Lectures on Quantum Mechanics

Leon A. Takhtajan

Department of Mathematics, Stony Brook University, Stony Brook, NY 11794-3651, USA
E-mail address: leontak@math.sunysb.edu

Contents

CHAPTER 1

Classical Mechanics

We use standard notations and basic facts from differential geometry. All manifolds, maps and functions are smooth (that is, C^∞) and real-valued, unless it is specified explicitly otherwise. Local coordinates $\mathbf{q} = (q_1, \dots, q_n)$ on a smooth n-dimensional manifold M at a point $q \in M$ are Cartesian coordinates on $\varphi(U) \subset \mathbb{R}^n$, where (U, φ) is a coordinate chart on M centered at $q \in U$. For $f : U \to \mathbb{R}^n$ we denote $(f \circ \varphi^{-1})(q_1, \dots, q_n)$ by $f(\mathbf{q})$. If U is a domain in $\mathbb{R}^n$ then for $f : U \to \mathbb{R}$ we denote by

$$\frac{\partial f}{\partial \mathbf{q}} = \left(\frac{\partial f}{\partial q_1}, \dots, \frac{\partial f}{\partial q_n} \right)$$

the gradient of a function f at a point $\mathbf{q} \in \mathbb{R}^n$ with Cartesian coordinates $(q_1, \dots, q_n)$. We denote by

$$\mathcal{A}^\bullet(M) = \bigoplus_{k=0}^{n} \mathcal{A}^k(M)$$

the graded algebra of smooth differential forms on M with respect to the wedge product, and by d the deRham differential — a graded derivation of $\mathcal{A}^\bullet(M)$ of degree 1 such that df is a differential of a function $f \in \mathcal{A}^0(M) = C^\infty(M)$. Let $\mathrm{Vect}(M)$ be the Lie algebra of smooth vector fields on M with the bracket $[\ ,\ ]$, given by a commutator of vector fields. For $X \in \mathrm{Vect}(M)$ we denote by $\mathcal{L}_X$ and i_X, respectively, the Lie derivative along X and the inner product with X. The Lie derivative is a degree 0 derivation of $\mathcal{A}^\bullet(M)$ which commutes with d and satisfies $\mathcal{L}_X(f) = X(f)$ for $f \in \mathcal{A}^0(M)$, and the inner product is a degree -1 derivation of $\mathcal{A}^\bullet(M)$ satisfying $i_X(f) = 0$ and $i_X(df) = X(f)$ for $f \in \mathcal{A}^0(M)$. They satisfy Cartan formulas

$$\begin{aligned} \mathcal{L}_X &= i_X \circ d + d \circ i_X = (d + i_X)^2, \\ i_{[X,Y]} &= \mathcal{L}_X \circ i_Y - i_Y \circ \mathcal{L}_X. \end{aligned}$$

For a smooth mapping of manifolds $f : M \to N$ we denote by $f_* : TM \to TN$ and $f^* : T^*N \to T^*M$, respectively, the induced mappings on tangent and cotangent bundles. Other notations, including those traditional for classical mechanics, will be introduced in the main text.

1. Lagrangian Mechanics

1.1. Generalized coordinates. Classical mechanics describes systems of finitely many interacting *particles*[1]. A system is called *closed* if its particles do not interact with the outside material bodies. The position of a system in space is specified by positions of its particles and defines a point in a smooth, finite-dimensional manifold M, the *configuration space* of a system. Coordinates on M are called *generalized coordinates* of a system, and the dimension $n = \dim M$ is called the number of *degrees of freedom*[2].

The *state* of a system at any instant of time is described by a point $q \in M$ and by a tangent vector $v \in T_qM$ at this point. The basic principle of classical mechanics is *Newton-Laplace determinacy principle* which asserts that a state of a system at a given instant completely determines its motion at all times t (in the future and in the past). The motion is described by the *classical trajectory* — a path $\gamma(t)$ in the configuration space M. In generalized coordinates $\gamma(t)$ is $(q_1(t), \dots, q_n(t))$ and corresponding derivatives $\dot{q}_i = \dfrac{dq_i}{dt}$ are called *generalized velocities*. Newton-Laplace principle is a fundamental experimental fact confirmed by our perception of everyday's experience. It implies that *generalized accelerations* $\ddot{q}_i = \dfrac{d^2q_i}{dt^2}$ are uniquely defined by generalized coordinates q_i and generalized velocities $\dot{q}_i$, so that classical trajectories satisfy a system of second order ordinary differential equations, called *equations of motion*. In the next section we formulate the most general principle governing the motion of mechanical systems.

1.2. The principle of the least action. In Lagrangian mechanics, a mechanical system with a configuration space M is completely characterized by its *Lagrangian* L — a smooth, real-valued function on $TM \times \mathbb{R}$ — the direct product of a tangent bundle TM of M and the time axis[3]. The motion of a Lagrangian system (M, L) is described by the *principle of the least action* (or *Hamilton's principle*), formulated as follows.

Let

$$PM(q_0, t_0, q_1, t_1) = \{\gamma : [t_0, t_1] \to M,\ \gamma(t_0) = q_0,\ \gamma(t_1) = q_1\}$$

be the space of smooth parametrized paths in M connecting points q_0 and q_1. The path space $PM = PM(q_0, t_0, q_1, t_1)$ is a infinite-dimensional real Fréchet manifold, and the tangent space $T_\gamma PM$ to PM at $\gamma \in PM$ consists of all smooth vector fields along the path γ in M which vanish at the endpoints q_0 and q_1. A smooth path Γ in PM, passing through $\gamma \in PM$ is

[1] A particle is a material body whose dimensions may be neglected in describing its motion.

[2] Systems with infinitely many degrees of freedom are described by classical field theory.

[3] It follows from Newton-Laplace principle that L could depend only on generalized coordinates and velocities, and on time.

called a *variation with fixed ends* of the path $\gamma(t)$ in M. A variation Γ is a family $\gamma_\varepsilon(t) = \Gamma(t, \varepsilon)$ of paths in M given by a smooth map

$$\Gamma : [t_0, t_1] \times [-\varepsilon_0, \varepsilon_0] \to M$$

such that $\Gamma(t, 0) = \gamma(t)$ for $t_0 \leq t \leq t_1$ and $\Gamma(t_0, \varepsilon) = q_0, \Gamma(t_1, \varepsilon) = q_1$ for $-\varepsilon_0 \leq \varepsilon \leq \varepsilon_0$. The tangent vector

$$\delta\gamma = \left.\frac{\partial \Gamma}{\partial \varepsilon}\right|_{\varepsilon=0} \in T_\gamma PM$$

corresponding to a variation $\gamma_\varepsilon(t)$ is traditionally called an *infinitesimal variation*. Explicitly,

$$\delta\gamma(t) = \Gamma_*(\tfrac{\partial}{\partial \varepsilon})(t, 0) \in T_{\gamma(t)}M, \;\; t_0 \leq t \leq t_1,$$

where $\frac{\partial}{\partial \varepsilon}$ is a tangent vector to the interval $[-\varepsilon_0, \varepsilon_0]$ at 0. Finally, a tangential lift of a path $\gamma(t)$ in M is the path $\gamma'(t)$ in TM defined by $\gamma'(t) = \gamma_*(\frac{\partial}{\partial t}) \in T_{\gamma(t)}M, \;\; t_0 \leq t \leq t_1$, where $\frac{\partial}{\partial t}$ is a tangent vector to $[t_0, t_1]$ at t. In other words, $\gamma'(t)$is the velocity vector of a path $\gamma(t)$ at time t.

Definition. The *action functional* $S : PM \to \mathbb{R}$ of a Lagrangian system (M, L) is defined by

$$S(\gamma) = \int_{t_0}^{t_1} L(\gamma'(t), t)dt.$$

Principle Of The Least Action (Hamilton's Prinicple). A path $\gamma \in PM$ describes the motion of a Lagrangian system (M, L) between the position $q_0 \in M$ at time t_0 and the position $q_1 \in M$ at time t_1 if and only if it is a critical point of the action functional S,

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\gamma_\varepsilon) = 0$$

for all variations $\gamma_\varepsilon(t)$ of $\gamma(t)$ with fixed ends.

The critical points of the action functional are called *extremals* and the principle of the least action states that a Lagrangian system (M, L) moves along the extremals[4]. The extremals are characterised by equations of motion — a system of second order differential equations in local coordinates on TM. For given local coordinates on M equations of motion have the most elegant form for the following choice of local coordinates on TM.

Definition. Let (U, φ) be a coordinate chart on M with local coordinates $\mathbf{q} = (q_1, \ldots, q_n)$. Coordinates

$$(\mathbf{q}, \mathbf{v}) = (q_1, \ldots, q_n, v_1, \ldots, v_n)$$

[4]The principle of the least action does not state that an extremal connecting points q_0 and q_1 is a minimum of S, nor that such an extremal is unique. It also does not state that any two points can be connected by an extremal.

on a chart TU on TM, where $\mathbf{v} = (v_1, \dots, v_n)$ are coordinates in the fibre corresponding to the basis $\dfrac{\partial}{\partial q_1}, \dots, \dfrac{\partial}{\partial q_n}$ for T_qM, are called *standard coordinates*.

Standard coordinates are Cartesian coordinates on $\varphi_*(TU) \subset T\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n$ and have the property that for $(q, v) \in TU$ and $f \in C^\infty(U)$,

$$v(f) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial q_i}(q) = \mathbf{v}\frac{\partial f}{\partial \mathbf{q}}.$$

The tangential lift $\gamma'(t)$ of a path $\gamma(t)$ in M in standard coordinates on TU is $(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = (q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t))$, where dot stands for the time derivative, so that

$$L(\gamma'(t), t) = L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t).$$

Following a centuries long tradition[5], we will denote standard coordinates by

$$(\mathbf{q}, \dot{\mathbf{q}}) = (q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n),$$

where the dot *does not* stand for a time derivative. Since we only consider paths in TM that are tangential lifts of paths in M, there will be no confusion[6] .

THEOREM 1.1. *The equations of motion of a Lagrangian system (M, L) in standard coordinates on TM are given by the Euler-Lagrange equations*

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

PROOF. Suppose first that an extremal $\gamma(t)$ lies in a coordinate chart U of M. Then a simple computation in standard coordinates, using integration by parts, gives

$$\begin{aligned}
0 &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\gamma_\varepsilon) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_{t_0}^{t_1} L\left(\mathbf{q}(t,\varepsilon), \dot{\mathbf{q}}(t,\varepsilon), t\right) dt \\
&= \sum_{i=1}^{n} \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\right) dt \\
&= \sum_{i=1}^{n} \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right) \delta q_i dt + \sum_{i=1}^{n} \left.\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right|_{t_0}^{t_1}.
\end{aligned}$$

The second sum in the last line vanishes due to the property $\delta q_i(t_0) = \delta q_i(t_1) = 0,\ i = 1, \dots, n$. The first sum is zero for arbitrary smooth functions

[5]Used in all texts on classical mechanics and theoretical physics.

[6]We reserve notation $(\mathbf{q}(t), \mathbf{v}(t))$ for general paths in TM.

δq_i on the interval $[t_0, t_1]$ which vanish at the endpoints. This implies that for each term in the sum the integrand is identically zero,

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

Since the restriction of an extremal of the action functional S to a coordinate chart on M is again an extremal, each extremal in standard coordinates on TM satisfies Euler-Lagrange equations. □

REMARK. In calculus of variations, the directional derivative of a functional S with respect to a tangent vector $V \in T_\gamma PM$ — the *Gato derivative*, is defined by

$$\delta_V S = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\gamma_\varepsilon),$$

where γ_ε is a path in PM with a tangent vector V at $\gamma_0 = \gamma$. The result of the above computation (when γ lies in a coordinate chart $U \subset M$) can be written as

$$\begin{aligned} \delta_V S &= \int_{t_0}^{t_1} \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\mathbf{q}(t), \dot{\mathbf{q}}(t), t) v_i(t) dt \\ &= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)\mathbf{v}(t) dt. \end{aligned} \tag{1.1}$$

Here $V(t) = \sum_{i=1}^{n} v_i(t)\frac{\partial}{\partial q_i}$ is a vector field along the path γ in M. Formula (1.1) is called formula for the first variation of the action with fixed ends. The principle of the least action is the statement that $\delta_V S(\gamma) = 0$ for all $V \in T_\gamma PM$.

REMARK. It is also convenient to consider the space $\widehat{PM} = \{\gamma : [t_0, t_1] \to M\}$ of all smooth parametrized paths in M. The tangent space $T_\gamma \widehat{PM}$ to $\widehat{PM}$ at $\gamma \in \widehat{PM}$ is the space of all smooth vector fields along the path γ in M (no conditions at the endpoints). The computation in the proof of Theorem 1.1 yields the following formula for the first variation of the action with free ends

$$\delta_V S = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\mathbf{v}\, dt + \left.\frac{\partial L}{\partial \dot{\mathbf{q}}}\mathbf{v}\right|_{t_0}^{t_1}. \tag{1.2}$$

PROBLEM 1.1. Show that the action functional is given by the evaluation of a 1-form Ldt on $TM \times \mathbb{R}$ over a 1-chain $\tilde{\gamma}$ on $TM \times \mathbb{R}$,

$$S(\gamma) = \int_{\tilde{\gamma}} Ldt,$$

where $\tilde{\gamma} = \{(\gamma'(t), t); t_0 \le t \le t_1\}$ and $Ldt\left(w, c\frac{\partial}{\partial t}\right) = cL(q, v)$, $w \in T_{(q,v)}TM$, $c \in \mathbb{R}$.

PROBLEM 1.2. Let $f \in C^\infty(M)$. Show that Lagrangians systems (M, L) and $(M, L + df)$ (where df is a fibre-wise linear function on TM) have the same equations of motion.

PROBLEM 1.3. Give examples of Lagrangian systems such that an extremal connecting two given points (i) is not a local minimum; (ii) is not unique; (iii) does not exist.

PROBLEM 1.4. For γ an extremal of the action functional S, the *second variation* of S is defined by

$$\delta^2_{V_1V_2}S = \frac{\partial^2}{\partial\varepsilon_1\partial\varepsilon_2}\bigg|_{\varepsilon_1=\varepsilon_2=0} S(\gamma_{\varepsilon_1,\varepsilon_2}),$$

where $\gamma_{\varepsilon_1,\varepsilon_2}$ is a smooth two-parameter family of paths in M such that the paths $\gamma_{\varepsilon_1,0}$ and γ_{0,ε_2} in PM at the point $\gamma_{0,0} = \gamma \in PM$ have tangent vectors V_1 and V_2. For a Lagrangian system (M, L) find the second variation of S and verify that it does not depend on the choice of $\gamma_{\varepsilon_1,\varepsilon_2}$.

1.3. Examples of Lagrangian systems. To describe mechanical phenomena it is necessary to choose a *frame of reference*. The properties of the *space-time* where the motion takes place depend on this choice. The space-time is characterized by the following postulates[7].

NEWTONIAN SPACE-TIME. The space is a three-dimensional affine Euclidean space E^3. A choice of the *origin* $0 \in E^3$ — a *reference point*, establishes the isomorphism $E^3 \simeq \mathbb{R}^3$, where the vector space $\mathbb{R}^3$ carries Euclidean inner product and has a fixed orientation. The time is one-dimensional — a time axis $\mathbb{R}$, and the space-time is a direct product $E^3 \times \mathbb{R}$. An *inertial* reference frame is a coordinate system with respect to the origin $0 \in E^3$, initial time t_0, and an orthonormal basis in $\mathbb{R}^3$. In inertial frame the space is *homogeneous* and *isotropic* and the time is *homogeneous*. The laws of motion are invariant with respect to the transformations

$$\mathbf{r} \mapsto g \cdot \mathbf{r} + \mathbf{r}_0, \quad t \mapsto t + t_0,$$

where $\mathbf{r}, \mathbf{r}_0 \in \mathbb{R}^3$ and g is an orthogonal linear transformation in $\mathbb{R}^3$. The time in classical mechanics is *absolute*.

The Galilean group is the group of all affine transformations of $E^3 \times \mathbb{R}$ which preserve time intervals and which for every $t \in \mathbb{R}$ are isometries in E^3. Every Galilean transformation is a composition of rotation, space-time translation, and a transformation

$$\mathbf{r} \mapsto \mathbf{r} + \mathbf{v}t, \quad t \mapsto t,$$

where $\mathbf{v} \in \mathbb{R}^3$. Any two inertial frames are related by a Galilean transformation.

GALILEO'S RELATIVITY PRINCIPLE. The laws of motion are invariant with respect to the Galilean group.

[7]Strictly speaking, these postulates are valid only in the non-relativistic limit of special relativity, when the speed of light in the vacuum is assumed to be infinite.

These postulate impose restrictions on Lagrangians of mechanical systems. Thus it follows from the first postulate that the Lagrangian L of a closed system does not explicitly depend on time. Physical systems are described by special Lagrangians, in agreement with the experimental facts about the motion of material bodies.

EXAMPLE 1.1 (Free particle). Configuration space for a free particle is $M = \mathbb{R}^3$, and it can deduced from Galileo's relativity principle that the Lagrangian for a free particle is

$$L = \tfrac{1}{2} m \dot{\mathbf{r}}^2.$$

Here $m > 0$ is the mass of a particle and $\dot{\mathbf{r}}^2 = |\dot{\mathbf{r}}|^2$ is the length square of the velocity vector $\dot{\mathbf{r}} \in T_{\mathbf{r}}\mathbb{R}^3 \simeq \mathbb{R}^3$. Euler-Lagrange equation gives *Newton's law of inertia*,

$$\ddot{\mathbf{r}} = 0.$$

EXAMPLE 1.2 (Interacting particles). Closed system of N interacting particles in $\mathbb{R}^3$ with masses $m_1, \ldots, m_N$, is described by a configuration space

$$M = \mathbb{R}^{3N} = \underbrace{\mathbb{R}^3 \times \cdots \times \mathbb{R}^3}_{N}$$

with a position vector $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$, where $\mathbf{r}_a \in \mathbb{R}^3$ is the position vector of a-th particle, $a = 1, \ldots, N$. It is found that the Lagrangian is given by

$$L = \sum_{a=1}^{N} \tfrac{1}{2} m_a \dot{\mathbf{r}}_a^2 - U(\mathbf{r}) = T - U,$$

where

$$T = \sum_{a=1}^{N} \tfrac{1}{2} m_a \dot{\mathbf{r}}_a^2$$

is called the *kinetic energy* of a system and $U(\mathbf{r})$ — the *potential energy*. The Euler-Lagrange equations give *Newton's equations*

$$m_a \ddot{\mathbf{r}}_a = \mathbf{F}_a,$$

where

$$\mathbf{F}_a = -\frac{\partial U}{\partial \mathbf{r}_a}$$

is a *force* on a-th particle, $a = 1, \ldots, N$. Forces of this form are called *conservative*. It follows from homogeneity of space that potential energy $U(\mathbf{r})$ of a closed system of N interacting particles with conservative forces depends only on relative positions of the particles, which leads to the equation

$$\sum_{a=1}^{N} \mathbf{F}_a = 0.$$

In particular, for a closed system of two particles $\mathbf{F}_1 + \mathbf{F}_2 = 0$, which is the equality of action and reaction forces, also called the *Newton's third law*.

The potential energy of a closed system with only pair-wise interaction between the particles has the form

$$U(\mathbf{r}) = \sum_{1\le a<b\le N} U_{ab}(\mathbf{r}_a - \mathbf{r}_b).$$

It follows from the isotropy of space that $U(\mathbf{r})$ depends only on relative distances between the particles, so that the Lagrangian of a closed system of N particles with pair-wise interaction has the form

$$L = \sum_{a=1}^{N} \tfrac{1}{2} m_a \dot{\mathbf{r}}_a^2 \; - \sum_{1\le a<b\le N} U_{ab}(|\mathbf{r}_a - \mathbf{r}_b|).$$

EXAMPLE 1.3 (Universal gravitation). According to the *Newton's law of gravitation*, the potential energy of the gravitational force between two particles with masses m_a and m_b is

$$U(\mathbf{r}_a - \mathbf{r}_b) = -G\frac{m_a m_a}{|\mathbf{r}_a - \mathbf{r}_b|},$$

where G is the gravitational constant. The configuration space of N particles with gravitational interaction is

$$M = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N} \,|\, \mathbf{r}_a \neq \mathbf{r}_b \text{ for } a \neq b,\, a, b = 1, \dots, N\}.$$

EXAMPLE 1.4 (Particle in an external potential field). Here $M = \mathbb{R}^3$ and

$$L = \tfrac{1}{2} m \dot{\mathbf{r}}^2 - U(\mathbf{r}, t),$$

where potential energy can explicitly depend on time. Equations of motion are Newton's equations

$$m\ddot{\mathbf{r}} = \mathbf{F} = -\frac{\partial U}{\partial \mathbf{r}}.$$

If $U = U(|\mathbf{r}|)$ — a function only of the distance $|\mathbf{r}|$, the potential field is called *central*.

EXAMPLE 1.5 (Charged particle in an electromagnetic field[8]). Consider a particle of charge e and mass m in a time-independent electromagnetic field with scalar and vector potentials $\varphi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r}) = (A_1(\mathbf{r}), A_2(\mathbf{r}), A_3(\mathbf{r}))$. The Lagrangian has the form

$$L = \frac{m\dot{\mathbf{r}}^2}{2} + e\left(\frac{\dot{\mathbf{r}}\,\mathbf{A}}{c} - \varphi\right),$$

where c is the speed of light. Corresponding Euler-Lagarange equation is Newton equation with the *Lorentz force*,

$$m\ddot{\mathbf{r}} = e\left(\mathbf{E} + \frac{\mathbf{r}}{c} \wedge \mathbf{H}\right),$$

[8]This is a non-relativistic limit of an example in classical electrodynamics.

where

$$\mathbf{E} = -\frac{\partial \varphi}{\partial \mathbf{r}} \quad \text{and} \quad \mathbf{H} = \operatorname{rot} \mathbf{A}$$

are electric and magnetic fields.

EXAMPLE 1.6 (Free particle on a Riemannian manifold). Let (M, ds^2) be a Riemannian manifold with the Riemannian metric ds^2. In local coordinates $x^1, \dots, x^n$ on M,

$$ds^2 = g_{\mu\nu}(x) dx^\mu \otimes dx^\nu,$$

where following tradition we are assuming summation over repeated indices. The Lagrangian of a free particle on M is

$$L(v) = \tfrac{1}{2}\langle v, v\rangle = \tfrac{1}{2}\|v\|^2,\ v \in T_x M,$$

where $\langle\ ,\ \rangle$ stands for the inner product in the fibres of TM. Corresponding functional S is called the action functional in Riemannian geometry. The Euler-Lagrange equations are

$$g_{\mu\nu}\ddot{x}^\mu + \frac{\partial g_{\mu\nu}}{\partial x^\lambda}\dot{x}^\mu \dot{x}^\lambda = \frac{1}{2}\frac{\partial g_{\mu\lambda}}{\partial x^\nu}\dot{x}^\mu \dot{x}^\lambda,$$

and after multiplying by the inverse metric tensor $g^{\sigma\nu}$ and summation over ν they take the form

$$\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0, \quad \sigma = 1, \dots, n,$$

where

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

are Christofel's symbols. The Euler-Lagrange equations of a free particle moving on a Riemannian manifold are geodesic equations.

Let ∇ be the Levi-Civita connection — the metric connection in the tangent bundle TM, and for $\xi \in \operatorname{Vect}(M)$ let ∇_ξ be a covariant derivative with respect to the vector field ξ. Explicitly,

$$(\nabla_\xi\, \eta)^\mu = \left(\frac{\partial \eta^\mu}{\partial x^\nu} + \Gamma^\mu_{\nu\lambda}\eta^\lambda\right)\xi^\nu, \quad \text{where} \quad \xi = \xi^\mu(x)\frac{\partial}{\partial x^\mu},\ \eta = \eta^\mu(x)\frac{\partial}{\partial x^\mu}.$$

For a path $\gamma = x(t)$ in M denote by $\nabla_{\dot{x}}$ the covariant derivative along γ,

$$(\nabla_{\dot{x}}\eta)^\mu(t) = \frac{d\eta^\mu(t)}{dt} + \Gamma^\mu_{\nu\lambda}(x(t))\dot{x}^\nu(t)\eta^\lambda(t), \quad \text{where} \quad \eta = \eta^\mu(t)\frac{\partial}{\partial x^\mu}$$

is a vector field along γ. The formula (1.1) can now be written in an invariant form

$$\delta S = -\int_{t_0}^{t_1} \langle \nabla_{\dot{x}}, \delta x\rangle dt,$$

which is known as the formula for the first variation of the action in Riemannian geometry.

EXAMPLE 1.7 (The rigid body). The configuration space of a rigid body in $\mathbb{R}^3$ with a fixed point is a Lie group $G = \mathrm{SO}(3)$ of orientation preserving orthogonal linear transformations in $\mathbb{R}^3$. Every left-invariant Riemannian metric $\langle\ ,\ \rangle_g$ on G defines a Lagrangian $L : TG \to \mathbb{R}$ by

$$L(v) = \tfrac{1}{2}\langle v, v\rangle_g, \quad v \in T_gG.$$

According to the previous example, equations of motion of a rigid body are geodesic equations on G. Let $\mathfrak{g} = \mathrm{so}(3)$ be the Lie algebra of G. A velocity vector $\dot g \in T_gG$ defines the *angular velocity of the body* by $\Omega = (L_{g^{-1}})_*\dot g \in \mathfrak{g}$, where $L_g : G \to G$ are left translations on G. In terms of the angular velocitiy, the Lagrangian takes the form

$$L = \tfrac{1}{2}\langle \Omega, \Omega\rangle,$$

where $\langle\ ,\ \rangle$ is an inner product on $\mathfrak{g} = T_eG$ given by the Riemannian metric. The Lie algebra $\mathfrak{g}$ — the Lie algebra of 3×3 skew-symmetric matrices, has the invariant inner product $\langle u, v\rangle_0 = -\frac{1}{2}\operatorname{Tr} uv$ (the Killing form), so that $\langle \Omega, \Omega\rangle = \langle \mathbf{A}\cdot\Omega, \Omega\rangle_0$. A symmetric linear operator $\mathbf{A} : \mathfrak{g} \to \mathfrak{g}$, positive-definite with respect to the Killing form, is called the *inertia tensor* of the body. The *principal axes of inertia* of the body are orthonormal eigenvectors e_1, e_2, e_3 of $\mathbf{A}$; corresponding eigenvalues I_1, I_2, I_3 are called *principal moments of inertia*. Setting $\Omega = \Omega_1 e_1 + \Omega_2 e_2 + \Omega_3 e_3$[9], we get

$$L = \tfrac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2).$$

In this parametrization, the Euler-Langrange equations become the *Euler's equations*

$$\begin{aligned} I_1\dot\Omega_1 &= (I_2 - I_3)\Omega_2\Omega_3,\\ I_2\dot\Omega_2 &= (I_3 - I_1)\Omega_1\Omega_3,\\ I_3\dot\Omega_3 &= (I_1 - I_2)\Omega_1\Omega_2. \end{aligned}$$

The Euler's equations describe the rotation of a free rigid body around a fixed point with the principal moments of inertia I_1, I_2, I_3 in the system of coordinates whose axes are the principal axes of inertia.

PROBLEM 1.5. Show that the Euler-Lagrange equations for the Lagrangian $L(v) = \|v\|$, $v \in T_qM$ coincide with the geodesic equations written with respect to a constant multiple of the natural parameter.

PROBLEM 1.6. Prove that the second variation of the action functional in Riemannian geometry, defined in Problem 1.4, is given by

$$\delta^2 S = -\int_{t_0}^{t_1} \langle \mathcal{J}\delta_1 x, \delta_2 x\rangle dt.$$

Here $\delta_1 x, \delta_2 x \in T_\gamma PM$ and $\mathcal{J} = \nabla^2_{\dot x} + R(\dot x,\ )\dot x$ is second-order linear differential operator acting on vector fields along the extremal γ, called the *Jacobi operator*, and R is a curvature operator — a fibre-wise linear mapping $R : TM \otimes TM \to \mathrm{End}(TM)$ of vector bundles, defined by $R(\xi, \eta) = \nabla_\eta\nabla_\xi - \nabla_\xi\nabla_\eta + \nabla_{[\xi,\eta]} : TM \to TM$, where $\xi, \eta \in \mathrm{Vect}(M)$.

[9]This establishes the isomorphism $\mathfrak{g} \simeq \mathbb{R}^3$ with the commutator given by the cross-product.

PROBLEM 1.7. Show that there exists a symmetric 3×3 matrix A such that $\mathbf{A} \cdot \Omega = A\Omega + \Omega A$, and find A for diagonal $\mathbf{A}$.

PROBLEM 1.8. Derive Euler's equations for a rigid body. (*Hint:* Use that $L = -\frac{1}{2} \operatorname{Tr} A\Omega^2$, where $\Omega = g^{-1}\dot{g}$, and derive the Euler-Lagrange equations in the matrix form $A\dot{\Omega} + \dot{\Omega}A = A\Omega^2 - \Omega^2 A$.)

1.4. Symmetries and Noether theorem. To describe the motion of a mechanical system one needs to solve corresponding Euler-Lagrange equations — a system of second order ordinary differential equations for the generalized coordinates. This could be a very difficult problem. Therefore of particular interest are those functions of generalized coordinates and velocities which remain constant during the motion.

DEFINITION. A smooth function $I : TM \to \mathbb{R}$ is called the *integral of motion* (*first integral*, or *conservation law*) for a Lagrangian system (M, L) if

$$\frac{d}{dt} I(\gamma'(t)) = 0$$

for all extremals γ of the action functional.

DEFINITION. The *energy* of a Lagrangian system (M, L) is a function E on $TM \times \mathbb{R}$ defined in standard coordinates on TM by

$$E(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}, t) - L(\mathbf{q}, \dot{\mathbf{q}}, t).$$

LEMMA 1.1. *The energy* $E = \dot{\mathbf{q}} \dfrac{\partial L}{\partial \dot{\mathbf{q}}} - L$ *is a well-defined function on* $TM \times \mathbb{R}$.

PROOF. Let (U, φ) and (U', φ') be coordinate charts on M with the transition function $f = \varphi' \circ \varphi^{-1} : \varphi(U \cap U') \to \varphi'(U \cap U')$. We have $\mathbf{q}' = f(\mathbf{q})$ and

$$dq'_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial q_j} dq_j, \quad i = 1, \dots, n,$$

or $d\mathbf{q}' = f_*(\mathbf{q}) d\mathbf{q}$, where $f_*(\mathbf{q}) = \left\{ \frac{\partial f_i}{\partial q_j} \right\}_{i,j=1}^{n}$ is a matrix-valued function on $\varphi(U \cap U')$. By the definition of standard coordinates $\dot{\mathbf{q}}' = f_*(\mathbf{q})\dot{\mathbf{q}}$, so that $\dot{\mathbf{q}} = (\dot{q}_1, \dots, \dot{q}_n)$ transform like components of a tangent vector on M. Therefore $d\dot{\mathbf{q}}' = g(\mathbf{q}, \dot{\mathbf{q}}) d\mathbf{q} + f_*(\mathbf{q}) d\dot{\mathbf{q}}$ for some matrix-valued function $g(\mathbf{q}, \dot{\mathbf{q}})$, and we compute

$$\begin{aligned} dL =& \frac{\partial L}{\partial \mathbf{q}'} d\mathbf{q}' + \frac{\partial L}{\partial \dot{\mathbf{q}}'} d\dot{\mathbf{q}}' \\ =& \left(\frac{\partial L}{\partial \mathbf{q}'} f_*(\mathbf{q}) + \frac{\partial L}{\partial \dot{\mathbf{q}}'} g(\mathbf{q}, \dot{\mathbf{q}}) \right) d\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}'} f_*(\mathbf{q}) d\dot{\mathbf{q}} \\ =& \frac{\partial L}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}} d\dot{\mathbf{q}}. \end{aligned}$$

Thus under the change of variables $\mathbf{q}' = f(\mathbf{q}), \dot{\mathbf{q}}' = f_*(\mathbf{q})\dot{\mathbf{q}}$,

$$\frac{\partial L}{\partial \dot{\mathbf{q}}'} f_*(\mathbf{q}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \quad \text{and} \quad \dot{\mathbf{q}}' \frac{\partial L}{\partial \dot{\mathbf{q}}'} = \dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}},$$

so that E is well-defined. □

COROLLARY 1.2. *Under a change of coordinates on M, components of $\dfrac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}) = \left(\dfrac{\partial L}{\partial \dot{q}_1}, \ldots, \dfrac{\partial L}{\partial \dot{q}_n} \right)$ transform like components of a 1-form on M.*

PROPOSITION 1.1 (Conservation of energy). *The energy of a closed system is an integral of motion.*

PROOF. For an extremal γ set $E(t) = E(\gamma(t))$. We have, according to the Euler-Lagrange equations,

$$\begin{aligned} \frac{dE}{dt} =& \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\dot{\mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}} - \frac{\partial L}{\partial \mathbf{q}}\dot{\mathbf{q}} - \frac{\partial L}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}} - \frac{\partial L}{\partial t} \\ =& \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial L}{\partial \mathbf{q}}\right)\dot{\mathbf{q}} - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t}. \end{aligned}$$

Since for a closed system $\dfrac{\partial L}{\partial t} = 0$, the energy is conserved. □

Conservation of energy for a closed mechanical system is a fundamental law of physics which follows from the homogeneity of time. For a general closed system of N interacting particles considered in Example 1.2,

$$E = \sum_{a=1}^{N} m_a \dot{\mathbf{r}}_a^2 - L = \sum_{a=1}^{N} \tfrac{1}{2} m_a \dot{\mathbf{r}}_a^2 + U(\mathbf{r}).$$

In other words, the total energy $E = T + U$ is a sum of the kinetic energy and the potential energy.

DEFINITION. A Lagrangian $L : TM \to \mathbb{R}$ is invariant with respect to the diffeomorphism $h : M \to M$, if $L(h_*(v)) = L(v)$ for all $v \in TM$. The diffeomorphism h is called the *symmetry* of a closed Lagrangian system (M, L).

Continuous symmetries give rise to conservation laws.

THEOREM 1.3 (Noether). *Suppose that a Lagrangian $L : TM \to \mathbb{R}$ is invariant under a one-parameter group $\{h^s\}_{s\in\mathbb{R}}$ of diffeomorphisms of M. Then Lagrangian system (M, L) admits an integral of motion I, given in standard coordinates on TM by*

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) \left(\left. \frac{dh_i^s(\mathbf{q})}{ds} \right|_{s=0} \right) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \mathbf{q}'.$$

The integral of motion I is called the Noether integral.

