi^E(fg)$ *iff* $\mathcal{D}_{fg}^E \subset \mathcal{D}_g^E$.

Proof. [see **?** , Unbounded Operators on a Hilbert Space, Resolution of Identity] □

Corollary 11.3.1.8. *If E is a spectral measure on real line, then for Ψ^E from Theorem 11.3.1.7, $\Psi^E(id)$ is self-adjoint.*

Proof. It follows from Therem 11.3.1.7 Moreover part point 1. □

We will use symbol $E(f) := \Psi^E(f)$ as defined in the theorem above. Note that in this convention $E(1_\omega) = E(\omega)$. We will use also another convention, if we consider operator $A = E(\mathrm{id})$, we will write $f(A) := \Psi^E(f)$.

Definition 11.3.1.9. *Let H be Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. Let $\psi \in H$.*

$$H_\psi^E := \{E(f)\psi : \int |f|^2 d\mu_\psi^E < +\infty\}. \tag{11.3.1.6}$$

Definition 11.3.1.10. *Let H be Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. Let $\psi \in H$. We will say that ψ is **cyclic** in E iff $H_\psi^E = H$.*

Definition 11.3.1.11. *Let H_1, H_2 be Hilbert spaces. $U : H_1 \to H_2$ is called a unitary mapping, if U is a linear bijection and*

$$\langle U(\psi), U(\phi)\rangle_{H_2} = \langle \psi, \phi\rangle_{H_1} \text{ for all } \psi, \phi \in H_1. \tag{11.3.1.7}$$

Definition 11.3.1.12. *Let H be Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. Let $\psi \in H$. We define*

$$U_\psi^E : H_\psi^E \to L^2(\mathbb{R}, \mu_\psi^E), \tag{11.3.1.8}$$

such that

$$U_\psi^E(E(f)\psi) := f \text{ for all } f \in L^2(\mathbb{R}, \mu_\psi^E). \tag{11.3.1.9}$$

Theorem 11.3.1.13. *Let H be Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. Let $\psi \in H$ Then U_ψ^E is a unitary mapping between H_ψ and $L^2(\mathbb{R}, \mu_\psi^E)$ and*

$$U_\psi^E(E(f)\phi) = f \cdot U_\psi^E(\phi) \text{ for any } f \in L^2(\mathbb{R}, \mu_\psi^E) \text{ and } \phi \in H_\psi \cap \mathcal{D}_f. \tag{11.3.1.10}$$

Moreover $U_f^E(\mathcal{D}_f \cap H_\psi) = \mathcal{D}(g \to f \cdot g)$ for any $f \in L^2(\mathbb{R}, \mu_\psi^E)$.

Proof. [see **?** , The Spectral Theorem] Note that to prove equation (11.3.1.10), you represent ϕ as $\phi = E(g)\psi$ for some $g \in L^2(\mathbb{R}, \mu_\psi^E)$ and then use Theorem 11.3.1.7 for

$$U_\psi^E(E(f)\phi) = U_\psi^E(E(f)E(g)\psi) = U_\psi^E(E(fg))\psi = fg = f \cdot U_\psi^E(\phi). \tag{11.3.1.11}$$

□

Definition 11.3.1.14. *Let H be a seperable Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. A set of vectors $\{\psi_i\}_{i\in J}$ is a spectral basis with respect to a spectral measure on real line E iff*

1. *$\|\psi_i\| = 1$ for each $i \in J$.*

2. $H^E_{\psi_i} \perp H^E_{\psi_j}$ *for all* $i, j \in J$ *where* $i \neq j$.

3. $H = \bigoplus H^E_{\psi_i}$.

Theorem 11.3.1.15. *Let H be a separable Hilbert space and let $E : \mathfrak{M} \to \mathcal{B}(H)$ be a spectral measure on real line. There exist at least countanable spectral basis $\{\psi_i\}_{i \in J}$ and an unitary mapping U^E such that*

$$U^E : H \to \bigoplus L^2(\mathbb{R}, \mu^E_{\psi_i}), \tag{11.3.1.12}$$

and for any measurable function $f : \mathbb{R} \to \mathbb{C}$ and $\psi \in \mathcal{D}_f$.

$$U^E(E(f)\psi) = f^M U^E(\psi) \text{ and } U^E(\mathcal{D}_f) = \mathcal{D}(f^M), \tag{11.3.1.13}$$

where $f^M := \bigoplus f^M_i$ and $f^M_i(g) := f \cdot g$ on $L^2(\mathbb{R}, \mu^E_{\psi_i})$, in particular

$$E(\omega)\psi = (U^E)^{-1}(1^M_\omega \cdot U^E(\psi)). \tag{11.3.1.14}$$

Proof. [see **?** , The Spectral Theorem] □

Treat the above theorem also as a definition of U^E. Note the equation (11.3.1.14) shows that having a unitary mapping U^E we can reconstruct spectral measure E.

Theorem 11.3.1.16. *Let H be a Hilbert space and $U : H \to L^2(\mathbb{R}, \nu)$ be a unitary mapping. If*

$$E(\omega)\psi = U^{-1}(1_\omega \cdot U(\psi)) \text{ for all } \psi \in H \tag{11.3.1.15}$$

and any Lebesgue measurable $\omega \subset \mathbb{R}$, then E is a spectral measure on real line and $A = E(id)$ is a densly defined self-adjoint linear operator. Moreover

$$\langle f(A)\psi, \phi\rangle = \int_{-\infty}^{\infty} f(\lambda)U(\psi)\overline{U(\phi)}d\nu(\lambda) \tag{11.3.1.16}$$

and

$$f(A)\psi = U^{-1}(f \cdot U(\psi)) \tag{11.3.1.17}$$

for any measurable $f : \mathbb{R} \to \mathbb{C}$ and all $\psi \in \mathcal{D}^E_f = \mathcal{D}(f(A))$ and $\phi \in H$.

11.3.2 Spectral Measure - Multidimentional representation

Theorem 11.3.2.1. *Let H be a Hilbert space and let $(X, \mathfrak{M}, \nu)$ be a measurable space. Let $U : H \to L^2(X, \nu)$ be a unitary mapping and $g : X \to \mathbb{R}$ be a measurable function.*

$$E(\omega)\psi = U^{-1}(1_\omega(g) \cdot U(\psi)) \text{ for all } \psi \in H, \tag{11.3.2.1}$$

and any Lebesgue measurable $\omega \subset \mathbb{R}$, then E is a spectral measure on real line and $A = E(id)$ is a densly defined self-adjoint linear operator. Moreover

$$\langle f(A)\psi, \phi\rangle = \int_{-\infty}^{\infty} f(g(x))U(\psi)\overline{U(\phi)}d\nu \tag{11.3.2.2}$$

and

$$f(A)\psi = U^{-1}(f(g) \cdot U(\psi)) \tag{11.3.2.3}$$

for any measurable $f : \mathbb{R} \to \mathbb{C}$ and all $\psi \in \mathcal{D}(f(A))$ and $\phi \in H$. Also

$$\mathcal{D}(f(A)) = D_f^E = \{\psi \in H : f(g) \cdot U(\psi) \in L^2(X, \nu)\}. \tag{11.3.2.4}$$

Proof. Note first that if $u \in L^1(X, \nu)$ then

$$\mu(\omega) = \int 1_\omega(g)ud\nu \tag{11.3.2.5}$$

is a finite measure on $\mathbb{R}$. From that show that $\langle E(\omega)\psi, \phi\rangle = \int 1_\omega(g)U(\psi)\overline{U(\phi)}d\nu$. is a complex valued measure on $\mathbb{R}$ for fixed $\psi, \phi \in H$. From that it is easy to show that E is a spectral measure and all other properties using Theorem 11.3.1.7. Self-adjointness of A follows easily from Corollary 11.3.1.8. To prove (11.3.2.4), recall that by Definition 11.3.1.4 and Theorem 11.3.1.7 we have

$$\mathcal{D}(f(A)) = \mathcal{D}_f^E = \{\psi \in H : \int |f|^2 d\mu_\psi^E < +\infty\}. \tag{11.3.2.6}$$

But since (11.3.2.5),

$$\int |f|^2 d\mu_\psi^E = \int |f(g)|^2 U(\psi)\overline{U(\psi)}d\nu. \tag{11.3.2.7}$$

And this proves (11.3.2.4). □

The following form of "inverse" theorem is possible.

Theorem 11.3.2.2. *(spectral theorem–multiplication operator form) Let H be a seperale Hilbert space and let A be a densly defined self-adjoint operator. Then, there exists a measurable space $(X, \mathfrak{M}, \nu)$ with $\nu(X) < +\infty$, a unitary mapping $U : H \to L^2(X, \nu)$ and a measurable function $g : X \to \mathbb{R}$ such that*

$$\psi \in \mathcal{D}(A) \text{ iff } g \cdot U(\psi) \in L^2(X, \nu) \tag{11.3.2.8}$$

and

$$A\psi = U^{-1}(g \cdot U(\psi)) \text{ for any } \psi \in \mathcal{D}(A). \tag{11.3.2.9}$$

Proof. [see **?** , VIII.3 The spectral theorem] □

Definition 11.3.2.3. *Let H be a Hilbert space. (X, ν, U, g) is L^2-representation of an self-adjoint operator A iff $U : H \to L^2(X, \nu)$ is a unitary mapping and $g : X \to \mathbb{R}$ is measurable, such that*

$$\mathcal{D}(A) = \{\psi \in H : g \cdot U(\psi) \in L^2(X, \nu)\} \tag{11.3.2.10}$$

and

$$A\psi = U^{-1}(g \cdot U(\psi)) \text{ for any } \psi \in \mathcal{D}(A). \tag{11.3.2.11}$$

Theorem 11.3.2.1 shows that if we have L^2-representation of an self-adjoint operator, it is densly defined and we know what its spectral measure is. Theorem 11.3.2.1 shows that if we are in seperable Hilbert space, L^2-representation exists for each densly defined self-adjoint operator.

11.4 Multidimensional Fourier Transform and Schwartz space

We will use multindexes as introduced e.g in [**?** , V.3] By $C_0(\mathbb{R}^n)$ we denote a space of of continous coplex valued functions which converge to 0 in infinity.

Definition 11.4.0.1. *S_n will denote all functions $f \in C^\infty(\mathbb{R}^n)$ such that*

$$\sup_{x \in \mathbb{R}^n} |x^\beta (D^\alpha f)(x)| < \infty \text{ for all multi-indicies } \alpha, \beta. \tag{11.4.0.1}$$

We call S_n Schwartz space.

Theorem 11.4.0.2. *If $f \in S_n$, then $D^\alpha f \in S_n$ for any multi-index α.*

Proof. This follows directly from the Definition 11.4.0.1. □

Theorem 11.4.0.3. *If $f, g \in S_n$, then $f \cdot g \in S_n$.*

Proof. We will show this by induction. It is trivial to note that

$$\sup x^\beta |f||g| < \infty. \tag{11.4.0.2}$$

Assume that for any $f, g \in S_n$ and each index β and index α such that $|\alpha| \leq m$, we have

$$\sup |x^\beta D^\alpha (f \cdot g)| < \infty. \tag{11.4.0.3}$$

Now take any index i.

$$\sup |x^\beta D^{\alpha+i}(f \cdot g)| \leq \sup |x^\beta D^\alpha (D^i f \cdot g)| + \sup |x^\beta D^\alpha (f \cdot D^i g)| < \infty. \tag{11.4.0.4}$$

The last inequality above follows from Theorem 11.4.0.3. □

Theorem 11.4.0.4. *S_n is dense in $L^p(\mathbb{R}^n)$ for any $p \in [1, \infty)$.*

Proof. [see **?** , 3.2 Fourier Transform] □

Definition 11.4.0.5. *Let $f \in L^1(\mathbb{R}^n)$.*

$$\mathcal{F}(f)(x) := (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix \cdot s} f(s) ds. \tag{11.4.0.5}$$

Theorem 11.4.0.6. *If $f \in L^1(\mathbb{R}^n)$, then $\mathcal{F}(f) \in C_0(\mathbb{R}^n)$ and $\|\mathcal{F}(f)\|_\infty \leq \|f\|_1$.*

Proof. [see **?** , II.7 Fourier Transforms] □

Definition 11.4.0.7. *Let $f \in L^1(\mathbb{R})$.*

$$\mathcal{F}^{-1}(f)(x) := \mathcal{F}(f)(-x). \tag{11.4.0.6}$$

Theorem 11.4.0.8. *If $f \in L^1(\mathbb{R})$ and $\mathcal{F}(f) \in L^1(\mathbb{R})$, then*

$$f \underset{a.e.}{=} \mathcal{F}^{-1}(\mathcal{F}(f)) \in C_0(\mathbb{R}). \tag{11.4.0.7}$$

Proof. [see **?** , II.7 Fourier Transforms] □

Theorem 11.4.0.9. *If $f \in L^1(\mathbb{R}^k)$ then $\mathcal{F}(\overline{f}) = \overline{\mathcal{F}^{-1}(f)}$*

Proof. Follows directly from definitions. □

Theorem 11.4.0.10. *If $f, g \in S_n$, then*

$$\int f\overline{g} = \int \mathcal{F}(f)\overline{\mathcal{F}(g)}. \tag{11.4.0.8}$$

Also $\|f\|_2 = \|\mathcal{F}(f)\|_2$.

Proof. [see **?** , 3.2 Fourier Transform] □

Definition 11.4.0.11. *We will extend $\mathcal{F}$ to $L^2(\mathbb{R}^n)$. For any $f \in L^2(\mathbb{R}^n)$*

$$\mathcal{F}(f) := \lim_{m\to\infty} \mathcal{F}(f_m), \tag{11.4.0.9}$$

where $S_n \ni f_m \to f \in L^2(\mathbb{R})$.

Sometimes we will use $\mathcal{F}_n$ if we need to keep track of $\mathbb{R}^n$ dimension. Theorems 11.4.0.4 and 11.4.0.10 guarantee the existence of $\lim_{m\to\infty} \mathcal{F}(f_m)$ for each f and its independence of choice of f_m.

Theorem 11.4.0.12. *$\mathcal{F}$ is a unitary mapping from $L^2(\mathbb{R}^n)$ onto $L^2(\mathbb{R}^n)$. Moreover*

$$\mathcal{F}(f)(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\cdot s} f(s)ds \text{ for } f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n). \tag{11.4.0.10}$$

Proof. [see **?** , 3.2 Fourier Transform] □

Definition 11.4.0.13. *Let Ω be an nonempty open set in $\mathbb{R}^n$. $g \in L^p(\Omega)$ iff $g \in L^p(K)$ for each compact $K \subset \Omega$.*

Theorem 11.4.0.14. *Let Ω be an nonempty open set in $\mathbb{R}^n$. Let $u \in C^m(|Omega)$ and $\phi \in C_0^m(\Omega)$ and let α be a multi-index with $|\alpha| \leq m$. Then*

$$\int_\Omega u D^\alpha \phi = (-1)^{|\alpha|} \int_\Omega (D^\alpha u)\phi. \tag{11.4.0.11}$$

Proof. For a very elegant proof of this theorem [see **?** , 2.4]. □

Definition 11.4.0.15. *J_ε with $\varepsilon > 0$ is said to be mollifier iff $J_\varepsilon \geq 0$, $J_\varepsilon(x) = 0$ for $|x| \geq \varepsilon$ and $\int_{\mathbb{R}^n} J_\varepsilon = 1$.*

For construction of mollifier see [**?** , 3.1].

Theorem 11.4.0.16. *Let Ω be an nonempty subset of $\mathbb{R}^n$. Let $p, q \in [0, \infty)$, $u \in L^p_{loc}(\Omega)$, $v \in L^q_{loc}(\Omega)$. Let $K \subset \Omega$ be a compact set. Let α be a multi-index. If*

$$(-1)^{|\alpha|} \int_\Omega u D^\alpha \phi = \int_\Omega v\phi \text{ for all } \phi \in C_0^\infty(\Omega) \tag{11.4.0.12}$$

and

$$u_\varepsilon = \int_{K+B(0,d)} J_\varepsilon(x-y)u(y)dy, \tag{11.4.0.13}$$

$$v_\varepsilon = \int_{K+B(0,d)} J_\varepsilon(x-y)v(y)dy \tag{11.4.0.14}$$

with such a $d > 0$ that $K + Clo(B(0,d)) \subset \Omega$, $x \in \mathbb{R}^n$ and J_ε mollifier, then the following statements are true.

1. $u_\varepsilon, v_\varepsilon \in C_0^\infty(\Omega)$ *for every* $\varepsilon > 0$.

2. $(D^\alpha u_\varepsilon)(x) = v(x)$ *for* $x \in K$ *and* $0 < \varepsilon < d$.

3. $\lim\limits_{\varepsilon\to 0} \int_K |u_\varepsilon - u|^p = \lim\limits_{\varepsilon\to 0} \int_K |D^\alpha u_\varepsilon - v|^q = 0$.

4. *If* $u = 0$, *then* $v \underset{a.e.}{=} 0$ *in* Ω.

Proof. [see **?** , 3.4] □

Definition 11.4.0.17. *Let* Ω *be an nonempty subset of* $\mathbb{R}^n$. *Let* $u \in L^1_{loc}(\Omega)$ *and* α *be a multi-index. We say that* u *has the* α*th weak derivative and* $\mathsf{D}^\alpha u = v$ *iff*

$$(-1)^{|\alpha|} \int_\Omega u D^\alpha \phi = \int_\Omega v\phi \text{ for all } \phi \in C_0^\infty(\Omega). \tag{11.4.0.15}$$

Theorem 11.4.0.16 guarantees the corectness of the above definition.

Theorem 11.4.0.18. *Let* Ω *be an nonempty subset of* $\mathbb{R}^n$. *If* $u \in C^m(\Omega)$ *and* α *is an multi-index with* $|\alpha| \le m$, *then*

$$\mathsf{D}^\alpha u \underset{a.e.}{=} D^\alpha u. \tag{11.4.0.16}$$

Proof. [see **?** , 3.4] □

Theorem 11.4.0.19. *Let* α *be an multi-index and let* $f \in L^2(\mathbb{R}^n)$. *Then*

1. $x^\alpha \mathcal{F}(f)(x) \in L^2(\mathbb{R}^n)$ *iff* $\mathsf{D}^\alpha f \in L^2(\mathbb{R}^n)$.

2. *If* $\mathsf{D}^\alpha f \in L^2(\mathbb{R}^n)$ *then*

$$\mathcal{F}(\mathsf{D}^\alpha f)(x) = i^{|\alpha|} x^\alpha (\mathcal{F}(f))(x). \tag{11.4.0.17}$$

Proof. [see **?** , 3.4] □

11.4.1 Some important Integrals

The following facts can be found e.g. in [**?**]

Fact 11.4.1.1.

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1. \tag{11.4.1.1}$$

Fact 11.4.1.2.

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{ixy} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \exp\left(i\mu y - \frac{\sigma^2 y^2}{2}\right). \tag{11.4.1.2}$$

11.5 Theory of Distributions

Majority of theorems and definitions in this section are citations from [**?**] and [**?**].

In the context of topological vector spaces local base will always mean local base of open neigbourhoods of 0.

11.5.1 Measure Theory Preliminaries

Theorem 11.5.1.1. *Let $(\Omega,,\mathcal{M},\mu)$ be a measurable space. If $\int_E f d\mu = 0$ for any $E \in \mathcal{M}$, then $f \underset{a.e.}{=} 0$.*

Proof. [See **?** , 1.39] □

Lemma 11.5.1.2. *Let $(\Omega,,\mathcal{M},\mu)$ be a measurable space with μ being $\sigma-$finite. If $\int_E f d\mu = 0$ for any $E \in \mathcal{M}$ such that $\mu(E) < +\infty$, then $f \underset{a.e.}{=} 0$.*

Proof. Since μ is $\sigma-$finite, then we have $\Omega = \bigcup_{n=1}^{\infty} E_n$ where E_n are pairwise disjoint and $\mu(E_n) < +\infty$ for $n = 1, 2, \ldots$. Take any $A \in \mathcal{M}$, we have

$$\int_A f d\mu = \sum_{n=1}^{\infty} \int_{A \cap E_n} f d\mu = 0. \tag{11.5.1.1}$$

Now we have thesis by Theorem 11.5.1.1. □

11.5.2 Topological Preliminaries

Definition 11.5.2.1. *Let V be a vector space over field K. A subset $B \subset V$ is called* ***ballanced*** *iff $\lambda B \subset B$ for any $\lambda \in K$ such that $|\lambda| \leq 1$.*

Definition 11.5.2.2. *Let V be a vector space over field K. A subset $A \subset V$ is called* ***absorbing*** *iff for any $x \in V$ there exists $c_x > 0$ such that for all $\lambda \in K$ such that $|\lambda| \leq c_x$, we have $\lambda x \in A$.*

Corollary 11.5.2.3. *Let V be a vector space. Let $A \subset V$. If $0 \in IntA$, then A is absorbing.*

Definition 11.5.2.4. *Let V be a vector space over field $K \in \{\mathbb{R}, \mathbb{C}\}$. A subset $A \subset V$ is* ***convex*** *iff for any $x, y \in A$, we have $\lambda x + (1-\lambda)y \in A$ for any $\lambda \in [0,1]$.*

Theorem 11.5.2.5. *Let V be a vector space and A be a subset of V. If A is convex, then for any $\lambda_i \geq 0$ for $i = 1, \ldots, n$ such that $\sum_{i=1}^n \lambda_i = 1$ and for any $x_i \in A$ for $i = 1, \ldots, n$, we have $\sum_{i=1}^n \lambda_i x_i \in A$.*

Proof. We will prove this by induction over n. For $n = 1$ the thesis is obvious. Let's assume that thesis holds for $n-1$. Take any $\lambda_i \geq 0$ for $i = 1, \ldots, n$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and take any $x_i \in A$ for $i = 1, \ldots, n$. Let $\lambda := \sum_{i=1}^{n-1} \lambda_i$. Note that $\lambda \in [0, 1]$. By induction hypothesis $\sum_{i=1}^{n-1} \lambda_i \lambda^{-1} x_i \in A$. But since A is convex, we have

$$A \ni \lambda(\sum_{i=1}^{n-1} \lambda_i \lambda^{-1} x_i) + (1 - \lambda)x_n = \sum_{i=1}^{n-1} \lambda_i x_i + \lambda_n x_n = \sum_{i=1}^{n} \lambda_i x_i. \quad (11.5.2.1)$$

□

Definition 11.5.2.6. *We say that X is a topological vector space iff X is equiped with topology in which addition and scalar multiplication are continuous.*

We will abreviate topological vector space as TVS. By $L(X, Y)$ we will denote all continous linear mappings between two TVS X, Y.

Theorem 11.5.2.7. *If X is TVS, then X is Hausdorff iff $\{x\}$ is closed for each $x \in X$.*

Proof. [See **?** , 1.12] □

Definition 11.5.2.8. *Let X be a TVS. We will say that X is a* ***locally convex space*** *iff there exists a local base of convex and open sets.*

Definition 11.5.2.9. *Let X be a TVS. We say that set $B \subset X$ is* ***bounded*** *iff for every open neighbourhood U of 0 there exists a scalar $\lambda \geq 0$ such that $B \subset \lambda U$.*

Theorem 11.5.2.10. *Let X be TVS. Every convex open neighbourhood of 0 contains a convex and balanced open neighbourhood of 0.*

Proof. [See **?** , 1.14] □

Corollary 11.5.2.11. *Let X be a locally convex space. Then X has a local base of convex and balanced open sets.*

Theorem 11.5.2.12. *If X is a Hausdorff TVS with countable base, then there exists a metric d in X such that*

1. *d is compatible with X topology,*

2. *the open balls centered at 0 are ballanced,*

3. *d is invariant: i.e. $d(x, y) = d(x + z, y + z)$ for each $x, y, z \in X$.*

   Moreover, if X is locally convex, then d can be chosen so as to satisfy (1), (2) and (3) and also

4. *Every open ball is convex.*

Proof. [see **?** , 1.24] □

Definition 11.5.2.13. *X is called an F-space iff X is TVS, where topology is generated by a complete and invariant metric.*

Theorem 11.5.2.14. ***(The closed graph theorem)*** *Let X, Y be $F-$ spaces. If $A : X \to Y$ is a linear mapping such that $\{(x, Ax) : x \in X \times Y\}$ is closed in $X \times Y$, then A is continuous.*

Proof. [See **?** , 2.15 The closed graph theorem] □

Definition 11.5.2.15. *X is called a Fréchet space iff X is locally convex TVS, where topology is generated by a complete and invariant metric.*

Definition 11.5.2.16. *A seminorm on a vector space X is a function $p : X \to \mathbb{R}$ such that*

1. *$p(x + y) \leq p(x) + p(y)$,*
2. *$p(\lambda x) = |\lambda| p(x)$ for all $x, y \in X$ and all scalars λ.*

Definition 11.5.2.17. *A family of seminorms $\mathcal{P}$ on a vector space X is said to be separating if for each $x \neq 0$ there exists $p \in \mathcal{P}$ such that $p(x) \neq 0$.*

Theorem 11.5.2.18. *Let X be a vector space and $\mathcal{P}$ be a separating family of seminorms. Let*

$$V(p, n) = \{x \in X : p(x) < \frac{1}{n}\} \tag{11.5.2.2}$$

for any positive integer n and $p \in \mathcal{P}$. If $\mathcal{B}$ is a collection of all finite intersections of the sets $V(p, n)$, then $\mathcal{B}$ is convex balanced local base for a topology on X, which turns X into a locally convex Hausdorff TVS such that every $p \in \mathcal{P}$ is continuous.

Proof. [see **?** , 1.37] □

We will say that such a family of seminorms $\mathcal{P}$ generates topology on X.

Corollary 11.5.2.19. *Let X be a vector space with topology generated by a countable family of separating seminorms $\mathcal{P}$, then $x_n \to x$ iff $p(x_n - x) \to 0$ for each $p \in \mathcal{P}$.*

Theorem 11.5.2.20. *Let X be Hausdorff TVS. For any $A \subset X$ we define*

$$\mu_A(x) := \inf\{\lambda > 0 : x \in \lambda A\}. \tag{11.5.2.3}$$

1. *If p is a seminorm on X, then set $B = \{x : p(x) < 1\}$ is convex, balanced, absorbing and $p = \mu_B$.*

2. *If A is convex absorbing and balanced, then μ_A is a seminorm.*

3. *If A is bounded then for any $x \in X$, $\mu_A(x) = 0$ implies $x = 0$.*

4. *If X is a locally convex space and $\mathcal{B}$ is a convex and balanced local base then $\{\mu_V : V \in \mathcal{B}\}$ is a family of separating seminorms generating an original topology of X.*

Proof. For 1., 2. and 4. [See **?** , Seminorms and Local Convexity] We will prove 3. Take $x \in X$ such that $\mu_A(x) = 0$. Let U be an arbitrary open neighbourhood of 0. Since A is bounded there exists $\lambda \geq 0$ such that $\lambda A \subset U$. If $\lambda = 0$, then $A = \{0\}$ and thesis is shown. Assume then that $\lambda > 0$. Since $\mu_A(x) = 0$, $x \in \lambda A$. Thus $x \in U$. But since U was arbitrary open neighbourhood of 0 and X is Hausdorff, then $x = 0$. □

Corollary 11.5.2.21. *If X is a locally convex Hausdorff TVS, then there exists a family of separating seminorms which generates an original topology on X.*

Proof. This folows from Theorem 11.5.2.10 and Theorem 11.5.2.20. □

Definition 11.5.2.22. *Let X be a locally convex space. A family of continuous seminorms $\mathcal{P}$ is called a* ***basis of continuous seminorms*** *iff for any continuous seminorm q on X, there exists a seminorm $p \in \mathcal{P}$ and $C > 0$ such that*

$$q(x) \leq Cp(x) \text{ for any } x \in X. \tag{11.5.2.4}$$

Theorem 11.5.2.23. *Let X be a locally convex space and $\mathcal{P}$ a family of seminorms which generates topology on X. Let*

$$p_B(x) := \max\{p(x) : p \in B\} \tag{11.5.2.5}$$

where B is a finite subset of $\mathcal{P}$. The family of all such seminorms p_B is a basis of continuous seminorms in X.

Proof. [See **?** , Locally convex spaces. Seminorms.] □

Theorem 11.5.2.24. *If X is a locally convex space and $\mathcal{P}$ is a basis of continuous seminorms, then the topology generated by $\mathcal{P}$ conincides with the original topology of X.*

Proof. It is enough to show that given two basis of continous seminorms $\mathcal{P}_1$ and $\mathcal{P}_2$ generates the same topologies on X. This is obvious considering that for any $p_1 \in \mathcal{P}_1$ there exists $p_2 \in \mathcal{P}_2$ and $C > 0$ such that

$$p_1 \leq Cp_2, \tag{11.5.2.6}$$

and on the other hand for any $q_2 \in \mathcal{P}_2$ there exists $q_1 \in \mathcal{P}_1$ and $K > 0$ such that

$$q_2 \leq Kq_1. \tag{11.5.2.7}$$

□

Definition 11.5.2.25. *Let X be a vector space. By X^* we denote a space of all linear functionals.*

Definition 11.5.2.26. *Let X be a TVS. By X' we denote a space of all* ***continuous*** *linear functionals.*

Note that this convention is oposit to [**?**].

Definition 11.5.2.27. *Let X be a TVS. A weak-* topology on X' is the weakest topology for which all mappings $X' \ni y \mapsto y(x)$ are continuous for any fixed $x \in X$.*

Theorem 11.5.2.28. *If X is a Hausdorff TVS, then X' with weak-* topology is a locally convex Hausdorff TVS.*

Proof. [See **?** , 3.14] □

Theorem 11.5.2.29. *If X and Y are locally convex spaces, then a linear operator $A : X \to Y$ is continous iff for any continuous seminorm q in Y, there exists a continous seminorm p in X such that*

$$q(Ax) \leq p(x) \text{ for any } x \in X. \tag{11.5.2.8}$$

Proof. [See **?** , Locally convex spaces. Seminorms.] □

Fact 11.5.2.30. *If X is a locally convex space, then for any $u \in X'$, $X \ni x \mapsto |u(x)|$ is a continuous seminorm.*

Corollary 11.5.2.31. *If X is a locally convex space and $\mathcal{P}$ is a basis of continuous seminorms, then for each $u \in X'$ there exists $C > 0$ and $p \in \mathcal{P}$ such that*

$$|u(x)| \leq Cp(x) \text{ for any } x \in X. \tag{11.5.2.9}$$

Theorem 11.5.2.32. *If X is a locally convex space and let $\|\cdot\|$ be a continuous norm and let $\mathcal{P}$ be a base of continous seminorms in X. Then there exists a base of continous seminorms which constists only of norms.*

Proof. Note that for any seminorm $p \in \mathcal{P}$, $p + \|\cdot\|$ is a norm. Take any continuous seminorm q in X. We have such $p \in \mathcal{P}$ and $C > 0$ that

$$q(x) \leq Cp(x) \leq C(p(x) + \|x\|). \tag{11.5.2.10}$$

Thus a family

$$\{p + \|\cdot\| : p \in \mathcal{P}\} \tag{11.5.2.11}$$

is a base of continuous seminorms on X. □

Definition 11.5.2.33. *Let X be a locally convex Hausdorff space and B be its bounded, convex and balanced subset. Let X_B be a linear subspace spanned by B. Let*

$$\mu_B(x) := \inf\{\lambda > 0 : x \in \lambda B\} \tag{11.5.2.12}$$

*We say that B is **infracomplete** iff X_B with norm μ_B is a Banach space.*

Theorem 11.5.2.34. *Let X be a locally convex Hausdorff space and B be its bounded, convex and balanced subset. Let X_B be a linear subspace spanned by B. Then μ_B is norm on X_B.*

Proof. We will first show that B is absorbing in X_B. Any $x \in X_B$ is of the form $x = \sum_{i=1}^n \lambda_i x_i$. Where $\lambda_i \neq 0$ and $x_i \in B$ for $i = 1, \ldots, n$. Let $z_i := \frac{|\lambda_i|}{\overline{\lambda_i}} x_i$. Since B is balanced, $z_i \in B$ for $i = 1, \ldots, n$. Thus

$$x = \sum_{i=1}^n \lambda_i \frac{\overline{\lambda_i}}{|\lambda_i|} z_i = \sum_{i=1}^n |\lambda_i| z_i. \tag{11.5.2.13}$$

Let $\lambda_0 := (\sum_{i=1}^n |\lambda_i|)^{-1}$. Since B is convex, by Theorem 11.5.2.5, we have $\lambda_0 x \in B$. Take any $\lambda \in \mathbb{C}$ such that $|\lambda| < \lambda_0$. Obvioulsy $|\frac{\lambda}{\lambda_0}| < 1$. Since B is balanced,

$$B \ni \frac{\lambda}{\lambda_0} \lambda_0 x = \lambda x. \tag{11.5.2.14}$$

We showed that B is absorbing in X_B. Now, by Theorem 11.5.2.20 μ_B is a norm on X_B. □

Definition 11.5.2.35. *Let X be a topological space and Y be a TVS. Let F be a set of functions from X to Y. We will say that the set F is* ***equicontinuous*** *at point $x_0 \in X$ iff for any open V neighbourhood of 0, we have an open U neighbourhood of x_0 such that*

$$f(x) - f(x_0) \in V \tag{11.5.2.15}$$

for any $f \in F$ and any $x \in U$.

