§ 1. The algebra of observables in classical mechanics

We consider the simplest problem in classical mechanics: the problem of the motion of a material point (a particle) with mass m in a force field $V(\mathbf{x})$, where $\mathbf{x}(x_1, x_2, x_3)$ is the radius vector of the particle. The force acting on the particle is

$$\mathbf{F} = -\operatorname{grad} V = -\frac{\partial V}{\partial \mathbf{x}}.$$

The basic physical characteristics of the particle are its coordinates x_1, x_2, x_3 and the projections of the velocity vector $\mathbf{v}(v_1, v_2, v_3)$. All the remaining characteristics are functions of $\mathbf{x}$ and $\mathbf{v}$; for example, the momentum $\mathbf{p} = m\mathbf{v}$, the angular momentum $\mathbf{l} = \mathbf{x} \times \mathbf{p} = m\mathbf{x} \times \mathbf{v}$, and the energy $E = m\mathbf{v}^2/2 + V(\mathbf{x})$.

The equations of motion of a material point in the Newton form are

$$(1) \qquad m\frac{d\mathbf{v}}{dt} = -\frac{\partial V}{\partial \mathbf{x}}, \qquad \frac{d\mathbf{x}}{dt} = \mathbf{v}.$$

It will be convenient below to use the momentum $\mathbf{p}$ in place of the velocity $\mathbf{v}$ as a basic variable. In the new variables the equations of motion are written as follows:

$$(2) \qquad \frac{d\mathbf{p}}{dt} = -\frac{\partial V}{\partial \mathbf{x}}, \qquad \frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}.$$

Noting that $\frac{\mathbf{p}}{m} = \frac{\partial H}{\partial \mathbf{p}}$ and $\frac{\partial V}{\partial \mathbf{x}} = \frac{\partial H}{\partial \mathbf{x}}$, where $H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$ is the Hamiltonian function for a particle in a potential field, we arrive at the equations in the Hamiltonian form

$$(3) \qquad \frac{d\mathbf{x}}{dt} = \frac{\partial H}{\partial \mathbf{p}}, \qquad \frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial x}.$$

It is known from a course in theoretical mechanics that a broad class of mechanical systems, and conservative systems in particular,

are described by the Hamiltonian equations

$$(4) \qquad \dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, 2, \ldots, n.$$

Here $H = H(q_1, \ldots, q_n; p_1, \ldots, p_n)$ is the Hamiltonian function, q_i and p_i are the generalized coordinates and momenta, and n is called the number of degrees of freedom of the system. We recall that for a conservative system, the Hamiltonian function H coincides with the expression for the total energy of the system in the variables q_i and p_i. We write the Hamiltonian function for a system of N material points interacting pairwise:

$$(5) \qquad H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \sum_{i<j}^{N} V_{ij}(\mathbf{x}_i - \mathbf{x}_j) + \sum_{i=1}^{N} V_i(\mathbf{x}_i).$$

Here the Cartesian coordinates of the particles are taken as the generalized coordinates q, the number of degrees of freedom of the system is $n = 3N$, and $V_{ij}(\mathbf{x}_i - \mathbf{x}_j)$ is the potential of the interaction of the ith and jth particles. The dependence of V_{ij} only on the difference $\mathbf{x}_i - \mathbf{x}_j$ is ensured by Newton's third law. (Indeed, the force acting on the ith particle due to the jth particle is $\mathbf{F}_{ij} = -\frac{\partial V_{ij}}{\partial \mathbf{x}_i} = \frac{\partial V_{ij}}{\partial \mathbf{x}_j} = -\mathbf{F}_{ji}$.) The potentials $V_i(\mathbf{x}_i)$ describe the interaction of the ith particle with the external field. The first term in (5) is the kinetic energy of the system of particles.

For any mechanical system all physical characteristics are functions of the generalized coordinates and momenta. We introduce the set $\mathfrak{A}$ of real infinitely differentiable functions $f(q_1, \ldots, q_n; p_1, \ldots, p_n)$, which will be called observables.[1] The set $\mathfrak{A}$ of observables is obviously a linear space and forms a real algebra with the usual addition and multiplication operations for functions. The real $2n$-dimensional space with elements $(q_1, \ldots, q_n; p_1, \ldots, p_n)$ is called the phase space and is denoted by $\mathcal{M}$. Thus, the algebra of observables in classical mechanics is the algebra of real-valued smooth functions defined on the phase space $\mathcal{M}$.

We shall introduce in the algebra of observables one more operation, which is connected with the evolution of the mechanical system.

---

[1] We do not discuss the question of introducing a topology in the algebra of observables. Fortunately, most physical questions do not depend on this topology.

For simplicity the exposition to follow is conducted using the example
of a system with one degree of freedom. The Hamiltonian equations
in this case have the form

$$(6) \qquad \dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}, \qquad H = H(q, p).$$

The Cauchy problem for the system (6) and the initial conditions

$$(7) \qquad q|_{t=0} = q_0, \qquad p|_{t=0} = p_0$$

has a unique solution

$$(8) \qquad q = q(q_0, p_0, t), \qquad p = p(q_0, p_0, t).$$

For brevity of notation a point (q, p) in phase space will sometimes
be denoted by μ, and the Hamiltonian equations will be written in
the form

$$(9) \qquad \dot{\mu} = v(\mu),$$

where $v(\mu)$ is the vector field of these equations, which assigns to each
point μ of phase space the vector v with components $\frac{\partial H}{\partial p}, \frac{-\partial H}{\partial q}$.

The Hamiltonian equations generate a one-parameter commuta-
tive group of transformations

$$G_t \colon \mathcal{M} \to \mathcal{M}$$

of the phase space into itself,[2] where $G_t\mu$ is the solution of the Hamil-
tonian equations with the initial condition $G_t\mu|_{t=0} = \mu$. We have the
equalities

$$(10) \qquad G_{t+s} = G_t G_s = G_s G_t, \qquad G_t^{-1} = G_{-t}.$$

In turn, the transformations G_t generate a family of transformations

$$U_t \colon \mathfrak{A} \to \mathfrak{A}$$

of the algebra of observables into itself, where

$$(11) \qquad U_t f(\mu) = f_t(\mu) = f(G_t\mu).$$

In coordinates, the function $f_t(q, p)$ is defined as follows:

$$(12) \qquad f_t(q_0, p_0) = f(q(q_0, p_0, t), p(q_0, p_0, t)).$$

---

[2]We assume that the Hamiltonian equations with initial conditions (7) have a
unique solution on the whole real axis. It is easy to construct examples in which
a global solution and, correspondingly, a group of transformations G_t do not exist.
These cases are not interesting, and we do not consider them.

