

UNIVERSAL
LIBRARY

OU_158027

UNIVERSAL
LIBRARY

OUP—881—5-8-74—15,000

OSMANIA UNIVERSITY LIBRARY

Call No. 539
I 93C

Accession No. G 18050

Author

Ivanov, B.

Title

contemporary physics.

This book should be returned on or before the date last marked below.

B. I V A N O V

**CONTEMPORARY
P H Y S I C S**

(A Review of Basic Principles)

**PEACE PUBLISHERS
M O S C O W**

TRANSLATED FROM THE RUSSIAN BY V. T A L M Y

Б. Иванов

Новая физика

На английском языке

CONTENTS

Author's Preface	5
Introduction	7
<i>Chapter 1.</i> Classical Mechanics	15
<i>Chapter 2.</i> Relativistic Mechanics	29
<i>Chapter 3.</i> Electromagnetic Field Theory	43
<i>Chapter 4.</i> Gravitational Field Theory	54
<i>Chapter 5.</i> Quantum Mechanics (Nonrelativistic Theory)	71
<i>Chapter 6.</i> Relativistic Quantum Theory	94
<i>Chapter 7.</i> Physics of Elementary Particles	102
<i>Chapter 8.</i> Nuclear Physics	107
<i>Chapter 9.</i> Statistical Physics	108
<i>Chapter 10.</i> Physical Kinetics	125
Index	128

AUTHOR'S PREFACE

Ours is a time of unprecedented progress of the natural sciences. Especially great are advances in physics, the fundamental science of nature, which embraces the most ambitious fields of human learning.

In the second half of the twentieth century, physics has developed into a comprehensive, universal study of nature. Its methods and conceptions are increasingly invading all the natural sciences, thereby changing and throwing open new broad vistas before them.

(Physics investigates the simplest but most fundamental "structures of the world". It studies the simplest but deepest-lying links in the "world harmony". This makes physical concepts extremely abstract and difficult to visualise.)

It is no easy task for man to study nature. His trouble is that, by virtue of his mutual relations with nature, he is forced to proceed in his investigations from the particular to the general. From this stem many of his frustrations in cognising the world, for very often a realisation of generalised concepts requires a drastic reappraisal of well-established and apparently self-evident notions.

(Physical concepts are of a quantitative nature and are therefore of necessity bound together by mathematical formalism. But then, mathem. is an excellent tool for coping with abstract notion: all kinds and its possibilities in this respect are virtually unlimited. The mathematical apparatus of physics enables the researcher to delve freely into regions where imagination alone is powerless.)

Contemporary physics comprises a beautiful, well-balanced system of principles and concepts whose depth and uniformity allow for a thorough investigation of both

the elementary "structures of the world" and the world as a whole.

Physics constitutes the mainstay of many new and future industries and technologies. It has had a tremendous impact on the advance of other natural sciences and given rise to wonderful new scientific spheres. Moreover, the very investigation of the systems of physical notions is a source of great intellectual and aesthetic satisfaction.

Mathematical methods, as mentioned above, play a tremendous part in physical researches. But, in the words of Paul Dirac, mathematics remains but a tool. It is physical ideas that determine the progress of physics. Sound physical ideas (often they are of a highly mathematical nature) are, as a rule, conceived in the analysis of experimental findings.

The whole science of physics rests on a comprehensive system of physical ideas, an understanding of which is essential for any understanding at all.

The available volume of physical data is virtually unlimited. A wonderful development in contemporary physics is the change that has taken place in its inner logical structure, a change which has served to simplify, systematise and provide a uniform method of approach to the sum-total of physical phenomena, from the properties of elementary particles to the structure of the universe.

Physicists in the U.S.S.R. and in other countries know Academician L. D. Landau's and Professor Y. M. Lifshitz's nine-volume *Course of Theoretical Physics*. The present book is an attempt at a popular outline of that fundamental work. In preparing this digest I adopted a purely qualitative approach to the subject matter, and accordingly some new points had to be introduced into the discourse. Small though this volume is, I have incorporated several excerpts from the *Course* which seemed to me especially vivid and important (this refers especially to the chapters on Relativistic Mechanics and Statistical Physics). Insofar as they are not quite verbatim quotations, I have omitted specific references to the volume and pages of the original work.

Introduction

All physical processes take place in *space* and in *time*. Every physical law in every physical domain contains, explicitly or otherwise, space-time relationships, viz., *length (distance)* and *time interval (duration)*.

We know from experience that space and time possess certain *symmetry properties*, which impose restrictions on physical processes. Among these properties are space and time homogeneity.

