PREFACE

The following material comprises a set of class notes in Introduction to Physics taken by math graduate students in Ann Arbor in 1995/96. The goal of this course was to introduce some basic concepts from theoretical physics which play so fundamental role in a recent intermarriage between physics and pure mathematics. No physical background was assumed since the instructor had none. I am thankful to all my students for their patience and willingness to learn the subject together with me.

There is no pretense to the originality of the exposition. I have listed all the books which I used for the preparation of my lectures.

I am aware of possible misunderstanding of some of the material of these lectures and numerous inaccuracies. However I tried my best to avoid it. I am grateful to several of my colleagues who helped me to correct some of my mistakes. The responsibility for ones which are still there is all mine.

I sincerely hope that the set of people who will find these notes somewhat useful is not empty. Any critical comments will be appreciated.

CONTENTS

Lecture 1. CLASSICAL MECHANICS

1.1 We shall begin with Newton's law in Classical Mechanics:

$$m\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}(\mathbf{x}, t). \tag{1.1}$$

It describes the motion of a single particle of mass m in the Euclidean space $\mathbb{R}^3$. Here $\mathbf{x} : [t_1, t_2] \to \mathbb{R}^3$ is a smooth parametrized path in $\mathbb{R}^3$, and $\mathbf{F} : [t_1, t_2] \times \mathbb{R}^3 \to \mathbb{R}$ is a map smooth at each point of the graph of $\mathbf{x}$. It is interpreted as a *force* applied to the particle. We shall assume that the force $\mathbf{F}$ is *conservative*, i.e., it does not depend on t and there exists a function $V : \mathbb{R}^3 \to \mathbb{R}$ such that

$$\mathbf{F}(\mathbf{x}) = -\mathrm{grad}\ V(\mathbf{x}) = (-\frac{\partial V}{\partial x_1}(\mathbf{x}), -\frac{\partial V}{\partial x_2}(\mathbf{x}), -\frac{\partial V}{\partial x_3}(\mathbf{x})).$$

The function V is called the *potential energy*. We rewrite (1.1) as

$$m\frac{d^2\mathbf{x}}{dt^2} + \mathrm{grad}\ V(\mathbf{x}) = 0. \tag{1.2}$$

Let us see that this equation is equivalent to the statement that the quantity

$$S(\mathbf{x}) = \int_{t_1}^{t_2} (\frac{1}{2}m||\frac{d\mathbf{x}}{dt}||^2 - V(\mathbf{x}(t)))dt, \tag{1.3}$$

called the *action*, is stationary with respect to variations in the path $\mathbf{x}(t)$. The latter means that the derivative of S considered as a function on the set of paths with fixed ends $\mathbf{x}(t_1) = \mathbf{x}_1, \mathbf{x}(t_2) = \mathbf{x}_2$ is equal to zero. To make this precise we have to define the derivative of a functional on the space of paths.

1.2 Let V be a normed linear space and X be an affine space whose associated linear space is V. This means that V acts transitively and freely on X. A choice of a point $a \in X$

allows us to identify V with X via the map $x \to x+a$. Let $F : X \to Y$ be a map of X to another normed linear space Y. We say that F is differentiable at a point $x \in X$ if there exists a linear map $L_x : V \to Y$ such that the function $\alpha_x : X \to Y$ defined by

$$\alpha_x(h) = F(x+h) - F(x) - L_x(h)$$

satisfies

$$\lim_{||h|| \to 0} ||\alpha_x(h)||/||h|| = 0. \tag{1.4}$$

In our situation, the space V is the space of all smooth paths $\mathbf{x} : [t_1, t_2] \to \mathbb{R}^3$ with $\mathbf{x}(t_1) = \mathbf{x}(t_2) = 0$ with the usual norm $||\mathbf{x}|| = \max_{t \in [t_1, t_2]} ||\mathbf{x}(t)||$, the space X is the set of smooth paths with fixed ends, and $Y = \mathbb{R}$. If F is differentiable at a point x, the linear functional L_x is denoted by $F'(x)$ and is called the *derivative* at x. We shall consider any linear space as an affine space with respect to the translation action on itself. Obviously, the derivative of an affine linear function coincides with its linear part. Also one easily proves the chain rule: $(F \circ G)'(x) = F'(G(x)) \circ G'(x)$.

Let us compute the derivative of the action functional

$$S(\mathbf{x}) = \int_{t_1}^{t_2} L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt, \tag{1.5}$$

where

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

is a smooth function (called a *Lagrangian*). We shall use coordinates $\mathbf{q} = (q_1, \ldots, q_n)$ in the first factor and coordinates $\dot{\mathbf{q}} = (\dot{q}_1, \ldots, \dot{q}_n)$ in the second one. Here the dots do not mean derivatives. The functional S is the composition of the linear functional $I : C^\infty(t_1, t_2) \to \mathbb{R}$ given by integration and the functional

$$\mathcal{L} : \mathbf{x}(t) \to L(\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

Since

$$L(\mathbf{x} + \mathbf{h}(t), \dot{\mathbf{x}} + \dot{\mathbf{h}}(t)) = L(\mathbf{x}, \dot{\mathbf{x}}) + \sum_{i=1}^n \frac{\partial L}{\partial q_i}(\mathbf{x}, \dot{\mathbf{x}}) h_i(t) + \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i}(\mathbf{x}, \dot{\mathbf{x}}) \dot{h}_i(t) + o(||\mathbf{h}||)$$

$$= L(\mathbf{x}, \dot{\mathbf{x}}) + \mathrm{grad}_{\mathbf{q}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \mathbf{h}(t) + \mathrm{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \dot{\mathbf{h}}(t) + o(||\mathbf{h}||),$$

we have

$$S'(\mathbf{x})(\mathbf{h}) = I'(L(\mathbf{x})) \circ \mathcal{L}'(\mathbf{x})(\mathbf{h}) = \int_{t_1}^{t_2} (\mathrm{grad}_{\mathbf{q}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \mathbf{h}(t) + \mathrm{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \dot{\mathbf{h}}(t)) dt. \tag{1.6}$$

Using the integration by parts we have

$$\int_{t_1}^{t_2} \mathrm{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \dot{\mathbf{h}}(t) dt = [\mathrm{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \mathbf{h}]\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \mathrm{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \dot{\mathbf{h}}(t) dt.$$

This allows us to rewrite (1.6) as

$$S'(\mathbf{x})(\mathbf{h}) = \int_{t_1}^{t_2} \Big(\text{grad}_{\mathbf{q}} L(\mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt}\text{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}})\Big) \cdot \mathbf{h}(t)dt + [\text{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}}) \cdot \mathbf{h}]\Big|_{t_1}^{t_2}. \quad (1.7)$$

Now we recall that $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$. This gives

$$S'(\mathbf{x})(\mathbf{h}) = \int_{t_1}^{t_2} \Big(\text{grad}_{\mathbf{q}} L(\mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt}\text{grad}_{\dot{\mathbf{q}}} L(\mathbf{x}, \dot{\mathbf{x}})\Big) \cdot \mathbf{h}(t)dt.$$

Since we want this to be identically zero, we get

$$\frac{\partial L}{\partial q_i}(\mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(\mathbf{x}, \dot{\mathbf{x}}) = 0, i = 1, \ldots, n. \quad (1.8)$$

This are the so called *Euler-Lagrange equations.*

In our situation of Newton's law, we have $n = 3, \mathbf{q} = (q_1, q_2, q_3) = \mathbf{x} = (x_1, x_2, x_3)$,

$$L(\mathbf{q}, \dot{\mathbf{q}}) = -V(q_1, q_2, q_3) + \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2) \quad (1.9)$$

and the Euler-Lagrange equations read

$$-\frac{\partial V}{\partial q_i}(\mathbf{x}) - m\frac{d^2\mathbf{x}}{dt^2} = 0, .$$

or

$$m\frac{d^2\mathbf{x}_0}{dt^2} = -\text{grad}V(\mathbf{x}) = \mathbf{F}(\mathbf{x}).$$

We can rewrite the previous equation as follows. Notice that it follows from (1.9) that

$$m\dot{q}_i = \frac{\partial L}{\partial \dot{q}_i}.$$

Hence if we set $p_i = \frac{\partial L}{\partial \dot{q}_i}$ and introduce the *Hamiltonian* function

$$H(\mathbf{p}, \mathbf{q}) = \sum_i p_i \dot{q}_i - L(\mathbf{q}, \dot{\mathbf{q}}), \quad (1.10)$$

then, using (1.8), we get

$$\dot{p}_i := \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}. \quad (1.11)$$

They are called *Hamilton's equations* for Newton's law.

1.3 Let us generalize equations (1.11) to an arbitrary Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}})$ on $\mathbb{R}^{2n}$. We view $\mathbb{R}^{2n}$ as the tangent bundle $T(\mathbb{R}^n)$ to $\mathbb{R}^n$. A path $\gamma : [t_1, t_2] \to \mathbb{R}^n$ extends to a path $\dot{\gamma} : [t_1, t_2] \to T(\mathbb{R}^n)$ by the formula

$$\dot{\gamma}(t) = (\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

We use $\mathbf{q} = (q_1, \ldots, q_n)$ as coordinates in the base $\mathbb{R}^n$ and $\dot{\mathbf{q}} = (\dot{q}_1, \ldots, \dot{q}_n)$ as coordinates in the fibres of $T(\mathbb{R}^n)$. The function L becomes a function on the tangent bundle. Let us see that L (under certain assumptions) can be transformed to a function on the cotangent space $T(\mathbb{R}^n)^*$. This will be taken as the analog of the Hamiltonian function (1.10). The corresponding transformation from L to H is called the *Legendre transformation.*

We start with a simple example. Let

$$Q = \sum_{i=1}^{n} a_{ii}x_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}x_ix_j$$

be a quadratic form on $\mathbb{R}^n$. Then $\frac{\partial Q}{\partial x_i} = 2\sum_{j=1}^n a_{ij}x_j$ is a linear form on $\mathbb{R}^n$. The form Q is non-degenerate if and only if the set of its partials forms a basis in the space $(\mathbb{R}^n)^*$ of linear functions on $\mathbb{R}^n$. In this case we can express each x_i as a linear expression in $p_j = \frac{\partial Q}{\partial x_j}, j = 1, \ldots, n$. Plugging in, we get a function on $(\mathbb{R}^n)^*$:

$$Q^*(p_1, \ldots, p_n) = Q(x_1(p_1, \ldots, p_n), \ldots, x_n(p_1, \ldots, p_n)).$$

For example, if $n = 1$ and $Q = x^2$ we get $p = 2x, x = p/2$ and $Q^*(p) = p^2/4$. In coordinate-free terms, let $b : V \times V \to K$ be the polar bilinear form $b(x, y) = Q(x + y) - Q(x) - Q(y)$ associated to a quadratic form $Q : V \to K$ on a vector space over a field K of characteristic 0. We can identify it with the canonical linear map $b : V \to V^*$ which is bijective if Q is non-degenerate. Let $b^{-1} : V^* \to V$ be the inverse map. It is the bilinear form associated to the quadratic form Q^* on V^*.

Since $2Q(x_1, \ldots, x_n) = \sum_{i=1}^n \frac{\partial Q}{\partial x_i}x_i$, we have $Q(\mathbf{x}) = \sum_{i=1}^n \frac{\partial Q}{\partial x_i}x_i - Q(\mathbf{x})$. Hence

$$Q^*(p_1, \ldots, p_n) = \sum_{i=1}^{n} p_ix_i - Q(\mathbf{x}(\mathbf{p})) = \mathbf{p} \bullet \mathbf{x} - Q(\mathbf{x}(\mathbf{p})).$$

Now let $f(x_1, \ldots, x_n)$ be any smooth function on $\mathbb{R}^n$ whose second differential is a positive definite quadratic form. Let $F(\mathbf{p}, \mathbf{x}) = \mathbf{p} \bullet \mathbf{x} - f(\mathbf{x})$. The *Legendre transform* of f is defined by the formula:

$$f^*(\mathbf{p}) = \max_{\mathbf{x}} F(\mathbf{p}, \mathbf{x}).$$

To find the value of $f^*(\mathbf{p})$ we take the partial derivatives of $F(\mathbf{p}, \mathbf{x})$ in $\mathbf{x}$ and eliminate x_i from the equation

$$\frac{\partial F}{\partial x_i} = p_i - \frac{\partial f}{\partial x_i} = 0.$$

This is possible because the assumption on f allows one to apply the Implicit Function Theorem. Then we plug in $x_i(\mathbf{p})$ in $F(\mathbf{p}, \mathbf{x})$ to get the value for $f^*(\mathbf{p})$. It is clear that applying this to the case when f is a non-degenerate quadratic form Q, we get the dual quadratic form Q^*.

Now let $L(\mathbf{q}, \dot{\mathbf{q}}) : T(\mathbb{R}^n) \to \mathbb{R}$ be a Lagrangian function. Assume that the second differential of its restriction $L(\mathbf{q}_0, \dot{\mathbf{q}})$ to each fibre $T(\mathbb{R}^n)_{\mathbf{q}_0} \cong \mathbb{R}^n$ is positive-definite. Then we can define its Legenedre transform $H(\mathbf{q}_0, \mathbf{p}) : T(\mathbb{R}^n)^*_{\mathbf{q}_0} \to \mathbb{R}$, where we use the coordinate

$\mathbf{p} = (p_1, \dots, p_n)$ in the fibre of the cotangent bundle. Varying $\mathbf{q}_0$ we get the *Hamiltonian function*

$$H(\mathbf{q}, \mathbf{p}) : T(\mathbb{R})^* \to \mathbb{R}.$$

By definition,

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p} \bullet \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}), \tag{1.12}$$

where $\dot{\mathbf{q}}$ is expressed via $\mathbf{p}$ and $\mathbf{q}$ from the equations $p_i = \frac{\partial L}{\partial \dot{q}_i}$. We have

$$\frac{\partial H}{\partial q_i} = \sum_{j=1}^{n} p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial L}{\partial q_i} = -\frac{\partial L}{\partial q_i};$$

$$\frac{\partial H}{\partial p_i} = \dot{q}_i + \sum_{j=1}^{n} p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_i} = \dot{q}_i.$$

Thus, the Euler-Lagrange equations

$$\frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) = 0, i = 1, \dots, n.$$

are equivalent to Hamilton's equations

$$\dot{p}_i := \frac{dp_i}{dt} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} = -\frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial q_i},$$

$$\dot{q}_i := \frac{dq_i}{dt} = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial p_i}. \tag{1.13}$$

1.4 The difference between the Euler-Lagrange equations (1.8) and Hamilton's equations (1.13) is the following. The first equation is a second order ordinary differential equation on $T(\mathbb{R}^n)$ and the second one is a first order ODE on $T(\mathbb{R}^n)^*$ which has a nice interpretation in terms of vector fields. Recall that a (smooth) *vector field* on a smooth manifold M is a (smooth) section ξ of its tangent bundle $T(M)$. For each smooth function $\phi \in C^\infty(M)$ one can differentiate ϕ along ξ by the formula

$$D_\xi(\phi)(m) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}\xi_i,$$

where $m \in M, (x_1, \dots, x_m)$ are local coordinates in a neighborhood of m, and ξ_i are the coordinates of $\xi(m) \in T(M)_m$ with respect to the basis $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ of $T(M)_m$. Given a smooth map $\gamma : [t_1, t_2] \to M$, and a vector field ξ we say that γ satisfies the differential equation defined by ξ (or is an *integral curve* of ξ) if

$$\frac{d\gamma}{dt} := (d\gamma)_\alpha(\frac{\partial}{\partial t}) = \xi(\gamma(\alpha)) \quad \text{for all } \alpha \in (t_1, t_2).$$

The vector field on the right-hand-side of Hamilton's equations (1.13) has a nice interpretation in terms of the canonical symplectic structure on the manifold $M = T(\mathbb{R}^n)^*$.

Recall that a *symplectic form* on a manifold M (as always, smooth) is a smooth closed 2-form $\omega \in \Omega^2(M)$ which is a non-degenerate bilinear form on each $T(M)_m$. Using ω one can define, for each $m \in M$, a linear isomorphism from the cotangent space $T(M)^*_m$ to the tangent space $T(M)_m$. Given a smooth function $F : M \to \mathbb{R}$, its differential dF is a 1-form on M, i.e., a section of the cotangent bundle $T(M)$. Thus, applying ω we can define a section $\imath_\omega(dF)$ of the tangent bundle, i.e., a vector field. We apply this to the situation when $M = T(\mathbb{R}^n)^*$ with coordinates $(\mathbf{q}, \mathbf{p})$ and F is the Hamiltonian function $H(\mathbf{q}, \mathbf{p})$. We use the symplectic form

$$\omega = \sum_{i=1}^{n} dq_i \wedge dp_i.$$

If we consider $(dq_1, \ldots, dq_n, dp_1, \ldots, dp_n)$ as the dual basis to the basis $(\frac{\partial}{\partial q_1}, \ldots, \frac{\partial}{\partial q_n}, \frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_n})$ of the tangent space $T(M)_m$ at some point m, then, for any $v, w \in T(M)_m$

$$\omega_m(v, w) = \sum_{i=1}^{n} dq_i(v) dp_i(w) - \sum_{i=1}^{n} dq_i(w) dp_i(v).$$

In particular,

$$\omega_m(\frac{\partial}{\partial q_i}, \frac{\partial}{\partial q_j}) = \omega_m(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}) = 0 \quad \text{for all } i, j,$$

$$\omega_m(\frac{\partial}{\partial q_i}, \frac{\partial}{\partial p_j}) = -\omega_m(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial q_j}) = \delta_{i,j}.$$

The form ω sends the tangent vector $\frac{\partial}{\partial q_i}$ to the covector dp_i and the tangent vector $\frac{\partial}{\partial p_i}$ to the covector $-dq_i$. We have

$$dH = \sum_{i=1}^{n} \frac{\partial H}{\partial q_i} dq_i + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} dp_i$$

hence

$$\imath_\omega(dH) = \sum_{i=1}^{n} \frac{\partial H}{\partial q_i}(-\frac{\partial}{\partial p_i}) + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i}.$$

So we see that the ODE corresponding to this vector field defines Hamilton's equations (1.13).

1.5 Finally let us generalize the previous Lagrangian and Hamiltonian formalism replacing the standard Euclidean space $\mathbb{R}^n$ with any smooth manifold M. Let $L : T(M) \to \mathbb{R}$ be a Lagrangian, i.e., a smooth function such that its restriction to any fibre $T(M)_m$ has positive definite second differential.

Definition. A *motion* (or *critical path*) in the Lagrange dynamical system with configuration space M and Lagrange function L is a smooth map $\gamma : [t_1, t_2] \to M$ which is a stationary point of the action functional

$$S(\gamma) = \int_{t_1}^{t_2} L(\gamma(t), \dot{\gamma}(t))dt.$$

Let us choose local coordinates $(q_1, \ldots, q_n)$ in the base M, and fibre coordinates $(\dot{q}_1, \ldots, \dot{q}_n)$. Then we obtain Lagrange's second order equations

$$\frac{\partial L}{\partial q_i}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(\gamma(t), \dot{\gamma}(t)) = 0, \quad i = 1, \ldots, n. \tag{1.14}$$

To get Hamilton's ODE equation, we need to equip $T(M)^*$ with a canonical symplectic form ω. It is defined as follows. First there is a natural 1-form α on $T(M)^*$. It is constructed as follows. Let $\pi : T(M)^* \to M$ be the bundle projection. Let $(m, \eta_m), m \in M, \eta_m \in T(M)^*_m = (T(M)_m)^*$ be a point of $T(M)^*$. The differential map

$$d\pi : T(T(M)^*)_{(m,\eta_m)} \to T(M)_m$$

sends a tangent vector v to a tangent vector ξ_m at m. If we evaluate η_m at ξ_m we get a scalar which we take as the value of our 1-form α at the vector v. In local coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ on $T(M)^*$ the expression for the form α is equal to

$$\alpha = \sum_{i=1}^{n} p_i dq_i.$$

Thus

$$\omega = -d\alpha = \sum_{i=1}^{n} dq_i \wedge dp_i$$

is a symplectic form on $T(M)^*$. Now if $H : T(M)^* \to \mathbb{R}$ is the Legendre transform of L, then $\imath_\omega(dH)$ is a vector field on $T(M)^*$. This defines Hamilton's equations on $T(M)^*$. Recall that the motion γ can be considered as a path in $T(M), t \mapsto (\gamma(t), \frac{d\gamma(t)}{dt})$, and the Lagrange function L allows us to transform it to a path on $T(M)^*$ (by expressing the fibre coordinates $\dot{q}_i$ of $T(M)$ in terms of the fibre coordinates p_i of $T(M)^*$).

1.6 A natural choice of the Lagrangian function is a metric on the manifold M. Recall that a *Riemannian metric* on M is a positive definite quadratic form $g(m) : T(M)_m \to \mathbb{R}$ which depends smoothly on m. In a basis $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ on $T(M)_m$ it is written as

$$g = \sum_{i,j=1}^{n} g_{ij} dx_i dx_j$$

where (g_{ij}) is a symmetric matrix whose entries are smooth functions on M. Changing local coordinates $(x_1, \ldots, x_n)$ to $(y_1, \ldots, y_n)$ replaces the basis $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ with $(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n})$. We have

$$\frac{\partial}{\partial y_j} = \sum_{i=1}^{n} \frac{\partial x_i}{\partial y_j} \frac{\partial}{\partial x_i}, \quad dx_i = \sum_{j=1}^{n} \frac{\partial x_i}{\partial y_j} dy_j.$$

This gives a new local description of g:

$$g = \sum_{i,j=1}^{n} g'_{ij}(m) dy_i dy_j$$

where

$$g'_{ij} = \sum_{a,b=1}^{n} g_{ab} \frac{\partial x_a}{\partial y_i} \frac{\partial x_b}{\partial y_j} = g_{ab} \frac{\partial x_a}{\partial y_i} \frac{\partial x_b}{\partial y_j}.$$

The latter expression is the "physicist's expression for summation".

One can view the Riemannian metric g as a function on the tangent bundle $T(M)$. For physicists the function $T = \frac{1}{2}g$ is the *kinetic energy* on M. A *potential energy* is a smooth function $U : M \to \mathbb{R}$. An example of a Lagrangian on M is the function

$$L = \frac{1}{2}g - U \circ \pi, \tag{1.15}$$

where $\pi : T(M) \to M$ is the bundle projection. What is its Hamiltonian function? If we change the notation for the coordinates $(x_1, \ldots, x_n), (\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ on $T(M)$ to $(q_1, \ldots, q_n), (\dot{q}_1, \ldots, \dot{q}_n)$, then $g = \sum_{i,j=1}^{n} g_{ij} \dot{q}_i \dot{q}_j$, and

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{1}{2} \frac{\partial g}{\partial \dot{q}_i} = \sum_{j=1}^{n} g_{ij} \dot{q}_j.$$

Thus

$$H = \sum_{i=1}^{n} p_i \dot{q}_i - (\frac{1}{2}g - U) = g - (\frac{1}{2}g - U) = \frac{1}{2}g + U = T + U.$$

The expression on the right-hand-side is called the *total energy*. Of course we view this function as a function on $T(M)^*$ by replacing the coordinates $\dot{q}_i$ with $p_i = \sum_{i=1}^{n} g^{ij} \dot{q}_j$, where

$$(g^{ij}) = (g_{ij})^{-1}.$$

A *Riemannian manifold* is a pair (M, g) where M is a smooth manifold and g is a metric on M. The Lagrangian of the form (1.15) is called the natural Lagrangian with potential function U. The simplest example of a Riemannian manifold is the Euclidean space $\mathbb{R}^n$ with the constant metric

$$g = m(\sum_i dx_i dx_i).$$

As a function on $T(M) = \mathbb{R}^{2n}$ it is the function $g = \sum_i \dot{q}_i^2$. Thus the corresponding natural Lagrangian is the same one as we used for Newton's law.

The critical paths for the natural Lagrangian with zero potential on a Riemannian manifold (M, g) are usually called *geodesics*. Of course, in the previous example of the Euclidean space, geodesics are straight lines.

1.7 Example. The advantage of using arbitrary manifolds instead of ordinary Euclidean space is that we can treat the cases when a particle is moving under certain constraints. For example, consider two particles in $\mathbb{R}^2$ with masses m_1 and m_2 connected by a (massless) rod of length l. Then the configuration space for this problem is the 3-fold

$$M = \{(x_1, x_2, y_1, y_2) \in \mathbb{R}^4 : (x_1 - y_1)^2 + (x_2 - y_2)^2 = l^2\}.$$

By projecting to the first two coordinates, we see that M is a smooth circle fibration with base $\mathbb{R}^2$. We can use local coordinates (x_1, x_2) on the base and the angular coordinate θ in the fibre to represent any point on $(x, y) \in M$ in the form

$$(x, y) = (x_1, x_2, x_1 + l\cos\theta, x_2 + l\sin\theta).$$

We can equip M with the Riemannian metric obtained from restriction of the Riemannian metric

$$m_1(dx_1dx_1 + dx_2dx_2) + m_2(dy_1dy_1 + dy_2dy_2)$$

on the Euclidean space $\mathbb{R}^4$ to M. Then its local expression is given by

$$g = m_1(dx_1dx_1 + dx_2dx_2) + m_2(dy_1dy_1 + dy_2dy_2) = m_1(dx_1dx_1 + dx_2dx_2) + l^2m_2d\theta d\theta.$$

The natural Lagrangian has the form

$$L(x_1(t), x_2(t), \theta(t)) = \frac{1}{2}(m_1\dot{x}_1^2 + m_1\dot{x}_2^2 + m_2l^2\dot{\theta}^2) - U(x_1, x_2, \theta).$$

The Euler-Lagrange equations are

$$m_1\frac{d^2x_i}{dt^2} = -\frac{\partial U}{\partial x_i}, \quad i = 1, 2,$$

$$m_2l^2\frac{d^2\theta}{dt} = -\frac{\partial U}{\partial\theta}.$$

Exercises.

1. Let g be a Riemannian metric on a manifold M given in some local coordinates $(x_1, \ldots, x_n)$ by a matrix $(g_{ij}(x))$. Let

$$\Gamma^i_{jk} = \sum_{l=1}^{n} \frac{g^{il}}{2}\left(\frac{\partial g_{lj}}{\partial x_k} + \frac{\partial g_{lk}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_l}\right).$$

Show the Lagrangian equation for a path $x_i = y_i(t)$ can be written in the form:

$$\frac{d^2 y_i}{dt^2} = -\sum_{j,k=1}^{n} \Gamma^i_{jk} \dot{y}_j \dot{y}_k.$$

2. Let M be the upper half-plane $\{z = x + iy \in \mathbb{C} : y > 0\}$. Consider the metric $g = y^{-2}(dxdx + dydy)$. Find the solutions of the Euler-Lagrange equation for the natural Lagrangian function with zero potential. What is the corresponding Hamiltonian function?

3. Let M be the unit sphere $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Write the Lagrangian function such that the corresponding action functional expresses the length of a path on M. Find the solutions of the Euler-Lagrange equation.

4. In the case $n = 1$, show that the Legendre transform $g(p)$ of a function $y = f(x)$ is equal to the maximal distance from the graph of $y = f(x)$ to the line $y = px$.

5. Show that the Legendre transformation of functions on $\mathbb{R}^n$ is involutive (i.e., it is the identity if repeated twice).

6. Give an example of a compact symplectic manifold.

7. A rigid rectangular triangle is rotated about its vertex at the right angle. Show that its motion is described by a motion in the configuration space M of dimension 3 which is naturally diffeomorphic to the orthogonal group $SO(3)$.

Lecture 2. CONSERVATION LAWS

2.1 Let G be a Lie group acting on a smooth manifold M. This means that there is given a smooth map $\mu : G \times M \to M, (g, x) \to g \cdot x$, such that $(g \cdot g')x = g \cdot (g' \cdot x)$ and $1 \cdot x = x$ for all $g \in G, x \in M$. For any $x \in M$ we have the map $\mu_x : G \to X$ given by the formula $\mu_x(g) = g \cdot x$. Its differential $(d\mu_x)_1 : T(G)_1 \to T(M)_x$ at the identity element $1 \in G$ defines, for any element ξ of the Lie algebra $\mathfrak{g} = T(G)_1$, the tangent vector $\xi_x^\natural \in T(M)_x$. Varying x but fixing ξ, we get a vector field $\xi^\natural \in \Theta(M)$. The map

$$\mu_* : \mathfrak{g} \to \Theta(M), \quad \xi \mapsto \xi^\natural$$

is a Lie algebra anti-homomorphism (that is, $\mu_*([\xi, \eta]) = [\mu_*(\eta), \mu_*(\xi)]$ for any $\xi, \eta \in \mathfrak{g}$). For any $\xi \in \mathfrak{g}$ the induced action of the subgroup $\exp(\mathbb{R}\xi)$ defines the action

$$\mu_\xi : \mathbb{R} \times M \to M, \mu_\xi(t, x) = \exp(t\xi) \cdot x.$$

This is called the *flow* associated to ξ.

Given any vector field $\eta \in \Theta(M)$ one may consider the differential equation

$$\dot{\gamma}(t) = \eta(\gamma(t)). \tag{2.1}$$

Here $\gamma : \mathbb{R} \to M$ is a smooth path, and $\dot{\gamma}(t) = (d\gamma)_t(\frac{\partial}{\partial t}) \in T(M)_{\gamma(t)}$. Let $\gamma_x : [-c, c] \to M$ be an integral curve of η with initial condition $\gamma_x(0) = x$. It exists and is unique for some $c > 0$. This follows from the standard theorems on existence and uniqueness of solutions of ordinary differential equations. We shall assume that this integral curve can be extended to the whole real line, i.e., $c = \infty$. This is always true if M is compact. For any $t \in \mathbb{R}$, we can consider the point $\gamma(t) \in M$ and define the map $\mu_\eta : \mathbb{R} \times M \to M$ by the formula

$$\mu_\eta(t, x) = \gamma_x(t).$$

For any $t \in \mathbb{R}$ we get a map $g_\eta^t : M \to M, x \to \mu_\eta(t, x)$. The fact that this map is a diffeomorphism follows again from ODE (a theorem on dependence of solutions on initial conditions). It is easy to check that

$$g_\eta^{t+t'} = g_\eta^t \circ g_\eta^{t'}, \quad g_\eta^{-t} = (g_\eta^t)^{-1}.$$

So we see that a vector field generates a one-parameter group of diffeomorphisms of M (called the *phase flow* of η). Of course if $\eta = \xi^\natural$ for some element ξ of the Lie algebra of a Lie group acting on M, we have

$$g^t_{\xi^\natural} = \exp(t\xi). \tag{2.2}$$

2.2. When a Lie group G acts on a manifold, it naturally acts on various geometric objects associated to M. For example, it acts on the space of functions $C^\infty(M)$ by composition: ${}^g f(x) = f(g^{-1}\cdot x)$. More generally it acts on the space of tensor fields $T^{p,q}(M)$ of arbitrary type (p,q) (p-covariant and q-contravariant) on M. In fact, by using the differential of $g \in G$ we get the map $dg : T(M) \to T(M)$ and, by transpose, the map $(dg)^* : T(M)^* \to T(M)^*$. Now it is clear how to define the image $g^*(Q)$ of any section Q of $T(M)^{\otimes p} \otimes T(M)^{*\otimes q}$. Let η be a vector field on M and let $\{g^t_\eta\}_{t\in\mathbb{R}}$ be the associated phase flow. We define the *Lie derivative* $\mathcal{L}_\eta : T^{p,q}(M) \to T^{p,q}(M)$ by the formula:

$$(\mathcal{L}_\eta Q)(x) = \lim_{t\to 0} \frac{1}{t}((g^t_\eta)^*(Q)(x) - Q(x)), \quad x \in M. \tag{2.3}$$

Explicitly, if $\eta = \sum_{i=1}^n a^i \frac{\partial}{\partial x_i}$ is a local expression for the vector field η and $Q^{i_1\ldots i_p}_{j_1\ldots j_q}$ are the coordinate functions for Q, we have

$$(\mathcal{L}_\eta Q)^{i_1\ldots i_p}_{j_1\ldots j_q} = \sum_{i=1}^n a^i \frac{\partial Q^{i_1\ldots i_p}_{j_1\ldots j_q}}{\partial x_i} - \sum_{\alpha=1}^p \sum_{i=1}^n \frac{\partial a^{i_\alpha}}{\partial x_i} Q^{i_1\ldots \hat{i}_\alpha i i_{\alpha+1}\ldots i_p}_{j_1\ldots j_q} +$$

$$+ \sum_{\beta=1}^q \sum_{j=1}^n \frac{\partial a^j}{\partial x_{j_\beta}} Q^{i_1\ldots i_p}_{j_1\ldots \hat{j}_\beta j j_{\beta+1}\ldots j_q}. \tag{2.4}$$

Note some special cases. If $p = q = 0$, i.e., $Q = Q(x)$ is a function on M, we have

$$\mathcal{L}_\eta(Q) = <\eta, dQ> = \eta(Q), \tag{2.5}$$

where we consider η as a derivation of $C^\infty(M)$. If $p = 1, q = 0$, i.e., $Q = \sum_{i=1}^n Q^i \frac{\partial}{\partial x_i}$ is a vector field, we have

$$\mathcal{L}_\eta(Q) = [\eta, Q] = \sum_{j=1}^n (\sum_{i=1}^n a^i \frac{\partial Q^j}{\partial x_i} - Q_i \frac{\partial a^j}{\partial x_i}) \frac{\partial}{\partial x_j}. \tag{2.6}$$

If $p = 0, q = 1$, i.e., $Q = \sum_{i=1}^n Q_i dx_i$ is a 1-form, we have

$$\mathcal{L}_\eta(Q) = \sum_{j=1}^n (\sum_{i=1}^n a^i \frac{\partial Q^j}{\partial x_i} + Q_i \frac{\partial a^i}{\partial x_j}) dx_j. \tag{2.7}$$

We list without proof the following properties of the Lie derivative:

(i) $\mathcal{L}_{\lambda\eta+\mu\xi}(Q) = \lambda\mathcal{L}_\eta(Q) + \mu\mathcal{L}_\xi(Q), \quad \lambda, \mu \in \mathbb{R}$;
(ii) $\mathcal{L}_{[\eta,\xi]}(Q) = [\mathcal{L}_\eta, \mathcal{L}_\xi] := \mathcal{L}_\eta \circ \mathcal{L}_\xi - \mathcal{L}_\xi \circ \mathcal{L}_\eta$;
(iii) $\mathcal{L}_\eta(Q \otimes Q') = \mathcal{L}_\eta(Q) \otimes Q' + Q \otimes \mathcal{L}_\eta(Q')$;
(iv) (Cartan's formula)

$$\mathcal{L}_\eta(\omega) = d < \eta, \omega > + < \eta, dw >, \quad \omega \in \mathcal{A}^q(M) = \Gamma(M, \bigwedge^q(T(M)^*));$$

(v) $\mathcal{L}_\eta \circ d(\omega) = d \circ \mathcal{L}_\eta(\omega), \quad \omega \in \mathcal{A}^q(M)$;
(vi) $\mathcal{L}_{f\eta}(\omega) = f\mathcal{L}_\eta(\omega) + df \wedge < \eta, \omega >, \quad f \in C^\infty(M), \omega \in \mathcal{A}^q(M)$.

Here we denote by $<,>$ the bilinear map $T^{1,0}(M) \times T^{0,q}(M) \to T^{0,q-1}(M)$ induced by the map $T(M) \times T(M)^* \to C^\infty(M)$.

Note that properties (i)-(iii) imply that the map $\eta \to \mathcal{L}_\eta$ is a homomorphism from the Lie algebra of vector fields on M to the Lie algebra of derivations of the tensor algebra $T^{*,*}(M)$.

Definition. Let $\tau \in T^{p,q}(M)$ be a smooth tensor field on M. We say that a vector field $\eta \in \Theta(M)$ is an *infinitesimal symmetry* of τ if

$$\mathcal{L}_\eta(\tau) = 0.$$

It follows from property (ii) of the Lie derivative that the set of infinitesimal symmetries of a tensor field τ form a Lie subalgebra of $\Theta(M)$.

The following proposition easily follows from the definitions.

Proposition. *Let η be a vector field and $\{g_\eta^t\}$ be the associated phase flow. Then η is an infinitesimal symmetry of τ if and only if $(g_\eta^t)^*(\tau) = \tau$ for all $t \in \mathbb{R}$.*

Example 1. Let (M, ω) be a symplectic manifold of dimension n and $H : M \to \mathbb{R}$ be a smooth function. Consider the vector field $\eta = \imath_\omega(dH)$. Then, applying Cartan's formula, we obtain

$$\mathcal{L}_\eta(\omega) = d < \eta, \omega > + < \eta, d\omega >= d < \eta, \omega >= d < \imath_\omega(dH), \omega >= d(dH) = 0.$$

This shows that the field η (called the *Hamiltonian vector field* corresponding to the function H) is an infinitesimal symmetry of the symplectic form. By the previous proposition, the associated flow of η preserves the symplectic form, i.e., is a one-parameter group of symplectic diffeomorphisms of M. It follows from property (iii) of $\mathcal{L}_\eta$ that the volume form $\omega^n = \omega \wedge \ldots \wedge \omega$ (n times) is also preserved by this flow. This fact is usually referred to as *Liouville's Theorem.*

2.3 Let $f : M \to M$ be a diffeomorphism of M. As we have already observed, it can be canonically lifted to a diffeomorphism of the tangent bundle $T(M)$. Now if η is a vector field on M, then there exists a unique lifting of η to a vector field $\tilde{\eta}$ on $T(M)$ such that its flow is the lift of the flow of η. Choose a coordinate chart U with coordinates $(q_1, \ldots, q_n)$

such that the vector field η is written in the form $\sum_{i=1}^n a_i(q)\frac{\partial}{\partial q_i}$. Let $(q_1,\ldots,q_n,\dot q_1,\ldots,\dot q_n)$ be the local coordinates in the open subset of $T(M)$ lying over U. Then local coordinates in $T(T(M))$ are $(\frac{\partial}{\partial q_1},\ldots,\frac{\partial}{\partial q_n},\frac{\partial}{\partial \dot q_1},\ldots,\frac{\partial}{\partial \dot q_n})$. Then the lift $\tilde\eta$ of η is given by the formula

$$\tilde\eta=\sum_{i=1}^n\Big(a_i\frac{\partial}{\partial q_i}+\sum_{j=1}^n\dot q_j\frac{\partial a_i}{\partial q_j}\frac{\partial}{\partial \dot q_i}\Big). \tag{2.8}$$

Later on we will give a better, more invariant, definition of $\tilde\eta$ in terms of connections on M.

Now we can define infinitesimal symmetry of any tensor field on $T(M)$, in particular of any function $F: T(M)\to\mathbb{R}$. This is the infinitesimal symmetry with respect to a vector field $\tilde\eta$, where $\eta\in\Theta(M)$. For example, we may take F to be a Lagrangian function.

Recall that the canonical symplectic form ω on $T(M)^*$ was defined as the differential of the 1-form $\alpha=-\sum_{i=1}^n p_i dq_i$. Let $L: T(M)\to\mathbb{R}$ be a Lagrangian function. The corresponding Legendre transformation is a smooth mapping: $i_L: T(M)\to T(M)^*$. Let

$$\omega_L=-i_L^*(\alpha)=\sum_{i=1}^n\frac{\partial L}{\partial \dot q_i}dq_i.$$

The next theorem is attributed to Emmy Noether:

Theorem. *If η is an infinitesimal symmetry of the Lagrangian L, then the function $\langle\tilde\eta,\omega_L\rangle$ is constant on $\dot\gamma:[a,b]\to T(M)$ for any L-critical path $\gamma:[a,b]\to M$.*

Proof. Choose a local system of coordinates $(\mathbf{q},\dot{\mathbf{q}})$ such that $\eta=\sum_{i=1}^n a_i(q)\frac{\partial}{\partial q_i}$. Then $\tilde\eta=a_i(q)\frac{\partial}{\partial q_i}+\dot q_j\frac{\partial a_i}{\partial q_j}\frac{\partial}{\partial \dot q_i}$ (using physicist's notation) and

$$\langle\omega_L,\tilde\eta\rangle=\sum_{i=1}^n a_i\frac{\partial L}{\partial \dot q_i}.$$

We want to check that the pre-image of this function under the map $\dot\gamma$ is constant. Let $\dot\gamma(t)=(\mathbf{q}(t),\dot{\mathbf{q}}(t))$. Using the Euler-Lagrange equations, we have

$$\frac{d}{dt}\sum_{i=1}^n a_i(q(t))\frac{\partial L(q(t),\dot q(t))}{\partial \dot q_i}=\sum_{i=1}^n\Big(\frac{da_i}{dt}\frac{\partial L(q(t),\dot q(t))}{\partial \dot q_i}+a_i(q(t))\frac{d}{dt}\frac{\partial L(q(t),\dot q(t))}{\partial \dot q_i}\Big)=$$

$$=\sum_{i=1}^n\Big(\sum_{j=1}^n\frac{\partial a_i}{\partial q_j}\dot q_j\frac{\partial L(q(t),\dot q(t))}{\partial \dot q_i}+a_i(q(t))\frac{\partial L(q(t),\dot q(t))}{\partial q_i}\Big).$$

Since $\tilde\eta$ is an infinitesimal symmetry of L, we have

$$0=\langle dL,\tilde\eta\rangle=\sum_{i=1}^n\Big(\sum_{j=1}^n a_i(q)\frac{\partial L}{\partial q_i}+\dot q_j\frac{\partial a_i}{\partial q_j}\frac{\partial L}{\partial \dot q_i}\Big).$$

Together with the previous equality, this proves the assertion.

Noether's Theorem establishes a very important principle:

Infinitesimal symmetry $\Longrightarrow$ Conservation Law.

By *conservation law* we mean a function which is constant on critical paths of a Lagrangian. Not every conservation law is obtained in this way. For example, we have for any critical path $\gamma(t) = (q(t), \dot{q}(t))$

$$\frac{dL}{dt} = \frac{dq_i}{dt}\frac{\partial L}{\partial q_i} + \frac{d\dot{q}_i}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{d}{dt}\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}.$$

This shows that the function

$$E = \dot{q}_i\frac{\partial L}{\partial \dot{q}_i} - L$$

(*total energy*) is always a conservation law. However, if we replace our configuration space M with the space-time space $M\times\mathbb{R}$, and postulate that L is independent of time, we obtain the previous conservation law as a corollary of existence of the infinitesimal symmetry of L corresponding to the vector field $\frac{\partial}{\partial t}$.

Example 2. Let $M = \mathbb{R}^n$. Let G be the Lie group of Galilean transformations of M. It is the subgroup of the affine group $\mathrm{Aff}(n)$ generated by the orthogonal group $O(n)$ and the group of parallel translations $\mathbf{q} \to \mathbf{q} + \mathbf{v}$. Assume that G is the group of symmetries of L. This means that $L(\mathbf{q}, \dot{\mathbf{q}})$ is a function of $||\dot{\mathbf{q}}||$ only. Write $L = F(X)$ where $X = ||\dot{\mathbf{q}}||^2/2 = (1/2)\sum_i \dot{q}_i^2$. Lagrange's equation becomes

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{d}{dt}(\dot{q}_i\frac{dF}{dX}).$$

This implies that $\dot{q}_iF'(X) = c$ does not depend on t. Thus either $\dot{q}_i = 0$ which means that there is no motion, or $F'(X) = c/\dot{q}_i$ is independent of $\dot{q}_j, j \neq i$. The latter obviously implies that $c = F'(X) = 0$. Thus $L = m\sum_{i=1}^n \dot{q}_i^2$ is the Lagrangian for Newton's law in absence of force. So we find that it can be obtained as a corollary of invariance of motion with respect to Galilean transformations.

Example 3. Assume that the Lagrangian $L(\mathbf{q}_1, \ldots, \mathbf{q}_N, \dot{\mathbf{q}}_1, \ldots, \dot{\mathbf{q}}_N)$ of a system of N particles in $\mathbb{R}^n$ is invariant with respect to the translation group of $\mathbb{R}^n$. This implies that the function

$$F(\mathbf{a}) = L(\mathbf{q}_1 + \mathbf{a}, \ldots, \mathbf{q}_N + \mathbf{a}, \dot{\mathbf{q}}_1, \ldots, \dot{\mathbf{q}}_N)$$

is independent of $\mathbf{a}$. Taking its derivative, we obtain

$$\sum_{i=1}^N \mathrm{grad}_{\mathbf{q}_i} L = 0.$$

The Euler-Lagrange equations imply that

$$\frac{d}{dt}\sum_{i=1}^N \mathrm{grad}_{\dot{\mathbf{q}}_i} L = \frac{d}{dt}\sum_{i=1}^N \mathbf{p}_i = 0,$$

where $\mathbf{p_i} = \mathrm{grad}_{\dot{\mathbf{q}}_i} L$ is the *linear momentum vector* of the i-th particle. This shows that the total linear momentum of the system is conserved.

Example 4. Assume that L is given by a Riemannian metric $g = (g_{ij})$. Let $\eta = a_i \frac{\partial}{\partial q_i}$ be an infinitesimal symmetry of L. According to Noether's theorem and Conservation of Energy, the functions $E_L = g$ and $\omega_L(\tilde{\eta}) = 2\sum_{i,j} g_{ij} a_i \dot{q}_j$ are constants on geodesics. For example, take M to be the upper half plane with metric $y^{-2}(dx^2 + dy^2)$, and let $\eta \in \mathfrak{g} = Lie(G)$ where $G = PSL(2,\mathbb{R})$ is the group of Moebius transformations $z \to \frac{az+b}{cz+d}$, $z = x + iy$. The Lie algebra of G is isomorphic to the Lie algebra of 2×2 real matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with trace zero. View η as a derivation of the space $C^\infty(G)$ of functions on G. Let $C^\infty(M) \to C^\infty(G)$ be the map $f \to f(g\cdot x)$ (x is a fixed point of M). Then the composition with the derivation η gives a derivation of $C^\infty(M)$. This is the value $\eta^\natural_x$ of the vector field $\eta^\natural$ at x. Choose $\frac{\partial}{\partial a} - \frac{\partial}{\partial d}, \frac{\partial}{\partial b}, \frac{\partial}{\partial c}$ as a basis of $\mathfrak{g}$. Then simple computations give

$$(\frac{\partial}{\partial a})^\natural(f(z)) = \frac{\partial}{\partial a} f(\frac{az+b}{cz+d})|_{(a,b,c,d)=(1,0,0,1)} = \frac{\partial f}{\partial x} x,$$

$$(\frac{\partial}{\partial b})^\natural(f(z)) = \frac{\partial}{\partial b} f(\frac{az+b}{cz+d})|_{(a,b,c,d)=(1,0,0,1)} = \frac{\partial f}{\partial x},$$

$$(\frac{\partial}{\partial c})^\natural(f(z)) = \frac{\partial}{\partial c} f(\frac{az+b}{cz+d})|_{(a,b,c,d)=(1,0,0,1)} = (y^2 - x^2)\frac{\partial f}{\partial x} - 2xy\frac{\partial f}{\partial y},$$

$$(\frac{\partial}{\partial d})^\natural(f(z)) = \frac{\partial}{\partial d} f(\frac{az+b}{cz+d})|_{(a,b,c,d)=(1,0,0,1)} = -x\frac{\partial f}{\partial x} - 2y\frac{\partial f}{\partial y}.$$

Thus we obtain that the image of $\mathfrak{g}$ in $\Theta(M)$ is the Lie subalgebra generated by the following three vector fields:

$$\eta_1 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y},\ \eta_2 = (y^2 - x^2)\frac{\partial}{\partial x} - 2xy\frac{\partial}{\partial y},\ \eta_3 = \frac{\partial}{\partial x}$$

They are lifted to the following vector fields on $T(M)$:

$$\tilde{\eta}_1 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + \dot{x}\frac{\partial}{\partial \dot{x}} + \dot{y}\frac{\partial}{\partial \dot{y}},$$

$$\tilde{\eta}_2 = (y^2 - x^2)\frac{\partial}{\partial x} - 2xy\frac{\partial}{\partial y} + (2y\dot{y} - 2x\dot{x})\frac{\partial}{\partial \dot{x}} - (2y\dot{x} + 2x\dot{y})\frac{\partial}{\partial \dot{y}};$$

$$\tilde{\eta}_3 = \frac{\partial}{\partial x}.$$

If we take for the Lagrangian the metric function $L = y^{-2}(\dot{x}^2 + \dot{y}^2)$ we get

$$\tilde{\eta}_1(L) = \langle \tilde{\eta}_1, dL \rangle = -2y^{-3}y(\dot{x}^2 + \dot{y}^2) + 2y^{-2}\dot{x}^2 + 2y^{-2}\dot{y}^2 = 0,$$

$$\tilde{\eta}_2(L) = 4xL + (2y\dot{y} - 2x\dot{x})y^{-2}2\dot{x} - (2y\dot{x} + 2x\dot{y})y^{-2}2\dot{y} = 0,$$

$$\tilde{\eta}_3(L) = 0.$$

Thus each $\eta = a\eta_1 + b\eta_2 + c\eta_3$ is an infinitesimal symmetry of L. We have

$$\omega_L = \frac{\partial L}{\partial \dot{x}}dx + \frac{\partial L}{\partial \dot{y}}dy = 2y^{-2}(\dot{x}dx + \dot{y}dy).$$

This gives us the following 3-parameter family of conservation laws:

$$\lambda_1\omega_L(\tilde{\eta}_1) + \lambda_2\omega_L(\tilde{\eta}_2) + \lambda_3\omega_L(\tilde{\eta}_3) = 2y^{-2}[\lambda_1(x\dot{x} + y\dot{y}) + \lambda_2(y^2\dot{x} - x^2\dot{x} - 2xy\dot{y}) + \lambda_3\dot{x}].$$

2.4 Let $\Psi : T(M) \to \mathbb{R}$ be any conservation law with respect to a Lagrangian function L. Recall that this means that for any L-critical path $\gamma : [a,b] \to M$ we have $d\Psi(\gamma(t), \dot{\gamma}(t))/dt = 0$. Using the Legendre transformation $i_L : T(M) \to T(M)^*$ we can view γ as a path $\mathbf{q} = \mathbf{q}(t), \mathbf{p} = \mathbf{p}(t)$ in $T(M)^*$. If i_L is a diffeomorphism (in this case the Lagrangian is called *regular*) we can view Ψ as a function $\Psi(\mathbf{q},\mathbf{p})$ on $T(M)^*$. Then, applying Hamilton's equations we have

$$\frac{d\Psi(\mathbf{q}(t),\mathbf{p}(t))}{dt} = \sum_{i=1}^{n} \frac{\partial \Psi}{\partial q_i}\dot{q}_i + \frac{\partial \Psi}{\partial p_i}\dot{p}_i = \sum_{i=1}^{n} \frac{\partial \Psi}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial \Psi}{\partial p_i}\frac{\partial H}{\partial q_i},$$

where H is the Hamiltonian function.

For any two functions F, G on $T(M)^*$ we define the *Poisson bracket*

$$\{F,G\} = \sum_{i=1}^{n} \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}. \tag{2.9}$$

It satisfies the following properties:

(i) $\{\lambda F + \mu G, H\} = \lambda\{F,H\} + \mu\{G,H\}, \quad \lambda,\mu \in \mathbb{R}$;
(ii) $\{F,G\} = -\{G,F\}$;
(iii) (Jacobi's identity) $\{E,\{F,G\}\} + \{F,\{G,E\}\} + \{G,\{E,F\}\} = 0$;
(iv) (Leibniz's rule) $\{F,GH\} = G\{F,H\} + \{F,G\}H$.

In particular, we see that the Poisson bracket defines the structure of a Lie algebra on the space $C^\infty(T(M)^*)$ of functions on $T(M)^*$. Also property (iv) shows that it defines derivations of the algebra $C^\infty(T(M)^*)$. An associative algebra with additional structure of a Lie algebra for which the Lie bracket acts on the product by the Leibniz rule is called a *Poisson algebra*. Thus $C^\infty(T(M)^*)$ has a natural structure of a Poisson algebra.

For the functions q_i, p_i we obviously have

$$\{q_i,q_j\} = \{p_i,p_j\} = 0, \{q_i,p_j\} = \delta_{ij}. \tag{2.10}$$

Using the Poisson bracket we can rewrite the equation for a conservation law in the form:

$$\{\Psi, H\} = 0. \tag{2.11}$$

In other words, the set of conservation laws forms the centralizer Cent(H), the subset of functions commuting with H with respect to Poisson bracket. Note that properties (iii) and (iv) of the Poisson bracket show that Cent(H) is a Poisson subalgebra of $C^\infty(T(M)^*)$.

We obviously have $\{H,H\}=0$ so that the function $\sum_{i=1}^n \dot{q}_i\frac{\partial L}{\partial \dot{q}_i} - L(\mathbf{q},\dot{\mathbf{q}})$ (from which H is obtained via the Legendre transformation) is a conservation law. Recall that when L is a natural Lagrangian on a Riemannian manifold M, we interpreted this function as the total energy. So the total energy is a conservation law; we have seen this already in section 2.3.

Recall that $T(M)^*$ is a symplectic manifold with respect to the canonical 2-form $\omega = \sum_i dp_i \wedge dq_i$. Let (X,ω) be any symplectic manifold. Then we can define the Poisson bracket on $C^\infty(X)$ by the formula:

$$\{F,G\} = \omega(\imath_\omega(dF), \imath_\omega(dG)).$$

It is easy to see that it coincides with the previous one in the case $X = T(M)^*$.

If the Lagrangian L is not regular, we should consider the Hamiltonian functions as not necessary coming from the Lagrangian. Then we can define its conservation laws as functions $F : T(M)^* \to \mathbb{R}$ commuting with H. We shall clarify this in the next Lecture. One can prove an analog of Noether's Theorem to derive conservation laws from infinitesimal symmetries of the Hamiltonian functions.

Exercises.

1. Consider the motion of a particle with mass m in a central force field, i.e., the force is given by the potential function V depending only on $r = ||\mathbf{q}||$. Show that the functions $m(q_j\dot{q}_i - q_i\dot{q}_j)$ (called *angular momenta*) are conservation laws. Prove that this implies Kepler's Law that "equal areas are swept out over equal time intervals".

2. Consider the two-body problem: the motion $\gamma : \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n$ with Lagrangian function

$$L(\mathbf{q}_1,\mathbf{q}_2,\dot{\mathbf{q}}_1,\dot{\mathbf{q}}_2) = \frac{1}{2}m_1(||\dot{\mathbf{q}}_1||^2 + ||\dot{\mathbf{q}}_2||^2 - V(||\mathbf{q}_1 - \mathbf{q}_2||^2).$$

Show that the Galilean group is the group of symmetries of L. Find the corresponding conservation laws.

3. Let $x = \cos t, y = \sin t, z = ct$ describe the motion of a particle in $\mathbb{R}^3$. Find the conservation law corresponding to the radial symmetry.

4. Verify the properties of Lie derivatives and Lie brackets listed in this Lecture.

5. Let $\{g_\eta^t\}_{t\in\mathbb{R}}$ be the phase flow on a symplectic manifold (M,ω) corresponding to a Hamiltonian vector field $\eta = \imath_\omega(dH)$ for some function $H : M \to \mathbb{R}$ (see Example 1). Let D be a bounded (in the sense of volume on M defined by ω) domain in M which is preserved under $g = g_\eta^1$. Show that for any $x \in D$ and any neighborhood U of x there exists a point $y \in U$ such that $g_\eta^n(y) \in U$ for some natural number $n > 0$.

6. State and prove an analog of Noether's theorem for the Hamiltonian function.

Lecture 3. POISSON STRUCTURES

3.1 Recall that we have already defined a Poisson structure on a associative algebra A as a Lie structure $\{a, b\}$ which satisfies the property

$$\{a, bc\} = \{a, b\}c + \{a, c\}b. \tag{3.1}$$

An associative algebra together with a Poisson structure is called a Poisson algebra. An example of a Poisson algebra is the algebra $\mathcal{O}(M)$ (note the change in notation for the ring of functions) of smooth functions on a symplectic manifold (M, ω). The Poisson structure is defined by the formula

$$\{f, g\} = \omega(\imath_\omega(df), \imath_\omega(dg)). \tag{3.2}$$

By a theorem of Darboux, one can choose local coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ in M such that ω can be written in the form

$$\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$$

(*canonical coordinates*). In these coordinates

$$\imath_\omega(df) = \imath_\omega(\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} dq_i + \sum_{i=1}^{n} \frac{\partial f}{\partial p_i} dp_i) = -\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial}{\partial p_i} + \sum_{i=1}^{n} \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q_i}.$$

This expression is sometimes called the *skew gradient*. From this we obtain the expression for the Poisson bracket in canonical coordinates:

$$\{f, g\} = \omega(\sum_{i=1}^{n} \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial}{\partial p_i}, \sum_{i=1}^{n} \frac{\partial g}{\partial p_i} \frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial g}{\partial q_i} \frac{\partial}{\partial p_i}) =$$

$$= \sum_{i=1}^{n} (\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}). \tag{3.3}$$

Definition. A *Poisson manifold* is a smooth manifold M together with a Poisson structure on its ring of smooth functions $\mathcal{O}(M)$.

An example of a Poisson manifold is a symplectic manifold with the Poisson structure defined by formula (3.2).

Let $\{,\}$ define a Poisson structure on M. For any $f \in \mathcal{O}(M)$ the formula $g \mapsto \{f, g\}$ defines a derivation of $\mathcal{O}(M)$, hence a vector field on M. We know that a vector field on a manifold M can be defined locally, i.e., it can be defined in local charts (or in other words it is a derivation of the sheaf $\mathcal{O}_M$ of germs of smooth functions on M). This implies that the Poisson structure on M can also be defined locally. More precisely, for any open subset U it defines a Poisson structure on $\mathcal{O}(U)$ such that the restriction maps $\rho_{U,V} : \mathcal{O}(U) \to \mathcal{O}(V), V \subset U$, are homomorphisms of the Poisson structures. In other words, the structure sheaf $\mathcal{O}_M$ of M is a sheaf of Poisson algebras.

Let $i : \mathcal{O}(M) \to \Theta(M)$ be the map defined by the formula $i(f)(g) = \{f, g\}$. Since constant functions lie in the kernel of this map, we can extend it to exact differentials by setting $W(df) = i(f)$. We can also extend it to all differentials by setting $W(adf) = aW(df)$. In this way we obtain a map of vector bundles $W : T^*(M) \to T(M)$ which we shall identify with a section of $T(M) \otimes T(M)$. Such a map is sometimes called a *cosymplectic structure*. In a local chart with coordinates $x_1, \ldots, x_n$ we have

$$\{f, g\} = W(df, dg) = W(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i, \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} dx_i) =$$

$$= \sum_{i,j=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} W(dx_i, dx_j) = \sum_{i,j=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} W^{jk}. \tag{3.4}$$

Conversely, given W as above, we can define $\{f, g\}$ by the formula

$$\{f, g\} = W(df, dg).$$

The Leibniz rule holds automatically. The anti-commutativity condition is satisfied if we require that W is anti-symmetric, i.e., a section of $\bigwedge^2(T(M)$. Locally this means that

$$W^{jk} = -W^{kj}.$$

The Jacobi identity is a more serious requirement. Locally it means that

$$\sum_{k=1}^{n} (W^{jk} \frac{\partial W^{lm}}{\partial x_k} + W^{lk} \frac{\partial W^{mj}}{\partial x_k} + W^{mk} \frac{\partial W^{jl}}{\partial x_k}) = 0. \tag{3.5}$$

Examples 1. Let (M, ω) be a symplectic manifold. Then $\omega \in \bigwedge^2(T(M)^*)$ can be thought as a bundle map $T(M) \to T(M)^*$. Since ω is non-degenerate at any point $x \in M$, we can invert this map to obtain the map $\omega^{-1} : T(M)^* \to T(M)$. This can be taken as the cosymplectic structure W.

2. Assume the functions $W^{jk}(x)$ do not depend on x. Then after a linear change of local parameters we obtain

$$(W^{jk}) = \begin{pmatrix} 0_k & -I_k & 0_{n-2k} \\ I_k & 0_k & 0_{n-2k} \\ 0_{n-2k} & 0_{n-2k} & 0_{n-2k} \end{pmatrix},$$

where I_k is the identity matrix of order k and 0_m is the zero matrix of order m. If $n = 2k$, we get a symplectic structure on M and the associated Poisson structure.

3. Assume $M = \mathbb{R}^n$ and

$$W^{jk} = \sum_{i=1}^{n} C^i_{jk} x_i$$

are given by linear functions. Then the properties of W^{jk} imply that the scalars C^i_{jk} are structure constants of a Lie algebra structure on $(\mathbb{R}^n)^*$ (the Jacobi identity is equivalent to (3.5)). Its Lie bracket is defined by the formula

$$[x_j, x_k] = \sum_{i=1}^{n} C^i_{jk} x_i.$$

For this reason, this Poisson structure is called the *Lie-Poisson structure*. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{g}^*$ be the dual vector space. By choosing a basis of $\mathfrak{g}$, the corresponding structure constants will define the Poisson bracket on $\mathcal{O}(\mathfrak{g}^*)$. In more invariant terms we can describe the Poisson bracket on $\mathcal{O}(\mathfrak{g}^*)$ as follows. Let us identify the tangent space of $\mathfrak{g}^*$ at 0 with $\mathfrak{g}^*$. For any $f : \mathfrak{g}^* \to \mathbb{R}$ from $\mathcal{O}(\mathfrak{g}^*)$, its differential df_0 at 0 is a linear function on $\mathfrak{g}^*$, and hence an element of $\mathfrak{g}$. It is easy to verify now that for linear functions f, g,

$$\{f, g\} = [df_0, dg_0].$$

Observe that the subspace of $\mathcal{O}(\mathfrak{g}^*)$ which consists of linear functions can be identified with the dual space $(\mathfrak{g}^*)^*$ and hence with $\mathfrak{g}$. Restricting the Poisson structure to this subspace we get the Lie structure on $\mathfrak{g}$. Thus the Lie-Poisson structure on $M = \mathfrak{g}^*$ extends the Lie structure on $\mathfrak{g}$ to the algebra of all smooth functions on $\mathfrak{g}^*$. Also note that the subalgebra of polynomial functions on $\mathfrak{g}^*$ is a Poisson subalgebra of $\mathcal{O}(\mathfrak{g}^*)$. We can identify it with the symmetric algebra $\mathrm{Sym}(\mathfrak{g})$ of $\mathfrak{g}$.

3.2 One can prove an analog of Darboux's theorem for Poisson manifolds. It says that locally in a neighborhood of a point $m \in M$ one can find a coordinate system $q_1, \ldots, q_k$, $p_1, \ldots, p_k, x_{2k+1}, \ldots, x_n$ such that the Poisson bracket can be given in the form:

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}\right).$$

This means that the submanifolds of M given locally by the equations $x_{2k+1} = c_1, \ldots, x_n = c_n$ are symplectic manifolds with the associated Poisson structure equal to the restriction of the Poisson structure of M to the submanifold. The number $2k$ is called the *rank* of the

Poisson manifold at the point P. Thus any Poisson manifold is equal to the union of its symplectic submanifolds (called *symplectic leaves*). If the Poisson structure is given by a cosymplectic structure $W : T(M)^* \to T(M)$, then it is easy to check that the rank $2k$ at $m \in M$ coincides with the rank of the linear map $W_m : T(M)^*_m \to T(M)_m$. Note that the rank of a skew-symmetric matrix is always even.

Example 4. Let G be a Lie group and $\mathfrak{g}$ be its Lie algebra. Consider the Poisson-Lie structure on $\mathfrak{g}^*$ from Example 3. Then the symplectic leaves of $\mathfrak{g}^*$ are coadjoint orbits. Recall that G acts on itself by conjugation $C_g(g') = g \cdot g' \cdot g^{-1}$. By taking the differential we get the adjoint action of G on $\mathfrak{g}$:

$$Ad_g(\eta) = (dC_{g^{-1}})_1(\eta), \quad \eta \in \mathfrak{g} = T(G)_1.$$

By transposing the linear maps $Ad_g : \mathfrak{g} \to \mathfrak{g}$ we get the coadjoint action

$$Ad_g^*(\phi)(\eta) = \phi(Ad_g(\eta)), \quad \phi \in \mathfrak{g}^*, \eta \in \mathfrak{g}.$$

Let O_ϕ be the coadjoint orbit of a linear function ϕ on $\mathfrak{g}$. We define the symplectic form ω_ϕ on it as follows. First of all let $G_\phi = \{g \in G : Ad_g^*(\phi) = \phi\}$ be the stabilizer of ϕ and let $\mathfrak{g}_\phi$ be its Lie algebra. If we identify $T(\mathfrak{g}^*)_\phi$ with $\mathfrak{g}^*$ then $T(O_\phi)_\phi$ can be canonically identified with $\mathrm{Lie}(G_\phi)^\perp \subset \mathfrak{g}^*$. For any $v \in \mathfrak{g}$ let $v^\natural \in \mathfrak{g}^*$ be defined by $v^\natural(w) = \phi([v, w])$. Then the map $v \to v^\natural$ is an isomorphism from $\mathfrak{g}/\mathfrak{g}_\phi$ onto $\mathfrak{g}_\phi^\perp = T(O_\phi)_\phi$. The value $\omega_\phi(\phi)$ of ω_ϕ on $T(O_\phi)_\phi$ is computed by the formula

$$\omega_\phi(v^\natural, w^\natural) = \phi([v, w]).$$

This is a well-defined non-degenerate skew-symmetric bilinear form on $T(O_\phi)_\phi$. Now for any $\psi = Ad_g^*(\phi)$ in the orbit O_ϕ, the value $\omega_\phi(\psi)$ of $\omega + \phi$ at ψ is a skew-symmetric bilinear form on $T(O_\phi)_\psi$ obtained from $\omega_\phi(\phi)$ via the differential of the translation isomorphism $Ad_g^*(v) : \mathfrak{g}^* \to \mathfrak{g}^*$. We leave to the reader to verify that the form ω_ϕ on O_ϕ obtained in this way is closed, and that the Poisson structure on the coadjoint orbit is the restriction of the Poisson structure on $\mathfrak{g}^*$. It is now clear that symplectic leaves of $\mathfrak{g}^*$ are the coadjoint orbits.

3.3 Fix a smooth function $H : M \to \mathbb{R}$ on a Poisson manifold M. It defines the vector field $f \to \{f, H\}$ (called the *Hamiltonian vector field* with respect to the function H) and hence the corresponding differential equation (=dynamical system, flow)

$$\frac{df}{dt} = \{f, H\}. \tag{3.6}$$

A solution of this equation is a function f such that for any path $\gamma : [a, b] \to M$ we have

$$\frac{df(\gamma(t))}{dt} = \{f, H\}(\gamma(t)).$$

This is called the *Hamiltonian dynamical system* on M (with respect to the Hamiltonian function H). If we take f to be coordinate functions on M, we obtain Hamilton's equations for the critical path $\gamma : [a, b] \to M$ in M. As before, the conservation laws here are the functions f satisfying $\{f, H\} = 0$.

The flow g^t of the vector field $f \to \{f, H\}$ is a one-parameter group of operators U_t on $\mathcal{O}(M)$ defined by the formula

$$U_t(f) = f_t := (g^t)^*(f) = f \circ g^{-t}.$$

If $\gamma(t)$ is an integral curve of the Hamiltonian vector field $f \to \{f, H\}$ with $\gamma(0) = x$, then

$$U_t(f) = f_t(x) = f(\gamma(t))$$

and the equation for the Hamiltonian dynamical system defined by H is

$$\frac{df_t}{dt} = \{f_t, H\}.$$

Here we use the Poisson bracket defined by the symplectic form of M.

Theorem 1. *The operators U_t are automorphisms of the Poisson algebra $\mathcal{O}(M)$.*

Proof. We have to verify only that U_t preserves the Poisson bracket. The rest of the properties are obvious. We have

$$\frac{d}{dt}\{f, g\}_t = \{\frac{df_t}{dt}, g_t\} + \{f_t, \frac{dg_t}{dt}\} = \{\{f_t, H\}, g_t\} + \{f_t, \{g_t, H\}\} = \{\{f_t, g_t\}, H\} =$$

$$= \frac{d}{dt}\{f_t, g_t\}.$$

Here we have used property (iv) of the Poisson bracket from Lecture 2. This shows that the functions $\{f, g\}_t$ and $\{f_t, g_t\}$ differ by a constant. Taking $t = 0$, we obtain that this constant must be zero.

Example 5. Many classical dynamical systems can be given on the space $M = \mathfrak{g}^*$ as in Example 3. For example, take $\mathfrak{g} = \mathfrak{so}(3)$, the Lie algebra of the orthogonal group SO(3). One can identify $\mathfrak{g}$ with the algebra of skew-symmetric 3×3 matrices, with the Lie bracket defined by $[A, B] = AB - BA$. If we write

$$A(\mathbf{x}) = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}, \quad \mathbf{x} = (x_1, x_2, x_3), \tag{3.7}$$

we obtain

$$[A(\mathbf{x}), A(\mathbf{v})] = \mathbf{x} \times \mathbf{v}, \quad A(\mathbf{x}) \cdot \mathbf{v} = \mathbf{x} \times \mathbf{v},$$

where $\times$ denotes the cross-product in $\mathbb{R}^3$. Take

$$e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

as a basis in $\mathfrak{g}$ and x_1, x_2, x_3 as the dual basis in $\mathfrak{g}^*$. Computing the structure constants we find

$$\{x_j, x_k\} = \sum_{i=1}^{n} \epsilon_{jki} x_i,$$

where ϵ_{jki} is totally skew-symmetric with respect to its indices and $\epsilon_{123} = 1$. This implies that for any $f, g \in \mathcal{O}(M)$ we have, in view of (3.4),

$$\{f,g\}(x_1,x_2,x_3) = (x_1,x_2,x_3) \bullet (\nabla f \times \nabla g) = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_3} \\ \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \frac{\partial g}{\partial x_3} \end{pmatrix}. \tag{3.8}$$

By definition, a *rigid body* is a subset B of $\mathbb{R}^3$ such that the distance between any two of its points does not change with time. The motion of any point is given by a path $Q(t)$ in $SO(3)$. If $\mathbf{b} \in B$ is fixed, its image at time t is $\mathbf{v} = Q(t) \cdot \mathbf{b}$. Suppose the rigid body is moving with one point fixed at the origin. At each moment of time t there exists a direction $\mathbf{x} = (x_1, x_2, x_3)$ such that the point $\mathbf{v}$ rotates in the plane perpendicular to this direction and we have

$$\dot{\mathbf{v}} = \mathbf{x} \times \mathbf{v} = A(\mathbf{x})\mathbf{v} = \dot{Q}(t) \cdot \mathbf{b} = \dot{Q}(t)Q^{-1}\mathbf{v}, \tag{3.9}$$

where $A(\mathbf{x})$ is a skew-symmetric matrix as in (3.7). The vector $\mathbf{x}$ is called the *spatial angular velocity.* The *body angular velocity* is the vector $\mathbf{X} = Q(t)^{-1}\mathbf{x}$. Observe that the vectors $\mathbf{x}$ and $\mathbf{X}$ do not depend on $\mathbf{b}$. The kinetic energy of the point $\mathbf{v}$ is defined by the usual function

$$K(\mathbf{b}) = \frac{1}{2} m||\dot{\mathbf{v}}||^2 = m||\mathbf{x} \times \mathbf{v}||^2 = m||Q^{-1}(\mathbf{x} \times \mathbf{v})|| = m||\mathbf{X} \times \mathbf{b}||^2 = X\Pi(\mathbf{b})X,$$

for some positive definite matrix $\Pi(\mathbf{b})$ depending on $\mathbf{b}$. Here m denotes the mass of the point $\mathbf{b}$. The total kinetic energy of the rigid body is defined by integrating this function over B:

$$K = \frac{1}{2} \int_B \rho(\mathbf{b}) ||\dot{\mathbf{v}}||^2 d^3\mathbf{b} = \mathbf{X}\Pi\mathbf{X},$$

where $\rho(\mathbf{b})$ is the density function of B and Π is a positive definite symmetric matrix. We define the *spatial angular momentum* (of the point $\mathbf{v}(t) = Q(t)\mathbf{b}$) by the formula

$$\mathbf{m} = m\mathbf{v} \times \dot{\mathbf{v}} = m\mathbf{v} \times (\mathbf{x} \times \mathbf{v}).$$

We have

$$\mathbf{M}(\mathbf{b}) = Q^{-1}\mathbf{m} = mQ^{-1}\mathbf{v} \times (Q^{-1}\mathbf{x} \times Q^{-1}\mathbf{v}) = m\mathbf{b} \times (\mathbf{X} \times \mathbf{b}) = \Pi(\mathbf{b})\mathbf{X}.$$

After integrating $\mathbf{M}(\mathbf{b})$ over B we obtain the vector $\mathbf{M}$, called the *body angular momentum.* We have

$$\mathbf{M} = \Pi\mathbf{X}.$$

If we consider the Euler-Lagrange equations for the motion $\mathbf{v}(t) = Q(t)\mathbf{b}$, and take the Lagrangian function defined by the kinetic energy K, then Noether's theorem will give us that the angular momentum vector $\mathbf{m}$ does not depend on t (see Problem 1 from Lecture 1). Therefore

$$\dot{\mathbf{m}} = \frac{d(Q(t)\cdot\mathbf{M}(\mathbf{b}))}{dt} = Q\cdot\dot{\mathbf{M}}(\mathbf{b})+\dot{Q}\cdot\mathbf{M}(\mathbf{b}) = Q\cdot\dot{\mathbf{M}}(\mathbf{b})+\dot{Q}Q^{-1}\cdot\mathbf{m} = Q\cdot\dot{\mathbf{M}}(\mathbf{b})+A(\mathbf{x})\cdot\mathbf{m} =$$

$$= Q\cdot\dot{\mathbf{M}}(\mathbf{b}) + \mathbf{x}\times\mathbf{m} = Q\cdot\dot{\mathbf{M}}(\mathbf{b}) + Q\mathbf{X}\times Q\mathbf{M}(\mathbf{b}) = Q(\dot{\mathbf{M}}(\mathbf{b}) + \mathbf{X}\times\mathbf{M}(\mathbf{b})) = 0.$$

After integrating over B we obtain the Euler equation for the motion of a rigid body about a fixed point:

$$\dot{\mathbf{M}} = \mathbf{M}\times\mathbf{X}. \tag{3.10}$$

Let us view $\mathbf{M}$ as a point of $\mathbb{R}^3$ identified with $\mathfrak{so}(3)^*$. Consider the Hamiltonian function

$$H(\mathbf{M}) = \frac{1}{2}\mathbf{M}\cdot(\Pi^{-1}\mathbf{M}).$$

Choose an orthogonal change of coordinates in $\mathbb{R}^3$ such that Π is diagonalized, i.e.,

$$\mathbf{M} = \Pi\mathbf{X} = (I_1X_1, I_2X_2, I_3X_3)$$

for some positive scalars I_1, I_2, I_3 (the *moments of inertia* of the rigid body B). In this case the Euler equation looks like

$$I_1\dot{\mathbf{X}}_1 = (I_2 - I_1)X_2X_3,$$

$$I_2\dot{\mathbf{X}}_2 = (I_3 - I_1)X_1X_3,$$

$$I_3\dot{\mathbf{X}}_3 = (I_1 - I_2)X_1X_2.$$

We have

$$H(\mathbf{M}) = \frac{1}{2}(\frac{M_1^2}{I_1} + \frac{M_2^2}{I_2} + \frac{M_3^2}{I_3})$$

and, by formula (3.8),

$$\{f(\mathbf{M}), H(\mathbf{M})\} = \mathbf{M}\cdot(\nabla f\times\nabla H) = \mathbf{M}\cdot(\nabla f\times(\frac{M_1}{I_1}, \frac{M_2}{I_2}, \frac{M_3}{I_3})) = \mathbf{M}\cdot(\nabla f\times(X_1, X_2, X_3)).$$

If we take for f the coordinate functions $f(\mathbf{M}) = M_i$, we obtain that Euler's equation (3.10) is equivalent to Hamilton's equation

$$\dot{M}_i = \{M_i, H\}, i = 1, 2, 3.$$

Also, if we take for f the function $f(\mathbf{M}) = ||\mathbf{M}||^2$, we obtain $\{f, H\} = \mathbf{M}\cdot(\mathbf{M}\times\mathbf{X}) = 0$. This shows that the integral curves of the motion of the rigid body B are contained in the level sets of the function $||\mathbf{M}||^2$. These sets are of course coadjoint orbits of $SO(3)$ in $\mathbb{R}^3$.

Also observe that the Hamiltonian function H is constant on integral curves. Hence each integral curve is contained in the intersection of two quadrics (an ellipsoid and a sphere)

$$\frac{1}{2}(\frac{M_1^2}{I_1}+\frac{M_2^2}{I_2}+\frac{M_3^2}{I_3})=c_1,$$

$$\frac{1}{2}(M_1^2+M_2^2+M_3^2)=c_2.$$

Notice that this is also the set of real points of the quartic elliptic curve in $\mathbb{P}^3(\mathbb{C})$ given by the equations

$$\frac{1}{I_1}T_1^2+\frac{1}{I_2}T_2^2+\frac{1}{I_3}T_3^2-2c_1T_0^2=0,$$

$$T_1^2+T_2^2+T_3^2-2c_2T_0^2=0.$$

The structure of this set depends on I_1, I_2, I_3. For example, if $I_1 \geq I_2 \geq I_3$, and $c_1 < c_2/I_3$ or $c_1 > c_2/I_3$, then the intersection is empty.

3.4 Let G be a Lie group and $\mu : G\times G \to G$ its group law. This defines the *comultiplication map*

$$\Delta : \mathcal{O}(G) \to \mathcal{O}(G\times G).$$

Any Poisson structure on a manifold M defines a natural Poisson structure on the product $M\times M$ by the formula

$$\{f(x,y),g(x,y)\}=\{f(x,y),g(x,y)\}_x+\{f(x,y),g(x,y)\}_y,$$

where the subscript indicates that we apply the Poisson bracket with the variable in the subscript being fixed.

Definition. A *Poisson-Lie group* is a Lie group together with a Poisson structure such that the comultiplication map is a homomorphism of Poisson algebras.

Of course any abelian Lie group is a Poisson-Lie group with respect to the trivial Poisson structure. In the work of Drinfeld quantum groups arose in the attempt to classify Poisson-Lie groups.

Given a Poisson-Lie group G, let $\mathfrak{g} = Lie(G)$ be its Lie algebra. By taking the differentials of the Poisson bracket on $\mathcal{O}(G)$ at 1 we obtain a Lie bracket on $\mathfrak{g}^*$. Let $\phi : \mathfrak{g} \to \mathfrak{g}\otimes\mathfrak{g}$ be the linear map which is the transpose of the linear map $\mathfrak{g}^*\otimes\mathfrak{g}^* \to \mathfrak{g}^*$ given by the Lie bracket. Then the condition that $\Delta : \mathcal{O}(G) \to \mathcal{O}(G\times G)$ is a homomorphism of Poisson algebras is equivalent to the condition that ϕ is a 1-cocycle with respect to the action of $\mathfrak{g}$ on $\mathfrak{g}\otimes\mathfrak{g}$ defined by the adjoint action. Recall that this means that

$$\phi([v,w]) = ad(v)(\phi(w)) - ad(w)(\phi(v)),$$

where $ad(v)(x\otimes y) = [v,x]\otimes[v,y]$.

Definition. A Lie algebra $\mathfrak{g}$ is called a *Lie bialgebra* if it is additionally equipped with a linear map $\mathfrak{g} \to \mathfrak{g}\otimes\mathfrak{g}$ which defines a 1-cocycle of $\mathfrak{g}$ with coefficients in $\mathfrak{g}\otimes\mathfrak{g}$ with respect to the adjoint action and whose transpose $\mathfrak{g}^*\otimes\mathfrak{g}^* \to \mathfrak{g}^*$ is a Lie bracket on $\mathfrak{g}^*$.

Theorem 2. *(V. Drinfeld). The category of connected and simply-connected Poisson-Lie groups is equivalent to the category of finite-dimensional Lie bialgebras.*

To define a structure of a Lie bialgebra on a Lie algebra $\mathfrak{g}$ one may take a cohomologically trivial cocycle. This is an element $r \in \mathfrak{g} \otimes \mathfrak{g}$ defined by the formula

$$\phi(v) = ad(v)(r) \quad \text{for any } v \in \mathfrak{g}.$$

The element r must satisfy the following conditions:
(i) $r \in \bigwedge^2 \mathfrak{g}$, i.e., r is a skew-symmetric tensor;
(ii)
$$[r^{12}, r^{13}] + [r^{12}, r^{23}] + [r^{13}, r^{23}] = 0. \qquad (CYBE)$$

Here we view r as a bilinear function on $\mathfrak{g}^* \times \mathfrak{g}^*$ and denote by r^{12} the function on $\mathfrak{g}^* \times \mathfrak{g}^* \times \mathfrak{g}^*$ defined by the formula $r^{12}(x, y, z) = r(x, y)$. Similarly we define r^{ij} for any pair $1 \leq i < j \leq 3$.

The equation (CYBE) is called the *classical Yang-Baxter equation.* The Poisson bracket on the Poisson-Lie group G corresponding to the element $r \in \mathfrak{g} \otimes \mathfrak{g}$ is defined by extending r to a G-invariant cosymplectic structure $W(r) \in T(G)^* \to T(G)$.

The classical Yang-Baxter equation arises by taking the classical limit of QYBE (*quantum Yang-Baxter equation*):

$$R^{12}R^{13}R^{23} = R^{23}R^{13}R^{12},$$

where $R \in U_q(\mathfrak{g}) \otimes U_q(\mathfrak{g})$ and $U_q(\mathfrak{g})$ is a deformation of the enveloping algebra of $\mathfrak{g}$ over $\mathbb{R}[[q]]$ (quantum group). If we set

$$r = \lim_{q \to 0}(\frac{R-1}{q}),$$

then QYBE translates into CYBE.

3.5 Finally, let us discuss completely integrable systems.

We know that the integral curves of a Hamiltonian system on a Poisson manifold are contained in the level sets of conservation laws, i.e., functions F which Poisson commute with the Hamiltonian function. If there were many such functions, we could hope to determine the curves as the intersection of their level sets. This is too optimistic and one gets as the intersection of all level sets something bigger than the integral curve. In the following special case this "something bigger" looks very simple and allows one to find a simple description of the integral curves.

Definition. A *completely integrable system* is a Hamiltonian dynamical system on a symplectic manifold M of dimension $2n$ such that there exist n conservation laws $F_i : M \to \mathbb{R}, i = 1, \ldots, n$ which satisfy the following conditions:
(i) $\{F_i, F_j\} = 0, \quad i, j = 1, \ldots, n;$
(ii) the map $\pi : M \to \mathbb{R}^n, x \mapsto (F_1(x), \ldots, F_n(x))$ has no critical points (i.e., $d\pi_x : T(M)_x \to \mathbb{R}^n$ is surjective for all $x \in M$).

The following theorem is a generalization of a classical theorem of Liouville:

Theorem 3. *Let $F_1, \ldots, F_n$ define a completely integrable system on M. For any $c \in \mathbb{R}^n$ denote by M_c the fibre $\pi^{-1}(c)$ of the map $\pi : M \to \mathbb{R}^n$ given by the functions F_i. Then*

(i) if M_c is non-empty and compact, then each of its connected components is diffeomorphic to a n-dimensional torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$;

(ii) one can choose a diffeomorphism $\mathbb{R}^n/\mathbb{Z}^n \to M_c$ such that the integral curves of the Hamiltonian flow defined by $H = F_1$ in M_c are the images of straight lines in $\mathbb{R}^n$;

(iii) the restriction of the symplectic form ω of M to each M_c is trivial.

Proof. We give only a sketch, referring to [**Arnold**] for the details. First of all we observe that by (ii) the non-empty fibres $M_c = \pi^{-1}(c)$ of π are n-dimensional submanifolds of M. For every point $x \in M_c$ its tangent space is the n-dimensional subspace of $T(M)_x$ equal to the set of common zeroes of the differentials dF_i at the point x.

Let $\eta_1, \ldots, \eta_n$ be the vector fields on M defined by the functions $F_1, \ldots, F_n$ (with respect to the symplectic Poisson structure on M). Since $\{F_i, F_j\} = 0$ we get $[\eta_i, \eta_j] = 0$. Let $g_{\eta_i}^t$ be the flow corresponding to η_i. Since each integral curve is contained in some level set M_c, and the latter is compact, the flow $g_{\eta_i}^t$ is complete (i.e., defined for all t). For each $t, t' \in \mathbb{R}$ and $i, j = 1, \ldots, n$, the diffeomorphisms $g_{\eta_i}^t$ and $g_{\eta_j}^{t'}$ commute (because $[\eta_i, \eta_j] = 0$). These diffeomorphisms generate the group $G \cong \mathbb{R}^n$ acting on M. Obviously it leaves M_c invariant. Take a point $x_0 \in M_c$ and consider the map $f : \mathbb{R}^n \to M_c$ defined by the formula

$$f(t_1, \ldots, t_n) = g_{\eta_n}^{t_n} \circ \ldots \circ g_{\eta_n}^{t_1}(x_0).$$

Let G_{x_0} be the isotropy subgroup of x_0, i.e., the fibre of this map over x_0. One can show that the map f is a local diffeomorphism onto a neighborhood of x_0 in M_c. This implies easily that the subgroup G_{x_0} is discrete in $\mathbb{R}^n$ and hence $G_{x_0} = \Gamma$ where $\Gamma = \mathbb{Z}\mathbf{v}_1 + \ldots + \mathbb{Z}\mathbf{v}_n$ for some linearly independent vectors $v_1, \ldots, v_n$. Thus $\mathbb{R}^n/\Gamma$ is a compact n-dimensional torus. The map f defines an injective proper map $\bar{f} : T^n \to M_c$ which is a local diffeomorphism. Since both spaces are n-dimensional, its image is a connected component of M_c. This proves (i).

Let $\gamma : \mathbb{R} \to M_c$ be an integral curve of η_i contained in M_c. Then $\gamma(t) = g_{\eta_i}^t(\gamma(0))$. Let $(a_1, \ldots, a_n) = f^{-1}(\gamma(0)) \in \mathbb{R}^n$. Then

$$f(a_1 + t, a_2, \ldots, a_n) = g_{\eta_1}^t(f(a_1, \ldots, a_n)) = g_{\eta_1}^t(\gamma(0)) = \gamma(t).$$

Now the assertion (ii) is clear. The integral curve is the image of the line in $\mathbb{R}^n$ through the point $(a_1, \ldots, a_n)$ and $(1, 0, \ldots, 0)$.

Finally, (iii) is obvious. For any $x \in M_c$ the vectors $\eta_1(x), \ldots, \eta_n(x)$ form a basis in $T(M_c)_x$ and are orthogonal with respect to $\omega(c)$. Thus, the restriction of $\omega(c)$ to $T(M_c)_x$ is identically zero.

Definition. A submanifold N of a symplectic manifold (M, ω) is called *Lagrangian* if the restriction of ω to N is trivial and $\dim N = \frac{1}{2}\dim(M)$.

Recall that the dimension of a maximal isotropic subspace of a non-degenerate skew-symmetric bilinear form on a vector space of dimension $2n$ is equal to n. Thus the tangent space of a Lagrangian submanifold at each of its points is equal to a maximal isotropic subspace of the symplectic form at this point.

So we see that a completely integrable system on a symplectic manifold defines a fibration of M with Lagrangian fibres.

One can choose a special coordinate system (*action-angular coordinates*)

$$(I_1, \ldots, I_n, \phi_1, \ldots, \phi_n)$$

in a neighborhood of the level set M_c such that the equations of integral curves look like

$$\dot{I}_j = 0, \quad \dot{\phi}_j = \ell_j(c_1, \ldots, c_n), j = 1, \ldots, n$$

for some linear functions ℓ_j on $\mathbb{R}^n$. The coordinates ϕ_j are called *angular coordinates.* Their restriction to M_c corresponds to the angular coordinates of the torus. The functions I_j are functionally dependent on $F_1, \ldots, F_n$. The symplectic form ω can be locally written in these coordinates in the form

$$\omega = \sum_{j=1}^{n} dI_j \wedge d\phi_j.$$

There is an explicit algorithm for finding these coordinates and hence writing explicitly the solutions. We refer for the details to [**Arnold**].

3.6 Definition. An *algebraically completely integrable system* $(F_1, \ldots, F_n)$ is a completely integrable system on (M, ω) such that

(i) $M = X(\mathbb{R})$ for some algebraic variety X,

(ii) the functions F_i define a smooth morphism $\Pi : X \to \mathbb{C}^n$ whose non-empty fibres are abelian varieties (=complex tori embeddable in a projective space),

(iii) the fibres of $\pi : M \to \mathbb{R}^n$ are the real parts of the fibres of Π;

(iv) the integral curves of the Hamiltonian system defined by the function F_1 are given by meromorphic functions of time.

Example 6. Let (M, ω) be $(\mathbb{R}^2, dx \wedge dy)$ and $H : M \to \mathbb{R}$ be a polynomial of degree d whose level sets are all nonsingular. It is obviously completely integrable with $F_1 = H$. Its integral curves are connected components of the level sets $H(x, y) = c$. The Hamiltonian equation is $\dot{y} = -\frac{\partial H}{\partial x}, \dot{x} = \frac{\partial H}{\partial y}$. It is algebraically integrable only if $d = 3$. In fact in all other cases the generic fibre of the complexification is a complex affine curve of genus $g = (d-1)(d-2)/2 \neq 1$.

Example 7. Let E be an ellipsoid

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \frac{x_i^2}{a_i^2} = 1\} \subset \mathbb{R}^n.$$

Consider $M = T(E)^*$ with the standard symplectic structure and take for H the Legendre transform of the Riemannian metric function. It turns out that the corresponding Hamiltonian system on $T(E)^*$ is always algebraically completely integrable. We shall see it in the simplest possible case when $n = 2$.

Let $x_1 = a_1 \cos t, x_2 = a_2 \sin t$ be a parametrization of the ellipse E. Then the arc length is given by the function

$$s(t) = \int_0^t (a_1^2 \sin^2 \tau + a_2^2 \cos^2 \tau)^{1/2} d\tau = \int_0^t (a_1^2 - (a_1^2 - a_2^2) \cos^2 \tau)^{1/2} d\tau =$$

$$= a_1 \int_0^t (1 - \varepsilon^2 \cos^2 \tau)^{1/2} d\tau$$

where $\varepsilon = \frac{(a_1^2 - a_2^2)^{1/2}}{a_1}$ is the eccentricity of the ellipse. If we set $x = \cos \tau$ then we can transform the last integral to

$$s(t) = \int_0^{\cos t} \frac{1 - \varepsilon^2 x^2}{\sqrt{(1 - x^2)(1 - \varepsilon^2 x^2)}} dx.$$

The corresponding indefinite integral is called an *incomplete Legendre elliptic integral of the second kind.* After extending this integral to complex domain, we can interpret the definite integral

$$F(z) = \int_0^z \frac{(1 - \varepsilon^2 x^2) du}{\sqrt{(1 - x^2)(1 - \varepsilon^2 x^2)}}$$

as the integral of the meromorphic differential form $(1 - \varepsilon^2 x^2) dx/y$ on the Riemann surface X of genus 1 defined by the equation

$$y^2 = (1 - x^2)(1 - \varepsilon^2 x^2).$$

It is a double sheeted cover of $\mathbb{C} \cup \{\infty\}$ branched over the points $x = \pm 1, x = \pm \varepsilon^{-1}$. The function $F(z)$ is a multi-valued function on the Riemann surface X. It is obtained from a single-valued meromorphic doubly periodic function on the universal covering $\tilde{X} = \mathbb{C}$. The fundamental group of X is isomorphic to $\mathbb{Z}^2$ and X is isomorphic to the complex torus $\mathbb{C}/\mathbb{Z}^2$. This shows that the function $s(t)$ is obtained from a periodic meromorphic function on $\mathbb{C}$.

It turns out that if we consider $s(t)$ as a natural parameter on the ellipse, i.e., invert so that $t = t(s)$ and plug this into the parametrization $x_1 = a_1 \cos t, x_2 = a_2 \sin t$, we obtain a solution of the Hamiltonian vector field defined by the metric. Let us check this. We first identify $T(\mathbb{R}^2)$ with $\mathbb{R}^4$ via $(x_1, x_2, y_1, y_2) = (q_1, q_2, \dot{q}_1, \dot{q}_2)$. Use the original parametrization $\phi : [0, 2\pi] \to \mathbb{R}^2$ to consider t as a local parameter on E. Then $(t, \dot{t})$ is a local parameter on $T(E)$ and the inclusion $T(E) \subset T(\mathbb{R}^2)$ is given by the formula

$$(t, \dot{t}) \to (q_1 = a_1 \cos t, q_2 = a_2 \sin t, \dot{q}_1 = -\dot{t} a_1 \sin t, \dot{q}_2 = \dot{t} a_2 \cos t).$$

The metric function on $T(\mathbb{R}^2)$ is

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2).$$

Its pre-image on $T(E)$ is the function

$$l(t,\dot{t}) = \frac{1}{2}\dot{t}^2(a_1^2\sin^2 t + a_2^2\cos^2 t).$$

The Legendre transformation $p = \frac{\partial l}{\partial \dot{t}}$ gives $\dot{t} = p/(a_1^2\sin^2 t + a_2^2\cos^2 t)$. Hence the Hamiltonian function H on $T(E)^*$ is equal to

$$H(t,p) = \dot{t}p - L(t,\dot{t}) = \frac{1}{2}p^2/(a_1^2\sin^2 t + a_2^2\cos^2 t).$$

A path $t = \psi(\tau)$ extends to a path in $T(E)^*$ by the formula $p = (a_1^2\sin^2 t + a_2^2\cos^2 t)\frac{dt}{d\tau}$. If τ is the arc length parameter s and $t = \psi(\tau) = t(s)$, we have $(\frac{ds}{dt})^2 = a_1^2\sin^2 t + a_2^2\cos^2 t$. Therefore

$$H(t(s), p(s)) \equiv 1/2.$$

This shows that the natural parametrization $x_1 = a_1\cos t(s), x_2 = a_2\sin t(s)$ is a solution of Hamilton's equation.

Let us see the other attributes of algebraically completely integrable systems. The Cartesian equation of $T(E)$ in $T(\mathbb{R}^2)$ is

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1, \quad \frac{x_1y_1}{a_1^2} + \frac{x_2y_2}{a_2^2} = 0.$$

If we use (ξ_1, ξ_2) as the dual coordinates of (y_1, y_2), then $T(E)^*$ has the equation in $\mathbb{R}^4$

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1, \quad \frac{x_2\xi_1}{a_1^2} - \frac{x_1\xi_2}{a_2^2} = 0. \tag{3.11}$$

For each fixed x_1, x_2, the second equation is the equation of the dual of the line $\frac{x_1y_1}{a_1^2} + \frac{x_2y_2}{a_2^2} = 0$ in $(\mathbb{R}^2)^*$.

The Hamiltonian function is

$$H(x,\xi) = \xi_1^2 + \xi_2^2.$$

Its level sets M_c are

$$\xi_1^2 + \xi_2^2 - c = 0. \tag{3.12}$$

Now we have an obvious complexification of $M = T(E)^*$. It is the subvariety of $\mathbb{C}^4$ given by equations (3.11). The level sets are given by the additional quadric equation (3.12). Let us compactify $\mathbb{C}^4$ by considering it as a standard open subset of $\mathbb{P}^4$. Then by homogenizing the equation (3.11) we get that $M = V(\mathbb{R})$ where V is a Del Pezzo surface of degree 4 (a complete intersection of two quadrics in $\mathbb{P}^4$). Since the complete intersection of three quadrics in $\mathbb{P}^4$ is a curve of genus 5, the intersection of M with the quadric (3.12) must be a curve of genus 5. We now see that something is wrong, as our complexified level sets must be elliptic curves. The solution to this contradiction is simple. The surface V is singular along the line $x_1 = x_2 = t_0 = 0$ where x_0 is the added homogeneous variable. Its intersection with the quadric (3.12) has 2 singular points $(x_1, x_2, \xi_1, \xi_2, x_0) = (0, 0, \pm\sqrt{-1}, 1, 0)$. Also the points $(1, a_1, \pm a_2\sqrt{-1}, 0, 0, 0)$ are singular points for all c from (3.12). Thus the genus of the complexified level sets is equal to $5 - 4 = 1$. To construct the right complexification of V we must first replace it with its normalization $\bar{V}$. Then the rational function

$$c = \frac{\xi_1^2 + \xi_2^2}{x_0^2}$$

defines a rational map $\bar{V} \to \mathbb{P}^1$. After resolving its indeterminacy points we find a regular map $\pi : X \to \mathbb{P}^1$. For all but finitely many points $z \in \mathbb{P}^1$, the fibre of this map is an elliptic curve. For real $z \neq \infty$ it is a complexification of the level set M_c. If we throw away all singular fibres from X (this will include the fibre over ∞) we finally obtain the right algebraization of our completely integrable system.

Exercises

1. Let θ, ϕ be polar coordinates on a sphere S of radius R in $\mathbb{R}^3$. Show that the Poisson structure on S (considered as a coadjoint orbit of $\mathfrak{so}_3$) is given by the formula

$$\{F(\theta,\phi), G(\theta,\phi)\} = \frac{1}{R\sin\theta}(\frac{\partial F}{\partial \theta}\frac{\partial G}{\partial \phi} - \frac{\partial F}{\partial \phi}\frac{\partial G}{\partial \theta}).$$

2. Check that the symplectic form defined on a coadjoint orbit is closed.

3. Find the Poisson structure and coadjoint orbits for the Lie algebra of the Galilean group.

4. Let (a_{ij}) be a skew-symmetric $n \times n$ matrix with coefficients in a field k. Show that the formula

$$\{P, Q\} = \sum_{i,j=1}^{n} a_{ij} X_i X_j \frac{\partial P}{\partial X_i}\frac{\partial Q}{\partial X_i}$$

defines a Poisson structure on the ring of polynomials $k[X_1, \ldots, X_n]$.

5. Describe all possible cases for the intersection of a sphere and an ellipsoid. Compare it with the topological classification of the sets $E(\mathbb{R})$ where E is an elliptic curve.

6. Let $\mathfrak{g}$ be the Lie algebra of real $n \times n$ lower triangular matrices with zero trace. It is the Lie algebra of the group of real $n \times n$ lower triangular matrices with determinant equal to 1. Identify $\mathfrak{g}^*$ with the set of upper triangular matrices with zero trace by means of the pairing $< A, B >= \mathrm{Trace}(AB)$. Choose $\phi \in \mathfrak{g}^*$ with all entries zero except $a_{ii+1} = 1, i = 1, \ldots, n-1$. Compute the coadjoint orbit of ϕ and its symplectic form induced by the Poisson-Lie structure on $\mathfrak{g}^*$.

7. Find all Lie bialgebra structures on a two-dimensional Lie algebra.

8. Describe the normalization of the surface $V \subset \mathbb{P}^4$ given by equations (3.11). Find all fibres of the elliptic fibration $X \to \mathbb{P}^1$.

Lecture 4. OBSERVABLES AND STATES

4.1 Let M be a Poisson manifold and $H : M \to \mathbb{R}$ be a Hamiltonian function. A smooth function on $f \in \mathcal{O}(M)$ on M will be called an *observable*. For example, H is an observable. The idea behind this definition is clear. The value of f at a point $x \in M$ measures some feature of a particle which happens to be at this point. Any point x of the configuration space M can be viewed as a linear function $f \mapsto f(x)$ on the algebra of observables $\mathcal{O}(M)$. However, in practice, it is impossible to make an experiment which gives the precise value of f. This for example happens when we consider $M = T(\mathbb{R}^{3N})^*$ describing the motion of a large number N of particles in $\mathbb{R}^3$.

So we should assume that f takes the value λ at a point x only with a certain probability. Thus for each subset E of $\mathbb{R}$ and an observable f there will be a certain probability attached that the value of f at x belongs to E.

Definition. A *state* is a function $f \to \mu_f$ on the algebra of observables with values in the set of probability measures on the set of real numbers. For any Borel subset E of $\mathbb{R}$ and an observable $f \in \mathcal{O}(M)$ we denote by $\mu_f(E)$ the measure of E with respect to μ_f. The set of states is denoted by $\mathcal{S}(M)$.

Recall the definition of a probability measure. We consider the set of Borel subsets in $\mathbb{R}$. This is the minimal subset of the Boolean $\mathcal{P}(\mathbb{R})$ which is closed under countable intersections and complements and contains all closed intervals. A *probability measure* is a function $E \to \mu(E)$ on the set of Borel subsets satisfying

$$0 \leq \mu(E) \leq 1, \quad \mu(\emptyset) = 0, \mu(\mathbb{R}) = 1, \quad \mu(\coprod_{n \in \mathbf{N}} E_n) = \sum_{n=1}^{\infty} \mu(E_n).$$

An example of a probability measure is the function:

$$\mu(A) = Lb(A \cap I)/Lb(I), \tag{4.1}$$

where Lb is the Lebesgue measure on $\mathbb{R}$ and I is any closed segment. Another example is the *Dirac probability measure*:

$$\delta_c(E) = \begin{cases} 1 & \text{if } c \in E \\ 0 & \text{if } c \notin E. \end{cases} \tag{4.2}$$

It is clear that each point x in M defines a state which assigns to an observable f the probability measure equal to $\delta_{f(x)}$. Such states are called *pure states.*

We shall assume additionally that states behave naturally with respect to compositions of observables with functions on $\mathbb{R}$. That is, for any smooth function $\phi : \mathbb{R} \to \mathbb{R}$, we have

$$\mu_{\phi\circ f}(E) = \mu_f(\phi^{-1}(E)). \tag{4.3}$$

Here we have to assume that ϕ is B-measurable, i.e., the pre-image of a Borel set is a Borel set. Also we have to extend $\mathcal{O}(M)$ by admitting compositions of smooth functions with B-measurable functions on $\mathbb{R}$.

By taking ϕ equal to the constant function $y = c$, we obtain that for the constant function $f(x) = c$ in $\mathcal{O}(M)$ we have

$$\mu_c = \delta_c. \tag{4.3}$$

For any two states μ, μ' and a nonnegative real number $a \leq 1$ we can define the convex combination $a\mu + (1-a)\mu'$ by

$$(a\mu + (1-a)\mu')(E) = a\mu_f(E) + (1-a)\mu'_f(E).$$

Thus the set of states $\mathcal{S}(M)$ is a convex set.

If $E = (-\infty, \lambda]$ then $\mu_f(E)$ is denoted by $\mu_f(\lambda)$. It gives the probability that the value of f in the state μ is less than or equal than λ. The function $\lambda \to \mu_f(\lambda)$ on $\mathbb{R}$ is called the *distribution function* of the observable f in the state μ.

