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

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

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

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

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

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

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

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

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

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

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

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

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

$\mathbf{p} = (p_1, \ldots, p_n)$ 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

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

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 $m \in M, (x_1, \ldots, x_m)$ 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 $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ 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:

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

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

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

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

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

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 $\xi^\natural_x \in T(M)_x$. 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 $g^t_\eta : M \to M, x \to \mu_\eta(t,x)$. 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

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

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

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

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

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

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

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

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

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 $\dot{q}_i F'(X) = c$ 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 $g^t_{\eta_i}$ 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 $g^t_{\eta_i}$ is complete (i.e., defined for all t). For each $t, t' \in \mathbb{R}$ and $i, j = 1, \ldots, n$, the diffeomorphisms $g^t_{\eta_i}$ and $g^{t'}_{\eta_j}$ 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

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

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 $\gamma(t) = g^t_{\eta_i}(\gamma(0))$. Let $(a_1, \ldots, a_n) = f^{-1}(\gamma(0)) \in \mathbb{R}^n$. Then

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

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

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

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

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

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

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

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

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

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

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 $\langle A|\omega\rangle = \int_{\mathbb{R}} x d\mu_A$ 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

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

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

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

where $A = S \cdot \mathrm{diag}[d_1, \dots, d_n] \cdot S^{-1}$. 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 $\mathrm{End}(V)$, 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

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

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

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

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

$$\left(\int_{\mathbb{R}} x d\mu\right) f = \lim_{n\to\infty} \sum_{i=1}^{2n} \frac{(-n+i-1)C}{n} \phi_{[\frac{(-n+i-1)C}{n}, \frac{(-n+i)C}{n}]} f = gf.$$

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

$$\langle \theta(\lambda - x)(A) | \omega \rangle = Tr(MP_A(\lambda)).$$

Comparing this with section 4.4 from Lecture 4, we see that the function

$$\mu_A(\lambda) = Tr(MP_A(\lambda))$$

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

$$\mu_A(\lambda) = \sum_{n:\lambda_n \le \lambda} (Me_n, e_n).$$

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

$$\int_{\mathbb{R}^3} |\psi(x, y, z)|^2 dx dy dz = 1, \quad \lim_{||(x,y,z)|| \to \infty} \psi(x, y, z) = 0.$$

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

$$f(X) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) e^{iXt} dt. \tag{6.5}$$

This agrees with the formula

$$f(X) = \int_{\mathbb{R}} f(\lambda) d\mu_X(\lambda)$$

we used in the previous Lecture. Indeed,

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) e^{iXt} dt = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) \big(\int_{\mathbb{R}} e^{ixt} d\mu_X \big) dt =$$

$$= \big(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) e^{ixt} dt) \big) d\mu_X = \int_{\mathbb{R}} f(x) d\mu_X = f(X).$$

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

$$g(x) = f_1 * f_2 := \int_{\mathbb{R}} f_1(t) f_2(x-t) dt$$

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

$$\phi(x) \to \int_{\mathbb{R}} \rho(x)\phi(x)dx,$$

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

$$\phi \to \int_{\mathbb{R}} \delta(x)\phi(x)dx.$$

More generally, the functional $\phi \to \phi(a)$ can be written in the form

$$\ell(\phi) = \int_{\mathbb{R}} \delta(x-a)\phi(x)dx = \int_{\mathbb{R}} \delta(x)\phi(x-a)dx.$$

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

$$\ell(\phi) = \int_{\mathbb{R}} \frac{1}{x}\phi(x)dx.$$

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

$$\int_{\mathbb{R}} \frac{1}{x}\phi(x)dx = \int_a^b \frac{1}{x}\phi(x)dx = \int_a^b \frac{\phi(x)-\phi(0)}{x}dx + \int_a^b \frac{\phi(0)}{x}dx.$$