We find a differential equation that the function $f_t(q,p)$ satisfies. To this end, we differentiate the identity $f_{s+t}(\mu) = f_t(G_s\mu)$ with respect to the variable s and set $s = 0$:

$$\frac{\partial f_{s+t}(\mu)}{\partial s}\bigg|_{s=0} = \frac{\partial f_t(\mu)}{\partial t},$$

$$\frac{\partial f_t(G_s\mu)}{\partial s}\bigg|_{s=0} = \nabla f_t(\mu) \cdot v(\mu) = \frac{\partial f_t}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial f_t}{\partial p}\frac{\partial H}{\partial q}.$$

Thus, the function $f_t(q,p)$ satisfies the differential equation

$$(13) \qquad \frac{\partial f_t}{\partial t} = \frac{\partial H}{\partial p}\frac{\partial f_t}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial f_t}{\partial p}$$

and the initial condition

$$(14) \qquad f_t(q,p)|_{t=0} = f(q,p).$$

The equation (13) with the initial condition (14) has a unique solution, which can be obtained by the formula (12); that is, to construct the solutions of (13) it suffices to know the solutions of the Hamiltonian equations.

We can rewrite (13) in the form

$$(15) \qquad \frac{df_t}{dt} = \{H, f_t\},$$

where $\{H, f_t\}$ is the Poisson bracket of the functions H and f_t. For arbitrary observables f and g the Poisson bracket is defined by

$$\{f, g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p},$$

and in the case of a system with n degrees of freedom

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

We list the basic properties of Poisson brackets:

1) $\{f, g + \lambda h\} = \{f, g\} + \lambda\{f, h\}$ (linearity);

2) $\{f, g\} = -\{g, f\}$ (skew symmetry);

3) $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$ (Jacobi identity);

4) $\{f, gh\} = g\{f, h\} + \{f, g\}h.$

The properties 1), 2), and 4) follow directly from the definition of the Poisson brackets. The property 4) shows that the "Poisson bracket" operation is a derivation of the algebra of observables. Indeed, the Poisson bracket can be rewritten in the form

$$\{f, g\} = X_f g,$$

where $X_f = \frac{\partial f}{\partial p}\frac{\partial}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial}{\partial p}$ is a first-order linear differential operator, and the property 4) has the form

$$X_f gh = (X_f g)h + gX_f h.$$

The property 3) can be verified by differentiation, but it can be proved by the following argument. Each term of the double Poisson bracket contains as a factor the second derivative of one of the functions with respect to one of the variables; that is, the left-hand side of 3) is a linear homogeneous function of the second derivatives. On the other hand, the second derivatives of h can appear only in the sum $\{f, \{g, h\}\} + \{g, \{h, f\}\} = (X_f X_g - X_g X_f)h$, but a commutator of first-order linear differential operators is a first-order differential operator, and hence the second derivatives of h do not appear in the left-hand side of 3). By symmetry, the left-hand side of 3) does not contain second derivatives at all; that is, it is equal to zero.

The Poisson bracket $\{f, g\}$ provides the algebra of observables with the structure of a real Lie algebra.[3] Thus, the set of observables has the following algebraic structure. The set $\mathfrak{A}$ is:

   1) a real linear space;
   2) a commutative algebra with the operation fg;
   3) a Lie algebra with the operation $\{f, g\}$.

The last two operations are connected by the relation

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

The algebra $\mathfrak{A}$ of observables contains a distinguished element, namely, the Hamiltonian function H, whose role is to describe the

---

[3]We recall that a linear space with a binary operation satisfying the conditions 1)–3) is called a Lie algebra.

variation of observables with time:

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

We show that the mapping $U_t \colon \mathfrak{A} \to \mathfrak{A}$ preserves all the operations in $\mathfrak{A}$:

$$h = f + g \to h_t = f_t + g_t,$$
$$h = fg \to h_t = f_t g_t,$$
$$h = \{f, g\} \to h_t = \{f_t, g_t\};$$

that is, it is an automorphism of the algebra of observables. For example, we verify the last assertion. For this it suffices to see that the equation and the initial condition for h_t is a consequence of the equations and initial conditions for the functions f_t and g_t:

$$\frac{\partial h_t}{\partial t} = \left\{ \frac{\partial f_t}{\partial t}, g_t \right\} + \left\{ f_t, \frac{\partial g_t}{\partial t} \right\}$$
$$= \{\{H, f_t\}, g_t\} + \{f_t, \{H, g_t\}\} = \{H, \{f_t, g_t\}\} = \{H, h_t\}.$$

Here we used the properties 2) and 4) of Poisson brackets. Furthermore,

$$h_t|_{t=0} = \{f_t, g_t\}|_{t=0} = \{f, g\}.$$

Our assertion now follows from the uniqueness of the solution of the equation (13) with the initial condition (14).

§ 2. States

The concept of a state of a system can be connected directly with the conditions of an experiment. Every physical experiment reduces to a measurement of the numerical value of an observable for the system under definite conditions that can be called the conditions of the experiment. It is assumed that these conditions can be reproduced multiple times, but we do not assume in advance that the measurement will give the same value of the observable when the experiment is repeated. There are two possible answers to the question of how to explain this uncertainty in the results of an experiment.

1) The number of conditions that are fixed in performing the experiments is insufficient to uniquely determine the results of the measurement of the observables. If the nonuniqueness arises only

for this reason, then at least in principle these conditions can be supplemented by new conditions, that is, one can pose the experiment more "cleanly", and then the results of all the measurements will be uniquely determined.

2) The properties of the system are such that in repeated experiments the observables can take different values independently of the number and choice of the conditions of the experiment.

Of course if 2) holds, then insufficiency of the conditions can only increase the nonuniqueness of the experimental results. We discuss 1) and 2) at length after we learn how to describe states in classical and quantum mechanics.

We shall consider that the conditions of the experiment determine the state of the system if conducting many repeated trials under these conditions leads to probability distributions for all the observables. In this case we speak of the measurement of an observable f for a system in the state ω. More precisely, a state ω on the algebra $\mathfrak{A}$ of observables assigns to each observable f a probability distribution of its possible values, that is, a measure on the real line $\mathbf{R}$.

Let f be an observable and E a Borel set on the real line $\mathbf{R}$. Then the definition of a state ω can be written as

$$f, E \xrightarrow{\omega} \omega_f(E).$$

We recall the properties of a probability measure:

(1) $$0 \leqslant \omega_f(E) \leqslant 1, \quad \omega_f(\varnothing) = 0, \quad \omega_f(\mathbf{R}) = 1,$$

and if $E_1 \cap E_2 = \varnothing$, then $\omega_f(E_1 \cup E_2) = \omega_f(E_1) + \omega_f(E_2)$.