PROOF. Since

$$\left(\frac{dh_1^s(\mathbf{q})}{ds}\bigg|_{s=0}, \dots, \frac{dh_n^s(\mathbf{q})}{ds}\bigg|_{s=0} \right) = \mathbf{q}'$$

are components of the vector field on M associated with the one-parameter group $\{h^s\}_{s\in\mathbb{R}}$, it follows from Corollary 1.2 that I is a well-defined function on TM. Now differentiating $L(h_*^s(v)) = L(v)$ with respect to s at $s = 0$ and using the Euler-Lagrange equations we get

$$0 = \frac{\partial L}{\partial \mathbf{q}}\mathbf{q}' + \frac{\partial L}{\partial \dot{\mathbf{q}}}\dot{\mathbf{q}}' = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\mathbf{q}' + \frac{\partial L}{\partial \dot{\mathbf{q}}}\frac{d\mathbf{q}'}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\mathbf{q}'\right).$$

□

REMARK. Noether theorem generalizes to time-dependent Lagrangians $L : TM \times \mathbb{R} \to \mathbb{R}$. Namely, on the *extended configuration space* $M_1 = M \times \mathbb{R}$ define a time-independent Lagrangian L_1 by

$$L_1(\mathbf{q}, \tau, \dot{\mathbf{q}}, \dot{\tau}) = L\left(\mathbf{q}, \frac{\dot{\mathbf{q}}}{\dot{\tau}}, \tau\right)\dot{\tau},$$

where $(\mathbf{q}, \tau)$ are local coordinates on M_1 and $(\mathbf{q}, \tau, \dot{\mathbf{q}}, \dot{\tau})$ are standard coordinates on TM_1. The Noether integral I_1 for a closed system (M_1, L_1) defines an integral of motion I for a system (M, L) by the formula

$$I(\mathbf{q}, \dot{\mathbf{q}}, t) = I_1(\mathbf{q}, t, \dot{\mathbf{q}}, 1).$$

When Lagrangian L does not depend on time, L_1 is invariant with respect to the one-parameter group of translations $\tau \mapsto \tau + s$, and the Noether integral $I_1 = \dfrac{\partial L_1}{\partial \dot{\tau}}\dot{\tau}$ gives $I = -E$.

Noether theorem can be generalized as follows. For $X \in \mathrm{Vect}(M)$ let X' be a vector field on TM defined by a local flow on TM induced from the corresponding local flow on M. In standard coordinates on TM,

$$X = \sum_{i=1}^n a_i(\mathbf{q})\frac{\partial}{\partial q_i} \quad \text{and} \quad X' = \sum_{i=1}^n a_i(\mathbf{q})\frac{\partial}{\partial q_i} + \sum_{i,j=1}^n \dot{q}^j \frac{\partial a_i}{\partial q_j}(\mathbf{q})\frac{\partial}{\partial \dot{q}_i}. \tag{1.3}$$

PROPOSITION 1.2. *Suppose that for the Lagrangian $L : TM \to \mathbb{R}$ there exist a vector field X on M and a function F on TM such that for every path γ in M*

$$dL(X')(\gamma(t)) = \frac{d}{dt}F(\gamma'(t)).$$

Then

$$I = \sum_{i=1}^n a_i(\mathbf{q})\,\frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) - F(\mathbf{q}, \dot{\mathbf{q}})$$

is an integral of motion for the Lagrangian system (M, L).

PROOF. Denoting $\mathbf{a}(t) = (a_1(\gamma(t)), \dots, a_n(\gamma(t))$ and using Euler-Lagrange equations, we have along the extremal γ,

$$\frac{d}{dt}\left(\mathbf{a}\,\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) = \mathbf{a}\,\frac{\partial L}{\partial \mathbf{q}} + \dot{\mathbf{a}}\,\frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{dF}{dt}.$$

□

EXAMPLE 1.8 (Conservation of momentum). Let $M = V$ — a vector space, and suppose that a Lagrangian L is invariant with respect to a one-parameter group $h^s(q) = q + sv$, $v \in V$. According to Noether's theorem,

$$I = \sum_{i=1}^{n} v_i \frac{\partial L}{\partial \dot{q}_i}$$

is an integral of motion. Now let (M, L) be a closed Lagrangian system of N interacting particles considered in Example 1.2. We have $M = V = \mathbb{R}^{3N}$ and Lagrangian L is invariant under simultaneous translation of coordinates $\mathbf{r}_a = (r_{a1}, r_{a2}, r_{a3})$ of all particles by the same vector $\mathbf{c} \in \mathbb{R}^3$. Thus $v = (\mathbf{c}, \dots, \mathbf{c}) \in \mathbb{R}^{3N}$ and for every $\mathbf{c} = (c_1, c_2, c_3) \in \mathbb{R}^3$,

$$I = \sum_{a=1}^{N} \left(c_1 \frac{\partial L}{\partial \dot{r}_{a1}} + c_2 \frac{\partial L}{\partial \dot{r}_{a2}} + c_3 \frac{\partial L}{\partial \dot{r}_{a3}} \right) = c_1 P^1 + c_2 P^2 + c_3 P^3$$

is an integral of motion. The integrals of motion P^1, P^2, P^3 define the vector

$$\mathbf{P} = \sum_{a=1}^{N} \frac{\partial L}{\partial \dot{\mathbf{r}}_a} \in \mathbb{R}^3$$

(or rather a vector in the dual space to $\mathbb{R}^3$), called the *momentum* of the system. Explicitly,

$$\mathbf{P} = \sum_{a=1}^{N} m_a \dot{\mathbf{r}}_a,$$

so that the total momentum of a closed system is the sum of momenta of individual particles. Conservation of momentum is a fundamental physical law which reflects the homogeneity of space.

In general, $p^i = \dfrac{\partial L}{\partial \dot{q}_i}$ are called *generalized momenta* corresponding to generalized coordinates q_i, and $F^i = \dfrac{\partial L}{\partial q_i}$ are called *generalized forces*. In these notations, the Euler-Lagrange equations have the same form

$$\dot{\mathbf{p}} = \mathbf{F}$$

as Newton's equations in Cartesian coordinates. Conservation of momentum implies Newton's third law.

EXAMPLE 1.9 (Conservation of angular momentum). Let $M = V$ be a vector space with Euclidean inner product. Let $G = \mathrm{SO}(V)$ be the connected Lie group of automorphisms of V preserving the inner product, and let

$\mathfrak{g} = \mathrm{so}(V)$ be the Lie algebra of G. Suppose that a Lagrangian L is invariant with respect to the action of a one-parameter subgroup $h^s(q) = e^{sx} \cdot q$ of G on V, where $x \in \mathfrak{g}$ and e^x is the exponential map. According to Noether's theorem,

$$I = \sum_{i=1}^{n} (x \cdot q)_i \, \frac{\partial L}{\partial \dot{q}_i}$$

is an integral of motion. Now let (M, L) be a closed Lagrangian system of N interacting particles considered in Example 1.2. We have $M = V = \mathbb{R}^{3N}$ and Lagrangian L is invariant under a simultaneous rotation of coordinates $\mathbf{r}_a$ of all particles by the same orthogonal transformation in $\mathbb{R}^3$. Thus $x = (u, \dots, u) \in \underbrace{\mathrm{so}(3) \oplus \cdots \oplus \mathrm{so}(3)}_{N}$, and for every $u \in \mathrm{so}(3)$

$$I = \sum_{a=1}^{N} \left((u \cdot \mathbf{r}_a)_1 \, \frac{\partial L}{\partial \dot{r}_{a1}} + (u \cdot \mathbf{r}_a)_2 \, \frac{\partial L}{\partial \dot{r}_{a2}} + (u \cdot \mathbf{r}_a)_3 \, \frac{\partial L}{\partial \dot{r}_{a3}} \right)$$

is an integral of motion. Using a basis in $\mathrm{so}(3) \simeq \mathbb{R}^3$ corresponding to the rotations with axes given by the vectors e_1, e_2, e_3 of the standard orthonormal basis in $\mathbb{R}^3$, we get the vector

$$\mathbf{M} = \sum_{a=1}^{N} \mathbf{r}_a \times \frac{\partial L}{\partial \dot{\mathbf{r}}_a} \in \mathbb{R}^3$$

(or rather a vector in the dual space to so(3)), called *angular momentum* of the system. Explicitly,

$$\mathbf{M} = \sum_{a=1}^{N} m_a \mathbf{r}_a \times \dot{\mathbf{r}}_a,$$

so that the total angular momentum of a closed system is the sum of angular momenta of individual particles. Conservation of angular momentum is a fundamental physical law which reflects the isotropy of space.

PROBLEM 1.9. Find how total momentum and total angular momentum transform under the Galilean transformations.

1.5. One-dimensional motion. The motion of systems with one degree of freedom is called one-dimensional. In terms of a Cartesian coordinate x on $M = \mathbb{R}$ the Lagrangian takes the form

$$L = \tfrac{1}{2} m \dot{x}^2 - U(x).$$

The conservation of energy

$$E = \frac{1}{2} m \dot{x}^2 + U(x)$$

allows to solve equation of motion in a closed form by formally solving

$$\frac{dx}{dt} = \sqrt{\frac{2}{m}(E - U(x))}$$

and integrating

$$t = \sqrt{\frac{m}{2}} \int \frac{dx}{\sqrt{E - U(x)}}.$$

This is the general solution of Newton's equation

$$m\ddot{x} = -\frac{dU}{dx}$$

with two arbitrary constants, the energy E and the constant of integration.

Since kinetic energy is non-negative, for a given value of E the actual motion takes place in the region of $\mathbb{R}$ where $U(x) \leq E$. The points where $U(x) = E$ are called *turning points*. The motion which is confined between two turning points is called *finite*. The finite motion is periodic — the particle oscillates between the turning points $x_1(E)$ and $x_2(E)$ with the period

$$T(E) = \sqrt{2m} \int_{x_1(E)}^{x_2(E)} \frac{dx}{\sqrt{E - U(x)}}.$$

If the region $U(x) \leq E$ is unbounded the motion is called *infinite*, the particle eventually goes to infinity.

On the phase plane with coordinates (x, y) the Newton's equation is given by a system

$$m\dot{x} = y, \quad \dot{y} = -\frac{dU}{dx}.$$

Trajectories correspond to the phase curves $(x(t), y(t))$, which lie on the level sets

$$\frac{y^2}{2m} + U(x) = E$$

of the energy function. The points $(x_0, 0)$, where is x_0 is a critical point of the potential energy $U(x)$, correspond to the equilibrium solutions. The local minima correspond to the stable solutions and local maxima — to the unstable solutions. For the values of E which do not correspond to the equilibrium solutions the level sets are smooth curves. These curves are closed if the motion is is finite.

The simplest example of a one-dimensional system is a *harmonic oscillator* with $U(x) = \frac{1}{2}kx^2$. For $k > 0$ the general solution of equations of motion is

$$x(t) = A\cos(\omega t + \alpha),$$

where A is the *amplitude*, $\omega = \sqrt{\frac{k}{m}}$ is the *circular frequency*, and α is the *phase* of a simple harmonic motion with the period $T = \frac{2\pi}{\omega}$. The energy is $E = \frac{1}{2}m\omega^2 A^2$ and the motion is finite for all $E \geq 0$ with the same period T for $E > 0$. For $k < 0$ the motion is infinite for all E.

PROBLEM 1.10. Show that for $U(x) = -x^4$ there are phase curves which do not exist for all times. Prove that if $U(x) \geq 0$ for all x than all phase curves exist for all times.

PROBLEM 1.11. The simple pendulum is a Lagrangian system with $M = S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and $L = \frac{1}{2}\dot{\theta}^2 + \cos\theta$. Find the period T of the pendulum as a function of the amplitude of the oscillations.

PROBLEM 1.12. Suppose that the potential energy $U(x)$ is even, $U(0) = 0$ and $U(x)$ is one-to-one monotonically increasing function for $x \geq 0$. Prove that the inverse function $x(U)$ and the period $T(E)$ are related by the Abel transform

$$T(E) = 2\sqrt{2m}\int_0^E \frac{dx}{dU}\frac{dU}{\sqrt{E-U}} \quad \text{and} \quad x(U) = \frac{1}{2\pi\sqrt{2m}}\int_0^U \frac{T(E)dE}{\sqrt{U-E}}.$$

1.6. The motion in a central field and the Kepler problem. The motion of a system of two interacting particles — the *two-body problem* — also can be solved completely. Namely, in this case (Example 1.2) $M = \mathbb{R}^6$ and

$$L = \frac{m_1\dot{\mathbf{r}}_1^2}{2} + \frac{m_2\dot{\mathbf{r}}_2^2}{2} - U(|\mathbf{r}_1 - \mathbf{r}_2|).$$

Introducing on $\mathbb{R}^6$ new coordinates

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \quad \text{and} \quad \mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2},$$

we get

$$L = \tfrac{1}{2}\mu\dot{\mathbf{R}}^2 + \tfrac{1}{2}m\dot{\mathbf{r}}^2 - U(|\mathbf{r}|),$$

where $\mu = m_1 + m_2$ is the *total mass* and $m = \dfrac{m_1 m_2}{m_1 + m_2}$ is the *reduced mass* of a two-body system. The Lagrangian L depends only on the velocity $\dot{\mathbf{R}}$ of the center of mass and not on its position $\mathbf{R}$. A generalized coordinate with this property is called *cyclic*. it follows from the Euler-Lagrange equations that generalized momentum corresponding to the cyclic coordinate is conserved. In our case it is a total momentum of the system,

$$\mathbf{P} = \frac{\partial L}{\partial \dot{\mathbf{R}}} = \mu\dot{\mathbf{R}},$$

so that the center of mass $\mathbf{R}$ moves uniformly. Thus in the frame of reference where $\mathbf{R} = 0$, the Lagrangian of a two-body problem is reduced to the Lagrangian of a single particle in the external central field $U(|\mathbf{r}|)$.

From conservation of angular momentum $\mathbf{M} = m\mathbf{r} \times \dot{\mathbf{r}}$ it follows that during the motion the position vector $\mathbf{r}$ lies in the plane P orthogonal to $\mathbf{M}$ in $\mathbb{R}^3$. Introducing polar coordinates (r, φ) in P we get[10]

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) - U(r).$$

The coordinate φ is cyclic and its generalized momentum $mr^2\dot{\varphi}$ coincides with $|\mathbf{M}|$ if $\dot{\varphi} > 0$ and with $-|\mathbf{M}|$ if $\dot{\varphi} < 0$. Denoting this quantity by M, we get the equation

$$mr^2\dot{\varphi} = \mathrm{M}, \tag{1.4}$$

[10]Note that here r is not the length of the position vector $\mathbf{r}$.

which is equivalent to the *Kepler's second law*[11]. Using (1.4) we get for the total energy

$$E = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) + U(r) = \tfrac{1}{2}m\dot{r}^2 + U(r) + \frac{\mathrm{M}^2}{2mr^2}. \tag{1.5}$$

Thus the radial motion reduces to a one-dimensional motion with the effective potential energy

$$U_{eff}(r) = U(r) + \frac{\mathrm{M}^2}{2mr^2},$$

where the second term is called the *centrifugal energy*. As in the previous section, is given by

$$t = \sqrt{\frac{m}{2}} \int \frac{dr}{\sqrt{E - U_{eff}(r)}}. \tag{1.6}$$

It follows from (1.4) that φ is monotonic function of t, given by another quadrature

$$\varphi = \frac{\mathrm{M}}{\sqrt{2m}} \int \frac{dr}{r^2\sqrt{E - U_{eff}(r)}}. \tag{1.7}$$

The set $U_{eff}(r) \le E$ is a union of annuli $0 \le r_{min} \le r \le r_{max} \le \infty$, and the motion is finite if $0 < r_{min} \le r \le r_{max} < \infty$. Though for a finite motion $r(t)$ oscillates between r_{min} and r_{max}, corresponding trajectories are not necessarily closed. The necessary and sufficient condition for a finite motion to have a close trajectory is that the angle

$$\Phi = \frac{\mathrm{M}}{\sqrt{2m}} \int_{r_{min}}^{r_{max}} \frac{dr}{r^2\sqrt{E - U_{eff}(r)}}$$

is commensurable with 2π, i.e., $\Phi = 2\pi\dfrac{m}{n}$ for some $m, n \in \mathbb{Z}$. If the angle Φ is not commensurable with 2π, the orbit is everywhere dense in the annulus $r_{min} \le r \le r_{max}$. If

$$\lim_{r\to\infty} U_{eff}(r) = \lim_{r\to\infty} U(r) = U < \infty,$$

the motion is inifnite for $E > U$ — the particle goes to ∞ with finite velocity $\sqrt{\frac{2}{m}(E - U)}$.

Very important special case is when

$$U(r) = -\frac{\alpha}{r}.$$

It describes Newton's gravitational attraction ($\alpha > 0$) and Coulomb electrostatic interaction (either attractive or repulsive). First consider the case when $\alpha > 0$ — the Kepler's problem. The effective potential energy is

$$U_{eff}(r) = -\frac{\alpha}{r} + \frac{\mathrm{M}^2}{2mr^2}$$

[11]It is the statement that *sectorial velocity* of a particle in a central field is constant.

and has the global minimum

$$U_0 = -\frac{\alpha^2 m}{2\mathrm{M}^2}$$

at $r_0 = \dfrac{\mathrm{M}^2}{\alpha m}$. The motion is infinite for $E \geq 0$ and is finite for $U_0 \leq E < 0$. The explicit form of trajectories can be determined by an elementary integration in (1.7), which gives

$$\varphi = \cos^{-1} \frac{\dfrac{\mathrm{M}}{r} - \dfrac{\mathrm{M}}{r_0}}{\sqrt{2m(E - U_0)}} + C.$$

Choosing a constant of integration $C = 0$ and introducing notations

$$p = r_0 \quad \text{and} \quad e = \sqrt{1 - \frac{E}{U_0}},$$

we get the equation of the orbit (trajectory)

$$\frac{p}{r} = 1 + e\cos\varphi. \tag{1.8}$$

This is the equation of a conic section with one focus at the origin. Quantity $2p$ is called the *latus rectum* of the orbit, and e is called the *eccentricity*. The choice $C = 0$ is such that the point with $\varphi = 0$ is the point nearest to the origin (called the *perihelion*). When $U_0 \leq E < 0$ the eccentricity $e < 1$ so that the orbit is the ellipse[12] with the major and minor semi-axes

$$a = \frac{p}{1 - e^2} = \frac{\alpha}{2|E|}, \quad b = \frac{p}{\sqrt{1 - e^2}} = \frac{|\mathrm{M}|}{\sqrt{2m|E|}}.$$

Correspondingly, $r_{min} = \dfrac{p}{1 + e}$, $r_{max} = \dfrac{p}{1 - e}$ and the period T of elliptic orbit is given by

$$T = \pi\alpha\sqrt{\frac{m}{2|E|^3}}.$$

The last formula is *Kepler's third law*. When $E > 0$ the eccentricity $e > 1$ and the motion is infinite — the orbit is a hyperbola with the origin as internal focus. When $E = 0$ the eccentricity $e = 1$ — the particle starts from rest at ∞ and the orbit is a parabola.

For the repulsive case $\alpha < 0$ the effective potential energy $U_{eff}(r)$ is always positive and decreases monotonically from ∞ to 0. The motion is always infinite and the trajectories are hyperbolas (parabola if $E = 0$)

$$\frac{p}{r} = -1 + e\cos\varphi$$

with

$$p = \frac{\mathrm{M}^2}{\alpha m} \quad \text{and} \quad e = \sqrt{1 + \frac{2E\mathrm{M}^2}{m\alpha^2}}.$$

[12]The statement that planets have elliptic orbits with a focus at the Sun is *Kepler's first law*.

The Kepler's problem is very special: for every $\alpha \in \mathbb{R}$ the Lagrangian system on $\mathbb{R}^3$ with

$$L = \tfrac{1}{2} m \dot{\mathbf{r}}^2 + \frac{\alpha}{r}$$

has three extra integrals of motion W_1, W_2, W_3 in addition to the components of the angular momentum $\mathbf{M}$. Corresponding vector $\mathbf{W} = (W_1, W_2, W_3)$, called *Laplace-Runge-Lenz vector*, is given by

$$\mathbf{W} = \dot{\mathbf{r}} \times \mathbf{M} - \frac{\alpha \mathbf{r}}{r}. \tag{1.9}$$

Indeed, using equations of motion $m\ddot{\mathbf{r}} = -\dfrac{\alpha \mathbf{r}}{r^3}$ and conservation of the angular momentum $\mathbf{M} = m\mathbf{r} \times \dot{\mathbf{r}}$, we get

$$\begin{aligned} \dot{\mathbf{W}} &= m\ddot{\mathbf{r}} \times (\mathbf{r} \times \dot{\mathbf{r}}) - \frac{\alpha \dot{\mathbf{r}}}{r} + \frac{\alpha (\dot{\mathbf{r}} \cdot \mathbf{r}) \mathbf{r}}{r^3} \\ &= (m\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}}) \mathbf{r} - (m\ddot{\mathbf{r}} \cdot \mathbf{r}) \dot{\mathbf{r}} - \frac{\alpha \dot{\mathbf{r}}}{r} + \frac{\alpha (\dot{\mathbf{r}} \cdot \mathbf{r}) \mathbf{r}}{r^3} \\ &= 0. \end{aligned}$$

The fact that all orbits are conic sections follows from this extra symmetry of the Kepler's problem.

PROBLEM 1.13. Prove all the statements made in this section.

PROBLEM 1.14. Show that if

$$\lim_{r \to 0} U_{eff}(r) = -\infty,$$

then there are orbits with $r_{min} = 0$ — "fall " of the particle to the center.

PROBLEM 1.15. Prove that all finite trajectories in central field are closed only when

$$U(r) = k r^2, \ \ k > 0 \quad \text{and} \quad U(r) = -\frac{\alpha}{r}, \ \ \alpha > 0.$$

PROBLEM 1.16. Find parametric equations for orbits in Kepler's problem.

PROBLEM 1.17. Prove that the Laplace-Runge-Lenz vector $\mathbf{W}$ points in the direction of the major axis of the orbit and that $|\mathbf{W}| = \alpha^2 e$, where e is the eccentricity of the orbit.

1.7. Legendre transformation. The equations of motion of a Lagrangian system (M, L) in standard coordinates on TM associated with a coordinate chart U on M are the Euler-Lagrange equations. In expanded form, they are given the following system of ordinary differential equations of second order,

$$\begin{aligned} \frac{\partial L}{\partial q_i}(\mathbf{q}, \dot{\mathbf{q}}) &= \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i}(\mathbf{q}, \dot{\mathbf{q}}) \right) \\ &= \sum_{j=1}^{n} \left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \, \ddot{q}_j + \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j}(\mathbf{q}, \dot{\mathbf{q}}) \, \dot{q}_j \right), \quad i = 1, \ldots, n. \end{aligned}$$

In order for this system to be solvable for the highest derivatives for all initial conditions in TU, the symmetric $n \times n$ matrix

$$H_L = \left\{ \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}(\mathbf{q}, \dot{\mathbf{q}}) \right\}_{i,j=1}^{n}$$

should be invertible on TU.

DEFINITION. A Lagrangian system (M, L) is called *non-degenerate* if for every coordinate chart U on M the matrix $H_L(q, v)$ is invertible at every point $(q, v) \in TU$.

For an invariant formulation, consider the 1-form θ_L on TM, defined in standard coordinates on TM associated with a coordinate chart U on M by

$$\theta_L = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} dq_i = \frac{\partial L}{\partial \dot{\mathbf{q}}} d\mathbf{q}.$$

It follows from Corollary 1.2 that θ_L is indeed a well-defined 1-form on TM.

LEMMA 1.2. *A Lagrangian system (M, L) is non-degenerate if and only if the 2-form $d\theta_L$ on TM is non-degenerate.*

PROOF. In standard coordinates,

$$d\theta_L = \sum_{i,j=1}^{n} \left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} d\dot{q}_j \wedge dq_i + \frac{\partial^2 L}{\partial \dot{q}_i \partial q_j} dq_j \wedge dq_i \right),$$

so that $2n \times 2n$ matrix corresponding to the 2-form $d\theta_L$ is non-degenerate if and only if the $n \times n$ matrix H_L is non-degenerate. □

DEFINITION. Let (U, φ) be a coordinate chart on M. Coordinates

$$(\mathbf{p}, \mathbf{q}) = (p^1, \dots, p^n, q_1, \dots, q_n)$$

on the chart $T^*U = \mathbb{R}^n \times U$ on the cotangent bundle T^*M are called *standard coordinates*[13] if for $(p, q) \in T^*U$ and $f \in C^\infty(U)$

$$p^i(df) = \frac{\partial f}{\partial q_i}, \quad i = 1, \dots, n.$$

Equivalently, standard coordinates on T^*U are uniquely characterized by the condition that $\mathbf{p} = (p^1, \dots, p^n)$ are coordinates in the fibre corresponding to the basis $dq_1, \dots, dq_n$ for T_q^*M dual to the basis $\dfrac{\partial}{\partial q_1}, \dots, \dfrac{\partial}{\partial q_n}$ for T_qM.

DEFINITION. The 1-form θ on T^*M, defined in standard coordinates by

$$\theta = \sum_{i=1}^{n} p^i dq_i = \mathbf{p} d\mathbf{q},$$

is called the *canonical Liouville 1-form*.

[13]Following tradition, the first n coordinates parametrize the fibre of T^*U and the last n coordinates — the base.

Corollary 1.2 shows that θ is a well-defined 1-form on T^*M. Invariantly, the 1-form θ is defined by $\theta(u) = p(\pi_*(u))$, where $u \in T_{(p,q)}T^*M$ and $\pi : T^*M \to M$ is the canonical projection.

DEFINITION. A fibre-wise mapping $\tau_L : TM \to T^*M$ is called a *Legendre's transformation* associated with the Lagrangian L, if

$$\theta_L = \tau_L^*(\theta).$$

In standard coordinates the Legendre's transformation is given by

$$\tau_L(\mathbf{q}, \dot{\mathbf{q}}) = (\mathbf{p}, \mathbf{q}), \quad \text{where} \quad \mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}).$$

The mapping τ_L is a local diffeomorphism if and only if Lagrangian L is non-degenerate.

DEFINITION. Suppose that the Legendre's transformation $\tau_L : TM \to T^*M$ is a diffeomorphism. The *Hamiltonian* $H : T^*M \to \mathbb{R}$, associated with the Lagrangian $L : TM \to \mathbb{R}$, is defined by

$$H \circ \tau_L = E = \dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} - L.$$

In standard coordinates,

$$H(\mathbf{p}, \mathbf{q}) = \left. (\mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}})) \right|_{\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}},$$

where $\dot{\mathbf{q}}$ is considered a function of $\mathbf{p}$ and $\mathbf{q}$ obtained from the equation $\mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}})$ by using the implicit function theorem. The cotangent bundle T^*M is called the *phase space* of the Lagrangian system (M, L). It turns out on the phase space the equations of motion take a very simple and symmetric form.

THEOREM 1.4. *Suppose that the Legendre transformation $\tau_L : TM \to T^*M$ is a diffeomorphism. Then the Euler-Lagrange equations in standard coordinates on TM,*

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

are equivalent to the following system of first order differential equations in standard coordinates on T^*M,

$$\dot{p}^i = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p^i}, \quad i = 1, \ldots, n.$$

PROOF. We have

$$\begin{aligned} dH &= \frac{\partial H}{\partial \mathbf{p}} d\mathbf{p} + \frac{\partial H}{\partial \mathbf{q}} d\mathbf{q} \\ &= \left(\mathbf{p} d\dot{\mathbf{q}} + \dot{\mathbf{q}} d\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}} d\mathbf{q} - \frac{\partial L}{\partial \dot{\mathbf{q}}} d\dot{\mathbf{q}} \right) \Big|_{\mathbf{p}=\frac{\partial L}{\partial \dot{\mathbf{q}}}} \\ &= \left(\dot{\mathbf{q}} d\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}} d\mathbf{q} \right) \Big|_{\mathbf{p}=\frac{\partial L}{\partial \dot{\mathbf{q}}}}. \end{aligned}$$

Thus under the Legendre transform,

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \quad \text{and} \quad \dot{\mathbf{p}} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{\partial L}{\partial \mathbf{q}} = -\frac{\partial H}{\partial \mathbf{q}}.$$

□

Corresponding first order differential equations on T^*M are called *Hamilton's equations* (*canonical equations*).

COROLLARY 1.5. *The Hamiltonian H is constant on the solutions of the Hamilton's equations.*

PROOF. For $H(t) = H(\mathbf{p}(t), \mathbf{q}(t))$ we have

$$\frac{dH}{dt} = \frac{\partial H}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial H}{\partial \mathbf{p}} \dot{\mathbf{p}} = \frac{\partial H}{\partial \mathbf{q}} \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial H}{\partial \mathbf{p}} \frac{\partial H}{\partial \mathbf{q}} = 0.$$

□

For the Lagrangian

$$L = \frac{m\dot{\mathbf{r}}^2}{2} - U(\mathbf{r}), \quad \mathbf{r} \in \mathbb{R}^3,$$

of a particle of mass m in a potential field $U(\mathbf{r})$, considered in Example 1.4, we have

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}}.$$

Thus the Legendre's transformation $\tau_L : T\mathbb{R}^3 \to T^*\mathbb{R}^3$ is a global diffeomorphism, linear on the fibres, and

$$H = (\mathbf{p}\dot{\mathbf{r}} - L)|_{\dot{\mathbf{r}}=\frac{\mathbf{p}}{m}} = \frac{\mathbf{p}^2}{2m} + U(\mathbf{r}).$$

The Hamilton's equations

$$\begin{aligned} \dot{\mathbf{r}} &= \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{m}, \\ \dot{\mathbf{p}} &= -\frac{\partial H}{\partial \mathbf{r}} = -\frac{\partial U}{\partial \mathbf{r}}, \end{aligned}$$

are equivalent to Newton's equations with the force $\mathbf{F} = -\dfrac{\partial U}{\partial \mathbf{r}}$.

In general, consider the Lagrangian

$$L = \sum_{i,j=1}^{n} \tfrac{1}{2} a^{ij}(\mathbf{q}) \dot{q}_i \dot{q}_j - U(\mathbf{q}),\ \mathbf{q} \in \mathbb{R}^n,$$

where $A(\mathbf{q}) = \{a^{ij}(\mathbf{q})\}_{i,j=1}^n$ is a symmetric $n \times n$ matrix. We have

$$p^i = \frac{\partial L}{\partial \dot{q}_i} = \sum_{j=1}^{n} a^{ij}(\mathbf{q}) \dot{q}_j, \quad i = 1, \dots, n,$$

and the Legendre's transformation is a global diffeomorphism, linear on the fibres, if and only if the matrix $A(\mathbf{q})$ is non-degenerate for all $\mathbf{q} \in \mathbb{R}^n$. In this case,

$$H(\mathbf{p}, \mathbf{q}) = (\mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}))\big|_{\mathbf{p}=\frac{\partial L}{\partial \dot{\mathbf{q}}}} = \sum_{i,j=1}^{n} \tfrac{1}{2} a_{ij}(\mathbf{q}) p^i p^j + U(\mathbf{q}),$$

where $\{a_{ij}(\mathbf{q})\}_{i,j=1}^n = A^{-1}(\mathbf{q})$ is the inverse matrix.

PROBLEM 1.18. Show that $\theta_L(v) = dL(\pi_*(v))$, where $v \in T(TM)$ and $\pi : TM \to M$ is the canonical projection.

PROBLEM 1.19. Prove that the path $\gamma(t)$ in M is a trajectory for the Lagrangian system (M, L) if and only if

$$i_{\dot{\gamma}'(t)}(d\theta_L) + dE_L(\gamma'(t)) = 0,$$

where $\dot{\gamma}'(t)$ is the velocity vector of the path $\gamma'(t)$ in TM.

PROBLEM 1.20. Suppose that for a Lagrangian system $(\mathbb{R}^n, L)$ the Legendre's transformation τ_L is a diffeomorphism and let H be the corresponding Hamiltonian. Prove that for fixed $\mathbf{q}$ and $\dot{\mathbf{q}}$ the function $\mathbf{p}\dot{\mathbf{q}} - H(\mathbf{p}, \mathbf{q})$ has a single critical point at $\mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}$.

2. Hamiltonian Mechanics

2.1. Hamilton's equations. With every function $H : T^*M \to \mathbb{R}$ on the phase space T^*M there are associated Hamilton's equations — a first-order system of ordinary differential equations, which in the standard coordinates on T^*U has the form

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \quad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}.$$

The corresponding vector field X_H on T^*U,

$$X_H = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p^i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p^i} \right) = \frac{\partial H}{\partial \mathbf{p}} \frac{\partial}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \frac{\partial}{\partial \mathbf{p}},$$

gives rise to a well-defined vector field X_H on T^*M, called the *Hamiltonian vector field*. Suppose now that the vector field X_H on T^*M is complete, i.e., its integral curves exist for all times. The corresponding one-parameter group $\{g^t\}_{t \in \mathbb{R}}$ of diffeomorphisms of T^*M generated by X_H is defined by

$g^t(p,q) = (p(t), q(t))$, where $p(t), q(t)$ is a solution of the Hamilton's equations satisfying $p(0) = p, q(0) = q$, and is called the *Hamiltonian phase flow*.

The canonical Liouville's 1-form θ on T^*M defines a 2-form $\omega = d\theta$. In standard coordinates on T^*M it is given by

$$\omega = \sum_{i=1}^{n} dp^i \wedge dq_i = d\mathbf{p} \wedge d\mathbf{q},$$

and is a non-degenerate 2-form. The form ω is called the *canonical symplectic form* on T^*M. The symplectic form ω for every $(p,q) \in T^*M$ defines an isomorphism $J : T^*_{(p,q)}T^*M \to T_{(p,q)}T^*M$ by

$$\omega(u_1, u_2) = J^{-1}(u_2)(u_1),\ u_1, u_2 \in T_{(p,q)}T^*M.$$

In standard coordinates,

$$J(d\mathbf{p}) = \frac{\partial}{\partial \mathbf{q}}, \quad J(d\mathbf{q}) = -\frac{\partial}{\partial \mathbf{p}} \quad \text{and} \quad X_H = J(dH).$$

THEOREM 2.1. *The Hamiltonian phase flow on T^*M preserves the canonical symplectic form.*

PROOF. We need to prove that $(g^t)^*\omega = \omega$. Since g^t is a one-parameter group of diffeomorphisms, it is sufficient to show that

$$\frac{d}{dt}(g^t)^*\omega\Big|_{t=0} = \mathcal{L}_{X_H}\omega = 0,$$

where $\mathcal{L}_{X_H}$ is the Lie derivative along the vector field X_H. Since for every vector field X,

$$d\mathcal{L}_X(f) = \mathcal{L}_X(df),$$

we have

$$\mathcal{L}_{X_H}(dp^i) = d(X_H(p_i)) = -d\left(\frac{\partial H}{\partial q^i}\right) \quad \text{and} \quad \mathcal{L}_{X_H}(dq_i) = d(X_H(q_i)) = d\left(\frac{\partial H}{\partial p^i}\right).$$

Thus

$$\begin{aligned}\mathcal{L}_{X_H}\omega &= \sum_{i=1}^{n} \left(\mathcal{L}_{X_H}(dp^i) \wedge dq_i + dp^i \wedge \mathcal{L}_{X_H}(dq_i)\right) \\ &= \sum_{i=1}^{n} \left(-d\left(\frac{\partial H}{\partial q_i}\right) \wedge dq_i + dp^i \wedge d\left(\frac{\partial H}{\partial p^i}\right)\right) = -d(dH) = 0.\end{aligned}$$

□

COROLLARY 2.2. $\mathcal{L}_{X_H}(\theta) = d(-H + i_{X_H}(\theta))$.