4.2 Now let us recall the definition of an integral with respect to a probability measure μ. It is first defined for the characteristic functions χ_E of Borel subsets by setting

$$\int_{\mathbb{R}} \chi_E d\mu := \mu(E).$$

Then it is extended to simple functions. A function $\phi : \mathbb{R} \to \mathbb{R}$ is simple if it has a countable set of values $\{y_1, y_2, \ldots\}$ and for any n the pre-image set $\phi^{-1}(y_n)$ is a Borel set. Such a function is called integrable if the infinite series

$$\int_{\mathbb{R}} \phi d\mu := \sum_{n=1}^{\infty} y_n \mu(\phi^{-1}(y_n))$$

converges. Finally we define an integrable function as a function $\phi(x)$ such that there exists a sequence $\{\phi_n\}$ of simple integrable functions with

$$\lim_{n\to\infty} \phi_n(x) = \phi(x)$$

for all $x \in \mathbb{R}$ outside of a set of measure zero, and the series

$$\int_{\mathbb{R}} f d\mu := \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n d\mu$$

converges. It is easy to prove that if ϕ is integrable and $|\psi(x)| \leq \phi(x)$ outside of a subset of measure zero, then $\psi(x)$ is integrable too. Since any continuous function on $\mathbb{R}$ is equal to the limit of simple functions, we obtain that any bounded function which is continuous outside a subset of measure zero is integrable. For example, if μ is defined by Lebesgue measure as in (4.1) where $I = [a, b]$, then

$$\int_{\mathbb{R}} \phi d\mu = \int_a^b \phi(x)dx$$

is the usual Lebesgue integral. If μ is the Dirac measure δ_c, then for any continuous function ϕ,

$$\int_{\mathbb{R}} \phi d\delta_c = \phi(c).$$

The set of functions which are integrable with respect to μ is a commutative algebra over $\mathbb{R}$. The integral

$$\int_{\mathbb{R}} : \phi \to \int_{\mathbb{R}} \phi d\mu$$

is a linear functional. It satisfies the positivity property

$$\int_{\mathbb{R}} \phi^2 d\mu \geq 0.$$

One can define a probability measure with help of a *density function*. Fix a positive valued function $\rho(x)$ on $\mathbb{R}$ which is integrable with respect to a probability measure μ and satisfies

$$\int_{\mathbb{R}} \rho(x)d\mu = 1.$$

Then define the new measure ρ_μ by

$$\rho_\mu(E) = \int_{\mathbb{R}} \chi_E \rho(x)d\mu.$$

Obviously we have

$$\int_{\mathbb{R}} \phi(x)d\rho\mu = \int_{\mathbb{R}} \phi\rho d\mu.$$

For example, one takes for μ the measure defined by (4.1) with $I = [a, b]$ and obtains a new measure μ'

$$\int_{\mathbb{R}} \phi d\mu' = \int_a^b \phi(x)\rho(x)dx.$$

Of course not every probability measure looks like this. For example, the Dirac measure does not. However, by definition, we write

$$\int_{\mathbb{R}} \phi d\delta_c = \int_{\mathbb{R}} \phi(x)\rho_c dx,$$

where $\rho_c(x)$ is the Dirac "function". It is zero outside $\{c\}$ and equal to ∞ at c.

The notion of a probability measure on $\mathbb{R}$ extends to the notion of a probability measure on $\mathbb{R}^n$. In fact for any measures $\mu_1, \ldots, \mu_n$ on $\mathbb{R}$ there exists a unique measure μ on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} f_1(x_1)\ldots f_n(x_n)d\mu = \prod_{i=1}^{n} \int_{\mathbb{R}} f_i(x)d\mu_i.$$

After this it is easy to define a measure on any manifold.

4.3 Let $\mu \in \mathcal{S}(M)$ be a state. We define

$$\langle f|\mu\rangle = \int_{\mathbb{R}} x d\mu_f.$$

We assume that this integral is defined for all f. This puts another condition on the state μ. For example, we may always restrict ourselves to probability measures of the form (4.1). The number $\langle f|\mu\rangle$ is called the *mathematical expectation of the observable f in the state μ*. It should be thought as the mean value of f in the state μ. This function completely determines the values $\mu_f(\lambda)$ for all $\lambda \in \mathbb{R}$. In fact, if we consider the *Heaviside function*

$$\theta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{4.4}$$

then, by property (4.2), we have

$$\mu'_f(E) = \mu_{\theta(\lambda - f)}(E) = \mu_f(\{x \in \mathbb{R} : \theta(\lambda - x) \in E\}) = \begin{cases} 0 & \text{if } \{0,1\} \cap E = \emptyset \\ 1 & \text{if } 0,1 \in E \\ \mu_f(\lambda) & \text{if } 1 \in E, 0 \notin E \\ 1 - \mu_f(\lambda) & \text{if } 0 \in E, 1 \notin E. \end{cases}$$

It is clear now that, for any simple function ϕ, $\int_{\mathbb{R}} \phi(x)d\mu'_f = 0$ if $\phi(0) = \phi(1) = 0$. Now writing the identity function $y = x$ as a limit of simple functions, we easily get that

$$\mu_f(\lambda) = \int_{\mathbb{R}} x d\mu'_f = \langle \theta(\lambda - f)|\mu\rangle.$$

Also observe that, knowing the function $\lambda \to \mu_f(\lambda)$, we can reconstruct μ_f (defining it first on intervals by $\mu_f([a,b]) = \mu_f(b) - \mu_f(a)$). Thus the state μ is completely determined by the values $\langle f|\mu\rangle$ for all observable f.

4.4 We shall assume some additional properties of states:

$$\langle af + bg|\mu\rangle = a\langle f|\mu > +b < g|\mu\rangle, a, b \in \mathbb{R}, \tag{S1}$$

$$\langle f^2|\mu\rangle \geq 0, \tag{S2}$$

$$\langle 1|\mu\rangle = 1. \tag{S3}$$

In this way a state becomes a linear function on the space of observables satisfying the positivity propertry (S2) and the normalization property (S3). We shall assume that each such functional is defined by some probability measure Υ_μ on M. Thus

$$\langle f|\mu\rangle = \int_M f d\Upsilon_\mu.$$

Property (S1) gives

$$\int_M d\Upsilon_\mu = \Upsilon_\mu(M) = 1$$

as it should be for a probability measure. An example of such a measure is given by choosing a volume form Ω on M and a positive valued integrable function $\rho(x)$ on M such that

$$\int_M \rho(x)\Omega = 1.$$

We shall assume that the measure Υ_μ is given in this way. If (M, ω) is a symplectic manifold, we can take for Ω the canonical volume form $\omega^{\wedge n}$.

Thus every state is completely determined by its density function ρ_μ and

$$\langle f|\mu\rangle = \int_M f(x)\rho_\mu(x)\Omega.$$

From now on we shall identify states μ with density functions ρ satisfying

$$\int_M \rho(x)\Omega = 1.$$

The expression $\langle f|\mu\rangle$ can be viewed as evaluating the function f at the state μ. When μ is a pure state, this is the usual value of f. The idea of using non-pure states is one of the basic ideas in quantum and statistical mechanics.

4.5 There are two ways to describe evolution of a mechanical system. We have used the first one (see Lecture 3, 3.3); it says that observables evolve according to the equation

$$\frac{df_t}{dt} = \{f_t, H\}$$

and states do not change with time, i.e.

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

Here we define ρ_t by the formula

$$\rho_t(x) = \rho(g^{-t} \cdot x), \quad x \in M.$$

Alternatively, we can describe a mechanical system by the equations

$$\frac{d\rho_t}{dt} = -\{\rho_t, H\}, \quad \frac{df_t}{dt} = 0.$$

Of course here we either assume that ρ is smooth or learn how to differentiate it in a generalized sense (for example if ρ is a Dirac function). In the case of symplectic Poisson structures the equivalence of these two approaches is based on the following

Proposition. *Assume M is a symplectic Poisson manifold. Let $\Omega = \omega^{\wedge n}$, where ω is a symplectic form on M. Define μ_t as the state corresponding to the density function ρ_t. Then*

$$\langle f_t|\mu\rangle = \langle f|\mu_t\rangle.$$

Proof. Applying Liouville's Theorem (Example 1, Lecture 2), we have

$$\langle f_t|\mu\rangle = \int_M f(g^t \cdot x)\rho_\mu(x)\Omega = \int_M f(x)\rho_\mu(g^{-t} \cdot x)(g^{-t})^*(\Omega) =$$

$$\int_M f(x)\rho_\mu(g^{-t} \cdot x)\Omega = \int_M f(x)\rho_{\mu_t}(x)\Omega = \langle f|\mu_t\rangle.$$

The second description of a mechanical system described by evolution of the states via the corresponding density function is called *Liouville's picture.* The standard approach where we evolve observables and leave states unchanged is called *Hamilton's picture.* In statistical mechanics Liouville's picture is used more often.

4.6 The introduction of quantum mechanics is based on Heisenberg's Uncertainty Principle. It tells us that in general one cannot measure simultaneously two observables. This makes impossible in general to evaluate a composite function $F(f, g)$ for any two observables f, g. Of course it is still possible to compute $F(f)$. Nevertheless, our definition of states allows us to define the sum $f + g$ by

$$\langle f + g|\mu\rangle = \langle f|\mu\rangle + \langle g|\mu\rangle \quad \text{for any state } \mu,$$

considering f as a function on the set of states. To make this definition we of course use that $f = f'$ if $\langle f|\mu\rangle = \langle f'|\mu\rangle$ for any state μ. This is verified by taking pure states μ.

However the same property does not apply to states. So we have to identify two states μ, μ' if they define the same function on observables, i.e., if $\langle f|\mu\rangle = \langle g|\mu'\rangle$ for all observables f. Of course this means that we replace $\mathcal{S}(M)$ with some quotient space $\bar{\mathcal{S}}(M)$.

We have no similar definition of product of observables since we cannot expect that the formula $\langle fg|\mu\rangle = \langle f|\mu\rangle\langle g|\mu\rangle$ is true for any definition of states μ. However, the following product

$$f * g = \frac{(f+g)^2 - (f-g)^2}{4} \tag{4.5}$$

makes sense. Note that $f * f = f^2$. This new binary operation is commutative but not associative. It satifies the following weaker axiom

$$x^2(yx) = (x^2y)x.$$

The addition and multiplication operations are related to each other by the distributivity axiom. An algebraic structure of this sort is called a *Jordan algebra*. It was introduced in 1933 by P. Jordan in his works on the mathematical foundations of quantum mechanics. Each associative algebra A defines a Jordan algebra if one changes the multiplication by using formula (4.5).

We have to remember also the Poisson structure on $\mathcal{O}(M)$. It defines the Poisson structure on the Jordan algebra of observables, i.e., for any fixed f, the formula $g \to \{f, g\}$ is a derivation of the Jordan algebra $\mathcal{O}(M)$

$$\{f, g * h\} = \{f, g\} * h + \{f, h\} * g.$$

We leave the verification to the reader.

Here comes our main definition:

Definition. A *Jordan Poisson algebra* $\mathcal{O}$ over $\mathbb{R}$ is called an *algebra of observables*. A linear map $\omega : \mathcal{O} \to \mathbb{R}$ is called a *state* on $\mathcal{O}$ if it satisfies the following properties:
(1) $\langle a^2|\omega\rangle \geq 0$ for any $a \in \mathcal{O}$;
(2) $\langle 1|\omega\rangle = 1$.

The set $\mathcal{S}(\mathcal{O})$ of states on $\mathcal{O}$ is a convex non-empty subset of states in the space of real valued linear functions on $\mathcal{O}$.

Definition. A state $\omega \in \mathcal{S}(\mathcal{O})$ is called *pure* if $\omega = c\omega_1 + (1-c)\omega_2$ for some $\omega_1, \omega_2 \in \mathcal{S}(\mathcal{O})$ and some $0 \leq c \leq 1$ implies $\omega = \omega_1$. In other words, ω is an extreme point of the convex set $\mathcal{S}(\mathcal{O})$.

4.7 In the previous sections we discussed one example of an algebra of observables and its set of states. This was the Jordan algebra $\mathcal{O}(M)$ and the set of states $\mathcal{S}(\mathcal{M})$. Let us give another example.

We take for $\mathcal{O}$ the set $\mathcal{H}(V)$ of self-adjoint (= Hermitian) operators on a finite-dimensional complex vector space V with unitary inner product $(,)$. Recall that a linear operator $A : V \to V$ is called self-adjoint if

$$(Ax, y) = (x, Ay) \quad \text{for any } x, y \in V.$$

If (a_{ij}) is the matrix of A with respect to some orthonormal basis of V, then the property of self-adjointness is equivalent to the property $\bar{a}_{ij} = a_{ji}$, where the bar denotes the complex conjugate.

We define the Jordan product on $\mathcal{H}(V)$ by the formula

$$A * B = \frac{(A+B)^2 - (A-B)^2}{4} = \frac{1}{2}(A \circ B + B \circ A).$$

Here, as usual, we denote by $A \circ B$ the composition of operators. Notice that since

$$\overline{\sum_{k=1}^{n}(a_{ik}b_{kj} + b_{ik}a_{kj})} = \sum_{k=1}^{n}(\bar{a}_{ik}\bar{b}_{kj} + \bar{b}_{ik}\bar{a}_{kj}) = \sum_{k=1}^{n}(a_{jk}b_{ki} + b_{jk}a_{ki})$$

the operator $A \circ B + B \circ A$ is self-adjoint. Next we define the Poisson bracket by the formula

$$\{A, B\}_\hbar = \frac{i}{\hbar}[A, B] = \frac{i}{\hbar}(A \circ B - B \circ A). \tag{4.6}$$

Here $i = \sqrt{-1}$ and $\hbar$ is any real constant. Note that neither the product nor the Lie bracket of self-adjoint operators is a self-adjoint operator. We have

$$\overline{\sum_{k=1}^{n}(a_{ik}b_{kj} - b_{ik}a_{kj})} = \sum_{k=1}^{n}(\bar{a}_{ik}\bar{b}_{kj} - \bar{b}_{ik}\bar{a}_{kj}) = \sum_{k=1}^{n}(b_{jk}a_{ki} - a_{jk}b_{ki}) = -\sum_{k=1}^{n}(a_{jk}b_{ki} - b_{jk}a_{ki}).$$

However, putting the imaginary scalar in, we get the self-adjointness. We leave to the reader to verify that

$$\{A, B * C\}_\hbar = \{A, B\}_\hbar * C + \{A, C\}_\hbar * B.$$

The Jacobi property of $\{,\}$ follows from the Jacobi property of the Lie bracket $[A, B]$. The structure of the Poisson algebra depends on the constant $\hbar$. Note that rescaling of the Poisson bracket on a symplectic manifold leads to an isomorphic Poisson bracket!

We shall denote by $Tr(A)$ the trace of an operator $A : V \to V$. It is equal to the sum of the diagonal elements in the matrix of A with respect to any basis of V. Also

$$Tr(A) = \sum_{i=1}^{n}(Ae_i, e_i)$$

where $(e_1, \ldots, e_n)$ is an orthonormal basis in V. We have, for any $A, B, C \in \mathcal{H}(V)$,

$$Tr(AB) = Tr(BA), \quad Tr(A + B) = Tr(A) + Tr(B). \tag{4.7}$$

Lemma. *Let $\ell : \mathcal{H}(V) \to \mathbb{R}$ be a state on the algebra of observables $\mathcal{H}(V)$. Then there exists a unique linear operator $M : V \to V$ with $\ell(A) = Tr(MA)$. The operator M satisfies the following properties:*
(i) (self-adjointness) $M \in \mathcal{H}(V)$;
(ii) (non-negativity) $(Mx, x) \geq 0$ for any $x \in V$;
(iii) (normalization) $Tr(M) = 1$.

Proof. Let us consider the linear map $\alpha : \mathcal{H}(V) \to \mathcal{H}(V)^*$ which associates to M the linear functional $A \to Tr(MA)$. This map is injective. In fact, if M belongs to the kernel, then by properties (4.6) we have $Tr(MBA) = Tr(AMB) = 0$ for all $A, B \in \mathcal{H}(V)$. For each $v \in V$ with $||v|| = 1$, consider the orthogonal projection operator P_v defined by

$$P_v x = (x, v)v.$$

It is obviously self-adjoint since $(P_v x, y) = (x, v)(v, y) = (x, P_v y)$. Fix an orthonormal basis $v = e_1, e_2, \ldots, e_n$ in V. We have

$$0 = Tr(MP_v) = \sum_{k=1}^{n} (MP_v e_k, e_k) = (Mv, v).$$

Since $(M\lambda v, \lambda v) = |\lambda|^2 (Mv, v)$, this implies that $(Mx, x) = 0$ for all $x \in V$. Replacing x with $x + y$ and $x + iy$, we obtain that $(Mx, y) = 0$ for all $x, y \in V$. This proves that $M = 0$. Since dim $\mathcal{H}(V) = \dim \mathcal{H}(V)^*$, we obtain the surjectivity of the map α. This shows that $\ell(A) = Tr(MA)$ for any $A \in \mathcal{H}(V)$ and proves the uniqueness of such M.

Now let us check the properties (i) - (iii) of M. Fix an orthonormal basis $e_1, \ldots, e_n$. Since $\ell(A) = Tr(MA)$ must be real, we have

$$Tr(MA) = \overline{Tr(MA)} = \sum_{i=1}^{n} \overline{(MA(e_i), e_i)} = \sum_{i=1}^{n} (e_i, MA(e_i)) = \sum_{i=1}^{n} ((MA)^* e_i, e_i) =$$

$$\sum_{i=1}^{n} (A^* M^* e_i, e_i) = Tr(A^* M^*) = Tr(M^* A^*) = Tr(M^* A).$$

Since this is true for all A, we get $M^* = M$. This checks (i).

We have $\ell(A^2) = Tr(A^2 M) \geq 0$ for all A. Take $A = P_v$ to be a projection operator as above. Changing the orthonormal basis we may assume that $e_1 = v$. Then

$$Tr(P_v^2 M) = Tr(P_v M) = (Mv, v) \geq 0.$$

This checks (ii). Property (iii) is obvious.

The operator M satisfying the properties of the previous lemma is called the *density operator*. We see that any state can be uniquely written in the form

$$\ell(A) = Tr(MA)$$

where M is a density operator. So we can identify states with density operators.

All properties of a density operator are satisfied by the projection operator P_v. In fact

(i) $(P_v x, y) = (x, v)(v, y) = \overline{(y, v)}(x, v) = (x, P_v y)$.
(ii) $(P_v x, x) = (x, v)(v, x) = |(x, y)|^2 \geq 0$
(iii) $Tr(P_v) = (v, v) = 1$.

Notice that the set of states M is a convex set.

Proposition. *An operator $M \in \mathcal{H}(V)$ is a pure state if and only if it is equal to a projection operator P_v.*

Proof. Choose a self-adjoint operator M and a basis of eigenvectors ξ_i (it exists because M is self-adjoint). Then

$$v = \sum_{i=1}^{n} (v, \xi_i)\xi_i, \quad Mv = \sum_{i=1}^{n} \lambda_i (v, \xi_i)\xi_i = \sum_{i=1}^{n} \lambda_i P_{\xi_i} v.$$

Since this is true for all v, we have

$$M = \sum_{i=1}^{n} \lambda_i P_{\xi_i}.$$

All numbers here are real numbers (again because M is self-adjoint). Also by property (ii) from the Lemma, the numbers λ_i are non-negative. Since $Tr(M) = 1$, we get that they add up to 1. This shows that only projection operators can be pure. Now let us show that P_v is always pure.

Assume $P_v = cM_1 + (1-c)M_2$ for some $0 < c < 1$. For any density operator M the new inner product $(x, y)_M = (Mx, y)$ has the property $(x, x)' \geq 0$. So we can apply the Cauchy-Schwarz inequality

$$|(Mx, y)|^2 \leq (Mx, x)(My, y)$$

to obtain that $Mx = 0$ if $(Mx, x) = 0$. Since M_1, M_2 are density operators, we have

$$0 \leq c(M_1 x, x) \leq c(M_1 x, x) + (1-c)(M_2 x, x) = (P_v x, x) = 0$$

for any x with $(x, v) = 0$. This implies that $M_1 x = 0$ for such x. Since M_1 is self-adjoint,

$$0 = (M_1 x, v) = (x, M_1 v) \quad \text{if } (x, v) = 0.$$

Thus M_1 and P_v have the same kernel and the same image. Hence $M_1 = cP_v$ for some constant c. By the normalization property $Tr(M_1) = 1$, we obtain $c = 1$; hence $P_v = M_1$.

Since $P_v = P_c v$ where $|c| = 1$, we can identify pure states with the points of the projective space $\mathbb{P}(V)$.

4.8 Recall that $\langle f, \mu \rangle$ is interpreted as the mean value of the state μ at an observable f. The expression

$$\Delta_\mu(f) = \langle (f - \langle f|\mu\rangle)^2|\mu\rangle^{\frac{1}{2}} = (\langle f^2|\mu\rangle - \langle f|\mu\rangle^2)^{1/2} \tag{4.8}$$

can be interpreted as the mean of the deviation of the value of μ at f from the mean value. It is called the *dispersion* of the observable $f \in \mathcal{O}$ with respect to the state μ. For example, if $x \in M$ is a pure state on the algebra of observables $\mathcal{O}(M)$, then $\Delta_\mu(f) = 0$.

We can introduce the inner product on the algebra of observables by setting

$$(f, g)_\mu = \frac{1}{2}\Big(\Delta_\mu(f+g)^2 - (\Delta_\mu(f))^2 - (\Delta_\mu(g))^2\Big). \tag{4.9}$$

It is not positive definite but satisfies

$$(f, f)_\mu = \Delta_\mu(f)^2 \geq 0.$$

By the parallelogram rule,

$$(f, g)_\mu = \frac{1}{4}[(f+g, f+g)_\mu - (f-g, f-g)_\mu] =$$

$$= \frac{1}{4}(\langle (f+g)^2|\mu\rangle - \langle (f-g)^2|\mu\rangle - \langle f+g, \mu\rangle^2 + \langle f-g|\mu\rangle^2)$$

$$= \langle \frac{1}{4}[(f+g)^2 - (f-g)^2]|\mu\rangle - \langle f|\mu\rangle\langle g|\mu\rangle = \langle f * g|\mu\rangle - \langle f|\mu\rangle\langle g|\mu\rangle. \tag{4.10}$$

This shows that

$$\langle f * g|\mu\rangle = \langle f|\mu\rangle\langle g|\mu\rangle \Longleftrightarrow (f, g)_\mu = 0.$$

As we have already observed, in general $\langle f*g|\mu\rangle \neq \langle f|\mu\rangle\langle g|\mu\rangle$. If this happens, the observables are called *independent* with respect to μ. Thus, two observables are independent if and only if they are orthogonal.

Let ϵ be a positive real number. We say that an observable f is *ϵ-measurable at the state μ* if $\Delta_\mu(f) < \epsilon$.

Applying the Cauchy-Schwarz inequality

$$|(f, g)_\mu| \leq (f, f)_\mu^{\frac{1}{2}}(g, g)_\mu^{\frac{1}{2}} = \Delta_\mu(f)\Delta_\mu(g), \tag{4.11}$$

and (4.8), we obtain that $f + g$ is 2ϵ-measurable if f, g are ϵ-measurable.

The next theorem expresses one of the most important principles of quantum mecahnics:

Theorem (Heisenberg's Uncertainty Principle). *Let $\mathcal{O} = \mathcal{H}(V)$ be the algebra of observables formed by self-associated operators with Poisson structure defined by (4.6). For any $A, B \in \mathcal{O}$ and any $\mu \in \mathcal{S}(V)$,*

$$\Delta_\mu(A)\Delta_\mu(B) \geq \frac{\hbar}{2}|\langle \{A, B\}_\hbar|\mu\rangle|. \tag{4.12}$$

Proof. For any $v \in V$ and $\lambda \in \mathbb{R}$, we have

$$0 \leq ((A+i\lambda B)v, (A+i\lambda B)v) = (A^2v, v) + \lambda^2(B^2v, v) - i\lambda((AB-BA)v, v).$$

Assume $||v|| = 1$ and let μ be state defined by the projection operator $M = P_v$. Then we can rewrite the previous inequality in the form

$$\langle A^2|\mu\rangle + \lambda^2\langle B^2|\mu\rangle - \lambda\hbar\langle\{A,B\}_\hbar|\mu\rangle \geq 0.$$

Since this is true for all λ, we get

$$\langle A^2|\mu\rangle\langle B^2|\mu\rangle \geq \frac{\hbar^2}{4}\langle\{A,B\}_\hbar|\mu\rangle^2.$$

Replacing A with $A - \langle A|\mu\rangle$, and B with $B - \langle B|\mu\rangle$, and taking the square root on both sides we obtain the inequality (4.11). To show that the same inequality is true for all states we use the following convexity property of the dispersion

$$\Delta_\mu(f)\Delta_\mu(g) \geq c\Delta_{\mu_1}(f)\Delta_{\mu_2}(g) + (1-c)\Delta_{\mu_1}(f)\Delta_{\mu_2}(g),$$

where $\mu = c\mu_1 + (1-c)\mu_2, 0 \leq c \leq 1$ (see Exercise 1).

Corollary. *Assume that $\{A,B\}_\hbar$ is the identity operator and A, B are ϵ-measurable in a state μ. Then*

$$\epsilon > \sqrt{\hbar/2}.$$

The Indeterminacy Principle says that there exist observables which cannot be simulateneously measured in any state with arbitrary small accuracy.

Exercises.

1. Let $\mathcal{O}$ be an algebra of observables and $\mathcal{S}(\mathcal{O})$ be a set of states of $\mathcal{O}$. Prove that mixing states increases dispersion, that is, if $\omega = c\omega_1 + (1-c)\omega_2, 0 \leq c \leq 1$, then
 (i) $\Delta_\omega f \geq c\Delta_{\omega_1} f + (1-c)\Delta_{\omega_2} f$.
 (ii) $\Delta_\omega f \Delta_\omega g \geq c\Delta_{\omega_1} f \Delta_{\omega_2} g + (1-c)\Delta_{\omega_1} f \Delta_{\omega_2} g$. What is the physical meaning of this?

2. Suppose that f is ϵ-measurable at a state μ. What can you say about f^2?

3. Find all Jordan structures on algebras of dimension ≤ 3 over real numbers. Verify that they are all obtained from associative algebras.

4. Let $\mathfrak{g}$ be the Lie algebra of real triangular 3×3-matrices. Consider the algebra of polynomial functions on $\mathfrak{g}$ as the Poisson-Lie algebra and equip it with the Jordan product defined by the associative multiplication. Prove that this makes an algebra of observables and describe its states.

5. Let $v_1, \ldots, v_n$ be a basis of eigenvectors of a self-adjoint operator A and $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. Assume that $\lambda_1 < \ldots < \lambda_n$. Define the spectral function of A as the operator

$$P_A(\lambda) = \sum_{\lambda_i \leq \lambda} P_{v_i}.$$

Let M be a density operator.

(i) Show that the function $\mu_A(\lambda) = Tr(MP_A(\lambda))$ defines a probability measure μ_A on $\mathbb{R}$.

(ii) Show that $\langle A|\omega\rangle = \int\limits_{\mathbb{R}} x d\mu_A$ where ω is the state with the density operator M.

Lecture 5. OPERATORS IN HILBERT SPACE

5.1 Let $\mathcal{O} = \mathcal{H}(V)$ be the algebra of observables consisting of self-adjoint operators in a unitary finite-dimensional space. The set of its states can be identified with the set of density operators, i.e., the set of self-adjoint positive definite operators M normalized by the condition $Tr(M) = 1$. Recall that the value $\langle A|\omega\rangle$ is defined by the formula

$$\langle A|M\rangle = Tr(MA).$$

Now we want to introduce dynamics on $\mathcal{O}$. By analogy with classical mechanics, it can be defined in two ways:

$$\frac{dA(t)}{dt} = \{A(t), H\}_\hbar, \quad \frac{dM}{dt} = 0 \tag{5.1}$$

(*Heisenberg's picture*), or

$$\frac{dM(t)}{dt} = -\{M(t), H\}_\hbar, \quad \frac{dA}{dt} = 0 \tag{5.2}$$

(*Schrödinger's picture*).

Let us discuss the first picture. Here $A(t)$ is a smooth map $\mathbb{R} \to \mathcal{O}$. If one chooses a basis of V, this map is given by $A(t) = (a_{ij}(t))$ where $a_{ij}(t)$ is a smooth function of t. Consider equation (5.1) with initial conditions

$$A(0) = A.$$

We can solve it uniquely using the formula

$$A(t) = A_t := e^{-\frac{i}{\hbar}Ht} A e^{\frac{i}{\hbar}Ht}. \tag{5.3}$$

Here for any diagonalizable operator A and any analytic function $f : \mathbb{C} \to \mathbb{C}$ we define

$$f(A) = S \cdot \mathrm{diag}[f(d_1), \ldots, f(d_n)] \cdot S^{-1}$$

where $A = S \cdot \text{diag}[d_1, \dots, d_n] \cdot S^{-1}$. Now we check

$$\frac{dA_t}{dt} = -\frac{i}{\hbar} He^{-\frac{i}{\hbar}Ht} Ae^{\frac{i}{\hbar}Ht} + e^{-\frac{i}{\hbar}Ht} A(\frac{i}{\hbar} He^{\frac{i}{\hbar}Ht}) = \frac{i}{H}(A_t H - HA_t) = \{A_t, H\}_\hbar$$

and the initial condition is obviously satisfied.

The operators

$$U(t) = e^{\frac{i}{\hbar}Ht}, t \in \mathbb{R}$$

define a one-parameter group of unitary operators. In fact,

$$U(t)^* = e^{-\frac{i}{\hbar}Ht} = U(t)^{-1} = U(-t), \quad U(t+t') = U(t)U(t').$$

This group plays the role of the flow g^t of the Hamiltonian. Notice that (5.3) can be rewritten in the form

$$A_t = U(t)^{-1} AU(t).$$

There is an analog of Theorem 4.1 from Lecture 4:

Theorem 1. *The map $\mathcal{O} \to \mathcal{O}$ defined by the formula*

$$U_t : A \to A_t = U(t)^{-1} AU(t)$$

is an automorphism of the algebra of observables $\mathcal{H}(V)$.

Proof. Since conjugation is obviously a homomorphism of the associative algebra structure on $\text{End}(V)$, we have

$$(A * B)_t = A_t B_t + B_t A_t = A_t * B_t.$$

Also

$$(\{A, B\}_\hbar)_t = \frac{i}{\hbar}[A, B]_t = \frac{i}{\hbar}(A_t B_t - B_t A_t) = \frac{i}{\hbar}\{A_t, B_t\}_\hbar.$$

The unitary operators

$$U(t) = e^{\frac{i}{\hbar}Ht}$$

are called the *evolution operators* with respect to the Hamiltonian H.

The Heisenberg and Schrödinger pictures are equivalent in the following sense:

Proposition 1. *Let M be a density operator and let $\omega(t)$ be the state with the density operator M_{-t}. Then*

$$\langle A_t | \omega \rangle = \langle A | \omega(t) \rangle.$$

Proof. We have

$$\langle A_t | \omega \rangle = Tr(MU(t)AU(-t)) = Tr(U(-t)MU(t)A) = Tr(M_{-t}A) = \langle A | \omega(t) \rangle.$$

5.2 Let us consider Schrödinger's picture and take for A_t the evolution of a pure state $(P_v)_t = U_t P_v$. We have

Lemma 1.

$$(P_v)_t = P_{U(-t)v}.$$

Proof. Let S be any unitary operator. For any $x \in V$, we have

$$(SP_vS^{-1})x = SP_v(S^{-1}x) = S((v, S^{-1}x)v) = (v, S^{-1}x)Sv = (Sv, x)Sv = P_{Sv}x. \quad (5.4)$$

Now the assertion follows when we take $S = U(-t)$.

The Lemma says that the evolution of the pure state P_v is equivalent to the evolution of the vector v defined by the formula

$$v_t = U(-t)v = e^{-\frac{i}{\hbar}Ht}v. \quad (5.5)$$

Notice that, since $U(t)$ is a unitary operator,

$$||v_t|| = ||v||.$$

We have

$$\frac{dv_t}{dt} = \frac{de^{-\frac{i}{\hbar}Ht}v}{dt} = -\frac{i}{\hbar}e^{-\frac{i}{\hbar}Ht}v = -\frac{i}{\hbar}v_t.$$

Thus the evolution $v(t) = v_t$ satisfies the following equation (called the *Schrödinger equation*)

$$i\hbar\frac{dv(t)}{dt} = Hv(t), \quad v(0) = v. \quad (5.6)$$

Assume that v is an eigenvector of H, i.e., $Hv = \lambda v$ for some $\lambda \in \mathbb{R}$. Then

$$v_t = e^{-\frac{i}{\hbar}Ht}v = e^{-\frac{i}{\hbar}\lambda t}v.$$

Although the vector changes with time, the corresponding state P_v does not change. So the problem of finding the spectrum of the operator H is equivalent to the problem of finding stationary states.

5.3 Now we shall move on and consider another model of the algebra of observables and the corresponding set of states. It is much closer to the "real world". Recall that in classical mechanics, pure states are vectors $(q_1, \ldots, q_n) \in \mathbb{R}^n$. The number n is the number of degrees of freedom of the system. For example, if we consider N particles in $\mathbb{R}^3$ we have $n = 3N$ degrees of freedom. If we view our motion in $T(\mathbb{R}^n)^*$ we have $2n$ degrees of freedom. When the number of degrees of freedom is increasing, it is natural to consider the space $\mathbb{R}^{\mathbb{N}}$ of infinite sequences $\mathbf{a} = (a_1, \ldots, a_n, \ldots)$. We can make the set $\mathbb{R}^{\mathbb{N}}$ a vector space by using operations of addition and multiplication of functions. The subspace $l_2(\mathbb{R})$ of $\mathbb{R}^{\mathbb{N}}$ which consists of sequences $\mathbf{a}$ satisfying

$$\sum_{n=1}^{\infty} a_i^2 < \infty$$

can be made into a Hilbert space by defining the unitary inner product

$$(\mathbf{a}, \mathbf{b}) = \sum_{n=1}^{\infty} a_i b_i.$$

Recall that a *Hilbert space* is a unitary inner product space (not necessary finite-dimensional) such that the norm makes it a complete metric space. In particular, we can define the notion of limits of sequences and infinite sums. Also we can do the same with linear operators by using pointwise (or better vectorwise) convergence. Every Cauchy (fundamental) sequence will be convergent.

Now it is clear that we should define observables as self-adjoint operators in this space. First, to do this we have to admit complex entries in the sequences. So we replace $\mathbb{R}^{\mathbb{N}}$ with $\mathbb{C}^{\mathbb{N}}$ and consider its subspace $V = l^2(\mathbb{C})$ of sequences satisfying

$$\sum_{i=1}^{n} |a_i|^2 < \infty$$

with inner product

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} a_i \bar{b}_i.$$

More generally, we can replace $\mathbb{R}^N$ with any manifold M and consider pure states as complex-valued functions $F : M \to \mathbb{C}$ on the configuration space M such that the function $|f(x)|^2$ is integrable with respect to some appropriate measure on M. These functions form a linear space. Its quotient by the subspace of functions equal to zero on a set of measure zero is denoted by $L^2(M, \mu)$. The unitary inner product

$$(f, g) = \int_M f \bar{g} d\mu$$

makes it a Hilbert space. In fact the two Hilbert spaces $l^2(\mathbb{C})$ and $L^2(M, \mu)$ are isomorphic. This result is a fundamental result in the theory of Hilbert spaces due to Fischer and Riesz.

One can take one step further and consider, instead of functions f on M, sections of some *Hermitian vector bundle* $E \to M$. The latter means that each fibre E_x is equiped with a Hermitian inner product $\langle , \rangle_x$ such that for any two local sections s, s' of E the function

$$\langle s, s' \rangle : M \to \mathbb{C}, x \to \langle s(x), s'(x) \rangle_x$$

is smooth. Then we define the space $L^2(E)$ by considering sections of E such that

$$\int_M \langle s, s \rangle d\mu < \infty$$

and defining the inner product by

$$(s, s') = \int_M \langle s, \bar{s} \rangle d\mu.$$

We will use this model later when we study quantum field theory.

5.4 So from now on we shall consider an arbitrary Hilbert space V. We shall also assume that V is *separable* (with respect to the topology of the corresponding metric space), i.e., it contains a countable everywhere dense subset. In all examples which we encounter, this condition is satisfied. We understand the equality

$$v = \sum_{n=1}^{\infty} v_n$$

in the sense of convergence of infinite series in a metric space. A *basis* is a linearly independent set $\{v_n\}$ such that each $v \in V$ can be written in the form

$$v = \sum_{n=1}^{\infty} \alpha_n v_n.$$

In other words, a basis is a linearly independent set in V which spans a dense linear subspace. In a separable Hilbert space V, one can construct a countable orthonormal basis $\{e_n\}$. To do this one starts with any countable dense subset, then goes to a smaller linearly independent subset, and then orthonormalizes it.

If (e_n) is an orthonormal basis, then we have

$$v = \sum_{n=1}^{\infty} (v, e_n) e_n \qquad \text{(Fourier expansion)},$$

and

$$(\sum_{n=1}^{\infty} a_n e_n, \sum_{n=1}^{\infty} b_n e_n) = \sum_{n=1}^{\infty} a_n \bar{b}_n. \tag{5.7}$$

The latter explains why any two separable infinite-dimensional Hilbert spaces are isomorphic. We shall use that any closed linear subspace of a Hilbert space is a Hilbert (separable) space with respect to the induced inner product.

We should point out that not everything which the reader has learned in a linear algebra course is transferred easily to arbitrary Hilbert spaces. For example, in the finite dimensional case, each subspace L has its orthogonal complement $L^{\perp}$ such that $V = L \oplus L^{\perp}$. If the same were true in any Hilbert space V, we get that any subspace L is a closed subset (because the condition $L = \{x \in V : (x, v) = 0\}$ for any $v \in L^{\perp}$ makes L closed). However, not every linear subspace is closed. For example, take $V = L^2([-1, 1], dx)$ and L to be the subspace of continuous functions.

Let us turn to operators in a Hilbert space. A linear operator in V is a linear map $T : V \to V$ of vector spaces. We shall denote its image on a vector $v \in V$ by Tv or $T(v)$. Note that in physics, one uses the notation

$$T(v) = \langle T | v \rangle.$$

In order to distinguish operators from vectors, they use notation $\langle T|$ for operators (*bra*) and $|v\rangle$ for vectors (*ket*). The reason for the names is obvious : bracket = bra+c+ket. We shall stick to the mathematical notation.

A linear operator $T : V \to V$ is called *continuous* if for any convergent sequence $\{v_n\}$ of vectors in V,

$$\lim_{n\to\infty} Tv_n = T(\lim_{n\to\infty} v_n).$$

A linear operator is continuous if and only if it is *bounded*. The latter means that there exists a constant C such that, for any $v \in V$,

$$||Tv|| < C||v||.$$

Here is a proof of the equivalence of these definitions. Suppose that T is continuous but not bounded. Then we can find a sequence of vectors v_n such that $||Tv_n|| > n||v_n||, n = 1, \ldots$. Then $w_n = v_n/n||v_n||$ form a sequence convergent to 0. However, since $||Tw_n|| > 1$, the sequence $\{Tw_n\}$ is not convergent to $T(0) = 0$. The converse is obvious.

We can define the *norm* of a bounded operator by

$$||T|| = \sup\{||Tv||/||v|| : v \in V \setminus \{0\}\} = \sup\{||Tv|| : ||v|| = 1\}.$$

With this norm, the set of bounded linear operators $\mathcal{L}(V)$ on V becomes a normed algebra.

A *linear functional* on V is a continuous linear function $\ell : V \to \mathbb{C}$. Each such functional can be written in the form

$$\ell(v) = (v, w)$$

for a unique vector $w \in V$. In fact, if we choose an orthonormal basis (e_n) in V and put $a_n = \ell(e_n)$, then

$$\ell(v) = \ell(\sum_{n=1}^{\infty} c_n e_n) = \sum_{n=1}^{\infty} c_n \ell(e_n) = \sum_{n=1}^{\infty} c_n a_n = (v, w)$$

where $w = \sum_{n=1}^{\infty} \bar{a}_n e_n$. Note that we have used the assumption that ℓ is continuous. We denote the linear space of linear functionals on V by V^*.

A linear operator T^* is called *adjoint* to a bounded linear operator T if, for any $v, w \in V$,

$$(Tv, w) = (v, T^*w). \tag{5.8}$$

Since T is bounded, by the Cauchy-Schwarz inequality,

$$|(Tv, w)| \leq ||Tv||||w|| \leq C||v||||w||.$$

This shows that, for a fixed w, the function $v \to (Tv, w)$ is continuous, hence there exists a unique vector x such that $(Tv, w) = (v, x)$. The map $w \mapsto x$ defines the adjoint operator T^*. Clearly, it is continuous.

A bounded linear operator is called *self-adjoint* if $T^* = T$. In other words, for any $v, w \in V$,

$$(Tv, w) = (v, Tw). \tag{5.8}$$

An example of a bounded self-adjoint operator is an *orthoprojection operator* P_L onto a closed subspace L of V. It is defined as follows. First of all we have $V = L \oplus L^\perp$. To see this, we consider L as a Hilbert subspace (L is complete because it is closed). Then for any fixed $v \in V$, the function $x \to (v, x)$ is a linear functional on L. Thus there exists $v_1 \in L$ such that $(v, x) = (v_1, x)$ for any $x \in L$. This implies that $v - v_1 \in L^\perp$. Now we can define P_L in the usual way by setting $P_L(v) = x$, where $v = x + y, x \in L, y \in L^\perp$. The boundedness of this operator is obvious since $||P_L v|| \le ||v||$. Clearly P_L is idempotent (i.e. $P_L^2 = P_L$). Conversely each bounded idempotent operator is equal to some orthoprojection operator P_L (take $L = \mathrm{Ker}(P - I_V)$).

Example 1. An example of an operator in $L^2(M, \mu)$ is a *Hilbert-Schmidt operator*:

$$Tf(x) = \int_M f(y) K(x, y) d\mu, \tag{5.9}$$

where $K(x, y) \in L^2(M \times M, \mu \times \mu)$. In this formula we integrate keeping x fixed. By Fubini's theorem, for almost all x, the function $y \to K(x, y)$ is μ-integrable. This implies that $T(f)$ is well-defined (recall that we consider functions modulo functions equal to zero on a set of measure zero). By the Cauchy-Schwarz inequality,

$$|Tf|^2 = \left| \int_M f(y) K(x, y) d\mu \right|^2 \le \int_M |K(x, y)|^2 dg \int_M |f(y)|^2 dy = ||f||^2 \int_M |K(x, y)|^2 d\mu.$$

This implies that

$$||Tf||^2 = \int_M |Tf|^2 d\mu \le ||f||^2 \int_M \int_M |K(x, y)|^2 d\mu d\mu,$$

i.e., T is bounded, and

$$||T||^2 \le \int_M \int_M |K(x, y)|^2 d\mu d\mu.$$

We have

$$(Tf, g) = \int_M (\int_M f(y) K(x, y) d\mu, \bar{g}(x) d\mu) = \int_M \int_M K(x, y) f(y) \bar{g}(x) d\mu d\mu.$$

This shows that the Hilbert-Schmidt operator (5.9) is self-adjoint if and only if

$$K(x, y) = \overline{K(y, x)}$$

outside a subset of measure zero in $M \times M$.

In quantum mechanics we will often be dealing with operators which are defined only on a dense subspace of V. So let us extend the notion of a linear operator by admitting linear maps $D \to V$ where D is a dense linear subspace of V (note the analogy with rational maps in algebraic geometry). For such operators T we can define the adjoint operator as follows. Let $\mathcal{D}(T)$ denote the domain of definition of T. The adjoint operator T^* will be defined on the set

$$\mathcal{D}(T^*) = \{y \in V : \sup_{0 \neq x \in \mathcal{D}(T)} \frac{|\langle T(x), y\rangle|}{||x||} < \infty\}. \tag{5.10}$$

Take $y \in \mathcal{D}(T^*)$. Since $\mathcal{D}(T)$ is dense in V the linear functional $x \to \langle T(x), y\rangle$ extends to a unique bounded linear functional on V. Thus there exists a unique vector $z \in V$ such that $\langle T(x), y\rangle = \langle x, z\rangle$. We take z for the value of T^* at x. Note that $\mathcal{D}(T^*)$ is not necessary dense in V. We say that T is *self-adjoint* if $T = T^*$. We shall always assume that T cannot be extended to a linear operator on a larger set than $\mathcal{D}(\mathcal{T})$. Notice that T cannot be bounded on $\mathcal{D}(T)$ since otherwise we can extend it to the whole V by continuity. On the other hand, a self-adjoint operator $T : V \to V$ is always bounded. For this reason linear operators T with $\mathcal{D}(T) \neq V$ are called *unbounded* linear operators.

Example 2. Let us consider the space $V = L^2(\mathbb{R}, dx)$ and define the operator

$$Tf = if' = i\frac{df}{dx}.$$

Obviously it is defined only for differentiable functions with integrable derivative. The functions $x^n e^{-x^2}$ obviously belong to $\mathcal{D}(T)$. We shall see later in Lecture 9 that the space of such functions is dense in V. Let us show that the operator T is self-adjoint. Let $f \in \mathcal{D}(T)$. Since $f' \in L^2(\mathbb{R}, dx)$,

$$\int_0^t f'(x)\overline{f(x)}dx = |f(t)|^2 - |f(0)|^2 - \int_0^t f(x)\overline{f'(x)}dx$$

is defined for all t. Letting t go to $\pm\infty$, we see that $\lim_{t\to\pm\infty} f(t)$ exists. Since $|f(x)|^2$ is integrable over $(-\infty, +\infty)$, this implies that this limit is equal to zero. Now, for any $f, g \in \mathcal{D}(T)$, we have

$$(Tf, g) = \int_0^t if'(x)\overline{g(x)}dx = if(t)\overline{g(x)}\Big|_{-\infty}^{+\infty} - \int_0^t if(x)\overline{g'(x)}dx =$$

$$= \int_0^t f(x)\overline{ig'(x)}dx = (f, Tg).$$

This shows that $\mathcal{D}(T) \subset \mathcal{D}(T^*)$ and T^* is equal to T on $\mathcal{D}(T)$. The proof that $\mathcal{D}(T) = \mathcal{D}(T^*)$ is more subtle and we omit it (see [**Jordan**]).

5.5 Now we are almost ready to define the new algebra of observables. This will be the algebra $\mathcal{H}(V)$ of self-adjoint operators on a Hilbert space V. Here the Jordan product is defined by the formula $A * B = \frac{1}{4}[(A+B)^2 - (A-B)^2]$ if A, B are bounded self-adjoint operators. If A, B are unbounded, we use the same formula if the intersection of the domains of A and B is dense. Otherwise we set $A * B = 0$. Similarly, we define the Poisson bracket

$$\{A, B\}_\hbar = \frac{i}{\hbar}(AB - BA).$$

Here we have to assume that the intersection of $B(V)$ with the domain of A is dense, and similarly for $A(V)$. Otherwise we set $\{A, B\}_\hbar = 0$.

We can define states in the same way as we defined them before: positive real-valued linear functionals $A \to \langle A|\omega\rangle$ on the algebra $\mathcal{H}(V)$ normalized with the condition $\langle 1_V|\omega\rangle = 1$. To be more explicit we would like to have, as in the finite-dimensional case, that

$$\langle A|\omega\rangle = Tr(MA)$$

for a unique non-negative definite self-adjoint operator. Here comes the difficulty: the trace of an arbitrary bounded operator is not defined.

We first try to generalize the notion of the trace of a bounded operator. By analogy with the finite-dimensional case, we can do it as follows. Choose an orthonormal basis (e_n) and set

$$Tr(T) = \sum_{n=1}^{\infty}(Te_n, e_n). \tag{5.11}$$

We say that T is *nuclear* if this sum absolutely converges for some orthonormal basis.

By the Cauchy-Schwarz inequality

$$\sum_{n=1}^{\infty}|(Te_n, e_n)| \leq \sum_{n=1}^{\infty}||Te_n||||e_n|| = \sum_{n=1}^{\infty}||Te_n||.$$

So

$$\sum_{n=1}^{\infty}||Te_n|| < \infty. \tag{5.12}$$

for some orthonormal basis (e_n) implies that T is nuclear. But the converse is not true in general.

Let us show that (5.11) is independent of the choice of a basis. Let (e'_n) be another orthonormal basis. It follows from (5.7), by writing $Te_n = \sum_k a_k e'_k$, that

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(Te_n, e'_m)(e'_m, e_n) = \sum_{n=1}^{\infty}(Te_n, e_n)$$

and hence the left-hand side does not depend on (e'_m). On the other hand,

$$\sum_{n=1}^{\infty}(Te_n, e_n) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(Te_n, e'_m)(e'_m, e_n) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(e_n, T^*e'_m)(e'_m, e_n) =$$

$$\overline{\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(T^*e'_m, e_n)(e_n, e'_m))} = \sum_{n=1}^{\infty}\overline{(T^*e'_n, e'_n)}$$

and hence does not depend on (e_n) and is equal to $\overline{Tr(T^*)}$. This obviously proves the assertion. It also proves that T is nuclear if and only if T^* is nuclear.

Now let us see that, for any bounded linear operators A, B,

$$Tr(AB) = Tr(BA), \tag{5.13}$$

provided that both sides are defined. We have, for any two orthonormal bases $(e_n), (e'_n)$,

$$Tr(AB) = \sum_{n=1}^{\infty}(ABe_n, e_n) = \sum_{n=1}^{\infty}(Be_n, A^*e_n) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(Be_n, e'_m)(e'_m, A^*e_n) =$$

$$= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(Be_n, e'_m)(Ae'_m, e_n) = Tr(BA).$$

As in the Lemma from the previous lecture, we want to define a state ω as a non-negative linear real-valued function $A \to \langle A|\omega\rangle$ on the algebra of observables $\mathcal{H}(V)$ normalized by the condition $\langle I_V|\omega\rangle = 1$. We have seen that in the finite-dimensional case, such a function looks like $\langle A|\omega\rangle = Tr(MA)$ where M is a unique density operator, i.e. a self-adjoint and non-negative (i.e., $(Mv, v) \geq 0$) operator with $Tr(M) = 1$. In the general case we have to take this as a definition. But first we need the following:

Lemma. *Let M be a self-adjoint bounded non-negative nuclear operator. Then for any bounded linear operator A both $Tr(MA)$ and $Tr(AM)$ are defined.*

Proof. We only give a sketch of the proof, referring to Exercise 8 for the details. The main fact is that V admits an orthonormal basis (e_n) of eigenvectors of M. Also the infinite sum of the eigenvalues λ_n of e_n is absolutely convergent. From this we get that

$$\sum_{n=1}^{\infty}|(MAe_n, e_n)| = \sum_{n=1}^{\infty}|(Ae_n, Me_n)| = \sum_{n=1}^{\infty}|(Ae_n, \lambda_n e_n)| = \sum_{n=1}^{\infty}|\lambda_n||(Ae_n, e_n)| \leq$$

$$\leq \sum_{n=1}^{\infty}|\lambda_n|||Ae_n|| \leq C\sum_{n=1}^{\infty}|\lambda_n| < \infty.$$

Similarly, we have

$$\sum_{n=1}^{\infty}|(AMe_n, e_n)| = \sum_{n=1}^{\infty}|(A\lambda_n e_n, e_n)| = \sum_{n=1}^{\infty}|\lambda_n||(Ae_n, e_n)| \leq C\sum_{n=1}^{\infty}|\lambda_n| < \infty.$$

Definition. A *state* is a linear real-valued non-negative function ω on the space $\mathcal{H}(V)_{bd}$ of bounded self-adjoint operators on V defined by the formula

$$\langle A|\omega\rangle = Tr(MA)$$

where M is a self-adjoint bounded non-negative nuclear operator with $Tr(M) = 1$. The operator M is called the *density operator* of the state ω. The state corresponding to the orthoprojector $M = P_{\mathbb{C}v}$ to a one-dimensional subspace is called a *pure state*.

We can extend the function $A \to \langle A|\omega\rangle$ to the whole of $\mathcal{H}(V)$ if we admit infinite values. Note that for a pure state with density operator $P_{\mathbb{C}v}$ the value $\langle A|\omega\rangle$ is infinite if and only if A is not defined at v.

Example 3. Any Hilbert-Schmidt operator from Example 1 is nuclear. We have

$$Tr(T) = \int_M K(x,x)d\mu.$$

It is non-negative if we assume that $K(x,y) \geq 0$ outside of a subset of measure zero.

5.6 Recall that a self-adjoint operator T in a finite-dimensional Hilbert space V always has eigenvalues. They are all real and V admits an orthonormal basis of eigenvectors of T. This is called the Spectral Theorem for a self-adjoint operator. There is an analog of this theorem in any Hilbert space V. Note that not every self-adjoint operator has eigenvalues. For example, consider the bounded linear operator $Tf = e^{-x^2}f$ in $L^2(\mathbb{R}, dx)$. Obviously this operator is self-adjoint. Then $Tf = \lambda f$ implies $(e^{-x^2} - \lambda)f = 0$. Since $\phi(x) = e^{-x^2} - \lambda$ has only finitely many zeroes, $f = 0$ off a subset of measure zero. This is the zero element in $L^2(M, \mu)$.

First we must be careful in defining the spectrum of an operator. In the finite- dimensional case, the following are equivalent:
(i) $Tv = \lambda v$ has a non-trivial solution;
(ii) $T - \lambda I_V$ is not invertible.
In the general case these two properties are different.

Example 4. Consider the space $V = l_2(\mathbb{C})$ and the (shift) operator T defined by

$$T(a_1, a_2, \ldots, a_n, \ldots) = (0, a_1, a_2, \ldots).$$

Clearly it has no inverse, in fact its image is a closed subspace of V. But $Tv = \lambda v$ has no non-trivial solutions for any λ.

Definition. Let $T : V \to V$ be a linear operator. We define
Discrete spectrum: the set $\{\lambda \in \mathbb{C} : Ker(T - \lambda I_V) \neq \{0\}\}$. Its elements are called *eigenvalues* of T.
Continuous spectrum: the set $\{\lambda \in \mathbb{C} : (T - \lambda I_V)^{-1}$ exists as an unbounded operator $\}$.
Residual spectrum: all the remaining λ for which $T - \lambda I_V$ is not invertible.
Spectrum: the union of the previous three sets.

Thus in Example 4, the discrete spectrum and the continuous spectrum are each the empty set. The residual spectrum is $\{0\}$.

Example 5. Let $T : L^2([0,1], dx) \to L^2([0,1], dx)$ be defined by the formula

$$Tf = xf.$$

Let us show that the spectrum of this operator is continuous and equals the set $[0, 1]$. In fact, for any $\lambda \in [0, 1]$ the function 1 is not in the image of $T - \lambda I_V$ since otherwise the improper integral

$$\int_0^1 \frac{dx}{(x-\lambda)^2}$$

converges. On the other hand, the functions $f(x)$ which are equal to zero in a neighborhood of λ are divisible by $x - \lambda$. The set of such functions is a dense subset in V (recall that we consider functions to be identical if they agree off a subset of measure zero). Since the operator $T - \lambda I_V$ is obviously injective, we can invert it on a dense subset of V. For any $\lambda \notin [0, 1]$ and any $f \in V$, the function $f(x)/(x - \lambda)$ obviously belongs to V. Hence such λ does not belong to the spectrum.

For any eigenvalue λ of T we define the *eigensubspace* $E_\lambda = \mathrm{Ker}(T - \lambda I_V)$. Its non-zero vectors are called *eigenvectors* with eigenvalue λ. Let E be the direct sum of these subspaces. It is easy to see that E is a closed subspace of V. Then $V = E \oplus E^\perp$. The subspace $E^\perp$ is T-invariant. The spectrum of $T|E$ is the discrete spectrum of T. The spectrum of $T|E^\perp$ is the union of the continuous and residual spectrum of T. In particular, if $V = E$, the space V admits an orthonormal basis of eigenvectors of T.

The operator from the previous example is self-adjoint but does not have eigenvalues. Nevertheless, it is possible to state an analog of the spectral theorem for self-adjoint operators in the general case.

For any vector v of norm 1, we denote by P_v the orthoprojector $P_{\mathbb{C}v}$. Suppose V has an orthonormal basis of eigenvectors e_n of a self-adjoint operator A with eigenvalues λ_n. For example, this would be the case when V is finite-dimensional. For any

$$v = \sum_{n=1}^{\infty} a_n e_n = \sum_{n=1}^{\infty} (v, e_n) e_n,$$

we have

$$Av = \sum_{n=1}^{\infty} (v, e_n) A e_n = \sum_{n=1}^{\infty} \lambda_n (v, e_n) = \sum_{n=1}^{\infty} \lambda_n P_{e_n} v.$$

This shows that our operator can be written in the form

$$A = \sum_{n=1}^{\infty} \lambda_n P_{e_n},$$

where the sum is convergent in the sense of the operator norm.

In the general case, we cannot expect that V has a countable basis of eigenvectors (see Example 5 where the operator T is obviously self-adjoint and bounded). So the sum above should be replaced with some integral. First we extend the notion of a real-valued measure on $\mathbb{R}$ to a measure with values in $\mathcal{H}(V)$ (in fact one can consider measures with values in any Banach algebra = complete normed ring). An example of such a measure is the map $E \to \sum_{\lambda_n \in E} P_{e_n} \in \mathcal{H}(V)$ where we use the notation from above.

Now we can state the Spectral Theorem for self-adjoint operators.

Theorem 2. *Let A be a self-adjoint operator in V. There exists a function $P_A : \mathbb{R} \to \mathcal{H}(V)$ satisfying the following properties:*

(i) for each $\lambda \in \mathbb{R}$, the operator $P_A(\lambda)$ is a projection operator;

(ii) $P_A(\lambda) \leq P_A(\lambda')$ (in the sense that $P_A(\lambda)P_A(\lambda') = P_A(\lambda)$ if $\lambda < \lambda'$);

(iii) $\lim_{\lambda\to-\infty} P_A(\lambda) = 0, \lim_{\lambda\to+\infty} P_A(\lambda) = 1_V$;

(iv) $P_A(\lambda)$ defines a unique $\mathcal{H}(V)$-measure μ_A on $\mathbb{R}$ with $\mu_A((-\infty,\lambda)) = P_A(\lambda)$;

(v)

$$A = \int_{\mathbb{R}} x d\mu_A$$

(vi) for any $v \in V$,

$$(Av, v) = \int_{\mathbb{R}} x d\mu_{A,v}$$

where $\mu_{A,v}(E) = (\mu_A(E)v, v)$.

(viii) λ belongs to the spectrum of A if and only if $P_A(x)$ increases at λ;

(vii) a real number λ is an eigenvalue of A if and only if $\lim_{x\to\lambda^-} P_A(x) < \lim_{x\to\lambda^+} P_A(x)$;

(viii) the spectrum is a bounded subset of $\mathbb{R}$ if and only if A is bounded.

Definition. The function $P_A(\lambda)$ is called the *spectral function* of the operator A.

Example 6. Suppose V has an orthonormal basis (e_n) of eigenvectors of a self-adjoint operator A. Let λ_n be the eigenvalue of e_n. Then the spectral function is

$$P_A(\lambda) = \sum_{\lambda_n \leq \lambda} P_{e_n}.$$

Example 7. Let $V = L^2([0,1], dx)$. Consider the operator $Af = gf$ for some μ-integrable function g. Consider the $\mathcal{H}(V)$-measure defined by $\mu(E)f = \phi_E f$, where $\phi_E(x) = 1$ if $g(x) \in E$ and $\phi_E(x) = 0$ if $g(x) \notin E$. Let $|g(x)| < C$ outside a subset of measure zero. We have

$$\lim_{n\to\infty} \sum_{i=1}^{2n} \frac{(-n+i-1)C}{n} \phi_{[\frac{(-n+i-1)C}{n}, \frac{(-n+i)C}{n}]} = g(x).$$

This gives

$$\Big(\int_{\mathbb{R}} x d\mu\Big) f = \lim_{n\to\infty} \sum_{i=1}^{2n} \frac{(-n+i-1)C}{n} \phi_{[\frac{(-n+i-1)C}{n}, \frac{(-n+i)C}{n}]} f = gf.$$

This shows that the spectral function $P_A(\lambda)$ of A is defined by $P_A(\lambda)f = \mu((-\infty,\lambda])f = \theta(\lambda - g)f$, where θ is the Heaviside function (4.4).

An important class of linear bounded operators is the class of compact operators.

Definition. A bounded linear operator $T : V \to V$ is called *compact* (or *completely continuous*) if the image of any bounded set is a relatively compact subset (equivalently, if the image of any bounded sequence contains a convergent subsequence).

For example any bounded operator whose image is a finite-dimensional is obviously compact (since any bounded subset in a finite-dimensional space is relatively compact).

We refer to the Exercises for more information on compact operators. The most important fact for us is that any density operator is a compact operator (see Exercise 9).

Theorem 3(Hilbert-Schmidt). *Assume that T is a compact operator and let $Sp(T)$ be its spectrum. Then $Sp(T)$ consists of eigenvalues of T. In particular, each $v \in H$ can be written in the form*

$$v = \sum_{n=1}^{\infty} c_n e_n$$

where e_n is a linearly independent set of eigenvectors of T.

5.7 The spectral theorem allows one to define a function of a self-adjoint operator. We set

$$f(A) = \int_{\mathbb{R}} f(x) d\mu_A.$$

The domain of this operator consists of vectors v such that

$$\int_{\mathbb{R}} |f(x)|^2 d\mu_{A,v} < \infty.$$

Take $f = \theta(\lambda - x)$ where θ is the Heaviside function. Then

$$\theta(\lambda - x)(A) = \int_{\mathbb{R}} \theta(\lambda - x) d\mu_A = \int_{-\infty}^{\lambda} d\mu_A = P_A(\lambda).$$

Thus, if M is the density operator of a state ω, we obtain

$$\langle \theta(\lambda - x)(A) | \omega \rangle = Tr(M P_A(\lambda)).$$

Comparing this with section 4.4 from Lecture 4, we see that the function

$$\mu_A(\lambda) = Tr(M P_A(\lambda))$$

is the distribution function of the observable A in the state ω. If (e_n) is a basis of V consisting of eigenvectors of A with eigenvalues λ_n, then we have

$$\mu_A(\lambda) = \sum_{n:\lambda_n \leq \lambda} (M e_n, e_n).$$

This formula shows that the probability that A takes the value λ at the state ω is equal to zero unless λ is equal to one of the eigenvalues of A.

Now let us consider the special case when M is a pure state P_v. Then

$$\langle A|\omega\rangle = Tr(P_v A) = Tr(AP_v) = (Av, v),$$

$$\mu_A(\lambda) = (P_A(\lambda)v, v).$$

In particular, A takes the eigenvalue λ_n at the pure state P_{e_n} with probability 1.

5.9 Now we have everything to extend the Hamiltonian dynamics in $\mathcal{H}(V)$ to any Hilbert space word by word. We consider $A(t)$ as a differentiable path in the normed algebra of bounded self-adjoint operators. The solution of the Schrödinger picture

$$\frac{dA(t)}{dt} = \{A, H\}_\hbar, \quad A(0) = A$$

is given by

$$A(t) = A_t := U(-t)AU(t), \quad A(0) = A,$$

where $U(t) = e^{\frac{i}{\hbar}Ht}$.

Let $v(t)$ be a differential map to the metric space V. The Schrödinger equation is

$$i\hbar\frac{dv(t)}{dt} = Hv(t), \quad v(0) = v. \tag{5.14}$$

It is solved by

$$v(t) = v_t := U(t)v = e^{\frac{i}{\hbar}Ht}v = \Big(\int_{\mathbb{R}} e^{\frac{ixt}{\hbar}} d\mu_H\Big)v.$$

In particular, if V admits an orthonormal basis (e_n) consistsing of eigenvectors of H, then

$$v_t = \sum_{n=1}^{\infty} e^{it\lambda_n/\hbar}(e_n, v)v. \tag{5.15}$$

Exercises.

1. Show that the norm of a bounded operator A in a Hilbert space is equal to the norm of its adjoint operator A^*.

2. Let A be a bounded linear operator in a Hilbert space V with $||A|| < |\alpha|^{-1}$ for some $\alpha \in \mathbb{C}$. Show that the operator $(A + \alpha I_V)^{-1}$ exists and is bounded.

3. Use the previous problem to prove the following assertions:
 (i) the spectrum of any bounded linear operator is a closed subset of $\mathbb{C}$;
 (ii) if λ belongs to the spectrum of a bounded linear operator A, then $|\lambda| < ||A||$.

4. Prove the following assertions about compact operators:
 (i) assume A is compact, and B is bounded; then A^*, BA, AB are compact.

(ii) a compact operator has no bounded inverse if V is infinite-dimensional;

(iii) the limit (with respect to the norm metric) of compact operators is compact;

(iv) the image of a compact operator T is a closed subspace and its cokernel $V/T(V)$ is finite-dimensional.

5.(i) Let $V = l_2(\mathbb{C})$ and $A(x_1, x_2, \ldots, x_n, \ldots) = (a_1x_1, a_2x_2, \ldots, a_nx_n, \ldots)$. For which $(a_1, \ldots, a_n, \ldots)$ is the operator A compact?

(ii) Is the identity operator compact?

6. Show that a bounded linear operator T satisfying $\sum_{n=1}^{\infty} ||Te_i||^2 < \infty$ for some orthonormal basis (e_n) is compact. [Hint: Write T as the limit of operators T_m of the form $\sum_{k=1}^{m} P_{e_k}T$ and use Exercise 4 (iii).]

7. Using the previous problem, show that the Hilbert-Schmidt integral operator is compact.

8.(i) Using Exercise 6, prove that any bounded non-negative nuclear operator T is compact. [Hint: Use the Spectral theorem to represent T in the form $T = A^2$.]

(ii) Show that the series of eigenvalues of a compact nuclear operator is absolutely convergent.

(iii) Using the previous two assertions prove that MA is nuclear for any density operator M and any bounded operator A.

9. Using the following steps prove the Hilbert-Schmidt Theorem:

(i) If $v = \lim_{n\to\infty} v_n$, then $\lim_{n\to\infty}(Av_n, v_n) = (Av, v)$;

(ii) If $|(Av, v))|$ achieves maximum at a vector v_0 of norm 1, then v_0 is an eigenvector of A;

(iii) Show that $|(Av, v)|$ achieves its maximum on the unit sphere in V and the corresponding eigenvector v_1 belongs to the eigenvalue with maximal absolute value.

(iv) Show that the orthogonal space to v_1 is T-invariant, and contains an eigenvector of T.

10. Consider the linear operator in $L^2([0, 1], dx)$ defined by the formula

$$Tf(x) = \int_0^x f(t)dt.$$

Prove that this operator is bounded self-adjoint. Find its spectrum.

Lecture 6. CANONICAL QUANTIZATION

6.1 Many quantum mechanical systems arise from classical mechanical systems by a process called a quantization. We start with a classical mechanical system on a symplectic manifold (M, ω) with its algebra of observables $\mathcal{O}(M)$ and a Hamiltonian function H. We would like to find a Hilbert space V together with an injective map of algebras of observables

$$\mathcal{Q}_1 : \mathcal{O}(M) \to \mathcal{H}(V), \quad f \mapsto A_f,$$

and an injective map of the corresponding sets of states

$$\mathcal{Q}_2 : \mathcal{S}(M) \to \mathcal{S}(V), \quad \rho \mapsto M_f.$$

The image of the Hamiltonian observable will define the Schrödinger operator H in terms of which we can describe the quantum dynamics. The map $\mathcal{Q}_1$ should be a homomorphism of Jordan algebras but not necessarily of Poisson algebras. But we would like to have the following property:

$$\lim_{\hbar \to 0} A_{\{f,g\}} = \lim_{\hbar \to 0} \{A_f, A_g\}_\hbar, \quad \lim_{\hbar \to 0} Tr(M_\rho A_f) = (f, \mu_\rho).$$

The main objects in quantum mechanics are some particles, for example, the electron. They are not described by their exact position at a point (x, y, z) of $\mathbb{R}^3$ but rather by some numerical function $\psi(x, y, z)$ (the *wave function*) such that $|\psi(x, y, z)|^2$ gives the distribution density for the probability to find the particle in a small neighborhood of the point (x, y, z). It is reasonable to assume that

$$\int_{\mathbb{R}^3} |\psi(x, y, z)|^2 dxdydz = 1, \quad \lim_{||(x,y,z)|| \to \infty} \psi(x, y, z) = 0.$$

This suggests that we look at ψ as a pure state in the space $L^2(\mathbb{R}^3)$. Thus we take this space as the Hilbert space V. When we are interested in describing not one particle but many, say N of them, we are forced to replace $\mathbb{R}^3$ by $\mathbb{R}^{3N}$. So, it is better to take now

$V = L^2(\mathbb{R}^n)$. As is customary in classical mechanics we continue to employ the letters q_i to denote the coordinate functions on $\mathbb{R}^n$. Then classical observables become functions on $M = T(\mathbb{R}^n)^*$ where we use the canonical coordinates q_i, p_i. In this lecture we are discussing only one possible approach to quantization (*canonical quantization*). There are others which apply to an arbitrary symplectic manifold (M, ω) (*geometric quantizations*).

6.2 Let us first define the operators corresponding to the coordinate functions q_i, p_i. We know from (2.10) that

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij}.$$

Let

$$Q_i = A_{q_i}, P_i = A_{p_i}, \quad i = 1, \ldots, n.$$

Then we must have

$$[Q_i, Q_j] = [P_i, P_j] = 0, \quad [P_i, Q_j] = -i\hbar\delta_{ij}. \tag{6.1}$$

So we have to look for such a set of $2n$ operators.

Define

$$Q_i\phi(\mathbf{q}) = q_i\phi(\mathbf{q}), \quad P_i\phi(\mathbf{q}) = -i\hbar\frac{\partial}{\partial q_i}\phi(\mathbf{q}), \quad i = 1, \ldots, n. \tag{6.2}$$

The operators Q_i are defined on the dense subspace of functions which go to zero at infinity faster than $||\mathbf{q}||^{-2}$ (e.g., functions with compact support). The second group of operators is defined on the dense subset of differentiable functions whose partial derivatives belong to $L^2(\mathbb{R}^n)$. It is easy to show that these operators satisfy $\mathcal{D}(T) \subset \mathcal{D}(T^*)$ and $T^* = T$ on $\mathcal{D}(T)$. For the operators Q_i this is obvious, and for operators P_i the proof is similar to that in Example 2 from Lecture 5. We skip the proof that $\mathcal{D}(T) = \mathcal{D}(T^*)$ (see [**Jordan**]).

Now let us check (6.1). Obviously $[Q_i, Q_j] = [P_i, P_j] = 0$. We have

$$P_iQ_j\phi(x) = -i\hbar\frac{\partial x_j\phi(x)}{\partial x_i} = -i\hbar(\delta_{ij}\phi(x) + x_j\frac{\partial\phi(x)}{\partial x_i}) - (-i\hbar x_j\frac{\partial\phi(x)}{\partial x_i}) = -i\hbar\delta_ij\phi(x).$$

The operators Q_i (resp. P_i) are called the *position* (resp. *momentum*) operators.

6.3 So far, so good. But we still do not know what to associate to any function $f(\mathbf{p}, \mathbf{q})$. Although we know the definition of a function of one operator, we do not know how to define the function of several operators. To do this we use the following idea. Recall that for a function $f \in L^2(\mathbb{R})$ one can define the *Fourier transform*

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)e^{-ixt}dt. \tag{6.3}$$

There is the *inversion formula*

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{f}(t)e^{ixt}dt. \tag{6.4}$$

Assume for a moment that we can give a meaning to this formula when x is replaced by some operator X on $L^2(\mathbb{R})$. Then

$$f(X) = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \hat{f}(t) e^{iXt} dt. \tag{6.5}$$

This agrees with the formula

$$f(X) = \int\limits_{\mathbb{R}} f(\lambda) d\mu_X(\lambda)$$

we used in the previous Lecture. Indeed,

$$\frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \hat{f}(t) e^{iXt} dt = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \hat{f}(t) \Big(\int\limits_{\mathbb{R}} e^{ixt} d\mu_X \Big) dt =$$

$$= \Big(\frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \hat{f}(t) e^{ixt} dt \Big) \Big) d\mu_X = \int\limits_{\mathbb{R}} f(x) d\mu_X = f(X).$$

To define (6.5) and also its generalization to functions in several variables requires more tools. First we begin with reminding the reader of the properties of Fourier transforms. We start with functions in one variable. The Fourier transform of a function absolutely integrable over $\mathbb{R}$ is defined by formula (6.3). It satisfies the following properties:

(F1) $\hat{f}(x)$ is a bounded continuous function which goes to zero when $|x|$ goes to infinity;

(F2) if f'' exists and is absolutely integrable over $\mathbb{R}$, then $\hat{f}$ is absolutely integrable over $\mathbb{R}$ and

$$f(-x) = \hat{\hat{f}}(x) \qquad (\textit{Inversion Formula});$$

(F3) if f is absolutely continuous on each finite interval and f' is absolutely integrable over $\mathbb{R}$, then

$$\widehat{f'}(x) = ix\hat{f}(x);$$

(F4) if $f, g \in L^2(\mathbb{R})$, then

$$(\hat{f}, \hat{g}) = (f, g) \qquad (\textit{Plancherel Theorem});$$

(F5) If

$$g(x) = f_1 * f_2 := \int\limits_{\mathbb{R}} f_1(t) f_2(x-t) dt$$

then

$$\hat{g} = \hat{f}_1 \cdot \hat{f}_2.$$

Notice that property (F3) tells us that the Fourier transform $\mathcal{F}$ maps the domain of definition of the operator $P = f \to -if'$ onto the domain of definition of the operator $Q = f \to xf$. Also we have

$$\mathcal{F}^{-1} \circ Q \circ \mathcal{F} = P. \tag{6.6}$$

6.4 Unfortunately the Fourier transformation is not defined for many functions (for example, for constants). To extend it to a larger class of functions we shall define the notion of a distribution (or a generalized function) and its Fourier transform.

Let K be the vector space of smooth complex-valued functions with compact support on $\mathbb{R}$. We make it into a topological vector space by defining a basis of open neighborhoods of 0 by taking the sets of functions $\phi(x)$ with support contained in the same compact subset and satisfying $|\phi^{(k)}(x)| < \epsilon_k$ for all $x \in \mathbb{R}, k = 0, 1, \ldots$.

Definition. A *distribution* (or a *generalized function*) on $\mathbb{R}$ is a continuous complex valued linear function on K. A distribution is called *real* if

$$\overline{\ell(\phi)} = \ell(\bar{\phi}) \qquad \text{for any } \phi \in K.$$

Let K' be the space of complex-valued functions which are absolutely integrable over any finite interval (*locally integrable functions*). Obviously, $K \subset K'$. Any function $f \in K'$ defines a distribution by the formula

$$\ell_f(\phi) = \int_{\mathbb{R}} \bar{f}(x)\phi(x)dx = (\phi, f)_{L^2(\mathbb{R})}. \tag{6.7}$$

Such distributions are called *regular*, the rest are *singular* distributions.

Since $(\bar{\phi}, \bar{f}) = \overline{(\phi, f)}$, we obtain that a regular distribution is real if and only if $f = \bar{f}$, i.e., f is a real-valued function.

It is customary to denote the value of a distribution $\ell \in K^*$ as in (6.7) and view $f(x)$ (formally) as a generalized function. Of course, the value of a generalized function at a point x is not defined in general. For simplicity we shall assume further, unless stated otherwise, that all our distributions are real. This will let us forget about the conjugates in all formulas. We leave it to the reader to extend everything to the complex-valued distributions.

For any affine linear transformation $x \to ax + b$ of $\mathbb{R}$, and any $f \in K'$, we have

$$\int_{\mathbb{R}} f(t)\phi(a^{-1}t + ba^{-1})dt = a\int_{\mathbb{R}} f(ax+b)\phi(x)dx.$$

This allows us to define an affine change of variables for a generalized function:

$$\phi \to \frac{1}{a}\ell(\phi(a^{-1}x + ba^{-1})) := \int_{\mathbb{R}} f(ax+b)\phi(x)dx. \tag{6.8}$$

Examples. 1. Let μ be a measure on $\mathbb{R}$, then

$$\phi(x) \to \int_{\mathbb{R}} \phi(x)d\mu$$

is a real-valued distribution (called positive distribution). It can be formally written in the form

$$\phi(x) \to \int\limits_{\mathbb{R}} \rho(x)\phi(x)dx,$$

where ρ is called the *density distribution.*

For example, we can introduce the distribution

$$\ell(\phi) = \phi(0).$$

It is called the *Dirac δ-function.* It is clearly real. Its formal expression (6.7) is

$$\phi \to \int\limits_{\mathbb{R}} \delta(x)\phi(x)dx.$$

More generally, the functional $\phi \to \phi(a)$ can be written in the form

$$\ell(\phi) = \int\limits_{\mathbb{R}} \delta(x-a)\phi(x)dx = \int\limits_{\mathbb{R}} \delta(x)\phi(x-a)dx.$$

The precise meaning for $\delta(x-a)$ is explained by (6.8).

2. Let $f(x) = 1/x$. Obviously it is not integrable on any interval containing 0. So

$$\ell(\phi) = \int\limits_{\mathbb{R}} \frac{1}{x}\phi(x)dx.$$

should mean a singular distribution. Let $[a,b]$ contain the support of ϕ. If $0 \notin [a,b]$ the right-hand side is well-defined. Assume $0 \in [a,b]$. Then, we write

$$\int\limits_{\mathbb{R}} \frac{1}{x}\phi(x)dx = \int_a^b \frac{1}{x}\phi(x)dx = \int_a^b \frac{\phi(x)-\phi(0)}{x}dx + \int_a^b \frac{\phi(0)}{x}dx.$$

The function $\frac{\phi(x)-\phi(0)}{x}$ has a removable discontinuity at 0, so the first integral exists. The second integral exists in the sense of Cauchy's principal value

$$\int_a^b \frac{\phi(0)}{x}dx := \lim_{\epsilon\to 0}\Big(\int\limits_a^{-\epsilon} \frac{\phi(0)}{x}dx + \int\limits_{\epsilon}^{b} \frac{\phi(0)}{x}dx\Big).$$

Now our intergral makes sense for any $\phi \in K$. We take it as the definition of the generalized function $1/x$.

Obviously distributions form an $\mathbb{R}$-linear subspace in K^*. Let us denote it by $\mathcal{D}(\mathbb{R})$. We can introduce the topology on $\mathcal{D}(\mathbb{R})$ by using pointwise convergence of linear functions.

By considering regular distributions, we define a linear map $K' \to \mathcal{D}(\mathbb{R})$. Its kernel is composed of functions which are equal to zero outside of a subset of measure zero. Also notice that $\mathcal{D}(\mathbb{R})$ is a module over the ring K. In fact we can make it a module over the larger ring K'_{sm} of smooth locally integrable functions. For any $\alpha \in K'_{sm}$ we set

$$(\alpha\ell)(\phi) = \ell(\alpha\phi).$$

Although we cannot define the value of a generalized function f at a point a, we can say what it means that f vanishes in an open neighborhood of a. By definition, this means that, for any function $\phi \in K$ with support outside of this neighborhood, $\int_{\mathbb{R}} f\phi dx = 0$. The *support* of a generalized function is the subset of points a such that f does not vanish in any neighborhood of a. For example, the support of the Dirac delta function $\delta(x-a)$ is equal to the set $\{a\}$.

We define the derivative of a distribution ℓ by setting

$$\frac{d\ell}{dx}(\phi) := -\ell(\frac{d\phi}{dx}).$$

In notation (6.7) this reads

$$\int_{\mathbb{R}} f'(x)\phi(x)dx := -\int_{\mathbb{R}} f(x)\phi'(x)dx. \tag{6.9}$$

The reason for the minus sign is simple. If f is a regular differentiable distribution, we can integrate by parts to get (6.9).

It is clear that a generalized function has derivatives of any order.

Examples. 3. Let us differentiate the Dirac function $\delta(x)$. We have

$$\int_{\mathbb{R}} \delta'(x)\phi(x)dx := -\int_{\mathbb{R}} \delta(x)\phi'(x)dx = -\phi'(0).$$

Thus the derivative is the linear function $\phi \to -\phi'(0)$.

4. Let $\theta(x)$ be the Heaviside function. It defines a regular distribution

$$\ell(\phi) = \int_{\mathbb{R}} \theta(x)\phi(x)dx = \int_0^\infty \phi(x)dx.$$

Obviously, $\theta'(0)$ does not exist. But, considered as a distribution, we have

$$\int_{\mathbb{R}} \theta'(x)\phi(x)dx = -\int_0^\infty \phi'(x)dx = \phi(0).$$

Thus

$$\theta'(x) = \delta(x). \tag{6.10}$$

So we know the derivative and the anti-derivative of the Dirac function!

6.5 Let us go back to Fourier transformations and extend them to distributions.

Definition. Let $\ell \in \mathcal{D}(\mathbb{R})$ be a distribution. Its Fourier transform is defined by the formula

$$\hat{\ell}(\phi) = \ell(\psi), \quad \text{where } \hat{\psi}(x) = \phi(x).$$

In other words,

$$\int_{\mathbb{R}} \hat{f}\phi dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)\Big(\int_{\mathbb{R}} \phi(t)e^{itx}dt\Big)dx.$$

It follows from the Plancherel formula that for regular distributions this definition coincides with the usual one.

Examples. 5. Take $f = 1$. Its usual Fourier transform is obviously not defined. Let us view 1 as a regular distribution. Then, letting $\phi = \hat{\psi}$, we have

$$\int_{\mathbb{R}} \hat{1}\phi(x)dx = \int_{\mathbb{R}} 1\psi(x)dx = \int_{\mathbb{R}} \psi(x)e^{i0x}dx = \sqrt{2\pi}\phi(0).$$

This shows that

$$\hat{1} = \sqrt{2\pi}\delta(x).$$

One can view this equality as

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-itx}dt = \sqrt{2\pi}\delta(x). \tag{6.11}$$

In fact, let $f(x)$ be any locally integrable function on $\mathbb{R}$ which has polynomial growth at infinity. The latter means that there exists a natural number m such that $f(x) = (x^2+1)^m f_0(t)$ for some integrable function $f_0(t)$. Let us check that

$$\lim_{\xi\to\infty} \int_{-\xi}^{+\xi} f(t)e^{-ixt}dt$$

exists. Since $f_0(t)$ admits a Fourier transform, we have

$$\frac{1}{\sqrt{2\pi}} \lim_{\xi\to\infty} \int_{-\xi}^{+\xi} f_0(t)e^{-ixt}dt = \hat{f}_0(x),$$

where the convergence is uniform over any bounded interval. This shows that the corresponding distributions converge too. Now applying the operator $(-\frac{d^2}{dx^2}+1)^m$ to both sides, we conclude that

$$\frac{1}{\sqrt{2\pi}} \lim_{\xi\to\infty} \int_{-\xi}^{+\xi} f(t)e^{-ixt}dt = (-\frac{d^2}{dx^2}+1)^m \hat{f}_0(x).$$

We see that $\hat{f}(x)$ can be defined as the limit of distributions $f_\xi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\xi}^{+\xi} f(t)e^{-ixt}dt$.

Let us recompute $\hat{1}$. We have

$$\lim_{\xi\to\infty} \int\limits_{-\xi}^{+\xi} e^{-ixt}dt = \lim_{\xi\to\infty} \frac{e^{-ix\xi} - e^{ix\xi}}{-ix} = \lim_{\xi\to\infty} \frac{2\sin x\xi}{x}.$$

One can show (see **[Gelfand]**, Chapter 2, §2, n° 5, Example 3) that this limit is equal to $2\pi\delta(x)$. Thus we have justified the formula (6.11).

Let $f = e^{iax}$ be viewed as a regular distribution. By the above

$$\widehat{e^{iax}} = \frac{1}{\sqrt{2\pi}} \lim_{\xi\to\infty} \int\limits_{-\xi}^{+\xi} e^{iat}e^{-ixt}dt = \frac{1}{\sqrt{2\pi}} \lim_{\xi\to\infty} \int\limits_{-\xi}^{+\xi} e^{-i(x-a)t}dt = \sqrt{2\pi}\delta(x-a). \tag{6.12}$$

6. Let us compute the Fourier transform of the Dirac function. We have

$$\hat{\delta}(\phi) = \delta(\psi) = \psi(0) = \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \phi(t)e^{i0t}dt.$$

This shows that

$$\hat{\delta} = \frac{1}{\sqrt{2\pi}}.$$

For any two distributions $f, g \in \mathcal{D}(\mathbb{R})$, we can define their *convolution* by the formula

$$\int\limits_{\mathbb{R}} f * g\phi(x)dx = \int\limits_{\mathbb{R}} f\Big(\int\limits_{\mathbb{R}} g\phi(x+y)dx\Big)dy. \tag{6.13}$$

Since, in general, the function $y \to \int\limits_{\mathbb{R}} g\phi(x+y)dx$ does not have compact support, we have to make some assumption here. For example, we may assume that f or g has compact support. It is easy to see that our definition of convolution agrees with the usual definition when f and g are regular distributions:

$$f * g(x) = \int\limits_{\mathbb{R}} f(t)g(x-t)dt.$$

The latter can be also rewritten in the form

$$f * g(x) = \int\limits_{\mathbb{R}} f(t)\big(\int\limits_{\mathbb{R}} \delta(t+t'-x)g(t')dt'\big)dt = \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} f(t)g(t')\delta(t+t'-x)dtdt'. \tag{6.14}$$

This we may take as an equivalent form of writing (8) for any distributions.

Using the inversion formula for the Fourier transforms, we can define the product of generalized functions by the formula

$$f \cdot g(-x) = \widehat{\hat{f} * \hat{g}}. \tag{6.15}$$

The property (F4) of Fourier transforms shows that the product of regular distributions corresponds to the usual product of functions. Of course, the product of two distributions is not always defined. In fact, this is one of the major problems in making rigorous some computations used by physicists.

Example 7. Take f to be a regular distribution defined by an integrable function $f(x)$, and g equal to 1. Then

$$\int\limits_{\mathbb{R}} f * 1\phi(x)dx = \int\limits_{\mathbb{R}} f\Big(\int\limits_{\mathbb{R}} \phi(x+y)dx\Big)dy = \Big(\int\limits_{\mathbb{R}} f(x)dx\Big)\Big(\int\limits_{\mathbb{R}} \phi(x)dx\Big).$$

This shows that

$$f * 1 = \Big(\int\limits_{\mathbb{R}} f(x)dx\Big) \cdot 1.$$

This suggests that we take $f * 1$ as the definition of $\int\limits_{\mathbb{R}} f(x)dx$ for any distribution f with compact support. For example,

$$\int\limits_{\mathbb{R}} \delta(x)\Big(\int\limits_{\mathbb{R}} \phi(x+y)dy\Big)dx = \int\limits_{\mathbb{R}} \phi(y)dy.$$

Hence

$$\int\limits_{\mathbb{R}} \delta(x)dx = \delta * 1 = 1, \tag{6.16}$$

where 1 is considered as a distribution.

6.6 Let $T : K \to K$ be a bounded linear operator. We can extend it to the space of generalized functions $\mathcal{D}(\mathbb{R})$ by the formula

$$T(\ell)(\phi) = \ell(T^*(\phi))$$

or, in other words,

$$\int\limits_{\mathbb{R}} \overline{T(f)}\phi dx = \int\limits_{\mathbb{R}} \bar{f}T^*(\phi)dx).$$

The reason for putting the adjoint operator is clear. The previous formula agrees with the formula

$$(T^*(\phi), f)_{L^2(\mathbb{R})} = (\phi, T(f))_{L^2(\mathbb{R})}$$

for $f, \phi \in K$.

Now we can define a *generalized eigenvector* of T as a distribution ℓ such that $T(\ell) = \lambda\ell$ for some scalar $\lambda \in \mathbb{C}$.

Examples. 8. The operator $\phi(x) \to \phi(x-a)$ has eigenvectors $e^{ixt}, t \in \mathbb{R}$. The corresponding eigenvalue of e^{ixt} is e^{iat}. In fact

$$\int_{\mathbb{R}} e^{ixt}\phi(x-a)dx = e^{iat}\int_{\mathbb{R}} e^{ix}\phi(x)dx.$$

9. The operator $T_g : \phi(x) \to g(x)\phi(x)$ has eigenvectors $\delta(x-a), a \in \mathbb{R}$, with corresponding eigenvalues $g(a)$. In fact,

$$\int_{\mathbb{R}} \delta(x-a)g(x)\phi(x)dx = g(a)\phi(a) = g(a)\int_{\mathbb{R}} \delta(x-a)\phi(x)dx.$$

Let $f_\lambda(x)$ be an eigenvector of an operator A in $\mathcal{D}(\mathbb{R})$. Then, for any $\phi \in K$, we can try to write

$$\phi(x) = \int_{\mathbb{R}} a(\lambda)f_\lambda(x)d\lambda, \tag{6.16}$$

viewing this expression as a decomposition of $\phi(x)$ as a linear combination of eigenvectors of the operator A. Of course this is always possible when V has a countable basis ϕ_n of eigenvectors of A.

For instance, take the operator $T_a : \phi \to \phi(x-a), a \neq 0$, from Example 8. Then we want to have

$$\psi(x) = \int_{\mathbb{R}} a(\lambda)e^{i\lambda x}d\lambda.$$

But, taking

$$a(\lambda) = \frac{1}{2\pi}\hat{\phi}(\lambda) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \psi(x)e^{-i\lambda x}dx\right),$$

we see that this is exactly the inversion formula for the Fourier transform. Note that it is important here that $a \neq 0$, otherwise any function is an eignevector.

Assume that $f_\lambda(x)$ are orthonormal in the following sense. For any $\phi(\lambda) \in K$, and any $\lambda' \in \mathbb{R}$,

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f_\lambda(x)\bar{f}_{\lambda'}(x)dx\Big)\phi(\lambda)d\lambda = \phi(\lambda-\lambda'),$$

or in other words,

$$\int_{\mathbb{R}} f_\lambda(x)\bar{f}_{\lambda'}(x)dx = \delta(\lambda-\lambda').$$

This integral is understood either in the sense of Example 7 or as the limit (in $\mathcal{D}(\mathbb{R})$) of integrals of locally integrable functions (as in Example 5). It is not always defined.

Using this we can compute the coefficients $a(\lambda)$ in the usual way. We have

$$\int_{\mathbb{R}} \psi(x)\bar{f}_{\lambda'}(x)dx = \Big(\int_{\mathbb{R}} a(\lambda)f_\lambda(x)d\lambda\Big)\bar{f}_{\lambda'}(x)dx = \int_{\mathbb{R}} a(\lambda)\delta(\lambda-\lambda')d\lambda = a(\lambda').$$

For example when $f_\lambda = e^{i\lambda x}$ we have, as we saw in Example 5,

$$\int_{\mathbb{R}} e^{i(\lambda-\lambda')x}dx = 2\pi\delta(\lambda-\lambda').$$

Therefore,

$$f_\lambda(x) = \frac{1}{\sqrt{2\pi}}e^{ix\lambda}$$

form an orthonormal basis of eigenvectors, and the coefficients $a(\lambda)$ coincide with $\hat{\psi}(\lambda)$.

6.7 It is easy to extend the notions of distributions and Fourier transforms to functions in several variables. The Fourier transform is defined by

$$\hat{\phi}(x_1,\ldots,x_n) = \frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} \phi(t_1,\ldots,t_n)e^{-i(t_1x_1+\ldots+t_nx_n)}dt_1\ldots dt_n. \tag{6.16}$$

In a coordinate-free way, we should consider ϕ to be a function on a vector space V with some measure μ invariant with respect to translation (a *Haar measure*), and define the Fourier transform $\hat{f}$ as a function on the dual space V^* by

$$\hat{\phi}(\xi) = \frac{1}{(\sqrt{2\pi})^n}\int_V \phi(x)e^{-i\xi(x)}d\mu.$$

Here ϕ is an absolutely integrable function on $\mathbb{R}^n$.

For example, let us see what happens with our operators P_i and Q_i when we apply the Fourier transform to functions $\phi(\mathbf{q})$ to obtain functions $\hat{\phi}(\mathbf{p})$ (note that the variables p_i are dual to the variables q_i). We have

$$\widehat{P_1\phi}(\mathbf{p}) = \frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} P_1\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q} = \frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} \frac{\hbar}{i}\frac{\partial\phi(\mathbf{q})}{\partial q_1}e^{-i(p_1q_1+\ldots+p_nq_n)}d\mathbf{q} =$$

$$= \frac{1}{(\sqrt{2\pi})^n}\left(\frac{\hbar}{i}\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}\Big|_{q_1=-\infty}^{q_1=\infty} + \hbar p_1\int_{\mathbb{R}^n} \phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q}\right) =$$

$$= \frac{1}{(\sqrt{2\pi})^n} \int\limits_{\mathbb{R}^n} P_1 \phi(\mathbf{q}) e^{-i\mathbf{p}\cdot\mathbf{q}} d\mathbf{q} = \hbar p_1 \hat{\phi}(\mathbf{p}).$$

Similarly we obtain

$$P_1 \hat{\phi}(\mathbf{p}) = \frac{\hbar}{i} \frac{\partial}{\partial p_1} \hat{\phi}(\mathbf{p}) = \frac{\hbar}{i} \frac{\partial}{\partial p_1} \frac{1}{(\sqrt{2\pi})^n} \int\limits_{\mathbb{R}^n} \phi(\mathbf{q}) e^{-i\mathbf{p}\cdot\mathbf{q}} d\mathbf{q} =$$

$$= \frac{\hbar}{i} \frac{1}{(\sqrt{2\pi})^n} \int\limits_{\mathbb{R}^n} \phi(\mathbf{q})(-iq_1) e^{-i\mathbf{p}\cdot\mathbf{q}} d\mathbf{q} = -\hbar \widehat{p_1 \phi} = -\hbar \widehat{Q_1 \phi}(\mathbf{p}).$$

This shows that under the Fourier transformation the role of the operators P_i and Q_i is interchanged.

We define distributions as continuous linear functionals on the space K_n of smooth functions with compact support on $\mathbb{R}^n$. Let $\mathcal{D}(\mathbb{R}^n)$ be the vector space of distributions. We use the formal expressions for such functions:

$$\ell(\phi) = \int\limits_{\mathbb{R}^n} \bar{f}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}.$$

Example 10. Let $\phi(\mathbf{x}, \mathbf{y}) \in K_{2n}$. Consider the linear function

$$\ell(\phi) = \int\limits_{\mathbb{R}^n} \phi(\mathbf{x}, \mathbf{x}) d\mathbf{x}.$$

This functional defines a real distribution which we denote by $\delta(\mathbf{x} - \mathbf{y})$. By definition

$$\int\limits_{\mathbb{R}^n} \delta(\mathbf{x} - \mathbf{y}) \phi(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} = \int\limits_{\mathbb{R}^n} \phi(\mathbf{x}, \mathbf{x}) d\mathbf{x}.$$

Note that we can also think of $\delta(\mathbf{x} - \mathbf{y})$ as a distribution in the variable $\mathbf{x}$ with parameter $\mathbf{y}$. Then we understand the previous integral as an iterated integral

$$\int\limits_{\mathbb{R}^n} \left(\int\limits_{\mathbb{R}^n} \phi(\mathbf{x}, \mathbf{y}) \delta(\mathbf{x} - \mathbf{y}) d\mathbf{x} \right) d\mathbf{y}.$$

Here $\delta(\mathbf{x} - \mathbf{y})$ denotes the functional

$$\phi(\mathbf{x}) \to \phi(\mathbf{y}).$$

The following Lemma is known as the *Kernel Theorem*:

Lemma. *Let $B : K_n \times K_n \to \mathbb{C}$ be a bicontinuous bilinear function. There exists a distribution $\ell \in \mathcal{D}(\mathbb{R}^{2n})$ such that*

$$B(\phi(\mathbf{x}),\psi(\mathbf{x})) = \ell(\phi(\mathbf{x})\psi(\mathbf{y})) = \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \bar{K}(\mathbf{x},\mathbf{y})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

This lemma shows that any continuous linear map $T : K_n \to \mathcal{D}(\mathbb{R}^n)$ from the space K_n to the space of distributions on $\mathbb{R}^n$ can be represented as an "integral operator" whose kernel is a distribution. In fact, we have that $(\phi,\psi) \to (T\phi,\bar{\psi})$ is a bilinear function satisfying the assumption of the lemma. So

$$(T\phi,\bar{\psi}) = \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \bar{K}(\mathbf{x},\mathbf{y})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

This means that $T\phi$ is a generalized function whose value on ψ is given by the right-hand side. This allows us to write

$$T\phi(\mathbf{y}) = \int\limits_{\mathbb{R}^n} \bar{K}(\mathbf{x},\mathbf{y})\phi(\mathbf{x})d\mathbf{x}.$$

Examples 11. Consider the operator $T\phi = g\phi$ where $g \in K_n$ is any locally integrable function on $\mathbb{R}^n$. Then

$$(T\phi,\bar{\psi}) = \int\limits_{\mathbb{R}^n} \bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{x})d\mathbf{x}.$$

Consider the function $\bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{x})$ as the restriction of the function $\bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{y})$ to the diagonal $\mathbf{x} = \mathbf{y}$. Then, by Example 10, we find

$$\int\limits_{\mathbb{R}^n} \bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{x})d\mathbf{x} = \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} \delta(\mathbf{x}-\mathbf{y})\bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

This shows that

$$K(\mathbf{x},\mathbf{y}) = \bar{g}(\mathbf{x})\delta(\mathbf{x}-\mathbf{y}).$$

Thus the operator $\phi \to g\phi$ is an integral operator with kernel $\delta(\mathbf{x}-\mathbf{y})\bar{g}(\mathbf{x})$. By continuity we can extend this operator to the whole space $L^2(\mathbb{R}^n)$.

6.8 By Example 10, we can view the operator $Q_i : \phi(\mathbf{q}) \to q_i\phi(\mathbf{q})$ as the integral operator

$$Q_i\phi(\mathbf{q}) = \int\limits_{\mathbb{R}^n} q_i\delta(\mathbf{q}-\mathbf{y})\phi(\mathbf{y})d\mathbf{y}.$$

More generally, for any locally integrable function f on $\mathbb{R}^n$, we can define the operator $f(Q_1,\ldots,Q_n)$ as the integral operator

$$f(Q_1,\ldots,Q_n)\phi(\mathbf{q}) = \int\limits_{\mathbb{R}^n} f(\mathbf{q})\delta(\mathbf{q}-\mathbf{y})\phi(\mathbf{y})d\mathbf{y}.$$

Using the Fourier transform of f, we can rewrite the kernel as follows

$$\int\limits_{\mathbb{R}^n} f(\mathbf{q})\delta(\mathbf{q}-\mathbf{y})\phi(\mathbf{y})d\mathbf{y} = \frac{1}{(\sqrt{2\pi})^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{v})e^{i\mathbf{v}\cdot\mathbf{q}}\delta(\mathbf{q}-\mathbf{y})\phi(\mathbf{y})d\mathbf{y}d\mathbf{v}.$$

Let us introduce the operators

$$V(\mathbf{v}) = V(v_1,\ldots,v_n) = e^{i(v_1Q_1+\ldots+v_nQ_n)}.$$

Since all Q_i are self-adjoint, the operators $V(\mathbf{v})$ are unitary. Let us find its integral expression. By 6.7, we have

$$Q(\mathbf{v})\phi(\mathbf{q}) = e^{i(\mathbf{v}\cdot\mathbf{q})}\phi = \int\limits_{\mathbb{R}^n} e^{i(\mathbf{v}\cdot\mathbf{q})}\delta(\mathbf{q}-\mathbf{y})\phi(\mathbf{y})d\mathbf{y}.$$

Comparing the kernels of $f(Q_1,\ldots,Q_n)$ and of $Q(\mathbf{v})$ we see that

$$f(Q_1,\ldots,Q_n) = \frac{1}{(\sqrt{2\pi})^n}\int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{v})Q(\mathbf{v})d\mathbf{v}.$$

Now it is clear how to define A_f for any $f(\mathbf{p},\mathbf{q})$. We introduce the operators

$$U(\mathbf{u}) = U(u_1,\ldots,u_n) = e^{i(u_1P_1+\ldots+u_nP_n)}.$$

Let $\phi(\mathbf{u},\mathbf{q}) = U(\mathbf{u})\phi(\mathbf{q})$. Differentiating with respect to the parameter $\mathbf{u}$, we have

$$\frac{\partial\phi(\mathbf{u},\mathbf{q})}{\partial u_i} = iP_i\phi(\mathbf{u},\mathbf{q}) = \hbar\frac{\partial\phi(\mathbf{u},\mathbf{q})}{\partial q_i}.$$

This has a unique solution with initial condition $\phi(0,\mathbf{q}) = \phi(\mathbf{q})$, namely

$$\phi(\mathbf{u},\mathbf{q}) = \phi(\mathbf{q}-\hbar\mathbf{u}).$$

Thus we see that

$$V(\mathbf{v})U(\mathbf{u})\phi(\mathbf{q}) = e^{i\mathbf{v}\cdot\mathbf{q}}\phi(\mathbf{q}-\hbar\mathbf{u}),$$
$$U(\mathbf{u})V(\mathbf{v})\phi(\mathbf{q}) = e^{i\mathbf{v}\cdot(\mathbf{q}-\hbar\mathbf{u})}\phi(\mathbf{q}-\hbar\mathbf{u}).$$

This implies

$$[U(\mathbf{u}),V(\mathbf{v})] = e^{-i\hbar\mathbf{v}\cdot\mathbf{u}}. \tag{6.18}$$

Now it is clear how to define A_f. We set for any $f \in K'_{2n}$ (viewed as a regular distribution),

$$A_f = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{u}, \mathbf{v}) V(\mathbf{v}) U(\mathbf{u}) e^{-\frac{i\hbar \mathbf{u}\cdot\mathbf{v}}{2}} d\mathbf{v} d\mathbf{u}. \tag{6.19}$$

The scalar exponential factor here helps to verify that the operator A_f is self-adjoint. We have

$$A_f^* = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \overline{\hat{f}(\mathbf{u}, \mathbf{v})} U(\mathbf{u})^* V(\mathbf{v})^* e^{\frac{i\hbar \mathbf{u}\cdot\mathbf{v}}{2}} d\mathbf{v} d\mathbf{u} =$$

$$= \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \hat{f}(-\mathbf{v}, -\mathbf{u}) U(-\mathbf{u}) V(-\mathbf{v}) e^{\frac{i\hbar \mathbf{u}\cdot\mathbf{v}}{2}} d\mathbf{v} d\mathbf{u} =$$

$$= \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{u}, \mathbf{v}) U(\mathbf{u}) V(\mathbf{v}) e^{\frac{i\hbar \mathbf{u}\cdot\mathbf{v}}{2}} d\mathbf{v} d\mathbf{u} =$$

$$= \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{u}, \mathbf{v}) V(\mathbf{v}) U(\mathbf{u}) e^{-\frac{i\hbar \mathbf{u}\cdot\mathbf{v}}{2}} d\mathbf{v} d\mathbf{u} = A_f.$$

Here, at the very end, we have used the commutator relation (6.18).