We will say that F is equicontinuous iff F is equicontinuous at each $x \in X$.

Definition 11.5.2.36. *Let X, Y be two locally convex Hausdorff spaces and $A \in L(X, Y)$. We will say that A is* ***nuclear*** *iff there is an equicontinous sequence $\{f_k\}$ in X', a sequence $\{y_k\}$ contained in a convex balanced infracomplete bouned subset B of Y and a complex sequence $\{\lambda_k\}$ with $\sum_{k=1}^{\infty} |\lambda_k| < +\infty$ such that*

$$Ax = \sum_{k=1}^{\infty} \lambda_k f_k(x) y_k. \tag{11.5.2.16}$$

Theorem 11.5.2.37. *Let H_1 and H_2 be Hilbert spaces and $A \in L(H_1, H_2)$. A is nuclear iff there is a sequence of orthonormal vectors $\{e_k\}$ in H_1, a sequence of orthonormal vectors $\{y_k\}$ in H_2 and a complex sequence $\{\lambda_k\}$ with $\sum_{k=1}^{\infty} |\lambda_k| < +\infty$ such that*

$$Ax = \sum_{k=1}^{\infty} \lambda_k (x|e_k) y_k. \tag{11.5.2.17}$$

Proof. [See **?** , 48.7] □

11.5.3 Regular Distributions

Let $C_0^\infty(\Omega)$ denotes all functions $\phi \in C^\infty(\Omega)$ such that its support i.e. $\mathrm{supp}(\phi) = \mathrm{Clo}(\{x \in \Omega : \phi(x) \neq 0\})$ is compact. We will use traditional notation $\mathcal{D}(\Omega) := C_0^\infty(\Omega)$.

We will construct certain topology on $\mathcal{D}(\Omega)$.

Definition 11.5.3.1. *Let Ω be an open subset of $\mathbb{R}^n$ and let $K \subset \Omega$ be compact.*

$$\mathcal{D}_K(\Omega) := \{\phi \in \mathcal{D}(\Omega) : supp(\phi) \subset K\}. \tag{11.5.3.1}$$

Let's introduce norms

Definition 11.5.3.2.

$$\|\phi\|_N := \max\{|D^\alpha \phi(x)| : x \in \Omega, |\alpha| \leq N\} \tag{11.5.3.2}$$

for any $\phi \in \mathcal{D}(\Omega)$ *and* $N = 0, 1, \ldots$.

Recall that for multindex $\alpha = (\alpha_1, \ldots, \alpha_2)$, $|\alpha| = \alpha_1 + \cdots + \alpha_2$.

Definition 11.5.3.3. *Let* Ω *be an open subset of* $\mathbb{R}^n$ *and let* $K \subset \Omega$ *be compact. We define a Hausdorff TVS* $(\mathcal{D}_K(\Omega), \tau_K)$*, where* τ_K *is a topology generated by a family of norms* $\|.\|_N$ *for* $N = 0, 1, \ldots$ *like in Theorem 11.5.2.18.*

Corollary 11.5.3.4. *Let* Ω *be an open subset of* $\mathbb{R}^n$ *and let* $K \subset \Omega$ *be compact.* $(\mathcal{D}_K(\Omega), \tau_K)$ *is a Fréchet space.*

Proof. [See **?** , 1.46] □

Definition 11.5.3.5. *Let* Ω *be an open subset of* $\mathbb{R}^n$*. We define a topological space* $(\mathcal{D}(\Omega), \tau)$*, where* τ *is a topology generated by a local base of all convex and balanced sets* $W \subset \mathcal{D}(\Omega)$*, such that* $W \cap K \in \tau_K$ *for every compact* $K \subset \Omega$.

Theorem 11.5.3.6. *Let* Ω *be an open subset of* $\mathbb{R}^n$*.* $(\mathcal{D}(\Omega), \tau)$ *is a locally convex Hausdorff TVS.*

Proof. [See **?** , 6.4] □

Definition 11.5.3.7. *Let* Ω *be an open subset of* $\mathbb{R}^n$*. By* $\mathcal{D}'(\Omega)$ *we denote a space of all linear functionals on* $\mathcal{D}(\Omega)$ *which are continuous in* τ*. Elements of* $\mathcal{D}'(\Omega)$ *are called* ***distributions****.*

Theorem 11.5.3.8. *Let* Ω *be an open subset of* $\mathbb{R}^n$*. If* Λ *is a linear functional on* $\mathcal{D}(\Omega)$ *the following conditions are equivalent:*

1. $\Lambda \in \mathcal{D}'(\Omega)$
2. *For every compact* $K \subset \Omega$ *there exists a positive integer* N *and a positive constant* $C < +\infty$ *such that*

   $$|\Lambda\phi| \leq C\|\phi\|_N \tag{11.5.3.3}$$

   for every $\phi \in \mathcal{D}_K(\Omega)$.

Proof. [see **?** , 6.8] □

Definition 11.5.3.9. *Let Ω be an open subset of $\mathbb{R}^n$. Let $\Lambda \in \mathcal{D}'(\Omega)$. The order of Λ is a minimal positive integer N for which for every compact $K \subset \Omega$ there exists a positive constant $C < +\infty$ such that*

$$|\Lambda\phi| \leq C\|\phi\|_N \tag{11.5.3.4}$$

for every $\phi \in \mathcal{D}_K(\Omega)$. If such N doesn't exists, we say that Λ is of infinite order.

By $L_{\text{loc}}(\Omega)$ we denote a space of all functions f for which $\int_K |f| < +\infty$ for any compact $K \subset \Omega$.

Theorem 11.5.3.10. *Let Ω be an open subset of $\mathbb{R}^n$. If $f \in L_{loc}(\Omega)$ and*

$$\Lambda_f(\phi) := \int_\Omega f(x)\phi(x)dx \text{ for any } \phi \in \mathcal{D}(\Omega), \tag{11.5.3.5}$$

then $\Lambda \in \mathcal{D}'(\Omega)$.

Proof. [see **?** , 6.11] □

We usually identify distribution Λ_f with function f and we say that such distributions "are" functions [see **?** , 6.11].

Theorem 11.5.3.11. *Let Ω be an open subset of $\mathbb{R}^n$. If $K \subset \Omega$ and K is compact, then there exists $\phi \in C_0^\infty(\Omega)$ such that $\phi(x) \in [0,1]$ for all $x \in \Omega$ and $\phi(z) = 1$ for all $z \in K$.*

Proof. [see **?** , Chapter 3. Sobolev Spaces. 3.1. Introduction] □

Definition 11.5.3.12. *Let Ω be an open subset of $\mathbb{R}^n$. We say that $\Lambda \in D'(\Omega)$ vanishes on an open $\omega \subset \Omega$ iff we have $\Lambda\phi = 0$ for every $\phi \in \mathcal{D}(\Omega)$ such that supp$(\phi) \subset \omega$.*

Definition 11.5.3.13. *Let Ω be an open subset of $\mathbb{R}^n$. Let $\Lambda \in D'(\Omega)$. We define*

$$supp(\Lambda) := \Omega \setminus \bigcup\{\omega \subset \Omega : \omega \text{ is open and } \Lambda \text{ vanishes on } \omega\}. \tag{11.5.3.6}$$

Definition 11.5.3.14. *(Dirac delta) Let $x \in \mathbb{R}^n$. We define a functional $\delta_x : \mathcal{D}(\mathbb{R}^n) \to \mathbb{C}$ as*

$$\delta_x(\phi) := \phi(x). \tag{11.5.3.7}$$

δ_0 is called Dirac delta.

Theorem 11.5.3.15. *Let $x \in \mathbb{R}^n$, then $\delta_x \in \mathcal{D}'(\Omega)$ and δ_x is of order 0.*

Proof. Take any compact $K \subset \mathbb{R}^n$. Take any $\phi \in D_K(\mathbb{R}^n)$. We have

$$|\delta_x(\phi)| = |\phi(x)| \leq \|\phi\|_0. \tag{11.5.3.8}$$

Thus by Theorem 11.5.3.8 $\delta_x \in \mathcal{D}'(\Omega)$ and by Definition 11.5.3.9 δ_x is of order 0. □

Theorem 11.5.3.16. *Let $x \in \mathbb{R}^n$, then $supp(\delta_x) = \{x\}$.*

Proof. It is enough to show that for any open $\omega \subset \mathbb{R}^n$, we have δ_x vanishes on ω iff $x \notin \omega$. Assume that δ_x vanishes on ω. If $x \in \omega$ we can choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such as $\phi(x) \neq 0$ and $supp(\phi) \subset \omega$, but as δ_x vanishes on ω, we must have $0 = \delta_x(\phi) = \phi(x)$. Contradiction, thus $x \neq \omega$. Assume on the other hand that $x \notin \omega$. Take any $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that $supp(\phi) \subset \omega$, then obviously $0 = \phi(x) = \delta_x(\phi)$. Thus δ_x vanishes on ω. □

11.5.4 Tempered Distributions

Definition 11.5.4.1.

$$\|\phi\|_N^S := \sup_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^N |D^\alpha \phi(x)| \tag{11.5.4.1}$$

for any $\phi \in S_n$.

Note that condition 11.4.0.1 in the Definition 11.4.0.1 of S_n is equivalent with the requirement that $\|\phi\|_N^S$ is finite for all $N = 0, 1 \ldots$.

Definition 11.5.4.2. *We equip S_n in topology generated by a family of norms $\|\phi\|_N^S$ for $N = 0, 1 \ldots$.*

Theorem 11.5.4.3. *If $p \geq 1$, then the identity mapping $i : S_n \to L^p(\mathbb{R}^n)$ is continuous.*

Proof. [See **?** , Problem 1.30] □

Theorem 11.5.4.4. *The following statements are true:*

1. *S_n is a Fréchet space.*
2. *The Fourier transform $\mathcal{F} : S_n \to S_n$ is a continuous linear transformation.*

Proof. [see **?** , 7.4] □

Theorem 11.5.4.5. *The following statemts are true:*

1. *$\mathcal{D}(\mathbb{R}^n)$ is dense in S_n.*

2. *Identity mapping from $\mathcal{D}(\mathbb{R}^n)$ to S_n is continuous.*

Proof. [see **?** , 7.10] □

Let S_n' be a space of all continous linear functionals on S_n.

Theorem 11.5.4.6. *For any $L \in S_n'$ there exists $u_L \in \mathcal{D}'(\mathbb{R}^n)$ such that*

$$u_L = L \circ i \tag{11.5.4.2}$$

where $i : \mathcal{D}(\mathbb{R}^n) \to S_n$ is an identity mapping.

Proof. [see **?** , 7.11] □

Definition 11.5.4.7. *Any u_L from Theorem 11.5.4.6 is called tempered distribution.*

Note that L is a unique extension of u_L to S_n. We will sometimes call tempered distributions elements of S_n'.

Theorem 11.5.4.8. *There exists an unique continous mapping $\Lambda : S_n \to S_n'$ (S_n' with weak-* topology) such that*

$$\Lambda_\phi(\psi) = \int_{R^n} \phi(x)\psi(x)dx \tag{11.5.4.3}$$

for any $\phi \in S_n$ and $\psi \in S_n$.

Proof. Let's assume equation (11.5.4.3) as definition of Λ. Note that $|\Lambda_\phi(\psi)| \le \|\phi\|_{L^2}\|\psi\|_{L^2}$ for any $\phi, \psi \in S_n$. By Theorem 11.5.4.3 S_n is continously embedded in $L^2(\mathbb{R}^n)$. Thus $\Lambda_\phi \in S_n'$ for any $\phi \in S_n$. Now take any $\phi_n \to \phi$ in S_n. We have $\phi_n \to \phi$ in L^2 and $|\Lambda_{\phi_n}(\psi) - \Lambda_\phi(\psi)| \le \|\phi_n - \phi\|_{L^2}\|\psi\|_{L^2}$. Thus $\Lambda_{\phi_n} \to \Lambda_\phi$ in S_n' with weak-* topology. □

Theorem 11.5.4.9. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $supp(u)$ is compact, then u is a tempered distribution.*

Proof. [see **?** , 7.12] □

Corollary 11.5.4.10. *Let $x \in \mathbb{R}^n$. δ_x is a tempered distribution and as extended to S_n it's still*

$$\delta_x(\phi) = \phi(x) \tag{11.5.4.4}$$

for every $\phi \in S_n$.

Example 11.5.4.11. *Let $\alpha \in R^n$ and $u_\alpha \in \mathcal{D}'(\mathbb{R}^n)$ be a distribution identified with the function $\mathbb{R}^n \ni x \mapsto e^{-i\alpha\cdot x}$. Then u_α is a tempered distribution. In this sense $e^{-i\alpha\cdot x} \in S'_n$.*

Proof. Let's define a linear functional $\tilde{u}_\alpha$ on S_n.

$$\tilde{u}_\alpha(\phi) := (2\pi)^{\frac{n}{2}}\mathcal{F}(\phi)(\alpha) = \int_{\mathbb{R}^n} e^{-i\alpha\cdot x}\phi(x)dx. \tag{11.5.4.5}$$

for all $\phi \in S_n$. $\tilde{u}_\alpha$ is well defined because Fourier tramsform exists for all ϕ in S_n. Because $\mathcal{F} : S_n \to S_n$ is continuous, for any $\phi_m \to \phi$, $\mathcal{F}(\phi_n) \to \mathcal{F}(\phi)$ uniformly. Thus $\tilde{u}_\alpha(\phi_n) \to \tilde{u}_\alpha(\phi)$. Hence $\tilde{u}_\alpha \in S'_n$. By definition $\tilde{u}_\alpha$ is an extension of u_α, thus u_α is a tempered distribution. □

Now we will extend Fourier transform on S'_n.

Definition 11.5.4.12. *Let $u \in S'_n$.*

$$\mathcal{F}(u)(\phi) := u(\mathcal{F}(\phi)) \tag{11.5.4.6}$$

for all $\phi \in S_n$.

Fact 11.5.4.13. *If $u, v \in S_n$, then*

$$\int \mathcal{F}(u)v = \int u\mathcal{F}(v). \tag{11.5.4.7}$$

Proof. Follows from Theorem 11.4.0.10 and Theorem 11.4.0.9. Indeed, $\int fg = \int f\bar{\bar{g}} = \int \mathcal{F}(f)\overline{\mathcal{F}(\bar{g})} = \int \mathcal{F}(f)\mathcal{F}^{-1}(g)$. Now substitute $f = v$ and $g = \mathcal{F}(u)$. □

From the fact above, Definition 11.5.4.12 is consistent with Fourier transform of function in case $u \in S'_n$ is a function, because for any $\phi \in S_n$

$$u(\mathcal{F}(\phi)) = \int u\mathcal{F}(\phi) = \int \mathcal{F}(u)\phi. \tag{11.5.4.8}$$

Theorem 11.5.4.14. *If $\mathcal{F}$ is a Fourier transform defined in Definition 11.5.4.12, $\mathcal{F} : S'_n \to S'_n$ is a linear mapping, continous in weak-* topology.*

Proof. [see **?** , 7.15] □

11.5.5 Schwartz Kernel Theorems

Definition 11.5.5.1. *Let X, Y be any vector spaces of functions valued in $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Let $U : X \times Y \to \mathbb{K}$. We will define operation:*

$$(U \bullet \phi)(\psi) := U(\phi \otimes \psi) \tag{11.5.5.1}$$

for any $\phi \in X$ and any $\psi \in Y$.

The following two theorems are corollaries from [**?** , 51. Examples of Nuclear Spaces. The Kernels Theorem]

Theorem 11.5.5.2. *Let Ω_1 be an open subset of $\mathbb{R}^{k_1}$ and Ω_2 be an open subset of $\mathbb{R}^{k_2}$. Let $L : \mathcal{D}(\Omega_1) \to \mathcal{D}'(\Omega_2)$ be linear. The following two statements are equivalent.*

1. *L is continuous with $\mathcal{D}'(\Omega_2)$ equiped with weak-* topology.*

2. *There exists unique $U \in D'(\Omega_1 \times \Omega_2)$ such that*

$$L\phi = U \bullet \phi. \tag{11.5.5.2}$$

Theorem 11.5.5.3. *Let $L : S_{k_1} \to S'_{k_2}$ be linear. The following two statements are equivalent.*

1. *L is continuous with S_{k_2} equiped with weak-* topology.*

2. *There exists unique $U \in S'_{k_1+k_2}$ such that*

$$L\phi = U \bullet \phi. \tag{11.5.5.3}$$

11.5.6 Properties of δ_0

Example 11.5.6.1. *Let $f_m : \mathbb{R}^n \to \mathbb{C}$ be a sequence of functions such that*

$$f_m(\alpha) = (2\pi)^{-n} \int_{K_m} e^{-ix\cdot\alpha} dx, \tag{11.5.6.1}$$

where K_m is a sequence of compact sets such that $K_m \nearrow \mathbb{R}^n$. Then

$$\lim_{m\to\infty} f_m = \delta_0 \text{ in } \mathcal{D}'(\mathbb{R}^n) \text{ with weak-* topology.} \tag{11.5.6.2}$$

Proof. Let $\Lambda_m \in \mathcal{D}'(\mathbb{R}^n)$ be a distribution corresponding to f_n. Note that

$$\int_K \int_{K_m} |e^{-ix\cdot\alpha}| dx d\alpha = \int_K dx \int_{K_m} d\alpha < +\infty, \tag{11.5.6.3}$$

for any compact $K \subset \mathbb{R}^n$. Thus by Theorem 11.2.0.10 (Fubini theorem), $f_n \in L_{\text{loc}}(\mathbb{R}^n)$.

Take any $\phi \in \mathcal{D}(\mathbb{R}^n)$. Note that

$$\Lambda_m(\phi) = (2\pi)^{-n} \int_{\mathbb{R}^n} \left(\int_{K_m} e^{-ix\cdot\alpha} dx \right) \phi(\alpha) d\alpha. \tag{11.5.6.4}$$

Since

$$\int_{\mathbb{R}^n} \left(\int_{K_m} |e^{-ix\cdot\alpha}| dx \right) \phi(\alpha) d\alpha = \int_{K_m} dx \int_{\mathbb{R}^n} \phi(\alpha) d\alpha < +\infty, \tag{11.5.6.5}$$

then again by Theorem 11.2.0.10 (Fubini theorem)

$$\Lambda_m(\phi) = (2\pi)^{-\frac{n}{2}} \int_{K_m} \mathcal{F}(\phi)(x) dx = (2\pi)^{-\frac{n}{2}} \int_{K_m} e^{ix\cdot 0} \mathcal{F}(\phi)(x) dx. \tag{11.5.6.6}$$

Hence

$$\lim_{m\to\infty} \Lambda_m(\phi) = \mathcal{F}^{-1}(\mathcal{F}(\phi))(0) = \phi(0). \tag{11.5.6.7}$$

Thus by Definition 11.5.3.14, we have thesis. □

Lemma 11.5.6.2. *If $f_n \in C(\mathbb{R}^k)$ is a sequence of non-negative real functions (i.e. $f_n \geq 0$) for which $\lim_{n\to\infty} f_n = \delta_0$ in $\mathcal{D}'(\mathbb{R}^k)$ with weak-* topology, then for each compact K such that $0 \notin K$ we have*

$$\lim_{n\to\infty} \int_K f_n = 0, \tag{11.5.6.8}$$

Proof. Since $\mathbb{R}^k$ is T_3, we have an open Ω such that $K \subset \Omega$ and $0 \notin \Omega$. By Theorem 11.5.3.11, we have $\phi \in C_0^\infty(\mathbb{R}^k)$ such that $0 \leq \phi \leq 1$, $\operatorname{supp}(\phi) \subset \Omega$ and $\phi(x) = 1$ for all $x \in K$. Since $f_n \geq 0$ and $\phi \geq 0$ we have

$$0 \leq \int_K f_n = \int_K f_n \phi \leq \int_{R^k} f_n \phi \to \phi(0) = 0. \tag{11.5.6.9}$$

□

Lemma 11.5.6.3. *If $f_n \in C(\mathbb{R}^k)$ is a sequence of non-negative real functions (i.e. $f_n \geq 0$) for which $\lim_{n\to\infty} f_n = \delta_0$ in $\mathcal{D}'(\mathbb{R}^k)$ with weak-* topology, then for each open $\Omega \subset \mathbb{R}^k$ such that $0 \in \Omega$ we have*

$$\lim_{n\to\infty} \int_\Omega f_n = 1, \tag{11.5.6.10}$$

Proof. Choose an open and bounded U such that $\mathrm{Clo}(\Omega) \subset U$. By Theorem 11.5.3.11 we have we have $\phi \in C_0^\infty(\mathbb{R}^k)$ such that $0 \leq \phi \leq 1$, $\mathrm{supp}(\phi) \subset U$ and $\phi(x) = 1$ for all $x \in \mathrm{Clo}(\Omega)$. Let $A_n = \int_{\mathbb{R}^k} f_n \phi$. Since $\lim_{n\to\infty} f_n = \delta_0$ in $\mathcal{D}'(\mathbb{R}^k)$ with weak-* topology, $A_n \to \phi(0) = 1$. Note that

$$A_n = \int_\Omega f_n + \int_{\mathrm{Clo}(U)\setminus\Omega} f_n \phi. \tag{11.5.6.11}$$

Thus

$$A_n - \int_\Omega f_n = \int_{\mathrm{Clo}(U)\setminus\Omega} f_n \phi \leq \int_{\mathrm{Clo}(U)\setminus\Omega} f_n \to 0. \tag{11.5.6.12}$$

We have the convergence above by Lemma 11.5.6.2. Thus

$$\lim_{n\to\infty} \int_\Omega f_n = 1. \tag{11.5.6.13}$$

□

11.6 Tensor Product of Hilbert Spaces (first aproach)

We will briefly summarize a construction of tensor product of Hilbert spaces from [?]. In this section, we are not very precise if Hilbert spaces are separable or not. This should be fixed in subsequent revisions and editions, but for now assume that all Hilbert spaces in this section are sparable.

Theorem 11.6.0.1. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces. Let $l : \Pi_{i=1}^n H_i \to \mathbb{C}$ be a bounded multi-lineral functional. then the sum (finite or infinite)*

$$\sum_{(y_1,\dots,y_n)\in Y_1\times\dots\times Y_n} |l(y_1,\dots,y_n)|^2 \tag{11.6.0.1}$$

doesn't depend on choice of orthonormal bases $Y_1, \dots, Y_n$, where Y_i is an orthonormal base of H_i.

Proof. [see **?** , 2.6 Constructions with Hilbert Spaces] □

In the context of the above theorem, we can correctly define

$$\|l\|_2 := \sum_{(y_1,\dots,y_n)\in Y_1\times\dots\times Y_n} |l(y_1,\dots,y_n)|^2. \tag{11.6.0.2}$$

Definition 11.6.0.2. *A bounded multi-lineral functional $l : \Pi_{i=1}^n H_i \to \mathbb{C}$ is a Hilbert-Schmidt functional if the sum in (11.6.0.1) is finite. $HS(\Pi_{i=1}^n H_i)$ is a space of all Hilbert-Schmidt functionals.*

Theorem 11.6.0.3. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces. $HS(\Pi_{i=1}^n H_i)$ is a Hilbert space with an inner product*

$$\langle l_1, l_2 \rangle := \sum_{(y_1,\ldots,y_n)\in Y_1\times\cdots\times Y_n} l_1(y_1,\ldots,y_n)\overline{l_2(y_1,\ldots,y_n)}, \tag{11.6.0.3}$$

where Y_i is an orthonormal base of H_i. The sum in (11.6.0.3) is absolutely convergent and doesn't depend on choice of orthonormal bases $Y_1,\ldots,Y_n$.

Proof. [see **?** , 2.6 Constructions with Hilbert Spaces] □

Definition 11.6.0.4. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces. Let H be a Hilbert space. A bounded multi-linear mapping $L : \Pi_{i=1}^n H_i \to H$ is a weak Hilbert-Schmidt mapping if the following conditions hold.*

1. *For each $u \in H$ a functional L_u, defined as*

$$L_u(y_1,\ldots,y_n) = \langle L(y_1,\ldots,y_n), u\rangle, \tag{11.6.0.4}$$

   is a Hilbert-Schmidt functional.

2. *There exists $c \geq 0$ such that $\|L_u\|_2 \leq c\|u\|$ for each $u \in H$.*

Moreover we define

$$\|L\|_2 := \inf\{c \geq 0 : \|L_u\|_2 \leq c\|u\| \text{ for each } u \in H\}. \tag{11.6.0.5}$$

Theorem 11.6.0.5. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces.*

1. *There exists a Hilbert space H and a weak Hilbert-Schmidt mapping $p : \Pi_{i=1}^n H_i \to H$, such that for any Hilbert space K and any weak Hilbert-Schmidt mapping $L : \Pi_{i=1}^n H_i \to K$, there exists $T \in \mathcal{B}(H,K)$ such that*

$$L = Tp. \tag{11.6.0.6}$$

2. *If there exists p' and H' with properties attributed in 1. to p and H, there exists an unitary mapping $U : H \to H'$ such that*

$$p' = Up. \tag{11.6.0.7}$$

3.

$$\langle p(x_1, \ldots, x_n), p(z_1, \ldots, z_n)\rangle = \Pi_{i=1}^n \langle x_i, z_i\rangle \tag{11.6.0.8}$$

for any $x_i, z_i \in H_i$ *for* $i = 1, \ldots, n$ *and*

$$\{p(y_1, \ldots, y_n) : (y_1, \ldots, y_n) \in Y_1 \times \cdots \times Y_n\} \tag{11.6.0.9}$$

is an orthonormal basis of H*, where* Y_i *is an orthonormal basis of* H_i *for* $i = 1, \ldots, n$*.*

Proof. [see **?** , 2.6 Constructions with Hilbert Spaces] □

Definition 11.6.0.6. *Being in the context of the above theorem, we define*

$$H_1 \hat{\otimes} \ldots \hat{\otimes} H_n := H \tag{11.6.0.10}$$

and

$$x_1 \otimes \cdots \otimes x_n := p(x_1, \ldots, x_n). \tag{11.6.0.11}$$

We will write also $\hat{\bigotimes}_{i=1}^n H_i$.

Theorem 11.6.0.7. *Let* $\{H_i\}_{i=1}^n$ *be Hilbert spaces.*

1. $\hat{\bigotimes}_{i=1}^n H_i$ *is a Hilbert space.*
2. *A mapping*

$$\Pi_{i=1}^n H_i \ni (x_1, \ldots x_n) \mapsto x_1 \otimes \cdots \otimes x_n \in \hat{\bigotimes}_{i=1}^n H_i \tag{11.6.0.12}$$

is multi-linear.

3. $$\langle x_1 \otimes \cdots \otimes x_n, z_1 \otimes \cdots \otimes z_n\rangle = \Pi_{i=1}^n \langle x_i, z_i\rangle \tag{11.6.0.13}$$

for for any $x_i, z_i \in H_i$ *for* $i = 1, \ldots, n$*.*

4. $$\{y_1 \otimes \cdots \otimes y_n : (y_1, \ldots, y_n) \in Y_1 \times \cdots \times Y_n\} \tag{11.6.0.14}$$

is an orthonormal basis of $\hat{\bigotimes}_{i=1}^n H_i$*, where* Y_i *is an orthonormal basis of* H_i *for* $i = 1, \ldots, n$*.*

There are many alternative ways of introducing tensor product of Hilbert Spaces (e.g [see **?** , 3.4 Tensor products of Hilbert spaces] – through algebraic tensor product and then completition), however they always lead to the Theorem 11.6.0.7. Let's reserve symbol $\otimes$ for algebraic tensor product, while we will use $\hat{\otimes}$ for completition. In this sense $\bigotimes_{i=1}^n H_i$ is a dense subspace of $\hat{\bigotimes}_{i=1}^n H_i$.

Definition 11.6.0.8. *Let* $\{H_i\}_{i=1}^n$ *be Hilbert spaces and let* A_i *with* $\mathcal{D}(A_i) \subset H_i$ *be linear operators.*

$$\mathcal{D}(A_1 \otimes \cdots \otimes A_n) := \mathcal{D}(A_1) \otimes \cdots \otimes \mathcal{D}(A_n) \tag{11.6.0.15}$$

and

$$(A_1 \otimes \cdots \otimes A_n)(\sum_{i=1}^{m} x_i^1 \otimes \cdots \otimes x_i^n) := \sum_{i=1}^{m} A_1 x_i^1 \otimes \cdots \otimes A_n x_i^n \tag{11.6.0.16}$$

for any integer m *and any* $(x_i^1, \ldots, x_i^n) \in \Pi_{k=1}^n \mathcal{D}(A_k)$ *for* $i = 1, \ldots, m$.

We can prove that the definition is correct by similiar argument as in [see **?** , 8.5 Analytic vectors and tensor products of self-adjoint operators].

Definition 11.6.0.9. *Let* $\{H_i\}_{i=1}^n$ *be Hilbert spaces and let* A_i *with* $\mathcal{D}(A_i) \subset H_i$ *be linear operators.*

$$A_1 \hat{\otimes} \ldots \hat{\otimes} A_n := \mathit{Clo}(A_1 \otimes \cdots \otimes A_n). \tag{11.6.0.17}$$

$$A_1 \hat{\otimes} I \hat{\otimes} \ldots \hat{\otimes} I + \cdots + I \hat{\otimes} \ldots I \hat{\otimes} A_n := \mathit{Clo}(A_1 \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \ldots I \otimes A_n) \tag{11.6.0.18}$$

Theorem 11.6.0.10. *Let* $\{H_i\}_{i=1}^n$ *be Hilbert spaces and* $\{K_i\}_{i=1}^n$ *be Hilbert spaces. Let* $A_i \in \mathcal{B}(H_i, K_i)$ *for* $i = 1, \ldots, n$*. Then there exists an unique* $A \in \mathcal{B}(\hat{\bigotimes}_{i=1}^n H_i, \hat{\bigotimes}_{i=1}^n K_i)$ *such that*

$$A(x_1 \otimes \cdots \otimes x_n) = A_1 x_1 \otimes \cdots \otimes A_n x_n \tag{11.6.0.19}$$

for all $(x_1 \otimes \cdots \otimes x_n) \in \Pi_{i=1}^n H_i$*. Moreover*

$$\|A\| = \|A_1\| \cdot \cdots \cdot \|A_n\|. \tag{11.6.0.20}$$

Proof. [see **?** , 2.6 Constructions with Hilbert Spaces] □

Theorem 11.6.0.11. *Let* $\{H_i\}_{i=1}^n$ *be Hilbert spaces and* $\{K_i\}_{i=1}^n$ *be Hilbert spaces. Let* $A_i, B_i \in \mathcal{B}(H_i, K_i)$ *for* $i = 1, \ldots, n$*. Then*

$$A_1 \hat{\otimes} \ldots \hat{\otimes} A_n \in \mathcal{B}(\hat{\bigotimes}_{i=1}^{n} H_i, \hat{\bigotimes}_{i=1}^{n} K_i), \tag{11.6.0.21}$$

$$\left\|A_1 \hat{\otimes} \ldots \hat{\otimes} A_n\right\| = \|A_1\| \cdot \cdots \cdot \|A_n\|. \tag{11.6.0.22}$$

$$(A_1 \hat{\otimes} \ldots \hat{\otimes} A_n)(B_1 \hat{\otimes} \ldots \hat{\otimes} B_n) = A_1 B_1 \hat{\otimes} \ldots \hat{\otimes} A_n B_n. \tag{11.6.0.23}$$

Theorem 11.6.0.12. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces and $\{K_i\}_{i=1}^n$ be Hilbert spaces. Let $U_i \in \mathcal{B}(H_i, K_i)$ for $i = 1, \dots, n$ be unitary mappings. Then*

$$U_1 \hat{\otimes} \dots \hat{\otimes} U_n : \hat{\bigotimes}_{i=1}^{n} H_i \to \hat{\bigotimes}_{i=1}^{n} K_i \tag{11.6.0.24}$$

is also a unitary mapping.