Among the observables there may be some that are functionally dependent, and hence it is necessary to impose a condition on the probability distributions of such observables. If an observable φ is a function of an observable f, $\varphi = \varphi(f)$, then this assertion means that a measurement of the numerical value of f yielding a value f_0 is at the same time a measurement of the observable φ and gives for it the numerical value $\varphi_0 = \varphi(f_0)$. Therefore, $\omega_f(E)$ and $\omega_{\varphi(f)}(E)$ are connected by the equality

(2) $$\omega_{\varphi(f)}(E) = \omega_f(\varphi^{-1}(E)),$$

where $\varphi^{-1}(E)$ is the inverse image of E under the mapping φ.

A convex combination

(3) $\omega_f(E) = \alpha\omega_{1f}(E) + (1 - \alpha)\omega_{2f}(E), \quad 0 < \alpha < 1,$

of probability measures has the properties (1) for any observable f and corresponds to a state which we denote by

(4) $\omega = \alpha\omega_1 + (1 - \alpha)\omega_2.$

Thus, the states form a convex set. A convex combination (4) of states ω_1 and ω_2 will sometimes be called a mixture of these states. If for some state ω it follows from (4) that $\omega_1 = \omega_2 = \omega$, then we say that the state ω is indecomposable into a convex combination of different states. Such states are called *pure* states, and all other states are called *mixed* states.

It is convenient to take E to be an interval $(-\infty, \lambda]$ of the real axis. By definition, $\omega_f(\lambda) = \omega_f((-\infty, \lambda])$, and this is the distribution function of the observable f in the state ω. Numerically, $\omega_f(\lambda)$ is the probability of getting a value not exceeding λ when measuring f in the state ω. It follows from (1) that the distribution function $\omega_f(\lambda)$ is a nondecreasing function of λ with $\omega_f(-\infty) = 0$ and $\omega_f(+\infty) = 1$.

The mathematical expectation (mean value) of an observable f in a state ω is defined by the formula[4]

$$\langle f \,|\, \omega \rangle = \int_{-\infty}^{\infty} \lambda \, d\omega_f(\lambda).$$

We remark that knowledge of the mathematical expectations for all the observables is equivalent to knowledge of the probability distributions. To see this, it suffices to consider the function $\theta(\lambda - f)$ of the observables, where $\theta(x)$ is the Heaviside function

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

It is not hard to see that

(5) $\omega_f(\lambda) = \langle \theta(\lambda - f) \,|\, \omega \rangle.$

---

[4]The notation $\langle f \,|\, \omega \rangle$ for the mean value of an observable should not be confused with the Dirac notation often used in quantum mechanics for the scalar product $\langle \varphi, \psi \rangle$ of vectors.

For the mean values of observables we require the following conditions, which are natural from a physical point of view:

(6)
$$
\begin{aligned}
&1) \quad \langle C \,|\, \omega \rangle = C, \\
&2) \quad \langle f + \lambda g \,|\, \omega \rangle = \langle f \,|\, \omega \rangle + \lambda \langle g \,|\, \omega \rangle, \\
&3) \quad \langle f^2 \,|\, \omega \rangle \geqslant 0.
\end{aligned}
$$

If these requirements are introduced, then the realization of the algebra of observables itself determines a way of describing the states. Indeed, the mean value is a positive linear functional on the algebra $\mathfrak{A}$ of observables. The general form of such a functional is

$$
\langle f \,|\, \omega \rangle = \int_{\mathcal{M}} f(p, q) \, d\mu_\omega(p, q), \tag{7}
$$

where $d\mu_\omega(p, q)$ is the differential of the measure on the phase space, and the integral is over the whole of phase space. It follows from the condition 1) that

$$
\int_{\mathcal{M}} d\mu_\omega(p, q) = \mu_\omega(\mathcal{M}) = 1. \tag{8}
$$

We see that a state in classical mechanics is described by specifying a probability distribution on the phase space. The formula (7) can be rewritten in the form

$$
\langle f \,|\, \omega \rangle = \int_{\mathcal{M}} f(p, q) \, \rho_\omega(p, q) \, dp \, dq; \tag{9}
$$

that is, we arrive at the usual description in statistical physics of a state of a system with the help of the distribution function $\rho_\omega(p, q)$, which in the general case is a positive generalized function. The normalization condition of the distribution function has the form

$$
\int_{\mathcal{M}} \rho(p, q) \, dp \, dq = 1. \tag{10}
$$

In particular, it is easy to see that to a pure state there corresponds a distribution function

$$
\rho(p, q) = \delta(q - p_0) \, \delta(p - p_0), \tag{11}
$$

where $\delta(x)$ is the Dirac δ-function. The corresponding measure on the phase space is concentrated at the point (q_0, p_0), and a pure state is defined by specifying this point of phase space. For this reason the

phase space $\mathcal{M}$ is sometimes called the state space. The mean value of an observable f in the pure state ω is

$$\langle f \,|\, \omega \rangle = f(q_0, p_0). \tag{12}$$

This formula follows immediately from the definition of the δ-function:

$$f(q_0, p_0) = \int_{\mathcal{M}} f(q, p)\, \delta(q - q_0)\, \delta(p - p_0)\, dq\, dp. \tag{13}$$

In mechanics courses one usually studies only pure states, while in statistical physics one considers mixed states, with distribution function different from (11). But an introduction to the theory of mixed states from the very start is warranted by the following circumstances. The formulation of classical mechanics in the language of states and observables is nearest to the formulation of quantum mechanics and makes it possible to describe states in mechanics and statistical physics in a uniform way. Such a formulation will enable us to follow closely the passage to the limit from quantum mechanics to classical mechanics. We shall see that in quantum mechanics there are also pure and mixed states, and in the passage to the limit, a pure quantum state can be transformed into a mixed classical state, so that the passage to the limit is most simply described when pure and mixed states are treated in a uniform way.

We now explain the physical meaning of mixed and pure states in classical mechanics, and we find out why experimental results are not necessarily determined uniquely by the conditions of the experiment.

Let us consider a mixture

$$\omega = \alpha\omega_1 + (1 - \alpha)\omega_2, \quad 0 < \alpha < 1,$$

of the states ω_1 and ω_2. The mean values obviously satisfy the formula

$$\langle f \,|\, \omega \rangle = \alpha\langle f \,|\, \omega_1 \rangle + (1 - \alpha)\langle f \,|\, \omega_2 \rangle. \tag{14}$$

The formulas (14) and (3) admit the following interpretation. The assertion that the system is in the state ω is equivalent to the assertion that the system is in the state ω_1 with probability α and in the state ω_2 with probability $(1 - \alpha)$. We remark that this interpretation is possible but not necessary.