6.9 Let us compute the Poisson bracket $\{A_f, A_g\}_\hbar$ and compare it with $A_{\{f,g\}}$. Let

$$K^f_{\mathbf{u},\mathbf{v}}(\mathbf{q}, \mathbf{y}) = \hat{f}(\mathbf{u}, \mathbf{v}) e^{i\mathbf{v}\cdot\mathbf{q}} \delta(\mathbf{q} - \hbar\mathbf{u} - \mathbf{y}) e^{-i\hbar\mathbf{u}\cdot\mathbf{v}/2},$$

$$K^g_{\mathbf{u}',\mathbf{v}'}(\mathbf{q}, \mathbf{y}) = \hat{g}(\mathbf{u}', \mathbf{v}') e^{i\mathbf{v}'\cdot\mathbf{q}} \delta(\mathbf{q} - \hbar\mathbf{u}' - \mathbf{y}) e^{-i\hbar\mathbf{u}'\cdot\mathbf{v}'/2},$$

Then, the kernel of $A_f A_g$ is

$$K_1 = \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} K^f_{\mathbf{u},\mathbf{v}}(\mathbf{q}, \mathbf{q}') K^g_{\mathbf{u}',\mathbf{v}'}(\mathbf{q}', \mathbf{y}) d\mathbf{q}' d\mathbf{u} d\mathbf{v} d\mathbf{u}' d\mathbf{v}' =$$

$$= \int\limits_{\mathbb{R}^{5n}} \hat{f}(\mathbf{u}, \mathbf{v}) e^{i(\mathbf{v}\cdot\mathbf{q} + \mathbf{v}'\cdot\mathbf{q}')} \delta(\mathbf{q} - \hbar\mathbf{u} - \mathbf{q}') \hat{g}(\mathbf{u}', \mathbf{v}') \delta(\mathbf{q}' - \hbar\mathbf{u}' - \mathbf{y}) e^{\frac{-i\hbar(\mathbf{u}'\cdot\mathbf{v}' + \mathbf{u}\cdot\mathbf{v})}{2}} d\mathbf{q}' d\mathbf{u} d\mathbf{v} d\mathbf{u}' d\mathbf{v}'$$

$$= \int\limits_{\mathbb{R}^{5n}} \hat{f}(\mathbf{u}, \mathbf{v}) \hat{g}(\mathbf{u}', \mathbf{v}') e^{i\mathbf{v}\cdot\mathbf{q}} e^{\frac{-i\hbar\mathbf{u}\cdot\mathbf{v}}{2}} e^{i\mathbf{v}'\cdot(\mathbf{q} - \hbar\mathbf{u})} \delta(\mathbf{q} - \hbar(\mathbf{u} + \mathbf{u}') - \mathbf{y}) e^{\frac{-i\hbar\mathbf{u}'\cdot\mathbf{v}'}{2}} d\mathbf{u} d\mathbf{v} d\mathbf{u}' d\mathbf{v}'.$$

Similarly, the kernel of $A_g A_f$ is

$$K_2 = \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} K^g_{\mathbf{u}',\mathbf{v}'}(\mathbf{q}, \mathbf{q}') K^f_{\mathbf{u},\mathbf{v}}(\mathbf{q}', \mathbf{y}) d\mathbf{q}' d\mathbf{u} d\mathbf{v} d\mathbf{u}' d\mathbf{v}' =$$

$$\int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} \hat{f}(\mathbf{u},\mathbf{v})\hat{g}(\mathbf{u}',\mathbf{v}')e^{i\mathbf{v}'\cdot\mathbf{q}}e^{-i\hbar\mathbf{u}'\cdot\mathbf{v}'/2}e^{i\mathbf{v}\cdot(\mathbf{q}-\hbar\mathbf{u}')}\delta(\mathbf{q}-\hbar(\mathbf{u}+\mathbf{u}')-\mathbf{y})e^{-i\hbar\mathbf{u}\cdot\mathbf{v}/2}d\mathbf{u}d\mathbf{v}d\mathbf{u}'d\mathbf{v}'.$$

We see that the integrands differ by the factor $e^{i\hbar(\mathbf{v}'\cdot\mathbf{u}-\mathbf{u}'\cdot\mathbf{v})}$. In particular we obtain

$$\lim_{\hbar\to 0} A_f A_g = \lim_{\hbar\to 0} A_g A_f.$$

Now let check that

$$\lim_{\hbar\to 0}\{A_f, A_g\}_\hbar = \lim_{\hbar\to 0} A_{\{f,g\}} \tag{6.20}$$

This will assure us that we are on the right track.

It follows from above that the operator $\{A_f, A_g\}_\hbar$ has the kernel

$$\frac{i}{\hbar}\int_{\mathbb{R}^{4n}} \hat{f}(u,v)\hat{g}(u',v')e^{i(v'+v)\cdot q-\frac{\hbar(u\cdot v+u'\cdot v')}{2}}\delta(q-\hbar(u+u')-y)(e^{-i\hbar v'\cdot u}-e^{-i\hbar u'\cdot v})dudvdu'dv',$$

where we "unbolded" the variables for typographical reasons. Since

$$\lim_{\hbar\to 0}\frac{e^{-i\hbar\mathbf{v}'\cdot\mathbf{u}}-e^{-i\hbar\mathbf{u}'\cdot\mathbf{v}}}{-i\hbar} = \mathbf{v}'\cdot\mathbf{u}-\mathbf{u}'\cdot\mathbf{v},$$

passing to the limit when $\hbar$ goes to 0, we obtain that the kernel of $\lim_{\hbar\to 0}\{A_f, A_g\}_\hbar$ is equal to

$$\int_{\mathbb{R}^{4n}} \hat{f}(\mathbf{u}',\mathbf{v}')\hat{g}(\mathbf{u}'',\mathbf{v}'')e^{i(\mathbf{v}'+\mathbf{v}'')\cdot\mathbf{q}}\delta(\mathbf{q}-\mathbf{y})(\mathbf{v}''\cdot\mathbf{u}'-\mathbf{u}''\cdot\mathbf{v})d\mathbf{u}'d\mathbf{v}'d\mathbf{u}''d\mathbf{v}''. \tag{6.21}$$

Now let us compute $A_{\{f,g\}}$. We have

$$\{f,g\} = \sum_{i=1}^n\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial p_i}\right).$$

Applying properties $(F3)$ and $(F5)$ of Fourier transforms, we get

$$\widehat{\{f,g\}}(\mathbf{u},\mathbf{v}) = \sum_{i=1}^n (iv_i\hat{f})*(iu_i\hat{g})-(iu_i\hat{f})*(iv_i\hat{g}_i).$$

By formula (6.14), we get

$$\widehat{\{f,g\}}(\mathbf{u},\mathbf{v}) =$$

$$= \int_{\mathbb{R}^{4n}} \hat{f}(\mathbf{u}',\mathbf{v}')\hat{g}(\mathbf{u}'',\mathbf{v}'')(\mathbf{v}''\cdot\mathbf{u}'-\mathbf{u}''\cdot\mathbf{v}')\delta(\mathbf{u}'+\mathbf{u}''-\mathbf{u})\delta(\mathbf{v}'+\mathbf{v}''-\mathbf{v})d\mathbf{u}'d\mathbf{v}'d\mathbf{u}''d\mathbf{v}''.$$

Thus, comparing with (6.21), we find (identifying the operators with their kernels):

$$\lim_{\hbar\to 0} A_{\{f,g\}} = \lim_{\hbar\to 0}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \widehat{\{f,g\}}(\mathbf{u},\mathbf{v})e^{i\mathbf{v}\cdot\mathbf{q}}\delta(\mathbf{q}-\hbar\mathbf{u}-\mathbf{y})e^{-\hbar\mathbf{u}\cdot\mathbf{v}/2}d\mathbf{u}d\mathbf{v} =$$

$$= \lim_{\hbar\to 0}\int\limits_{\mathbb{R}^{4n}} \hat{f}(\mathbf{u}',\mathbf{v}')\hat{g}(\mathbf{u}'',\mathbf{v}'')e^{i(\mathbf{v}'+\mathbf{v}'')\cdot\mathbf{q}}\delta(\mathbf{q}-\hbar(\mathbf{u}'+\mathbf{u}'')-\mathbf{y})\delta(\mathbf{q}-\mathbf{y})(\mathbf{v}''\cdot\mathbf{u}'-\mathbf{u}''\cdot\mathbf{v})\times$$

$$\times e^{-\hbar\mathbf{u}\cdot\mathbf{v}/2}d\mathbf{u}'d\mathbf{v}'d\mathbf{u}''d\mathbf{v}'' = \lim_{\hbar\to 0}\{A_f, A_g\}_\hbar.$$

6.10 We can also apply the quantization formula (14) to density functions $\rho(\mathbf{q},\mathbf{p})$ to obtain the density operators M_ρ. Let us check the formula

$$\lim_{\hbar\to 0} Tr(M_\rho A_f) = \int\limits_M f(\mathbf{p},\mathbf{q})d\mathbf{p}d\mathbf{q}. \tag{6.22}$$

Let us compute the trace of the operator A_f using the formula

$$Tr(T) = \sum_{n=1}^{\infty}(Te_n, e_n)$$

By writing each e_n and Te_n in the form

$$e_n(\mathbf{q}) = \frac{1}{(\sqrt{2\pi})^n}\int\limits_{\mathbb{R}} \hat{e}_n(\mathbf{t})e^{i\mathbf{q}\cdot\mathbf{t}}d\mathbf{t},\quad Te_n(\mathbf{q}) = \frac{1}{(\sqrt{2\pi})^n}\int\limits_{\mathbb{R}} \hat{e}_n(\mathbf{t})Te^{i\mathbf{q}\cdot\mathbf{t}}d\mathbf{t},$$

and using the Plancherel formula, we easily get

$$Tr(T) = \frac{1}{(2\pi)^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}(Te^{i\mathbf{q}\cdot\mathbf{t}}, e^{i\mathbf{q}\cdot\mathbf{t}})d\mathbf{q}d\mathbf{t} \tag{6.23}$$

Of course the integral (6.23) in general does not converge. Let us compute the trace of the operator

$$V(\mathbf{v})U(\mathbf{u}) : \phi \to e^{i\mathbf{v}\cdot\mathbf{q}}\phi(\mathbf{q}-\hbar\mathbf{u}).$$

We have

$$Tr(V(\mathbf{v})U(\mathbf{u})) = \frac{1}{(2\pi)^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} e^{i\mathbf{v}\cdot\mathbf{q}}e^{i(\mathbf{q}-\hbar\mathbf{u})\cdot\mathbf{t}}e^{-i\mathbf{q}\cdot\mathbf{v}}d\mathbf{q}d\mathbf{t} =$$

$$= \frac{1}{(\sqrt{2\pi})^n}\Big(\int\limits_{\mathbb{R}^n} e^{i\mathbf{v}\cdot\mathbf{q}}d\mathbf{v}\Big)\frac{1}{(\sqrt{2\pi})^n}\Big(\int\limits_{\mathbb{R}^n} e^{-i\hbar\mathbf{t}\cdot\mathbf{u}}d\mathbf{t}\Big) = (\frac{2\pi}{\hbar})^n\delta(\mathbf{v})\delta(\mathbf{u}).$$

Thus

$$Tr(A_f) = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{u},\mathbf{v})e^{-i\hbar\mathbf{u}\cdot\mathbf{v}/2}Tr(V(\mathbf{v})U(\mathbf{u}))d\mathbf{u}d\mathbf{v} =$$

$$= \frac{1}{\hbar^n} \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \hat{f}(\mathbf{u},\mathbf{v})e^{-i\hbar\mathbf{u}\cdot\mathbf{v}/2}\delta(\mathbf{v})\delta(\mathbf{u})d\mathbf{u}d\mathbf{v} =$$

$$= \frac{1}{\hbar^n}\hat{f}(0,0) = \frac{1}{(2\pi\hbar)^n} \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} f(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v}.$$

This formula suggests that we change the definition of the density operator $M_\rho = A_\rho$ by replacing it with

$$M_\rho = (2\pi\hbar)^n A_\rho.$$

Then

$$Tr(M_\rho) = \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \rho(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v} = 1,$$

as it should be. Now, appplying (6.21), we obtain

$$\lim_{\hbar\to 0} Tr(M_\rho A_f) = (2\pi\hbar)^n \lim_{\hbar\to 0} Tr(A_\rho f) = (2\pi\hbar)^n \frac{1}{(2\pi\hbar)^n} \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \rho(\mathbf{u},\mathbf{v})f(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v} =$$

$$= \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} \rho(\mathbf{u},\mathbf{v})f(\mathbf{u},\mathbf{v})d\mathbf{u}d\mathbf{v}.$$

This proves formula (6.22).

Exercises.

1. Define the direct product of distributions $f, g \in \mathcal{D}(\mathbb{R}^n)$ as the distribution $f \times g$ from $\mathcal{D}(\mathbb{R}^{2n})$ with

$$\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} (f\times g)\phi(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} = \int\limits_{\mathbb{R}^n} f(\mathbf{x})\Big(\int\limits_{\mathbb{R}^n} g(\mathbf{y})\phi(\mathbf{x},\mathbf{y})d\mathbf{y}\Big)d\mathbf{x}.$$

(i) Show that $\delta(\mathbf{x}) \times \delta(\mathbf{y}) = \delta(\mathbf{x},\mathbf{y})$.
(ii) Find the relationship of this operation with the operation of convolution.
(ii) Prove that the support of $f \times g$ is equal to the product of the supports of f and g.

2. Let $\phi(x)$ be a function absolutely integrable over $\mathbb{R}$. Show that $\phi \to \int\limits_{\mathbb{R}} f(x)\phi(x,y)dxdy$ defines a generalized function ℓ_f in two variables x, y. Prove that the Fourier transform of ℓ_f is equal to the generalized function $\sqrt{2\pi}\hat{f}(u) \times \delta(v)$.

3. Define partial derivatives of distributions and prove that $D(f * g) = Df * g = f * Dg$ where D is any differential operator.

4 Find the Fourier transform of a regular distribution defined by a polynomial function.

5. Let $f(x)$ be a function in one variable with discontinuity of the first kind at a point $x = a$. Assume that $f(x)'$ is continuous at $x \neq a$ and has discontinuity of the first kind at $x = a$. Show that the derivative of the distribution $f(x)$ is equal to $f'(x)+(f(a+)-f(a-))\delta(x-a)$.

6. Prove the Poisson summation formula $\sum_{n=-\infty}^{+\infty} \hat{\phi}(n) = 2\pi \sum_{n=-\infty}^{+\infty} \phi(n)$.

7. Find the generalized kernel of the operator $\phi(x) \to \phi'(x)$.

8. Verify that $A_{p_i} = P_i$.

9. Verify that the density operators M_ρ are non-negative operators as they should be.

10. Show that the pure density operator P_ψ corresponding to a normalized wave function $\psi(\mathbf{q})$ is equal to M_ρ where $\rho(\mathbf{p}, \mathbf{q}) = |\psi(\mathbf{q})|^2\delta(\mathbf{p})$.

11. Show that the Uncertainty Principle implies $\Delta_\mu P \Delta_\mu Q \geq \hbar/2$.

Lecture 7. HARMONIC OSCILLATOR

7.1 A (one-dimensional) *harmonic oscillator* is a classical mechanical system with Hamiltonian function

$$H(p,q) = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2. \tag{7.1}$$

The parameters m and ω are called the mass and the frequency. The corresponding Newton equation is

$$m\frac{d^2x}{dt^2} = -m\omega x.$$

The Hamilton equations are

$$\dot{p} = -\frac{\partial H}{\partial q} = -qm\omega^2, \quad \dot{q} = \frac{\partial H}{\partial p} = p/m.$$

After differentiating, we get

$$\ddot{p} + \omega^2 p = 0, \quad \ddot{q} + \omega^2 q = 0.$$

This can be easily integrated. We find

$$p = cm\omega\cos(\omega t + \phi), \quad q = c\sin(\omega t + \phi)$$

for some constants c, ϕ determined by the initial conditions.

The Schrödinger operator

$$H = \frac{1}{2m}P^2 + \frac{m\omega^2}{2}Q^2 = -\frac{\hbar^2}{2m}\frac{d^2}{dq^2} + \frac{m\omega^2}{2}q^2.$$

Let us find its eigenvectors. We shall assume for simplicity that $m = 1$. We now consider the so-called *annihilation operator* and the *creation operator*

$$a = \frac{1}{\sqrt{2\omega}}(\omega Q - iP), \quad a^* = \frac{1}{\sqrt{2\omega}}(\omega Q + iP). \tag{7.2}$$

They are obviously adjoint to each other. We shall see later the reason for these names. Using the commutator relation $[Q,P] = -i\hbar$, we obtain

$$\omega aa^* = \frac{1}{2}(\omega^2 Q^2 + P^2) + \frac{i\omega}{2}(-QP + PQ) = H + \frac{\hbar\omega}{2},$$

$$\omega a^* a = \frac{1}{2}(\omega^2 Q^2 + P^2) + \frac{i\omega}{2}(QP - PQ) = H - \frac{\hbar\omega}{2}.$$

From this we deduce that

$$[a, a^*] = \hbar, \quad [H, a] = -\hbar\omega a, \quad [H, a^*] = \hbar\omega a^*. \tag{7.3}$$

This shows that the operators $1, H, a, a^*$ form a Lie algebra $\mathcal{H}$, called the *extended Heisenberg algebra*. The *Heisenberg algebra* is the subalgebra $N \subset \mathcal{H}$ generated by $1, a, a^*$. It is isomorphic to the Lie algebra of matrices

$$\begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}.$$

N is also an ideal in $\mathcal{H}$ with one-dimensional quotient. The Lie algebra $\mathcal{H}$ is isomorphic to the Lie algebra of matrices of the form

$$A(x,y,z,w) = \begin{pmatrix} 0 & x & y & z \\ 0 & w & 0 & y \\ 0 & 0 & -w & -x \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad x, y, z, w \in \mathbb{R}.$$

The isomorphism is defined by the map

$$1 \to A(0,0,1,0), \quad a \to A(1,0,0,0), \quad a^* \to \frac{\hbar}{2}A(0,1,0,0), \quad H \to \frac{\omega\hbar}{2}A(0,0,0,1).$$

So we are interested in the representation of the Lie algebra $\mathcal{H}$ in $L^2(\mathbb{R})$.

Suppose we have an eigenvector ψ of H with eigenvalue λ. Since a^* is obviously adjoint to a, we have

$$\hbar\lambda||\psi||^2 = (\psi, H\psi) = (\psi, \omega a^* a\psi) + (\psi, \frac{\hbar\omega}{2}\psi) = \omega||a\psi||^2 + \frac{\hbar\omega}{2}||\psi||^2.$$

This implies that all eigenvalues λ are real and satisfy the inequality

$$\lambda \geq \frac{\hbar\omega}{2}. \tag{7.4}$$

The equality holds if and only if $a\psi = 0$. Clearly any vector annihilated by a is an eigenvector of H with minimal possible absolute value of its eigenvalue. A vector of norm one with such a property is called a *vacuum vector*.

Denote a vacuum vector by $|0\rangle$. Because of the relation $[H,a] = -\hbar\omega a$, we have

$$Ha\psi = aH\psi - \hbar\omega a\psi = (\lambda - \hbar\omega)a\psi. \tag{7.5}$$

This shows that $a\psi$ is a new eigenvector with eigenvalue $\lambda - \hbar\omega$. Since eigenvalues are bounded from below, we get that $a^{n+1}\psi = 0$ for some $n \geq 0$. Thus $a(a^n\psi) = 0$ and $a^n\psi$ is a vacuum vector. Thus we see that the existence of one eigenvalue of H is equivalent to the existence of a vacuum vector.

Now if we start applying $(a^*)^n$ to the vacuum vector $|0\rangle$, we get eigenvectors with eigenvalue $\hbar\omega/2 + n\hbar\omega$. So we are getting a countable set of eigenvectors

$$\psi_n = a^{*n}|0\rangle$$

with eigenvalues $\lambda_n = \frac{2n+1}{2}\hbar\omega$. We shall use the following commutation relation

$$[a,(a^*)^n] = n\hbar(a^*)^{n-1}$$

which follows easily from (7.3). We have

$$||\psi_1||^2 = ||a^*|0\rangle||^2 = (|0\rangle, aa^*|0>) = (|0\rangle, a^*a|0\rangle + \hbar|0>) = (|0\rangle, \hbar|0\rangle) = \hbar,$$

$$||\psi_n|| = ||a^{*n}|0\rangle||^2 = (a^{*n}|0\rangle, a^{*n}|0>) = (|0\rangle, a^{n-1}(a(a^*)^n)|0\rangle) =$$
$$= (|0\rangle, a^{n-1}(a^{*n}a)|0\rangle) + n\hbar(|0\rangle, a^{n-1}(a^*)^{n-1}|0\rangle) = n\hbar||(a^*)^{n-1}|0\rangle||^2.$$

By induction, this gives

$$||\psi_n||^2 = n!\hbar^n.$$

After renormalization we obtain a countable set of orthonormal eigenvectors

$$|n\rangle = \frac{1}{\sqrt{\hbar^n n!}}(a^*)^n|0\rangle, \quad n = 0,1,2,\ldots. \tag{7.6}$$

The orthogonality follows from self-adjointness of the operator H. It is easy to see that the subspace of $L^2(\mathbb{R})$ spanned by the vectors (7.6) is an ireducible representation of the Lie algebra $\mathcal{H}$. For any vacuum vector $|0\rangle$ we find an irreducible component of $L^2(\mathbb{R})$.

7.2 It remains to prove the existence of a vacuum vector, and to prove that $L^2(\mathbb{R})$ can be decomposed into a direct sum of irreducible components corresponding to vacuum vectors. So we have to solve the equation

$$\sqrt{2\omega}a\psi = (\omega Q - iP)\psi = \omega q\psi + \hbar\frac{d\psi}{dq} = 0.$$

It is easy to do by separating the variables. We get a unique solution

$$\psi = |0\rangle = (\frac{\omega}{\hbar\pi})^{\frac{1}{4}}e^{-\frac{\omega q^2}{2\hbar}},$$

where the constant $(\frac{\omega}{\hbar\pi})^{\frac{1}{4}}$ was computed by using the condition $||\psi|| = 1$. We have

$$|n\rangle = \frac{1}{\sqrt{\hbar^n n!}}(a^*)^n|0\rangle = (\frac{\omega}{\hbar\pi})^{\frac{1}{4}}\frac{(2\omega)^{\frac{-n}{2}}}{\sqrt{\hbar^n n!}}(\omega q - \hbar\frac{d}{dq})^n e^{-\frac{\omega q^2}{2\hbar}} =$$

$$= (\frac{\omega}{\hbar\pi})^{\frac{1}{4}}\frac{1}{\sqrt{2^n n!}}(q\sqrt{\frac{\omega}{\hbar}} - \sqrt{\frac{\hbar}{\omega}}\frac{d}{dq})^n e^{-\frac{\omega q^2}{2\hbar}} =$$

$$= (\frac{\omega}{\hbar\pi})^{\frac{1}{4}}\frac{1}{\sqrt{2^n n!}}(x - \frac{d}{dx})^n e^{\frac{-x^2}{2}} = (\frac{\omega}{\hbar})^{\frac{1}{4}} H_n(x)e^{\frac{-x^2}{2}}. \tag{7.7}$$

Here $x = q\frac{\sqrt{\omega}}{\sqrt{\hbar}}$, $H_n(x)$ is a (normalized) *Hermite polynomial* of degree n. We have

$$\int_{\mathbb{R}} |n\rangle|m\rangle dq = \int_{\mathbb{R}} (\frac{\omega}{\hbar\pi})^{\frac{1}{2}} e^{-x^2} H_n(x)H_m(x)(\frac{\omega}{\hbar\pi})^{-\frac{1}{2}} dx = \int_{\mathbb{R}} e^{-x^2} H_n(x)H_m(x)dx = \delta_{mn}. \tag{7.8}$$

This agrees with the orthonormality of the functions $|n\rangle$ and the known orthonormality of the *Hermite functions* $H_n(x)e^{-\frac{x^2}{2}}$. It is known also that the orthonormal system of functions $H_n(x)e^{\frac{-x^2}{2}}$ is complete, i.e., forms an orthonormal basis in the Hilbert space $L^2(\mathbb{R})$. Thus we constructed an irreducible representation of $\mathcal{H}$ with unique vacuum vector $|0\rangle$. The vectors (7.7) are all orthonormal eigenvectors of H with eigenvalues $(n + \frac{1}{2})\hbar\omega$.

7.3 Let us compare the classical and quantum pictures. The pure states $P_{|n\rangle}$ corresponding to the vectors $|n\rangle$ are stationary states. They do not change with time. Let us take the observable Q (the quantum analog of the coordinate function q). Its spectral function is the operator function

$$P_Q(\lambda) : \psi \to \theta(\lambda - q)\psi,$$

where $\theta(x)$ is the Heaviside function (see Lecture 5, Example 6). Thus the probability that the observable Q takes value $\leq \lambda$ in the state $|n\rangle$ is equal to

$$\mu_Q(\lambda) = Tr(P_Q(\lambda)P_{|n\rangle}) = (\theta(\lambda - q)|n\rangle, |n\rangle) = \int_{\mathbb{R}} \theta(\lambda - q)\big||n\rangle\big|^2 dq =$$

$$= \int_{-\infty}^{\lambda} \big||n\rangle\big|^2 dq = \int_{\mathbb{R}} H_n(x)^2 e^{-x^2} dx. \tag{7.9}$$

The density distribution function is $\big||n\rangle\big|^2$. On the other hand, in the classical version, for the observable q and the pure state $q_t = A\sin(\omega t + \phi)$, the density distribution function is $\delta(q - A\sin(\omega t + \phi))$ and

$$\omega_q(\lambda) = \int_{\mathbb{R}} \theta(\lambda - q)\delta(q - A\sin(\omega t + \phi))dq = \theta(\lambda - A\sin(\omega t + \phi)),$$

as expected. Since $|q| \leq A$ we see that the classical observable takes values only between $-A$ and A. The quantum particle, however, can be found with a non-zero probability in any point, although the probability goes to zero if the particle goes far away from the equilibrium position 0.

The mathematical expectation of the observable Q is equal to

$$Tr(QP_{|n\rangle}) = \int_{\mathbb{R}} Q|n\rangle|n\rangle dq = \int_{\mathbb{R}} q|n\rangle|n\rangle dq = \int_{\mathbb{R}} xH_n(x)^2(x)e^{-x^2}dx = 0. \tag{7.10}$$

Here we used that the function $H_n(x)^2$ is even. Thus we expect that, as in the classical case, the electron oscillates around the origin.

Let us compare the values of the observable H which expresses the total energy. The spectral function of the Schrödinger operator H is

$$P_H(\lambda) = \sum_{\hbar\omega(n+\frac{1}{2})\leq\lambda} P_{|n\rangle}.$$

Therefore,

$$\omega_H(\lambda) = Tr(P_H(\lambda)P_{|n\rangle}) = \begin{cases} 0 & \text{if } \lambda < \hbar\omega(2n+1)/2, \\ 1 & \text{if } \lambda \geq \hbar\omega(2n+1)/2. \end{cases}$$

$$\langle H|\omega_{P_{|n\rangle}}\rangle = Tr(HP_{|n\rangle}) = (H|n\rangle, |n\rangle) = \hbar\omega(2n+1)/2. \tag{7.11}$$

This shows that H takes the value $\hbar\omega(2n+1)/2$ at the pure state $P_{|n\rangle}$ with probability 1 and the mean value of H at $|n\rangle$ is its eigenvalue $\hbar\omega(2n+1)/2$. So $P_{|n\rangle}$ is a pure stationary state with total energy $\hbar\omega(2n+1)/2$. The energy comes in quanta. In the classical picture, the total energy at the pure state $q = A\sin(\omega t + \phi)$ is equal to

$$E = \frac{1}{2}m\dot{q}^2 + \frac{m\omega^2}{2}q^2 = \frac{1}{2}A^2\omega^2.$$

There are pure states with arbitrary value of energy.

Using (7.8) and the formula

$$H_n(x)' = \sqrt{2n}H_{n-1}(x), \tag{7.12}$$

we obtain

$$\int_{\mathbb{R}} x^2H_n(x)^2e^{-x^2}dx = \frac{1}{2}\int_{\mathbb{R}} (xH_n(x)^2)'e^{-x^2}dx =$$

$$= \frac{1}{2}\int_{\mathbb{R}} H_n(x)^2e^{-x^2}dx + \int_{\mathbb{R}} e^{-x^2}xH_n(x)H_n(x)'dx =$$

$$= \frac{1}{2} + \frac{1}{2}\int_{\mathbb{R}} e^{-x^2}(H_n(x)')^2dx + \frac{1}{2}\int_{\mathbb{R}} e^{-x^2}(H_n(x)H_n(x)''dx =$$

$$= \frac{1}{2} + \frac{1}{2}2n \int_{\mathbb{R}} e^{-x^2} H_{n-1}(x)^2 dx + n(n-1) \int_{\mathbb{R}} e^{-x^2} H_n(x) H_{n-2}(x) dx = n + \frac{1}{2}. \tag{7.13}$$

Consider the classical observable $h(p,q) = \frac{1}{2}(p^2 + \omega^2 q^2)$(total energy) and the classical density

$$\rho(p,q) = |n\rangle^2(q)\delta(p - \omega q). \tag{7.14}$$

Then, using (7.13), we obtain

$$\int_{\mathbb{R}} \int_{\mathbb{R}} h(p,q)\rho(p,q)dpdq = \frac{1}{2} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} (p^2 + \omega^2 q^2)\delta(p - \omega q)dp \Big) |n\rangle^2(q)dq =$$

$$= \omega^2 \int_{\mathbb{R}} (q^2 |n\rangle^2(q)dq = \omega^2 \int_{\mathbb{R}} \frac{\hbar}{\omega} x^2 H_n(x)^2 e^{-x^2} dx = \hbar\omega(n + \frac{1}{2}).$$

Comparing the result with (7.13), we find that the classical analog of the quantum pure state $|n\rangle$ is the state defined by the density function (7.14). Obviously this state is not pure. The mean value of the observable q at this state is

$$\langle q|\mu_\rho\rangle = \int_{\mathbb{R}} \int_{\mathbb{R}} q \big|n\rangle(q)\big|^2 \delta(p - q\omega)dpdq = \int_{\mathbb{R}} q\big|n\rangle\big|^2 dq = \int_{\mathbb{R}} xH_n(x)^2 e^{-x^2} dx = 0.$$

This agrees with (7.10).

To explain the relationship between the pure classical states $A\sin(\omega t + \phi)$ and the mixed classical states $\rho_n(p,q) = |n\rangle^2(q)\delta(p - \omega q)$ we consider the function

$$F(q) = \frac{1}{2\pi} \int_0^{2\pi} \delta(A\sin(\omega t + \phi) - x)d\phi.$$

It should be thought as a convex combination of the pure states $A\sin(\omega t + \phi)$ with random phases ϕ. We compute

$$F(q) = \frac{1}{\pi} \int_{\frac{\pi}{2}}^{\frac{\pi}{2}} \delta(A\sin(\phi - q)d\phi =$$

$$= \frac{1}{\pi} \int_{-A}^{A} \delta(y - q) \frac{1}{\sqrt{A^2 - y^2}} dy = \frac{1}{\pi} \int_{\mathbb{R}} \delta(y - q) \frac{\theta(A^2 - q^2)}{\sqrt{A^2 - y^2}} dy = \frac{\theta(A^2 - q^2)}{\pi\sqrt{A^2 - q^2}}.$$

One can now show that

$$\lim_{n \to \infty} \big||n\rangle(q)\big|^2 = F(q).$$

Since we keep the total energy $E = A^2\omega^2/2$ of classical pure states $A\sin(\omega t + \phi)$ constant, we should keep the value $\hbar\omega(n + \frac{1}{2})$ of H at $|n\rangle$ constant. This is possible only if $\hbar$ goes to 0. This explains why the stationary quantum pure states and the classical stationary mixed states $F(q)\delta(p-\omega q)$ correspond to each other when $\hbar$ goes to 0.

7.4 Since we know the eigenvectors of the Schrödinger operator, we know everything. In Heisenberg's picture any observable A evolves according to the law:

$$A(t) = e^{iHt/\hbar}A(0)e^{-iHt/\hbar} = \sum_{m,n=0}^{\infty} e^{-it\omega(m-n)}P_{|m\rangle}A(0)e^{-it\omega(m-n)}P_{|n\rangle}.$$

Any pure state can be decomposed as a linear combination of the eigenevectors $|n\rangle$:

$$\psi(q) = \sum_{n=0}^{\infty} c_n|n\rangle, \quad c_n = \int_{\mathbb{R}} \psi|n\rangle dx.$$

It evolves by the formula (see (5.15)):

$$\psi(q;t) = e^{\frac{-iHt}{\hbar}}\psi(q) = \sum_{n=0}^{\infty} e^{-it\omega(n+\frac{1}{2})}c_n|n\rangle.$$

The mean value of the observable Q in the pure state P_ψ is equal to

$$Tr(QP_\psi) = \sum_{n=0}^{\infty} c_n Tr(Q|n\rangle = 0.$$

The mean value of the total energy is equal to

$$Tr(HP_\psi) = \sum_{n=0}^{\infty} c_n Tr(H|n\rangle = \sum_{n=0}^{\infty} c_n\hbar\omega(n+\frac{1}{2}) = (H\psi,\psi).$$

7.5 The Hamiltonian function of the harmonic oscillator is of course a very special case of the following Hamiltonian function on $T(\mathbb{R}^N)^*$:

$$H = A(p_1,\ldots,p_n) + B(q_1,\ldots,q_n),$$

where $A(p_1,\ldots,p_n)$ is a positive definite quadratic form, and B is any quadratic form. By simultaneously reducing the pair of quadratic forms (A,B) to diagonal form, we may assume

$$A(p_1,\ldots,p_n) = \frac{1}{2}\sum_{i=1}^{n}\frac{p_i^2}{m_i},$$

$$B(q_1,\ldots,q_n) = \frac{1}{2}\sum_{i=1}^{n}\varepsilon_i m_i\omega_i^2 q_i^2,$$

where $1/m_1, \dots, 1/m_n$ are the eigenvalues of A, and $\varepsilon_i \in \{-1, 0, 1\}$. Thus the problem is divided into three parts. After reindexing the coordinates we may assume that

$$\varepsilon_i = 1, i = 1, \dots, N_1, \quad \varepsilon_i = -1, i = N_1 + 1, \dots, N_1 + N_2, \quad \varepsilon_i = 0, i = N_1 + N_2 + 1, \dots, N.$$

Assume that $N = N_1$. This can be treated as before. The normalized eigenfunctions of H are

$$|n_1, \dots, n_N\rangle = \prod_{i=1}^{N} |n_i\rangle_i,$$

where

$$|n_i\rangle = (\frac{m_i\omega_i}{\pi\hbar})^{\frac{1}{4}} H_{n_i}(q_i\sqrt{\frac{m_i\omega_i}{\hbar}})e^{-\frac{m_i\omega_i}{2\hbar}q^2}.$$

The eigenvalue of $|n_1, \dots, n_N\rangle$ is equal to

$$\hbar\sum_{i=1}^{N}\omega_i(n_i + \frac{1}{2}) = \hbar\sum_{i=1}^{n}\omega_i(n_i + \frac{1}{2}).$$

So we see that the eigensubspaces of the operator H are no longer one-dimensional.

Now let us assume that $N = N_3$. In this case we are describing a system of a free particle in potential zero field. The Schrödinger operator H becomes a *Laplace operator*:

$$H = \frac{1}{2}\sum_{i=1}^{n}\frac{P_i^2}{m_i} = -\frac{1}{2}\sum_{i=1}^{n}\frac{\hbar^2}{m_i}\frac{d^2}{dq_i^2}.$$

After scaling the coordinates we may assume that

$$H = -\sum_{i=1}^{n}\frac{d^2}{dx_i^2}.$$

If $N = N_1$, the eigenfunctions of the operator H with eigenvalue E are solutions of the differential equation

$$(\frac{d^2}{dx^2} + E)\psi = 0.$$

The solutions are

$$\psi = C_1 e^{\sqrt{-E}x} + C_2 e^{-\sqrt{-E}x}.$$

It is clear that these solutions do not belong to $L^2(\mathbb{R})$. Thus we have to look for solutions in the space of distributions in $\mathbb{R}$. Clearly, when $E < 0$, the solutions do not belong to this space either. Thus, we have to assume that $E \geq 0$. Write $E = k^2$. Then the space of eigenfunctions of H with eigenvalue E is two-dimensional and is spanned by the functions e^{ikx}, e^{-ikx}.

The orthonormal basis of generalized eigenfunctions of H is continuous and consists of functions

$$\psi_k(x) = \frac{1}{\sqrt{2\pi}}e^{ikx}, \quad k \in \mathbb{R}.$$

(see Lecture 6, Example 9). However, this generalized state does not have any physical interpretation as a state of a one-particle system. In fact, since $\int_{\mathbb{R}} e^{ikx}e^{-ikx}dx = \int_{\mathbb{R}} dx$ has no meaning even in the theory of distributions, we cannot normalize the wave function e^{ikx}. Because of this we cannot define $Tr(QP_{\psi_k})$ which, if it would be defined, must be equal to

$$\frac{1}{2\pi}\int_{\mathbb{R}} xe^{ikx}e^{-ikx}dx = \frac{1}{2\pi}\int_{\mathbb{R}} xdx = \infty.$$

Similarly we see that $\mu_Q(\lambda)$ is not defined.

One can reinterpret the wave function $\psi_k(x)$ as a beam of particles. The failure of normalization means simply that the beam contains infinitely many particles.

For any pure state $\psi(x)$ we can write (see Example 9 of Lecture 6):

$$\psi(x) = \int_{\mathbb{R}} \phi(k)\psi_k dk = \int_{\mathbb{R}} \phi(k)e^{ikx}dk = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \hat{\psi}(k)e^{ikx}dk.$$

It evolves according to the law

$$\psi(x;t) = \int_{\mathbb{R}} \hat{\psi}(k)e^{kx-tk^2}dk.$$

Now, if we take $\psi(x)$ from $L^2(\mathbb{R})$ we will be able to normalize $\psi(x,t)$ and obtain an evolution of a pure state of a single particle. There are some special wave functions (*wave packets*) which minimizes the product of the dispersions $\Delta_\mu(Q)\Delta_\mu(P)$.

Finally, if $N = N_3$, we can find the eigenfunctions of H by reducing to the case $N = N_1$. To do this we replace the unknown x by ix. The computations show that the eigenfunctions look like $P_n(x)e^{x^2/2}$ and do not belong to the space of distributions $\mathcal{D}(\mathbb{R})$.

Exercises.

1. Consider the quantum picture of a free one-dimensional particle with the Schrödinger operator $H = P^2/2m$ confined in an interval $[0, L]$. Solve the Schrödinger equation with boundary conditions $\psi(0) = \psi(L) = 0$.

2. Find the decomposition of the delta function $\delta(q)$ in terms of orthonormal eigenfunctions of the Schrödinger operator $\frac{1}{2}(P^2 + \hbar\omega^2 Q^2)$.

3. Consider the one-dimensional harmonic oscillator. Find the density distribution for the values of the momentum observable P at a pure state $|n\rangle$.

4. Compute the dispersions of the observables P and Q (see Problem 1 from lecture 5) for the harmonic oscillator. Find the pure states ω for which $\Delta_\omega Q\Delta_\omega P = \hbar$ (compare with Problem 2 in Lecture 5).

5. Consider the quantum picture of a particle in a potential well, i.e., the quantization of the mechanical system with the Hamiltonian $h = p^2/2m + U(q)$, where $U(q) = 0$ for

$q \in (0,a)$ and $U(q) = 1$ otherwise. Solve the Schrödinger equation, and describe the stationary pure states.

6. Prove the following equality for the generating function of Hermite polynomials:

$$\Phi(t,x) = \sum_{m=1}^{\infty} \frac{1}{m!} H_m(x) t^m = e^{-t^2+2tx}.$$

7. The Hamiltonian of a three-dimensional isotropic harmonic oscillator is

$$H = \frac{1}{2m} \sum_{i=1}^{3} p_i^2 + \frac{m^2\omega^2}{2} \sum_{i=1}^{3} q_i^2.$$

Solve the Schrödinger equation $H\psi(x) = E\psi(x)$ in rectangular and spherical coordinates.

Lecture 8. CENTRAL POTENTIAL

8.1 The Hamiltonian function of the harmonic oscillator (7.1) is a special case of the Hamiltonian functions of the form

$$H = \frac{1}{2m}\sum_{i=1}^{n} p_i^2 + V(\sum_{i=1}^{n} q_i^2). \tag{8.1}$$

After quantization we get the Hamiltonian operator

$$H = -\frac{\hbar^2}{2m}\Delta + V(r), \tag{8.2}$$

where

$$r = ||\mathbf{x}||$$

and

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \tag{8.3}$$

is the Laplace operator.

An example of the function $V(r)$ (called the *radial potential*) is the function c/r used in the model of hydrogen atom, or the function

$$V(r) = g\frac{e^{-\mu r}}{r},$$

called the *Yukawa potential.*

The Hamiltonian of the form (8.1) describes the motion of a particle in a central potential field. Also the same Hamiltonian can be used to describe the motion of two particles with Hamiltonian

$$H = \frac{1}{2m_1}||\mathbf{p}_1||^2 + \frac{1}{2m_2}||\mathbf{p}_2||^2 + V(||\mathbf{q}_1 - \mathbf{q}_2||).$$

Here we consider the configuration space $T(\mathbb{R}^{2n})^*$. Let us introduce new variables

$$\mathbf{q}_1' = \mathbf{q}_1 - \mathbf{q}_2, \quad \mathbf{q}_2' = \frac{m_1\mathbf{q}_1 + m_2\mathbf{q}_2}{m_1 + m_2},$$

$$\mathbf{p}_1' = \mathbf{p}_1 - \mathbf{p}_2, \quad \mathbf{p}_2' = \frac{m_1\mathbf{p}_1 + m_2\mathbf{p}_2}{m_1 + m_2}.$$

Then

$$H = \frac{1}{2M_1}||\mathbf{p}_1'||^2 + \frac{1}{2M_2}||\mathbf{p}_2'||^2 + V(||\mathbf{q}_1'||), \tag{8.4}$$

where

$$M_1 = m_1 + m_2, \quad M_2 = \frac{m_1 m_2}{m_1 + m_2}.$$

The Hamiltonian (8.4) describes the motion of a free particle and a particle moving in a central field. We can solve the corresponding Schrödinger equation by separation of variables.

8.2 We shall assume from now on that $n = 3$. By quantizing the angular momentum vector $\mathbf{m} = m\mathbf{q} \times \mathbf{p}$ (see Lecture 3), we obtain the operators of angular momentums

$$\mathbf{L} = (L_1, L_2, L_3) = (Q_2P_3 - Q_3P_2, Q_3P_1 - Q_1P_3, Q_1P_2 - Q_2P_1). \tag{8.5}$$

Let

$$L = L_1^2 + L_2^2 + L_3^2.$$

In coordinates

$$L_1 = i\hbar(x_3\frac{\partial}{\partial x_2} - x_3\frac{\partial}{\partial x_2}), \quad L_2 = i\hbar(x_1\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_1}), \quad L_3 = i\hbar(x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}),$$

$$L = \hbar^2[(\sum_{i=1}^3 x_i^2)\Delta - \sum_{i=1}^3 x_i^2\frac{\partial^2}{\partial x_i^2} - 2\sum_{1\le i<j\le 3} x_i x_j\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} - 2\sum_{i=1}^3 x_i\frac{\partial}{\partial x_i}]. \tag{8.6}$$

The operators L_i, L satisfy the following commutator relations

$$[L_1, L_2] = iL_3, \quad [L_2, L_3] = iL_1, \quad [L_3, L_1] = iL_2, \quad [L_i, L] = 0, \quad i = 1, 2, 3. \tag{8.7}$$

Note that the Lie algebra $\mathfrak{so}(3)$ of skew-symmetric 3×3 matrices is generated by matrices e_i satisfying the commutation relations

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2$$

(see Example 5 from Lecture 3). If we assign iL_i to e_i, we obtain a linear representation of the algebra $\mathfrak{so}(3)$ in the Hilbert space $L^2(\mathbb{R}^3)$. It has a very simple interpretation. First observe that any element of the orthogonal group $SO(3)$ can be written in the form

$$g(\phi) = \exp(\phi_1 e_1 + \phi_2 e_2 + \phi_3 e_3)$$

for some real numbers ϕ_1, ϕ_2, ϕ_3. For example,

$$\exp(\phi_1 e_1) = \exp\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\phi_1 \\ 0 & \phi_1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_1 & -\sin\phi_1 \\ 0 & \sin\phi_1 & \cos\phi_1 \end{pmatrix}.$$

If we introduce the linear operators

$$T(\phi) = e^{i(\phi_1 L_1 + \phi_2 L_2 + \phi_3 L_3)},$$

then $g(\phi) \to T(\phi)$ will define a linear unitary representation of the orthogonal group

$$\rho : SO(3) \to GL(L^2(\mathbb{R}^3))$$

in the Hilbert space $L^2(\mathbb{R}^3)$.

Lemma. *Let $T = \rho(g)$. For any $f \in L^2(\mathbb{R}^3)$ and $x \in \mathbb{R}^3$,*

$$Tf(\mathbf{x}) = f(g^{-1}\mathbf{x}).$$

Proof. Obviously it is enough to verify the asssertion for elements $g = \exp(te_i)$. Without loss of generality, we may assume that $i = 1$. The function $f(\mathbf{x}, t) = e^{-itL_1} f(\mathbf{x})$ satisfies the equation

$$\frac{\partial f(\mathbf{x}, \phi)}{\partial t} = -iL_1 f(\mathbf{x}, t) = x_3 \frac{\partial f(\mathbf{x}, t)}{\partial x_2} - x_2 \frac{\partial f(\mathbf{x}, t)}{\partial x_3}.$$

One immediately checks that the function

$$\psi(\mathbf{x}, t) = f(g^{-1}\mathbf{x}) = f(x_1, \cos tx_2 + \sin tx_3, -\sin tx_2 + \cos tx_3)$$

satisfies this equation. Also, both $f(\mathbf{x}, t)$ and $f(g^{-1}\mathbf{x})$ satisfy the same initial condition $f(\mathbf{x}, 0) = \psi(\mathbf{x}, 0) = f(x)$. Now the assertion follows from the uniqueness of a solution of a partial linear differential equation of order 1.

The previous lemma shows that the group $SO(3)$ operates naturally in $L^2(\mathbb{R}^3)$ via its action on the arguments. The angular moment operators L_i are infinitesimal generators of this action.

8.3 The Schrödinger differential equation for the Hamiltonian (8.2) is

$$H\psi(\mathbf{x}) = [-\frac{\hbar^2}{2m}\Delta + V(r)]\psi(\mathbf{x}) = E\psi(\mathbf{x}). \tag{8.8}$$

It is clear that the Hamiltonian is invariant with respect to the linear representation ρ of $SO(3)$ in $L^2(\mathbb{R}^3)$. More precisely, for any $g \in SO(3)$, we have

$$\rho(g) \circ H \circ \rho(g)^{-1} = H.$$

This shows that each eigensubspace of the operator H is an invariant subspace for all operators $T \in \rho(SO(3))$. Using the fact that the group $SO(3)$ is compact one can show that any linear unitary representation of $SO(3)$ decomposes into a direct sum of finite-dimensional irreducible representations. This is called the Peter-Weyl Theorem (see, for example, [**Sternberg**]). Thus the space V_E of solutions of equation (5.9) decomposes into the diret sum of irreducible representations. The eigenvalue E is called *non-degenerate* if V_E is an irreducible representation. It is caled *degenerate* otherwise.

Let us find all irreducible representations of $SO(3)$ in $L^2(\mathbb{R}^3)$. First, let us use the canonical homomorphism of groups

$$\tau : SU(2) \to SO(3).$$

Here the group $SU(2)$ consists of complex matrices

$$U = \begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}, \quad |z_1|^2 + |z_2|^2 = 1.$$

The homomorphism τ is defined as follows. The group $SU(2)$ acts by conjugation on the space of skew-Hermitian matrices of the form

$$A(\mathbf{x}) = \begin{pmatrix} ix_1 & x_2 + ix_3 \\ -(x_2 - ix_3) & -ix_1 \end{pmatrix}.$$

They form the Lie algebra of $SU(2)$. We can identify such a matrix with the 3-vector $\mathbf{x} = (x_1, x_2, x_3)$. The determinant of $A(\mathbf{x})$ is equal to $||\mathbf{x}||^2$. Since the determinant of $UA(\mathbf{x})U^{-1}$ is equal to the determinant of $A(\mathbf{x})$, we can define the homomorphism τ by the property

$$\tau(U) \cdot \mathbf{x} = U \cdot A(\mathbf{x}) \cdot U^{-1}.$$

It is easy to see that τ is surjective and $\mathrm{Ker}(\tau) = \{\pm I_2\}$. If $\rho : SO(3) \to GL(V)$ is a linear representation of $SO(3)$, then $\rho \circ \tau : SU(2) \to GL(V)$ is a linear representation of $SU(2)$. Conversely, if $\rho' : SU(2) \to GL(V)$ is a linear representation of $SU(2)$ satisfying

$$\rho'(-I_2) = \mathbf{id}_V, \tag{8.9}$$

then $\rho' = \rho \circ \tau$, where ρ is defined by $\rho(g) = \rho'(U)$ with $\tau(U) = g$. So, it is enough to classify irreducible representations of $SU(2)$ in $L^2(\mathbb{R}^2)$ which satisfy (8.9).

Being a subgroup of $SL(2, \mathbb{C})$, the group $SU(2)$ acts naturally on the space $V(d)$ of complex homogeneous polynomials $P(z_1, z_2)$ of degree d. If $d = 2k$ is even, then this representation satisfies (8.9). The representation $V(2k)$ can be made unitary if we introduce the inner product by

$$(P, Q) = \frac{1}{\pi^2} \int_{\mathbb{C}^2} P(z_1, z_2)\overline{Q(z_1, z_2)} e^{-|z_1|^2 - |z_2|^2} dz_1 dz_2.$$

We shall deal with this inner product in the next Lecture. One verifies that the functions

$$e_s = \frac{z_1^{k+s} z_2^{k-s}}{[(s+k)!(k-s)!]^{1/2}}, \quad s = -k, -k+1, \ldots, k-1, k$$

form an orthonormal basis in $V(2k)$. It is not difficult to check that the space $V(2k)$ is irreducible.

Let us use spherical coordinates (r,θ,ϕ) in $\mathbb{R}^3$. For each fixed r, the function $f(r,\theta,\phi)$ is a function on the unit sphere S^2 which belongs to the space $L^2(S^2,\mu)$, where $\mu = \sin\theta d\theta d\phi$. Now let us identify $V(2k)$ with a subspace $\mathcal{H}^k$ of $L^2(S^2)$ in such a way that the action of $SU(2)$ on $V(2k)$ corresponds to the action of $SO(3)$ on $\mathcal{H}^k$ under the homomorphism τ. Let $\mathcal{P}^k$ be the space of complex-valued polynomials on $\mathbb{R}^3$ which are homogeneous of degree k. Since each such a function is completely determined by its values on S^2 we can identify $\mathcal{P}^k$ with a linear subspace of $L^2(S^2)$. Its dimension is equal to $(k+1)(k+2)/2$. We set

$$\mathcal{H}^k = \{P \in \mathcal{P}^k : \Delta(P) = 0\}.$$

Elements of $\mathcal{H}^k$ are called *harmonic polynomials* of degree k. Obviously the Laplace operator Δ sends $\mathcal{P}^k$ to $\mathcal{P}^{k-2}$. Looking at monomials, one can compute the matrix of the map

$$\Delta : \mathcal{P}^k \to \mathcal{P}^{k-2}$$

and deduce that this map is surjective. Thus

$$\dim\mathcal{H}^k = \dim\mathcal{P}^k - \dim\mathcal{P}^{k-2} = 2k+1. \tag{8.10}$$

Obviously, $\mathcal{H}^k$ is invariant with respect to the action of $SO(3)$. We claim that $\mathcal{H}^k$ is an irreducible representation isomorphic to $V(2k)$. Since each polynomial from $V(2k)$ satisfies $F(-z_1,-z_2) = F(z_1,z_2)$, we can write it uniquely as a polynomial of degree k in

$$x_1 = z_1^2 - z_2^2, \quad x_2 = i(z_1^2 + z_2^2), \quad x_3 = -2z_1z_2.$$

This defines a surjective linear map

$$\alpha : \mathcal{P}^k \to V(2k).$$

It is easy to see that it is compatible with the actions of $SU(2)$ on $\mathcal{P}^k$ and $SO(3)$ on $V(2k)$. We claim that its restriction to H_k is an isomorphism. Notice that

$$P^2 = H^2 \oplus \mathbb{C}Q,$$

where $Q = x_1^2 + x_2^2 + x_3^2$. We can prove by induction on k that

$$\mathcal{P}^k = \mathcal{H}^k \oplus Q\mathcal{P}^{k-2} = \bigoplus_{i=0}^{[k/2]} Q^i\mathcal{H}^{k-2i}. \tag{8.11}$$

Since $\alpha(Q) = 0$, we obtain $\alpha(Q\mathcal{P}^{k-2}) = 0$. This shows that $\alpha(\mathcal{H}^k) = \alpha(\mathcal{P}^k) = V(2k)$. Since $\dim\mathcal{H}^k = \dim V(2k)$, we are done.

Using the spherical coordinates we can write any harmonic polynomial $P \in \mathcal{H}^k$ in the form

$$P = r^k Y(\theta,\phi),$$

where $Y \in L^2(S^2)$. The functions $Y(\theta,\phi)$ on S^2 obtained in this way from harmonic polynomials of degree k are called *spherical harmonics* of degree k. Let $\tilde{\mathcal{H}}^k$ denote the space of such functions.

The expression for the Laplacian Δ in polar coordinates is

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_{S^2}, \tag{8.12}$$

where

$$\Delta_{S^2} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi} \tag{8.13}$$

is the *spherical Laplacian*. We have

$$0 = \Delta r^k Y(\theta,\phi) = (k(k+1) + \Delta_{S^2})r^{k-2}Y(\theta,\phi).$$

This shows that $Y(\theta,\phi)$ is an eigenvector of Δ_{S^2} with eigenvalue $-k(k+1)$.

In spherical coordinates the operators L_i have the form

$$L_1 = i(\sin\phi\frac{\partial}{\partial\theta} + \mathrm{cot}\theta\cos\phi\frac{\partial}{\partial\phi}), \quad L_2 = i(\cos\phi\frac{\partial}{\partial\theta} - \mathrm{cot}\theta\sin\phi\frac{\partial}{\partial\phi}),$$

$$L_3 = -i\frac{\partial}{\partial\phi}, \quad L = L_1^2 + L_2^2 + L_3^2 = -\Delta_{S^2}. \tag{8.14}$$

These operators act on $L^2(S^2)$ leaving each subspace $\mathcal{H}^k$ invariant. Let

$$L_+ = L_1 + iL_2 = e^{i\phi}(\frac{\partial}{\partial\theta} + i\mathrm{cot}\theta\frac{\partial}{\partial\phi}), \quad L_- = L_1 - iL_2 = e^{i\phi}(-\frac{\partial}{\partial\theta} + i\mathrm{cot}\theta\frac{\partial}{\partial\phi}).$$

Let $Y_{kk}(\theta,\phi) \in \tilde{\mathcal{H}}^k$ be an eigenvector of L_3 with eigenvalue k. Let us see how to find it. We have

$$-i\frac{\partial Y_{kk}(\theta,\phi)}{\partial\partial\phi} = kY_{kk}(\theta,\phi).$$

This gives

$$Y_{kk}(\theta,\phi) = e^{ik\phi}F_k(\theta).$$

Now we use

$$L_-\circ L_+ = L_1^2 + i[L_1,L_2] + L_2^2 = L_1^2 - L_3 + L_2^2 = L - L_3^2 - L_3,$$

to get

$$(L_-\circ L_+)Y_{kk}(\theta,\phi) = (L - L_3^2 - L_3)Y_{kk}(\theta,\phi) = (k(k+1) - k^2 - k)Y_{kk}(\theta,\phi) = 0.$$

So we can find $F_k(\theta)$ by requiring that $L_+Y_{kk}(\theta,\phi) = 0$. This gives the equation

$$\frac{\partial F_k(\theta)}{\partial\theta} - k\mathrm{cot}\theta F_K(\theta) = 0.$$

Solving this equation we get

$$F_k(\theta) = C \sin^k \theta.$$

Thus

$$Y_{kk}(\theta, \phi) := e^{ik\phi} \sin^k \theta$$

Now we use that

$$[L_-, L_3] = L_-.$$

Thus, applying L_- to Y_{kk}, we get

$$L_- Y_{kk} = L_- L_3 Y_{kk} - L_3 L_- Y_{kk} = k L_- Y_{kk} - L_3 L_- Y_{kk},$$

hence

$$L_3(L_- Y_{kk}) = (k-1) L_- Y_{kk}.$$

This shows that

$$Y_{k\ k-1} := L_- Y_{kk}$$

is an eigenvector of L_3 with eigenvalue $k-1$. Continuing applying L_-, we find a basis of $\tilde{\mathcal{H}}^k$ formed by the functions $Y_{km}(\theta, \phi)$ which are eigenvectors of the operator L_3 with eigenvalue $m = k, k-1, \ldots, -k+1, -k$. After appropriate normalization, we find explictly the formulas

$$Y_{km}(\theta, \phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} P_k^m(\cos\theta),$$

where

$$P_k^m(t) = \sqrt{\frac{(2k+1)(k+m)!}{2(k-m)!}} \frac{1}{2^k k!} (1-t^2)^{-\frac{m}{2}} \frac{d^{k-m}(t^2-1)^k}{dt^{k-m}}. \tag{8.15}$$

They are the so-called normalized *associated Legendre polynomials* (see **[Whittaker]**, Chapter 15). The spherical harmonics $Y_{km}(\theta, \phi)$ from (8.12) with fixed k form an orthonormal basis in the space $\tilde{\mathcal{H}}^k$.

We now claim that the direct sum $\oplus \tilde{\mathcal{H}}^k$ is dense in $L^2(S)$. We use that the set of continuous functions on S^2 is dense in $L^2(S^2)$ and that we can approximate any continuous function on S^2 by a polynomial on $\mathbb{R}^3$. Now any polynomial can be written as a sum of its homogeneous components. Finally, by (8.11), any homogeneous polynomial can be written in the form $f + Qf_1 + Q^2 f_2 + \ldots$, where $f_i \in \mathcal{H}^{k-2i}$. This proves our claim.

As a corollary of our claim, we obtain that any function from $L^2(\mathbb{R}^3)$ has an expansion in spherical harmonics

$$f = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=-k}^{k} c_{nkm} f_n(r) Y_{km}(\theta, \phi), \tag{8.16}$$

where $(f_0(r), f_1(r), \ldots)$ is a basis in the space $L^2(\mathbb{R}_{\geq 0})$.

8.4 Let us go back to the study of the motion of a particle in a central field. We are looking for the space $W(E)$ of solutions of the Schrödinger equation (8.8). It is invariant with respect to the natural representation of $SO(3)$ in $L^2(\mathbb{R}^3)$. Assume that E is non-degenerate, i.e., the space $W(E)$ is an irreducible representation of $SO(3)$. Then we can write each function from $W(r)$ in the form

$$f(\mathbf{x}) = f(r,\theta,\phi) = g(r)\sum_{m=-k}^{k} c_m Y_{km}(\theta,\phi), \tag{8.17}$$

where $g(r) \in L^2(\mathbb{R}_{\geq 0})$. Using the expression (8.12) for the Laplacian in spherical coordinates, we get

$$0 = [-\frac{\hbar^2}{2m}\Delta + V(r) - E]g(r)Y_{km} = [-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) - \Delta_{S^2} + V(r) - E]f =$$

$$= Y_{km}(\theta,\phi)[-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) + k(k+1) + V(r) - E]g(r).$$

If we set

$$h(r) = g(r)r,$$

we get the following equation for $h(r)$:

$$[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2 k(k+1)}{2mr^2} + V(r)]h(r) = Eh(r). \tag{8.18}$$

It is called the *radial Schrödinger equation*. It is the Schrödinger equation for a particle moving in one-dimensional space $\mathbb{R}$ with central potential $V(r)' = \frac{k(k+1)\hbar^2}{2mr^2} + V(r)$. One can prove that for each E the space $W_k(E)$ of solutions of (8.18) is of dimension ≤ 1. This shows that E is non-degenerate unless $W_k(E)$ is not zero for different k. Notice that the number E for which $W_k(E) \neq \{0\}$ can be interpreted as the energy of the particle in the pure state defined by the wave function

$$\psi(r,\theta,\phi) = g(r)Y_{km}(\theta,\phi) = h(r)Y_{km}(\theta,\phi)/r.$$

The set of such numbers is the union of the spectra of the operators

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2 k(k+1)}{2mr^2} + V(r)$$

for all $k = 0, 1 \ldots$. The number m is the eigenvalue of the moment operator L_3. It is the quantum number describing the moment of the particle with respect to the z-axis. The number $k(k+1)$ is the eigenvalue of the operator L. It describes the norm of the moment vector of the particle. The solutions of (8.15) depend very much on the property of the potential function $V(r)$. In the next section we shall consider one special case.

8.5 Let us consider the case of the *hydrogen atom*. In this case

$$V(r) = -\frac{e^2 Z}{r}, \quad m = m_e M/(m_e + M), \tag{8.19}$$

where $e > 0$ is the absolute value of the charge of the electron, Z is the *atomic number*, m_e is the mass of the electron and M is the mass of the nucleus. The atomic number is equal to the number of positively charged protons in the nucleus. It determines the position of the atom in the periodic table. Each proton is matched by a negatively charged electron so that the total charge of the atom is zero. For example, the hydrogen has one electron and one proton but the atom of helium has two protons and two electrons. The mass of the nucleus M is equal to Am_p, where m_p is the mass of proton, and A is equal to $Z+N$, where N is the number of neutrons. Under the high-energy condition one may remove or add one or more electrons, creating *ions* of atoms. For example, there is the helium ion He^+ with $Z = 2$ but with only one electron. There is also the lithium ion Li^{++} with one electron and $Z = 3$. So our potential (8.19) describes the structure of one-electron ions with atomic number Z.

We shall be solving the problem in *atomic units*, i.e., we assume that $\hbar = 1, m = 1, e = 1$. Then the radial Schrödinger equation has the form

$$h(r)'' - \frac{k(k+1) - 2Zr - 2r^2E}{r^2}h(r) = 0.$$

We make the substitution $h(r) = e^{-ar}v(r)$, where $a^2 = -2E$, and transform the equation to the form

$$r^2v(r)'' - 2ar^2v(r)' + [2Zr - k(k+1)]v(r) = 0. \tag{8.20}$$

It is a second order ordinary differential equation with regular singular point $r = 0$. Recall that a linear ODE of order n

$$a_0(x)v^{(n)} + a_1(x)v^{(n-1)} + \ldots + a_n(x) = 0 \tag{8.21}$$

has a *regular singular point* at $x = c$ if for each $i = 0, \ldots, n$, $a_i(x) = (x-c)^{-i}b_i(x)$ where $b_i(x)$ is analytic in a neighborhood of c. We refer for the theory of such equations to [**Whittaker**], Chapter 10. In our case n equals 2 and the regular singular point is the origin. We are looking for a formal solution of the form

$$v(r) = r^\alpha(1 + \sum_{n=1}^{\infty} c_n r^n). \tag{8.22}$$

After substituting this in (8.20), we get the equation

$$\alpha^2 + (b_1(0) - 1)\alpha + b_2(0) = 0.$$

Here we assume that $b_0(0) = 1$. In our case

$$b_1(0) = 0, b_2(0) = -k(k+1).$$

Thus

$$\alpha = k+1, -k.$$

This shows that $v(r) \sim Cr^{k+1}$ or $v(r) \sim Cr^{-k}$ when r is close to 0. In the second case, the solution $f(r,\theta,\phi) = h(r)Y_{km}(\theta,\phi)/r$ of the original Schrödinger equation is not continuous at 0. So, we should exclude this case. Thus,

$$\alpha = k + 1.$$

Now if we plug $v(r)$ from (8.22) in (8.20), we get

$$\sum_{i=0}^{\infty} c_i(i(i-1)r^i + 2(k+1)ir^i - 2iar^i + (2Z - 2ak - 2a)r^{i+1}] = 0.$$

This easily gives

$$c_{i+1} = 2\frac{a(i+k+1) - Z}{(i+1)(i+2k+2)}c_i. \tag{8.23}$$

Applying the known criteria of convergence, we see that the series $v(r)$ converges for all r. When i is large enough, we have

$$c_{i+1} \sim \frac{2a}{i+1}c_i.$$

This means that for large i we have

$$c_i \sim C\frac{(2a)^k}{i!},$$

i.e.,

$$v(r) \sim Cr^{k+1}e^{-ar}e^{2ar} = Cr^{k+1}e^{ar}.$$

Since we want our function to be in $L^2(\mathbb{R})$, we must have $C = 0$. This can be achieved only if the coefficients c_i are equal to zero starting from some number N. This can happen only if

$$a(i+k+1) = Z$$

for some i. Thus

$$a = \frac{Z}{i+k+1}$$

for some i. In particular, $-2E = a^2 > 0$ and

$$E = E_{ki} := -\frac{Z^2}{2(k+i+1)^2}.$$

This determines the discrete spectrum of the Schrödinger operator. It consists of numbers of the form

$$E_n = -\frac{Z^2}{n^2}. \tag{8.24}$$

The corresponding eigenfunctions are linear combinations of the functions

$$\psi_{nkm}(r,\theta,\phi) = r^k e^{-Zr/n}L_{nk}(r)Y_{km}(\theta,\phi), \quad 0 \le k < n, -k \le m \le k, \tag{8.25}$$

where $L_{nk}(r)$ is the *Laguerre* polynomial of degree $n-k-1$ whose coefficients are found by formula (8.23). The coefficient c_0 is determined by the condition that $||\psi_{nkm}||_{L_2} = 1$.

We see that the dimension of the space of solutions of (8.8) with given $E = E_n$ is equal to

$$q(n) = \sum_{k=0}^{n-1}(2k+1) = n^2.$$

Each $E_n, n \neq 1$ is a degenerate eigenvalue. An explanation of this can be found in Exercise 8. If we restore the units, we get

$$m_e = 9.11 \times 10^{-28}\text{g}, \quad m_p = 1.67 \times 10^{-24}\text{g},$$

$$e = 1.6 \times 10^{-9}C, \quad \hbar = 4.14 \times 10^{-15}\text{eV} \cdot \text{s},$$

$$E_n = -\frac{me^4}{2n^2\hbar^2} = -27.21\frac{Z^2}{2n^2}\text{eV}. \tag{8.25}$$

In particular,

$$E_1 = -13.6\text{eV}.$$

The absolute value of this number is called the *ionization potential.* It is equal to the work that is required to pull out the electron from the atom.

The formula (8.25) was discovered by N. Bohr in 1913. In quantum mechanics it was obtained first W. Pauli and, independently, by E. Schrödinger in 1926.

The number

$$\omega_{mn} = \frac{E_n - E_m}{\hbar} = \frac{me^4}{2\hbar^3}(\frac{1}{n^2} - \frac{1}{m^2}), \quad n < m.$$

is equal to the frequency of spectral lines (the frequency of electromagnetic radiation emitted by the hydrogen atom when the electron changes the possible energy levels). This is known as Balmer's formula, and was known a long time before the discovery of quantum mechanics.

Recall that $|\psi_{nkm}(r, \theta, \phi)|$ is interpreted as the density of the distribution function for the coordinates of the electron. For example, if we consider the *principal state* of the electron corresponding to $E = E_1$, we get

$$|\psi_{100}(\mathbf{x})| = \frac{1}{\sqrt{\pi}}e^{-r}.$$

The integral of this function over the ball $B(r)$ of radius r gives the probability to find the electron in $B(r)$. This shows that the density function for the distribution of the radius r is equal to

$$\rho(r) = 4\pi|\psi_{100}(\mathbf{x})|^2 r^2 = 4e^{-2r}r^2.$$

The maximum of this function is reached at $r = 1$. In the restored physical units this translates to

$$r = \hbar^2/me^2 = .529 \times 10^{-8}\text{cm}. \tag{8.26}$$

This gives the approximate size of the hydrogen atom.

What we have described so far are the "bound states" of the hydrogen atom and of hydrogen type ions. They correspond to the discrete spectrum of the Schrödinger operator. There is also the continuous spectrum corresponding to $E > 0$. If we enlarge the Hilbert space by admitting distributions, we obtain the solutions of (8.4) which behave like "plane waves" at infinity. It represents a "scattering state" corresponding to the ionization of the hydrogen atom.

Finally, let us say a few words about the case of atoms with $N > 1$ electrons. In this case the Hamiltonian operator is

$$\frac{\hbar}{2m}\sum_{i=1}^{N}\Delta_i - \sum_{i=1}^{N}\frac{Ze^2}{r_i} + \sum_{1\le i<j\le N}\frac{e^2}{|r_i - r_j|},$$

where r_i are the distances from the i-th electron to the nucleus. The problem of finding an exact solution to the corresponding Schrödinger equation is too complicated. It is similar to the corresponding n-body problem in celestial mechanics. A possible approach is to replace the potential with the sum of potentials $V(r_i)$ for each electron with

$$V(r_i) = -\frac{Ze^2}{r_i} + W(r_i).$$

The Hilbert space here becomes the tensor product of N copies of $L^2(\mathbb{R}^3)$.

Exercises.

1. Show that the the polynomials $(x+iy)^k$ are harmonic.

2. Compute the mathematical expectation for the moment operators L_i at the states ψ_{nkm}.

3. Find the pure states for the helium atom.

4. Compute ψ_{nkm} for $n = 1, 2$.

5. Find the probability distribution for the values of the impulse operators P_i at the state $\psi_{1,0,0}$.

6. Consider the vector operator

$$\mathbf{A} = \frac{1}{r}\mathbf{Q} - \frac{1}{2}(\mathbf{P}\times\mathbf{L} - \mathbf{L}\times\mathbf{Q}),$$

where $\mathbf{Q} = (Q_1, Q_2, Q_3), \mathbf{P} = (P_1, P_2, P_3)$. Show that each component of $\mathbf{A}$ commutes with the Hamiltonian $H = \frac{1}{2}\mathbf{P}\cdot\mathbf{P} - \frac{1}{r}$.

7. Show that the operators L_i together with operators $U_i = \frac{1}{\sqrt{-2E_n}}A_i$ satisfy the commutator relations $[L_j, U_k] = ie_{jks}U_s, [U_j, U_k] = ie_{jks}L_s$, where e_{jks} is skew-symmetric with values $0, 1, -1$ and $e_{123} = 1$.

8. Using the previous exercise show that the space of states of the hydrogen atom with energy E_n is an irreducible representation of the orthogonal group $SO(4)$.

Lecture 9. THE SCHRÖDINGER REPRESENTATION

9.1 Let V be a real finite-dimensional vector space with a skew-symmetric nondegenerate bilinear form $S : V \times V \to \mathbb{R}$. For example, $V = \mathbb{R}^n \oplus \mathbb{R}^n$ with the bilinear form

$$S((\mathbf{u}', \mathbf{v}'), (\mathbf{u}'', \mathbf{v}'')) = \mathbf{u}'\mathbf{v}'' - \mathbf{u}''\mathbf{v}'. \tag{9.1}$$

We define the *Heisenberg group* $\tilde{V}$ as the central extension of V by the circle $\mathbb{C}_1^* = \{z \in \mathbb{C} : |z| = 1\}$ defined by S. More precisely, it is the set $V \times \mathbb{C}_1^*$ with the group law defined by

$$(x, \lambda) \cdot (x', \lambda') = (x + x', e^{iS(x,x')}\lambda\lambda'). \tag{9.2}$$

There is an obvious complexification $\tilde{V}_{\mathbb{C}}$ which is the extension of $V_{\mathbb{C}} = V \otimes \mathbb{C}$ with the help of $\mathbb{C}^*$.

We shall describe its Schrödinger representation. Choose a complex structure on V

$$J : V \to V$$

(i.e., an $\mathbb{R}$-linear operator with $J^2 = -I_V$) such that
(i) $S(Jx, Jx') = S(x, x')$ for all $x, x' \in V$;
(ii) $S(Jx, x') > 0$ for all $x, x' \in V$.

Since $J^2 = -I_V$, the space $V_{\mathbb{C}} = V \otimes \mathbb{C} = V + iV$ decomposes into the direct sum $V_{\mathbb{C}} = A \oplus \bar{A}$ of eigensubspaces with eigenvalues i and $-i$, respectively. Here the conjugate is an automorphism of $V_{\mathbb{C}}$ defined by $v + iw \to v - iw$. Of course,

$$A = \{v - iJv : v \in V\}, \quad \bar{A} = \{v + iJv : v \in V\}. \tag{9.3}$$

Let us extend S to $V_{\mathbb{C}}$ by $\mathbb{C}$-linearity, i.e.,

$$S(v + iw, v' + iw') := S(v, v') - S(w, w') + i(S(w, v') + S(v, w')).$$

Since

$$S(v \pm iJv, v' \pm iJv') = S(v, v') - S(Jv, Jv') \pm i(S(v, Jv') + S(Jv, v'))$$

$$= \pm i(S(Jv, J^2v') + S(Jv, v')) = \pm i(S(Jv, -v') + S(Jv, v'))$$
$$= \pm i(-S(Jv, v') + S(Jv, v')) = 0, \quad (9.4)$$

we see that A and $\bar{A}$ are complex conjugate isotropic subspaces of S in $V_{\mathbb{C}}$. Observe that

$$S(Jv, w) = S(J^2v, Jw) = S(-v, Jw) = -S(v, Jw) = S(Jw, v), \quad v, w \in V.$$

This shows that $(v, w) \to S(Jv, w)$ is a symmetric bilinear form on V. Thus the function

$$H(v, w) = S(Jv, w) + iS(v, w) \quad (9.5)$$

is a Hermitian form on (V, J). Indeed,

$$H(w, v) = S(Jw, v) + iS(w, v) = S(J^2w, Jv) - iS(v, w) = S(Jv, w) - iS(v, w) = \overline{H(v, w)},$$

$$H(iv, w) = H(Jv, w) = S(J^2v, w) + iS(Jv, w) = -S(v, w) + iS(Jv, w) =$$
$$= i(S(Jv, w) + iS(v, w)) = iH(v, w).$$

By property (ii) of S, this form is positive definite.

Conversely, given a complex structure on V and a positive definite Hermitian form H on V, we write

$$H(v, w) = Re(H(v, w)) + iIm(H(v, w))$$

and immediately verify that the function

$$S(v, w) = Im(H(v, w))$$

is a skew-symmetric real bilinear form on V, and

$$S(iv, w) = Re(H(v, w))$$

is a positive definite symmetric real bilinear form on V. The function $S(v, w)$ satisfies properties (i) and (ii).

We know from (9.4) that A and $\bar{A}$ are isotropic subspaces with respect to $S : V_{\mathbb{C}} \to \mathbb{C}$. For any $a = v - iJv, b = w - iJw \in A$, we have

$$S(a, \bar{b}) = S(v - iJv, w + iJw) = S(v, w) + S(Jv, Jw) - i(S(Jv, w) - S(v, Jw))$$

$$= 2S(v, w) - i(S(Jv, w) + S(Jw, v)) = 2S(v, w) - 2iS(Jv, w) = -2iH(v, w).$$

We set for any $a = v - iJv, b = w - iJw \in A$

$$\langle a, \bar{b}\rangle = 4H(v, w) = 2iS(a, \bar{b}). \quad (9.6)$$

9.2 The standard representation of $\tilde{V}$ associated to J is on the Hilbert space $\tilde{S}(A)$ obtained by completing the symmetric algebra $S(A)$ with respect to the inner product on $S(A)$ defined by

$$(a_1a_2 \cdots a_n | b_1b_2 \cdots b_n) = \sum_{\sigma \in \Sigma_n} \langle a_1, \bar{b}_{\sigma(1)}\rangle \cdots \langle a_n, \bar{b}_{\sigma(n)}\rangle. \quad (9.7)$$

We identify A and $\bar{A}$ with subgroups of $\tilde{V}$ by $a \to (a,1)$ and $\bar{a} \to (\bar{a},1)$. Consider the space $Hol(\bar{A})$ of holomorphic functions on $\bar{A}$. The algebra $S(A)$ can be identified with the subalgebra of polynomial functions in $Hol(\bar{A})$ by considering each a as the linear function $z \to \langle a,z\rangle$ on $\bar{A}$. Let us make $\bar{A}$ act on $Hol(\bar{A})$ by translations

$$\bar{a}\cdot f(z) = f(z-\bar{a}), \quad z \in \bar{A},$$

and A by multiplication

$$a\cdot f(z) = e^{\langle a,z\rangle} f(z) = e^{2iS(a,z)} f(z).$$

Since

$$[a,\bar{b}] = [(a,1),(\bar{b},1)] = (0, e^{2iS(a,\bar{b})}) = (0, e^{\langle a,\bar{b}\rangle}),$$

we should check that

$$a\circ \bar{b} = e^{\langle a,\bar{b}\rangle}\bar{b}\circ a.$$

We have

$$\bar{b}\circ a\cdot f(z) = e^{\langle a,z-\bar{b}\rangle} f(z-\bar{b}) = e^{-\langle a,\bar{b}\rangle} e^{\langle a,z\rangle} f(z-\bar{b}),$$

$$a\circ \bar{b}\cdot f(z) = e^{\langle a,z\rangle} f(z-\bar{b}) = e^{\langle a,\bar{b}\rangle}\bar{b}\circ a\cdot f(z),$$

and the assertion is verified.

This defines a representation of the group $\tilde{V}_{\mathbb{C}}$ on $Hol(\bar{A})$. We get representation of $\tilde{V}$ on $Hol(\bar{A})$ by restriction. Write $v = a + \bar{a}, \quad a \in A$. Then

$$(v,1) = (a+\bar{a},1) = (a,1)\cdot(\bar{a},1)\cdot(0,e^{-iS(a,\bar{a})}). \tag{9.8}$$

This gives

$$v\cdot f(z) = e^{-iS(a,\bar{a})} a\cdot(\bar{a}\cdot f(z)) = e^{-\frac{1}{2}\langle a,\bar{a}\rangle} e^{\langle a,z\rangle} f(z-\bar{a}). \tag{9.9}$$

The space $\tilde{S}(A)$ can be described by the following:

Lemma 1. *Let W be the subspace of $\mathbb{C}^A$ spanned by the characteristic functions χ_a of $\{a\}, a \in A$, with the Hermitian inner product given by*

$$\langle \chi_a, \chi_b\rangle = e^{\langle a,b\rangle},$$

where χ_a is the characteristic function of $\{a\}$. Then this Hermitian product is positive definite, and the completion of W with respect to the corresponding norm is $\tilde{S}(A)$.

Proof. To each $\xi \in A$ we assign the element $e^a = 1 + a + \frac{a^2}{2} + \ldots$ from $\tilde{S}(A)$. Now we define the map by $\chi_a \to e^a$. Since

$$\langle e^a | e^b\rangle = \langle \sum_{n=1}^{\infty}\frac{a^n}{n!}\Big|\sum_{n=1}^{\infty}\frac{b^n}{n!}\rangle = \sum_{n=1}^{\infty}(\frac{1}{n!})^2\langle a^n|b^n\rangle$$

$$= \sum_{n=1}^{\infty} (\frac{1}{n!})^2 n! \langle a, \bar{b} \rangle^n = \sum_{n=1}^{\infty} \frac{1}{n!} \langle a, \bar{b} \rangle^n = e^{\langle a, \bar{b} \rangle}, \tag{9.10}$$

we see that the map $W \to \tilde{S}(A)$ preserves the inner product. We know from (9.7) that the inner product is positive definite. It remains to show that the functions e^a span a dense subspace of $\tilde{S}(A)$. Let F be the closure of the space they span. Then, by differentiating e^{ta} at $t = 0$, we obtain that all a^n belong to F. By the theorem on symmetric functions, every product $a_1 \cdots a_n$ belongs to F. Thus F contains $\tilde{S}(A)$ and thus must coincide with $\tilde{S}(A)$.

We shall consider e^a as a holomorphic function $z \to e^{\langle a, z \rangle}$ on $\bar{A}$.

Let us see how the group $\tilde{V}$ acts on the basis vectors e^a. Its center $\mathbb{C}_1^*$ acts by

$$(0, \lambda) \cdot e^a = e^{\lambda a}. \tag{9.11}$$

We have

$$\bar{b} \cdot e^a = \bar{b} \cdot e^{\langle a, z \rangle} = e^{\langle a, z - \bar{b} \rangle} = e^{-\langle a, \bar{b} \rangle} e^{\langle a, z \rangle} = e^{-\langle a, \bar{b} \rangle} e^a,$$

$$b \cdot e^a = b \cdot e^{\langle a, z \rangle} = e^{\langle b, z \rangle} e^{\langle a, z \rangle} = e^{a+b}.$$

Let $v = \bar{b} + b \in V$, then

$$(v, 1) = (b + \bar{b}, 1) = (b, 1) \cdot (\bar{b}, 1) \cdot (0, e^{-iS(b, \bar{b})}).$$

Hence

$$v \cdot e^a = e^{-iS(b, \bar{b})} b \cdot (\bar{b} \cdot e^a) = e^{-\frac{1}{2}\langle b, \bar{b} \rangle} e^{-\langle a, \bar{b} \rangle} e^{a+b} = e^{-\frac{1}{2}\langle b, \bar{b} \rangle - \langle a, \bar{b} \rangle} e^{a+b}. \tag{9.12}$$

In particular, using (9.10), we obtain

$$||v \cdot e^a||^2 = \langle e^a | e^a \rangle = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} ||e^{a+b}||^2 = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} \langle e^{a+b} | e^{a+b} \rangle =$$

$$= e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} e^{\langle a+b, \bar{a} + \bar{b} \rangle} = e^{-\langle b, \bar{b} \rangle - 2\langle a, \bar{b} \rangle} e^{\langle a, \bar{a} \rangle + 2\langle a, \bar{b} \rangle + \langle b, \bar{b} \rangle} = e^{\langle a, \bar{a} \rangle} = ||e^a||^2.$$

This shows that the representation of $\tilde{V}$ on $\tilde{S}(A)$ is unitary. This representation is called the *Schrödinger representation* of $\tilde{V}$.

Theorem 1. *The Schrödinger representation of the Heisenberg group $\tilde{V}$ is irreducible.*

Proof. The group $\mathbb{C}^*$ acts on $\tilde{S}(A)$ by scalar multiplication on the arguments and, in this way defines a grading:

$$\tilde{S}(A)_k = \{f(z) : f(\lambda \cdot z) = \lambda^k f(z)\}, \quad k \geq 0.$$

The induced grading on $S(A)$ is the usual grading of the symmetric algebra. Let W be an invariant subspace in $\tilde{S}(A)$. For any $w \in W$ we can write

$$w(z) = \sum_{k=0}^{\infty} w_k(z), \quad w_k(z) \in S(A)_k. \tag{9.13}$$

Replacing $w(z)$ with $w(\lambda z)$, we get

$$w(\lambda z) = \sum_{k=0}^{\infty} w_k(\lambda z) = \sum_{k=0}^{\infty} \lambda^k w_k(z) \in W.$$

By continuity, we may assume that this is true for all $\lambda \in \mathbb{C}$. Differentiating in t, we get

$$\frac{dw(\lambda z)}{d\lambda}(0) = \lim_{t\to 0} \frac{1}{t}[w((\lambda+t)z) - w(\lambda z)] = w_1(z) \in W.$$

Continuing in this way, we obtain that all w_k belong to W. Now let $\tilde{S}(A) = W \perp W'$, where W' is the orthogonal complement of W. Then W' is also invariant. Let $1 = w + w'$, where $w \in W$, $w' \in W'$. Then writing w and w' in the form (9.13), we obtain that either $1 \in W$, or $1 \in W'$. On the other hand, formula (9.12) shows that all basis functions e^a can be obtained from 1 by operators from $\tilde{V}$. This implies $W = \tilde{S}(A)$ or $W' = \tilde{S}(A)$. This proves that $W = \{0\}$ or $W = \tilde{S}(A)$.

9.3 From now on we shall assume that V is of dimension $2n$. The corresponding complex space (V, J) is of course of dimension n. Pick some basis $e_1, \ldots, e_n$ of the complex vector space (V, J) such that $e_1, Je_1, \ldots, e_n, Je_n$ is a basis of the real vector space V. Since $e_i + iJe_i = i(Je_i + iJJe_i)$ we see that the vectors $e_i + iJe_i, i = 1, \ldots, n$, form a basis of the linear subspace $\bar{A}$ of $V_{\mathbb{C}}$. Let us identify $\bar{A}$ with $\mathbb{C}^n$ by means of this basis. Thus we shall use $\mathbf{z} = (z_1, \ldots, z_n)$ to denote both a general point of this space as well as the set of coordinate holomorphic functions on $\mathbb{C}^n$. We can write

$$\mathbf{z} = \mathbf{x} + i\mathbf{y} = (x_1, \ldots, x_n) + i(y_1, \ldots, y_n)$$

for some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We denote by $\mathbf{z}\cdot\mathbf{z}'$ the usual dot product in $\mathbb{C}^n$. It defines the standard unitary inner product $\mathbf{z} \cdot \bar{\mathbf{z}}'$. Let

$$||\mathbf{z}||^2 = |z_1|^2 + \ldots |z_n|^2 = (x_1^2 + y_1^2) + \ldots + (x_n^2 + y_n^2) = ||\mathbf{x}||^2 + |\mathbf{y}||^2$$

be the corresponding norm squared.

Let us define a unitary isomorphism between the space $L^2(\mathbb{R}^n)$ and $\tilde{S}(A)$. First let us identify the space $\tilde{S}(A)$ with the space of holomorphic functions on on $\bar{A}$ wich are square-integrable with respect to some gaussian measure on $\bar{A}$. More precisely, consider the measure on $\mathbb{C}^n$ defined by

$$\int_{\mathbb{C}^n} f(\mathbf{z})d\mu = \int_{\mathbb{C}^n} f(\mathbf{z})e^{-\pi||\mathbf{z}||^2}d\mathbf{z} := \int_{\mathbb{R}^{2n}} e^{-\pi(||\mathbf{x}||^2+||\mathbf{y}||^2)}d\mathbf{x}d\mathbf{y}. \tag{9.14}$$

Notice that the factor π is chosen in order that

$$\int_{\mathbb{C}^n} e^{-\pi||\mathbf{z}||^2}d\mathbf{z} = \Big(\int_{\mathbb{R}} e^{-\pi(x^2+y^2)}dxdy\Big)^n = \Big(\int_0^{2\pi}\int_{\mathbb{R}} e^{-r^2}rdrd\theta\Big)^n = \Big(\int_{\mathbb{R}} e^{-t}dt\Big)^n = 1.$$

Thus our measure μ is a probability measure (called the *Gaussian measure* on $\mathbb{C}^n$).

Let

$$H_n = \{f(\mathbf{z}) \in \mathrm{Hol}(\mathbb{C}^n) : \int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 e^{-\pi||\mathbf{z}||^2} d\mathbf{z} \quad \text{converges}\}.$$

Let

$$f(\mathbf{z}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{z}^{\mathbf{i}} = \sum_{i_1,\ldots,i_n=0} a_{i_1,\ldots,i_n} z_1^{i_1} \cdots z_n^{i_n}$$

be the Taylor expansion of $f(\mathbf{z})$. We have

$$\int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 d\mu = \lim_{r\to\infty} \int\limits_{\max\{|z_i|\}\le r} |f(\mathbf{z})|^2 d\mu =$$

$$= \sum_{\mathbf{i},\mathbf{j}=1}^{\infty} a_{\mathbf{i}} \bar{a}_{\mathbf{j}} \lim_{r\to\infty} \int\limits_{\max\{|z_i|\}\le r} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} d\mu = \sum_{\mathbf{i},\mathbf{j}} a_{\mathbf{i}} \bar{a}_{\mathbf{j}} \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{j}} d\mu$$

Now, one can show that

$$\int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} d\mu = (i)^n \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} e^{-\pi \mathbf{z}\bar{\mathbf{z}}} d\mathbf{z} d\bar{\mathbf{z}} = (i)^n \int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{j}} e^{-\pi \mathbf{z}\bar{\mathbf{z}}} d\mathbf{z} d\bar{\mathbf{z}} = 0 \quad \text{if} \quad \mathbf{i} \neq \mathbf{j},$$

and

$$\int\limits_{\mathbb{C}^n} \mathbf{z}^{\mathbf{i}} \bar{\mathbf{z}}^{\mathbf{i}} d\mu = \prod_{k=1}^{n} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} (x_k^2 + y_k^2)^{i_k} e^{-\pi(x_k^2+y_k^2)} dx dy = \prod_{k=1}^{n} \int\limits_{0}^{2\pi} \int\limits_{\mathbb{R}} r^{2i_k} e^{-\pi r^2} r dr d\theta =$$

$$= \prod_{k=1}^{n} \pi \int\limits_{\mathbb{R}} t^{i_k} e^{-\pi t} dt = \prod_{k=1}^{n} i_k!/\pi^{i_k} = \mathbf{i}! \pi^{-|\mathbf{i}|}.$$

Here the absolute value denotes the sum $i_1 + \ldots + i_n$ and $\mathbf{i}! = i_1! \cdots i_n!$. Therefore, we get

$$\int\limits_{\mathbb{C}^n} |f(\mathbf{z})|^2 d\mu = \sum_{\mathbf{i}} |a_{\mathbf{i}}|^2 \pi^{-|\mathbf{i}|} \mathbf{i}!.$$

From this it follows easily that

$$\int\limits_{\mathbb{C}^n} f(\mathbf{z}) \bar{\phi}(\mathbf{z}) d\mu = \sum_{\mathbf{i}} a_{\mathbf{i}} \bar{b}_{\mathbf{i}} \pi^{-|\mathbf{i}|} \mathbf{i}!,$$

where $\bar{b}_{\mathbf{i}}$ are the Taylor coefficients of $\phi(\mathbf{z})$. Thus, if we use the left-hand-side for the definition of the inner product in H_n, we get an orthonormal basis formed by the functions

$$\phi_{\mathbf{i}}(\mathbf{z}) = \left(\frac{\pi^{|\mathbf{i}|}}{\mathbf{i}!}\right)^{\frac{1}{2}} \mathbf{z}^{\mathbf{i}}. \tag{9.15}$$

Lemma 2. *H_n is a Hilbert space and the ordered set of functions $\psi_{\mathbf{i}}$ is an orthonormal basis.*

Proof. By using Taylor expansion we can write each $\psi(\mathbf{z}) \in H_n$ as an infinite series

$$\psi(\mathbf{z}) = \sum_{\mathbf{i}} c_{\mathbf{i}} \phi_{\mathbf{i}}. \tag{9.16}$$

It converges absolutely and uniformly on any bounded subset of $\mathbb{C}^n$. Conversely, if ψ is equal to such an infinite series which converges with respect to the norm in H_n, then $c_{\mathbf{i}} = (\psi, \phi_{\mathbf{i}})$ and

$$||\psi||^2 = \sum_{\mathbf{i}} |c_p|^2 < \infty.$$

By the Cauchy-Schwarz inequality,

$$\sum_{\mathbf{i}} |c_{\mathbf{i}}||\phi_{\mathbf{i}}| \leq (\sum_{\mathbf{i}} |c_{\mathbf{i}}|^2)^{\frac{1}{2}} e^{\pi ||\mathbf{z}||^2/2}.$$

This shows that the series (9.16) converges absolutely and uniformly on every bounded subset of $\mathbb{C}^n$. By a well-known theorem from complex analysis, the limit is an analytic function on $\mathbb{C}^n$. This proves the completeness of the basis $(\psi_{\mathbf{i}})$.

Let us look at our functions e^a from $\tilde{S}(\bar{A})$. Choose coordinates $\zeta = (\zeta_1, \ldots, \zeta_n)$ in A such that, for any $\zeta = (\zeta_1, \ldots, \zeta_n) \in A, z = (z_1, \ldots, z_n) \in \bar{A}$,

$$\langle \zeta, z \rangle = \pi(\bar{\zeta}_1 z_1 + \ldots + \bar{\zeta}_n z_n) = \pi \bar{\zeta} \cdot \mathbf{z}.$$

We have

$$e^{\zeta} = e^{\pi(\zeta_1 z_1 + \ldots \zeta_n z_n)} = e^{\pi \zeta \cdot \mathbf{z}} = \sum_{n=1}^{\infty} \frac{\pi^n (\zeta \cdot \mathbf{z})^n}{n!} = \sum_{n=1}^{\infty} \frac{\pi^n}{n!} \Big(\sum_{|\mathbf{i}|=n} \frac{n!}{\mathbf{i}!} \zeta^{\mathbf{i}} \mathbf{z}^{\mathbf{i}} \Big) = \sum_{\mathbf{i}} \phi_{\mathbf{i}}(\zeta) \phi_{\mathbf{i}}(\mathbf{z}).$$

Comparing the norms, we get

$$(e^{\pi \zeta \cdot \mathbf{z}}, e^{\pi \zeta \cdot \mathbf{z}})_{H_n} = \sum_{\mathbf{i}} |\phi_{\mathbf{i}}(\zeta)|^2 = \sum_{\mathbf{i}} \frac{\pi^{|\mathbf{i}|}}{\mathbf{i}!} |\zeta_1|^{2i_1} \ldots |\zeta_n|^{2i_n} =$$

$$= \sum_{n=1}^{\infty} \frac{\pi^n ||\zeta||^{2n}}{n!} = e^{\pi \bar{\zeta} \cdot \zeta} = e^{\langle \bar{\zeta}, \zeta \rangle} = \langle e^{\zeta} | e^{\zeta} \rangle_{\tilde{S}(\bar{A})}.$$

So our space $\tilde{S}(A)$ is maped isometrically into H_n. Since its image contains the basis functions (9.16) and $\tilde{S}(A)$ is complete, the isometry is bijective.

9.4 Let us find now an isomorphism between H_n and $L^2(\mathbb{R}^n)$. For any $\mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{C}^n$, let

$$k(\mathbf{x}, \mathbf{z}) = 2^{\frac{n}{4}} e^{-\pi ||\mathbf{x}||^2} e^{2\pi i \mathbf{x} \cdot \mathbf{z}} e^{\frac{\pi}{2} ||\mathbf{z}||^2}. \tag{9.17}$$

Lemma 3.

$$\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\zeta)}d\mathbf{x} = e^{\pi\bar{\zeta}\cdot\mathbf{z}}.$$

Proof. We have

$$\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\zeta)}d\mathbf{x} = 2^{\frac{n}{2}} e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)} \int_{\mathbb{R}^n} e^{-2\pi||\mathbf{x}||^2} e^{-2\pi i\mathbf{x}\cdot(\zeta-\mathbf{z})}d\mathbf{x} =$$

$$= e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)} \frac{1}{\sqrt{(2\pi^n}} \int_{\mathbb{R}^n} e^{-||\mathbf{y}||^2/2} e^{-i\sqrt{\pi}\mathbf{y}\cdot(\zeta-\mathbf{z})}d\mathbf{y} = e^{\frac{\pi}{2}(||\mathbf{z}||^2+||\zeta||^2)} F(\sqrt{\pi}(\mathbf{z}-\zeta)).$$

where $F(\mathbf{t})$ is the Fourier transform of the function $e^{-||\mathbf{y}||^2/2}$. It is known that $\widehat{e^{-x^2/2}} = e^{-t^2/2}$. This easily implies that $F(\mathbf{t}) = e^{-||\mathbf{t}||^2/2}$. Plugging in $\mathbf{t} = \sqrt{\pi}(\mathbf{z}-\zeta)$, we get the assertion.

Lemma 4. *Let*

$$k(\mathbf{x},\mathbf{z}) = \sum_{\mathbf{i}} h_{\mathbf{i}}(\mathbf{x})\phi_{\mathbf{i}}(\mathbf{z})$$

be the Taylor expansion of $k(\mathbf{x},\mathbf{z})$ (recall that $\phi_{\mathbf{i}}$ are monomials in $\mathbf{z}$). Then $(h_{\mathbf{i}}(\mathbf{x}))_{\mathbf{i}}$ forms a complete orthonormal system in $L^2(\mathbb{R}^n)$. In fact,

$$h_{\mathbf{i}}(x_1,\ldots,x_n) = h_{i_1}(x_1)\ldots h_{i_n}(x_n),$$

where $h_k(x_i) = H_k(\sqrt{2\pi}x_i)$ with $H_k(x)$ being a Hermite function.

Proof. We will check this in the case $n=1$. Since both $k(\mathbf{x},\mathbf{z})$ and $\phi_{\mathbf{i}}$ are products of functions in one variable, the general case is easily reduced to our case. Since $k(x,z) = 2^{\frac{1}{4}}e^{\pi x^2}e^{-2\pi(x-iz/2)^2}$, we get

$$h_i(x) = \int_{\mathbb{R}} k(x,z)\phi_i(x)dx = 2^{\frac{1}{4}}(\pi/i!)^{1/2} \int_{\mathbb{R}} e^{\pi x^2}e^{-2\pi(x-iz/2)^2}x^i dx.$$

We omit the computation of this integral which lead to the desired answer.

Now we can define the linear map

$$\Phi : L^2(\mathbb{R}^n) \to \mathrm{Hol}(\mathbb{C}^n), \quad \psi(\mathbf{x}) \to \int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\psi(\mathbf{x})d\mathbf{x}. \tag{9.18}$$

By Lemma 3 and the Cauchy-Schwarz inequality,

$$|\Phi(\psi)(\mathbf{z})| \le \Big(\int_{\mathbb{R}^n} k(\mathbf{x},\mathbf{z})\overline{k(\mathbf{x},\mathbf{z})}d\mathbf{x}\Big)||\psi(\mathbf{x})|| = e^{(\pi||\mathbf{z}||^2/2)}||\psi(\mathbf{x})||.$$

This implies that the integral is uniformly convergent with respect to the complex parameter $\mathbf{z}$ on every bounded subset of $\mathbb{C}^n$. Thus $\Phi(\psi)$ is a holomorphic function; so the map is well-defined. Also, by Lemma 4, since $\phi_{\mathbf{i}}$ is an orthonormal system in $\mathrm{Hol}(\mathbb{C}^n)$ and $h_{\mathbf{i}}$ is an orthonormal basis in $L^2(\mathbb{R}^n)$, we get

$$\int_{\mathbb{C}^n} |\Phi(\psi)(\mathbf{z})|^2 d\mu = \sum_{\mathbf{i}} |(h_{\mathbf{i}}, \bar{\psi})|^2 = ||\psi||^2 < \infty.$$

This shows that the image of Φ is contained in H_n, and at the same time, that Φ is a unitary linear map. Under the map Φ, the basis $(h_{\mathbf{i}}(\mathbf{x}))_{\mathbf{i}}$ of $L^2(\mathbb{R}^n)$ is mapped to the basis $(\phi_{\mathbf{i}})_{\mathbf{i}}$ of H_n. Thus Φ is an isomorphism of Hilbert spaces.

9.5 We know how the Heisenberg group $\tilde{V}$ acts on H_n. Let us see how it acts on $L^2(\mathbb{R}^n)$ via the isomorphism $\Phi : L^2(\mathbb{R}^n) \cong H_n$.

Recall that we have a decomposition $V_{\mathbb{C}} = A \oplus \bar{A}$ of $V_{\mathbb{C}}$ into the sum of conjugate isotropic subspaces with respect to the bilinear form S. Consider the map $V \to A, v \to v - iJv$. Since $Jv - iJ(Jv) = i(v - iJv)$, this map is an isomorphism of complex vector spaces $(V, J) \to A$. Similarly we see that the map $v \to v + iJv$ is an isomorphism of complex vector spaces $(V, -J) \to \bar{A}$. Keep the basis $e_1, \ldots, e_n$ of (V, J) as in 9.3 so that $V_{\mathbb{C}}$ is identified with $\mathbb{C}^n$. Then $e_i - iJe_i, i = 1, \ldots, n$ is a basis of A, $e_i + iJe_i, i = 1, \ldots, n$, is a basis of $\bar{A}$, and the pairing $A \times \bar{A} \to \mathbb{C}$ from (9.6) has the form

$$\langle (\zeta_1, \ldots, \zeta_n), (z_1, \ldots, z_n) \rangle = \pi(\sum_{i=1}^n \zeta_i z_i) = \pi\zeta \cdot \mathbf{z}.$$

Let us identify (V, J) with $\mathbb{C}^n$ by means of the basis (e_i). And similarly let us do it for A and $\bar{A}$ by means of the bases $(e_i - iJe_i)$ and $(e_i + iJe_i)$, respectively. Then $V_{\mathbb{C}}$ is identified with $A \oplus \bar{A} = \mathbb{C}^n \oplus \mathbb{C}^n$, and the inclusion $(V, J) \subset V_{\mathbb{C}}$ is given by $\mathbf{z} \to (\mathbf{z}, \bar{\mathbf{z}})$.

The skew-symmetric bilinear form $S : V_{\mathbb{C}} \times V_{\mathbb{C}} \to \mathbb{C}$ is now given by the formula

$$S((\mathbf{z}, \mathbf{w}), (\mathbf{z}', \mathbf{w}')) = \frac{\pi}{2i}(\mathbf{z} \cdot \mathbf{w}' - \mathbf{w} \cdot \mathbf{z}').$$

Its restriction to V is given by

$$S(\mathbf{z}, \mathbf{z}') = S((\mathbf{z}, \bar{\mathbf{z}}), (\mathbf{z}', \bar{\mathbf{z}}')) = \frac{\pi}{2i}(\mathbf{z} \cdot \bar{\mathbf{z}}' - \bar{\mathbf{z}} \cdot \mathbf{z}') = \frac{\pi}{2i}(2i\mathrm{Im})(\mathbf{z} \cdot \bar{\mathbf{z}}') = \pi\mathrm{Im}(\mathbf{z} \cdot \bar{\mathbf{z}}') = \pi(\mathrm{yx}' - \mathrm{xy}'),$$

where $\mathbf{z} = \mathbf{x} + i\mathbf{y}, \mathbf{z}' = \mathbf{x}' - i\mathbf{y}'$. Let

$$V_{re} = \{\mathbf{z} \in V : \mathrm{Im}(\mathbf{z}) = 0\} = \{\mathbf{z} = \mathbf{x} \in \mathbb{R}^n\},$$
$$V_{im} = \{\mathbf{z} \in V : \mathrm{Re}(\mathbf{z}) = 0\} = \{\mathbf{z} = i\mathbf{y} \in i\mathbb{R}^n\}.$$

Then the decomposition $V = V_{re} \oplus V_{im}$ is a decomposition of V into the sum of two maximal isotropic subspaces with respect to the bilinear form S.