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

$$\int_a^b \frac{\phi(0)}{x}dx := \lim_{\epsilon\to 0}\Big(\int_a^{-\epsilon} \frac{\phi(0)}{x}dx + \int_{\epsilon}^{b} \frac{\phi(0)}{x}dx\Big).$$

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, $\int\limits_{\mathbb{R}} f\phi dx = 0$. 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

$$\int\limits_{\mathbb{R}} f'(x)\phi(x)dx := -\int\limits_{\mathbb{R}} f(x)\phi'(x)dx. \tag{6.9}$$

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

$$\int\limits_{\mathbb{R}} \delta'(x)\phi(x)dx := -\int\limits_{\mathbb{R}} \delta(x)\phi'(x)dx = -\phi'(0).$$

Thus the derivative is the linear function $\phi \to -\phi'(0)$.

4. Let $\theta(x)$ be the Heaviside function. It defines a regular distribution

$$\ell(\phi) = \int\limits_{\mathbb{R}} \theta(x)\phi(x)dx = \int_0^\infty \phi(x)dx.$$

Obviously, $\theta'(0)$ does not exist. But, considered as a distribution, we have

$$\int\limits_{\mathbb{R}} \theta'(x)\phi(x)dx = -\int_0^\infty \phi'(x)dx = \phi(0).$$

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

$$\int_{\mathbb{R}} f * 1\phi(x)dx = \int_{\mathbb{R}} f\Big(\int_{\mathbb{R}} \phi(x+y)dx\Big)dy = \Big(\int_{\mathbb{R}} f(x)dx\Big)\Big(\int_{\mathbb{R}} \phi(x)dx\Big).$$

This shows that

$$f * 1 = \Big(\int_{\mathbb{R}} f(x)dx\Big) \cdot 1.$$

This suggests that we take $f * 1$ as the definition of $\int_{\mathbb{R}} f(x)dx$ for any distribution f with compact support. For example,

$$\int_{\mathbb{R}} \delta(x)\Big(\int_{\mathbb{R}} \phi(x+y)dy\Big)dx = \int_{\mathbb{R}} \phi(y)dy.$$

Hence

$$\int_{\mathbb{R}} \delta(x)dx = \delta * 1 = 1, \tag{6.16}$$

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,

$$\int_{\mathbb{R}} \overline{T(f)}\phi dx = \int_{\mathbb{R}} \bar{f}T^*(\phi)dx).$$

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 $\lambda' \in \mathbb{R}$,

$$\int_{\mathbb{R}} (\int_{\mathbb{R}} f_\lambda(x)\bar{f}_{\lambda'}(x)dx)\phi(\lambda)d\lambda = \phi(\lambda-\lambda'),$$

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

$$\widehat{P_1\phi}(\mathbf{p}) = \frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} P_1\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q} = \frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} \frac{\hbar}{i}\frac{\partial\phi(\mathbf{q})}{\partial q_1}e^{-i(p_1q_1+\ldots+p_nq_n)}d\mathbf{q} =$$

$$= \frac{1}{(\sqrt{2\pi})^n}\left(\frac{\hbar}{i}\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}\Big|_{q_1=-\infty}^{q_1=\infty} + \hbar p_1\int_{\mathbb{R}^n}\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q}\right) =$$

$$= \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} P_1\phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q} = \hbar p_1\hat{\phi}(\mathbf{p}).$$

Similarly we obtain

$$P_1\hat{\phi}(\mathbf{p}) = \frac{\hbar}{i}\frac{\partial}{\partial p_1}\hat{\phi}(\mathbf{p}) = \frac{\hbar}{i}\frac{\partial}{\partial p_1}\frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} \phi(\mathbf{q})e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q} =$$

$$= \frac{\hbar}{i}\frac{1}{(\sqrt{2\pi})^n}\int_{\mathbb{R}^n} \phi(\mathbf{q})(-iq_1)e^{-i\mathbf{p}\cdot\mathbf{q}}d\mathbf{q} = -\hbar\widehat{p_1\phi} = -\hbar\widehat{Q_1\phi}(\mathbf{p}).$$