The simplest mixed state is a convex combination of two pure states:

$$\rho(q,p) = \alpha\delta(q-q_1)\,\delta(p-p_1) + (1-\alpha)\,\delta(q-q_2)\,\delta(p-p_2).$$

Mixtures of n pure states are also possible:

$$\rho(q,p) = \sum_{i=1}^{n}\alpha_i\delta(q-q_i)\,\delta(p-p_i), \quad \alpha_i > 0, \quad \sum_{i=1}^{n}\alpha_i = 1.$$

Here α_i can be interpreted as the probability of realization of the pure state given by the point q_i, p_i of phase space. Finally, in the general case the distribution function can be written as

$$\rho(q,p) = \int_{\mathcal{M}} \rho(q_0,p_0)\,\delta(q-q_0)\,\delta(p-p_0)\,dq_0\,dp_0.$$

Such an expression leads to the usual interpretation of the distribution function in statistical physics: $\int_{\Omega}\rho(q,p)\,dq\,dp$ is the probability of observing the system in a pure state represented by a point in the domain Ω of phase space. We emphasize once more that this interpretation is not necessary, since pure and mixed states can be described in the framework of a unified formalism.

One of the most important characteristics of a probability distribution is the variance

$$(15) \qquad \Delta_\omega^2 f = \langle (f - \langle f \,|\, \omega\rangle)^2 \,|\, \omega\rangle = \langle f^2 \,|\, \omega\rangle - \langle f \,|\, \omega\rangle^2.$$

We show that mixing of states leads to an increase in the variance. A more precise formulation of this assertion is contained in the inequalities

$$(16) \qquad \Delta_\omega^2 f \geqslant \alpha\Delta_{\omega_1}^2 f + (1-\alpha)\Delta_{\omega_2}^2 f,$$

$$(17) \qquad \Delta_\omega f \Delta_\omega g \geqslant \alpha\Delta_{\omega_1} f \Delta_{\omega_1} g + (1-\alpha)\Delta_{\omega_2} f \Delta_{\omega_2} g,$$

with equalities only if the mean values of the observables in the states ω_1 and ω_2 coincide:

$$\langle f \,|\, \omega_1\rangle = \langle f \,|\, \omega_2\rangle, \quad \langle g \,|\, \omega_1\rangle = \langle g \,|\, \omega_2\rangle.$$

In a weakened form, (16) and (17) can be written as

$$(18) \qquad \Delta_\omega^2 f \geqslant \min(\Delta_{\omega_1}^2 f, \Delta_{\omega_2}^2 f),$$

$$(19) \qquad \Delta_\omega f \Delta_\omega g \geqslant \min(\Delta_{\omega_1} f \Delta_{\omega_1} g, \Delta_{\omega_2} f \Delta_{\omega_2} g).$$

The proof uses the elementary inequality

(20) $\varphi(x) = \alpha + (1-\alpha)x^2 - (\alpha + (1-\alpha)x)^2 \geqslant 0, \quad -\infty < x < \infty,$

with

(21) $\varphi(1) = 0, \quad \varphi(x) > 0 \quad \text{for } x \neq 1.$

Using (14) and (20), we get that

$$
\begin{aligned}
\Delta_\omega^2 f &= \langle f^2 \,|\, \omega \rangle - \langle f \,|\, \omega \rangle^2 \\
&= \alpha \langle f^2 \,|\, \omega_1 \rangle + (1-\alpha)\langle f^2 \,|\, \omega_2 \rangle - [\alpha \langle f \,|\, \omega_1 \rangle + (1-\alpha)\langle f \,|\, \omega_2 \rangle]^2 \\
&\geqslant \alpha \langle f^2 \,|\, \omega_1 \rangle + (1-\alpha)\langle f^2 \,|\, \omega_2 \rangle - \alpha \langle f \,|\, \omega_1 \rangle^2 - (1-\alpha)\langle f \,|\, \omega_2 \rangle^2 \\
&= \alpha \Delta_{\omega_1}^2 f + (1-\alpha)\Delta_{\omega_2}^2 f.
\end{aligned}
$$

The inequality (16) is proved. The inequality (17) is a consequence of (16). Indeed,

$$
\begin{aligned}
\Delta_\omega^2 f \Delta_\omega^2 g &\geqslant (\alpha \Delta_{\omega_1}^2 f + (1-\alpha)\Delta_{\omega_2}^2 f)(\alpha \Delta_{\omega_1}^2 g + (1-\alpha)\Delta_{\omega_2}^2 g) \\
&\geqslant [\alpha \Delta_{\omega_1} f \Delta_{\omega_1} g + (1-\alpha)\Delta_{\omega_2} f \Delta_{\omega_2} g]^2.
\end{aligned}
$$

For pure states,

$$
\rho(q, p) = \delta(q - q_0)\,\delta(p - p_0),
$$
$$
\langle f \,|\, \omega \rangle = f(q_0, p_0),
$$
$$
\Delta_\omega^2 f = f^2(q_0, p_0) - f^2(q_0, p_0) = 0;
$$

that is, for pure states in classical mechanics the variance is zero. This means that for a system in a pure state, the result of a measurement of any observable is uniquely determined. A state of a classical system will be pure if by the time of the measurement the conditions of the experiment fix the values of all the generalized coordinates and momenta. It is clear that if a macroscopic body is regarded as a mechanical system of N molecules, where N usually has order 10^{23}, then no conditions in a real physical experiment can fix the values of q_0 and p_0 for all molecules, and the description of such a system with the help of pure states is useless. Therefore, one studies mixed states in statistical physics.

Let us summarize. In classical mechanics there is an infinite set of states of the system (pure states) in which all observables have completely determined values. In real experiments with systems of a

huge number of particles, mixed states arise. Of course, such states are possible also in experiments with simple mechanical systems. In this case the theory gives only probabilistic predictions.

§ 3. Liouville's theorem, and two pictures of motion in classical mechanics

We begin this section with a proof of an important theorem of Liouville. Let Ω be a domain in the phase space $\mathcal{M}$. Denote by $\Omega(t)$ the image of this domain under the action of a phase flow, that is, the set of points $G_t\mu$, $\mu \in \Omega$. Let $V(t)$ be the volume of $\Omega(t)$. Liouville's theorem asserts that

$$\frac{dV(t)}{dt} = 0.$$

Proof.

$$V(t) = \int_{\Omega(t)} d\mu = \int_{\Omega} \left| \frac{D(G_t\mu)}{D(\mu)} \right| d\mu, \quad d\mu = dq\, dp.$$

Here $D(G_t\mu)/D(\mu)$ denotes the Jacobi determinant of the transformation G_t. To prove the theorem, it suffices to show that