As in (9.8), we have, for any $v = \mathbf{x} + i\mathbf{y} \in V$,

$$v = \mathbf{x} + i\mathbf{y} = a + \bar{a} = (\mathbf{x} + i\mathbf{y}, 0) + (0, \mathbf{x} - i\mathbf{y}) \in A \oplus \bar{A}.$$

The Heisenberg group V acts on $\mathrm{Hol}(\bar{A})$ by the formulae

$$\mathbf{x} + i\mathbf{y} \cdot f(\mathbf{z}) = e^{-S(a,\bar{a})} a \cdot \bar{a} f(\mathbf{z}) = e^{-S(a,\bar{a})} e^{\langle \bar{a}, \mathbf{z} \rangle} f(\mathbf{z} - \bar{a}) =$$
$$= e^{-\frac{\pi}{2} a \cdot \bar{a}} e^{\pi a \cdot \mathbf{z}} f(z - \bar{a}) = e^{-\frac{\pi}{2}(\mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y})} e^{\pi(\mathbf{x} + i\mathbf{y}) \cdot \mathbf{z}} f(\mathbf{z} - \mathbf{x} + i\mathbf{y}). \tag{9.19}$$

Theorem 2. *Under the isomorphism $\Phi : H_n \to L^2(\mathbb{R}^n)$, the Schrödinger representation of $\tilde{V}$ on H_n is isomorphic to the representation of $\tilde{V}$ on $L^2(\mathbb{R}^n)$ defined by the formula*

$$(\mathbf{v} + i\mathbf{u}, t)\psi(\mathbf{x}) = te^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x} - \mathbf{u}). \tag{9.20}$$

Proof. In view of (9.9) and (9.10), we have

$$(\mathbf{v} + i\mathbf{u}) \cdot \Phi(\psi(\mathbf{x})) = \int_{\mathbb{R}^n} (\mathbf{v} + i\mathbf{u})k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})}e^{\pi(\mathbf{v}+i\mathbf{u})\cdot\mathbf{z}}k(\mathbf{z} - \mathbf{v} + i\mathbf{u})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})}e^{\pi(\mathbf{v}+i\mathbf{u})\cdot\mathbf{z}}e^{2\pi i\mathbf{x}(-\mathbf{v}+i\mathbf{u})}e^{\frac{\pi}{2}(2\mathbf{z}(-\mathbf{v}+i\mathbf{u})+(-\mathbf{v}+i\mathbf{u})\cdot(-\mathbf{v}+i\mathbf{u}))}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\pi\mathbf{u}\cdot\mathbf{u}+2\pi i\mathbf{u}\cdot\mathbf{z}-2\pi i\mathbf{x}\mathbf{v}-2\pi\mathbf{x}\mathbf{u}-\pi i\mathbf{v}\cdot\mathbf{u}}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}.$$

$$\Phi((\mathbf{v} + i\mathbf{u}) \cdot \psi(\mathbf{x})) = \int_{\mathbb{R}^n} k(\mathbf{x}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x} - \mathbf{u})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} k(\mathbf{t} + \mathbf{u}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i(\mathbf{t}+\mathbf{u})\cdot\mathbf{v}}\psi(\mathbf{t})d\mathbf{t} = \int_{\mathbb{R}^n} k(\mathbf{x} + \mathbf{u}, \mathbf{z})e^{\pi i\mathbf{v}\cdot\mathbf{u}}e^{-2\pi i(\mathbf{x}+\mathbf{u})\cdot\mathbf{v}}\psi(\mathbf{x})d\mathbf{x}$$

$$= \int_{\mathbb{R}^n} e^{-\pi\mathbf{u}\cdot\mathbf{u}-2\pi\mathbf{u}\mathbf{x}+2\pi i\mathbf{u}\cdot\mathbf{z}-\pi i\mathbf{v}\cdot\mathbf{u}-2\pi i\mathbf{x}\mathbf{v}}k(\mathbf{x}, \mathbf{z})\psi(\mathbf{x})d\mathbf{x}.$$

By comparing, we observe that

$$(\mathbf{v} + i\mathbf{u}) \cdot \Phi(\psi(\mathbf{x})) = \Phi((\mathbf{v} + i\mathbf{u}) \cdot \psi(\mathbf{x})).$$

This checks the assertion.

9.6 It follows from the proof that the formula (9.20) defines a representation of the group $\tilde{V}$ on $L^2(\mathbb{R}^n)$. Let us check this directly. We have

$$(\mathbf{v} + i\mathbf{u}, t)(\mathbf{v}' + \mathbf{u}', t') \cdot \psi(\mathbf{x}) = ((\mathbf{v} + \mathbf{v}') + i(\mathbf{u} + \mathbf{u}'), tt'e^{i\pi(\mathbf{u}\cdot\mathbf{v}'-\mathbf{v}\cdot\mathbf{u}')} \cdot \psi(\mathbf{x}) =$$

$$= tt'e^{i\pi(\mathbf{u}\cdot\mathbf{v}'-\mathbf{v}\cdot\mathbf{u}')}e^{\pi i(\mathbf{v}+\mathbf{v}')\cdot(\mathbf{u}+\mathbf{u}')}e^{-2\pi i\mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')}\psi(\mathbf{x} - \mathbf{u} - \mathbf{u}') =$$

$$= tt'e^{i\pi(\mathbf{v}\cdot\mathbf{u}+\mathbf{v}'\cdot\mathbf{u}'+2\mathbf{v}'\cdot\mathbf{u})}e^{-2\pi i\mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')}\psi(\mathbf{x} - \mathbf{u} - \mathbf{u}').$$

$$(\mathbf{v} + i\mathbf{u}, t) \cdot ((\mathbf{v}' + \mathbf{u}', t') \cdot \psi(\mathbf{x})) = (\mathbf{v} + i\mathbf{u}, t) \cdot (t'e^{i\pi\mathbf{u}'\mathbf{v}'}e^{-2\pi i\mathbf{x}\cdot\mathbf{v}'}\psi(\mathbf{x} - \mathbf{u}')) =$$

$$= tt' e^{i\pi \mathbf{u}\mathbf{v}} e^{-2\pi i \mathbf{x}\cdot\mathbf{v}} e^{i\pi \mathbf{u}'\mathbf{v}'} e^{-2\pi i(\mathbf{x}-\mathbf{u})\cdot\mathbf{v}'} \psi(\mathbf{x}-\mathbf{u}-\mathbf{u}') =$$

$$= tt' e^{i\pi(\mathbf{v}\cdot\mathbf{u}+\mathbf{v}'\cdot\mathbf{u}'+2\mathbf{v}'\cdot\mathbf{u})} e^{-2\pi i \mathbf{x}\cdot(\mathbf{v}+\mathbf{v}')} \psi(\mathbf{x}-\mathbf{u}-\mathbf{u}').$$

So this matches.

Let us alter the definition of the Heisenberg group $\tilde{V}$ by setting

$$(\mathbf{z}, t)\cdot(\mathbf{z}', t') = (\mathbf{z}+\mathbf{z}', tt' e^{2\pi i \mathbf{x}\cdot\mathbf{y}'}),$$

where $\mathbf{z} = \mathbf{x} + i\mathbf{y}, \mathbf{z}' = \mathbf{x}' + i\mathbf{y}'$. It is immediately checked that the map $(\mathbf{z}, t) \to (\mathbf{z}, te^{\pi \mathbf{x}\cdot\mathbf{y}})$ is an isomorphism from our old group to the new one. Then the new group acts on $L^2(\mathbb{R}^n)$ by the formula

$$(\mathbf{v} + i\mathbf{u}, t)\cdot\psi(\mathbf{x}) = te^{-2\pi i \mathbf{x}\cdot\mathbf{v}}\psi(\mathbf{x}-\mathbf{u}), \tag{9.21}$$

and on $\mathrm{Hol}(\bar{A})$ by the formula

$$(\mathbf{v} + i\mathbf{u}, t)\cdot f(\mathbf{z}) = te^{-\frac{\pi}{2}(\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v})+\pi\mathbf{z}\cdot(\mathbf{v}+i\mathbf{u})+i\pi\mathbf{v}\cdot\mathbf{u}} f(\mathbf{z}-\mathbf{u}) =$$

$$= e^{\pi((\mathbf{z}-\frac{1}{2}\mathbf{v})\cdot\mathbf{v}-\frac{1}{2}\mathbf{u}\cdot\mathbf{u})+i\pi(\mathbf{z}+\mathbf{v})\cdot\mathbf{u}} f(\mathbf{z}-\mathbf{u}). \tag{9.22}$$

This agrees with the formulae from **[Igusa]**, p. 35.

9.7 Let us go back to Lecture 6, where we defined the operators $V(\mathbf{v})$ and $U(\mathbf{u})$ on the space $L^2(\mathbb{R}^n)$ by the formula:

$$U(\mathbf{u})\psi(\mathbf{q}) = e^{i(u_1P_1+\ldots+u_nP_n)}\psi(\mathbf{q}) = \psi(\mathbf{q}-\hbar\mathbf{u}),$$

$$V(\mathbf{v})\psi(\mathbf{q}) = e^{i(v_1Q_1+\ldots+v_nQ_n)}\psi(\mathbf{q}) = e^{i\mathbf{v}\cdot\mathbf{q}}\psi(\mathbf{q}).$$

Comparing this with the formula (9.19), we find that

$$U(\mathbf{u})\psi(\mathbf{q}) = i\hbar\mathbf{u}\cdot\psi(\mathbf{q}), \quad V(\mathbf{v})\psi(\mathbf{q}) = -\frac{1}{2\pi}\mathbf{v}\cdot\psi(\mathbf{q}),$$

where we use the Schrödinger representation of the (redefined) Heisenberg group $\tilde{V}$ in $L^2(\mathbb{R}^n)$. The commutator relation (6.18)

$$[U(\mathbf{u}), V(\mathbf{v})] = e^{-i\hbar\mathbf{u}\cdot\mathbf{v}}$$

agrees with the commutation relation

$$[i\hbar\mathbf{u}, -\frac{1}{2\pi}\mathbf{v}] = [(i\hbar\mathbf{u}, 1), (-\frac{1}{2\pi}\mathbf{v}, 1)] = (0, e^{2\pi i(\mathbf{u}\cdot-\frac{1}{2\pi}\mathbf{v})}) = (0, e^{-i\mathbf{u}\cdot\mathbf{v}}).$$

In Lecture 7 we defined the Heisenberg algebra $\mathcal{H}$. Its subalgebra $\mathcal{N}$ generated by the operators a and a^* coincides with the Lie subalgebra of self-adjoint (unbounded) operators in $L^2(\mathbb{R})$ generated by the operators $P = i\hbar\frac{d}{dq}$ and $Q = q$. If we exponentiate these operators we find the linear operators e^{ivQ}, e^{iuP}. It follows from the above that they form a group isomorphic to the Heinseberg group $\tilde{V}$, where $\dim V = 2$. The Schrödinger representation of this group in $L^2(\mathbb{R})$ is the exponentiation of the representation of $\mathcal{N}$ described in Lecture 7. Adding the exponent e^{iH} of the Schrödinger operator $H = \omega aa^* - \frac{\hbar\omega}{2}$ leads to an extension

$$1 \to \tilde{V} \to G \to \mathbb{R}^* \to 1.$$

Here the group G is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & x & y & z \\ 0 & w & 0 & y \\ 0 & 0 & w^{-1} & -x \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Of course, we can generalize this to the case of n variables. We consider similar matrices

$$\begin{pmatrix} 1 & \mathbf{x} & \mathbf{y} & z \\ 0 & wI_n & 0 & \mathbf{y}^t \\ 0 & 0 & w^{-1}I_n & -\mathbf{x}^t \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, z, w \in \mathbb{R}$. More generally, we can introduce the *full Heisenberg group* as the extension

$$1 \to \tilde{V} \to \tilde{G} \to \mathrm{Sp}(2n, \mathbb{R}) \to 1,$$

where $\mathrm{Sp}(2n, \mathbb{R})$ is the symplectic group of $2n \times 2n$ matrices X satisfying

$$X \begin{pmatrix} 0_n & -I_n \\ I_n & 0_n \end{pmatrix} X^t = \begin{pmatrix} 0_n & -I_n \\ I_n & 0_n \end{pmatrix}.$$

The group $\tilde{G}$ is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & \mathbf{x} & \mathbf{y} & z \\ 0 & A & B & \mathbf{y}^t \\ 0 & C & D & -\mathbf{x}^t \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is an $n \times n$-block presentation of a matrix $X \in \mathrm{Sp}(2n, \mathbb{R})$.

Exercises.

1. Let V be a real vector space of dimension $2n$ equipped with a skew-symmetric bilinear form S. Show that any decomposition $V_{\mathbb{C}} = A \oplus A$ into a sum of conjugate isotropic subspaces with the property that $(a, b) \to 2iS(a, \bar{b})$ is a positive definite Hermitian form on A defines a complex structure J on V such that $S(Jv, Jw) = S(v, w), S(Jv, w) > 0$.

2. Extend the Schrödinger representation of $\tilde{V}$ on $\tilde{S}(\bar{A})$ to a projective representation of the full Heisenberg group in $\mathbb{P}(\tilde{S}(\bar{A}))$ preserving (up to a scalar factor) the inner product in $\tilde{S}(\bar{A})$. This is called the *metaplectic representation* of order n. What will be the corresponding representation in $\mathbb{P}(L^2(\mathbb{R}^n))$?

3. Consider the natural representation of $SO(2)$ in $L_2(\mathbb{R}^2)$ via action on the arguments. Desribe the representation of $SO(2)$ in H_2 obtained via the isomorphism $\Phi : H_2 \cong L_2(\mathbb{R}^2)$.

4. Consider the natural representation of $SU(2)$ in H_2 via action on the arguments. Describe the representation of $SU(2)$ in $L_2(\mathbb{R}^2)$ obtained via the isomorphism $\Phi : H_2 \cong L_2(\mathbb{R}^2)$.

Lecture 10. ABELIAN VARIETIES AND THETA FUNCTIONS

10.1 Let us keep the notation from Lecture 9. Let Λ be a lattice of rank $2n$ in V. This means that Λ is a subgroup of V spanned by $2n$ linearly independent vectors. It follows that Λ is a free abelian subgroup of V and thus it is isomorphic to $\mathbb{Z}^{2n}$. The orbit space V/Λ is a compact torus. It is diffeomorphic to the product of $2n$ circles $(\mathbb{R}/\mathbb{Z})^{2n}$. If we view V as a complex vector space (V,J), then T has a canonical complex structure such that the quotient map

$$\pi : V \to V/\Lambda = T$$

is a holomorphic map. To define this structure we choose an open cover $\{U_i\}_{i\in I}$ of V such that $U_i \cap U_i + \gamma = \emptyset$ for any $\gamma \in \Lambda$. Then the restriction of π to U_i is an isomorphism and the complex structure on $\pi(U_i)$ is induced by that of U_i.

The skew-symmetric bilinear form S on V defines a complex line bundle L on T which can be used to embed T into a projective space so that T becomes an algebraic variety. A compact complex torus which is embeddable into a projective space is called an *abelian variety.*

Let us describe the construction of L. Start with any holomorphic line bundle L over T. Its pre-image $\pi^*(L)$ under the map π is a holomorphic line bundle over the complex vector space V. It is known that all holomorphic vector bundles on V are (holomorphically) trivial. Let us choose an isomorphism $\phi : \pi^*(L) = L \times_T V \to V \times \mathbb{C}$. Then the group Λ acts naturally on $\pi^*(L)$ via its action on V. Under the isomorphism ϕ it acts on $V \times \mathbb{C}$ by a formula

$$\gamma \cdot (v,t) = (v+\gamma, \alpha_\gamma(v)t), \quad \gamma \in \Lambda, v \in V, t \in \mathbb{C}. \tag{10.1}$$

Here $\alpha_\gamma(v)$ is a non-zero constant depending on γ and v. Since the action is by holomorphic automorphisms, $\alpha_\gamma(v)$ depends holomorphically on v and can be viewed as a map

$$\alpha : \Lambda \to \mathcal{O}(V)^*, \quad \gamma \to \alpha_\gamma(v),$$

where $\mathcal{O}(V)$ denotes the ring of holomorphic functions on V and $\mathcal{O}(V)^*$ is its group of invertible elements. It follows from the definition of action that

$$\alpha_{\gamma+\gamma'}(v) = \alpha_\gamma(v+\gamma')\alpha_{\gamma'}(v). \tag{10.2}$$

Denote by $Z^1(\Lambda, \mathcal{O}(V)^*)$ the set of functions α as above satisfying (10.2). These functions are called *theta factors* associated to Λ. Obviously $Z^1(\Lambda, \mathcal{O}(V)^*)$ is a commutative group with respect to pointwise multiplication. For any $g(v) \in \mathcal{O}(V)^*$ the function

$$\alpha_\gamma(v) = g(v+\gamma)/g(v) \tag{10.3}$$

belongs to $Z^1(\Lambda, \mathcal{O}(V)^*)$. It is called the *trivial theta factor.* The set of such functions forms a subgroup $B^1(\Lambda, \mathcal{O}(V)^*)$ of $Z^1(\Lambda, \mathcal{O}(V)^*)$. The quotient group is denoted by $H^1(\Lambda, \mathcal{O}(V)^*)$. The reader familiar with the notion of group cohomology will recognize the latter group as the first cohomology group of the group Λ with coefficients in the abelian group $\mathcal{O}(V^*)$ on which Λ acts by translation in the argument.

The theta factor $\alpha \in Z^1(\Lambda, \mathcal{O}(V)^*)$ defined by the line bundle L depends on the choice of the trivialization ϕ. A different choice leads to replacing ϕ with $g \circ \phi$, where $g : V \times \mathbb{C} \to V \times \mathbb{C}$ is an automorphism of the trivial bundle defined by the formula $(v, t) \to (v, g(v)t)$ for some function $g \in \mathcal{O}(V)^*$. This changes α_γ to $\alpha_\gamma(v)g(v+\gamma)/g(v)$. Thus the coset of α in $H^1(\Lambda, \mathcal{O}(V)^*)$ does not depend on the trivialization ϕ. This defines a map from the set $\mathrm{Pic}(T)$ of isomorphism classes of holomorphic line bundles on T to the group $H^1(\Lambda, \mathcal{O}(V)^*)$. In fact, this map is a homomorphism of groups, where the operation of an abelian group on $\mathrm{Pic}(T)$ is defined by tensor multiplication and taking the dual bundle.

Theorem 1. *The homomorphism*

$$\mathrm{Pic}(T) \to H^1(\Lambda, \mathcal{O}(V)^*)$$

is an isomorphism of abelian groups.

Proof. It is enough to construct the inverse map $H^1(\Lambda, \mathcal{O}(V)^*) \to \mathrm{Pic}(T)$. We shall define it, and leave to the reader to verify that it is the inverse.

Given a representative α of a class from the right-hand side, we consider the action of Λ on $V \times \mathbb{C}$ given by formula (10.2). We set L to be the orbit space $V \times \mathbb{C}/\Lambda$. It comes with a canonical projection $p : L \to T$ defined by sending the orbit of (v, t) to the orbit of v. Its fibres are isomorphic to $\mathbb{C}$. Choose a cover $\{U_i\}_{i \in I}$ of V as in the beginning of the lecture and let $\{W_i\}_{i \in I}$ be its image cover of T. Since the projection π is a local analytic isomorphism, it is an open map (so that the image of an open set is open). Because T is compact, we may find a finite subcover of the cover $\{W_i\}_{i \in I}$. Thus we may assume that I is finite, and $\pi^{-1}(W_i) = \coprod_{\gamma \in \Lambda}(U_i + \gamma)$. Let $\pi_i : U_i \to W_i$ be the analytic isomorphism induced by the projection π. We may assume that $\pi_i^{-1}(W_i \cap W_j) = \pi_j^{-1}(W_i \cap W_j) + \gamma_{ij}$ for some $\gamma_{ij} \in \Lambda$, provided that $W_i \cap W_j \neq \emptyset$. Since each $U_i \times \mathbb{C}$ intersects every orbit of Λ in $V \times \Lambda$ at a unique point (v, t), we can identify $p^{-1}(W_i)$ with $W_i \times \mathbb{C}$. Also

$$W_i \times \mathbb{C} \supset p_i^{-1}(W_i \cap W_j) \cong p_j^{-1}(W_i \cap W_j) \subset W_j \times \mathbb{C},$$

where the isomorphism is given explicitly by $(v, t) \to (v, \alpha_{\gamma_{ij}}(v)t)$. This shows that L is a holomorphic line bundle with transition functions $g_{W_i, W_j}(x) = \alpha_{\gamma_{ij}}(\pi_i^{-1}(x))$. We also leave to the reader to verify that replacing α by another representative of the same class in H^1 changes the line bundle L to an isomorphic line bundle.

10.2 So, in order to construct a line bundle L we have to construct a theta factor. Recall from (9.11) that V acts on $\tilde{S}(A)$ by the formula

$$v \cdot e^a = e^{-\langle b,\bar{b}\rangle/2-\langle a,\bar{b}\rangle} e^{a+b},$$

where $v = b + \bar{b}$. Let us identify V with A by means of the isomorphism $v \to b = v - iJv$. From (9.6) we have

$$\langle v - iJv, w + iJw\rangle = 4H(v,w) = 4S(Jv,v) + 4iS(v,w).$$

Thus

$$e^{-\langle b,\bar{b}\rangle/2-\langle a,\bar{b}\rangle} = e^{-2H(\gamma,\gamma)-4H(v,\gamma)}.$$

Set

$$\alpha_\gamma(v) = e^{-2H(\gamma,\gamma)-4H(v,\gamma)}.$$

We have

$$\alpha_{\gamma+\gamma'}(v) = e^{-2H(\gamma+\gamma',\gamma+\gamma')-4H(v,\gamma+\gamma')} = e^{-2H(\gamma,\gamma)-2H(\gamma',\gamma')-4ReH(\gamma,\gamma')-4H(v,\gamma)-4H(v,\gamma')}$$

$$= e^{4iImH(\gamma,\gamma')-2H(\gamma,\gamma)-4H(v+\gamma',\gamma)-2H(\gamma',\gamma')-4H(v,\gamma')} = e^{-4iImH(\gamma,\gamma')}\alpha_\gamma(v+\gamma')\alpha_{\gamma'}(v).$$

We see that condition (10.2) is satisfied if, for any $\gamma,\gamma' \in \Lambda$,

$$ImH(\gamma,\gamma') = S(\gamma,\gamma') \in \frac{\pi}{2}\mathbb{Z}.$$

Let us redefine α by replacing it with

$$\alpha_\gamma(v) = e^{-\pi H(\gamma,\gamma)-2\pi H(v,\gamma)} \tag{10.4}$$

Of course this is equivalent to multiplying our bilinear form S by $-\frac{\pi}{2}$. Then the previous condition is replaced with

$$S(\Lambda \times \Lambda) \subset \mathbb{Z}. \tag{10.5}$$

Assuming that this is true we have a line bundle L_α. As we shall see in a moment it is not yet the final definition of the line bundle associated to S. We have to adjust the theta factor (10.4) a little more. To see why we should do this, let us compute the first Chern class of the obtained line bundle L_α.

Since V is obviously simply connected and $\pi : V \to T$ is a local isomorphism, we can idenify V with the universal cover of T and the group Λ with the fundamental group of T. Since it is abelian, we can also identify it with the first homology group $H_1(T,\mathbb{Z})$. Now, we have

$$H_k(T,\mathbb{Z}) \cong \bigwedge^k(H_1(T,\mathbb{Z})$$

because T is topologically the product of circles. In particular,

$$H^2(T,\mathbb{Z}) = Hom(H_2(T,\mathbb{Z}),\mathbb{Z}) = \bigwedge^2(H_1(T,\mathbb{Z}))^* = (\bigwedge^2\Lambda)^*. \tag{10.6}$$

This allows one to consider $S : \Lambda \times \Lambda \to \mathbb{Z}$ as an element of $H^2(T,\mathbb{Z})$. The latter group is where the first Chern class takes its value.

Recall that we have a canonical isomorphism between $\mathrm{Pic}(T)$ and $H^1(T,\mathcal{O}_T^*)$ by computing the latter groups as the Čech cohomology group and assigning to a line bundle L the set of the transition functions with respect to an open cover. The exponential sequence of sheaves of abelian groups

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_T \overset{e^{2\pi i}}{\longrightarrow} \mathcal{O}_T^* \longrightarrow 1 \tag{10.7}$$

defines the coboundary homomorphism

$$H^1(T,\mathcal{O}_T^*) \to H^2(T,\mathbb{Z}).$$

The image of $L \in \mathrm{Pic}(T)$ is the first Chern class $c_1(L)$ of L. In our situation, the coboundary homomorphism coincides with the coboundary homomorphism for the exact sequence of group cohomology

$$H^1(\Lambda,\mathcal{O}(V)) \to H^1(\Lambda,\mathcal{O}(V)^*) \overset{\delta}{\longrightarrow} H^2(\Lambda,\mathbb{Z}) \cong H^2(T,\mathbb{Z})$$

arising from the exponential exact sequence

$$1 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}(V) \overset{e^{2\pi i}}{\longrightarrow} \mathcal{O}(V)^* \longrightarrow 1. \tag{10.8}$$

Here the isomorphism $H^2(\Lambda,\mathbb{Z}) \cong H^2(T,\mathbb{Z})$ is obtained by assigning to a $\mathbb{Z}$-valued 2-cocycle $\{c_{\gamma,\gamma'}\}$ of Λ the alternating bilinear form $\tilde{c}(\gamma,\gamma') = c_{\gamma,\gamma'} - c_{\gamma',\gamma}$. The condition for a 2-cocycle is

$$c_{\gamma_2,\gamma_3} - c_{\gamma_2+\gamma_1,\gamma_3} + c_{\gamma_1,\gamma_2+\gamma_3} - c_{\gamma_1,\gamma_2} = 0. \tag{10.9}$$

It is not difficult to see that this implies that $\tilde{c}$ is an alternating bilinear form.

Let us compute the first Chern class of the line bundle L_α defined by the theta factor (10.4). We find $\beta_\gamma(v) : \Lambda \times V \to \mathbb{C}$ such that $\alpha_\gamma(v) = e^{2\pi i\beta_\gamma(v)}$. Then, using

$$1 = \alpha_{\gamma+\gamma'}(v)/\alpha_{\gamma'}(v+\gamma')\alpha_{\gamma'}(v),$$

we get

$$c_{\gamma,\gamma'}(v) = \beta_{\gamma+\gamma'}(v) - \beta_{\gamma'}(v+\gamma') - \beta_{\gamma'}(v) \in \mathbb{Z}.$$

By definition of the coboundary homomorphism $\delta(L_\alpha)$ is given by the 2-cocycle $\{c_{\gamma,\gamma'}\}$. Returning to our case when $\alpha_\gamma(v)$ is given by (10.4), we get

$$\beta_\gamma(v) = \frac{i}{2}(H(\gamma,\gamma) + 2H(v,\gamma)),$$

$$c_{\gamma,\gamma'}(v) = \beta_{\gamma+\gamma'}(v) - \beta_{\gamma'}(v+\gamma') - \beta_{\gamma'}(v) = \frac{i}{2}(2ReH(\gamma,\gamma') - 2H(\gamma',\gamma)) = -ImH(\gamma,\gamma').$$

Thus

$$c_1(L_\alpha) = \{c_{\gamma,\gamma'} - c_{\gamma',\gamma'}\} = \{-2ImH(\gamma,\gamma')\} = -2S.$$

We would like to have $c_1(L) = S$. For this we should change H to $-H/2$. However the corresponding function $\alpha_\gamma(v)' = e^{\frac{\pi}{2}H(\gamma,\gamma)+\pi H(v,\gamma)}$ is not a theta factor. It satisfies

$$\alpha_{\gamma+\gamma'}(v)' = e^{i\pi S(\gamma,\gamma')}\alpha_\gamma(v+\gamma')'\alpha_{\gamma'}(v)'.$$

We correct the definition by replacing $\alpha_\gamma(v)'$ with $\alpha_\gamma(v)'\chi(\gamma)$, where the map

$$\chi : \Lambda \to \mathbb{C}_1^*$$

has the property

$$\chi(\gamma+\gamma') = \chi(\gamma)\chi(\gamma')e^{i\pi S(\gamma,\gamma')}, \quad \forall \gamma,\gamma' \in \Lambda. \tag{10.10}$$

We call such a map a *semi-character* of Λ. An example of a semi-character is the map $\chi(\gamma) = e^{i\pi S'(\gamma,\gamma)}$, where S' is any bilinear form on Λ with values in $\mathbb{Z}$ such that

$$S'(\gamma,\gamma') - S'(\gamma',\gamma) = S(\gamma,\gamma'). \tag{10.11}$$

Now we can make the right definition of the theta factor associated to S. We set

$$\alpha_\gamma(v) = e^{\frac{\pi}{2}H(\gamma,\gamma)+\pi H(v,\gamma)}\chi(\gamma). \tag{10.12}$$

Clearly $\gamma \to \chi^2(\gamma)$ is a character of Λ (i.e., a homomorphism of abelian groups $\Lambda \to \mathbb{C}_1^*$). Obviously, any character defines a theta factor whose values are constant functions. Its first Chern class is zero. Also note that two semi-characters differ by a character and any character can be given by a formula

$$\chi(\gamma) = e^{2\pi i l(\gamma)},$$

where $l : V \to \mathbb{R}$ is a real linear form on V.

We define the line bundle $L(H,\chi)$ as the line bundle corresponding to the theta factor (10.12). It is clear now that

$$c_1(L(H,\chi)) = S. \tag{10.13}$$

10.3 Now let us interpret global sections of any line bundle constructed from a theta factor $\alpha \in Z^1(V, \mathcal{O}(V)^*)$. Recall that L is isomorphic to the line bundle obtained as the orbit space $V \times \mathbb{C}/\Lambda$ where Λ acts by formula (10.2). Let $s : V \to V \times \mathbb{C}$ be a section of the trivial bundle. It has the form $s(v) = (v, \phi(v))$ for some holomorphic function ϕ on V. So we can identify it with a holomorphic function on V. Assume that, for any $\gamma \in \Lambda$, and any $v \in V$,

$$\phi(v+\gamma) = \alpha_\gamma(v)\phi(v). \tag{10.14}$$

This means that $\gamma \cdot s(v) = s(v+\gamma)$. Thus s descends to a holomorphic section of $L = V \times \mathbb{C}/\Lambda \to V/\Lambda = T$. Conversely, every holomorphic section of L can be lifted to a holomorphic section of $V \times \mathbb{C}$ satisfying (10.14).

We denote by $\Gamma(T, L_\alpha)$ the complex vector space of global sections of the line bundle L_α defined by a theta factor α. In view of the above,

$$\Gamma(T, L_\alpha) = \{\phi \in \mathrm{Hol}(V) : \phi(z+\gamma) = \alpha_\gamma(v)\phi(v)\}. \tag{10.15}$$

Applying this to our case $L = L(H,\chi)$ we obtain

$$\Gamma(T, L(H,\chi)) \cong \{\phi \in \mathrm{Hol}(V) : \phi(v+\gamma) = e^{\frac{\pi}{2}H(\gamma,\gamma')+\pi H(v,\gamma)}\chi(\gamma)\phi(v),\ \forall v \in V, \gamma \in \Lambda\}, \tag{10.16}$$

Let $\beta_\gamma(v) = g(v+\gamma)/g(v)$ be a trivial theta factor. Then the multiplication by g defines an isomorphism

$$\Gamma(T, L_\alpha) \cong \Gamma(T, L_{\alpha\beta}).$$

We shall show that the vector space $\Gamma(T, L(H,\chi))$ is finite-dimensional and compute its dimension.

But first we need some lemmas.

Lemma 1. *Let $S : \Lambda \times \Lambda \to \mathbb{Z}$ be a non-degenerate skew-symmetric bilinear form. Then there exists a basis $\omega_1, \ldots, \omega_{2n}$ of Λ such that $S(\omega_i, \omega_j) = d_i\delta_{i+n,j}, i = 1, \ldots, n$. Moreover, we may assume that the integers d_i are positive and $d_1|d_2|\ldots|d_n$. Under this condition they are determined uniquely.*

Proof. This is well-known, nevertheless we give a proof. We use induction on the rank of Λ. The assertion is obvious for $n = 2$. For any $\gamma \in \Lambda$ the subset of integers $\{S(\gamma,\gamma'), \gamma' \in \Lambda\}$ is a cyclic subgroup of $\mathbb{Z}$. Let d_γ be its positive generator. We set d_1 to be the minimum of the d_γ's and choose ω_1, ω_{n+1} such that $S(\omega_1, \omega_{n+1}) = d_1$. Then for any $\gamma \in \Lambda$, we have $d_1|S(\gamma,\omega_1), S(\gamma,\omega_{n+1})$. This implies that

$$\gamma - \frac{S(\gamma,\omega_1)}{d_1}\omega_{n+1} - \frac{S(\gamma,\omega_{n+1})}{d_1}\omega_1 \in \Lambda' = (\mathbb{Z}\omega_1 + \mathbb{Z}\omega_{n+1})^\perp.$$

Now we use the induction assumption on Λ'. There exists a basis $\omega_2, \ldots, \omega_n, \omega_{n+2}, \ldots, \omega_{2n}$ of Λ' satisfying the properties from the statement of the lemma. Let $d_2, \ldots, d_n$ be the corresponding integers. We must have $d_1|d_2$, since otherwise $S(k\omega_1 + \omega_2, \omega_{n+1} + \omega_{n+2}) = kd_1 + d_2 < d_1$ for some integer k. This contradicts the choice of d_1. Thus $\omega_1, \ldots, \omega_{2n}$ is the desired basis of Λ.

Lemma 2. *Let H be a positive definite Hermitian form on a complex vector space V and let Λ be a lattice in V such that $S = Im(H)$ satisfies (10.5). Let $\omega_1, \ldots, \omega_{2n}$ be a basis of Λ chosen as in Lemma 1 and let Δ be the diagonal matrix $diag[d_1, \ldots, d_n]$. Then the last n vectors ω_i are linearly independent over $\mathbb{C}$ and, if we use these vectors to identify V with $\mathbb{C}^n$, the remaining vectors $\omega_{n+1}, \ldots, \omega_{2n}$ form a matrix $\Omega = X + iY \in M_n(\mathbb{C})$ such that*

(i) $\Omega\Delta^{-1}$ *is symmetric;*
(ii) $Y\Delta^{-1}$ *is positive definite.*

Proof. Let us first check that the vectors $\omega_{n+1},\dots,\omega_{2n}$ are linearly independent over $\mathbb{C}$. Suppose $\lambda_1\omega_{n+1}+\dots+\lambda_n\omega_{2n}=0$ for some $\lambda_i=x_i+iy_i$. Then

$$w=-\sum_{i=1}^{n}x_i\omega_{n+i}=i\sum_{i=1}^{n}y_i\omega_{n+i}=iv.$$

We have $S(iv,v)=S(w,v)=0$ because the restriction of S to $\mathbb{R}\omega_{n+1}+\dots+\mathbb{R}\omega_{2n}$ is trivial. Since $S(iv,v)=H(v,v)$, and H was assumed to be positive definite, this implies $v=w=0$ and hence $x_i=y_i=\lambda_i=0, i=1,\dots,n$.

Now let us use $\omega_{n+1},\dots,\omega_{2n}$ to identify V with $\mathbb{C}^n$. Under this identification, $\omega_{n+i}=e_i$, the i-th unit vector in $\mathbb{C}^n$. Write

$$\omega_j=(\omega_{1j},\dots,\omega_{nj})=(x_{1j},\dots,x_{nj})+i(y_{1j},\dots,y_{nj})=\mathrm{Re}(\omega_j)+i\mathrm{Im}(\omega_j),\quad j=1,\dots,n.$$

We have

$$d_i\delta_{ij}=S(\omega_i,e_j)=\sum_{k=1}^{n}\big(x_{ki}S(e_k,e_j)+y_{ki}S(ie_k,e_j)\big)=\sum_{k=1}^{n}y_{ki}S(ie_k,e_j)=\sum_{k=1}^{n}S(ie_j,e_k)y_{ki}.$$

Let $A=(S(ie_j,e_k))_{j,k=1,\dots,n}$ be the matrix defining $H:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$. Then the previous equality translates into the following matrix equality

$$\Delta=A\cdot Y. \tag{10.17}$$

Since A is symmetric and positive definite, we deduce from this property (ii). To check property (i) we use that

$$0=S(\omega_i,\omega_j)=S(\sum_{k=1}^{n}x_{ki}e_k+y_{ki}ie_k,\sum_{k=1}^{n}x_{kj}e_k+y_{kj}ie_k)=$$

$$=\sum_{k'=1}^{n}x_{k'j}(\sum_{k=1}^{n}y_{ki}S(ie_k,e_{k'}))-\sum_{k=1}^{n}x_{ki}(\sum_{k'=1}^{n}y_{k'j}S(ie_k',e_k))=$$

$$=\sum_{k'=1}^{n}x_{k'j}d_i\delta_{k'i}-\sum_{k=1}^{n}x_{ki}d_j\delta_{kj}=x_{ij}d_i-x_{ji}d_j=d_id_j(x_{ij}d_j^{-1}-x_{ji}d_i^{-1}).$$

In matrix notation this gives

$$X\cdot\Delta^{-1}=(X\cdot\Delta^{-1})^t.$$

This proves the lemma.

Corollary. *Let $H : V \times V \to \mathbb{C}$ be a Hermitian positive definite form on V such that $Im(H)(\Lambda \times \Lambda) \subset \mathbb{Z}$. Let Π be a $2n \times n$-matrix whose columns form a basis of Λ with respect to some basis of V over $\mathbb{R}$. Let A be the matrix of $Im(H)$ with respect to this basis. Then the following Riemann-Frobenius relations hold*
(i) $\Pi A^{-1}\Pi^t = 0$;
(ii) $-i\Pi A^{-1}\bar{\Pi}^t > 0$.

Proof. It is easy to see that the relations do not depend on the choice of a basis. Pick a basis as in Lemma 2. Then, in block matrix notation,

$$\Pi = \begin{pmatrix} \Omega & I_n \end{pmatrix}, \quad A = \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix}. \tag{10.18}$$

This gives

$$\Pi A^{-1}\Pi^t = \begin{pmatrix} \Omega & I_n \end{pmatrix} \begin{pmatrix} 0_n & \Delta^{-1} \\ -\Delta^{-1} & 0_n \end{pmatrix} \begin{pmatrix} \Omega^t \\ I_n \end{pmatrix} = -\Delta^{-1}\Omega^t + \Omega\Delta^{-1} = 0,$$

$$-i\Pi A^{-1}\bar{\Pi}^t = -i \begin{pmatrix} \Omega & I_n \end{pmatrix} \begin{pmatrix} 0_n & \Delta^{-1} \\ -\Delta^{-1} & 0_n \end{pmatrix} \begin{pmatrix} \bar{\Omega}^t \\ I_n \end{pmatrix} =$$
$$= -i(-\Delta^{-1}\bar{\Omega}^t + \Omega\Delta^{-1}) = -i(2iY\Delta^{-1}) = 2Y\Delta^{-1} > 0.$$

Lemma 3. *Let $\lambda : \Gamma \to \mathbb{C}_1^*$ be a character of Λ. Let $\omega_1, \ldots, \omega_{2n}$ be a basis of Λ. Define the vector c_γ by the condition:*

$$\lambda(\omega_i) = e^{2\pi i S(\omega_i, c_\lambda)}, i = 1, \ldots, 2n.$$

Note that this is possible because S is non-degenerate. Then

$$\phi(v) \to \phi(v + c_\gamma)e^{\pi H(v, c_\lambda)}$$

defines an isomorphism from $\Gamma(T, L(H, \chi))$ to $\Gamma(T, L(H, \chi \cdot \lambda))$.

Proof. Let

$$\tilde{\phi}(v) = \phi(v + c_\lambda)e^{\pi H(v, c_\lambda)}.$$

Then

$$\tilde{\phi}(v+\gamma) = e^{\pi H(v+\gamma, c_\lambda)}\phi(v+c_\lambda+\gamma) = e^{\pi H(v, c_\lambda)}\phi(v+c_\lambda)\chi(\gamma)e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v+c_\lambda,\gamma))}e^{\pi H(\gamma, c_\lambda)}$$

$$= \tilde{\phi}(v)e^{\pi(H(\gamma,c_\lambda)-H(c_\lambda,\gamma))}e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma) = \tilde{\phi}(v)e^{2\pi i S(\gamma,c_\lambda)}e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma)$$
$$= \tilde{\phi}(v)e^{\pi(\frac{1}{2}H(\gamma,\gamma)+H(v,\gamma))}\chi(\gamma)\lambda(\gamma).$$

This shows that $\tilde{\phi} \in \Gamma(T, L(H, \chi \cdot \lambda)$. Obviously the map $\phi \to \tilde{\phi}$ is invertible.

10.4 From now on we shall use the notation of Lemma 2. In this notation V is identified with $\mathbb{C}^n$ so that we can use $\mathbf{z} = (z_1, \ldots, z_n)$ instead of v to denote elements of V. Our lattice Λ looks like

$$\Lambda = \mathbb{Z}\omega_1 + \ldots + \mathbb{Z}\omega_n + \mathbb{Z}e_1 + \ldots + \mathbb{Z}e_n.$$

The matrix

$$\Omega = [\omega_1, \ldots, \omega_n] \tag{10.19}$$

satisfies properties (i),(ii) from Lemma 2. Let

$$V_1 = \mathbb{R}e_1 + \ldots + \mathbb{R}e_n = \{\mathbf{z} \in \mathbb{C}^n : Im(\mathbf{z}) = 0\}.$$

We know that the restriction of S to V_1 is trivial. Therefore the restriction of H to V_1 is a symmetric positive definite quadratic form. Let $B : V \times V \to \mathbb{C}$ be a quadratic form such that its restriction to $V_1 \times V_1$ coincides with H (just take B to be defined by the matrix $(H(e_i, e_j))$. Then

$$\alpha'_\gamma(v) = \alpha_\gamma(v)e^{-\pi B(\mathbf{z},\gamma)-\frac{\pi}{2}B(\gamma,\gamma)} = \alpha_\gamma(v)\big(e^{-\frac{\pi}{2}B(\mathbf{z}+\gamma,\mathbf{z}+\gamma)}/e^{-\frac{\pi}{2}B(\mathbf{z},\mathbf{z})}\big) =$$

$$= \chi(\gamma)e^{\pi(H-B)(\mathbf{z},\gamma)+\frac{\pi}{2}(H-B)(\gamma,\gamma)}.$$

Since α and α' differ by a trivial theta factor, they define isomorphic line bundles.

Also, by Lemma 3, we may choose any semi-character χ since the dimension of the space of sections does not depend on its choice. Choose χ in the form

$$\chi_0(\gamma) = e^{i\pi S'(\gamma,\gamma)}, \tag{10.20}$$

where S' is defined in (10.11) and its restriction to V_1 and to $V_2 = \mathbb{R}\omega_1 + \ldots + \mathbb{R}\omega_n$ is trivial, For example, one may take S' to be defined in the basis $\omega_1, \ldots, \omega_n, e_1, \ldots, e_n$ by the matrix $\begin{pmatrix} 0_n & \Delta + I_n \\ I_n & 0 \end{pmatrix}$. We have

$$\chi_0(\gamma) = 1, \gamma \in V_1 \cup V_2.$$

So we will be computing the dimension of $\Gamma(T, L_{\alpha^\sharp})$ where

$$\alpha^\sharp_\gamma(\mathbf{z}) = \chi_0(\gamma)e^{\pi(H-B)(\mathbf{z},\gamma)+\frac{\pi}{2}(H-B)(\gamma,\gamma)}, \tag{10.21}$$

Using (10.17), we have, for any $\mathbf{z} \in V$ and $k = 1, \ldots, n$

$$(H-B)(\mathbf{z}, e_k) = \sum_{i=1}^{n} z_i(H-B)(e_i, e_k) = 0,$$

$$(H-B)(\mathbf{z}, \omega_k) = \mathbf{z} \cdot (H(e_i, e_j)) \cdot \bar{\omega}_k - \mathbf{z} \cdot ((B(e_i, e_j)) \cdot \omega_k = \mathbf{z}(H(e_i, e_j))(\bar{\omega}_k - \omega_k) =$$

$$-2i\mathbf{z} \cdot (S(ie_i, e_j)) \cdot Im\bar{\omega}_k = -2i\mathbf{z} \cdot (AY)e_k = -2i\mathbf{z} \cdot \Delta \cdot e_k = -2id_k z_k. \tag{10.22}$$

Theorem 2.

$$\dim_{\mathbb{C}}\Gamma(T, L(H,\chi)) = |\Delta| = d_1 \cdots d_n.$$

Proof. By (10.22), any $\phi \in \Gamma(T, L_\alpha^\sharp)$ satisfies

$$\phi(\mathbf{z}+e_k) = \alpha_{e_k}(\mathbf{z})\phi(\mathbf{z}) = \phi(\mathbf{z}), \quad k=1,\dots,n,$$

$$\phi(\mathbf{z}+\omega_k) = \alpha_{\omega_k}(\mathbf{z})\phi(\mathbf{z}) = e^{-2\pi i d_k(z_k+\frac{1}{2}\omega_{kk})}\phi(\mathbf{z}), \quad k=1,\dots,n. \tag{10.23}$$

The first equation allows us to expand $\phi(\mathbf{z})$ in Fourier series

$$\phi(\mathbf{z}) = \sum_{\mathbf{r}\in\mathbb{Z}^n} a_{\mathbf{r}} e^{2\pi i \mathbf{r}\cdot\mathbf{z}}.$$

By comparing the coefficients of the Fourier series, the second equality allows us to find the recurrence relation for the coefficients

$$a_{\mathbf{r}} = e^{-\pi i(2\mathbf{r}\cdot\omega_k + d_k\omega_{kk})} a_{\mathbf{r}-d_k e_k}. \tag{10.24}$$

Let

$$M = \{\mathbf{m} = (m_1,\dots,m_n) \in \mathbb{Z}^n : 0 \le m_i < d_i\}.$$

Using (10.24), we are able to express each $a_{\mathbf{r}}$ in the form

$$a_{\mathbf{r}} = e^{2\pi i \lambda(\mathbf{r})} a_{\mathbf{m}},$$

where $\lambda(\mathbf{r})$ is a function in $\mathbf{r}$, $\mathbf{r} \equiv \mathbf{m} \bmod (d_1,\dots,d_n)$, satisfying

$$\lambda(\mathbf{r} - d_k e_k) = \lambda(\mathbf{r}) - \mathbf{r}\cdot\omega_k - \frac{1}{2} d_k\omega_{kk}.$$

This means that the difference derivative of the function λ is a linear function. So we should try some quadratic function to solve for λ. We can take

$$\lambda(\mathbf{r}) = \frac{1}{2}\mathbf{r}\cdot\Omega\cdot(\Delta^{-1}\cdot\mathbf{r}).$$

We have

$$\lambda(\mathbf{r} - d_k e_k) = \frac{1}{2}(-d_k e_k + \mathbf{r})\cdot\Omega\cdot(-e_k + \Delta^{-1}\cdot\mathbf{r}) =$$

$$= \lambda(\mathbf{r}) - \frac{1}{2}\Big(- d_k e_k\cdot\Omega\Delta^{-1}\cdot\mathbf{r} - \mathbf{r}\cdot\Omega\cdot e_k + d_k e_k\cdot\Omega\cdot e_k\Big) = \lambda(\mathbf{r}) - \mathbf{r}\cdot\omega_k - \frac{1}{2}d_k\omega_{kk}.$$

This solves our recurrence. Clearly two solutions of the recurrence (10.24) differ by a constant factor. So we get

$$\phi(\mathbf{z}) = \sum_{\mathbf{m}\in M} c_{\mathbf{m}}\theta_{\mathbf{m}}(\mathbf{z}),$$

for some constants $c_{\mathbf{m}}$, and

$$\theta_{\mathbf{m}}(\mathbf{z}) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{2\pi i\lambda(\mathbf{m}+\Delta\cdot\mathbf{r})}e^{2\pi i\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r})} = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i((\mathbf{m}\cdot\Delta^{-1}+\mathbf{r})\cdot(\Delta\Omega)\cdot(\Delta^{-1}\mathbf{m}+\mathbf{r})+2\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r}))}. \tag{10.25}$$

The series converges uniformly on any bounded subset. In fact, since $\Omega' = \Omega\cdot\Delta^{-1}$ is a positive definite symmetric matrix, we have

$$||Im(\Omega\cdot\Delta^{-1})\cdot\mathbf{r}|| \geq C||\mathbf{r}||^2,$$

where C is the minimal eigenvalue of Ω'. Thus

$$\sum_{\mathbf{r}\in\mathbb{Z}^n} |e^{2\pi i\lambda(\mathbf{m}+\Delta\cdot\mathbf{r})}||e^{2\pi i\mathbf{z}\cdot(\mathbf{m}+\Delta\cdot\mathbf{r})}| = \sum_{\mathbf{r}\in\mathbb{Z}^n} |e^{-\pi Im(\Omega')||(\mathbf{m}+\Delta\cdot\mathbf{r})||^2}|\mathbf{q}^{\mathbf{m}+\Delta\cdot\mathbf{r}}|$$

$$\leq \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{-C\pi||\mathbf{m}+\Delta\cdot\mathbf{r}||^2}|\mathbf{q}|^{\mathbf{m}+\Delta\cdot\mathbf{r}},$$

where $\mathbf{q} = e^{2\pi i\mathbf{z}}$. The last series obviously converges on any bounded set.

Thus we have shown that $\dim\Gamma(T, L_\alpha) \leq \#M = |\Delta|$. One can show, using the uniqueness of Fourier coefficients for a holomorphic function, that the functions $\theta_{\mathbf{m}}$ are linearly independent. This proves the assertion.

10.5 Let us consider the special case when

$$d_1 = \ldots = d_n = d.$$

This means that $\frac{1}{d}S|\Lambda\times\Lambda$ is a unimodular bilinear form. We shall identify the set of residues M with $(\mathbb{Z}/d\mathbb{Z})^n$. One can rewrite the functions $\theta_{\mathbf{m}}$ in the following way:

$$\theta_{\mathbf{m}}(\mathbf{z}) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d\mathbf{z}\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})}. \tag{10.26}$$

Definition. Let $(\mathbf{m},\mathbf{m}') \in (\mathbb{Z}/d\mathbb{Z})^n \oplus (\mathbb{Z}/d\mathbb{Z})^n$ and let Ω be a symmetric complex $n\times n$-matrix with positive definite imaginary part. The holomorphic function

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)}$$

is called the *Riemann theta function of order d with theta characteristic* $(\mathbf{m},\mathbf{m}')$ with respect to Ω.

A similar definition can be given in the case of arbitrary Δ. We leave it to the reader.

So, we see that $\Gamma(T, L_{\alpha^\sharp}) \cong \Gamma(T, L(H,\chi_0))$ has a basis formed by the functions

$$\theta_{\mathbf{m}}(\mathbf{z}) = \theta_{\mathbf{m},0}(d\mathbf{z}; d\Omega), \quad \mathbf{m}\in(\mathbb{Z}/d\mathbb{Z})^n. \tag{10.27}$$

Using Lemma 3, we find a basis of $\Gamma(T, L(H,\chi))$. It consists of functions

$$e^{-\frac{\pi}{2}\mathbf{z}\cdot A\cdot\mathbf{z}}\theta_{\mathbf{m},0}(d(\mathbf{z}+\mathbf{c}; d\Omega)),$$

where A is the symmetric matrix $(H(e_i,e_j)) = (S(ie_i,e_j))$ and $\mathbf{c}\in V$ is defined by the condition

$$\chi(\gamma) = e^{2\pi i S(c,\gamma)}, \gamma = e_i, \omega_i, i = 1,\ldots,n.$$

Let us see how the functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z},\Omega)$ change under translation of the argument by vectors from the lattice Λ. We have

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}+e_k;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+e_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} e^{\frac{2\pi i m_i}{d}} = e^{\frac{2\pi i m_k}{d}}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega),$$

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}+\omega_k;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)+2(\mathbf{z}+\omega_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r}-e_k)\right)} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i\left((\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot\Omega\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2(\mathbf{z}+\omega_k+\frac{1}{d}\mathbf{m}')\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})\right)} e^{-\pi(2z_k+\omega_{kk})+2\frac{\pi i m_k}{d}} =$$

$$= e^{\frac{2\pi i m'_k}{d}} e^{-i\pi(2z_k+\omega_{kk})}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega). \tag{10.28}$$

Comparing this with (10.23), this shows that

$$\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)^d \in \Gamma(T, L_{\alpha^\sharp}) \cong \Gamma(T, L(H,\chi_0)). \tag{10.29}$$

This implies that $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)$ generates a one-dimensional vector space $\Gamma(T,L)$, where $L^{\otimes d}\cong L(H,\chi_0)$. This line bundle is isomorphic to the line bundle $L(\frac{1}{d}H,\chi)$ where $\chi^n = \chi_0$. A line bundle over T with unimodular first Chern class is called a *principal polarization.* Of course, it does not necessary exist. Observe that we have d^{2n} functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)^d$ in the space $\Gamma(T, L_{\alpha^\sharp})$ of dimension d^n. Thus there must be some linear relations between the d-th powers of theta functions with characteristic. They can be explicitly found.

Let us put $\mathbf{m} = \mathbf{m}' = 0$ and consider the Riemann theta function (without characteristic)

$$\Theta(\mathbf{z};\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi(\mathbf{r}\cdot\Omega\cdot\mathbf{r}+2\mathbf{z}\cdot\mathbf{r})}. \tag{10.30}$$

We have

$$\Theta(\mathbf{z}+\frac{1}{d}\mathbf{m}'+\frac{1}{d}\mathbf{m}\cdot\Omega;\Omega) = \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi(\mathbf{r}\cdot\Omega\cdot\mathbf{r}+2(\mathbf{z}+\frac{1}{d}\mathbf{m}'+\frac{1}{d}\mathbf{m}\cdot\Omega)\cdot\mathbf{r})} =$$

$$= \sum_{\mathbf{r}\in\mathbb{Z}^n} e^{i\pi((\mathbf{r}+\frac{1}{d}\mathbf{m})\cdot\Omega\cdot(\mathbf{r}+\frac{1}{d}\mathbf{m})+2(\mathbf{z}+\frac{1}{d}\mathbf{m}')(\frac{1}{d}\mathbf{m}+\mathbf{r})-\frac{1}{d^2}(\mathbf{m}\cdot\mathbf{m}'+\mathbf{m}\cdot\Omega\frac{1}{d}\mathbf{m})} =$$

$$= e^{-i\pi\frac{1}{d^2}(\mathbf{m}\cdot\mathbf{m}'+\mathbf{m}\cdot\Omega\frac{1}{d}\mathbf{m})}\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega). \tag{10.31}$$

Thus, up to a scalar factor, $\theta_{\mathbf{m},\mathbf{m}'}$ is obtained from Θ by translating the argument by the vector

$$\frac{1}{d}(\mathbf{m}'+\mathbf{m}\Omega) \in \frac{1}{d}\Lambda/\Lambda = T_n = \{a \in T : na = 0\}. \tag{10.32}$$

This can be interpreted as follows. For any $a \in T$ denote by t_a the translation automorphism $x \to x+a$ of the torus T. If $v \in V$ represents a, then t_a is induced by the translation map $t_v : V \to V, \mathbf{z} \to \mathbf{z}+v$. Let L_α be the line bundle defined by a theta factor α. Then $\beta : \gamma \to \alpha_\gamma(\mathbf{z}+v)$ defines a theta factor such that $L_\beta = t_a^*(L)$. Its sections are functions $\phi(\mathbf{z}+v)$ where ϕ represents a section of L. Thus the theta functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z};\Omega)$ are sections of the bundle $t_a^*(L^\sharp)$ where $\Theta \in \Gamma(T, L^\sharp)$ and a is an n-torsion point of T represented by the vector (10.32).

10.6 Let L be a line bundle over T. Define the group

$$G(L) = \{a \in T : t_a^*(L) \cong L\}.$$

If L is defined by a theta factor $\alpha_\gamma(\mathbf{z})$, then $t_a^*(L)$ is defined by the theta factor $\alpha_\gamma(\mathbf{z}+v)$, where v is a representative of a in V. Thus, $t_a^*(L) \cong L$ if and only if $\alpha_\gamma(\mathbf{z}+v)/\alpha_\gamma(\mathbf{z})$ is a trivial theta factor, i.e.,

$$\alpha_\gamma(\mathbf{z}+v)/\alpha_\gamma(\mathbf{z}) = g(\mathbf{z}+\gamma)/g(v)$$

for some function $g \in \mathcal{O}(V)^*$. Take $L = L(H,\chi)$. Then

$$g(\mathbf{z}+\gamma)/g(\mathbf{z}) = e^{\frac{\pi}{2}(H(\gamma,\gamma)+2H(\mathbf{z}+v,\gamma))}/e^{\frac{\pi}{2}(H(\gamma,\gamma)+2H(\mathbf{z},\gamma))} = e^{\pi H(v,\gamma)}.$$

Multiplying $g(\mathbf{z})$ by $e^{\pi H(\mathbf{z},v)} \in \mathcal{O}(V)^*$, we get that

$$e^{\pi H(v,\gamma)}e^{\pi H(\mathbf{z}+\gamma,v)}/e^{\pi H(\mathbf{z},v)} = e^{\pi(H(v,\gamma)+H(\gamma,v))} = e^{2\pi i \mathrm{Im} H(v,\gamma)} = e^{2\pi i S(v,\gamma)} \tag{10.33}$$

is the trivial theta factor. This happens if and only if $e^{2\pi i S(v,\gamma)} = 1$. This is true for any theta factor which is given by a character $\chi : \Lambda \to \mathbb{C}_1^*$. In fact $\chi(\lambda) = g(\mathbf{z}+\gamma)/g(\mathbf{z})$ implies that $|g(\mathbf{z})|$ is periodic with respect to Λ, hence is bounded on the whole of V. By Liouville's theorem this implies that g is constant, hence $\chi(\lambda) = 1$. Now the condition $e^{2\pi i S(v,\gamma)} = 1$ is equivalent to

$$S(v,\gamma) \in \mathbb{Z}, \quad \forall \gamma \in \Lambda.$$

So

$$G(L(H,\chi)) \cong \Lambda_S := \{v \in V : S(v,\gamma) \in \mathbb{Z}, \quad \forall \gamma \in \Lambda\}/\Lambda \cong \bigoplus_{i=1}^{n}(\mathbb{Z}/d_i\mathbb{Z})^2. \tag{10.34}$$

If the invariants $d_1, \ldots, d_n$ are all equal to d, this implies

$$G(L(H,\chi)) = \frac{1}{d}\Lambda/\Lambda = T_n \cong (\frac{1}{d}\mathbb{Z}/\mathbb{Z})^{2n} \cong (\mathbb{Z}/d\mathbb{Z})^{2n}.$$

Now let us define the *theta group* of L by

$$\tilde{G}(L) = \{(a, \psi) : a \in G(L), \psi : t_a^*(L) \to L \quad \text{is an isomorphism}\}.$$

Here the pairs $(a, \psi), (a', \psi')$ are multiplied by the rule

$$((a, \psi) \cdot (a', \psi') = (a + a', \psi \circ t_a^*(\psi') : t_{a+a'}(L) = t_a^*(t_{a'}^*(L)) \xrightarrow{t_a^*(\psi')} t_a^*(L) \xrightarrow{\psi} L).$$

The map $\tilde{G}(L) \to G(L), (a, \psi) \to a$ is a surjective homomorphism of groups. Its kernel is the group of automorphisms of the line bundle L. This can be identified with $\mathbb{C}^*$. Thus we have an extension of groups

$$1 \to \mathbb{C}^* \to \tilde{G}(L) \to G(L) \to 1.$$

Let us take $L = L(H, \chi)$. Then the isomorphism $t_a^*(L) \to L$ is determined by the trivial theta factor $e^{\pi H(v,\gamma)}$ and an isomorphism $\psi : t_a^*(L) \to L$ by a holomorphic invertible function $g(\mathbf{z})$ such that

$$g(\mathbf{z} + \gamma)/g(\mathbf{z}) = e^{\pi H(v,\gamma)}. \tag{10.35}$$

It follows from (10.33) that $\psi(\mathbf{z}) = e^{\pi H(\mathbf{z},v)}$ satisfies (10.35). Any other solution will differ from this by a constant factor. Thus

$$\tilde{G}(L) = \{(v, \lambda e^{\pi H(\mathbf{z},v)}), v \in G(L), \lambda \in \mathbb{C}^*\}.$$

We have

$$((v, \lambda e^{\pi H(\mathbf{z},v)}) \cdot (v', \lambda' e^{\pi H(\mathbf{z},v')}) = (v + v', \lambda\lambda' e^{\pi H(\mathbf{z}+v,v')} e^{\pi H(\mathbf{z},v)}) =$$

$$= (v + v', \lambda\lambda' e^{\pi H(\mathbf{z},v+v')} e^{\pi H(v,v')}).$$

This shows that the map

$$(v, \lambda e^{\pi H(\mathbf{z},v)}) \to (v, \lambda)$$

is an isomorphism from the group $\tilde{G}(L)$ onto the group $\tilde{\Lambda}_S$ which consists of pairs $(v, \lambda), v \in G(\Lambda, S), \lambda \in \mathbb{C}$ which are multiplied by the rule

$$(v, \lambda) \cdot (v', \lambda') = (v + v', \lambda\lambda' e^{\pi H(v,v')}).$$

This group defines an extension of groups

$$1 \to \mathbb{C}^* \to \tilde{\Lambda}_S \to \Lambda_S \to 1.$$

The map $(v, \lambda) \to (v, \lambda/|\lambda|)$ is a homomorphism from $\tilde{\Lambda}_S$ into the quotient of the Heisenberg group $\tilde{V}$ by the subgroup $\tilde{\Lambda}$ of elements $(\gamma, \lambda) \in \Lambda \times \mathbb{C}_1^*$. Its image is the subgroup of elements $(v, \lambda), v \in G(L)$ modulo $\tilde{\Lambda}$. Its kernel is the subgroup $\{(0, \lambda) : \lambda \in \mathbb{R}\} \cong \mathbb{R}$.

Let us see that the group $\tilde{G}(L)$ acts naturally in the space $\Gamma(T, L)$. Let s be a section of L. Then we have a canonical identification between $t_a^*(L)_{x-a}$ and L_x. Take $(a, \psi) \in \tilde{G}(L)$.

Then $x \to \psi(s(x))$ is a section of $t_a^*(L)$, and $x \to \psi(s(x-a))$ is a section of L. It is easily checked that the formula

$$((a,\psi)\cdot s)(x) = \psi(s(x-a))$$

is a representation of $\tilde{G}(L)$ in $\Gamma(T,L)$. This is sometimes called the *Schrödinger representation of the theta group.* Take now $L_{\alpha^\sharp} \cong L(H,\chi_0)$. Identify s with a function $\phi(\mathbf{z})$ satisfying

$$\phi(\mathbf{z}+\gamma) = \chi_0(\gamma)e^{\pi(\frac{1}{2}(H-B)(\gamma,\gamma)+(H-B)(\mathbf{z},v))}\phi(\mathbf{z}).$$

Represent (a,ψ) by $(v,\lambda e^{\pi H(\mathbf{z},v)})\frac{g(\mathbf{z}+v)}{g(\mathbf{z})}$. Then

$$\tilde{\phi}(\mathbf{z}) = (a,\psi)\cdot\phi(\mathbf{z}) = \lambda e^{\pi(H-B)(\mathbf{z}-v,v)}\phi(\mathbf{z}-v). \tag{10.36}$$

We have, by using (10.22),

$$\frac{1}{d}e_i\cdot\theta_{\mathbf{m}}(\mathbf{z}) = e^{\pi(H-B)(\mathbf{z}-\frac{1}{d}e_i,\frac{1}{d}e_i)}\theta_{\mathbf{m}}(z-\frac{1}{d}e_i) = \theta_{\mathbf{m}}(z-\frac{1}{d}e_i) =$$

$$\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d(\mathbf{z}-\frac{1}{d}e_i)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})} = e^{\frac{2\pi i m_i}{d}}\theta_{\mathbf{m}}(z),$$

$$\frac{1}{d}\omega_i\cdot\theta_{\mathbf{m}}(\mathbf{z}) = e^{\pi(H-B)(\mathbf{z}-\frac{1}{d}\omega_i,\frac{1}{d}\omega_i)}\theta_{\mathbf{m}}(z-\frac{1}{d}\omega_i) =$$

$$= e^{-2\pi i(z_i+\frac{\omega_i}{2d})}\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}\mathbf{m}+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})+2\pi i d(\mathbf{z}-\frac{1}{d}\omega_i)\cdot(\frac{1}{d}\mathbf{m}+\mathbf{r})} =$$

$$e^{-2\pi i(z_i+\frac{\omega_i}{2d})}\sum_{\mathbf{r}\in\mathbb{Z}^n} e^{\pi i(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})\cdot(d\Omega)\cdot(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})+2\pi i d(\mathbf{z}\cdot(\frac{1}{d}(\mathbf{m}-e_i)+\mathbf{r})} = \theta_{\mathbf{m}-e_i}(z). \tag{10.37}$$

10.7 Recall that given $N+1$ linearly independent sections $s_0,\ldots,s_N$ of a line bundle L over a complex manifold X they define a holomorphic map

$$f: X' \to \mathbb{P}^N, x \to (s_0(x).\ldots,s_N(x)),$$

where X' is an open subset of points at which not all s_i vanish. Applying this to our case we take $L_{\alpha^\sharp} \cong L(H,\chi_0)$ and $s_i = \theta_{\mathbf{m}}(\mathbf{z})$ (with some order in the set of indices M) and obtain a map

$$f: T' \to \mathbb{P}^N, \quad N = d_1\cdots d_n - 1.$$

where $d_1|\ldots|d_n$ are the invariants of the bilinear form $Im(H)|\Lambda\times\Lambda$.

Theorem (Lefschetz). *Assume $d_1 \geq 3$, then the map $f: T \to \mathbb{P}^N$ is defined everywhere and is an embedding of complex manifolds.*

Proof. We refer for the proof to [**Lange**].

Corollary. *Let $T = \mathbb{C}^n/\Lambda$ be a complex torus. Assume that its period matrix Π composed of a basis of Λ satisfies the Riemann-Frobenius conditions from the Corollary to Lemma 2 in section 9.3. Then T is isomorphic to a complex algebraic variety. Conversely, if T has a structure of a complex algebraic variety, then the matrix Π of Λ satisfies the Riemann-Frobenius conditions.*

This follows from a general fact (Chow's Theorem) that a closed complex submanifold of projective space is an algebraic variety. Assuming this the proof is clear. The Riemann-Frobenius conditions are equaivalent to the existence of a positive definite Hermitian form H on $\mathbb{C}^n$ such that $Im(H)|\Lambda\times\Lambda \subset \mathbb{Z}$. This forms defines a line bundle $L(H,\chi)$. Replacing H by $3H$, we may assume that the condition of the Lefschetz theorem is satified. Then we apply the Chow Theorem.

Observe that the group $G(L)$ acts on T via its action by translation. The theta group $\tilde{G}(L)$ acts in $\Gamma(T,L)^*$ by the dual to the Schrödinger representation. The map $T \to \mathbb{P}(\Gamma(T,L)^*)$ given by the line bundle L is equivariant.

Example 1. Let $n = 1$, so that $E = \mathbb{C}/\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ is a Riemann surface of genus 1. Choose complex coordinates such that $\omega_2 = 1, \tau = \omega_1 = a + bi$. Replacing τ by $-\tau$, if needed, we may assume that $b = \mathrm{Im}(\tau) > 0$. Thus the period matrix $\Pi = [\tau, 1]$ satisfies the Riemann-Frobenius conditions when we take $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. The corresponding Hermitian form on $V = \mathbb{C}$ is given by

$$H(z,z') = z\bar{z}'H(1,1) = z\bar{z}'S(i,1) = z\bar{z}'S(\frac{1}{b}(\tau - a),1) = \frac{z\bar{z}'}{b}S(\tau,1) = \frac{z\bar{z}'}{\mathrm{Im}\tau}. \qquad (10.38)$$

The Riemann theta function of order 1 is

$$\Theta(\tau,z) = \sum_{r\in\mathbb{Z}} e^{i\pi(\tau r^2+2rz)}. \qquad (10.39)$$

The Riemann theta functions of order d with theta characteristic are

$$\theta_{m,m'}(\tau,z) = \sum_{r\in\mathbb{Z}} e^{i\pi((r+\frac{m}{d})^2\tau+2(r+\frac{m}{d})(z+\frac{m'}{d}))}, \qquad (10.40)$$

where $(m,m') \in (\mathbb{Z}/d\mathbb{Z})^2$. Take the line bundle L with $c_1(L) = dH$ whose space of sections has a basis formed by the functions

$$\theta_m(\mathbf{z}) = \theta_{m,0}(d\tau, dz) = \sum_{r\in\mathbb{Z}} e^{i\pi((r+\frac{m}{d})^2 d\tau+2(dr+m)z)}, \qquad (10.41)$$

where $m \in \mathbb{Z}/d\mathbb{Z}$. It follows from above that $L = L_{\alpha^\sharp}$, where

$$\alpha^\sharp_{a+b\tau}(z) = e^{-2\pi i(bdz+db^2\tau)}.$$

The theta group $\tilde{G}(L)$ is isomorphic to an extension

$$0 \to \mathbb{C}^* \to \tilde{G} \to (\mathbb{Z}/d\mathbb{Z})^2 \to 1$$

Let $\sigma_1 = (\frac{1}{d} + \Lambda, 1), \sigma_2 = (\frac{\tau}{d} + \Lambda, 1)$. Then σ_1, σ_2 generate $\tilde{G}$ with relation

$$[\sigma_1, \sigma_2] = (0, e^{2\pi i d \mathrm{Im}(H)(\frac{1}{d}, \frac{\tau}{d})}) = (0, e^{-\frac{2\pi i}{d}}). \tag{10.42}$$

It acts linearly in the space $\Gamma(E, L(dH, \chi_0^d))$ by the formulae (10.36)

$$\sigma_1(\theta_m(\mathbf{z})) = e^{-\frac{2\pi i m}{d}} \theta_m(\mathbf{z}), \quad \sigma_2(\theta_m(\mathbf{z})) = \theta_{m-1}(\mathbf{z}).$$

Let us take $d = 2$ and consider the map

$$f : E \to \mathbb{P}^1, \quad z \to (\theta_{0,0}(2\tau, 2z), \theta_{1,0}(2\tau, 2z)).$$

Notice that the functions $\theta_{m,0}(2\tau, 2z)$ are even since

$$\theta_{0,0}(2\tau, 2z) = \sum_{r \in \mathbb{Z}} e^{i\pi((r^2 2\tau) + 2(-2z))} = \sum_{r \in \mathbb{Z}} e^{i\pi((-r)^2 2\tau + 2((-r)2z)} = \theta_{0,0}(2\tau, 2z)$$

$$\theta_{1,0}(2\tau, -2z) = \sum_{r \in \mathbb{Z}} e^{i\pi((-r-\frac{1}{2})^2 2\tau + 2((-r-\frac{1}{2})2z)} =$$

$$= \sum_{r \in \mathbb{Z}} e^{i\pi((-r+1-\frac{1}{2})^2 2\tau) + 2((-r+1-\frac{1}{2})2z)} = \theta_{1,0}(2\tau, 2z).$$

Thus the map f is constant on the pairs $(a, -a) \in E$. This shows that the degree of f is divisible by 2. In fact the degree is equal to 2. This can be seen by counting the number of zeros of the function $\theta_m(z) - c$ in the fundamental parallelogram $\{\lambda + \mu\tau : 0 \le \lambda, \mu \le 1\}$. Or, one can use elementary algebraic geometry by noticing that the degree of the line bundle is equal to 2.

Take $d = 3$. This time we have an embedding $f : E \to \mathbb{P}^2$. The image is a cubic curve. Let us find its equation. We use that the equation is invariant with respect to the Schrödinger representation. If we choose the coordinates (t_0, t_1, t_2) in $\mathbb{P}^2$ which correspond to the functions $\theta_0(z), \theta_1(z), \theta_2(z)$ then the action of $\tilde{G}$ is given by the formulae

$$\sigma_1 : (t_0, t_1, t_2) \to (t_0, \zeta t_1, \zeta^2 t_2), \quad \sigma_2 : (t_0, t_1, t_2) \to (t_2, t_0, t_1),$$

where $\zeta = e^{\frac{2\pi i}{3}}$. Let W be the space of homogeneous cubic polynomials. It decomposes into eigensubspaces with respect to the action of σ_1: $W = W_0 + W_2 + W_3$. Here W_i is spanned by monomials $t_0^a t_1^b t_2^c$ with $a + b + c = 3, b + 2c \equiv i \mod 3$. Solving these congruences, we find that $(a, b, c) = (c, c, c) \mod 3$. This implies that the equation of $f(E)$ is in the *Hesse form*

$$\lambda_0 t_0^3 + \lambda_1 t_1^3 + \lambda_2 t_2^3 + \lambda_4 t_0 t_1 t_2 = 0.$$

Since it is also invariant with respect to σ_2, we obtain the equation

$$F(\lambda) = t_0^3 + t_1^3 + t_2^3 + \lambda t_0 t_1 t_2 = 0. \tag{10.43}$$

Since $f(E)$ is a nonsingular cubic,

$$\lambda^3 \neq -27.$$

we can also throw in more symmetries by considering the metaplectic representation. The group $Sp(2, \mathbb{Z}/3\mathbb{Z})$ is of order 24, so that together with the group $E_3 \cong (\mathbb{Z}/3\mathbb{Z})^2$ it defines a subgroup of the group of projective transformations of projective plane whose order is 216. This group is called the *Hesse group*. The elements g of $Sp(2, \mathbb{Z}/3\mathbb{Z})$ transform the curve $F(\lambda) = 0$ to the curve $F(\lambda') = 0$ such that the transformation $g : \lambda \to \lambda'$ defines an isomorphism from the group $Sp(2, \mathbb{Z}/3\mathbb{Z})$ onto the octahedron subgroup of automorphisms of the Riemann sphere $\mathbb{P}^1(\mathbb{C})$.

Example 2. Take $n = 2$ and consider the lattice with the period matrix

$$\Pi = \begin{pmatrix} \sqrt{-2} & \sqrt{-5} & 1 & 0 \\ \sqrt{-3} & \sqrt{-7} & 0 & 1. \end{pmatrix}$$

Suppose there exists a skew-symmetric matrix

$$A = \begin{pmatrix} 0 & b_{12} & b_{13} & b_{14} \\ -b_{12} & 0 & b_{23} & b_{24} \\ -b_{13} & -b_{23} & 0 & b_{34} \\ -b_{14} & -b_{24} & -b_{34} & 0 \end{pmatrix}$$

such that $\Pi \cdot A \cdot \Pi^t = 0$. Then

$$b_{12} + b_{13}\sqrt{-3} + b_{14}\sqrt{-7} - b_{23}\sqrt{-2} + b_{34}\sqrt{-14} - b_{24}\sqrt{-5} - b_{34}\sqrt{-15} = 0.$$

Since the numbers $1, \sqrt{-3}, \sqrt{-7}, \sqrt{-14} - \sqrt{-15}, \sqrt{-2}$ are linearly independent over $\mathbb{R}$, we obtain $A = 0$. This shows that our torus is not algebraic.

Example 3. Let $n = d = 2$ and T be a torus admiting a principal polarization. Then the theta functions $\theta_{\mathbf{m},0}(2\mathbf{z}; 2\Omega), \mathbf{m} \in (\mathbb{Z}/2\mathbb{Z})^2$ define a map of degree 2 from T to a surface of degree 4 in $\mathbb{P}^3$. The surface is isomorphic to the quotient of T by the involution $a \to -a$. It has 16 singular points, the images of 16 fixed points of the involution. The surface is called a *Kummer surface*. Similar to example 1, one can write its equation:

$$\lambda_0(t_0^4 + t_1^4 + t_2^4 + t_3^4) + \lambda_1(t_0^2t_1^2 + t_2^2t_3^2) + \lambda_2(t_0^2t_2^2 + t_1^2t_3^2) + \lambda_3(t_0^2t_3^2 + t_1^2t_2^2) + \lambda_4 t_0t_1t_2t_3 = 0.$$

Here

$$\lambda_0^3 - \lambda_0(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - \lambda_4^2) + 2\lambda_1\lambda_2\lambda_3 = 0.$$

If $n = 2, d = 3$, the map $f : T \to \mathbb{P}^8$ has its image a certain surface of degree 18. One can write its equations as the intersection of 9 quadric equations.

10.8 When we put the period matrix Π satisfying the Riemann-Frobenius relation in the form $\Pi = [\Omega\ I_n]$, the matrix $\Omega\Delta^{-1}$ belongs to the *Siegel domain*

$$\mathcal{Z}_n = \{Z = X + iY \in M_n(\mathbb{C}) : \Omega = \Omega^t, Y > 0\}.$$

It is a domain in $\mathbb{C}^{n(n+1)/2}$ which is homogeneous with respect to the action of the group

$$Sp(2n, \mathbb{R}) = \{M \in GL(2n, \mathbb{R}) : M \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix} \cdot M^t = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}\}.$$

The group $Sp(2n, \mathbb{R})$ acts on $\mathcal{Z}_n$ by the formula

$$M : Z \to M \cdot Z := (A\Omega + B) \cdot (C\Omega + D)^{-1}, \tag{10.44}$$

where we write M as a block matrix of four $n \times n$-matrices

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

(one proves that $C\Omega + D$ is always invertible). So we see that any algebraic torus defines a point in the Siegel domain $\mathcal{Z}_n$ of dimension $n(n+1)/2$. This point is defined modulo the transformaton $Z \to M \cdot Z$, where $M \in \Gamma_{n,\Delta} := Sp(2n, \mathbb{Z})_\Delta$ is the subgroup of $Sp(2n, \mathbb{R})$ of matrices with integer entries satisfying

$$M \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix} \cdot M^t = \begin{pmatrix} 0_n & \Delta \\ -\Delta & 0_n \end{pmatrix}$$

This corresponds to changing the basis of Λ preserving the property that $S(\omega_i, \omega_{j+n}) = d_i\delta_{ij}$. The group $\Gamma_{n,\Delta}$ acts discretely on $\mathcal{Z}_n$ and the orbit space

$$\mathcal{A}_{g,d_1,\ldots,d_n} = \mathcal{Z}_n/\Gamma_{n,\Delta}$$

is the coarse moduli space of abelian varieties of dimension n together with a line bundle with $c_1(L) = S$ where S has the invariants $d_1, \ldots, d_n$. In the special case when $d_1 = \ldots = d_n = d$, we have $\Gamma_{n,\Delta} = Sp(2n, \mathbb{Z})$. It is called the *Siegel modular group* of order n. The corresponding moduli space $\mathcal{A}_{g,d,\ldots,d} \cong \mathcal{A}_{n,1,\ldots,1}$ is denoted by $\mathcal{A}_n$. It parametrizes principal polarized abelian varieties of dimension n. When $n = 1$, we obtain

$$\mathcal{Z}_1 = H := \{z = x + iy : y > 0\},\ Sp(2, \mathbb{R}) = SL(2, \mathbb{R}).$$

The group Γ_1 is the modular group $\Gamma = SL(2, \mathbb{Z})$. It acts on the upper half-plane H by Moebius transformations $z \to az + b/cz + d$. The quotient H/Γ is isomorphic to the complex plane $\mathbb{C}$.

The theta functions $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}; \Omega)$ can be viewed as holomorphic function in the variable $\Omega \in \mathcal{Z}_n$. When we plug in $z = 0$, we get

$$\theta_{\mathbf{m},\mathbf{m}'}(0; \Omega) = \sum_{\mathbf{r} \in \mathbb{Z}^n} e^{\pi i \left((\frac{1}{d}\mathbf{m}+\mathbf{r}) \cdot \Omega \cdot (\frac{1}{d}\mathbf{m}+\mathbf{r}) + 2(\frac{1}{d}\mathbf{m}') \cdot (\frac{1}{d}\mathbf{m}+\mathbf{r}) \right)}$$

These functions are *modular forms* with respect to some subgroup of Γ_n. For example in the case $n = 1, d = 2k$

$$\theta_0(d\tau, 0) = \sum_{r\in\mathbb{Z}} e^{i\pi 2kr^2\tau}$$

is a modular form of weight k with respect to Γ (see [**Serre**]).

Exercises.

1. Prove that the cohomology group $H^k(\Lambda, \mathbb{Z})$ is isomorphic to $\bigwedge^k(\Lambda)^*$.

2. Let Γ be any group acting holomorphically on a complex manifold M. Define automorphy factors as 1-cocycles from $Z^1(\Gamma, \mathcal{O}(M)^*)$ where Γ acts on $\mathcal{O}(M)^*$ by translation in the argument. Show that the function $\alpha_\gamma(x) = det(d\gamma_x)^k$ is an example of an automorphy factor.

3. Show that any theta factor on $\mathbb{C}^n$ with respect to a lattice Λ is equivalent to a theta factor of the form $e^{a_\gamma\cdot\mathbf{z}+b_\gamma}$ where $a : \Gamma \to \mathbb{C}^n, b : \Lambda \to \mathbb{C}$ are certain functions.

4. Assuming that the previous is true, show that
 (i) a is a homomorphism,
 (ii) $E(\gamma, \gamma') = a_\gamma \cdot \gamma' - a_{\gamma'} \cdot \gamma$ is a skew-symmetric bilinear form,
 (iii) the imaginary part of the function $c(\gamma) = b_\gamma - \frac{1}{2}a_\gamma \cdot \gamma$ is a homomorphism.

5. Show that the group of characters $\chi : T \to \mathbb{C}^*$ is naturally isomorphic to the torus V^*/Λ^*, where $\Lambda^* = \{\phi : V \to \mathbb{R} : \phi(\Lambda) \subset \mathbb{Z}\}$. Put the complex structure on this torus by defining the complex structure on V^* by $J\phi(v) = \phi(Jv)$, where J is a complex structure on V. Prove that
 (i) the complex torus $T^* = V^*/\Lambda^*$ is algebraic if $T = V/\Lambda$ is algebraic;
 (ii) $T^* = V^*/\Lambda^* \cong T$ if T admits a principal polarization;
 (iii) T^* is naturally isomorphic to the kernel of the map $c_1 : Pic(T) \to H^2(T, \mathbb{Z})$.

6. Show that a theta function $\theta_{\mathbf{m},\mathbf{m}'}(\mathbf{z}; \Omega)$ is even if $\mathbf{m} \cdot \mathbf{m}'$ is even and odd otherwise. Compute the number of even functions and the number of odd theta functions $\theta_{\mathbf{m},\mathbf{m}'}(z; \Omega)$ of order d.

7. Find the equations of the image of an elliptic curve under the map given by theta functions $\theta_{\mathbf{m}}(\mathbf{z})$ of order 4.

Lecture 11. FIBRE G-BUNDLES

11.1 From now on we shall study *fields*, like electric, or magnetic or gravitation fields. This will be either a scalar function $\psi(\mathbf{x}, t)$ on the space-time $\mathbb{R}^4$, or, more generally, a section of some vector bundle over this space, or a connection on such a bundle. Many such functions can be interpreted as caused by a particle, like an electron in the case of an electric field. The origin of such functions could be different. For example, it could be a wave function $\psi(x)$ from quantum mechanics. Recall that its absolute value $|\psi(x)|$ can be interpreted as the distribution density for the probability of finding the particle at the point x. If we multiply ψ by a constant $e^{i\theta}$ of absolute value 1, it will not make any difference for the probability. In fact, it does not change the corresponding state, which is the projection operator P_ψ. In other words, the "phase" θ of the particle is not observable. If we write $\psi = \psi_1 + i\psi_2$ as the sum of the real and purely imaginary parts, then we may think of the values of ψ as vectors in a 2-plane $\mathbb{R}^2$ which can be thought of as the "internal space" of the particle. The fact that the phase is not observable implies that no particular direction in this plane has any special physical meaning. In quantum field theory we allow the interpretation to be that of the internal space of a particle. So we should think that the particle is moving from point to point and carrying its internal space with it. The geometrical structure which arises in this way is based on the notion of a fibre bundle. The disjoint union of the internal spaces forms a fibre bundle $\pi : S \to M$. Its fibre over a point $x \in M$ of the space-time manifold M is the internal space S_x. It could be a vector space (then we speak about a vector bundle) or a group (this leads to a principal bundle). For example, we can interpret the phase angle θ as an element of the group $U(1)$. It is important that there is no common internal space in general, so one cannot identify all internal spaces with the same space F. In other words, the fibre bundle is not trivial, in general. However, we allow ourselves to identify the fibres along paths in M. Thus, if $x(t)$ depends on time t, the internal state $\tilde{x}(t) \in S_{x(t)}$ describes a path in S lying over the original path. This leads to the notion of "parallel transport" in the fibre bundle, or equivalently, a " connection ". In general, there is no reason to expect that different paths from x to y lead to the same parallel transport of the internal state. They could differ by application of a symmetry group acting in the fibres (the structure group of the fibre bundle). Physically, this is viewed as a "phase shift". It is produced by the external field.

So, a connection is a mathematical interpretation of this field. Quantitatively, the phase shift is described by the "curvature" of the connection .

If $\phi(x) \in \mathbb{R}^n$ is a section of the trivial bundle, one can define the trivial transport by the differential operators $\phi(x) \to \phi(x + \Delta_\mu x) = \phi(x) + \frac{\partial}{\partial x_\mu}\phi\Delta_\mu x$. In general, still taking the trivial bundle, we have to replace $\partial_\mu = \frac{\partial}{\partial x_\mu}$ with

$$\nabla_\mu = \partial_\mu + A_\mu,$$

where $A_\mu(x) : S_x \to S_x$ is the operator acting in the internal state space S_x (an element of the structure group G). The phase shift is determined by the commutator of the operators ∇_μ:

$$F_{\mu\nu} = [\nabla_\mu, \nabla_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Here the group G is a Lie group, and the commutator takes its value in the Lie algebra $\mathfrak{g}$ of G. The expression $\{F_{\mu\nu}\}$ can be viewed as a differential 2-form on M with values in $\mathfrak{g}$. It is the curvature form of the connection $\{A_\mu\}$.

In general our state bundle $S \to M$ is only *locally trivial*. This means that, for any point $x \in M$, one can find a coordinate system in its neighborhood U such that $\pi^{-1}(U_i)$ can be identified with $U \times F$, and the connection is given as above. If V is another neighborhood, then we assume that the new identification $\pi^{-1}(V) = V \times F$ differs over $U \cap V$ from the old one by the "gauge transformation" $g(x) \in G$ so that $(x, s) \in U \times F$ corresponds to $(x, g(x)s)$ in $V \times F$. We shall see that the connection changes by the formula

$$A_\mu \to g^{-1}A_\mu + g^{-1}\partial_\mu g.$$

Note that we may take $V = U$ and change the trivialization by a function $g : U \to G$. The set of such functions forms a group (infinite-dimensional). It is called the gauge group.

This constitutes our introduction for this and the next lecture. Let us go to mathematics and give precise definitions of the mathematical structures involved in definitions of fields.

11.2 We shall begin with recalling the definition of a fibre bundle.

Definition. Let F be a smooth manifold and G be a subgroup of its group of diffeomorphisms. Let $\pi : S \to M$ be a smooth map. A family $\{(U_i, \phi_i)\}_{i\in I}$ of pairs (U_i, ϕ_i), where U_i is an open subset of M and $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ is a diffeomorphism, is called a *trivializing family* of π if

(i) for any $i \in I$, $\pi \circ \phi_i = pr_1$;

(ii) for any $i, j \in I$, $\phi_j^{-1} \circ \phi_i : (U_i \cap U_j) \times F \to (U_i \cap U_j) \times F$ is given by $(x, a) \to (x, g_{ij}(x)(a))$, where $g_{ij}(x) \in G$.

The open cover $(U_i)_{i\in I}$ is called the *trivializing cover*. The corresponding diffeomorphisms ϕ_i are called the *trivializing diffeomorphisms*. The functions $g_{ij} : U_i \cap U_j \to G, x \to g_{ij}(x)$, are called the *transition functions* of the trivializing family. Two trivializing families $\{(U_i, \phi_i)\}_{i\in I}$ and $\{(V_j, \psi_j)\}_{j\in J}$ are called equivalent if for any $i \in I, j \in J$, the map $\psi_j^{-1} \circ \phi_i : (U_i \cap V_j) \times F \to \pi^{-1}(U_i \cap V_j)$ is given by a function $(x, a) \to (x, g(x)(a))$, where $g(x) \in G$.

A *fibre G-bundle with typical fibre F* is a smooth map $\pi : S \to M$ together with an equivalence class of trivializing families.

Let $\pi : S \to M$ be a fibre G-bundle with typical fibre F. For any $x \in X$ we can find a pair (U, ϕ) from some trivializing family such that $x \in U$. Let $\phi_x : F \to S_x := \pi^{-1}(x)$ be the restriction of ϕ to $\{x\} \times F$. We call such a map a *fibre marking.* If $\psi_x : F \to F$ is another fibre marking, then, by definition, $\psi_x^{-1} \circ \phi_x : F \to F$ belongs to the group G. In particular, if two markings are defined by (U_i, ϕ_i) and (U_j, ϕ_j) from the same family of trivializing diffeomorphisms, then, for any $a \in F$,

$$(\phi_i)_x(a) = (\phi_j)_x(g_{ij}(x)(a)), \tag{11.1}$$

where g_{ij} is the transition function of the trivializing family.

Fibre G-bundles form a category with respect to the following definition of morphism:

Definition. Let $\pi : S \to M$ and $\pi' : S' \to M$ be two fibre G-bundles with the same typical fibre F. A smooth map $f : S \to S'$ is called a *morphism* of fibre G-bundles if $\pi = \pi' \circ f$ and, for any $x \in M$ and fibre markings $\phi_x : F \to S_x, \psi_x : F \to S'_x$, the composition $\psi_x^{-1} \circ \phi_x : F \to F$ belongs to G.

A fibre G-bundle is called *trivial* if it is isomorphic to the fibre G-bundle $pr_1 : M \times F \to M$ with the trivializing family $\{(M, id_{M\times F})\}$.

Let $f : S \to S'$ be a morphism of two fibre G-bundles. One may always find a common trivializing open cover $(U_i)_{i\in I}$ for π and π'. Let $\{\phi_i\}_{i\in I}$ and $\{\psi_i\}_{i\in I}$ be the corresponding sets of trivializing diffeomorphisms. Then a morphism $f : S \to S'$ is defined by the maps $f_i : U_i \to G$ such that the map $(x, a) \to (x, f_i(x)(a))$ is a diffeomorphism of $U_i \times F$, and

$$f_i(x) = g'_{ij}(x)^{-1} \circ f_j(x) \circ g_{ij}(x), \quad \forall x \in U_i \cap U_j, \tag{11.2}$$

where g_{ij} and g'_{ij} are the transition functions of $\{(U_i, \phi_i)\}$ and $\{(U_i, \psi_i)\}$, respectively.

Remark 1. The group $\mathrm{Diff}(F)$ has a natural structure of an infinite dimensional Lie group. We refer for the definition to [**Pressley**]. To simplify our life we shall work with fibre G-bundles, where G is a finite-dimensional Lie subgroup of $\mathrm{Diff}(F)$. One can show that the functions g_{ij} are smooth functions from $U_i \cap U_j$ to G.

Let $(U_i)_{i\in I}$ be a trivializing cover of a fibre G-bundle $S \to M$ and g_{ij} be its transition functions. They satisfy the following obvious properties:

(i) $g_{ii} \equiv 1 \in G$,
(ii) $g_{ij} = g_{ji}^{-1}$,
(iii) $g_{ij} \circ g_{jk} = g_{ik}$ over $U_i \cap U_j \cap U_k$.

Let g'_{ij} be the set of transition functions corresponding to another trivializing family of $S \to M$ with the same trivializing cover. Then for any $x \in U_i \cap U_j$,

$$g'_{ij}(x) = h_i(x)^{-1} \circ g_{ij}(x) \circ h_j(x),$$

where

$$h_i(x) = (\psi_i)_x^{-1} \circ (\phi_i)_x : U_i \to G, i \in I.$$

We can give the following cohomological interpretation of transition functions. Let $\mathcal{O}_M(G)$ denote the sheaf associated with the pre-sheaf of groups $U \to \{$smooth functions from U to $G\}$. Then the set of transition functions $\{g_{ij}\}_{i,j\in I}$ with respect to a trivializing family $\{(U_i, \phi_i)\}$ of a fibre G-bundle can be identified with a Čech 1-cocycle

$$\{g_{ij}\} \in Z^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G)).$$

Recall that two cocycles $\{g_{ij}\}$ are called equivalent if they satisfy (11.2) for some collection of smooth functions $f_i : U_i \to G$. The set of equivalence classes of 1-cocycles is denoted by $\check{H}^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G))$. It has a marked element 1 corresponding to the equivalence class of the cocycle $g_{ij} \equiv 1$. This element corresponds to a trivial G-bundle. By taking the inductive limit of the sets $\check{H}^1(\{U_i\}_{i\in I}, \mathcal{O}_M(G))$ with respect to the inductive set of open covers of M, we arrive at the set $H^1(M, \mathcal{O}_M(G))$. It follows from above that a choice of a trivializing family and the corresponding transition functions g_{ij} defines an injective map from the set $FIB_M(F;G)$ of isomorphism classes of fibre G-bundles over M with typical fibre F to the cohomology set $H^1(M, \mathcal{O}_M(G))$. Conversely, given a cocycle $g_{ij} : U_i \cap U_j \to G$ for some open cover $(U_i)_{i\in I}$ of M, one may define the fibre G-bundle as the set of equivalence classes

$$S = \coprod_{i\in I} U_i \times F/R, \tag{11.3}$$

where $(x_i, a_i) \in U_i \times F$ is equivalent to $(x_j, a_j) \in U_j \times F$ if $x_i = x_j = x \in U_i \cap U_j$ and $a_i = g_{ij}(x)(a_j)$. The properties (i),(ii),(iii) of a cocycle are translated into the definition of equivalence relation. The structure of a smooth manifold on S is defined in such a way that the factor map is a local diffeomorphism. The projection $\pi : S \to M$ is defined by the projections $U_i \times F \to U_i$. It is clear that the cover $(U_i)_{i\in I}$ is a trivializing cover of S and the trivializing diffeomorphisms are the restrictions of the factor map onto $U_i \times F$.

Summing up, we have the following:

Theorem 1. *A choice of transition functions defines a bijection*

$$FIB_M(F;G) \longleftrightarrow H^1(M, \mathcal{O}_M(G))$$

between the set of isomorphism classes of fibre G-bundles with typical fibre F and the set of Čech 1-cohomology with coefficients in the sheaf $\mathcal{O}_M(G)$. Under this map the isomorphism class of the trivial fibre G-bundle $M \times F$ corresponds to the cohomology class of the trivial cocycle.

A *cross-section* (or just a *section*) of a fibre G-bundle $\pi : S \to M$ is a smooth map $s : M \to S$ such that $p \circ s = id_M$. If $\{U_i, \phi_i\}_{i\in I}$ is a trivializing family, then a section s is defined by the sections $s_i = \phi^{-1} \circ s|U_i : U_i \to U_i \times F$ of the trivial fibre bundle $U_i \times F \to U_i$. Obviously, we can identify s_i with a smooth function $s_i : U_i \to F$ such that $s_i(x) = (x, s_i(x))$. It follows from (11.1) that, for any $x \in U_i \cap U_j$,

$$s_j(x) = g_{ij}(x) \circ s_i(x), \tag{11.4}$$

where $g_{ij} : U_i \cap U_j \to G$ are the transition functions of $\pi : S \to M$ with respect to $\{U_i, \phi_i\}_{i\in I}$.

Finally, for any smooth map $f : M' \to M$ and a fibre G-bundle $\pi : S \to M$ we define the *induced fibre bundle* (or the *inverse transform*) $f^*(S) = S \times_M M' \to M'$. It has a natural structure of a fibre G-bundle with typical fibre F. Its transition families are inverse transforms

$$f^*(\phi_i) = \phi_i \times id : f^{-1}(U_i) \times F = (U_i \times F) \times_{U_i} f^{-1}(U_i) \to \pi^{-1}(U_i) \times_{U_i} f^{-1}(U_i).$$

of the transition functions of $S \to M$. If $M' \to M$ is the identity map of an open submanifold M' of M, we denote the induced bundle by $S|M'$ and call it the *restriction* of the bundle to M'.

11.3 A sort of universal example of a fibre G-bundle is a *principal G-bundle*. Here we take for typical fibre F a Lie group G. The structure group is taken to be G considered as a subgroup of Diff(F) consisting of left translations $L_g : a \to g \cdot a$. Let $PFIB_M(G)$ denote the set of isomorphism classes of principal G-bundles. Applying Theorem 1, we obtain a natural bijection

$$PFIB_M(G) \longleftrightarrow H^1(M, \mathcal{O}_M(G)). \tag{11.5}$$

Let $s_U : U \to P$ be a section of a principal G-bundle over an open subset U of M. It defines a trivialization of $\phi_U : U \times G \to \pi^{-1}(U)$ by the formula

$$\phi_U((x, g)) = s_U(x) \cdot g.$$

Since $\phi_U((x, g \cdot g')) = s_U(x) \cdot (g \cdot g') = (s_U(x) \cdot g) \cdot g' = \phi_U((x, g)) \cdot g'$, the map ϕ_U is an isomorphism from the trivial principal G-bundle $U \times G$ to $P|U$. Let $\{(U_i, \phi_i)\}_{i\in I}$ be a trivializing family of P. If $s_i : U_i \to P$ is a section over U_i, then for any $x \in U_i \cap U_j$, we can write $s_i(x) = s_j(x) \cdot u_{ij}(x)$ for some $u_{ij} : U_i \cap U_j \to G$. The corresponding trivializations are related by

$$\phi_{U_i}((x, g)) = s_i(x) \cdot g = (s_j(x) \cdot u_{ij}(x)) \cdot g = \phi_{U_j}((x, u_{ij}(x) \cdot g)). \tag{11.6}$$

Comparing this with formula (11.1), we find that the function u_{ij} coincides with the transition function g_{ij} of the fibration P.

Let $\pi : P \to M$ be a principal G-bundle. The group G acts naturally on P by right translations along fibres $g' : (x, g) \to (x, g \cdot g')$. This definition does not depend on the local trivialization of P because left translations commute with right translations. It is clear that orbits of this action are equal to fibres of $P \to M$, and $P/G \cong M$.

Let $P \to M$ and $P' \to M$ be two principal G-bundles. A smooth map (over M) $f : P \to P'$ is a morphism of fibre G-bundles if and only if, for any $g \in G, p \in P$,

$$f(p \cdot g) = f(p) \cdot g.$$

This follows easily from the definitions. One can use this to extend the notion of a morphism of G-bundles for principal bundles with not necessarily the same structure group.

Definition. Let $\pi : P \to M$ be a principal G-bundle and $\pi' : P' \to M$ be a principal G'-bundle. A smooth map $f : P \to P'$ is a morphism of principal bundles if $\pi = \pi' \circ f$ and there exists a morphism of Lie groups $\phi : G \to G''$ such that, for any $p \in P, g \in G$,

$$f(p \cdot g) = f(p) \cdot \phi(g).$$

Any fibre G-bundle with typical fibre F can be obtained from some principal G-bundle by the following construction. Recall that a homomorphism of groups $\rho : G \to \mathrm{Diff}(F)$ is called a *smooth action* of G on F if for any $a \in F$, the map $G \to F, g \to \rho(g)(x)$ is a smooth map. A smooth action is called *faithful* if its kernel is trivial. Fix a principal G-bundle $P \to M$ and a smooth faithful action $\rho : G \to \mathrm{Diff}(F)$ and consider a fibre $\rho(G)$-bundle with typical fibre F over M defined by (11.4) using the transition functions $\rho(g_{ij})$, where g_{ij} is a set of transition functions of P. This bundle is called the fibre G-bundle associated to a principal G-bundle by means of the action $\rho : G \to \mathrm{Diff}(F)$ and transition functions g_{ij}. Clearly, a change of transition functions changes the bundle to an isomorphic G-bundle. Also, two representations with the same image define the same G-bundles. Finally, it is clear that any fibre G-bundle is associated to a principal G'-bundle, where $G' \to G$ is an isomorphism of Lie groups.

There is a canonical construction for an associated G-bundle which does not depend on the choice of transition functions.

Let $\rho : G \to \mathrm{Diff}(F)$ be a faithful smooth action of G on F. The group G acts on $P \times F$ by the formula $g : (p, a) \to (p \cdot g, \rho(g^{-1})(a))$. We introduce the orbit space

$$P \times_\rho F := P \times F/G, \tag{11.7}$$

Let $\pi' : P \times_\rho F \to M$ be defined by sending the orbit of (p, a) to $\pi(p)$. It is clear that, for any open set $U \subset M$, the subset $\pi^{-1}(U) \times F$ of $P \times F$ is invariant with respect to the action of G, and $\pi'^{-1}(U) = \pi^{-1}(U) \times F/G$. If we choose a trivializing family $\phi_i : U_i \times G \to \pi^{-1}(U_i)$, then the maps

$$(U_i \times G) \times F \to U_i \times F, ((x, g), a) \to (x, \rho(g)(x))$$

define, after passing to the orbit space, diffeomorphisms $\pi'^{-1}(U) \to U_i \times F$. Taking inverses we obtain a trivializing family of $\pi' : P \times_\rho F \to M$. This defines a structure of a fibre G-bundle on $P \times_\rho F$.

The principal bundle P acts on the associated fibre bundle $P \times_\rho F$ in the following sense. There is a map $(P \times_\rho F) \times P \to P \times_\rho F$ such that its restriction to the fibres over a point $x \in M$ defines the map $F \times P_x \to F$ which is the action ρ of $G = P_x$ on F.

Let $S \to M$ be a fibre G-bundle and G' be a subgroup of G. We say that the structure group of S can be *reduced* to G' if one can find a trivializing family such that its transition functions take values in G'. For example, the structure group can be reduced to the trivial group if and only if the bundle is trivial.

Let G' be a Lie subgroup of a Lie group G. It defines a natural map of the cohomology sets

$$i : H^1(M, \mathcal{O}_M(G')) \to H^1(M, \mathcal{O}_M(G)).$$

A principal G-bundle P defines an element in the image of this map if and only if its structure group can be reduced to G'.

11.4 We shall deal mainly with the following special case of fibre G-bundles:

Definition. A rank n *vector bundle* $\pi : E \to M$ is a fibre G-bundle with typical fibre F equal to $\mathbb{R}^n$ with its natural structure of a smooth manifold. Its structure group G is the group $GL(n, \mathbb{R})$ of linear automorphisms of $\mathbb{R}^n$.

The local trivializations of a vector bundle allow one to equip its fibres with the structure of a vector space isomorphic to $\mathbb{R}^n$. Since the transition functions take values in $GL(n, \mathbb{R})$, this definition is independent of the choice of a set of trivializing diffeomorphisms. For any section $s : M \to E$ its value $s(x) \in E_x = \pi^{-1}(x)$ cannot be considered as an element in $\mathbb{R}^n$ since the identification of E_x with $\mathbb{R}^n$ depends on the trivialization. However, the expression $s(x) = 0 \in E_x$ is well-defined, as well as the sum of the sections $s + s'$ and the scalar product λs. The section s such that $s(x) = 0$ for all $x \in M$ is called the *zero section*. The set of sections is denoted by $\Gamma(E)$. It has the structure of a vector space. It is easy to see that

$$\mathcal{E} : U \to \Gamma(E|U)$$

is a sheaf of vector spaces. We shall call it the *sheaf* of sections of E.

If we take $E = M \times \mathbb{R}$ the trivial vector bundle, then its sheaf of sections is the structure sheaf $\mathcal{O}_M$ of the manifold M.

One can define the category of vector bundles. Its morphisms are morphisms of fibre bundles such that restriction to a fibre is a linear map. This is equivalent to a morphism of fibre $GL(n, \mathbb{R})$-bundles.

We have

$$\{\text{ rank } n \text{ vector bundles over } M/\text{isomorphism}\} \longleftrightarrow H^1(M, \mathcal{O}_M(GL(n, \mathbb{R}))).$$

It is clear that any rank n vector bundle is associated to a principal $GL(n, \mathbb{R})$-bundle by means of the identity representation id. Now let G be any Lie group. We say that E is a G-bundle if E is isomorphic to a vector bundle associated with a principal G-bundle by means of some linear representation $G \to GL(n, \mathbb{R})$. In other words, the structure group of E can be reduced to a subgroup $\rho(G)$ where $\rho : G \to GL(n, \mathbb{R}) \subset \text{Diff}(\mathbb{R}^n)$ is a faithful smooth (linear) action. It follows from the above discussion that

$$\{\text{rank } n \text{ vector } G\text{-bundles/isomorphism}\} \longleftrightarrow H^1(M, \mathcal{O}_M(G)) \times Rep_n(G).$$

where $\text{Rep}_n(G)$ stands for the set of equivalence classes of faithful linear representations $G \to GL(n, \mathbb{R})$ ($\rho \sim \rho'$ if there exists $A \in GL(n, \mathbb{R})$ such that $\rho(g) = A\rho'(g)A^{-1}$ for all $g \in G$).