The canonical symplectic form ω on T^*M defines the volume form $\dfrac{\omega^n}{n!} = \dfrac{1}{n!}\underbrace{\omega \wedge \cdots \wedge \omega}_{n}$ on T^*M, called the *Liouville volume form*.

COROLLARY 2.3 (Liouville's Theorem). *The Hamiltonian phase flow on T^*M preserves the Liouville volume form.*

The restriction of the symplectic form ω on T^*M to the configuration space M is 0. Generalizing this property we have the following

DEFINITION. A submanifold $\mathscr{L}$ of the phase space T^*M is called a *Lagrangian submanifold* if $\dim \mathscr{L} = \dim M$ and $\omega|_{\mathscr{L}} = 0$.

It follows from Theorem 2.1 that the image of a Lagrangian submanifold under the Hamiltonian phase flow is a Lagrangian submanifold.

PROBLEM 2.1. Verify that X_H is a well-defined vector field on T^*M.

PROBLEM 2.2. Show that if all level sets of the Hamiltonian H are compact submanifolds of T^*M, then the Hamiltonian vector field X_H is complete.

PROBLEM 2.3. Let $\pi : T^*M \to M$ be the canonical prjection, and let $\mathscr{L}$ be a Lagrangian submanifold. Show that if the the mapping $\pi|_{\mathscr{L}} : \mathscr{L} \to M$ is a diffeomorphism, then $\mathscr{L}$ is a graph of a smooth function on M. Give examples when for some $t > 0$ the corresponding projection of $g^t(\mathscr{L})$ onto M is no longer a diffeomorphism.

2.2. The action functional in the phase space. With every function H on the phase space T^*M there is an associated 1-form

$$\theta - H dt = \mathbf{p} d\mathbf{q} - H dt$$

on the extended phase space $T^*M \times \mathbb{R}$, called the *Poincaré-Cartan form*. Let $\pi_1 : T^*M \times \mathbb{R} \to M$ and $\pi_2 : T^*M \times \mathbb{R} \to \mathbb{R}$ be the canonical projections, and let $P(T^*M \times \mathbb{R})$ be the space of smooth parametrized paths $\sigma : [t_0, t_1] \to T^*M \times \mathbb{R}$ such that $\pi_1(\sigma(t_0)) = q_0$, $\pi_1(\sigma(t_1)) = q_1$, and $\pi_2(\sigma(t)) = t$ for all $t \in [t_0, t_1]$. Such paths called are *admissible* paths in $T^*M \times \mathbb{R}$. A variation of an admissible path σ is a smooth family of admissible paths σ_ε, where $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ and $\sigma_0 = \sigma$, and the corresponding infinitesimal variation is

$$\delta\sigma = \left.\frac{\partial \sigma_\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0} \in T_\sigma P(T^*M \times \mathbb{R})$$

(cf. Section 1.2). The principle of the least action in the phase space is the following statement.

THEOREM 2.4 (Poincaré). *The admissible path σ in $T^*M \times \mathbb{R}$ is an extremal for the action functional*

$$S(\sigma) = \int_\sigma (\mathbf{p} d\mathbf{q} - H dt) = \int_{t_0}^{t_1} (\mathbf{p}\dot{\mathbf{q}} - H) dt$$

if and only if its projection onto T^*M is a solution of the canonical Hamilton's equations

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}, \ \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}.$$

PROOF. As in the proof of Theorem 1.1, for an admissible family $\sigma_\varepsilon(t) = (\mathbf{p}(t,\varepsilon), \mathbf{q}(t,\varepsilon), t)$ in standard coordinates we compute, using integration by parts,

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\sigma_\varepsilon) = \sum_{i=1}^{n} \int_{t_0}^{t_1} \left(\dot{q}_i \delta p^i - \dot{p}^i \delta q_i - \frac{\partial H}{\partial q_i}\delta q_i - \frac{\partial H}{\partial p^i}\delta p^i \right) dt + \sum_{i=1}^{n} p^i \ \delta q_i|_{t_0}^{t_1}.$$

Since $\delta\mathbf{q}(t_0) = \delta\mathbf{q}(t_1) = 0$, we conclude that the path σ is critical if and only if $\mathbf{p}(t)$ and $\mathbf{q}(t)$ satisfy canonical Hamilton's equations. □

REMARK. For a Lagrangian system (M, L), every path $\gamma(t) = (\mathbf{q}(t))$ in configuration space M connecting points q_0 and q_1, defines an admissible path $\hat{\gamma}(t) = (\mathbf{p}(t), \mathbf{q}(t), t)$ in the phase space T^*M by setting $\mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}$. If the Legendre's transformation $\tau_L : TM \to T^*M$ is a diffeomorphism, then

$$S(\hat{\gamma}) = \int_{t_0}^{t_1} (\mathbf{p}\dot{\mathbf{q}} - H)dt = \int_{t_0}^{t_1} L(\gamma'(t), t)dt.$$

Thus the principle of the least action in a configuration space — Hamilton's principle, follows from the principle of the least action in a phase space. In fact, it follows from Problem 1.20 that in this case both principles are equivalent.

From Corollary 1.5 we immediately get

COROLLARY 2.5. *Solutions of canonical Hamilton's equations lying on the hypersurface $H(\mathbf{p}, \mathbf{q}) = E$ are extremals of the functional $\int_\sigma \mathbf{p}d\mathbf{q}$ in the class of admissible paths σ lying on this hypersurface.*

COROLLARY 2.6 (Maupertuis' Principle). *The trajectory $\gamma = (\mathbf{q}(\tau))$ of a closed Lagrangian system (M, L) connecting points q_0 and q_1 and having the energy E is the extremal of the functional*

$$\int_\gamma \mathbf{p}d\mathbf{q} = \int_\gamma \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}(\tau), \dot{\mathbf{q}}(\tau))\dot{\mathbf{q}}(\tau)d\tau$$

on the space of all paths in the configuration space M connecting points q_0 and q_1 and parametrized such that $H(\frac{\partial L}{\partial \dot{\mathbf{q}}}(\tau), \mathbf{q}(\tau)) = E$.

The functional

$$S_0(\gamma) = \int_\gamma \mathbf{p}d\mathbf{q}$$

is called the *abbreviated action*. The precise formulation of Maupertuis' princple is due to Euler and Lagrange.

PROOF. Every path $\gamma = \mathbf{q}(\tau)$, parametrized such that $H(\frac{\partial L}{\partial \dot{\mathbf{q}}}, \mathbf{q}) = E$, lifts to an admissible path $\sigma = (\frac{\partial L}{\partial \dot{\mathbf{q}}}(\tau), \mathbf{q}(\tau), \tau)$, $a \le \tau \le b$, lying on the hypersurface $H(\mathbf{p}, \mathbf{q}) = E$. □

PROBLEM 2.4 (Jacobi). On a Riemannian manifold (M, ds^2) consider a Lagrangian system with $L(q, v) = \frac{1}{2}\|v\|^2 + U(q)$. Let $E > U(q)$ for all $q \in M$. Show that the trajectories of a closed Lagrangian system (M, L) with total energy E are geodesics for the Riemannian metric $d\hat{s}^2 = (E - U(q))ds^2$ on M.

2.3. The action as a function of coordinates. For a Lagrangian system (M, L) denote by $\gamma(t; q_0, v_0)$ the solution of Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0$$

with the initial conditions $\gamma(t_0) = q_0 \in M$, $\dot{\gamma}(t_0) = v_0 \in T_{q_0}M$. Fix q_0, v_0 and t_0, and suppose that there exist a neighborhood V_0 of $v_0 \in T_{q_0}M$, a neighborhood U of $q_0 \in M$, and times t such that the mapping

$$T_{q_0}M \ni v \mapsto q = \gamma(t; q_0, v) \in M$$

is a diffeomorphism between V_0 and U. For such t, for every $q \in U$ there is a unique extremal $\gamma(\tau; q_0, v)$ with $v \in V_0$, passing through q_0 and q at times t_0 and t. Such extremals are said to form a *central field*. Basic theorems in the theory of ordinary differential equations guarantee that for times t sufficiently close to t_0, every extremal $\gamma(t)$ can be included into a central field of extremals.

In standard coordinates this mapping is denoted by $\dot{\mathbf{q}}_0 \mapsto \mathbf{q}(t) = \gamma(t; \mathbf{q}_0, \dot{\mathbf{q}}_0)$. We define the *action as function of coordinates and time* (*classical action*) by

$$S(\mathbf{q}, t; \mathbf{q}_0, t_0) = \int_{t_0}^{t} L(\gamma'(\tau))d\tau,$$

where γ is the extremal connecting $\mathbf{q}_0$ and $\mathbf{q}$. For a fixed energy E,

$$S(\mathbf{q}, t; \mathbf{q}_0, t_0) = S_0(\mathbf{q}, t; \mathbf{q}_0, t_0) + E(t - t_0). \tag{2.1}$$

THEOREM 2.7. *Under the above assumptions, the differential of the action as a function of coordinates with fixed initial point is given by*

$$dS = \mathbf{p}d\mathbf{q} - Hdt,$$

where $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$ *and* $H = \mathbf{p}\dot{\mathbf{q}} - L$ *are determined by the velocity* $\dot{\mathbf{q}}$ *of the trajectory* γ *at time* t.

PROOF. In standard coordinates, let $\mathbf{q}_\varepsilon$ be a path in M passing through $\mathbf{q}$ at $\varepsilon = 0$ with the tangent vector $\mathbf{v} \in T_{\mathbf{q}}M \simeq \mathbb{R}^n$. For the corresponding family of extremals $\gamma_\varepsilon(\tau) = \gamma(\tau; \mathbf{q}_0, \mathbf{q}_\varepsilon)$ the infinitesimal variation $\delta\gamma$ satisfies $\delta\gamma(t_0) = 0$ and $\delta\gamma(t) = \mathbf{v}$, and it follows from the variation with free ends formula (1.2) that for fixed t,

$$dS(\mathbf{v}) = \frac{\partial L}{\partial \dot{\mathbf{q}}}\mathbf{v}.$$

This shows that $\dfrac{\partial S}{\partial \mathbf{q}} = \mathbf{p}$. Now along the extremal $\gamma(t; \mathbf{q}_0, \mathbf{q})$,

$$\frac{d}{dt} S(\mathbf{q}(t), t; \mathbf{q}_0, t_0) = \frac{\partial S}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial S}{\partial t} = L,$$

so that $\dfrac{\partial S}{\partial t} = L - \mathbf{p}\dot{\mathbf{q}} = -H$. □

COROLLARY 2.8. *The classical action satisfies the following nonlinear partial differential equation*

$$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial \mathbf{q}}, \mathbf{q}\right) = 0. \tag{2.2}$$

This equation is called the *Hamilton-Jacobi equation*. Hamilton's equations can be used for solving the Cauchy problem

$$S(\mathbf{q}, t)|_{t=t_0} = s(\mathbf{q}) \tag{2.3}$$

for Hamilton-Jacobi equation (2.2) by the method of characteristics, assuming that the Hamiltonian phase flow g^t on the phase space $\mathscr{M} = T^*M$ exists. Namely, consider the Lagrangian submanifold

$$\mathscr{L} = \left\{ (\mathbf{p}, \mathbf{q}) \in T^*M : \mathbf{p} = \frac{\partial s(\mathbf{q})}{\partial \mathbf{q}} \right\}$$

— a graph of the section ds of the cotangent bundle $\pi : T^*M \to M$. The mapping $\pi|_{\mathscr{L}}$ is one to one and for sufficiently small $t - t_0$ the restriction of the projection π to the Lagrangian submanifold $\mathscr{L}^t = g^{t-t_0}\mathscr{L}$ remains to be one to one. For such t the mapping $\pi^t = \pi \circ g^t \circ (\pi|_{\mathscr{L}})^{-1} : M \to M$ is a diffeomorphism and the extremals $\gamma(\tau, \mathbf{q}_0, \dot{\mathbf{q}}_0)$ in the extended configuration space $M \times \mathbb{R}$, where $\dot{\mathbf{q}}_0 = \dfrac{\partial H}{\partial \mathbf{p}}(\mathbf{p}_0, \mathbf{q}_0)$ and $(\mathbf{p}_0, \mathbf{q}_0) \in \mathscr{L}$, do not intersect for $t_0 \le \tau \le t$. Such extremals are called the *characteristics* of the Hamilton-Jacobi equation.

PROPOSITION 2.1. *Under the above assumptions, the solution $S(\mathbf{q}, t)$ to the Cauchy problem* (2.2)–(2.3) *is given by*

$$S(\mathbf{q}, t) = s(\mathbf{q}_0) + \int_{t_0}^{t} L(\gamma'(\tau)) d\tau,$$

where $\gamma(\tau)$ is the characteristic which ends at a given point $(\mathbf{q}, t) \in M \times \mathbb{R}$ and starts at a point $(\mathbf{q}_0, t_0) \in M \times \mathbb{R}$, uniquely determined by $\mathbf{q} \in M$.

PROOF. Using again formula (1.2), where now $\mathbf{q}_0$ depends on $\mathbf{q}$, we get that along the characteristic,

$$\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}) = \frac{\partial s}{\partial \mathbf{q}_0}(\mathbf{q}_0) \frac{\partial \mathbf{q}_0}{\partial \mathbf{q}} + \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}}) - \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}_0, \dot{\mathbf{q}}_0) \frac{\partial \mathbf{q}_0}{\partial \mathbf{q}} = \mathbf{p},$$

since $\dfrac{\partial s}{\partial \mathbf{q}}(\mathbf{q}_0) = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}_0, \dot{\mathbf{q}}_0) = \mathbf{p}_0$. Now as in the proof of Theorem 2.7, we get along the characteristic, using that $\mathbf{q}_0(t)$ depends on $\mathbf{q}(t)$,

$$\frac{d}{dt}S(\mathbf{q}(t), t) = \frac{\partial S}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial S}{\partial t} = \frac{\partial s}{\partial \mathbf{q}}(\mathbf{q}_0)\dot{\mathbf{q}}_0 - \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}_0, \dot{\mathbf{q}}_0)\dot{\mathbf{q}}_0 + L(\mathbf{q}, \dot{\mathbf{q}}),$$

so that

$$\frac{\partial S}{\partial t} = -H(\mathbf{p}, \mathbf{q})$$

and S satisfies the Hamilton-Jacobi equation. □

We can also consider the action $S(\mathbf{q}, t; \mathbf{q}_0, t_0)$ as a function of both variables $\mathbf{q}$ and $\mathbf{q}_0$. The analog of Theorem 2.3 is the following statement.

PROPOSITION 2.2. *Differential of the action as a function of initial and final points is given by*

$$dS = \mathbf{p}d\mathbf{q} - \mathbf{p}_0 d\mathbf{q}_0 - H(\mathbf{p}, \mathbf{q})dt + H(\mathbf{p}_0, \mathbf{q}_0)dt_0.$$

PROBLEM 2.5. Prove that solution to the Cauchy problem for the Hamilton-Jacobi equation is unique.

2.4. Classical observables and Poisson bracket. Smooth real-valued functions on the phase space T^*M are called *classical observables*. The vector space $C^\infty(T^*M)$ is an $\mathbb{R}$-algebra — an associative algebra over $\mathbb{R}$ with a unit, given by the constant function 1, and with a multiplication given by the point-wise product of functions. The commutative algebra $C^\infty(T^*M)$ is called the *algebra of classical observables*. Assuming that the Hamiltonian phase flow g^t exists for all times, the time evolution of every observable $f \in C^\infty(T^*M)$ is given by

$$f_t(p, q) = f(g^t(p, q)) = f(p(t), q(t)), \quad q \in M, p \in T_qM.$$

Equivalently, the time evolution is described by the differential equation

$$\begin{aligned}\frac{df_t}{dt} &= \left.\frac{df_{s+t}}{ds}\right|_{s=0} = \left.\frac{d(f_t \circ g^s)}{ds}\right|_{s=0} = X_H(f_t) \\ &= \sum_{i=1}^{n}\left(\frac{\partial H}{\partial p^i}\frac{\partial f_t}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f_t}{\partial p^i}\right) = \frac{\partial H}{\partial \mathbf{p}}\frac{\partial f_t}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}}\frac{\partial f_t}{\partial \mathbf{p}},\end{aligned}$$

called Hamilton's equation for classical observables. Setting for $f, g \in C^\infty(T^*M)$,

$$\{f, g\} = X_f(g) = \frac{\partial f}{\partial \mathbf{p}}\frac{\partial g}{\partial \mathbf{q}} - \frac{\partial f}{\partial \mathbf{q}}\frac{\partial g}{\partial \mathbf{p}}, \tag{2.4}$$

we can rewrite the Hamilton's equation in the following concise form

$$\frac{df}{dt} = \{H, f\}, \tag{2.5}$$

where it is understood that (2.5) is a differential equation for a family of functions f_t on T^*M with the initial condition $f_t(\mathbf{p},\mathbf{q})|_{t=0} = f(\mathbf{p},\mathbf{q})$. The properties of the bilinear mapping

$$\{\ ,\ \} : C^\infty(T^*M) \times C^\infty(T^*M) \to C^\infty(T^*M)$$

are summarized below.

THEOREM 2.9. *The mapping* $\{\ ,\ \}$ *satisfies the following properties.*

(i) *(Relation with the symplectic form)*

$$\{f,g\} = \omega(Jdf, Jdg) = \omega(X_f, X_g).$$

(ii) *(Skew-symmetry)*

$$\{f,g\} = -\{g,f\}.$$

(iii) *(Leibniz rule)*

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

(iv) *(Jacobi identity)*

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$$

for all $f,g,h \in C^\infty(T^*M)$.

PROOF. Property (i) immediately follows from the definitions of ω and J in Section 2.1. Properties (ii)-(iii) are obvious. The Jacobi identity is verified by a straightforward computation, or by the following elegant argument. Observe that $\{f,g\}$ is a bilinear form in first partial derivatives of f and g, and every term in the left hand side of the Jacobi identity is a linear homogenous function of second partial derivatives of f, g and h. Now the only terms in the Jacobi identity which could actually contain second partial derivatives of a function h are the following

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} = (X_f X_g - X_g X_f)(h).$$

However, this expression does not contain second partial derivatives of h since it is commutator of two differential operators of the first order which is again an operator of the first order! □

The observable $\{f,g\}$ is called *canonical Poisson bracket* of the observables f and g. The Poisson bracket map $\{\ ,\ \}$ turns the algebra of classical observables $C^\infty(T^*M)$ into a Lie algebra with a Lie bracket given by the Poisson bracket. It has an important property that the Lie bracket is a biderivation with respect to the multiplication in $C^\infty(T^*M)$. The algebra of classical observables $C^\infty(T^*M)$ provides an example of the *Poisson algebra* — a commutative algebra over $\mathbb{R}$ with a Lie algebra structure satisfying the derivation property.

2.5. Canonical transformations and generating functions.

DEFINITION. A diffeomorphism g of the phase space T^*M is called a *canonical transformation*, if it preserves the canonical symplectic form ω on T^*M, i.e., $g^*(\omega) = \omega$. By Theorem 2.1, the Hamiltonian phase flow g^t is a one-parameter group of canonical transformations.

PROPOSITION 2.3. *Canonical transformations preserve Hamilton's equations.*

PROOF. If an observable f satisfies Hamilton's equation

$$\frac{df}{dt} = \omega(JdH, Jdf) = \{H, f\},$$

then, since $g^*(\omega) = \omega$, the observable $\tilde{f} = f \circ g$ satisfies

$$\frac{d\tilde{f}}{dt} = \omega(Jd\tilde{H}, Jd\tilde{f}) = \{\tilde{H}, \tilde{f}\},$$

where $\tilde{H} = H \circ g$. In other words, if in the standard coordinates $\mathbf{p}, \mathbf{q}$ we have $\omega = d\mathbf{p} \wedge d\mathbf{q}$ and

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{p}, \mathbf{q}), \quad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{p}, \mathbf{q}),$$

then in the new coordinates $(\mathbf{P}, \mathbf{Q}) = g(\mathbf{p}, \mathbf{q})$ we have $\omega = d\mathbf{P} \wedge d\mathbf{Q}$ and

$$\dot{\mathbf{P}} = -\frac{\partial H}{\partial \mathbf{Q}}(\mathbf{P}, \mathbf{Q}), \quad \dot{\mathbf{Q}} = \frac{\partial H}{\partial \mathbf{P}}(\mathbf{P}, \mathbf{Q}).$$

□

Consider now the classical case $M = \mathbb{R}^n$. For a canonical transformation $(\mathbf{P}, \mathbf{Q}) = g(\mathbf{p}, \mathbf{q})$ set $\mathbf{P} = \mathbf{P}(\mathbf{p}, \mathbf{q})$ and $\mathbf{Q} = \mathbf{Q}(\mathbf{p}, \mathbf{q})$. Since $d\mathbf{P} \wedge d\mathbf{Q} = d\mathbf{p} \wedge d\mathbf{q}$ on $T^*M \simeq \mathbb{R}^{2n}$, the 1-form $\mathbf{p}d\mathbf{q} - \mathbf{P}d\mathbf{Q}$ — the difference between the canonical Liouville 1-form and its pullback by the mapping g — is closed. From Poincaré lemma it follows that there exists a function $F(\mathbf{p}, \mathbf{q})$ on $\mathbb{R}^{2n}$ such that

$$\mathbf{p}d\mathbf{q} - \mathbf{P}d\mathbf{Q} = dF(\mathbf{p}, \mathbf{q}).$$

Now assume that at some point $(\mathbf{p}_0, \mathbf{q}_0)$ the $n \times n$ matrix $\dfrac{\partial \mathbf{Q}}{\partial \mathbf{p}} = \left\{ \dfrac{\partial Q_i}{\partial p^j} \right\}_{i,j=1}^{n}$ is nondegenerate. Canonical transformation g with this property is called *free*. By the inverse function theorem, there exists a neighborhood U of $(\mathbf{p}_0, \mathbf{q}_0)$ in $\mathbb{R}^{2n}$ for which the functions $\mathbf{Q}, \mathbf{q}$ are coordinate functions. The function

$$S(\mathbf{Q}, \mathbf{q}) = F(\mathbf{p}, \mathbf{q})$$

is called a *generating function* of the canonical transformation g in U. In new coordinates $\mathbf{Q}, \mathbf{q}$ on U we have

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{Q}, \mathbf{q}) \quad \text{and} \quad \mathbf{P} = -\frac{\partial S}{\partial \mathbf{Q}}(\mathbf{Q}, \mathbf{q}).$$

The converse statement easily follows from the implicit function theorem.

PROPOSITION 2.4. *Let $S(\mathbf{Q},\mathbf{q})$ be a function in some neighborhood U of a point $(\mathbf{Q}_0,\mathbf{q}_0)\in\mathbb{R}^{2n}$ such that the $n\times n$ matrix*

$$\frac{\partial^2 S}{\partial\mathbf{Q}\partial\mathbf{q}}(\mathbf{Q}_0,\mathbf{q}_0)=\left\{\frac{\partial^2\mathbf{Q}}{\partial Q_i\partial q_j}(\mathbf{Q}_0,\mathbf{q}_0)\right\}_{i,j=1}^{n}$$

is nondegenerate. Then S is a generating function of a local (i.e., defined in some neighborhood of $(\mathbf{Q}_0,\mathbf{q}_0)$ in $\mathbb{R}^{2n}$) free canonical transformation.

Suppose there is a canonical transformation $(\mathbf{P},\mathbf{Q})=g(\mathbf{p},\mathbf{q})$ such that $H(\mathbf{p},\mathbf{q})=K(\mathbf{Q})$ for some function K. Then in the new coordinates Hamilton's equations have the form

$$\dot{\mathbf{Q}}=0,\quad \dot{\mathbf{P}}=-\frac{\partial K}{\partial\mathbf{Q}},$$

and are trivially solved as follows

$$\mathbf{Q}(t)=\mathbf{Q}(0),\quad \mathbf{P}(t)=\mathbf{P}(0)-t\frac{\partial K}{\partial\mathbf{Q}}(\mathbf{Q}(0)).$$

Assuming that canonical transformation is free, its generating function $S(\mathbf{Q},\mathbf{q})$ satisfies the differential equation

$$H\Big(\frac{\partial S}{\partial\mathbf{q}}(\mathbf{Q},\mathbf{q}),\mathbf{q}\Big)=K(\mathbf{Q}), \tag{2.6}$$

where after the differentiation one should substitute $\mathbf{q}=\mathbf{q}(\mathbf{P},\mathbf{Q})$, defined by the canonical transformation g^{-1}. Differential equation (2.6) for fixed $\mathbf{Q}$, as it follows from (2.1), coincides with the Hamilton-Jacobi equation for the abbreviated action $S_0=S-Et$ with $E=K(\mathbf{Q})$,

$$H\Big(\frac{\partial S_0}{\partial\mathbf{q}}(\mathbf{Q},\mathbf{q}),\mathbf{q}\Big)=E.$$

THEOREM 2.10 (Jacobi). *Suppose that there exists a function $S(\mathbf{Q},\mathbf{q})$ depending on n parameters $\mathbf{Q}=(Q_1,\dots,Q_n)$ and satisfying the Hamilton-Jacobi equation (2.6) for some function $K(\mathbf{Q})$ and that the $n\times n$ matrix $\dfrac{\partial^2 S}{\partial\mathbf{Q}\partial\mathbf{q}}$ is non-degenerate. Then (locally) the Hamilton's equations*

$$\dot{\mathbf{p}}=-\frac{\partial H}{\partial\mathbf{q}},\quad \dot{\mathbf{q}}=\frac{\partial H}{\partial\mathbf{p}}$$

reduce to

$$\dot{\mathbf{Q}}=0,\quad \dot{\mathbf{P}}=-\frac{\partial K}{\partial\mathbf{Q}}$$

and are solved explicitly.

PROOF. Set $\mathbf{p}=\dfrac{\partial S}{\partial\mathbf{q}}(\mathbf{Q},\mathbf{q})$ and $\mathbf{P}=-\dfrac{\partial S}{\partial\mathbf{Q}}(\mathbf{Q},\mathbf{q})$. By the inverse function theorem, $g(\mathbf{p},\mathbf{q})=(\mathbf{P},\mathbf{Q})$ is a (local) canonical transformation with the generated function S, and it follows from (2.6) that $H(\mathbf{p}(\mathbf{P},\mathbf{Q}),\mathbf{q}(\mathbf{P},\mathbf{Q}))=K(\mathbf{Q})$. □

It seems that finding n-parameter solution of the Hamilton-Jacobi equation, which is a nonlinear partial differential equation, is a more difficult problem then solving Hamilton's equations, which is a system of ordinary differential equations. It is quite remarkable that for many problems of classical mechanics one can find n-parameter solution of the Hamilton-Jacobi equation by the method of separation of variables. By the Jacobi theorem, this solves the corresponding Hamilton's equations.

PROBLEM 2.6. Prove Proposition 2.4.

PROBLEM 2.7. Suppose that a canonical transformation $g(\mathbf{p},\mathbf{q}) = (\mathbf{P},\mathbf{Q})$ is such that locally $(\mathbf{P},\mathbf{q})$ can be considered as new coordinates. Prove that $S_1(\mathbf{P},\mathbf{q}) = \mathbf{PQ} + F(\mathbf{p},\mathbf{q})$, also called a *generation function*, satisfies

$$\mathbf{p} = \frac{\partial S_1}{\partial \mathbf{q}} \quad \text{and} \quad \mathbf{Q} = \frac{\partial S_1}{\partial \mathbf{P}}.$$

Find the generating function for identity transformation $\mathbf{P} = \mathbf{p}, \mathbf{Q} = \mathbf{q}$.

2.6. Symplectic manifolds. The notion of a symplectic manifold is a generalization of the example of a cotangent bundle T^*M.

DEFINITION. A non-degenerate, closed 2-form ω on a manifold $\mathscr{M}$ is called a *symplectic form*, and the pair $(\mathscr{M},\omega)$ is called a *symplectic manifold*.

Since symplectic form is non-degenerate, a symplectic manifold $\mathscr{M}$ is necessarily even-dimensional, $\dim \mathscr{M} = 2n$. Generalizing further the example $\mathscr{M} = T^*M$, we get the following

DEFINITION. A submanifold $\mathscr{L}$ of a symplectic manifold $(\mathscr{M},\omega)$ is called a *Lagrangian submanifold*, if $\dim \mathscr{L} = \frac{1}{2}\dim \mathscr{M}$ and the restriction of the symplectic form ω to $\mathscr{L}$ is 0.

Symplectic manifolds form a category. A morphism between $(\mathscr{M}_1,\omega_1)$ and $(\mathscr{M}_2,\omega_2)$, also called a *symplectomorphism*, is a mapping $f : \mathscr{M}_1 \to \mathscr{M}_2$ such that $\omega_1 = f^*(\omega_2)$. When $\mathscr{M}_1 = \mathscr{M}_2$ and $\omega_1 = \omega_2$, the notion of a symplectomorphism generalizes the notion of a canonical transformation. The direct product of symplectic manifolds $(\mathscr{M}_1,\omega_1)$ and $(\mathscr{M}_2,\omega_2)$ is a symplectic manifold

$$(\mathscr{M}_1 \times \mathscr{M}_2, \pi_1^*(\omega_1) + \pi_2^*(\omega_2)),$$

where π_1 and π_2 are, respectively, projections of $\mathscr{M}_1 \times \mathscr{M}_2$ onto the first and second factors in the Cartesian product. In addition to tangent bundles, other examples of symplectic manifolds are given by the real forms of complex Kähler manifolds, with a symplectic form being the Kähler form. In particular, for the case of complex projective varieties, a symplectic form is the pull-back of the Kähler form of the Fubini-Study metric on $\mathbb{CP}^n$.

The simplest example of a non-compact symplectic manifold is a symplectic vector space — the pair (V,ω), where ω is a non-degenerate, skew-symmetric bilinear form on a vector space V. It is the basis fact of linear

algebra that every symplectic vector space V has a *symplectic basis* — a basis $e_1, \dots, e_n, f^1, \dots, f^n$ of V, where $2n = \dim V$, such that

$$\omega(e_i, e_j) = \omega(f^i, f^j) = 0 \quad \text{and} \quad \omega(e_i, f^j) = \delta_i^j, \;\; i, j = 1, \dots, n.$$

In coordinates $(\mathbf{p}, \mathbf{q}) = (p^1, \dots, p^n, q_1, \dots, q_n)$ with respect to this basis, $V \simeq \mathbb{R}^{2n}$ and

$$\omega = d\mathbf{p} \wedge d\mathbf{q} = \sum_{i=1}^{n} dp^i \wedge dq_i.$$

Every symplectic vector space is isomorphic to a direct product of the phase planes $\mathbb{R}^2$ with the canonical symplectic form $dp \wedge dq$.

It turns out that every symplectic manifold locally looks like a symplectic vector space.

THEOREM 2.11 (Darboux' Theorem). *Let $(\mathscr{M}, \omega)$ be a $2n$-dimensional symplectic manifold. For every point $x \in \mathscr{M}$ there is a neighborhood U of x with local coordinates $(\mathbf{p}, \mathbf{q}) = (p^1, \dots, p^n, q_1, \dots, q_n)$ such that on U*

$$\omega = d\mathbf{p} \wedge d\mathbf{q} = \sum_{i=1}^{n} dp^i \wedge dq_i.$$

Coordinates $\mathbf{p}, \mathbf{q}$ are called *canonical* or Darboux coordinates. The proof proceeds by induction on n with two main steps stated as Problems 2.9 and 2.10.

A non-degenerate 2-form ω for every $x \in \mathscr{M}$ defines an isomorphism $J : T_x^*\mathscr{M} \to T_x\mathscr{M}$ by

$$\omega(u_1, u_2) = J^{-1}(u_2)(u_1), \;\; u_1, u_2 \in T_x\mathscr{M}.$$

In local coordinates $\mathbf{x} = (x_1, \dots, x_{2n})$ for the coordinate chart (U, φ) on $\mathscr{M}$, the 2-form ω is given by

$$\omega = \tfrac{1}{2} \sum_{i,j=1}^{2n} \omega^{ij}(\mathbf{x})\, dx_i \wedge dx_j,$$

where $\{\omega^{ij}(\mathbf{x})\}_{i,j=1}^{2n}$ is a non-degenerate, skew-symmetric matrix-function on $\varphi(U)$. Denoting the inverse matrix by $\{\omega_{ij}(\mathbf{x})\}_{i,j=1}^{2n}$, we have

$$J(dx_i) = -\sum_{j=1}^{2n} \omega_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j}, \quad i = 1, \dots, 2n.$$

DEFINITION. A *Hamiltonian system* is a pair consisting of a symplectic manifold $(\mathscr{M}, \omega)$, called a *phase space*, and a smooth real-valued function H on $\mathscr{M}$, called a *Hamiltonian*. The motion of a points on the phase space is described by the vector field

$$X_H = J(dH),$$

called a *Hamiltonian vector field*.

The trajectories of a Hamiltonian system $((\mathscr{M},\omega),H)$ are the integral curves of a Hamiltonian vector field X_H on $\mathscr{M}$. In canonical coordinates $(\mathbf{p},\mathbf{q})$ they are described by the canonical Hamilton's equations

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \quad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}.$$

Suppose now that the Hamiltonian vector field X_H on $\mathscr{M}$ is complete. The *phase flow* on $\mathscr{M}$ associated with a Hamiltonian H is a one-parameter group $\{g^t\}_{t\in\mathbb{R}}$ of diffeomorphisms of $\mathscr{M}$ generated by X_H. The following statement generalizes Theorem 2.1.

THEOREM 2.12. *The Hamiltonian phase flow on the phase space preserves the symplectic form.*

PROOF. It is sufficient to show that $\mathcal{L}_{X_H}\omega = 0$. Using Cartan's formula

$$\mathcal{L}_X = i_X \circ d + d \circ i_X$$

and $d\omega = 0$, we get for every $X \in \mathrm{Vect}(\mathscr{M})$,

$$\mathcal{L}_X\omega = (d \circ i_X)(\omega).$$

Since $i_X(\omega)(Y) = \omega(X,Y)$, we have for $X = X_H$ and every $Y \in \mathrm{Vect}(\mathscr{M})$ that

$$i_{X_H}(\omega)(Y) = \omega(J(dH),Y) = -dH(Y).$$

Thus $i_{X_H}(\omega) = -dH$, and the statement follows from $d^2 = 0$. □

COROLLARY 2.13. *A vector field X on $\mathscr{M}$ is a Hamiltonian vector field if and only if the 1-form $i_X(\omega)$ is exact.*

DEFINITION. A vector field X on a symplectic manifold $(\mathscr{M},\omega)$ is called *symplectic* vector field, if the 1-form $i_X(\omega)$ is closed, which is equivalent to $\mathscr{L}_X\omega = 0$.