Because of *time homogeneity* a physical phenomenon is always the same, if the conditions are the same, whenever it is observed. Two millennia have passed since Archimedes discovered the laws of buoyancy, yet today they can be readily developed, provided the conditions of observation are the same. The physical equivalence of different instants of time—the homogeneity of time—imposes certain restrictions on physical phenomena which find expression in the *law of conservation of energy*.

Because of *space homogeneity* a physical phenomenon is always the same, if the conditions are the same, wherever it is observed: the same physical experiment staged in Moscow or New York yields the same results. The physical equivalence of different points of space—the homogeneity of space—imposes certain restrictions on physical phenomena which find expression in the *law of conservation of (linear) momentum*.

If these symmetry properties of space and time, which seem so obvious, did not exist, it would be useless to conduct scientific research and attempt to cognise the world. Imagine what would happen if there were no spatial homogeneity: one set of physical laws would hold in Moscow, another in London, yet another in New York! Non-

homogeneity of time would make it impossible for people to advance their knowledge. The buoyancy principle discovered the other day would no longer be valid today, and a new set of experiments would have to be carried out, which would have to be repeated again tomorrow.)

Insofar as the laws of conservation of energy and momentum follow from the very general symmetry properties of space and time the conservation laws are *universal*: they are equally valid in the physics of elementary particles and outer space, nuclear and solid-state physics. In fact, they are among the several most general laws which constitute the bedrock of contemporary physics.

(It is impossible to investigate physical phenomena without introducing a *frame of reference*, reference bodies or landmarks relative to which observations are carried out. Imagine a highway crossing a very flat and monotonous plain. A car travelling along it in the distance at first seems to be standing still. But when one finds a landmark—a telegraph pole or a tree (usually this is done subconsciously)—and continues to observe the car for a period of time one detects a change in the mutual positions of the car and the landmark. (The concept of frame of reference (coordinate system) is a fundamental one in physics.)

(A researcher embarking on physical investigations is free to choose any reference body he likes. It is natural to assume that the laws governing a given phenomenon do not depend on the choice of reference system. Experience, however, confirms this only with respect to a very special class of reference systems which move at a uniform velocity relative to each other. This is what is known as the *relativity principle*.)

A space-time description of phenomena in terms of length and time interval is possible only when a reference system has been selected. (The facts of life tell us that in studying physical phenomena lengths and time intervals *do not change* in passing over from one reference system to another. A navigator plotting a forthcoming flight on a map uses a ruler and a clock. He continues to use them quite confidently in flight without fearing that the length of the ruler or the rhythm of the clock might have changed, and he is quite justified in this.)

One of the greatest contributions to twentieth-century physics was the postulate of the existence of a *limiting velocity* of propagation of anything whatsoever in nature, equal to the velocity of light in vacuum.

(Experience indicates that as long as the velocities of investigated bodies are small compared with the limiting velocity, time intervals and lengths do not change in passing from one reference system to another. The science which treats of the motion of bodies in these conditions is called *classical mechanics*.)

(The velocities of bodies on earth and in the solar system (artificial earth satellites included) are incomparably smaller than the limiting velocity, and their motions are governed by the laws of classical mechanics to a high degree of accuracy.)

(When speeds are comparable with the limiting velocity, however, the properties of motion change: lengths and time intervals begin to *depend* on velocity. The mechanics based on these concepts is called *relativistic mechanics*.)

Relativistic mechanics is, of course, more general, and in the special case of small velocities it turns into classical mechanics. Relativity theory does not reject classical mechanics: it brings in the necessary adjustments for high velocities.

Velocities comparable with the speed of light are found in the study of fast atomic particles. Experiments reveal that their motion is in fact subject to the laws of relativistic mechanics.

(The physical entities of the real world include bodies and *fields*. A field can be described consistently *only* in terms of relativistic concepts. Unlike bodies, fields cannot generally be used as reference frames. In fact, for the first time since man began to investigate nature he is confronted with a physical entity which does not affect any of his senses directly.)

The properties of fields are studied from the behaviour of bodies in them. Two fields are said to be physically identical if their action on the same body is the same. Thus, by studying the behaviour of "sensible", ponderable entities—bodies—man has been able to cognise the properties of intangible, imponderable entities, fields,

What better example of man's amazing aptitude for cognition?

In ordinary conditions we deal with two types of field, electromagnetic and gravitational. The properties of the electromagnetic field are more or less familiar to people who have taken a secondary-school course in physics. The properties of the *gravitational field*, however, are less familiar.

It should be noted, (first of all, that gravitational fields are of importance only on the cosmic scale. The ratio of the intensity of gravitational interactions between elementary particles to their electromagnetic interactions is of the order of 10^{-40} .)