For example, we have the notion of an orthogonal vector bundle, or a unitary vector bundle.

Example 1. The tangent bundle $T(M)$ is defined by transition functions g_{ij} equal to the Jacobian matrices $J = (\frac{\partial x_\alpha}{\partial y_\beta})$, where $(x_1, \ldots, x_n), (y_1, \ldots, y_n)$ are local coordinate functions in U_i and U_j, respectively. The choice of local coordinates defines the trivialization $\phi_i : U_i \times \mathbb{R}^n \to \pi^{-1}(U_i)$ by sending $(x, (a_1, \ldots, a_n))$ to $(x, \sum_{i=1}^n a_i \frac{\partial}{\partial x_i})$. A section of the tangent bundle is a vector field on M. It is defined locally by $\eta = \sum_{i=1}^n a_i(x) \frac{\partial}{\partial x_i}$, where $a_i(x)$ are smooth functions on U_i. If $\eta = \sum_{i=1}^n b_i(y) \frac{\partial}{\partial y_i}$ represents η in U_j, we have

$$a_j(y) = \sum_{i=1}^n b_i(x) \frac{\partial x_j}{\partial y_i}.$$

One can prove (or take it as a definition) that the structure group of the tangent bundle can be reduced to $SL(n, \mathbb{R})$ if and only if M is orientable. Also, if M is a symplectic manifold of dimension $2n$, then using Darboux's theorem one checks that the structure group of $T(M)$ can be reduced to the symplectic group $Sp(2n, \mathbb{R})$.

Let $F(M)$ be the principal $GL(n, \mathbb{R})$-bundle over M with transition functions defined by the Jacobian matrices J. In other words, the tangent bundle $T(M)$ is associated to $F(M)$. Let $e_1, \ldots, e_n$ be the standard basis in $\mathbb{R}^n$. The correspondence $A \to (Ae_1, \ldots, Ae_n)$ is a bijective map from $GL(n, \mathbb{R})$ to the set of bases (or frames) in $\mathbb{R}^n$. It allows one to interpret the principal bundle $F(M)$ as the bundle of frames in the tangent spaces of M.

Example 2. Let W be a complex vector space of dimension n and $W_{\mathbb{R}}$ be the corresponding real vector space. We have $GL_{\mathbb{C}}(W) = \{g \in GL(W_{\mathbb{R}}) : g \circ J = J \circ g\}$, where $J : W \to W$ is the operator of multiplication by $i = \sqrt{-1}$. In coordinate form, if $W = \mathbb{C}^n, W_{\mathbb{R}} = \mathbb{R}^{2n} = \mathbb{R}^n + i(\mathbb{R}^n)$, $GL(W) = GL(n, \mathbb{C})$ can be identified with the subgroup of $GL(W_{\mathbb{R}}) = GL(2n, \mathbb{R})$ consisting of invertible matrices of the form $\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, where $A, B \in Mat_n(\mathbb{R})$. The identification map is

$$X = A + iB \in GL(n, \mathbb{C}) \to \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in GL(2n, \mathbb{R}).$$

One defines a rank n *complex vector bundle* over M as a real rank $2n$ vector bundle whose structure group can be reduced to $GL(n, \mathbb{C})$. Its fibres acquire a natural structure of a complex n-dimensional vector space, and its space of sections is a complex vector space.

Suppose M can be equipped with a structure of an n-dimensional complex manifold. Then at each point $x \in M$ we have a system of local complex coordinates $z_1, \ldots, z_n$. Let $x_i = (z_i + \bar{z}_i)/2, y_i = (z_i - \bar{z}_i)/2i$ be a system of local real parameters. Let $z'_1, \ldots, z'_n$ be another system of local complex parameters, and $(x'_i, y'_i), i = 1, \ldots, n$, be the corresponding system of real parameters. Since $(z'_1, \ldots, z'_n) = (f_1(z_1, \ldots, z_n), \ldots, f_n(z_1, \ldots, z_n))$ is a holomorphic coordinate change, the functions $f_i(z_1, \ldots, z_n)$ satisfy the Cauchy-Riemann conditions

$$\frac{\partial x'_i}{\partial x_j} = \frac{\partial y'_i}{\partial y_j}, \quad \frac{\partial x'_i}{\partial y_j} = -\frac{\partial y'_i}{\partial x_j},$$

and the Jacobian matrix has the form

$$\frac{\partial(x'_1, \ldots, x'_n, y'_1, \ldots, y'_n)}{\partial(x_1, \ldots, x_n, y_1, \ldots, y_n)} = \begin{pmatrix} A & B \\ -B & A \end{pmatrix},$$

where

$$A = \frac{\partial(x'_1, \dots, x'_n)}{\partial(x_1, \dots, x_n)}, \quad B = \frac{\partial(x'_1, \dots, x'_n)}{\partial(y_1, \dots, y_n)}.$$

This shows that $T(M)$ can be equipped with the structure of a complex vector bundle of dimension n.

11.5 Let E be a rank n vector bundle and $P \to M$ be the corresponding principal $GL(n, \mathbb{R})$-bundle. For any representation $\rho : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ we can construct the associated vector bundle $E(\rho)$ with typical fibre $\mathbb{R}^n$ and the structure group $\rho(G) \subset GL(n, \mathbb{R})$. This allows us to define many operations on bundles. For example, we can define the exterior product $\bigwedge^k(E)$ as the bundle $E(\rho)$, where

$$\rho : GL(n, \mathbb{R}) \to GL(\bigwedge^k(\mathbb{R}^n)), A \to \bigwedge^k A.$$

Similarly, we can define the tensor power $E^{\otimes k}$, the symmetric power $S^k(E)$, and the dual bundle E^*. Also we can define the tensor product $E \otimes E'$ (resp. direct sum $E \oplus E'$) of two vector bundles. Its typical fibre is $\mathbb{R}^{nm}$ (resp. $\mathbb{R}^{n+m}$), where $n = \text{rank } E, m = \text{rank } E'$. Its transition functions are the tensor products (resp. direct sums) of the transition matrix functions of E and E'.

The vector bundle $E^* \otimes E$ is denoted by $\text{End}(E)$ and can be viewed as the bundle of endomorphisms of E. Its fibre over a point $x \in M$ is isomorphic (under a trivialization map) to $V^* \otimes V = \text{End}(V, V)$. Its sections are morphisms of vector bundles $E \to E$.

Example 3. Let $E = T(M)$ be the tangent bundle. Set

$$T^p_q(M) = T(M)^{\otimes p} \otimes T^*(M)^{\otimes q}. \tag{11.8}$$

A section of $T^p_q(M)$ is called a p-covariant and q-contravariant *tensor* over M (or just a tensor of type (p, q)). We have encountered them earlier in the lectures. Tensors of type $(0, 0)$ are scalar functions $f : M \to \mathbb{R}$. Tensors of type $(1, 0)$ are vector fields. Tensors of type $(0, 1)$ are differential 1-forms. They are locally represented in the form

$$\theta = \sum_{i=1}^n a_i(x) dx_i,$$

where $dx_j(\frac{\partial}{\partial x_i}) = \delta_{ij}$ and $a_i(x)$ are smooth functions in U_i. If $\theta = \sum_{i=1}^n b_i(y) dy_i$ represents θ in U_j, then

$$a_j(x) = \sum_{i=1}^n \frac{\partial y_i}{\partial x_j} b_i(y).$$

In general, a tensor of type (p, q) can be locally given by

$$\tau = \sum_{i_1, \dots, i_p = 1}^n \sum_{j_1, \dots, j_q = 1}^n a^{i_1 \dots i_p}_{j_1 \dots j_q} \frac{\partial}{\partial x_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x_{i_p}} \otimes dx_{j_1} \otimes \dots \otimes dx_{j_q}.$$

So such a tensor is determined by collections of smooth functions $a^{i_1\dots i_p}_{j_1\dots j_q}$ assigned to an open subset U_i with local cooordinates $x_1,\dots,x_n$. If the same tensor is defined by a collection $b^{i_1\dots i_p}_{j_1\dots j_q}$ in an open subset U_j with local coordinates $y_1,\dots,y_n$, we have

$$a^{i_1\dots i_p}_{j_1\dots j_q} = \frac{\partial x_{i_1}}{\partial y_{i'_1}}\cdots\frac{\partial x_{i_p}}{\partial y_{i'_p}}\frac{\partial y_{j'_1}}{\partial x_{j_1}}\cdots\frac{\partial y_{j'_q}}{\partial x_{j_q}}b^{i'_1\dots i'_p}_{j'_1\dots j'_q},$$

where we use the "Einstein convention" for summation over the indices $i'_1,\dots,i'_p,j'_1,\dots,j'_q$.

A section of $\bigwedge^k T^*(M)$ is a smooth differential k-form on M. We can view it as an anti-symmetric tensor of type $(0,k)$. It is defined locally by $\sum_{i_1,\dots,i_k=1}^n a_{i_1\dots i_k}dx_{i_1}\otimes\dots\otimes dx_{i_k}$, where

$$a_{i_{\sigma(1)}\dots i_{\sigma(k)}} = \epsilon(\sigma)a_{i_1\dots i_k}$$

where $\epsilon(\sigma)$ is the sign of the permutation σ. This can be rewritten in the form

$$\sum_{1\le i_1<\dots<i_k\le n} a_{i_1\dots i_k}dx_{i_1}\wedge\dots\wedge dx_{i_k},$$

where

$$dx_{i_1}\wedge\dots\wedge dx_{i_k} = \sum_{1\le i_1<\dots<i_k\le n}\epsilon(\sigma)dx_{i_{\sigma(1)}}\otimes\dots\otimes dx_{i_{\sigma(k)}}.$$

In particular, $\bigwedge^n(T^*(M))$ is a rank 1 vector bundle. It is called the *volume bundle*. Its sections over a trivializing open set U_i can be identified with scalar functions $a_{12\dots n}$. Over $U_i\cap U_j$ these functions transform according to the formula

$$a_{12\dots n}(x) = \left|\frac{\partial(y_1,\dots,y_n)}{\partial(x_1,\dots,x_n)}\right|b_{12\dots n}(y(x)).$$

A manifold M is *orientable* if one can find a section of its volume bundle which takes positive values at any point. Such a section is called a *volume form*. We shall always assume that our manifolds are orientable. Obviously, any two volume forms on M are obtained from each other by multiplying with a global scalar function $f: M\to\mathbb{R}_{>0}$. A volume form Ω defines a distribution on M, i.e., a continuous linear functional on the space $C_0^\infty(M)$ of smooth functions on M with compact support. Its values on $\phi\in C_0^\infty(M)$ is the integral $\int_M\phi(x)\Omega$. For example, if $M=\mathbb{R}^n$, we may take $\Omega = dx_1\wedge\dots\wedge dx_n$. In this case, physicists use the notation

$$\int_{\mathbb{R}^n}\phi(x)dx_1\wedge\dots\wedge dx_n = \int d^nx\phi(x).$$

Example 4. Recall that a structure of a *pseudo-Riemannian manifold* on a connected smooth manifold M is defined by a smooth function on $Q: T(M)\to\mathbb{R}$ such that its restriction to each fibre $T(M)_x$ is a non-degenerate quadratic form. The signature of $Q|T(M)_x$ is independent of x. It is called the *signature* of the pseudo-Riemannian manifold

(M,Q). We say that (M,Q) is a Riemannian manifold if the signature is $(n,0)$, or in other words, $Q_x = Q|T(M)_x$ is positive definite. We say that (M,Q) is *Lorentzian* (or *hyperbolic*) if the signature is of type $(n-1,1)$ (or $(1,n-1)$). For example, the space-time $\mathbb{R}^4$ is a Lorentzian pseudo-Riemannian manifold. One can define a pseudo-metric Q by the corresponding symmetric bilinear form $B : T(M) \times T(M) \to \mathbb{R}$. It can be viewed as a section of $T^*(M) \otimes T^*(M)$. It is locally defined by

$$B(x) = \sum_{i,j=1}^{n} a_{ij}(x) dx_i \otimes dx_j, \tag{11.9}$$

where $a_{ij}(x) = a_{ji}(x)$. The associated quadratic form is given by

$$Q_x = \sum_{i=1}^{n} a_{ii} dx_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}(x) dx_i dx_j.$$

Its value on a tangent vector $\eta_x = \sum_{i=1}^{n} c_i \frac{\partial}{\partial x_i} \in T(M)_x$ is equal to

$$Q_x(\eta_x) = \sum_{i=1}^{n} a_{ii} c_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij}(x) c_i c_j.$$

Using a pseudo-Riemannian metric one can define a subbundle of the frame bundle of $T(M)$ whose fibre over a point $x \in M$ consists of orthonormal frames in $T(M)_x$. This allows us to reduce the structure of the frame bundle to the orthogonal group $O(k,n-k)$ of the quadratic form $x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_n^2$. The structure group of the tangent bundle cannot be reduced to this subgroup unless the curvature of the metric is zero (see below).

11.6 The group of automorphisms of a fibre G-bundle P is called the *gauge group* of P and is denoted by $\mathcal{G}(P)$.

Theorem 2. *Let P be a principal G-bundle and let P^c be the associated bundle with typical fibre G and representation $G \to Aut(G)$ given by the conjugation action $g \cdot h = g \cdot h \cdot g^{-1}$ of G on itself. Then*

$$\mathcal{G}(P) \cong \Gamma(P^c).$$

Proof. This is obviously true if P is trivial. In this case the both groups are isomorphic to the group of smooth functions $M \to G$. Let $(U_i)_{i \in I}$ be a trivializing cover of P with trivializing diffeomorphims $\phi_i : U_i \times G \to P|U_i$. Any element $g \in \mathcal{G}(P)$ defines, after restrictions, the automorphisms $g_i : P|U_i \to P|U_i$ and hence the automorphisms $s_i = \phi_i^{-1} \circ g \circ \phi_i$ of $U_i \times G$. By the above, s_i is a section of $\mathcal{O}_M(G)(U_i)$. If $\psi_j : V_j \times G \to P|V_j$ is another set of automorphisms, and $s'_j \in \mathcal{O}_M(V_j)$ is the corresponding automorphism of $V_j \times G$, then $s'_j = \psi_j^{-1} \circ g \circ \psi_j$. Comparing these two sections over $U_i \cap U_j$, we get

$$s_i = (\phi_i^{-1} \circ \psi_j) \circ s'_j \circ (\psi_j^{-1} \circ \psi_i) = g_{ij}^{-1} \circ s'_j \circ g_{ij}.$$

This shows that g defines a section of P^c. Conversely, every section (s_i) of P^c defines automorphisms of trivialized bundles $P|U_i$ which agree on the intersections.

Let $S \to M$ be a fibre G-bundle associated to a principal G-bundle P. Let $\rho : G \to \mathrm{Diff}(F)$ be the corresponding representation. Then it defines an isomomorphism of the gauge groups $\mathcal{G}(P) \to \mathcal{G}(S)$. Locally it is given by $g(x) \to \rho(g(x))$.

Exercises.

1. Find a non-trivial vector bundle E over circle S^1 such that $E \oplus E$ is trivial.

2. Let $\mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{P}^n(\mathbb{C})$ be the quotient map from the definition of projective space. Show that it is a principal $\mathbb{C}^*$-bundle over $\mathbb{P}^n(\mathbb{C})$. It is called the Hopf bundle (in the case $n = 1$ its restriction to the sphere $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ is the usual Hopf bundle $S^3 \to S^2$). Find its transition functions with respect to the standard open cover of projective space.

3. A manifold M is called *parallelizable* if its tangent bundle is trivial. Show that S^1 is parallelizable (a much deeper result due to Milnor and Kervaire is that S^1, S^3 and S^7 are the only parallelizable spheres).

4. Show that a rank n vector bundle is trivial if and only if it admits n sections whose values at any point are linearly independent.

5. Find transition functions of the tangent bundle to $\mathbb{P}^n(\mathbb{C})$. Show that its direct sum with the trivial rank 2 vector bundle is isomorphic to $L^{\oplus n+1}$, where L is the rank 2 vector bundle associated to the Hopf bundle by means of the standard action of $\mathbb{C}^*$ in $\mathbb{R}^2$.

6. Prove that any fibre $\mathbb{R}$-bundle over a compact manifold is trivial.

7. Using the previous problem prove that the structure group of any fibre $\mathbb{C}^*$-bundle (resp. $\mathbb{R}^*$) can be reduced to $U(1)$ (resp. $\mathbb{Z}/2\mathbb{Z}$).

8. Show that any principal G-bundle with finite group G is isomorphic to a covering space $N \to M$ with the group of deck transformations isomorphic to G. Using the previous two problems, prove that any line bundle over a simply-connected compact manifold is trivial.

Lecture 12. GAUGE FIELDS

In this lecture we shall define connection (or *gauge field*) on a vector bundle.

12.1 First let us fix some notation from the theory of Lie groups. Recall that a *Lie group* is a group object G in the category of smooth manifolds. For any $g \in G$ we denote by L_g (resp. R_g) the left (resp. right) translation action of the group G on itself. The differentials of the maps R_g, L_g transform a vector field η to the vector field $(L_g)_*(\eta), (R_g)_*(\eta)$, respectively. The *Lie algebra* of G is the vector space $\mathfrak{g}$ of vector fields which are invariant with respect to left translations. Its Lie algebra structure is the bracket operation of vector fields. For any $\eta \in \mathfrak{g}$, the vector $(dL_{g^{-1}})_g(\eta_g)$ belongs to the tangent space $T(G)_1$ of G at the identity element 1. It is independent of $g \in G$. This defines a bijective linear map $\omega_G : \mathrm{Lie}(G) \to T(G)_1$. By transferring the Lie bracket, we may identify the Lie algebra of G with the tangent space of G at the identity element $1 \in G$. Let $v \in T(G)_1$ and $\tilde{v}$ be the corresponding vector field. We have $\tilde{v}_g = (dL_g)_1(v)$, so that $(dR_{g^{-1}})_g(\tilde{v}_g) = dc(g)_1(v)$, where $c(g) : G \to G$ is the conjugacy action $h \to g \cdot h \cdot g^{-1}$. The action $v \to dc(g)_1(v)$ of G on $T(G)_1$ is called the *adjoint representation* of G and is denoted by $Ad : G \to GL(\mathfrak{g})$. We have for any $v \in T(G)_1$,

$$(R_g)_*(\tilde{v}) = Ad(g^{-1})(v). \tag{12.1}$$

From now on we shall identify vectors $v \in T(G)_1$ with left-invariant vector fields $\tilde{v}$ on G. The adjoint representation preserves the Lie bracket, i.e.,

$$Ad(g)([\eta, \tau]) = [Ad(g)(\eta), Ad(g)(\tau)]. \tag{12.2}$$

Let $\eta \in \mathfrak{g}$ and $G \to \mathfrak{g}$ be the map $g \to Ad(g)(\eta)$. Its differential at $1 \in G$ is the linear map $\mathrm{Lie}(G) \to \mathfrak{g}$. It coincides with the map $\tau \to [\eta, \tau]$. The latter is a homomorphism of the Lie algebra to itself. It is also called the *adjoint representation* (of the Lie algebra) and is denoted by $ad : \mathfrak{g} \to \mathfrak{g}$.

Most of the Lie groups we shall deal with will be various subgroups of the Lie group $G = GL(n, \mathbb{R})$. Since $GL(n, \mathbb{R})$ is an open submanifold of the vector space $\mathrm{Mat}_n(\mathbb{R})$ of real $n \times n$-matrices, we can identify $T(G)_1$ with $\mathrm{Mat}_n(\mathbb{R})$. The matrix functions $x_{ij} : A =$

$(a_{ij}) \to a_{ij}$ are a system of local parameters. We can write any vector field on $\mathrm{Mat}_n(\mathbb{R})$ in the form

$$\eta = \sum_{i,j=1}^{n} \phi_{ij}(X) \frac{\partial}{\partial x_{ij}},$$

where $\phi_{ij}(X)$ are some smooth functions on $\mathrm{Mat}_n(\mathbb{R})$. The left translation map $L_A : X \to A \cdot X$ is a linear map on $\mathrm{Mat}_n(\mathbb{R})$, so it coincides with its differential. It maps $\frac{\partial}{\partial x_{ij}}$ to $\sum_{k=1}^{n} a_{ik} \frac{\partial}{\partial x_{kj}}$, where a_{ij} is the ij-th entry of A. From this we easily get

$$(L_A)_*(\eta) = \sum_{i,j=1}^{n} \phi_{ij}(A \cdot X)(\sum_{k=1}^{n} a_{ik} \frac{\partial}{\partial x_{kj}}).$$

If we assign to η the matrix function $\Phi(X) = (\phi_{ij}(X))$, then we can rewrite the previous equation in the form

$$(L_A)_*(\Phi)(X) = \Phi(AX)A^{-1}.$$

In particular, η is left-invariant, if and only if $\Phi(AX) = \Phi(X)A$ for all $A, X \in G$. By taking $X = E_n$, we obtain that $\Phi(A) = A\Phi(I_n)$. This implies that $\eta \to X_\eta = \Phi(I_n) \in \mathrm{Mat}_n(\mathbb{R})$ is a bijective linear map from $\mathrm{Lie}(G)$ to $\mathrm{Mat}_n(\mathbb{R})$. Also we verify that

$$[\eta, \tau] = [X_\eta, X_\tau] := X_\eta X_\tau - X_\tau X_\eta.$$

The Lie algebra of $GL(n, \mathbb{R})$ is denoted by $\mathfrak{gl}(n, \mathbb{R})$. We list some standard subgroups of $GL(n, \mathbb{R})$. They are

$$SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) : \det(A) = 1\}$$

(*special linear group*).

Its Lie algebra is the subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices with trace zero. It is denoted by $\mathfrak{sl}(n, \mathbb{R})$.

$$O(n-k, k) = \{A \in GL(n, \mathbb{R}) : A^t \cdot J_{n-k,k} \cdot A = J_{n-k,k} := \begin{pmatrix} I_{n-k} & 0 \\ 0 & -I_k \end{pmatrix}\}$$

(*orthogonal group* of type $(n-k, k)$).

Its Lie algebra is the Lie subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices A satisfying $A^t J_{n-k,k} + J_{n-k,k} A^t = 0$. It is denoted by $\mathfrak{o}(n-k, k)$.

$$Sp(2n, \mathbb{R}) = \{A \in GL(2n, \mathbb{R}) : A^t \cdot J \cdot A = J := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}\}$$

(*symplectic group*).

Its Lie algebra is the Lie subalgebra of $\mathrm{Mat}_n(\mathbb{R})$ of matrices A such that $A^t \cdot J + J \cdot A = 0$. It is denoted by $\mathfrak{sp}(2n, \mathbb{R})$.

$$GL(n, \mathbb{C}) = \{A = \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix} \in GL(2n, \mathbb{R}), X, Y \in \mathrm{Mat}_n(\mathbb{R})\}$$

(*complex general linear group*).

Its Lie algebra is isomorphic to the Lie algebra of complex $(n \times n)$-matrices and is denoted by $\mathfrak{gl}(n, \mathbb{C})$.

$$U(n) = \{A \in GL(n, \mathbb{C}) : A^t \cdot A = I_{2n}\}$$

(*unitary group*).

Its Lie algebra consists of skew-symmetric matrices A from $\mathfrak{gl}(n, \mathbb{C})$). It is denoted by $\mathfrak{u}(n)$.

Other groups are realized as subgroups of $GL(n, \mathbb{R})$ by means of a morphism of Lie groups $G \to GL(n, \mathbb{R})$ which are called linear representations. For example, $GL(n, \mathbb{C})$ is isomorphic to the Lie group of complex invertible $n \times n$-matrices $X + iY$ under the representation

$$X + iY \to \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}.$$

Under this map, the unitary group $U(n)$ is isomorphic to the group of complex invertible $n \times n$-matrices A satisfying $A \cdot \bar{A}^t = I_n$, where $\overline{X + iY} = X - iY$.

12.2 We start with a principal G-bundle $P \to M$. For any point $p = (x, g) \in P$, the tangent space $T(P)_p$ contains $T(P_x)_p$ as a linear subspace. Its elements are called *vertical tangent vectors*. We denote the subspace of vertical vectors $T(P_x)_p$ by $T(P)_p^v$.

A *connection* on P is a choice of a subspace $H_p \subset T(P)_p$ for each $p \in P$ such that

$$T(P)_p = H_p \oplus T(P)_p^v.$$

This choice must satisfy the following two properties:

(i) (smoothness) the map $T(P) \to T(P)$ given by the projection maps $T(P)_p \to T(P)_p^v \subset T(P)_p$ is smooth;

(ii) (equivariance) for any $g \in G$, considered as an automorphism of P, we have

$$(dg)_p(H_p) = H_{p \cdot g}.$$

Now observe that P_x is isomorphic to the group G although not canonically (up to left translation isomorphism of G). To any tangent vector $\xi \in T(P_x)_p$, we can associate a left invariant vector field on G. This is independent of the choice of an isomorphism $P_x \cong G$. This allows one to define an isomorphism

$$\alpha_p : T(P)_p^v \to \mathfrak{g}, \tag{12.3}$$

where $\mathfrak{g}$ is the Lie algebra of G. If $p' = g \cdot p$, then

$$\alpha_{p'} = Ad(g^{-1}) \circ \alpha_p.$$

Let $\mathfrak{g}_P = P \times \mathfrak{g}$ be the trivial vector bundle with fibre $\mathfrak{g}$. Then the projection maps $T(P)_p \to T(P)_p^v$ define a linear map of vector bundles $T(P) \to \mathfrak{g}_P$. This is equivalent to a section of the bundle $T(P)^* \otimes \mathfrak{g}_P$ and can be viewed as a differential 1-form A with values in the Lie algebra $\mathfrak{g}$. Its restriction to the vertical space $T(P)_p$ coincides with the isomorphism α_p. It is zero on the horizontal part H_p of $T(P)_p$. From this we deduce that the form A satisfies

$$A(dg_p(\xi_p)) = Ad(g^{-1})A(\xi_p), \quad \forall \xi_p \in T(P)_p, \forall g \in G. \tag{12.4}$$

Conversely, given $A \in \Gamma(T(P)^* \otimes \mathfrak{g}_P)$ which satisfies (12.4) and whose restriction to $T(P)_p^v$ is equal to the map $\alpha_p : T(P)_p^v \to \mathfrak{g}$, it defines a connection. We just set

$$H_p = \mathrm{Ker}(A_p).$$

Let $\{H_p\}_{p \in P}$ be a connection on P. Under the projection map $\pi : P \to M$, the differential map $d\pi$ maps H_p isomorphically onto the tangent space $T(M)_{\pi(p)}$. Thus we can view the connection as a linear map $A : \pi^*(T(M)) \to T(P)$ such that its composition with the differential map $T(P) \to \pi^*(T(M))$ is the identity. Now given a section ξ of $T(M)$ (a vector field), it defines a section of $\pi^*(T(M))$ by the formula $\pi^*(\xi)(p) = \xi(\pi(p))$, where we identify $\pi^*(T(M))_p$ with $T(M)_{\pi(p)}$. Using A we get a section $\xi^* = A(\pi^*(\xi))$ of $T(P)$. For any $p \in P$,

$$\xi_p^* \in H_p.$$

The latter is expressed by saying that ξ^* is a *horizontal vector field.*

Thus a connection allows us to lift any vector field on M to a unique horizontal vector field on P. In particular one can lift an integral curve $\gamma : [0,1] \to M$ of vector fields to a horizontal path $\gamma^* : [0,1] \to P$. The latter means that $\pi \circ \gamma^* = \gamma$ and the velocity vectors $\frac{d\gamma^*}{dt}$ are horizontal. If we additionally fix the initial condition $\gamma^*(0) = p \in P_{\gamma(0)}$, this path will be unique. Let us prove the existence of such a lifting for any path $\gamma : [0,1] \to M$.

First, we choose some curve $p(t) (0 \le t \le 1)$ in P lying over $\gamma(t)$. This is easy, since P is locally trivial. Now we shall try to correct it by applying to each $p(t)$ an element $g(t) \in G$ which acts on fibres of P by right translation. Consider the map $t \to q(t) = p(t) \cdot g(t)$ as the composition $[0,1] \to P \times G \to P, t \to (p(t), g(t)) \to p(t) \cdot g(t)$. Let us compute its differential. For any fixed $(p_0, g_0) \in P \times G$, we have the maps: $\mu_1 : P \to P, p \to p \cdot g_0$, and $\mu_2 : G \to P, g \to p_0 \cdot g$. The restriction of the map $\mu : P \times G \to P$ to $P \times \{g_0\}$ (resp. $\{p_0\} \times G$) coincides with μ_1 (resp. μ_2). This easily implies that

$$dq(t) = dp(t) \cdot g(t) + p(t) \cdot dg(t) = dp(t) \cdot g(t) + q(t) \cdot g(t)^{-1} \cdot dg(t). \tag{12.5}$$

Here the left-hand side is a tangent vector at the point $q(t)$, the term $dp(t) \cdot g(t)$ is the translate of the tangent vector $dp(t)$ at $p(t)$ by the right translation $R_{g(t)}$, the term $q(t) \cdot g(t)^{-1} \cdot dg(t)$ is the vertical tangent vector at $q(t)$ corresponding to the element $g(t)^{-1}dg(t)$ of the Lie algebra $\mathfrak{g}$ under the isomorphism $\alpha_{q(t)}$ from (12.3). Using (12.5), we are looking for a path $g(t) : [0,1] \to G$ satisfying

$$A_{q(t)}(dq(t)) = Ad(g(t)^{-1})A_{p(t)}(dp(t)) + g(t)^{-1}dg(t) = 0.$$

This is equivalent to the following differential equation on G:

$$A_{p(t)}(dp(t)) = -Ad(g(t))(g(t)^{-1}dg(t)) = -g(t)^{-1}dg(t).$$

This can be solved.

Thus the connection allows us to move from the fibre $P_{\gamma(0)}$ to the fibre $P_{\gamma(t)}$ along the horizontal lift of a path connecting the points $\gamma(0)$ and $\gamma(t)$.

12.3 There are other nice things that a connection can do. First let introduce some notation. For any vector bundle E over a manifold X we set

$$\Lambda^k(E) = \bigwedge^k T^*(X) \otimes E,$$

$$\mathcal{A}^k(E)(X) = \Gamma(\Lambda^k(E))$$

The elements of this space are called differential k-forms on X with values in E. In the special case when E is the trivial bundle of rank 1, we drop E from the notation. Thus

$$\mathcal{A}^k(X) = \Gamma(\bigwedge^k T^*(X))$$

is the space of smooth differential k-forms on X. Also, if $E = V_X$ is the trivial vector bundle with fibre V, then we set

$$\mathcal{A}^k(V)(X) = \mathcal{A}^k(V_X)(X).$$

Thus a connection A on a principal G-bundle is an element of the space $\mathcal{A}^1(L(G))(P)$.

Let us define the *covariant derivative*. Let P be a principal G-bundle over M. For any differential k-form ω on P with values in a vector space V (i.e. $\omega \in \mathcal{A}^k(V)(P)$) we can define its covariant derivative $d^A(\omega)$ with respect to connection A by

$$d^A(\omega)(\tau_1, \dots, \tau_{k+1}) = d\omega(\tau_1^h, \dots, \tau_{k+1}^h), \tag{12.6}$$

where the superscript is used to denote the horizontal part of the vector field. Here we also use the usual operation of exterior derivative satisfying the following *Cartan's formula*:

$$d\omega(\xi_1 \dots, \xi_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i+1} \xi_i(\omega(\xi_1, \dots, \hat{\xi}_i, \dots, \xi_{k+1})) +$$

$$+ \sum_{1 \leq i < j \leq k+1} (-1)^{i+j} \omega([\xi_i, \xi_j]), \xi_1, \dots, \hat{\xi}_i, \dots, \hat{\xi}_j, \dots, \xi_{k+1}). \tag{12.7}$$

In the first part of this formula we view vector fields as derivations of functions.

We set

$$F_A = d^A(A), \tag{12.8}$$

where the connection A is viewed as an element of the space $\mathcal{A}^1(L(G))(P)$. This is a differential 2-form on P with values in $\mathfrak{g}$. It is called the *curvature form* of A. By definition

$$F_A(\xi_1, \xi_2) = \xi_1^h(A(\xi_2^h)) - \xi_1^h(A(\xi_1^h)) - A([\xi_1^h, \xi_2^h]) = -A([\xi_1^h, \xi_2^h]). \tag{12.9}$$

Here we have used that the value of A on a horizontal vector is zero. This gives a geometric interpretation of the curvature form we alluded to in the introduction.

In particular,

$$F_A = 0 \Longleftrightarrow [\xi, \tau] \quad \text{is horizontal for any horizontal vector fields } \xi, \tau.$$

Assume $\omega \in \mathcal{A}^p(L(G))(P)$ satisfies the following two conditions:
(a) $\omega(\xi_1, \ldots, \xi_k) = 0$ if at least one of the vector fields ξ_i^h is vertical.
(b) for any $g \in G$ acting on P by right translatons, $(R_g)^*(\omega)(\xi) = Ad(g^{-1})(\omega(\xi))$.

For example, the curvature form F_A satisfies these properties. If A, A' are two connections, then the difference $A - A'$ satisfies these properties (since their restrictions to vertical vector fields coincide).

Let us try to descend such ω to M by setting

$$\omega(\xi_1, \ldots, \xi_k)(x) = \omega(ds_U(\xi_1), \ldots, ds_U(\xi_k)), \tag{12.10}$$

where we choose a local section $s_U : U \to P$. A different choice of trivialization gives $s_U = s_V \cdot g_{UV}$, where g_{UV} is the transition function of P (see (11.3)). Differentiating, we have (similar to (12.5))

$$ds_U = ds_V \cdot g_{UV} + s_U \cdot g_{UV}^{-1} dg_{UV}. \tag{12.11}$$

Recall that, for any $\tau_x \in T(M)_x$, $ds_V \cdot g_{UV}(\tau_x)$ equals the image of the tangent vector $ds_V(\tau_x) \in T(P)_{s_V(x)}$ under the differential of the right translation $g_{UV} : P \to P$. The second term $s_U \cdot g_{UV}^{-1} dg_{UV}(\tau_x)$ is the image of $(dg_{UV})_x(\tau_x) \in T(G)_{g_{UV}(x)}$ in $T(P_{s_U(x)})^v$ under the isomorphism $T(G)_{g_{UV}(x)} \cong \mathfrak{g} \cong T(P_x)_{s_U(x)}$. Since ω vanishes on vertical vectors and also satisfies property (b), we obtain that the right-hand side of (12.10) changes, when we replace s_U with s_V, by applying to the value the transformation $Ad(g_{UV}^{-1})$. Thus, if we introduce the vector bundle $Ad(P)$ associated to P by means of the adjoint representation of G in $\mathfrak{g}$, the left-hand side of (12.11) is a well-defined section of the bundle $\mathcal{A}^1(Ad(P))$.

The covariant derivative operator d^A obviously preserves properties (a) and (b), hence descends to an operator

$$d^A : \mathcal{A}^k(Ad(P))(M) \to \mathcal{A}^{k+1}(Ad(P))(M).$$

Also its definition is local, so we obtain a homomorphism of sheaves

$$d^A : \mathcal{A}^k(Ad(P)) \to \mathcal{A}^{k+1}(Ad(P)). \tag{12.12}$$

Taking $k = 0$, we get a homomorphism

$$\nabla_A : Ad(P) \to \mathcal{A}^1(Ad(P)). \tag{12.13}$$

It satisfies Leibnitz's rule: for any section $s \in Ad(P)(U)$, and a function $f : U \to \mathbb{R}$,

$$\nabla_A(fs) = df \otimes s + f\nabla_A(s). \tag{12.14}$$

12.4 Let us consider the natural pairing

$$\mathcal{A}^k(\mathfrak{g})(P) \times \mathcal{A}^k(\mathfrak{g})(P) \to \mathcal{A}^{k+s}(\mathfrak{g})(P), \quad (\omega, \omega') \to [\omega, \omega'] \tag{12.15}$$

It is defined by writing any $\omega \in \mathcal{A}^k(\mathfrak{g})(P)$ as a sum of the expressions $\alpha \otimes \xi$, where $\alpha \in \mathcal{A}^k(X)$ and $\xi \in \mathfrak{g}$, This allows us to define the pairing and to set

$$[\alpha \otimes \xi, \alpha' \otimes \xi'] = (\alpha \wedge \alpha') \otimes [\xi, \xi'].$$

We leave to the reader to check the following:

Lemma 1. *For any* $\phi \in \Lambda^p(X, \mathfrak{g}), \psi \in \Lambda^q(X, \mathfrak{g}), \rho \in \Lambda^r(X, \mathfrak{g})$,
(i) $d[\phi, \beta] = [d\phi, \psi] + (-1)^p[\phi, d\psi]$,
(ii) $(-1)^{pr}[[\phi, \psi], \rho] + (-1)^{rq}[[\rho, \phi], \psi] + (-1)^{qp}[[\psi, \rho], \phi] = 0$,
(iii) $[\psi, \phi] = -(-1)^{pq}[\phi, \psi]$.

Let us prove the following

Theorem 1. *Let* $A \in \Lambda^1(P, \mathfrak{g})$ *be a connection on* P *and* F_A *be its curvarure form.*
(i) (The structure equation):

$$F_A = dA + \frac{1}{2}[A, A].$$

(ii) (Bianchi's identity):

$$d^A(F_A) = 0.$$

(iii) for any $\beta \in \mathcal{A}^k(Ad(P))(M)$

$$d^A(\beta) = d\beta + [A, \beta].$$

Proof. (i) We want to compare the values of each side on a pair of vector fields ξ, τ. By linearity, it is enough to consider three cases: ξ, τ are horizontal; ξ is vertical, τ is horizontal; both fields are vertical. In the first case we have, applying (12.7),

$$(dA + \frac{1}{2}[A, A])(\xi, \tau) = dA(\xi^h, \tau^h) + \frac{1}{2}([A(\xi^h), A(\tau^h)] - [A(\tau^h), A(\xi^h)]) =$$

$$= \xi^h(A(\tau^h) - \tau^h A(\xi^h) - A([\xi^h, \tau^h]) = -A([\xi^h, \tau^h]) = F_A(\xi, \tau).$$

In the second case,

$$(dA + \frac{1}{2}[A, A])(\xi, \tau)) = \xi^v(A(\tau^h)) - \tau^h(A(\xi^v)) - A([\xi^v, \tau^h]) +$$

$$+\frac{1}{2}([A(\xi^v), A(\tau^h)] - [A(\tau^h), A(\xi^v)]) = -\tau^h A(\xi^v) - A([\xi^v, \tau^h]) = 0 = F_A(\xi, \tau).$$

Here we use that at any point ξ^v can be extended to a vertical vector field $\eta^\sharp$ for some vector $\eta \in \mathfrak{g}$ (recall that G acts on P). Hence $\tau(A(\xi^v)) = \tau(A(\eta^\sharp)) = 0$ since $A(\eta^\sharp) = \eta$ is constant. Also, we use that $[\xi^v, \tau^h] = 0$. To see this, we may assume that $\xi^v = \eta^\sharp$. Then, using the property of Lie derivative from Lecture 2, we obtain that

$$[\xi^v, \tau^h] = \mathcal{L}_{\eta^\sharp}(\tau^h) = \lim_{t\to 0} \frac{R^*_{\eta^t}(\tau^h) - \tau^h}{t}.$$

The last expression shows that $[\xi^v, \tau^h]$ is a horizontal vector.

Finally, let us take $\xi = \xi^v, \tau = \tau^v$. As above we may assume that $\xi = \eta^\sharp, \tau = \theta^\sharp$. Then $F_A(\xi, \tau) = 0$, and

$$(dA + \frac{1}{2}[A, A])(\xi^v, \tau^v)) = \xi^v(A(\tau^v)) - \tau^v(A(\xi^v)) + \frac{1}{2}([A(\xi^v), A(\tau^v)] - [A(\tau^v), A(\xi^v)]) -$$

$$-A([\xi^v, \tau^v]) = \xi^v(A(\tau^v)) - \tau^v(A(\xi^v)) - A([\xi^v, \tau^v]) + [A(\xi^v), A(\tau^v)] - [\eta, \theta)] + [\eta, \theta)] = 0.$$

(ii) This follows easily from Lemma 1 and (i). We leave the details to the reader.

(iii) We check only in the case $k = 1$, and leave the general case to the reader. Since $\beta(\xi) = \beta(\xi^h)$, we have

$$d^A(\beta)(\xi_1, \xi_2) = \xi_1^h(\beta(\xi_2)) - \xi_1^h(\beta(\xi_1)) - \beta([\xi_1, \xi_2]).$$

On the other hand,

$$(d\beta + [A, \beta])(\xi_1, \xi_2) = d\beta(\xi_1, \xi_2) + [A, \beta](\xi_1, \xi_2) =$$

$$= \xi_1(\beta(\xi_2)) - \xi_2(\beta(\xi_1)) - \beta([\xi_1, \xi_2] + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)] =$$

$$= d^A(\beta)(\xi_1, \xi_2) + \xi_1^v(\beta(\xi_2)) - \xi_2^v(\beta(\xi_1)) + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)].$$

As in the proof of (i), we may assume that $\xi_i^v = \eta_i^\sharp$ for some $\eta_i^\sharp \in \mathfrak{g}$. since $R_g^*(\beta) = Ad(g^{-1}(\beta))$, we obtain $\eta_i^\sharp(\beta(\xi_j)) = -[\eta_i, \beta(\xi_j)]$. Also, $[A(\xi_i), \beta(\xi_j)] = [\eta_i, \beta(\xi_j)]$. This implies that

$$\xi_1^v(\beta(\xi_2)) - \xi_2^v(\beta(\xi_1)) + [A(\xi_1), \beta(\xi_2)] - [A(\xi_2), \beta(\xi_1)] = 0$$

and proves the assertion.

Corollary 1.

$$d^A \circ d^A(\omega) = [F_A, \omega].$$

Proof. We have

$$d^A(d^A(\omega)) = d^A(d\omega + [A, \omega]) = d(d\omega + [A, \omega]) + [A, d\omega + [A, \omega]] =$$

$$= d[A, \omega] + [A, d\omega] + [A, [A, \omega]].$$

Applying Lemma 1, we get

$$d^A(d^A(\omega)) = [dA, \omega] - [A, d\omega] + [A, d\omega] + \frac{1}{2}[[A, A], \omega] = [dA + \frac{1}{2}[A, A], \omega].$$

Remark 1. If G is a subgroup of $GL(n,\mathbb{R})$, then we may identify A with a matrix whose entries are 1-forms. Recall that the Lie bracket in $\mathfrak{g}$ is the matrix bracket $[X,Y] = XY - YX$. Then $[A,A] = A\wedge A - (-1)A\wedge A = 2A\wedge A = 2A^2$. Thus

$$F_A = dA + A^2.$$

Definition. A connection is called *flat* if $F_A = 0$.

12.5 Now let E be a rank d vector G-bundle over M with typical fibre a real vector space V. We may assume that E is associated to a principal G-bundle. Let $\mathcal{E}$ be the sheaf of sections of E and $\mathcal{A}^k(E)$ be the sheaf of sections of $\Lambda^k(E)$.

Definition. A *connection on a vector bundle* E is a map of sheaves

$$\nabla : \mathcal{E} \to \mathcal{A}^1(E),$$

satisfying the Leibnitz rule: for any section $s : U \to E$ and $f \in \mathcal{O}_M(U)$,

$$\nabla(fs) = df \otimes s + f\nabla(s).$$

Let us explain how to construct a connection on E by using a connection A on the corresponding principal bundle P. Choose a local section $e : U \to P$ (which is equivalent to a trivialization $U \times G \to \pi^{-1}(U)$) and consider $\theta_e = e^*(A)$ as a 1-form on U with coefficients in $\mathfrak{g}$. Thus, for any tangent vector $\tau_x \in T(M)_x$, we have

$$e^*(A)(\tau_x) = A(de_x(\tau_x)).$$

If $e' : U' \to P$ is another section, then by (12.13), for any $x \in U \cap U'$,

$$\theta_e(\tau_x) = Ad(g)^{-1}(\theta_{e'}(\tau_x)) + g^{-1}dg,$$

where $g : U \cap U' \to G$ is a smooth map such that $e(x) = e'(x)\cdot g(x)$. Let $(U_i)_{i\in I}$ be an open cover of M trivializing P, and let $e_i : U_i \to P$ be some trivializing diffeomorphisms. The forms $\theta_i = \theta_{e_i}$ are related by

$$\theta_i = Ad(g_{ij})^{-1}\cdot\theta_j + g_{ij}^{-1}dg_{ij},$$

where g_{ij} are the transition functions. We see that, because of the second term in the right-hand side, $\{\theta_i\}_{i\in I}$ do not form a section of any vector bundle. However, the difference $A - A'$ of two connections defines a section $\{\theta_i - \theta_i'\}_{i\in I}$ of $\mathcal{A}^1(Ad(P))$. Applying the representation $d\rho : \mathfrak{g} \to \mathrm{End}(V) = \mathfrak{gl}(V))$, we obtain a set of 1-forms θ_e^ρ on U with values in $d\rho(\mathfrak{g}) \subset \mathrm{End}(V)$. They satisfy

$$\theta_e^\rho = \rho(g)^{-1}\cdot\theta_{e'}^\rho\cdot\rho(g) + \rho(g)^{-1}d\rho(g). \tag{12.16}$$

Note that the trivialization $e : U \to P$ defines a trivialization of $U \times V \to E|U$. In fact,

$$E_U = P_U \times_G V \cong (U \times G) \times_G V \cong U \times (G \times_G V) \cong U \times V.$$

Now we can define the connection ∇^ρ_A on E as follows. Using the trivialization, we can consider $s : U \to E$ as a $\mathbb{R}^n$-valued function. This means that $s = (s_1, \ldots, s_n)$, where s_i are scalar functions. Then we set

$$\nabla^\rho_A(s) = ds + \theta^\rho_e \cdot s, \tag{12.17}$$

where we view θ^ρ_e as an endomorphism in the space of sections of E which depends on a vector field. It obviously satisfies the Leibnitz rule. To check that our formula is well-defined, we compare $\nabla^\rho_A(s)$ with $\nabla^\rho_A(s')$, where $s'(x) : U' \to E$ and $s' = \rho(g)s$ for some transition function $g : U \cap U' \to G$ of the prinicipal G-bundle. We have

$$ds + \theta^\rho_e \cdot s = d(\rho(g)^{-1}s') + \rho(g)^{-1}\theta^\rho_{e'}\rho(g) \cdot s + \rho(g)^{-1}d\rho(g) \cdot s =$$

$$d\rho(g)^{-1}\rho(g) \cdot s + \rho(g)^{-1}ds' + \rho(g)^{-1}\theta^\rho_{e'} \cdot s' + \rho(g)^{-1}d\rho(g) \cdot s =$$

$$= \rho(g)^{-1}(ds' + \theta^\rho_{e'} \cdot s') + [d\rho(g)^{-1}\rho(g) + \rho(g)^{-1}d\rho(g)] \cdot s = \rho(g)^{-1}(ds' + \theta^\rho_{e'} \cdot s').$$

Here, at the last step, we have used that $0 = d(\rho(g)^{-1} \cdot \rho(g)) = d\rho(g)^{-1}\rho(g) + \rho(g)^{-1}d\rho(g)$. This shows that $\nabla^\rho_A(s)$ takes values in $\mathcal{A}^1(E)$ so that our map ∇^ρ_A is well-defined.

Remark 2. Let $\mathcal{T}_M$ be the sheaf of sections of $T(M)$. Its sections over U are vector fields on U. There is a canonical pairing

$$\mathcal{T}_M \otimes \mathcal{A}^1(E) \to \mathcal{E}.$$

Composing it with $\nabla^\rho_A \otimes 1 : \mathcal{E} \otimes \mathcal{T}_M \to \mathcal{A}^1(E) \to \mathcal{E}$, we get a map

$$\mathcal{E} \otimes \mathcal{T}_M \to \mathcal{E},$$

or, equivalently,

$$\mathcal{T}_M \to \mathcal{E}nd(E), \tag{12.18}$$

where $\mathcal{E}nd(E)$ is the sheaf of sections of $E^* \otimes E = End(E)$. It follows from our definition of ∇^ρ_A that the image of this map lies in the sheaf $Ad(E)$ which is the sheaf of sections of the image of the natural homomorphism $d\rho : Ad(P) \to \mathcal{E}nd(E)$.

We can also define the curvature of the connection ∇^ρ_A. We already know that F_A can be considered as a section of the sheaf $\mathcal{A}^2(Ad(P))$ on M. We can take its image in $\mathcal{A}^2(Ad(E))$ under the map $Ad(P) \to Ad(E)$. This section is denoted by F^ρ_A and is called the curvature form of E. Locally, with respect to the trivialization $e : U \to P$, it is given by

$$(F^\rho_A)_e = d\theta^\rho_e + \frac{1}{2}[\theta^\rho_e, \theta^\rho_e] = d\theta^\rho_e + \theta^\rho_e \wedge \theta^\rho_e. \tag{12.19}$$

Here for two 1-forms α, β with values in End(V),

$$[\alpha,\beta](\xi_1,\xi_2) = [\alpha(\xi_1),\beta(\xi_2)] - [\alpha(\xi_2),\beta(\xi_1)] = 2(\alpha(\xi_1)\beta(\xi_2) - \beta(\xi_2)\alpha(\xi_1)) := 2\alpha\wedge\beta(\xi_1,\xi_2).$$

Finally, we can extend ∇_A^ρ to

$$d_A^\rho : \mathcal{A}^k(E) \to \mathcal{A}^{k+1}(E)$$

by requiring $d_A^\rho = \nabla_A^\rho$ for $k = 0$ and

$$d_A^\rho(\omega \wedge \alpha) = d\omega \wedge \alpha + (-1)^{deg(\omega)}\omega \wedge \nabla_A^\rho(\alpha).$$

for $\omega \in \mathcal{A}^k(U), \alpha \in \mathcal{E}(U)$. Using Theorem 1, we can easily check that

$$d_A^\rho \circ d_A^\rho(\beta) = [F_A^\rho, \beta].$$

12.6 Let us consider the special case when $G = GL(n,\mathbb{R})$ and E is a rank n vector bundle. In this case the Lie algebra $\mathfrak{g}$ is equal to the algebra of matrices $\mathrm{Mat}_n(\mathbb{R})$ with the bracket $[X,Y] = XY - YX$. We omit ρ in the notation ∇_A^ρ. A trivialization $e : U \to P$ is a choice of a basis $(e_1(x),\ldots,e_n(x))$ in E_x. This defines the trivialization of $E|U$ by the formula

$$U \times \mathbb{R}^n \to E|U, (x,(a_1,\ldots,a_n)) \to a_1e_1(x) + \ldots + a_ne_n(x).$$

A section $s : U \to E$ is given by $s(x) = a_1(x)e_1(x) + \ldots + a_n(x)e_n(x)$. The 1-form θ_e is a matrix $\theta_e = (\theta_{ij})$ with entries $\theta_{ij} \in \mathcal{A}^1(U)$. If we fix local coordinates $(x_1,\ldots,x_d)$ on U, then

$$\theta_{ij} = \sum_{k=1}^{d} \Gamma_{ij}^k dx_k.$$

The connection ∇_A is given by

$$\nabla_A(s) = \sum_{i=1}^{n} d(a_i(x)e_i(x)) = \sum_{i=1}^{n} da_i(x)e_i(x) + \sum_{j=1}^{n} a_j(x)\nabla_A(e_j(x)) =$$

$$= \sum_{i,j=1}^{n} (\sum_{k=1}^{d} (\frac{\partial a_i}{\partial x_k} + a_j(x)\Gamma_{ij}^k dx_k)e_i(x). \tag{12.20}$$

If we view ∇_A in the sense (12.18), then we can put (12.19) in the following equivalent forms:

$$\nabla_A(\frac{\partial}{\partial x_k})(s) = \sum_{i=1}^{n} (\frac{\partial a_i}{\partial x_k} + \sum_{j=1}^{n} a_j(x)\Gamma_{ij}^k)e_i(x), \tag{12.21}$$

$$\nabla_{A,k} = \nabla_A(\frac{\partial}{\partial x_k}) = \frac{\partial}{\partial x_k} + (\Gamma_{ij}^k)_{1\le i,j\le n} = \frac{\partial}{\partial x_k} + A_k^e. \tag{12.22}$$

This is an operator which acts on the section $a_1(x)e_1(x)+\ldots+a_n(x)e_n(x)$ by transforming it into the section $b_1(x)e_1(x)+\ldots+b_n(x)e_n(x)$, where

$$\begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial a_1(x)}{\partial x_k} \\ \vdots \\ \frac{\partial a_n(x)}{\partial x_k} \end{pmatrix} + \begin{pmatrix} \Gamma^k_{11} & \ldots & \Gamma^k_{1n} \\ \vdots & \vdots & \vdots \\ \Gamma^k_{n1} & \ldots & \Gamma^k_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_1(x) \\ \vdots \\ a_n(x) \end{pmatrix}.$$

If we change the trivialization to $e(x) = e'(x) \cdot g(x)$, then we can view $g(x)$ as a matrix function $g : U \to GL(n, \mathbb{R})$, and get

$$A^e_k = g^{-1} A^{e'}_k g + g^{-1} \frac{\partial g}{\partial x_k}. \tag{12.23}$$

By taking the Lie bracket of operators (12.22), we obtain

$$[\nabla_{A,i}, \nabla_{A,j}] = [\frac{\partial}{\partial x_i} + A^e_i, \frac{\partial}{\partial x_j} + A^e_j] = \frac{\partial}{\partial x_i} A^e_j - \frac{\partial}{\partial x_j} A^e_i + [A^e_i, A^e_j]. \tag{12.24}$$

Comparing it with (12.19), we get

$$F_A = \sum_{\mu,\nu=1}^{d} [\nabla^A_\mu, \nabla^A_\nu] dx_\mu \wedge dx_\nu. \tag{12.25}$$

Example 1. Let us consider a very special case when E is a rank 1 vector bundle (a *line bundle*). In this case everything simplifies drastically. We have

$$\nabla_{A,k} = \frac{\partial}{\partial x_k} + \Gamma^k.$$

In different trivialization $e(x) = e'(x)g(x)$, we have

$$\Gamma^k = \Gamma'^k + \frac{\partial \log g(x)}{\partial x_k}.$$

If we put $\Gamma = \sum_{k=1}^d \Gamma^k dx_k$, then

$$\Gamma = \Gamma' + d \log g(x). \tag{12.26}$$

Thus a connection on a line bundle defined by transition functions g_{ij} is a section of the principal $\mathbb{R}^{\dim M}$-bundle on M with transition functions $\{d\log(g_{ij}(x))\}$ belonging to the affine group $\mathrm{Aff}(\mathbb{R}^n)$. The difference of two connections is a section of the cotangent bundle $T^*(M)$. The curvature is a section of $\bigwedge^2(T^*(M))$ given by 2-forms $d\Gamma$.

Example 2. Let $E = T(M)$ be the tangent bundle of M. The Lie derivative $\mathcal{L}_\eta$ coincides with the connection $\nabla_A(\eta)$ such that all sections (vector fields) are horizontal. This means that $\Gamma^k_{ij} \equiv 0$. We shall discuss connections on the tangent bundle in more detail later.

12.7 Let us summarize the description of the set of all connections on a principal G-bundle P. It is the set $\mathrm{Con}(P)$ of all differential 1-forms A on P with values in $\mathfrak{g} = Lie(G)$ satisfying

(i) $(R_g)^*(A) = Ad(g^{-1}) \cdot (A)$;

(ii) $A_p : T(P)_p^v \to \mathfrak{g}$ is the canonical map $\alpha_p : T(P_{\pi(p)}) \to \mathfrak{g}$.

Because of (ii), the difference of two connections is a 1-form which satisfies (i) and is identically zero on vertical vector fields. Using a horizontal lift we can descend this form to M to obtain a section of the vector bundle $T^*(M) \otimes Ad(P)$. In this way we obtain that $\mathrm{Con}(P)$ is an affine space over the linear space $\mathcal{A}^1(Ad(P))(M)$. It is not empty. This is not quite trivial. It follows from the existence of a metric on P which is invariant with respect to G. We refer to [**Nomizu**] for the proof.

The gauge group $\mathcal{G}(P)$ acts naturally on the set of sections $\Gamma(P)$ by the rule:

$$(g \cdot s)(x) = g^{-1} \cdot s(x).$$

It also acts on the set of connections on P (or on an associated vector bundle) by the rule:

$$\nabla_{g(A)}(s) = g\nabla_A(g^{-1}s).$$

Locally, if A is given by a $\mathfrak{g}$-valued 1-forms θ_e, it acts by changing the trivialization $e : U \times G \to P|U$ to $e' = g \cdot e$. As we know this changes θ_e to $Ad(g) \cdot \theta_e + g^{-1}dg$.

Exercises.

1. A section $s : U \to E$ of a vector bundle E is called *horizontal* with respect to a connection A if $\nabla_A(s) = 0$. Show that horizontal sections form a sheaf of sections of some vector bundle whose transition functions are constant functions.

2. Let $\pi : E \to M$ be a vector bundle. Show that a connection ∇ on E is obtained from a connection A on some principal G-bundle to which E is associated.

3. Show that there is a bijective correspondence between connections on a vector bundle E and linear morphisms of vector bundles $\pi^*(T(M) \to T(E)$ which are right inverses of the differential map $d\pi : T(E) \to \pi^*(T(M))$.

4. Let ∇ be a connection on a vector bundle E. A tangent vector $\xi_z \in T(E)_z$ is called horizontal if there exists a horizontal section $s : U \to E$ such that ξ_z is tangent to $s(U)$ at the point $z = s(x)$. Let $E = T(M)$ be the tangent bundle of a manifold M. Show that the natural lift $\tilde{\gamma}(t) = (\gamma(t), \frac{d\gamma}{dt})$ of paths in M is a horizontal lift for some connection on $T(M)$.

5. Let ∇ be a connection on a vector bundle. Show that it defines a canonical connection on its tensor powers, on its dual, on its exterior powers. Define the tensor product of connections.

6. Let $P \to M$ be a principal G-bundle. Fix a point $x_0 \in M$. For any closed smooth path $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x_0$ and a point $p \in P_{x_0}$ consider a horizontal lift

$\tilde{\gamma} : [0,1] \to P$ with $\tilde{\gamma}(0) = p \in P_{x_0}$. Write $\tilde{\gamma}(1) = p \cdot g$ for a unique $g \in G$. Show that g is independent of the choice of p and the map $\gamma \to g$ defines a homomorphism of groups $\rho : \pi_1(M; x_0) \to G$. It is called the *holonomy representation* of the fundamental group $\pi_1(M; x_0)$.

7. Let $E = E(\rho) = P \times_\rho V$ be the vector bundle associated to a principal G-bundle by means of a linear representation $\rho : G \to GL(V)$. Let $a : P \times V \to E$ be the canonical projection. For any $v \in V$ let $\varphi : P \to E$ be the map $p \to a(p, v)$. Use the differential of this map and a connection A on P to define, for each point $z \in E$, a unique subspace H_z of $T(E)_z$ such that $T(E_{\pi(z)}) \oplus H_z = T(E)_z$. Show that the map $\pi^*(T(M)) \to T(E)$ from Exercise 3 maps each $T(M)_x$ isomorphically onto H_z, where $z \in E_x$.

8. Let $G = U(1)$. Show that the torus $T = \mathbb{R}^4/\mathbb{Z}^4$ admits a non-flat self-dual $U(1)$ connection if and only if it has a structure of an abelian variety.

Lecture 13. KLEIN-GORDON EQUATION

This equation describes classical fields and is an analog of the Euler-Lagrange equations for classical systems with a finite number of degrees of freedom.

13.1 Let us consider a system of finitely many harmonic oscillators which we view as a finite set of masses arranged on a straight line, each connected to the next one via springs of length ϵ. The Lagrangian of this system is

$$L = \frac{1}{2}\sum_{i=1}^{N}\left[m\dot{x}_i^2 - k(x_i - x_{i+1})^2\right] = \frac{\epsilon}{2}\sum_{i=1}^{N}\left[\frac{m}{\epsilon}\dot{x}_i^2 - k\epsilon\frac{(x_i - x_{i+1})^2}{\epsilon^2}\right]. \qquad (13.1)$$

Now let us take N go to infinity, or equivalently, replace ϵ with infinitely small Δx. We can write

$$\lim_{N\to\infty} L = \frac{1}{2}\int_a^b \left[\mu\left(\frac{\partial\phi(x,t)}{\partial t}\right)^2 - Y\left(\frac{\partial\phi(x,t)}{\partial x}\right)^2\right]dx, \qquad (13.2)$$

where m/ϵ is replaced with the mass density μ, and the coordinate x_i is replaced with the function $\phi(x,t)$ of displacement of the particle located at position x and time t. The function Y is the limit of $k\epsilon$. It is a sort of elasticity density function for the spring, the so-called "Young's modulus". The right-hand side of equation (13.2) is a functional on the space of functions $\phi(x,t)$ with values in the space of functions on $\mathbb{R}^3$. We can generalize the Euler-Lagrange equation to find its extremum point. Let us introduce the *action functional*

$$S = \int_{t_0}^{t_1}\int_a^b L(\phi, \partial_\mu\phi)dxdt.$$

Here L is a smooth function in two variables, and

$$\partial_\mu\phi = \left(\frac{\partial\phi}{\partial t}, \frac{\partial\phi}{\partial x}\right),$$

and the expression

$$\mathcal{L}(\phi) = L(\phi, \partial_\mu \phi)$$

is a functional on the space of functions ϕ. In our case

$$\mathcal{L}(\phi) = \frac{1}{2}\mu(\frac{\partial \phi}{\partial t})^2 - \frac{1}{2}Y(\frac{\partial \phi}{\partial x})^2. \tag{13.3}$$

We shall look for the functions ϕ which are critical points of the action S.

13.2. Recall from Lecture 1 that the critical points can be found as the zeroes of the derivative of the functional S. We shall consider a more general situation. Let $L : \mathbb{R}^{n+1} \to \mathbb{R}$ be a smooth function (it will be called the *Lagrangian*). Consider the functional

$$\mathcal{L} : L_2(\mathbb{R}^n) \to L_2(\mathbb{R}^n), \quad \phi(x,t) \to L(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}). \tag{13.4}$$

Then the action S is the composition of $\mathcal{L}$ and the integration functional over $\mathbb{R}^n$;

$$S(\phi) = \int\limits_{\mathbb{R}^n} \mathcal{L}(\phi) d^n x = \int L(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) d^n x.$$

By the chain rule, its derivative $S'(\phi)$ at the function ρ is equal to the linear functional $L_2(\mathbb{R}^n) \to \mathbb{R}$ defined by

$$S'(\phi)(h) = \int\limits_{\mathbb{R}^n} \mathcal{L}'(\phi) h d^n x.$$

The functional $\mathcal{L}$ (also often called the *Lagrangian*) is the composition of the linear operator

$$L_2(\mathbb{R}) \to L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) = L_2(\mathbb{R}^n)^{n+1}, \quad \phi \to \Phi = (\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n})$$

and the functional

$$L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) \to L_2(\mathbb{R}^n), \quad \Phi \to L \circ \Phi.$$

The derivative of the latter functional at $\Phi = (\Phi_1, \ldots, \Phi_{n+1})$ is equal to the linear functional $L_2(\mathbb{R}^n, \mathbb{R}^{n+1}) \to L_2(\mathbb{R}^n)$

$$(h_1, \ldots, h_{n+1}) \to \sum_{i=0}^{n} \frac{\partial L}{\partial y_i}(\Phi(x)) h_i,$$

where we use coordinates $y_0, \ldots, y_n$ in $\mathbb{R}^{n+1}$. This implies that the derivative of $\mathcal{L}$ at ϕ is equal to the linear functional

$$\mathcal{L}'(\phi) : h \to \frac{\partial L}{\partial y_0}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) h + \sum_{\mu=1}^{n} \frac{\partial L}{\partial y_\mu}(\phi, \frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}) \frac{\partial h}{\partial x_i}. \tag{13.5}$$

Let us set

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{\partial L}{\partial y_0}(\phi, \frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n}), \quad \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = \frac{\partial L}{\partial y_\nu}(\phi, \frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n}), \quad \nu = 1, \dots, n \tag{13.6}$$

These are of course functions on $\mathbb{R}^n$. In this notation we can rewrite (13.5) as

$$\mathcal{L}'(\phi)h = \frac{\delta \mathcal{L}}{\delta \phi} h + \sum_{\nu=1}^{n} \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \frac{\partial h}{\partial x_\nu}.$$

This easily implies that

$$S(\phi)(h) = \int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} h + \frac{\partial \mathcal{L}}{\partial \partial_1 \phi} \frac{\partial h}{\partial x_1} + \dots + \frac{\partial \mathcal{L}}{\partial \partial_n \phi} \frac{\partial h}{\partial x_n}) d^n x.$$

In the physicists notation, we write $\delta\phi$ instead of h, and $\delta\partial_i\phi = \partial_i\delta\phi$ instead of $\frac{\partial h}{\partial x_i}$ and find that, if the action has a critical point at ϕ,

$$\int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_1 \phi} \delta\partial_1\phi + \dots + \frac{\partial \mathcal{L}}{\partial \partial_n \phi} \delta\partial_n\phi) d^n x = 0, \tag{13.7}$$

or, using the Einstein convention (to which we shall switch from time to time without warning),

$$\int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \delta\partial_\nu\phi) d^n x = 0.$$

Let us restrict the action to the affine subspace of functions ϕ with fixed boundary condition. This means that ϕ has support on some open submanifold $V \subset \mathbb{R}^n$ with boundary δV, and $\phi|\delta V$ is fixed. For example, $V = [0, \infty) \times \mathbb{R}^3 \subset \mathbb{R}^4$ with $\delta V = \{0\} \times \mathbb{R}^3$. If we are looking for the minimum of the action S on such a subspace, we have to take $\delta\phi$ equal to zero on δV. Integrating by parts, we find (using ∂_ν to denote $\frac{\partial}{\partial x_\nu}$)

$$0 = \int_{\mathbb{R}^n} (\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} \delta\partial_\nu\phi) d^n x = \int_V \partial_\nu (\delta\phi \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) d^n x + \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x =$$

$$= \int_{\delta V} (\delta\phi \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) d^{n-1} x + \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x = \int_V (\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi}) \delta\phi d^n x.$$

From this we deduce the following *Euler-Lagrange equation*:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = \frac{\partial \mathcal{L}}{\partial \phi} - \sum_{\nu=1}^{n} \partial_\nu \frac{\partial \mathcal{L}}{\partial \partial_\nu \phi} = 0. \tag{13.8}$$

Observe the analogy with the Euler-Lagrange equations from Lecture 1:

$$\frac{\partial L}{\partial q_i}(\mathbf{x}_0, \frac{d\mathbf{x}_0}{dt}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{x}_0, \frac{d\mathbf{x}_0}{dt}) = 0, i = 1, \dots, n.$$

In our case $(q_1(t), \dots, q_n(t))$ is replaced with $\phi(x,t)$, and $(\dot{q}_1(t), \dots, \dot{q}_n(t))$ is replaced with $\frac{\partial \phi(x,t)}{\partial t}$.

Returning to our situation where $\mathcal{L}$ is given by (13.3), we get from (13.8)

$$\frac{\partial^2 \phi}{\partial x^2} = \left(\frac{\mu}{Y}\right)\frac{\partial^2 \phi}{\partial t^2}. \tag{13.9}$$

This is the standard wave equation for a one-dimensional system traveling with velocity $\sqrt{Y/\mu}$.

Let us take $n = 4$ and consider the space-time $\mathbb{R}^4$ with coordinates (t, x_1, x_2, x_3). That is, we shall consider the functions $\phi(t,x) = \phi(t, x_1, x_2, x_3)$ defined on $\mathbb{R}^4$. We use $x^\nu, \nu = 0,1,2,3$ to denote t, x_1, x_2, x_3, respectively. Set

$$\partial_\nu = \frac{\partial}{\partial x^\nu}, \nu = 0,1,2,3,$$

$$\partial^\nu = \frac{\partial}{\partial x^\nu}, \nu = 0, \quad \partial^\nu = -\frac{\partial}{\partial x^\nu}, \nu = 1,2,3.$$

The operator

$$\Box = \sum_{\nu=0}^{3} \partial_\nu \partial^\nu = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

is called the *D'Alembertian*, or *relativistic Laplacian.*

If we take

$$\mathcal{L} = \frac{1}{2}\left(\sum_{\nu=0}^{3} \partial_\nu \phi \partial^\nu \phi - m^2 \phi^2\right), \tag{13.10}$$

the Euler-Lagrange equation gives

$$\Box \phi + m^2 \phi = 0. \tag{13.11}$$

This is called the *Klein-Gordon equation.*

13.3 We can generalize the Euler-Lagrange equation to fields which are more general than scalar functions on $\mathbb{R}^n$, namely sections of some vector bundle E over a smooth n-dimensional manifold M with some volume form $d\mu$ equipped with a connection $\nabla : \Gamma(E) \to \mathcal{A}^1(E)(M)$. For this we have to consider the Lagrangian as a function $L : E \oplus \Lambda^1(E) \to \mathbb{R}$. In a trivialization $E \cong U \times \mathbb{R}^r$, our field is represented by a vector function $\phi = (\phi_1, \dots, \phi_r)$, and the Lagrangian by a scalar function on $U \times \mathbb{R}^r \times \mathbb{R}^{nr}$. Let $\Gamma = \Gamma(E) \oplus \mathcal{A}^1(E)(M)$ be the space of sections of the vector bundle $E \oplus \Lambda^1(E)$. We have a canonical map $\Gamma(E) \to \Gamma$ which sends ϕ to $(\phi, \nabla(\phi))$. Composing it with L, we get the map

$$\mathcal{L} : \Gamma(E) \to C^\infty(M), \phi \to \mathcal{L}(\phi) = L(\phi(x), \nabla\phi(x)).$$

Then we can define the action functional on X by the formula

$$S(\phi) = \int_M \mathcal{L}(\phi) d\mu = \int_M L(\phi, d\phi) d\mu.$$

The condition for ϕ to be an extremum point is

$$\int_M \Big(\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \nabla\phi}\nabla\delta\phi\Big) d^n x = 0 \tag{13.12}$$

Here $\delta\phi$ is any function in $L_2(M,\mu)$ with $||\delta\phi|| < \epsilon$ for sufficiently small ϵ. The partial derivative $\frac{\partial \mathcal{L}}{\partial \phi}$ is the partial derivative of $\mathcal{L}$ with respect to the summand $\Gamma(E)$ of Γ at the point $(\phi, \nabla\phi)$. The partial derivative $\frac{\partial \mathcal{L}}{\partial \nabla\phi}$ is the partial derivative of $\mathcal{L}$ with respect to the summand $\mathcal{A}^1(E)(M)$ of Γ computed at the point $(\phi, \nabla\phi)$. This partial derivative is a linear map $\mathcal{A}^1(E)(M) \to C^\infty(M)$ which we apply at the point $\nabla\delta\phi$. We leave it to the reader to deduce the corresponding Euler-Lagrange equation.

The Klein-Gordon Lagrangian can be extended to the general situation we have just described. Let us fix some function $q : E \to \mathbb{R}$ whose restriction to each fibre E_x is a quadratic form on E_x. Let g be a pseudo-Riemannian metric on M, considered as a function $g : T(M) \to \mathbb{R}$ whose restriction to each fibre is a non-degenerate quadratic form on $T(M)_x$. If we view g as an isomorphism of vector bundles $g : T(M) \to T^*(M)$, then its inverse is an isomorphism $g^{-1} : T^*(M) \to T(M)$ which can be viewed as a metric $g^{-1} : T^*(M) \to \mathbb{R}$ on $T^*(M)$. We define the Klein-Gordon Lagrangian

$$L = -u \circ q + \lambda g^{-1} \otimes q : E \oplus \Lambda^1(E) \to \mathbb{R},$$

Its restriction to the fibre over a point $x \in M$ is given by the formula

$$(a_x, \omega_x \otimes b_x) \to -u(x)q(a_x) + \lambda(x)q(b_x)g^{-1}(\omega_x) \tag{13.13}$$

for some appropriate non-negative valued scalar functions λ and u. Usually q is taken to be a positive definite form. It is an analog of potential energy. The term $\lambda q g^{-1}$ is the analog of kinetic energy.

The simplest case we have considered above corresponds to the situation when $M = \mathbb{R}^n, E = M \times \mathbb{R}$ is the trivial bundle with the trivial connection $\phi \to d\phi$. The Lagrangian $\mathcal{L} : M \times \mathbb{R}^{n+1} \to \mathbb{R}$ does not depend on $x \in M$. The Lagrangian from (13.10) corresponds to taking g to be the Lorentz metric, the quadratic function q equals $(x, z) \to -m^2 z^2$, and the scalar λ is equal to $1/2$.

A little more general situation frequently considered in physics is the case when E is the trivial vector bundle of rank r with trivial connection. Then ϕ is a vector function $(\phi_1, \ldots, \phi_n)$, $d\phi = (d\phi_1, \ldots, d\phi_n)$ and the Euler-Lagrange equation looks as follows:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} = \sum_{\mu=1}^{n} \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_i}, \quad i = 1, \ldots, r.$$

13.4 The Lagrangian (13.10) and the corresponding Euler-Lagrange equation (13.11) admits a large group of symmetries. Let us explain what this means. Let G be any Lie group which acts smoothly on a manifold M (see Lecture 2).

Let $\pi : E \to M$ be a vector bundle on M. A *lift* of the action $G \times M \to M$ is a smooth action $G \times E \to E$ such that, for any $z \in E_x$, $g \cdot z \in E_{g \cdot x}$. Given such a lift, G acts naturally on the space of sections of E by the formula

$$(g \cdot s)(x) = g \cdot s(g^{-1} \cdot x). \tag{13.14}$$

If, additionally, the restriction of any $g \in G$ to E_x is a linear map $E_x \to E_{g \cdot x}$ (such a lift is called a *linear lift*) then the previous formula defines a linear representation of G in the space $\Gamma(E)$.

Suppose, for example, that $E = M \times \mathbb{R}^n$ is the trivial bundle. Let G be a Lie group of diffeomorphisms of M and $\rho : G \to GL(n, \mathbb{R})$ be its linear representation. Then the formula

$$g \cdot (x, a) = (g \cdot x, \rho(g)(a))$$

defines a linear lift of the action of G on M to E. In particular, if $n = 1$ and ρ is trivial, then $g \in G$ transforms a scalar function ϕ into a functon $\phi' = g^{-1} \circ \phi$. It satisfies

$$\phi'(g \cdot x) = \phi(x). \tag{13.15}$$

If we identify a section of E with a smooth vector function $\underline{\phi}(x) = (\phi_1(x), \ldots, \phi_n(x))$ on M, then g acts on $\Gamma(E)$ by the formula

$$g \cdot \underline{\phi}(x) = (\psi_1(x), \ldots, \psi_n(x)),$$

where

$$\psi_i(x) = \sum_{j=1}^{n} \alpha_{ij} \phi_j(g^{-1} \cdot x), \quad \rho(g) = (\alpha_{ij}) \in GL(n, \mathbb{R}), \quad i = 1, \ldots, n.$$

The tangent bundle $T(M)$ admits the *natural lift* of any action of G on M. For any $\tau_x \in T(M)_x$ we set

$$g \cdot \tau_x = (dg)_x(\tau_x) \in T(M)_{g \cdot x}.$$

A section of $T(M)$ is a vector field $\tau : x \to \tau_x$. We have

$$(g \cdot \tau)_x = g \cdot \tau_{g^{-1} \cdot x} = (dg)_{g^{-1} \cdot x}(\tau_{g^{-1} \cdot x}).$$

If we view a vector field τ as a derivation $C^\infty(M) \to C^\infty(M)$, this translates to

$$(g \cdot \tau)_x(\phi) = \tau_{g^{-1} \cdot x}(g^*(\phi)).$$

Let $g \in G$ define a diffeomorphism from an open set $U \subset M$ with coordinate functions $\underline{x} = (x_1, \ldots, x_n) : U \to \mathbb{R}^n$ to an open set $V = g(U)$ with coordinate functions $\underline{y} = (y_1, \ldots, y_n) : V \to \mathbb{R}^n$. For any $a \in U$ we have

$$y_i(g(a)) = \psi_i(x_1(a), \ldots, x_n(a)), i = 1, \ldots, n, \tag{13.16}$$

for some smooth function $\psi : \underline{x}(U) \to \mathbb{R}^n$. Thus

$$(g \cdot \frac{\partial}{\partial x_j})(y_i) = \frac{\partial}{\partial x_j}(\psi_i(x_1, \ldots, x_n)) = \frac{\partial \psi_i}{\partial x_j}.$$

This implies

$$g_*(\frac{\partial}{\partial x_j})_{g\cdot x} := (g \cdot \frac{\partial}{\partial x_j})_{g\cdot x} = \sum_{i=1}^{n} \frac{\partial \psi_i}{\partial x_j}(x)\frac{\partial}{\partial y_i}, \tag{13.17}$$

$$g \cdot \sum_{j=1}^{n} a^j(x_1, \ldots, x_n)\frac{\partial}{\partial x_j} = \sum_{i=1}^{n} b^i(y)\frac{\partial}{\partial y_i},$$

where

$$b^i(y) = \sum_{j=1}^{n} a^j(g^{-1} \cdot y)\frac{\partial \psi_i}{\partial x_j}(g^{-1} \cdot y).$$

The cotangent bundle $T^*(M)$ also admits the *natural lift*. For any $\omega_x \in T^*(M)_x = T(M)_x^*$ and $\tau_{g\cdot x} \in T(M)_{g\cdot x}$,

$$(g \cdot \omega_x)(\tau_{g\cdot x}) = \omega_x(g^{-1} \cdot \tau_{g\cdot x}).$$

A section of $T^*(M)$ is a differential 1-form $\omega : x \to \omega_x$. We have

$$(g \cdot \omega)_x(\tau_x) = g \cdot \omega_{g^{-1}\cdot x}(\tau_x) = \omega_{g^{-1}\cdot x}(g^{-1} \cdot \tau_x).$$

On the other hand, if $f : M \to N$ is any smooth map of manifolds, one defines the inverse image of a differential 1-form ω on N by

$$f^*(\omega)_x(\tau) = \omega_{f(x)}(df_x(\tau)).$$

If $f = g : M \to M$, we have

$$g \cdot \omega = (g^{-1})^*(\omega).$$

If $g : U \to V$ as above, then

$$g^{-1} \cdot dy_i(\frac{\partial}{\partial x_j}) = g^*(dy_i)(\frac{\partial}{\partial x_j}) = dy_i(g_*(\frac{\partial}{\partial x_j})) = dy_i(\sum_{k=1}^{n} \frac{\partial \psi_k}{\partial x_j}\frac{\partial}{\partial y_k}) = \frac{\partial \psi_i}{\partial x_j}.$$

Therefore

$$g^{-1} \cdot (dy_i) := g^*(dy_i) = \sum_{j=1}^{n} \frac{\partial \psi_i}{\partial x_j} dx_j = d(g^*(y_i)).$$

$$g^{-1} \cdot (\sum_{i=1}^{n} b_i(y)dy_i) = \sum_{i=1}^{n} b_i(g \cdot x)\frac{\partial \psi_i}{\partial x_j} dx_j.$$

Now assume that G consists of linear transformations of $M = \mathbb{R}^n$. Fix coordinate functions $x_1, \ldots, x_n$ such that $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$ gives a basis in each $T(M)_{\mathbf{a}}$, and $dx_1, \ldots, dx_n$ gives a basis in $T^*(M)_{\mathbf{a}}$. Let $g \in G$ be identified with a matrix $A = (\alpha_{ij})$. It acts on $\mathbf{a} = (a_1, \ldots, a_n)$ by transforming it to

$$\mathbf{b} = A \cdot \mathbf{a}.$$

We have

$$x_i(A \cdot \mathbf{a}) = b_i = \sum_{j=1}^{n} \alpha_{ij} a_j = \sum_{j=1}^{n} \alpha_{ij} x_j(\mathbf{a}).$$

Thus (13.16) takes the form

$$g^*(x_i) = \phi_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} \alpha_{ij} x_j, i = 1, \ldots, n,$$

and

$$g \cdot \frac{\partial}{\partial x_j} = g_*(\frac{\partial}{\partial x_j}) = \sum_{i=1}^{n} \alpha_{ij} \frac{\partial}{\partial x_i},$$

$$g^{-1} \cdot dx_i = g^*(dx_i) = \sum_{j=1}^{n} \alpha^{ij} dx_j,$$

where

$$A^{-1} = (\alpha^{ij}).$$

13.5 Let us identify the Lagrangian $\mathcal{L}$ with a map $\mathcal{L} : C^\infty(M) \to C^\infty(M), \phi(x) \to \mathcal{L}(\phi)(x)$. We use the unknown x to emphasize that the values of $\mathcal{L}$ are smooth functions on M. Let $g : M \to M$ be a diffeomorphism. It acts on the source space and on the target space of $\mathcal{L}$, transforming $\mathcal{L}$ to $g\mathcal{L}$, where

$$(g \cdot \mathcal{L})(\phi) = g \cdot \mathcal{L}(g^{-1} \cdot \phi) = \mathcal{L}(g^*(\phi)((g^{-1})^* \cdot x).$$

We say that the *Lagrangian is invariant* with respect to g if $g \cdot \mathcal{L} = \mathcal{L}$, or, equivalently,

$$\mathcal{L}(g^*(\phi))(x) = \mathcal{L}(\phi)(g \cdot x). \tag{13.18}$$

Example. Let $M = \mathbb{R}$ and $\mathcal{L}(\phi) = \phi'(x)$. Take $g : x \to 2x$. Then

$$\mathcal{L}(g^*(\phi)) = \phi(2x)' = 2\phi'(2x) = 2\mathcal{L}(\phi)(2x) = 2\mathcal{L}(\phi)(g \cdot x).$$

So, $\mathcal{L}$ is not invariant with respect to g. However, if we take $h : x \to x + c$, then

$$\mathcal{L}(h^*(\phi)) = \phi(x + c)' = \phi(x + c) = \mathcal{L}(\phi)(x + c) = \mathcal{L}(\phi)(h \cdot x).$$

So, $\mathcal{L}$ is invariant with respect to h.

Assume $\mathcal{L}$ is invariant with respect to g and g leaves the volume form $d^n x$ invariant. Then

$$S(g^*(\phi)) = \int_M \mathcal{L}(g^*(\phi))d^n x = \int_M \mathcal{L}(\phi))(g\cdot x)g^*(d^n x) = \int_M \mathcal{L}(\phi)(x)d^n x = S(\phi).$$

This shows that g leaves the level sets of the action S invariant. In particular, it transforms critical fields to critical fields. Or, equivalently, it leaves invariant the set of solutions of the Euler-Lagrange equation.

Let η be a vector field on M and g^t_η be the associated one-parameter group of diffeomorphisms of M. We say that η is an *infinitesimal symmetry* of the Lagrangian $\mathcal{L}$ if $\mathcal{L}$ is invariant with respect to the diffeomorphisms g^t_η.

Let L_η be the Lie derivative associated to the vector field η (see Lecture 2). Assume η is an infinitesimal symmetry of $\mathcal{L}$. Then $\mathcal{L}(\phi)(g^t_\eta\cdot x) = \mathcal{L}(\phi(g^t_\eta\cdot x))$. This implies

$$L_\eta(\mathcal{L}(\phi,d\phi)(x)) = \lim_{t\to 0}\frac{\mathcal{L}(\phi)(g^t_\eta\cdot x)-\mathcal{L}(\phi)(x)}{t} = \frac{\partial\mathcal{L}}{\partial\phi}L_\eta(\phi)+\frac{\partial\mathcal{L}}{\partial d\phi}L_\eta(d\phi), \tag{13.19}$$

where L_η denotes the Lie derivative with respect to the vector field η. By Cartan's formula,

$$L_\eta(\phi) = \eta(\phi), \quad L_\eta(d\phi) = d\eta(\phi).$$

Let us assume $M = \mathbb{R}^n$ and $\mathcal{L}(\phi,d\phi) = F(\phi,\partial_1\phi,\ldots,\partial_n\phi)$ for some smooth function F in $n+1$ variables. Let $\eta = \sum_i a^i(x)\frac{\partial}{\partial x_i}$, then

$$\eta(\phi) = \sum_i a^i(x)\frac{\partial\phi}{\partial x_i},$$

$$d\eta(\phi) = d(\sum_i a^i\frac{\partial\phi}{\partial x_i}) = \sum_{i,j=1}^n(\frac{\partial a^i}{\partial x_j}\frac{\partial\phi}{\partial x_i}+a^i(x)\frac{\partial^2\phi_i}{\partial x_j\partial x_i})dx_j.$$

We can now rewrite (13.19) as

$$a^\mu\partial_\mu\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}a^\nu(x)\partial_\nu\phi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(\partial_\mu a^\nu\partial_\nu\phi+a^\nu(x)\partial_\nu\partial_\mu\phi).$$

Assume now that ϕ satisfies the Euler-Lagrange equation. Then we can rewrite the previous equality in the form

$$a^\mu\partial_\mu\mathcal{L} = \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}a^\nu(x)\partial_\nu\phi+\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(\partial_\mu a^\nu\partial_\nu\phi+a^\nu(x)\partial_\nu\partial_\mu\phi).$$

Since we assume that g^t_η preserves the volume form $\omega = dx_1\wedge\ldots\wedge dx_n$, we have

$$L_\eta(\omega) = d\langle\eta,\omega\rangle = (\sum_{i=1}^n\frac{\partial a^i}{\partial x_i})\omega = 0.$$

This shows that

$$a^\mu \frac{\partial \mathcal{L}}{\partial x_\mu} = \partial_\mu (a^\mu \mathcal{L}) - \mathcal{L}\partial_\mu (a^\mu) = \partial_\mu (a^\mu \mathcal{L}).$$

Now combine everything under ∂_μ to obtain

$$0 = -\partial_\mu (a^\mu \mathcal{L}) + \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} a^\nu (x) \partial_\nu \phi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} (\partial_\mu a^\nu \partial_\nu \phi + a^\nu (x) \partial_\nu \partial_\mu \phi) =$$

$$= \partial_\mu \Big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} a^\nu \partial_\nu \phi - \delta^\mu_\nu a^\nu \mathcal{L} \Big) = \sum_{\nu,\mu=1}^n \partial_\mu \Big(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} a^\nu \partial_\nu \phi - \delta^\nu_\mu a^\nu \mathcal{L} \Big), \tag{13.20}$$

where $\delta^\nu_\mu = \delta_{\nu\mu}$ is the Kronecker symbol.

Set

$$J^\mu = \sum_{\nu=1}^n (a^\nu (\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi} \partial_\nu \phi - \delta_{\mu\nu} \mathcal{L}), \quad \mu = 1, \ldots, n, \tag{13.21}$$

The vector-function

$$\mathbf{J} = (J^1, \ldots, J^n)$$

is called the *current* of the field ϕ with respect to the vector field η. From (13.21) we now obtain

$$\mathrm{div}\mathbf{J} = \partial_\mu J^\mu = \sum_{\mu=1}^n \frac{\partial J^\mu}{\partial x_\mu} = 0. \tag{13.22}$$

For example, let $n = 4, (x_1, x_2, x_3, x_4) = (t, x_1, x_2, x_3)$. We can write $(J^1, J^2, J^3, J^4) = (j^0, \mathbf{j})$ and rewrite (13.21) in the form

$$\frac{\partial j^0}{\partial t} = -\mathrm{div}(\mathbf{j}) := -\Big(\sum_{i=1}^3 \frac{\partial j^i}{\partial x_i} \Big).$$

Thus, if we introduce the *charge*

$$Q(t) = \int_{\mathbb{R}^3} j^0(t, x) d^3 x,$$

and assume that $\mathbf{j}$ vanishes fast enough at infinity, then the divergence theorem implies the conservation of the charge

$$\frac{dQ(t)}{dt} = 0.$$

Remarks. 1. If we do not assume that the infinitesimal transformation of M preserves the volume form, we have to require that the action function is invariant but not the Lagrangian. Then, we can deduce equation (13.21) for any field ϕ satisfying the Euler-Lagrange equation.

2. We can generalize equation (13.22) to the case of fields more general than scalar fields. In the notation of 13.3, we assume that E is equipped with a lift with respect to some

one-parameter group of diffeomorphisms g_η^t of M. This will define a canonical lift of g_η^t to $E \oplus \Lambda^1(E)$. At this point we have to further assume that the connection on E is invariant with respect to this lift. Now we consider the Lagrangian $\mathcal{L} : E \oplus (\Lambda^1(E)) \to \mathbb{R}$ as the operator $\mathcal{L} : \Gamma(E) \to \Gamma(E), \phi \to \mathcal{L}(\phi, \nabla\phi)$. We say that η is an infinitesimal symmetry of $\mathcal{L}$ if $g_\eta^t \cdot \mathcal{L}(g_\eta^{-t} \cdot \phi) = \mathcal{L}(\phi)$ for any $\phi \in \Gamma(E)$. Here we consider the action of g_η^t on sections as defined in 13.5. Now equation (13.19) becomes

$$\nabla(\tau)(\mathcal{L}(\phi)) = \frac{\partial \mathcal{L}}{\partial \phi}\nabla(\tau)(\phi) + \frac{\partial \mathcal{L}}{\partial \nabla(\phi)}\nabla(\nabla(\tau)(\phi)).$$

Here we consider the connection ∇ both as a map $\Gamma(E) \to \Gamma(\Lambda^1(E))$ and as a map $\Gamma(T(M)) \to \Gamma(End(E))$. We leave to the reader to find its explicit form analogous to equation (13.20), when we trivialize E and choose a connection matrix Γ_{ij}^k.

13.6 Now we shall consider various special cases of equation (13.20). Let us assume that $M = \mathbb{R}^n$ and $G = \mathbb{R}^n$ acts on itself by translations. Since ∂_ν commute with translations, any Lagrangian $\mathcal{L}(\phi, d\phi)$ is invariant with respect to translations. Also translations preserve the standard volume in $\mathbb{R}^n$. We can identify the Lie algebra of G with $\mathbb{R}^n$. For any $\eta = (a^1, \ldots, a^n) \in \mathbb{R}^n$, the vector field $\eta^\sharp$ is the constant vector field $\sum_i a^i \frac{\partial}{\partial x_i}$. Thus we get from (13.20) that

$$\sum_{\nu,\mu=1}^{n} a^\nu \partial_\mu \Big(\sum_{\nu=1}^{n} \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\nu \phi - \delta_{\nu\mu}\mathcal{L}\Big) = 0.$$

Since this is true for all η, we get

$$\sum_{\mu=1}^{n} \partial_\mu (\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\nu \phi - \delta_\nu^\mu \mathcal{L}) = 0, \quad \nu = 1, \ldots, n.$$

The tensor

$$(T_\nu^\mu) = (\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\nu \phi - \delta_\nu^\mu \mathcal{L}) := \sum_{\nu,\mu}^{n} (\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi}\partial_\nu \phi - \delta_\nu^\mu \mathcal{L})\frac{\partial}{\partial x_\mu} \otimes dx_\nu \tag{13.23}$$

is called the *energy-momentum tensor*. For example, consider the Klein-Gordon action where

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2).$$

We have

$$(T_\nu^\mu) = \begin{pmatrix} (\partial_0\phi)^2 - \mathcal{L} & \partial_0\phi\partial_1\phi & \partial_0\phi\partial_2\phi & \partial_0\phi\partial_3\phi \\ -\partial_0\phi\partial_1\phi & -(\partial_1\phi)^2 - \mathcal{L} & -\partial_1\phi\partial_2\phi & -\partial_1\phi\partial_3\phi \\ -\partial_0\phi\partial_2\phi & -\partial_1\phi\partial_2\phi & -(\partial_2\phi)^2 - \mathcal{L} & -\partial_2\phi\partial_3\phi \\ -\partial_0\phi\partial_3\phi & -\partial_1\phi\partial_3\phi & -\partial_2\phi\partial_3\phi & -(\partial_3\phi)^2 - \mathcal{L} \end{pmatrix}.$$

Since ϕ satisfies Klein-Gordon equation (13.11), we easily check that each row of the matrix T is a vector whose divergence is equal to zero. the frequency

We have

$$T_0^0 = \frac{1}{2}(m^2\phi^2 + \sum_{i=0}^{3}(\partial_i\phi)^2).$$

If we integrate T_0^0 over $\mathbb{R}^3$ with coordinates (x_1, x_2, x_3), we obtain the analog of the total energy function

$$\frac{1}{2}\sum_{i=1}^{n}(m^2q_i^2 + \dot{q}_i^2).$$

We also have $T^{0k} = \partial_0\phi\partial_k\phi = \frac{\partial\mathcal{L}}{\partial\partial_0\phi}\partial_k\phi, k = 1,2,3$. In classical mechanics $p_i = \frac{\partial\mathcal{L}}{\partial\dot{q}_i}$ is the momentum coordinate. Thus the integral over $\mathbb{R}^3$ of the vector function (T^{01}, T^{02}, T^{03}) represents the analog of the momentum.

13.7 Consider the orthogonal group $G = O(1,3)$ of linear transformations $A : \mathbb{R}^4 \to \mathbb{R}^4$ which preserve the quadratic form $x_0^2 - x_1^2 - x_2^2 - x_3^2$. Take the Klein-Gordon Lagrangian $\mathcal{L}$. Since $(\partial_0, \partial_1, \partial_2, \partial_3)$ transforms in the same way as the vector (x_0, x_1, x_2, x_3), we obtain that $\partial_\nu\partial^\nu(\phi(A\cdot x)) = \partial_\nu\partial^\nu(\phi)(A\cdot x)$. Thus

$$\mathcal{L}(\phi(A\cdot x)) = \partial_\nu\partial^\nu(\phi(A\cdot x)) - m^2\phi(A\cdot x)^2 = \partial_\nu\partial^\nu(\phi)(A\cdot x) - m^2\phi^2(A\cdot x) = \mathcal{L}(\phi)(A\cdot x).$$

This shows that the Lagrangian is invariant with respect to all transformations from the group $O(1,3)$. Let us find the corresponding currents.

As we saw in Lecture 12, the Lie algebra of $O(1,3)$ consists of matrices $A \in Mat_4(\mathbb{R})$ satisfying $A\cdot J + J\cdot A^t = 0$, where

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

If $A = (a_i^j)$, this means that

$$a_0^0 = 0, a_0^i = a_i^0, \quad i \neq 0, \quad a_j^i = -a_i^j, i,j \neq 0.$$

This can also be expresed as follows. Let us write

$$J = (g_{ij}),$$

and $A\cdot J = (a_{ij})$, where

$$a_{ij} = \sum_{k=0}^{3} a_i^k g_{kj}.$$

Then $A = (a_i^j)$ belongs to the Lie algebra of $O(1,3)$ if and only if the matrix $A\cdot J = (a_{ij})$ is a skew-symmetric matrix. In fact, this generalizes to any orthogonal group. Let $O(f)$

be the group of linear transformations preserving the quadratic form $f = g_{\nu\mu}x^\nu x^\mu$. Here we switched to superscript indices for the coordinate functions x_i. Then $A = (a_i^j)$ belongs to the Lie algebra of $O(f)$ if and only if $A \cdot (g_{\nu\mu}) = (a_i^\nu g_{\nu j})$ is skew-symmetric. The reason that we change the usual notation for the entries of A is that we consider A as an endomorphism of the linear space V, i.e., a tensor of type $(1,1)$. On the other hand, the quadratic form f is considered as a tensor of type $(0,2)$ (x_i forming a basis of the dual space). The contraction map $(V^*\otimes V)\otimes(V^*\otimes V^*) \to V^*\otimes V^*$ corresponds to the product $A\cdot(g_{ij})$ and has to be considered as a quadratic form, hence the notation a_{ij}.

Under the natural action of $GL(n,\mathbb{R}) \to \mathbb{R}^n, A \to A\cdot \mathbf{v}$, the matrix $A = (a_i^j)$ goes to

$$(v_1,\ldots,v_n)\cdot A^t = (\sum_{j=1}^n a_1^j v_j, \ldots, \sum_{j=1}^n a_n^j v_j).$$

Under this map the coordinate function x^i on $\mathbb{R}^n$ is pulled back to the function $\sum_{j=1}^n x_i^j c_j$ on $\mathrm{Mat}_n(\mathbb{R})$, where x_i^j is the coordinate function on $Mat_n(\mathbb{R})$ whose value on the matrix (δ_{ij}^{kl}) is equal to δ_{ij}^{kl}. Using (13.17), we see that the vector field $\frac{\partial}{\partial x_i^j}$ on $GL_n(\mathbb{R})$ goes to the vector field $x^j\frac{\partial}{\partial x_i}$ (no summation here!). Therefore, the matrix A^t, identified with the vector field $\eta = \sum_{i,j=1}^n a_j^i \frac{\partial}{\partial x_i^j}$, goes to the vector field

$$\eta^\sharp = \sum_{i=1}^n(\sum_{j=1}^n a_j^i x^j)\frac{\partial}{\partial x_i} = a_\mu^\nu x^\mu \partial_\nu = a^\nu\partial_\nu,$$

where $a^\nu = a_\rho^\nu x^\rho$, $a_{\rho\nu} = -a_{\nu\rho}$.

Note that

$$\sum_{\nu=0}^3 \partial_\nu(a^\nu) = \sum_{\rho=0}^3 a_\rho^\rho = 0,$$

so that the group $O(1,3)$ preserves the standard volume in $\mathbb{R}^3$.

We can rewrite equation (13.20) as follows

$$0 = \partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}x^\alpha a_\alpha^\nu\partial_\nu\phi - \delta_\nu^\mu x^\alpha a_\alpha^\nu\mathcal{L}\big) = a_\alpha^\nu\partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}x^\alpha\partial_\nu\phi - \delta_\nu^\mu x^\alpha\mathcal{L}\big) =$$

$$= g_{\beta\nu}a_\alpha^\nu\partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}g^{\beta\nu}x^\alpha\partial_\nu\phi - g^{\beta\nu}\delta_\nu^\mu x^\alpha\mathcal{L}\big) = a_{\beta\alpha}\partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}x^\alpha\partial^\beta\phi - \delta^{\beta\mu}x^\alpha\mathcal{L}\big) =$$

$$= \sum_{0\le\beta<\alpha\le 3}^n a_{\beta\alpha}\partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(x^\alpha\partial^\beta\phi - x^\beta\partial^\alpha\phi) - (\delta^{\beta\mu}x^\alpha - \delta^{\alpha\mu}x^\beta)\mathcal{L}\big), \tag{13.24}$$

where $(g^{\mu\nu}) = (g_{\mu\nu})^{-1}$. Since the coefficients $a_{\beta\alpha}, \beta<\alpha$, are arbitrary numbers, we deduce that

$$\partial_\mu\big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}(x^\alpha\partial^\beta\phi - x^\beta\partial^\alpha\phi) - (\delta^{\beta\mu}x^\alpha - \delta^{\alpha\mu}x^\beta)\mathcal{L}\big) = 0.$$

Let

$$T^{\beta\mu} = g^{\beta\nu}T^{\mu}_{\nu} = \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial^\beta\phi - \delta^{\beta\mu}\mathcal{L}$$

where (T^{μ}_{ν}) is the energy-momentum tensor (13.23). Set

$$\mathcal{M}^{\mu,\alpha\beta} = T^{\beta\mu}x^\alpha - T^{\alpha\mu}x^\beta. \tag{13.25}$$

Now the equation $\mathrm{div}(\mathbf{J}) = 0$ is equiavlent to

$$\partial_\mu\mathcal{M}^{\mu,\alpha\beta} = \sum_{\mu=1}^{n}\partial_\mu\mathcal{M}^{\mu,\alpha\beta} = 0.$$

We can introduce the charge

$$M^{\alpha\beta} = \int_{\mathbb{R}^3}\mathcal{M}^{0,\alpha\beta}d^3x.$$

Then it is conserved

$$\frac{dM^{\alpha\beta}}{dt} = 0.$$

When we restrict the values α, β to $1, 2, 3$, we obtain the conservation of angular momentum. For this reason, the tensor $(\mathcal{M}^{\mu,\alpha\beta})$ is called the *angular momentum tensor.*

13.8 The Klein-Gordon equation was proposed as a relativistic version of Schrödinger equation by E. Schrödinger, Gordon and O. Klein in 1926-1927. The original idea was to use it to construct a relativistic 1-particle theory. However, it was abandoned very soon due to several difficulties.

The first difficulty is that the equation admits obvious solutions

$$\phi(x,t) = e^{i(k\cdot x + Et)}$$

with

$$|k|^2 + m^2 = k_1^2 + k_2 + k_3^2 + m^2 - E^2 = 0. \tag{13.26}$$

Since the Schrödinger equation (5.6) should give us

$$i\frac{\partial\phi}{\partial t} = -E\phi = H\phi,$$

this implies that $-E$ is the eigenvalue of the Hamiltonian. Thus we have to interpret $-E$ as the energy. However, equation (13.26) admits positive solutions for E, so this leads to the negative energy of a system consisting of one particle. We, of course, have seen this already in the case of a particle in central field, for example, the electron in the hydrogen atom. However in that case, the energy was bounded from below. In our case, the energy spectrum is not bounded from below. This means we can extract an arbitrarily large amount of energy from our system. This is difficult to assume. In the next Lectures we shall show how to overcome this difficulty by applying second quantization.

Exercises.

1. Let $\tilde{g} : E \to E$ be a lift of a diffeomorphism $g : M \to M$ to a vector bundle over M. Show that the dual bundle (as well as the exterior and tensor powers) admits a natural lift too.

2. Generalize the Euler-Lagrange equation (13.8) to the case when the Lagrangian depends on $x \in \mathbb{R}^n$.

3. Let $\mathbf{J} = (\frac{\partial \mathcal{L}}{\partial \partial_t \phi}\partial_t \phi - \mathcal{L}, \frac{\partial \mathcal{L}}{\partial \partial_{x_1} \phi}\partial_t \phi, \frac{\partial \mathcal{L}}{\partial \partial_{x_2} \phi}\partial_t \phi, \frac{\partial \mathcal{L}}{\partial \partial_{x_3} \phi}\partial_t \phi)$, where $\mathcal{L}$ is the Klein-Gordon Lagrangian (13.10). Verify directly that $\mathrm{div}(\mathbf{J}) = 0$. What symmetry transformation in the Lagrangian gives us this conserved current?

4. Let $M = \mathbb{R}$ and $\mathcal{L}(\phi)(x) = \phi(x) + 2\log|\phi'(x)| - \log|\phi''(x)|$. Find the Euler-Lagrange equation for this Lagrangian. Note that it depends on the second derivative so this case was discussed in the lecture. Observing that $\mathcal{L}$ is invariant with respect to transformations $x \to \lambda x$, where $\lambda \in \mathbb{R}^*$, find the corresponding conservation law.

5. The Klein-Gordon Lagrangian for complex scalar fields $\phi : \mathbb{R}^4 \to \mathbb{C}$ has the form $\mathcal{L}(\phi) = \partial_\mu \phi \partial^\mu \bar{\phi} - m^2|\phi|^2$. Obviously it is invariant with respect to transformations $\phi \to \lambda\phi$, where $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. Find the corresponding current vector $\mathbf{J}$ and the conserved charge.

Lecture 14. YANG-MILLS EQUATIONS

These equations arise as the Euler-Lagrange equations for gauge fields with a special Lagrangian. A particular case of these equations is the Maxwell equations for electromagnetic fields.