The commutative algebra $C^\infty(\mathscr{M})$ with a multiplication given by the point-wise product of functions, is called the *algebra of classical observables*. Assuming that the Hamiltonian phase flow g^t exists for all times, the time evolution of every observable $f \in C^\infty(\mathscr{M})$ is given by

$$f_t(x) = f(g^t(x)), \quad x \in \mathscr{M},$$

and is described by the differential equation

$$\frac{df_t}{dt} = X_H(f_t)$$

— the Hamilton's equation for classical observables. The Hamilton's equations for observables on $\mathscr{M}$ have the same form as the Hamilton's equations on $\mathscr{M} = T^*M$, considered in Section 2.3. Since

$$X_H(f) = df(X_H) = \omega(X_H, J(df)) = \omega(X_H, X_f),$$

we have the following

DEFINITION. A Poisson bracket on the algebra $C^\infty(\mathscr{M})$ of classical observables on a symplectic manifold $(\mathscr{M},\omega)$ is a bilinear mapping $\{\ ,\ \}: C^\infty(\mathscr{M})\times C^\infty(\mathscr{M})\to C^\infty(\mathscr{M})$, defined by

$$\{f,g\}=\omega(X_f,X_g),\quad f,g\in C^\infty(\mathscr{M}).$$

Now the Hamilton's equation takes the following concise form

$$\frac{df}{dt}=\{H,f\}, \tag{2.7}$$

understood as a differential equation for a family of functions f_t on $\mathscr{M}$ with the initial condition $f_t|_{t=0}=f$. In local coordinates $\mathbf{x}=(x_1,\dots,x_{2n})$ on $\mathscr{M}$,

$$\{f,g\}(x)=-\sum_{i,j=1}^{2n}\omega_{ij}(\mathbf{x})\frac{\partial f(\mathbf{x})}{\partial x_i}\frac{\partial g(\mathbf{x})}{\partial x_j}.$$

THEOREM 2.14. *The Poisson bracket $\{\ ,\ \}$ on a symplectic manifold $(\mathscr{M},\omega)$ is skew-symmetric, satisfies Leibniz rule and the Jacobi identity.*

PROOF. The first two properties are obvious. It follows from the definition of a Poisson bracket and the formula

$$[X_f,X_g](h)=(X_gX_f-X_fX_g)(h)=\{g,\{f,h\}\}-\{f,\{g,h\}\},$$

that the Jacobi identity is equivalent to the property

$$[X_f,X_g]=X_{\{f,g\}}. \tag{2.8}$$

Let X and Y be symplectic vector fields. Using Cartan's formulas we get

$$\begin{aligned} i_{[X,Y]}(\omega)&=\mathcal{L}_X(i_Y(\omega))-i_Y(\mathcal{L}_X(\omega))\\ &=d(i_X\circ i_Y(\omega))+i_Xd(i_Y(\omega))\\ &=d(\omega(Y,X))=i_{X_{\omega(X,Y)}}(\omega), \end{aligned}$$

where $X_{\omega(X,Y)}$ is a Hamiltonian vector field corresponding to $\omega(X,Y)\in C^\infty(\mathscr{M})$. Since 2-form ω is non-degenerate, this implies that

$$[X,Y]=X_{\omega(X,Y)},$$

and setting $X=X_f, Y=X_g$, we get (2.8). □

From (2.8) we immediately get

COROLLARY 2.15. *The subspace* $\mathrm{Ham}(\mathscr{M})$ *of Hamiltonian vector fields on $\mathscr{M}$ is a Lie subalgebra of* $\mathrm{Vect}(\mathscr{M})$. *The mapping $C^\infty(\mathscr{M})\to\mathrm{Ham}(\mathscr{M})$, given by $f\mapsto X_f$, is a Lie algebra homomorphism with the kernel consisting of locally constant functions on $\mathscr{M}$*

In Lagrangian mechanics, a function I on TM is an integral of motion for the Lagrangian system (M,L), if it is constant along the trajectories. In Hamiltonian mechanics, an observable I — a function on the phase space $\mathscr{M}$ — is called an integral of motion (first integral) for the Hamiltonian system

$((\mathscr{M},\omega),H)$, if it is constant along the Hamiltonian phase flow. According to (2.7), this is equivalently to the condition

$$\{H,I\}=0.$$

It is said that the observables H and I are *in involution* (*Poisson commute*). From the Jacobi identity for the Poisson bracket we get the following

COROLLARY 2.16 (Poisson's Theorem). *The Poisson bracket of two integrals of motion is an integral of motion.*

PROOF. If $\{H,I_1\}=\{H,I_2\}=0$, then

$$\{H,\{I_1,I_2\}\}=\{\{H,I_1\},I_2\}-\{\{H,I_2\},I_1\}=0.$$

□

It follows from the Poisson theorem that integrals of motion form a Lie algebra and, by (2.8), corresponding Hamiltonian vector fields form a Lie subalgebra in $\mathrm{Vect}(\mathscr{M})$. Since $\{I,H\}=dH(X_I)=0$, the vector fields X_I are tangent to submanifolds $H=E$ of $\mathscr{M}$ — the level sets of the Hamiltonian H. This defines a Lie algebra of integrals of motion for the Hamiltonian system $((\mathscr{M},\omega),H)$ at the level set $H=E$.

Let G be a finite-dimensional Lie group that acts on a connected symplectic manifold $(\mathscr{M},\omega)$ by symplectomorphisms. The Lie algebra $\mathfrak{g}$ of G acts on $\mathscr{M}$ by vector fields and the linear mapping $\mathfrak{g}\ni\xi\mapsto X_\xi\in\mathrm{Vect}(\mathscr{M})$ is a homorphism of Lie algebras,

$$[X_\xi,X_\eta]=X_{[\xi,\eta]},\quad \xi,\eta\in\mathfrak{g}.$$

The G-action is called *Hamiltonian*, if X_ξ are Hamiltonian vector fields, i.e., for every $\xi\in\mathfrak{g}$ there is $\Phi_\xi\in C^\infty(\mathscr{M})$, defined up to an additive constant, such that $X_\xi=X_{\Phi_\xi}=J(d\Phi_\xi)$. The action is called *Poisson*, if there is a choice of functions Φ_ξ such that the linear mapping $\Phi:\mathfrak{g}\to C^\infty(\mathscr{M})$ is a homorphism of Lie algebras,

$$\{\Phi_\xi,\Phi_\eta\}=\Phi_{[\xi,\eta]},\quad \xi,\eta\in\mathfrak{g}. \tag{2.9}$$

DEFINITION. A Lie group G is a *symmetry group* of the Hamiltonian system $((\mathscr{M},\omega),H)$, if there is a Hamiltonian action of G on $\mathscr{M}$ such that

$$H(g\cdot x)=H(x),\quad g\in G,\ x\in\mathscr{M}.$$

THEOREM 2.17 (Noether theorem with symmetries). *If G is a symmetry group of the Hamiltonian system $((\mathscr{M},\omega),H)$, then the functions Φ_ξ, $\xi\in\mathfrak{g}$, are the integrals of motion. If the action of G is Poisson, the integrals of motion satisfy* (2.9).

PROOF. By definition of the Hamiltonian action, for every $\xi\in\mathfrak{g}$,

$$0=X_\xi(H)=X_{\Phi_\xi}(H)=\{\Phi_\xi,H\}.$$

□

COROLLARY 2.18. *Suppose that for a Lagrangian system (M, L) a Lie group G acts on the configuration space M such that $L(g_*(v)) = L(v)$ for $g \in G$, $v \in TM$. If the Legendre transformation $\tau_L : TM \to T^*M$ is a diffeomorphism, the Lie group G is a symmetry group of Hamiltonian system $((T^*M, \omega), H)$ corresponding to (M, L), and the G-action on the phase space T^*M is Poisson. In particular, $\Phi_\xi = -I_\xi \circ \tau_L$, where I_ξ is a Noether integral of motion for the one-parameter subgroup of G generated by $\xi \in \mathfrak{g}$.*

PROOF. According to Theorem 1.3, Noether integrals are linear in momenta, so that for $\Phi_\xi = -I_\xi \circ \tau_L$ we get, using (1.3),

$$X_{\Phi_\xi} = J(d\Phi_\xi) = X_\xi \quad \text{and} \quad \Phi_\xi = -i_{X_\xi}(\theta),$$

where θ is a canonical Liouville 1-form on T^*M. Using Cartan's formula and Corollary 2.2, we obtain

$$\begin{aligned}\Phi_{[\xi,\eta]} &= -i_{[X_\xi, X_\eta]}(\theta) = -\mathcal{L}_{X_\xi} i_{X_\eta}(\theta) - i_{X_\eta}\mathcal{L}_{X_\xi}(\theta) \\ &= X_{\Phi_\xi}(\Phi_\eta) = \{\Phi_\xi, \Phi_\eta\}.\end{aligned}$$

□

EXAMPLE 2.1. The Lagrangian

$$L = \tfrac{1}{2} m\dot{\mathbf{r}}^2 - U(r)$$

for a particle in $\mathbb{R}^3$ moving in a central field (see Section 1.6), is invariant with respect to the action of the group $SO(3)$ of orthogonal transformations of the Euclidean space $\mathbb{R}^3$. Let u_1, u_2, u_3 be a basis for the Lie algebra so(3) corresponding to the rotations with the axes given by the vectors of the standard basis e_1, e_2, e_3 for $\mathbb{R}^3$, (see Example 1.9 in Section 1.4). These generators satisfy the commutation relations

$$[u_i, u_j] = \varepsilon_{ijk} u_k,$$

where $i, j, k = 1, 2, 3$, and ε_{ijk} is totally antisymmetric tensor, $\varepsilon_{123} = 1$. Corresponding Noether integrals of motion are given by $\Phi_{u_i} = -M_i$, where

$$\begin{aligned}M_1 &= (\mathbf{r} \times \mathbf{p})_1 = r_2 p_3 - r_3 p_2, \\ M_2 &= (\mathbf{r} \times \mathbf{p})_2 = r_3 p_1 - r_1 p_3, \\ M_3 &= (\mathbf{r} \times \mathbf{p})_3 = r_1 p_2 - r_2 p_1.\end{aligned}$$

(It is convenient to lower the indices of the momenta p^i by the Euclidean metric on $\mathbb{R}^3$.) According to Theorem 2.17 and Corollary 2.18, their Poisson brackets satisfy

$$\{M_i, M_j\} = -\varepsilon_{ijk} M_k,$$

which is also easy to verify directly using (2.4),

$$\{f, g\}(\mathbf{r}, \mathbf{p}) = \frac{\partial f}{\partial \mathbf{p}} \frac{\partial g}{\partial \mathbf{r}} - \frac{\partial f}{\partial \mathbf{r}} \frac{\partial g}{\partial \mathbf{p}}.$$

EXAMPLE 2.2 (The Kepler's problem). For every $\alpha \in \mathbb{R}$ the Lagrangian system on $\mathbb{R}^3$ with

$$L = \tfrac{1}{2} m \dot{\mathbf{r}}^2 + \frac{\alpha}{r}$$

has three extra integrals of motion — the components W_1, W_2, W_3 of the Laplace-Runge-Lenz vector, given by

$$\mathbf{W} = \frac{\mathbf{p}}{m} \times \mathbf{M} - \frac{\alpha \mathbf{r}}{r}$$

(see Section 1.6). Using Poisson brackets from the previous example, together with $\{r_i, M_j\} = -\varepsilon_{ijk} r_k$ and $\{p_i, M_j\} = -\varepsilon_{ijk} p_k$, we get by a straightforward computation,

$$\{W_i, M_j\} = -\varepsilon_{ijk} W_k, \quad \text{and} \quad \{W_i, W_j\} = \frac{2H}{m} \varepsilon_{ijk} M_k,$$

where $H = \dfrac{\mathbf{p}^2}{2m} - \dfrac{\alpha}{r}$ is a Hamiltonian of the Kepler's problem.

PROBLEM 2.8 (Coadjont orbits). Let G be a finite-dimensional Lie group, $\mathfrak{g}$ be its Lie algebra, and let $\mathfrak{g}^*$ be the dual vector space to $\mathfrak{g}$. For $u \in \mathfrak{g}^*$ let $\mathscr{M} = \mathcal{O}_u$ be the orbit of u under the coadjoint action of G on $\mathfrak{g}^*$. Show that the formula

$$\omega(u_1, u_2) = u([x_1, x_2]),$$

where $u_1 = \mathrm{ad}^* x_1(u), u_2 = \mathrm{ad}^* x_2(u) \in \mathcal{O}_u$ and ad^* stands for the coadjoint action of a Lie algebra $\mathfrak{g}$ on $\mathfrak{g}^*$, gives rise to a well-defined 2-form on $\mathscr{M}$, which is closed and non-degenerate. (The 2-form ω is called the *Kirillov-Kostant* symplectic form.)

PROBLEM 2.9. Let $(\mathscr{M}, \omega)$ be a symplectic manifold. For $x \in \mathscr{M}$ choose a function q_1 on $\mathscr{M}$ such that $q_1(x) = 0$ and dq_1 does not vanish at x, and set $X = -X_{q_1}$. Show that there is a neighborhood U of $x \in \mathscr{M}$ and a function p^1 on U such that $X(q_1) = 1$ on U, and there exist coordinates $p^1, q_1, z_1, \ldots, z_{2n-2}$ on U such that

$$X = \frac{\partial}{\partial p^1} \quad \text{and} \quad Y = X_{p^1} = \frac{\partial}{\partial q_1}.$$

PROBLEM 2.10. Continuing Problem 2.9, show that the 2-form $\omega - dp^1 \wedge dq_1$ on U depends only on the variables $z_1, \ldots, z_{2n-2}$ and is non-degenerate.

PROBLEM 2.11. Do the computation in Example 2.2 and show that the Lie algebra of the integrals $M_1, M_2, M_3, W_1, W_2, W_3$ in Kepler's problem at $H(\mathbf{p}, \mathbf{r}) = E$ is isomorphic to the Lie algebra so(4), if $E < 0$, to the Lie algebra e(3), if $E = 0$, and to the Lie algebra so(1, 3), if $E > 0$.

PROBLEM 2.12 (Symplectic quotients). For a Poisson action of a Lie group G on a symplectic manifold $(\mathscr{M}, \omega)$, define the *moment map* $P : \mathscr{M} \to \mathfrak{g}^*$ by

$$P(x)(\xi) = \Phi_\xi(x), \ \xi \in \mathfrak{g}, \ x \in \mathscr{M},$$

where $\mathfrak{g}$ is the Lie algebra of G. For every regular value $p \in \mathfrak{g}^*$ of the moment map P such that a stabilizer G_p of p acts freely and proper on $\mathscr{M}_p = P^{-1}(p)$, the quotient $M_p = G_p \backslash \mathscr{M}_p$ is called a *reduced phase space*. Show that M_p is a symplectic manifold with the symplectic form uniquely characterized by the condition that its pull-back to $\mathscr{M}_p$ coincides with the restriction to $\mathscr{M}_p$ of the symplectic form ω.

2.7. Poisson manifolds. The notion of a Poisson manifold generalizes the notion of a symplectic manifold.

DEFINITION. *A Poisson manifold* is a manifold $\mathscr{M}$ equipped with a *Poisson structure* — a skew-symmetric bilinear mapping

$$\{\ ,\ \} : C^\infty(\mathscr{M}) \times C^\infty(\mathscr{M}) \to C^\infty(\mathscr{M})$$

which satisfies the Leibniz rule and Jacobi identity.

Equivalently, $\mathscr{M}$ is a Poisson manifold if the algebra $\mathcal{A} = C^\infty(\mathscr{M})$ of classical observables is a Poisson algebra — a Lie algebra such that the Lie bracket is a biderivation with respect to the multiplication in $\mathcal{A}$ (a pointwise product of functions). It follows from the derivation property that in local coordinates $\mathbf{x} = (x_1, \dots, x_n)$ on $\mathscr{M}$, the Poisson bracket has the form

$$\{f, g\}(x) = \sum_{i,j=1}^{N} \eta_{ij}(\mathbf{x}) \frac{\partial f(\mathbf{x})}{\partial x_i} \frac{\partial g(\mathbf{x})}{\partial x_j}.$$

The 2-tensor $\eta_{ij}(\mathbf{x})$ defines a global section η of the vector bundle $T\mathscr{M} \wedge T\mathscr{M}$ over $\mathscr{M}$, called a *Poisson tensor*.

The evolution of classical observables on a Poisson manifold is given by the Hamilton's equations, which have the same form as (2.7),

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

The phase flow g^t for a complete Hamiltonian vector field $X_H = \{H, \cdot\}$ defines the *evolution operator* $U_t : \mathcal{A} \to \mathcal{A}$ by

$$U_t(f)(x) = f(g^t(x)),\ f \in \mathcal{A}.$$

THEOREM 2.19. *Suppose that every Hamiltonian vector field on a Poisson manifold $(\mathscr{M}, \{\ ,\ \})$ is complete. Then for every $H \in \mathcal{A}$, the corresponding evolution operator U_t is an automorphism of the Poisson algebra $\mathcal{A}$, i.e.,*

$$U_t(\{f, g\}) = \{U_t(f), U_t(g)\} \quad \textit{for all} \quad f, g \in \mathcal{A}. \tag{2.10}$$

Conversely, if a skew-symmetric bilinear mapping $\{\ ,\ \} : C^\infty(\mathscr{M}) \times C^\infty(\mathscr{M}) \to C^\infty(\mathscr{M})$ is such that $X_H = \{H, \cdot\}$ are complete vector fields for all $H \in \mathcal{A}$ and corresponding evolution operators U_t satisfy (2.10)*, then $(\mathscr{M}, \{\ ,\ \})$ is a Poisson manifold.*

PROOF. Let $f_t = U_t(f)$, $g_t = U_t(g)$ and $h_t = U_t(\{f, g\})$. By definition,

$$\frac{d}{dt}\{f_t, g_t\} = \{\{H, f_t\}, g_t\} + \{f_t, \{H, g_t\}\} \quad \text{and} \quad \frac{dh_t}{dt} = \{H, h_t\}.$$

If $(\mathscr{M}, \{\ ,\ \})$ is a Poisson manifold, then it follows from the Jacobi identity that

$$\{\{H, f_t\}, g_t\} + \{f_t, \{H, g_t\}\} = \{H, \{f_t, g_t\}\},$$

so that h_t and $\{f_t, g_t\}$ satisfy satisfy the same differential equation (2.7). Since these functions coincide at $t = 0$, (2.10) follows from the uniqueness theorem for the ordinary differential equations.

Conversely, we get Jacobi identity for the functions f, g and H by differentiating (2.10) with respect to t at $t = 0$. □

COROLLARY 2.20. *A global section η of $T\mathscr{M} \wedge T\mathscr{M}$ is a Poisson tensor if and only if*

$$\mathcal{L}_{X_f}\eta = 0 \quad \textit{for all} \quad f \in \mathcal{A}.$$

DEFINITION. The *center* of a Poisson algebra $\mathcal{A}$ is

$$\mathcal{Z}(\mathcal{A}) = \{f \in \mathcal{A} : \{f, g\} = 0 \quad \text{for all} \quad g \in \mathcal{A}\}.$$

A Poisson manifold $(\mathscr{M}, \{\ ,\ \})$ is called *non-degenerate*, if a center of the Poisson algebra of classical observables $\mathcal{A} = C^\infty(\mathscr{M})$ consists only of locally constant functions ($\mathcal{Z}(\mathcal{A}) = \mathbb{R}$ for connected $\mathscr{M}$).

Equivalently, a Poisson manifold $(\mathscr{M}, \{\ ,\ \})$ is non-degenerate if the Poisson tensor η for every $x \in \mathscr{M}$ defines an isomorphism $J : T_x^*\mathscr{M} \to T_x\mathscr{M}$ by

$$\eta(u_1, u_2) = u_2(J(u_1)),\ u_1, u_2 \in T_x^*\mathscr{M}.$$

In local coordinates $\mathbf{x} = (x_1, \dots, x_n)$ for the coordinate chart (U, φ) on $\mathscr{M}$, we have

$$J(dx_i) = \sum_{j=1}^{n} \eta_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j}, \quad i = 1, \dots, n.$$

Poisson manifolds form a category. A morphism between $(\mathscr{M}_1, \{\ ,\ \}_1)$ and $(\mathscr{M}_2, \{\ ,\ \}_2)$ is a mapping $\varphi : \mathscr{M}_1 \to \mathscr{M}_2$ of smooth manifolds such that

$$\{f \circ \varphi, g \circ \varphi\}_1 = \{f, g\}_2 \circ \varphi \quad \forall f, g \in C^\infty(\mathscr{M}_2).$$

A direct product of Poisson manifolds $(\mathscr{M}_1, \{\ ,\ \}_1)$ and $(\mathscr{M}_1, \{\ ,\ \}_1)$ is a Poisson manifold $(\mathscr{M}_1 \times \mathscr{M}_2, \{\ ,\ \})$ defined by the property that natural projections maps $\pi_1 : \mathscr{M}_1 \times \mathscr{M}_2 \to \mathscr{M}_1$ and $\pi_2 : \mathscr{M}_1 \times \mathscr{M}_2 \to \mathscr{M}_2$ are Poisson mappings. For $f \in C^\infty(\mathscr{M}_1 \times \mathscr{M}_2)$ and $(x_1, x_2) \in \mathscr{M}_1 \times \mathscr{M}_2$ denote, respectively, by $f_{x_2}^{(1)}$ and $f_{x_1}^{(2)}$ restrictions of f to $\mathscr{M} \times \{x_2\}$ and $\{x_1\} \times \mathscr{M}_2$. Then for $f, g \in C^\infty(\mathscr{M}_1 \times \mathscr{M}_2)$,

$$\{f, g\}(x_1, x_2) = \{f_{x_2}^{(1)}, g_{x_2}^{(1)}\}_1(x_1) + \{f_{x_1}^{(2)}, g_{x_1}^{(2)}\}_2(x_2).$$

Non-degenerate Poisson manifolds form a subcategory of the category of Poisson manifolds.

THEOREM 2.21. *The category of symplectic manifolds is (anti-) isomorphic to the category of non-degenerate Poisson manifolds.*

PROOF. According to Theorem 2.14, every symplectic manifold carries a non-degenerate Poisson structure. Conversely, let $(\mathscr{M}, \{\ ,\ \})$ be a non-degenerate Poisson manifold. Define the 2-form ω on $\mathscr{M}$ by

$$\omega(X, Y) = J^{-1}(Y)(X), \quad X, Y \in \mathrm{Vect}(\mathscr{M}),$$

where the isomorphism $J : T^*\mathscr{M} \to T\mathscr{M}$ is defined by the Poisson tensor η. In local coordinates $\mathbf{x} = (x_1, \dots, x_n)$ on $\mathscr{M}$,

$$\omega = -\sum_{1 \le i < j \le n} \eta^{ij}(\mathbf{x})\, dx_i \wedge dx_j,$$

where $\{\eta^{ij}(\mathbf{x})\}_{i,j=1}^n$ is the inverse matrix to $\{\eta_{ij}(\mathbf{x})\}_{i,j=1}^n$. The 2-form ω is skew-symmetric and non-degenerate. For every $f \in \mathcal{A}$ let $X_f = \{f, \cdot\}$ be the corresponding vector field on $\mathscr{M}$. The Jacobi identity for the Poisson bracket $\{\ ,\ \}$ is equivalent to $\mathcal{L}_{X_f}\eta = 0$ for every $f \in \mathcal{A}$, so that

$$\mathcal{L}_{X_f}\omega = 0.$$

Since $X_f = Jdf$, we have $\omega(X, Jdf) = df(X)$ for every $X \in \mathrm{Vect}(\mathscr{M})$, so that

$$\omega(X_f, X_g) = \{f, g\}.$$

By Cartan's formula,

$$\begin{aligned} d\omega(X,Y,Z) =& \tfrac{1}{3}\left(\mathcal{L}_X\omega(X,Y) - \mathcal{L}_Y\omega(X,Z) + \mathcal{L}_Z\omega(X,Y)\right. \\ & \left. -\omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X)\right), \end{aligned}$$

where $X, Y, Z \in \mathrm{Vect}(\mathscr{M})$. Now setting $X = X_f, Y = X_g, Z = X_h$, we get

$$\begin{aligned} d\omega(X_f,X_g,X_h) =& \tfrac{1}{3}\left(\omega(X_h,[X_f,X_g]) + \omega(X_f,[X_g,X_h]) + \omega(X_g,[X_h,X_f])\right) \\ =& \tfrac{1}{3}\left(\omega(X_h,X_{\{f,g\}}) + \omega(X_f,X_{\{g,h\}}) + \omega(X_g,X_{\{h,f\}})\right) \\ =& \tfrac{1}{3}\left(\{h,\{f,g\}\} + \{f,\{g,h\}\} + \{g,\{h,f\}\}\right) \\ =& 0. \end{aligned}$$

The exact 1-forms df, $f \in \mathcal{A}$, generate the vector space of 1-forms $\mathcal{A}^1(\mathscr{M})$ as a module over $\mathcal{A}$, so that Hamiltonian vector vector fields $X_f = Jdf$ generate the vector space $\mathrm{Vect}(\mathscr{M})$ as a module over $\mathcal{A}$. Thus $d\omega = 0$ and $(\mathscr{M}, \omega)$ is a symplectic manifold associated with the Poisson manifold $(\mathscr{M}, \{\ ,\ \})$. It follows from the definitions that Poisson mappings of non-degenerate Poisson manifolds correspond to symplectomorphisms of associated Poisson manifolds. □

Remark. One can also prove the theorem by a straightforward computation in local coordinates $\mathbf{x} = (x_1, \dots, x_n)$ on $\mathscr{M}$. Just observe that the condition

$$\frac{\partial \eta^{ij}(\mathbf{x})}{\partial x_l} + \frac{\partial \eta^{jl}(\mathbf{x})}{\partial x_i} + \frac{\partial \eta^{li}(\mathbf{x})}{\partial x_j} = 0, \quad i, jl = 1, \dots, n,$$

which is a coordinate form of $d\omega = 0$, follows from the condition

$$\sum_{j=1}^{n} \left(\eta_{ij}(\mathbf{x}) \frac{\partial \eta_{kl}(\mathbf{x})}{\partial x_j} + \eta_{lj}(\mathbf{x}) \frac{\partial \eta_{ik}(\mathbf{x})}{\partial x_j} + \eta_{kj}(\mathbf{x}) \frac{\partial \eta_{li}(\mathbf{x})}{\partial x_j} \right) = 0,$$

which is a coordinate form of the Jacobi identity, by multiplying it three times by the inverse matrix using

$$\sum_{p=1}^{n}\left(\eta_{ip}(\mathbf{x})\frac{\partial\eta^{pk}(\mathbf{x})}{\partial x_j}+\frac{\partial\eta_{ip}(\mathbf{x})}{\partial x_j}\eta^{pk}(\mathbf{x})\right)=0.$$

PROBLEM 2.13 (Dual space to a Lie algebra). Let $\mathfrak{g}$ be a finite-dimensional Lie algebra with a Lie bracket $[\ ,\ ]$, and let $\mathfrak{g}^*$ be its dual space. For $f,g\in C^\infty(\mathfrak{g}^*)$ define

$$\{f,g\}(u)=u\left([df,dg]\right),$$

where $u\in\mathfrak{g}^*$ and $T^*_u\mathfrak{g}^*\simeq\mathfrak{g}$. Prove that $\{\ ,\ \}$ is a Poisson bracket. (It was introduced by Sophus Lie and is called *linear*, or *Lie-Poisson* bracket.) Show that bracket is degenerate and determine the center of $\mathcal{A}=C^\infty(\mathfrak{g}^*)$.

PROBLEM 2.14. A Poisson bracket $\{\ ,\ \}$ on $\mathscr{M}$ restricts to a Poisson bracket $\{\ ,\ \}_0$ on a submanifold $\mathscr{N}$, if the inclusion $\imath:\mathscr{N}\to\mathscr{M}$ is a Poisson mapping. Show that the Lie-Poisson bracket on $\mathfrak{g}^*$ restricts to a non-degenerate Poisson bracket ona coadjoints orbit, associated with Kirillov-Kostant symplectic form.

PROBLEM 2.15 (Lie-Poisson groups). A finite-dimensional Lie group is called a *Lie-Poisson group* if it has a structure of a Poisson manifold $(G,\{\ ,\ \})$ such that the group multiplication $G\times G\to G$ is a Poisson mapping, where $G\times G$ is a direct product of Poisson manifolds. Using a basis $\partial_1,\dots,\partial_n$ of left-invariant vector fields on G corresponding to a basis $x_1,\dots,x_n$ of the Lie algebra $\mathfrak{g}$, the Poisson bracket $\{\ ,\ \}$ can be written as

$$\{f_1,f_2\}(g)=\sum_{i,j=1}^{n}\eta^{ij}(g)\partial_i f_1\partial_j f_2,$$

where 2-tensor $\eta^{ij}(g)$ defines a mapping $\eta:G\to\Lambda^2\mathfrak{g}$ by $\eta(g)=\sum_{i,j=1}^n\eta^{ij}(g)x_i\otimes x_j$. Show that the bracket $\{\ ,\ \}$ equips G with a Lie-Poisson structure if and only if the following conditions are satisfied: (i) for all $g\in G$,

$$\begin{aligned}\xi^{ijk}(g)=&\sum_{l=1}^{n}\left(\eta^{il}(g)\partial_l\eta^{jk}(g)+\eta^{jl}(g)\partial_l\eta^{ki}(g)+\eta^{kl}(g)\partial_l\eta^{ij}(g)\right)\\&+\sum_{l,p=1}^{n}\left(c^i_{lp}\eta^{pj}(g)\eta^{kl}(g)+c^j_{lp}\eta^{pk}(g)\eta^{il}(g)+c^k_{lp}\eta^{pi}(g)\eta^{jl}(g)\right)=0,\end{aligned}$$

where $[x_i,x_j]=\sum_{k=1}^n c^k_{ij}x_k$; (ii) the mapping η is a group 1-cocycle with the adjoint action on $\Lambda^2\mathfrak{g}$, i.e., $\eta(g_1g_2)=\mathrm{Ad}^{-1}g_2\cdot\eta(g_1)+\eta(g_2),\ g_1,g_2\in G$.

PROBLEM 2.16. Show that the second condition in the previous problem trivially holds when η is a coboundary, $\eta(g)=-r+\mathrm{Ad}^{-1}g\cdot r$ for some $r=\sum_{i,j=1}^n r^{ij}x_i\otimes x_j\in\Lambda^2\mathfrak{g}$, and then the first condition is satisfied if and only if the element

$$\xi(r)=[r_{12},r_{13}+r_{23}]+[r_{13},r_{23}]\in\Lambda^3\mathfrak{g}$$

is invariant under the adjoint action of $\mathfrak{g}$ on $\Lambda^3\mathfrak{g}$. Here $r_{12}=\sum_{i,j=1}^n r^{ij}x_i\otimes x_j\otimes 1$, $r_{13}=\sum_{i,j=1}^n r^{ij}x_i\otimes 1\otimes x_j$ and $r_{23}=\sum_{i,j=1}^n r^{ij}1\otimes x_i\otimes x_j$ are corresponding elements in the universal enveloping algebra $U\mathfrak{g}$ of a Lie algebra $\mathfrak{g}$. In particular, G is a Lie-Poisson group if $\xi(r)=0$, which is called the *classical Yang-Baxter equation*.

PROBLEM 2.17. Suppose that $r=\sum_{i,j=1}^n r^{ij}x_i\otimes x_j\in\Lambda^2\mathfrak{g}$ is such that the matrix $\{r^{ij}\}$ is non-degenerate, and let $\{r_{ij}\}$ be the inverse matrix. Show r satisfies the classical Yang-Baxter equation if and only if the map $c:\Lambda^2\mathfrak{g}\to\mathbb{C}$, defined by

$c(x,y) = \sum_{i,j=1}^n r_{ij}u^iv^j$, where $x = \sum_{i=1}^n u^ix_i$, $y = \sum_{i=1}^n v^ix_i$, is a non-degenerate Lie algebra 2-cocycle, i.e., it satisfies

$$c(x,[y,z]) + c(z,[x,y]) + c(y,[z,x]) = 0, \quad x,y,z, \in \mathfrak{g}.$$

2.8. Hamilton's and Liouville's representations. We complete the formulation of classical mechanics by describing the process of *measurement*.

In physics, by a measurement of a classical system we understand the result of a physical experiment which gives numerical values for classical observables. The experiment consists of creating certain conditions for the system and it is always assumed that these conditions can be repeated over and over. The conditions of the experiment define a *state* of the system, if repeating these conditions results in probability distributions for the values of all observables of the system.

Mathematically, a state μ on the algebra $\mathcal{A} = C^\infty(\mathscr{M})$ of classical observables on the phase space $\mathscr{M}$ is the assignment

$$\mathsf{A} \ni f \mapsto \mu_f \in \mathscr{P}(\mathbb{R}),$$

where $\mathscr{P}(\mathbb{R})$ is a set of probability measures on $\mathbb{R}$ — Borel measures on $\mathbb{R}$ such that the total measure of $\mathbb{R}$ is 1. For every Borel subset $E \subset \mathbb{R}$ the quantity $0 \le \mu_f(E) \le 1$ is a probability that in the state μ the value of the observable f belongs to E. By definition, the *expectation value* of an observable f in the state μ is given by the Lebesgue-Stieltijes integral

$$\mathsf{E}_\mu(f) = \int_{-\infty}^{\infty} \lambda d\mu_f(\lambda),$$

where $\mu_f(\lambda) = \mu_f\left((-\infty,\lambda)\right)$ is a distribution function of the measure $d\mu_f$. The correspondce $f \mapsto \mu_f$ should should satisfy the following natural properties.

S1. $|\mathsf{E}_\mu(f)| < \infty$ for $f \in \mathcal{A}_0$ — the subalgebra of bounded observables.

S2. $\mathsf{E}_\mu(1) = 1$, where 1 is the unit in $\mathcal{A}$.

S3. For all $a,b \in \mathbb{R}$ and $f,g \in \mathcal{A}$,

$$\mathsf{E}_\mu(af + bg) = a\mathsf{E}_\mu(f) + b\mathsf{E}_\mu(g),$$

if both $\mathsf{E}_\mu(f)$ and $\mathsf{E}_\mu(g)$ exist.

S4. If $f_1 = \varphi \circ f_2$ with smooth $\varphi : \mathbb{R} \to \mathbb{R}$, then for every Borel subset $E \subset \mathbb{R}$,

$$\mu_{f_1}(E) = \mu_{f_2}(\varphi^{-1}(E)).$$

It follows from the property **S4** and a definition of Lebesgue-Stieltijes integral, that

$$\mathsf{E}_\mu(\varphi(f)) = \int_{-\infty}^{\infty} \varphi(\lambda) d\mu_f(\lambda).$$

In particular, $\mathsf{E}_\mu(f^2) \ge 0$ for all $f \in \mathcal{A}$, so that the states define normalized, positive, linear functionals on the subalgebra $\mathcal{A}_0$.