Proof. Unitary mapping map orthonormal basis into orthonormal basis. Thus, by Theorem 11.6.0.7 we can show thesis. □

Theorem 11.6.0.13. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces and A_i be self-adjoint with $\mathcal{D}(A_i) \subset H_i$ for $i = 1, \dots, n$. Then $A_1 \hat{\otimes} \dots \hat{\otimes} A_n$ is self-adjoint and*

$$A_1 \hat{\otimes} I \hat{\otimes} \dots \hat{\otimes} I + \dots + I \hat{\otimes} \dots I \hat{\otimes} A_n. \tag{11.6.0.25}$$

is self-adjoint.

Proof. [see **?** , 8.5 Analytic vectors and tensor products of self-adjoint operators]. □

The following theorem is an convinient instruction, what minimal domain we need egzamine if we want to establish equality between some self-adjoint operator and closed product of self-adjoint operators.

Lemma 11.6.0.14. *Let H be a hilbert space and B be a self adjoint-operator. If $A \subset B$ and Clo(A) is self-adjoint, then Clo$(A) = B$.*

Proof. Since B is self-adjoint, B is closed. So, $\mathrm{Clo}(A) \subset B$. But $\mathrm{Clo}(A)$ is self-adjoint, hence it's maximal symetric operator, thus $\mathrm{Clo}(A) = B$. □

Theorem 11.6.0.15. *Let $\{H_i\}_{i=1}^n$ be Hilbert spaces, and let A_i be self-adjoint with $\mathcal{D}(A_i) \subset H_i$ for $i = 1, \dots, n$ and let B be a self-adjoint operator on $\hat{\bigotimes}_{i=1}^n H_i$. If*

$$\mathcal{D}(A_1) \otimes \dots \otimes \mathcal{D}(A_n) \subset \mathcal{D}(B) \tag{11.6.0.26}$$

and

$$B(x_1 \otimes \dots \otimes x_n) = A_1 x_1 \otimes \dots \otimes A_n x_n \tag{11.6.0.27}$$

for all $(x_1, \dots, x_n) \in \Pi_{k=1}^n \mathcal{D}(A_k)$, then

$$B = A_1 \hat{\otimes} \dots \hat{\otimes} A_n. \tag{11.6.0.28}$$

Proof. Let $A = A_1 \otimes \dots \otimes A_n$. By linearity we get $A \subset B$. By Theorem 11.6.0.13, $\mathrm{Clo}(A)$ is self-adjoint. Now, by Lemma 11.6.0.14, we get $\mathrm{Clo}(A) = B$. □

Theorem 11.6.0.16. *If* (X_1, μ_1) *and* (X_2, μ_2) *are measurable spaces, then*

$$L^2(X_1, \mu_1)\hat{\otimes}L^2(X_1, \mu_2) = L^2(X_1 \times X_2, \mu_1 \times \mu_2), \tag{11.6.0.29}$$

where

$$(\psi_1 \otimes \psi_2)(x_1, x_2) = \psi_1(x_1) \cdot \psi_2(x_2) \text{ for } \psi_i \in L^2(X_i, \mu_i) \text{ for } i = 1, 2. \tag{11.6.0.30}$$

Proof. [see **?** , 2.6 Constructions with Hilbert Spaces] □

Example 11.6.0.17. *Let* $\mathcal{F}_k$ *be a k-dimensional Fourier transform extended to* $L^2(\mathbb{R}^k)$*.*

$$\mathcal{F}_n\hat{\otimes}\mathcal{F}_m = \mathcal{F}_{n+m}. \tag{11.6.0.31}$$

Proof. Take any $\psi_1 \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $\psi_2 \in L^1(\mathbb{R}^m) \cap L^2(\mathbb{R}^m)$.

$$(\mathcal{F}_n\hat{\otimes}\mathcal{F}_m)(\psi_1 \otimes \psi_2)(x_1, x_2) = \mathcal{F}_n\psi_1(x_1) \otimes \mathcal{F}_m\psi_2(x_2) =$$
$$\left((2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \psi_1(s_1)\exp(-ix_1 \cdot s_1)ds_1\right)\left((2\pi)^{-\frac{m}{2}} \int_{\mathbb{R}^m} \psi_2(s_2)\exp(-ix_2 \cdot s_2)ds_2\right) =$$
$$= (2\pi)^{-\frac{n+m}{2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \psi_1(s_1)\psi_1(s_2)\exp(-i(x_1 \cdot s_1 + x_2 \cdot s_2))ds_1ds_2 =$$
$$\mathcal{F}_{n+m}(\psi_1 \otimes \psi_2)(x_1, x_2). \tag{11.6.0.32}$$

By continuity we have shown $\mathcal{F}_n\hat{\otimes}\mathcal{F}_m(\psi_1 \otimes \psi_2) = \mathcal{F}_{n+m}(\psi_1 \otimes \psi_2)$ for any $\psi_1 \in L^2(\mathbb{R}^n)$ and $\psi_2 \in L^2(\mathbb{R}^m)$. Now by Theorem 11.6.0.10, we have (11.6.0.31). □

In the theorem below we will use L^2-representations of self-adjoint operators as defined by Definition 11.3.2.3.

Theorem 11.6.0.18. *Let* H *be a Hilbert space and* A, B *be self-adjoint densly defined operators on* H *and let* (X_A, ν_A, U_A, g_A) *and* (X_B, ν_B, U_B, g_B) *be respectively their* L^2*-representations, then*

1. $$(X_A \times X_B, \nu_A \times \nu_B, U_A\hat{\otimes}U_B, g_A + g_B)$$
   is L^2*-representation of* $A\hat{\otimes}I + I\hat{\otimes}B$*.*

2. $$(X_A \times X_B, \nu_A \times \nu_B, U_A\hat{\otimes}U_B, g_A \cdot g_B)$$
   is L^2*-representation of* $A\hat{\otimes}B$*.*

Proof. By Theorem 11.3.2.1, it's easy to prove that

$$C\psi := (U_A \hat{\otimes} U_B)^{-1}((g_A + g_B) \cdot (U_A \hat{\otimes} U_B)(\psi)) \tag{11.6.0.33}$$

is densly defined and self-adjoint. Take any $\psi_1 \in \mathcal{D}(A)$ and $\psi_2 \in \mathcal{D}(B)$. Note that

$$\begin{gathered} C(\psi_1 \otimes \psi_2) = (U_A \hat{\otimes} U_B)^{-1}(g_A \cdot U_A(\psi_1) \cdot U_B(\psi_2) + U_A(\psi_1) \cdot g_B \cdot U_B(\psi_2)) = \\ (U_A \hat{\otimes} U_B)^{-1}((g_A \cdot U_A(\psi_1)) \otimes U_B(\psi_2) + U_A(\psi_1) \otimes (g_B \cdot U_B(\psi_2))) = \\ (U_A^{-1}(g_A \cdot U_A(\psi_1)) \otimes U_B^{-1} U_B(\psi_2) + U_A^{-1} U_A(\psi_1) \otimes U_B^{-1}(g_B \cdot U_B(\psi_2)) = \\ A\psi_1 \otimes \psi_2 + \psi_1 \otimes B\psi_2. \end{gathered} \tag{11.6.0.34}$$

By linearity we showed $A \otimes I + I \otimes B \subset C$. By Lemma 11.6.0.14, we have $\operatorname{Clo}(A \otimes I + I \otimes B) = C$. Point 2 can be showed analogously. $\square$

Chapter 12

Mathematical Methods

12.1 Vector Analysis in $\mathbb{R}^n$

12.1.1 $\mathbb{R}^3$ Case

We will try to give quite precise but still a heuristic formulation of the Stokes's Theorem in $\mathbb{R}^3$ case, which is a ceratin compromise between being realitive easy to comprehend and rigorous enough to be comfortably used in most physical applications. The serious mathematical treatment of Stokes theorem for differentiable manifolds can be found in e.g. [**?**] or [**?**]. However, in my opinion, it is propedeutically recommended to understand first the $\mathbb{R}^3$ case in a heuristic way, as presented here, to be ready for more advanced and rigorous treatment of this topic in differentiable manifolds theory.

Definition 12.1.1.1. *We say that the picewise smooth surface $(S, \vec{n})$ is oriented when $\vec{n} : S \to \mathbb{R}^3$ is a picewise C^1 continous field of normed vectors perpendicular to S in each point.*

Definition 12.1.1.2. *We say that the piecewise smooth curve $(\Gamma, \vec{l})$ is oriented when $\vec{l} : \Gamma \to \mathbb{R}^3$ is a picewise C^1 field of normed vectors tanget to Γ in each point.*

Definition 12.1.1.3. *Let $(S, \vec{n})$ be an oriented surface and let $\Gamma \subset \partial S$ be a picewise smooth closed oriented curve $(\Gamma, \vec{l})$. We say that these orientations are consistent if $\vec{n} \times \vec{l}$ is directed towards S for any point of Γ.*

Theorem 12.1.1.4. ***(Stokes's Theorem)*** *Let $(S, \vec{n})$ be a bounded, picewise smooth, oriented surface in $\mathbb{R}^3$. Let $(\partial S, \vec{l})$ be a picewise C^1 edge-boundary of S consisting of finitelly many closed curves oriented consistently with S. Let*

Ω be on open set such that $S \subset \Omega$. Let ds symbolise an infinitesimal element of ∂S and dS an infinitesimal element of surface S. If $\vec{F} \in C^1(\Omega, \mathbb{R}^3)$, then

$$\oint_{\partial S} \vec{F} \cdot (ds)\vec{l} = \int_S (\nabla \times \vec{F}) \cdot (dS)\vec{n}. \tag{12.1.1.1}$$

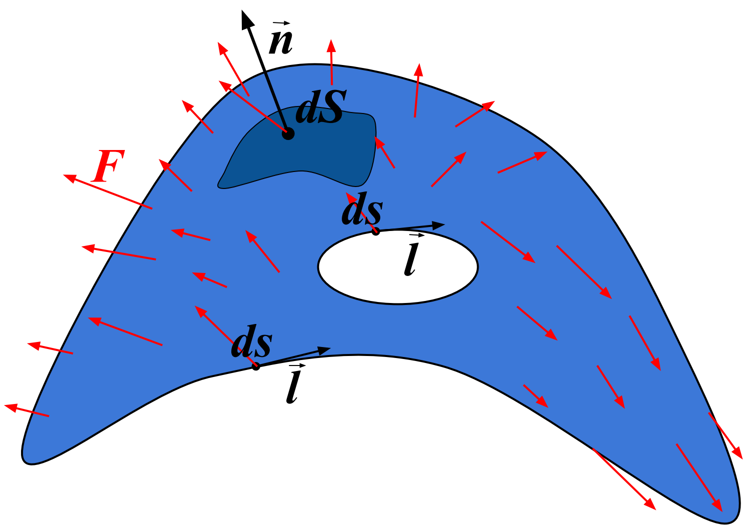

Figure 12.1: Ilustration for Stokes Theorem

Proof. [see **?** , 7.3 Stokes's and Gauss's Theorems] □

Theorem 12.1.1.5. *(**Gauss's Theorem**) Let Ω be an open an connected subset of $\mathbb{R}^3$. Let $(\partial\Omega, \vec{n})$ be a picewise smooth directed closed surface, such that $\vec{n}$ points outside Ω. Let dV symbolise an infinitesimal element of volume Ω and let dS symbolise an infinitesimal element of surface $\partial\Omega$. If $\vec{F} \in C^1(Clo(\Omega), \mathbb{R}^3)$, then*

$$\oint_{\partial\Omega} \vec{F} \cdot (dS)\vec{n} = \int_\Omega \nabla \cdot \vec{F} dV. \tag{12.1.1.2}$$

Proof. [see **?** , 7.3 Stokes's and Gauss's Theorems] □

12.1.2 Introduction to Differential Forms on Manifolds Embedded in Real Coordinate Space

In this subsection we will consider only parametrised smooth manifolds with one map. This is an obvious simplification, but we will treat this only as a traning before doing serious differentiable manifolds theory. This is a propedeutical way taken after [**?**] with only slight modifications. We will first introduce a concept of differentiable k-form.

Definition 12.1.2.1. *A basis k-form in $\mathbb{R}^n$*

$$dx_{i_1}\wedge\ldots\wedge dx_{i_k}, \tag{12.1.2.1}$$

where $i_1,\ldots,i_k = 1,\ldots,n$, is a multi-linear function from $\mathbb{R}^n$ to $\mathbb{R}$ such that

$$dx_{i_1}\wedge\ldots\wedge dx_{i_k}(\vec{a_1},\ldots,\vec{a_k}) := \det\begin{bmatrix} dx_{i_1}(\vec{a_1}) & \ldots & dx_{i_1}(\vec{a_k}) \\ \vdots & \ddots & \vdots \\ dx_{i_k}(\vec{a_1}) & \ldots & dx_{i_k}(\vec{a_k}) \end{bmatrix} \tag{12.1.2.2}$$

where for $\vec{a} = [a_1,\ldots,a_n]$, $dx_j(\vec{a}) = a_j$.

Definition 12.1.2.2. *Let U be an open subset of $\mathbb{R}^n$. A differentiable k-form on U*

$$\omega(x) = \sum_{i_1,\ldots,i_k} F_{i_1,\ldots,i_k}(x)dx_{i_1}\wedge\ldots\wedge dx_{i_k} \tag{12.1.2.3}$$

is a mapping from U to the space of multi-linear functions from $\mathbb{R}^n$ to $\mathbb{R}$, where $F_{i_1,\ldots,i_k} \in C^1(U)$ and

$$\omega(x)(\vec{a_1},\ldots,\vec{a_k}) = \sum_{i_1,\ldots,i_k} F_{i_1,\ldots,i_k}(x)dx_{i_1}\wedge\ldots\wedge dx_{i_k}(\vec{a_1},\ldots,\vec{a_k}). \tag{12.1.2.4}$$

A differentiable 0-form is each function from $C^1(U,\mathbb{R})$. For formal reasons we assume that basis 0-form is constant function $= 1$.

We treat $\wedge$ as a natural operator on basis forms with assumed associativity. It follows from the definition of determinant that if any dx_i appears twice the basis form is equal to 0 and it changes sign when we swap two of its elements.

Definition 12.1.2.3. *Let $\omega_1 = \sum_i f_i^1\Omega_i^1$ be a differentiable k-form and $\omega_2 = \sum_j f_j^2\Omega_j^2$ be a differentiable l-form where Ω_i^1 are basis k-forms and Ω_j^2 are basis l-forms.*

$$\omega_1\wedge\omega_2 := \sum_i\sum_j f_i^1 f_j^2(\Omega_i^1\wedge\Omega_j^2). \tag{12.1.2.5}$$

Theorem 12.1.2.4. *Let $\omega, \omega_1, \omega_2$ be differentiable k-forms, ν be differentiable l-form and τ be a p-form. Let f be 0-form Then*

1. $(\omega_1 + \omega_2)\wedge\nu = \omega_1\wedge\nu + \omega_2\wedge\nu$.
2. $\omega\wedge\nu = (-1)^{kl}\nu\wedge\omega$.
3. $(\omega\wedge\nu)\wedge p = \omega\wedge(\nu\wedge p)$.
4. $(f\omega)\wedge\nu = f(\omega\wedge\nu) = \omega\wedge(f\nu)$.

Definition 12.1.2.5. *Let $\omega = \sum_j f_j\Omega_j$ be a differentiable k-form on $U \subset \mathbb{R}^n$ where Ω_i are basis k-forms.*

$$d\omega := \sum_j \sum_{i=1}^{n} \frac{\partial f_j}{\partial x_i} dx_i \wedge \Omega_j. \tag{12.1.2.6}$$

Now we will introduce a simplified concept of manifold.

Definition 12.1.2.6. *Let $D \subset \mathbb{R}^k$ be a region consists of some open set U and any parts of its clousure. X is parametrised k-manifold if $X : D \to \mathbb{R}^n$ and*

1. $X|_U$ *is 1-to-1 and C^1 class.*
2. $\frac{\partial X}{\partial u_1}, \ldots, \frac{\partial X}{\partial u_k}$ *are linearly independent.*

Definition 12.1.2.7. *Let X be a parametrised k-manifold.*

$$T_X^i := \frac{\partial X}{\partial u_i}. \tag{12.1.2.7}$$

We say that T_X^i is a tangent vector to i-th coordinate curve.

Definition 12.1.2.8. *Let X be parametrised k-manifold. k-form Ω is called an orientation of X, iff*

$$\Omega(T_X^1, \ldots, T_X^k) > 0. \tag{12.1.2.8}$$

at each point of X.

Proposition 12.1.2.9. *Let $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$, then*

$$\det\left(\vec{a}, \vec{b}, \vec{c}\right) = \vec{a} \cdot (\vec{b} \times \vec{c}). \tag{12.1.2.9}$$

Example 12.1.2.10. *Let X be a parametrised 2-manifold and*

$$\vec{n} = \frac{T_X^1 \times T_X^2}{\|T_X^1 \times T_X^2\|}. \tag{12.1.2.10}$$

Note that $(X, \vec{n})$ is an orientation of X in terms of Definition 12.1.1.1. Indeed $\vec{n}$ is a unit vector perpendicular to the tanget space. The corresponding orientation in terms of Definition 12.1.2.8 is

$$\Omega(\vec{a}^1, \vec{a}^2) = \det\bigl(\vec{n}, \vec{a}^1, \vec{a}^2\bigr). \tag{12.1.2.11}$$

Proof. Note that $\Omega(T_X^1, T_X^2) > 0$ because

$$\Omega(T_X^1, T_X^2) = \frac{1}{\|T_X^1 \times T_X^2\|}(T_X^1 \times T_X^2) \cdot (T_X^1 \times T_X^2) > 0. \tag{12.1.2.12}$$

□

Definition 12.1.2.11. *Let X be a parametrised k-manifold and Y be a parametrised $k-1$-manifold such that Y is contained in a boundary of X. Let Ω_X be an orientation of X. We say that orientation of Y is induced by orientation of X iff*

$$\Omega_Y(\vec{a}^1, \dots, \vec{a}^{k-1}) := \Omega(\vec{v}, \vec{a}^1, \dots, \vec{a}^{k-1}) \tag{12.1.2.13}$$

is an orientation of Y where $\vec{v}$ is a vector field on X such that $\vec{v}$ is tangent to X (i.e. is a linear combination of $T_X^1, \dots, T_X^k$), normal to Y (i.e. perpendicular to each $T_Y^1, \dots, T_Y^{k-1}$) and $-\vec{v}$ on Y points towards X.

Proposition 12.1.2.12. *Induced orientation from Definition 12.1.2.11 coincides with consistent orientation from Definition 12.1.1.3*

Proof. Let X be parametrised 2-manifold. Let

$$\vec{n} := \frac{T_X^1 \times T_X^2}{\|T_X^1 \times T_X^2\|}. \tag{12.1.2.14}$$

As it was previously noted, $(X, \vec{n})$ is an orientation of X in terms of Definition 12.1.1.1. We have shown in Example 12.1.2.10 that $\Omega_X(\vec{a}^1, \vec{a}^2) = \det(\vec{n}, \vec{a}^1, \vec{a}^2)$ is orientation of X in terms of Definition 12.1.2.8. Let Y be a parametrised 1-manifold (i.e. curve) such that Y is contained in a boundary of X. Let

$$\vec{l} := \frac{T_Y^1}{\|T_Y^1\|}. \tag{12.1.2.15}$$

Note that $\vec{l}$ is a unit vector tangent to Y at each point. Thus $(Y, \vec{l})$ is an orientation of Y in terms of Definition 12.1.1.2.

Let $\vec{v}$ be some vector filed on Y wich is tangent to X and perpendicular to Y and $-\vec{v}$ points towards X.

$$\Omega(\vec{a}) := \Omega_X(\vec{v}, \vec{a}) = \det(\vec{n}, \vec{v}, \vec{a}) = \vec{n} \cdot (\vec{v} \times \vec{a}) = \vec{v} \cdot (\vec{a} \times \vec{n}). \qquad (12.1.2.16)$$

Ω is obviously a good candidate for orientation of Y. Now all we need to do is to check the sign of $\Omega(\vec{l})$.

$$\Omega(\vec{l}) = \vec{v} \cdot (\vec{l} \times \vec{n}) = -\vec{v} \cdot (\vec{n} \times \vec{l}). \qquad (12.1.2.17)$$

Notice that because $\vec{v}$ is tangent to X and perpendicular to Y, we have $-\vec{v} \parallel \vec{n} \times \vec{l}$. Thus, $\Omega(\vec{l}) > 0$ if and only if $\vec{n} \times \vec{l}$ points in the same direction as $-\vec{v}$, i.e. towards X. And this means that Ω is an orientation of Y induced by X if and only if orientation $(Y, \vec{l})$ is consistent with orientation $(X, \vec{n})$ in terms of Definition 12.1.1.3. □

Definition 12.1.2.13. *Let $X : D \to \mathbb{R}^n$ be parametrised k-manifold. Let ω be a C^1 k-form defined on an open set U such that $X(D) \subset U$.*

$$\int_X \omega := \int_D \omega(T_X^1, \ldots, T_X^k) du, \qquad (12.1.2.18)$$

where du symbolise an infinitesimal volume element of D.

The central theorem in this section is Stokes's Theorem.

Theorem 12.1.2.14. ***(Stokes's Theorem)*** *Let $X : D \to \mathbb{R}^n$ be parametrised k-manifold. Let ∂X be $k-1$-manifold which image is equal to the boundary of X. Let ω be a C^1 k-form defined on an open set U such that $X(D) \subset U$. If orientation of ∂X is induced by X, then*

$$\int_X d\omega = \int_{\partial X} \omega. \qquad (12.1.2.19)$$

Lemma 12.1.2.15. *Let $X : D \to \mathbb{R}^3$ be a parametrised 2-manifold. Let du be an infinitesimal element of D. Let $dS = X(du)$ be a corresponding infinitesimal surface element of X. Then*

$$dS = \left\| T_X^1 \times T_X^2 \right\| du. \qquad (12.1.2.20)$$

Proof. Let $u = (u_1, u_2)$ and $\Delta u = (\Delta u_1, \Delta u_2)$ with $\Delta u_1, \Delta u_2$ being infinitesimals of the first order. Imagine an infinitesimal volume element du as

$[u_1, u_1 + \Delta u_1] \times [u_2, u_2 + \Delta u_2]$. Let $\theta = (\theta_1, \theta_2)$ Neglecting second-order infinitesimals we have

$$X(u+\theta\cdot\Delta u) = X(u) + \begin{bmatrix} \frac{\partial X}{\partial u_1} & \frac{\partial X}{\partial u_2} \end{bmatrix} \begin{bmatrix} \theta_1 \Delta u_1 \\ \theta_2 \Delta u_2 \end{bmatrix} = X(u) + \theta_1 \Delta u_1 T_X^1 + \theta_2 \Delta u_2 T_X^2. \tag{12.1.2.21}$$

Note that

$$du = \mu([u_1, u_1 + \Delta u_1] \times [u_2, u_2 + \Delta u_2]) = \mu\{\theta \cdot \Delta u : \theta \in [0,1]^2\}. \tag{12.1.2.22}$$

Thus

$$dS = X(du) = \mu\{\theta_1 \Delta u_1 T_X^1 + \theta_2 \Delta u_2 T_X^2 : \theta \in [0,1]^2\}. \tag{12.1.2.23}$$

But this equal to the area of the parallelogram spanned by $\Delta u_1 T_X^1$ and $\Delta u_2 T_X^2$, which is equal to $\|T_X^1 \times T_X^2\| \Delta u_1 \Delta u_2 = \|T_X^1 \times T_X^2\| du$. □

Example 12.1.2.16. *Let X be a parametrised 2-manifold. Let $\vec{n} = \dfrac{T_X^1 \times T_X^2}{\|T_X^1 \times T_X^2\|}$. Let $F = [F^1, F^2, F^3] \in C^1(U, \mathbb{R}^3)$, where $X \subset U$. Let*

$$\omega = F^1 dx_2 \wedge dx_3 + F^2 dx_3 \wedge dx_1 + F^3 dx_1 \wedge dx_2. \tag{12.1.2.24}$$

Then

$$\int_X \omega = \int_X \begin{bmatrix} F^1 & F^2 & F^3 \end{bmatrix} \cdot \vec{n} dS. \tag{12.1.2.25}$$

Proof.

$$\omega(T_X^1, T_X^2) = F^1 dx_2 \wedge dx_3 (T_X^1, T_X^2) + \cdots = F_1 \begin{vmatrix} dx_2(T_X^1) & dx_2(T_X^2) \\ dx_3(T_X^1) & dx_3(T_X^2) \end{vmatrix} + \cdots =$$

$$\begin{bmatrix} F^1 & F^2 & F^3 \end{bmatrix} \cdot (T_X^1 \times T_X^2) = \begin{bmatrix} F^1 & F^2 & F^3 \end{bmatrix} \cdot \vec{n} \left\| T_X^1 \times T_X^2 \right\| \tag{12.1.2.26}$$

Thus by Lemma 12.1.2.15 we have

$$\int_X \omega = \int_D \begin{bmatrix} F^1 & F^2 & F^3 \end{bmatrix} \cdot \vec{n} \left\| T_X^1 \times T_X^2 \right\| du = \int_X \begin{bmatrix} F^1 & F^2 & F^3 \end{bmatrix} \cdot \vec{n} dS. \tag{12.1.2.27}$$

□

Theorem 12.1.2.17. *(**Gauss's Theorem** n**-dimensional case**) Let Ω be an open an connected subset of $\mathbb{R}^n$. Let $\vec{n}$ be a fector field of unit vectors, normal to a picewise smooth closed surface of $\partial\Omega$, such that $\vec{n}$ points outside Ω. Let dV symbolise an infinitesimal element of n-dimentional volume of Ω and let dS symbolise an infinitesimal element of surface $\partial\Omega$. If $\vec{F} \in C^1(Clo(\Omega), \mathbb{R}^n)$, then*

$$\oint_{\partial\Omega} \vec{F} \cdot (dS)\vec{n} = \int_\Omega \nabla \cdot \vec{F} dV. \tag{12.1.2.28}$$

Proof. Follows from Theorem 12.1.2.14 (Stokes's Theorem). □

12.2 Group Theory

12.2.1 Generators

In this subsection we will use Einstein summation convention. Let's assume we have certain continous symetry $u_{(\varepsilon)}(x)$ of $\mathbb{R}^n$ with a generator g (i.e. $\left.\frac{du_\varepsilon}{d\varepsilon}\right|_{\varepsilon=0} = g$). This symetry on domain induce a transformation of functions.

$$f \mapsto f \circ u_{(\varepsilon)} \tag{12.2.1.1}$$

We will show that the generator of this transformation (at least in a pointwise sense) is

$$f \mapsto \nabla f \cdot g. \tag{12.2.1.2}$$

We would like to calculate $\left.\frac{d(f \circ u_{(\varepsilon)})}{d\varepsilon}\right|_{\varepsilon=0}$.

$$\left.\frac{d(f \circ u_{(\varepsilon)})}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{\partial f}{\partial x^i}\frac{du^i_{(\varepsilon)}}{d\varepsilon}\right|_{\varepsilon=0} = \nabla f \cdot g. \tag{12.2.1.3}$$

Although it is not shown here, it seems it could be changed into a rigorous theorem, that if we consider the space $L^2(\mathbb{R}^n)$, the convergence in deriviative above is in L^2 norm sense.

Example 12.2.1.1. *Consider one dimensional case. Let $u_{(\varepsilon)}(x) = x + \varepsilon$. In this case generator of $u_{(\varepsilon)}$ is just 1. Thus the generator of $f \mapsto f \circ u_{(\varepsilon)}$ is $\frac{d}{dx}$.*

Example 12.2.1.2. *Rotation on plane. Let*

$$u_{(\theta)}(x) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}. \tag{12.2.1.4}$$

The generator of $f \mapsto f \circ u_{(\theta)}$ is

$$x^1\frac{\partial}{\partial x^2} - x^2\frac{\partial}{\partial x^1}. \tag{12.2.1.5}$$

Proof.

$$g(x) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \begin{bmatrix} -x^2 \\ x^1 \end{bmatrix}, \tag{12.2.1.6}$$

thus

$$\nabla f(x) \cdot g(x) = \begin{bmatrix} \frac{\partial f}{\partial x^1} & \frac{\partial f}{\partial x^2} \end{bmatrix} \begin{bmatrix} -x^2 \\ x^1 \end{bmatrix} = x^1\frac{\partial f}{\partial x^2} - x^2\frac{\partial f}{\partial x^1}. \tag{12.2.1.7}$$

□

Proposition 12.2.1.3. *If $u_{(\theta)}$ is a rotation about an axis specified by unit vector n with an angle θ measured couterclockwise if seen from the tip of n, then*

$$\left.\frac{du_{(\theta)}(x)}{d\theta}\right|_{\theta=0} = n \times x. \tag{12.2.1.8}$$

Proof. Choose an orthonormal basis n, a, b (hint: they corespond to x, y, z respectively in stanard setup) such that

$$a \times n = -b \text{ and } b \times n = a. \tag{12.2.1.9}$$

Then

$$u_{(\varepsilon)}(x) = (u \cdot x)u + \\ ((a \cdot x)\cos\theta - (b \cdot x)\sin\theta)a + ((a \cdot x)\sin\theta + (b \cdot x)\cos\theta)b. \tag{12.2.1.10}$$

Thus

$$\left.\frac{du_{(\theta)}(x)}{d\theta}\right|_{\theta=0} = -(b \cdot x)a + (a \cdot x)b. \tag{12.2.1.11}$$

We will show that

$$n \times x = -(b \cdot x)a + (a \cdot x)b. \tag{12.2.1.12}$$

It's obvious that

$$n \cdot (n \times x) = 0. \tag{12.2.1.13}$$

By Theorem 11.1 and (12.2.1.9), we have

$$a \cdot (n \times x) = x \cdot (a \times n) = -x \cdot b. \tag{12.2.1.14}$$

and

$$b \cdot (n \times x) = x \cdot (b \times n) = x \cdot a. \tag{12.2.1.15}$$

The 3 equations above are equivalent to (12.2.1.12), which completes the proof. □

Example 12.2.1.4. *Rotation in $\mathbb{R}^3$. Let $u_{(\theta)}$ be a rotation about an axis specified by unit vector n with an angle θ measured couterclockwise if seen from the tip of n. The generator of $f \mapsto f \circ u_{(\theta)}$ is*

$$f \mapsto n \cdot (x \times \nabla f). \tag{12.2.1.16}$$

Proof. Note that by Theorem 11.1 $\nabla f \cdot (n \times x) = n \cdot (x \times \nabla f)$. □

For the benefit of the reader, let's write explicitly generator from (12.2.1.16):

$$n \cdot \left(x^2\frac{\partial}{\partial x^3} - x^3\frac{\partial}{\partial x^2},\ x^3\frac{\partial}{\partial x^1} - x^1\frac{\partial}{\partial x^3},\ x^1\frac{\partial}{\partial x^2} - x^2\frac{\partial}{\partial x^1}\right). \tag{12.2.1.17}$$

12.3 Tensor Analysis

12.3.1 Vectors and dual vectors

Let V be a vector space and V^* be a dual space to the V. Let $\{e_i\}$ be an arbitrary base of V. $\{e^i\} \subset V^*$ is a dual base to $\{e_i\}$ iff

$$\boxed{e^i(e_j) = \delta^i_j} \tag{12.3.1.1}$$

Let $\{\hat{e}_i\}$ be another base of V and $\{\hat{e}^i\}$ be its dual base. We have two transition matrices $[X^i_{\hat{j}}]$ and $[X^{\hat{i}}_j]$ such as

$$\hat{e}_k = \sum_i X^i_{\hat{k}} e_i, \tag{12.3.1.2}$$

$$\hat{e}^k = \sum_i X^{\hat{k}}_i e^i. \tag{12.3.1.3}$$

Fact 12.3.1.1. $\sum_i X^l_{\hat{i}} X^{\hat{i}}_k = \delta^l_k$.

Proof. $\delta^l_k = \hat{e}^l(\hat{e}_k) = (\sum_i X^{\hat{l}}_i e^i)(\sum_j X^j_{\hat{k}} e_j) = \sum_i \sum_j X^{\hat{l}}_i X^j_{\hat{k}} e^i(e_j)$
$= \sum_i \sum_j X^{\hat{l}}_i X^j_{\hat{k}} \delta^i_j = \sum_i X^{\hat{l}}_i X^i_{\hat{k}}$. □

Corollary 12.3.1.2.

$$e_j = \sum_k X^{\hat{k}}_j \hat{e}_k,$$

$$e^j = \sum_k X^j_{\hat{k}} \hat{e}^k.$$

Corollary 12.3.1.3. *If* $\sum_j v^j e_j = \sum_j \hat{v}^j \hat{e}_j$ *then* $\hat{v}^k = \sum_j X^{\hat{k}}_j v^j$.

Corollary 12.3.1.4. *If* $\sum_j v_j e^j = \sum_j \hat{v}_j \hat{e}^j$ *then* $\hat{v}_k = \sum_j X^j_{\hat{k}} v_j$.