14.1 Let G be a Lie group and P be a principal G-bundle over a d-dimensional manifold M. Recall that a *gauge field* is a connection on P. It is given by a 1-form A on M with values in the adjoint affine bundle $Ad(P) \subset End(\mathfrak{g})$. Here affine means that the transition functions belong to the affine group of the Lie algebra $\mathfrak{g}$ of G. More precisely, the transition functions are given by

$$A_e = g \cdot A_{e'} \cdot g^{-1} + g^{-1} dg,$$

where A_e and $A_{e'}$ are the $\mathfrak{g}$-valued 1-forms on M corresponding to A via two different trivializations of P (*gauges*).

The curvature of a gauge field A is a section F_A of the vector bundle $\Lambda^2(ad(E)) = \bigwedge^2 T^*(M) \otimes Ad(E)$. It can be defined by the formula

$$F_A = dA + \frac{1}{2}[A, A].$$

If we fix a gauge, then $F = F_A$ is given by an $n \times n$ matrix whose entries F_i^j are smooth 2-forms on U. If $(x^1, \ldots, x^d)$ is a system of local parameters in U, then we can write

$$F_i^j = \sum_{\mu,\nu=1}^{d} F_{i,\mu\nu}^j(x) dx^\mu \wedge dx^\nu$$

Equivalently we can write

$$F = (F_{\mu\nu})_{1 \leq \mu,\nu \leq d},$$

where $F_{\mu\nu}$ is the matrix with (ij)-entry equal to $F_{i,\mu\nu}^j(x)$. We can start with any $Ad(P)$-valued 2-form F as above, and say that a gauge field A is a *potential* gauge field for F if $F = F_A$.

Let $E = E(\rho) \to M$ be an associated vector bundle on M with respect to some linear representation $\rho : G \to GL(V)$. The corresponding representation of the Lie algebra $d\rho : \mathfrak{g} \to End(V)$ defines a homomorphism of vector bundles $Ad(P) \to End(E)$. Applying this homomorphism to the values of A and F_A, we obtain the notion of a connection and its curvature in E.

For example, let us take $G = U(1)$ and let P be the trivial principal G-bundle. In this case $Ad(P)$ is the trivial bundle $M \times \mathbb{R}$. So a gauge field is a 1 form $\sum_{\mu=0}^{3} A_\mu dx^\mu$ which can be identified with a vector function

$$A = (A_0, A_1, A_2, A_3).$$

For any smooth function $\phi : M \to U(1)$, the 1-form

$$A' = A + d\log\phi$$

defines the same connection. Its curvature is a smooth 2-form

$$F = \sum_{\lambda,\nu=0}^{3} F_{\mu\nu} dx^\mu \wedge dx^\nu.$$

We shall identify it with the skew symmetric matrix $(F_{\mu\nu})$ whose entries are smooth functions in (t, x):

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & H_3 & -H_2 \\ E_2 & -H_3 & 0 & H_1 \\ E_3 & H_2 & -H_1 & 0 \end{pmatrix}. \tag{14.1}$$

For a reason which will be clear later, this is called the *electromagnetic tensor*. Since $[A, A] = 0$, we have

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{14.2}$$

Let us write

$$A = (\phi, \mathbf{A}) \in C^\infty(\mathbb{R}^4) \times (C^\infty(\mathbb{R}^4))^3.$$

This gives

$$F_{0\nu} = \partial_0 A_\nu - \partial_\nu \phi, \nu = 0, 1, 2, 3,$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad \mu, \nu = 1, 2, 3.$$

If we set

$$\mathbf{E} = (E_1, E_2, E_3), \quad \mathbf{H} = (H_1, H_2, H_3),$$

we can rewrite the previous equalities in the form

$$\mathbf{E} = \mathrm{grad}_x \phi - \partial_0 \mathbf{A} = (\partial_1\phi, \partial_2\phi, \partial_3\phi) - (\partial_0 A_1, \partial_0 A_2, \partial_0 A_3), \tag{14.3}$$

$$\mathbf{H} = \mathrm{curl}\ \mathbf{A} = \nabla \times (A_1, A_2, A_3), \tag{14.4}$$

where

$$\nabla = (\partial_1, \partial_2, \partial_3).$$

Of course, equation (14.2) means that the differential form F_A satisfies

$$F_A = d(A_0 dx^0 + A_1 dx^1 + A_2 dx^2 + A_3 dx^3).$$

Since $\mathbb{R}^4$ is simply-connected, this occurs if and only if

$$dF = d(-E_1 dx^0 \wedge dx^1 - E_2 dx^0 \wedge dx^2 - E_3 dx^0 \wedge dx^3 + H_3 dx^1 \wedge dx^2 - H_2 dx^1 \wedge dx^3 + H_1 dx^2 \wedge dx^3)$$

$$= (-\partial_2 E_1 + \partial_1 E_2 + \partial_0 H_3) dx^0 \wedge dx^1 \wedge dx^2 + (-\partial_3 E_1 + \partial_1 E_3 - \partial_0 H_2) dx^0 \wedge dx^1 \wedge dx^3 +$$

$$+ (-\partial_3 E_2 + \partial_1 E_3 + \partial_0 H_3) dx^0 \wedge dx^2 \wedge dx^3 + \mathrm{div} H dx^1 \wedge dx^2 \wedge dx^3 = 0.$$

This is equivalent to

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{H}}{\partial t} = 0, \quad (M1)$$

$$\mathrm{div}\mathbf{H} = \nabla \cdot \mathbf{H} = 0. \quad (M2)$$

This is the first pair of Maxwell's equations. The second pair will follow from the Euler-Lagrange equation for gauge fields.

14.2 Let us now introduce the Yang-Mills Lagrangian defined on the set of gauge fields. Its definition depends on the choice of a pseudo-Riemannian metric g on M which is locally given by

$$g = \sum_{\nu,\mu=1}^{n} g_{\mu\nu} dx^\nu \otimes dx^\mu.$$

Its value at a point $x \in M$ is a non-degenerate quadratic form on $T(M)_x$. The dual quadratic form on $T^*(M)_x$ is given by the inverse matrix $(g^{\mu\nu}(x))$. It defines a symmetric tensor

$$g^{-1} = \sum_{\nu,\mu=1}^{n} g^{\mu\nu} \frac{\partial}{\partial x^\nu} \otimes \frac{\partial}{\partial x^\mu}.$$

We can use g to transform the $\mathfrak{g}$-valued 2-form $F = (F_{\mu\nu})$ to the $\mathfrak{g}$-valued vector field

$$\hat{F} = (F^{\mu\nu}) = (g^{\mu\alpha} g^{\nu\beta} F_{\beta\alpha}) = \sum_{\mu,\nu=1}^{n} F^{\mu\nu} \frac{\partial}{\partial x^\nu} \otimes \frac{\partial}{\partial x^\mu}.$$

Let $\mathfrak{g} \otimes \mathfrak{g} \to \mathbb{R}$ be the Killing form on $\mathfrak{g}$. It is defined by

$$\langle A, B \rangle = -Tr(ad(A) \circ ad(B)),$$

where $ad : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ is the adjoint representation, $ad(A) : X \to [A, X]$. This is a bilinear form which is invariant with respect to the adjoint representation $Ad : G \to GL(\mathfrak{g})$. In the case when G is semisimple, the Killing form is non-degenerate (in fact, this is one of many equivalent definitions of the semi-simplicity of a Lie group). If furthermore, G is compact (e.g. $G = SU(n)$), it is also positive definite. The latter explains the minus sign in the formula.

Now we can form the scalar function

$$\langle F, \hat{F} \rangle := \sum_{\mu,\nu=1}^{n} \langle F_{\mu\nu}, F^{\mu\nu} \rangle. \tag{14.5}$$

It is easy to see that this expression does not depend on the choice of coordinate functions. Neither does it depend on the choice of the gauge. Now if we choose the volume form $\mathrm{vol}(g)$ on M associated to the metric g (we recall its definition in the next section), we can integrate (14.5) to get a functional on the set of gauge fields

$$S_{YM}(A) = \int_M \langle F, \hat{F} \rangle \mathrm{vol}(g). \tag{14.6}$$

This is called the *Yang-Mills action* functional.

14.3 One can express (14.5) differently by using the star operator $*$ defined on the space of differential forms. Recall its definition. Let V be a real vector space of dimension n, and $g : V \times V \to \mathbb{R}$ (equivalently, $g : V \to V^*$) be a metric on V, i.e. a symmetric non-degenerate form on V. Let $e_1, \ldots, e_n$ be an orthonormal basis in V with respect to $\mathbb{R}$. This means that $g(e_i, e_j) = \pm\delta_{ij}$. Let $e^1, \ldots, e^n$ be the dual basis in V^*. The n-form $\mu = e^1 \wedge \ldots \wedge e^n$ is called the *volume form* associated to g. It is defined uniquely up to ± 1. A choice of the two possible volume elements is called an orientation of V. We shall choose a basis such that $\mu(e_1, \ldots, e_n) > 0$. Such a basis is called positively oriented.

Let W be a vector space with a bilinear form $B : W \times W \to \mathbb{R}$. We extend the operation of exterior product $\bigwedge^k V \times \bigwedge^s V \to \bigwedge^{k+s} V$ by defining

$$(\bigwedge^k V \otimes W) \times (\bigwedge^s V \otimes W) \to \bigwedge^{k+s} V$$

using the formula

$$(\alpha \otimes w) \wedge (\alpha' \otimes w') = B(w, w')\alpha \wedge \alpha'.$$

Also, if $g : V \to V$ is a bilinear form on V, we extend it to a bilinear form on $\bigwedge^k V \otimes W$ by setting

$$g(\alpha \otimes w, \alpha' \otimes w') = \wedge^k(g)(\alpha, \alpha')B(w, w'),$$

where we consider g as a linear map $V \to V^*$ so that $\wedge^k(g)$ is the linear map $\bigwedge^k V \to \bigwedge^k V^*$ which is considered as a bilinear form on $\bigwedge^k V$.

Lemma 1. *Let g be a metric on V and $\mu \in \bigwedge^n(V^*)$ be a volume form associated to g. Let W be a finite-dimensional vector space with a non-degenerate bilinear form $B : W \times W \to \mathbb{R}$. Then there exists a unique linear isomorphism $* : \bigwedge^k V^* \otimes W \to \bigwedge^{n-k} V^* \otimes W$, such that, for all $\alpha, \beta \in Hom(\bigwedge^k V, W)$,*

$$\alpha \wedge *\beta = g^{-1}(\alpha, \beta)\mu. \tag{14.7}$$

Here $g^{-1} : V^* \times V^* \to \mathbb{R}$ is the inverse metric.

Proof. Consider $\gamma \in \bigwedge^{n-k} V^* \otimes W$ as a linear function $\phi_\gamma : \bigwedge^k V^* \otimes W \to \mathbb{R}$ which sends $\omega \in \bigwedge^k V^* \otimes W$ to the number $\phi_\gamma(\omega)$ such that $\gamma \wedge \omega = \phi_\gamma(\omega)\mu$. It is easy to see that $\beta \to \phi_\beta$ establishes a linear isomorphism from $\bigwedge^{n-k} V^* \otimes W$ to $(\bigwedge^k V^* \otimes W)^* \cong \bigwedge^k V \otimes W$. On the other hand, we have an isomorphism $\wedge^k(g^{-1}) : \bigwedge^k V^* \otimes W \to \bigwedge^k V \otimes W$. Thus if, for a fixed $\beta \in \bigwedge^k V^* \otimes W$ we define a linear function $\alpha \to g^{-1}(\alpha, \beta)$, there exists a unique $*\beta \in \bigwedge^{n-k} V^* \otimes W$ such that (14.7) holds.

The assertion of the previous lemma can be "globalized" by considering metrics $g : T(M) \to T^*(M)$ and vector bundles over oriented manifolds. We obtain the definition of the associated volume form $\mathrm{vol}(g)$, and the star operator

$$* : \Lambda^k(E) \to \Lambda^{n-k}(E)$$

satisfies (14.7) at each point $x \in M$. Here we assume that the vector bundle E is equipped with a bilinear map $B : E \times E \to 1_M$ such that the corresponding morphism $E \to E^*$ is bijective. We also can define the star operator on the sheaves of sections of $\Lambda^k(E)$ to get a morphism of sheaves:

$$* : \mathcal{A}^k(E) \to \mathcal{A}^{n-k}(E).$$

We shall apply this to the special case when E is the adjoint bundle. Its typical fibre is a Lie algebra $\mathfrak{g}$ and the bilinear form is the Killing form. The non-degeneracy assumption requires us to assume further that $\mathfrak{g}$ is a semi-simple Lie algebra.

In particular, taking $\mathfrak{g} = \mathbb{R}$, we obtain that $\mathcal{A}^k(E)(M) = \mathcal{A}^k(M)$ is the space of k-differential forms and $*$ is the usual star operator introduced in analysis on manifolds. Note that in this case

$$*1 = \mathrm{vol}(g). \tag{14.8}$$

In our situation, $k = 2$ and we have $F_A \in \mathcal{A}^2(Ad(P)(M), *F_A \in \mathcal{A}^{n-2}(Ad(P)(M)$ and

$$F_A \wedge *F_A = g^{-1}(F_A, F_A)\mathrm{vol}(g) = \langle F, \hat{F} \rangle \mathrm{vol}(g).$$

Here we use the following explicit expression for $g^{-1}(\alpha, \beta), \alpha, \beta \in \bigwedge^k(V^*) \otimes W$:

$$g^{-1}(\alpha, \beta) = \sum g^{i_1 j_1} \ldots g^{i_k j_k} B(\alpha_{i_1 \ldots i_k}, \beta_{j_1 \ldots j_k}),$$

where $\alpha = \sum e^{i_1} \wedge \ldots e^{i_k} \otimes \alpha_{i_1 \ldots i_k} \in \bigwedge^k(V^*) \otimes W$, and there is a similar expression for β.

Assume that we have a coordinate system where $g_{ij} = \pm\delta_{ij}$ (a *flat* coordinate system). Then

$$g^{-1}(dx^{i_1} \wedge \ldots \wedge dx^{i_k}, dx^{j_1} \wedge \ldots \wedge dx^{j_k}) = g^{i_1 j_1} \ldots g^{i_k j_k}.$$

This implies that

$$*(F_{i_1 \ldots i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k}) = g^{i_1 i_1} \ldots g^{i_k i_k} F_{i_1 \ldots i_k} dx^{j_1} \wedge \ldots \wedge dx^{j_{n-k}}, \tag{14.9}$$

where

$$dx^{i_1} \wedge \ldots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_{n-k}} = dx^1 \wedge \ldots \wedge dx^n.$$

If $g^{ij} = \delta_{ij}$ then this formula can be written in the form

$$(*F)_{j_1 \ldots j_{n-k}} = \epsilon(i_1, \ldots, i_k, j_1, \ldots, j_{n-k}) F_{i_1 \ldots i_k} dx^{j_1} \wedge \ldots \wedge dx^{j_{n-k}},$$

where $\epsilon(i_1, \ldots, i_k, j_1, \ldots, j_{n-k})$ is the sign of the permutation $(i_1, \ldots, i_k, j_1, \ldots, j_{n-k})$.

If G is a compact semi-simple group we may consider the unitary inner product in the space $\mathcal{A}^k(Ad(P))(M)$

$$\langle F_1, F_2 \rangle = \int_M F_1 \wedge *F_2. \tag{14.10}$$

Lemma 2. *Assume M is compact, or F, G vanish on the boundary of M. Let A be a connection on P. Then, for any $F_1 \in \mathcal{A}^k(Ad(P))(M), F_2 \in \mathcal{A}^{k+1}(Ad(P))(M)$,*

$$\langle d^A(F_1), F_2 \rangle = (-1)^{k+1} \langle F_1, d^A * F_2 \rangle.$$

Proof. By Theorem 1 from Lecture 12, we have $d^A(F_1) = dF_1 + [A, F_1]$. Since $d(F_1 \wedge *F_2)$ is a closed n-form, by Stokes' theorem, its integral over M is equal to the integral of $F_1 \wedge *F_2$ over the boundary of M. By our assumption, it must be equal to zero. Also, since the Killing form $\langle \ , \ \rangle$ is invariant with respect to the adjoint representation, we have $\langle [a, b], c \rangle = \langle [c, a], b \rangle$ (check it in the case $\mathfrak{g} = \mathfrak{gl}_n$ where we have $Tr((ab - ba)c) = Tr(abc - bac) = Tr(abc) - Tr(bac) = Tr(bca) - Tr(bac) = Tr(b(ca - ac)) = \langle b, [c, a] \rangle$). Using this, we obtain

$$[A, F_1] \wedge *F_2 = (-1)^{k+1} F_1 \wedge [A, *F_2].$$

Collecting this information, we get

$$\langle d^A(F_1), F_2 \rangle = \langle dF_1 + [A, F_1], F_2 \rangle = \int_M dF_1 \wedge *F_2 + [A, F_1] \wedge *F_2 =$$

$$\int_M d(F_1 \wedge *F_2) - (-1)^k (F_1 \wedge d * F_2 + F_1 \wedge [A, *F_2])$$

$$= \int_M dF_1 \wedge *F_2 - (-1)^k \int_M F_1 \wedge d * F_2 + F_1 \wedge [A, *F_2] =$$

$$= (-1)^{k+1} \int_M F_1 \wedge d^A * F_2 = (-1)^{k+1} \langle F_1, d^A * F_2 \rangle.$$

14.4 Let us find the Euler-Lagrange equation for the Yang-Mills action functional

$$S_{YM}(A) = \int_M \langle F_A, \hat{F}_A \rangle \mathrm{vol}(g) = \int_M F_A \wedge *F_A = ||F_A||^2,$$

where the norm is taken in the sense of (14.10). We know that the difference of two connections is a 1-form with values in $Ad(P)$. Applying Theorem 1 (iii) from lecture 12, we have, for any $h \in \mathcal{A}^1(Ad(P))$ and any $s \in \Gamma(E)$,

$$F_{A+h} = d(A+h) + \frac{1}{2}[A+h, A+h] = F_A + dh + [A,h] + o(||h||) = F_A + d^A(h).$$

Now, ignoring terms of order $o(||h||)$, we get

$$S_{YM}(A+h) - S_{YM}(A) = ||F_{A+h}||^2 - ||F_A||^2 = ||F_{A+h} - F_A||^2 + 2\langle F_A, F_{A+h} - F_A\rangle =$$

$$= 2\int_M d^A h \wedge *F_A = -2\int_M h \wedge d^A(*F_A).$$

Here, at the last step, we have used Lemma 2. This implies the equation for a critical connection A:

$$d^A(*F_A) = 0 \tag{14.11}$$

It is called the *Yang-Mills* equation. Note, by Bianchi's identity, we always have $d^A(F_A) = 0$.

In flat coordinates we can use (14.9) to find an explicit expression for the coordinate functions $F_{\mu\nu}$ of $*F$. Then explicitly equation (14.11) reads

$$\sum_\mu \frac{\partial F_{\mu\nu}}{\partial x^\nu} + [A_\nu, F_{\mu\nu}] = 0. \tag{14.12}$$

This is a second-order differential equation in the coordinate functions A_ν of the connection A.

Definition. A connection satisfying the Yang-Mills equation (14.11) is called a *Yang-Mills connection.*

14.5 Note the following property of the star operator $* : \bigwedge^k V^* \to \bigwedge^{n-k} V^*$:

$$* \circ * = \epsilon(g)(-1)^{k(n-k)}\mathbf{id}, \tag{14.13}$$

where $\epsilon(g)$ is equal to the determinant of the Gram-Schmidt matrix of the metric g with respect to some orthonormal basis.

In particular, if $n = 4$ and $k = 2$, and g is a Riemannian metric, we have

$$*^2 = \mathbf{id}.$$

This allows us to decompose the space $\mathcal{A}^2(Ad(P))(M)$ into two eigensubspaces

$$\mathcal{A}^2(Ad(P))(M) = \mathcal{A}^2(Ad(P))(M)_+ \oplus \mathcal{A}^2(Ad(P))(M)_-.$$

A connection A with $F_A \in \mathcal{A}^2(Ad(P))(M)_+$ (resp. $F_A \in \mathcal{A}^2(Ad(P))(M)_-$) is called *self-dual* (resp. *anti-self-dual*) connection. Let us write $F_A = F_A^+ + F_A^-$, where

$$*(F_A^{\pm}) = \pm F_A.$$

Thus we have

$$S_{YM}(A) = ||F_A^+ + F_A^-||^2 = ||F_A^+||^2 + ||F_A^-||^2 + 2\langle F_A^+, F_A^-\rangle = ||F_A^+||^2 + ||F_A^-||^2, \quad (14.14)$$

where we use that $\langle F_A^+, F_A^-\rangle = -\langle F_A^+, *F_A^-\rangle = -\langle *F_A^+, F_A^-\rangle = -\langle F_A^+, F_A^-\rangle$.

Recall that by the De Rham theorem

$$H^k(M,\mathbb{R}) = \mathrm{Ker}(d : \mathcal{A}^k(M) \to \mathcal{A}^{k+1}(M))/\mathrm{Im}(d : \mathcal{A}^{k-1}(M) \to \mathcal{A}^k(M)).$$

Let E be a complex rank r vector bundle associated to a principal G-bundle P. Pick a connection form A on P and let Let F_A be the curvature form of the associated connection. When we trivialize E over an open subset U, we can view F_A as a $r \times r$ matrix X with coefficients in $\mathcal{A}^2(U)$. Let $T = (T_{ij}), i, j = 1, \ldots, r$ be a matrix whose entries are variables T_{ij}. Consider the characteristic polynomial

$$\det(T - \lambda I_r) = (-\lambda)^r + a_1(T)(-\lambda)^{r-1} + \ldots + a_r(T).$$

Its coefficients are homogeneous polynomials in the variables T_{ij} which are invariant with respect to the conjugation transformation $T \to C \cdot T \cdot C^{-1}, C \in GL(n, K)$, where K is any field. If we plug in the entries of the matrix F_A^{ρ} in $a_k(T)$ we get a differential form $a_k(F_A)$ of degree $2k$. By the invariance of the polynomials a_k, the form $a_k(F_A)$ does not depend on the choice of trivialization of E. Moreover, it can be proved that the form $a_k(F_A)$ is closed. After we rescale F_A by multiplying it by $\frac{i}{2\pi}$ we obtain the cohomology class

$$c_k(E) = [a_k(\frac{i}{2\pi}F_A)] \in H^{2k}(M,\mathbb{R}).$$

It is called the kth Chern class of E and is denoted by $c_k(E)$. One can prove that it is independent of the choice of the connection A, so it can be denoted by $c_k(E)$. Also one proves that

$$c_k(E) \in H^{2k}(M,\mathbb{Z}) \subset H^{2k}(M,\mathbb{R}).$$

Assume $G = SU(2)$ and take $E = E(\rho)$, where ρ is the standard representation of G in $\mathbb{C}^2$. Then $\mathrm{Lie}(G)$ is isomorphic to the algebra of matrices $X \in M_2(\mathbb{C})$ with $X^t + \bar{X} = 0$. The Killing form is equal to $-4Tr(AB)$. Then

$$F_A^{\rho} = \begin{pmatrix} F_1^1 & F_1^2 \\ F_2^1 & F_2^2 \end{pmatrix},$$

where $F_1^1 + F_2^2 = 0, F_1^2 = \bar{F}_2^1, F_1^1 \in i\mathbb{R}$. This shows that

$$c_1(E) = 0,$$

$$c_2(E) = -\frac{1}{4\pi^2}[\det F_A] = -\frac{1}{32\pi^2}[F_A \wedge F_A],$$

Here we use that $Tr(X^2) = -2\det(X)$ for any 2×2 matrix X with zero trace. Now, omitting the subscript A,

$$F \wedge F = (F^+ + F^-) \wedge (F^+ + F^-) = F^+ \wedge F^+ + F^- \wedge F^- + F^+ \wedge F^- + F^- \wedge F^+ =$$

$$= F^+ \wedge F^+ + F^- \wedge F^- = F^+ \wedge *F^+ - F^- \wedge *F^-.$$

Here we use that

$$F^+ = F_{01}(dx^0 \wedge dx^1 + dx^2 \wedge dx^3) + F_{02}(dx^0 \wedge dx^2 + dx^3 \wedge dx^1) + F_{03}(dx^0 \wedge dx^3 + dx^1 \wedge dx^2),$$

$$F^- = F_{01}(dx^0 \wedge dx^1 - dx^2 \wedge dx^3) + F_{02}(dx^0 \wedge dx^2 - dx^3 \wedge dx^1) - F_{03}(dx^0 \wedge dx^3 + dx^1 \wedge dx^2),$$

which easily implies that $F^+ \wedge F^- = 0$. From this we deduce

$$-32\pi^2 k(A) = \int_M (F_A^+ \wedge *F_A^+ - F_A^- \wedge *F_A^-) = ||F_A^+||^2 - ||F_A^-||^2,$$

where

$$k(E) = \int_M c_2(E). \tag{14.15}$$

Hence, using (14.14), we obtain

$$S_{YM}(A) = 2||F_A^+||^2 + (||F_A^-||^2 - ||F_A^+||^2) \geq 32\pi^2 k(A) \quad \text{if } k(E) > 0$$

with equality iff $F^+ = 0$. Similarly,

$$S_{YM}(A) = 2||F_A^-||^2 + (||F_A^+||^2 - ||F_A^-||^2) \geq -32\pi^2 \quad \text{if } k(E) < 0$$

with equality iff $F^- = 0$. Also, we see that A is a Yang-Mills connection if it is self-dual or anti-self-dual. The condition for this is the first order *self-dual (anti-self-dual) Yang-Mills equation*

$$F = \pm * F. \tag{14.16}$$

Definition. An *instanton* is a principal $SU(2)$-bundle P over a four-dimensional Riemannian manifold with $k(Ad(P)) > 0$ together with an anti-self-dual connection. The number $k(Ad(P)))$ is called the *instanton number.*

Note that all complex vector G-bundles of the same rank and the same Chern classes are isomorphic. So, let us fix one principal $SU(2)$-bundle P with given instanton number k and consider the set of all anti-self-dual connections on it. The group of gauge transformations acts naturally on this set, we can consider the moduli space

$$\mathcal{M}_k(M) = \{\text{anti-self-dual connections on } P\}/\mathcal{G}.$$

The main result of Donaldson's theory is that this moduli space is independent of the choice of a Riemannian metric and hence is an invariant of the smooth structure on M. Also, when M is a smooth structure on a nonsingular algebraic surface, the moduli space $\mathcal{M}_k(M)$ can be identified with the moduli space of stable holomorphic rank 2 bundles with first Chern class zero and second Chern class equal to k.

14.6 It is time to consider an example. Let

$$M = S^4 := \{(x_1, \ldots, x_5) \in \mathbb{R}^5 : \sum_{i=1}^{5} x_i^2 = 1\}$$

be the four-dimensional unit sphere with its natural metric induced by the standard Euclidean metric of $\mathbb{R}^5$. Let $N_+ = (0,0,0,0,1)$ be its "north pole" and $N_- = (0,0,0,0,-1)$ be its "south pole". By projecting from the poles to the hyperplane $x_5 = 0$, we obtain a bijective map $S^4 \setminus \{N_\pm\} \to \mathbb{R}^4$. The precise formulae are

$$(u_1^\pm, u_2^\pm, u_3^\pm, u_4^\pm) = \frac{1}{1 \pm x_5}(x_1, x_2, x_3, x_4).$$

This immediately implies that

$$(\sum_{i=1}^{4}(u_i^+)^2)(\sum_{i=1}^{4}(u_i^-)^2 = \frac{\sum_{i=1}^4 x_i^2}{(1-x_5^2)} = \frac{1-x_5^2}{1-x_5^2} = 1.$$

After simple transformations, this leads to the formula

$$u_i^- = \frac{u_i^+}{\sum_{i=1}^4 (u_i^+)^2}, \quad i = 1,2,3,4. \tag{14.17}$$

Thus we can cover S^4 with two open sets $U_+ = S^4 \setminus \{N_+\}$ and $U_- = S^4 \setminus \{N_-\}$ diffeomeorphic to $\mathbb{R}^4$. The transition functions are given by (14.17). We can think of S^4 as a compactification of $\mathbb{R}^4$. If we identify the latter with U_+, then the point N_+ corresponds to the point at infinity ($u_i^+ \to \infty$). This point is the origin in the chart U_-.

Let P be a principal $SU(2)$-bundle over S^4. It trivializes over U_+ and U_-. It is determined by a transition function $g : U_+ \cap U_- \to SU(2)$. Restricting this function to the sphere $S^3 = S^4 \cap \{x_5 = 0\}$, we obtain a smooth map

$$g : S^3 \to SU(2).$$

Both S^3 and $SU(2)$ are compact diffeomorphic three-dimensional manifolds. The homotopy classes of such maps are classified by the degree $k \in \mathbb{Z}$ of the map. Now, $H^2(S^4, \mathbb{Z}) = 0$, so the first Chern class of P must be zero. One can verify that the number k can be identified with the second Chern number defined by (14.15).

Let us view $\mathbb{R}^4$ as the algebra $\mathbb{H}$ of quaternions $x = x_1 + x_2\mathbf{i} + x_2\mathbf{j} + x_4\mathbf{k}$. The Lie algebra of $SU(2)$ is equal to the space of complex 2×2-matrices X satisfying $A^* = -A$. We can write such a matrix in the form

$$\begin{pmatrix} ix_2 & x_3 + ix_4 \\ -(x_3 - ix_4) & -ix_2 \end{pmatrix}$$

and identify it with the pure quaternion $x_2\mathbf{i} + x_2\mathbf{j} + x_4\mathbf{k}$. Then we can view the expression

$$dx = dx_1 + dx_2\mathbf{i} + df\mathbf{j} + dx_4\mathbf{k}$$

as quaternion valued differential 1-form. Then

$$d\bar{x} = dx_1 - dx_2\mathbf{i} - df\mathbf{j} - dx_4\mathbf{k}$$

has clear meaning. Now we define the gauge potential (=connection) A as the 1-form on U_+ given by the formula

$$A(x) = \mathrm{Im}\Big(\frac{xd\bar{x}}{1 + |x|^2}\Big) = \frac{1}{2}\frac{xd\bar{x} - dx\bar{x}}{1 + |x|^2}. \tag{14.18}$$

Here we use the coordinates x_i instead of u_i^+. The coordinates A_μ of A are given by

$$A_1 = \frac{x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}}{1 + |x|^2}, \quad A_2 = \frac{x_1\mathbf{i} + x_4\mathbf{j} - x_3\mathbf{k}}{1 + |x|^2},$$

$$A_3 = \frac{-x_4\mathbf{i} - x_1\mathbf{j} + x_2\mathbf{k}}{1 + |x|^2}, \quad A_4 = \frac{x_3\mathbf{i} + x_2\mathbf{j} - x_1\mathbf{k}}{1 + |x|^2}.$$

Computing the curvature F_A of this potential, we get

$$F = \frac{d\bar{x} \wedge dx}{(1 + |x|^2)^2},$$

where

$$d\bar{x} \wedge dx = -2[(dx_1 \wedge dx_2 + dx_3 \wedge dx_4)\mathbf{i} + (dx_1 \wedge dx_3 + dx_4 \wedge dx_2)\mathbf{j} + (dx_1 \wedge dx_4 + dx_2 \wedge dx_3)\mathbf{k}].$$

Now formula (14.9) shows that F is anti-self-dual.

On the open set U_- the gauge potential is given by

$$A'(y) = \mathrm{Im}\Big(\frac{yd\bar{y}}{1 + |y|^2}\Big).$$

In quaternion notation, the coordinate change (14.7) is

$$x = \bar{y}^{-1}.$$

It is easy to verify the following identity

$$\bar{x}\Big(\frac{xd\bar{x}}{1+|x|^2}\bar{x}\Big)+\bar{x}d\bar{x}^{-1}=\frac{yd\bar{y}}{1+|y|^2}. \tag{14.19}$$

Let us define the map $U_+\cap U_-\to SU(2)$ by the formula

$$\phi(x)=\frac{\bar{x}}{|x|}$$

Then

$$\mathrm{Im}(\bar{x}d\bar{x}^{-1})=\phi(x)^{-1}d\phi(x)$$

and taking the imaginary part in (14.19), we obtain

$$A'(x)=\phi(x)\circ A(x)\circ\phi^{-1}(x)+\phi(x)^{-1}d\phi(x).$$

Note that the stereographic projection $S^4\to\mathbb{R}^4$ which we have used maps the subset $x_5=0$ bijectively onto the unit 3-sphere $S^3=\{x:|x|=1\}$. The restriction of the gauge map ϕ to S^3 is the identity map. Thus, its degree is equal to 1, and hence we have constructed an instanton over S^4 with instanton number $k=1$.

There is interesting relation between instantons over S^4 and algebraic vector bundles over $\mathbf{P}^3(\mathbb{C})$. To see it, we have to view S^4 as the quaternionic projective line $\mathbb{P}^1(\mathbb{H})$. For this we identify $\mathbb{C}^2$ with $\mathbb{H}$ by the map $(a+bi,c+di)\to(a+bi)+(c+di)j=a+bi+cj+dk$, then

$$S^4=\mathbb{H}^2\setminus\{0\}/\mathbb{H}^*=\mathbb{H}\cup\{\infty\}=\mathbb{R}^4\cup\{\infty\},$$

using this, we construct the map

$$\pi:\mathbb{P}^3(\mathbb{C})=\mathbb{C}^4\setminus\{0\}/\mathbb{C}^*=\mathbb{H}^2\setminus\{0\}/\mathbb{C}^*\to S^4=\mathbb{H}^2\setminus\{0\}/\mathbb{H}^*$$

with the fibre

$$\mathbb{H}^*/\mathbb{C}^*=\mathbb{C}^2\setminus\{0\}/\mathbb{C}^*=\mathbb{P}^1(\mathbb{C})=S^2.$$

Now the pre-image of the adjoint vector bundle $Ad(P)$ of an instanton $SU(2)$-bundle over S^4 is a rank 3 complex vector bundle over $\mathbb{P}^3(\mathbb{C})$. One shows that this bundle admits a structure of an algebraic vector bundle. We refer to [**Atiyah**] for details and further references.

14.7 Let us take $G=U(1),M=\mathbb{R}^4$. Take the Lorentzian metric g on M. Let

$$F=dA=\sum(\partial_\mu A_\nu-\partial_\nu A_\mu)dx^\mu\wedge dx^\nu,$$

and, by (14.7)

$$*F=F_{23}dx^0\wedge dx^1-F_{01}dx^2\wedge dx^3+F_{31}dx^0\wedge dx^2-F_{02}dx^3\wedge dx^1+F_{12}dx^0\wedge dx^3-F_{03}dx^1\wedge dx^2=$$

$$= H_1 dx^0 \wedge dx^1 + E_1 dx^2 \wedge dx^3 + H_2 dx^0 \wedge dx^2 + E_2 dx^3 \wedge dx^1 + H_3 dx^0 \wedge dx^3 + E_3 dx^1 \wedge dx^2.$$

It corresponds to the matrix

$$*F = \begin{pmatrix} 0 & H_1 & H_2 & H_3 \\ -H_1 & 0 & E_3 & -E_2 \\ -H_2 & -E_3 & 0 & E_1 \\ -H_3 & E_2 & -E_1 & 0 \end{pmatrix}.$$

It is obtained from matrix (14.1) of F by replacing $(\mathbf{E}, \mathbf{H})$ with $(\mathbf{H}, -\mathbf{E})$. Equivalently, if we form the complex electromagnetic tensor $\mathbf{H} + i\mathbf{E}$, then $*F$ corresponds to $i(\mathbf{H} + i\mathbf{E})$. The Yang-Mills equation

$$0 = d^A(*F) = d(*F)$$

gives the second pair of Maxwell equations (in vacuum):

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{E}}{\partial t} = 0, \tag{M3}$$

$$\text{div}\mathbf{E} = 0, \tag{M4}$$

Remark 1. The two equations $dF = 0$ and $d(*F) = 0$ can be stated as one equation

$$(dd^* + d^*d)F = 0,$$

where d^* is the operator adjoint to the operator d with respect to the unitary metric on the space of forms defined by (14.10).

Let us try to solve the Maxwell equations in vacuum. Since we can always add the form $d\psi$ to the potential A, we may assume that the scalar potential $\phi = A_0 = 0$. Thus equation (14.3) gives us $\mathbf{E} = -\frac{\partial A}{\partial t}$. Substituting this and (14.4) into (M3), we get

$$\text{curl curl } \mathbf{A} = -\Delta\mathbf{A} + \text{grad div } \mathbf{A} = -\frac{\partial^2 \mathbf{A}}{\partial t^2},$$

where Δ is the Laplacian operator applied to each coordinate of $\mathbf{A}$. We still have some freedom in the choice of $\mathbf{A}$. By adding to $\mathbf{A}$ the gradient of a suitable time-independent function, we may assume that $\text{div}\mathbf{A} = 0$. Thus we arrive at the equation

$$-\Box\mathbf{A} = \Delta\mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0. \tag{14.20}$$

Thus in the gauge where $\text{div}\mathbf{A} = 0, A_0 = 0$, the Maxwell equations imply that each coordinate function of the gauge potential $\mathbf{A}$ satisfies the d'Alembert equation from the previous lecture. Conversely, it is easy to see that (14.20) implies the Maxwell equations in vacuum (in the chosen gauge). Let us use this opportunity to explain how the d'Alembert equation can be solved if $\mathbf{A}$ depends only on one coordinate (plane waves). We rewrite (14.20) in the form

$$(\frac{\partial}{\partial x} - \frac{\partial}{\partial x})(\frac{\partial}{\partial x} + \frac{\partial}{\partial x})f = 0.$$

After introducing new variables $\xi = x - t, \eta = x + t$, it transforms to the equation

$$\frac{\partial^2 f}{\partial \xi \partial \eta} = 0.$$

Integrating this equation with respect to ξ, and then with respect to η, we get

$$f = f_1(\xi) + f_2(\eta) = f_1(x - t) + f_2(x + t).$$

This represents two plane waves moving in opposite directions.

To get the general Maxwell equations, we change the Yang-Mills action. The new action is

$$S(A) = S_{YM}(A) - j_e \cdot A,$$

where

$$j_e = (\rho_e, \mathbf{j}_e)$$

is the electric 4-vector current. The Euler-Lagrange equation is

$$d(*F_A) = \mathbf{j}_e.$$

This gives

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{j}_e, \qquad (M3')$$

$$\mathrm{div}\mathbf{E} = \rho_e, \qquad (M4')$$

Remark 2. If we change the metric $g = dt^2 - \sum_\mu (dx^\mu)^2$ to $c^2 dt^2 - \sum_\mu (dx^\mu)^2$, where c is the light speed, the Maxwell equations will change to

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = 0, \qquad (M1)$$

$$\mathrm{div}\mathbf{H} = \nabla \cdot \mathbf{H} = 0, \qquad (M2)$$

$$\nabla \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c}\mathbf{j}_e, \qquad (M3)$$

$$\mathrm{div}\mathbf{E} = \rho_e. \qquad (M4)$$

14.8 Let $\mathbf{q}(\tau) = x^\mu(t), \mu = 0, 1, 2, 3$ be a path in the space $\mathbb{R}^4$. Consider the natural (classical) Lagrangian on $T(M)$ defined by

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{m}{2}\sum_{i,j=1}^{4} g_{ij}\dot{q}_i\dot{q}_j = \frac{m}{2} g_{\mu\nu} dx^\mu dx^\nu.$$

Consider the action defined on the set of particles with *charge* e in the electromagnetic field by

$$S(\mathbf{q}) = \frac{m}{2}\int_{\mathbb{R}}(g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau})dt - e\int_{\mathbb{R}}\frac{dx^\mu}{d\tau}A_\mu(x^\mu(\tau))dt + \frac{1}{2}||F_A||^2.$$

Here we think of j_e as the derivative of the delta-function $\delta(x - x(\tau))$. The Euler Lagrange equation gives the equation of motion

$$m\frac{d^2x^\mu}{d\tau^2} = eg_{\nu\sigma}\frac{dx^\nu}{d\tau}F^{\sigma\mu}. \tag{14.21}$$

Set

$$\mathbf{v}(\tau) = \beta(\tau)^{-1}(\frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}),$$

where

$$\beta(\tau) = \frac{dx^0}{d\tau}.$$

If $g(\mathbf{x}', \mathbf{x}') = 1(\tau$ is proper time), then

$$\beta(\tau) = (1 - ||v||^2)^{-1/2}.$$

The we can write

$$\mathbf{x}(\tau) = \beta(1, \mathbf{v})$$

and rewrite (14.21) in the form

$$\frac{dm\beta}{d\tau} = e\mathbf{E}\cdot\mathbf{v}, \quad \frac{d(m\beta\mathbf{v})}{d\tau} = e(\mathbf{E} + \mathbf{v}\times\mathbf{H}). \tag{14.22}$$

the first equation asserts that the rate of change of energy of the particle is the rate at which work is done on it by the electric field. The second equation is the relativistic analogue of Newton's equation, where the force is the so-called *Lorentz force.*

14.9 Let E be a vector bundle with a connection A. In Lecture 13 we have defined the general Lagrangian $\mathcal{L} : E \oplus (E \otimes T^*(M)) \to \mathbb{R}$ by the formula

$$\mathcal{L} = -u \circ q + \lambda g^{-1} \otimes q.$$

Here $g : T(M) \to \mathbb{R}$ is the non-degenerate quadratic form defined by a metric on M, and q is a positive definite quadratic form on E, u and λ are some scalar non-negative valued functions on $\mathbb{R}$ and M, respectively. The Lagrangian $\mathcal{L}$ defines an action on $\Gamma(E)$ by the formula

$$S(\phi) = \int_M (\lambda(x)g^{-1} \otimes q(\nabla^A(\phi)) - u(q(\phi))d^nx.$$

Now we can extend this action by throwing in the Yang-Mills functional. The new action is

$$S(\phi) = \int_M (\lambda(x)g^{-1} \otimes q(\nabla^A(\phi)) - u(q(\phi)))\mathrm{vol}(g) + c\int_M F_A \wedge *F_A \tag{14.23}$$

for some positive constant c. For example, if we choose a gauge such that $g^{-1}(dx_\mu, dx_\nu) = g^{\mu\nu}$ and $||\phi|| = \sum_{i,j=1}^r a^{ij}\phi_i\phi_j$, then

$$\nabla(\phi)_i = \partial_\mu \phi_i dx^\mu + A_i^j{}_\mu \phi_j dx^\mu$$

and

$$\mathcal{L}(\phi)(x) = \lambda(x)a^{ij}(g^{\mu\nu}(\partial_\mu\phi_i + A_i^k{}_\mu\phi_k)(\partial_\nu\phi_j + A_i^l{}_\nu\phi_l) - u(a^{ij}\phi_i\phi_j) - cTr(F_{\mu\nu}F^{\mu\nu}). \quad (14.24)$$

Exercises.

1. Show that the introduction of the complex electromagnetic tensor defines a homomorphism of Lie groups $O(1,3) \to GL(3,\mathbb{C})$. Show that the pre-image of $O(3,\mathbb{C})$ is a subgroup of index 4 in $O(1,3)$.

2. Show that the star operator is a unitary operator with respect to the inner product $\langle F, G\rangle$ given by (14.10).

3. Show that the operator $\delta^A = (-1)^k * d^A * : \mathcal{A}^k(Ad(P)) \to \mathcal{A}^{k-1}(Ad(P))$ is the adjoint to the operator $d^A : \mathcal{A}^k(Ad(P)) \to \mathcal{A}^{k+1}(Ad(P))$.

4. The operator $\Delta^A = d^A \circ \delta^A + \delta^A \circ d^A$ is called the *covariant Laplacian*. Show that A is a Yang-Mills connection if and only if $\Delta^A(F_A) = 0$.

5. Prove that the quantities $||\mathbf{E}||^2 - ||\mathbf{H}||^2$ and $(\mathbf{H}\cdot\mathbf{E})^2$ are invariant with respect to the choice of coordinates.

6. Show that applying a Lorentzian transformation one can reduce the electromagnetic tensor $F \neq 0$ to the form where
 (i) $\mathbf{E} \times \mathbf{H} = 0$ if $\mathbf{H}\cdot\mathbf{E} \neq 0$;
 (ii) $\mathbf{H} = 0$ if $\mathbf{H}\cdot\mathbf{H} = \mathbf{0}, ||\mathbf{E}||^{\mathbf{2}} > ||\mathbf{H}||^{\mathbf{2}}$;
 (iii) $\mathbf{E} = 0$ if $\mathbf{H}\cdot\mathbf{H} = \mathbf{0}, ||\mathbf{E}||^{\mathbf{2}} < ||\mathbf{H}||^{\mathbf{2}}$.

7. Show that the Lagrangian (20) is invariant with respect to gauge transformations. Find the corresponding conservation laws.

8. Replacing the quaternions with complex numbers follow section 14.6 to construct a non-trivial principal $SU(1)$-bundle over two-dimensional sphere S^2. Show that its total space is diffeomorphic to S^3.

9. Show that the total space of the $SU(2)$-instanton over S^4 constructed in section 14.6 is diffeomorphic to S^7.

Lecture 15. SPINORS

15.1 Before we embark on the general theory, let me give you the idea of a spinor. Suppose we are in three-dimensional complex space $\mathbb{C}^3$ with the standard quadratic form $Q(x) = x_1^2 + x_2^2 + x_3^2$. The set of zeroes (=isotropic vectors) of Q considered up to proportionality is a conic in $\mathbb{P}^2(\mathbb{C})$. We know that it must be isomorphic to the projective line $\mathbb{P}^1(\mathbb{C})$. The isomorphism is given, for example, by the formulas

$$x_1 = t_0^2 - t_1^2, \quad x_2 = i(t_0^2 + t_1^2), \quad x_3 = -2t_0t_1.$$

The inverse map is given by the formulas

$$t_0 = \pm\sqrt{\frac{x_1 - ix_2}{2}}, \quad t_1 = \pm\sqrt{\frac{-x_1 - ix_2}{2}}.$$

It is not possible to give a consistent choice of sign so that these formulas define an isomorphism from $\mathbb{C}^2$ to the set I of isotropic vectors in $\mathbb{C}^3$. This is too bad because if it were possible, we would be able to find a linear representation of $O(3,\mathbb{C})$ in $\mathbb{C}^2$, using the fact that the group acts naturally on the set I. However, in the other direction, if we start with the linear group $GL(2,\mathbb{C})$ which acts on (t_0, t_1) by linear transformations

$$(t_0, t_1) \to (at_0 + bt_1, ct_0 + dt_1),$$

then we get a representation of $GL(2,\mathbb{C})$ in $\mathbb{C}^3$ which will preserve I. This implies that, for any $g \in GL(2,\mathbb{C})$, we must have $g^*(Q) = \xi(g)Q$ for some homomorphism $\xi : GL(2,\mathbb{C}) \to \mathbb{C}^*$. Since the restriction of ξ to $SL(2,\mathbb{C})$ is trivial, we get a homomorphism

$$s : SL(2,\mathbb{C}) \to SO(3,\mathbb{C}).$$

It is easy to see that the homomorphism s is surjective and its kernel consists of two matrices equal to $\pm$ the identity matrix. The vectors (t_0, t_1) are called *spinors* representing isotropic vectors in $\mathbb{C}^3$. Although the group $SO(3,\mathbb{C})$ does not act on spinors, its double cover $SL(2,\mathbb{C})$ does. The action is called the *spinor representation* of $SO(3,\mathbb{C})$.

Now let us look at the Lorentz group $G = O(1,3)$. Via its action on $\mathbb{R}^4$, it acts naturally on the space $\bigwedge^2(\mathbb{R}^4)$ of skew-symmetric 4×4-matrices $A = (a_{ij})$. If we set

$$\mathbf{v} = (ia_{12} + a_{34}, ia_{13} + a_{24}, ia_{14} + a_{23}) = (a_1, a_2, a_3) + i(b_1, b_2, b_3) \in \mathbb{C}^3,$$

we obtain an isomorphism of real vector spaces $\bigwedge^2(\mathbb{R}^4) \to \mathbb{C}^3$. Since

$$\mathbf{v}\cdot\mathbf{v} = \mathbf{a}\cdot\mathbf{a} - \mathbf{b}\cdot\mathbf{b} + i\mathbf{a}\cdot\mathbf{b},$$

the real part of $\mathbf{v}\cdot\mathbf{v}$ coincides with $\langle A, A\rangle$, where $\langle\ ,\ \rangle$ is the scalar product on $\bigwedge^2(\mathbb{R}^4)$ induced by the Lorentz metric in $\mathbb{R}^4$. Now if we switch from $O(1,3)$ to $SO(1,3)$, we obtain that the determinant of A is preserved. But the determinant of A is equal to $(\mathbf{a}\cdot\mathbf{b})^2$ (the number $\mathbf{a}\cdot\mathbf{b}$ is the Pfaffian of A). To preserve $\mathbf{a}\cdot\mathbf{b}$ we have to replace $O(1,3)$ with $SO(1,3)$ which guarantees the preservation of the square of the Pfaffian. To preserve the Pfaffian itself we have to go further and choose a subgroup of index 2 of $SO(1,3)$. This is called the *proper Lorentz group.* It is denoted by $SO(1,3)_0$. Thus under the isomorphism $\bigwedge^2(\mathbb{R}^4) \cong \mathbb{C}^3$, the group $SO(1,3)_0$ is mapped to $SO(3,\mathbb{C})$. One can show that this is an isomorphism. Combining with the above, we obtain a homomorphism

$$s' : SL(2,\mathbb{C}) \to SO(1,3)_0$$

which is the spinor representation of the proper Lorentz group.

15.2 The general theory of spinors is based on the theory of Clifford algebras. These were introduced by the English mathematician William Clifford in 1876. The notion of a spinor was introduced by the German mathematician R. Lipschitz in 1886. In the twenties, they were rediscovered by B. L. van der Waerden to give a mathematical explanation of Dirac's equation in quantum mechanics. Let us first explain Dirac's idea. He wanted to solve the D'Alembert equation

$$(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2})\phi = 0.$$

This would be easy if we knew that the left-hand side were the square of an operator of first order. This is equivalent to writing the quadratic form $t^2 - x_1^2 - x_2^2 - x_3^2$ as the square of a linear form. Of course this is impossible as the latter is an irreducible polynomial over $\mathbb{C}$. But it is possible if we leave the realm of commutative algebra. Let us look for a solution

$$t^2 - x_1^2 - x_2^2 - x_3^2 = (tA_1 + x_2A_2 + x_3A_3 + x_4A_4)^2.$$

where A_i are square complex matrices, say of size two-by-two. By expanding the right-hand side, we find

$$A_1^2 = I_2, \quad A_2^2 = A_3^2 = A_4^2 = -I_2, \quad A_iA_j + A_jA_i = 0, \quad i \neq j. \qquad (15.1)$$

One introduces a non-commutative unitary algebra C given by generators A_1, A_2, A_3, A_4 and relations (15.1). This is the Clifford algebra of the quadratic form $t^2 - x_1^2 - x_2^2 - x_3^2$ on

$\mathbb{R}^4$. Now the map $X_i \to A_i$ defines a linear (complex) two-dimensional representation of the Clifford algebra (the spinor representation). The vectors in $\mathbb{C}^2$ on which the Clifford algebra acts are spinors.

Let us first recall the definition of a Clifford algebra. It generalizes the concept of the exterior algebra of a vector space.

Definition. Let E be a vector space over a field K and $Q : E \to K$ be a quadratic form. The *Clifford algebra* of (E, Q) is the quotient algebra

$$C(Q) = T(E)/I(Q)$$

of the tensor algebra $T(E) = \oplus_n T^n E$ by the two-sided ideal $I(Q)$ generated by the elements $v \otimes v - Q(v), v \in T^1 = E$.

It is clear that the restrictions of the factor map $T(E) \to C(Q)$ to the subspaces $T^0(E) = K$ and $T^1(E) = E$ are bijective. This allows us to identify K and E with subspaces of $C(Q)$. By definition, for any $v \in E$,

$$v \cdot v = Q(v). \tag{15.2}$$

Replacing $v \in V$ with $v + w$ in this identity, we get

$$v \cdot w + w \cdot v = B(v, w), \tag{15.3}$$

where $B : E \times E \to K, (x, y) \to Q(x + y) - Q(x) - Q(y)$, is the symmetric bilinear form associated to Q. Using the *anticommutator* notation

$$\{a, b\} = ab + ba$$

one can rewrite (15.3) as

$$\{v, w\} = B(v, w). \tag{15.3$'$}$$

Example 1. Take Q equal to the zero function. By definition $v \cdot v = 0$ for any $v \in E$. Let us show that this implies that $C(Q)$ is isomorphic to the exterior (or Grassmann) algebra of the vector space E

$$\bigwedge(E) = \bigoplus_{n \geq 0} \bigwedge^n(E).$$

In fact, the latter is defined as the quotient of $T(E)$ by the two-sided ideal J generated by elements of the form $v \otimes x \otimes v$ where $v \in E, x \in T(E)$. It is clear that $I(Q) \subset J$. It is enough to show that $J \subset I(Q)$. For this it suffices to show that $v \otimes x \otimes v \in I(Q)$ for any $v \in E, x \in T^n(E)$, where n is arbitrary. Let us prove it by induction on n. The assertion is obvious for $n = 0$. Write $x \in T^n(E)$ in the form $x = w \otimes y$ for some $y \in T^{n-1}(E), w \in E$. Then

$$v \otimes w \otimes y \otimes v = (v + w) \otimes (v + w) \otimes y \otimes v - v \otimes v \otimes y \otimes v - w \otimes v \otimes y \otimes v - w \otimes w \otimes y \otimes v.$$

By induction, each of the four terms in the right-hand side belongs to the ideal $I(Q)$.

Example 2. Let dim $E = 1$. Then $T(E) \cong K[t]$ and $C(Q) \cong K[t]/(t^2 - Q(e))$, where E is spanned by e. In particular, any quadratic extension of K is a Clifford algebra.

Theorem 1. *Let $(e_i)_{i\in I}$ be a basis of E. For any finite ordered subset $S = (i_1, \dots, i_k)$ of I, set*

$$e_S = e_{i_1} \dots e_{i_k} \in C(Q).$$

Then there exists a basis of $C(Q)$ formed by the elements e_S. In particular, if $n = \dim\ E$,

$$\dim_K C(Q) = 2^n.$$

Proof. Using (15.3), we immediately see that the elements e_S span $C(Q)$. We have to show that they are linearly independent. This is true if $Q = 0$, since they correspond to the elements $e_{i_1} \wedge \dots \wedge e_{i_k}$ of the exterior algebra. Let us show that there is an isomorphism of vector spaces $C(Q) \cong C(0)$. Take a linear function $f \in E^*$ on E. Define a linear map

$$\phi_f : T(E) \to T(E) \tag{15.4}$$

as follows. By linearity it is enough to define $\phi_f(x)$ for decomposable tensors $x \in T^n$. For $n = 0$, set $\phi_f = 0$. For $n = 1$, set $\phi_f(v) = f(v) \in T^0(E)$. Now for any $x = v \otimes y \in T^n(E)$, set

$$\phi_f(v \otimes y) = f(v)y - v \otimes \phi_f(y). \tag{15.5}$$

We have $\phi_f(v \otimes v - Q(v)) = f(v)v - v \otimes f(v) = 0$. We leave to the reader to check that moreover $\phi_f(x \otimes (v \otimes v - Q(x)) \otimes y) = 0$ for any $x, y \in T(E)$. This implies that the map ϕ_f factors to a linear map $C(Q) \to C(Q)$ which we also denote by ϕ_f. Notice that the map $E^* \to \mathrm{End}(T(E)), f \to \phi_f$ is linear.

Now let $F : E \times E \to K$ be a bilinear form, and let $i_F : E \to E^*$ be the corresponding linear map. Define a linear map

$$\lambda_F : T(E) \to T(E) \tag{15.6}$$

by the formula

$$\lambda_F(v \otimes y) = v \otimes \lambda_F(y) + \phi_{i_F(v)}(\lambda_F(y)), \tag{15.7}$$

where $v \in E, y \in T^{n-1}(E)$ and $\lambda_F(1) = 1$.

Obviously the restriction of this map to $E = T^0(E) \oplus T^1(E)$ is equal to the identity. Also the map λ_F is the identity when $F = 0$. We need the following

Lemma 1.

(i) *for any $f \in E^*$,*

$$\phi_f^2 = 0;$$

(ii) *for any $f, g \in E^*$,*

$$\{\phi_f, \phi_g\} = 0;$$

(iii) *for any $f \in E^*$,*

$$[\lambda_F, \phi_f] = 0.$$

Proof. (i) Use induction on the degree of the decomposable tensor. We have

$$\phi_f^2(v \otimes y) = \phi_f(\phi_f(v)y - v \otimes \phi_f(y)) = \phi_f(v)\phi_f(y) - \phi_f(v)\phi_f(y) - v \otimes \phi_f^2(y) = 0.$$

(ii) We have

$$\phi_f \circ \phi_g(v \otimes y) = \phi_f(g(v)y - v \otimes \phi_g(y)) = \phi_f(g(v))y) - g(v)\phi_f(y) + v \otimes \phi_f(\phi_g(y)) =$$
$$= v \otimes \phi_f(\phi_g(y)).$$

(iii) Using (ii) and induction, we get

$$\lambda_F(\phi_f(v\otimes y)) = \lambda_F(f(v)y - v\otimes\phi_f(y)) = f(v)\lambda_F(y) - v\otimes\lambda_F(\phi_f(y)) - \phi_{i_F(v)}(\lambda_F(\phi_f(y)) =$$
$$= f(v)\lambda_F(y) - v\otimes\phi_f(\lambda_F(y)) - \phi_{i_F(v)}\phi_f(\lambda_F(y)) =$$
$$= \phi_f(v\otimes\lambda_F(y)) + \phi_f(\phi_{i_F(v)}(\lambda_F(y)) = \phi_f(\lambda_F(v\otimes y)).$$

Corollary. *For any other bilinear form G on E, we have*

$$\lambda_{F+G} = \lambda_F \circ \lambda_G,$$

Proof. Using (iii) and induction, we have

$$\lambda_G \circ \lambda_F(v\otimes y) = \lambda_G(v\otimes\lambda_F(y) + \phi_{i_F(v)}(\lambda_F(y)) = v\otimes\lambda_G(\lambda_F(y)) + \phi_{i_G}(v)(\lambda_F(y)))+$$
$$+\lambda_G(\phi_{i_F(v)}(\lambda_F(y)) = v\otimes\lambda_G(\lambda_F(y)) + (\phi_{i_G(v)} + \phi_{i_G(v)})(\lambda_G(\lambda_F(y))) =$$
$$= v\otimes\lambda_{G+F}(y) + \phi_{i_{F+G}(v)}(\lambda_{G+F}(y)) = \lambda_{F+G}(v\otimes y).$$

It is clear that

$$\lambda_{-F} = (\lambda_F)^{-1}. \tag{15.8}$$

Let Q' be another quadratic form on E defined by $Q'(v) = Q(v) + F(v,v)$. Let us show that λ_F defines a linear isomorphism $C(Q) \cong C(Q')$. By (15.8), it suffices to verify that $\lambda_F(I(Q')) \subset I(Q)$. Since $\phi_{i_F(v)}(I(Q)) \subset I(Q)$, formula (15.5) shows that the set of $x \in T(E)$ such that $\lambda_F(x) \in I(Q)$ is a left ideal. So, it is enough to verify that $\lambda_F(v\otimes v\otimes x - Q'(v)x) \in I(Q)$ for any $v \in E, x \in T(E)$. We have

$$\lambda_F(v\otimes v\otimes x - Q'(v)x) = v\otimes\lambda_F(v\otimes x) + \phi_{i_F(v)}(\lambda_F(v\otimes x)) - Q'(v)\lambda_F(x) =$$
$$= v\otimes v\otimes\lambda_F(x) + v\otimes\phi_{i_F(v)}(\lambda_F(x)) + \lambda_F(\phi_{i_F(v)}(v\otimes x)) - Q'(v)\lambda_F(x) =$$
$$= v\otimes v\otimes\lambda_F(x) + v\otimes\phi_{i_F(v)}(\lambda_F(x)) + \lambda_F(F(v,v)x - v\otimes\phi_{i_F(v)}(x)) - Q'(v)\lambda_F(x) =$$
$$= (v\otimes v - F(v,v) - Q'(v))\otimes\lambda_F(x) = v\otimes v - Q(v).$$

To finish the proof of the theorem, it suffices to find a bilinear form F such that $Q(x) = -F(x,x)$. If $\mathrm{char}(K) \neq 2$, we may take $F = -\frac{1}{2}B$ where B is the associated symmetric bilinear form of Q. In the general case, we may define F by setting

$$F(e_i,e_j) = \begin{cases} -B(e_i,e_j) & \text{if } i > j, \\ Q(e_i) & \text{if } i = j, \\ 0 & \text{if } i < j. \end{cases}$$

15.3 The structure of the Clifford algebra $C(Q)$ depends very much on the property of the quadratic form. We have seen already that $C(Q)$ is isomorphic to the Grassmann algebra when Q is trivial. If we know the classification of quadratic forms over K we shall find out the classification of Clifford algebras over K.

Let $C^+(Q), C^-(Q)$ be the subspaces of $C(Q)$ equal to the images of the subspaces $T^+(E) = \oplus_k T^{2k}(E)$ and $T^-(E) = \oplus_k T^{2k+1}(E)$ of $T(E)$, respectively. Since $T^+(E)$ is a subalgebra of $T(E)$, the subspace $C^+(Q)$ is a subalgebra of $C(Q)$. Since $I(Q)$ is spanned by elements of $T^+(E)$, $I(Q) = I(Q) \cap T^+(E) \oplus I(Q) \cap T^-(E)$. This implies that

$$C(Q) = C^+(Q) \oplus C^-(Q). \tag{15.9}$$

We call elements of $C^+(Q)$ (resp. $C^-(Q)$) *even* (resp. *odd*). Let $C(Q_1)$ and $C(Q_2)$ be two Clifford algebras. We define their tensor product using the formula

$$(a \otimes b) \cdot (a' \otimes b') = \epsilon a a' \otimes b b', \tag{15.10}$$

where $\epsilon = 1$ unless a, a' are not both even or odd, and b, b' are not both even or odd. In the latter case $\epsilon = -1$. We assume here that each a, a', b, b' is either even or odd.

Theorem 2. *Let $Q_i : E_i \to K, i = 1, 2$ be two quadratic forms, and $Q = Q_1 \oplus Q_2 : E = E_1 \oplus E_2 \to K$. Then there exists an algebra isomorphism*

$$C(Q_1) \otimes C(Q_2) \cong C(Q).$$

Proof. We have a canonical map $p : C(Q_1) \otimes C(Q_2) \to C(Q)$ induced by the bijective canonical map $T(E_1) \otimes T(E_2) \to T(E)$. Since the basis of E is equal to the union of bases in E_1 and E_2, we can apply Theorem 1 to obtain that p is bijective. We have $vw + wv = B(v, w) = 0$ if $v \in E_1, w \in E_2$. From this we easily find that

$$(v_1 \dots v_k)(w_1 \dots w_s) = (-1)^{ks}(w_1 \dots w_s)(v_1 \dots v_k),$$

where $v_i \in E_1, w_j \in E_2$. This implies that the map $C(Q_1) \otimes C(Q_2) \to C(Q)$ defined by $x \otimes y \to xy$ is an isomorphism. Here is where we use the definition of the tensor product given in (15.10).

Corollary. *Assume $n =$ dim E. Let $(e_1, \dots, e_n)$ be a basis in E such that the matrix of Q is equal to a diagonal matrix diag$(d_1, \dots, d_n)$ (it always exists). Then $C(Q)$ is an associative algebra with a unit, generated by $1, e_1, \dots, e_n$, with relations*

$$e_i^2 = d_i, \quad e_i e_j = -e_j e_i, \quad i, j = 1, \dots, n, i \neq j.$$

Example 3. Let $E = \mathbb{R}^2$ and Q have signature $(0, 2)$. Then there exists a basis such that $Q(\sum x_1 e_1 + x_2 e_2) = -x_1^2 - x_2^2$. Then, setting $e_1 = \mathbf{i}, e_2 = \mathbf{j}$, we obtain that $C(Q)$ is generated by $1, \mathbf{i}, \mathbf{j}$ with relations

$$\mathbf{i}^2 = \mathbf{j}^2 = -1, \quad \mathbf{ij} = -\mathbf{ji}.$$

Thus $Q(E)$ is isomorphic to the algebra of quaternions $\mathbb{H}$.

Recall from linear algebra that a non-degenerate quadratic form $Q : V \to K$ on a vector space V of dimension $2n$ is called *neutral* (or of *Witt index* n) if $V = E \oplus F$ with $Q|E \equiv 0, Q|F \equiv 0$. It follows that dim $E =$ dim $F = n$. For example, every non-degenerate quadratic form on a complex vector space of dimension $2n$ is neutral.

Theorem 3. *Let* dim $E = 2r, Q : E \to K$ *be a neutral quadratic form on* V. *There exists an isomorphism of algebras*

$$s : C(Q) \cong \mathrm{Mat}_{2^r}(K).$$

The image of $C^+(Q)$ *under* s *is isomorphic to the sum of two left ideals in* $\mathrm{Mat}_{2^r}(K)$, *each isomorphic to* $\mathrm{Mat}_{2^{r-1}}(K)$.

Proof. Let $E = F \oplus F'$, where $Q|F \equiv 0, Q|F' \equiv 0$. Let S be the subalgebra of $C(Q)$ generated by F. Then $S \cong \bigwedge(F)$. We can choose a basis $(f_1, \ldots, f_r)$ of F and a basis $f_1^*, \ldots, f_r^*$ of F' such that $B(f_i, f_j^*) = \delta_{ij}$. Thus we can consider F' as the dual space F^* and $(f_1^*, \ldots, f_r^*)$ as the dual basis to $(f_1, \ldots, f_r)$. For any $f \in F$ let $l_f : x \to fx$ be the left multiplication endomorphism of S (*creation operator*). For any $f' \in F'$ let $\phi_{f'} : S \to S$ be the endomorphism defined in the proof of Theorem 1 (*annihilation operator*). Using formula (15.5), we have

$$l_f \circ \phi_{f'} + \phi_{f'} \circ l_f = B(f, f')\mathbf{id}. \tag{15.11}$$

For $v = f + f' \in E$ set

$$s_v = l_f + \phi_{f'} \in \mathrm{End}(F).$$

Then $v \to s_v$ is a linear homomorphism from V to $\mathrm{End}(S) \cong \mathrm{Mat}_{2^r}(K)$. From the equality

$$s_v^2(x) = (l_f + \phi_{f'})^2(x) = e \otimes e \otimes x + B(f, f')x = Q(f)x + B(f, f')x = Q(v)x,$$

we get that $v \to s_v$ can be extended to a homomorphism $s : C(Q) \to \mathrm{End}(S)$. Since both algebras have the same dimension ($= 2^n$), it is enough to show that the obtained homomorphism is surjective. We use a matrix representation of $\mathrm{End}(S)$ corresponding to the natural basis of S formed by the products $e_I = e_{i_1} \wedge \ldots \wedge e_{i_k}, i_1 < \ldots < i_k$. We need to construct an element $x \in C(Q)$ such that $s(x)$ is the unit matrix $E_{KH} = (\delta_{KH})$, where $K = (1 \le k_1 < \ldots < k_s \le r), H = (1 \le h_1 < \ldots < h_t \le r)$ are ordered subsets of $[r] = \{1, \ldots, r\}$. One takes $x = e_H f_{[r]} e_K$, where $e_H = e_{k_1} \ldots e_{k_s}, e_K = e_{h_1} \ldots e_{h_s}, f_{[r]} = f_1 \ldots f_r \in C(Q)$. We skip the computation.

For the proof of the second assertion, we set $S^+ = S \cap C^+(Q)$ and $S^- = S \cap C^-(Q)$. Each of them is a subspace of S of dimension 2^{r-1}. Then it is easy to see that S^+ and S^- are left invariant under $s(C^+(Q))$. Since s is injective, s maps $C^+(Q)$ isomorphically onto the subalgebra of $\mathrm{End}(S)$ isomorphic to $\mathrm{End}(S^+) \times \mathrm{End}(S^-)$.

Example 4. Let $E = \mathbb{R}^2, Q(x_1, x_2) = x_1 x_2$. Then we can take $F = Ke_1 = \{x_2 = 0\}$, $F' = Ke_2 = \{x_1 = 0\}$. The algebra S is two-dimensional with basis consisting of $1 \in \bigwedge^0(F)$ and $e_1 \in \bigwedge^1(F) = F$. The Clifford algebra is spanned by $1, e_1, e_2, e_1e_2$ with $e_i^2 = 0, e_1e_2 = -e_2e_1$. The map $C(Q) \to \mathrm{End}(S) \cong \mathrm{Mat}_2(K)$ is defined by

$$s(e_1) = l_{e_1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad s(e_2) = \phi_{e_2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$s(e_1e_2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad s(1 - e_1e_2) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This checks the theorem.

Corollary 1. *Assume $n = 2r$, Q is non-degenerate and K is algebraically closed, or $K = \mathbb{R}$ and Q is of signature (r, r). Then*

$$C(Q) \cong \mathrm{Mat}_{2^r}(K).$$

Corollary 2. *Let $E = \mathbb{R}^{2r}$ and Q be of signature $(r, r+2)$. Then*

$$C(Q) \cong \mathrm{Mat}_r(\mathbb{H}),$$

where H is the algebra of quaternions.

Proof. We can write $E = E_1 \oplus E_2$, where E_1 is orthogonal to E_2 with respect to B, and $Q|E_1$ is of signature $(1,1)$ and $Q|E_2$ is of signature $(0,2)$. By Theorem 2 $C(Q) \cong C(Q_1) \otimes C(Q_2)$. By Theorem 3, $C(Q_1) \cong \mathrm{Mat}_2(\mathbb{R})$. By Example 3, $C(Q_2) \cong \mathbb{H}$. It is easy to check that $\mathrm{Mat}_2(\mathbb{R}) \otimes \mathbb{H} \cong \mathrm{Mat}_2(\mathbb{H})$.

Corollary 3. *Assume Q is non-degenerate and $n = 2r$. Then there exists a finite extension K'/K of the field K such that*

$$C(Q) \otimes_K K' \cong \mathrm{Mat}_{2^r}(K').$$

In particular, $C(Q)$ is a central simple algebra (i.e., has no non-trivial two-sided ideals and its center is equal to K). The center Z of $C^+(Q)$ is a quadratic extension of K. If Z is a field $C^+(Q)$ is simple. Otherwise, $C^+(Q)$ is the direct sum of two simple algebras.

Proof. Let $e_1, \ldots, e_n$ be a basis in E such that the matrix of B is equal to $d_i\delta_{ij}$ for some $d_i \in K$. Let K' be obtained from K by adjoining square roots of d_i's and $\sqrt{-1}$. Then we can find a basis $f_1, \ldots, f_n$ of $E \otimes_K K'$ such that $B(f_i, f_j) = \delta_{ii+r}$. Then we can apply Theorem 3 to obtain the assertion.

The algebra $\mathrm{Mat}_{2^r}(K')$ is central simple. This implies that $C(Q)$ is also central simple.

In the case when dim E is odd, the structure of $C(Q)$ is a little different.

Theorem 4. *Assume* dim $E = 2r + 1$ *and there exists a hyperplane H in E such that $(H, Q|H)$ is neutral. Then*

$$C^+(Q) \cong \mathrm{Mat}_{2^r}(K).$$

The center Z of $C(Q)$ is a quadratic extension of K and

$$C(Q) \cong \mathrm{Mat}_{2^r}(Z).$$

Proof. Let $v_0 \in E$ be a non-zero vector orthogonal to H. It is non-isotropic since otherwise Q is degenerate. Define the quadratic form Q' on H by $Q'(h) = -Q(v_0)Q(h)$. Since $v_0 h = -h v_0$ for all $h \in H$, we have $(v_0 h)^2 = -v_0^2 h^2 = -Q(v_0)Q(h) = Q'(h)$. This easily implies that the homomorphism $T(H) \to T(E), h_1 \otimes \ldots \otimes h_n \to (v_0 \otimes h_1) \otimes \ldots \otimes (v_0 \otimes h_n)$ factors to define a homomorphism $f : C(Q') \to C^+(Q)$. Since dim H is even, by Theorem 3, the algebra $C(Q')$ is a matrix algebra, hence does not have non-trivial two-sided ideals. This implies that the homomorphism f is injective. Since the dimension of both algebras is the same, it is also bijective. Applying again Theorem 3, we obtain the first assertion.

Let $v_1, \ldots, v_{2r}$ be an orthogonal basis of $Q|H$. Then, it is immediately verified that $z = v_0 v_1 \ldots v_{2r}$ commutes with any $v_i, i = 0, \ldots, 2r$, hence belongs to the center of $C(Q)$. We have $z^2 = (-1)^r Q(v_0) \ldots Q(v_{2r}) \in K^*$. So $Z' = K + Kz$ is a quadratic extension of K contained in the center. Since $z \in C^-(Q)$ and is invertible, $C(Q) = zC^+(Q) + C^+(Q)$. This implies that the image of the homomorphism $\alpha : Z' \otimes C^+(Q) \to C(Q), x \otimes y \to xy$, contains $C^+(Q)$ and $C^-(Q)$, hence is surjective. Thus, $Z' \otimes C^+(Q) \cong \mathrm{Mat}_{2^r}(Z')$. Since the center of $\mathrm{Mat}_{2^r}(Z')$ can be identified with Z', we obtain that $Z' = Z$. and the center hence α is injective. Thus $C(Q) \cong \mathrm{Mat}_{2^r}(Z)$.

Corollary. *Assume n is odd and Q is non-degenerate. Then $C(Q)$ is either simple (if Z is a field), or the product of two simple algebras (if Z is not a field). The algebra $C^+(Q)$ is a central simple algebra.*

If n is even, $C(Q)$ is simple and hence has a unique (up to an isomorphism) irreducible linear representation. It is called *spinor representation*. The elements of the corresponding space are called *spinors*. In the case when Q satisfies the assumptions of Theorem 3, we may realize this representation in the maximal isotropic subspace S of E. We can do the same if we allow ourselves to extend the field of scalars K. The restriction of the spinor representation to $C^+(Q)$ is either irreducible or isomorphic to the direct sum of two non-isomorphic irreducible representations. The latter happens when Q is neutral. In this case the two representations are called the *half-spinor representations* of $C^+(Q)$. The elements of the corresponding spaces are called *half-spinors*.

If n is odd, $C^+(Q)$ has a unique irreducible representation which is called a spinor representation. The elements of the corresponding space are spinors.

15.4 Now we can define the spinor representations of orthogonal groups. First we introduce the *Clifford group* of the quadratic form Q. By definition, it is the group $G(Q)$ of invertible elements x of the Clifford algebra of $C(Q)$ such that $xvx^{-1} \in E$ for any $v \in E$. The subgroup $G^+(Q) = G(Q) \cap C^+(Q)$ of $G(Q)$ is called the *special Clifford group*.

Let $\phi : G(Q) \to GL(E)$ be the homomorphism defined by

$$\phi(x)(v) = xvx^{-1}. \tag{15.12}$$

Since

$$Q(xvx^{-1}) = xvx^{-1}xvx^{-1} = xvvx^{-1} = xQ(v)x^{-1} = Q(v),$$

we see that the image of ϕ is contained in the orthogonal group $O(Q)$.

From now on, we shall assume that $\mathrm{char}(K) \neq 2$ and Q is a nondegenerate quadratic form on E. We continue to denote the associated symmetric bilinear form by B.

Let $v \in E$ be a non-isotropic vector (i.e., $Q(v) \neq 0$). The linear transformation of E defined by the formula

$$r_v(w) = w - \frac{B(v,w)}{Q(v)} v$$

is called the *reflection* in the vector v. Since

$$B(r_v(w), r_v(w)) = B(w, w) - 2\frac{B(v, w)}{Q(v)}B(w, v) + 2(\frac{B(v, w)}{Q(v)})^2 Q(v) = B(w, w),$$

the transformation r_v is orthogonal. It is immediately seen that the restriction of r_v to the hyperplane H_v orthogonal to v is the identity, and $r_v(v) = -v$. Conversely, any orthogonal transformation of (E, Q) with this property is equal to r_v.

We shall use the following result from linear algebra:

Theorem 5. *$O(Q)$ is generated by reflections.*

Proof. Induction on dim E. Let $T : E \to E$ be an element of $O(Q)$. Let v be a non-isotropic vector. Assume $T(v) = v$. Then T leaves invariant the orthogonal complement H of v and induces an orthogonal transformation of H. It is clear that the restriction of Q to H is a non-degenerate quadratic form. By induction, $T|H = r_1 \dots r_k$ is the product of reflections in some vectors $v_1, \dots, v_k$ in H. Let $L_i = (Kv_i)^{\perp}_{H}$. If we extend r_i to an orthogonal transformation $\tilde{r}_i$ of E which fixes v, then $\tilde{r}_i$ fixes pointwisely the hyperplane $Kv_i + L_i = (Kv_i)^{\perp}_{E}$ and coincides with the reflection in E in the vector v_i. It is clear that $T = \tilde{r}_1 \dots \tilde{r}_k$.

Now assume that $T(v) = -v$. Then $r_v \circ T(v) = v$. By the previous case, $r_v \circ T$ is the product of reflections, so T is also.

Finally, let us consider the general case. Take $w = T(v)$, then $Q(w) = Q(v)$ and $Q(v+w) + Q(v-w) = 2Q(v) + 2Q(w) = 4Q(v) \neq 0$. So, one of the two vectors $v + w$ and $v - w$ is non-isotropic. Assume $h = v - w$ is non-isotropic. Then $B(w, h) = Q(w + h) - Q(h) - Q(w) = Q(v) - Q(v) - Q(h) = -Q(h)$. Thus $r_h(w) = w - \frac{B(w,x)}{Q(h)} h = w + h = v$. This implies that $T \circ r_h(w) = T(v) = w$. By the first case, $T \circ r_h$ is the product of reflections, so is T. Similarly we consider the case when $v + w$ is non-isotropic. We leave this to the reader.

Proposition 1. *The set $G(Q) \cap E$ is equal to the set of vectors $v \in E$ such that $Q(v) \neq 0$. For any $v \in G(Q) \cap E$, the map $-\phi(v) : E \to E$ is equal to the reflection r_v.*

Proof. Since $v^2 = Q(v)$, the vector v is invertible in $C(Q)$ if and only if $Q(v) \neq 0$. So, if $v \in G(Q)$, we must have $v^{-1} = Q(v)^{-1}v$, hence, for any $w \in E$,

$$vwv^{-1} = vwQ(v)^{-1}v = Q(v)^{-1}vwv = Q(v)^{-1}v(B(v, w) - vw) =$$

$$= Q(v)^{-1}vB(v, w) - Q(v)^{-1}v^2 w = Q(v)^{-1}vB(v, w) - w.$$

Corollary. *Any element from $G(Q)$ can be written in the form*

$$g = zv_1 \dots v_k,$$

where $v_1, \dots, v_k \in E \setminus Q^{-1}(0)$ and z is an invertible element from the center Z of $C(Q)$.

Proof. Let $\phi : G(Q) \to O(Q)$ be the map defined in (15.12). Its kernel consists of elements x which commute with all elements from E. Since $C(Q)$ is generated by elements from E, we obtain that $\mathrm{Ker}(\phi)$ is a subset of the group Z^* of invertible elements of Z. Obviously, the converse is true. So

$$\mathrm{Ker}(\phi) = Z^*.$$

By Theorem 5, for any $x \in G(Q)$, the image $\phi(x)$ is equal to the product of reflections r_i. By Proposition 1, each $r_i = -\phi(v_i)$ for some non-isotropic vector $v_i \in E$. Thus x differs from the product of v_i's by an element from $\mathrm{Ker}(\phi) = Z^*$. This proves the assertion.

Theorem 6. *Assume char$(K) \neq 2$. If $n = \dim\ E$ is even, the homomorphism $\phi : G(Q) \to O(Q)$ is surjective. In this case the image of the subgroup $G^+(Q)$ is equal to $SO(Q)$. If n is odd, the image of $G(Q)$ and $G^+(Q)$ is $SO(Q)$. The kernel of ϕ is equal to the group Z^* of invertible elements of the center of $C(Q)$.*

Proof. The assertion about the kernel has been proven already. Also we know that the image of ϕ consists of transformations $(-r_{v_1}) \circ \ldots \circ (-r_{v_k}) = (-1)^k r_{v_1} \circ \ldots \circ r_{v_k}$, where $v_1, \ldots, v_k$ are non-isotropic vectors from E.

Assume n is even. By Theorem 5, any element T from $O(Q)$ is a product of N reflections. If N is even, we take $k = N$ and find that T is in the image. If N is odd, we write $-\mathbf{id}_E$ as the product of even number n of reflections. Then we take $k = N + n$ to obtain that T is in the image.

Assume n is odd. As above we show that each product of even number of reflections in $O(Q)$ belongs to the image of ϕ. Now the product of odd number of reflections cannot be in the image because its determinant is -1 (the determinant of any reflection equals -1.) But the determinant of $(-r_{v_1}) \circ \ldots \circ (-r_{v_k})$ is always one. However, $-\mathbf{id}_E$ is equal to the product of n reflections. So, multiplying a product of odd number of reflections by $-\mathbf{id}_E$ we get a product of even reflections which, by the above, belongs to the image. This proves the assertion about the image of ϕ.

If n is even, the restriction of the spinor representation of $C(Q)$ to $G(Q)$ is a linear representation of the Clifford group $G(Q)$. It is called the *spinor representation*. One can prove that it is irreducible. Abusing the terminology this is also called the spinor representation of $O(Q)$. The restriction of the spinor representation to $G^+(Q)$ is isomorphic to the direct sum of two non-isomorphic irreducible representations. They are called the *half-spinor representations* of $G^+(Q)$ (or of $SO(Q)$).

If n is odd, the restriction of the spinor representation of $C^+(Q)$ to $G^+(Q)$ is an irreducible representation. It is called the spinor representation of $G^+(Q)$ (or of $SO(Q)$).

Example 5. Take $E = \mathbb{C}^3$ and $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$. This is the odd case. Changing the basis, we can transform Q to the form $-x_1x_2 + x_3^2$. Obviously, Q satisfies the assumption of Theorem 4 (take $H = \mathbb{C}e_1 + \mathbb{C}e_2$.) Thus $C^+(Q) \cong M_2(\mathbb{C})$. Hence $G^+(Q) \subset M_2(\mathbb{C})^* = GL(2, \mathbb{C})$. By the proof of Theorem 4, $C^+(Q) \cong C(Q')$ where $Q' : \mathbb{C}^2 \to \mathbb{C}, (z_1, z_2) \to z_1z_2$. The maximal isotropic subspace of Q' is $F = \mathbb{C}e_1$. The Grassmann algebra $S = \bigwedge^0(F) \oplus \bigwedge^1(F) = \mathbb{C} + \mathbb{C}e_1 \cong \mathbb{C}^2$. Since $e_1e_2 + e_2e_1 = B(e_1, e_2) = 1, e_i^2 = Q'(e_i) = 0$, we can write any element $x \in C(Q')$ in the form $x = a + be_1 + ce_2 + de_1e_2$.

Let us find the corresponding matrix. It is enough to find it for $x = e_1, e_2$, and e_1e_2. By the proof of Theorem 3,

$$s_{e_1}(1) = e_1, \quad s_{e_1}(e_1) = e_1e_1 = 0;$$

$$s_{e_2}(1) = \phi_{e_2}(1) = 0, \quad s_{e_2}(e_1) = \phi_{e_2}(e_1) = B'(e_1, e_2) = 1;$$

$$s_{e_1e_2}(1) = s_{e_1}(s_{e_2}(1)) = 0, \quad s_{e_1e_2}(e_1) = s_{e_1}(s_{e_2}(e_1)) = s_{e_1}(1) = e_1.$$

So, in the basis $(1, e_1)$ of S, we get the matrices

$$s_{e_1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad s_{e_2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad s_{e_1e_2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad s_x = \begin{pmatrix} a+d & c \\ b & a \end{pmatrix}.$$

At this point it is better to write each x in the form $x = ae_1e_2 + be_2 + ce_1 + de_2e_1$ to be able to identify x with the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Now $C^+(Q)$ consists of elements

$$y = a(e_3e_1)(e_3e_2) + be_3e_2 + ce_3e_1 + d(e_3e_2)(e_3e_1) = -ae_1e_2 + be_3e_2 + ce_3e_1 - de_2e_1.$$

and $y \in G^+(Q)$ iff $y \in C(Q)^*$ and $ye_iy^{-1} \in E, i = 1, 2, 3$. Before we check these conditions, notice that

$$e_3^2 = 1, (e_1e_2)^2 = e_1e_2(-1 - e_2e_1) = -e_1e_2, \quad e_2e_1e_2 = e_2(-1 - e_2e_1)$$

$$= -e_2, \quad e_1e_2e_1 = e_1(-1 - e_1e_2) = -e_1.$$

Notice that here we use the form B whose restriction to H equals $-B'$. Clearly,

$$y^{-1} = (ad - bc)^{-1}(-de_1e_2 - be_3e_2 - ce_3e_1 - ae_2e_1),$$

where, of course, we have to assume that $ad - bc \neq 0$. We have

$$(-ae_1e_2 + be_3e_2 + ce_3e_1 - de_2e_1)e_3(-de_1e_2 - be_3e_2 - ce_3e_1 - ae_2e_1) =$$

$$= (-ae_1e_2e_3 - be_2 - ce_1 - de_2e_1e_3)(-de_1e_2 - be_3e_2 - ce_3e_1 - ae_2e_1) =$$

$$= -(ad + bc)e_1e_2e_3 - (ad + bc)e_2e_1e_3 - 2bde_2 - 2ace_1 = -(ad + bc)e_3 - 2bde_2 - 2ace_1.$$

Just to confirm that it is correct, notice that

$$Q(\phi(y)(e_3)) = \frac{1}{(ad - bc)^2}((ad + bc)^2 - 2bdac) = \frac{1}{(ad - bc)^2}(ad - bc)^2 = 1 = Q(e_3),$$

as it should be, because $\phi(y)$ is an orthogonal transformation. Similarly, we check that $ye_iy^{-1} \in E, i = 1, 2$. Thus

$$G^+(Q) \cong GL(2, \mathbb{C}).$$

The kernel of $\phi : G^+(Q) \to SO(Q)$ consists of scalar matrices. Restricting this map to the subgroup $SL(2,\mathbb{C})$ has the kernel ± 1. This agrees with our discussion in 15.1.

By Theorem 4, $C(Q) \cong C^+(Q) \otimes_{\mathbb{C}} Z$, where $Z \cong \mathbb{C} \oplus \mathbb{C}$. Hence $C(Q) \cong C^+(Q) \oplus C^+(Q) \cong \mathrm{Mat}_2(\mathbb{C})^2$. This shows that $C(Q)$ admits two isomorphic 2-dimensional representations. To find one of them explicitly, we need to assign to each $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{C}^3$ a matrix $A(\mathbf{x})$ such that

$$A(\mathbf{x})^2 = ||\mathbf{x}||^2 I_2 = (x_1^2 + x_2^2 + x_3^2) I_2, \quad A(\mathbf{x})A(\mathbf{y}) + A(\mathbf{y})A(\mathbf{x}) = 2\mathbf{x}\cdot\mathbf{y} I_2.$$

Here it is:

$$A(\mathbf{x}) = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}.$$

Let

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = A(e_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = A(e_2) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = A(e_3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{15.13}$$

In the physics literature these matrices are called the *Pauli matrices*. Then

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I_2, \quad \sigma_i\sigma_j = -\sigma_j\sigma_i, \quad i = 1,2,3, \quad i \neq j.$$

Also, if we write

$$a_0\sigma_0 + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 = \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix} \tag{15.14}$$

with real a_0, a_1, a_2, a_3, we obtain an isomorphism from $\mathbb{R}^4$ to $\mathrm{Mat}_2(\mathbb{C})$ such that

$$a_0^2 - a_1^2 - a_2^2 - a_3^2 = \det(a_0\sigma_0 + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3). \tag{15.15}$$

Notice that $SL(2,\mathbb{C})$ acts $\mathbb{R}^4$ by the formula

$$X \cdot A(\mathbf{x}) = X \cdot A(\mathbf{x}) \cdot X^*. \tag{15.16}$$

Notice that the set of matrices of the form $A(\mathbf{x})$ is the subset of Hermitean matrices in $\mathrm{Mat}_2(\mathbb{C})$. So the action is well defined. We will explain this homomorphism in the next lecture.

15.5 Finally let us define the *spinor group* $Spin(Q)$. This is a subgroup of the Clifford group $G^+(Q)$ such the kernel of the restriction of the canonical homomorphism $G^+(Q) \to SO(Q)$ to $Spin(Q)$ is equal to ± 1.

Consider the natural anti-automorphism of the tensor algebra $\rho' : T(E) \to T(E)$ defined on decomposable tensors by the formula

$$v_1 \otimes \ldots \otimes v_k \to v_k \otimes \ldots \otimes v_1.$$

Recall that an anti-homomorphism of rings $R \to R'$ is a homomorphism from R to the opposite ring R'^o (same ring but with the redefined multiplication law $x * y = y \cdot x$). It is clear that $\rho'(I(Q)) = I(Q)$ so ρ' induces an anti-automorphism $\rho : C(Q) \to C(Q)$. For any $x \in G(Q)$, we set

$$N(x) = x \cdot \rho(x). \tag{15.17}$$

By Corollary to Proposition 1, each $x \in G(Q)$ can be written in the form $x = zv_1 \dots v_k$ for some $z \in Z(C(Q))^*$ and non-isotropic vectors $v_1, \dots, v_k$ from E. We have

$$N(x) = zv_1 \dots v_k \rho(zv_1 \dots v_k) = z^2 v_1 \dots v_k v_k \dots v_1 = z^2 Q(v_1) \dots Q(v_k) \in K^*. \tag{15.18}$$

Also

$$N(x \cdot y) = xy\rho(xy) = xy\rho(y)\rho(x) = xN(y)\rho(x) = N(y)x\rho(x) = N(y)N(x).$$

This shows that the map $x \to N(x)$ defines a homomorphism of groups

$$N : G(Q) \to K^*. \tag{15.19}$$

It is called the *spinor norm* homomorphism. We define the *spinor group* (or *reduced Clifford group*) $Spin(Q)$ by setting

$$Spin(Q) = Ker(N) \cap G^+(Q). \tag{15.20}$$

Let

$$SO(Q)_0 = \mathrm{Im}(G^+(Q) \xrightarrow{\phi} SO(Q)). \tag{15.21}$$

This group is called the *reduced orthogonal group.*

From now on we assume that $\mathrm{char}(K) \neq 2$.

Obviously, $\mathrm{Ker}(\phi) \cap Spin(Q)$ consists of central elements z with $z^2 = 1$. This gives us the exact sequence

$$1 \to \{\pm 1\} \to Spin(Q) \to SO(Q)_0 \to 1. \tag{15.22}$$

It is clear that $N(K^*) \subset (K^*)^2$ so $\mathrm{Im}(N)$ contains the subgroup $(K^*)^2$ of K^*. If $K = \mathbb{R}$ and Q is definite, then (15.18) shows that $N(G^+) \subset (\mathbb{R}^*)^2 = \mathbb{R}_{>0}$. Since $N(\rho(x)) = N(x)$, we have $N(x/\sqrt{N(x)}) = N(x)/N(x) = 1$. Thus we can always represent an element of $SO(Q)$ by an element $x \in G^+(Q)$ with $N(x) = 1$. This shows that the canonical homomorphism $Spin(Q) \to SO(Q)$ is surjective, i.e., $SO(Q)_0 = SO(Q)$. In this case, if $n = \dim\ E$, the group $Spin(Q)$ is denoted by $Spin(n)$, and we obtain an exact sequence

$$1 \to \{\pm 1\} \to Spin(n) \to SO(n) \to 1. \tag{15.23}$$

One can show that $Spin(n), n \geq 3$ is the universal cover of the Lie group $SO(n)$.

If K is arbitrary, but E contains an isotropic vector $v \neq 0$, we know that $Q(w) = a$ has a solution for any $a \in K^*$ (because E contains a hyperbolic plane $Kv + Kv'$, with $Q(v) = Q(v') = 0, B(v, v') = b \neq 0$, and we can take $w = av + b^{-1}v'$). Thus if we choose $w, w' \in E$ with $Q(w) = a, Q(w') = 1$, we get $N(ww') = Q(w)Q(w') = a$. This shows that

$\mathrm{Im}(N(G^+(Q))) = K^*$. Under the homomorphism N the factor group $G^+(Q)/Spin(Q)$ is mapped isomorphically onto K^*. On the other hand, under ϕ this group is mapped surjectively to $SO(Q)/SO(Q)_0$ with kernel $(K^*)^2$. This shows that

$$SO(Q)/SO(Q)_0 \cong K^*/(K^*)^2. \tag{15.24}$$

In particular, when $K = \mathbb{R}$, and Q is indefinite, we obtain that $SO(Q)_0$ is of index 2 in $SO(Q)$. For example, when $E = \mathbb{R}^4$ with Lorentzian quadratic form, we see that $SO(1,3)_0$ is the proper Lorentz group, so even our notation agrees. As we shall see in the next lecture, $Spin(1,3) \cong SL(2,\mathbb{C})$. This agrees with section 15.1 of this lecture.

Exercises.

1. Show that the Lorentz group $O(1,3)$ has 4 connected components. Identify $SO(1,3)_0$ with the connected component containing the identity.

2. Assume $Q = 0$ so that $C(Q) \cong \bigwedge(E)$. Let $f \in E^*$ and $i_f : \bigwedge(E) \to \bigwedge(E)$ be the map defined in (15.3). Show that

$$i_f(x_1 \wedge \ldots \wedge x_k) = \sum_{i=1}^{k} (-1)^{i-1} f(x_i)(x_1 \wedge \ldots \wedge x_{i-1} \wedge x_{i+1} \wedge \ldots \wedge x_k).$$

3. Using the classification of quadratic forms over a finite field, classify Clifford algebras over a finite field.

4. Show that $C(Q)$ can be defined as follows. For any linear map $f : E \to D$ to some unitary algebra $C(Q)$ satisfying $f(v) = Q(v) \cdot 1$, there exists a unique homomorphism f' of algebras $C(Q) \to D$ such that its restriction to $E \subset C(Q)$ is equal to f.

5. Let Q be a non-degenerate quadratic form on a vector space of dimension 2 over a field K of characteristic different from 2. Prove that $C(Q) \cong M_2(K)$ if and only if there exists a vector $x \neq 0$ and two vectors y, z such that $Q(x) + Q(y)Q(z) = 0$.

6. Using the theory of Clifford algebras find the known relationship between the orthogonal group $O(3)$ and the group of pure quaternions $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, $a^2 + b^2 + c^2 \neq 0$.

7. Show that $Spin(2) \cong SU(1)$, $\quad Spin(3) \cong SU(2)$.

8. Prove that $Spin(1,2) \cong SL(2,\mathbb{R})$.

Lecture 16. THE DIRAC EQUATION

As we have seen in Lecture 13, the Klein-Gordon equation does not give the right relativistic picture of the 1-particle theory. The right approach is via the Dirac equation. Instead of a scalar field we shall consider a section of complex rank 4 bundle over $\mathbb{R}^4$ whose structure group is the spinor group $Spin(1,3)$ of the Lorentzian orthogonal group.

16.1 Let $O(1,3)$ be the Lorentz group. The corresponding Clifford algebra C is isomorphic to $\mathrm{Mat}_2(\mathbb{H})$ (Corollary 2 to Theorem 3 from Lecture 15). If we allow ourselves to extend $\mathbb{R}$ to $\mathbb{C}$, we obtain a 4-dimensional complex spinor representation of C. It is given by the *Dirac matrices*:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix}, \quad \gamma^j = \begin{pmatrix} 0 & -\sigma_j \\ \sigma_j & 0 \end{pmatrix}, \quad j = 1,2,3, \tag{16.1}$$

where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

are the Pauli matrices (15.13). It is easy to see that

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4,$$

where $g^{\mu\nu}$ is the inverse of the standard Lorentzian metric. Using (15.3), this implies that $e_i \to \gamma^i$ is a 4-dimensional complex representation $S = \mathbb{C}^4$ of C. Let us restrict it to the space $V = \mathbb{R}^4$. Then the image of V consists of matrices of the form

$$\begin{pmatrix} 0 & X \\ \mathrm{adj}(X) & 0 \end{pmatrix}, \tag{16.2}$$

where

$$X = \begin{pmatrix} x_0 - x_1 & ix_3 - x_2 \\ -x_2 - ix_3 & x_0 + x_1 \end{pmatrix} \tag{16.3}$$

is a Hermitian 2×2-complex matrix (i.e., $X^* = X$) and $\mathrm{adj}(X)$ is its adjugate matrix (i.e., $\mathrm{adj}(X) \cdot X = \det(X) I_2$). The Clifford group $G(Q) \subset C(Q)^*$ acts on the set of such matrices by conjugation (see (15.12)). Its subgroup $G^+(Q)$ must preserve the subspaces $S_L = \mathbb{R}e_1 + \mathbb{R}e_2$ and $S_R = \mathbb{R}e_3 + \mathbb{R}e_4$. Hence any $g \in G^+(Q)$ can be given by a block-diagonal matrix of the form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}. \tag{16.4}$$

It must satisfy

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} 0 & X \\ \mathrm{adj}(X) & 0 \end{pmatrix} \cdot \begin{pmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{pmatrix} = \begin{pmatrix} 0 & Y \\ \mathrm{adj}(Y) & 0 \end{pmatrix}$$

for some Hermitian matrix Y. This immediately implies that

$$A = \begin{pmatrix} B & 0 \\ 0 & (B^*)^{-1} \end{pmatrix}, \quad Y = BXB^*, \tag{16.5}$$

where $B \in GL(2,\mathbb{C})$ satisfies $|\det B| = 1$. Now it is easy to see that any $A \in G^+$ must satisfy the additional property that $\det B = 1$. This follows from the fact that G^+ is mapped surjectively onto the group $SO(1,3)$ and the kernel consists only of the matrices ± 1. Thus

$$G^+ = Spin(1,3) \cong SL(2,\mathbb{C}). \tag{16.6}$$

Also, we see that, if we identify vectors from V with Hermitian matrices X, the group $G^+ = SL(2,\mathbb{C})$ acts on V via transformations $X \to B \cdot X \cdot B^*$, which is in accord with equality (15.15) from Lecture 15.

It is clear that the restriction of the spinor representation S to G^+ splits into the direct sum $S_L \oplus S_R$ of two complex representations of dimension 2. One is given by $g \to A$, another one by $g \to (A^*)^{-1}$. The elements of S_L (resp. of S_R) are called *left-handed half-spinors* (resp. *right-handed half-spinors*). They are dual to each other and not isomorphic (as real 4-dimensional representations of G^+). We denote the vectors of the 4-dimensional space of spinors $S = \mathbb{C}^4$ by

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix},$$

where $\phi_L \in S_L, \phi_R \in S_R$. Note that the spinor representation S is self-dual, since

$$(g^*)^{-1} = \gamma^0 g (\gamma^0)^{-1}.$$

We can express the projector operators $p_R : S \to S_R$, $p_L : S \to S_L$ by the matrices

$$p_R = \frac{1+\gamma^5}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad p_L = \frac{1-\gamma^5}{2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

where

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{16.7}$$

(the reason for choosing superscript 5 but not 4 is to be able to use the same notation for this element even if we start indexing the Dirac matrices by numbers $1, 2, 3, 4$). The spinor group $Spin(1, 3)$ acts in S leaving S_R and S_L invariant. These are the half-spinor representations of G^+.

16.2 Now we can introduce the *Dirac equation*

$$(\gamma^\mu \partial_\mu + imI_4)\psi = 0. \tag{16.8}$$

The left-hand side, is a 4×4-matrix differential operator, called the *Dirac operator.* The constant m is the mass (of the electron). The function ψ is a vector function on the space-time with values in the spinor representation S. If we multiply both sides of (16.8) by $(\gamma^\mu \partial_\mu - imI_4)$, we get

$$(\gamma^\mu \partial_\mu)^2 \psi = (\partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2) I_4 \psi + m^2 \psi = 0.$$

This means that each component of ψ satisfies the Klein-Gordon equation.

In terms of the right and left-handed half-spinors the Dirac equation reads

$$\begin{pmatrix} (\sigma_0 \partial_0 - \sigma_j \partial_j)\psi_R \\ (\sigma_0 \partial_0 + \sigma_j \partial_j)\psi_L \end{pmatrix} = -im \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \tag{16.9}$$

Notice that the Dirac operator maps S_R to S_L.

Now let us see the behavior of the Dirac equation (16.8) under the proper Lorentz group $SO(1, 3)_0$. We let it act on $V = \mathbb{R}^4$ via its spinor cover $Spin(1, 3)$ by identifying $\mathbb{R}^4$ with the space of matrices (16.2). Now we let $g \in Spin(1, 3)$ act on the solutions $\psi : \mathbb{R}^4 \to S = \mathbb{C}^4$ by acting on the source by means of $\phi(g) \in SO(1, 3)_0$ and on the target by means of the spinor representation. We identify the arguments $x = (t = x_0, x_1, x_2, x_3)$ with matrices $M(x) = x_i \gamma^i$. An element $g \in Spin(1, 3)$ is identified with a matrix $A(g)$ from (16.4). If $\Lambda(g)$ is the corresponding to $SO(1, 3)_0$, then we have

$$A(g) \cdot M(x) \cdot A(g)^{-1} = M(\Lambda(g) \cdot x).$$

Thus the function $\psi(x)$ is transformed to the function ψ' such that

$$\psi'(x') = A(g) \cdot \psi(x), \quad x' = \Lambda(g) \cdot x \tag{16.10}$$

We have

$$\partial'_\mu \phi'(x') = \frac{\partial \phi'(x')}{\partial x'_\mu} = A(g) \Lambda(g)_{\mu\nu} \partial_\nu \phi(x),$$

$$A(g)^{-1}(\gamma^\mu \partial'_\mu + imI_4)\phi'(x') = A(g)^{-1} M(e_\mu) A(g)(\Lambda(g)^{-1} \partial_\mu)\phi + mA(g)^{-1}\phi'(x') =$$

$$= M(\Lambda^{-1}(g) \cdot e_\mu)(\Lambda(g)^{-1} \partial_\mu)\partial_\mu \phi + m\phi(x) = \gamma^\mu \partial_\mu \phi + m\phi(x) = 0.$$

This shows that the Dirac equation is invariant under the Lorentz group when we transform the function $\psi(x)$ according to (16.10). There are other nice properties of invariance. For example, let us set

$$\psi^* = (\psi_L, \psi_R)^* = (\bar{\psi}_R, \bar{\psi}_L) = \gamma^0 \bar{\psi} \tag{16.11}$$

Then

$$\psi^* \cdot \psi = \bar{\psi}_R \cdot \psi_L + \bar{\psi}_L \cdot \psi_R$$

is invariant. Indeed

$$\psi'(x')^* \cdot \psi'(x') = \bar{\psi}'(x')\gamma^0\psi'(x') = \bar{\psi}(x)A(g)^*\gamma^0 A(g) \cdot \psi =$$

$$= \bar{\psi}(x)\gamma^0 A(g)^*\gamma^0 A(g) \cdot \psi = \psi(x)^*\psi(x).$$

Here we used that

$$\gamma^0 A(g)^*\gamma^0 A(g) = \begin{pmatrix} 0 & I \\ 1 & 0 \end{pmatrix} \begin{pmatrix} B^* & 0 \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} 0 & I \\ 1 & 0 \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & B^{*-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

16.3 Now we shall forget about physics for a while and consider the analog of the Dirac operator when we replace the Lorentizian metric with the usual Euclidean metric. That is, we change the group $O(1,3)$ to $O(4)$, i.e. we consider $V = \mathbb{R}^4$ with the standard quadratic form $Q = x_1^2 + x_2^2 + x_3^2 + x_4^2$. The spinor complex representation $s : C(Q) \to M_4(\mathbb{C})$ is given by the matrices

$$\gamma'^0 = \begin{pmatrix} 0 & \sigma_0' \\ \sigma_0' & 0 \end{pmatrix}, \quad \gamma'^j = \begin{pmatrix} 0 & -\sigma_j' \\ \sigma_j' & 0 \end{pmatrix}, \quad j = 1,2,3, \tag{16.12}$$

obtained by replacing the Pauli matrices $\sigma_i, i = 1,2,3$ with

$$\sigma_0' = \sigma_0, \quad \sigma_1' = i\sigma_1, \quad \sigma_2' = i\sigma_2, \quad \sigma_3' = i\sigma_3.$$

The image $s(V)$ of V consists of matrices of the form

$$A = \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix},$$

where

$$X = \begin{pmatrix} a - bi & -c - di \\ c - di & a + bi \end{pmatrix} = a\sigma_0' + b\sigma_1' + c\sigma_2' + d\sigma_3', \quad a,b,c,d \in \mathbb{R}. \tag{16.13}$$

These are characterized by the condition $X^* = \mathrm{adj}(X)$. Note that

$$a^2 + b^2 + c^2 + d^2 = \det(s(a,b,c,d)).$$

If we identify (a,b,c,d) with the quaternion $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, this corresponds to the well-known identification of quaternions with complex matrices (16.13).