Assuming that the functional E_μ can be extended to a bounded, piecewise continuous functions on $\mathscr{M}$, one can recover the distribution function from the expectation values by the formula

$$\mu_f(\lambda) = \mathsf{E}_\mu\left(\theta(\lambda - f)\right),$$

where $\theta(x)$ is Heavyside step function,

$$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0. \end{cases}$$

Indeed, setting $\theta_\lambda(x) = \theta(\lambda - x)$, we get

$$\mu_{\theta_\lambda(f)}((-\infty, s)) = \mu_f\left(\theta_\lambda^{-1}(-\infty, s)\right) = \begin{cases} 1, & s \ge 1, \\ \mu_f([\lambda, \infty)), & 0 < s < 1, \\ 0, & s \le 0, \end{cases}$$

so that

$$\mathsf{E}_\mu(\theta(\lambda - f)) = \int_{-\infty}^{\infty} s d\mu_{\theta_\lambda(f)}(s) = 1 - \mu_f([\lambda, \infty)) = \mu_f(\lambda).$$

A probability measure $d\mu$ on $\mathscr{M}$ defines the state μ on $\mathcal{A}$ by assigning[14] to every observable f a probability measure μ_f on $\mathbb{R}$ with the distribution function

$$\mu_f(\lambda) = \int\limits_{\mathscr{M}} \theta(\lambda - f) d\mu = \int\limits_{\mathscr{M}_\lambda(f)} d\mu,$$

where $\mathscr{M}_\lambda(f) = \{x \in \mathscr{M} : f(x) < \lambda\}$. It follows from the Fubini theorem that

$$E_\mu(f) = \int_{-\infty}^{\infty} \lambda d\mu_f(\lambda) = \int\limits_{\mathscr{M}} f d\mu. \tag{2.11}$$

Conversely, for locally compact $\mathscr{M}$ the Riesz-Markov theorem states that for every positive, linear functional l on the space $C_{\mathrm{c}}(\mathscr{M})$ of continuous functions on $\mathscr{M}$ with compact support, there exists a unique regular Borel measure $d\mu$ on $\mathscr{M}$ such that

$$l(f) = \int\limits_{\mathscr{M}} f d\mu \quad \text{for all} \quad f \in C_{\mathrm{c}}(\mathscr{M}).$$

This leads to the following definition.

Definition. The set of states $\mathcal{S}$ for a Hamiltonian system with the phase space $\mathscr{M}$ is a set $\mathscr{P}(\mathscr{M})$ of probability measures on $\mathscr{M}$. For every $\mu \in \mathcal{S}$ and $f \in \mathcal{A}$ the distribution function μ_f is defined by

$$\mu_f(\lambda) = \int\limits_{\mathscr{M}} \theta(\lambda - f) d\mu = \int\limits_{\mathscr{M}_\lambda(f)} d\mu.$$

[14]There should be no confusion in denoting the state and the measure by μ.

The expectation values of classical observables are given by (2.11). The states corresponding to Dirac measures $d\mu_x$ supported at points $x \in \mathscr{M}$ are called *pure states*; all other states are called *mixed states*.

Physically, pure states are characterized by the property that a measurement of every observable in the pure state gives a well-defined result. Mathematically this can be expressed as follows. Let

$$\sigma_\mu^2(f) = \mathsf{E}_\mu\left((f - \mathsf{E}_\mu(f))^2\right) = \mathsf{E}_\mu(f^2) - \mathsf{E}_\mu(f)^2 \geq 0$$

be the *dispersion* of the observable f in the state μ.

LEMMA 2.1. *Thus pure states are the only states in which every observable has zero dispersion.*

PROOF. It follows from the Cauchy-Bunyakovskii-Schwarz inequality that $\sigma_\mu^2(f) = 0$ if only if f is constant on the support of a probability measure $d\mu$. □

In particular, a *mixture* of a pure states $d\mu_x$ and $d\mu_y$, $x, y \in \mathscr{M}$, is a mixed state with the measure

$$d\mu = \alpha d\mu_x + (1-\alpha)d\mu_y, \quad 0 < \alpha < 1,$$

and $\sigma_\mu^2(f) > 0$ for every observable f such that $f(x) \neq f(y)$.

For a system consisting of few interacting particles (say, a motion of planets in celestian mechanics) it is possible to measure all coordinates and momenta, so one considers only pure states. Mixed states necessarily appear for *macroscopic* systems, when it is impossible to measure all coordinates and momenta[15].

We end this chapter by presenting two equivalent ways of describing the dynamics of a Hamiltonian system $((\mathscr{M}, \{\,,\,\}), H)$ with the algebra of observables $\mathcal{A} = C^\infty(\mathscr{M})$ and the set of states $\mathcal{S} = \mathscr{P}(\mathscr{M})$.

1. **Hamilton's description of dynamics.** States do not depend on time, and time evolution of observables is given by Hamilton's equations of motion,

$$\frac{d\mu}{dt} = 0, \;\; \mu \in \mathcal{S}, \quad \text{and} \quad \frac{df}{dt} = \{H, f\} \;\; f \in \mathcal{A}.$$

The expectation value of an observable f in the state μ at time t is given by[16]

$$\mathsf{E}_\mu(f_t) = \int\limits_{\mathscr{M}} f(g^t(x))d\mu(x).$$

[15]Typically, a macroscopic system consists of $N \sim 10^{23}$ molecules. Macroscopic systems are studied in classical statistical mechanics.

[16]Assuming that Hamiltonian vector field X_H is complete so that the phase flow g^t exists.

In particular, the expectation value of f in the pure state $d\mu_x$ corresponding to the point $x \in \mathscr{M}$ is given by $f(g^t(x))$. Hamilton's picture is commonly used for mechanical systems consisting of few interacting particles.

2. **Liouville's description of dynamics.** The observables do not depend on time
$$\frac{df}{dt} = 0, \quad f \in \mathcal{A},$$
and states $d\mu(x) = \rho(x)dx$ satisfy Liouville's equation.
$$\frac{d\rho}{dt} = -\{H, \rho\}, \quad \rho(x)dx \in \mathcal{S}.$$
Here dx is a volume form on $\mathscr{M}$ invariant under the phase flow, whose existence is assumed[17], $\rho(x) = \dfrac{d\mu}{dx}$ is the distributional Radon-Nikodim derivative, and the Liouville equation is understood in the distributional sense. The expectation value of an observable f in the state μ at time t is given by
$$\mathsf{E}_{\mu_t}(f) = \int\limits_{\mathscr{M}} f(x)\rho(g^{-t}(x))dx.$$

Liouiville's picture, where states are given by generalized distribution functions on $\mathscr{M}$, is commonly used in statistical mechanics. The equality
$$\mathsf{E}_{\mu}(f_t) = \mathsf{E}_{\mu_t}(f) \quad \text{for all} \quad f \in \mathcal{A},\, \mu \in \mathcal{S},$$
which follows from the invariance of the volume form dx and the change of variables, expresses the equivalence between Liouville's and Hamilton's descriptions of dynamics.

3. Notes and references

Classical references are the textbooks [**Arn89**] and [**LL76**], which are written, respectively, from mathematics and physics perspectives. The treatise [**AM78**] and the encyclopaedia surveys [**AG90**], [**AKN97**] provide a comprehensive exposition, including the history and the references to classical works and recent contributions. Monographs [**DFN84**], [**DFN85**] and lecture notes [**Bry95**] contain all necessary material from differential geometry and theory of Lie groups, and the reference to other sources. Most of the problems are fairly standard and are taken from [**Arn89**],[**LL76**], [**Bry95**] and [**DFN84**]. Problems 2.8 and 2.14 introduce the reader to the orbit method [**Kir04**], and Problem 2.12 — to the method of symplectic reduction (see [**Arn89**], [**Bry95**] and references therein). Problems 2.15 –2.17 introduce the reader to the theory of Lie-Poisson groups (see [**Dri86**], [**Dri87**],[**STS85**], and [**Tak90**] for an elementary exposition).

[17] It is the Liouville volume form when Poisson structure on $\mathscr{M}$ is non-degenerate.

CHAPTER 2

Foundations of Quantum Mechanics

We recall the standard notations and basic facts from the theory of self-adjoint operators on Hilbert spaces. Let $\mathscr{H}$ be a separable Hilbert space with an inner product $(\ ,\ )$ and let A be a linear operator in $\mathscr{H}$ with the domain $D(A) \subset \mathscr{H}$ — a linear subset of $\mathscr{H}$. Operator A is called closed if its graph $\Gamma(A) = \{(\varphi, A\varphi) \in \mathscr{H} \times \mathscr{H} : \varphi \in D(A)\}$ is a closed subspace in $\mathscr{H} \times \mathscr{H}$. If domain of A is dense[1] in $\mathscr{H}$, i.e., $\overline{D(A)} = \mathscr{H}$, the domain $D(A^*)$ of the adjoint operator A^* consists of $\varphi \in \mathscr{H}$ such that there is $\eta \in \mathscr{H}$ with the property that

$$(A\psi, \varphi) = (\psi, \eta) \quad \text{for all} \quad D(A),$$

and the operator A^* is defined by $A^*\varphi = \eta$. Operator A is called symmetric if

$$(A\varphi, \psi) = (\varphi, A\psi) \quad \text{for all} \quad \varphi, \psi \in D(A).$$

By definition, the regular set of a closed operator A with a dense domain $D(A)$ is the set

$$\rho(A) = \{\lambda \in \mathbb{C} \,|\, A - \lambda I : D(A) \to \mathscr{H} \text{ is a bijection with a bounded inverse}^2\},$$

and for $\lambda \in \rho(A)$, the bounded operator $R_\lambda(A) = (A - \lambda I)^{-1}$ is called the resolvent of A at λ. The regular set $\rho(A) \subset \mathbb{C}$ is open and its complement $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is the spectrum of A. The subset $\sigma_p(A)$ of $\sigma(A)$ consisting of eigenvalues of A is called the point spectrum.

An operator A is self-adjoint (or Hermitian) if $A = A^*$. Equivalently, A is symmetric and $D(A) = D(A^*)$, and for such operators $\sigma(A) \subset \mathbb{R}$. A symmetric operator A is called essentially self-adjoint if its closure $\bar{A} = A^{**}$ is self-adjont. For a symmetric operator A the following conditions are equivalent:

(i) A is essentially self-adjoint.
(ii) $\ker(A^* + iI) = \ker(A^* - iI) = \{0\}$.
(iii) $\overline{\mathrm{Im}(A + iI)} = \overline{\mathrm{Im}(A - iI)} = \mathscr{H}$.

A symmetric operator A with $D(A) = \mathscr{H}$ is bounded and self-adjoint. An operator A is positive if $(A\varphi, \varphi) \geq 0$ for all $\phi \in D(A)$, which we denote by $A \geq 0$. Positive operators satisfy the Cauchy-Bunyakovski-Schwarz

[1] We consider only linear operators with dense domains.

[2] By the closed graph theorem, the last condition is redundant.

inequality

$$|(A\varphi,\psi)|^2 \le (A\varphi,\varphi)(A\psi,\psi) \quad \text{for all} \quad \varphi,\psi \in D(A). \tag{0.1}$$

In particular, $(A\varphi,\varphi)=0$ implies that $A\varphi=0$. Every bounded positive operator is self-adjoint[3]. We denote by $\mathscr{L}(\mathscr{H})$ the Banach algebra of bounded linear operators on $\mathscr{H}$. Compact operator A is of trace class, if

$$\sum_{n=1}^{\infty} \mu_n(A) < \infty,$$

where $\mu_n(A)$ are singular values of A, $\mu_n(A)=\sqrt{\lambda_n(A)} \ge 0$, where $\lambda_n(A)$ are eigenvalues for A^*A. A bounded operator A is of trace class if and only if for every orthonormal basis $\{e_n\}_{n=1}^{\infty}$ for $\mathscr{H}$,

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

Since a permutation of an orthonormal basis is again an orthonormal basis, this condition is equivalent to

$$\sum_{n=1}^{\infty} (Ae_n,e_n) < \infty$$

for every orthonormal basis $\{e_n\}_{n=1}^{\infty}$ for $\mathscr{H}$. The trace of a trace class operator A is defined by

$$\operatorname{Tr} A = \sum_{n=1}^{\infty} (Ae_n,e_n),$$

and does not depend on the choice of an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ for $\mathscr{H}$. Operators of trace class form a two-sided ideal $\mathscr{S}_1$ (von Neumann-Schatten ideal) in the Banach algebra $\mathscr{L}(\mathscr{H})$ and

$$\operatorname{Tr} AB = \operatorname{Tr} BA \quad \text{for all} \quad A \in \mathscr{S}_1,\ B \in \mathscr{L}(\mathscr{H})$$

— the cyclic property of the trace. Bounded positive operator A is of trace class if there is an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ for $\mathscr{H}$ such that

$$\sum_{n=1}^{\infty} (Ae_n,e_n) < \infty.$$

An operator $A \in \mathscr{L}(\mathscr{H})$ is Hilbert-Schmidt if $AA^* \in \mathscr{S}_1$. The vector space $\mathscr{S}_2$ of Hilbert-Schmidt operators in $\mathscr{H}$ is a Hilbert space with the inner product $(A,B)_2 = \operatorname{Tr} AB^*$. The Hilbert-Schmidt space $\mathscr{S}_2 \subset \mathscr{S}_1$ is also a two-sided ideal in the Banach algebra $\mathscr{L}(\mathscr{H})$.

[3]This is true only for complex Hilbert spaces.

1. Observables and States

1.1. Physical principles. Quantum mechanics studies the microworld — the physical laws at the atomic scale. The properties of the microworld are so different from our everyday's experience that there is no surprise that its laws seem to contradict the common sense. The need for quantum mechanics is a breakdown of classical mechanics at atomic level, its inadequacy to describe the properties of microscopic systems. Thus classical mechanics and classical electrodynamics can not explain stability of atoms and molecules. Neither can these theories reconcile different properties of light, its wave-like behavior in interference and diffraction phenomena and its particle-like behavior in photo-electric emission and scattering by free photons.

We will not discuss here these and other basic experimental facts, referring the interested reader to physics textbooks. Nor will we follow the historic path of the theory. Instead, we show how to formulate quantum mechanics using the general notions of states, observables and time evolution, described in the previous chapter. The departure from classical mechanics is that we will realize these notions differently. The fundamental difference between microworld and the perceived world around us is that in the microworld every experiment results in interaction with the system and thus disturbs its properties, whereas in classical physics it is always assumed that one can neglect the disturbances the measurement brings upon a system. This imposes a limitation on our powers of observation and leads to a conclusion that there exist observables which can not be measured simultaneously.

Mathematically, this means that observables in quantum mechanics no longer commute. Indeed, according to Gelfand-Naimark theorem, every semi-simple commutative Banach algebra with unit is an algebra of continuous functions on a compact topological space, the spectrum of the algebra. This is the situation we have in classical mechanics, where the spectrum of the algebra of classical observables is the phase space. An example of a non-commutative algebra is given by the Banach algebra of bounded operators on a complex Hilbert space, and it is this algebra which plays a fundamental role in quantum mechanics. Here we formulate the basic principles of quantum mechanics in the precise mathematical form. At this point it should be noted that one can not verify directly the principles lying in the foundation of quantum mechanics. Nevertheless, the validity of quantum mechanics, whenever it is applicable, is continuously being confirmed by numerous experimental facts which perfectly agree with the predictions of the theory[4].

1.2. Basic axioms.

[4]This refers to non-relativistic phenomena at atomic scale.

A1. With every quantum system there is an associated separable complex Hilbert space $\mathscr{H}$, in physics terminology called the *space of states*[5].

A2. The set of *observables* $\mathscr{A}$ of a quantum system with the Hilbert space $\mathscr{H}$ consists of all self-adjoint operators on $\mathscr{H}$. The subset $\mathscr{A}_0 = \mathscr{A} \cap \mathscr{L}(\mathscr{H})$ of bounded observables is a vector space over $\mathbb{R}$.

A3. Set of *states* $\mathscr{S}$ of a quantum system with a Hilbert space $\mathscr{H}$ consists of all positive (and hence self-adjoint) $M \in \mathscr{S}_1$ such that $\operatorname{Tr} M = 1$. *Pure states* are projection operators onto one-dimensional subspaces of $\mathscr{H}$. For $\psi \in \mathscr{H}$, $\|\psi\| = 1$, the corresponding projection is denoted by P_ψ. All other states are called *mixed states*[6].

A4. The measurement is a correspondence

$$\mathscr{A} \times \mathscr{S} \ni (A, M) \mapsto \mu_A \in \mathscr{P}(\mathbb{R}),$$

which to every observable $A \in \mathscr{A}$ and state $M \in \mathscr{S}$ assigns a probability measure μ_A on $\mathbb{R}$. For every Borel subset $E \subset \mathbb{R}$, the quantity $0 \le \mu_A(E) \le 1$ is the probability that for a quantum system in the state M the result of a measurement of an observable A belongs to E. The expectation value (the mean-value) of an observable $A \in \mathscr{A}$ in a state $M \in \mathscr{S}$ is

$$\langle A|M\rangle = \int_{-\infty}^{\infty} \lambda d\mu_A(\lambda),$$

where $\mu_A(\lambda) = \mu_A((-\infty, \lambda))$ is a distribution function for the probability measure μ_A.

The set of states $\mathscr{S}$ is a convex set. According to the Hilbert-Schmidt theorem on the canonical decomposition for compact self-adjoint operators, for every $M \in \mathscr{S}$ there exists (finite or infinite) orthonormal set $\{\psi_n\}_{n=1}^N$ in $\mathscr{H}$ such that

$$M = \sum_{n=1}^{N} \alpha_n P_{\psi_n} \quad \text{and} \quad \operatorname{Tr} M = \sum_{n=1}^{N} \alpha_n = 1, \tag{1.1}$$

where $\alpha_n > 0$ are non-zero eigenvalues of M. Thus every mixed state is a convex linear combination of pure states. The following result characterizes the pure states.

LEMMA 1.1. *A state $M \in \mathscr{S}$ is a pure state if and only if it can not be represented as a non-trivial convex linear combination in $\mathscr{S}$.*

PROOF. Clearly, if $M \in \mathscr{S}$ can not be represented as a non-trivial convex linear combination of states, then $M = P_\psi$ for some $\psi \in \mathscr{H}$, $\|\psi\| = 1$. Conversely, suppose that

$$P_\psi = aM_1 + (1-a)M_2, \quad 0 < a < 1,$$

[5]Space of pure states, to be precise.

[6]In physics terminology, operator M is called the *density operator*.

and let $\mathscr{H} = \mathbb{C}\psi \oplus \mathscr{H}_1$ be the orthogonal sum decomposition. Since M_1 and M_2 are positive operators, for $\varphi \in \mathscr{H}_1$ we have

$$a(M_1\varphi, \varphi) \le (P_\psi \varphi, \varphi) = 0,$$

so that $(M_1\varphi, \varphi) = 0$ for all $\varphi \in \mathscr{H}_1$ and by (0.1) we get $M_1|_{\mathscr{H}_1} = 0$. Since M_1 is self-adjoint, it leaves the complimentary subspace $\mathbb{C}\psi$ invariant, and from $\operatorname{Tr} M_1 = 1$ it follows that $M_1 = P_\psi$, and therefore, $M_2 = P_\psi$. □

Explicit construction of the correspondence $\mathscr{A} \times \mathscr{S} \to \mathscr{P}(\mathbb{R})$ is based on the general spectral theorem of von Neumann, which emphasizes the fundamental role self-adjoint operators play in quantum mechanics.

Definition. A projection-valued measure on $\mathbb{R}$ is a mapping $\mathsf{P} : \mathscr{B}(\mathbb{R}) \to \mathscr{L}(\mathscr{H})$ of the σ-algebra $\mathscr{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ into the Banach algebra of bounded operators on $\mathscr{H}$, satisfying the following properties,

PM1. For every Borel subset $E \subset \mathbb{R}$, $\mathsf{P}(E)$ is an orthogonal projection, i.e., $\mathsf{P}(E) = \mathsf{P}(E)^2$ and $\mathsf{P}(E) = \mathsf{P}(E)^*$.

PM2. $\mathsf{P}(\emptyset) = 0$, $\mathsf{P}(\mathbb{R}) = I$, the identity operator on $\mathscr{H}$.

PM3. For every disjoint union of Borel subsets,

$$E = \coprod_{n=1}^{\infty} E_n, \quad \mathsf{P}(E) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathsf{P}(E_i)$$

in the strong topology on $\mathscr{L}(\mathscr{H})$.

It follows from properties **PM1-PM3** that

$$\mathsf{P}(E_1)\mathsf{P}(E_2) = \mathsf{P}(E_1 \cap E_2) \quad \text{for all } E_1, E_2 \in \mathscr{B}(\mathbb{R}). \tag{1.2}$$

With every projection-valued measure P we associate a projection-valued function

$$\mathsf{P}(\lambda) = \mathsf{P}((-\infty, \lambda)),$$

called the projection-valued resolution of the identity. It is characterized by the following properties.

PD1.

$$\mathsf{P}(\lambda)\mathsf{P}(\mu) = \mathsf{P}(\min\{\lambda, \mu\}).$$

PD2.

$$\lim_{\lambda\to-\infty} \mathsf{P}(\lambda) = 0, \quad \lim_{\lambda\to\infty} \mathsf{P}(\lambda) = I.$$

PD3.

$$\lim_{\mu\to\lambda-0} \mathsf{P}(\mu) = \mathsf{P}(\lambda).$$

For every $\varphi \in \mathscr{H}$ the resolution of the identity $\mathsf{P}(\lambda)$ defines a distribution function $(\mathsf{P}(\lambda)\varphi, \varphi)$ of the bounded measure on $\mathbb{R}$ (probability measure when $\|\varphi\| = 1$). By the polarization identity

$$\begin{aligned}(\mathsf{P}(\lambda)\varphi, \psi) = \tfrac{1}{4}\{&(\mathsf{P}(\lambda)(\varphi+\psi), \varphi+\psi) - (\mathsf{P}(\lambda)(\varphi-\psi), \varphi-\psi) \\ &+ i(\mathsf{P}(\lambda)(\varphi+i\psi), \varphi+i\psi) - i(\mathsf{P}(\lambda)(\varphi-i\psi), \varphi-i\psi)\},\end{aligned}$$

so that $(\mathsf{P}(\lambda)\varphi, \psi)$ corresponds to a complex measure on $\mathbb{R}$ — a complex linear combination of measures.

A measurable function f on $\mathbb{R}$ is said to be finite almost everywhere (a.e.) with respect to the projection-valued measure P, if it is finite a.e. with respect to all measures $(\mathsf{P}\psi, \psi)$, $\psi \in \mathscr{H}$. For separable $\mathscr{H}$ a theorem of von Neumann states that for every projection-valued measure P there exists $\varphi \in \mathscr{H}$ such that a function f is finite a.e. with respect to P if and only if it is finite a.e. with respect to the measure $(\mathsf{P}\varphi, \varphi)$.

The next statement is the celebrated general spectral theorem of von Neumann.

THEOREM 1.1 (von Neumann). *For every self-adjoint operator A on the Hilbert space $\mathscr{H}$ there exists a unique resolution of the identity $\mathsf{P}(\lambda)$, satisfying the following properties.*

(i)
$$D(A) = \left\{ \varphi \in \mathscr{H} \;\middle|\; \int_{-\infty}^{\infty} \lambda^2 d(\mathsf{P}(\lambda)\varphi, \varphi) < \infty \right\},$$
and for every $\varphi \in D(A)$
$$A\varphi = \int_{-\infty}^{\infty} \lambda \, d\mathsf{P}(\lambda)\varphi,$$
defined as a limit of Riemann-Stieltjes sums in the strong topology on $\mathscr{H}$. The support of corresponding projection-valued measure P coincides with the spectrum of A.

(ii) *For every continuous function f on $\mathbb{R}$, $f(A)$ is a linear operator on $\mathscr{H}$ with a dense domain*
$$D(f(A)) = \left\{ \varphi \in \mathcal{H} \;\middle|\; \int_{-\infty}^{\infty} f(\lambda)^2 d(\mathsf{P}(\lambda)\varphi, \varphi) < \infty \right\},$$
defined for $\varphi \in D(f(A))$ as
$$f(A)\varphi = \int_{-\infty}^{\infty} f(\lambda) d\mathsf{P}(\lambda)\varphi,$$
and understood as in part (i). The operator $f(A)$ satisfies
$$f(A)^* = \bar{f}(A),$$
where $\bar{f}$ is the complex conjugate function to f, and the operator $f(A)$ is bounded if and only if the function f is bounded. For bounded continuous functions f and g,
$$f(A)g(A)\varphi = \int_{-\infty}^{\infty} f(\lambda)g(\lambda) d\mathsf{P}(\lambda)\varphi, \quad \varphi \in \mathscr{H}.$$

(iii) *For every measurable function f on $\mathbb{R}$, finite a.e. with respect to the projection-valued measure P, $f(A)$ is a linear operator on $\mathscr{H}$*

with a dense domain $D(f(A))$ defined as in (ii), understood in the weak sense: for every $\varphi \in D(f(A))$ and $\psi \in \mathscr{H}$,

$$(f(A)\varphi, \psi) = \int_{-\infty}^{\infty} f(\lambda) d(\mathsf{P}(\lambda)\varphi, \psi),$$

— a Lebesgue-Stieltjes integral with respect to a complex measure. The correspondence $f \mapsto f(A)$ satisfies the same properties as in (ii).

(iv) *For every bounded operator B which commutes with A, that is, $B(D(A)) \subset D(A)$ and $AB = BA$ on $D(A)$, operator B commutes with the resolution of the identity $\mathsf{P}(\lambda)$ and, therefore, with every operator $f(A)$.*

We will denote the resolution of the identity for a self-adjoint operator A, given by the spectral theorem, by $\mathsf{P}_A(\lambda)$. Conversely, every resolution of the identity $\mathsf{P}(\lambda)$, as defined by properties **PD1**-**PD3**, by virtue of (i)-(ii) is a resolution of the identity for a self-adjoint operator. It follows from the spectral theorem that the spectrum of a self-adjoint A coincides with the support of its projection-valued measure P_A, i.e., $\lambda \in \sigma(A)$ if and only if $\mathsf{P}_A((\lambda - \varepsilon, \lambda + \varepsilon)) \neq 0$ for all $\varepsilon > 0$.

Now the correspondence $\mathscr{A} \times \mathscr{S} \to \mathscr{P}(\mathbb{R})$, postulated in **A4**, associates to every $M \in \mathscr{S}$ and $A \in \mathscr{A}$ a probability measure μ_A on $\mathbb{R}$, defined by the celebrated Born-von Neumann formula

$$\mu_A(E) = \operatorname{Tr} \mathsf{P}_A(E)M, \quad E \in \mathscr{B}(\mathbb{R}). \tag{1.3}$$

It follows from the Hilbert-Schmidt decomposition (1.1) that

$$\mu_A(E) = \sum_{n=1}^{N} \alpha_n (\mathsf{P}_A(E)\psi_n, \psi_n),$$

so indeed $0 \leq \mu_A(E) \leq 1$. We denote by $\mu_A(\lambda)$ the distribution function of the probability measure μ_A, $\mu_A(\lambda) = (\mathsf{P}_A(\lambda)\psi, \psi)$ for $M = P_\psi$.

LEMMA 1.2. *Suppose that an observable $A \in \mathscr{A}$ and a state $M \in \mathscr{S}$ are such that $\langle A|M\rangle$ exists and $AM \in \mathscr{S}_1$ (this is always the case when $A \in \mathscr{A}_0$). Then*

$$\langle A|M\rangle = \operatorname{Tr} AM.$$

In particular, if $M = P_\psi$ and $\psi \in D(A)$, then

$$\langle A|M\rangle = (A\psi, \psi) \quad \textit{and} \quad \langle A^2|M\rangle = \|A\psi\|^2.$$

PROOF. Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal basis for $\mathscr{H}$. It follows from the spectral theorem that

$$\begin{aligned}\langle A|M\rangle &= \int_{-\infty}^{\infty} \lambda d\mu_A(\lambda) = \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \lambda d(\mathsf{P}_A(\lambda)Me_n, e_n) \\ &= \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \lambda d(\mathsf{P}_A(\lambda)Me_n, e_n) = \sum_{n=1}^{\infty} (AMe_n, e_n) = \operatorname{Tr} AM,\end{aligned}$$

where the interchange of the summation and the integration is legitimate since $\int_{-\infty}^{\infty} |\lambda| d\mu_A(\lambda) < \infty$. In particular, when $M = P_\psi$ and $\psi \in D(A)$,

$$\langle A|M\rangle = \int_{-\infty}^{\infty} \lambda d(\mathsf{P}_A(\lambda)\psi, \psi) = (A\psi, \psi).$$

Finally, it follows from the spectral theorem and the change of variables formula that

$$\|A\psi\|^2 = \int_{-\infty}^{\infty} \lambda^2 d(\mathsf{P}_A(\lambda)\psi, \psi) = \int_0^{\infty} \lambda d(\mathsf{P}_{A^2}(\lambda)\psi, \psi) = \langle A^2|M\rangle.$$

□

REMARK. It is convenient to approximate a unbounded self-adjoint operator A by bounded operators $A_n = f_n(A)$, where $f_n = \chi_{[-n,n]}$ — a characteristic function of the interval $[-n, n]$. Assuming that $\langle A|M\rangle$ exists, we have

$$\langle A|M\rangle = \int_{-\infty}^{\infty} \lambda d\mu_A(\lambda) = \lim_{n\to\infty} \int_{-n}^{n} \lambda d\mu_A(\lambda) = \lim_{n\to\infty} \langle A_n|M\rangle.$$

DEFINITION. Self-adjoint operators A and B commute if the corresponding projection-valued measures P_A and P_B commute, $\mathsf{P}_A(E_1)\mathsf{P}_B(E_2) = \mathsf{P}_B(E_2)\mathsf{P}_A(E_1)$ for all $E_1, E_2 \in \mathscr{B}(\mathbb{R})$.

The following two results, which follow from the spectral theorem, are very useful in applications.

PROPOSITION 1.1. *The following statements are equivalent.*

(i) *Self-adjoint operators A and B commute.*

(ii) *For all $\lambda, \mu \in \mathbb{C}$, $\operatorname{Im}\lambda, \operatorname{Im}\mu \neq 0$,*

$$R_\lambda(A)R_\mu(B) = R_\mu(B)R_\lambda(A).$$

(iii) *For all $u, v \in \mathbb{R}$,*

$$e^{iuA}e^{ivB} = e^{ivB}e^{iuA}.$$

(iv) *For all $u \in \mathbb{R}$, the operators e^{iuA} and B commute.*

Slightly abusing notations[7], we will often write $[A, B] = AB - BA = 0$ for commuting self-adjoint operators A and B.

[7]In general, for unbounded self-adjoint operators A and B the commutator $[A, B] = AB - BA$ is not necessarily closed, i.e., it could be defined only for $\varphi = 0$.

PROPOSITION 1.2. *Let $\mathbf{A} = \{A_1, \dots, A_n\}$ be a finite set of self-adjoint, pair-wise commuting operators on $\mathscr{H}$. Then there exists a unique projection-valued measure $\mathsf{P}_{\mathbf{A}}$ on the Borel subsets of $\mathbb{R}^n$ having the following properties.*

(i) *In the strong operator topology,*

$$A_k = \int_{\mathbb{R}^n} \lambda_k d\mathsf{P}_{\mathbf{A}}, \quad k = 1, \dots, n,$$

where λ_k is the k-th coordinate function on $\mathbb{R}^n$.

(ii) *For every measurable function f on $\mathbb{R}^n$, finite a.e. with respect to the projection-valued measure $\mathsf{P}_{\mathbf{A}}$, $f(A_1, \dots, A_n)$ is a linear operator on $\mathscr{H}$ defined by*

$$f(A_1, \dots, A_n) = \int_{\mathbb{R}^n} f d\mathsf{P}_{\mathbf{A}},$$

where the integral us understood in the weak operator topology. The correspondence $f \mapsto f(A_1, \dots, A_n)$ satisfies the same properties as in the part (ii) of the spectral theorem.

The support of the projection-valued measure $\mathsf{P}_{\mathbf{A}}$ on $\mathbb{R}^n$ is called the joint spectrum of the commutative family $\mathbf{A} = \{A_1, \dots, A_n\}$.

REMARK. According to von Neumann theorem on a generating operator, for every commutative family $\mathbf{A}$ of self-adjoint operators (not necessarily finite) on a separable Hilbert space $\mathscr{H}$ there is a generating operator — a self-adjoint operator R on $\mathscr{H}$ such that all operators in $\mathbf{A}$ are functions of R.

It seems natural that simultaneous measurement of a finite set of observables $\mathbf{A} = \{A_1, \dots, A_n\}$ in the state $M \in \mathscr{S}$ should be described by the probability measure $\mu_{\mathbf{A}}$ on $\mathbb{R}^n$ given by the following generalization of the Born-von Neumann formula,

$$\mu_{\mathbf{A}}(\mathbf{E}) = \mathrm{Tr}(\mathsf{P}_{A_1}(E_1) \dots \mathsf{P}_{A_n}(E_n) M), \quad \mathbf{E} = E_1 \times \dots \times E_n \in \mathscr{B}(\mathbb{R}^n). \tag{1.4}$$

However, formula (1.4) defines a probability measure on $\mathbb{R}^n$ if and only if $\mathsf{P}(\mathbf{E}) = \mathsf{P}_{A_1}(E_1) \dots \mathsf{P}_{A_n}(E_n)$ is a projection-valued measure on $\mathbb{R}^n$. Since a product of orthogonal projections is an orthogonal projection only when the projection operators commute, we conclude that the operators $A_1, \dots, A_n$ form a commutative family. This result agrees with the requirement that simultaneous measurement of several observables should be independent of the order of the measurements of individual observables. We summarize these arguments as the following axiom.

A5. A finite set of observables $\mathbf{A} = \{A_1, \dots, A_n\}$ can be measured simultaneously (*simultaneously measured observables*) if and only if they form a commutative family. Simultaneous measurement of the commutative family $\mathbf{A} \subset \mathscr{A}$ in the state $M \in \mathscr{S}$ is described

by the probability measure $\mu_{\mathbf{A}}$ on $\mathbb{R}^n$ given by

$$\mu_{\mathbf{A}}(\mathbf{E}) = \operatorname{Tr} \mathsf{P}_{\mathbf{A}}(\mathbf{E})M, \quad \mathbf{E} \in \mathscr{B}(\mathbb{R}^n),$$

where $\mathsf{P}_{\mathbf{A}}$ is the projection-valued measure from Proposition 1.2. Explicitly, $\mathsf{P}_{\mathbf{A}}(\mathbf{E}) = \mathsf{P}_{A_1}(E_1) \dots \mathsf{P}_{A_n}(E_n)$ for $\mathbf{E} = E_1 \times \cdots \times E_n \in \mathscr{B}(\mathbb{R}^n)$. For every Borel subset $\mathbf{E} \subset \mathbb{R}^n$ the quantity $0 \le \mu_{\mathbf{A}}(\mathbf{E}) \le 1$ is the probability that for a quantum system in the state M the result of the simultaneous measurement of observables $A_1, \dots, A_n$ belongs to $\mathbf{E}$.