12.3.2 Tensors

Definition 12.3.2.1. *A tensor of type* $\binom{m}{n}$ *is a multilinear function* $T : (V^*)^m \times V^n \to \mathbb{R}$.

Let $\mathcal{T}\binom{m}{n}$ be a space of all tensors of the type $\binom{m}{n}$. It is trivial to notice that $\mathcal{T}\binom{m}{n}$ is a vector space.

Note that $\mathcal{T}\binom{0}{1} = V^*$. And as we can define for $v \in V$ and $w^* \in V^*$, $v(w^*) := w^*(v)$, we will identify V with $\mathcal{T}\binom{1}{0}$.

Notice that because of multilinearity tensor is uniquely defined by the set of numbers $\{T^{\mu_1\ldots\mu_m}_{\nu_1\ldots\nu_n}\}$, where

$$T^{\mu_1\ldots\mu_m}_{\nu_1\ldots\nu_n} = T(e^{\mu_1},\ldots,e^{\mu_m};e_{\nu_1},\ldots,e_{\nu_n}). \tag{12.3.2.1}$$

It is now obvious that $\dim \mathcal{T}\binom{m}{n} = dim(V)^{m+n}$. Let

$$\hat{T}^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} = T(\hat{e}^{\alpha_1},\ldots,\hat{e}^{\alpha_m};\hat{e}_{\beta_1},\ldots,\hat{e}_{\beta_n}) \tag{12.3.2.2}$$

Fact 12.3.2.2. $\hat{T}^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} = \sum_{\mu_1,\ldots\mu_m;\nu_1\ldots\mu_n} X^{\hat{\alpha_1}}_{\mu_1}\ldots X^{\hat{\alpha_m}}_{\mu_m} X^{\nu_1}_{\hat{\beta_1}}\ldots X^{\nu_n}_{\hat{\beta_n}} T^{\mu_1\ldots\mu_m}_{\nu_1\ldots\nu_n}$.

Proof. $\hat{T}^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} = T(\hat{e}^{\alpha_1},\ldots,\hat{e}^{\alpha_m};\hat{e}_{\beta_1},\ldots,\hat{e}_{\beta_n})$
$= T(\sum_{\mu_1} X^{\hat{\alpha_1}}_{\mu_1} e^{\mu_1},\ldots;\sum_{\nu_1} X^{\nu_1}_{\hat{\beta_1}} e_{\nu_1},\ldots)$
$= \sum_{\mu_1,\ldots\mu_m;\nu_1\ldots\mu_n} X^{\hat{\alpha_1}}_{\mu_1}\ldots X^{\hat{\alpha_m}}_{\mu_m} X^{\nu_1}_{\hat{\beta_1}}\ldots X^{\nu_n}_{\hat{\beta_n}} T(e^{\mu_1},\ldots,e^{\mu_m};e_{\nu_1},\ldots,e_{\nu_n})$. □

Note that this is enough to define tensor product and contraction on base vectors.

Definition 12.3.2.3. $(S\otimes T)(e^{\mu_1},\ldots,e^{\mu_m},e^{\mu'_1},\ldots,e^{\mu'_{m'}};e_{\nu_1},\ldots,e_{\nu_n},e_{\nu'_1},\ldots,e_{\nu'_{n'}}) =$
$S(e^{\mu_1},\ldots,e^{\mu_m};e_{\nu_1},\ldots,e_{\nu_n})T(e^{\mu'_1},\ldots,e^{\mu'_{m'}};e_{\nu'_1},\ldots,e_{\nu'_{n'}})$.

Corollary 12.3.2.4. *If $T \in \mathcal{T}\binom{m}{n}$, then*

$$T = \sum_{\mu_1,\ldots\mu_m;\nu_1\ldots\mu_n} T(e^{\mu_1},\ldots,e^{\mu_m};e_{\nu_1},\ldots,e_{\nu_n}) e_{\mu_1}\otimes\cdots\otimes e_{\mu_m}\otimes e^{\nu_1}\otimes\cdots\otimes e^{\nu_n}. \tag{12.3.2.3}$$

Using coefficients from 12.3.2.1 we have:

$$\boxed{T = \sum_{\mu_1,\ldots\mu_m;\nu_1\ldots\mu_n} T^{\mu_1\ldots\mu_m}_{\nu_1\ldots\nu_n} e_{\mu_1}\otimes\cdots\otimes e_{\mu_m}\otimes e^{\nu_1}\otimes\cdots\otimes e^{\nu_n}.} \tag{12.3.2.4}$$

Remark 12.3.2.5. $e_i \otimes e^j = e^j \otimes e_i$.

Definition 12.3.2.6. $C\binom{i}{j} : \mathcal{T}\binom{m}{n} \to \mathcal{T}\binom{m-1}{n-1}$ *such as*
$(C\binom{i}{j}T)(e^{\mu_1},\ldots,e^{\mu_{m-1}};e_{\nu_1},\ldots,e_{\nu_{n-1}}) = \sum_\lambda T(e^{\mu_1},\ldots,e^\lambda,\ldots,e^{\mu_m};e_{\nu_1},\ldots,e_\lambda,\ldots e_{\nu_n})$.

It is easy to show that tensor product and contraction are base independent. We will use "base-less" tensor notation with Latin indices as introduced in [**?**].

Lemma 12.3.2.7. *If a tensor $T^{a_1\ldots a_m}_{b_1\ldots b_n}(v_c)$ depends linearly on v_c, then there exists a tensor $S^{ca_1\ldots a_m}_{b_1\ldots b_n}$ such that $T^{a_1\ldots a_m}_{b_1\ldots b_n}(v_c) = S^{ca_1\ldots a_m}_{b_1\ldots b_n} v_c$.*

Proof. Let $T^{a_1\dots a_m}_{b_1\dots b_n}(v_c) = \sum_{\mu_1,\dots\mu_m;\nu_1\dots\mu_n} T^{\mu_1\dots\mu_m}_{\nu_1\dots\nu_n}(v_c) e_{\mu_1} \otimes \dots \otimes e_{\mu_m} \otimes e^{\nu_1} \otimes \dots \otimes e^{\nu_n}$. $T^{\mu_1\dots\mu_m}_{\nu_1\dots\nu_n}(v_c)$ depends linearly on v_c thus we can choose such coefficients $S^{\lambda\mu_1\dots\mu_m}_{\nu_1\dots\nu_n}$ that $T^{\mu_1\dots\mu_m}_{\nu_1\dots\nu_n}(v_c) = \sum_\lambda S^{\lambda\mu_1\dots\mu_m}_{\nu_1\dots\nu_n} v_\lambda$. □

Corollary 12.3.2.8. *If a tensor $T^{a_1\dots a_m}_{b_1\dots b_n}(v^c)$ depends linearly on v^c, then there exists a tensor $S^{a_1\dots a_m}_{cb_1\dots b_n}$ such that $T^{a_1\dots a_m}_{b_1\dots b_n}(v^c) = S^{a_1\dots a_m}_{cb_1\dots b_n} v^c$.*

12.3.3 Manifolds

Let M be a manifold, $p \in M$. Let $x = (x_1, ..., x_l)$ be a coordinates system in M, then assosiated base of a tangent space V_p is

$$\boxed{e_i(f) = \frac{\partial}{\partial x^i}(f \circ x^{-1})} \tag{12.3.3.1}$$

Let $\hat{x}$ be another coordinate system. Then $\hat{e}_i(f) = \frac{\partial}{\partial \hat{x}^i}(f \circ \hat{x}^{-1})$. With the application of chain rule, we get:

$$\hat{e}_k(f) = \frac{\partial}{\partial \hat{x}^k}(f \circ \hat{x}^{-1}) = \sum_i \frac{\partial x^i}{\partial \hat{x}^k} \frac{\partial}{\partial x^i}(f \circ x^{-1}) = \sum_i \frac{\partial x^i}{\partial \hat{x}^k} e_k(f). \tag{12.3.3.2}$$

Comparing with equation 12.3.1.2, we get transition matrix:

$$\boxed{X^i_{\hat{k}} = \frac{\partial x^i}{\partial \hat{x}^k}} \tag{12.3.3.3}$$

and

$$\boxed{X^{\hat{k}}_i = \frac{\partial \hat{x}^k}{\partial x^i}} \tag{12.3.3.4}$$

Let $\mathcal{T}_p\binom{m}{n}$ be a space of tensors associated with tangent space V_p. A tensor filed on manifold M is a following mapping:

$$M \ni p \to T \in \mathcal{T}_p\binom{m}{n}. \tag{12.3.3.5}$$

Let $T^{a_1\dots a_m}_{b_1\dots b_n}$ be a tensor field on manifold M. Assume that

$$T^{a_1\dots a_m}_{b_1\dots b_n} = \sum_{\mu_1,\dots\mu_m;\nu_1\dots\mu_n} T^{\mu_1\dots\mu_m}_{\nu_1\dots\nu_n} e_{\mu_1} \otimes \dots \otimes e_{\mu_m} \otimes e^{\nu_1} \otimes \dots \otimes e^{\nu_n}. \tag{12.3.3.6}$$

Then

$$\partial_c T^{a_1\dots a_m}_{b_1\dots b_n} = \sum_{\lambda,\mu_1,\dots\mu_m;\nu_1\dots\mu_n} \frac{\partial T^{\mu_1\dots\mu_m}_{\nu_1\dots\nu_n}}{\partial x^\lambda} e^\lambda \otimes e_{\mu_1} \otimes \dots \otimes e_{\mu_m} \otimes e^{\nu_1} \otimes \dots \otimes e^{\nu_n}. \tag{12.3.3.7}$$

Note that ∂ depends on coordinates system x.

Fact 12.3.3.1. *$\partial_c(TS) = (\partial_c T)S + T\partial_c S$.*

Fact 12.3.3.2. *$\partial_c(C\binom{i}{j}T) = C\binom{i}{j}(\partial_c T)$.*

In this formalism we treat scalar field as field of tensors of type $\binom{0}{0}$. We treat this scalar field as dependent on coordinates. In this formalism $\partial_c f$ is just a gradient of f.

$$\boxed{\partial_c f = \sum_\mu \frac{\partial f}{\partial x^\mu} e^\mu} \tag{12.3.3.8}$$

Note that for $v^a \in V_p$ formally

$$v^a(f) = v^a \partial_a f. \tag{12.3.3.9}$$

In case of a vector field it's just its matrix derivative.

$$\boxed{\partial_c v^a = \sum_{\mu,\nu} \frac{\partial v^\mu}{\partial x^\nu} e^\nu \otimes e_\mu} \tag{12.3.3.10}$$

12.3.4 Covariant derivative

Definition 12.3.4.1. *Operator ∇ mapping smooth tensor field of type $\binom{m}{n}$ into tensor field of type $\binom{m}{n+1}$ is a covariant derivative iff the following conditions holds:*

$$\nabla(\alpha S + \beta T) = \alpha \nabla T + \beta \nabla S, \tag{12.3.4.1}$$

$$\nabla(S \otimes T) = \nabla S \otimes T + S \otimes \nabla T, \tag{12.3.4.2}$$

$$\nabla(C\binom{i}{j}T) = C\binom{i}{j}(\nabla T), \tag{12.3.4.3}$$

$$\nabla f = \partial f \textit{ for each smooth scalar field } f, \tag{12.3.4.4}$$

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f \textit{ for each smooth scalar field } f. \tag{12.3.4.5}$$

Definition 12.3.4.2. *Let $T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n}$ be coefficients of a tensor field T in a coordinates system x, then coefficients of a tensor field ∇T in a coordinates system x will be denoted as $\nabla_\lambda T^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n}$.*

Remark 12.3.4.3. *In established coordinate system x, derivative ∂ is a special case of covariant derivative.*

12.4 Introduction to Matrix Calculus

We will be using tensor calculus to prove swiftly equations which are written in matrix notation.

It doesn't really matter at this point to know what a tensor is in physics. For us all which matters is that they are variables indexed by bunch of lower or/and upper indices like that:

$$T_i, A^i_j, B^{ij}_{klm}, etc.$$

We will use Einstein summation convention. Meaning, we will omit sigma sign for the sum of a number of multiplications over index which is lower in one factor and upper in the other. Like this:

$$A^i_k B^k_j = \sum_{k=1}^{n} A^i_k B^k_j$$

or this (in case of summation over more than one index):

$$A^{lk}_{pqj} B^{jr}_{mk} = \sum_{k=1}^{N} \sum_{j=1}^{n} A^{lk}_{pqj} B^{jr}_{mk}.$$

So be careful, the left side of the above equations is easy to confuse with only one multiplication. You must check if the same index is not in the upper and the lower position at a time. If it is - this is not just one multiplication, it is a summation of all such multiplications over this index and other indices if they are also in upper and lower positions (assuming the number of such multiplications is known from context, or the precise value of this number is irrelevant.).

Contravariant vectors will be column vectors and we will denote them with upper index (i.e. x^i), thus we will be operating in a scope of a numerator layout. We will call them vectors as long as not stated otherwise. Covariant vectors will be row vectors and we will denote them with lower index (i.e. x_i).

The identity between matrix and tensor is established as follows for matrix A of dim $n \times m$

$$A = \begin{bmatrix} A^1_1 & A^1_2 & A^1_3 & \dots & A^1_m \\ A^2_1 & A^2_2 & A^2_3 & \dots & A^2_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^n_1 & A^n_2 & A^n_3 & \dots & A^n_m \end{bmatrix}. \tag{12.4.0.1}$$

It is easy to remember than in numerator layout gradient of a function $f : \mathbb{R}^m \to R$, i. e. $\dfrac{\partial f}{\partial x^j}$ is a row vector.

More generally if f is a n dimensional function of m dimensional vector x, in numerator layout we have

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} & \frac{\partial f^1}{\partial x^3} & \dots & \frac{\partial f^1}{\partial x^m} \\ \frac{\partial f^2}{\partial x^1} & \frac{\partial f^2}{\partial x^2} & \frac{\partial f^2}{\partial x^3} & \dots & \frac{\partial f^2}{\partial x^m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial x^1} & \frac{\partial f^n}{\partial x^2} & \frac{\partial f^n}{\partial x^3} & \dots & \frac{\partial f^n}{\partial x^m} \end{bmatrix}. \tag{12.4.0.2}$$

As a side note we add that in denominator layout the above matrix would be transposed. But we will be not using denominator layout in this document.

For convenience we will use Kronecker deltas as follows

$$\delta_{ij} = \delta^i_j = \begin{cases} 1 \text{ for } i = j, \\ 0 \text{ for } i \neq j. \end{cases} \tag{12.4.0.3}$$

We will now define the transposition of a tensor as a generalization of matrix transposition.

Definition 12.4.0.1. *Let A be a tensor.*

$$(A^T)^{q_1 \dots q_m}_{p_1 \dots p_n} = A^{p_1 \dots p_n}_{q_1 \dots q_m}. \tag{12.4.0.4}$$

Note that:

$$(A^T)^{q_1 \dots q_m}_{p_1 \dots p_n} = \delta_{i_1 p_1} \dots \delta_{i_n p_n} \delta^{j_1 q_1} \dots \delta^{j_m q_m} A^{i_1 \dots i_n}_{j_1 \dots j_m}. \tag{12.4.0.5}$$

And in particular:

$$(A^T)^l_k = \delta_{ik} \delta^{jl} A^i_j, \tag{12.4.0.6}$$

and

$$(x^T)_j = \delta_{ij} x^i \tag{12.4.0.7}$$

Now we are going to see how matrix multiplication looks like when using tensors (mind Einstein summation convention!)

Theorem 12.4.0.2. *If A, B are matricies, then*

$$(AB)^i_j = A^i_k B^k_j. \tag{12.4.0.8}$$

And how the dot product looks like when using tensors (mind Einstein summation convention - but this is really last time I am reminding this.)

Theorem 12.4.0.3. *If x, y are vectors, then*

$$x^T y = \delta_{ij} x^i y^j. \tag{12.4.0.9}$$

Here are some basic facts about matrix and vector transposition that we will be heavily exploiting:

$$(AB)^T = B^T A^T, \tag{12.4.0.10}$$

$$(Ax)^T = x^T A^T, \tag{12.4.0.11}$$

$$(A+B)^T = A^T + B^T, \tag{12.4.0.12}$$

$$(x+y)^T = x^T + y^T, \tag{12.4.0.13}$$

where A, B are matrices and x, y are vectors.

As an example, we will prove $(AB)^T = B^T A^T$.

$$(B^T A^T)^i_j = (B^T)^i_k (A^T)^k_j = B^k_i A^j_k = A^j_k B^k_i = (AB)^j_i = ((AB)^T)^i_j. \tag{12.4.0.14}$$

As we will be dealing with different kind of derivatives, might be matrix over vector, vector over vector, vector over matrix etc. It is good to give one definition of derivative tensor over tensor. However, while you are relatively safe on the ground of our applications, I am not giving you any authorization to use this definition in the physical context without deeper understanding what tensor is in physics.

Definition 12.4.0.4. *Let U be a tensor function of a tensor X, then*

$$(\frac{\partial U}{\partial X})^{i_1 \ldots i_n p_1 \ldots p_k}_{j_1 \ldots j_m q_1 \ldots q_l} = \frac{\partial U^{i_1 \ldots i_n}_{j_1 \ldots j_m}}{\partial X^{q_1 \ldots q_l}_{p_1 \ldots p_k}}. \tag{12.4.0.15}$$

With the definition above we can at least understand the following concepts:

$$\frac{\partial u}{\partial x}, \frac{\partial u^T v}{\partial x}, \frac{\partial U}{\partial x}, etc.$$

where u, v are vector functions of vector x and U is a matrix function of vector x,

as well as,

$$\frac{\partial u}{\partial X}, \frac{\partial U}{\partial X}, etc.$$

,

where u is a vector function of a matrix X and U is a matrix function of a matrix X.

Now we will prove a few usefull equations, which we will be using later.

Theorem 12.4.0.5. *If x is a vector, then*

$$\frac{\partial x^T}{\partial x} = \delta_{ij}, \tag{12.4.0.16}$$

$$\frac{\partial x}{\partial x} = \delta^i_j = I. \tag{12.4.0.17}$$

Theorem 12.4.0.6. *If A is a constant matrix and u is a vector function of vector x.*

$$\frac{\partial A u}{\partial x} = A\frac{\partial u}{\partial x}. \tag{12.4.0.18}$$

Proof.

$$\frac{\partial A^i_j u^j}{\partial x^k} = \frac{\partial A^i_j}{\partial x^k} u^j + A^i_j \frac{\partial u^j}{\partial x^k} = 0 + A^i_j \frac{\partial u^j}{\partial x^k} = A\frac{\partial u}{\partial x}. \tag{12.4.0.19}$$

□

Corollary 12.4.0.7. *If A is a constant matrix, then*

$$\frac{\partial A x}{\partial x} = A. \tag{12.4.0.20}$$

Theorem 12.4.0.8. *If u, v are vector functions of vector x, then*

$$\frac{\partial u^T v}{\partial x} = u^T\frac{\partial v}{\partial x} + v^T\frac{\partial u}{\partial x}. \tag{12.4.0.21}$$

We will use $[\cdot]^i_j$ as a formal operator, which acts on a matrix and returns an element from $i-th$ row and $j-th$ column.

Proof.

$$\frac{\partial u^T v}{\partial x} = \frac{\partial \delta_{ij} u^i v^j}{\partial x^k} = \delta_{ij} u^i \frac{\partial v^j}{\partial x^k} + \delta_{ij} v^j \frac{\partial u^i}{\partial x^k} = u^T \frac{\partial v}{\partial x} + v^T \frac{\partial u}{\partial x}. \tag{12.4.0.22}$$

□

Theorem 12.4.0.9. *Let A be a matrix of size $N \times M$ partitioned in matrices A^i_j (of size $n_i \times m_j$) in the following way*

$$A = \begin{bmatrix} A^1_1 & \dots & A^1_q \\ \vdots & \ddots & \vdots \\ A^p_1 & \dots & A^p_q \end{bmatrix} \tag{12.4.0.23}$$

where $N = \sum_{i=1}^p n_i$ and $M = \sum_{j=1}^q m_j$,

and let B be a matrix of size $M \times L$ partitioned in matrices B^j_k (of size $m_j \times l_k$) in the following way

$$B = \begin{bmatrix} B^1_1 & \dots & B^1_r \\ \vdots & \ddots & \vdots \\ B^q_1 & \dots & B^q_r \end{bmatrix}, \tag{12.4.0.24}$$

where $L = \sum_{k=1}^r l_k$

then

$$AB = \begin{bmatrix} \sum_{u=1}^q A^1_u B^u_1 & \dots & \sum_{u=1}^q A^1_u B^u_r \\ \vdots & \ddots & \vdots \\ \sum_{u=1}^q A^p_u B^u_1 & \dots & \sum_{u=1}^q A^p_u B^u_r \end{bmatrix}, \tag{12.4.0.25}$$

where sum and multiplications in expressions $\sum_{u=1}^q A^i_u B^u_k$ are operations on matrices, hence these expressions denotes also matrices of size $n_i \times l_k$ as partitions of a matrix AB (of size $N \times L$).

Proof. Note that in this context $[A]^i_j$ is as element of matrix A which is something very different from A^i_j which is just one of matrices into which A is partitioned.

Let us denote by C a matrix of size $N \times L$ partitioned in the matrices $C^i_k = \sum_{u=1}^q A^i_u B^u_k$ of size $n_i \times l_k$ in the following way

$$C = \begin{bmatrix} C^1_1 & \dots & C^1_r \\ \vdots & \ddots & \vdots \\ C^p_1 & \dots & C^p_r \end{bmatrix}. \tag{12.4.0.26}$$

We will show that $AB = C$.

Take arbitrary indecies $e = 1, 2, \ldots, N$ and $f = 1, 2 \ldots, L$. Note that we have

$$[AB]^e_f = \sum_{h=1}^{M} [A]^e_h [B]^h_f \tag{12.4.0.27}$$

Note that we can express index e as $e = \sum_{i=1}^{d-1} n_i + s$ where $d = 1, \ldots, p$ and $s = 1, \ldots, n_d$, we can express index f as $f = \sum_{k=1}^{g-1} l_k + t$ where $g = 1, \ldots, r$ and $t = 1, \ldots, l_g$ and finally, we can express index h as $h = \sum_{j=1}^{u-1} m_j + v$ where $u = 1, \ldots, q$ and $v = 1, \ldots, m_u$.

With the above substitutions, we have

$$[AB]^e_f = \sum_{h=1}^{M} [A]^e_h [B]^h_f = \sum_{u=1}^{q} \sum_{v=1}^{m_u} [A^d_u]^s_v [B^u_g]^v_t$$

$$= \sum_{u=1}^{q} (\sum_{v=1}^{m_u} [A^d_u]^s_v [B^u_g]^v_t) = \sum_{u=1}^{q} [A^d_u B^u_g]^s_t = [\sum_{u=1}^{q} A^d_u B^u_g]^s_t = [C^d_g]^s_t = [C]^e_f.$$

□

For tensor $g_{\mu\nu}$ we will represent it as matrix in a following way $[g]^\mu_\nu = g_{\nu\sigma}\delta^{\mu\sigma}$. In this context we will treat $[g]$ as matrix representation of $g_{\mu\nu}$. Don't confuse this with lowering and raising indices.

Theorem 12.4.0.10. *If $g_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho = g'_{\sigma\rho}$ then*

$$\Lambda^T[g]\Lambda = [g']. \tag{12.4.0.28}$$

Proof. By assumption we have

$$g_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho\delta^{\rho\eta} = g'_{\sigma\rho}\delta^{\rho\eta} \tag{12.4.0.29}$$

and thus by expanding $g_{\mu\nu}$ with identity.

$$g_{\mu\gamma}\delta^{\gamma\zeta}\delta_{\zeta\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho\delta^{\rho\eta} = g'_{\sigma\rho}\delta^{\rho\eta}. \tag{12.4.0.30}$$

Thus

$$[\Lambda^T]^\eta_\zeta[g]^\zeta_\mu\Lambda^\mu_\sigma = [g']^\eta_\sigma. \tag{12.4.0.31}$$

□

12.4.1 Determinant

Let's denote by S_n a set of all permutations of $[1,n]\cap\mathbb{N}$. Note that $|S_n| = n!$.

Definition 12.4.1.1. *Let $\sigma \in S_n$.*

$$dis(\sigma) = |\{(i,j) : i < j \text{ and } \sigma(j) < \sigma(i)\}|, \tag{12.4.1.1}$$

$$sgn\sigma = (-1)^{dis(\sigma)} = (-1)^{\sigma}. \tag{12.4.1.2}$$

A pair $i < j$ in permutation for which $\sigma(j) < \sigma(i)$ will be called disorder and number $\mathrm{dis}(\sigma)$ is number of disorders for σ.

Definition 12.4.1.2. *For $i, j = 1, \ldots, n$, we define a permutation $[i \leftrightarrow j] \in S_n$*

$$[i \leftrightarrow j](k) = \begin{cases} i \text{ for } k = j, \\ j \text{ for } k = i, \\ k \text{ for } k \notin \{i,j\}. \end{cases} \tag{12.4.1.3}$$

Definition 12.4.1.3. *Let A be a matrix of size $n \times n$*

$$detA = \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} [A]^i_{\sigma(i)}. \tag{12.4.1.4}$$

Let's define operation on column of matrices.

Definition 12.4.1.4. *Let A be a matrix of size $n \times n$*

$$[[w^i]_j \to A]^p_q = \begin{cases} w^p \text{ for } q = j, \\ [A]^i_j \text{ otherwise}. \end{cases} \tag{12.4.1.5}$$

In simple words $[w^i]_j \to A$ replaces a column j with a column vector w^i. From Definition 12.4.1.3 immediately follows:

Theorem 12.4.1.5. *If A is a matrix of size $n \times n$, then*

$$\det\big([\alpha A^i_j + w^i]_j \to A\big) = \alpha \det(A) + \det\big([w^i]_j \to A\big). \tag{12.4.1.6}$$

Proof. Let $E = ([\alpha A^i_j + w^i]_j \to A)$. Consider

$$detE = \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} [E]^i_{\sigma(i)}. \tag{12.4.1.7}$$

We can rearrange terms of multiplication sorting by lower index an get

$$\begin{aligned} detE &= \sum_{\sigma\in S_n}(-1)^\sigma\prod_{i=1}^n[E]_i^{\sigma^{-1}(i)} \\ &= \sum_{\sigma\in S_n}(-1)^\sigma\prod_{i=1}^{j-1}[A]_i^{\sigma^{-1}(i)}(\alpha A_j^{\sigma^{-1}(i)}+w^{\sigma^{-1}(i)})\prod_{i=j+1}^n[A]_i^{\sigma^{-1}(i)} \\ &= \alpha\det(A)+\det\big([w^i]_j\to A\big). \end{aligned}$$

□

Theorem 12.4.1.6. *If A is a matrix of size $n\times n$ and $j<k\le n$ such that $[A]_j^i=[A]_k^i$ for $i=1,\dots,n$, then*

$$\det A=0. \tag{12.4.1.8}$$

Proof. Let $S_n^\star=\{\sigma\in S_n:\sigma^{-1}(j)<\sigma^{-1}(k)\}$. Note that $S_n=S_n^\star\,\dot\cup\,([j\leftrightarrow k]S_n^\star)$. Then note

$$\begin{aligned} \det A &= \sum_{\sigma\in S_n^\star}(-1)^\sigma\prod_{i=1}^n[A]^i_{\sigma(i)}+\sum_{\sigma\in S_n^\star}(-1)^{[j\leftrightarrow k]\sigma}\prod_{i=1}^n[A]^i_{[j\leftrightarrow k]\sigma(i)} \\ &= \sum_{\sigma\in S_n^\star}(-1)^\sigma\prod_{i=1}^n[A]^i_{\sigma(i)}-\sum_{\sigma\in S_n^\star}(-1)^\sigma\prod_{i=1}^n[A]^i_{\sigma(i)}=0. \end{aligned} \tag{12.4.1.9}$$

□

Lemma 12.4.1.7. *For any sequence of integers $a_1,\cdots,a_n$ parity of the number*

$$\delta_1a_1+\cdots+\delta_na_n \tag{12.4.1.10}$$

doesn't depend on a choice of factors $\delta_1,\dots,\delta_n\in\{-1,1\}$.

Proof. Note that

$$2\Big|\delta_ia_i-a_i.$$

And since

$$2\Big|\sum_{i=1}^{n}(\delta_i a_i - a_i), \tag{12.4.1.11}$$

we have

$$2\Big|\sum_{i=1}^{n}\delta_i a_i - \sum_{i=1}^{n} a_i. \tag{12.4.1.12}$$

□

Theorem 12.4.1.8. *For any $\sigma_1, \sigma_2 \in S_n$, we have*

$$sgn(\sigma_1\sigma_2) = sgn(\sigma_1)sgn(\sigma_2). \tag{12.4.1.13}$$

Proof. Let's consider two sets of pairs

$$A^- = \{(i,j) : i < j \text{ and } \sigma_2(j) < \sigma_2(i)\}$$

and

$$A^+ = \{(i,j) : i < j \text{ and } \sigma_2(i) < \sigma_2(j)\}.$$

Note that

$$\Big|\{(i,j) \in A^- : \sigma_1(\sigma_2(j)) < \sigma_1(\sigma_2(i))\Big| = |A^-| - k_1$$

and

$$\Big|\{(i,j) \in A^+ : \sigma_1(\sigma_2(j)) < \sigma_1(\sigma_2(i))\Big| = k_2$$

where $k_1 + k_2$ is a total number of disorders for σ_1. Thus the total number of disorders for $\sigma_1\sigma_2$ is $|A^-| - k_1 + k_2$. Note that

$$\begin{aligned} \mathrm{sgn}(\sigma_1\sigma_2) &= (-1)^{|A^-|-k_1+k_2} = (-1)^{|A^-|+k_1+k_2} \\ &= (-1)^{|A^-|}(-1)^{k_1+k_2} = \mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2), \end{aligned}$$

where equality $(-1)^{|A^-|-k_1+k_2} = (-1)^{|A^-|+k_1+k_2}$ holds because of Lemma 12.4.1.7. □

As an easy corollary we will formulate the following.

Theorem 12.4.1.9. *For any $\sigma \in S_n$, we have*

$$sgn(\sigma) = sgn(\sigma^{-1}). \tag{12.4.1.14}$$

Sometimes it is convenient to consider a permutation $\sigma \in S_n$ as $\sigma : \mathbb{N} \setminus \{0\} \xrightarrow{1-1} \mathbb{N} \setminus \{0\}$ with $\sigma(k) = k$ for $k > n$.

In that interpretation, for any $\sigma : \mathbb{N} \setminus \{0\} \xrightarrow{1-1} \mathbb{N} \setminus \{0\}$ we will define $\mathrm{ord}(\sigma) = \max\{k : \sigma(k) \neq k\}$ with $\mathrm{ord}(I) = 0$ where I is identity. In this context we consider all $\sigma : \mathbb{N} \setminus \{0\} \xrightarrow{1-1} \mathbb{N} \setminus \{0\}$ for which $\mathrm{ord}(\sigma) < +\infty$ as permutations.

With this interpretation in mind let's define $S_{k,n}$ as set of all permutations σ for which $\sigma(i) = i$ for $i < k$ and $s(i) = i$ for $i > n$.