Similar to (16.5) in the previous section we see that elements of $s(G^+(Q))$ are matrices of the form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where $A_1, A_2 \in U(2), \det(A_1) = \det(A_2)$. In particular, the two half-spinor representations in this case are isomorphic. This is explained by the general theory. The center of the algebra $C^+(Q)$ is equal to $Z = \mathbb{R} + \gamma'^5\mathbb{R}$, where

$$\gamma'^5 = \gamma'_0\gamma'_1\gamma'_2\gamma'_3.$$

Since $(\gamma'^5)^2 = -I_4$, the center Z is a field. This implies that $C^+(Q)$ is a simple algebra, hence the restriction of the spinor representation to $C^+(Q)$ is isomorphic to the direct sum of two isomorphic 2-dimensional representations. Notice that in the Lorentzian case, the center is spanned by 1 and $\gamma^5 = i\gamma'_0\gamma'_1\gamma'_2\gamma'_3$ so that $(\gamma^5)^2 = I_4$. This implies that C^+ is not simple.

Now, arguing as in the Lorentzian case, we easily get that

$$Spin(4) \cong G^+(Q) \cong SU(2) \times SU(2) \cong Sp(1) \times Sp(1), \tag{16.14}$$

where $Sp(1)$ is the group of quaternions of norm 1. The isomorphism $Sp(1) \cong SU(2)$ is induced by the map $\mathbb{R}^4 \to GL_2(\mathbb{C})$ given by formula (16.13).

Let

$$S^+ = \mathbb{C}e_1 + \mathbb{C}e_2, \quad S^- = \mathbb{C}e_3 + \mathbb{C}e_4$$

They are our half-spinor representations of $G^+(Q)$. The Dirac matrices γ'^μ define a linear map

$$s : V \to \mathrm{Hom}_{\mathbb{C}}(S^+, S^-).$$

In the standard bases of $S^\pm$ this is given by the Pauli matrices

$$e_i \to \sigma'_i, \quad i = 0, \ldots, 3.$$

Let us view $S^\pm$ as the space of quaternions by writing

$$(a + bi, c + di) \to a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = (a + bi) + j(c - di).$$

Then the image $s(V)$ is the set of linear maps over $\mathbb{H}$. It is a line over the quaternion skew field. Equivalently, if we define the (anti-linear) operator $J : S^\pm \to S^\pm$ by

$$J(z_1, z_2) = (-\bar{z}_2, \bar{z}_1),$$

then $s(V)$ consists of linear maps satisfying $f(Jx) = Jf(x)$.

Let us equip S^+ and S^- with the standard Hermitian structure

$$\langle (z_1, z_2), (w_1, w_2) \rangle = z_1\bar{w}_1 + z_2\bar{w}_2.$$

Then for any $\psi : S^+ \to S^-$ we can define the adjoint map $\psi^* : S^- \to S^+$. It is equal to the composition $S^- \to S^+ \to S^- \to S^+$, where the first and the third arrow are defined by using the Hermitian pairing. The middle arrow is the map ψ. It is easy to see that

$$s(e_i) + s(e_i)^* : S^+ \oplus S^- \to S^- \oplus S^+$$

is given by the Dirac matrix γ'^i. This implies that, for any $v \in V$, the image of v in the spinor representation $\tilde{s} : C(Q) \to \mathrm{Mat}_4(\mathbb{C})$ is given by $s(v) + s(v)^*$. Since s is a homomorphism, we get, for any $v, v' \in C(Q)$,

$$\tilde{s}(vv') = \tilde{s}(v)\tilde{s}(v') = (s(v) + s(v)^*) \circ (s(v') + s(v')^*) = s(v)^* s(v') + s(v')^* s(v) = s(vv').$$

In particular, if v is orthogonal to v', we have $vv' = -v'v$, hence

$$s(v)^* s(v') + s(v')^* s(v) = 0 \quad \text{if } v \cdot v' = 0. \tag{16.15}$$

Also, if $||v|| = Q(v) = 1$, we must have $v^2 = 1$, hence

$$s(v)^* s(v) = 1 \quad \text{if } v \cdot v = 1. \tag{16.16}$$

Let $\phi : \bigwedge^2(V) \to \mathrm{End}(S^+)$ be defined by sending $v \wedge v'$ to $s(v)^* s(v') - s(v')^* s(v)$. If v, v' are orthogonal, we have by (8),

$$(s(v)^* s(v') - s(v')^* s(v))^* = -(s(v)^* s(v') - s(v')^* s(v)).$$

Since any $v \wedge v'$ can be written in the form $w \wedge w'$, where w, w' are orthogonal, we see that the image of ϕ is contained in the subspace of $\mathrm{End}(S^+)$ which consists of operators A such that $A^* + A = 0$. From Lecture 12, we know that this subspace is the Lie algebra of the Lie group of unitary transformations of S^+. So, we denote it by $\mathfrak{su}(S^+)$. Thus we have defined a linear map

$$\phi^+ : \bigwedge^2 V \to \mathfrak{su}(S^+). \tag{16.17}$$

Similarly we define the linear map

$$\phi^- : \bigwedge^2 V \to \mathfrak{su}(S^-). \tag{16.18}$$

To summarize, we make the following definition

Definition. Let V be a four-dimensional real vector space with the Euclidean inner product. A *spinor structure* on V is a homomorphism

$$s : V \to \mathrm{Hom}_{\mathbb{C}}(S^+, S^-)$$

where $S^{\pm}$ are two-dimensional complex vector spaces equipped with a Hermitian inner product. Additionally, it is required that the homomorphism preserves the inner products.

Here the inner product in $\mathrm{Hom}(S^+, S^-)$ is defined naturally by

$$||\phi||^2 = \left| \frac{\phi(e_1) \wedge \phi(e_2)}{e_1' \wedge e_2'} \right|,$$

where e_1, e_2 (resp. (e_1', e_2')) is an orthonormal basis in S^+ (resp. S^-).

Given a spin-structure on V, we can consider the action of $Spin(4) \cong SU(2) \times SU(2)$ on $\mathrm{Hom}_{\mathbb{C}}(S^+, S^-)$ via the formula

$$(A, B) \cdot f(\sigma) = A \cdot f(B^{-1} \cdot \sigma).$$

This action leaves $s(V)$ invariant and induces the action of $SO(4)$ on V.

16.4 Now let us globalize. Let M be an oriented smooth Riemannian 4-manifold. The tangent space $T(M)_x$ of any point $x \in M$ is equipped with a quadratic form $g_x : T(M)_x \to \mathbb{R}$ which is given by the metric on M. Thus we are in the situation of the previous lecture. First of all we define a *Hermitian vector bundle* over a manifold M. This is a complex vector bundle E together with a Hermitian form $Q : E \to 1_M$. The structure group of a Hermitian bundle can be reduced to $U(n)$, where n is the rank of E.

Definition. A *spin structure* on M is the data which consists of two complex Hermitian rank 2 vector bundles S^+ and S^- over M together with a morphism of vector bundles

$$s : T(M) \to \mathrm{Hom}(S^+, S^-)$$

such that for any $x \in M$, the induced map of fibres $s_x : T(M)_x \to Hom(S^+_x, S^-_x)$ preserves the inner products.

The Riemannian metric on M plus its orientation allows us to reduce the structure group of $T(M)_{\mathbb{C}}$ to the group $SO(4)$. A choice of a spin-structure allows us to lift it to the group $Spin(4) \cong SU(2) \times SU(2)$. Conversely, given such a lift, the two obvious homomorphisms $Spin(4) \to SU(2)$ define the associated vector bundles $S^{\pm}$ and the morphism $s : T(M) \to \mathrm{Hom}(S^+, S^-)$.

Theorem 1. *A spin-structure exists on M if and only if for any $\gamma \in H^2(M, \mathbb{Z}/2\mathbb{Z})$*

$$\gamma \cap \gamma = 0.$$

Moreover, it is unique if $H^1(M, \mathbb{Z}/2\mathbb{Z}) = 0$.

Proof. We give only a sketch. Recall that there is a natural bijective correspondence between the isomorphism classes of principal G-bundles and the cohomology set $H^1(M, \mathcal{O}_M(G))$. Let

$$1 \to (\mathbb{Z}/2\mathbb{Z})_M \to \mathcal{O}_M(Spin(4)) \to \mathcal{O}_M(SO(4)) \to 1$$

be the exact sequence of sheaves corresponding to the exact sequence of groups

$$1 \to \{\pm 1\} \to Spin(4) \to SO(4) \to 1.$$

Applying cohomology, we get an exact sequence

$$H^1(M, \mathbb{Z}/2\mathbb{Z}) \to H^1(M, \mathcal{O}_M(Spin(4)) \to H^1(M, \mathcal{O}_M(SO(4))) \to H^2(M, \mathbb{Z}/2\mathbb{Z}). \quad (16.18)$$

If $T(M)$ corresponds to an element $t \in H^1(M, \mathcal{O}_M(SO(4)))$, then its image $\delta(t)$ in $H^2(M, \mathbb{Z}/2\mathbb{Z})$ is equal to zero if and only if t is the image of an element t' from the cohomology group $\in H^1(M, \mathcal{O}_M(Spin(4))$ representing the isomorphism class of a principal $Spin(4)$-bundle. Clearly, this means that the structure group of $T(M)$ can be lifted to $Spin(4)$. One shows that $\delta(t)$ is the Stieffel-Whitney characteristic class $w_2(M)$ of $T(M)$. Now, by Wu's formula

$$\gamma \cap \gamma = \gamma \cap w_2(M).$$

This shows that the condition $w_2(M) = 0$ is equivalent to the condition from the assertion of Theorem 1. The uniqueness assertion follows immediately from the exact sequence (16.18).

Example. A compact complex 2-manifold M is called a K3-surface if $b_1(M) = 0, c_1(M) = 0$, where $c_1(M)$ is the Chern class of the complex tangent bundle $T(M)$. An example of a K3 surface is a smooth projective surface of degree 4 in $\mathbb{P}^3(\mathbb{C})$ of degree 4. One shows that M is simply-connected, so that $H^1(M, \mathbb{Z}) = 0$. It is also known that for any complex surface M, $w_2(M) \equiv c_1(M)$ mod 2. This implies that $w_2(M) = 0$, hence every K3 surface together with a choice of a Riemannian metric admits a unique spin structure.

16.5 To define the global Dirac operator, we have to introduce the Levi-Civita connection on the tangent bundle $T(M)$. Let A be any connection on the tangent bundle $T(M)$. It defines a covariant derivative

$$\nabla_A : \Gamma(T(M)) \to \Gamma(T(M)) \otimes T^*(M)).$$

Thus for any vector field τ, we have an endomorphism

$$\nabla_A^\tau : \Gamma(T(M)) \to \Gamma(T(M)), \quad \eta \to \nabla_\tau^A(\eta),$$

defined by

$$\nabla_A^\tau(\eta) = \langle \nabla_A(\eta), \tau \rangle,$$

where $\langle\ ,\ \rangle$ denotes the contraction $\Gamma(T(M) \otimes T^*(M)) \times \Gamma(T(M)) \to \Gamma(T(M)$. Recall now that we have an additional structure on $\Gamma(T(M))$ defined by the Lie bracket of vector fields. For any $\tau, \eta \in \Gamma(T(M))$ define the *torsion operator* by

$$T_A(\tau, \eta) = \nabla_A^\tau(\eta) - \nabla_A^\eta(\tau) - [\tau, \eta] \in \Gamma(T(M)). \qquad (16.20)$$

Obviously, T_A is a skew-symmetric bilinear map on $\Gamma(T(M)) \times \Gamma(T(M))$ with values in $\Gamma(T(M))$. In other words,

$$T_A \in \Gamma(\bigwedge^2(T^*(M) \otimes T(M)) = \mathcal{A}^2(T(M))(M).$$

One should not confuse it with the curvature

$$F_A \in \mathcal{A}^2(\mathrm{End}(T(M)))(M)$$

defined by (see (12.9))

$$F_A(\tau,\eta) = [\nabla_A^\tau, \nabla_\eta] - \nabla_A^{[\tau,\eta]}.$$

The connection ∇_A on $T(M)$ naturally yields a connection on $T^*(M)$ and on $T^*(M)^* \otimes T^*(M)$ (use the formula $\nabla(s\otimes s') = \nabla(s)\otimes s' + s\otimes\nabla(s')$). Locally, if $(\partial_1,\ldots,\partial_n)$ is a basis of vector fields on M, and $dx^1,\ldots,dx^n$ is the dual basis of 1-forms, we can write

$$T_A = T^i_{jk}\partial_i\otimes dx^j\otimes dx^k, \quad F_A = R^i_{j\ kl}\partial_i\otimes dx^j\otimes dx^k\otimes dx^l. \tag{16.21}$$

We shall keep the same notation ∇_A for it. Now, we can view a Riemannian metric g on M as a section of the vector bundle $T^*(M)\otimes T^*(M)$.

Lemma-Definition. Let g be a Riemannian metric on M. There exists a unique connection A on $T(M)$ satisfying the properties
(i) $T_A = 0$;
(ii) $\nabla_A(g) = 0$.
This connection is called the *Levi-Civita* (or *Riemannian*) connection.

The Levi-Civita connection ∇_{LC} is defined by the formula

$$g(\eta,\nabla^\tau_{LC}(\xi)) = -\eta g(\tau,\xi)+\xi g(\eta,\tau)+\tau g(\xi,\eta)+g(\xi,[\eta,\tau])+g(\eta,[\xi,\tau])-g(\tau,[\xi,\eta]), \tag{16.22}$$

where η,τ,ξ are arbitrary vector fields on M. Once checks that $\xi\to\nabla^\tau_{LC}(\xi)$ defined by this formula satisfies the Leibnitz rule, and hence is a covariant derivative. Using (16.20) we easily obtain

$$g(\xi,\nabla^\tau_{LC}(\eta)) + g(\eta,\nabla^\tau_{LC}(\xi)) = \tau g(\eta,\xi). \tag{16.23}$$

If locally

$$\nabla^{\partial_i}_{LC}(\partial_j) = \sum_k \Gamma^k_{ij}\partial_k, \quad g_{ij} = g(\partial_i,\partial_j),$$

then we get (*Christoffel's identity*)

$$\Gamma^l_{ij} = \frac{1}{2}\sum_k(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij})g^{kl}. \tag{16.24}$$

The vanishing of the torsion tensor is equivalent to

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

The meaning of (16.23) is the following. Let $\gamma:(a,b)\to M$ be an integral curve of a vector field τ. We say that a vector field ξ is *parallel* over γ if $\nabla^\tau_{LC}(\xi)_x = 0$ for any $x\in\gamma((0,1))$. Let $(x_1,\ldots,x_n)$ be a system of local parameters in an open subset $U\subset M$, and let $\gamma(t) = (x_1(t),\ldots,x_n(t)), \xi = \sum_j a^j\partial_j, \tau = \sum_i b^i\partial_i$. Since $\gamma(T)$ is an integral curve of τ, we have $b^i = \frac{dx^i}{dt}$. Then ξ is parallel over $\gamma(t)$ if and only if

$$\nabla^\tau_{LC}(\xi)_{\gamma(t)} = \nabla^\tau_{LC}(a^j\partial_j)_{\gamma(t)} = \frac{da^k(\gamma(t))}{dt}\partial_k + a^j\frac{dx^i(t)}{dt}\Gamma^k_{ij}\partial_k = 0.$$

This is equivalent to the following system of differential equations:

$$\frac{da^k(\gamma(t))}{dt} + a^j\frac{dx^i(t)}{dt}\Gamma^k_{ij} = 0, \quad k = 1, \ldots, n$$

In particular, τ is parallel over its integral curve γ if and only if

$$\frac{d^2x^k(t)}{dt^2} + \frac{dx^j(t)}{dt}\frac{dx^i(t)}{dt}\Gamma^k_{ij} = 0 \quad k = 1, \ldots, n.$$

Comparing this and (16.24) with Exercise 1 in Lecture 1, we find that this happens if and only if γ is a geodesic, or a critical path for the natural Lagrangian on the Riemannian manifold M.

Now formula (16.23) tells us that for any two vector fields η and ξ parallel over $\gamma(t)$, we have

$$\frac{d(g(\eta_{\gamma(t)}, \xi_{\gamma(t)}))}{dt} = 0.$$

In other words, the scalar product $g(\eta_x, \xi_x)$ is constant along $\gamma(t)$.

Remark. Let

$$F \in \Gamma(\mathrm{End}(T(M)) \otimes \bigwedge^2(T^*(M)) \subset \Gamma(T^*(M) \otimes T(M) \otimes T^*(M) \otimes T^*(M))$$

be the curvature tensor of the Levi-Civita connection. Let $Tr : T(M) \otimes T^*(M) \to 1_M$ be the natural contraction map. If we apply it to the product of the second and the third tensor factor in above, we obtain the linear map $\Gamma(\mathrm{End}(T(M) \otimes \bigwedge^2(T^*(M)) \to \Gamma(\bigwedge^2(T^*(M))$. The image of F is a 2-form R on M. It is called the *Ricci tensor* of the Riemannian manifold M. One proves that R is a symmetric tensor. In coordinate expression (16.21), we have

$$R = R_{kl}dx^k \otimes dx^l, \quad R_{kl} = \sum_i R^i_{k\ il}.$$

We can do other things with the curvature tensor. For example, we can use the inverse metric $g^{-1} \in \gamma(T(M) \otimes T(M))$ to contract R to get the scalar function $K : M \to \mathbb{R}$. It is called the *scalar curvature* of the Riemannian manifold M. We have

$$K = \sum_{k,l=1}^{n} g^{kl}R_{kl}. \tag{16.25}$$

When

$$R = \lambda g \tag{16.26}$$

for some scalar function $\lambda : M \to \mathbb{R}$, the manifold M is called an *Einstein space*. By contracting both sides, we get

$$K = g^{ij}R_{ij} = \lambda g^{ij}g_{ij} = n\lambda$$

One can show that K is constant for an Einstein space of dimension $n \geq 3$. The equation

$$R - \frac{K}{n} g = \frac{8\pi G}{c^4} T \tag{16.27}$$

is the *Einstein equation* of general relativity. Here c is the speed of light, G is the gravitational constant, and T is the energy-momentum tensor of the matter.

16.6 The notion of the Levi-Civita connection extends to any Hermitian bundle E. We require that $\nabla_A(h) = 0$, where $h : E \to 1_M$ is the positive definite Hermitian form on E. Such a connection is called a *unitary connection*.

Let us use the linear map (16.18) to define a unitary connection on the spinor bundles $S^{\pm}$ such that the induced connection on $\mathrm{Hom}(S^+, S^-)$ coincides with the Levi-Civita connection on $T(M)$ (after we identify $T(M)$ with its image $s(T(M))$. Let $\mathfrak{so}(n)$ be the Lie algebra of $SO(n)$. From Lecture 12 we know that $\mathfrak{so}(n)$ is isomorphic to the space of skew-symmetric real $n \times n$-matrices. This allows us to define an isomorphism of Lie algebras

$$\mathfrak{so}(n) \cong \bigwedge^2(V).$$

More explicitly, if we take the standard basis $(e_1, \ldots, e_n)$ of $\mathbb{R}^n$, then $e_i \wedge e_j, i < j$, is identified with the skew-symmetric matrix $E_{ij} - E_{ji}$, so that

$$[e_i \wedge e_j, e_k \wedge e_l] = [E_{ij} - E_{ji}, E_{kl} - E_{lk}] = [E_{ij}, E_{kl}] - [E_{ji}, E_{kl}] - [E_{ij}, E_{lk}] + [E_{ji}, E_{lk}]$$

$$= \begin{cases} 0 & \text{if } \#\{i,j\} \cap \{k,l\} \text{ is even,} \\ -E_{jl} & \text{if } i = k, j \neq l, \\ -E_{ik} & \text{if } j = l, i \neq k, \\ E_{jk} & \text{if } i = l, \\ E_{il} & \text{if } j = k. \end{cases}$$

We use this to verify that the linear maps (16.17) and (16.18) are homomorphism of Lie algebras

$$\psi^{\pm} : \mathfrak{so}(4) \to \mathfrak{su}(S^{\pm}).$$

Now the half-spinor representations of $Spin(4)$ define two associated bundles over M which we denote by $S^{\pm}$. If we identify the Lie algebra of $Spin(4)$ with the Lie algebra of $SO(4)$, then $\psi^{\pm}$ becomes the corresponding representation of the Lie algebras. Let A be the connection on the principal bundle of orthonormal frames which defines the Levi-Civita connection on the associated tangent bundle $T(M)$. Then A defines a connection on the associated vector bundles $S^{\pm}$.

Given a spin structure on M we can define the *Dirac operator*.

$$D : \Gamma(S^+) \to \Gamma(S^-). \tag{16.28}$$

Locally, we choose an orthonormal frame $e_1, \ldots, e_4$ in $T(M)$, and define, for any $\sigma \in \Gamma(S^+)$,

$$D\sigma = \sum_i s(e_i) \nabla_i(\sigma). \tag{16.28}$$

where $\nabla_i = \partial_i + A_i$ is the local expression of the Levi-Civita connection. We leave to the reader to verify that this definition is independent of the choice of a trivialization. We can also define the *adjoint Dirac operator*

$$D^* : \Gamma(S^-) \to \Gamma(S^+),$$

by setting

$$D^*\sigma = -\sum_i s(e_i)^* \nabla_i(s).$$

When M is equal to $\mathbb{R}^4$ with the standard Euclidean metric, the Christoffel identity (16.24) tells us that $\Gamma_{ij}^k = 0$, hence

$$\nabla_i(a^\mu \partial_\mu) = \partial_i(a^\mu)\partial_\mu.$$

We can identify $S^\pm$ with trivial Hermitian bundles $M \times \mathbb{C}^2$. A section of $S^\pm$ is a function $\psi^\pm : M \to \mathbb{C}^2$. We have

$$D\phi^+ = \sum_{i=1}^4 s(e_i)\partial_i(\phi^+) = (\sigma_0'\partial_0 + \sigma_1'\partial_1 + \sigma_2'\partial_2 + \sigma_3'\partial_3)\phi^+$$

Similarly,

$$D^*\phi^- = -\sum_{i=1}^4 s(e_i)^*\partial_i(\phi^-) = (-\sigma_0'\partial_0 + \sigma_1'\partial_1 + \sigma_2'\partial_2 + \sigma_3'\partial_3)\phi^-.$$

We verify immediately that

$$D^* \circ D = D \circ D^* = -\sum_{i=1}^4 (\frac{\partial}{\partial x^i})^2. \tag{16.30}$$

A solution of the *Dirac equation*

$$D\psi = 0 \tag{16.31}$$

is called a *harmonic spinor*. It follows from (16.30) that each coordinate of a harmonic spinor satisfies the Laplace equation.

An equivalent way to see the Dirac operator is to consider the composition

$$D : \Gamma(S^+) \xrightarrow{\nabla} \Gamma(S^+) \otimes T^*(M) \xrightarrow{c} \Gamma(S^-)$$

Here c is the Clifford multiplication arising from the spin-structure $T(M) \to Hom(S^+, S^-)$.

Finally, we can generalize the Dirac operator by throwing in another Hermitian vector bundle E with a unitary connection A and defining the Dirac operators

$$D_A : \Gamma(E \otimes S^+) \to \Gamma(E \otimes S^-), \quad D_A^* : \Gamma(E \otimes S^-) \to \Gamma(E \otimes S^+).$$

We shall state the next theorem without proof (see [**Lawson**]).

Theorem (Weitzenböck's Formula). *Let A be a unitary connection on a bundle over a 4-manifold M with a fixed spinor structure. For any section σ of $E \otimes S^+$ we have*

$$D_A^* D_A \sigma = (\nabla_A)^* \nabla_A \sigma - F_A^+ \sigma + \frac{1}{4} K \sigma,$$

where K is the scalar curvature of M.

16.7 As we saw not every Riemannian manifold admits a spin-structure. However, there is no obstruction to defining a *complex spin-structure* on any Riemannian manifold of even dimension. Let $Spin^c(4)$ denote the group of complex 4×4 matrices of the form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where $A_1, A_2 \in U(2), \det(A_1) = \det(A_2)$. The map $A \to \det(A_1)$ defines an exact sequence of groups

$$1 \to Spin(4) \to Spin^c(4) \to U(1) \to 1. \tag{16.32}$$

As we saw in section 15.3, the group $Spin^c(4)$ acts on the space $V = \mathbb{R}^4$ identified with the subspace $s(V)$, where $s : V \to \mathrm{Mat}_4(\mathbb{C})$ is the complex spinor representation. This defines a homomorphism of groups $Spin^c(4) \to SO(4)$. Since $Spin^c(4)$ contains $Spin(4)$, it is surjective. Its kernel is isomorphic to the subgroup of scalar matrices isomorphic to $U(1)$. The exact sequence

$$1 \to U(1) \to Spin^c(4) \to SO(4) \to 1 \tag{16.33}$$

gives an exact cohomology sequence

$$H^0(M, \underline{Spin^c}(4)) \to H^0(M, \underline{SO}(4)) \to H^1(M, \underline{U}(1)) \to$$

$$\to H^1(M, \underline{Spin^c}(4)) \to H^1(M, \underline{SO}(4)) \to H^2(M, \underline{U}(1)).$$

Here the underline means that we are considering the sheaf of smooth functions with values in the corresponding Lie group.

One can show that the map

$$H^0(M, \underline{Spin^c}(4)) \to H^0(M, \underline{SO}(4))$$

is surjective. This gives us the exact sequence

$$1 \to H^1(M, \underline{U}(1)) \to H^1(M, \underline{Spin^c}(4)) \to H^1(M, \underline{SO}(4)) \to H^2(M, \underline{U}(1)). \tag{16.34}$$

Consider the homomorphism $U(1) \to U(1)$ which sends a complex number $z \in U(1)$ to its square. It is surjective, and its kernel is isomorphic to the group $\mathbb{Z}/2$. The exact sequence

$$1 \to \mathbb{Z}/2\mathbb{Z} \to U(1) \to U(1) \to 1$$

defines an exact sequence of cohomology groups

$$H^1(M,\underline{U}(1)) \to H^2(M,\mathbb{Z}/2\mathbb{Z}) \to H^2(M,\underline{U}(1)).$$

Consider the inclusion $Spin(4) \subset Spin^c(4)$. We have the following commutative diagram

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & \mathbb{Z}/2 & \to & Spin(4) & \to & SO(4) & \to & 0 \\
 & & \downarrow & & \downarrow & & \| & & \\
0 & \to & U(1) & \to & Spin^c(4) & \to & SO(4) & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & U(1) & = & U(1) & \to & 0 & & \\
 & & \downarrow & & \downarrow & & & & \\
 & & 0 & & 0 & & & &
\end{array}$$

Here the middle vertical exact sequence is the sequence (16.32) and the middle horizontal sequence is the exact sequence (16.33). Applying the cohomology functor we get the following commutative diagram

$$\begin{array}{ccccc}
 & & & & H^1(M,\underline{U}(1)) \\
 & & & & \downarrow \\
 & & H^1(M,\underline{SO}(4)) & \to & H^2(M,\mathbb{Z}/2) \\
 & & \| & & \downarrow \\
H^1(M,\underline{Spin^c}(4))) & \to & H^1(M,\underline{SO}(4)) & \to & H^2(M,\underline{U}(1)).
\end{array}$$

From this we infer that the cohomology class $c \in H^1(M,\underline{SO}(4))$ of the principal $SO(4)$-bundle defining $T(M)$ lifts to a cohomology class $\tilde{c} \in H^1(M,\underline{Spin^c}(4))$ if and only if its image in $H^2(M,\underline{U}(1))$ is equal to zero. On the other hand, if we look at the previous horizontal exact sequence, we find that c is mapped to the second Stiefel-Whitney class $w_2(M) \in H^2(M,\mathbb{Z}/2\mathbb{Z})$. Thus, $\tilde{c}$ exists if and only if the image of $w_2(M)$ in $H^2(M,\underline{U}(1))$ in the right vertical exact sequence is equal to zero.

Now let us use the isomorphism $U(1) \cong \mathbb{R}/\mathbb{Z}$ induced by the map $\phi \to e^{2\pi i\phi}$. It defines the exact sequence of abelian groups

$$1 \to \mathbb{Z} \to \mathbb{R} \to U(1) \to 1, \tag{16.35}$$

and the corresponding exact sequence of sheaves

$$1 \to \mathbb{Z} \to \mathcal{O}_M \to \underline{U}(1)) \to 1, \tag{16.36}$$

It is known that $H^i(M, \mathcal{O}_M) = 0, i > 0$. This implies that

$$H^i(M, \underline{U}(1)) \cong H^{i+1}(M, \mathbb{Z}), \quad i > 0. \tag{16.37}$$

The right vertical exact sequence in the previous commutative diagram is equivalent to the exact sequence

$$H^2(M, \mathbb{Z}) \to H^2(M, \mathbb{Z}/2) \to H^3(M, \mathbb{Z})$$

defined by the exact sequence of abelian groups

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/2 \to 0.$$

It is clear that the image of $w_2(M)$ in $H^3(M, \mathbb{Z})$ is equal to zero if and only if there exists a class $w \in H^2(M, \mathbb{Z})$ such that its image in $H^2(M, \mathbb{Z}/2)$ is equal to $w_2(M)$. In particular, this is always true if $H^3(M, \mathbb{Z}) = 0$. By Poincaré duality, $H^3(M, \mathbb{Z}) \cong H^1(M, \mathbb{Z})$. So, any compact oriented 4-manifold with $H^1(M, \mathbb{Z}) = 0$ admits a spin^c-structure.

The exact sequence (16.34) and isomorphism $H^2(M, \mathbb{Z}) \cong H^1(M, \underline{U}(1))$ from (16.38) tells us that two lifts $\tilde{c}$ of c differ by an element from $H^2(M, \mathbb{Z})$. Thus the set of spin^c-structures is a principal homogeneous space with respect to the group $H^2(M, \mathbb{Z})$.

This can be interpreted as follows. The two homomorphisms $Spin^c(4) \to U(2)$ define two associated rank 2 Hermitian bundles $W^\pm$ over M and an isomorphism

$$c : T(M) \otimes \mathbb{C} \to Hom(W^+, W^-).$$

We have

$$\bigwedge^2 (W^+) \cong \bigwedge^2 (W^-) \cong L,$$

where L is a complex line bundle over M. The Chern class $c_1(L) \in H^2(M, \mathbb{Z})$ is the element $\tilde{c}$ which lifts $w_2(M)$. If one can find a line bundle M such that $M^{\otimes 2} \cong L$, then

$$W^\pm \cong S^\pm \otimes M.$$

Otherwise, this is true only locally. A unitary connection A on L allows one to define the Dirac operator

$$D_A : \Gamma(W^+) \to \Gamma(W^-). \tag{16.38}$$

Locally it coincides with the Dirac operator $D : \Gamma(S^+) \to \Gamma(S^-)$.

The *Seiberg-Witten equations* for a 4-manifold with a spin^c-structure $W^\pm$ are equations for a pair (A, ψ), where

(i) A is a unitary connection on $L = \bigwedge^2(W^+)$;
(ii) ψ is a section of W^+.

They are

$$D_A\psi = 0, \quad F_A^+ = -\tau(\psi), \tag{16.39}$$

where $\tau(\psi)$ is the self-dual 2-form corresponding to the trace-free part of the endomorphism $\psi^* \otimes \psi \in W^+ \otimes W^+ \cong \mathrm{End}(W^+)$ under the natural isomorphism $\Lambda^+ \to \mathrm{End}(W^+)$ which is analogous to the isomorphism from Exercise 4.

Exercises.

1. Show that the reversing the orientation of M interchanges D and D^*.

2. Show that the complex plane $\mathbb{P}^2(\mathbb{C})$ does not admit a spinor structure.

3. Prove the Christoffel identity.

4. Let $\Lambda^\pm$ be the subspace of $\bigwedge^2 V$ which consists of symmetric (resp. anti-symmetric) forms with respect to the star operator $*$. Show that the map ϕ^+ defined in (16.17) induces an isomorphism $\Lambda^+ \to \mathfrak{su}(S^+)$.

5. Show that the contraction $\sum_i R^i_{i\ kl}$ for the curvature tensor of the Levi-Civita connection must be equal to zero.

6. Prove that $Spin^c(4) \cong (U(1) \times Spin(4)/\{\pm 1\}$, where $\{\pm\}$ acts on both factors in the obvious way.

7. Show that the group $SO(4)$ admits two homomorphisms into $SO(3)$, and the associated rank 3 vector bundles are isomorphic to the bundles $\Lambda^\pm$ of self-dual and anti-self-dual 2-forms on M.

8. Show that the Dirac equation corresponds to the Lagrangian $\mathcal{L}(\psi) = \psi^*(i\gamma^\mu\partial_\mu - m)\psi$.

9. Prove the formula $Tr(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho})$.

10. Let $\sigma_{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. Show that the matrices of the spinor representation of $Spin(3,1)$ can be written in the form $A = e^{-i\sigma_{\mu\nu}a_{\mu\nu}/4}$, where $(a_{\mu\nu}) \in \mathfrak{so}(1,3)$.

11. Show that the current $J = (j^0, j^1, j^2, j^3)$, where $j^\mu = \bar{\psi}\gamma^\mu\psi$ is conserved with respect to Lorentz transformations.

Lecture 17. QUANTIZATION OF FREE FIELDS

In the previous lectures we discussed classical fields; it is time to make the transition to the quantum theory of fields. It is a generalization of quantum mechanics where we considered states corresponding to a single particle. In quantum field theory (QFT) we will be considering multiparticle states. There are different approaches to QFT. We shall start with the canonical quantization method. It is very similar to the canonical quantization in quantum mechanics. Later on we shall discuss the path integral method. It is more appropriate for treating such fields as the Yang-Mills fields. Other methods are the Gupta-Bleuler covariant quantization or the Becchi-Rouet-Stora-Tyutin (BRST) method, or the Batalin-Vilkovsky (BV) method. We shall not discuss these methods. We shall start with quantization of free fields defined by Lagrangians without interaction terms.

17.1 We shall start for simplicity with real scalar fields $\psi(t,\mathbf{x})$ described by the Klein-Gordon Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\psi)^2 - \frac{1}{2}m^2\psi^2.$$

By analogy with classical mechanics we introduce the conjugate *momentum field*

$$\pi(t,\mathbf{x}) = \frac{\delta\mathcal{L}}{\delta\partial_0\psi(t,\mathbf{x})} = \partial_0\psi(t,\mathbf{x}) := \dot{\psi}.$$

We can also introduce the Hamiltonian functional in two function variables ψ and π:

$$H = \frac{1}{2}\int(\pi\dot{\psi} - \mathcal{L})d^3x. \tag{17.1}$$

Then the Euler-Lagrange equation is equivalent to the *Hamilton equations* for fields

$$\dot{\psi}(t,\mathbf{x}) = \frac{\delta H}{\delta\pi(t,\mathbf{x})}, \quad \dot{\pi}(t,\mathbf{x}) = -\frac{\delta H}{\delta\psi(t,\mathbf{x})}. \tag{17.2}$$

Here we use the partial derivative of a functional $F(\psi,\pi)$. For each fixed ψ_0,π_0, it is a linear functional F'_ψ defined as

$$F(\psi_0+h,\pi_0)-F(\pi_0)=F'_\psi(\psi_0,\pi_0)(h)+o(||h||),$$

where h belongs to some normed space of functions on $\mathbb{R}^3$, for example $L_2(\mathbb{R}^3)$. We can identify it with the kernel $\frac{\delta F}{\delta\phi}$ in its integral representation

$$F'_\psi(\psi_0,\pi_0)(h)=\int\frac{\delta F}{\delta\phi}(\psi_0,\pi_0)h(\mathbf{x})d^3x.$$

Note that the kernel function has to be taken in the sense of distributions. We can also define the Poisson bracket of two functionals $A(\psi,\pi),B(\psi,\pi)$ of two function variables ψ and π (the analogs of $\mathbf{q},\mathbf{p}$ in classical mechanics):

$$\{A,B\}=\int_{\mathbb{R}^3}\Big(\frac{\delta A}{\delta\psi(\mathbf{z})}\frac{\delta B}{\delta\pi(\mathbf{z})}-\frac{\delta A}{\delta\pi(\mathbf{z})}\frac{\delta B}{\delta\psi(\mathbf{z})}\Big)d^3z. \tag{17.3}$$

To give it a meaning we have to understand $A(\psi,\pi),B(\psi,\pi)$ as bilinear distributions on some space of test functions K defined by

$$(f,g)\to\int_{\mathbb{R}^3}A(\psi,\pi)(\mathbf{z})f(z)g(\mathbf{z})d^3z,$$

and similar for $B(\psi,\pi)$. The product of two generalized functionals is understood in a generalized sense, i.e., as a bilinear distribution.

In particular, if we take B equal to the Hamiltonian functional H, and use (17.2), we obtain

$$\{A,H\}=\int_{\mathbb{R}^3}\Big(\frac{\delta A}{\delta\psi(\mathbf{z})}\dot{\psi}+\frac{\delta A}{\delta\pi(\mathbf{z})}\dot{\pi}(\mathbf{z})\Big)d^3z=\dot{A}. \tag{17.4}$$

Here

$$\dot{A}(\psi,\pi):=\frac{dA(\psi(t,\mathbf{x}),\pi(t,\mathbf{x}))}{dt}.$$

This is the Poisson form of the Hamilton equation (17.2).

Let us justify the following identity:

$$\{\psi(t,\mathbf{x}),\pi(t,\mathbf{y})\}=\delta(\mathbf{x}-\mathbf{y}). \tag{17.5}$$

Here the right-hand-side is the Dirac delta function considered as a bilinear functional on the space of test functions K

$$(f(\mathbf{x}),g(\mathbf{y}))\to\int_{\mathbb{R}^3}f(\mathbf{z})g(\mathbf{z})d^3z$$

(see Example 10 from Lecture 6). The functional $\psi(\mathbf{x})$ is defined by $F(\psi,\pi) = \psi(t,\mathbf{x})$ for a fixed $(t,\mathbf{x}) \in \mathbb{R}^4$. Since it is linear, its partial derivative is equal to itself. Similar interpretation holds for $\pi(\mathbf{y})$. If we denote the functional by $\psi(\mathbf{a})$, then we get

$$\frac{\delta\psi(\mathbf{x})}{\delta\psi(\mathbf{z})}(\psi,\pi) = \psi(t,\mathbf{x}) = \int_{\mathbb{R}^3} \psi(t,\mathbf{z})\delta(\mathbf{z}-\mathbf{x})d^3z.$$

Thus we obtain

$$\frac{\delta\psi(\mathbf{x})}{\delta\psi(\mathbf{z})} = \delta(\mathbf{z}-\mathbf{x}), \quad \frac{\delta\pi(\mathbf{y})}{\delta\pi(\mathbf{z})} = \delta(\mathbf{z}-\mathbf{y}), \quad \frac{\delta\psi(\mathbf{x})}{\delta\pi(\mathbf{z})} = \frac{\delta\pi(\mathbf{y})}{\delta\psi(\mathbf{z})} = 0.$$

Following the definition (17.3), we get

$$\{\psi(\mathbf{x}),\pi(\mathbf{y})\} = \int \delta(\mathbf{z}-\mathbf{x})\delta(\mathbf{z}-\mathbf{y})d^3z = \delta(\mathbf{x}-\mathbf{y}).$$

The last equality is the definition of the product of two distributions $\delta(\mathbf{z}-\mathbf{x})$ and $\delta(\mathbf{z}-\mathbf{y})$. Similarly, we get

$$\{\psi(\mathbf{x}),\psi(\mathbf{y})\} = \{\pi(\mathbf{x}),\pi(\mathbf{y})\} = 0.$$

Since it is linear, its partial derivative is equal to itself. Similar interpretation holds for $\pi(\mathbf{y})$. If we denote the functional by $\Psi(\mathbf{a})$, then we get

$$\frac{\delta\psi(\mathbf{x})}{\delta\psi(\mathbf{z})} \equiv \delta(\mathbf{z}-\mathbf{x}), \quad \frac{\delta\pi(\mathbf{y})}{\delta\Pi(\mathbf{z})} \equiv \delta(\mathbf{z}-\mathbf{y}), \quad \frac{\delta\Psi(\mathbf{x})}{\delta\Pi(\mathbf{z})} = \frac{\delta\Pi(\mathbf{y})}{\delta\Psi(\mathbf{z})} \equiv 0.$$

where we understand the delta function $\delta(\mathbf{z}-\mathbf{x})$ as the linear functional $h(\mathbf{z}) \to h(\mathbf{x})$. Thus we have

$$\{\psi(\mathbf{x}),\pi(\mathbf{y})\} = \int \delta(\mathbf{z}-\mathbf{x})\delta(\mathbf{z}-\mathbf{y})d^3z = \delta(\mathbf{x}-\mathbf{y})$$

Similarly, we get

$$\{\psi(\mathbf{x}),\psi(\mathbf{y})\} = \{\pi(\mathbf{x}),\pi(\mathbf{y})\} = 0.$$

17.2 To quantize the fields ψ and π we have to reinterpret them as Hermitian operators in some Hilbert space $\mathcal{H}$ which satisfy the commutation relations

$$[\Psi(t,\mathbf{x}),\Pi(t,\mathbf{y})] = \frac{i}{\hbar}\delta(\mathbf{x}-\mathbf{y}),$$

$$[\Psi(t,\mathbf{x}),\Psi(t,\mathbf{y})] = [\Pi(t,\mathbf{x}),\Pi(t,\mathbf{y})] = 0. \tag{17.6}$$

This is the quantum analog of (17.5). Here we have to consider Ψ and Π as operator valued distributions on some space of test functions $K \subset C^\infty(\mathbb{R}^3)$. If they are regular distributions, i.e., operators dependent on t and $\mathbf{x}$, they are defined by the formula

$$\Psi(t,f) = \int f(\mathbf{x})\Psi(t,\mathbf{x})d^3x.$$

Otherwise we use this formula as the formal expression of the linear continuous map $K \to \mathrm{End}(\mathcal{H})$. The operator distribution $\delta(\mathbf{x}-\mathbf{y})$ is the bilinear map $K \times K \to \mathrm{End}(\mathcal{H})$ defined by the formula

$$(f,g) \to (\int_{\mathbb{R}^3} f(\mathbf{x})g(\mathbf{x})d^3x)\mathbf{id}_{\mathcal{H}}.$$

Thus the first commutation relation from (17.6) reads as

$$[\Psi(t,f),\Pi(t,g)] = \Big(i\int_{\mathbb{R}^3} f(\mathbf{x})g(\mathbf{x})d^3x\Big)\mathbf{id}_{\mathcal{H}},$$

The quantum analog of the Hamilton equations is straightforward

$$\dot{\Pi} = \frac{i}{\hbar}[\Pi, H], \quad \dot{\Psi} = \frac{i}{\hbar}[\Psi, H], \tag{17.7}$$

where H is an appropriate Hamiltonian operator distribution.

Let us first consider the discrete version of quantization. Here we assume that the operator functions $\phi(t,\mathbf{x})$ belong to the space $L_2(B)$ of square integrable functions on a box $B = [-l_1,l_1]\times[-l_2,l_2]\times[-l_3,l_3]$ of volume $\Omega = 8l_1l_2l_3$. At fixed time t we expand the operator functions $\Psi(t,\mathbf{x})$ in Fourier series

$$\Psi(t,\mathbf{x}) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}q_{\mathbf{k}}(t). \tag{17.8}$$

Here

$$\mathbf{k} = (k_1,k_2,k_3) \in \frac{\pi}{l_1}\mathbb{Z}\times\frac{\pi}{l_2}\mathbb{Z}\times\frac{\pi}{l_3}\mathbb{Z}.$$

The reason for inserting the factor $\frac{1}{\sqrt{\Omega}}$ instead of the usual $\frac{1}{\Omega}$ is to accommodate with our definition of the Fourier transform: when $l_i \to \infty$ the expansion (17.8) passes into the Fourier integral

$$\Psi(t,\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3} e^{i\mathbf{k}\cdot\mathbf{x}}q_{\mathbf{k}}(t)d^3k. \tag{17.9}$$

Since we want the operators Ψ and Π to be Hermitian, we have to require that

$$q_{\mathbf{k}}(t) = q^*_{-\mathbf{k}}(t).$$

Recalling that the scalar fields $\psi(t,\mathbf{x})$ satisfied the Klein-Gordon equation, we apply the operator $(\partial_\mu\partial^\mu + m^2)$ to the operator functions Ψ and Π to obtain

$$(\partial_\mu\partial^\mu + m^2)\Psi = \sum_{\mathbf{k}}(\partial_\mu\partial^\mu + m^2)e^{i\mathbf{k}\cdot\mathbf{x}}q_{\mathbf{k}}(t) =$$

$$= \sum_{\mathbf{k}}(|\mathbf{k}|^2 + m^2 - E_{\mathbf{k}}^2)e^{i\mathbf{k}\cdot\mathbf{x}}q_{\mathbf{k}}(t) = 0,$$

provided that we assume that

$$q_{\mathbf{k}}(t) = q_{\mathbf{k}} e^{-iE_{\mathbf{k}}t}, \quad q_{-\mathbf{k}}(t) = q_{-\mathbf{k}} e^{iE_{\mathbf{k}}t},$$

where

$$E_{\mathbf{k}} = \sqrt{|\mathbf{k}|^2 + m^2}. \tag{17.10}$$

Similarly we consider the Fourier series for $\Pi(t, \mathbf{x})$

$$\Pi(t, \mathbf{x}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} p_{\mathbf{k}}(t)$$

to assume that

$$p_{\mathbf{k}}(t) = p_{\mathbf{k}} e^{-iE_{\mathbf{k}}t}, \quad p_{-\mathbf{k}}(t) = p_{-\mathbf{k}} e^{iE_{\mathbf{k}}t}.$$

Now let us make the transformation

$$q_{\mathbf{k}} = \frac{1}{\sqrt{2E_{\mathbf{k}}}}(a(\mathbf{k}) + a(-\mathbf{k})^*),$$

$$p_{-\mathbf{k}} = -i\sqrt{E_{\mathbf{k}}/2}(a(\mathbf{k}) - a(-\mathbf{k})^*).$$

Then we can rewrite the Fourier series in the form

$$\Psi(t, \mathbf{x}) = \sum_{\mathbf{k}} (2\Omega E_k)^{-1/2} (e^{i(\mathbf{k}\cdot\mathbf{x} - E_k t)} a(\mathbf{k}) + e^{-i(\mathbf{k}\cdot\mathbf{x} - E_k t)} a(\mathbf{k})^*), \tag{17.11a}$$

$$\Pi(t, \mathbf{x}) = \sum_{\mathbf{k}} -i(E_k/2\Omega)^{1/2} (e^{i(\mathbf{k}\cdot\mathbf{x} - E_k t)} a(\mathbf{k}) - e^{-i(\mathbf{k}\cdot\mathbf{x} - E_k t)} a(\mathbf{k})^*). \tag{17.11b}$$

Using the formula for the coefficients of the Fourier series we find

$$q_{\mathbf{k}}(t) = \frac{1}{\Omega^{1/2}} \int_B e^{-i\mathbf{k}\cdot\mathbf{x}} \Psi(t, \mathbf{x}) d^3x,$$

$$p_{\mathbf{k}}(t) = \frac{1}{\Omega^{1/2}} \int_B e^{i\mathbf{k}\cdot\mathbf{x}} \Pi(t, \mathbf{x}) d^3x.$$

Then

$$[q_{\mathbf{k}}, p_{\mathbf{k}'}] = \frac{1}{\Omega} \int_B \int_B e^{-i(\mathbf{k}\cdot\mathbf{x} - \mathbf{k}'\cdot\mathbf{x}')} [\Psi(t, \mathbf{x}), \Pi(t, \mathbf{x}')] d^3x d^3x' =$$

$$= \frac{i}{\Omega} \int_B e^{-i(\mathbf{k}\cdot\mathbf{x} - \mathbf{k}'\cdot\mathbf{x})} \delta(\mathbf{x} - \mathbf{x}') d^3x d^3x' = \frac{i}{\Omega} \int_B e^{-i(\mathbf{k} - \mathbf{k}')\cdot\mathbf{x}} d^3\mathbf{x} = i\delta_{\mathbf{k}\mathbf{k}'},$$

where

$$\delta_{\mathbf{k}\mathbf{k}'} = \begin{cases} 1 & \text{if } \mathbf{k} = \mathbf{k}', \\ 0 & \text{if } \mathbf{k} \neq \mathbf{k}'. \end{cases}$$

is the Kronecker symbol. Similarly, we get

$$[p_{\mathbf{k}}, p_{\mathbf{k}'}] = [q_{\mathbf{k}}, q_{\mathbf{k}'}] = 0.$$

This immediately implies that

$$[a(\mathbf{k}), a(\mathbf{k}')^*] = \delta_{\mathbf{k}\mathbf{k}'},$$
$$[a(\mathbf{k}), a(\mathbf{k}')] = [a(\mathbf{k})^*, a(\mathbf{k}')^*] = 0. \tag{17.12}$$

This is in complete analogy with the case of the harmonic oscillator, where we had only one pair of operators a, a^* satisfying $[a, a^*] = 1, [a, a] = [a^*, a^*] = 0$. We call the operators $a(\mathbf{k}), a(\mathbf{k})^*$ the *harmonic oscillator annihilation* and *creation operators*, respectively.

Conversely, if we assume that (17.12) holds, we get

$$[\Psi(t, \mathbf{x}), \Pi(t, \mathbf{y})] = \frac{i}{\Omega^{1/2}} \sum_{\mathbf{k}} e^{i(\mathbf{x}-\mathbf{y})\mathbf{k}}.$$

Of course this does not make sense. However, if we replace the Fourier series with the Fourier integral (17.9), we get (17.6)

$$[\Psi(t, \mathbf{x}), \Pi(t, \mathbf{y})] = \frac{i}{(2\Pi)^{3/2}} \int_{\mathbb{R}^3} e^{i(\mathbf{x}-\mathbf{y})\mathbf{k}} d^3k = i\delta(\mathbf{x} - \mathbf{y})$$

(see Lecture 6, Example 5).

The Hamiltonian (17.1) has the form

$$H = \int (\Pi\dot{\Psi} - \mathcal{L}) d^3x = \int (\Pi^2 - \frac{1}{2}\Pi^2 + \frac{1}{2}(\partial_1^2 + \partial_2^2 + \partial_3^2)\Psi + \frac{1}{2}m^2\Psi^2) d^3x =$$

$$= \int (\Pi^2 + \frac{1}{2}(\partial_1^2 + \partial_2^2 + \partial_3^2)\Psi + \frac{1}{2}m^2\Psi^2) d^3x.$$

To quantize it, we plug in the expressions for Π and Ψ from (17.11) and use (17.10). We obtain

$$H = \sum_{\mathbf{k}} E_k (a(\mathbf{k})^* a(\mathbf{k}) + \frac{1}{2}).$$

However, it obviously diverges (because of the $\frac{1}{2}$). Notice that

$$\frac{1}{2} : a(\mathbf{k})^* a(\mathbf{k}) + a(\mathbf{k}) a(\mathbf{k})^* := a(\mathbf{k})^* a(\mathbf{k}), \tag{17.13}$$

where $: P(a^*, a) :$ denotes the *normal ordering*; we put all $a(\mathbf{k})^*$'s on the left pretending that the a^* and a commute. So we can rewrite, using (17.12),

$$: a(\mathbf{k})^* a(\mathbf{k}) + \frac{1}{2} :=: a(\mathbf{k})^* a(\mathbf{k}) + \frac{1}{2}(a(\mathbf{k})^* a(\mathbf{k}) - a(\mathbf{k}) a(\mathbf{k})^*) :=$$

$$\frac{1}{2} : a(\mathbf{k})^* a(\mathbf{k}) + a(\mathbf{k}) a(\mathbf{k})^* := a(\mathbf{k})^* a(\mathbf{k}).$$

This gives

$$: H := \sum_{\mathbf{k}} E_{\mathbf{k}} a(\mathbf{k})^* a(\mathbf{k}). \tag{17.14}$$

Let us take this for the definition of the Hamiltonian operator H. Notice the commutation relations

$$[H, a(\mathbf{k})^*] = E_{\mathbf{k}} a(\mathbf{k})^*, \quad [H, a(\mathbf{k})] = E_{\mathbf{k}} a(\mathbf{k}). \tag{17.15}$$

This is analogous to equation (7.3) from Lecture 7.

We leave to the reader to check the Hamilton equations

$$\dot{\Psi}(t, \mathbf{x}) = i[\Psi(t, \mathbf{x}), H].$$

Similarly, we can define the *momentum vector operator*

$$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k} a(\mathbf{k})^* a(\mathbf{k}) = (P_1, P_2, P_3). \tag{17.16}$$

We have

$$\partial_\mu \Psi(t, \mathbf{x}) = i[P_\mu, \Psi]. \tag{17.17}$$

We can combine H and $\mathbf{P}$ into one 4-momentum operator $P = (H, \mathbf{P})$ to obtain

$$[\Psi, P^\mu] = i\partial^\mu \Psi, \tag{17.18}$$

where as always we switch to superscripts to denote the contraction with $g^{\mu\nu}$. Let us now define the *vacuum state* $|0\rangle$ as follows

$$a(\mathbf{k})|0\rangle = 0, \quad \forall \mathbf{k}.$$

We define a *one-particle state* by

$$a(\mathbf{k})^*|0\rangle = |\mathbf{k}\rangle.$$

We define an *n-particle state* by

$$|\mathbf{k}_n \ldots \mathbf{k}_1\rangle := a(\mathbf{k}_n)^* \ldots a(\mathbf{k}_1)^*|0\rangle$$

The states $|\mathbf{k}_1 \ldots \mathbf{k}_n\rangle$ are eigenvectors of the Hamiltonian operator:

$$H|\mathbf{k}_n \ldots \mathbf{k}_1\rangle = (E_{\mathbf{k}_1} + \ldots + E_{\mathbf{k}_n})|\mathbf{k}_n \ldots \mathbf{k}_1\rangle. \tag{17.19}$$

This is easily checked by induction on n using the commutation relations (17.12). Similarly, we get

$$P|\mathbf{k}_n \ldots \mathbf{k}_1\rangle = (\mathbf{k}_1 + \ldots + \mathbf{k}_n)|\mathbf{k}_n \ldots \mathbf{k}_1\rangle. \tag{17.20}$$

This tells us that the total energy and momentum of the state $|\mathbf{k}_n \dots \mathbf{k}_1\rangle$ are exactly those of a collection of n free particles of momenta $\mathbf{k}_i$ and energy $E_{\mathbf{k}_i}$. The operators $a(\mathbf{k})^*$ and $a(\mathbf{k})$ create and annihilate them. It is the stationary state corresponding to n particles with momentum vectors $\mathbf{k}_i$.

If we apply the *number operator*

$$N = \sum_{\mathbf{k}} a(\mathbf{k})^* a(\mathbf{k}) \tag{17.21}$$

we obtain

$$N|\mathbf{k}_n \dots \mathbf{k}_1\rangle = n|\mathbf{k}_n \dots \mathbf{k}_1\rangle.$$

Notice also the relation (17.10) which, after the physical units are restored, reads as

$$E_{\mathbf{k}} = \sqrt{m^2c^4 + |\mathbf{k}|^2c^2}.$$

The Klein-Gordon equations which we used to derive our operators describe spinless particles such as *pi mesons*.

In the continuous version, we define

$$\Psi(t,\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} (\frac{1}{\sqrt{2E_{\mathbf{k}}}} a(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{x} - E_{\mathbf{k}}t)} + a(\mathbf{k})^* e^{-i(\mathbf{k}\cdot\mathbf{x} - E_{\mathbf{k}}t)}) d^3k, \tag{17.22a}$$

$$\Pi(t,\mathbf{x}) = \frac{-i}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} (\sqrt{E_{\mathbf{k}}/2}(a(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{x} - E_{\mathbf{k}}t)} - a(\mathbf{k})^* e^{-i(\mathbf{k}\cdot\mathbf{x} - E_{\mathbf{k}}t)}) d^3k. \tag{17.22b}$$

Then we repeat everything replacing $\mathbf{k}$ with a continuous parameter from $\mathbb{R}^3$. The commutation relation (17.12) must be replaced with

$$[a(\mathbf{k}), a(\mathbf{k}')^*] = i\delta(\mathbf{k} - \mathbf{k}').$$

The Hamiltonian, momentum and number operators now appear as

$$H = \int_{\mathbb{R}^3} E_{\mathbf{k}} a(\mathbf{k})^* a(\mathbf{k}) d^3k, \tag{17.23}$$

$$\mathbf{P} = \int_{\mathbb{R}^3} \mathbf{k} a(\mathbf{k})^* a(\mathbf{k}) d^3k, \tag{17.24}$$

$$N = \int a(\mathbf{k})^* a(\mathbf{k}) d^3k. \tag{17.25}$$

17.3 We have not yet explained how to construct a representation of the algebra of operators $a(\mathbf{k})^*$ and $a(\mathbf{k})$, and, in particular, how to construct the vacuum state $|0\rangle$. The Hilbert space $\mathcal{H}$ where these operators act admits different realizations. The most frequently used one is the realization of $\mathcal{H}$ as the Fock space which we encountered in Lecture 9. For simplicity we consider its discrete version. Let $V = l_2(\mathbb{Z}^3)$ be the space of square integrable complex valued functions on $\mathbb{Z}^3$. We can identify such a function with a complex formal power series

$$f = \sum_{\mathbf{k}\in\mathbb{Z}^3} c_{\mathbf{k}}\mathbf{t}^{\mathbf{k}}, \quad \sum_{\mathbf{k}} |c_{\mathbf{k}}|^2 < \infty. \tag{17.26}$$

Here $\mathbf{t}^{\mathbf{k}} = t_1^{k_1}t_2^{k_2}t_3^{k_3}$ and $c_{\mathbf{k}} \in \mathbb{C}$. The value of f at $\mathbf{k}$ is equal to $c_{\mathbf{k}}$.

Consider now the complete tensor product $T^n(V) = V\hat{\otimes}\ldots\hat{\otimes}V$ of n copies of V. Its elements are convergent series of the form

$$F_n = \sum_{i_1,\ldots,i_n=0}^{\infty} c_{i_1\ldots i_n} f_{i_1}\otimes\ldots\otimes f_{i_n}, \tag{17.27}$$

where the convergence is taken with respect to the norm defined by $\sum |c_{i_1\ldots i_n}|^2$. Let

$$T(V) = \bigoplus_{n=0}^{\infty} T^n(V)$$

be the tensor algebra built over V. Its homogeneous elements $F_n \in T^n(V)$ are functions in n non-commuting variables $\mathbf{k}_1,\ldots,\mathbf{k}_n \in \mathbb{Z}^3$. Let $\hat{T}(V)$ be the completion of $T(V)$ with respect to the norm defined by the inner product

$$\langle F, G\rangle = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\mathbf{k}_1,\ldots,\mathbf{k}_n} F_n^*(\mathbf{k}_1,\ldots,\mathbf{k}_n)G_n(\mathbf{k}_1,\ldots,\mathbf{k}_n). \tag{17.28}$$

Its elements are the Cauchy equivalence classes of convergent sequences of complex functions of the form

$$F = \{F_n(\mathbf{k}_1,\ldots,\mathbf{k}_n)\}.$$

We take

$$\mathcal{H} = \{F = \{F_n\} \in \hat{T}(V) : F_n(\mathbf{k}_{\sigma(1)},\ldots,\mathbf{k}_{\sigma(n)}) = F_n(\mathbf{k}_1,\ldots,\mathbf{k}_n), \forall\sigma\in S_n\} \tag{17.29}$$

to be the subalgebra of $\hat{T}(V)$ formed by the sequences where each F_n is a symmetric function on $(\mathbb{Z}^3)^n$. This will be the bosonic case. There is also the fermionic case where we take $\mathcal{H}$ to be the subalgebra of $\hat{T}(V)$ which consists of anti-symmetric functions. Let $\mathcal{H}_n$ denote the subspaces of constant sequences of the form $(0,\ldots,0,F_n,0,\ldots)$. If we interpret functions $f\in V$ as power series (17.26), we can write $F_n \in \mathcal{H}_n$ as the convergent series

$$F_n = \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_n\in\mathbb{Z}^3} c_{\mathbf{k}_1\ldots\mathbf{k}_n}\mathbf{t}_1^{\mathbf{k}_1}\ldots\mathbf{t}_n^{\mathbf{k}_n}.$$

The operators $a(\mathbf{k})^*, a(\mathbf{k})$ are, in general, operator valued distributions on $\mathbb{Z}^3$, i.e., linear continious maps from some space $K \subset V$ of test sequences to the space of linear operators in $\mathcal{H}$. We write them formally as

$$A(\phi) = \sum_{\mathbf{k}} \phi(\mathbf{k}) A(\mathbf{k}),$$

where $A(\mathbf{k})$ is a generalized operator valued function on $\mathbb{Z}^3$. We set

$$a(\phi)^*(F)_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) = \sum_{i=1}^{n} \phi(\mathbf{k}_j) F_{n-1}(\mathbf{k}_1, \ldots, \hat{\mathbf{k}}_j, \ldots, \mathbf{k}_n), \tag{17.30a}$$

$$a(\phi)(F)_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) = \sum_{\mathbf{k}} \phi(\mathbf{k}) F_{n+1}(\mathbf{k}, \mathbf{k}_1, \ldots, \mathbf{k}_n), \tag{17.30b}$$

In particular, we may take ϕ equal to the characteristic function of $\{\mathbf{k}\}$ (corresponding to the monomial $\mathbf{t}^{\mathbf{k}}$). Then we denote the corresponding operators by $a(\mathbf{k}), a(\mathbf{k})^*$.

The vacuum vector is

$$|0\rangle = (1, 0, \ldots, 0, \ldots). \tag{17.31}$$

Obviously

$$a(\mathbf{k})|0\rangle = 0,$$

$$a(\mathbf{k})^*|0\rangle = (0, \mathbf{t}^{\mathbf{k}}, 0, \ldots, 0, \ldots),$$

$$a(\mathbf{k}_2)^* a(\mathbf{k}_1)^*|0\rangle(\mathbf{p}_1, \mathbf{p}_2) = (0, 0, \delta_{\mathbf{k}_2\mathbf{p}_1}\mathbf{t}_2^{k_1} + \delta_{\mathbf{k}_2\mathbf{p}_2}\mathbf{t}_1^{k_1}, 0, \ldots) = (0, 0, \mathbf{t}_1^{k_2}\mathbf{t}_2^{k_1} + \mathbf{t}_1^{k_1}\mathbf{t}_2^{k_2}, 0, \ldots).$$

Similarly, we get

$$a(\mathbf{k}_n)^* \ldots a(\mathbf{k}_1)^*|0\rangle = |\mathbf{k}_n \ldots \mathbf{k}_1\rangle = (0, \ldots, 0, \sum_{\sigma \in S_n} \mathbf{t}_1^{\mathbf{k}_{\sigma(1)}} \cdots \mathbf{t}_n^{\mathbf{k}_{\sigma(n)}}, 0 \ldots).$$

One can show that the functions of the form

$$F_{\mathbf{k}_1, \ldots, \mathbf{k}_n} = \sum_{\sigma \in S_n} \mathbf{t}_1^{\mathbf{k}_{\sigma(1)}} \cdots \mathbf{t}_n^{\mathbf{k}_{\sigma(n)}}$$

form a basis in the Hilbert space $\mathcal{H}$. We have

$$\langle F_{\mathbf{k}_1, \ldots, \mathbf{k}_n}, F_{\mathbf{q}_1, \ldots, \mathbf{q}_m} \rangle = \delta_{mn} \frac{1}{n!} \sum_{\mathbf{p}_1, \ldots, \mathbf{p}_n} (\sum_{\sigma \in S_n} \prod_{i=1}^{n} \delta_{\mathbf{k}_{\sigma(i)}\mathbf{p}_i})(\sum_{\sigma \in S_n} \prod_{i=1}^{n} \delta_{\mathbf{q}_{\sigma(i)}\mathbf{p}_i}) =$$

$$= n! \delta_{mn} \prod_{i=1}^{n} \delta_{\mathbf{k}_i \mathbf{q}_i}. \tag{17.32}$$

Thus the functions

$$\frac{1}{\sqrt{n!}} |\mathbf{k}_n \ldots \mathbf{k}_1\rangle$$

form an orthonormal basis in the space $\mathcal{H}$.

Let us check the commutation relations. We have for any homogeneous $F_n(\mathbf{k}_1,\ldots,\mathbf{k}_n)$

$$a(\phi)^*a(\phi')(F_n)(\mathbf{k}_1,\ldots,\mathbf{k}_n) = a(\phi)^*(\sum_{\mathbf{k}}\phi'(\mathbf{k})F_n(\mathbf{k},\mathbf{k}_1,\ldots,\mathbf{k}_{n-1})) =$$

$$=\sum_{i=1}^{n}\phi(\mathbf{k}_i)\sum_{\mathbf{k}}\phi'(\mathbf{k})F_n(\mathbf{k},\mathbf{k}_1,\ldots,\hat{\mathbf{k}}_i,\ldots,\mathbf{k}_n),$$

$$a(\phi')a(\phi)^*(F_n)(\mathbf{k}_1,\ldots,\mathbf{k}_n) = a(\phi')(\sum_{i=1}^{n+1}\phi(\mathbf{k}_j)F_n(\mathbf{k}_1,\ldots,\hat{\mathbf{k}}_j,\ldots,\mathbf{k}_{n+1})) =$$

$$=\sum_{\mathbf{k}}(\phi'(\mathbf{k})\phi(\mathbf{k})F_n(\mathbf{k}_1,\ldots,\mathbf{k}_n)) + \sum_{i=1}^{n}\phi'(\mathbf{k})\phi(\mathbf{k}_i)F_n(\mathbf{k},\mathbf{k}_1\ldots,\hat{\mathbf{k}}_i,\ldots,\mathbf{k}_n)).$$

This gives

$$[a(\phi)^*,a(\phi')](F_n)(\mathbf{k}_1,\ldots,\mathbf{k}_n) = \sum_{\mathbf{k}}\phi'(\mathbf{k})\phi(\mathbf{k})F_n(\mathbf{k}_1,\mathbf{k}_2,\ldots,\mathbf{k}_n)),$$

so that

$$[a(\phi)^*,a(\phi')] = \Big(\sum_{\mathbf{k}}\phi'(\mathbf{k})\phi(\mathbf{k})\Big)\mathbf{id}_{\mathcal{H}}.$$

In particular,

$$[a(\mathbf{k})^*,a(\mathbf{k}')] = \delta_{\mathbf{k}\mathbf{k}'}\mathbf{id}_{\mathcal{H}}.$$

Similarly, we get the other commutation relations.

We leave to the reader to verify that the operators $a(\phi)^*$ and $a(\phi)$ are adjoint to each other, i.e.,

$$\langle a(\phi)F,G\rangle = \langle F,a(\phi)^*G\rangle$$

for any $F,G\in\mathcal{H}$.

Remark 1. One can construct the Fock space formally in the following way. We consider a Hilbert space with an orthonormal basis formed by symbols $|a(\mathbf{k}_n)\ldots a(\mathbf{k}_1)\rangle$. Any vector in this space is a convergent sequnece $\{F_n\}$, where

$$F_n = \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_n} c_{\mathbf{k}_1,\ldots,\mathbf{k}_n}|\mathbf{k}_n\ldots\mathbf{k}_1\rangle.$$

We introduce the operators $a(\mathbf{k})^*$ and $a(\mathbf{k})$ via their action on the basis vectors

$$a(\mathbf{k})^*|\mathbf{k}_n\ldots\mathbf{k}_1\rangle = |\mathbf{k}\mathbf{k}_n\ldots\mathbf{k}_1\rangle,$$

$$a(\mathbf{k})|\mathbf{k}_n\ldots\mathbf{k}_1\rangle = |\mathbf{k}_n\ldots\hat{\mathbf{k}}_j\ldots\mathbf{k}_1\rangle$$

where $\mathbf{k} = \mathbf{k}_j$ for some j and zero otherwise. Then it is possible to check all the needed commutation relations.

17.4 Now we have to learn how to quantize generalized functionals $A(\Psi, \Pi)$.

Consider the space $\mathcal{H}\hat{\otimes}\mathcal{H}$. Its elements are convergent double sequences

$$F(\mathbf{k}, \mathbf{p}) = \{F_{nm}(\mathbf{k}_1, \ldots, \mathbf{k}_n; \mathbf{p}_1, \ldots, \mathbf{p}_m)\}_{m,n},$$

where

$$F_{nm}(\mathbf{k}_1, \ldots, \mathbf{k}_n; \mathbf{p}_1, \ldots, \mathbf{p}_m) = \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_n,\mathbf{p}_1,\ldots,\mathbf{p}_m \in \mathbb{Z}^3} c_{\mathbf{k}_1\ldots\mathbf{k}_n;\mathbf{p}_1\ldots\mathbf{p}_m} \mathbf{t}_1^{\mathbf{k}_1} \ldots \mathbf{t}_n^{\mathbf{k}_n} \mathbf{s}_1^{\mathbf{p}_1} \ldots \mathbf{s}_m^{\mathbf{p}_m}.$$

This function defines the following operator in $\mathcal{H}$:

$$A(F_{nm}) = \sum_{\mathbf{k}_i, \mathbf{p}_j} c_{\mathbf{k}_1\ldots\mathbf{k}_n;\mathbf{p}_1\ldots\mathbf{p}_m} a(\mathbf{k}_1)^* \ldots a(\mathbf{k}_n)^* a(\mathbf{p}_1) \ldots a(\mathbf{p}_m).$$

More generally, we can define the generalized operator valued $(m+n)$-multilinear function

$$A(F_{nm})(f_1, \ldots, f_n, g_1, \ldots, g_m) = \sum_{\mathbf{k}_i, \mathbf{p}_j} c_{\mathbf{k}_1\ldots\mathbf{k}_n;\mathbf{p}_1\ldots\mathbf{p}_m} a(f_1)^* \ldots a(f_n)^* a(g_1) \ldots a(g_m).$$

Here we have to assume that the series converges on a dense subset of $\mathcal{H}$. For example, the Hamiltonian operator $H = A(F_{11})$, where

$$F_{11} = \sum_{\mathbf{k}} E_{\mathbf{k}} \mathbf{t}^{\mathbf{k}} \mathbf{s}^{\mathbf{k}}.$$

Now we define for any $F = \{F_{nm}\} \in \mathcal{H}\hat{\otimes}\mathcal{H}$

$$A(F) = \sum_{m,n=0}^{\infty} A(F_{nm}),$$

where the sum must be convergent in the operator topology.

Now we attempt to assign an operator to a functional $P(\phi, \Pi)$. First, let us assume that P is a polynomial functional $\sum_{ij} a_{ij} \phi^i \Pi^j$. Then we plug in the expressions (6), and transform the expression using the normal ordering of the operators a^*, a. Similarly, we can try to define the value of any analytic function in ϕ, Π, or analytic differential operator

$$P(\partial_\mu \phi, \partial_\mu \Pi) = \sum P_{\mathbf{i},\mathbf{j}}(\phi, \Pi) \partial^{\mathbf{i}}(\phi) \partial^{\mathbf{j}}(\Pi).$$

We skip the discussion of the difficulties related to the convergence of the corresponding operators.

17.5 Let us now discuss how to quantize the Dirac equation. The field in this case is a vector function $\psi = (\psi_0, \dots, \psi_3) : \mathbb{R}^4 \to \mathbb{C}^4$. The Lagrangian is taken in such a way that the corresponding Euler-Lagrange equation coincides with the Dirac equation. It is easy to verify that

$$\mathcal{L} = \psi_\beta^*(i(\gamma^\mu)_{\beta\alpha}\partial_\mu - m\delta_{\beta\alpha})\psi_\alpha, \tag{17.33}$$

where $(\gamma^\mu)_{\beta\alpha}$ denotes the $\beta\alpha$-entry of the Dirac matrix γ^μ and ψ^* is defined in (16.11). The momentum field is $\pi = (\pi_0, \dots, \pi_3)$, where

$$\pi_\alpha = \frac{\delta \mathcal{L}}{\delta \dot{\psi}_\alpha} = i\psi_\beta^*(\gamma^0)_{\beta\alpha} = i\psi_\alpha^*.$$

The Hamiltonian is

$$H = \int_{\mathbb{R}^3} (\pi\dot{\psi} - \mathcal{L})d^3x = \int_{\mathbb{R}^3} (\psi^*(-i\sum_{\mu=1}^{3}\gamma^\mu\partial_\mu + m\gamma^0)\psi)d^3x. \tag{17.34}$$

For any $k = (k_0, \mathbf{k}) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^3$ consider the operator

$$A(k) = \sum_{i=0}^{3}\gamma^i k_i = \begin{pmatrix} 0 & 0 & -k_0 - k_1 & -k_3 + ik_2 \\ 0 & 0 & -k_2 i - k_3 & -k_0 + k_1 \\ k_0 + k_1 & k_3 - ik_2 & 0 & 0 \\ k_2 i + k_3 & k_0 - k_1 & 0 & 0 \end{pmatrix} \tag{17.35}$$

in the spinor space $\mathbb{C}^4$. Its characteristic polynomial is equal to $\lambda^4 - 2|k|^2 + |k|^4 = (\lambda^2 - |k|^2)^2$, where

$$|k|^2 = k_0^2 - k_1^2 - k_2^2 - k_3^2.$$

We assume that

$$|k| = m. \tag{17.36}$$

Hence we have two real eigenvalues of $A(k)$ equal to $\pm m$. We denote the corresponding eigensubspaces by $V_\pm(k)$. They are of dimension 2. Let

$$(\vec{u}_\pm(\mathbf{k}), \vec{v}_\pm(\mathbf{k}))$$

denote an orthogonal basis of $V_\pm(\mathbf{k})$ normalized by

$$\vec{\bar{u}}_\pm(\mathbf{k}) \cdot \vec{u}_\mp(\mathbf{k}) = -\vec{\bar{v}}_\pm(\mathbf{k}) \cdot \vec{v}_\mp(\mathbf{k}) = 2m,$$

where the dot-product is taken in the Lorentz sense. We take indices $\mathbf{k}$ in $\mathbb{R}^3$ since the first coordinate $k_0 \geq 0$ in the vector $k = (k_0, \mathbf{k})$ is determined by $\mathbf{k}$ in view of the relation $|k| = m$. Now we can solve the Dirac equation, using the Fourier integral expansions

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\sum_{\pm}\int_{\mathbb{R}^3}(b_\pm(\mathbf{k})\vec{u}_\pm(\mathbf{k})e^{-k\cdot x} + d_\pm(\mathbf{k})^*\vec{v}_\pm(\mathbf{k})e^{ik\cdot x})d^3x, \tag{17.37a}$$

$$\pi(\mathbf{x}) = \frac{i}{(2\pi)^{3/2}} \sum_{\pm} \int_{\mathbb{R}^3} (b_\pm(\mathbf{k})^* \vec{u}_\pm(\mathbf{k}) e^{ik\cdot x} + d_\pm(\mathbf{k}) \vec{v}_\pm(\mathbf{k}) e^{-ik\cdot x}) d^3x. \tag{17.37b}$$

where

$$E_{\mathbf{k}} = k_0$$

satisfies (17.10). Here $x = (t, \mathbf{x}) = (x_0, x_1, x_2, x_3)$ and the dot product is taken in the sense of the Lorentzian metric. To quantize ψ, π, we replace the coefficients in this expansion with operator-valued functions $\hat{b}_\pm(\mathbf{k}), \hat{b}_\pm(\mathbf{k})^*, \hat{d}_\pm(\mathbf{k}), \hat{d}_\pm(\mathbf{k})^*$ to obtain

$$\Psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{\pm} \int_{\mathbb{R}^3} (\hat{b}_\pm(\mathbf{k}) \vec{u}_\pm(\mathbf{k}) e^{-k\cdot x} + \hat{d}_\pm(\mathbf{k})^* \vec{v}_\pm(\mathbf{k}) e^{ik\cdot x}) d^3x, \tag{17.38a}$$

$$\Pi(\mathbf{x}) = \frac{i}{(2\pi)^{3/2}} \sum_{\pm} \int_{\mathbb{R}^3} (\hat{b}_\pm(\mathbf{k})^* \vec{u}_\pm(\mathbf{k}) e^{ik\cdot x} + \hat{d}_\pm(\mathbf{k}) \vec{v}_\pm(\mathbf{k}) e^{-ik\cdot x}) d^3x. \tag{17.38b}$$

The operators $\hat{b}_\pm(\mathbf{k}), \hat{b}_\pm(\mathbf{k})^*, \hat{d}_\pm(\mathbf{k}), \hat{d}_\pm(\mathbf{k})^*$ are the analogs of the annihilation and creation operators. They satisfy the anticommutator relations:

$$\{\hat{b}_\pm(\mathbf{k}), \hat{b}_\pm(\mathbf{k}')^*\} = \{\hat{d}_\pm(\mathbf{k}), \hat{d}_\pm(\mathbf{k}')^*\} = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\pm\pm}, \tag{17.39}$$

while other anticommutators among them vanish. These anticommutator relations guarantee that

$$[\Psi(t, \mathbf{x})_\alpha, \Pi(t, \mathbf{x})_\beta] = i\delta(\mathbf{x} - \mathbf{y})\delta_{\alpha\beta}.$$

This is analogous to (17.6) in the case of vector field. The Hamiltonian operator is obtained from (17.34) by replacing ψ and π with Ψ, Π given by (17.38). We obtain

$$H = \sum_{\pm} \int_{\mathbb{R}^3} E_{\mathbf{k}} (\hat{b}_\pm(\mathbf{k})^* \hat{b}_\pm(\mathbf{k}) - \hat{d}_\pm(\mathbf{k}) \hat{d}_\pm(\mathbf{k})^*) d^3x. \tag{17.40}$$

Let us formally introduce the vacuum vector $|0\rangle$ with the property

$$\hat{b}_\pm(\mathbf{k})|0\rangle = \hat{d}_\pm(\mathbf{k})|0\rangle = 0$$

and define a state of n particles and m anti-particles by setting

$$\hat{b}(\mathbf{k}_n)^* \ldots \hat{b}(\mathbf{k}_1)^* \hat{d}(\mathbf{p}_m)^* \ldots \hat{d}(\mathbf{p}_1)^* |0\rangle := |\mathbf{k}_n \ldots \mathbf{k}_1 \mathbf{p}'_m \ldots \mathbf{p}'_1\rangle. \tag{17.41}$$

Then we see that the energy of the state $|\mathbf{p}'_m \ldots \mathbf{p}'_1\rangle$ is equal to

$$H|\mathbf{p}'_m \ldots \mathbf{p}'_1\rangle = -(E_{\mathbf{p}_1} + \ldots + E_{\mathbf{p}_m}) < 0$$

but the energy of the state $|\mathbf{k}_n \ldots \mathbf{k}_1\rangle$ is equal to

$$H|\mathbf{k}_n \ldots \mathbf{k}_1\rangle = E_{\mathbf{k}_1} + \ldots + E_{\mathbf{k}_m} > 0.$$

However, we have a problem. Since

$$\hat{d}_\pm(\mathbf{k})\hat{d}_\pm(\mathbf{k})^* = 1 - \hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k}),$$

$$H = \sum_\pm \int_{\mathbb{R}^3} E_\mathbf{k}(\hat{b}_\pm(\mathbf{k})^*\hat{b}_\pm(\mathbf{k}) + (\hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k})) - 1)d^3x.$$

This shows that $|0\rangle$ is an eigenvector of H with eigenvalue

$$-2\int_{\mathbb{R}^3} E_\mathbf{k} = -\infty.$$

To solve this contradiction we "subtract the negative infinity" from the Hamiltonian by replacing it with the normal ordered Hamiltonian

$$: H := \sum_\pm \int_{\mathbb{R}^3} E_\mathbf{k}(\hat{b}_\pm(\mathbf{k})^*\hat{b}_\pm(\mathbf{k}) + \hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k}))d^3x, \tag{17.42}$$

Similarly we introduce the momentum operators and the number operators

$$: \mathbf{P} := \sum_\pm \int_{\mathbb{R}^3} \mathbf{k}(\hat{b}_\pm(\mathbf{k})^*\hat{b}_\pm(\mathbf{k}) + \hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k}))d^3x. \tag{17.43}$$

$$: N := \sum_\pm \int_{\mathbb{R}^3} (\hat{b}_\pm(\mathbf{k})^*\hat{b}_\pm(\mathbf{k}) + \hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k}))d^3x, \tag{17.44}$$

There is also the *charge operator*

$$: Q := e\sum_\pm \int_{\mathbb{R}^3} (\hat{b}_\pm(\mathbf{k})^*\hat{b}_\pm(\mathbf{k}) - \hat{d}_\pm(\mathbf{k})^*\hat{d}_\pm(\mathbf{k}))d^3x. \tag{17.45}$$

It is obtained from normal ordering of the operator

$$Q = \int_{\mathbb{R}^3} \Psi^*\Psi d^3x = e\sum_{\mathbf{k},\pm}(b_\pm(\mathbf{k})^*b_\pm(\mathbf{k}) + d_\pm(\mathbf{k})^*d_\pm(\mathbf{k})).$$

The operators $: N :, : \mathbf{P} :, : Q :$ commute with $: H :$, i.e., they represent the quantized conserved currents.

We conclude:

1) The operator $\hat{d}^*_\pm(\mathbf{k})$ (resp. $\hat{d}_\pm(\mathbf{k})$) creates (resp. annihilates) a positive-energy particle (electron) with *helicity* $\pm\frac{1}{2}$, momentum $(E_\mathbf{k}, \mathbf{k})$ and negative electric charge $-e$.

2) The operator $b^*_\pm(\mathbf{k})$ (resp. $b_\pm(\mathbf{k})$) creates (resp. annihilates) a positive-energy *anti-particle* (*positron*) with helicity $\pm\frac{1}{2}$, momentum $(E_\mathbf{k}, \mathbf{k})$ and charge $-e$.

Note that because of the anti-commutation relations we have

$$b_\pm(\mathbf{k})b_\pm(\mathbf{k}) = d_\pm(\mathbf{k})d_\pm(\mathbf{k}) = 0.$$

This shows that in the package $|\mathbf{k}_n \dots \mathbf{k}_1 \mathbf{p}'_m \dots \mathbf{p}'_1\rangle$ all the vectors $\mathbf{k}_i, \mathbf{p}_j$ must be distinct. This is explained by saying that the particles satisfy the *Pauli exclusion principle*: no two particles with the same helicity and momentum can appear in same state.

17.6 Finally let us comment about the Poincaré invariance of canonical quantization. Using (17.35), we can rewrite the expression (17.38) for $\Psi(t,\mathbf{x})$ as

$$\Psi(t,\mathbf{x}) = \int\limits_{\mathbb{R}_{\geq 0}}\int\limits_{\mathbb{R}^3} (2(2\pi)^3 k_0)^{-1/2}\theta(k_0)\delta(|k|^2 - k_0^2 + m^2)\times$$

$$\times(e^{-i(k_0,\mathbf{k})\cdot(t,\mathbf{x})}a(\mathbf{k}) + e^{i(k_0,\mathbf{k})\cdot(t,\mathbf{x})}a(\mathbf{k})^*)d^3k dk_0. \tag{17.46}$$

Recall that the Poincaré group is the semi-direct product $SO(3,1)\tilde{\times}\mathbb{R}^4$ of the Lorentz group $SO(3,1)$ and the group $\mathbb{R}^4$, where $SO(3,1)$ acts naturally on $\mathbb{R}^4$. We have

$$\Psi(g(t,\mathbf{x})) = \int (2(2\pi)^3 k_0)^{-1/2}\theta(k_0)\delta(|k|^2 - k_0^2 + m^2)\times$$

$$\times(e^{-i(k_0,\mathbf{k})\cdot g(t,\mathbf{x})}a(\mathbf{k}) + e^{i(k_0,\mathbf{k})\cdot(g\cdot(t,\mathbf{x}))}a(\mathbf{k})^*)d^3k dk_0.$$

Now notice that if we take g to be the translation $g : (t,\mathbf{x}) \to (t,\mathbf{x}) + (a_0,\mathbf{a})$, we obtain, with help of formula (17.18),

$$\Psi(g\dot{(}t,\mathbf{x})) = e^{ia_\mu P^\mu}\Psi(t,\mathbf{x})e^{-ia_\mu P^\mu},$$

where $P = (P^\mu)$ is the momentum operator in the Hilbert space $\mathcal{H}$. This shows that the map $(t,\mathbf{x}) \to \Psi(t,\mathbf{x})$ is invariant with respect to the translation group acting on the arguments $(t,\mathbf{x})$ and on the values by means of the representation

$$a = (a_0,\mathbf{a}) \to (A \to e^{i(a_\mu P^\mu)} \circ A \circ e^{-i(a_\mu P^\mu)})$$

of $\mathbb{R}^4$ in the space of operators $\mathrm{End}(\mathcal{H})$.

To exhibit the Lorentzian invariance we argue as follows.

First, we view the operator functions $a(\mathbf{k}), a(\mathbf{k})^*$ as operator valued distributions on $\mathbb{R}^4$ by assigning to each test function $f \in C_0^\infty(\mathbb{R}^4)$ the value

$$a(f) = \int f(k_0,\mathbf{k})\theta(k_0)\delta(|k|^2 - k_0^2 + m^2)a(\mathbf{k})dk_0 d^3k.$$

Now we define the representation of $SO(3,1)$ in $\mathcal{H}$ by introducing the following operators

$$M_{0j} = i\int a(\mathbf{k})^*(E_k\frac{\partial}{\partial k_j})a(\mathbf{k})d^3k, \quad j = 0,1,2,3,$$

$$M_{ij} = i\int a(\mathbf{k})^*(k_i\frac{\partial}{\partial k_j} - k_j\frac{\partial}{\partial k_i})a(\mathbf{k})d^3k, \quad i,j = 1,2,3.$$

Here the partial derivative of the operator $a(\mathbf{k})$ is taken in the sense of distributions

$$\frac{\partial}{\partial k_j}a(\mathbf{k})(f) = \int \frac{\partial f}{\partial k_j}a(\mathbf{k})d^3x.$$

We have

$$[\Psi, M^{\mu\nu}] = i(x_\mu\partial^\nu - x_\nu\partial^\mu)\Psi, \tag{17.47}$$

where $M^{\mu\nu} = g^{\mu\nu}M_{\mu\nu}, \quad x_0 = t.$

Now we can construct the representation of the Lie algebra $\mathfrak{so}(3,1)$ of the Lorentz group in $\mathcal{H}$ by assigning the operator $M^{\mu\nu}$ to the matrix

$$E^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu] = (g^{\mu\nu}I_4 - \gamma^\nu\gamma^\mu).$$

It is easy to see that these matrices generate the Lie algebra of $SO(3,1)$. Now we exponentiate it to define for any matrix $\Lambda = c_{\mu\nu}E^{\mu\nu} \in \mathfrak{so}(3,1)$ the one-parameter group of operators in $\mathcal{H}$

$$U(\Lambda\tau) = \exp(i(c_{\mu\nu}M^{\mu\nu})\tau).$$

This group will act by conjugation in $\mathrm{End}(\mathcal{H})$

$$A \to U(\Lambda\tau)\circ A\circ U(-\Lambda\tau).$$

Using formula (17.47) we check that for any $g = \exp(i\Lambda\tau) \in SO(3,1)$,

$$\Psi(g(t,\mathbf{x})) = U(\Lambda\tau)\Psi(t,\mathbf{x})U(-\Lambda\tau).$$

this shows that $(t,\mathbf{x}) \to \Psi(t,\mathbf{x})$ is invariant with respect to the Poincaré group.

In the case of the spinor fields the Poincaré group acts on Ψ by acting on the arguments $(t,\mathbf{x})$ via its natural representation, and also on the values of Ψ by means of the spinor representation σ of the Lorentz group. Again, one checks that the fields $\Psi(t,\mathbf{x})$ are invariant.

What we described in this lecture is a quantization of *free fields* of spin 0 (pi-mesons) and spin $\frac{1}{2}$ (electron-positron). There is a similar theory of quantization of other free fields. For example, spin 1 fields correspond to the Maxwell equations (photons). In general, spin $n/2$-fields corresponds to irreducible linear representations of the group $Spin(3,1) \cong SL(2,\mathbb{C})$. As is well-known, they are isomorphic to symmetric powers $S^n(\mathbb{C}^2)$, where $\mathbb{C}^2$ is the space of the standard representation of $SL(2,\mathbb{C})$. So, spin $n/2$-fields correspond to the representation $S^n(\mathbb{C}^2), n = 0,1,2,\ldots$.

A non-free field contains some additional contribution to the Lagrangian responsible for interaction of fields. We shall deal with non-free fields in the following lectures.

Exercises.

1. Show that the momentum operators : $\mathbf{P}$: from (17.43) is obtained from the operator $\mathbf{P}$ whose coordinates P_μ are given by

$$P_\mu = -i \int\limits_{\mathbb{R}^3} \Psi(t,\mathbf{x})^* \partial_\mu \Psi(t,\mathbf{x}) d^3x.$$

2. Check that the Hamiltonian and momentum operators are Hermitian operators.

3. Show that the charge operator Q for quantized spinor fields corresponds to the conserved current $J^\mu = \bar{\Psi}\gamma^\mu\Psi$. Check that it is indeed conserved under Lorentizian transformations. Show that $Q = \int J^0 d^3x = \int : \Psi^*\Psi : d^3x$.

4. Verify that the quantization of the spinor fields is Poincaré invariant.

5. Prove that the vacuum vector $|0\rangle$ is invariant with respect to the representation of the Poincaré group in the Fock space given by the operators (P^μ) and $(M^{\mu\nu})$.

6. Consider $\Psi(t,\mathbf{x})$ as an operator-valued distribution on $\mathbb{R}^4$. Show that $[\Psi(f),\Psi(g)] = 0$ if the supports of f and g are spacelike separated. This means that for any $(t,\mathbf{x}) \in \text{Supp}(f)$ and any $(t',\mathbf{x}') \in \text{Supp}(g)$, one has $(t-t')^2 - |\mathbf{x}-\mathbf{x}'|^2 < 0$. This is called *microscopic causality.*

7. Find a discrete version of the quantized Dirac field and an explicit construction of the corresponding Fock space.

8. Explain why the process of quantizing fields is called also second quantization.

Lecture 18. PATH INTEGRALS

The path integral formalism was introduced by R. Feynman. It yields a simple method for quantization of non-free fields, such as gauge fields. It is also used in conformal field theory and string theory.

18.1 The method is based on the following postulate:

The probability $P(a,b)$ of a particle moving from point a to point b is equal to the square of the absolute value of a complex valued function $K(a,b)$:

$$P(a,b) = |K(a,b)|^2. \tag{18.1}$$

The complex valued function $K(a,b)$ is called the *probability amplitude* . Let γ be a path leading from the point a to the point b. Following a suggestion from P. Dirac, Feynman proposed that the amplitude of a particle to be moved along the path γ from the point a to b should be given by the expression $e^{iS(\gamma)/\hbar}$, where

$$S(\gamma) = \int_{t_a}^{t_b} L(\mathbf{x}, \dot{\mathbf{x}}, t)dt$$

is the action functional from classical mechanics. The total amplitude is given by the integral

$$K(a,b) = \int\limits_{\mathcal{P}(a,b)} e^{iS(\gamma)/\hbar} D[\gamma], \tag{18.2}$$

where $\mathcal{P}(a,b)$ denotes the set of all paths from a to b and $D[\gamma]$ is a certain measure on this set. The fact that the absolute value of the amplitude $e^{iS(\gamma)/\hbar}$ is equal to 1 reflects Feynman's principle of the democratic equality of all histories of the particle.

The expression (18.2) for the probability amplitude is called the *Feynman propagator.* It is an analog of the integral

$$F(\lambda) = \int_\Omega f(x)e^{i\lambda S(x)}dx,$$

where $x \in \Omega \subset \mathbb{R}^n, \lambda >> 0$, and $f(x), S(x)$ are smooth real-valued functions. Such integrals are called the integrals of rapidly oscillating functions. If f has compact support and $S(x)$ has no critical points on the support of $f(x)$, then $F(\lambda) = o(\lambda^{-n})$, for any $n > 0$, when $\lambda \to \infty$. So, the main contribution to the integral is given by critical (or stationary) points, i.e. points $x \in \Omega$ where $S'(x) = 0$. This idea is called the method of stationary phase. In our situation, x is replaced with a path γ. When $\hbar$ is sufficiently small, the analog of the method of stationary phase tells us that the main contribution to the probablity $P(a, b)$ is given by the paths with $\frac{\delta S}{\delta \gamma} = 0$. But these are exactly the paths favored by classical mechanics!

Let $\psi(t, x)$ be the wave function in quantum mechanics. Recall that its absolute value is interpreted as the probability for a particle to be found at the point $x \in \mathbb{R}^3$ at time t. Thus, forgetting about the absolute value, we can think that $K((t_b, x_b), (t, x))\psi(t, x)$ is equal to the probability that the particle will be found at the point x_b at time t_b if it was at the point x at time t. This leads to the equation

$$\psi(t_b, x_b) = \int\limits_{\mathbb{R}^3} K((t_b, x_b), (t, x))\psi(t, x)d^3x.$$

From this we easily deduce the following important property of the amplitude function

$$K(a, c) = \int\limits_{t_a}^{t_b}\int\limits_{\mathbb{R}^3} K((t_c, x_c), (t, x))K((t, x), (t_a, x_a))d^3x dt. \tag{18.3}$$

18.2 Before we start computing the path integral (18.2) let us try to explain its meaning. The space $\mathcal{P}(a, b)$ is of course infinite-dimensional and the integration over such a space has to be defined. Let us first restrict ourselves to some special finite-dimensional subspaces of $\mathcal{P}(a, b)$. Fix a positive integer N and subdivide the time interval $[t_a, t_b]$ into N equal parts by inserting intermediate points $t_1 = t_a, t_2, \ldots, t_N, t_{N+1} = t_b$. Choose some points $x_1 = x_a, x_2, \ldots, x_N, x_{N+1} = x_b$ in $\mathbb{R}^n$ and consider the path $\gamma : [t_a, t_b] \to \mathbb{R}^n$ such that its restriction to each interval $[t_i, t_{i+1}]$ is the linear function

$$\gamma_i(t) = \gamma_i(t) = x_i + \frac{x_{i+1} - x_i}{t_{i+1} - t_i}(t - t_i).$$

It is clear that the set of such paths is bijective with $(\mathbb{R}^n)^{N-1}$ and so can be integrated over. Now if we have a function $F : \mathcal{P}(a, b) \to \mathbb{R}$ we can integrate it over this space to get the number J_N. Now we can define (18.2) as the limit of integrals J_N when N goes to infinity. However, this limit may not exist. One of the reasons could be that J_N contains a factor C^N for some constant C with $|C| > 1$. Then we can get the limit by redefining J_N, replacing it with $C^{-N}J_N$. This really means that we redefine the standard measure on $(\mathbb{R}^n)^{N-1}$ replacing the measure d^nx on $\mathbb{R}^n$ by $C^{-1}d^nx$. This is exactly what we are going to do. Also, when we restrict the functional to the finite-dimensional space $(\mathbb{R}^n)^{N-1}$ of piecewise linear paths, we shall allow ourselves to replace the integral $\int_{t_a}^{t_b} S(\gamma)dt$ by its

Riemann sum. The result of this approximation is by definition the right-hand side in (18.2). We should immediately warn that the described method of evaluating of the path integrals is not the only possible. It is only meanigful when we specify the approximation method for its computation.

Let us compute something simple. Assume that the action is given by

$$S = \int_{t_a}^{t_b} \frac{1}{2} m\dot{x}^2 dt.$$

As we have explained before we subdivide $[t_a, t_b]$ into N parts with $\epsilon = (t_b - t_a)/N$ and put

$$K(a,b) = \lim_{N\to\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp[\frac{im}{2\epsilon}\sum_{i=1}^{N}(x_i - x_{i+1})^2] C^{-N} dx_2 \dots dx_N. \tag{18.4}$$

Here the number C should be chosen to guarantee convergence in (18.4). To compute it we use the formula

$$\int_{-\infty}^{\infty} x^{2s-1} e^{-zx^2} dx = \frac{\Gamma(s)}{z^s}. \tag{18.5}$$

This integral is convergent when $\mathrm{Re}(z) > 0, \mathrm{Re}(s) > 0$, and is an analytic function of z when s is fixed (see, for example, [**Whittaker**], 12.2, Example 2). Taking $s = \frac{1}{2}$, we get, if $\mathrm{Re}(z) > 0$,

$$\int_{-\infty}^{\infty} e^{-zx^2} dx = \frac{\Gamma(1/2)}{z^{1/2}} = \sqrt{\pi/z}. \tag{18.6}$$

Using this formula we can define the integral for any $z \neq 0$. We have

$$\int_{-\infty}^{\infty} \exp[-a(x_{i-1} - x_i)^2 - a(x_i - x_{i+1})^2] dx_i =$$

$$= \int_{-\infty}^{\infty} \exp[-2a(x_i + \frac{x_{i-1} + x_{i+1}}{2})^2 - \frac{a}{2}(x_{i-1} - x_{i+1})^2] dx_i =$$

$$= \exp[-\frac{a}{2}(x_{i-1} - x_{i+1})^2] \int_{-\infty}^{\infty} \exp(-2ax^2) dx = \sqrt{\frac{\pi}{2a}} \exp[-\frac{a}{2}(x_{i-1} - x_{i+1})^2].$$

After repeated integrations, we find

$$\int_{-\infty}^{\infty} \exp[-a\sum_{i=1}^{N}(x_i - x_{i+1})^2] dx_2 \dots dx_N = \sqrt{\frac{\pi^{N-1}}{N a^{N-1}}} \exp[-\frac{a}{N}(x_1 - x_{N+1})^2],$$

where $a = m/2i\epsilon$. If we choose the constant C equal to

$$C = \left(\frac{m}{2\pi i\epsilon}\right)^{\frac{1}{2}},$$

then we can rewrite (18.4) in the form

$$K(a,b) = \left(\frac{m}{2\pi iN\epsilon}\right)^{\frac{1}{2}} e^{\frac{mi(x_b-x_a)^2}{2N\epsilon}} = \left(\frac{m}{2\pi i(t_b-t_a)}\right)^{\frac{1}{2}} e^{\frac{mi(x_b-x_a)^2}{2(t_b-t_a)}}, \tag{18.7}$$

where $a = (t_a, x_a), b = (t_b, x_b)$.

The function $b = (t,x) \to K(a,b) = K(a;t,x)$ satisfies the *heat equation*

$$\frac{1}{2m}\frac{\partial^2}{\partial x^2}K(b,a) = i\frac{\partial}{\partial t}K(a,b), \quad b \neq a.$$

This is the Schrödinger equation (5.6) from Lecture 5 corresponding to the harmonic oscillator with zero potential energy.

18.3 The propagator function $K(a,b)$ which we found in (18.7) is the Green function for the Schrödinger equation. The *Green function* of this equation is the (generalized) function $G(x,y)$ which solves the equation

$$(i\frac{\partial}{\partial t} - \frac{1}{2m}\sum_\mu \frac{\partial^2}{\partial x_\mu^2})G(x,y) = \delta(x-y). \tag{18.8}$$

Here we view x as an argument and y as a parameter. Notice that the homogeneous equation

$$(i\frac{\partial}{\partial t} - \frac{1}{2m}\sum_\mu \frac{\partial^2}{\partial x_\mu^2})\phi = 0$$

has no non-trivial solutions in the space of test functions $C_0^\infty(\mathbb{R}^3)$ so it will not have solutions in the space of distributions. This implies that a solution of (18.8) will be unique. Let us consider a more general operator

$$D = \frac{\partial}{\partial t} - P(i\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}), \tag{18.9}$$

where P is a polynomial in n variables. Assume we find a function $F(x)$ defined for $t \geq 0$ such that

$$DF(t,\mathbf{x}) = 0, \quad F(0,\mathbf{x}) = \delta(\mathbf{x}).$$

Then the function

$$F'(x) = \theta(t)F(x)$$

(where $\theta(t)$ is the Heaviside function) satisfies

$$DF'(x) = \delta(x). \tag{18.10}$$

Taking the Fourier transform of both sides in (18.10), it suffices to show that

$$(-ik_0 - P(\mathbf{k}))\hat{F}'(k_0,\mathbf{k}) = (1/2\pi)^{n+1}. \tag{18.11}$$

Let us first take the Fourier transform of F' in coordinates x to get the function $V(t,k)$ satisfying

$$V(t,\mathbf{k}) = 0 \quad \text{if} \quad t < 0,$$

$$V(t,\mathbf{k}) = (1/\sqrt{2\pi})^n \quad \text{if} \quad t = 0,$$

$$\frac{dV(t,\mathbf{k})}{dt} - P(\mathbf{k})V(t,\mathbf{k}) = 0 \quad \text{if} \quad t > 0.$$

Then, the last equality can be extended to the following equality of distributions true for all t:

$$\frac{dV(t,\mathbf{k})}{dt} - P(\mathbf{k})V(t,\mathbf{k}) - (1/\sqrt{2\pi})^n\delta(t) = 0.$$

(see Exercise 5 in Lecture 6). Applying the Fourier transform in t, we get (18.11).