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:

$$\ell(\phi) = \int_{\mathbb{R}^n} \bar{f}(\mathbf{x})\phi(\mathbf{x})d\mathbf{x}.$$

Example 10. Let $\phi(\mathbf{x},\mathbf{y}) \in K_{2n}$. Consider the linear function

$$\ell(\phi) = \int_{\mathbb{R}^n} \phi(\mathbf{x},\mathbf{x})d\mathbf{x}.$$

This functional defines a real distribution which we denote by $\delta(\mathbf{x}-\mathbf{y})$. By definition

$$\int_{\mathbb{R}^n} \delta(\mathbf{x}-\mathbf{y})\phi(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} = \int_{\mathbb{R}^n} \phi(\mathbf{x},\mathbf{x})d\mathbf{x}.$$

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

$$\int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} \phi(\mathbf{x},\mathbf{y})\delta(\mathbf{x}-\mathbf{y})d\mathbf{x}\Big)d\mathbf{y}.$$

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*

$$B(\phi(\mathbf{x}), \psi(\mathbf{x})) = \ell(\phi(\mathbf{x})\psi(\mathbf{y})) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \bar{K}(\mathbf{x}, \mathbf{y})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

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

$$(T\phi, \bar{\psi}) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \bar{K}(\mathbf{x}, \mathbf{y})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

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

$$T\phi(\mathbf{y}) = \int_{\mathbb{R}^n} \bar{K}(\mathbf{x}, \mathbf{y})\phi(\mathbf{x})d\mathbf{x}.$$

Examples 11. Consider the operator $T\phi = g\phi$ where $g \in K_n$ is any locally integrable function on $\mathbb{R}^n$. Then

$$(T\phi, \bar{\psi}) = \int_{\mathbb{R}^n} \bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{x})d\mathbf{x}.$$

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

$$\int_{\mathbb{R}^n} \bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{x})d\mathbf{x} = \int_{\mathbb{R}}\int_{\mathbb{R}} \delta(\mathbf{x} - \mathbf{y})\bar{g}(\mathbf{x})\phi(\mathbf{x})\psi(\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

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

$$Q_i\phi(\mathbf{q}) = \int_{\mathbb{R}^n} q_i\delta(\mathbf{q} - \mathbf{y})\phi(\mathbf{y})d\mathbf{y}.$$

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

$$K_2 = \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} K^g_{\mathbf{u}',\mathbf{v}'}(\mathbf{q},\mathbf{q}') K^f_{\mathbf{u},\mathbf{v}}(\mathbf{q}',\mathbf{y}) d\mathbf{q}' d\mathbf{u} d\mathbf{v} d\mathbf{u}' d\mathbf{v}' =$$

$$\int\limits_{\mathbb{R}^n}\dots\int\limits_{\mathbb{R}^n}\hat{f}(\mathbf{u},\mathbf{v})\hat{g}(\mathbf{u}',\mathbf{v}')e^{i\mathbf{v}'\cdot\mathbf{q}}e^{-i\hbar\mathbf{u}'\cdot\mathbf{v}'/2}e^{i\mathbf{v}\cdot(\mathbf{q}-\hbar\mathbf{u}')}\delta(\mathbf{q}-\hbar(\mathbf{u}+\mathbf{u}')-\mathbf{y})e^{-i\hbar\mathbf{u}\cdot\mathbf{v}/2}d\mathbf{u}d\mathbf{v}d\mathbf{u}'d\mathbf{v}'.$$

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

$$\lim_{\hbar\to 0}A_fA_g=\lim_{\hbar\to 0}A_gA_f.$$

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

$$\frac{i}{\hbar}\int\limits_{\mathbb{R}^{4n}}\hat{f}(u,v)\hat{g}(u',v')e^{i(v'+v)\cdot q-\frac{\hbar(u\cdot v+u'\cdot v')}{2}}\delta(q-\hbar(u+u')-y)(e^{-i\hbar v'\cdot u}-e^{-i\hbar u'\cdot v})dudvdu'dv',$$