(1) $$\frac{D(G_t\mu)}{D(\mu)} = 1$$

for all t. The equality (1) is obvious for $t = 0$. Let us now show that

(2) $$\frac{d}{dt} \frac{D(G_t\mu)}{D(\mu)} = 0.$$

For $t = 0$ the formula (2) can be verified directly:

$$\frac{d}{dt} \frac{D(G_t\mu)}{D(\mu)} \bigg|_{t=0} = \left[\frac{D(\dot{q}, p)}{D(q, p)} + \frac{D(q, \dot{p})}{D(q, p)} \right] \bigg|_{t=0}$$

$$= \left(\frac{\partial \dot{q}}{\partial q} + \frac{\partial \dot{p}}{\partial p} \right) \bigg|_{t=0} = \frac{\partial^2 H}{\partial q\, \partial p} - \frac{\partial^2 H}{\partial p\, \partial q} = 0.$$

For $t \neq 0$ we differentiate the identity

$$\frac{D(G_{t+s}\mu)}{D(\mu)} = \frac{D(G_{t+s}\mu)}{D(G_t\mu)} \frac{D(G_t\mu)}{D(\mu)}$$

with respect to s and set $s = 0$, getting

$$\frac{d}{dt} \frac{D(G_t\mu)}{D(\mu)} = \left[\frac{d}{dt} \frac{D(G_s G_t\mu)}{D(G_t\mu)} \right]\bigg|_{t=0} \frac{D(G_t\mu)}{D(\mu)} = 0.$$

Thus, (2) holds for all t. The theorem is proved. □

We now consider the evolution of a mechanical system. We are interested in the time dependence of the mean values $\langle f \,|\, \omega \rangle$ of the observables. There are two possible ways of describing this dependence, that is, two pictures of the motion. We begin with the formulation of the so-called Hamiltonian picture. In this picture the time dependence of the observables is determined by the equation (1.15),[5] and the states do not depend on time:

$$\frac{df_t}{dt} = \{H, f_t\}, \qquad \frac{d\rho}{dt} = 0.$$

The mean value of an observable f in a state ω depends on the time according to the formula

$$(3) \qquad \langle f_t \,|\, \omega \rangle = \int_{\mathcal{M}} f_t(\mu)\, \rho(\mu)\, d\mu = \int_{\mathcal{M}} f(G_t\mu)\, \rho(\mu)\, d\mu,$$

or, in more detail,

$$(4) \qquad \langle f_t \,|\, \omega \rangle = \int_{\mathcal{M}} f(q(q_0, p_0, t), p(q_0, p_0, t))\, \rho(q_0, p_0)\, dq_0\, dp_0.$$

We recall that $q(q_0, p_0, t)$ and $p(q_0, p_0, t)$ are solutions of the Hamiltonian equations with initial conditions $q|_{t=0} = q_0$, $p|_{t=0} = p_0$.

For a pure state $\rho(q, p) = \delta(q - q_0)\delta(p - p_0)$, and it follows from (4) that

$$\langle f_t \,|\, \omega \rangle = f(q(q_0, p_0, t), p(q_0, p_0, t)).$$

This is the usual classical mechanics formula for the time dependence of an observable in a pure state.[6] It is clear from the formula (4) that a state in the Hamiltonian picture determines the probability distribution of the initial values of q and p.

---

[5] In referring to a formula in previous sections the number of the corresponding section precedes the number of the formula.

[6] In courses in mechanics it is usual to consider only pure states. Furthermore, no distinction is made between the dependence on time of an abstract observable in the Hamiltonian picture and the variation of its mean value.

An alternative way of describing the motion is obtained if in (3) we make the change of variables $G_t\mu \to \mu$. Then

$$\int_{\mathcal{M}} f(G_t\mu)\,\rho(\mu)\,d\mu = \int_{\mathcal{M}} f(\mu)\,\rho(G_{-t}\mu)\left|\frac{D(G_{-t}\mu)}{D(\mu)}\right|d\mu$$

$$= \int_{\mathcal{M}} f(\mu)\,\rho_t(\mu)\,d\mu = \langle f \,|\, \omega_t \rangle.$$

Here we have used the equality (1) and we have introduced the notation $\rho_t(\mu) = \rho(G_{-t}\mu)$. It is not hard to see that $\rho_t(\mu)$ satisfies the equation

(5) $$\frac{d\rho_t}{dt} = -\{H, \rho_t\},$$

which differs from (1.15) by the sign in front of the Poisson bracket. The derivation of the equation (5) repeats that word-for-word for the equation (1.15), and the difference in sign arises because $G_{-t}\mu$ satisfies the Hamiltonian equations with reversed time. The picture of the motion in which the time dependence of the states is determined by (5), while the observables do not depend on time, is called the Liouville picture:

$$\frac{df}{dt} = 0, \qquad \frac{d\rho_t}{dt} = -\{H, \rho_t\}.$$

The equation (5) is called Liouville's equation. It is obvious from the way the Liouville picture was introduced that

$$\langle f_t \,|\, \omega \rangle = \langle f \,|\, \omega_t \rangle.$$

This formula expresses the equivalence of the two pictures of motion. The mean values of the observables depend on time in the same way, and the difference between the pictures lies only in the way this dependence is described. We remark that it is common in statistical physics to use the Liouville picture.

§ 4. Physical bases of quantum mechanics

Quantum mechanics is the mechanics of the microworld. The phenomena it studies lie mainly beyond the limits of our perception, and therefore we should not be surprised by the seemingly paradoxical nature of the laws governing these phenomena.

It has not been possible to formulate the basic laws of quantum mechanics as a logical consequence of the results of some collection of fundamental physical experiments. In other words, there is so far no known formulation of quantum mechanics that is based on a system of axioms confirmed by experiment. Moreover, some of the basic statements of quantum mechanics are in principle not amenable to experimental verification. Our confidence in the validity of quantum mechanics is based on the fact that all the physical results of the theory agree with experiment. Thus, only consequences of the basic tenets of quantum mechanics can be verified by experiment, and not its basic laws. The main difficulties arising upon an initial study of quantum mechanics are apparently connected with these circumstances.

The creators of quantum mechanics were faced with difficulties of the same nature, though certainly much more formidable. Experiments most definitely pointed to the existence of peculiar quantum laws in the microworld, but gave no clue about the form of quantum theory. This can explain the truly dramatic history of the creation of quantum mechanics and, in particular, the fact that its original formulations bore a purely prescriptive character. They contained certain rules making it possible to compute experimentally measurable quantities, but a physical interpretation of the theory appeared only after a mathematical formalism of it had largely been created.

In this course we do not follow the historical path in the construction of quantum mechanics. We very briefly describe certain physical phenomena for which attempts to explain them on the basis of classical physics led to insurmountable difficulties. We then try to clarify what features of the scheme of classical mechanics described in the preceding sections should be preserved in the mechanics of the microworld and what can and must be rejected. We shall see that the rejection of only one assertion of classical mechanics, namely, the assertion that observables are functions on the phase space, makes it possible to construct a scheme of mechanics describing systems with behavior essentially different from the classical. Finally, in the following sections we shall see that the theory constructed is more general than classical mechanics, and contains the latter as a limiting case.