The axioms **A1**-**A5** are known as Dirac-von Neumann axioms.

PROBLEM 1.1. Prove property (1.2).

PROBLEM 1.2. Prove that the state M is a pure state if and only if $\operatorname{Tr} M^2 = 1$.

PROBLEM 1.3. Prove all the remaining statements in this section.

1.3. Heisenberg's uncertainty relations. The variance of the observable A in the state M, which measures the mean deviation of A from its expectation value, is defined by

$$\sigma_M^2(A) = \langle (A - \langle A|M\rangle I)^2 | M\rangle = \langle A^2|M\rangle - \langle A|M\rangle^2 \ge 0,$$

provided the expectation values $\langle A^2|M\rangle$ and $\langle A|M\rangle$ exist. It follows from Lemma 1.2 that for $M = P_\psi$, where $\psi \in D(A)$,

$$\sigma_M^2(A) = \|(A - \langle A|M\rangle I)\psi\|^2 = \|A\psi\|^2 - (A\psi, \psi)^2.$$

LEMMA 1.3. *For $A \in \mathscr{A}$ and $M \in \mathscr{S}$ the variance $\sigma_M(A) = 0$ if and only if $\operatorname{Im} M$ is an eigenspace for the operator A with the eigenvalue $a = \langle A|M\rangle$. In particular, if $M = P_\psi$, then ψ is an eigenvector of A, $A\psi = a\psi$.*

PROOF. It follows from the spectral theorem that

$$\sigma_M^2(A) = \int_{-\infty}^{\infty} (\lambda - a)^2 d\mu_A(\lambda),$$

so that $\sigma_M(A) = 0$ if and only if the probability measure μ_A is supported at the point $a \in \mathbb{R}$, i.e., $\mu_A(\{a\}) = 1$. Since $\mu_A(\{a\}) = \operatorname{Tr} P_A(\{a\})M$ and $\operatorname{Tr} M = 1$, we conclude that this is equivalent to $\operatorname{Im} M$ being an invariant subspace for $P_A(\{a\})$. □

Now we formulate generalized *Heisenberg's uncertainty relations.*

PROPOSITION 1.3 (H. Weyl). *Let $A, B \in \mathscr{A}$ and let $M = P_\psi$ be the pure state such that $\psi \in D(A) \cap D(B)$ and $A\psi, B\psi \in D(A) \cap D(B)$. Then*

$$\sigma_M^2(A)\sigma_M^2(B) \ge \tfrac{1}{4}\langle i[A, B]|M\rangle^2.$$

The same inequality holds for all $M \in \mathscr{S}$, where by definition $\langle i[A, B]|M\rangle = \lim_{n\to\infty}\langle i[A_n, B_n]|M\rangle$.

PROOF. Let $M = P_\psi$. Since

$$[A - \langle A|M\rangle I, B - \langle B|M\rangle I] = [A, B],$$

it is sufficient to prove the inequality

$$\langle A^2|M\rangle\langle B^2|M\rangle \geq \tfrac{1}{4}\langle i[A,B]|M\rangle^2.$$

We have for all $\alpha \in \mathbb{R}$,

$$\begin{aligned} 0 \leq \|(A + i\alpha B)\psi\|^2 =& \alpha^2(B\psi, B\psi) - i\alpha(A\psi, B\psi) + i\alpha(B\psi, A\psi) + (A\psi, A\psi) \\ =& \alpha^2(B^2\psi, \psi) + \alpha(i[A,B]\psi, \psi) + (A^2\psi, \psi), \end{aligned}$$

so that necessarily $4(A^2\psi, \psi)(B^2\psi, \psi) \geq (i[A,B]\psi, \psi)$. The same argument works for the mixed states. Since

$$\sigma_M^2(A)\sigma_M^2(B) = \lim_{n\to\infty} \sigma_M^2(A_n)\sigma_M^2(B_n)$$

(see the remark in the previous section), it is sufficient to prove the inequality for bounded A and B. Then using the cyclic property of the trace we have for all $\alpha \in \mathbb{R}$,

$$\begin{aligned} 0 \leq& \operatorname{Tr}((A + i\alpha B)M(A + i\alpha B)^*) = \operatorname{Tr}((A + i\alpha B)M(A - i\alpha B)) \\ =& \alpha^2 \operatorname{Tr} BMB + i\alpha \operatorname{Tr} BMA - i\alpha \operatorname{Tr} AMB + \operatorname{Tr} AMA \\ =& \alpha^2 \operatorname{Tr} MB^2 + \alpha \operatorname{Tr}(i[A,B]M) + \operatorname{Tr} MA^2, \end{aligned}$$

and the inequality follows. □

Heisenberg's uncertainty relations express quantitatively the fact that observables which do not commute can not be measured simultaneously, even in a pure state. This manifests a fundamental difference between the classical mechanics and the quantum mechanics.

1.4. Dynamics. Though quantum observables $\mathscr{A}$ do not form an algebra with respect to an operator product[8], a real vector space $\mathscr{A}_0$ of bounded obsevables has a structure of a Lie algebra with the bracket

$$i[A,B] = i(AB - BA), \quad A, B \in \mathscr{A}_0.$$

In analogy with classical mechanics, the time evolution of a quantum system is determined by a quantum observable $H \in \mathscr{A}$, called a *Hamiltonian operator* (Hamiltonian for brevitiy). The analog of the Hamilton's picture in classical mechanics (see Section 2.5 in Chapter 1) is the *Heisenberg's picture* in quantum mechanics, where the states do not depend on time and bounded observables satisfy *Heisenberg's equations of motion*

$$\frac{dA}{dt} = \{H, A\}_\hbar, \quad A \in \mathscr{A}_0, \tag{1.5}$$

where

$$\{\ ,\ \}_\hbar = \frac{i}{\hbar}[\ ,\ ] \tag{1.6}$$

[8]The product of two non commuting self-adjoint operators is not self-adjoint.

is called the *quantum bracket*. The positive number $\hbar$ — the *Planck constant*, is one of the fundamental constants in physics[9]. Introducing the one-parameter strongly continuous group $U(t)$ of unitary operators associated with the self-adjoint operator H,

$$U(t) = e^{-\frac{i}{\hbar}tH}, \quad t \in \mathbb{R}, \tag{1.7}$$

the solution to the Heisenberg's equations of motion with the initial condition $A(0) = A$ can be written as

$$A(t) = U(t)^{-1} A U(t). \tag{1.8}$$

The evolution operator $U_t : \mathscr{A}_0 \to \mathscr{A}_0$ is defined by $U_t(A) = U^{-1}(t)AU(t)$, and extends to the whole set $\mathscr{A}$ of quantum observables. The quantum dynamics of an observable $A \in \mathscr{A}$ is given by $A(t) = U_t(A)$, and in this sense all quantum observables (not necessarily bounded) satisfy Heisenberg's equations of motion. An observable A whose time evolution (1.8) does not depend on t is called a *quantum integral of motion*, or a *constant of motion*. An observable A is an integral of motion if and only if it commutes with the Hamiltonian H, so that, in agreement with (1.5),

$$\{A, H\}_\hbar = 0.$$

By Stone theorem, every strongly-continuous one-parameter group of uinitary operators[10] $U(t)$ is of the form (1.7), where

$$D(H) = \{\varphi \in \mathscr{H} : \lim_{t\to 0} \frac{U(t) - I}{t}\varphi \text{ exists}\} \quad \text{and} \quad H\varphi = i\hbar \lim_{t\to 0} \frac{U(t) - I}{t}\varphi.$$

Thus quantum dynamics is described by the strongly continuous one-parameter group of unitary operators.

The analog of the Liouville's picture in classical mechanics (see Section 2.5 in Chapter 1) is the *Schrödinger's picture* in quantum mechanics, where observables do not depend on time and the evolution of states is given by

$$\frac{dM}{dt} = -\{H, M\}_h, \quad M \in \mathscr{S}. \tag{1.9}$$

The solution to (1.9) with the initial condition $M(0) = M$ can be written as

$$M(t) = U(t)MU^{-1}(t). \tag{1.10}$$

It follows from the cyclic property of the trace that for $A \in \mathscr{A}_0$,

$$\langle A(t)|M\rangle = \mathrm{Tr}(U^{-1}(t)AU(t)M) = \mathrm{Tr}(AU(t)MU^{-1}(t)) = \langle A|M(t)\rangle,$$

which establishes the equivalence between Heisenberg's and Schrödinger's pictures. A quantum integral of motion A in Schrödinger's picture is defined

[9]The Planck constant has a physical dimension of the action. Its value $\hbar = 1.054 \times 10^{-27}$ erg $\times$ sec, which is determined from the experiment, manifests that quantum mechanics is a microscopic theory.

[10]According to von Neumann theorem, on a separable Hilbert space every weakly measurable one-parameter group of unitary operators is strongly continuous.

the property that the expectation value $\langle A|M(t)\rangle$, where $M(t)$ is given by (1.10), does not depend on t for any $M \in \mathscr{S}$.

It follows from (1.9) that the time evolution of a pure state $M = P_\psi$ is given by $M(t) = P_{\psi(t)}$, where $\psi(t) = U(t)\psi$. The vector $\psi(t)$ satisfies the *time-dependent Schrödinger equation*

$$i\hbar\frac{d\psi}{dt} = H\psi \tag{1.11}$$

with the initial condition $\psi(0) = \psi$. A state $M \in \mathscr{S}$ is called *stationary* if $M(t) = U(t)MU^{-1}(t)$ does not depend on time, i.e., $[M, U(t)] = 0$ for all t. According to Proposition 1.1, this is equivalent to $[M, H] = 0$.

LEMMA 1.4. *The pure state $M = P_\psi$ is stationary if and only if ψ is the eigenvector for H,*

$$H\psi = E\psi.$$

Corresponding eigenvalue E is called the energy and

$$\psi(t) = e^{-\frac{i}{\hbar}E}\psi.$$

PROOF. It follows from $U(t)P_\psi = P_\psi U(t)$ that ψ is a common eigenvector for unitary operators $U(t)$ for all t, $U(t)\psi = c(t)\psi$, $|c(t)| = 1$. Since $U(t)$ is strongly continuous one-parameter group of unitary operators, continuous function $c(t) = (U(t)\psi, \psi)$ satisfies the equation $c(t_1 + t_2) = c(t_1)c(t_2)$ for all $t_1, t_2 \in \mathbb{R}$, so that $c(t) = e^{-\frac{i}{\hbar}tE}$ for some $E \in \mathbb{R}$. Thus by the Stone theorem $\psi \in D(H)$ and $H\psi = E\psi$. □

The eigenvalue equation $H\psi = E\psi$ is called the *stationary Schrödinger equation.*

PROBLEM 1.4. Show that if an observable A is such that for every state M the expectation value $\langle A|M(t)\rangle$ does not depend on t, then A is a quantum integral of motion.

PROBLEM 1.5. Show that the solution of the initial value problem for the time-dependent Schrödinger equation (1.11) is given by

$$\psi(t) = \int_{-\infty}^{\infty} e^{-\frac{i}{\hbar}t\lambda} d\mathsf{P}(\lambda)\psi,$$

where $\mathsf{P}(\lambda)$ is the resolution of identity for the Hamiltonian H.

2. Quantization

To study the quantum system one needs to describe its Hilbert space of states $\mathscr{H}$ and the Hamiltonian H — a self-adjoint operator in $\mathscr{H}$ which defines the evolution of a system. When the quantum system has a classical analog, the procedure of constructing the corresponding Hilbert space $\mathscr{H}$ and the Hamiltonian H is called a *quantization* of a classical system. Heuristically, a quantization of a classical Hamiltonian system[11] $((\mathscr{M}, \{\ ,\ \}), h)$ is

[11] In this chapter we denote classical observables, including the Hamiltonian, by lowercase roman letters.

a one-to-one mapping $\mathrm{Q} : \mathcal{A} \to \mathscr{A}$ from the set of classical observables $\mathcal{A} = C^\infty(\mathscr{M})$ to the set $\mathscr{A}$ of quantum observables — the set of all self-adjoint operators on a Hilbert space $\mathscr{H}$. The map Q depends on a parameter $\hbar > 0$, is a linear mapping between the subspaces of bounded classical and quantum observables, and for all $f, g \in \mathcal{A}$ satisfies

$$\tfrac{1}{2}\mathrm{Q}^{-1}(\mathrm{Q}(f)\mathrm{Q}(g) + \mathrm{Q}(g)\mathrm{Q}(f)) \to fg, \quad \mathrm{Q}^{-1}(\{\mathrm{Q}(f), \mathrm{Q}(g)\}_\hbar) \to \{f, g\}$$

as $\hbar \to 0$. The latter property is the so-called *correspondence principle*. In particular $h \mapsto \mathrm{Q}(h) = H$ — the Hamiltonian for a quantum system.

Since quantum mechanics is different from classical mechanics, the correspondence $f \mapsto \mathrm{Q}(f)$ can not be an isomorphism between the Lie algebras of bounded classical and quantum observables with respect to classical and quantum brackets. It becomes an isomorphism only in the limit $\hbar \to 0$ when, according to the correspondence principle, quantum mechanics turns into the classical mechanics. Since quantum mechanics provides a more accurate and refine description then classical mechanics, quantization of a classical system may not be unique. However, for many "real" quantum systems — the systems describing actual physical phenomena, the corresponding Hamiltonian H is defined uniquely by its classical analog.

2.1. Heisenberg's commutation relations. The simplest classical system with one degree of freedom is described by the phase space $\mathbb{R}^2$ with coordinates p, q and the Poisson bracket $\{\ ,\ \}$, associated with the canonical symplectic form $\omega = dp \wedge dq$. In particular, the Poisson bracket between the classical observables p and q — the momentum and the coordinate of a particle, has the following simple form

$$\{p, q\} = 1. \tag{2.1}$$

It is another postulate of quantum mechanics that under the quantization, the classical observables p and q correspond to the quantum observables P and Q — self-adjoint operators P and Q on a Hilbert space $\mathscr{H}$, satisfyng the following properties.

C1. There is a dense $D \subset \mathscr{H}$ such that $P : D \to D$ and $Q : D \to D$.

C2. For all $\psi \in D$,

$$(PQ - QP)\psi = -i\hbar\psi.$$

C3. Every bounded operator on $\mathscr{H}$ which commutes with P and Q is a multiple of the identity operator I.

Property **C2** is called *Heisenberg's (canonical) commutation relation* for one degree of freedom. In terms of the quantum bracket (1.6) it take the form

$$\{P, Q\}_\hbar = I, \tag{2.2}$$

which is exactly the same as the Poisson bracket (2.1). The operators P and Q are called, respectively, the *momentum operator* and the *coordinate*

operator. The correspondence $p \mapsto P$, $q \mapsto Q$ with P and Q satisfying **C1-C3** is the cornerstone for the quantization of classical systems. The validity of (2.2), as well as of quantum mechanics as a whole, is confirmed by the agreement of the theory with numerous experiments.

It follows from the Heisenberg's uncertainty relations (see Proposition 1.3), that for any pure state $M = P_\psi$ with $\psi \in D$,

$$\sigma_M(P)\sigma_M(Q) \geq \frac{\hbar}{2}.$$

This is a fundamental result saying that it is impossible to measure the coordinate and the momentum of quantum particle simultaneously: the more accurate is the measurement of one quantity, the less accurate is the value of the other. It is often said that quantum particle has no observed path, so that "quantum motion" differs dramatically from the motion in classical mechanics.

Now it is straightforward to consider a classical system with n degrees of freedom, described by the phase space $\mathbb{R}^{2n}$ with coordinates $\mathbf{p} = (p^1, \dots, p^n)$ and $\mathbf{q} = (q_1, \dots, q_n)$, and the Poisson bracket $\{\ ,\ \}$, associated with the canonical symplectic form $\omega = d\mathbf{p} \wedge d\mathbf{q}$. The Poisson brackets between the classical observables $\mathbf{p}$ and $\mathbf{q}$ — the momenta and the coordinates of a particle, have the following form

$$\{q_k, q_l\} = 0, \quad \{p^k, p^l\} = 0, \quad \{p^k, q_l\} = \delta^k_l, \quad k, l = 1, \dots, n. \tag{2.3}$$

Corresponding quantum momenta and coordinate operators $\mathbf{P} = (P^1, \dots, P^n)$ and $\mathbf{Q} = (Q_1, \dots, Q_n)$ leave a dense domain $D \subset \mathscr{H}$ invariant and satisfy *Heisenberg's commutation relations* for n degrees of freedom,

$$\{Q_k, Q_l\}_\hbar = 0, \quad \{P^k, P^l\}_\hbar = 0, \quad \{P^k, Q_l\}_\hbar = \delta^k_l I, \quad k, l = 1, \dots, n. \tag{2.4}$$

It is also assumed that every bounded operator on $\mathscr{H}$ which commutes with all operators $\mathbf{P}$ and $\mathbf{Q}$ is a multiple of the identity operator I.

Algebraically, Heisenberg's commutation relations give rise to an irreducible representation of a *Heisenberg algebra* $\mathfrak{h}_n$ with n degrees of freedom, defined as follows.

DEFINITION. The Heisenberg algebra $\mathfrak{h}_n$ with n degrees of freedom is a Lie algebra with the generators $x^k, \dots, x^n, y_1, \dots, y_n, c$ and the relations

$$[x^k, c] = 0, \quad [y_k, c] = 0, \quad [x^k, y_l] = \delta^k_l c, \quad k, l = 1, \dots, n.$$

Equivalently, the Heisenberg algebra $\mathfrak{h}_n$ is a one-dimensional central extension of the abelian Lie algebra $\mathbb{R}^{2n}$,

$$0 \to \mathbb{R} \to \mathfrak{h}_n \to \mathbb{R}^{2n} \to 0,$$

by the Lie algebra 2-cocycle given by the canonical symplectic form ω on $\mathbb{R}^{2n}$. The corresponding element c is called the central element of the Heisenberg

algebra $\mathfrak{h}_n$. Explicitly, the Heisenberg algebra is a nilpotent subalgebra of the Lie algebra of $(n+2)\times(n+2)$ matrices with the elements

$$\sum_{k=1}^{n}(u_k x^k + v^k y_k) + \alpha c = \begin{bmatrix} 0 & u_1 & u_2 & \dots & u_n & \alpha \\ 0 & 0 & 0 & \dots & 0 & v^1 \\ 0 & 0 & 0 & \dots & 0 & v^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & v^n \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

It follows from (2.4) that Heisenberg's commutation relations correspond to the irreducible representation ρ of $\mathfrak{h}_n$ on the Hilbert space $\mathscr{H}$ by skew-Hermitian operators, defined by $\rho(x^k) = -iP^k$, $\rho(y_k) = -iQ_k$, $k = 1, \dots, n$, and $\rho(c) = i\hbar I$. The operators P^k and Q_k are necessarily unbounded (see Problem 2.1), so that rigorous definition of the mapping ρ requires a caution. To exclude the "pathological" representations, we will assume that ρ is an integrable representation, i.e., it can be integrated to an irreducible unitary representation of the Heisenberg group $\mathbf{H}_n$ — a connected, simply-connected Lie group with the Lie algebra $\mathfrak{h}_n$. Explicitly, the Heisenberg group is a unipotent subgroup of the Lie algebra $\mathrm{SL}(n+2,\mathbb{R})$ with the elements

$$g = \begin{bmatrix} 1 & u_1 & u_2 & \dots & u_n & \alpha \\ 0 & 1 & 0 & \dots & 0 & v^1 \\ 0 & 0 & 1 & \dots & 0 & v^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & v^n \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Abstractly, the Heisenberg group $\mathbf{H}_n$ is generated by two n-parameter abelian subgroups $e^{\mathbf{u}X} = e^{\sum_{k=1}^n u_k x^k}$, $e^{\mathbf{v}Y} = e^{\sum_{k=1}^n v^k y_k}$ and a one-parameter center $e^{\alpha c}$, which satisfy the relations

$$e^{\mathbf{u}X}e^{\mathbf{v}Y} = e^{\mathbf{u}\mathbf{v}c}e^{\mathbf{v}Y}e^{\mathbf{u}X}, \quad \text{where} \quad \mathbf{u}\mathbf{v} = \sum_{k=0}^{n} u_k v^k. \tag{2.5}$$

Indeed, it readily follows from the commutation relations of the Heisenberg algebra and the Campbell-Baker-Hausdorff formula that

$$e^{\mathbf{u}X}e^{\mathbf{v}Y} = e^{\frac{1}{2}[\mathbf{u}X,\mathbf{v}Y]}e^{(\mathbf{u}X+\mathbf{v}Y)}$$

and

$$e^{\mathbf{v}Y}e^{\mathbf{u}X} = e^{-\frac{1}{2}[\mathbf{u}X,\mathbf{v}Y]}e^{(\mathbf{u}X+\mathbf{v}Y)},$$

which imply (2.5). In the matrix realization, $e^{\mathbf{u}X} = I + \mathbf{u}X$, $e^{\mathbf{v}Y} = I + \mathbf{v}Y$ and $e^{\alpha c} = I + \alpha c$, where I is the $(n+2)\times(n+2)$ identity matrix.

The exponential map $\mathfrak{h}_n \to \mathbf{H}_n$ is onto, so that an irreducible unitary representation R of the Heisenberg group $\mathbf{H}_n$ in the Hilbert space $\mathscr{H}$ is

defined by two strongly continuous n-parameter abelian groups of unitary operators

$$U(\mathbf{u}) = R(e^{\mathbf{u}X}), \;\; V(\mathbf{v}) = R(e^{\mathbf{v}Y}),$$

satisfying *H. Weyl relations*

$$U(\mathbf{u})V(\mathbf{v}) = e^{i\hbar\mathbf{u}\mathbf{v}}V(\mathbf{v})U(\mathbf{u}), \tag{2.6}$$

where by the Schur lemma $R(e^{\alpha c}) = e^{i\hbar\alpha}I$. Irreducible representation ρ of the Heisenberg algebra is integrable if $\rho = dR$, where R is an irreducible integrable representation of the Heisenberg group. According to the Stone theorem,

$$P_k = i\,\frac{\partial U(\mathbf{u})}{\partial u_k}\bigg|_{\mathbf{u}=0} \quad \text{and} \quad Q_k = i\,\frac{\partial V(\mathbf{v})}{\partial v_k}\bigg|_{\mathbf{v}=0}, \quad k = 1, \dots, n.$$

Remark. Not every irreducible representation of Heisenberg algebra is integrable, so that H. Weyl relations can not be obtained from the Heisenberg's commutation relations. However, the following heuristic argument (which ignores the subtleties of dealing with unbounded operators) is commonly used in physics textbooks. Consider the case of one degree of freedom and start with

$$\{P, Q\}_\hbar = I.$$

Since quantum bracket satisfies the Leibniz rule, i.e., it is a derivation, we have (for a "suitable" function f)

$$\{f(P), Q\}_h = f'(P).$$

In particular, choosing $f(P) = e^{-iuP} = U(u)$, we obtain

$$U(u)Q - QU(u) = \hbar u U(u) \quad \text{or} \quad U(u)QU(u)^{-1} = Q + \hbar u I.$$

This implies (for a "suitable" function g)

$$U(u)g(Q) = g(Q + \hbar u I)U(u),$$

and setting $g(Q) = e^{-ivQ} = V(v)$, we get H. Weyl relation.

We will prove in Section 3.1 that all integrable irreducible representations of the Heisenberg algebra $\mathfrak{h}_n$ having the same value on the central element c are unitary equivalent. This justifies the following mathematical formulation of the Heisenberg's commutation relations for n degrees of freedom.

Heisenberg's Commutation Relations. Momenta and coordinate operators $\mathbf{P}$ and $\mathbf{Q}$ for a quantum particle with n degrees of freedom correspond to the integrable irreducible representation ρ of the Heisenberg algebra $\mathfrak{h}_n$ with the property $\rho(c) = i\hbar I$.

Problem 2.1. Prove that there are no bounded operators on the Hilbert space $\mathscr{H}$ satisfying $[A, B] = I$.

2.2. Coordinate and momentum representations. We start with the case of one degree of freedom and consider two natural realizations of the Heisenberg's commutation relation. They are defined by the property that one of the self-adjoint operators P and Q is "diagonal" (i.e., is a multiplication by a function operator in the corresponding Hilbert space).

In the *coordinate representation*, $\mathscr{H} = L^2(\mathbb{R}, dq)$ is the Hilbert L^2-space on the configuration space $\mathbb{R}$ with the coordinate q, which is a Lagrangian subspace of $\mathbb{R}^2$ defined by the equation $p = 0$. Set

$$D(Q) = \left\{ \varphi \in \mathscr{H} : \int_{-\infty}^{\infty} q^2 |\varphi(q)|^2 dq < \infty \right\}$$

and for $\varphi \in D(Q)$ define the operator Q as a "multiplication by q operator",

$$(Q\varphi)(q) = q\varphi(q), \ q \in \mathbb{R},$$

justifying the name coordinate representation. Coordinate operator Q is obviously self-adjoint and its projection-valued measure is given by

$$(\mathsf{P}(E)\varphi)(q) = \chi_E(q)\varphi(q), \tag{2.7}$$

where χ_E is the characteristic function of a Borel subset $E \subset \mathbb{R}$.

Recall that a self-adjoint operator A has an absolutely continuous spectrum if for every $\psi \in \mathscr{H}$, $\|\psi\| = 1$, the probability measure

$$\nu_\psi(E) = (\mathsf{P}_A(E)\psi, \psi), \quad E \in \mathscr{B}(\mathbb{R}),$$

absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.

LEMMA 2.1. *Coordinate operator Q has an absolutely continuous spectrum $\sigma(Q) = \mathbb{R}$, and every bounded operator B which commutes with Q is a function of Q, $B = f(Q)$ with $f \in L^\infty(\mathbb{R})$.*

PROOF. Clearly $\operatorname{supp} \mathsf{P}_A = \mathbb{R}$ and $\nu_\psi(E) = \int_E |\psi(q)|^2 dq$, which proves the first statement. Now a bounded operator B on $\mathscr{H}$ commutes with Q if and only if $\mathrm{i}B\mathsf{P}(E) = \mathsf{P}(E)B$ for all $E \in \mathscr{B}(\mathbb{R})$. Using (2.7) we get that for every $E \in \mathscr{B}(\mathbb{R})$ with finite Lebesgue measure,

$$B(\chi_E) = f_E,$$

where measurable f_E satisfies $\operatorname{supp} f_E \subseteq E$. Using the commutativity property once again, we get that for $E_1 \subset E_2$, $f_{E_2}|_{E_1} = f_{E_1}|_{E_1}$. Thus there exists a measurable function f on $\mathbb{R}$ such that $f|_E = f_E|_E$ for every $E \in \mathscr{B}(\mathbb{R})$. The linear subspace spanned by all $\chi_E \in L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, and we get that

$$(B\varphi)(q) = f(q)\varphi(q) \quad \text{for all} \quad \varphi \in L^2(\mathbb{R}).$$

Since B is a bounded operator, $\|B\| = \|f\|_\infty < \infty$ and $f \in L^\infty(\mathbb{R})$. □

We say that the coordinate operator Q has simple, absolutely continuous spectrum $\mathbb{R}$.

REMARK. By the Schwartz kernel theorem, the operator B can be represented an integral operator with a distributional kernel $K(q, q')$. Then the commutativity $BQ = QB$ implies that, in the distributional sense,

$$(q - q')K(q, q') = 0,$$

so that K is "proportional" to the Dirac delta-function, i.e.,

$$K(q, q') = f(q)\delta(q - q'),$$

with some $f \in L^\infty(\mathbb{R})$. This argument is usually given in the physics textbooks.

REMARK. The operator Q has no eigenvectors — the eigenvalue equation

$$Q\varphi = \lambda\,\varphi$$

has no solutions in $L^2(\mathbb{R})$. However, in the distributional sense, this equation for every $\lambda \in \mathbb{R}$ has a unique (up to a constant factor) solution $\varphi_\lambda(q) = \delta(q-\lambda)$, and these "generalized eigenfunctions" combine to a Schwarz kernel of the identity operator I on $L^2(\mathbb{R})$. This reflects the fact that operator Q is diagonal in the coordinate representation.

For a pure state $M = P_\psi$, $\|\psi\| = 1$, corresponding probability measure μ_Q on $\mathbb{R}$ is given by

$$\mu_Q(E) = \nu_\psi(E) = \int_E |\psi(q)|^2 dq, \quad E \in \mathscr{B}(\mathbb{R}).$$

Physically, this is interpreted that in the state P_ψ with the "wave function" $\psi(q)$, the probability of finding a quantum particle between q and $q + dq$ is $|\psi(q)|^2 dq$. In other words, the modulus square of a wave function is the probability distribution for the coordinate of a particle.

Corresponding momentum operator P is given by a differential operator

$$P = \frac{\hbar}{i}\frac{d}{dq}$$

with $D(P) = W^{1,2}(\mathbb{R})$ — a Sobolev space of absolutely continuous functions f on $\mathbb{R}$ such that f and its derivative f' (defined a.e.) are in $L^2(\mathbb{R})$. The operator P is self-adjoint and it is straightforward to verify that on $D = C_c^\infty(\mathbb{R})$ — the space of smooth functions on $\mathbb{R}$ with compact support,

$$QP - PQ = i\hbar I.$$

REMARK. Operator P on $\mathscr{H}$ has no eigenvectors — the eigenvalue equation

$$P\varphi = p\,\varphi,\ p \in \mathbb{R},$$

has a solution

$$\varphi(q) = \text{const} \times e^{ipq/\hbar}$$

which does not belong to $L^2(\mathbb{R})$. The family of "normalized generalized eigenfunctions"

$$\varphi_p(q) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar}$$

combines to a Schwartz kernel of the inverse $\hbar$-dependent Fourier transform operator, which diagonalizes the momentum operator P. Indeed, in the distributional sense,

$$\int_{-\infty}^{\infty} \varphi_p(q)\overline{\varphi_{p'}(q)}dq = \delta(p-p').$$

PROPOSITION 2.1. *Coordinate representation defines an irreducible, integrable representation of the Heisenberg algebra.*

PROOF. To show that coordinate representation is integrable, let $U(u) = e^{-iuP}$ and $V(v) = e^{-ivQ}$ be the corresponding one-parameter groups of unitary operators. Clearly, $(V(v)\varphi)\psi(q) = e^{-ivq}\varphi(q)$ and it easily follows from the Stone theorem (or by the definition of a derivative) that $(U(u)\varphi)(q) = \varphi(q-\hbar u)$, so that the operators $U(u)$ and $V(v)$ satisfy H. Weyl relation (2.6). Such realization of H. Weyl relation is called *Schrödinger representation*.

To prove that coordinate representation is irreducible, let B be a bounded operator commuting with P and Q. By Lemma 2.1, $T = f(Q)$ for some $f \in L^\infty(\mathbb{R})$. Now commutativity between T and P implies that

$$TU(u) = U(u)T \quad \text{for all} \quad u \in \mathbb{R},$$

which is equivalent to $f(q-\hbar u) = f(q)$ for all $q, u \in \mathbb{R}$, so that $f = \text{const}$ a.e. on $\mathbb{R}$. □

To summarize, the coordinate representation is characterized by the property that the coordinate operator Q is a multiplication by q operator and the momentum operator P is a differentiation operator,

$$Q = q \quad \text{and} \quad P = \frac{\hbar}{i}\frac{d}{dq}.$$

Similarly, *momentum representation* is defined by the property that the momentum operator P is a multiplication by p operator. Namely let $\mathscr{H} = L^2(\mathbb{R}, dp)$ be the Hilbert L^2-space on the "momentum space" $\mathbb{R}$ with the coordinate p, which is a Lagrangian subspace of $\mathbb{R}^2$ defined by the equation $q = 0$. The coordinate and momentum operators are given by

$$\hat{Q} = i\hbar\frac{d}{dp} \quad \text{and} \quad \hat{P} = p,$$

and satisfy the Heisenberg's commutation relation. In the momentum representation, the modulus square of the wave function $\psi(p)$ of a pure state $M = P_\psi$, $\|\psi\| = 1$, is the probability distribution for the momentum of the

quantum particle, i.e., the probability that a quantum particle has momentum between p and $p+dp$ is $|\psi(p)|^2 dp$.

Let $\mathscr{F}_\hbar : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be the $\hbar$-dependent Fourier transform operator, defined by

$$\hat{\varphi}(p) = \mathscr{F}_\hbar(\varphi)(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipq/\hbar} \varphi(q) dq.$$

Here integral is understood as the limit $\hat{\varphi} = \lim_{n\to\infty} \hat{\varphi}_n$ in the strong topology on $L^2(\mathbb{R})$, where

$$\hat{\varphi}_n(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-n}^{n} e^{-ipq/\hbar} \varphi(q) dq.$$

It is well-known that $\mathscr{F}$ is a unitary operator on $L^2(\mathbb{R})$ and

$$\hat{Q} = \mathscr{F}_\hbar Q \mathscr{F}_\hbar^{-1}, \quad \hat{P} = \mathscr{F}_\hbar P \mathscr{F}_\hbar^{-1},$$

so that coordinate and momentum representations are unitary equivalent. In particular, since the operator $\hat{P}$ is obviously self-adjoint, this immediately shows that the operator P is self-adjoint, and that the momentum representation is integrable.

For n degrees of freedom, the coordinate representation is defined by setting $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{q})$, where $d^n\mathbf{q} = dq_1 \dots dq_n$ is the Lebesgue measure on $\mathbb{R}^n$, and

$$\mathbf{Q} = \mathbf{q} = (q_1, \dots, q_n), \quad \mathbf{P} = \frac{\hbar}{i}\frac{\partial}{\partial \mathbf{q}} = \left(\frac{\hbar}{i}\frac{\partial}{\partial q_1}, \dots, \frac{\hbar}{i}\frac{\partial}{\partial q_n}\right).$$

Here $\mathbb{R}^n$ is the configuration space with the coordinates $\mathbf{q}$ — a Lagrangian subspace of $\mathbb{R}^{2n}$ defined by the equations $\mathbf{p} = 0$. The coordinate and momenta operators are self-adjoint and satisfy Heisenberg's commutation relations. Projection-valued measured for the operators Q_k are given by

$$(\mathsf{P}_k(E)\varphi)(\mathbf{q}) = \chi_{\pi_k^{-1}(E)}(\mathbf{q})\varphi(\mathbf{q}),$$

where $E \in \mathscr{B}(\mathbb{R})$ and $\pi_k : \mathbb{R}^n \to \mathbb{R}$ is a canonical projection onto the k-th component, $k = 1, \dots, n$. Correspondingly, the projection-valued measure P for the commutative family $\mathbf{Q} = (Q_1, \dots, Q_n)$ (see Proposition 1.2) is defined on the Borel subsets $\mathbf{E} \subset \mathbb{R}^n$ by

$$(\mathsf{P}(\mathbf{E})\varphi)(\mathbf{q}) = \chi_{\mathbf{E}}(\mathbf{q})\varphi(\mathbf{q}).$$

The family $\mathbf{Q}$ has absolutely continuous joint spectrum $\mathbb{R}^n$.