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

$$\int\limits_{\mathbb{R}^{4n}}\hat{f}(\mathbf{u}',\mathbf{v}')\hat{g}(\mathbf{u}'',\mathbf{v}'')e^{i(\mathbf{v}'+\mathbf{v}'')\cdot\mathbf{q}}\delta(\mathbf{q}-\mathbf{y})(\mathbf{v}''\cdot\mathbf{u}'-\mathbf{u}''\cdot\mathbf{v})d\mathbf{u}'d\mathbf{v}'d\mathbf{u}''d\mathbf{v}''. \tag{6.21}$$

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

$$\widehat{\{f,g\}}(\mathbf{u},\mathbf{v})=$$

$$=\int\limits_{\mathbb{R}^{4n}}\hat{f}(\mathbf{u}',\mathbf{v}')\hat{g}(\mathbf{u}'',\mathbf{v}'')(\mathbf{v}''\cdot\mathbf{u}'-\mathbf{u}''\cdot\mathbf{v}')\delta(\mathbf{u}'+\mathbf{u}''-\mathbf{u})\delta(\mathbf{v}'+\mathbf{v}''-\mathbf{v})d\mathbf{u}'d\mathbf{v}'d\mathbf{u}''d\mathbf{v}''.$$

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

$$Tr(V(\mathbf{v})U(\mathbf{u})) = \frac{1}{(2\pi)^n}\int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} e^{i\mathbf{v}\cdot\mathbf{q}}e^{i(\mathbf{q}-\hbar\mathbf{u})\cdot\mathbf{t}}e^{-i\mathbf{q}\cdot\mathbf{v}}d\mathbf{q}d\mathbf{t} =$$

$$= \frac{1}{(\sqrt{2\pi})^n}\left(\int\limits_{\mathbb{R}^n} e^{i\mathbf{v}\cdot\mathbf{q}}d\mathbf{v}\right)\frac{1}{(\sqrt{2\pi})^n}\left(\int\limits_{\mathbb{R}^n} e^{-i\hbar\mathbf{t}\cdot\mathbf{u}}d\mathbf{t}\right) = (\frac{2\pi}{\hbar})^n\delta(\mathbf{v})\delta(\mathbf{u}).$$

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

$$\int_{\mathbb{R}} x^2H_n(x)^2e^{-x^2}dx = \frac{1}{2}\int_{\mathbb{R}} (xH_n(x)^2)'e^{-x^2}dx =$$

$$= \frac{1}{2}\int_{\mathbb{R}} H_n(x)^2e^{-x^2}dx + \int_{\mathbb{R}} e^{-x^2}xH_n(x)H_n(x)'dx =$$

$$= \frac{1}{2} + \frac{1}{2}\int_{\mathbb{R}} e^{-x^2}(H_n(x)')^2dx + \frac{1}{2}\int_{\mathbb{R}} e^{-x^2}(H_n(x)H_n(x)''dx =$$

$$= \frac{1}{2} + \frac{1}{2}2n \int_{\mathbb{R}} e^{-x^2} H_{n-1}(x)^2 dx + n(n-1) \int_{\mathbb{R}} e^{-x^2} H_n(x) H_{n-2}(x) dx = n + \frac{1}{2}. \qquad (7.13)$$

Consider the classical observable $h(p,q) = \frac{1}{2}(p^2 + \omega^2 q^2)$(total energy) and the classical density

$$\rho(p,q) = |n\rangle^2(q)\delta(p - \omega q). \qquad (7.14)$$

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

$$\langle q|\mu_\rho\rangle = \int_{\mathbb{R}}\int_{\mathbb{R}} q\big|n\rangle(q)\big|^2 \delta(p - q\omega)dpdq = \int_{\mathbb{R}} q\big|n\rangle\big|^2 dq = \int_{\mathbb{R}} x H_n(x)^2 e^{-x^2} dx = 0.$$