Historically, the first quantum hypothesis was proposed by Planck in 1900 in connection with the theory of equilibrium radiation. He succeeded in getting a formula in agreement with experiment for the spectral distribution of the energy of thermal radiation, conjecturing that electromagnetic radiation is emitted and absorbed in discrete portions, or quanta, whose energy is proportional to the frequency of the radiation:

$$(1) \qquad\qquad E = h\omega,$$

where $\omega = 2\pi\nu$ with ν the frequency of oscillations in the light wave, and where $h = 1.05 \times 10^{-27}$ erg-sec is Planck's constant.[7]

Planck's hypothesis about light quanta led Einstein to give an extraordinarily simple explanation for the photoelectric effect (1905). The photoelectric phenomenon consists in the fact that electrons are emitted from the surface of a metal under the action of a beam of light. The basic problem in the theory of the photoelectric effect is to find the dependence of the energy of the emitted electrons on the characteristics of the light beam. Let V be the work required to remove an electron from the metal (the work function). Then the law of conservation of energy leads to the relation

$$h\omega = V + T,$$

where T is the kinetic energy of the emitted electron. We see that this energy is linearly dependent on the frequency and is independent of the intensity of the light beam. Moreover, for a frequency $\omega < V/h$ (the red limit of the photoelectric effect), the photoelectric phenomenon becomes impossible since $T \geq 0$. These conclusions, based on the hypothesis of light quanta, are in complete agreement with experiment. At the same time, according to the classical theory, the energy of the emitted electrons should depend on the intensity of the light waves, which contradicts the experimental results.

Einstein supplemented the idea of light quanta by introducing the momentum of a light quantum by the formula

$$(2) \qquad\qquad \mathbf{p} = h\mathbf{k}.$$

---

[7]In the older literature this formula is often written in the form $E = h\nu$, where the constant h in the latter formula obviously differs from the h in (1) by the factor 2π.

Here **k** is the so-called wave vector, which has the direction of propagation of the light waves. The length k of this vector is connected with the wavelength λ, the frequency ω, and the velocity c of light by the relations

$$(3) \qquad k = \frac{2\pi}{\lambda} = \frac{\omega}{c} \, .$$

For light quanta we have the formula

$$E = pc,$$

which is a special case of the relativity theory formula

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

for a particle with rest mass $m = 0$.

We remark that the historically first quantum hypotheses involved the laws of emission and absorption of light waves, that is, electrodynamics, and not mechanics. However, it soon became clear that discreteness of the values of a number of physical quantities was typical not only for electromagnetic radiation but also for atomic systems. The experiments of Franck and Hertz (1913) showed that when electrons collide with atoms, the energy of the electrons changes in discrete portions. The results of these experiments can be explained by the assumption that the energy of the atoms can have only definite discrete values. Later experiments of Stern and Gerlach in 1922 showed that the projection of the angular momentum of atomic systems on a certain direction has an analogous property. It is now well known that the discreteness of the values of a number of observables, though typical, is not a necessary feature of systems in the microworld. For example, the energy of an electron in a hydrogen atom has discrete values, but the energy of a freely moving electron can take arbitrary positive values. The mathematical apparatus of quantum mechanics had to be adapted to the description of observables taking both discrete and continuous values.

In 1911 Rutherford discovered the atomic nucleus and proposed a planetary model of the atom (his experiments on scattering of α particles on samples of various elements showed that an atom has a positively charged nucleus with charge Ze, where Z is the number of the element in the Mendeleev periodic table and e is the charge of an

electron, and that the size of the nucleus does not exceed 10^{-12} cm, while the atom itself has linear size of order 10^{-8} cm). The planetary model contradicts the basic tenets of classical electrodynamics. Indeed, when moving around the nucleus in classical orbits, the electrons, like all charged particles that are accelerating, must radiate electromagnetic waves. Thus, they must be losing their energy and must eventually fall into the nucleus. Therefore, such an atom cannot be stable, and this, of course, does not correspond to reality. One of the main problems of quantum mechanics is to account for the stability and to describe the structure of atoms and molecules as systems consisting of positively charged nuclei and electrons.

The phenomenon of diffraction of microparticles is completely surprising from the point of view of classical mechanics. This phenomenon was predicted in 1924 by de Broglie, who suggested that to a freely moving particle with momentum $\mathbf{p}$ and energy E there corresponds (in some sense) a wave with wave vector $\mathbf{k}$ and frequency ω, where

$$\mathbf{p} = h\mathbf{k}, \qquad E = h\omega;$$

that is, the relations (1) and (2) are valid not only for light quanta but also for particles. A physical interpretation of de Broglie waves was later given by Born, but we shall not discuss it for the present. If to a moving particle there corresponds a wave, then, regardless of the precise meaning of these words, it is natural to expect that this implies the existence of diffraction phenomena for particles. Diffraction of electrons was first observed in experiments of Davisson and Germer in 1927. Diffraction phenomena were subsequently observed also for other particles.

We show that diffraction phenomena are incompatible with classical ideas about the motion of particles along trajectories. It is most convenient to argue using the example of a thought experiment concerning the diffraction of a beam of electrons directed at two slits,[8] the scheme of which is pictured in Figure 1. Suppose that the electrons

---

[8]Such an experiment is a thought experiment, so the wavelength of the electrons at energies convenient for diffraction experiments does not exceed 10^{-7} cm, and the distance between the slits must be of the same order. In real experiments diffraction is observed on crystals, which are like natural diffraction lattices.
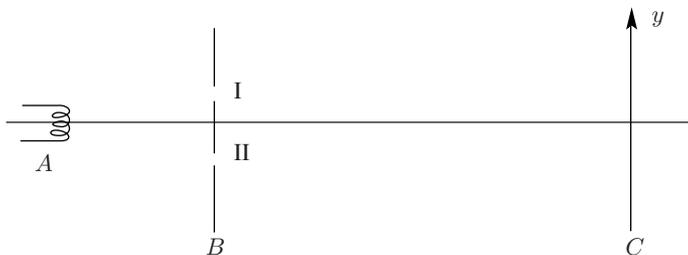
Figure 1

from the source A move toward the screen B and, passing through the slits I and II, strike the screen C.