Theorem 12.4.1.10. *If $ord(\sigma) < +\infty$ then there exists an integer $m > 0$ and integers $i_k \neq j_k$ for $k = 1, \ldots, m$ such that*

$$\sigma = [i_1 \leftrightarrow j_1] \ldots [i_m \leftrightarrow j_m] \tag{12.4.1.15}$$

Proof. We will prove this by induction over $\mathrm{ord}(\sigma)$. The thesis is true for $\mathrm{ord}(\sigma) = 0$, because $I = [1 \leftrightarrow 2][1 \leftrightarrow 2]$. Assume that thesis holds for $\mathrm{ord}(\sigma) \leq n - 1$. Take any σ for which $\mathrm{ord}(\sigma) = n$. Note that $\mathrm{ord}(\sigma[n \leftrightarrow \sigma^{-1}(n)]) \leq n - 1$, thus by induction $\sigma[n \leftrightarrow \sigma^{-1}(n)] = [i_1 \leftrightarrow j_1] \ldots [i_m \leftrightarrow j_m]$ and

$$\sigma = \sigma[n \leftrightarrow \sigma^{-1}(n)][n \leftrightarrow \sigma^{-1}(n)] = [i_1 \leftrightarrow j_1] \ldots [i_m \leftrightarrow j_m][n \leftrightarrow \sigma^{-1}(n)]. \tag{12.4.1.16}$$

□

Theorem 12.4.1.11. *If $ord(\sigma) < +\infty$ and there exists an integer $m > 0$ and integers $i_k \neq j_k$ for $k = 1, \ldots, m$ such that*

$$\sigma = [i_1 \leftrightarrow j_1] \ldots [i_m \leftrightarrow j_m], \tag{12.4.1.17}$$

then

$$(-1)^m = sgn(\sigma). \tag{12.4.1.18}$$

Proof. It is enough to note that $\mathrm{sgn}([i \leftrightarrow j]) = -1$. □

Theorem 12.4.1.12. *If A is a matrix of size $n \times n$, then $\det A = \det A^T$.*

Proof. Consider

$$detA = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} [A]^i_{\sigma(i)}. \tag{12.4.1.19}$$

We can rearrange terms of multiplication sorting by lower index an get

$$detA = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} [A]^{\sigma^{-1}(i)}_i = \sum_{\sigma \in S_n} (-1)^{\sigma^{-1}} \prod_{i=1}^{n} [A^T]^i_{\sigma^{-1}(i)} = detA^T. \tag{12.4.1.20}$$

□

Theorem 12.4.1.13. *Let A be a matrix of size $n \times n$ and let B be a matrix of $m \times m$ and C be a matrix of size $n \times m$, then*

$$\det \begin{bmatrix} A & 0 \\ C & B \end{bmatrix} = \det A \det B. \tag{12.4.1.21}$$

Proof. Let

$$E = \det \begin{bmatrix} A & 0 \\ C & B \end{bmatrix}.$$

Consider

$$\det E = \sum_{\sigma \in S_{n+m}} (-1)^{\sigma} \prod_{i=1}^{n+m} [E]^{i}_{\sigma(i)}. \tag{12.4.1.22}$$

Take any σ such that $\sigma(i) \leq n$ and $i > n$ for some index i (i.e. $(i, \sigma(i))$ is in C area). Sine σ is $1-1$ and "onto", for some $k \leq n$ we must have $\sigma(k) > n$ (i.e. $(k, \sigma(k))$ is in 0 area), but that means that $[E]^{k}_{\sigma(k)} = 0$.
On the other hand if we take any σ such that $\sigma(i) > n$ and $i \leq n$ for some index i, then immediately $[E]^{i}_{\sigma(i)} = 0$.
Therefore only perturbations σ for which $\sigma(i) \leq n$ for all $i \leq n$ and $\sigma(i) > n$ for all $i > n$ can have non-zero contribution to $\det E$. Note that such a perturbation is a composition of perturbations $\sigma = \sigma_1\sigma_2$ where $\sigma_1 \in S_n$ and $\sigma_2 \in S_{n+1,n+m}$.
Thus

$$\begin{aligned}
\det E &= \sum_{\sigma_1 \in S_n} \sum_{\sigma_2 \in S_{n+1,n+m}} (-1)^{\sigma_1\sigma_2} \prod_{i=1}^{n+m} [E]^{i}_{\sigma_1\sigma_2(i)} \\
&= \sum_{\sigma_1 \in S_n} \sum_{\sigma_2 \in S_{n+1,n+m}} (-1)^{\sigma_1}(-1)^{\sigma_2} \prod_{i=1}^{n} [E]^{i}_{\sigma_1(i)} \prod_{i=n+1}^{n+m} [E]^{i}_{\sigma_2(i)} \\
&= \left(\sum_{\sigma_1 \in S_n} (-1)^{\sigma_1} \prod_{i=1}^{n} [E]^{i}_{\sigma_1(i)} \right) \left(\sum_{\sigma_2 \in S_{n+1,n+m}} (-1)^{\sigma_2} \prod_{i=n+1}^{n+m} [E]^{i}_{\sigma_2(i)} \right) \\
&= \det A \det B.
\end{aligned}$$

□

The idea of the proof of the theorem below comes from [**?**]

Theorem 12.4.1.14. *If $\eta = diag(a_1, \ldots, a_n)$ where $a_i \neq 0$ for $i = 1, \ldots, n$ and A is a matrix of size $n \times n$ such that $A^T \eta A = \eta$, then for the partition*

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}, \tag{12.4.1.23}$$

where B and E are square matrices, we have $\det B = \det E$.

Proof. Let's assume that

$$\eta = \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix}, \tag{12.4.1.24}$$

where η_1 and η_2 are required diagonal matrices.
Let's calculate

$$\eta = A^T \eta A = \begin{bmatrix} B^T & D^T \\ C^T & E^T \end{bmatrix} \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$
$$= \begin{bmatrix} B^T \eta_1 & D^T \eta_2 \\ C^T \eta_1 & E^T \eta_2 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix} = \begin{bmatrix} B^T \eta_1 B + D^T \eta_2 D & B^T \eta_1 C + D^T \eta_2 E \\ C^T \eta_1 B + E^T \eta_2 D & C^T \eta_1 C + E^T \eta_2 E \end{bmatrix}.$$

Then from the above, we have $B^T \eta_1 B + D^T \eta_2 D = \eta_1$ and $B^T \eta_1 C + D^T \eta_2 E = 0$. Therefore the following holds

$$\begin{bmatrix} B^T \eta_1 & D^T \eta_2 \\ 0 & I \end{bmatrix} A = \begin{bmatrix} \eta_1 & 0 \\ D & E \end{bmatrix} \tag{12.4.1.25}$$

By Theorem 12.4.1.13, we have $\det\left(B^T \eta_1\right) = \det(\eta_1) \det(E)$ and thus $\det B = \det E$. ☐

Theorem 12.4.1.15. *If $\eta = diag(\delta_1, \ldots, \delta_n)$ and $\eta' = diag(\delta'_1, \cdots, \delta'_n)$ where $\delta_i, \delta'_i \in \{-1, 1\}$ for $i = 1, \ldots, n$ and A is a matrix of size $n \times n$ such that $A^T \eta A = \eta'$, then for the partition*

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}, \tag{12.4.1.26}$$

where B and E are square matrices, we have $|\det B| = |\det E|$.

Proof. Let's assume that

$$\eta = \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix}, \tag{12.4.1.27}$$

and

$$\eta' = \begin{bmatrix} \eta_1' & 0 \\ 0 & \eta_2' \end{bmatrix}, \tag{12.4.1.28}$$

where η_1 and η_2 are required diagonal matrices.
Let's calculate

$$\eta' = A^T \eta A = \begin{bmatrix} B^T & D^T \\ C^T & E^T \end{bmatrix} \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$
$$= \begin{bmatrix} B^T\eta_1 & D^T\eta_2 \\ C^T\eta_1 & E^T\eta_2 \end{bmatrix} \begin{bmatrix} B & C \\ D & E \end{bmatrix} = \begin{bmatrix} B^T\eta_1 B + D^T\eta_2 D & B^T\eta_1 C + D^T\eta_2 E \\ C^T\eta_1 B + E^T\eta_2 D & C^T\eta_1 C + E^T\eta_2 E \end{bmatrix}.$$

Then from the above, we have $B^T\eta_1 B + D^T\eta_2 D = \eta_1'$ and $B^T\eta_1 C + D^T\eta_2 E = 0$. Therefore the following holds

$$\begin{bmatrix} B^T\eta_1 & D^T\eta_2 \\ 0 & I \end{bmatrix} A = \begin{bmatrix} \eta_1' & 0 \\ D & E \end{bmatrix} \tag{12.4.1.29}$$

By Theorem 12.4.1.13, we have $\det\big(B^T\eta_1\big) = \det(\eta_1')\det(E)$ and thus $|\det B| = |\det E|$. $\square$

We will give a nice geometrical interpretation of the above theorem. We will use $|\cdot|_S$ to denote volume relative to subspace S and P_S to denote an orthogonal projection onto a subspace S.

Theorem 12.4.1.16. *Let X be a real vector space where $\dim X = n$ with a metric tensor g. Let $e_1, \ldots, e_n$ be orthonormal basis of X and let $\hat{e}_1, \ldots, \hat{e}_n$ be an orthonormal basis of X. Let $\hat{V} = span\{\hat{e}_1, \ldots, \hat{e}_k\}$ and $V = span\{e_1, \ldots, e_k\}$ for some $k \in \{1, \ldots, n-1\}$. For $A \subset \hat{V}$ and $B \subset \hat{V}^\perp$, we have*

$$\frac{|A|_{\hat{V}}}{|P_V(A)|_V} = \frac{|B|_{\hat{V}^\perp}}{|P_{V^\perp}(B)|_{V^\perp}}. \tag{12.4.1.30}$$

Proof. Let's write transition matrix from basis e to basis $\hat{e}$.

$$\hat{e}_\mu = E^\nu_{\hat{\mu}} e_\nu. \tag{12.4.1.31}$$

Note that

$$P_V(\hat{e}_\mu) = \sum_{\nu=1}^{k} E^\nu_{\hat{\mu}} e_\nu \tag{12.4.1.32}$$

and

$$P_{V^\perp}(\hat{e}_\mu) = \sum_{\nu=k+1}^{n} E^{\nu}_{\hat{\mu}} e_\nu. \tag{12.4.1.33}$$

Let's divide matrix $E_{\hat{\cdot}}$ into 4 partitions as follows:

$$\begin{bmatrix} E^1_{\hat{1}} & \dots & E^1_{\hat{k}} & E^1_{\widehat{k+1}} & \dots & E^1_{\hat{n}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ E^k_{\hat{1}} & \dots & E^k_{\hat{k}} & E^k_{\widehat{k+1}} & \dots & E^k_{\hat{n}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ E^n_{\hat{1}} & \dots & E^n_{\hat{k}} & E^n_{\widehat{k+1}} & \dots & E^n_{\hat{n}} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}, \tag{12.4.1.34}$$

where size of E_{11} is $k \times k$ and size of E_{22} is $(n-k) \times (n-k)$.

Note that $|(P_V(\hat{e}_1), \dots, P_V(\hat{e}_k))| = |\det E_{11}|$ and $|(P_{V^\perp}(\hat{e}_{k+1}), \dots, P_{V^\perp}(\hat{e}_n))| = |\det E_{22}|$.

But according to the laws of tensor transformation we have $\hat{g}_{\mu\nu} = g_{\rho\sigma} E^{\rho}_{\hat{\mu}} E^{\sigma}_{\hat{\nu}}$ (in this particular case you can use also Theorem 11.1.2.17) and by Theorem 12.4.0.10, we have $E^T_{\hat{\cdot}}[g]E_{\hat{\cdot}} = [\hat{g}]$. Thus by Theorem 12.4.1.15, we have $|\det E_{11}| = |\det E_{22}|$. Hence, thesis holds. $\square$

12.5 Review of Statistical Learning Methods

12.5.1 Linear Regression

Assume that we have our data in a $n \times m$ matrix X, where our data vectors are in rows and columns represent features (note that to deal with an intercept we might require that the initial column is set to 1.) Assume that we have also some column vector y with corresponding "results" for all row vectors from X. Our goal is to find linear coefficients to predict results from the data. Formally, we are looking for a column vector β, which minimizes the quadratic error (residual sum of squares):

$$\mathrm{RSS}(\beta) = \sum_{i=1}^{n} (y^i - \sum_{j=1}^{m} X^i_j \beta^j)^2. \tag{12.5.1.1}$$

(We abandon Einstein summation convention here at our convenience). The above might be put very nicely using a matrix algebra.

$$\mathrm{RSS}(\beta) = (y - X\beta)^T (y - X\beta). \tag{12.5.1.2}$$

Given that our data X and result vector y are already established, RSS is in fact a scalar value function from the vector β. Our goal is to find a minimum of this function.

Let's start with the auxiliary definition of positive-defined matrix.

Definition 12.5.1.1. *Let A be a $m \times m$ matrix. Matrix A is positive-defined if $x^T(Ax) > 0$ for all non-zero vectors x.*

Now we are ready to formulate in a very elegant way a local minimum theorem.

Theorem 12.5.1.2. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a scalar function of a vector x with its first and second derivative continuous. If $\frac{\partial f}{\partial x} = 0$ at vector x_0 and a matrix $\frac{\partial}{\partial x}(\frac{\partial f}{\partial x})^T$ is positive-defined at x_0, then f has its local minimum at vector x_0.*

We will need also necessary condition for an extremum of f.

Theorem 12.5.1.3. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a scalar function of a vector x with its first derivative continuous. If x_0 is an extremum of the function f, then $\frac{\partial f}{\partial x} = 0$.*

Note that we use $\frac{\partial}{\partial x}(\frac{\partial f}{\partial x})^T$ to be consistent with matrix calculus.

By our tensor definition of derivative (in a numerator layout)

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x^i} = [\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^m}].$$

Thus $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^i \partial x^j}$ is in our formalism a tensor, but not a matrix. It has exactly the same coefficients as Hessian, but formally it is a tensor with two lower indices. Hessian matrix is equal exactly to $\frac{\partial}{\partial x}(\frac{\partial f}{\partial x})^T$ in our formalism.This careful distinction will be needed later when we do our matrix calculus to minimaze RSS function.

Let's now calculate $\frac{\partial \mathrm{RSS}}{\partial \beta}$. By Theorem 12.4.0.7 and Corollary 12.4.0.8, we have:

$$\frac{\partial \mathrm{RSS}}{\partial \beta} = -2(y - X\beta)^T X. \tag{12.5.1.3}$$

and

$$\frac{\partial}{\partial \beta}(\frac{\partial \text{RSS}}{\partial \beta})^T = -2\frac{\partial}{\partial \beta}(X^T(y - X\beta)) = 2X^TX. \qquad (12.5.1.4)$$

Last equality is granted by Theorem 12.4.0.19.

For simplicity, we need to assume that matrix X has a full column rank. This means that features are not linearly depended, which is a pretty valid assumption in majority of real application.

We will show that X^TX is positive-defined. Take any non-zero m-dimensional vector x.

$$x^TX^TXx = (Xx)^TXx = ||Xx||^2 > 0. \qquad (12.5.1.5)$$

The last inequality is granted by the fact that when X has full column rank and x is non-zero, Xx is also non-zero (otherwise columns would have been linearly dependent).

Thus $\frac{\partial}{\partial \beta}(\frac{\partial \text{RSS}}{\partial \beta})^T = 2X^TX$ is positive-defined everywhere.

It's enough now to find a solution of the equation $\frac{\partial \text{RSS}}{\partial \beta} = 0$, i.e.

$$(y - X\beta)^TX = 0. \qquad (12.5.1.6)$$

Which is the same as

$$X^T(y - X\beta) = 0, \qquad (12.5.1.7)$$

$$X^Ty = X^TX\beta, \qquad (12.5.1.8)$$

Since X^TX is positive-defined it's also invertible, therefore

$$\beta = (X^TX)^{-1}X^Ty. \qquad (12.5.1.9)$$

Since the solution is unique and $\frac{\partial}{\partial \beta}(\frac{\partial \text{RSS}}{\partial \beta})^T$ is positive-defined, by Theorem 12.5.1.2 and Theorem 12.5.1.3, we have found the only one local minimum that RSS has, thus the global minimum.

Note that with the linear regression we are in this lucky unique position, that we can analytically find a global minimum of quadratic error function, which means we are able to find literally the best ever coefficients β. With many other estimators, we will be quite happy when we find a few local minimums and choose merely the minimum of them.

12.5.2 Ridge Regression

Now when we are equipped in our tool-set for calculating minimum of the error function, we might want to alter slightly the error function. In cases when we have strongly correlated features in the data, it sometimes happens that some big coefficients might be chosen for them, which nearly cancel themselves in an average but unfortunately produce great variance. To avoid such situation we are adding penalties for too big coefficients.

This could be achieved by the following error function:

$$\mathrm{RSS}(\beta) = (y - X\beta)^T(y - X\beta) + \lambda\beta^T\beta, \tag{12.5.2.1}$$

where $\lambda > 0$ is a hyper-parameter. By the calculation from the previous sub-section and Theorem 12.4.0.8, we are getting

$$\frac{\partial \mathrm{RSS}}{\partial \beta} = -2(y - X\beta)^T X + 2\lambda\beta^T. \tag{12.5.2.2}$$

and with

$$\frac{\partial}{\partial \beta}(\frac{\partial \mathrm{RSS}}{\partial \beta})^T = 2(X^T X + \lambda I). \tag{12.5.2.3}$$

It is easy to show that the above matrix is positive-defined. $x^T(X^T X + \lambda I)x = (Xx)^T(Xx) + \lambda x^T x > 0$ for any non-zero vector x.

Note that this time, we don't need additional assumption that X has a full column rank. $X^T X + \lambda I$ is positive-defined even if X has no full rank because of $\lambda x^T x$ part which is always strictly grater than 0 for each non-zero vector x.

Now all we need to do is to solve $\frac{\partial \mathrm{RSS}}{\partial \beta} = 0$ equation.

$$-(y - X\beta)^T X + \lambda\beta^T = 0, \tag{12.5.2.4}$$

$$-X^T(y - X\beta) + \lambda\beta = 0, \tag{12.5.2.5}$$

$$(X^T X + \lambda I)\beta = X^T y \tag{12.5.2.6}$$

$$\beta = (X^T X + \lambda I)^{-1} X^T y. \tag{12.5.2.7}$$

Thus the above β for analogous reasons as in case of ordinary linear regression is a global minimum of the error function RSS. And this time $X^T X + \lambda I$ is reversible for an arbitrary data matrix X (because it's positive-defined).

As a side note it's worth to notice that adding and epsilon penalty for the size of coefficients removes mathematical problem of dealings with X when its columns are linearly dependent - and that was in fact a reason why Rige Regression was introduced in a first place.

Chapter 13

Physical Diary

Because not all material that I go through is ready in my head for the form of a nice logically written section in my notes, I decided to start writing this diary to summarise some pieces of information that I am gaining together with bibliography of the subject, to be able to return to this later.

13.1 Leeds, Saturday, 13 October 2018

I am trying to formulate my version of Quantum Mechanis postulates. I am not sure jet, if they are equivalent to generally accepted (e.g [**?**]). I am also not sure if they are original, they resemble a bit those from Feynman lectures. Let me first write them here and I will investigate it later. Anyway, the working name for these set of postulates is Leeds Version of Quantum Postulates, because I started to work on them in Leeds and this is only for my internal nomenclature perpouses.

All L^2 spaces in this subsection are spaces of complex valued functions.

Definition 13.1.0.1. *We will say that a (probabilistic) measurable space* $(X, \mathfrak{M}, \mu)$ *is a description of a physical system. We will say that* $x \in X$ *is a simple description of a state of physical system.*

Definition 13.1.0.2. *Descriptions* (X_1, μ_1) *and* (X_2, μ_2) *are equivalent iff there exists an unitary mapping*

$$U : L^2(X_1, \mu_1) \to L^2(X_1, \mu_1). \tag{13.1.0.1}$$

Definition 13.1.0.3. *Let* (X, μ) *be a description of physical system. An observable is any measurable function* $g : X \to \mathbb{R}$.

Definition 13.1.0.4. *Let (X, μ) be a description of physical system. Any $\psi \in L^2(X, \mu)$ is called description of physical state or superposition of simple descriptions.*

Definition 13.1.0.5. *We will say that (X, μ, K, E) $K : \mathbb{R} \to X^X$ is simple dynamical system iff*

1. *(X, μ) is a description of physical system.*

2. *$E : X \to \mathbb{R}$ is a measurable function*

3. $$K(t + s)x = K(t)K(s)x \tag{13.1.0.2}$$

   for each $t, s \in \mathbb{R}$ and

4. $$E(K(t)x) = const. \tag{13.1.0.3}$$

Axiom 13.1.0.6. *Time evolution U of description ψ under a simple dynamical system (X, μ, K, E) is given by*

$$(U(t)\psi)(x) = \exp(-itE(x))\psi(x). \tag{13.1.0.4}$$

13.2 Leeds, Saturday, 20 October 2018

13.2.1 Phase Space Interpretation

I was thinking about quantum postulates that give more natural reasons for finding hamiltonian in Schrödinger equation. That's why I started to write (unsuccesfull for now) Section 13.1. My idea is to be more inspired by quasi-distributions of momentum and position in a phase space. It turned out, as very often in such cases, that all of that is already done. We will use Poisson brackets in phase space.

$$\{f, g\} = \sum_i \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial x_i}\frac{\partial f}{\partial p_i}. \tag{13.2.1.1}$$

In classical phase space, (x, p) are goverened by Hamilton equations

$$\frac{\partial H}{\partial x} = -\dot{p}, \tag{13.2.1.2}$$

$$\frac{\partial H}{\partial p} = \dot{x}. \tag{13.2.1.3}$$

If we assume that probability density $f_t(x,p)$ is conserved in time, the Liouville's equation holds

$$\frac{\partial f_t}{\partial t} = -\{f_t, H\}. \tag{13.2.1.4}$$

And by that way the time evolution of probability density is described (See 2.3.2). Then I started to think about Wiegner quasi-probability distribution. Which is supposed to be the only candidate for joint distribution (such joit distribution usually doesn't exist) of momentum and position in state ψ. In one dimentional case:

$$W(x,p) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{isp}\psi(x+\frac{s}{2})\overline{\psi}(x-\frac{s}{2})ds. \tag{13.2.1.5}$$

The problem with the above function is that it's not everywhere positive. Mentioned in [**?**] and [see **?** , 2.1.4 Joint probabilities in quantum mechanics], more details in [see **?** , 3.4 Characteristic functions; The Wigner Distribution Function].

It turned out that time evolution of Wiegner function in phase space is a well researched topic. In case of one dimentional particle in potential field V, which means $H = \frac{p^2}{2m} + V$. The evolution of $W_t(x,p)$ takes form:

$$\frac{\partial W_t}{\partial t} = -\left(\frac{p}{m}\right)\frac{\partial W_t}{\partial x} + \sum_{n=0}^{\infty}\left(\frac{i}{2}\right)^{2n}\frac{1}{(2n+1)!}\cdot\frac{\partial^{2n+1}V}{\partial x^{2n+1}}\cdot\frac{\partial^{2n+1}W_t}{\partial p^{2p+1}}. \tag{13.2.1.6}$$

Note that for any V for which $\frac{\partial^3 V}{\partial x^3} = 0$, we have

$$\frac{\partial W_t}{\partial t} = -\{W_t, H\}. \tag{13.2.1.7}$$

Which holds for 2 important cases: 1. Free particle 2. Harmonic Oscilator. The above is discribed in [see **?** , 3.2 Time Dependence of the Wigner Function]. If we define Moyal bracket

$$\{\{f,g\}\} = \frac{2}{\hbar}\, f(x,p)\ \sin\left(\tfrac{\hbar}{2}(\overleftarrow{\partial}_x\overrightarrow{\partial}_p - \overleftarrow{\partial}_p\overrightarrow{\partial}_x)\right)\ g(x,p). \tag{13.2.1.8}$$

The general equation (Moyal's evolution equation) for a time evolution has a form:

$$\frac{\partial W_t}{\partial t} = -\{\{W_t, H\}\}. \tag{13.2.1.9}$$

[**?** , see].

13.2.2 Spin

I have found quite interesting book to read as an intruduction to spin: [?].

13.3 Thursday, 1 November 2018

13.3.1 Formal derivation of Euler–Lagrange equation

Today I wanted to examine derivation of Euler-Lagrange equation for classical mechanics before I derive the Euler-Lagrange equation for classical field in the Special Relativity context. We want to characterise $u : [a, b] \to \mathbb{R}^n$ which is stationary for the functional

$$u \mapsto \int_a^b L(t, u, \dot{u})dt. \tag{13.3.1.1}$$

where $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is "nice enough" (we don't consider now in details assumptions about L). This means we look for u for which

$$\delta \int_a^b L(t, u, \dot{u})dt = 0, \tag{13.3.1.2}$$

for variations δu vanishing at a and b. We assume that L is nice enough that we can move δ under integral. Note that

$$0 = \delta \int_a^b L(t, u, \dot{u})dt = \int_a^b \delta L(t, u, \dot{u})dt = \int_a^b \delta u \frac{\partial L}{\partial u} + \delta \dot{u} \frac{\partial L}{\partial \dot{u}} dt. \tag{13.3.1.3}$$

On the other hand,

$$\frac{d}{dt}(\delta u \frac{\partial L}{\partial \dot{u}}) = \delta \dot{u} \frac{\partial L}{\partial \dot{u}} + \delta u \frac{d}{dt} \frac{\partial L}{\partial \dot{u}}. \tag{13.3.1.4}$$

Since δu vanishes at a and b, we get:

$$0 = \int_a^b \delta \dot{u} \frac{\partial L}{\partial \dot{u}} dt + \int_a^b \delta u \frac{d}{dt} \frac{\partial L}{\partial \dot{u}} dt. \tag{13.3.1.5}$$

which means that

$$\int_a^b \delta \dot{u} \frac{\partial L}{\partial \dot{u}} dt = -\int_a^b \delta u \frac{d}{dt} \frac{\partial L}{\partial \dot{u}} dt. \tag{13.3.1.6}$$

Now, by (13.3.1.3) and (13.3.1.6) we have

$$0 = \int_a^b \delta u \left(\frac{\partial L}{\partial u} - \frac{d}{dt} \frac{\partial L}{\partial \dot{u}} \right) dt. \tag{13.3.1.7}$$

Since δu is arbitrary, we got Euler–Lagrange equation:

$$\boxed{\frac{\partial L}{\partial u} - \frac{d}{dt}\frac{\partial L}{\partial \dot{u}} = 0} \tag{13.3.1.8}$$

I have found derivation of the Euler–Lagrange equations for classic field theory (Special Relativity) in [see **?** , 8.5.2 The Hamilton Principle of Stationary Action]. The equation for a field ϕ^α is as follows:

$$\frac{\partial \mathcal{L}}{\partial \phi^\alpha} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\alpha)} \right) = 0. \tag{13.3.1.9}$$

In this context $\mathcal{L} = \mathcal{L}(\phi^\alpha, \partial_\mu \phi^\alpha, x^\mu)$. The proof requires first to understand spacetime version of Gauss's Theorem (because of special metric tensor it's not simply Theorem 12.1.2.17 for $\mathbb{R}^n$ case).

13.4 Saturday, 3 November 2018

13.4.1 Spacetime version of Gauss's Theorem

In this section we will use Einstein summation convention. Based on [see **?** , Appendix B] we can formulate the following:

Theorem 13.4.1.1. *Let M be a n-dimentional compact manifold with some metric tensor $g_{\mu\nu}$. Integrals below are over volume element induced by g. Let ∂M be a $(n-1)$-dimentional manifold, which is a boundary of M. Let n^μ be a continous vector field of unit normal vectors which are "pointing outward" if n^μ is spacelike ($g_{\mu\nu} n^\mu n^\nu < 0$) and "pointing inward" if n^μ is timelike ($g_{\mu\nu} n^\mu n^\nu > 0$). If v^μ is C^1 vector field on M, then*

$$\int_M \nabla_\mu v^\mu = \int_{\partial M} n_\mu v^\mu. \tag{13.4.1.1}$$

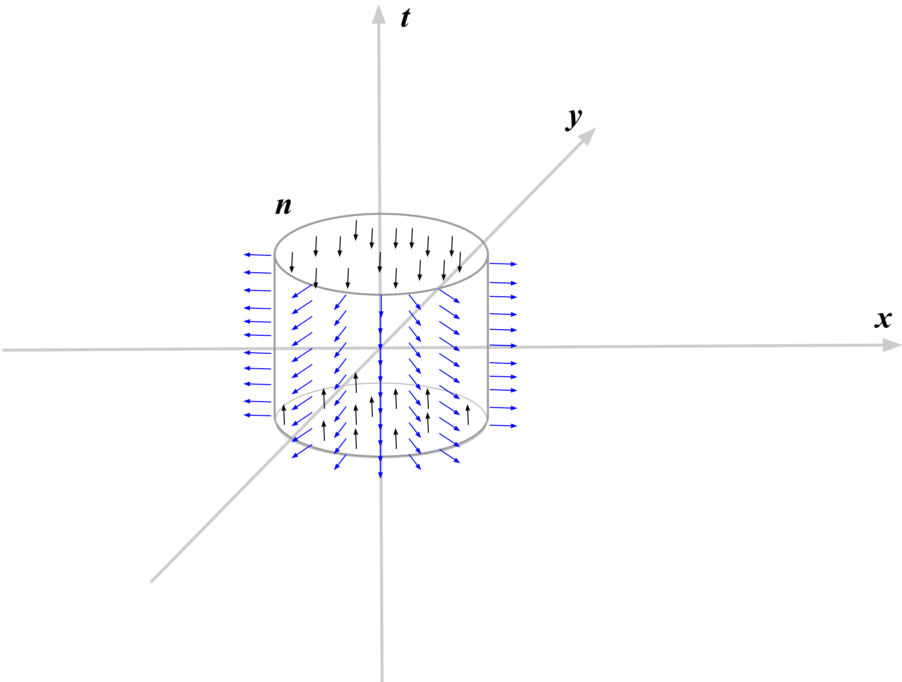

Figure 13.1: Normal vectors to the oriented surface in spacetime. Blue arrows indicate spacelike vectors. Black arrows indicate timelike vectors.