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 $1/m_1, \ldots, 1/m_n$ 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

$$\varepsilon_i = 1, i = 1, \ldots, N_1, \quad \varepsilon_i = -1, i = N_1 + 1, \ldots, N_1 + N_2, \quad \varepsilon_i = 0, i = N_1 + N_2 + 1, \ldots, N.$$

Assume that $N = N_1$. This can be treated as before. The normalized eigenfunctions of H are

$$|n_1, \ldots, n_N\rangle = \prod_{i=1}^{N} |n_i\rangle_i,$$

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 $|n_1, \ldots, n_N\rangle$ 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

$$\psi = C_1e^{\sqrt{-E}x} + C_2e^{-\sqrt{-E}x}.$$

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 $\int\limits_{\mathbb{R}} e^{ikx}e^{-ikx}dx = \int\limits_{\mathbb{R}} dx$ 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

$$\frac{1}{2\pi}\int\limits_{\mathbb{R}} xe^{ikx}e^{-ikx}dx = \frac{1}{2\pi}\int\limits_{\mathbb{R}} xdx = \infty.$$

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):

$$\psi(x) = \int\limits_{\mathbb{R}} \phi(k)\psi_k dk = \int\limits_{\mathbb{R}} \phi(k)e^{ikx}dk = \frac{1}{\sqrt{2\pi}}\int\limits_{\mathbb{R}} \hat{\psi}(k)e^{ikx}dk.$$

It evolves according to the law

$$\psi(x;t) = \int\limits_{\mathbb{R}} \hat{\psi}(k)e^{kx-tk^2}dk.$$

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

$$L_1 = i\hbar(x_3\frac{\partial}{\partial x_2} - x_3\frac{\partial}{\partial x_2}), \quad L_2 = i\hbar(x_1\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_1}), \quad L_3 = i\hbar(x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}),$$

$$L = \hbar^2[(\sum_{i=1}^3 x_i^2)\Delta - \sum_{i=1}^3 x_i^2\frac{\partial^2}{\partial x_i^2} - 2\sum_{1\le i<j\le 3} x_ix_j\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} - 2\sum_{i=1}^3 x_i\frac{\partial}{\partial x_i}]. \tag{8.6}$$

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

$$L_1 = i(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}), \quad L_2 = i(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}),$$

$$L_3 = -i\frac{\partial}{\partial\phi}, \quad L = L_1^2 + L_2^2 + L_3^2 = -\Delta_{S^2}. \tag{8.14}$$

These operators act on $L^2(S^2)$ leaving each subspace $\mathcal{H}^k$ invariant. Let

$$L_+ = L_1 + iL_2 = e^{i\phi}(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}), \quad L_- = L_1 - iL_2 = e^{i\phi}(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}).$$

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

$$\frac{\partial F_k(\theta)}{\partial\theta} - k\cot\theta F_K(\theta) = 0.$$

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

$$v(r) = r^\alpha(1 + \sum_{n=1}^{\infty} c_nr^n). \tag{8.22}$$

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

$$m_e = 9.11 \times 10^{-28}\mathrm{g}, \quad m_p = 1.67 \times 10^{-24}\mathrm{g},$$

$$e = 1.6 \times 10^{-9}C, \quad \hbar = 4.14 \times 10^{-15}\mathrm{eV}\cdot\mathrm{s},$$

$$E_n = -\frac{me^4}{2n^2\hbar^2} = -27.21\frac{Z^2}{2n^2}\mathrm{eV}. \tag{8.25}$$

In particular,

$$E_1 = -13.6\mathrm{eV}.$$

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

$$\rho(r) = 4\pi|\psi_{100}(\mathbf{x})|^2r^2 = 4e^{-2r}r^2.$$

The maximum of this function is reached at $r = 1$. In the restored physical units this translates to

$$r = \hbar^2/me^2 = .529 \times 10^{-8}\mathrm{cm}. \tag{8.26}$$

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 $SO(4)$.