We are interested in the distribution of electrons hitting the screen C with respect to the coordinate y. The diffraction phenomena on one and two slits have been thoroughly studied, and we can assert that the electron distribution $\rho(y)$ has the form a pictured in Figure 2 if only the first slit is open, the form b (Figure 2) if only the second slit is open, and the form c occurs when both slits are open. If we assume that each electron moves along a definite classical trajectory, then the electrons hitting the screen C can be split into two groups, depending on the slit through which they passed. For electrons in the first group it is completely irrelevant whether the second slit was open or not, and thus their distribution on the screen should be represented by the
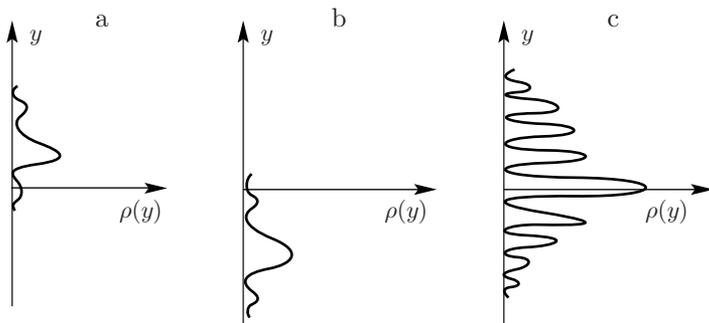

Figure 2

curve a; similarly, the electrons of the second group should have the distribution b. Therefore, in the case when both slits are open, the screen should show the distribution that is the sum of the distributions a and b. Such a sum of distributions does not have anything in common with the interference pattern in c. This contradiction indicates that under the conditions of the experiment described it is not possible to divide the electrons into groups according to the test of which slit they went through. Hence, we have to reject the concept of trajectory.

The question arises at once as to whether one can set up an experiment to determine the slit through which an electron has passed. Of course, such a formulation of the experiment is possible; for this it suffices to put a source of light between the screens B and C and observe the scattering of the light quanta by the electrons. In order to attain sufficient resolution, we have to use quanta with wavelength of order not exceeding the distance between the slits, that is, with sufficiently large energy and momentum. Observing the quanta scattered by the electrons, we are indeed able to determine the slit through which an electron passed. However, the interaction of the quanta with the electrons causes an uncontrollable change in their momenta, and consequently the distribution of the electrons hitting the screen must change. Thus, we arrive at the conclusion that we can answer the question as to which slit the electron passed through only at the cost of changing both the conditions and the final result of the experiment.

In this example we encounter the following general peculiarity in the behavior of quantum systems. The experimenter does not have the possibility of following the course of the experiment, since to do so would lead to a change in the final result. This peculiarity of quantum behavior is closely related to peculiarities of measurements in the microworld. Every measurement is possible only through an interaction of the system with the measuring device. This interaction leads to a perturbation of the motion of the system. In classical physics one always assumes that this perturbation can be made arbitrarily small, as is the case for the duration of the measurement. Therefore, the

simultaneous measurement of any number of observables is always possible.

A detailed analysis (which can be found in many quantum mechanics textbooks) of the process of measuring certain observables for microsystems shows that an increase in the precision of a measurement of observables leads to a greater effect on the system, and the measurement introduces uncontrollable changes in the numerical values of some of the other observables. This leads to the fact that a simultaneous precise measurement of certain observables becomes impossible in principle. For example, if one uses the scattering of light quanta to measure the coordinates of a particle, then the error of the measurement has the order of the wavelength of the light: $\Delta x \sim \lambda$. It is possible to increase the accuracy of the measurement by choosing quanta with a smaller wavelength, but then with a greater momentum $p = 2\pi h/\lambda$. Here an uncontrollable change Δp of the order of the momentum of the quantum is introduced in the numerical values of the momentum of the particle. Therefore, the errors Δx and Δp in the measurements of the coordinate and the momentum are connected by the relation

$$\Delta x \, \Delta p \sim 2\pi h.$$

A more precise argument shows that this relation connects only a coordinate and the momentum projection with the same index. The relations connecting the theoretically possible accuracy of a simultaneous measurement of two observables are called the Heisenberg uncertainty relations. They will be obtained in a precise formulation in the sections to follow. Observables on which the uncertainty relations do not impose any restrictions are simultaneously measurable. We shall see that the Cartesian coordinates of a particle are simultaneously measurable, as are the projections of its momentum, but this is not so for a coordinate and a momentum projection with the same index, nor for two Cartesian projections of the angular momentum. In the construction of quantum mechanics we must remember the possibility of the existence of quantities that are not simultaneously measurable.

After our little physical preamble we now try to answer the question posed above: what features of classical mechanics should be kept, and which ones is it natural to reject in the construction of a mechanics of the microworld? The main concepts of classical mechanics were the concepts of an observable and a state. The problem of a physical theory is to predict results of experiments, and an experiment is always a measurement of some characteristic of the system or an observable under definite conditions which determine the state of the system. Therefore, the concepts of an observable and a state must appear in any physical theory. From the point of view of the experimenter, to determine an observable means to specify a way of measuring it. We denote observables by the symbols $a, b, c, \ldots$, and for the present we do not make any assumptions about their mathematical nature (we recall that in classical mechanics the observables are functions on the phase space). As before, we denote the set of observables by $\mathfrak{A}$.

It is reasonable to assume that the conditions of the experiment determine at least the probability distributions of the results of measurements of all the observables, and therefore it is reasonable to keep the definition in § 2 of a state. The states will be denoted by ω as before, the probability measure on the real axis corresponding to an observable a by $\omega_a(E)$, the distribution function of a in the state ω by $\omega_a(\lambda)$, and finally, the mean value of a in the state ω by $\langle \omega \,|\, a \rangle$.

The theory must contain a definition of a function of an observable. For the experimenter, the assertion that an observable b is a function of an observable a ($b = f(a)$) means that to measure b it suffices to measure a, and if the number a_0 is obtained as a result of measuring a, then the numerical value of the observable b is $b_0 = f(a_0)$. For the probability measures corresponding to a and $f(a)$, we have

$$(4) \qquad \omega_{f(a)}(E) = \omega_a(f^{-1}(E))$$

for any states ω.

We note that all possible functions of a single observable a are simultaneously measurable, since to measure these observables it suffices to measure a. Below we shall see that in quantum mechanics this

example exhausts the cases of simultaneous measurement of observables; that is, if the observables $b_1, b_2, \ldots$ are simultaneously measurable, then there exist an observable a and functions $f_1, f_2, \ldots$ such that $b_1 = f_1(a), b_2 = f_2(a), \ldots$.

The set of functions $f(a)$ of an observable a obviously includes $f(a) = \lambda a$ and $f(a) = \text{const}$, where λ is a real number. The existence of the first of these functions shows that observables can be multiplied by real numbers. The assertion that an observable is a constant means that its numerical value in any state coincides with this constant.