Coordinate operators $Q_1, \dots, Q_n$ form a *complete system of commuting observables*. This means, by definition, that none of these operators is a function of the other operators, and that every bounded operator commuting with $Q_1, \dots, Q_n$ is a function of $Q_1, \dots, Q_n$, i.e., is a multiplication by $f(\mathbf{q})$ operator for some $f \in L^\infty(\mathbb{R}^n)$. The proof repeats verbatim the proof of Lemma 2.1. For the pure state $M = P_\psi$, $\|\psi\| = 1$, the modulus square $|\psi(\mathbf{q})|^2$ of the wave function is the density of a joint distribution function

$\mu_{\mathbf{Q}}$ for the commutative family $\mathbf{Q}$, i.e., the probability of finding a quantum particle in a Borel subset $\mathbf{E} \subset \mathbb{R}^n$ is given by

$$\mu_{\mathbf{Q}}(\mathbf{E}) = \int\limits_{\mathbf{E}} |\psi(\mathbf{q})|^2 d^n\mathbf{q}.$$

The coordinate repsentation defines an irreducible, integrable representation of the Heisenberg algebra $\mathfrak{h}_n$. Indeed, corresponding n-parameter groups of unitary operators $U(\mathbf{u}) = e^{-i\mathbf{u}\mathbf{P}}$ and $V(\mathbf{v}) = e^{-i\mathbf{v}\mathbf{Q}}$ are given by

$$(U(\mathbf{u})\varphi)(\mathbf{q}) = \varphi(\mathbf{q} - \hbar\mathbf{u}), \quad (V(\mathbf{v})\varphi)(\mathbf{q}) = e^{-i\mathbf{v}\mathbf{q}}\varphi(\mathbf{q}),$$

and satisfy H. Weyl relations (2.6). The same argument as in the proof of Proposition 2.2 shows that this representation of the Heisenberg group $\mathbf{H}_n$, called the Schrödinger representation for n degrees of freedom, is irreducible.

In the momentum representation, $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{p})$, where $d^n\mathbf{p} = dp^1 \dots dp^n$ is the Lebesgue measure on $\mathbb{R}^n$, and

$$\hat{\mathbf{Q}} = i\hbar\frac{\partial}{\partial\mathbf{p}} = \left(i\hbar\frac{\partial}{\partial p_1}, \dots, i\hbar\frac{\partial}{\partial p_n}\right), \quad \hat{\mathbf{P}} = \mathbf{p} = (p_1, \dots, p_n).$$

Here $\mathbb{R}^n$ is the momentum space with the coordinates $\mathbf{p}$ — a Lagrangian subspace of $\mathbb{R}^{2n}$ defined by the equations $\mathbf{q} = 0$.

The coordinate and momentum representations are unitary equivalent by the Fourier transform. As in the case $n = 1$, the Fourier transform $\mathscr{F}_\hbar : L^2(\mathbb{R}^n, d^n\mathbf{q}) \to L^2(\mathbb{R}^n, d^n\mathbf{p})$ is a unitary operator defined by

$$\begin{aligned}\hat{\varphi}(\mathbf{p}) = \mathscr{F}_\hbar(\varphi)(\mathbf{p}) =& (2\pi\hbar)^{-n/2} \int\limits_{\mathbb{R}^n} e^{-i\mathbf{p}\mathbf{q}/\hbar}\varphi(\mathbf{q})d^n\mathbf{q} \\ &= \lim_{N\to\infty} (2\pi\hbar)^{-n/2} \int\limits_{|\mathbf{q}|\le N} e^{-i\mathbf{p}\mathbf{q}/\hbar}\varphi(\mathbf{q})d^n\mathbf{q},\end{aligned}$$

where the limit is understood in the strong topology on $L^2(\mathbb{R}^n, d^n\mathbf{p})$. As in the case $n = 1$, we have

$$\hat{Q}_k = \mathscr{F}_\hbar Q_k \mathscr{F}_\hbar^{-1}, \quad \hat{P}_k = \mathscr{F}_\hbar P_k \mathscr{F}_\hbar^{-1}, \quad k = 1, \dots, n.$$

In particular, since operators $\hat{P}_1, \dots, \hat{P}_n$ are obviously self-adjoint, this immediately shows that $P_1, \dots, P_n$ are also self-adjoint.

Remark. Following Dirac, physicists use to denote a vector $\psi \in \mathscr{H}$ by a *ket vector* $|\psi\rangle$, a vector $\varphi \in \mathscr{H}^*$ in the dual space to $\mathscr{H}$ ($\mathscr{H}^* \simeq \mathscr{H}$ is a complex anti-linear isomorphism) — by a *bra vector* $\langle\varphi|$, and their inner product — by $\langle\varphi|\psi\rangle$. In standard mathematics notation,

$$(\psi, \varphi) = \langle\varphi|\psi\rangle \quad \text{and} \quad (A\psi, \varphi) = \langle\varphi|A|\psi\rangle.$$

where A is a linear operator. From physics point of view, Dirac's notations are intuitive and convenient for working with coordinate and momentum representations. Denoting by $|\mathbf{q}\rangle = \delta(\mathbf{q} - \mathbf{q}')$ and $|\mathbf{p}\rangle = (2\pi\hbar)^{-n/2}e^{i\mathbf{p}\mathbf{q}/\hbar}$

the set of generalized common eigenfunctions for the operators $\mathbf{Q}$ and $\mathbf{P}$ respectively, we formally get

$$\mathbf{Q}|\mathbf{q}\rangle = \mathbf{q}|\mathbf{q}\rangle, \quad \mathbf{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle,$$

where operators $\mathbf{Q}$ act on $\mathbf{q}'$, and

$$\langle \mathbf{q}|\psi\rangle = \int\limits_{\mathbb{R}^n} \delta(\mathbf{q}-\mathbf{q}')\psi(\mathbf{q}')d^n\mathbf{q}' = \psi(\mathbf{q}),$$

$$\langle \mathbf{p}|\psi\rangle = (2\pi\hbar)^{-n/2} \int\limits_{\mathbb{R}^n} e^{-i\mathbf{p}\mathbf{q}/\hbar}\psi(\mathbf{q})d^n\mathbf{q} = \hat{\psi}(\mathbf{p}),$$

as well as $\langle \mathbf{q}|\mathbf{q}'\rangle = \delta(\mathbf{q}-\mathbf{q}')$, $\langle \mathbf{p}|\mathbf{p}'\rangle = \delta(\mathbf{p}-\mathbf{p}')$. Though in our exposition we are not using Dirac's notations, this could help the interested reader to "translate" the notations used in physics textbooks to the standard mathematics notations.

PROBLEM 2.2. Give an example of non-integrable representation of the Heisenberg algebra.

PROBLEM 2.3. Prove that there exists $\varphi \in \mathscr{H} = L^2(\mathbb{R}, dq)$ such the vectors $\mathsf{P}(E)\varphi$, $E \in \mathscr{B}(\mathbb{R})$, where P is a projection-valued measure for the coordinate operator Q, are dense in $\mathscr{H}$.

PROBLEM 2.4. Find the generating operator for the commutative family $\mathbf{Q} = (Q_1, \dots, Q_n)$. Does it have a physical interpretation?

PROBLEM 2.5. Find the projection-valued measure for the commutative family $\mathbf{P} = (P_1, \dots, P_n)$ in the coordinate representation.

2.3. Free quantum particle. Free classical particle with one degree of freedom is described by the phase space $\mathbb{R}^2$ with coordinates p, q and the Poisson bracket (2.1), and by the Hamiltonian function $h(p,q) = \dfrac{p^2}{2m}$. The Hamiltonian operator of a free quantum particle with one degree of freedom is

$$H = \frac{P^2}{2m},$$

and in coordinate representation is given by

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dq^2}.$$

It is a self-adjoint operator on $\mathscr{H}$ with $D(H) = W^{2,2}(\mathbb{R})$ — a Sobolev space of functions in $L^2(\mathbb{R})$, whose generalized first and second derivatives are in $L^2(\mathbb{R})$.

The operator H is positive with absolutely continuous spectrum $[0, \infty)$ of multiplicity two. Indeed, let $\mathfrak{H} = L^2(\mathbb{R}_{>0}, \mathbb{C}^2; d\sigma)$ be the Hilbert space of

$\mathbb{C}^2$-valued measurable functions Ψ on the semi-line $\mathbb{R}_{>0} = (0, \infty)$, which are square-integrable with respect to the measure $d\sigma(E) = \sqrt{\frac{m}{2E}}\, dE$,

$$\mathfrak{H} = \left\{ \Psi(E) = \begin{pmatrix} \psi_1(E) \\ \psi_2(E) \end{pmatrix} : \|\Psi\|^2 = \int_0^\infty (|\psi_1(E)|^2 + |\psi_2(E)|^2) d\sigma(E) < \infty \right\}.$$

Since in momentum representation H is a multiplication by $\dfrac{p^2}{2m}$ operator, under the isomorphism $L^2(\mathbb{R}, dp) \simeq \mathfrak{H}$,

$$L^2(\mathbb{R}, dp) \ni \psi(p) \mapsto \Psi(E) = \begin{pmatrix} \psi(\sqrt{2mE}) \\ \psi(-\sqrt{2mE}) \end{pmatrix} \in \mathfrak{H},$$

the operator H becomes a multiplication by E operator, $H\Psi(E) = E\Psi(E)$.

Remark. The Hamiltonian operator H has no eigenvectors — the eigenvalue equation

$$H\psi = E\psi$$

has no solutions in $L^2(\mathbb{R})$. However, for every $E > 0$ this differential equation has two linear independent bounded solutions

$$\psi_\pm(q) = \frac{1}{\sqrt{2\pi\hbar}}\, e^{\pm i\sqrt{2mE}\, q/\hbar},$$

which do not belong to $L^2(\mathbb{R})$. In the distributional sense, these solutions combine to a Schwartz kernel of the unitary operator establishing the isomorphism between $\mathscr{H} = L^2(\mathbb{R}, dp)$ and the Hilbert space $\mathfrak{H}$ in which H acts as a multiplication by E operator.

The Schrödinger equation for a free particle,

$$i\hbar \frac{d\psi(t)}{dt} = H\psi(t), \quad \psi(0) = \psi,$$

is easily solved by the Fourier transform. Indeed, in the momentum representation it takes the form

$$i\hbar \frac{\partial \hat\psi(p,t)}{\partial t} = \frac{p^2}{2m}\hat\psi(p,t), \quad \hat\psi(p,0) = \hat\psi(p),$$

so that

$$\hat\psi(p,t) = e^{-\frac{ip^2 t}{2m\hbar}}\, \hat\psi(p).$$

Therefore in the coordinate representation solution is given by

$$\psi(q,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{i\left(qp - \frac{p^2}{2m}t\right)/\hbar} \hat\psi(p) dp. \tag{2.8}$$

The formula (2.8) describes the motion of a quantum particle and admits the following physical interpretation. Suppose that the initial condition ψ is

such that its Fourier transform $\hat{\psi} = \mathcal{F}_\hbar(\psi)$ is a smooth function supported in a neighborhood U_0 of $p_0 \in \mathbb{R} \setminus \{0\}$ such that $0 \notin U_0$ and

$$\int_{-\infty}^{\infty} |\hat{\psi}(p)|^2 dp = 1.$$

Such states are called "wave packets". Then for every compact subset $E \subset \mathbb{R}$ we have

$$\lim_{|t|\to\infty} \int_E |\psi(q,t)|^2 dq = 0. \tag{2.9}$$

Since

$$\int_{-\infty}^{\infty} |\psi(q,t)|^2 dq = 1$$

for all t, it follows from (2.9) that as $|t| \to \infty$, quantum particle "leaves" every compact subset of $\mathbb{R}$, so that the motion is infinite. To prove (2.9), observe that the function $\chi(p,q,t) = -\frac{p^2}{2m} + \frac{qp}{t}$ — the "phase" in integral representation (2.8), has the property that $|\frac{\partial \chi}{\partial p}| > C > 0$ for all $p \in U_0$, $q \in E$ and large enough $|t|$. Integrating by parts we get

$$\begin{aligned}\psi(q,t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{U_0} e^{it\chi(q,p,t)/\hbar} \hat{\psi}(p) dp \\ &= -\frac{1}{it}\sqrt{\frac{\hbar}{2\pi}} \int_{U_0} \frac{\partial}{\partial p}\left(\frac{\hat{\psi}(p)}{\frac{\partial \chi(q,p,t)}{\partial p}}\right) e^{it\chi(q,p,t)/\hbar} dp,\end{aligned}$$

so that uniformly on E,

$$\psi(q,t) = O\left(\frac{1}{|t|}\right) \quad \text{as} \quad |t| \to \infty.$$

By repeated integration by parts, we obtain that for every $n \in \mathbb{N}$ uniformly on E,

$$\psi(q,t) = O\left(\frac{1}{|t|^n}\right).$$

To describe the motion of a free quantum particle in unbounded regions, we use the stationary phase method. In its simplest form it is stated as follows.

The Method of Stationary Phase. Let $f, g \in C^\infty(\mathbb{R})$, where f is real-valued and g has compact support, and suppose that f has a single non-degenerate critical point x_0, i.e., $f'(x_0) = 0$ and $f''(x_0) \neq 0$. Then

$$\int_{-\infty}^{\infty} e^{iNf(x)} g(x) dx = \left(\frac{2\pi}{N|f''(x_0)|}\right)^{1/2} e^{iNf(x_0) + \frac{i\pi}{4}\operatorname{sgn} f''(x_0)} g(x_0) + O\left(\frac{1}{N}\right)$$

as $N \to \infty$.

Applying the stationary phase method to the integral representation (2.8) (and setting $N = t$), we find that the critical point of $\chi(q,p,t)$ is $p_0 = \frac{mq}{t}$ with $\chi''(p_0) = -\frac{1}{m} \neq 0$, and

$$\begin{aligned}\psi(q,t) &= \sqrt{\frac{m}{t}}\hat{\psi}\left(\frac{mq}{t}\right) e^{\frac{imq^2}{\hbar t} - \frac{\pi i}{4}} + O(t^{-1}) \\ &= \psi_0(q,t) + O(t^{-1}) \quad \text{as} \quad t \to \infty.\end{aligned}$$

Thus as $t \to \infty$, the wave function $\psi(q,t)$ is supported on $\frac{t}{m}U_0$ — a domain where the probability of finding a particle is asymptotically different from zero. At large t the points in this domain move with constant velocities $v = \frac{p}{m}$, $p \in U_0$. In this sense, the classical relation $p = mv$ remains valid in the quantum picture. Moreover, the asymptotic wave function ψ_0 satisfies

$$\int_{-\infty}^{\infty} |\psi_0(q,t)|^2 dq = \sqrt{\frac{m}{t}} \int_{-\infty}^{\infty} \left|\hat{\psi}\left(\frac{mq}{t}\right)\right|^2 dq = 1,$$

and, therefore, describes the asymptotic probability distribution. Similarly, setting $N = -|t|$, we can describe behavior of the wave function $\psi(q,t)$ as $t \to -\infty$.

Remark. In the weak topology on $\mathscr{H}$ the vector $\psi(t) \to 0$ as $|t| \to \infty$. Indeed, for every $\varphi \in \mathscr{H}$ we get by the Parseval identity

$$(\psi(t), \varphi) = \int_{-\infty}^{\infty} \hat{\psi}(p)\overline{\hat{\varphi}(p)} e^{-\frac{ip^2 t}{2m\hbar}} dp,$$

and the integral goes to zero as $|t| \to \infty$ by the Riemann-Lebesgue lemma.

Classical free particle with n degrees of freedom is described by the phase space $\mathbb{R}^{2n}$ with coordinates $\mathbf{p} = (p^1, \dots, p^n)$ and $\mathbf{q} = (q_1, \dots, q_n)$, the Poisson bracket (2.3), and the Hamiltonian function

$$h(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} = \frac{(p^1)^2 + \dots + (p^n)^2}{2m}.$$

The Hamiltonian operator of a free quantum particle with n degrees of freedom is

$$H = \frac{\mathbf{P}^2}{2m} = \frac{(P^1)^2 + \dots + (P^n)^2}{2m},$$

and in the coordinate representation is

$$H = -\frac{\hbar^2}{2m}\Delta,$$

where

$$\Delta = \frac{\partial^2}{\partial q_1^2} + \dots + \frac{\partial^2}{\partial q_n^2}$$

is the Laplace operator[12] in the Cartesian coordinates on $\mathbb{R}^n$. The Hamiltonian H is a self-adjoint operator on $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{q})$ with $D(H) = W^{2,2}(\mathbb{R}^n)$

[12]It is negative of the Laplace-Beltrami operator of the standard Euclidean metric on $\mathbb{R}^n$.

— the Sobolev space on $\mathbb{R}^n$. In the momentum representation,

$$H = \frac{\mathbf{p}^2}{2m}$$

— a multiplication by a function operator on $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{p})$.

The operator H is positive with absolutely continuous spectrum $[0, \infty)$ of infinite multiplicity. Namely, let $S^{n-1} = \{\mathbf{n} \in \mathbb{R}^n : \mathbf{n}^2 = 1\}$ be the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$, $d\mathbf{n}$ be the measure on S^{n-1} induced by the Lebesgue measure on $\mathbb{R}^n$, and let

$$\mathfrak{h} = \{f : S^{n-1} \to \mathbb{C} : \|f\|_{\mathfrak{h}}^2 = \int_{S^{n-1}} |f(\mathbf{n})|^2 d\mathbf{n} < \infty\}.$$

Let $\mathfrak{H}_n = L^2(\mathbb{R}_{>0}, \mathfrak{h}; d\sigma_n)$ be the Hilbert space of $\mathfrak{h}$-valued measurable functions[13] Ψ on $\mathbb{R}_{>0}$, square-integrable on $\mathbb{R}_{>0}$ with respect to the measure $d\sigma_n(E) = (2mE)^{\frac{n}{2}} \frac{dE}{2E}$,

$$\mathfrak{H}_n = \left\{ \Psi : \mathbb{R}_{>0} \to \mathfrak{h},\ \|\Psi\|^2 = \int_0^\infty \|\Psi(E)\|_{\mathfrak{h}}^2 \, d\sigma_n(E) < \infty \right\}.$$

Since in the momentum represntation H is a multiplication by $\dfrac{\mathbf{p}^2}{2m}$ operator, under the isomorphism $L^2(\mathbb{R}^n, d^n\mathbf{p}) \simeq \mathfrak{H}_n$,

$$L^2(\mathbb{R}^n, d^n\mathbf{p}) \ni \psi(\mathbf{p}) \mapsto \Psi(E)(\mathbf{n}) = \psi(\sqrt{2mE}\,\mathbf{n}) \in \mathfrak{H}_n,$$

the operator H becomes a multiplication by E operator, $H\Psi(E) = E\Psi(E)$.

Remark. As in the case $n = 1$, the Hamiltonian operator H has no eigenvectors — the eigenvalue equation

$$H\psi = E\psi$$

has no solutions in $L^2(\mathbb{R}^n)$. However, for every $E > 0$ this differential equation has an infinitely many linear independent bounded solutions

$$\psi_{\mathbf{n}}(\mathbf{q}) = (2\pi\hbar)^{-\frac{n}{2}} \, e^{i\sqrt{2mE}\,\mathbf{nq}/\hbar}$$

parametrized by the unit sphere S^{n-1}. These solutions do not belong to $L^2(\mathbb{R}^n)$, but in the distributional sense they combine to a Schwartz kernel of the unitary operator establishing the isomorphism between the Hilbert spaces $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{p})$ and $\mathfrak{H}_n$.

As in the case $n = 1$, the Schrödinger equation for free particle,

$$i\hbar \frac{d\psi(t)}{dt} = H\psi(t), \quad \psi(0) = \psi,$$

is solved by the Fourier transform,

$$\psi(\mathbf{q}, t) = (2\pi\hbar)^{-n/2} \int\limits_{\mathbb{R}^n} e^{i\left(\mathbf{qp} - \frac{\mathbf{p}^2}{2m}t\right)/\hbar} \hat{\psi}(\mathbf{p}) d^n\mathbf{p}.$$

[13] That is, for every $f \in \mathfrak{h}$ the function (f, Ψ) is measurable on $\mathbb{R}_{>0}$.

For a wave packet — an initial condition ψ such that its Fourier transform $\hat{\psi} = \mathscr{F}_\hbar(\psi)$ is a smooth function supported on a neighborhood U_0 of $\mathbf{p}_0 \in \mathbb{R}^n \setminus \{0\}$ such that $0 \notin U_0$ and

$$\int\limits_{\mathbb{R}^n} |\hat{\psi}(\mathbf{p})|^2 d^n\mathbf{p} = 1,$$

quantum particle "leaves" every compact subset of $\mathbb{R}^n$ and the motion is infinite. Asymptotically as $|t| \to \infty$, the wave function $\psi(\mathbf{q}, t)$ is different from 0 only when $\mathbf{q} = \frac{\mathbf{p}}{m}t$, $\mathbf{p} \in U_0$.

2.4. Examples of quantum systems. Here we describe quantum systems that correspond to classical Lagrangian systems introduced in Section 1.3 of Chapter 1. In Hamiltonian formulation, the phase space of these systems is a symplectic vector space $\mathbb{R}^{2n}$ with canonical coordinates $\mathbf{p}, \mathbf{q}$ and symplectic form $\omega = d\mathbf{p} \wedge d\mathbf{q}$.

Example 2.1 (Newtonian particle). Classical particle in $\mathbb{R}^n$ moving a potential field $v(\mathbf{q})$ is described by the Hamiltonian function

$$h(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} + v(\mathbf{q}).$$

Assume that the Hamiltonian operator for the quantum system is given by

$$H = \frac{\mathbf{P}^2}{2m} + V,$$

so that coordinate and momenta operators satisfy Heisenberg equations of motion

$$\dot{\mathbf{P}} = \{H, \mathbf{P}\}_\hbar, \ \dot{\mathbf{Q}} = \{H, \mathbf{Q}\}_\hbar. \tag{2.10}$$

To determine the operator V, we require that the classical relation $\dot{\mathbf{q}} = \dfrac{\mathbf{p}}{m}$ between the velocity and the momentum of a particle is preserved under the quantization, i.e.,

$$\dot{\mathbf{Q}} = \frac{\mathbf{P}}{m}.$$

Since $\{\mathbf{P}^2, \mathbf{Q}\}_\hbar = 2\mathbf{P}$, it follows (2.10) that this condition is equivalent to

$$[V, Q_k] = 0, \quad k = 1, \dots, n.$$

According to Section 2.2, V is a function of commuting operators $Q_1, \dots, Q_n$, and the natural choice[14] is $V = v(\mathbf{Q})$. Thus the Hamiltonian operator of a Newtonian particle is

$$H = \frac{\mathbf{P}^2}{2m} + v(\mathbf{Q}),$$

and in coordinate representation it is a *Schrödinger operator*

$$H = -\frac{\hbar^2}{2m}\Delta + v(\mathbf{q}) \tag{2.11}$$

[14]Confirmed by the agreement of the theory with the experiments.

with the real-valued potential $v(\mathbf{q})$.

The sum of two unbounded, self-adjoint operators is not necessarily self-adjoint, and one needs to describe admissible potentials $v(\mathbf{q})$ for which H is a self-adjoint operator on $L^2(\mathbb{R}^n, d^n\mathbf{q})$. If a potential $v(\mathbf{q})$ is real-valued, locally integrable function on $\mathbb{R}^n$, differential operator (2.11) defines a symmetric operator H with the domain $C_0^2(\mathbb{R}^n)$ — twice continuously differentiable functions on $\mathbb{R}^n$ with compact support. Potentials for which the operator H has no self-adjoint extensions are clearly non-physical. It may also happen that H has several self-adjoint extensions which are specified by some boundary conditions at infinity, which are again non-physical. The only physical case is when the symmetric operator H admits a unique self-adjoint extension, i,e., H is essentially self-adjoint[15]. In Chapter 3 we present necessary conditions for the essential self-adjointness. Here we only mention the von Neumann criterion that if A is a closed operator and $\overline{D(A)} = \mathscr{H}$, then $H = A^*A$ is a positive self-adjoint operator.

EXAMPLE 2.2 (Interacting particles). A closed system of N interacting particles on $\mathbb{R}^3$ is described by the canonical coordinates $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$, the canonical momenta $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_{N)}$, $\mathbf{r}_a, \mathbf{p}_a \in \mathbb{R}^3$, and by the Hamiltonian function

$$h(\mathbf{p}, \mathbf{r}) = \sum_{a=1}^{N} \frac{\mathbf{p}_a^2}{2m_a} + v(\mathbf{r}), \tag{2.12}$$

where m_a is the mass of the a-th particle, $a = 1, \dots, N$. Corresponding Hamiltonian operator H in the coordinate representation has the form

$$H = -\sum_{a=1}^{N} \frac{\hbar^2}{2m_a} \Delta_a + v(\mathbf{r}). \tag{2.13}$$

In particular, when

$$v(\mathbf{r}) = \sum_{1 \le a < b \le N} v(\mathbf{r}_a - \mathbf{r}_b),$$

the Schrödinger operator (2.13) describes the N-body problem in quantum mechanics. The fundamental quantum system is the complex atom, formed by a nuclei of charge Ne and mass M and by N electrons of charge $-e$ and mass m. Denoting by $\mathbf{R} \in \mathbb{R}^3$ the position of a nuclei, and by $\mathbf{r}_1, \dots, \mathbf{r}_N$ positions of the electrons and assuming that the interaction is given by the Coulomb attraction, we get for the Hamiltonian function (2.12)

$$h(\mathbf{P}, \mathbf{p}, \mathbf{R}, \mathbf{r}) = \frac{\mathbf{P}^2}{2M} + \sum_{a=1}^{N} \frac{\mathbf{p}_a^2}{2m} - \sum_{a=1}^{N} \frac{Ne^2}{|\mathbf{R} - \mathbf{r}_a|} + \sum_{1 \le a < b \le N} \frac{e^2}{|\mathbf{r}_a - \mathbf{r}_b|},$$

[15]The closure $\bar{H} = H^{**}$ is self-adjoint.

where $\mathbf{P}$ is the canonical momentum of a nuclei. Corresponding Schrödinger operator H in the coordinate representation has the form[16]

$$H = -\frac{\hbar^2}{2M}\Delta - \sum_{a=1}^{N}\frac{\hbar^2}{2m}\Delta_a - \sum_{a=1}^{N}\frac{Ne^2}{|\mathbf{R}-\mathbf{r}_a|} + \sum_{1\le a<b\le N}\frac{e^2}{|\mathbf{r}_a-\mathbf{r}_b|}.$$

In the case of the hydrogen atom, $N=1$, and the Hamiltonian is

$$H = -\frac{\hbar^2}{2M}\Delta_p - \frac{\hbar^2}{2m}\Delta_e - \frac{e^2}{|\mathbf{r}_p-\mathbf{r}_e|},$$

where $\mathbf{r}_p$ is the position of the proton, and $\mathbf{r}_e$ is the position of the electron. As the first approximation, the proton can be considered as infinitely heavy, so that the hydrogen atom is described by an electron in an attractive Coulomb field $-e^2/|\mathbf{r}|$, where now $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_p$. The corresponding Hamiltonian operator takes the form

$$H = -\frac{\hbar^2}{2m}\Delta - \frac{e^2}{|\mathbf{r}|}. \tag{2.14}$$

We will solve the Schrödinger equation with this Hamiltonian H in Chapter 3.

Example 2.3 (Charged particle in an electromagnetic field). Classical particle of charge e and mass m moving in the time-independent electromagnetic field with scalar and vector potentials $\varphi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$, $\mathbf{r}\in\mathbb{R}^3$, is described by the Hamiltonian function

$$h(\mathbf{p},\mathbf{r}) = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{r})\right)^2 + e\varphi(\mathbf{r}).$$

Corresponding Hamiltonian operator in the coordinated representation is given by

$$H = \frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial\mathbf{r}} - \frac{e}{c}\mathbf{A}(\mathbf{r})\right)^2 + e\varphi(\mathbf{r}) \tag{2.15}$$

— the Schrödinger operator of a particle in an electromagnetic field.

2.5. Harmonic oscillator. The simplest classical system with one degree of freedom is a harmonic oscillator, described by the phase space $\mathbb{R}^2$ with the canonical coordinates p, q, and the Hamiltonian function

$$h(p,q) = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}.$$

Here positive $\omega = \sqrt{\frac{k}{m}}$ (see Section 1.5 in Chapter 1) has a physical meaning of the frequency of oscillations. Namely, the Hamilton's equations

$$\dot p = \{h,p\} = -m\omega^2 q, \quad \dot q = \{h,q\} = \frac{p}{m}$$

[16]Ignoring the fact that electron has spin, see Chapter 6.

with the initial conditions p_0, q_0 are readily solved

$$p(t) = p_0 \cos \omega t - m\omega q_0 \sin \omega t, \tag{2.16}$$

$$q(t) = q_0 \cos \omega t + \frac{1}{m\omega} p_0 \sin \omega t, \tag{2.17}$$

and describe the harmonic motion.

Corresponding Hamiltonian operator is

$$H = \frac{P^2}{2m} + \frac{m\omega^2 Q^2}{2},$$

and in the coordinate representation $\mathscr{H} = L^2(\mathbb{R}, dq)$ is a Schrödinger operator with a quadratic potential,

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dq^2} + \frac{m\omega^2 q^2}{2}.$$

The harmonic oscillator is the simplest non-trivial quantum system whose Schrödinger equation can be solved explicitly. It appears in all problems involving quantized oscillations, namely in molecular and crystalline vibrations. The exact solution of the harmonic oscillator, described below, has remarkable[17] algebraic and analytic properties.

Set temporarily $m = 1$ and consider the operators

$$a = \frac{1}{\sqrt{2\omega\hbar}} \left(\omega Q + iP\right), \quad a^* = \frac{1}{\sqrt{2\omega\hbar}} \left(\omega Q - iP\right), \tag{2.18}$$

defined on $W^{1,2}(\mathbb{R}) \cap \widehat{W}^{1,2}(\mathbb{R})$. It is not difficult to show that a^* is the adjoint operator to a and $a^{**} = a$, so that a is a closed operator. From the Heisenberg's commutation relation (2.2) we get

$$[a, a^*] = I \tag{2.19}$$

on $W^{2,2}(\mathbb{R}) \cap \widehat{W}^{2,2}(\mathbb{R})$. Indeed,

$$aa^* = \frac{P^2 + \omega^2 Q^2}{2\omega\hbar} + \frac{i\omega}{2\omega\hbar}[P, Q] = \frac{P^2 + \omega^2 Q^2}{2\omega\hbar} + \frac{1}{2} I,$$

and

$$a^* a = \frac{P^2 + \omega^2 Q^2}{2\omega\hbar} - \frac{i\omega}{2\omega\hbar}[P, Q] = \frac{P^2 + \omega^2 Q^2}{2\omega\hbar} - \frac{1}{2} I,$$

so that (2.19) holds on $W^{2,2}(\mathbb{R}) \cap \widehat{W}^{2,2}(\mathbb{R})$, and

$$H = \omega\hbar \left(a^* a + \tfrac{1}{2} I\right) = \omega\hbar \left(aa^* - \tfrac{1}{2} I\right).$$

In particular, it follows from the von Neumann criterion that the Hamiltonian operator H is self-adjoint.

[17] Algebraic structure of the exact solution of the harmonic oscillator plays a fundamental role in quantum elecrodynamics and in quantum field theory in general.

The operators a, a^* and $N = a^*a$ satisfy the commutation relations of a nilpotent Lie algebra:

$$[N, a] = -a, \quad [N, a^*] = a^*, \quad [a, a^*] = I. \tag{2.20}$$

It is due to this Lie-algebraic structure that the Heisenberg's equations of motion for and the Schrödinger equation for the harmonic oscillator can be solved exactly.

Indeed, the Heisenberg's equations of motion for the operators a and a^* have the form

$$\dot{a} = \{H, a\}_\hbar = -i\omega a \quad \text{and} \quad \dot{a}^* = \{H, a^*\}_\hbar = i\omega a^*,$$

and are readily solved

$$a(t) = e^{-\omega t} a_0, \quad a^*(t) = e^{i\omega t} a_0^*,$$

where a_0, a_0^* are the initial conditions. Returning to the operators

$$P = i\sqrt{\frac{\omega\hbar}{2}}\,(a^* - a), \quad Q = \sqrt{\frac{\hbar}{2\omega}}\,(a^* + a),$$

we get

$$P(t) = P_0 \cos\omega t - \omega Q_0 \sin\omega t, \tag{2.21}$$

$$Q(t) = Q_0 \cos\omega t + \frac{1}{\omega} P_0 \sin\omega t. \tag{2.22}$$

Therefore (note that we have set $m = 1$), solutions of classical and quantum equations of motion for the harmonic oscillator have the same form!

Now consider the Schrödinger equation

$$H\psi = E\psi$$

for the harmonic oscillator. We will show that the Hamiltonian H has purely discrete spectrum by determining explicitly the eigenvalues — the energy levels of the harmonic oscillator, and the corresponding complete system of the eigenvectors. First, suppose that the following properties hold.

I. There exists a non-zero $\psi \in \mathscr{H}$ such that

$$H\psi = E\psi.$$

II. For all $n \in \mathbb{N}$, $\psi \in D(a^n) \cap D((a^*)^n)$.

Then the following statements hold.

(a) There exists $\psi_0 \in \mathscr{H}$, $\|\psi_0\| = 1$, such that

$$H\psi_0 = \tfrac{\hbar\omega}{2}\psi_0$$

— the *ground state vector* for the harmonic oscillator.

(b) The vectors

$$\psi_n = \frac{(a^*)^n}{\sqrt{n!}}\psi_0 \in \mathscr{H}, \quad n = 0, 1, 2, \ldots, \tag{2.23}$$

satisfy

$$a^*\psi_n = \sqrt{n+1}\,\psi_{n+1}, \quad a_n\psi_n = \sqrt{n}\,\psi_{n-1}, \tag{2.24}$$

and are the orthonormal eigenvectors for H with the eigenvalues $\hbar\omega(n+\frac{1}{2})$,

$$H\psi_n = \hbar\omega(n+\tfrac{1}{2})\psi_n.$$

(c) Operator H is essentially self-adjoint on the Hilbert space $\mathscr{H}_0$ — a closed subspace of $\mathscr{H}$, spanned by the orthonormal set $\{\psi_n\}_{n=0}^{\infty}$.