13.4.2 Euler–Lagrange equation for Classic Field Theory

Now we are ready to derive an equation (13.3.1.9) in the context of special relativity. For a vector field ϕ^α, let's define an action functional as

$$S(M, \phi^\alpha) := \int_M \mathcal{L}(\phi^\alpha, \partial_\mu \phi^\alpha, x^\mu) d^4x. \tag{13.4.2.1}$$

We want to find a characteristic of a field ϕ^α, for which for any variation $\delta\phi^\alpha$ which vanishes at ∂M, we have

$$\delta S = 0. \tag{13.4.2.2}$$

We assume that $\mathcal{L}$ is "nice enough" that we can go with δ under the integral. Thus we have:

$$0 = \int_M \frac{\partial \mathcal{L}}{\partial \phi^\alpha} \delta\phi^\alpha + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\alpha)} \delta\partial_\mu \phi^\alpha. \tag{13.4.2.3}$$

Note that

$$\partial_\mu\left(\delta\phi^\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right)=(\partial_\mu\delta\phi^\alpha)\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}+\delta\phi^\alpha\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}. \tag{13.4.2.4}$$

Since $\delta\phi^\alpha$ which vanishes at ∂M, by Theorem 13.4.1.1, we have

$$\int_M\partial_\mu\left(\delta\phi^\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right)=\int_{\partial M}n_\mu\delta\phi^\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}=0. \tag{13.4.2.5}$$

Thus by (13.4.2.4), we have

$$\int_M(\partial_\mu\delta\phi^\alpha)\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}=-\int_M\delta\phi^\alpha\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}. \tag{13.4.2.6}$$

Since $\partial_\mu\delta\phi^a=\delta\partial_\mu\phi^a$ and by (13.4.2.3) and (13.4.2.6), we get

$$0=\int_M\left(\frac{\partial\mathcal{L}}{\partial\phi^\alpha}-\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}\right)\delta\phi^\alpha. \tag{13.4.2.7}$$

Now, since $\delta\phi^\alpha$ is arbitrary, we get

$$\boxed{\frac{\partial\mathcal{L}}{\partial\phi^\alpha}-\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^\alpha)}=0} \tag{13.4.2.8}$$

13.5 Saturday, 10 November 2018

Today, I am investigating various variational principles in special relativity classic field theory following [? , 17 The Classical Theory of Fields]

13.5.1 Classical Field Theory – Free particle with no field

Let x^μ will be a curve which connects points in a spacetime a and b. Assume that b is "accessible" from a and that we can paremetrise the curve by proper time τ. Define an action along the curve as

$$S=-m\int_a^b d\tau. \tag{13.5.1.1}$$

We are going to show that the above action is extremised by stright line in a spacetime. We will consider an arbitrary variation δx^μ which vanishes at a and b.

$$\delta S=-m\delta\int_a^b d\tau=m\delta\int_a^b\sqrt{g_{\mu\nu}dx^\mu dx^\nu}. \tag{13.5.1.2}$$

Note that

$$\delta F(f \cdot g) \; = \; F(f \cdot g) - F(f \cdot g + \delta(f \cdot g)) \; = \; (\delta f \cdot g + f \cdot \delta g)\dot{F}(f \cdot g). \tag{13.5.1.3}$$

Thus

$$\begin{aligned} \delta S &= -\frac{1}{2}m \int_a^b \frac{1}{\sqrt{g_{\mu\nu} dx^\mu dx^\nu}} g_{\mu\nu}(\delta dx^\mu dx^\nu + dx^\mu \delta dx^\nu) \\ &= -m \int_a^b \frac{dx_\mu}{d\tau} \delta dx^\mu = -m \int_a^b u_\mu \delta dx^\mu = -m \int_a^b u_\mu d\delta x^\mu. \end{aligned} \tag{13.5.1.4}$$

Note that

$$d(u_\mu \delta x^\mu) = du_\mu \delta x^\mu + u_\mu d\delta x^\mu. \tag{13.5.1.5}$$

Since δx^μ vanishes at a and b, we have

$$\int_a^b du_\mu \delta x^\mu = -\int_a^b u_\mu d\delta x^\mu. \tag{13.5.1.6}$$

And thus

$$\delta S = m \int_a^b du_\mu \delta x^\mu = m \int_a^b \frac{du_\mu}{d\tau} \delta x^\mu d\tau. \tag{13.5.1.7}$$

Since we require $\delta S = 0$ for arbitrary δx^μ, we have

$$\boxed{\frac{du_\mu}{d\tau} = 0} \tag{13.5.1.8}$$

13.5.2 Classical Field Theory – Particle in a field with a vector potential

Assume now that we have a vector potential A_μ. We are now in the same context as in 13.5.1. The only difference is that we need to update action by the field-interacts-with-particle component. Let's define the action as

$$S = \int_a^b -m d\tau + q A_\mu dx^\mu. \tag{13.5.2.1}$$

By what was shown to get (13.5.1.7) we have

$$\delta S = \int_a^b m du_\mu \delta x^\mu + q(\delta A_\mu dx^\mu + A_\mu \delta dx^\mu). \tag{13.5.2.2}$$

Note that

$$d(A_\mu \delta x^\mu) = dA_\mu \delta x^\mu + A_\mu d\delta x^\mu. \tag{13.5.2.3}$$

Since δx^μ vanishes in a and b.

$$\int_a^b A_\mu d\delta A_\mu = -\int_a^b dA_\mu \delta x^\mu. \tag{13.5.2.4}$$

Going back to 13.5.2.2, we get

$$\delta S = \int_a^b m du_\mu \delta x^\mu + q(\delta A_\mu dx^\mu - dA_\mu \delta x^\mu). \tag{13.5.2.5}$$

Let's calculate δA_μ and dA_μ.

$$\delta A_\mu = \frac{\partial A_\mu}{\partial x^\nu}\delta x^\nu \tag{13.5.2.6}$$

and

$$dA_\mu = \frac{\partial A_\mu}{\partial x^\nu} dx^\nu. \tag{13.5.2.7}$$

Substitute the above in 13.5.2.5

$$\delta S = \int_a^b m du_\mu \delta x^\mu + q(\frac{\partial A_\mu}{\partial x^\nu}\delta x^\nu dx^\mu - \frac{\partial A_\mu}{\partial x^\nu} dx^\nu \delta x^\mu). \tag{13.5.2.8}$$

As summations in each part of the above equation is independent, we can exchange indecies in the 2nd part, i. e.

$$\delta S = \int_a^b m du_\mu \delta x^\mu + q(\frac{\partial A_\nu}{\partial x^\mu}\delta x^\mu dx^\nu - \frac{\partial A_\mu}{\partial x^\nu} dx^\nu \delta x^\mu). \tag{13.5.2.9}$$

Now

$$\delta S = \int_a^b \left(m \frac{du_\mu}{d\tau} + q(\partial_\mu A_\nu - \partial_\nu A_\mu) u^\nu \right) \delta x^\mu d\tau. \tag{13.5.2.10}$$

Since δx^μ is arbitrary and we require $\delta S = 0$, we get

$$\boxed{m \frac{du_\mu}{d\tau} = q(\partial_\nu A_\mu - \partial_\mu A_\nu) u^\nu} \tag{13.5.2.11}$$

Which is an equation of motion of a particle with charge q and mass m in a field with a vector potential A_μ. Let

$$F_{\nu\mu} := \partial_\nu A_\mu - \partial_\mu A_\nu. \tag{13.5.2.12}$$

Note that from definition F is antisymmetric, i.e. $F_{\mu\mu} = 0$ and $F_{\nu\mu} = -F_{\mu\nu}$. The equation (13.5.2.11) has now a form

$$m\frac{du_\mu}{d\tau} = qF_{\nu\mu}u^\nu. \tag{13.5.2.13}$$

Let's consider a particle $t \mapsto (\phi^0(t) := t, \phi^1(t), \phi^2(t), \phi^3(t))$ in a ceratain frame of refernce x^μ. Define

$$w^\mu := \frac{d\phi^\mu}{dt}. \tag{13.5.2.14}$$

Note that $\vec{\omega} := (\omega^1, \omega^2, \omega^3)$ is an ordinary 3-dimentional veliocity of the particle in the frame of referebnce x^μ. Now assuming that covariant 4-velocity of the particle is u_μ and recalling that $\frac{d\tau}{dt} = u_0^{-1}$, we may write an equation (13.5.2.13) is a form

$$m\frac{du_\mu}{dt} = qF_{\nu\mu}u^\nu u_0^{-1} = qF_{\nu\mu}\omega^\nu. \tag{13.5.2.15}$$

Which translates into

$$\frac{dp_1}{dt} = qF_{01} + q(\omega^2 F_{21} - \omega^3 F_{13}). \tag{13.5.2.16}$$

$$\frac{dp_2}{dt} = qF_{02} + q(\omega^3 F_{32} - \omega^1 F_{21}). \tag{13.5.2.17}$$

$$\frac{dp_3}{dt} = qF_{03} + q(\omega^1 F_{13} - \omega^2 F_{32}). \tag{13.5.2.18}$$

Now, if we put

$$\vec{E} := (F_{01}, F_{02}, F_{03}) \tag{13.5.2.19}$$

and

$$\vec{B} := (F_{32}, F_{13}, F_{21}). \tag{13.5.2.20}$$

we get

$$\frac{d\vec{p}}{dt} = q\vec{E} + q\vec{\omega} \times \vec{B}. \tag{13.5.2.21}$$

As a reference

$$F_{\nu\mu} = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{bmatrix}. \tag{13.5.2.22}$$

13.6 Saturday, 17 November 2018

13.6.1 The Millikan Oil Drop Experiment

Studing *The Millikan Oil Drop Experiment* from [?] and wrote a bit of python code for it

Listing 13.1: How to calculate elementary charge from experiment data

```
import numpy as np

mu = 1.849e-5
d = 7.63e-4
rho = 1.184
sigma = 883
rho_prim = rho - sigma
g = 9.802 #New York Value
b = 6.17e-6
P = 76.01
V = 500
s = 4.71e-3

def _mu(a0, mu0=mu, P=P):
    return mu0 * ((1 + b/(a0*P))**-1)

def _a0(B, mu=mu):
    return ((9*B*mu*d)/(2*rho_prim*g))**0.5

def _e(A, a, mu):
    return (6*s*np.pi*a*mu*d*A)/V

B1 = [-27.9, -29.6, -28.2, -29.3, -29.4]

a = list(map(lambda B: _a0(1/B, _mu(_a0(1/B))), B1))
better_a = np.array(a).mean()
better_mu = _mu(better_a, mu)
```

13.7 Saturday, 8 December 2018

13.7.1 Classical picture for studing Einstein-de Haas effect.

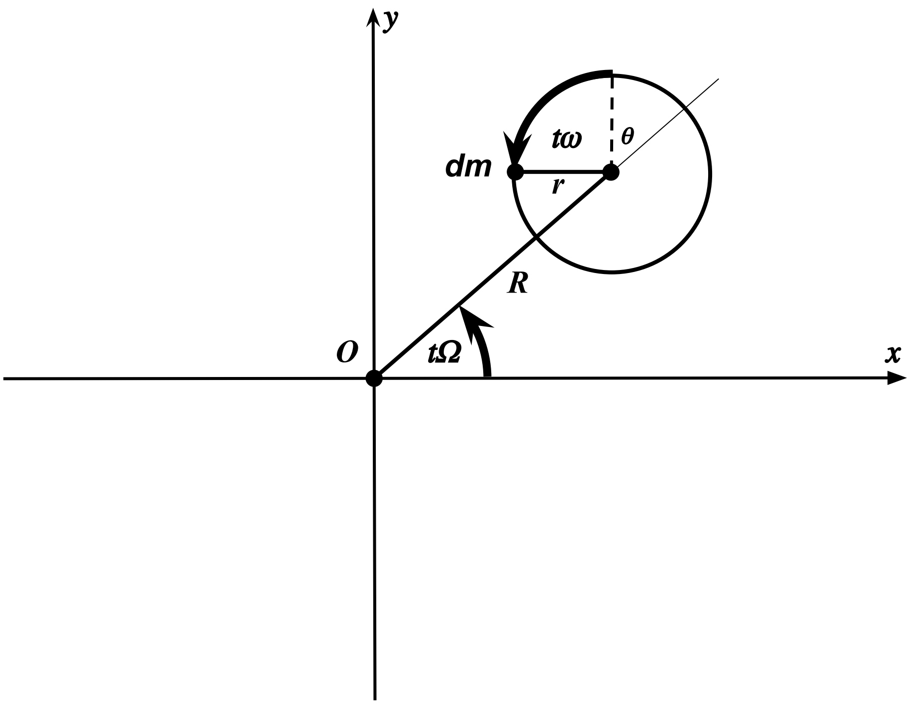

Figure 13.2: A circular hoop rotating around its center, which itself rotate around the central point O.

Our goal will be to calculate an angular momentum $\vec{L}$ (relative to the central point O) of a system consisting of a circular hoop rotating around its center, which itself also rotates around the central point O.

Assume that the circular hoop (lying is a plane xy) has a mass m and radious r, it's rotating around its center (axis of rotation perpendicular to the plane xy), which is fixed at one end of an arm of length R (lying in the plane xy) which again rotates with angular velocity Ω around its other end O in a plane of the hoop. Moreover assume that circular hoop rotates with an angular velocity ω relative to the arm. The situation is symbolically pictured on Figure 13.2).

Assume that at the moment $t = 0$ the center of the hoop was in the point $(R, 0)$. Consider an infinitesimal element of the hoop dm which at the moment $t = 0$ was at an angle θ to the extention of the arm (marked on Figure 13.2). We may assume that $dm = \rho d\theta$ where ρ is a constant angular density of the hoop. Let $\vec{q_\theta}(t)$ dentote position of the infinitesimal element of the hoop in time t.

$$\vec{q_\theta}(t) = (Rcos(t\Omega) + rcos(t\omega + \theta + t\Omega), Rsin(t\Omega) + rsin(t\omega + \theta + t\Omega), 0) \tag{13.7.1.1}$$

The angular momentum $d\vec{L}$ of the infinitesimal element of the hoop is equal to:

$$d\vec{L} = \vec{q_\theta} \times dm\dot{\vec{q_\theta}}. \tag{13.7.1.2}$$

After simplifications done by Mathematica 11.3 ([**?**]).

α[t_]:=$\Omega * t$

β[t_]:=$\omega * t$

θ[t_]:=$\alpha[t] - \beta[t]$

$x[\theta$_, t_]:={$R *$ Cos[$\alpha[t]$] + $r *$ Cos[$\alpha[t] + \beta[t] + \theta$],

$R *$ Sin[$\alpha[t]$] + $r *$ Sin[$\alpha[t] + \beta[t] + \theta$], 0}

xdot[θ_, t_]:=$D[x[\theta, t], t]$

TrigReduce[$x[\theta, t] \times$ xdot[θ, t]]

$\{0, 0, r^2\omega + r^2\Omega + R^2\Omega + rR\omega\mathrm{Cos}[\theta + t\omega] + 2rR\Omega\mathrm{Cos}[\theta + t\omega]\}$

$$d\vec{L} = \rho d\theta\big(0, 0, r^2(\omega + \Omega) + R^2\Omega + (rR\omega + 2rR\Omega)cos(\theta + t\omega)\big). \tag{13.7.1.3}$$

Now

$$\vec{L} = \int_0^{2\pi} d\vec{L} = \big(0, 0, mr^2(\omega + \Omega) + mR^2\Omega\big). \tag{13.7.1.4}$$

Let ω' be an angular velocity of the hoop relative to the plane xy, then $\omega' = \omega + \Omega$. Asumme that $\vec{L} = (L_x, L_y, L_z)$. Thus $L_x = L_y = 0$ and

$$L_z = mr^2\omega' + mR^2\Omega. \tag{13.7.1.5}$$

Assume now that we have a crystal structure in a shape of a cylindrical bar located perpendicularly to the plane xy with its hight parallel to $z - axis$. The bar rotates with an angular velocity Ω around its symmetry axis which lies along $z - axis$. Assume that the crystal structure has N nodes indexed from 1 to N, where Δm is a mass of each node. Assume that each node is a small hoop (lying in the plane parallel to xy plane) of some small radius r, rotating around its center with an angular velocity ω_i' (relative to plane xy) in a plane parallel to xy plane (axis of rotation goes through the center of the hoop and is perpendicular to the plane xy). From this point $\vec{L} = (L_x, L_y, L_z)$ will denote a total angular momentum of the crystal structure (relative to $z - axis$). Let $\vec{L_i} = (0, 0, \Delta m r^2 \omega_i')$ be an "intristic" angular momentum of $i - th$ node (as if ignoring the movement of its center due to the rotation of the bar). Let R_i be a distance of $i - th$ node from the bar's rotation axis.

By (13.7.1.5) we have that total angular momentum of the bar is

$$L_z = \sum_{i=1}^{N} (L_i)_z + \sum_{i=1}^{N} \Delta m R_i^2 \Omega = \sum_{i=1}^{N} (L_i)_z + \Omega \sum_{i=1}^{N} \Delta m R_i^2. \qquad (13.7.1.6)$$

Thus

$$\vec{L} = \sum_{i=1}^{N} \vec{L}_i + \Omega \vec{I}, \qquad (13.7.1.7)$$

where $\vec{I}$ is a moment of inertia of the solid cylinder of the shape of our cylindrical bar and exactly the same mass.

13.8 Thursday, 20 December 2018

13.8.1 Lifetime of excited states in Hydrogen

Reviewing Schrödinger model of hydrogen atom, I was also considering theoretical ways to calculate lifetimes of excited states. I have found an interesting paper [**?**] from 1982 which summarises theoretical and experimental results up to date. It is a valuable source of references. Also key word here is: Fermi's golden rule.

13.9 Sunday, 27 January 2019

13.9.1 Momentum conservation in the quantum two-body problem

In this subsection, we will work with Planck units. Consider Schrödinger equation of two particles, where interaction depends only on the distance between them.

$$i \frac{\partial \psi}{\partial t} = H\psi, \qquad (13.9.1.1)$$

where

$$H\psi = \frac{P_x^2}{2m_x}\psi + \frac{P_y^2}{2m_y}\psi + V(x-y)\psi. \tag{13.9.1.2}$$

and where ψ depends on (t, x, y) and $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3)$.

$$P_x^2 := P_{x,1}^2 + P_{x,2}^2 + P_{x,3}^2, \tag{13.9.1.3}$$

where

$$P_{x,i} := -i\frac{\partial}{\partial x_i}, \tag{13.9.1.4}$$

and

$$P_y^2 := P_{y,1}^2 + P_{y,2}^2 + P_{y,3}^2, \tag{13.9.1.5}$$

where

$$P_{y,i} := -i\frac{\partial}{\partial y_i}. \tag{13.9.1.6}$$

We are going to prove that

$$[P_{x,i} + P_{y,i}, H] = 0. \tag{13.9.1.7}$$

Let's calculate

$$\begin{aligned}(P_{x,i} + P_{y,i})(V(x-y)\psi) = \\ -i\dot{V}(x-y)\psi + V(x-y)P_{x,i}\psi + (-i)(-\dot{V}(x-y))\psi + V(x-y)P_{y,i}\psi \\ = V(x-y)(P_{x,i} + P_{y,i})\psi.\end{aligned} \tag{13.9.1.8}$$

On the other hand it's obvious that

$$[P_{x,i} + P_{y,i}, \frac{P_x^2}{2m_x} + \frac{P_y^2}{2m_y}] = 0. \tag{13.9.1.9}$$

Thus by (13.9.1.8) and (13.9.1.9), we have (13.9.1.7).

Found an interesting paper on two-body problem in quantum mechanics [?].

13.10 Saturday, 23 February 2019

13.10.1 Reduced mass in Quantum Mechanics

In this subsection we will continue to work with Hamiltonian introduced in 13.9.1.2 and with annotations introduced in Subsection 13.9.1. Let's introduce new system of coordinates

$$\begin{cases} R := \frac{m_x}{m_x + m_y} x + \frac{m_y}{m_x + m_y} y, \\ r := x - y. \end{cases} \tag{13.10.1.1}$$

We can now express ψ in t, r, R. R is a position of mass centre and r is a vector between two positions x, y. Check that

$$\begin{cases} x = R + \frac{m_y}{m_x + m_y} r, \\ y = R - \frac{m_x}{m_x + m_y} r, \end{cases} \tag{13.10.1.2}$$

Define operators

$$P_{r,i} := -i \frac{\partial}{\partial r_i}. \tag{13.10.1.3}$$

and

$$P_{R,i} := -i \frac{\partial}{\partial R_i}. \tag{13.10.1.4}$$

Note that

$$\frac{\partial \psi}{\partial r_i} = \frac{\partial \psi}{\partial x} \frac{\partial x}{\partial r_i} + \frac{\partial \psi}{\partial y} \frac{\partial y}{\partial r_i} = \frac{m_y}{m_x + m_y} \frac{\partial \psi}{\partial x_i} - \frac{m_x}{m_x + m_y} \frac{\partial \psi}{\partial y_i}. \tag{13.10.1.5}$$

Thus

$$P_{r,i} = \frac{m_y}{m_x + m_y} P_{x,i} - \frac{m_x}{m_x + m_y} P_{y,i}. \tag{13.10.1.6}$$

On the other hand

$$P_{R,i} = P_{x,i} + P_{y,i}. \tag{13.10.1.7}$$

Let

$$M = m_x + m_y, \tag{13.10.1.8}$$

which is total mass of the system and

$$\mu = \frac{m_x m_y}{m_x + m_y}, \tag{13.10.1.9}$$

which is reduced mass. Next, note that

$$\frac{P_{r,i}^2}{2\mu} = \frac{m_y}{2m_x(m_x+m_y)}P_{x,i}^2 - \frac{1}{m_x+m_y}P_{x,i}P_{y,i} + \frac{m_x}{2m_y(m_x+m_y)}P_{y,i}^2. \tag{13.10.1.10}$$

On the other hand

$$\frac{P_{R,i}^2}{2M} = \frac{1}{2(m_x+m_y)}(P_{x,i}^2 + 2P_{x,i}P_{y,i} + P_{y,i}^2). \tag{13.10.1.11}$$

Note that

$$\frac{m_y}{2m_x(m_x+m_y)} + \frac{1}{2(m_x+m_y)} = \frac{1}{2m_x}, \tag{13.10.1.12}$$

and

$$\frac{m_x}{2m_x(m_x+m_y)} + \frac{1}{2(m_x+m_y)} = \frac{1}{2m_y}. \tag{13.10.1.13}$$

Thus, eventually we get

$$\frac{P_{r,i}^2}{2\mu} + \frac{P_{R,i}^2}{2M} = \frac{P_x^2}{2m_x} + \frac{P_y^2}{2m_y}. \tag{13.10.1.14}$$

Thus if we introduce two hamiltonians:

$$H_r\psi = \frac{P_{r,i}^2}{2\mu}\psi + V(r)\psi, \tag{13.10.1.15}$$

and

$$H_R\psi = \frac{P_{R,i}^2}{2M}, \tag{13.10.1.16}$$

we will get

$$H = H_r + H_R. \tag{13.10.1.17}$$

It is crucial to note at this point that if we look at ψ as dependent on t, r, R, H_r acts seperately on coordinate r and H_R acts seperately on coordinate R and that $P_{r,i}$ and $P_{R,i}$ are respectively momentum operators.

Thus we might find solution of stationary equation $H\psi = E\psi$ by $\psi(r,R) = \phi_1(r)\phi_2(R)$, where $H_r\phi_1 = E_r\phi_1$ and $H_R\phi_2 = E_R\phi_2$.

13.11 Thursday, 16 May 2019

13.11.1 The electromagnetic field as an infinite system of harmonic oscilators

The idea of continuum limit of harmonic oscilators is mentioned in [**?** , 5.3]. It is described in details in [see **?** , 6 Quantization of the Electromagnetic Field].

Non-relativistic probability current in Quantum Mechanics, mentioned in [**?** , 6.2 Probability currents and densities] is introduced e.g. in [**?** , 3.6 Probability conservation].

13.12 Sunday, 26 May 2019

13.12.1 Derivation of the Schrödinger equation from the Ehrenfest theorems

Through wikipedia's entry on Ehrenfest theorem [see e.g. **?** , 3.7], I have found a paper [**?**] in which authors shows how to derive Schrödinger equation from the thesis of Ehrenfest theorem. The key point is the use of canonical comutation between momentum and position operators. It is showed as well that if the operators commute, we get Koopman–von Neumann classical mechanics. Recently they managed to generalise their result to Dirac Equation in [**?**].

13.12.2 Ehrenfest theorem

Assume that $\psi(t)$ is a solution of Schrödinger equation

$$i\hbar\frac{d\psi}{dt}(t) = H\psi(t). \tag{13.12.2.1}$$

Consider observalble A which is composed of momentum and position operators.

$$i\hbar\frac{d}{dt}(\psi(t), A\psi(t)) = (i\hbar\frac{d\psi}{dt}(t), A\psi(t)) - (\psi(t), i\hbar\frac{d}{dt}A\psi(t)) \tag{13.12.2.2}$$

As A is composed of momentum and position operators, it is clear that it comutes with $\frac{d}{dt}$, thus we may continue the equation 13.12.2.2 as follows:

$$\begin{gathered}(i\hbar\frac{d\psi}{dt}(t), A\psi(t)) - (\psi(t), Ai\hbar\frac{d}{dt}\psi(t)) = \\ (H\psi(t), A\psi(t)) - (\psi(t), AH\psi(t)) = \\ (\psi(t), HA\psi(t)) - (\psi(t), AH\psi(t)) = \\ (\psi(t), (HA - AH)\psi(t)) = (\psi(t), [H, A]\psi(t)).\end{gathered} \tag{13.12.2.3}$$

In bra-ket notation:

$$i\hbar\frac{d}{dt}\langle\psi(t)|A|\psi(t)\rangle = \langle\psi(t)|[A, H]|\psi(t)\rangle\,. \tag{13.12.2.4}$$

Assume that we have hamiltonian

$$H = \sum_{i=0}^{3}\frac{P_i^2}{2m} + V(Q_1, Q_2, Q_3). \tag{13.12.2.5}$$

Note that

$$[Q_i, H] = \frac{1}{2m}[P_i^2, Q_i] = \frac{i\hbar P_i}{m}. \tag{13.12.2.6}$$

Thus by 13.12.2.4

$$\frac{d}{dt}\langle\psi(t)|Q_i|\psi(t)\rangle = \frac{1}{m}\langle\psi(t)|P_i|\psi(t)\rangle\,. \tag{13.12.2.7}$$

Note that

$$[P_i, H] = [P_i, V] = -i\hbar\frac{\partial V}{\partial x_i}. \tag{13.12.2.8}$$

Thus by 13.12.2.4

$$\frac{d}{dt}\langle\psi(t)|P_i|\psi(t)\rangle = \langle\psi(t)|-\frac{\partial V}{\partial x_i}|\psi(t)\rangle\,. \tag{13.12.2.9}$$

13.13 Monday, 27 May 2019

13.13.1 Hamiltonian of electromagnetic field

The hamiltonian is itroduced without proof in [**?** , 6.2]. The deriviation can be found in [**?** , Examples 9.1 and 9.2].

13.14 Tuesday, 28 May 2019

13.14.1 Interaction with the orbital angular momentum

In [? , 6.5], we need to assume $\vec{B} = \text{const}$, to have $A = \frac{1}{2}(\vec{B} \times \vec{r})$ and $\nabla \times A = B$.

13.15 Tuesday, 2 July 2019

13.15.1 Dirac Delta Function in the Context of Fourier Transform in Physical Texts

We will calculate Fourier Transform (as defined in Definition 11.4.0.5) of 1 understood as distribution.

$$\mathcal{F}(1)(\phi) = 1(\mathcal{F}(\phi)) = \int_{\mathbb{R}^n} \mathcal{F}(\phi) \cdot 1 = \int_{\mathbb{R}^n} \mathcal{F}(\phi) \cdot e^{ix\cdot 0} dx = (2\pi)^{\frac{n}{2}} \mathcal{F}^{-1}(\mathcal{F}(\phi))(0) = (2\pi)^{\frac{n}{2}} \phi(0). \tag{13.15.1.1}$$

Thus from the definition of Dirac delta-function understood as distribution

$$\boxed{\mathcal{F}(1) = (2\pi)^{\frac{n}{2}} \delta_0.} \tag{13.15.1.2}$$

Equivalently

$$\delta_0 = (2\pi)^{-\frac{n}{2}} \mathcal{F}(1). \tag{13.15.1.3}$$

Let's calculate now Fourier Transform of Dirac delta:

$$\mathcal{F}(\delta_0)(\phi) = \delta_0(\mathcal{F}(\phi)) = \mathcal{F}(\phi)(0) = \int_{\mathbb{R}^n} \phi(x) e^{-ix\cdot 0} dx = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \phi(x) \cdot 1 dx = (2\pi)^{-\frac{n}{2}} 1(\phi). \tag{13.15.1.4}$$

Thus

$$\boxed{\mathcal{F}(\delta_0) = (2\pi)^{-\frac{n}{2}} \cdot 1.} \tag{13.15.1.5}$$

If we apply inverse Fourier transform to the equation (13.15.1.5), we obtain

$$\delta_0 = (2\pi)^{-\frac{n}{2}} \mathcal{F}^{-1}(1). \tag{13.15.1.6}$$

From equations (13.15.1.2) and (13.15.1.6) it is apparent that

$$\mathcal{F}(1) = \mathcal{F}^{-1}(1) \tag{13.15.1.7}$$

and

$$\mathcal{F}(\delta_0) = \mathcal{F}^{-1}(\delta_0), \tag{13.15.1.8}$$

which is intuitive as both distributions are symmetric. In physics texts (e.g. [?]), equations such as the equation (13.15.1.6) are usually written as

$$\delta(k) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ikx} dx, \tag{13.15.1.9}$$

despite the fact that the integral $\int_{\mathbb{R}^n} e^{ikx} dx$ doesn't exist in classical measure theory sense. If one knows distribution theory, it is easy to understand that the meaning of this is symbolic and has certain algebraic sense, however not experienced reader might encounter great dificulties, when one sees such an intergral for the first time without necessery commentary.

For distribution theory see e.g [? , Distributions and Fourier Transforms] or [? , 3.3 Distributions].

13.16 Saturday, 3 August 2019

13.16.1 Futher investigation in mathematical rigor in Dirac notation

My problem is that physists use equations like

$$\langle \psi_\alpha | \psi_\beta \rangle = \delta(\alpha - \beta), \tag{13.16.1.1}$$

even where $\int \psi_\alpha^* \psi_\beta$ isn't defined properly. As I understand, this requires to see at least the mapping $\alpha \mapsto \langle \psi_\alpha | \psi_\beta \rangle$ not as a function but distribution. Thus definition of $\langle \psi_\alpha | \psi_\beta \rangle$ can't be any longer $\int \psi_\alpha^* \psi_\beta$ in terms of integral over complex functions.

As I understand the proper mathematical rigor is achived through rigged Hilbert spaces. The nice summary of mathematics basis for Dirac formalism is here [?] also a good strating point is [?].

I tried to make a shortcut, because in lucky case we can easily understand $\alpha \mapsto \langle \psi_\alpha | \psi_\beta \rangle$ as distribution for $\psi_\alpha(x) = e^{-i\alpha \cdot x}$ as Fourier transform extended on the space of tempered distribution, I was thinking about the following generalisation:

Theorem 13.16.1.1. *Let $\{u_\alpha\}_{\alpha\in\mathbb{R}^n} \subset S'_n$. If a transformation $\mathcal{T}_u$ defined as*

$$(\mathcal{T}_u(\phi))(\alpha) = u_\alpha(\phi) \text{ for } \phi \in S_n \text{ and } \alpha \in \mathbb{R}^n \tag{13.16.1.2}$$

is a continuous mapping $\mathcal{T}_u : S_n \to S_n$, then the transformation $\hat{\mathcal{T}}_u$ defined as

$$\hat{\mathcal{T}}_u(v)(\phi) = v(\mathcal{T}_u(\phi)) \text{ for } v \in S'_n \text{ and } \phi \in S_n, \tag{13.16.1.3}$$

is a continuous mapping $\hat{\mathcal{T}}_u : S'_n \to S'_n$ in S'_n with weak topology.*

Proof. This version or very similiar should be proved. Not done yet. □

However, even if above theorem is true, it doesn't solve the main issue because opposite to the special case of Fourier transform, transformation $\hat{\mathcal{T}}_u$ in general case is not an extension of $\mathcal{T}_u$ in sense of $S_n \subset S'_n$ (when we consider functions from S_n as distributions).

Now I will try to study if I can't achive something similiar using integral in space of measurable functions into the space of distribuitions. But most likely this aproach will be not successul as Schwartz showed that multiplication of distributions is not associative ([see **?** , XIX. Theory of Distributions 7. Operations of distributions])

13.16.2 Plan for rigorous theory which includes Dirac notation

We will need a notion of integral of distribution $u \in \mathcal{D}'(\mathbb{R}^n)$

$$\int_A u d\mu \tag{13.16.2.1}$$

defined as some kind of limit $\lim_{n\to\infty} u(\psi_n)$ where $D(\mathbb{R}^n) \ni \psi_n \to 1_A$ in certain sense. With this kind of definition we will have

$$\int_{\mathbb{R}^n} \delta_z dx = 1, \tag{13.16.2.2}$$

for any $z \in \mathbb{R}^n$, where δ_z is Dirac delta in point z. We will also need to define the integral of functions with values in the space of distributions. Something like that:

Definition 13.16.2.1. *Let Ω be an open subset of $\mathbb{R}^k$ and let μ be any Borel measure on Ω and let A be any borel set. Let $u : A \to \mathcal{D}'(\mathbb{R}^k)$ be a function* *for which a function $A \ni \alpha \mapsto u_\alpha(\psi)$ is measurable for any $\phi \in \mathcal{D}(\Omega)$. We will call such u measurable. We define that*

$$\int_A u_\alpha d\mu(\alpha) := v \tag{13.16.2.3}$$

iff $v \in \mathcal{D}'(\Omega)$ and

$$v(\phi) = \int_A u_\alpha(\phi) d\mu(\alpha) \text{ for any } \phi \in \mathcal{D}(\Omega). \tag{13.16.2.4}$$

We say that the integral from (13.16.2.3) exists if there exists such a v.

Note that with this setup for any measurable function $f : A \to \mathbb{C}$ the integral $\int_A f(\alpha)u_\alpha d\mu(\alpha)$ has sense, as we treat $f(\alpha)$ as scalar value by which we multiply distribution u_α.

When in our considerations we treat distributions $\mathcal{D}'(\Omega)$ as vectors, the equivalent of matrix are distributions $\mathcal{D}'(\Omega \times \Omega)$. In that sense we might want to concatenate values of $u : \Omega \to \mathcal{D}'(\Omega)$ into matrix $U \in \mathcal{D}'(\Omega \times \Omega)$. In order to do that we introduce operation $(.)_{\alpha\in\Omega}$.

Definition 13.16.2.2. ***(Generalised matrix)*** *Let Ω be an open subset of $\mathbb{R}^k$ and $u : \Omega \to \mathcal{D}'(\Omega)$ be Borel measurable. We define*

$$(u_\alpha)_{\alpha\in\Omega} := U \tag{13.16.2.5}$$

iff $U \in \mathcal{D}'(\Omega \times \Omega)$ and

$$U(\phi) = \int_\Omega u_\alpha(\phi(\alpha, \cdot)) d\alpha, \tag{13.16.2.6}$$

for any $\phi \in \mathcal{D}(\Omega \times \Omega)$. We say that $(u_\alpha)_{\alpha\in\Omega}$ exists if there is such U.

For any function $f : A \times B \to C$ defined on the product of sets, we can define transposition as $f^T : B \times A \to C$ such that $f^T(b, a) = f(a, b)$ for any $a \in A$ and $b \in B$. Note that for any $\phi \in \mathcal{D}(\Omega \times \Omega)$, we have $\phi^T \in \mathcal{D}(\Omega \times \Omega)$.