We will be looking for $G(x,y)$ of the form $G(x-y)$, where $G(z)$ is a distribution on $\mathbb{R}^4$. According to the above, first we want to find the solution of the equation

$$(i\frac{\partial}{\partial t} - \frac{1}{2m}\sum_\mu \frac{\partial^2}{\partial x_\mu^2})G'(x-y) = 0, \quad t \geq t', \quad x_i \geq y_i \tag{18.12}$$

satisfying the boundary condition

$$G(0,\mathbf{x}-\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y}). \tag{18.13}$$

Consider the function

$$G'(x-y) = i\Big(\frac{m}{2\pi i(t-t')}\Big)^{3/2} e^{im||\mathbf{x}-\mathbf{y}||^2/2(t-t')}.$$

Differentiating, we find that for $x > y$ this function satisfies (18.11). Since we are considering distributions we can ignore subsets of measure zero. So, as a distribution, this function is a solution for all $x \geq y$. Now we use that

$$\lim_{t\to 0} \frac{1}{2\sqrt{\pi t}} e^{-x^2/4t} = \delta(x)$$

(see **[Gelfand]**, Chapter 1, §2, n°2, Example 2). Its generalization to functions in 3 variables is

$$\lim_{t\to 0} \Big(\frac{1}{2\sqrt{\pi t}}\Big)^3 e^{-||\mathbf{x}||^2/4t} = \delta(\mathbf{x}).$$

After the change of variables, we see that $G'(x-y)$ satisfies the boundary condition (18.13). Thus, we obtain that the function

$$G(x-y) = i\theta(t-t')e^{\frac{im|\mathbf{x}-\mathbf{y}|^2}{2(t-t')}}\Big(\frac{m}{2\pi i(t-t')}\Big)^{3/2} \tag{18.14}$$

is the Green function of the equation (18.8). We see that it is the three-dimensional analog of $K(a,b)$ which we computed earlier.

The advantage of the Green function is that it allows one to find a solution of the inhomogeneous Schrödinger equation

$$S\phi(x) = (i\frac{\partial}{\partial t} - \frac{1}{2m}\sum_{\mu}\frac{\partial}{\partial x_{\mu}}^2)\phi(x) = J(x).$$

If we consider $G(x-y)$ as a distribution, then its value on the left-hand side is equal to the value of $\delta(x-y)$ on the right-hand side. Thus we get

$$\int G(x-y)J(y)d^4y = \int G(x-y)S(\phi(y))d^4y =$$

$$= \int S(G(x-y))\phi(y)d^4y = \int \delta(x-y)\phi(y)d^4y = \phi(x).$$

Using the formula (18.6) for the Gaussian integral we can rewrite $G(x-y)$ in the form

$$G(x-y) = i\theta(t-t')\int_{\mathbb{R}^3}\phi_p(x)\phi_p^*(y)d^3p,$$

where

$$\phi_p(x) = \frac{1}{(2\pi)^{3/2}}e^{-ik\cdot x}, \tag{18.15}$$

where $x = (t,\mathbf{x}), k = (|\mathbf{p}|^2/2m,\mathbf{p})$. We leave it to the reader.

Similarly, we can treat the Klein-Gordon inhomogeneous equation

$$(\partial_\mu\partial^\mu + m^2)\phi(x) = J(x).$$

We find that the Green function $\Delta(x-y)$ is equal to

$$\Delta(x-y) = i\theta(t-t')\int\frac{\phi_k(x)\phi_k^*(y)}{2E_k}d^3k + i\theta(t'-t)\int\frac{\phi_k^*(x)\phi_k(y)}{2E_k}d^3k, \tag{18.16}$$

where $k = (E_k,\mathbf{k}), |k|^2 = m^2$, and ϕ_k are defined by (18.15).

Remark. There is also an expression for $\Delta(x-y)$ in terms of its Fourier transform. Let us find it. First notice that

$$\theta(t-t') = \int_{-\infty}^{\infty}\frac{e^{-ix(t-t')}}{2\pi(x+i\epsilon)}dx.$$

for some small positive ϵ. To see this we extend the contour from the line $\mathrm{Im}\ z = \epsilon$ into the complex z plane by integrating along a semi-circle in the lower-half plane when $t > t'$. There will be one pole inside, namely $z = -i/\epsilon$ with the residue 1. By applying the Cauchy residue theorem, and noticing that the integral over the arc of the semi-circle goes to zero when the radius goes to infinity, we get

$$\int_{-\infty}^{\infty} \frac{e^{-ix(t-t')}}{2\pi(x+i\epsilon)}dx = -i.$$

The minus sign is explained by the clockwise orientation of the contour. If $t < t'$ we can take the contour to be the upper half-circle. Then the residue theorem tells us that the integral is equal to zero (since the pole is in the lower half-plane). Inserting this expression in (18.6), we get

$$\Delta(x-y) = -\frac{1}{(2\pi)^4}\int \frac{e^{-i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})+E_k(t-t')+s(t-t'))}d^3kds}{2E_k(s+i\epsilon)}+$$

$$+\frac{1}{(2\pi)^4}\int \frac{e^{i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})+E_k(t-t')-s(t-t'))}d^3kds}{2E_k(s-i\epsilon)}.$$

Change now s to $k_0 - E_k$ (resp. $k_0 + E_k$) in the first integral (resp. the second one). Also replace $\mathbf{k}$ with $-\mathbf{k}$ in the second integral. Then we can write

$$\Delta(x-y) = -\frac{1}{(2\pi)^4}\int \frac{e^{-i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})+k_0(t-t'))}d^3kdk_0}{2E_k(k_0-E_k+i\epsilon)} - \frac{1}{(2\pi)^4}\int \frac{e^{-i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})+k_0(t-t'))}d^3kdk_0}{2E_k(k_0+E_k-i\epsilon)}$$

$$= -\frac{1}{(2\pi)^4}\int \frac{e^{-i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})+k_0(t-t'))}d^3kdk_0}{k_0^2-E_k^2+i\epsilon} = -\frac{1}{(2\pi)^4}\int \frac{e^{-i(k\cdot(x-y)+k_0(t-t'))}d^3kdk_0}{|k|^2-m^2+i\epsilon},$$

where $k = (k_0, \mathbf{k}), |k| = k_0^2 - ||\mathbf{k}||^2$. Here the expression in the denominator really means $|k|^2 - m^2 + \epsilon^2 + i\epsilon(m^2 + ||\mathbf{k}||^2)^{1/2}$. Thus we can formally write the following formula for $\Delta(x-y)$ in terms of the Fourier transform:

$$-\Delta(x-y) = (2\pi)^2 F(1/(|k|^2 - m^2 + i\epsilon)). \tag{18.17}$$

Note that if we apply the Fourier transform to both sides of the equation

$$(\partial_\mu\partial^\mu + m^2)G(x-y) = \delta(x-y)$$

which defines the Green function, we obtain

$$(-k_0^2 + k_1^2 + k_2^2 + k_3^2 + m^2)F(G(x-y)) = F(\delta(x-y)) = (2\pi)^2.$$

Applying the inverse Fourier transform, we get

$$-\Delta(x-y) = (2\pi)^2 F(1/(|k|^2 - m^2)).$$

The problem here is that the function $1/(|k|^2 - m^2)$ is not defined on the hypersurface $|k|^2 - m^2 = 0$ in $\mathbb{R}^4$. It should be treated as a distribution on $\mathbb{R}^4$ and to compute its

Fourier transform we replace this function by adding $i\epsilon$. Of course to justify formally this trick, we have to argue as in the proof of the formula (18.14).

18.4 Let us arrive at the path integral starting from the Heisenberg picture of quantum mechanics. In this picture the states ϕ originally do not depend on time. However, in Lecture 5, we defined their time evolution by

$$\phi(x,t) = e^{-iHt}\phi(x).$$

Here we have scaled the Planck constant $\hbar$ to 1. The probability that the state $\psi(x)$ changes to the state $\phi(x)$ in time $t'-t$ is given by the inner product in the Hilbert space

$$\langle \phi(x,t)|\psi(x,t')\rangle := \langle \phi(x)e^{-iHt}, e^{-iHt'}\psi(x)\rangle = \langle \phi(x), e^{-iH(t-t')}\psi(x)\rangle. \tag{18.18}$$

Let $|x\rangle$ denote the eigenstate of the position (vector) operator $Q = (Q_1, Q_2, Q_3) : \psi \to (x_1, x_2, x_3)\psi$ with eigenvalue $x \in \mathbb{R}^3$. By Example 9, Lecture 6, $|x\rangle$ is the Dirac distribution $\delta(x-y) : f(y) \to f(x)$. Then we denote $\langle |x'\rangle, t'||x\rangle, t\rangle$ by $\langle x', t'|x, t\rangle$. It is interpreted as the probability that the particle in position x at time t will be found at position x' at a later time t'. So it is exactly $P((t,x),(t',x'))$ that we studied in section 1. We can write any ϕ as a Fourier integral of eigenvectors

$$\psi(x,t) = \int \langle \psi|x\rangle dx,$$

where following physicist's notation we denote the inner product $\langle \psi, |x\rangle\rangle$ as $\langle \psi|x\rangle$. Thus we can rewrite (18.18) in the form

$$\langle \phi|\psi\rangle = \int\int \langle \phi|x'\rangle\langle x'|e^{\frac{-iH(t'-t)}{\hbar}}|x\rangle\langle x|\psi\rangle dx dx'.$$

We assume for simplicity that H is the Hamiltonian of the harmonic oscillator

$$H = (P^2/2m) + V(Q).$$

Recall that the momentum operator $P = i\frac{d}{dx}$ has eigenstates

$$|p\rangle = e^{-ipx}$$

with eigenvalues p. Thus

$$e^{-iP^2}|p\rangle = e^{-ip^2}|p\rangle.$$

We take $\Delta t = t' - t$ small to be able to write

$$e^{-iH(t'-t)} = e^{-i\frac{P^2}{2m}\Delta t}e^{-iV(Q)\Delta t}[1 + O(\Delta t^2)]$$

(note that $\exp(A+B) \neq \exp(A)\exp(B)$ if $[A,B] \neq 0$ but it is true to first approximation). From this we obtain

$$\langle x',t'|x,t\rangle := \langle \delta(y-x'(t))|\delta(y-x(t))\rangle = e^{-iV(x)\Delta t}\langle x'|e^{-iP^2\Delta t/2m}|x\rangle =$$

$$= e^{iV(x)\Delta t}\int\int \langle x'|p'\rangle\langle p'|e^{-iP^2\Delta t/2m}|p\rangle\langle p|x\rangle dpdp' =$$

$$= \frac{1}{2\pi}e^{iV(x)\Delta t}\int\int e^{-i(p^2\Delta t/2m)}\delta(p'-p)e^{ip'x'}e^{-ipx}dpdp' =$$

$$= \frac{1}{2\pi}e^{-iV(x)\Delta t}\int e^{-ip^2\Delta t/2m}e^{ip(x-x')}dp = \Big(\frac{m}{2\pi i\Delta t}\Big)^{\frac{1}{2}}e^{im(x'-x)^2/2\Delta t}e^{i\Delta t V(x)}.$$

The last expression is obtained by completing the square and applying the formula (18.6) for the Gaussian integral. When $V(x)=0$, we obtain the same result as the one we started with in the beginning. We have

$$\langle x',t'|x,t\rangle = \Big(\frac{m}{2\pi i\Delta t}\Big)^{\frac{1}{2}}\exp(iS),$$

where

$$S(t,t') = \int_t^{t'}[\frac{1}{2}m\dot{x}^2 - V(x)]dt.$$

So far, we assumed that $\Delta t = t'-t$ is infinitesimal. In the general case, we subdivide $[t,t']$ into N subintervals $[t=t_1,t_2]\cup\ldots\cup[t_{N-1},t_N=t']$ and insert corresponding eigenstates $x(t_i)$ to obtain

$$\langle x',t'|x,t\rangle = \int\langle x_1,t_1|\ldots|x_N,t_N\rangle dx_2\ldots dx_{N-1} = \int[Dx]\exp(iS(t,t')), \tag{18.19}$$

where $[Dx]$ means that we have to understand the integral in the sense described in the beginning of section 18.2

We can insert additional functionals in the path integral (18.19). For example, let us choose some moments of time $t < t_1 < \ldots < t_n < t'$ and consider the functional which assigns to a path $\gamma : [t,t'] \to \mathbb{R}$ the number $\gamma(t_i)$. We denote such functional by $x(t_i)$. Then

$$\int[Dx]x(t_1)\ldots x(t_n)\exp(iS(t,t')) = \langle x',t'|Q(t_1)\ldots Q(t_n)|x,t\rangle. \tag{18.20}$$

We leave the proof of (18.20) to the reader (use induction on n and complete sets of eigenstates for the operators $Q(t_i)$). For any function $J(t) \in L_2([t,t'])$ let us denote by S^J the action defined by

$$S^J(\gamma) = \int_t^{t'}(\mathcal{L}(x,\dot{x}) + xJ(t))dt.$$

Let us consider the functional

$$J \to Z[J] = k \int [Dx] e^{iS^J}.$$

Then

$$i \frac{\delta^n}{\delta J(t_1) \dots \delta J(t_n)}|_{J=0} \int [Dx] e^{iS^J} = \langle x', t' | Q(t_1) \dots Q(t_n) | x, t \rangle. \tag{18.21}$$

Here $\frac{\delta}{\delta J(x_i)}$ denotes the (weak) functional derivative of the functional $J \to Z[J]$ computed at the function

$$\alpha_a(x) = \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{if } x \neq a. \end{cases}$$

Recall that the weak derivative of a functional $F : C^\infty(\mathbb{R}^n) \to \mathbb{C}$ at the point $\phi \in C^\infty(\mathbb{R}^n)$ is the functional $DF(\phi) : C^\infty(\mathbb{R}^n) \to \mathbb{C}$ defined by the formula

$$DF(\phi)(h) = \frac{dF(\phi + th)}{dt} = \lim_{t \to 0} \frac{F(\phi + th) - F(\phi)}{t}.$$

If F has Frechét derivative in the sense that we used in Lecture 13, then the weak derivative exists too, and the derivatives coincide.

18.4 Now recall from the previous lecture the formula (17.22a) for the quantized free scalar field

$$\Psi(t, \mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} \frac{1}{\sqrt{2E_k}} (a(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - E_{\mathbf{k}} t)} + a(\mathbf{k})^* e^{-i(\mathbf{k} \cdot \mathbf{x} - E_{\mathbf{k}} t)}) d^3k.$$

We can rewrite it in the form

$$\Psi(t, \mathbf{x}) = \int_{\mathbb{R}^3} \frac{1}{\sqrt{2E_k}} (a(\mathbf{k}) \phi_k(x) + a(\mathbf{k})^* \phi_k^*(x)) d^3k,$$

where ϕ_k are defined in (18.15).

Define the *time-ordering* operator

$$T[\Psi(t, \mathbf{x}) \Psi(t', \mathbf{x}')] = \begin{cases} \Psi(t, \mathbf{x}) \circ \Psi(t', \mathbf{x}') & \text{if } t > t' \\ \Psi(t', \mathbf{x}') \circ \Psi(t, \mathbf{x}) & \text{if } t < t'. \end{cases}$$

Then, using (18.16), we immediately verify that

$$\langle 0 | T[\Psi(t, \mathbf{x}) \Psi(t', \mathbf{x}')] | 0 \rangle = -i\Delta(x - y). \tag{18.22}$$

This gives us another interpretation of the Green function for the Klein-Gordon equation.

For any operator A the expression

$$\langle 0 | A | 0 \rangle$$

is called the *vacuum expectation value* of A. When $A = T[\Psi(t, \mathbf{x}) \Psi(t', \mathbf{x}')]$ as above, we interpret it as the probability amplitude for the system being in the state $\Psi(t, \mathbf{x})$ at time t to change into state $\Psi(t', \mathbf{x}')$ at time t'.

If we write the operators $\Psi(t,\mathbf{x})$ in terms of the creation and annihilation operators, it is clear that the normal ordering $:\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}'):$ differs from the time-ordering $T[\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}')]$ by a scalar operator. To find this scalar, we compare the vacuum expectation values of the two operators. Since the operator in the normal ordering kills $|0\rangle$, the vacuum expectation value of $:\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}'):$ is equal to zero. Thus we obtain

$$T[\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}')] - :\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}'):= \langle 0|T[\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}')]|0\rangle.$$

The left-hand side is called the *Wick contraction* of the operators $\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}')$ and is denoted by

$$\overbrace{\Psi(t,\mathbf{x})\Psi(t',\mathbf{x}')}\,. \tag{18.23}$$

It is a scalar operator.

If we now have three field operators $\Psi(x_1), \Psi(x_2), \Psi(x_3)$, we obtain

$$T[\Psi(x_1),\Psi(x_2),\Psi(x_3)] - :\Psi(x_1),\Psi(x_2),\Psi(x_3):=$$

$$= \frac{1}{2}\sum_{\sigma\in S_3}\langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle\Psi(x_{\sigma(3)}).$$

For any n we have the following Wick's theorem:

Theorem 1.

$$T[\Psi(x_1),\dots,\Psi(x_n)] - :\Psi(x_1)\cdots\Psi(x_n):=$$

$$= \sum_{\text{perm}}\langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle :\Psi(x_{\sigma(3)}\cdots\Psi(x_{\sigma(n)}): +$$

$$+\sum_{\text{perm}}\langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle\langle 0|T[\Psi(x_{\sigma(3)})\Psi(x_{\sigma(4)})]|0\rangle :\Psi(x_{\sigma(5)}\cdots\Psi(x_{\sigma(n)}: +\dots$$

$$+\sum_{\text{perm}}\langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle\cdots\langle 0|T[\Psi(x_{\sigma(n-1)})\Psi(x_{\sigma(n)})]|0\rangle,$$

where, if n is odd, the last term is replaced with

$$\sum_{\text{perm}}\langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle\cdots\langle 0|T[\Psi(x_{\sigma(n-2)})\Psi(x_{\sigma(n-1)})]|0\rangle\Psi(x_{\sigma(n)}).$$

Here we sum over all permutations that lead to different expressions.

Proof. Induction on n, multiplying the expression for $n-1$ on the right by the operator $\Psi(x_n)$ with smallest time argument.

Corollary.

$$\langle 0|T[\Psi(x_1),\ldots,\Psi(x_n)]|0\rangle = \sum_{\text{perm}} \langle 0|T[\Psi(x_{\sigma(1)})\Psi(x_{\sigma(2)})]|0\rangle \cdots \langle 0|T[\Psi(x_{\sigma(n-1)})\Psi(x_{\sigma(n)})]|0\rangle$$

if n is even. Otherwise $\langle 0|T[\Psi(x_1),\cdots,\Psi(x_n)]|0\rangle = 0$.

The assertion of the previous theorem can be easily generalized to any operator valued distribution which we considered in Lecture 17, section 4.

The function

$$G(x_1,\ldots,x_n) = \langle 0|T[\Psi(x_1),\ldots,\Psi(x_n)]|0\rangle$$

is called the *n-point function* (for scalar field). Given any function $J(x)$ we can form the generating function

$$Z[J] = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int J(x_1)\ldots J(x_n)\langle 0|T[\Psi(x_1),\ldots,\Psi(x_n)]|0\rangle d^4x_1\ldots d^4x_n. \tag{18.24}$$

We can write down this expression formally as

$$Z[J] := \langle 0|T[\exp(i\int J(x)\Psi(x)d^4x)]|0\rangle.$$

We have

$$i^n G(x_1,\ldots,x_n) = \frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)} Z[J]|_{J=0}. \tag{18.25}$$

18.5 In section 18.3 we used the path integrals to derive the propagation formula for the states $|\mathbf{x}\rangle$ describing the probability of a particle to be at a point $\mathbf{x}\in\mathbb{R}^3$. We can try to do the same for arbitrary pure states $\phi(\mathbf{x})$. We shall use functional integrals to interpret the propagation formula (18.18)

$$\langle\phi(t,\mathbf{x})|\psi(t',\mathbf{x})\rangle = \langle\phi(\mathbf{x}), e^{-iH(t-t')}\psi(\mathbf{x})\rangle. \tag{18.26}$$

Here we now view $\phi(t,\mathbf{x}),\psi(t,\mathbf{x})$ as arbitrary fields, not necessary coming from evolved wave functions. We shall try to motivate the following expression for this formula:

$$\langle\phi_1(t,\mathbf{x})|\phi_2(t',\mathbf{x})\rangle = \int_{f_1}^{f_2}[D\phi]e^{iS(\phi)}, \tag{18.27}$$

where we integrate over some space of functions (or distributions) ϕ on $\mathbb{R}^4$ such that $\phi(t,\mathbf{x}) = f_1(\mathbf{x}), \phi(t',\mathbf{x}) = f_2(\mathbf{x})$. The integration is *functional integration.*

For example, assume that

$$S(\phi) = \int_t^{t'} \Big(\int_{\mathbb{R}^3} \mathcal{L}(\phi) d^3x \Big) dt,$$

where

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2\phi^2) + J\phi$$

corresponds to the Klein-Gordon Lagrangian with added source term $J\phi$. We can rewrite $S(\phi)$ in the form

$$S(\phi) = \int \mathcal{L}(\phi) d^4x = \int (\frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2 + J\phi) d^4x =$$

$$= \int (-\frac{1}{2}\phi(\partial_\mu \partial^\mu + m^2)\phi + J\phi) d^4x = \int (-\frac{1}{2}\phi A \phi + J\phi) d^4x, \tag{18.28}$$

where

$$A = \partial_\mu \partial^\mu + m^2$$

is the D'Alembertian linear operator. Now the action functional $S(\phi(t, \mathbf{x}))$ looks like a quadratic function in ϕ. The right-hand side of (18.27) recalls the Gaussian integral (18.6). In fact the latter has the following generalization to quadratic functions on any finite-dimensional vector spaces:

Lemma 1. *Let $Q(x) = \frac{1}{2}x \cdot Ax + b \cdot x$ be a complex valued quadratic function on $\mathbb{R}^n$. Assume that the matrix A is symmetric and its imaginary part is positive definite. Then*

$$\int_{\mathbb{R}^n} e^{iQ(x)} d^n x = (2\pi i)^{n/2} \exp[-\frac{1}{2} i b \cdot A^{-1} b](\det A)^{-1/2}.$$

Proof. First make the variable change $x \to x + A^{-1}b$, then make an orthogonal variable change to reduce the quadratic part to the sum of squares, and finally apply (18.6).

As in the case of (18.6) we use the right-hand side to define the integral for any invertible A.

So, as soon as our action $S(\phi)$ looks like a quadratic function in ϕ we can define the functional integral provided we know how to define the determinant of a linear operator on a function space. Recall that the determinant of a symmetric matrix can be computed as the product of its eigenvalues

$$\det(A) = \lambda_1 \ldots \lambda_n.$$

Equivalently,

$$\ln \det(A) = \sum_{i=1}^n \ln(\lambda_i) = -\frac{d(\sum_{i=1}^n \lambda_i^{-s})}{ds}\Big|_{s=0}$$

This suggests to define the determinant of an operator A in a Hilbert space V which admits a complete basis $(v_i), i = 0, 1\ldots$ composed of eigenvectors of A with eigenvalues λ_i as

$$\det(A) = e^{-\zeta'_A(0)}, \tag{18.28}$$

where

$$\zeta_A(s) = \sum_{i=1}^{n} \lambda_i^{-s}. \tag{18.29}$$

The function $\zeta_A(s)$ of the complex variable s is called the *zeta function* of the operator A. In order that it makes sense we have to assume that no λ_i are equal to zero. Otherwise we omit 0 from expression (18.19). Of course this will be true if A is invertible. For example, if A is an elliptic operator of order r on a compact manifold of dimension d, the zeta function ζ_A converges for Re $s > d/r$. It can be analytically continued into a meromorphic function of s holomorphic at $s = 0$. Thus (18.29) can be used to define the determinant of such an operator.

Example. Unfortunately the determinant of the D'Alembertian operator is not defined. Let us try to compute instead the determinant of the operator $A = -\Delta + m^2$ obtained from the D'Alembertian operator by replacing t with it. Here Δ is the 4-dimensional Laplacian. To get a countable complete set of eigenvalues of A we consider the space of functions defined on a box B defined by $-l_i \leq x_i \leq l_i, i = 1, \ldots, 4$. We assume that the functions are integrable complex-valued fucntions and also periodic, i.e. $f(-l) = f(l)$, where $l = (l_1, l_2, l_3, l_4)$. Then the orthonormal eigenvectors of A are the functions

$$f_k(x) = \frac{1}{\Omega} e^{ik\cdot x}, \quad k = 2\pi(n_1/l_1, \ldots, n_4/l_4), \quad n_i \in \mathbb{Z},$$

where Ω is the volume of B. The eigenvalues are

$$\lambda_k = ||k||^2 + m^2 > 0$$

Introduce the *heat function*

$$K(x, y, t) = \sum e^{-\lambda_k t} f_k(x) \bar{f}_k(x).$$

It obeys the heat equation

$$A_x K(x, y, t) = -\frac{\partial}{\partial t} K(x, y, t).$$

Here A_x denotes the application of the operator A to the function $K(x, y, t)$ when y, t are considered constants. Since $K(x, y, 0) = \delta(x - y)$ because of orthonormality of f_k, we obtain that $\theta(t)K(x, y, t)$ is the Green function of the heat equation. In our case we have

$$K(x, y, t) = \theta(t) \frac{1}{16\pi^2 t^2} e^{-m^2 t - (x-y)^2/4t} \tag{18.30}$$

(compare with (18.16)).

Now we use that

$$\zeta_A(s) = \sum_k \lambda_k^{-s} = \sum_{i=1}^{n} \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} e^{-\lambda_k t} dt =$$

$$= \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} (\sum_k e^{-\lambda_k t}) dt = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} Tr[e^{-At}] dt.$$

Since

$$Tr(e^{-At}) = \int_B K(x,x,t)dx,$$

we have

$$Tr(e^{-At}) = \frac{\Omega}{16\pi^2 t^2} e^{-m^2 t}.$$

This gives for Re $s > 4$,

$$\zeta_A(s) = \frac{\Omega}{16\pi^2 \Gamma(s)} \int_0^\infty t^{s-3} e^{-m^2 t} dt = \frac{\Omega (m^2)^{2-s} \Gamma(s-2)}{16\pi^2 \Gamma(s)}.$$

After we analytically continue ζ to a meromorphic function on the whole complex plane, we will be able to differentiate at 0 to obtain

$$\ln \det(-\Delta + m^2) = -\zeta'(0) = \Omega \frac{1}{32\pi^2} m^4 (-\frac{3}{2} + \ln m^2).$$

We see that when the box B increases its size, the determinant goes to infinity.

So, even for the Laplacian operator we don't expect to give meaning to the functional integral (18. 27) in the infinite-dimensional case. However, we located the source of the infinity in the functional integral. If we set

$$N = \int e^{i\phi A \phi} D[\phi] = \det(A)^{-1/2}$$

then

$$N^{-1} \int e^{i\phi A\phi + J\phi} D[\phi] = \exp[-\frac{1}{2} i J(x) G(x,y) J(y)]$$

has a perfectly well-defined meaning. Here $G(x,y)$ is the inverse operator to A represented by its Green function. In the case of the operator $A = \Box + m^2$ it agrees with the propagator formula:

$$\exp(-\frac{i}{2} \int J(x)J(y) \langle 0 | T[\Psi(\mathbf{x})\Psi(\mathbf{x})] | 0 \rangle d^3x d^3y = Z[J] = N \int e^{i\phi A\phi + J\phi} D[\phi] \qquad (18.32)$$

Exercises.

1. Consider the quadratic Lagrangian

$$L(x,\dot{x},t) = a(t(x^2 + b(t)\dot{x}^2 + c(t)x \cdot x + d(t)x + e(t) \cdot x + f(t).$$

Show that the Feynman propagator is given by the formula

$$K(a,b) = A(t_a,t_b)\exp[\frac{i}{\hbar}\int_{t_a}^{t_b} L(x_{cl},\dot{x}_{cl};t)dt],$$

where A is some function, and x_{cl} denotes the classical path minimizing the action.

2. Let $K(a,b)$ be propagator given by formula (18.14). Verify that it satisfies property (18.3).

3. Give a full proof of Wick's theorem.

4. Find the analog of the formula (18.24) for the generating function $Z[J]$ in ordinary quantum mechanics.

5. Show that the zeta function of the Hamiltonian operator for the harmonic oscillator is equal to the Riemann zeta function.

Lecture 19. FEYNMAN DIAGRAMS

19.1 So far we have succeeded in quantizing fields which were defined by Lagrangians for which the Euler-Lagrange equation was linear (like the Klein-Gordon equation, for example). In the non-linear case we no longer can construct the Fock space and hence cannot give a particle interpretation of quantized fields. One of the ideas to solve this difficulty is to try to introduce the quantized fields $\Psi(t,\mathbf{x})$ as obtained from free "input" fields $\Psi_{\rm in}(t,\mathbf{x})$ by a unitary automorphism $U(t)$ of the Hilbert space

$$\Psi(t,\mathbf{x}) = U^{-1}(t)\Psi_{\rm in}U(t). \tag{19.1}$$

We assume that $\Psi_{\rm in}$ satisfies the usual (quantum) Schrödinger equation

$$\frac{\partial}{\partial t}\Psi_{\rm in} = i[H_0,\Psi_{\rm in}], \tag{19.2}$$

where H_0 is the "linear part" of the Hamiltonian. We want Ψ to satisfy

$$\frac{\partial}{\partial t}\Psi = i[H,\Psi],$$

where H is the "full" Hamiltonian. We have

$$\frac{\partial}{\partial t}\Psi_{\rm in} = \frac{\partial}{\partial t}[U(t)\Psi U(t)^{-1}] = \dot U(t)\Psi U^{-1} + U(t)(\frac{\partial}{\partial t}\Psi)U^{-1}(t) + U(t)\Psi\dot U^{-1}(t) =$$

$$= \dot U(U^{-1}\Psi_{\rm in}U)U^{-1} + U[iH(\Psi,\Pi),\Psi]U^{-1} + U(U^{-1}\Psi_{\rm in}U)\dot U^{-1} =$$

$$= \dot U U^{-1}\Psi_{\rm in} + iU[H(\Psi,\Pi),\Psi]U^{-1} + \Psi_{\rm in}U\dot U^{-1}. \tag{19.3}$$

At this point we assume that the Hamiltonian satisfies

$$U(t)H(\Psi,\Pi)U(t)^{-1} = H(U(t)\Psi U(t)^{-1}, U(t)\Pi U(t)^{-1}) = H(\Psi_{\rm in},\Pi_{\rm in}).$$

Using the identity

$$0 = \frac{d}{dt}[U(t)U^{-1}(t)] = [\frac{d}{dt}U(t)]U(t)^{-1} + U(t)\frac{d}{dt}U^{-1}(t),$$

allows us to rewrite (19.3) in the form

$$[\dot{U}U^{-1} + iH(\Psi_{\text{in}}, \pi_{\text{in}}), \Psi_{\text{in}}] = i[H_0, \Psi_{\text{in}}].$$

This implies that

$$[\dot{U}U^{-1} + i(H(\Psi_{\text{in}}, \Pi_{\text{in}}) - H_0), \Psi_{\text{in}}] = 0.$$

Similarly, we get

$$[\dot{U}U^{-1} + i(H(\Psi_{\text{in}}, \Pi_{\text{in}}) - H_0), \Pi_{\text{in}}] = 0.$$

It can be shown that the Fock space is an irreducible representation of the algebra formed by the operators Ψ_{in} and Π_{in}. This implies that

$$\dot{U}U^{-1} + i(H(\Psi_{\text{in}}, \Pi_{\text{in}}) - H_0) = f(t)\mathbf{id}$$

for some function of time t. Note that the Hamiltonian does not depend on $\mathbf{x}$ but depends on t since Ψ_{in} and Π_{in} do. Let us denote the difference $H(\Psi_{\text{in}}, \Pi_{\text{in}}) - H_0 - f(t)\mathbf{id}$ by $H_{\text{int}}(t)$. Then multiplying the previous equation by $U(t)$ on the right, we get

$$i\frac{\partial}{\partial t}U(t) = H_{\text{int}}(t)U(t). \tag{19.4}$$

We shall see later that the correction term $f(t)\mathbf{id}$ can be ignored for all purposes.

Recall that we have two pictures in quantum mechanics. In the Heisenberg picture, the states do not vary in time, and their evolution is defined by $\phi(t) = e^{iHt}\phi(0)$. In the Schrödinger picture, the states depend on time, but operators do not. Their evolution is described by $A(t) = e^{iHt}Ae^{-iHt}$. Here H is the Hamiltonian operator. The equations of motion are

$$[A(t), H] = i\dot{A}(t), \quad \dot{\phi} = 0 \quad \text{(Heisenberg picture)},$$

$$H\phi = i\frac{\partial\phi}{\partial t}, \quad \dot{A} = 0 \quad \text{(Schrödinger picture)}.$$

We can mix them as follows. Let us write the Hamiltonian as a sum of two Hermitian operators

$$H = H_0 + H_{\text{int}}, \tag{19.5}$$

and introduce the equations of motion

$$[A(t), H_0] = \dot{A}(t), \quad H_{\text{int}}(t)\phi = i\frac{\partial\phi}{\partial t}.$$

One can show that this third picture (called the *interaction picture*) does not depend on the decomposition (19.5) and is equivalent to the Heisenberg and Schrödinger pictures. If we denote by Φ_S, Φ_H, Φ_I (resp. A_S, A_H, A_I) the states (resp. operators) in the corresponding pictures, then the transformations between the three different pictures are given as follows

$$\Phi_S = e^{-iHt}\Phi_H, \quad \Phi_I(t) = e^{iH_0t}\Phi_S(t),$$

$$A_S = e^{-iHt}A_H(t)e^{iHt}, \quad A_I = e^{iH_0t}A_S e^{-iH_0t}.$$

Let us consider the family of operators $U(t,t'), t, t' \in \mathbb{R}$, satisfying the following properties

$$U(t_1,t_2)U(t_2,t_3) = U(t_1,t_3),$$

$$U^{-1}(t_1,t_2) = U(t_2,t_1),$$

$$U(t,t) = \mathbf{id}.$$

We assume that our operator $U(t)$ can be expressed as

$$U(t) = U(t,-\infty) := \lim_{t'\to-\infty} U(t,t').$$

Then, for any fixed t_0, we have

$$U(t) = U(t,-\infty) = U(t,t_0)U(t_0,-\infty) = U(t,t_0)U(t_0).$$

Thus equation (19.4) implies

$$(H_{\text{int}}(t) - i\frac{\partial}{\partial t})U(t,t_0) = 0 \tag{19.6}$$

with the initial condition $U(t_0,t_0) = \mathbf{id}$. Then for any state $\phi(t)$ (in the interaction picture) we have

$$U(t,t_0)\phi(t_0) = \phi(t)$$

(provided the uniqueness of solutions of the Schrödinger equation holds). Taking the Hermitian conjugate we get

$$U(t,t_0)^*H_{\text{int}}(t) + i\frac{\partial}{\partial t}U(t,t_0)^* = 0,$$

and hence, multiplying the both sides by $U(t,t_0)$ on the right,

$$U(t,t_0)^*H_{\text{int}}(t)U(t,t_0) + i\frac{\partial}{\partial t}U(t,t_0)^*U(t,t_0) = 0.$$

Multiplying both sides of (19.6) by $U(t,t_0)^*$ on the left, we get

$$U(t,t_0)^*H_{\text{int}}(t)U(t,t_0) + iU(t,t_0)^*\frac{\partial}{\partial t}U(t,t_0) = 0.$$

Comparing the two equations, we obtain

$$\frac{\partial}{\partial t}(U(t,t_0)^*U(t,t_0)) = 0.$$

Since $U(t,t_0)^*U(t,t_0) = \mathbf{id}$, this tells us that the operator $U(t,t_0)$ is unitary. We define the *scattering matrix* by

$$S = \lim_{t_0 \to -\infty, t \to \infty} U(t,t_0). \tag{19.7}$$

Of course, this double limit may not exist. One tries to define the decomposition (19.5) in such a way that this limit exists.

Example 1. Let H be the Hamiltonian of the harmonic oscillator (with $\omega/2$ subtracted). Set

$$H_0 = \omega_0 a^* a, H_{\text{int}} = (\omega - \omega_0)a^*a.$$

Then

$$U(t,t_0) = e^{-i(\omega-\omega_0)(t-t_0)a^*a}.$$

It is clear that the S-matrix S is not defined.

In fact, the only partition which works is the trivial one where $H_{\text{int}}(t) = 0$.

We will be solving for $U(t,t_0)$ by iteration. Replace $H_{\text{int}}(t)$ with $\lambda H_{\text{int}}(t)$ and look for a solution in the form

$$U(t,t_0) = \sum_{n=0}^{\infty} \lambda^n U_n(t,t_0).$$

Replacing $H_{\text{int}}(t)$ with $\lambda H_{\text{int}}(t)$, we obtain

$$i\frac{\partial}{\partial t}U(t,t_0) = \sum_{n=0}^{\infty}\lambda^n\frac{\partial}{\partial t}U_n(t,t_0) = \lambda H_{\text{int}}(t)\sum_{n=0}^{\infty}\lambda^n\frac{\partial}{\partial t}U_n(t,t_0) = \sum_{n=0}^{\infty}\lambda^{n+1}H_{\text{int}}\frac{\partial}{\partial t}U_n(t,t_0).$$

Equating the coefficients at λ^n, we get

$$\frac{\partial}{\partial t}U_0(t,t_0) = 0,$$

$$i\frac{\partial}{\partial t}U_n(t,t_0) = H_{\text{int}}(t)U_{n-1}(t,t_0), \quad n \geq 1.$$

Together with the initial condition $U_0(t_0,t_0) = 1, U_n(t,t_0) = 0, n \geq 1$, we get

$$U_0(t,t_0) = 1, \quad U_1(t,t_0) = -i\int_{t_0}^{t} H_{\text{int}}(\tau)d\tau,$$

$$U_2(t,t_0) = (-i)^2\int_{t_0}^{t} H_{\text{int}}(\tau_2)\Big(\int_{\tau_1}^{\tau_2} H_{\text{int}}(\tau_1)d\tau_1\Big)d\tau_2,$$

$$U_n(t,t_0) = (-i)^n \int_{t_0}^{t} H_{\rm int}(\tau_n)\Big(\int_{t_0}^{\tau_n} \dots \Big(\int_{t_0}^{\tau_3} H_{\rm int}(\tau_2)\Big(\int_{t_0}^{\tau_2} H_{\rm int}(\tau_1)d\tau_1\Big)d\tau_2\Big)\dots\Big)d\tau_n =$$

$$= (-i)^n \int_{t_0\le\tau_1\le\dots\le\tau_n\le t} T[H_{\rm int}(\tau_n)\dots H_{\rm int}(\tau_1)]d\tau_1\dots d\tau_n =$$

$$= \frac{(-i)^n}{n!}\int_{t_0}^{t}\dots\int_{t_0}^{t} T[H_{\rm int}(\tau_n)\dots H_{\rm int}(\tau_1)]d\tau_1\dots d\tau_n.$$

After setting $\lambda = 1$, we get

$$U(t,t_0) = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{t_0}^{t}\dots\int_{t_0}^{t} T[H_{\rm int}(t_n)\dots H_{\rm int}(t_1)]dt_1\dots dt_n. \qquad (19.8)$$

Taking the limits $t_0 \to -\infty, t \to +\infty$, we obtain the expression for the scattering matrix

$$S = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} T[H_{\rm int}(t_n)\dots H_{\rm int}(t_1)]dt_1\dots dt_n. \qquad (19.9)$$

We write the latter expression formally as

$$S = T\exp\Big(-i\int_{-\infty}^{\infty} : H_{\rm int}(t) : dt\Big).$$

Let ϕ be a state, we interpret $S\phi$ as the result of the scattering process. It is convenient to introduce two copies of the same Hilbert space $\mathcal{H}$ of states. We denote the first copy by $\mathcal{H}_{\rm IN}$ and the second copy by $\mathcal{H}_{\rm OUT}$. The operator S is a unitary map

$$S : \mathcal{H}_{\rm IN} \to \mathcal{H}_{\rm OUT}.$$

The elements of the each space will be denoted by $\phi_{\rm IN}, \phi_{\rm OUT}$, respectively. The inner product $P = \langle\alpha_{\rm OUT}, \beta_{\rm IN}\rangle$ is the transition amplitude for the state $\beta_{\rm IN}$ transforming to the state $\alpha_{\rm OUT}$. The number $|P|^2$ is the probability that such a process occurs (recall that the states are always normalized to have norm 1). If we take $(\alpha_i), i \in I$, to be an orthonormal basis of $\mathcal{H}_{\rm IN}$, then the "matrix element" of S

$$S_{ij} = \langle S\alpha_i|\alpha_j\rangle$$

can be viewed as the probability amplitude for α_i to be transformed to α_j (viewed as an element of $\mathcal{H}_{\rm OUT}$).

19.2 Let us consider the so-called Ψ^4-*interaction process.* It is defined by the Hamiltonian

$$H = H_0 + H_{\rm int},$$

where

$$H_0 = \frac{1}{2}\int_{\mathbb{R}^3} :\Pi^2 + (\partial_1\Psi)^2 + (\partial_2\Psi)^2 + (\partial_3\Psi)^2) + m^2\Psi^2 : d^3x,$$

$$H_{\text{int}} = \frac{f}{4!}\int :\Psi^4 : d^3x.$$

Here Ψ denotes Ψ_{in}; it corresponds to the Hamiltonian H_0. Recall from Lecture 17 that

$$\Psi = \frac{1}{(2\pi)^{3/2}}\int (2E_k)^{-1/2}(e^{ik\cdot x}a(\mathbf{k}) + e^{-ik\cdot x}a(\mathbf{k})^*)d^3k,$$

where $k = (E_k, \mathbf{k}), |E_k| = m^2 + |\mathbf{k}|^2$.

We know that

$$H_0 = \int E_{\mathbf{k}}a(\mathbf{k})^*a(\mathbf{k})d^3k.$$

Take the initial state $\phi = |\mathbf{k}_1\mathbf{k}_2\rangle \in \mathcal{H}_{\text{IN}}$ and the final state $\phi' = |\mathbf{k}_1'\mathbf{k}_2'\rangle \in \mathcal{H}_{\text{OUT}}$. We have

$$S = 1 + (-if)/4!\int :\Psi(x)^4 : d^4x + (-if)^2/2(4!)^2\int T[:\Psi(x)^4 :: \Psi(y)^4 :]d^4xd^4y + \dots.$$

The approximation of first order for the amplitude $\langle\mathbf{k}_1\mathbf{k}_2|S|\mathbf{k}_1'\mathbf{k}_2'\rangle$ is

$$\langle\mathbf{k}_1\mathbf{k}_2|S|\mathbf{k}_1'\mathbf{k}_2'\rangle_1 = -i\int\langle\mathbf{k}_1\mathbf{k}_2| :\Psi^4 : |\mathbf{k}_1'\mathbf{k}_2'\rangle d^4k,$$

To compute $:\Psi^4:$ we have to select only terms proportional to $a(\mathbf{k}_1)^*a(\mathbf{k}_2)^*a(\mathbf{k}_1')a(\mathbf{k}_2')$. There are 4! such terms, each comes with the coefficient

$$e^{i(k_1+k_2-k_1'-k_2')\cdot x}/(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}_1'}E_{\mathbf{k}_2'})^{\frac{1}{2}}(2\pi)^6,$$

where $k = (E_{\mathbf{k}}, \mathbf{k})$ is the 4-momentum vector. After integration we get

$$\langle\mathbf{k}_1\mathbf{k}_2|S|\mathbf{k}_1'\mathbf{k}_2'\rangle = \mathcal{M}\delta(k_1 + k_2 - k_1' - k_2'),$$

where

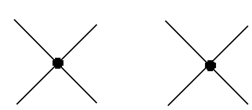
$$\mathcal{M} = -if/4(2\pi)^2(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}_1'}E_{\mathbf{k}_2'})^{\frac{1}{2}}. \tag{19.10}$$

So, up to first order in f, the probability that the particles $\langle\mathbf{k}_1\mathbf{k}_2|$ with 4-momenta $\mathbf{k}_1, \mathbf{k}_2$ will produce the particles $\langle\mathbf{k}_1'\mathbf{k}_2'|$ with momenta $\mathbf{k}_1', \mathbf{k}_2'$ is equal to zero, unless the total momenta are conserved:

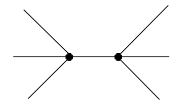
$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_1' + \mathbf{k}_2'.$$

since the energy $E_{\mathbf{k}}$ is determined by $\mathbf{k}$, we obtain that the total energy does not change under the collision.

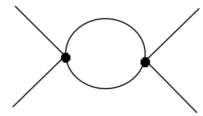
$$T[:\Psi^4(x) :: \Psi^4(y) :] =: \Psi^4(x)\Psi^4(y) : \tag{19.11a}$$



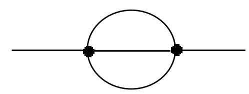
$$+16\,\overbrace{\Psi(x)\Psi(y)}\,:\Psi^3(x)\Psi^3(y): \tag{19.11b}$$



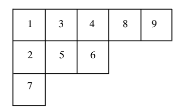
$$+72(\overbrace{\Psi(x)\Psi(y)})^2:\Psi^2(x)\Psi^2(y): \tag{19.11c}$$



$$+96(\overbrace{\Psi(x)\Psi(y)})^3:\Psi(x)\Psi(y): \tag{19.11d}$$



$$24(\overbrace{\Psi(x)\Psi(y)})^4: \tag{19.11e}$$



where the overbrace denotes the Wick contraction of two operators

$$\overbrace{AB} = T[AB] - :AB:$$

Here the diagrams are written according to the following rules:
(i) the number of vertices is equal to the order of approximation;
(ii) the number of internal edges is equal to the number of contractions in a term;
(iii) the number of external edges is equal to the number of uncontracted operators in a term.

We know from (18.22) that

$$\overbrace{\Psi(x)\Psi(y)} = \langle 0|T[\Psi(x)\Psi(y)]|0\rangle = -i\Delta(x-y), \tag{19.12}$$

where $\Delta(x-y)$ is given by (18.16) or, in terms of its Fourier transform, by (18.17). This allows us to express the Wick contraction as

$$\overbrace{\Psi(x)\Psi(y)} = \frac{i}{(2\pi)^4}\int \frac{e^{-ik\cdot(x-y)}}{|k|^2 - m^2 + i\epsilon} d^4k. \tag{19.13}$$

It is clear that only the diagrams (19.11c) compute the interaction of two particles. We have

$$\langle \mathbf{k}_1\mathbf{k}_2|S|\mathbf{k}_1'\mathbf{k}_2'\rangle_2 =$$

$$= \frac{f^2}{2!4!4!}\int \frac{d^4q_1 d^4q_2}{(2\pi)^8}\int d^4x d^4y \frac{e^{-i(q_1+q_2)\cdot(x-y)}\langle \mathbf{k}_3\mathbf{k}_4|:\Psi(x)\Psi(x)::\Psi(y)\Psi(y):|\mathbf{k}_1\mathbf{k}_2\rangle}{(q_1^2 - m^2 + i\epsilon)(q_2^2 - m^2 + i\epsilon)}.$$

The factor $\langle \mathbf{k}_1\mathbf{k}_2| : \Psi(x)\Psi(x) :: \Psi(y)\Psi(y) : |\mathbf{k}'_1\mathbf{k}'_2\rangle$ contributes terms of the form

$$e^{i(k_1-k'_1)x-(k_2-k'_2)\cdot y}/4(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}'_1}E_{\mathbf{k}'_2})^{\frac{1}{2}}(2\pi)^6.$$

Each such term contributes

$$\frac{f^2}{2!4!4!}\int\frac{d^4q_1d^4q_2}{(2\pi)^8 4(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}'_1}E_{\mathbf{k}'_2})^{\frac{1}{2}}(2\pi)^6}\frac{1}{(q_1^2-m^2+i\epsilon)}\frac{1}{(q_2^2-m^2+i\epsilon)}\times$$

$$\int e^{i(k_1-k'_1-q_1-q_2)}d^4x\int e^{i(k_2-k'_2-q_1-q_2)}d^4y.$$

The last two integrals can be expressed via delta functions:

$$(2\pi)^8\delta(k_1-k'_1-q_1-q_2)\delta(k_2-k'_2+q_1+q_2).$$

This gives us again that

$$k_1+k_2=k'_1+k'_2$$

is satisfied if the collision takes place. Also we have $q_2=k_2-k'_2-q_1$, and, changing the variable q_1 by k, we get the term

$$\frac{f^2}{2!4!4!4(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}'_1}E_{\mathbf{k}'_2})^{\frac{1}{2}}(2\pi)^6}\delta(k_1+k_2-k'_1-k'_2)\int\frac{d^4k}{(k^2-m^2+i\epsilon)((k-q)^2-m^2+i\epsilon)},$$

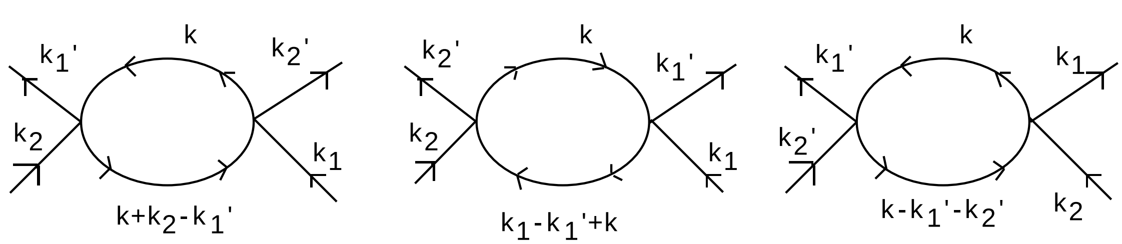
where $q=k_2-k'_2$. Other terms look similar but $q=k_2-k'_1$, $q=k'_1+k'_2$. The total contribution is given by

$$f^2[I(k'_1-k_1)+I(k'_1-k_2)+I(k'_1+k'_2)]\delta(k_1+k_2-k'_1-k'_2)/(E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}'_1}E_{\mathbf{k}'_2})^{\frac{1}{2}}(2\pi)^6,$$

where

$$I(q)=\frac{1}{2(2\pi)^4}\int\frac{d^4k}{(k^2-m^2+i\epsilon)((k-q)^2-m^2+i\epsilon)}. \tag{19.14}$$

Each term can be expressed by one of the following diagram:



Unfortunately, the last integral is divergent. So, in order for this formula to make sense, we have to apply the renormalizationprocess. For example, one subtracts from this expression the integral $I(q_0)$ for some fixed value q_0 of q. The difference of the integrals will be convergent. The integrals (19.14) are examples of *Feynman integrals.* They are integrals of some rational differential forms over algebraic varieties which vary with a parameter. This is studied in algebraic geometry and singularity theories. Many topological and algebraic-geometrical methods were introduced into physics in order to study these integrals. We view integrals (19.14) as special cases of integrals depending on a parameter q:

$$I(q) = \int_{\text{real } k} \frac{d^4k}{\prod_i S_i(k,q)}, \tag{19.15}$$

where S_i are some irreducible complex polynomials in k, q. The set of zeroes of $S_i(k,q)$ is a subvariety of the complex space $\mathbb{C}^4$ which varies with the parameter q. The cycle of integration is $\mathbb{R}^4 \subset \mathbb{C}^4$. Both the ambient space and the cycle are noncompact. First of all we compactify everything. We do this as follows. Let us embed $\mathbb{R}^4$ in $\mathbb{C}^6 \setminus \{0\}$ by sending $\mathbf{x} \in \mathbb{R}^4$ to $(x_0, \ldots, x_5) = (1, \mathbf{x}, |\mathbf{x}|^2) \in \mathbb{C}^6$. Then consider the complex projective space $\mathbb{P}^5(\mathbb{C}) = \mathbb{C}^6 \setminus \{0\}/\mathbb{C}^*$ and identify the image Γ' of $\mathbb{R}^4$ with a subset of the quadric

$$Q := z_5 z_0 - \sum_{i=1}^{4} z_i^2 = 0.$$

Clearly Γ' is the set of real points of the open Zariski subset of Q defined by $z_0 \neq 0$. Its closure Γ in Q is the set $Q(\mathbb{R}) = \Gamma' \cup \{(0,0,0,0,0,1)\}$. Now we can view the denominator in the integrand of (19.14) as the product of two linear polynomials $G = (x_5 - m^2 + i\epsilon)(x_5 - 2\sum_{i=1}^4 x_i q_i + q_5 - m^2 + i\epsilon)$. It is obtained from the homogeneous polynomial

$$F(z;q) = (z_5 - (m^2 - i\epsilon)z_0)(z_5 - 2\sum_{i=1}^{4} z_i q_i + q_5 z_0 - (m^2 - i\epsilon)z_0)$$

by dehomogenization with respect to the variable z_0. Thus integral (19.14) has the form

$$I(q) = \int_{\Gamma} \omega_q, \tag{19.16}$$

where ω_q is a rational differential 4-form on the quadric Q with poles of the first order along the union of two hyperplanes $F(z;q) = 0$. In the affine coordinate system $x_i = z_i/z_0$, this form is equal to

$$\omega_q = \frac{dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4}{(x_5 - m^2 + i\epsilon)(x_5 - 2\sum_{i=1}^4 x_i q_i + q_5 - m^2 + i\epsilon)}.$$

The 4-cycle Γ does not intersect the set of poles of ω_q so the integration makes sense. However, we recall that the integral is not defined if we do not make a renormalization. The reason is simple. In the affine coordinate system

$$y_i = z_i/z_5 = x_i/(x_1^2 + x_2^2 + x_3^2 + x_4^2) = x_i(y_1^2 + y_2^2 + y_3^2 + y_4^2) = x_i|y|^2, \quad i = 1, \ldots 4,$$

the form ω_q is equal to

$$\omega_q = \frac{d(y_1/|y|^2) \wedge d(y_2/|y|^2 \wedge d(y_3/|y|^2) \wedge d(y_4/|y|^2)}{(1-(m^2+i\epsilon)(\sum_{i=1}^4 y_i^2)(1-2\sum_{i=1}^4 y_i q_i + (q_5 - m^2 + i\epsilon)(\sum_{i=1}^4 y_i^2))}.$$

It is not defined at the boundary point $(0,\ldots,0,1) \in \Gamma$. This is the source of the trouble we have to overcome by using a renormalization. We refer to [**Pham**] for the theory of integrals of the form (19.16) which is based on the Picard-Lefschetz theory.

19.3 Let us consider the n-point function for the interacting field Ψ

$$G(x_1,\ldots,x_n) = \langle 0|T[\Psi(x_1)\cdots\Psi(x_n)]|0\rangle.$$

Replacing $\Psi(x)$ with $U(t)^{-1}\Psi_{\rm in}U(t)$ we can rewrite it as

$$G(x_1,\ldots,x_n) = \langle 0|T[U(t_1)^{-1}\Psi_{in}(x_1)U(t_1)\cdots U(t_n)^{-1}\Psi)in(x_n)U(t_n)]|0\rangle =$$

$$= \langle 0|T[U(t)^{-1}U(t,t_1)\Psi_{\rm in}(x_1)U(t_1,t_2)\cdots U(t_{n-1},t_n)\Psi_{\rm in}(x_n)U(t_n,-t)U(-t)]|0\rangle,$$

where we inserted $1 = U(t)^{-1}U(t)$ and $1 = U(-t)^{-1}U(-t)$. When $t > t_1$ and $-t < t_n$, we can pull $U(t)^{-1}$ and $U(-t)$ out of time ordering and write

$$G(x_1,\ldots,x_n) =$$

$$\langle 0|U(t)^{-1}T[U(t,t_1)\Psi_{\rm in}(x_1)U(t_1,t_2)\cdots U(t_{n-1},t_n)\Psi_{\rm in}(x_n)U(t_n,-t)]U(-t)]|0\rangle,$$

when t is sufficiently large.

Take $t \to \infty$. Then $U(-\infty)$ is the identity operator, hence $\lim_{t\to\infty} U(-t)|0\rangle = |0\rangle$. Similarly $U(\infty) = S$ and

$$\lim_{t\to\infty} U(t)|0\rangle = \alpha|0\rangle$$

for some complex number with $|\alpha| = 1$. Taking the inner product with $|0\rangle$, we get that

$$\alpha = \langle 0|S|0\rangle = \langle 0|T\exp(i\int_{-\infty}^{\infty} H_{\rm int}(\tau))d\tau|0\rangle.$$

Using that $U(t,t_1)\cdots U(t_n,-t) = U(t,-t)$, we get

$$G(x_1,\ldots,x_n) = \alpha^{-1}\langle 0|T[\Psi_{\rm in}(x_1)\cdots\Psi_{\rm in}(x_n)\exp(-i\int_{-\infty}^{\infty} H_{\rm int}(\tau)d\tau)]|0\rangle =$$

$$= \frac{\langle 0|T[\Psi_{\rm in}(x_1)\cdots\Psi_{\rm in}(x_n)\exp(-i\int_{-\infty}^{\infty} H_{\rm int}(\tau))d\tau)]|0\rangle}{\langle 0|T\exp(i\int_{-\infty}^{\infty} H_{\rm int}(\tau)d\tau)|0\rangle} =$$

$$= \frac{\sum_{m=0}^{\infty} \frac{(-i)^m}{m!} \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \langle 0|T[\Psi_{\mathrm{in}}(x_1)\cdots\Psi_{\mathrm{in}}(x_n)H_{\mathrm{int}}(t_1)\cdots H_{\mathrm{int}}(t_m)]|0\rangle dt_1\ldots dt_m}{\sum_{m=0}^{\infty} \frac{(-i)^m}{m!} \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \langle 0|T[H_{\mathrm{int}}(t_1)\cdots H_{\mathrm{int}}(t_m)]|0\rangle dt_1\ldots dt_m}. \tag{19.17}$$

Using this formula we see that adding to H_{int} the operator $f(t)\mathbf{id}$ will not change the n-point function $G(x_1,\ldots,x_n)$.

19.4 To compute the correlation function $G(x_1,\ldots,x_n)$ we use Feynman diagrams again. We do it in the example of the Ψ^4-interaction. First of all, the denominator in (19.17) is equal to 1 since $:\Psi^4:$ is normally ordered and hence kills $|0\rangle$. The m-th term in the numerator series is equal to $G(x_1,\ldots,x_n)$ is

$$G(x_1,\ldots,x_n)_m = \frac{(-i)^m}{m!}\langle 0| \int\cdots\int T[\prod_{i=1}^{n}\Psi_{\mathrm{in}}(x_i)\prod_{i=1}^{m}:\Psi_{\mathrm{in}}(y_i)^4:]d^4y_1\ldots d^4y_m]|0\rangle. \tag{19.18}$$

Using Wick's theorem, we write $\langle 0|T[\Psi_{\mathrm{in}}(x_1)\ldots\Psi_{\mathrm{in}}(x_n):\Psi_{\mathrm{in}}(y_1)^4:\ldots\Psi_{\mathrm{in}}(y_m)]|0\rangle$ as the sum of products of all possible Wick contractions

$$\overbrace{\Psi_{\mathrm{in}}(x_i)\Psi_{\mathrm{in}}(x_j)},\quad \overbrace{\Psi_{\mathrm{in}}(x_i)\Psi_{\mathrm{in}}(y_j)},\quad \overbrace{\Psi_{\mathrm{in}}(y_i)\Psi_{\mathrm{in}}(y_j)}.$$

Each x_i occurs only once and y_j occurs exactly 4 times. To each such product we associate a Feynman diagram. On this diagram $x_1,\ldots,x_n$ are the endpoints of the external edges (tails) and $y_1,\ldots,y_m$ are interior vertices. Each vertex is of valency 4, i.e. four edges originate from it. There are no loops (see Exercise 6). Each edge joining z_i with z_j ($z = x_i$ or y_j) represents the Green function $-i\Delta(z_i,z_j)$ equal to the Wick contraction of $\Psi_{\mathrm{in}}(z_i)$ and $\Psi_{\mathrm{in}}(z_j)$.

Note that there is a certain number $A(\Gamma)$ of possible contractions which lead to the same diagram Γ. For example, we may attach to an interior vertex 4 endpoints in 4! different way. It is clear that $A(\Gamma)$ is equal to the number of symmetries of the graph when the interior verices are fixed.

We shall summarize the *Feynman rules* for computing (19.19). First let us introduce more notation.

Let Γ be a graph as described in above (a *Feynman diagram*). Let E be the set of its endpoints, V the set of its interior vertices, and I the set of its internal edges. Let I_v denote the set of internal edges adjacent to $v \in V$, and let E_v be the set of tails coming into v. We set

$$i\Delta_F(p) = \frac{i}{|p|^2 - m^2 + i\epsilon}.$$

1. Draw all distinct diagrams Γ with $n = 2k$ endpoints $x_1,\ldots,x_n$ and m interior vertices $y_1,\ldots,y_m$ and sum all the contributions to (19.19) according to the next rules.

2. Put an orientation on all edges of Γ, all tails must be incoming at the interior vertex (the result will not depend on the orientation.) To each tail with endpoint $e \in E$ assign the 4-vector variable p_e. To each interior edge $\ell \in I$ assign the 4-variable variable k_ℓ.

3. To each vertex $v \in V$, assign the weight

$$A(v) = (-i)(2\pi)^4\delta^4(\sum_{\ell \in I_v, e \in E_v} \pm p_\ell \pm k_e).$$

The sign + is taken when the edge is incoming at vertex, and − otherwise.
4. To each tail with endpoint $e \in E$, assign the factor

$$B(e) = i\Delta_F(p_e).$$

5. To each interior edge $\ell \in I$, assign the factor

$$C(\ell) = \frac{i}{(2\pi)^4}\Delta_F(k_\ell)d^4k_\ell.$$

6. Compute the expression

$$\mathcal{I}(\Gamma) = \frac{1}{A(\Gamma)}\prod_{e \in E} B(e) \int \prod_{v \in V'} A(v) \prod_{\ell \in I} C(\ell). \tag{19.20}$$

Here V' means that we are taking the product over the maximal set of vertices which give different contributions $A(v)$.
7. If Γ contains a connected component which consists of two endpoints x_i, x_j joined by an edge, we have to add to $\mathcal{I}(\Gamma)$ the factor $i\Delta_F(p_i)(2\pi)^4\delta^4(p_i + p_j)$. Let $\mathcal{I}(\Gamma)'$ be obtained from $\mathcal{I}(\Gamma)$ by adding all such factors.

Then the Fourier transform of $G(x_1, \ldots, x_n)_m$ is equal to

$$\hat{G}(p_1, \ldots, p_n)_m = \sum_\Gamma (2\pi)^4\delta^4(p_1 + \ldots + p_n)\mathcal{I}(\Gamma)'. \tag{19.20}$$

The factor in from of the sum expresses momentum conservation of incoming momenta. Here we are summing up with respect to the set of all topologicall distinct Feynman diagrams with the same number of endpoints and interior vertices.

As we have observed already the integral $\mathcal{I}(\Gamma)$ usually diverges. Let L be the number of cycles in γ. The $|I|$ internal momenta must satisfy $|V| - 1$ relations. This is because of the conservation of momenta coming into a vertex; we subtract 1 because of the overall momentum conservation. Let us assume that instead of 4-variable momenta we have d-variable momenta (i.e., we are integrating over $\mathbb{R}^d$. The number

$$D = d(|I| - |V| + 1) - 2|I|$$

is the degree of divergence of our integral. This is equal to the dimension of the space we are integrating over plus the degree of the rational function which we are integrating. Obviously the integral diverges if $D > 0$. We need one more relation among $|V|, |E|$ and $|I|$. Let N be the valency of an interior vertex ($N = 4$ in the case of Ψ^4 interaction). Since each interior edge is adjacent to two interior vertices, we have $N|V| = |E| + 2|I|$. This gives

$$D = d(\frac{N|V| - |E|}{2} - |V| + 1) - N|V| + |E| = d - \frac{1}{2}(d-2)|E| + (\frac{N-2}{2}d - N).$$

In four dimension, this reduces to

$$D = 4 - |E| + (N-4)|V|.$$

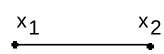
In particular, when $N = 4$, we obtain

$$D = 4 - |E|.$$

Thus we have only two cases with $D \geq 0$: $|E| = 4$ or 2. However, this analysis does not prove the integral converges when $D < 0$. It depends on the Feynman diagram. To give a meaning for $\mathcal{I}(G)$ in the general case one should apply the renormalization process. This is beyond of our modest goals for these lectures.

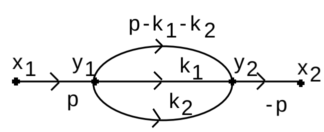
Examples 1. $n = 2$. The only possible diagrams with low-order contribution $m \leq 2$ and their corresponding contributions $\mathcal{I}(\Gamma)'$ for computing $\hat{G}(p, -p)$ are

$m = 0$:



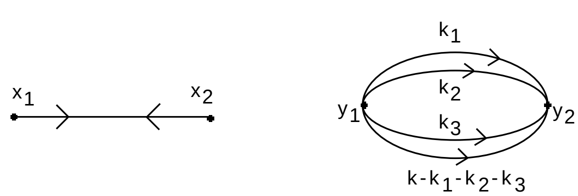
$$I(\Gamma)' = \frac{i}{|p_1|^2 - m^2 + i\epsilon}\delta(p_1 + p_2).$$

$m = 2$:



$$I(\Gamma)' = \frac{(-i\lambda)^2}{3!}\Big(\frac{i}{|p|^2 - m^2 + i\epsilon}\Big)^2 \int \frac{d^4k_1 d^4k_2}{(2\pi)^8}$$
$$\times \frac{i^3}{(|k_1|^2 - m^2 + i\epsilon)(|k_2|^2 - m^2 + i\epsilon)(|p - k_1 - k_2|^2 - m^2 + i\epsilon)}.$$

$m = 2$:

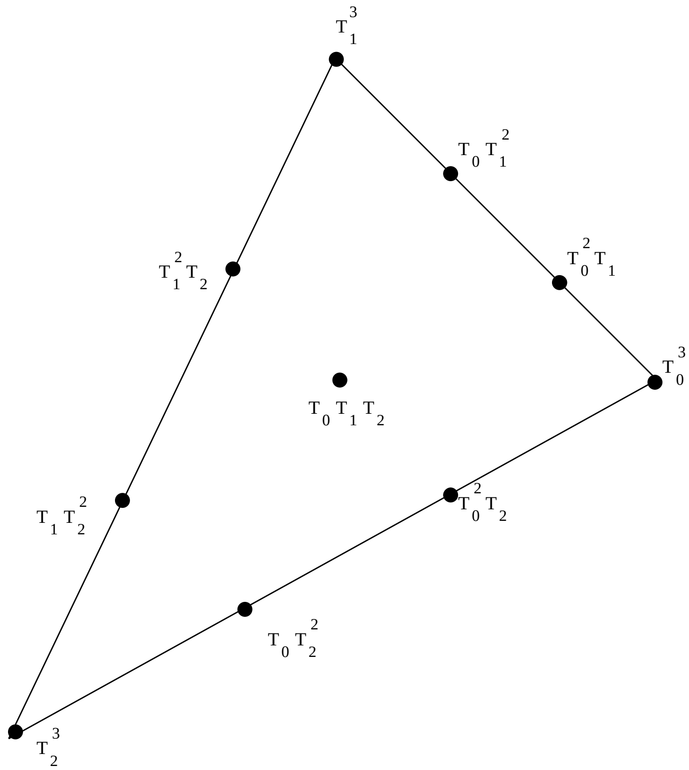


$$\mathcal{I}(\Gamma)' = \frac{(-i\lambda)^2}{4!} \frac{i}{|p|^2 - m^2 + i\epsilon} \int \frac{d^4k_1 d^4k_2 d^4k_3}{(2\pi)^{12}}$$

$$\times \frac{i^4}{(|k_1|^2 - m^2 + i\epsilon)(|k_2|^2 - m^2 + i\epsilon)(|k_3|^2 - m^2 + i\epsilon)(|k_1 + k_2 + k_3|^2 - m^2 + i\epsilon)}.$$

2. $n = 4$. The only possible diagrams with low-order contribution $m \leq 2$ and their corresponding contributions $\mathcal{I}(\Gamma)'$ for computing $\hat{G}(p_1, p_2, p_3, p_4)$ are

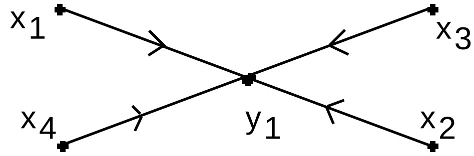
$m = 0:$



$$\mathcal{I}(\Gamma)' = i^2 (2\pi)^2 \frac{\delta(p_1 + p_2)\delta(p_3 + p_4)}{(|p_1|^2 - m^2 + i\epsilon)(|p_3|^2 - m^2 + i\epsilon)}.$$

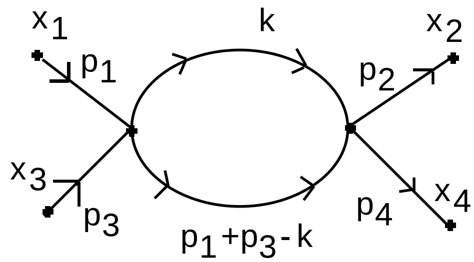
and similarly for other two diagrams.

$m = 1:$



$$\mathcal{I}(\Gamma)' = \frac{(2\pi)^4 \delta^4(p_1 + p_2 + p_3 + p_4) i^4 (-i)}{(|p_1|^2 - m^2 + i\epsilon)(|p_2|^2 - m^2 + i\epsilon)(|p_3|^2 - m^2 + i\epsilon)(|p_4|^2 - m^2 + i\epsilon)}.$$

$m = 2$: Here we write only one diagram of this kind and leave to the reader to draw the other diagrams and the corresponding contributions:



$$\mathcal{I}(\Gamma)' = \frac{1}{2}\prod_{i=1}^{4}\frac{i(-i)^2}{(|p_i|^2 - m^2 + i\epsilon)}\int\frac{d^4k(i)^4\delta(p_1+p_2-p_3-p_4)}{(2\pi)^4(|k|^2-m^2+i\epsilon)(|p_1+p_2-k|^2-m^2+i\epsilon)}$$

and similarly for other two diagrams.

Exercises.

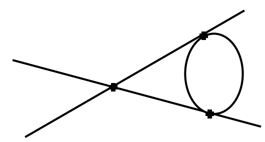
1. Write all possible Feynman diagrams describing the S-matrix up to order 3 for the free scalar field with the interaction given by $H_{\text{int}} = g/3!\Psi^3$. Compute the S-matrix up to order 3.

2. Describe the Feynman rules for the theory $\mathcal{L}(x) = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m_0^2\phi^2) - g\phi^3/3! - f\phi^4/4!$.

3. Describe Feynman diagrams for computing the 4-point correlation function for ψ^4 interaction of the free scalar of order 3.

4. Find the second order matrix elements of scattering matrix S for the Ψ^4-interaction.

5. Find the Feynman integrals which are contributed by the following Feynman diagram:



6. Explain why the normal ordering in $:\Psi_{\text{in}}^4(y_i):$ leads to the absence of loops in Feynman diagrams.

Lecture 20. QUANTIZATION OF YANG-MILLS FIELDS

20.1 We shall begin with quantization of the electro-magnetic field. Recall from Lecture 14 that it is defined by the 4-vector potential (A_μ) up to gauge transformation $A'_\mu = A_\mu + \partial_\mu \phi$. The Lagrangian is the Yang-Mills Lagrangian

$$\mathcal{L}(x) = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} = \frac{1}{2}(|\mathbf{E}|^2 - |\mathbf{H}|^2), \tag{20.1}$$

where $(F_{\mu\nu}) = (\partial_\mu A_\nu - \partial_\nu A_\mu)_{0\leq\mu,\nu\leq 3}$, and $\mathbf{E}, \mathbf{H}$ are the 3-vector functions of the electric and magnetic fields. The generalized conjugate momentum is defined as usual by

$$\pi_0(x) = \frac{\delta\mathcal{L}}{\delta(\delta_t A_0)} = 0, \qquad \pi_i(x) = \frac{\delta\mathcal{L}}{\delta(\delta_t A_i)} = -E_i, \quad i > 0. \tag{20.2}$$

We have

$$\frac{\partial A_\mu}{\partial t} = \frac{\partial A_0}{\partial x_\mu} + E_\mu.$$

Hence the Hamiltonian is

$$H = \int (\dot{A}_\mu \Pi_\mu - \mathcal{L})d^3x = \int \frac{1}{2}(|\mathbf{E}|^2 + |\mathbf{H}|^2) + \mathbf{E}\cdot\nabla A_0)d^3x. \tag{20.3}$$

Now we want to quantize it to convert A_μ and π_μ into operator-valued functions $\Psi_\mu(t,\mathbf{x})$, $\Pi_\mu(t,x)$ with commutation relations

$$[\Psi_\mu(t,\mathbf{x}), \Pi_\nu(t,\mathbf{x})] = i\delta_{\mu\nu}\delta(\mathbf{x}-\mathbf{y}),$$

$$[\Psi_\mu(t,\mathbf{x}), \Psi_\nu(t,\mathbf{x})] = [\Pi_\mu(t,\mathbf{x}), \Pi_\nu(t,\mathbf{x})] = 0. \tag{20.4}$$

As $\Pi_0 = 0$, we are in trouble. So, the solution is to eliminate the A_0 and $\Pi_0 = 0$ from consideration. Then A_0 should commute with all Ψ_μ, Π_μ, hence must be a scalar operator.

We assume that it does not depend on t. By Maxwell's equation $\text{div}\mathbf{E} = \nabla \cdot \mathbf{E} = \rho_e$. Integrating by parts we get

$$\int \mathbf{E} \cdot \nabla A_0 d^3x = \int A_0 \text{div}\mathbf{E} d^3x = \int A_0 \rho_e d^3x.$$

Thus we can rewrite the Hamiltonian in the form

$$H = \frac{1}{2}\int (|\mathbf{E}|^2 + |\mathbf{H}|^2) d^3x + \frac{1}{2}\int A_0 \rho_e d^3x. \tag{20.5}$$

Here $\mathbf{H} = \text{curl } \mathbf{\Psi}$.

We are still in trouble. If we take the equation $\text{div}\mathbf{E} = -\text{div}\Pi = -\rho_e \mathbf{id}$ as the operator equation, then the commutator relation (20.4) gives, for any fixed $\mathbf{y}$,

$$[A_j(t, \mathbf{y}), \nabla \mathbf{\Pi}(t, \mathbf{x})] = -[A_j(t, \mathbf{x}), \rho_e \mathbf{id}] = i \sum_i \partial_i \delta(\mathbf{x} - \mathbf{y}) = 0.$$

But the last sum is definitely non-zero. A solution for this is to replace the delta-function with a function $f(\mathbf{x} - \mathbf{y})$ such that $\sum_i \partial_i f(\mathbf{x} - \mathbf{y}) = 0$. Since we want to imitate the canonical quantization procedure as close as possible, we shall choose $f(\mathbf{x} - \mathbf{y})$ to be the *transverse δ-function*,

$$\delta_{ij}^{\text{tr}}(\mathbf{x} - \mathbf{y}) = \frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}(\delta_{ij} - \frac{k_i k_j}{|\mathbf{k}|^2}) d^3k = i(\delta_{ij}\delta(\mathbf{x} - \mathbf{y}) - \partial_i \partial_j) G_\triangle(\mathbf{x} - \mathbf{y}). \tag{20.6}$$

Here $G_\triangle(\mathbf{x} - \mathbf{y})$ is the Green function of the Laplace operator $\triangle$ whose Fourier transform equals $\frac{1}{|\mathbf{k}|^2}$. We immediately verify that

$$\sum_i \partial_i \delta_{ij}^{\text{tr}}(\mathbf{x} - \mathbf{y}) = \partial_j \delta(\mathbf{x} - \mathbf{y}) - \triangle \partial_j G_\triangle(\mathbf{x} - \mathbf{y}) = \partial_j \delta(\mathbf{x} - \mathbf{y}) - \partial_j \triangle G_\triangle(\mathbf{x} - \mathbf{y}) =$$

$$= \partial_j \delta(\mathbf{x} - \mathbf{y}) - \partial_j \delta(\mathbf{x} - \mathbf{y}) = 0.$$

So we modify the commutation relation, replacing the delta-function with the transverse delta-function. However the equation

$$[\Psi_i(t, \mathbf{x}), \Pi_j(t, \mathbf{y})] = i\delta_{ij}^{\text{tr}}(\mathbf{x} - \mathbf{y}) \tag{20.7}$$

also implies

$$[\nabla \cdot \mathbf{\Psi}(\mathbf{t}, \mathbf{x}), \Pi_j(t, \mathbf{y})] = 0.$$

This will be satisfied if initially $\nabla \cdot \Psi(\mathbf{x}) = 0$. This can be achieved by changing (A_μ) by a gauge transformation. The corresponding choice of A_μ is called the *Coulomb gauge*. In the Coulumb gauge the equation of motion is

$$\partial_\mu \partial^\mu \mathbf{A} = 0$$

(see (14.20)). Thus each coordinate $A_\mu, \mu = 1,2,3$, satisfies the Klein-Gordon equation. So we can decompose it into a Fourier integral as we did with the scalar field:

$$\mathbf{\Psi}(t,x) = \frac{1}{(2\pi)^{3/2}} \int \frac{1}{(2E_k)^{1/2}} \sum_{\lambda=1}^{2} \vec{\epsilon}_{\mathbf{k},\lambda}(a(\mathbf{k},\lambda)e^{-i\mathbf{k}\cdot\mathbf{x}} + a(\mathbf{k},\lambda)^* e^{i\mathbf{k}\cdot\mathbf{x}}).$$

Here the energy E_k satisfies $E_k = |\mathbf{k}|$ (this corresponds to massless particles (*photons*)), and $\vec{\epsilon}_{\mathbf{k},\lambda}$ are 3-vectors. They are called the *polarization vectors.* Since div$\mathbf{\Psi} = 0$, they must satisfy

$$\vec{\epsilon}_{\mathbf{k},\lambda} \cdot \mathbf{k} = 0, \quad \vec{\epsilon}_{\mathbf{k},\lambda} \cdot \vec{\epsilon}_{\mathbf{k},\lambda'} = \delta_{\lambda\lambda'}. \tag{20.8}$$

If we additionally impose the condition

$$\vec{\epsilon}_{\mathbf{k},1} \times \vec{\epsilon}_{\mathbf{k},2} = \frac{\mathbf{k}}{|\mathbf{k}|}, \qquad \vec{\epsilon}_{\mathbf{k},1} = -\vec{\epsilon}_{-\mathbf{k},1}, \quad \vec{\epsilon}_{\mathbf{k},2} = \vec{\epsilon}_{-\mathbf{k},2}, \tag{20.9}$$

then the commutator relations for Ψ_μ, Π_μ will follow from the relations

$$[a(\mathbf{k},\lambda), a(\mathbf{k}',\lambda')^*] = \delta(\mathbf{k}-\mathbf{k}')\delta_{\lambda\lambda'},$$
$$[a(\mathbf{k},\lambda), a(\mathbf{k}',\lambda')] = [a(\mathbf{k},\lambda)^*, a(\mathbf{k}',\lambda')^*] = 0. \tag{20.10}$$

So it is very similar to the Dirac field but our particles are massless. This also justifies the choice of $\delta_{ij}^{\mathrm{tr}}(\mathbf{x}-\mathbf{y})$ to replace $\delta(\mathbf{x}-\mathbf{y})$.

Again one can construct a Fock space with the vacuum state $|0\rangle$ and interpret the operators $a(\mathbf{k},\lambda)$, $a(\mathbf{k},\lambda)^*$ as the annihilation and creation operators. Thus

$$|0\rangle = a(\mathbf{k},\lambda)^*|0\rangle$$

is the state describing a photon with linear polarization $\vec{\epsilon}_{\mathbf{k},\lambda}$, energy $E_{\mathbf{k}} = |\mathbf{k}|^2$, and momentum vector $\mathbf{k}$.

The Hamiltonian can be put in the normal form:

$$H = \frac{1}{2}\int :|\mathbf{E}|^2 + |\mathbf{H}|^2: d^3x = \sum_{\lambda=1}^{2} \int a(\mathbf{k},\lambda)^* a(\mathbf{k},\lambda) d^3k.$$

Similarly we get the expression for the momentum operator

$$\mathbf{P} = \frac{1}{2}\int :\mathbf{E}\times\mathbf{H}: d^3x = \frac{1}{2}\int \mathbf{k}a(\mathbf{k},\lambda)^* a(\mathbf{k},\lambda) d^3k.$$

We have the *photon propagator*

$$iD_{\mu\nu}^{\mathrm{tr}}(x',x) = \langle 0|T[\Psi_\mu(x')\Psi_\nu(x)]|0\rangle. \tag{20.11}$$

Following the computation of the similar expression in the scalar case, we obtain

$$D_{\mu\nu}^{\mathrm{tr}}(x',x) = \frac{1}{(2\pi)^4}\int d^4k \frac{e^{-ik\cdot(x'-x)}}{k^2+\epsilon i} \sum_{\lambda=1}^{2} \vec{\epsilon}_{\mathbf{k},\lambda}^{\nu} \cdot \vec{\epsilon}_{\mathbf{k},\lambda}^{\mu}. \tag{20.12}$$

To compute the last sum, we use the following simple fact from linear algebra.

Lemma. *Let $v_i = (v_{1i}, \dots, v_{ni})$ be a set of orthonormal vectors in n-dimensional pseudo-Euclidean space. Then, for any fixed a, b,*

$$\sum_{i=1}^{n} g_{ii} v_{ai} v_{bi} = g_{ab}.$$

Proof. Let X be the matrix with i-th column equal to v_i. Then X^{-1} is the (pseudo)-orthogonal matrix, i.e., $X^{-1}(g_{ij})(X^{-1})^t = (g_{ij})$. Thus

$$X(g_{ij})X^t = X(X^{-1}(g_{ij})(X^{-1})^t)X^t = (g_{ij}).$$

However, the ab-entry of the matrix in the left-hand side is equal to the sum $\sum_{i=1}^{n} g_{ii} v_{ai} v_{bi}$.

We apply this lemma to the following vectors $\eta = (1,0,0,0), \vec{\epsilon}_{\mathbf{k},1}, \vec{\epsilon}_{\mathbf{k},2}, \hat{k} = (0, \mathbf{k}/|\mathbf{k}|)$ in the Lorentzian space $\mathbb{R}^4$. We get, for any $\mu, \nu \geq 1$,

$$\sum_{\lambda=1}^{2} \vec{\epsilon}^{\,\nu}_{\mathbf{k},\lambda} \cdot \vec{\epsilon}^{\,\mu}_{\mathbf{k},\lambda} = -g_{\mu\nu} + \eta_\mu \eta_\nu - \hat{k}_\mu \hat{k}_\nu = \delta_{\mu\nu} - \frac{k_\mu k_\nu}{|\mathbf{k}|^2}.$$

Plugging this expression in (20.12), we obtain

$$D^{\text{tr}}_{\mu\nu}(x', x) = \frac{1}{(2\pi)^4} \int \frac{e^{-ik\cdot(x'-x)}}{k^2 + \epsilon i} (\delta_{\mu\nu} - \frac{k_\mu k_\nu}{|\mathbf{k}|^2}) d^4k. \tag{20.13}$$

This gives us the following formula for the Fourier transform of $D^{\text{tr}}_{\mu\nu}(x', x)$:

$$D^{\text{tr}}_{\mu\nu}(x', x) = F\Big(\frac{1}{k^2 + \epsilon i} (\delta_{\mu\nu} - \frac{k_\mu k_\nu}{|\mathbf{k}|^2})\Big)(x' - x). \tag{20.14}$$

We can extend the definition of $D^{\text{tr}}_{\mu\nu}$ by adding the components $D^{\text{tr}}_{\mu\nu}$, where μ or ν equals zero. This corresponds to including the component Ψ_0 of $\mathbf{\Psi}$ which we decided to drop in the begining. Then the formula $\text{div}\mathbf{E} = \triangle A_0 = \rho_e$ gives

$$\Psi_0(t, \mathbf{x}) = \int \frac{\rho(t, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d^3y.$$

From this we easily get

$$D^{\text{tr}}_{00}(x, x') = \int e^{-ik\cdot(x'-x)} \frac{1}{|\mathbf{k}|^2} d^4k = \frac{\delta(t - t')}{4\pi|\mathbf{x} - \mathbf{x}'|}. \tag{20.15}$$

Other components $D^{\text{tr}}_{0\mu}(x, x')$ are equal to zero.

20.2 So far we have only considered the non-interaction picture. As an example of an interaction Hamiltonian let us consider the following Lagrangian

$$\mathcal{L} = \frac{1}{2}(|\mathbf{E}|^2 - |\mathbf{H}|^2) + \psi^*(iD - m)\psi, \tag{20.16}$$

where $D = \gamma^\mu(i\partial_\mu + A_\mu)$ is the Dirac operator for the trivial line bundle equipped with the connection (A_μ). Recall also that ψ^* denote the fermionic conjugate defined in (16.11). We can write the Lagrangian as the sum

$$\mathcal{L} = \frac{1}{2}(|\mathbf{E}|^2 - |\mathbf{H}|^2) + (i\psi^*(\gamma^\mu\partial_\mu - m)\psi) + (i\psi^*\gamma^\mu A_\mu\psi).$$

Here we change the notation Ψ to A. The first two terms represent the Lagrangians of the zero charge electromagnetic field (photons) and of the Dirac field (electron-positron). The third part is the interaction part.

The corresponding Hamiltonian is

$$H = \frac{1}{2}\int : |\mathbf{E}|^2 + |\mathbf{H}|^2 : d^3x - \int \psi^*(\gamma^\mu(i\partial_\mu + A_\mu) - m)\psi d^3x, \tag{20.17}$$

Its interaction part is

$$H_{\text{int}} = \int (i\psi^*\gamma^\mu A_\mu\psi)d^3x. \tag{20.18}$$

The interaction correlation function has the form

$$G(x_1, \ldots, x_n; x_{n+1}, \ldots, x_{2n}; y_1, \ldots, y_m) =$$

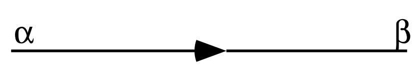
$$= \langle 0|T[\psi(x_1)\ldots\psi(x_n)\psi^*(x_{n+1})\ldots\psi^*(x_{2n})A_{\mu_1}(y_1)\ldots A_{\mu_m}(y_m)]|0\rangle. \tag{20.19}$$

Because of charge conservation, the correlation functions have an equal number of ψ and ψ^* fields. For the sake of brevity we omitted the spinor indices. It is computed by the formula

$$G(x_1, \ldots, x_n; x_{n+1}, \ldots, x_{2n}; y_1, \ldots, y_m) =$$

$$\frac{\langle 0|T[\psi_{\text{in}}(x_1)\ldots\psi^*_{\text{in}}(x_{n+1})\ldots\psi^*_{\text{in}}(x_{2n})A^{\text{in}}_{\mu_1}(y_1)\ldots A^{\text{in}}_{\mu_m}(y_m)\exp(iH_{\text{in}})|0\rangle}{\langle 0|T[\exp(iH_{\text{in}})]|0\rangle}. \tag{20.20}$$

Again we can apply Wick's formula to write a formula for each term in terms of certain integrals over Feynman diagrams. Here the 2-point Green functions are of different kind. The first kind will correspond to solid edges in Feynman diagrams:



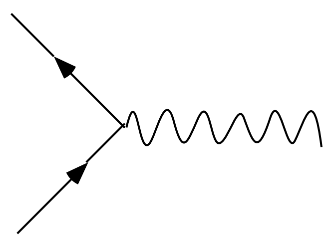
$$\langle 0|T[\psi_\alpha(x)\psi^*_\beta(x')|0\rangle = \overbrace{\psi_\alpha(x)\psi^*_\beta(x')} = \frac{i}{(2\pi)^4}\int e^{-ik\cdot(x'-x)}(k_\mu\gamma^\mu + i\epsilon)^{-1}_{\alpha\beta}d^4k.$$

We have not discussed the Green function of the Dirac equation but the reader can deduce it by analogy with the Klein-Gordon case. Here the indices mean that we take the $\alpha\beta$

entry of the inverse of the matrix $k_\mu \gamma^\mu - m + i\epsilon I_4$. The propagators of the second field correspond to wavy edges

$$\langle 0|T[A_\mu(x)A_\nu(x')|0\rangle = \overbrace{A_\mu(x)A_\nu(x')} = iD^{\mathrm{tr}}_{\mu\nu}(x',x).$$

Now the Feynman diagrams contain two kinds of tails. The solid ones correspond to incoming or outgoing fermions. The corresponding endpoint possesses one vector and two spinor indices. The wavy tails correspond to photons. The endpoint has one vector index $\mu = 1,\ldots,4$. It is connected to an interior vertex with vector index ν. The diagram is a 3-valent graph.



We leave to the reader to state the rest of Feynman rules similar to these we have done in the previous lecture.

20.3 The disadvantage of the quantization which we introduced in the previous section is that it is not Lorentz invariant. We would like to define a relativistic n-point function. We use the path integral approach. The formula for the generating function should look like

$$Z[J] = \int [d[(A_\mu)] \exp(i(\mathcal{L} + J^\mu A_\mu), \tag{20.21}$$

where $\mathcal{L}$ is a Lagrangian defined on gauge fieleds. Unfortunately, there is no measure on the affine space $\mathcal{A}(\mathfrak{g})$ of all gauge potentials (A_μ) on a principal G-bundle which gives a finite value to integral (20.21). The reason is that two connections equivalent under gauge transformation give equal contributions to the integral. Since the gauge group is non-compact, this leads to infinity. The idea is to cancel the contribution coming from integrating over the gauge group.

Let us consider a simple example which makes clear what we should do to give a meaning to the integral (20.21). Let us consider the integral

$$Z = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x-y)^2} dy dx.$$

Obviously the integral is divergent since the function depends only on the difference $x - y$. The latter means that the group of translations $x \to x + a, y \to y + a$ leave the integrand

and the measure invariant. The orbits of this group in the plane $\mathbb{R}^2$ are lines $x - y = c$. The orbit space can be represented by a curve in the plane which intersects each orbit at one point (a gauge slice). For example we can choose the line $x + y = 0$ as such a curve. We can write each point (x, y) in the form

$$(x,y) = (\frac{x-y}{2}, \frac{y-x}{2}) + (\frac{x+y}{2}, \frac{x+y}{2}).$$

Here the first point lies on the slice, and the second corresponds to the translation by (a, a), where $a = \frac{x+y}{2}$. Let us make the change of variables $(x, y) \to (u, v)$, where $u = x - y, v = x + y$. Then $x = (u + v)/2$, $y = (u - v)/2$, and the integral becomes

$$Z = \frac{1}{2}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} e^{-u^2}du\Big)dv = \sqrt{\pi}\int_{-\infty}^{\infty} da. \tag{20.22}$$

The latter integral represents the integral over the group of transformations. So, if we "divide" by this integral (equal to ∞), we obtain the right number.

Let us rewrite (20.22) once more using the delta-function:

$$Z = \int_{-\infty}^{\infty} da \int_{-\infty}^{\infty} 2\delta(x+y)e^{-(x-y)^2}dxdy = \int_{-\infty}^{\infty} da \int_{-\infty}^{\infty} 2e^{-4x^2}dx = \int_{-\infty}^{\infty} da \int_{-\infty}^{\infty} e^{-z^2}dz.$$

We can interpret this by saying that the delta-function fixes the gauge; it selects one representative of each orbit. Let us explain the origin of the coefficient 2 before the delta-function. It is related to the choice of the specific slice, in our case the line $x + y = 0$. Suppose we take some other slice $S : f(x, y) = 0$. Let $(x, y) = (x(s), y(s))$ be a parametrization of the slice. We choose new coordinates (s, a), where a is is uniquely defined by $(x, y) - (a, a) = (x', y') \in S$. Then we can write

$$Z = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-h(s)}|J|dads, \tag{20.23}$$

where $x - y = x'(s) - y'(s) = h(s)$ and $J = \frac{\partial(x,y)}{\partial(s,a)}$ is the Jacobian. Since $x = x'(s) + a, y = y'(s) + a$, we get

$$J = \begin{vmatrix} \frac{\partial x}{\partial s} & 1 \\ \frac{\partial y}{\partial s} & 1 \end{vmatrix} = \frac{\partial x}{\partial s} - \frac{\partial y}{\partial s} = \frac{\partial x'}{\partial s} - \frac{\partial y'}{\partial s}.$$

Now we use that

$$0 = \frac{\partial f}{\partial x'}\frac{\partial x'}{\partial s} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial s} = (\frac{\partial f}{\partial x'}, \frac{\partial f}{\partial y'}) \times (-\frac{\partial y'}{\partial s}, \frac{\partial x'}{\partial s}).$$

Let us choose the parameter s in such a way that

$$(\frac{\partial f}{\partial x'}, \frac{\partial f}{\partial y'}) = (-\frac{\partial y'}{\partial s}, \frac{\partial x'}{\partial s}).$$

One can show that this implies that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(f(x,y))dxdy=\int_{-\infty}^{\infty}ds.$$

This allows us to rewrite (20.23)

$$Z=\int_{-\infty}^{\infty}da\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(f(x,y))\Big|\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\Big|e^{-(x-y)^2}dxdy.$$

Now our integral does not depend on the choice of parametrization. We succeeded in separating the integral over the group of transformations. Now we can redefine Z to set

$$Z=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(f(x,y))\Big|\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\Big|e^{-(x-y)^2}dxdy.$$

Now let us interpret the Jacobian in the following way. Fix a point (x_0, y_0) on the slice, and consider the function

$$f(a)=f(x_0+a,y_0+a) \tag{20.24}$$

as a function on the group. Then

$$\frac{df}{da}\Big|_{f=0}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\Big|_{f=0}.$$

20.4 After the preliminary discussion in the previous section, the following definition of the path integral for Yang-Mills theory must be clear:

$$Z[J]=\int d[A_\mu]\Delta_{FP}\delta(F(A_\mu))\exp\Big[i\int(\mathcal{L}(A)+J^\mu A_\mu)d^4x\Big]. \tag{20.25}$$

Here $F:\mathcal{A}(\mathfrak{g})\to \mathrm{Maps}(\mathbb{R}^4\to\mathfrak{g})$ is a functional on the space of gauge potentials analogous to the function $f(x,y)$. It defines a slice in the space of connections by fixing a gauge (e.g. the Coulomb gauge for the electromagnetic field). The term Δ_{FP} is the *Faddeev-Popov* determinant. It is equal to

$$\Delta_{FP}=\Big|\frac{\delta F(g(x)A_\mu(x))}{\delta g(x)}\Big|_{g(x)=1}\Big|, \tag{20.26}$$

where $g(x):\mathbb{R}^4\to G$ is a gauge transformation.

To compute the determinant, we choose some local parameters θ^a on the Lie group G. For example, we may choose a basis η_a of the Lie algebra, and write $g(x)=\exp(\theta^a(x)\eta_a)$. Fix a gauge $F(A)$. Now we can write

$$F(g(x)A(x)))_a=F(A)_a+\int\sum_a M(x,y)_{ab}\theta^b(y)d^4y+\ldots, \tag{20.27}$$

where $M(x,y)$ is some matrix function with

$$M(x,y)_{ab} = \frac{\delta F_a(x)}{\delta \theta^b(y)}.$$

To compute the determinant we want to use the functional analog of the Gaussian integral

$$(\det M)^{-1/2} = \int\int d\phi e^{-\phi(x)M(x)\phi(x)} d\phi.$$

For our purpose we have to use its anti-commutative analog. We shall have

$$\det(M) = \int DcDc^* \exp\Big(i \int c^*(x)M(x,y)c(y)d^4x\Big), \tag{20.28}$$

where $c(x)$ and $c^*(y)$ are two fermionic fields (*Faddeev-Popov ghost fields*), and integration is over Grassmann variables. Let us explain the meaning of this integral. Consider the *Grassmann variables* α satisfying

$$\alpha^2 = \{\alpha, \alpha\} = 0.$$

Note that $\frac{dx}{dx} = 1$ is equivalent to $[x, \frac{d}{dx}] = 1$, where x is considered as the operator of multiplication by x. Then the Grassmann analog is the identity

$$\{\frac{d}{d\alpha}, \alpha\} = 1.$$

For any function $f(\alpha)$ its Taylor expansion looks like $F(\alpha) = a + b\alpha$ (because $\alpha^2 = 0$). So $\frac{d^2 f}{d\alpha^2} = 0$, hence

$$\{\frac{d}{d\alpha}, \frac{d}{d\alpha}\} = 0.$$

We have $\frac{df}{d\alpha} = -b$ if a, b are anti-commuting numbers. This is because we want to satisfy the previous equality.

There is also a generalization of integral. We define

$$\int d\alpha = 0, \qquad \int \alpha d\alpha = 1.$$

This is because we want to preserve the property

$$\int\limits_{-\infty}^{\infty} \phi(x)dx = \int\limits_{-\infty}^{\infty} \phi(x+c)dx$$

of an ordinary integral. For $\phi(\alpha) = a + b\alpha$, we have

$$\int \phi(\alpha)d\alpha = \int (a+b\alpha)d\alpha = a\int d\alpha + b\int \alpha d\alpha =$$

$$= \int (a + (\alpha + c)b)d\alpha = (a + bc) \int d\alpha + b \int \alpha d\alpha.$$

Now suppose we have N variables α_i. We want to perform the integration

$$I(A) = \int \exp(\sum_{i,j=1}^{N} \alpha_i A_{ij} \beta_j) \prod_{i=1}^{N} d\alpha_i d\beta_i,$$

where α_i, β_i are two sets of N anti-commuting variables. Since the integral of a constant is zero, we must have

$$I(A) = \int \frac{1}{N!} \Big(\sum_{i,j=1}^{N} \alpha_i A_{ij} \beta_j \Big)^N) \prod_{i=1}^{N} d\alpha_i d\beta_i.$$

The only terms which survive are given by

$$I(A) = \int \Big(\sum_{\sigma \in S_N} \epsilon(\sigma) A_{1\sigma(1)} \ldots A_{N\sigma(N)} \Big) \prod_{i=1}^{N} d\alpha_i d\beta_i = \det(A).$$

Now we take $N \to \infty$. In this case the variables α_i must be replaced by functions $\psi(x)$ satisfying

$$\{\phi(x), \phi(y)\} = 0.$$

Examples of such functions are fermionic fields like half-spinor fields. By analogy, we obtain integral (20.28).

20.5 Let us consider the simplest case when $G = U(1)$, i.e., the Maxwell theory. Let us choose the Lorentz-Landau gauge

$$F(A) = \partial^\mu A_\mu = 0.$$

We have $(gA)_\mu = A_\mu + \frac{\partial \theta}{\partial x_\mu}$ so that

$$F(gA) = \partial^\mu A_\mu + \partial^\mu \partial_\mu \theta.$$

Then the matrix M becomes the kernel of the D'Alembertian operator $\partial^\mu \partial_\mu$. Applying formula (20.28), we see that we have to add to the action the following term

$$\int \int c^*(x) \partial^\mu \partial_\mu c(y) d^4x d^4y,$$

where $c^*(x), c(x)$ are scalar Grassmann ghost fields. Since this does not depend on A_μ, it gives only a multiplicative factor in the S-matrix, which can be removed for convenience.

On the other hand, let us consider a non-abelian gauge potential $(A_\mu) : \mathbb{R}^4 \to \mathfrak{g}$. Then the gauge group acts by the formula $gA = gAg^{-1} + g^{-1}dg$. Then

$$M(x,y)_{ab} = \frac{\delta F(gA(x))}{\delta\theta^b(y)}\Big|_{g(x)=1} = \frac{\delta}{\delta\theta^b(y)}\Big[\frac{\delta F(A)_a}{\delta A^c_\mu(x)} D^\mu_{cd}\theta^d(x)\Big]_{\theta=0} =$$

$$= \frac{\delta F(A)_a}{\delta A^c_\mu(x)} D^\mu_{cb}\Delta(x-y),$$

where

$$D^\mu_{cd} = \partial^\mu\delta_{ab} + C_{abc}A^c_\mu,$$

and C_{abc} are the structure constants of the Lie algebra $\mathfrak{g}$. Choose the gauge such that

$$\partial^\mu A^a_\mu = 0.$$

Then

$$\frac{\delta F(A)}{\delta A^c_\mu} = \delta^\mu.$$

This gives, for any fixed y,

$$M(x,y)_{ab} = [\partial_\mu\partial^\mu\delta_{ab} + C_{abc}A^c_\mu\partial^\mu]\delta(x-y), \tag{20.29}$$

which we can insert as an additional term in the action:

$$\int d^4x(\alpha^*)^a(\delta_{ab}\partial_\mu\partial^\mu + C_{abc}A^c_\mu)\alpha^b.$$

Here α^*, α are ghost fermionic fields with values in the Lie algebra.

LITERATURE

The following references were used in preparations of my lecture notes.

V. I. Arnold, *Mathematical methods of classical mechanics.* Springer-Verlag. 1989

M. Atiyah, *Geometry of Yang-Mills fields.* Pisa. 1979

F. Berezin, M. Shubin, *The Schrödinger equation.* Kluwer. 1991

F. Berezin, *The method of second quantization.* Academic Press. 1966

Braid groups, knot theory and statistical mechanics. World Scient. 1989.

N. Bourbaki, *Algebra*, Springer-Verlag. 1989

D. Bleecker, *Gauge theory and variational principles.* Addison Wesley. 1981

N. N. Bogolyubov, D.V. Shirkov, *Introduction to the theory of quantized filelds.* Wiley-Interscience. 1959.

E. Cartan, *The theory of spinors.* Dover, 1981.

L. Faddeev, O. Yakubovsky. Lecture on quantum mechanics for mathematics students. Lecture Notes at St. Petersburg University.

F. Pham, *Introduction a l'etude topologique des singularités de Landau¿* Paris. Gauthiere-Villars. 1967.

S. Fomin. A. Kolmogorov, *Elements of the theory of functions and functional analysis.* Graylock Press. 1961.

I. Gelfand, G. Shilov, *Generalized functions* . Academic Press. 1961.

Geometry and Quantum field theory, ed. D. Freed, K. Uhlenbeck, Park City Lecture Notes. AMS. Providence. 1991.

M. Guidry, *Gauge fields theories.* Joh Wiley and Sons, Inc. 1991.

B. Hatfield, *Quantum field theory of point partciles and strings.* Addison-Wesley. 1992.

J.-I. Igusa, *Theta functions.* Springer-Verlag. 1972.

C. Itzykson, J.-B.Zuber,*Quantum Field Theory.* McGraw-Hill. 1980.

T. Jordan, Linear operators for Quantum mechanics, John Wiley and Sons, Inc, 1969.

M. Kaku, *Quantum field theory.* Oxford University Press. 1993.

H. Lange, Ch. Birkenhake, *Complex abelian varieties.* Sprin ger-Verlag.1992.

I. D. Laurie, *A unified grand tour of theoretical physics.* Institute of Physics Publishing. 1990.

B. Lawson, *Spin geometry*, Princeton Univ. Press. 1989.

G. Mackey, *The mathematical foundations of quantum mechanics.* Benjamin. 1963.

J. Marsden, *Lectures on Mechanics.* Cambridge Univ. Press, 1992.

D. Martin,*Manifold theory, An introduction for mathematical physicists.* Ellis Horwood, 1991,

K. Nomizu, *Lie groups and differential geometry.* Math. Soc. Japan. 1956.

A. Perelomov, *Integral systems of Classical mechanics and Lie algebras.* Birkhauser, 1990

A. Pressley, G. Segal, *Loop groups.* Oxford Science Publ. 1986.

E. Prugovecki, *Quantum mechanics in Hilbert spaces.* Academic Press. 1981.

P. Ramond, *Field theory: A Modern primer.* Addison-Wesley. 1989.

J.L. Scholma, *A Lie algebraic study of some integrable systems associated with root systems.* CWI Tracts, Amsterdam. 1993.

A. Schwarz, *Quantum field theory and topology.* Springer-Verlag. 1993.

J.-P. Serre, *A course in arithmetic.* Sp[ronger-Verlag, 1973.

S. Sternberg, *Group Theory and physics.* Cambridge University Press. 1994.

E. T. Whittaker, G. N. Watson, *A Course of modern analysis.* Cambridge Univ. Press. 1927.