It is easy to prove these statements. If a non-zero $\psi \in \mathscr{H}$ satisfies properties **I-II**, then rewriting commutation relations (2.20) as

$$Na = a(N-I) \quad \text{and} \quad Na^* = a^*(N+I),$$

we get for all $n \geq 0$,

$$Na^n\psi = (E-n)a^n\psi \quad \text{and} \quad N(a^*)^n\psi = (E+n)(a^*)^n\psi. \tag{2.25}$$

Since $N \geq 0$ on $D(N)$, it follows from the first equation in (2.25) that there exists $n_0 \geq 0$ such that $a^{n_0}\psi \neq 0$ but $a^{n_0+1}\psi = 0$. Setting $\psi_0 = \dfrac{a^{n_0}\psi}{\|a^{n_0}\psi\|} \in \mathscr{H}$ we get

$$a\psi_0 = 0 \quad \text{and} \quad N\psi_0 = 0. \tag{2.26}$$

Since $H = \hbar\omega(N + \frac{1}{2}I)$, this proves (a). To prove (b) we observe that it follows from the commutation relation (2.19) and the Leibniz rule that

$$[a, (a^*)^n] = n(a^*)^{n-1}, \tag{2.27}$$

and we get

$$\begin{aligned}\|(a^*)^n\psi_0\|^2 &= ((a^*)^n\psi_0, (a^*)^n\psi_0) = (\psi_0, a^{n-1}a(a^*)^n\psi_0)\\ &= n(\psi_0, a^{n-1}(a^*)^{n-1}\psi_0) + (\psi_0, a^{n-1}(a^*)^n a\psi_0)\\ &= n\|(a^*)^{n-1}\psi_0\|^2 = \cdots = n!\|\psi_0\|^2 = n!\end{aligned}$$

The first equation in (2.24) is trivial, the second equation follows from (2.26)-(2.27), and we conclude that ψ_n are normalized eigenvectors of H with the eigenvalues $\hbar\omega(n+\frac{1}{2})$. The last fact also follows from the second equation in (2.25). The eigenvectors ψ_n are orthogonal since the corresponding eigenvalues are distinct and the operator H is symmetric. Finally, property (c) immediately follows from the fact that, according to (b), the subspaces $\operatorname{Im}(H \pm iI)|_{\mathscr{H}_0}$ are dense in $\mathscr{H}_0$.

Remark. Since coordinate representation of the Heisenberg's commutation relations is irreducible, it is tempting to conclude from the properties (a)-(c) that $\mathscr{H}_0 = \mathscr{H}$. Namely, it follows from the construction that the linear span of vectors ψ_n — a dense subspace of $\mathscr{H}_0$, is invariant for the operators P and Q. However, we can not immediately conclude that the projection operator Π_0 onto the subspace $\mathscr{H}_0$ commutes with P and Q.

By using the coordinate representation, we can prove properties **I-II** and that $\mathscr{H}_0 = \mathscr{H}$. Namely, equation $a\psi_0 = 0$ becomes a first order linear differential equation

$$\left(\hbar\frac{d}{dq} + \omega q\right)\psi_0 = 0,$$

so that

$$\psi_0(q) = \sqrt[4]{\frac{\omega}{\pi\hbar}}\, e^{-\frac{\omega}{2\hbar}q^2}$$

and

$$\|\psi_0\|^2 = \sqrt{\frac{\omega}{\pi\hbar}}\int_{-\infty}^{\infty} e^{-\frac{\omega}{\hbar}q^2}\, dq = 1.$$

Correspondingly, the eigenfunctions

$$\psi_n(q) = \frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2\omega\hbar}}\left(\omega q - \hbar\frac{d}{dq}\right)\right)^n \psi_0$$

are of the form $P_n(q)e^{-\frac{\omega}{2\hbar}q^2}$, where $P_n(q)$ are polynomials of degree n. Thus to prove that the functions $\{\psi_n\}_{n=0}^{\infty}$ form an orthonormal basis in $L^2(\mathbb{R})$, it is sufficient to show that the system of functions $\{q^n e^{-q^2}\}_{n=0}^{\infty}$ is complete in $L^2(\mathbb{R})$. Namely, suppose that $f \in L^2(\mathbb{R})$ is such that

$$\int_{-\infty}^{\infty} f(q)q^n e^{-q^2} dq = 0 \quad \text{for all} \quad n = 0, 1, 2, \dots$$

For $z \in \mathbb{C}$ let

$$F(z) = \int_{-\infty}^{\infty} f(q)e^{iqz-q^2} dq.$$

Clearly, the integral is absolutely convergent for all $z \in \mathbb{C}$ and defines an entire function. We have

$$F^{(n)}(0) = i^n \int_{-\infty}^{\infty} f(q)q^n e^{-q^2} dq = 0 \quad \text{for all} \quad n = 0, 1, 2, \dots,$$

so that $F(z) = 0$ for all $z \in \mathbb{C}$. This implies that for the function $g(q) = f(q)e^{-q^2} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ we have $\mathscr{F}(g) = 0$, where $\mathscr{F}$ is the "ordinary" ($\hbar = 1$) Fourier transform. Thus $g = 0$ and, therefore, $f = 0$.

The polynomials P_n are expressed through classical Hermite-Tchebyscheff polynomials H_n, defined by

$$H_n(q) = (-1)^n e^{q^2}\frac{d^n}{dq^n}\, e^{-q^2}, \quad n = 0, 1, 2 \dots$$

Namely, using the identity

$$\begin{aligned} e^{\frac{q^2}{2}}\frac{d^n}{dq^n}\, e^{-q^2} &= -\left(q - \frac{d}{dq}\right)\left[e^{\frac{q^2}{2}}\frac{d^{n-1}}{dq^{n-1}}\, e^{-q^2}\right] \\ &= \cdots = (-1)^n\left(q - \frac{d}{dq}\right)^n e^{-\frac{q^2}{2}} \end{aligned}$$

we obtain

$$\psi_n(q) = \sqrt[4]{\frac{\omega}{\pi\hbar}}\frac{1}{\sqrt{2^n n!}}\, e^{-\frac{\omega}{2\hbar}q^2} H_n\left(\sqrt{\frac{\omega}{\hbar}}\, q\right).$$

We summarize these results as follows.

THEOREM 2.1. *The Hamiltonian*

$$H = -\frac{\hbar^2}{2}\frac{d^2}{dq^2} + \frac{m\omega^2 q^2}{2}$$

of a harmonic oscillator with one degree of freedom is a self-adjoint operator on $\mathscr{H} = L^2(\mathbb{R})$ *with the domain* $D(\bar{H}) = W^{2,2}(\mathbb{R}) \cap \widehat{W}^{2,2}(\mathbb{R})$, *and with pure point spectrum*

$$H\psi_n = E_n\psi_n, \quad n = 0, 1, 2, \ldots$$

with the energy levels $E_n = \hbar\omega(n + \frac{1}{2})$. *The corresponding eigenfunctions* ψ_n *form an orthonormal basis for* $\mathscr{H}$ *and are given by*

$$\psi_n(q) = \sqrt[4]{\frac{m\omega}{\pi\hbar}}\frac{1}{\sqrt{2^n n!}}\, e^{-\frac{m\omega}{2\hbar}q^2} H_n\left(\sqrt{\frac{m\omega}{\hbar}}\, q\right), \tag{2.28}$$

where $H_n(q)$ *are classical Hermite-Tchebyscheff polynomials.*

PROOF. Consider the operator H defined on the Schwartz space $\mathscr{S}(\mathbb{R})$ of rapidly decreasing functions. Since the operator H is symmetric and has a complete system of eigenvectors in $\mathscr{S}(\mathbb{R})$, the subspaces $\mathrm{Im}(H \pm iI)$ are dense $\mathscr{H}$, so that H is essentially self-adjoint. The proof that its self-adjoint closure (which we continue to denote by H) has the domain $W^{2,2}(\mathbb{R}) \cap \widehat{W}^{2,2}(\mathbb{R})$, is left to the reader. □

REMARK. Since the energy levels of the Hamiltonian H are equidistant by $\hbar\omega$, the quantum harmonic oscillator describes the system of identical "quanta" with the energy $\hbar\omega$. The state $|0\rangle = \psi_0$, in Dirac's notations, is the *vacuum state* with no quanta present and with the vacuum energy $\frac{1}{2}\hbar\omega$, and the states $|n\rangle = \psi_n$ consist of n quanta with the energy $\hbar\omega(n + \frac{1}{2})$. According to (2.24), the operator a^* adds one quant to the state $|n\rangle$ and is called a *creation operator*, and the operator a destroys one quant in the state $|n\rangle$ and is called an *annihilation operator*.

Example of harmonic oscillator illustrates the dramatic difference between the quantum motion and the motion in classical mechanics. The classical motion in the potential $V(q) = \frac{1}{2}m\omega^2 q^2$ is finite: a particle with energy E moves in the region $|\omega q| \le \sqrt{\frac{2E}{m}}$, whereas there is always an non-zero probability of finding quantum particle outside the classical region. Thus for the ground state energy $E = \frac{1}{2}\hbar\omega$ this probability is

$$\int_{|q|\ge\sqrt{\frac{\hbar}{m\omega}}} |\psi_0(q)|^2 dq = \frac{2}{\sqrt{\pi}}\int_1^\infty e^{-x^2} dx \simeq 0.1572992070.$$

Harmonic oscillator with n degrees of freedom is a classical system described by the phase space $\mathbb{R}^{2n}$ with the canonical coordinates $\mathbf{p}, \mathbf{q}$, and by the Hamiltonian function

$$h(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} + v(\mathbf{q}),$$

where $v(\mathbf{q})$ is a positive-definite quadratic form on $\mathbb{R}^n$. By applying an orthogonal transformation to both coordinates $\mathbf{q}$ and momenta $\mathbf{p}$, we can assume that the form $v(\mathbf{q})$ is diagonal and

$$h(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} + \sum_{j=1}^{n} \frac{m\omega_j^2 q_j^2}{2}$$

with $\omega_1, \dots, \omega_n > 0$. The motion of the system is described by an independent harmonic motions in (p_j, q_j)-planes with frequences ω_j, $j = 1, \dots, n$.

Corresponding Hamiltonian operator is

$$H = \frac{\mathbf{P}^2}{2m} + \sum_{j=1}^{n} \frac{m\omega_j^2 Q_j^2}{2}$$

and in the coordinate representation $\mathscr{H} = L^2(\mathbb{R}^n, d^n\mathbf{q})$ is a Schrödinger operator with quadratic potential,

$$H = -\frac{\hbar^2}{2m}\Delta + \sum_{j=1}^{n} \frac{m\omega_j^2 q_j^2}{2}.$$

Hamiltonian H is a self-adjoint operator with $D(H) = W^{2,2}(\mathbb{R}^n) \cap \widehat{W}^{2,2}(\mathbb{R}^n)$ and a pure discrete spectrum. Corresponding eigenfunctions

$$\psi_{\mathbf{k}}(\mathbf{q}) = \psi_{k_1}(q_1) \dots \psi_{k_n}(q_n),$$

where $\mathbf{k} = (k_1, \dots, k_n)$ and $\psi_{k_j}(q_j)$ are eigenfunctions (2.28) with $\omega = \omega_j$, form an orthonormal basis for $L^2(\mathbb{R}^n)$. Corresponding energy levels are given by

$$E_{\mathbf{k}} = \hbar\omega_1(k_1 + \tfrac{1}{2}) + \dots + \hbar\omega_n(k_n + \tfrac{1}{2}).$$

The spectrum of H is simple if and only if $\hbar\omega_1, \dots, \hbar\omega_n$ are linear independent over $\mathbb{Z}$. The highest degeneracy case is $\omega_1 = \dots = \omega_n = \omega$, when the multiplicity of the eigenvalue

$$E_{\mathbf{k}} = \hbar\omega \sum_{j=1}^{n} (k_j + \tfrac{1}{2})$$

is the partition function $p_n(|\mathbf{k}|)$ — the number of representations of the integer $|\mathbf{k}| = k_1 + \dots + k_n$ as a sum of n non-negaitve integers. Setting $m = 1$ and introducing the operators[18]

$$a_j = \frac{1}{\sqrt{2\omega\hbar}} \left(\omega Q_j + iP_j\right), \quad a_j^* = \frac{1}{\sqrt{2\omega\hbar}} \left(\omega Q_j - iP_j\right), \quad j = 1, \dots, n, \tag{2.29}$$

[18]Here using the standard Euclidean metric on $\mathbb{R}^n$, we lowered the indices for P^j.

we get commutation relations for creation and annihilation operators for n degrees of freedom,

$$[a_j, a_l] = 0, \quad [a_j^*, a_l^*] = 0, \quad [a_j, a_l^*] = \delta_{jl} I, \quad j, l = 1, \dots, n, \tag{2.30}$$

which generalize relation (2.18) for the one degree of freedom. The operators a_j, a_j^* and $N_j = a_j^* a_j$, $j = 1, \dots, n$, satisfy commutation relations of a nilpotent Lie algebra, a direct sum of n copies of the nilpotent Lie algebra (2.20). In particular, the operator

$$N = \sum_{j=1}^{n} a_j^* a_j$$

satisfies

$$[N, a_j] = -a_j, \quad [N, a_j^*] = a_j^*, \quad j = 1, \dots, n,$$

and $H = \hbar\omega(N + \frac{n}{2} I)$.

PROBLEM 2.6. Show that $\langle H|M\rangle \geq \frac{1}{2}\hbar\omega$ for every $M \in \mathscr{S}$, where H is the Hamiltonian of harmonic oscillator with one degree of freedom.

PROBLEM 2.7. Let $q(t) = A\cos(\omega t + \alpha)$ be the classical trajectory of the harmonic oscillator with $m = 1$ and the energy $E = \frac{\omega^2 A^2}{2}$, and let μ_α be the probability measure on $\mathbb{R}$ supported at the point $q(t)$. Show that the convex linear combination of the measures μ_α, $0 \leq \alpha \leq 2\pi$, is the probability measure on $\mathbb{R}$ with the distribution function $\mu(q) = \frac{\theta(A^2 - q^2)}{\pi\sqrt{A^2 - q^2}}$, where $\theta(q)$ is the Heavyside step function.

PROBLEM 2.8. Show that when $n \to \infty$ and $\hbar \to 0$ such that $\hbar\omega(n + \frac{1}{2}) = \frac{\omega^2 A^2}{2}$ remains fixed, the envelope of the distribution function $|\psi_n(q)|^2$ on the interval $|q| \leq A$ coincides with the classical distribution function $\mu(q)$ from the previous problem. (*Hint:* Prove the integral representation

$$e^{-q^2} H_n(q) = \frac{2^{n+1}}{\sqrt{\pi}} \int_0^\infty e^{-y^2} y^n \cos(2qy - \tfrac{n\pi}{2}) dy,$$

and derive the asymptotic formula

$$\psi_n(q) = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt[4]{A^2 - q^2}} \cos\Big\{ \frac{\omega}{2\hbar} \Big(A^2 \sin^{-1} \frac{q}{A} + q\sqrt{A^2 - q^2} - \frac{A^2 \pi}{2} \Big) + O(1) \Big\}$$

when $\hbar \to 0$ and $n + \frac{1}{2} = \frac{\omega A^2}{2\hbar}$, $|q| < A$.)

PROBLEM 2.9. Complete the proof of Theorem 2.1.

PROBLEM 2.10. Show that the operators $E_{ij} = a_i^* a_j$, $i, j = 1, \dots, n$, satisfy the commutation relations of the Lie algebra $\mathrm{sl}(n, \mathbb{C})$.

2.6. Holomorphic representation and Wick symbols.

Let

$$\ell^2 = \left\{ c = \{c_n\}_{n=0}^\infty \,:\, \|c\|^2 = \sum_{n=0}^\infty |c_n|^2 < \infty \right\}$$

be the Hilbert ℓ^2-space. The choice of an orthonormal basis $\{\psi_n\}_{n=0}^\infty$ for $L^2(\mathbb{R}, dq)$, given by the eigenfunctions (2.28) of the Schrödinger operator for

the harmonic oscillator, defines the Hilbert space isomorphism $L^2(\mathbb{R}, dq) \simeq \ell^2$,

$$L^2(\mathbb{R}, dq) \ni \psi = \sum_{n=0}^{\infty} c_n \psi_n \mapsto c = \{c_n\}_{n=0}^{\infty} \in \ell^2,$$

where

$$c_n = (\psi, \psi_n) = \int_{-\infty}^{\infty} \psi(q)\psi_n(q)dq,$$

since the functions ψ_n are real-valued. Using (2.24) we get

$$a^*\psi = \sum_{n=0}^{\infty} c_n a^*\psi_n = \sum_{n=0}^{\infty} \sqrt{n+1}\, c_n \psi_{n+1} = \sum_{n=1}^{\infty} \sqrt{n}\, c_{n-1}\psi_n, \quad \psi \in D(a^*),$$

and

$$a\psi = \sum_{n=0}^{\infty} c_n a\psi_n = \sum_{n=0}^{\infty} \sqrt{n}\, c_n \psi_{n-1} = \sum_{n=0}^{\infty} \sqrt{n+1}\, c_{n+1}\psi_n, \quad \psi \in D(a),$$

so that in ℓ^2 creation and annihilation operators a^* and a are represented by the following semi-infinite matrices:

$$a = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & . \\ 0 & 0 & \sqrt{2} & 0 & . \\ 0 & 0 & 0 & \sqrt{3} & . \\ . & . & . & . & . \end{bmatrix}, \quad a^* = \begin{bmatrix} 0 & 0 & 0 & 0 & . \\ \sqrt{1} & 0 & 0 & 0 & . \\ 0 & \sqrt{2} & 0 & 0 & . \\ 0 & 0 & \sqrt{3} & 0 & . \\ . & . & . & . & . \end{bmatrix}.$$

As the result,

$$N = a^*a = \begin{bmatrix} 0 & 0 & 0 & 0 & . \\ 0 & 1 & 0 & 0 & . \\ 0 & 0 & 2 & 0 & . \\ 0 & 0 & 0 & 3 & . \\ . & . & . & . & . \end{bmatrix},$$

so that the Hamiltonian of the harmonic oscillator is represented by a diagonal matrix,

$$H = \hbar\omega(N + \tfrac{1}{2}) = \mathrm{diag}\{\tfrac{1}{2}\hbar\omega, \tfrac{3}{2}\hbar\omega, \tfrac{5}{2}\hbar\omega, \dots\}.$$

This representation of the Heisenberg's commutation relations is called *representation by the occupation numbers*, and has the property that in this representation the Hamiltonian H of the harmonic operator is diagonal.

Another representation where H is diagonal is constructed as follows. Let $\mathscr{D}$ be the space of entire functions $f(z)$ with the inner product

$$(f, g) = \frac{1}{\pi} \iint_{\mathbb{C}} f(z)\overline{g(z)}e^{-|z|^2} d^2z,$$

where $d^2z = \frac{i}{2} dz \wedge d\bar{z}$ is the Lebesgue measure on $\mathbb{C} \simeq \mathbb{R}^2$. It is easy to check that $\mathscr{D}$ is a Hilbert space with the orthonormal basis

$$f_n(z) = \frac{z^n}{\sqrt{n!}}, \quad n = 0, 1, 2, \dots$$

The correspondence

$$\ell^2 \ni c = \{c_n\}_{n=0}^{\infty} \mapsto f(z) = \sum_{n=0}^{\infty} c_n f_n(z) \in \mathscr{D},$$

establishes the Hilbert space isomorphism $\ell^2 \simeq \mathscr{D}$. The realization of a Hilbert space $\mathscr{H}$ as the Hilbert space $\mathscr{D}$ of entire functions is called a *holomorphic representation*. In the holomorphic representation,

$$a^* = z, \quad a = \frac{d}{dz}, \quad \text{and} \quad H = \hbar\omega\left(z\frac{d}{dz} + \frac{1}{2}\right),$$

and it is very easy to show that a^* is the adjoint operator to a. The mapping

$$\mathscr{H} \ni \psi = \sum_{n=0}^{\infty} c_n \psi_n \mapsto f(z) = \sum_{n=0}^{\infty} c_n f_n(z) \in \mathscr{D},$$

establishes the isomorphism between the coordinate and holomorphic representations. It follows from the formula for the generating function for Hermite-Tchebyscheff polynomials,

$$\sum_{n=0}^{\infty} H_n(q)\frac{z^n}{n!} = e^{2qz - z^2},$$

that the corresponding unitary operator $U : \mathscr{H} \to \mathscr{D}$ is an integral operator

$$U\psi(z) = \int_{-\infty}^{\infty} U(z, q)\psi(q)dq$$

with the kernel

$$U(z, q) = \sum_{n=0}^{\infty} \psi_n(q) f_n(z) = \sqrt[4]{\frac{m\omega}{\pi\hbar}}\, e^{\frac{m\omega}{2\hbar}q^2 - \left(\sqrt{\frac{m\omega}{\hbar}}q - \frac{1}{\sqrt{2}}z\right)^2}.$$

Another useful realization is a representation in the Hilbert space $\bar{\mathscr{D}}$ of anti-holomorphic functions $f(\bar{z})$ on $\mathbb{C}$ with the inner product

$$(f, g) = \frac{1}{\pi}\iint\limits_{\mathbb{C}} f(\bar{z})\overline{g(\bar{z})}e^{-|z|^2} d^2z,$$

given by

$$a^* = \bar{z}, \quad a = \frac{d}{d\bar{z}}.$$

This representation is used to introduce the so-called *Wick symbols* of the operators. Namely, let A be an operator in $\bar{\mathscr{D}}$ which is a polynomial with constant coefficients in creation and annihilation operators a^* and a. Using commutation relation (2.19), we can move all operators a^* to the left of the

operators a, and represent A in the *Wick normal form* as the following finite sum

$$A = \sum_{l,m} A_{lm} (a^*)^l a^m. \tag{2.31}$$

By definition, the Wick symbol $A(\bar{z}, z)$ of the operator A is

$$A(\bar{z}, z) = \sum_{l,m} A_{lm} \bar{z}^l z^m. \tag{2.32}$$

It is a restriction of a polynomial $A(v, z)$ in variables v and z to $v = \bar{z}$.

In order to define Wick symbols of bounded operators in $\bar{\mathscr{D}}$, we consider the family of *coherent states*, or *Poisson vectors* $\Phi_v \in \bar{\mathscr{D}}$, $v \in \mathbb{C}$, defined by

$$\Phi_v(\bar{z}) = e^{v\bar{z}}, \quad z \in \mathbb{C}.$$

They satisfy the properties

$$a\Phi_v = v\Phi_v \quad \text{and} \quad f(\bar{v}) = (f, \Phi_v), \quad f \in \bar{\mathscr{D}},\ v \in \mathbb{C}. \tag{2.33}$$

Indeed, the first property is trivial, whereas the "reproduction property" immediately follows from the formula

$$\Phi_v(\bar{z}) = \sum_{n=0}^{\infty} f_n(v) \bar{f}_n(\bar{z}), \tag{2.34}$$

where $\bar{f}_n(\bar{z}) = \overline{f_n(z)}$, $n = 0, 1, 2 \ldots$, is the orthonormal basis for $\bar{\mathscr{D}}$.

We also have

$$(f, g) = \frac{1}{\pi} \iint_{\mathbb{C}} (f, \Phi_v) \overline{(g, \Phi_v)} e^{-|v|^2} d^2 v. \tag{2.35}$$

Now for the operator A in the Wick normal form (2.31) we get, using the first property in (2.33),

$$\begin{aligned} (A\Phi_z, \Phi_{\bar{v}}) &= \sum_{l,m} A_{lm} ((a^*)^l a^m \Phi_z, \Phi_{\bar{v}}) = \sum_{l,m} A_{lm} (a^m \Phi_z, a^l \Phi_{\bar{v}}) \\ &= A(v, z)(\Phi_z, \Phi_{\bar{v}}). \end{aligned}$$

Therefore,

$$A(v, z) = \frac{(A\Phi_z, \Phi_{\bar{v}})}{(\Phi_z, \Phi_{\bar{v}})} = e^{-vz} (A\Phi_z, \Phi_{\bar{v}}),$$

since by the reproduction property, $(\Phi_z, \Phi_{\bar{v}}) = \Phi_z(v) = e^{vz}$.

Definition. The Wick symbol $A(\bar{z}, z)$ of a bounded operator A in the Hilbert space $\bar{\mathscr{D}}$ is a restriction to $v = \bar{z}$ of an entire function $A(v, z)$ in variables v and z, defined by

$$A(v, z) = e^{-vz} (A\Phi_z, \Phi_{\bar{v}}).$$

In the next theorem, we summarize the properties of the Wick symbols.

THEOREM 2.2. *Wick symbols of the bounded operators on $\bar{\mathscr{D}}$ have the following properties.*

(i) *If $A(\bar{z}, z)$ is the Wick symbol an operator A, then the Wick symbol of the operator A^* we have $A^*(\bar{z}, z) = \overline{A(\bar{z}, z)}$ and*

$$A^*(v, z) = \overline{A(\bar{z}, \bar{v})}.$$

(ii) *For $f \in \bar{\mathscr{D}}$,*

$$(Af)(\bar{z}) = \frac{1}{\pi} \iint_{\mathbb{C}} A(\bar{z}, v) f(\bar{v}) e^{-v(\bar{v}-\bar{z})} d^2 v.$$

(iii) *A real-analytic function $A(\bar{z}, z)$ if a Wick symbol of a bounded operator A in $\bar{\mathscr{D}}$ if and only if it is a restriction to $v = \bar{z}$ of an entire function $A(v, z)$ in variables v and z with the property that for every $f \in \bar{\mathscr{D}}$ the integral in part (ii) is absolutely convergent and defines a function in $\bar{\mathscr{D}}$.*

(iv) *If $A_1(\bar{z}, z)$ and $A_2(\bar{z}, z)$ are the Wick symbols of operators A_1 and A_2, then the Wick symbol of the operator $A = A_1 A_2$ is given by*

$$A_1(\bar{z}, z) = \frac{1}{\pi} \iint_{\mathbb{C}} A_1(\bar{z}, v) A_2(\bar{v}, z) e^{-(v-z)(\bar{v}-\bar{z})} d^2 v.$$

PROOF. We have

$$A^*(v, z) = e^{-vz}(A^* \Phi_z, \Phi_{\bar{v}}) = e^{-vz}(\Phi_z, A\Phi_{\bar{v}}) = e^{-vz}\overline{(A\Phi_{\bar{v}}, \Phi_z)} = \overline{A(\bar{z}, \bar{v})},$$

which proves (i). To prove (ii), we use the reproduction property to get

$$(Af)(\bar{z}) = (Af, \Phi_z) = (f, A^* \Phi_z) = \frac{1}{\pi} \iint_{\mathbb{C}} f(\bar{v}) \overline{(A^* \Phi_z)(\bar{v})} e^{-|v|^2} d^2 v.$$

Using the reproduction property once again we have

$$(A^* \Phi_z)(\bar{v}) = (A^* \Phi_z, \Phi_v) = A^*(\bar{v}, z)(\Phi_z, \Phi_v) = \overline{e^{v\bar{z}} A(\bar{z}, v)},$$

which proves (ii). Property (iii) follows from the definition and the uniform boundness principle, which is needed to show that the operator A on $\bar{\mathscr{D}}$, defined by the integral in (ii), is bounded. We leave the standard details to the reader. Finally, to prove (iv) we get, using (2.35) and (i),

$$\begin{aligned} A(\bar{z}, z) &= e^{-|z|^2}(A_1 A_2 \Phi_z, \Phi_z) = e^{-|z|^2}(A_2 \Phi_z, A_1^* \Phi_z) \\ &= \frac{1}{\pi} \iint_{\mathbb{C}} (A_2 \Phi_z, \Phi_v) \overline{(A_1^* \Phi_z, \Phi_v)} e^{-(|v|^2+|z|^2)} d^2 v \\ &= \frac{1}{\pi} \iint_{\mathbb{C}} A_1(\bar{z}, v) A_2(\bar{v}, z) e^{-(v-z)(\bar{v}-\bar{z})} d^2 v. \end{aligned}$$

□

REMARK. Properties (i) and (iv) remain valid for the operators of the form (2.31).

A *matrix symbol* $\tilde{A}(\bar{z}, z)$ of a bounded operator A in the Hilbert space $\bar{\mathscr{D}}$ is a restriction to $v = \bar{z}$ of an entire function $\tilde{A}(v, z)$ in variables v and z, defined by the following absolutely convergent series

$$\tilde{A}(v, z) = \sum_{m,n=0}^{\infty} (A\bar{f}_m, \bar{f}_n) f_n(v) f_m(z). \tag{2.36}$$

The matrix and normal symbols are related as follows.

LEMMA 2.2. *For a bounded operator A in the Hilbert space $\bar{\mathscr{D}}$,*

$$\tilde{A}(v, z) = e^{vz} A(v, z).$$

PROOF. Using (2.34) we get

$$\begin{aligned}\tilde{A}(v, z) &= \frac{1}{\pi} \sum_{m,n=0}^{\infty} f_n(v) f_m(z) \iint_{\mathbb{C}} (A\bar{f}_m)(\bar{u}) f_n(u) e^{-|u|^2} d^2u \\ &= \frac{1}{\pi} \iint_{\mathbb{C}} (A\Phi_z)(\bar{u}) \overline{\Phi_{\bar{v}}(\bar{u})} e^{-|u|^2} d^2u = (A\Phi_z, \Phi_{\bar{v}}) = e^{vz} A(v, z).\end{aligned}$$

Changing the order of summation and integration is justified by the absolute convergence. □

It is straightforward to generalize these constructions to n degrees of freedom. The Hilbert space $\mathscr{D}_n$ defining the holomorphic representation is the space of entire functions $f(\mathbf{z})$ of n complex variables $\mathbf{z} = (z_1, \dots, z_n)$ with the inner product

$$(f, g) = \frac{1}{\pi^n} \int_{\mathbb{C}^n} f(\mathbf{z}) \overline{g(\mathbf{z})} e^{-|\mathbf{z}|^2} d^{2n}\mathbf{z} < \infty,$$

where $|\mathbf{z}|^2 = z_1^2 + \dots + z_n^2$ and $d^{2n}\mathbf{z} = d^2z_1 \cdots d^2z_n$ is the Lebesgue measure on $\mathbb{C}^n \simeq \mathbb{R}^{2n}$. The functions

$$f_{\mathbf{m}}(\mathbf{z}) = \frac{z_1^{m_1} \dots z_n^{m_n}}{\sqrt{m_1! \dots m_n!}}, \qquad m_1, \dots, m_n = 0, 1, 2, \dots,$$

form an orthonormal basis for $\mathscr{D}_n$. Corresponding creation and annihilation operators are given by

$$a_j^* = z_j, \quad a_j = \frac{\partial}{\partial z_j}, \quad j = 1, \dots, n.$$

The Hilbert space $\bar{\mathscr{D}}_n$ of anti-holomorphic functions $f(\bar{\mathbf{z}})$ on $\mathbb{C}^n$ is defined by the inner product

$$(f, g) = \frac{1}{\pi^n} \int_{\mathbb{C}^n} f(\bar{\mathbf{z}}) \overline{g(\bar{\mathbf{z}})} e^{-|\mathbf{z}|^2} d^{2n}\mathbf{z} < \infty,$$

and the creation and annihilation operators are given by

$$a_j^* = \bar{z}_j, \quad a_j = \frac{\partial}{\partial \bar{z}_j}, \quad j = 1, \dots, n.$$

The coherent states are $\Phi_{\mathbf{v}}(\bar{\mathbf{z}}) = e^{\mathbf{v}\bar{\mathbf{z}}}$, where $\mathbf{v}\bar{\mathbf{z}} = v_1\bar{z}_1 + \cdots + v_n z_n$, and satisfy the reproduction property

$$f(\bar{\mathbf{v}}) = (f, \Phi_{\mathbf{v}}), \quad f \in \bar{\mathscr{D}}_n, \ \ \mathbf{v} \in \mathbb{C}^n.$$

The Wick symbol $A(\bar{\mathbf{z}}, \mathbf{z})$ of a bounded operator A on $\bar{\mathscr{D}}_n$ is defined as a restriction to $\mathbf{v} = \bar{\mathbf{z}}$ of an entire function $A(\mathbf{v}, \mathbf{z})$ of $2n$ variables $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{z} = (z_1, \dots, z_n)$, given by

$$A(\mathbf{v}, \mathbf{z}) = e^{-\mathbf{v}\mathbf{z}}(A\Phi_{\mathbf{z}}, \Phi_{\bar{\mathbf{v}}}).$$

We have

$$(Af)(\bar{\mathbf{z}}) = \frac{1}{\pi^n} \int\limits_{\mathbb{C}^n} A(\bar{\mathbf{z}}, \mathbf{v}) f(\bar{\mathbf{v}}) e^{-\mathbf{v}(\bar{\mathbf{v}} - \bar{\mathbf{z}})} d^{2n}\mathbf{v}, \quad f \in \bar{\mathscr{D}}_n,$$

and the Wick symbol $A(\bar{\mathbf{z}}, \mathbf{z})$ of the operator $A = A_1 A_2$ is given by

$$A_1(\bar{\mathbf{z}}, \mathbf{z}) = \frac{1}{\pi^n} \int\limits_{\mathbb{C}^n} A_1(\bar{\mathbf{z}}, \mathbf{v}) A_2(\bar{\mathbf{v}}, \mathbf{z}) e^{-(\mathbf{v} - \mathbf{z})(\bar{\mathbf{v}} - \bar{\mathbf{z}})} d^{2n}\mathbf{v},$$

where $A_1(\bar{\mathbf{z}}, \mathbf{z})$ and $A_2(\bar{\mathbf{z}}, \mathbf{z})$ are the Wick symbols of the operators A_1 and A_2.

The matrix symbol $\tilde{A}(\bar{\mathbf{z}}, \mathbf{z})$ of a bounded operator A on $\bar{\mathscr{D}}_n$ is defined as a restriction to $\mathbf{v} = \bar{\mathbf{z}}$ of an entire function $\tilde{A}(\mathbf{v}, \mathbf{z})$ of $2n$ variables $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{z} = (z_1, \dots, z_n)$, given by the following absolutely convergent series

$$\tilde{A}(\mathbf{v}, \mathbf{z}) = \sum_{\mathbf{k}, \mathbf{m} = 0}^{\infty} (A\bar{f}_{\mathbf{k}}, \bar{f}_{\mathbf{m}}) f_{\mathbf{m}}(\mathbf{v}) f_{\mathbf{k}}(\mathbf{z}),$$

where $\mathbf{k} = (k_1, \dots, k_n)$, $\mathbf{m} = (m_1, \dots, m_n)$ are multi-indices, and $\bar{f}_{\mathbf{m}}(\bar{\mathbf{z}}) = \overline{f_{\mathbf{m}}(\mathbf{z})}$. The matrix and Wick symbols of a bounded operator A are related by

$$\tilde{A}(\mathbf{v}, \mathbf{z}) = e^{\mathbf{v}\mathbf{z}} A(\mathbf{v}, \mathbf{z}).$$

PROBLEM 2.11. Find an explicit formula for the unitary operator establishing the Hilbert space isomorphism $\bar{\mathscr{D}}_n \simeq L^2(\mathbb{R}^n, d^n\mathbf{q})$.

PROBLEM 2.12. Prove that for bounded operator A the functions $A(\mathbf{v}, \mathbf{z})$ and $\tilde{A}(\mathbf{v}, \mathbf{z})$ are entire functions of $2n$ variables.

PROBLEM 2.13. Let A be a trace class operator on $\bar{\mathscr{D}}_n$ with the Wick symbol $A(\bar{\mathbf{z}}, \mathbf{z})$. Prove that

$$\operatorname{Tr} A = \frac{1}{\pi^n} \int\limits_{\mathbb{C}^n} A(\bar{\mathbf{z}}, \mathbf{z}) e^{-|\mathbf{z}|^2} d^n\mathbf{z}.$$