We now try to make clear what meaning can be assigned to a sum $a + b$ and product ab of two observables. These operations would be defined if we had a definition of a function $f(a, b)$ of two variables. However, there arise fundamental difficulties here connected with the possibility of observables that are not simultaneously measurable. If a and b are simultaneously measurable, then the definition of $f(a, b)$ is completely analogous to the definition of $f(a)$. To measure the observable $f(a, b)$, it suffices to measure the observables a and b leading to the numerical value $f(a_0, b_0)$, where a_0 and b_0 are the numerical values of the observables a and b, respectively. For the case of observables a and b that are not simultaneously measurable, there is no reasonable definition of the function $f(a, b)$. This circumstance forces us to reject the assumption that observables are functions $f(q, p)$ on the phase space, since we have a physical basis for regarding q and p as not simultaneously measurable, and we shall have to look for observables among mathematical objects of a different nature.

We see that it is possible to define a sum $a + b$ and a product ab using the concept of a function of two observables only in the case when they are simultaneously measurable. However, another approach is possible for introducing a sum in the general case. We know that all information about states and observables is obtained by measurements; therefore, it is reasonable to assume that there are sufficiently many states to distinguish observables, and similarly, that there are sufficiently many observables to distinguish states.

More precisely, we assume that if

$$\langle a \,|\, \omega \rangle = \langle b \,|\, \omega \rangle$$

for all states ω, then the observables a and b coincide,[9] and if

$$\langle a \,|\, \omega_1 \rangle = \langle a \,|\, \omega_2 \rangle$$

for all observables a, then the states ω_1 and ω_2 coincide.

The first of these assumptions makes it possible to define the sum $a + b$ as the observable such that

(5) $$\langle a + b \,|\, \omega \rangle = \langle a \,|\, \omega \rangle + \langle b \,|\, \omega \rangle$$

for any state ω. We remark at once that this equality is an expression of a well-known theorem in probability theory about the mean value of a sum only in the case when the observables a and b have a common distribution function. Such a common distribution function can exist (and in quantum mechanics does exist) only for simultaneously measurable quantities. In this case the definition (5) of a sum coincides with the earlier definition. An analogous definition of a product is not possible, since the mean of a product is not equal to the product of the means, not even for simultaneously measurable observables.

The definition (5) of the sum does not contain any indication of a way to measure the observable $a + b$ involving known ways of measuring the observables a and b, and is in this sense implicit.

To give an idea of how much the concept of a sum of observables can differ from the usual concept of a sum of random variables, we present an example of an observable that will be studied in detail in what follows. Let

$$H = \frac{P^2}{2} + \frac{\omega^2 Q^2}{2} \,.$$

The observable H (the energy of a one-dimensional harmonic oscillator) is the sum of two observables proportional to the squares of the momentum and the coordinate. We shall see that these last observables can take any nonnegative numerical values, while the values of the observable H must coincide with the numbers $E_n = (n + 1/2)\omega$, where $n = 0, 1, 2, \ldots$; that is, the observable H with discrete numerical values is a sum of observables with continuous values.

Thus, two operations are defined on the set $\mathfrak{A}$ of observables: multiplication by real numbers and addition, and $\mathfrak{A}$ thereby becomes

---

[9]This assumption allows us to regard an observable as specified if a real number is associated with each state.

a linear space. Since real functions are defined on $\mathfrak{A}$, and in particular the square of an observable, there arises a natural definition of a product of observables:

$$(6) \qquad a \circ b = \frac{(a+b)^2 - (a-b)^2}{4}.$$

We note that the product $a \circ b$ is commutative, but it is not associative in general. The introduction of the product $a \circ b$ turns the set $\mathfrak{A}$ of observables into a real commutative algebra.

Recall that the algebra of observables in classical mechanics also contained a Lie operation: the Poisson bracket $\{f, g\}$. This operation appeared in connection with the dynamics of the system. With the introduction of such an operation each observable H generates a family of automorphisms of the algebra of observables:

$$U_t \colon \mathfrak{A} \to \mathfrak{A},$$

where $U_t f = f_t$; and f_t satisfies the equation

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

and the initial condition

$$f_t|_{t=0} = f.$$

We recall that the mapping U_t is an automorphism in view of the fact that the Poisson bracket has the properties of a Lie operation. The fact that the observables in classical mechanics are functions on phase space does not play a role here. We assume that the algebra of observables in quantum mechanics also has a Lie operation; that is, associated with each pair of observables a, b is an observable $\{a, b\}$ with the properties

$$\{a, b\} = -\{b, a\},$$
$$\{\lambda a + b, c\} = \lambda\{a, c\} + \{b, c\},$$
$$\{a, b \circ c\} = \{a, b\} \circ c + b \circ \{a, c\},$$
$$\{a, \{b, c\}\} + \{b, \{c, a\}\} + \{c, \{a, b\}\} = 0.$$

Moreover, we assume that the connection of the Lie operation with the dynamics in quantum mechanics is the same as in classical mechanics. It is difficult to imagine a simpler and more beautiful way of describing the dynamics. Moreover, the same type of description of the dynamics

in classical and quantum mechanics allows us to hope that we can construct a theory that contains classical mechanics as a limiting case.

In fact, all our assumptions reduce to saying that in the construction of quantum mechanics it is reasonable to preserve the structure of the algebra of observables in classical mechanics, but to reject the realization of this algebra as functions on the phase space since we admit the existence of observables that are not simultaneously measurable.

Our immediate problem is to see that there is a realization of the algebra of observables that is different from the realization in classical mechanics. In the next section we present an example of such a realization, constructing a finite-dimensional model of quantum mechanics. In this model the algebra $\mathfrak{A}$ of observables is the algebra of self-adjoint operators on the n-dimensional complex space $\mathbf{C}^n$. In studying this simplified model, we can follow the basic features of quantum theory. At the same time, after giving a physical interpretation of the model constructed, we shall see that it is too poor to correspond to reality. Therefore, the finite-dimensional model cannot be regarded as a definitive variant of quantum mechanics. However, it will seem very natural to improve this model by replacing $\mathbf{C}^n$ by a complex Hilbert space.

§ 5. A finite-dimensional model of quantum mechanics

We show that the algebra $\mathfrak{A}$ of observables can be realized as the algebra of self-adjoint operators on the finite-dimensional complex space $\mathbf{C}^n$.

Vectors in $\mathbf{C}^n$ will be denoted by the Greek letters $\xi, \eta, \varphi, \psi, \ldots$. Let us recall the basic properties of a scalar product:

$$
\begin{aligned}
&1) \quad (\xi, \psi) = \overline{(\psi, \xi)}, \\
(1) \quad &2) \quad (\xi + \lambda\eta, \psi) = (\xi, \psi) + \lambda(\eta, \psi), \\
&3) \quad (\xi, \xi) > 0 \quad \text{if } \xi \neq 0.
\end{aligned}
$$

Here λ is a complex number.