Definition 13.16.2.3. ***(Generilised transposition)*** *Let Ω be an open subset of $\mathbb{R}^k$ and let $U \in \mathcal{D}'(\Omega \times \Omega)$. We define a transposition of U as $U^T \in \mathcal{D}'(\Omega \times \Omega)$ such that*

$$U^T(\phi) := U(\phi^T) \text{ for any } \phi \in \mathcal{D}(\Omega \times \Omega). \tag{13.16.2.7}$$

In this terms for any measurable $\Omega \ni \alpha \mapsto u_\alpha \in \mathcal{D}'(\Omega)$ and $\psi \in C^\infty(\Omega)$ we can define Dirac bra-ket understood as a distribution with domain indicated by argument α

$$\langle \psi | u_\alpha \rangle := \int_\Omega \psi(\beta)^* u^\beta d\alpha. \tag{13.16.2.8}$$

where $\Omega \ni \beta \mapsto u^\beta \in \mathcal{D}'(\Omega)$ is a measurable function such that $(u^\beta)_{\beta\in\Omega} = \big((u_\alpha)_{\alpha\in\Omega}\big)^T$, provided that the integral exists. Next we can extend this as

$$\langle v | u_\alpha \rangle := \lim_{n\to\infty} \langle \psi_n | u_\alpha \rangle , \tag{13.16.2.9}$$

if the above limit (understood in $\mathcal{D}'(\Omega)$ with weak-* topology) exists and is the same for any $\mathcal{D}(\Omega)\psi_n \to v$ in $\mathcal{D}'(\Omega)$ with weak-* topology.
Note that if $\alpha \to \langle v|u_\alpha\rangle$ happen to be a function, each complex value $\langle v|u_\alpha\rangle$ has sense for any α. Thus this definition defines Dirac bra-kets as scalars whenever they have sense as scalars. It is possible that for this all to work additional condition that u_α is in certain sense orthogonal basis might be required.

There is a channce this construction will evolve into generalised vector spaces with vectors generelised to measures with notation $\int_A f(\alpha) u_\alpha d\mu(\alpha)$ which unifies finite and infinite dimension vector spaces.

When we will write $\langle v|u_\alpha\rangle$ in the context of distribution, it will always mean distribution with the domain indicated by the index in bra.

Remark 13.16.2.4. *If we are in the context of tempered distributions, we need to replace $\mathcal{D}(\Omega)$ by S_k, $\mathcal{D}'(\Omega)$ by S'_k, $\mathcal{D}(\Omega\times\Omega)$ by S_{2k} and $\mathcal{D}'(\Omega\times\Omega)$ by S'_{2k} in the above definitions.*

13.17 Sunday, 18 August 2019

13.17.1 Rigorous Dirac Formulation

We will continue considerations that we bagan in subsection 13.16.2. We will use dirac delta δ_α as defined in Definition 11.5.3.14. We will also define

Definition 13.17.1.1. *Let $\alpha \in \mathbb{R}^k$. Let's define $e_\alpha : \mathbb{R}^k \to \mathbb{C}$ as*

$$e_\alpha(x) = (2\pi)^{-\frac{k}{2}} e^{-i\alpha\cdot x}. \tag{13.17.1.1}$$

By Example 11.5.4.11 we can treat e_α as an element of S'_k.

Fact 13.17.1.2. *Let $\alpha \in \mathbb{R}^k$, then $\mathcal{F}(\delta_\alpha) = e_\alpha$.*

Proof.

$$\mathcal{F}(\delta_\alpha)(\phi) = \delta_\alpha(\mathcal{F}(\phi)) = \mathcal{F}(\phi)(\alpha) = \int e_\alpha \phi = e_\alpha(\phi) \tag{13.17.1.2}$$

for any $\phi \in S_n$. □

Fact 13.17.1.3. *Let $\alpha \in \mathbb{R}^k$, then $\mathcal{F}(e_\alpha) = \delta_{-\alpha}$.*

Proof.

$$\mathcal{F}(e_\alpha)(\phi) = e_\alpha(\mathcal{F}(\phi)) = \int e_\alpha \mathcal{F}(\phi) = \phi(-\alpha) = \delta_{-\alpha}(\phi) \tag{13.17.1.3}$$

for any $\phi \in S_n$. □

Proposition 13.17.1.4. $(\delta_\alpha)_{\alpha \in \mathbb{R}^k} = \left((\delta_\alpha)_{\alpha \in \mathbb{R}^k}\right)^T$.

Proof. Let $U = (\delta_\alpha)_{\alpha \in \mathbb{R}^k}$. Note that by Definition 13.19.1.1 $U(\phi) = \int \phi(\alpha, \alpha) d\alpha$. Thus $U^T = U$. □

Proposition 13.17.1.5. $(e_\alpha)_{\alpha \in \mathbb{R}^k} = \left((e_\alpha)_{\alpha \in \mathbb{R}^k}\right)^T$.

Proof. Let $U = (e_\alpha)_{\alpha \in \mathbb{R}^k}$. By Definition 13.19.1.1

$$U(\phi) = (2\pi)^{-\frac{k}{2}} \int \int e^{-i\alpha \cdot \beta} \phi(\alpha, \beta) d\alpha d\beta. \tag{13.17.1.4}$$

Thus $U^T = U$. □

Definition 13.17.1.6. *Let Ω be an open subset of $\mathbb{R}^k$. Define $(.)^* : D'(\Omega) \to D'(\Omega)$ as*

$$u^*(\phi) := (u(\phi^*))^*, \tag{13.17.1.5}$$

where $u \in D'(\Omega)$ and $\phi \in D(\Omega)$.

Corollary 13.17.1.7. *$(.)^* : D'(\Omega) \to D'(\Omega)$ is continuous in $\mathcal{D}'(\Omega)$ with weak-* topology.*

Corollary 13.17.1.8. *$\delta_\alpha^* = \delta_\alpha$ for any $\alpha \in \mathbb{R}^k$.*

Corollary 13.17.1.9. *$e_\alpha^* = e_{-\alpha}$ for any $\alpha \in \mathbb{R}^k$.*

Theorem 13.17.1.10. *If $u \in \mathcal{D}'(\mathbb{R}^k)$, then*

$$\langle u | \delta_\alpha \rangle = u^*. \tag{13.17.1.6}$$

Proof. Take any $D(\mathbb{R}^k) \ni f_n \to u$ in $\mathcal{D}'(\mathbb{R}^k)$ with weak-* topology. We should have

$$\langle f_n|\delta_\alpha\rangle = \int f_n^*(\alpha)\delta_\alpha d\alpha. \tag{13.17.1.7}$$

Take any $\phi \in \mathcal{D}'(\mathbb{R}^k)$.

$$\int f_n^*(\alpha)\delta_\alpha(\phi)d\alpha = \int f_n^*(\alpha)\phi(\alpha)d\alpha \to u^*(\phi). \tag{13.17.1.8}$$

Thus $\langle f_n|\delta_\alpha\rangle = f_n^* \to u^*$. □

The above proof could be easily modified to give the following:

Theorem 13.17.1.11. *(In the context of tempered distributions) If $u \in S_n'$, then*

$$\langle u|\delta_\alpha\rangle = u^*. \tag{13.17.1.9}$$

Corollary 13.17.1.12. $\langle \delta_\beta|\delta_\alpha\rangle = \delta_\beta$.

Which is in our convention exactly equivalent to what physicists mean by $\langle \delta_\beta|\delta_\alpha\rangle = \delta(\alpha-\beta)$.

Theorem 13.17.1.13. *(In the context of tempered distributions) If $u \in S_k'$, then*

$$\langle u|e_\alpha\rangle = \mathcal{F}^{-1}(u)^*. \tag{13.17.1.10}$$

Proof. Take any $S_k \ni f_n \to u$ in $\mathcal{D}'(\mathbb{R}^k)$ with weak-* topology. We should have

$$\langle f_n|e_\alpha\rangle = \int f_n^*(\alpha)e_\alpha d\alpha. \tag{13.17.1.11}$$

Take any $\phi \in S_k'$.

$$\int f_n^*(\alpha)e_\alpha(\phi)d\alpha = \int f_n^*(\alpha)\mathcal{F}(\phi)(\alpha)d\alpha = (\mathcal{F}^{-1}(f_n))^*(\phi) \to (\mathcal{F}^{-1}(u))^*(\phi). \tag{13.17.1.12}$$

Thus $\langle f_n|e_\alpha\rangle = (\mathcal{F}^{-1}(f_n))^* \to (\mathcal{F}^{-1}(u))^*$. □

Corollary 13.17.1.14. $\langle e_\beta|e_\alpha\rangle = \delta_\beta$.

Proof. $\langle e_\beta|e_\alpha\rangle = \mathcal{F}^{-1}(e_\beta)^* = \delta_\beta^* = \delta_\beta$. □

This is in our convention, again exactly equivalent to what physicists mean by $\langle e_\beta|e_\alpha\rangle = \delta(\alpha-\beta)$.

13.18 Monday, 19 August 2019

13.18.1 Rigorous Dirac Formulation - Matrix as Distribution

In Definition 13.19.1.1 we built an analogy between matrix and distribution. Let's now define matrix-vector multiplication.

Fact 13.18.1.1. *If $\phi \in S_{k_1}$ and $\psi \in S_{k_2}$, then*

$$\|\phi \otimes \psi\|_N^S \leq \|\phi\|_N^S \|\psi\|_N^S \tag{13.18.1.1}$$

Proof. We will split each multindex α into $\alpha = \alpha_1, \alpha_2$ where α_1 is multindex consists from first k_1 indecies and α_2 consits from k_2 subsequent indecies.

$$\begin{aligned} \|\phi \otimes \psi\|_N^S &= \sup_{(x,y)\in\mathbb{R}^{k_1}\times\mathbb{R}^{k_2}} \sup_{|\alpha_1,\alpha_2|\leq N} (1+|x|^2+|y|^2)^N |D^{\alpha_1}\phi(x)||D^{\alpha_2}\psi(y)| \\ &\leq \sup_{(x,y)\in\mathbb{R}^{k_1}\times\mathbb{R}^{k_2}} \sup_{|\alpha_1|\leq N, |\alpha_2|\leq N} (1+|x|^2)^N(1+|y|^2)^N |D^{\alpha_1}\phi(x)||D^{\alpha_2}\psi(y)| \\ &\leq \|\phi\|_N^S \|\psi\|_N^S. \end{aligned} \tag{13.18.1.2}$$

□

Corollary 13.18.1.2. *If $\phi \in S_{k_1}$ and $\psi \in S_{k_2}$, then $\phi \otimes \psi \in S^{k_1+k_2}$.*

Corollary 13.18.1.3. *Let $\phi \in S_{k_1}$, then the mapping $S_{k_2} \ni \psi \mapsto \phi \otimes \psi \in S^{k_1+k_2}$ is continuous.*

Proof. Take any $\psi_n \to 0$ in S_{k_2}. Thus $\|\psi_n\|_N^S \to 0$ for $N = 0, 1, \ldots$. Hence $\|\phi \otimes \psi_n\|_N^S \leq \|\phi\|_N^S \|\psi_n\|_N^S \to 0$ for $N = 0, 1, \ldots$. Therefore the mapping $\psi \mapsto \phi \otimes \psi$ is continous in $S_{k_1+k_2}$ □

Definition 13.18.1.4. *Let $U \in S'_{2k}$. For any $\phi \in S_k$, we will define $U \bullet \phi$ as linear functional on S_k such that*

$$(U \bullet \phi)(\psi) := U(\phi \otimes \psi) \tag{13.18.1.3}$$

for any $\psi \in S_k$.

Theorem 13.18.1.5. *Let $U \in S'_{2k}$. If $\phi \in S_k$, then $U \bullet \phi \in S'_k$.*

Proof. Since by Corollary 13.18.1.3 the mapping $\psi \mapsto \phi \otimes \psi$ is continous and U is a continuous linear functional on S_{2k}, $\psi \mapsto U(\phi \otimes \psi)$ is a continuous linear functional on S_k. Therefore, $U \bullet \phi \in S'_k$. □

Corollary 13.18.1.6. *If $U \in S'_{2k}$, then $S_k \ni \phi \mapsto U \bullet \phi \in S'_k$ is a linear mapping.*

We proved that $U \bullet S_k \subset S'_k$. We will consider situations in which the image of S_k are tempered distributions which are functions from S_k. We will write simply $U \bullet S_k \subset S_k$, treating S_k at the right side of inclusion as embedded in S'_k.

In the following theorems and proofs for any Λ_u for any $u \in S_k$ will denote the coresponding element of S'_k exactly as in Theorem 11.5.4.8.

Theorem 13.18.1.7. *Let $U \in S'_{2k}$. If $U \bullet S_k \subset S_k$, then the mapping $S_k \ni \phi \mapsto U \bullet \phi \in S_k$ is continuous.*

Proof. Since by Theorem 11.5.4.4 S_n is a Fréchet space and by this F-space, it is enough to prove that the graph of the linear mapping $S_k \ni \phi \mapsto U \bullet \phi \in S_k$ is closed and then by Theorem 11.5.2.14 (The closed graph theorem) it will be immediately proven that the mapping is continuous.
Take a sequence $\phi_n \to \phi$ in S_k such that $U \bullet \phi_n \to v$ in S_k (note that $U \bullet S_k \subset S_k$ in sense of embedding S_k in S'_k). Writing this in more rigorous way, we have a sequence $f_n \in S_n$ such that $f_n \to v \in S_k$ in topology S_k such that $\Lambda_{f_n} = U \bullet \phi_n$.
By Corollary 13.18.1.3 the mapping $S_k \ni \phi \mapsto \phi \otimes \psi \in S_{2k}$ is continous and U is a continuous linear functional on S_{2k}, thus $S_k \ni \phi \mapsto U(\phi \otimes \psi) = (U \bullet \phi)(\psi) \in \mathbb{C}$ is a continuous linear functional on S_k for any fixed $\psi \in S_k$. Therefore $(U \bullet \phi_n)(\psi) \to (U \bullet \phi)(\psi)$ for any fixed $\psi \in S_k$, which means that $U \bullet \phi_n \to U \bullet \phi$ in S'_k with weak-* topology. But by Theorem 11.5.4.8, $U \bullet \phi_n = \Lambda_{f_n} \to \Lambda_v$ in S'_k with weak-* topology. Since S'_k with weak-* topology it is a Hausdorff TVS and hence Hausdorff, we have $\Lambda_v = U \bullet \phi$. Thus in sense of embedding S_k in S'_k, we have $v = U \bullet \phi$. This establishes that the graph of the linear mapping $S_k \ni \phi \mapsto U \bullet \phi \in S_k$ is closed and by this completes the proof. □

Theorem 13.18.1.8. *If $U \in S'_{2k}$ such that $U^T \bullet S_k \subset S_k$, then*

$$(U \bullet \phi)(\psi) = \Lambda_\phi(U^T \bullet \psi) \tag{13.18.1.4}$$

for any $\phi \in S_k$ and $\psi \in S_k$.

Proof.

$$\begin{aligned}\Lambda_\phi(U^T \bullet \psi) = \int \phi(U^T \bullet \psi) = (U^T \bullet \psi)(\phi) \\ = U^T(\psi \otimes \phi) = U((\psi \otimes \phi)^T) = U(\phi \otimes \psi) = (U \bullet \phi)(\psi)\end{aligned} \tag{13.18.1.5}$$

□

In spite of the above theorem, we can extend $U \bullet (\cdot)$ to S'_k.

Definition 13.18.1.9. *Let $U \in S'_{2k}$ such that $U^T \bullet S_k \subset S_k$. For any $v \in S'_k$, we will define $U \bullet v$ as a linear functional on S_k such that*

$$(U \bullet v)(\psi) := v(U^T \bullet \psi) \tag{13.18.1.6}$$

for any $\psi \in S_k$.

Theorem 13.18.1.10. *Let $U \in S'_{2k}$ such that $U^T \bullet S_k \subset S_k$. If $v \in S'_k$, then $U \bullet v \in S'_k$.*

Proof. By Theorem 13.18.1.7, the mapping $S_k \ni \psi \mapsto U^T \bullet \psi \in S_k$ is continuous. Thus by Definition 13.18.1.9, $U \bullet v \in S'_k$. □

Theorem 13.18.1.11. *Let $U \in S'_{2k}$ such that $U^T \bullet S_k \subset S_k$. The linear mapping $S'_k \ni v \mapsto U \bullet v \in S'_k$ is continuous in S'_k with weak-* topology.*

Proof. Take any $v_n \to v$ in S'_k with weak-* topology. Obviously $v_n(U^T \bullet \psi) \to v(U^T \bullet \psi)$ for any fixed $\psi \in S_k$. Thus by Definition 13.18.1.9, the mapping $S'_k \ni v \mapsto U \bullet v \in S'_k$ is continuous in S'_k with weak-* topology. □

13.19 Friday, 23 August 2019

Question of the day: What is a dual to S'_k with strong topology? Is this possible that S_k? If so, can't we somehow deduce that for each linear continous $L : S_k \to S'_k$ (for starters in S'_k strong dual), $L^T : S'_k \to S_k$ and thus $L^T(S_k) \subset S_k$? But that would mean that convergence in S_k in strong dual S'_k topology implies convergence in S_k. Can this be true??

13.19.1 Rigorous Dirac Formulation - continuation (1)

I have noticed that we may apply Schwartz kernel theorem to improve the reasoning from subsection (13.18.1). As completly different thing, we will improve a bit general matrix definition. We inverted it compared with Definition 13.18.1.4.

Definition 13.19.1.1. *(**Generalised matrix**) Let $u : \mathbb{R}^k \to S'_k$ and μ be a complex or positive Borel measure on $\mathbb{R}^k$ be Borel measurable. We define*

$$(u_\alpha)_{\mu;\alpha\in\Omega} := U \tag{13.19.1.1}$$

iff $U \in \mathcal{S}'_{2k}$ and

$$U^T \bullet \phi = \int_{\mathbb{R}^k} \phi(\alpha) u_\alpha d\mu(\alpha), \tag{13.19.1.2}$$

for any $\phi \in S_k$. We say that $(u_\alpha)_{\mu;\alpha\in\mathbb{R}^k}$ exists if there is such U. If μ is a Lebesgue measure, we will omit it writing just $(u_\alpha)_{\alpha\in\Omega}$.

We will demonstrate an example to show why arbitrary Borel measure μ is needed. Let's first introduce certain usefull abreviation $\phi^y(x) := \phi(x, y)$.

Example 13.19.1.2. *Let $u : \mathbb{R}^k \to S'_k$. Let μ be concentrated on set $\{y_1, y_2\} \subset \mathbb{R}^k$ such that $\mu(y_1) = \mu(y_2) = 1$. There exists $U \in S'_{2k}$ such that*

$$U(\phi) = u_{y_1}(\phi^{y_1}) + u_{y_2}(\phi^{y_2}) = (u_\alpha)_{\mu;\alpha\in\mathbb{R}^k} \tag{13.19.1.3}$$

Theorem 13.19.1.3. *Let $U \in S'_{2k}$ and $u : \mathbb{R}^k \to S'_k$ be a family of tempered distributions such that $U = (u_\alpha)_{\alpha\in\mathbb{R}^k}$.*

$$U \bullet \phi \underset{a.e.}{=} [\mathbb{R}^k \ni \alpha \mapsto u_\alpha(\phi)] \text{ for any } \phi \in S_k. \tag{13.19.1.4}$$

Proof. Let's fix $\phi \in S_k$. By Definition 13.19.1.1 we have

$$(U \bullet \phi)(\psi) = (U^T \bullet \psi)(\phi) = \int \psi(\alpha) u_\alpha(\phi) d\alpha \tag{13.19.1.5}$$

for any $\psi \in S_k$. Thus thesis. □

Definition 13.19.1.4. *We will call a family $u : \mathbb{R}^k \to S'_k$ weakly continuous iff a function $\mathbb{R}^k \ni \alpha \mapsto u_\alpha(\phi)$ is continuous for every $\phi \in S'_k$.*

Theorem 13.19.1.5. *If $U \in S_{2k}$ and $U \bullet S_k \subset S_k$ and $u : \mathbb{R}^k \to S'_k$ is a weakly continuous family of tempered distributions such that $(u_\alpha)_{\alpha\in\mathbb{R}^k} = U$, then*

$$U^T \bullet \delta_\alpha = u_\alpha \text{ for every } \alpha \in \mathbb{R}^k. \tag{13.19.1.6}$$

Proof. Take any $\phi \in S_k$. We have

$$(U^T \bullet \delta_\alpha)(\phi) = \delta_\alpha(U \bullet \phi) = u_\alpha(\phi). \tag{13.19.1.7}$$

The last equality is by Theorem 13.19.1.3 and the fact that family u_α is weakly continuous. □

Definition 13.19.1.6. *Let $U \in S'_{2k}$ and $U \bullet S_k \subset S_k$ and $u : \mathbb{R}^k \to S'_k$ is a family of tempered distributions such that $(u_\alpha)_{\alpha \in \mathbb{R}^k} = U$. Let $\lambda \in S'_k$.*

$$\int \lambda u_\alpha d\alpha := U^T \bullet \lambda. \tag{13.19.1.8}$$

Theorem 13.19.1.7. *If $U \in S'_{2k}$ and $U^T \bullet S_k \subset S_k$, then $(U \bullet \delta_\alpha)_{\alpha \in \mathbb{R}^k} = U^T$.*

Proof. Take any $\phi, \psi \in S_k$

$$\left(\int \phi(\alpha)(U \bullet \delta_\alpha) d\alpha\right)(\psi) = \int \phi(\alpha)(U \bullet \delta_\alpha)(\psi) d\alpha =$$
$$\int \phi(\alpha)\delta_\alpha(U^T \bullet \psi) d\alpha = \left(\int \phi(\alpha)\delta_\alpha d\alpha\right)(U^T \bullet \psi) = \Lambda_\phi(U^T \bullet \psi) = (U \bullet \phi)(\psi). \tag{13.19.1.9}$$

Thus by Definition 13.19.1.1 thesis. □

Corollary 13.19.1.8. *If $U \in S'_{2k}$ and $U^T \bullet S_k \subset S_k$ and $\lambda \in S'_k$, then*

$$\int \lambda(U \bullet \delta_\alpha) d\alpha = U \bullet \lambda. \tag{13.19.1.10}$$

Theorem 13.19.1.9. *If $U \in S'_{2k}$, $U^T \bullet S_k \subset S_k$ and $V \in S'_{2k}$, $V^T \bullet S_k \subset S_k$ then there exists unique $W \in S'_{2k}$, such that*

$$W \bullet \phi = U \bullet (V \bullet \phi) \textit{ for every } \phi \in S_k. \tag{13.19.1.11}$$

Proof. Let's define a mapping $L(\phi) := U \bullet (V \bullet \phi)$ for any $\phi \in S_k$. By Theorem 13.18.1.7 applied to V and Theorem 13.18.1.11 applied to U, we have $L : S_k \to S'_k$ is continuous with weak-* topology on S'_k. Therefore by Theorem 11.5.5.3, there exists an unique $W \in S'_{2k}$ such that $W \bullet \phi = L(\phi)$. □

The above theorem enables us to formulate the following definition.

Definition 13.19.1.10. *Let* $U \in S'_{2k}$, $U^T \bullet S_k \subset S_k$ *and* $V \in S'_{2k}$, $V^T \bullet S_k \subset S_k$.

$$U \bullet V := W \tag{13.19.1.12}$$

iff

$$W \bullet \phi = U \bullet (V \bullet \phi) \text{ for every } \phi \in S_k. \tag{13.19.1.13}$$

Theorem 13.19.1.11. *If* $U \in S'_{2k}$, $U^T \bullet S_k \subset S_k$ *and* $V \in S'_{2k}$, $V^T \bullet S_k \subset S_k$ *then*

$$(U \bullet V)^T = V^T \bullet U^T. \tag{13.19.1.14}$$

Proof. Take any $\phi, \psi \in S_k$

$$\begin{aligned}((U \bullet V) \bullet \phi)(\psi) = (U \bullet (V \bullet \phi))(\psi) = (V \bullet \phi)(U^T \bullet \psi) = \Lambda_\phi(V^T \bullet (U^T \bullet \psi))\\ = \Lambda_\phi((V^T \bullet U^T) \bullet \psi)) = (((V^T \bullet U^T)^T) \bullet \phi)(\psi).\end{aligned} \tag{13.19.1.15}$$

□

Corollary 13.19.1.12. *If* $U \in S'_{2k}$, $U^T \bullet S_k \subset S_k$ *and* $V \in S'_{2k}$, $V^T \bullet S_k \subset S_k$ *then* $(V \bullet U)^T \bullet S_k \subset S_k$

Theorem 13.19.1.13. *Let* $V \in S'_{2k}$ *such that* $V \bullet S_k \subset S_k$ *and* $v : \mathbb{R}^k \to S'_k$ *be a family of tempered distributions such that* $(v_\alpha)_{\alpha \in \mathbb{R}^k} = V$. *If* $U \in S'_{2k}$ *and* $U^T \bullet S_k \subset S_k$ *and* $\lambda \in S'_k$, *then*

$$(U \bullet v_\alpha)_{\alpha \in \mathbb{R}^k} = (U \bullet V^T)^T \text{ and } (U \bullet V^T)^T \bullet S_k \subset S_k \tag{13.19.1.16}$$

and

$$U \bullet \int \lambda v_\alpha d\alpha = \int \lambda (U \bullet v_\alpha) d\alpha. \tag{13.19.1.17}$$

Proof. By Definition 13.19.1.6

$$U \bullet \int \lambda v_\alpha d\alpha = U \bullet (V^T \bullet \lambda) = (U \bullet V^T) \bullet \lambda. \tag{13.19.1.18}$$

First, we will show that

$$(U \bullet v_\alpha)_{\alpha \in \mathbb{R}^k} = (U \bullet V^T)^T. \tag{13.19.1.19}$$

Take any $\phi, \psi \in S_k$.

$$\begin{aligned}\left(\int \phi(\alpha)(U \bullet v_\alpha) d\alpha\right)(\psi) = \int \phi(\alpha)(U \bullet v_\alpha)(\psi) d\alpha = \left(\int \phi(\alpha) v_\alpha d\alpha\right)(U^T \bullet \psi)\\ = (V^T \bullet \phi)(U^T \bullet \psi) = (U \bullet (V^T \bullet \phi)(\psi) = ((U \bullet V^T) \bullet \phi)(\psi).\end{aligned} \tag{13.19.1.20}$$

We showed that (13.19.1.19). Note that by Theorem 13.19.1.11 and assumptions about U and V, we have $(U \bullet V^T)^T \bullet S_k = (V \bullet U^T) \bullet S_k \subset S_k$. Now, by Definition 13.19.1.6,

$$\int \lambda(U \bullet v_\alpha) d\alpha = (V \bullet U^T)^T \bullet \lambda = (U \bullet V^T) \bullet \lambda. \tag{13.19.1.21}$$

The above together with (13.19.1.18) gives thesis. □

13.20 Monday, 26 August 2019

Definition 13.20.0.1. *Define* $\overline{(.)} : S'_k \to S'_k$ *as*

$$\overline{u}(\phi) := \overline{u(\overline{\phi})}, \tag{13.20.0.1}$$

where $u \in S'_k$ *and* $\phi \in S_k$.

Note that $\overline{\Lambda_f} = \Lambda_{\overline{f}}$. Indeed

$$\overline{\Lambda_f}(\phi) = \overline{\Lambda_f(\overline{\phi})} = \overline{\int f\overline{\phi}} = \int \overline{f}\phi = \Lambda_{\overline{f}}(\phi). \tag{13.20.0.2}$$

This justifies the above definition.

Lemma 13.20.0.2. *If* $U \in S'_{2k}$ *and* $\phi \in S_k$, *then*

$$\overline{U \bullet \phi} = \overline{U} \bullet \overline{\phi}. \tag{13.20.0.3}$$

Proof. Take any $\psi \in S_k$.

$$\overline{U \bullet \phi}(\psi) = \overline{U(\phi \otimes \overline{\psi})} = \overline{U}(\overline{\phi} \otimes \psi) = \overline{U} \bullet \overline{\phi}(\psi). \tag{13.20.0.4}$$

□

Theorem 13.20.0.3. *If* $U \in S'_{2k}$,

$$\overline{U^T} = \overline{U}^T. \tag{13.20.0.5}$$

Proof. Take any $\phi \in S_{2k}$

$$\overline{U^T}(\phi) = \overline{U^T(\overline{\phi})} = \overline{U(\overline{\phi}^T)} = \overline{U(\overline{\phi^T})} = \overline{U}(\phi^T) = \overline{U}^T(\phi). \tag{13.20.0.6}$$

□

Theorem 13.20.0.4. *If $U \in S'_{2k}$ and $U^T \bullet S_k \subset S_k$ and $v \in S'_k$, then*

$$\overline{U \bullet v} = \overline{U} \bullet \overline{v}. \tag{13.20.0.7}$$

Proof. Take any $\phi \in S_k$.

$$\overline{U \bullet v}(\phi) = \overline{v(U^T \bullet \overline{\phi})} = \overline{v}(\overline{U}^T \bullet \phi) = (\overline{U} \bullet \overline{v})(\phi). \tag{13.20.0.8}$$

□

Theorem 13.20.0.5. *Let $U \in S'_{2k}$ and $u : \mathbb{R}^k \to S'_k$ be a family of tempered distributions such that $U = (u_\alpha)_{\alpha \in \mathbb{R}^k}$, then*

$$\overline{U} = (\overline{u}_\alpha)_{\alpha \in \mathbb{R}^k}. \tag{13.20.0.9}$$

Proof. Take any $\psi \in S_k$.

$$\begin{aligned}\int_{\mathbb{R}^k} \phi(\alpha)\overline{u}_\alpha(\psi)d\alpha = \int_{\mathbb{R}^k} \phi(\alpha)\overline{u_\alpha(\overline{\psi})}d\alpha = \overline{\int_{\mathbb{R}^k} \overline{\phi(\alpha)}u_\alpha(\overline{\psi})d\alpha} \\ = \overline{(U^T \bullet \overline{\phi})(\overline{\psi})} = (\overline{U}^T \bullet \phi)(\psi).\end{aligned} \tag{13.20.0.10}$$

□

Theorem 13.20.0.6. *Let $U \in S'_{2k}$ and $u : \mathbb{R}^k \to S'_k$ be a family of tempered distributions such that $U = (u_\alpha)_{\alpha \in \mathbb{R}^k}$ and $U \bullet S_k \subset S_k$. Let $\lambda \in S'_k$. Then*

$$\overline{\int \lambda u_\alpha d\alpha} = \int \overline{\lambda}\overline{u}_\alpha d\alpha. \tag{13.20.0.11}$$

Proof. We get this immediately by Theorem 13.20.0.5 and by Theorem 13.20.0.4. □

Definition 13.20.0.7. *Let $U \in S'_k$ and $U^T \bullet S_k \subset S_k$ and $u : \mathbb{R}^k \to S'_k$ is a family of tempered distributions such that $(u_\alpha)_{\alpha \in \mathbb{R}^k} = U$. For any $v \in S'_k$ we define*

$$\langle v|u_\alpha\rangle := U \bullet \overline{v}. \tag{13.20.0.12}$$

Theorem 13.20.0.8. *Let $U \in S'_k$ and $U^T \bullet S_k \subset S_k$ and $u : \mathbb{R}^k \to S'_k$ is a family of tempered distributions such that $(u_\alpha)_{\alpha \in \mathbb{R}^k} = U$.*

Let $V \in S'_k$ and $V \bullet S_k \subset S_k$ and $v : \mathbb{R}^k \to S'_k$ is a family of tempered distributions such that $(v_\beta)_{\beta \in \mathbb{R}^k} = V$. If $\lambda \in S'_k$, then

$$\int \overline{\lambda} \langle v_\beta|u_\alpha\rangle \, d\beta = \left\langle \int \lambda v_\beta \middle| u_\alpha \right\rangle. \tag{13.20.0.13}$$

Proof.

$$\int \overline{\lambda} \langle v_\beta | u_\alpha \rangle \, d\beta = \int \overline{\lambda}(U \bullet \overline{v_\beta}) d\beta = U \bullet \int \overline{\lambda} \overline{v_\beta} d\beta = U \bullet \overline{\int \lambda v_\beta d\beta} = \left\langle \int \lambda v_\beta \middle| u_\alpha \right\rangle. \tag{13.20.0.14}$$

The second equality in the above equation is by Theorem 13.19.1.13, the third equality is by Theorem 13.20.0.6. □

Definition 13.20.0.9. *A family of tempered distributions $u : \mathbb{R}^k \to S'_k$ is called an orthonormal continous basis of S'_k iff*

1. *There exists $U \in S'_{2k}$ such that $U \bullet S_k \subset S_k$ and $(u_\alpha)_{\alpha \in \mathbb{R}^k} = U$.*
2. *For any $v \in S'_k$ there exists $\lambda \in S'_k$ such that*
$$v = \int \lambda u_\alpha d\alpha. \tag{13.20.0.15}$$
3. *we have*
$$\langle u_\beta | u_\alpha \rangle = \delta_\beta \tag{13.20.0.16}$$
for any $\beta \in \mathbb{R}^k$.

Theorem 13.20.0.10. *Let $u : \mathbb{R}^k \to S'_k$ be a family of tempered distributions. u_α is an orthonormal continous basis of S'_k iff there exists an $U \in S'_{2k}$ such that $U \bullet S_k \subset S_k$, $U \bullet \overline{U}^T = \overline{U}^T \bullet U = I$ and $(u_\alpha)_{\alpha \in \mathbb{R}^k} = U$.*

13.21 Wednesday, 28 August 2019

13.21.1 Rigorous Dirac Formulation - continuation (2)

We will rewrite reasonings from 13.19.1 in more generic way. That will hopefully enable us later to apply this apparatus to mix of contiunous and discrete orthonormal basis.

Topological preparation We will use symbol $\hookrightarrow$ to denote continuous embedding.

Definition 13.21.1.1. *Let S be TVS. For any $y \in S'$ and $x \in S$, we define $(\cdot, \cdot) : S \times S' \to \mathbb{C}$ as*
$$(x, y) := y(x). \tag{13.21.1.1}$$

Definition 13.21.1.2. *Let S be a TVS. We say that S is embeded continously and communicativly in S' and we denote this by $S \hookrightarrow_C S'$ if for identity mapping $i : S \to S'$ holds*

$$(\phi, i(\psi)) = (\psi, i(\phi)) \tag{13.21.1.2}$$

for any $\phi, \psi \in S$.

Usually we will omit i writing only $(\phi, \psi) = (\psi, \phi)$.

Corollary 13.21.1.3. *Let S be a TVS such that $S \hookrightarrow_C S'$. Then*

$$(\phi, \psi) = (\psi, \phi) \text{ for all } \phi, \psi \in S. \tag{13.21.1.3}$$

Proof. This follows directly from Definition 13.21.1.2 □

Unless stated otherwise, S' in this section carries weak-* topology.

We define trasposition of continuous linear map after [**?**]. We will use A^t in subscript instead of inconvenient for edition in latex ${}^tA$.

Definition 13.21.1.4. *Let X, Y be TVS. For any continuous and linear $A : X \to Y$ we define $A^t : Y' \to X'$ as follows*

$$(A^t u) := u \circ A. \tag{13.21.1.4}$$

Corollary 13.21.1.5. *Let X, Y be TVS. If $A : X \to Y$ is a continuous linear mapping, $A^t : Y' \to X'$ is continous in weak-* topologies.*

Theorem 13.21.1.6. *If E is a locally convex Hausdorff TVS, then $(E')'$ is isomorphic to E with isomorphism $(E')' \ni L \mapsto x_L \in E$ where $x'(x_L) = L(x')$ for any $x' \in E'$.*

Proof. [See **?** , 35.1] □

After [**?**] for topological space E, we will denote weak-* topology on E' as $\sigma(E', E)$. Note that topology $\sigma(E', E)$ doesn't depend on topology of E. Obviously E' as set depends on topology of E.

Theorem 13.21.1.7. *Let E be a metrizable locally convex TVS and F be a locally convex Hausdorff TVS, and $A : E \to F$ is linear. If A is continous when E carries topology $\sigma(E, E')$ and F carries topology $\sigma(F, F')$, then A is continous in original topologies of E and F.*

Proof. [See **?** , 37.6] □

Theorem 13.21.1.8. *Let S be a Fréchet space. For any continuous linear mapping $A : S \to S'$, a mapping $A^t : (S'' = S) \to S'$ is continuous as a mapping $S \to S'$ (S' with weak-* topology) and*

$$(\psi, A\phi) = (\phi, A^t\psi) \tag{13.21.1.5}$$

for any $\phi, \psi \in S$.

Proof. By Corollary 13.21.1.5 $A^t : S'' \to S'$ is continous in weak-* topologies. By Theorem 13.21.1.6 we may substitute $S = S''$. Then $A^t : S \to S'$ is continuous when S carries topology $\sigma(S, S') = \sigma(S'', S')$ and S' carries $\sigma(S', S) = \sigma(S', S'')$. By Theorem 11.5.2.28 S' is a locally convex Hausdorff space, thus by Theorem 13.21.1.7, $A^t : S \to S'$ is continuous in original topologies S in S' (for S' this is still weak-* topology). The equation (13.21.1.5) follows from Corollary 13.21.1.5. Indeed $(\psi, A\phi) = (A\phi)(\psi) = (A^t\psi)(\phi) = (\phi, A^t\psi)$. □

Corollary 13.21.1.9. *Let S be a Fréchet space. If $A : S \to S'$ and $B : S \to S'$ are continuous linear maps such as $(\psi, A\phi) = (\phi, B\psi)$ for all $\phi, \psi \in S$ then $A^t = B$.*

Corollary 13.21.1.10. *Let S be a Fréchet space. If $A : S \to S'$ is a continuous linear map, then $(A^t)^t = A$.*

Theorem 13.21.1.11. *Let S be a Fréchet space and $S \hookrightarrow S'$. If $A : S \to S'$ is a continuous linear mapping such that $A(S) \subset S$, then A is continuous as $A : S \to S$.*

Proof. It is enough to prove that the graph of the linear mapping $S \ni \phi \mapsto A\phi \in S$ is closed and then by Theorem 11.5.2.14 (The closed graph theorem) it will be immediately proven that the mapping is continuous. Take any $\phi_n \to \phi$ in S, such that $A\phi_n \to v$ in S. Since $A : S \to S'$ is continuous, we have $A\phi_n \to A\phi$ in S'. But since $S \hookrightarrow S'$, $A\phi_n \to v$ in S'. By Theorem 11.5.2.28 S' is a locally convex Hausdorff space, thus $A\phi = v$. We showed that the graph of A is closed what completes the proof. □

Theorem 13.21.1.12. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A : S \to S'$ is a continuous linear mapping and $A^t(S) \subset S$, we can extend A to a continuous linear mapping $A_\bullet : S' \to S'$ in the following way*

$$(A_\bullet u)(\phi) = (A^t\phi, u). \tag{13.21.1.6}$$

Proof. It's clear that $A_\bullet : S' \to S'$ is continuous in weak-* topology. It is enough to show that extension is consistent with $A : S \to S'$. Indeed, take any $\phi, \psi \in S$. By Corollary 13.21.1.3 and Theorem 13.21.1.5, we have $(A^t\phi, \psi) = (\psi, A^t\phi) = (\phi, A\psi) = (A\psi)(\phi)$. □

Definition 13.21.1.13. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A : S \to S'$ is a continuous linear mapping and $A^t(S) \subset S$. by $A_\bullet$ we will denote the extension of A from Theorem 13.21.1.12.*

Remark 13.21.1.14. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A : S \to S'$ is a continuous linear mapping and $A^t(S) \subset S$, then*

$$(\phi, A_\bullet u) = (A^t\phi, u) \tag{13.21.1.7}$$

for any $u \in S'$ and any $\phi \in S$.

Proof. Follows directly from Theorem 13.21.1.12. □

Theorem 13.21.1.15. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A \in L(S, S')$ with $A^t(S) \subset S$ and $B \in L(S, S')$ with $B^t(S) \subset S$, then $(A_\bullet B)^t = B^t A^t$.*

Proof. Take any $\psi, \phi \in S$.

$$(\psi, A_\bullet B\phi) = (A^t\psi, B\phi) = (\phi, B^t A^t\psi). \tag{13.21.1.8}$$

The first equality is by Theorem 13.21.1.5 and the second eqaulity is by Therem 13.21.1.10. Now by Corollary 13.21.1.10 we have thesis. □

Theorem 13.21.1.16. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A \in L(S, S')$ with $A^t(S) \subset S$ and $B \in L(S, S')$ with $B^t(S) \subset S$, then $(A_\bullet B)_\bullet = A_\bullet B_\bullet$.*

Proof. Take any $u \in S'$ and $\phi \in S$.

$$(\phi, A_\bullet B_\bullet u) = (A^t\phi, B_\bullet u) = (B^t A^t\phi, u) = ((A_\bullet B)^t\phi, u) = (\phi, (A_\bullet B)_\bullet). \tag{13.21.1.9}$$

The first two equalities are by Remark 13.21.1.14. Third equality is by Theorem 13.21.1.15. And again the last equality is by Remark 13.21.1.14. □

Theorem 13.21.1.17. *Let S be a Fréchet space and $S \hookrightarrow_C S'$. If $A : S \to S'$ is a continuous linear mapping such that $A : S \underset{onto}{\overset{1-1}{\to}} S$ and $A^t : S \underset{onto}{\overset{1-1}{\to}} S$, then $(A^{-1})^t = (A^t)^{-1}$.*

Proof. Take any $\psi, \phi \in S$. Consider

$$(A^{-1}\psi, \phi) = (A^{-1}\psi, A^t(A^t)^{-1}\phi) = ((A^t)^{-1}\phi, AA^{-1}\psi) = ((A^t)^{-1}\phi, \psi). \tag{13.21.1.10}$$

Since $S \hookrightarrow_C S'$, we have

$$(\phi, A^{-1}\psi) = (\psi, (A^t)^{-1}\phi). \tag{13.21.1.11}$$

Thus by Corollary 13.21.1.9 we proved thesis. □

We will assume axiomatic definition of vector complex conjugate

Definition 13.21.1.18. *Let S be a TVS such that $S \hookrightarrow_C S'$. Mapping $\overline{(.)} : S \to S$ will be called a vector complex conjugate iff the following conditions are satisfied*

1. $\overline{(.)} : S \to S$ *is continuous antilinear mapping*
2. $\overline{\overline{\phi}} = \phi$ *for any* $\phi \in S$.
3. $(\phi, \overline{\psi}) = \overline{(\overline{\phi}, \psi)}$ *for any* $\phi, \psi \in S$.

Definition 13.21.1.19. *Let S be a TVS such that $S \hookrightarrow_C S'$. For any $u \in S'$ we define $\overline{u}$ such that*

$$\overline{u}(\phi) := \overline{u(\overline{\phi})}. \tag{13.21.1.12}$$

Corollary 13.21.1.20. *Let S be a TVS such that $S \hookrightarrow_C S'$. If $u \in S'$, then $\overline{u} \in S'$.*

Corollary 13.21.1.21. *Let S be a TVS such that $S \hookrightarrow_C S'$. Then $\overline{(.)} : S' \to S'$ is a continuous antilinear mapping such that $\overline{\overline{u}} = u$ for any $u \in S'$.*

Corollary 13.21.1.22. *Let S be a TVS such that $S \hookrightarrow_C S'$.*

$$(\phi, \overline{u}) = \overline{(\overline{\phi}, u)} \text{ for any } \phi \in S \text{ and } u \in S'. \tag{13.21.1.13}$$

Corollary 13.21.1.23. *Let S be a TVS such that $S \hookrightarrow_C S'$ and $i : S \to S'$ is identity mapping.*

$$i\overline{\phi} = \overline{i\phi} \text{ for any } \phi \in S. \tag{13.21.1.14}$$

The above means simply that the definition of $\overline{(\cdot)}$ coincide on elements from S and S'.

Definition 13.21.1.24. *Let S be a TVS such that $S \hookrightarrow_C S'$.*

$$\langle \phi | u \rangle := (\overline{\phi}, u). \tag{13.21.1.15}$$

Definition 13.21.1.25. *Let S be a TVS such that $S \hookrightarrow_C S'$ and $A \in L(S, S')$.*

$$A^* \phi = \overline{A^t \overline{\phi}} \text{ for any } \phi \in S. \tag{13.21.1.16}$$

Theorem 13.21.1.26. *Let S be a Fréchet space such that $S \hookrightarrow_C S'$. If $A \in L(S, S')$, then $A^* \in L(S, S')$ and*

$$\langle \phi | A\psi \rangle = \overline{\langle \psi | A^* \phi \rangle} \text{ for any } \phi, \psi \in S. \tag{13.21.1.17}$$

Proof.

$$\langle \phi | A\psi \rangle = (\overline{\phi}, A\psi) = (\psi, A^t \overline{\phi}) = \overline{\overline{(\psi, A^t \overline{\phi})}} = \overline{(\overline{\psi}, \overline{A^t \overline{\phi}})} = \overline{\langle \psi | A^* \phi \rangle}. \tag{13.21.1.18}$$

□

In context of measurable spaces

Definition 13.21.1.27. *Let (Ω, μ) be a measurable space. We will denote by $F(\Omega)$ a vector space of all complex valued measurable functions and by $F_0(\Omega)$ a vector space of all meaurable functions equal 0 almost everywhere in μ.*

Let S be a vector subspace of $F(\Omega)/F_0(\Omega)$. We define

$$F_\mu(\Omega; S) := \{u \in F(\Omega)/F_0(\Omega) : S \ni \phi \mapsto \int u\phi d\mu \in \mathbb{C} \text{ is continuous}\} \tag{13.21.1.19}$$

We will endow $F_\mu(\Omega; S)$ with weak- topology generated by S.*

In the context of measurable space (Ω, μ) for any $f \in F(\Omega)$

$$[f] := \{g \in F(\Omega) : f - g \in F_0(\Omega)\}. \tag{13.21.1.20}$$

It is important to note that in general it might be not true that $F_\mu(\Omega; S) \hookrightarrow S'$, because you might find $u \in F_\mu(\Omega; S)$ such that $u \neq 0$ for which $\int u\phi d\mu = 0$ for all $\psi \in S$. That's because S might be "too small" to seperate all elements from $F_\mu(\Omega; S)$. Hence the need of the following definition.

Definition 13.21.1.28. *Let S be as in Definition 13.21.1.27. We assume that S is TVS with its own topology. Let E be a vector subspace of $F_\mu(\Omega; S)$ and E carries weak-* topology generated by S We will say that S is strongly L^2 embedded in E and denote this by $S \hookrightarrow_{L^2} E$ iff $S \hookrightarrow L^2(\Omega; \mathbb{C}) \hookrightarrow E \hookrightarrow S'$*

The following theorem will show that such E might exist.

Theorem 13.21.1.29. *Let (Ω, μ) be a measurable space. Let S be as in Definition 13.21.1.27. Let S be TVS with its own topology. If $S \hookrightarrow L^2(\Omega; \mathbb{C})$ and S is dense in $L^2(\Omega; \mathbb{C})$, then $S \hookrightarrow_{L^2} L^2(\Omega; \mathbb{C})$.*

Proof. Since S is embeded continuously in $L^2(\Omega)$, it is easy to show that $L^2(\Omega) \subset F_\mu(\Omega; S)$. Indeed, consider the inequality below for any $\phi, \phi_n, \psi \in L^2(\Omega)$.

$$\left| \int \psi\phi_n d\mu - \int \psi\phi d\mu \right| = \left| \int \psi(\phi_n - \phi) d\mu \right| \leq \|\psi\|_{L^2} \|\phi_n - \phi\|_{L^2} \tag{13.21.1.21}$$

Assume that $\psi \in L^2(\Omega)$ and $\phi_n \to \phi$ in S. Then $\phi_n \to \phi$ in $L^2(\Omega)$ and therefore $\psi \in F_\mu(\Omega; S)$. Let $E := L^2(\Omega)$. We will endow E with weak-* topology generated by S. It is quite clear how to show that $L^2(\Omega; \mathbb{C}) \hookrightarrow E$. Consider again inequality (13.21.1.21) with $\psi \in S$ and $\phi_n \to \phi$ in $L^2(\Omega)$. By inequality (13.21.1.21) they converges in weak-* topology generated by S, thus $L^2(\Omega)$ is embeded continuously in E. We will show that $E \hookrightarrow S'$.

Let $j : E \to S'$ in a way that $j(u)(\phi) = \int u\phi d\mu$ for any $u \in E$ and $\phi \in S$. Since both E and S' carries weak-* topologies generated by S, it's clear that j is continuous. To prove that $E \hookrightarrow S'$ it is enough to show that j is an injection. Take $u \in E$ such that $j(u) = 0$, we will show that $u = 0$. Since S is dense in $L^2(\Omega)$ we have $S \ni \phi_n \to \overline{u}$ in $L^2(\Omega)$. Thus

$$0 = \int u\phi_n d\mu \to \int u\overline{u} d\mu = \|u\|_{L^2}^2. \tag{13.21.1.22}$$

Hence we have $u = 0$. We completed the reasoning that $E \hookrightarrow S'$. □

Definition 13.21.1.30. *Let (Ω, μ) be a measurable space and let S be a vector subspace of $F(\Omega)/F_0(\Omega)$. We will say that $\delta : \Omega \to S^*$ is a family of Dirac deltas iff for any $f \in F(\Omega)$ such that $[f] \in S$, we have*

$$\delta_\alpha(f) = f(\alpha) \text{ for almost all } \alpha \in \Omega. \tag{13.21.1.23}$$

Definition 13.21.1.31. *Let (Ω, μ) be measurable space and let S be TVS. Let $u : \Omega \to S'$.*

1. *We will say that u is measurable iff $\Omega \ni \alpha \mapsto u_\alpha(\phi)$ is measurable for any $\phi \in S$.*

2. $$\int_\Omega u_\alpha \mu(d\alpha) := v \in S' \tag{13.21.1.24}$$

   if and only if

   $$\int_\Omega u_\alpha(\phi)\mu(d\alpha) = v(\phi) \text{ for any } \phi \in S. \tag{13.21.1.25}$$

We say that integral from (13.21.1.24) exists if an only if such $v \in S'$ exists.

Theorem 13.21.1.32. *Let S be TVS and let E be a vector subspace of $F_\mu(\Omega; S)$. If $S \hookrightarrow_{L^2} E$, then $S \hookrightarrow_C S'$.*

Proof. Since E carries is weak-* topology, we have $E \hookrightarrow S'$. Thus $S \hookrightarrow S'$. By Definition 13.21.1.28, for any $\phi \in S$ and $u \in E$ we have

$$(\phi, u) = \int u\phi d\mu. \tag{13.21.1.26}$$

Thus, for any $\phi, \psi \in S$, we have $(\phi, \psi) = (\psi, \phi)$. □

Definition 13.21.1.33. *Let (Ω, μ) be a measurable space, let S be as in Definition 13.21.1.27 and let S be a TVS. Let $u : \Omega \to S'$ and let $U : S \to S'$ be a continuous linear mapping. We define that*

$$(u_\alpha)_{\mu;\alpha\in\Omega} := U \tag{13.21.1.27}$$

iff

$$U^t\phi = \int \phi(\alpha)u_\alpha\mu(d\alpha) \text{ for any } \phi \in S. \tag{13.21.1.28}$$

Definition 13.21.1.34. *Let (Ω, μ) be a measurable space, let S be as in Definition 13.21.1.27 and let S be a TVS. Let E be a vector subspace of $F_\mu(\Omega; S)$. We will say that $u : \Omega \to S'$ is measurable of class E iff*

$$[\Omega \ni \alpha \mapsto u_\alpha(\phi) \in \mathbb{C}] \in E \tag{13.21.1.29}$$

for any $\phi \in S$.

Lemma 13.21.1.35. *Let (Ω, μ) be a measurable space, let S be as in Definition 13.21.1.27 and let S be a Fréchet space. Let E be a vector subspace of $F_\mu(\Omega; S)$ and $S \hookrightarrow_{L^2} E$. Let $u : \Omega \to S'$ be measurable of class E. If $U \in L(S, S')$ such that $(u_\alpha)_{\mu;\alpha\in\Omega} = U$, then*

$$U\psi = [\Omega \ni \alpha \mapsto u_\alpha(\psi)] \text{ for any } \psi \in S. \tag{13.21.1.30}$$

Proof. Take any $\psi \in S$. Consider a linear functional

$$S \ni \phi \mapsto (\phi, U\psi) = (\psi, U^t\phi) = (\psi, \int \phi(\alpha)u_\alpha\mu(d\alpha)) = \int \phi(\alpha)u_\alpha(\psi)\mu(d\alpha) \tag{13.21.1.31}$$

This functional is clearly continous as $U\psi \in S'$. Thus, by Definition 13.21.1.27, we have $[\Omega \ni \alpha \mapsto u_\alpha(\psi)] \in F_\mu(\Omega; S)$. But $u : \Omega \to S'$ is measurable of class E, thus $[\Omega \ni \alpha \mapsto u_\alpha(\psi)] \in E$. Since $E \hookrightarrow S'$, it is apparent from (13.21.1.31) that $[\Omega \ni \alpha \mapsto u_\alpha(\psi)] = U\psi$. □

Theorem 13.21.1.36. *Let (Ω, μ) be a measurable space, let S be as in Definition 13.21.1.27 and let S be a Fréchet space. Let E be a vector subspace of $F_\mu(\Omega; S)$ and $S \hookrightarrow_{L^2} E$. Let $u : \Omega \to S'$ be measurable of class E. Let $U \in L(S, S')$ and $U(S) \subset S$ such that $(u_\alpha)_{\mu;\alpha\in\Omega} = U$. If $f \in E$, then*

$$\int f(\alpha)u_\alpha\mu(d\alpha) = (U^t)_\bullet f. \tag{13.21.1.32}$$

Proof. By Lemma 13.21.1.35, we have $[\Omega \ni \alpha \mapsto u_\alpha(\psi)] = U\psi$. But because $U(S) \subset S$, we have $U\psi \in S$. Consider

$$(\psi, (U^t)_\bullet f) = (U\psi, f) = \int f(\alpha) u_\alpha(\psi)\mu(d\alpha). \tag{13.21.1.33}$$

□

Theorem 13.21.1.36 justifies the following definition.

Definition 13.21.1.37. *Let (Ω, μ) be a measurable space and let S be a Fréchet space and $S \hookrightarrow_C S'$. Let $U \in L(S, S')$ such that $U(S) \subset S$ and let $u : \Omega \to S'$ such that $(u_\alpha)_{\mu;\alpha\in\Omega} = U$. Let $\lambda \in S'$. We define*

$$\int \lambda u_\alpha \mu(d\alpha) := (U^t)_\bullet \lambda. \tag{13.21.1.34}$$

Theorem 13.21.1.38. *Let (Ω, μ) be a measurable space and let S be a Fréchet space and $S \hookrightarrow_C S'$. Let $V \in L(S, S')$ such that $V(S) \subset S$ and let $v : \Omega \to S'$ be such that $(v_\alpha)_{\mu;\alpha\in\Omega} = V$. If $U \in L(S, S')$ with $U^t(S) \subset S$ and $\lambda \in S'$, then*

$$(U_\bullet v_\alpha)_{\mu;\alpha\in\Omega} = (U_\bullet V^t)^t \text{ and } (U_\bullet V^t)^t(S) \subset S. \tag{13.21.1.35}$$

and

$$U_\bullet \int \lambda v_\alpha \mu(d\alpha) = \int \lambda(U_\bullet v_\alpha)\mu(d\alpha). \tag{13.21.1.36}$$

Proof. First we will show that $(U_\bullet v_\alpha)_{\mu;\alpha\in\Omega} = (UV^t)^t$. Take any $\phi, \psi \in S$.

$$\begin{aligned}(\psi, \int \phi(\alpha) U_\bullet v_\alpha \mu(d\alpha)) &= \int \phi(\alpha)(\psi, U_\bullet v_\alpha)\mu(d\alpha) = \int \phi(\alpha)(U^t\phi, v_\alpha)\mu(d\alpha) \\ &= (U^t\psi, \int \phi(\alpha) v_\alpha \mu(d\alpha)) = (U^t\psi, V^t\phi) = (\psi, U_\bullet V^t \phi).\end{aligned} \tag{13.21.1.37}$$

Now by Theorem 13.21.1.15, we have $(U_\bullet V^t)^t = VU^t$. This completes showing (13.21.1.35). Now, we will show (13.21.1.36).

$$U_\bullet \int \lambda v_\alpha \mu(d\alpha) = U_\bullet (V^t)_\bullet \lambda = (U_\bullet V^t)_\bullet = \int \lambda(Uv_\alpha)\mu(d\alpha). \tag{13.21.1.38}$$

Where first equality is by Definition 13.21.1.37, second equality is by Theorem 13.21.1.16 and the last equality is by (13.21.1.35) and again Definition 13.21.1.37. □

Corollary 13.21.1.39. *Let (Ω, μ) be a measurable space, let S be as in Definition 13.21.1.27 and let S be a Fréchet space. Let E be a vector subspace of $F_\mu(\Omega; S)$ and $S \hookrightarrow_{L^2} E$. Let $v : \Omega \to S'$ be measurable of class E. If $V \in L(S, S')$ such that $V : S \underset{onto}{\overset{1-1}{\to}} S$ and $V^t : S \underset{onto}{\overset{1-1}{\to}} S$ and $(v_\alpha)_{\mu;\alpha\in\Omega} = V$, then $\Omega \ni \alpha \mapsto ((V^t)^{-1})_\bullet v_\alpha$ is a family of Dirac deltas.*

Proof. By Theorem 13.21.1.17 we have $((V^t)^{-1})^t = ((V^t)^t)^{-1}$. Thus by Corollary 13.21.1.10 $((V^t)^{-1})^t = V^{-1}$. Thus by Theorem 13.21.1.12 $((V^t)^{-1})_\bullet$ exists. Let $u_\alpha := ((V^t)^{-1})_\bullet v_\alpha$. Now by (13.21.1.36) $(u_\alpha)_{\mu;\alpha\in\Omega} = (((V^t)^{-1})_\bullet V^t)^t = I$. Now, by Lemma 13.21.1.35, we have

$$\psi = [\Omega \ni \alpha \mapsto u_\alpha(\psi)] \text{ for any } \psi \in S. \tag{13.21.1.39}$$

Thus by Definition 13.21.1.30, u_α is a family of Dirac deltas. □

13.22 Sunday, 8 September 2019

13.22.1 Rigged Hilbert Space

Definitions for many terms used in this subsection are from [**?**].

Definition 13.22.1.1. *Let $\Phi \subset H \subset \Phi'$ be a rigged Hilbert space and (Ω, μ) be measurable space where $L^2_\mu(\Omega; \mathbb{C})$ is isomporhic with H through unitary operator $U : H \to L^2_\mu(\Omega; \mathbb{C})$. We call U a realisation of $\Phi \subset H \subset \Phi'$ as a space of functions via isomorphism $L^2_\mu(\Omega; \mathbb{C}) \cong H$.*

When it doesn't cause disambiguity for any element $\phi \in \Phi$ we will use the same symbol ϕ to denote function $U\phi$ and consequently Φ to denote $U(\Phi)$. In that cases we will denote U symbolically as $\phi \mapsto \phi(\alpha)$.

Definition 13.22.1.2. *Let $\phi \mapsto \phi(\alpha)$ be a realisation of a rigged Hilbert space $\Phi \subset H \subset \Phi'$ as a space of functions via isomorphism $L^2_\mu(\Omega; \mathbb{C}) \cong H$. We will call $\delta : \Omega \to \Phi'$ a family of Dirac deltas iff for any measurable $f : \Omega \to \mathbb{C}$ such that $[f] \in \Phi$ we have*

$$\delta_\alpha(f) = f(\alpha) \textit{ for almost all } \alpha \in \Omega. \tag{13.22.1.1}$$

Theorem 13.22.1.3. *Let $\phi \mapsto \phi(\alpha)$ be a realisation of a rigged Hilbert space $\Phi \subset H \subset \Phi'$ as a space of functions via isomorphism $L^2_\mu(\Omega; \mathbb{C}) \cong H$. There exists a family of Dirac deltas $\delta : \Omega \to \Phi'$.*

Proof. [See **?** , Ch. I.4.3] □

From now in context of a a rigged Hilbert space $\Phi \subset H \subset \Phi'$ on we will always use representation of $[\phi] \in \Phi$ for which $\phi(\alpha) = \delta_\alpha(\phi)$ for any $\alpha \in \Omega$.

13.23 Sunday, 23 January 2022

This is about alternative formulation of relativistic quantum theory, where we will introduce observables related to 4-momentum. Energy operator will be then

$$E\phi = i\hbar \frac{\partial}{\partial x^0}\phi, \tag{13.23.0.1}$$

(Hamiltonian in this interpretation will not have meaning identical with energy operator but will be related exactly as you would expect from Schrödinger's equation $E\phi = H\phi$) and momentum as usual.

$$P_k\phi = -i\hbar \frac{\partial}{\partial x^k}\phi. \tag{13.23.0.2}$$

Consequently, we will have T time operator or X_0 if one likes.

$$T\phi = x_0\phi. \tag{13.23.0.3}$$

Note that

$$[T, E] = -i\hbar. \tag{13.23.0.4}$$

Indeed,

$$[T, E]\phi = (TE - ET)\phi = x_0 i\hbar \frac{\partial}{\partial x^0}\phi - i\hbar \frac{\partial}{\partial x^0}(x_0\phi)$$

$$= x_0 i\hbar \frac{\partial}{\partial x^0}\phi - i\hbar\phi - x_0 i\hbar \frac{\partial}{\partial x^0}\phi = -i\hbar\phi.$$

The above is obviously a time-energy Heisenberg uncertainty principle.

Note that we immediately get an invariance of uncertainty principle under Lorentz transformation. Assume $\beta \in (0, 1)$ is an arbitrary velocity ($c = 1$) and $\gamma = (1 - \beta^2)^{-1/2}$. Let's take a Lorentz transformation of 4-vector observables

$$X' = \gamma X - \beta\gamma T, \tag{13.23.0.5}$$

$$T' = \gamma T - \beta\gamma X. \tag{13.23.0.6}$$

Let's take also Lorentz transformation of 4-momentum observables

$$P'_x = \gamma P_x - \beta\gamma E, \tag{13.23.0.7}$$

$$E' = \gamma E - \beta\gamma P_x. \tag{13.23.0.8}$$

For this moment we may forget definitions of our operators. All that matters are their commutators $[X, P_x] = i\hbar$ and $[T, E] = -i\hbar$ and $[X, T] = [P_x, E] = [T, P_X] = [X, E] = 0$ (Note that commutation relations are analogous to Minkowski's metric tensor which corresponds nicely with why energy operator is defined with opposite sign than momentum operator).

Indeed,

$$[X', P'_x] = \gamma^2[X, P_x] + \beta^2\gamma^2[T, E] = i\hbar\frac{1-\beta^2}{1-\beta^2} = i\hbar. \tag{13.23.0.9}$$

$$[T', E'] = \gamma^2[T, E] + \beta^2\gamma^2[X, P_x] = i\hbar\frac{\beta^2-1}{1-\beta^2} = -i\hbar. \tag{13.23.0.10}$$

Invariance of $[X, P_x] = i\hbar$ and $[T, E] = -i\hbar$ has fundamental meaning given the defining role of these commutators in quantum mechanics. We can express it that we got exactly the same quantum mechanics in each Lorentz frame of reference.

Interpretation of this formulation of quantum mechanics, which will give us the same results as at least classical quantum mechanics for non-relativistic case is following. We treat ϕ as a quantum state which describes the whole history of a particle. When we want to know what are the properties of the particle in time t for a given frame of reference, we collapse state ϕ by projecting it on the subspace $\{\psi : T\psi = t\psi\}$, I. e, we first make an assumption that we have just measured time T of the state ϕ an got measurement t, then we can do any other sort of things on the collapsed state to establish its property at time t.

13.24 Sunday, 07 March 2022

Let assume we have a group $U(\varepsilon)$ of linear transformations of $\mathbb{R}^n$. Let A be it's generator, i.e:

$$\left.\frac{d}{d\varepsilon}U(\varepsilon)x\right|_{\varepsilon=0} = Ax. \tag{13.24.0.1}$$

or using Einstein summation convention:

$$\left.\frac{d}{d\varepsilon}U(\varepsilon)^{\mu}_{\nu}x^{\nu}\right|_{\varepsilon=0} = A^{\mu}_{\nu}x^{\nu}. \tag{13.24.0.2}$$

or in exponential form

$$U(\varepsilon) = e^{\varepsilon A}. \tag{13.24.0.3}$$

Consider transformation on $\phi : \mathbb{R}^n \to \mathbb{C}$ defined as

$$\big(\hat{U}(\varepsilon)\phi\big)(x) := \phi(U(-\varepsilon)x). \tag{13.24.0.4}$$

Note that $\hat{U}(\varepsilon)$ moves the shape of function ϕ against its domain exactly in the same direction as $U(\varepsilon)$ transforms domain. Note that

$$\begin{aligned} \left.\frac{d}{d\varepsilon}\big(\hat{U}(\varepsilon)\phi\big)(x)\right|_{\varepsilon=0} &= \left.\frac{d}{d\varepsilon}\phi(U(-\varepsilon)x)\right|_{\varepsilon=0} \\ = \partial_{\mu}\phi(U(-\varepsilon)x)\left.\frac{d}{d\varepsilon}U(-\varepsilon)^{\mu}_{\nu}x^{\nu}\right|_{\varepsilon=0} &= -\partial_{\mu}\phi(x)A^{\mu}_{\nu}x^{\nu}. \end{aligned} \tag{13.24.0.5}$$

Thus generator of $\hat{U}(\varepsilon)$ denoted as $\hat{A}$ is given by

$$(\hat{A}\phi)(x) = -\partial_{\mu}\phi(x)A^{\mu}_{\nu}x^{\nu}. \tag{13.24.0.6}$$