Gravitational fields occupy a unique position in nature. For one, every physical thing produces a gravitational field: particles, bodies and fields, even gravitational fields themselves. Secondly, gravitational fields shape and determine, as it were, the geometrical structure of space and time.

The properties of gravitational fields viewed on the vast scale of the observable universe seem to indicate that the spatial structure of the universe is changing with time in the direction of expansion. In fact, the spectacular effect of the expanding universe has been observed.

The field of physical research ranges from the boundless regions of the macrocosmos to the minute spaces of the microcosmos. When the observable phenomena involve very small masses (particles) and take place in very small regions of space, with dimensions of the order of 10^{-8} cm, these are *atomic phenomena*.

In investigating atomic phenomena man has once again displayed the might of his intellect. The atomic world is invisible, inaudible and intangible to our senses; atomic dimensions are so tiny as to defy the imagination, yet man has developed a picture of the atom and fathomed its laws. How was this achieved?

Physical science deals only with things that are directly or indirectly *observable*. Any observation of a physical thing presumes *interaction* with its environment. A physical entity displays its properties only in interactions with something external with respect to it. In order to obtain a comprehensive description of the properties of such an

entity it must be placed in different conditions. Not infrequently, especially where atomic phenomena are concerned, an object may display apparently contradictory, mutually exclusive properties. In fact, in atomic phenomena such seemingly paradoxical combinations of incompatible properties are rather the rule than the exception.

(The foregoing can be illustrated by the following analogy. A stranger is always a sort of "thing in itself", an unknown entity, until one comes into contact with him. In order to get to know the man, his disposition and other properties, one must observe him in different circumstances and situations. One often finds that the behaviour ("properties") of a person ("object") in various circumstances is so different that one finds it hard to believe that it is the same person. This is just about what happened when the properties of the atom were investigated.)

(An observation of a microscopic object is an interaction which inevitably involves a disturbance, or perturbation, of the investigated system. If the disturbance is small and can be neglected, the observed phenomenon is said to be of a macroscopic nature. If, on the other hand, it is impossible to neglect some limiting perturbation—associated with the existence of a *quantum of action*—the phenomenon is said to belong to the microscopic sphere.)

The discovery that there exists a *universal minimum perturbation*, which it is impossible to avoid no matter how refined the observations, was a tremendous triumph of the human intellect. A submicroscopic system (known also as a *quantum object*) cannot be observed without causing appreciable perturbations on it. A consequence of this is that quantum objects, such as atoms and molecules, are subject to *uncertainty relationships*.

(In classical mechanics, when a certain quantity characterising similar objects is measured in identical conditions, the results are, quite naturally, the same. When quantum objects of the same kind are measured in the same conditions (compatible with their submicroscopic physical nature) the results are specific, but different. The law of nature in such cases is that the various results are always obtained in the same proportion to the total number of measurements.)

Quantum phenomena tax the imagination much more than classical processes. Depending on the conditions of observation, quantum systems may behave either as waves or as classical particles. (In classical mechanics a moving body is characterised by a specific position in space and velocity at any given time. With quantum objects, however, it is either the position or the velocity that can be specified at a given instant, but not both at the same time. Try to visualise such a situation!)

(The very small regions of space that have been studied so far display all the basic symmetry properties of space and time, in particular, homogeneity. Hence, the laws of conservation of momentum and energy are valid for quantum systems.)

(Homogeneity is one of the most important symmetry properties of space but not the only one. Thus, space possesses symmetry between right and left. The mirror image of space is wholly equivalent to the original. As we have seen, the symmetry properties of space and time are associated with certain conservation laws. In nonquantum physics mirror symmetry introduces no new conservation laws. In quantum physics, however, the *law of conservation of parity* comes into effect. Parity is a concept which characterises the state of quantum objects.)

(Our daily experience tells us that time flows in one direction only. It is impossible to bring back one's youth, yet such things were permissible in the framework of non-quantum physics. Only quantum theory offers a sound explanation of the *anisotropy* of time by virtue of which the direct and reverse passages of time are not equivalent.

(In very small regions of space, fields, like particles, are subject to the laws of quantum. The *quantised field*, like all other fields, is a real relativistic entity. The particle and the field constitute the two fundamental concepts of physics. The synthesis of relativistic and quantum ideas has resulted in a merger of the two and contributed to the remarkable discovery of so-called *antiparticles*, which were first predicted theoretically and whose existence was later confirmed by laboratory experiment.)

One of the most wonderful properties of particle-antiparticle pairs is their ability for production and annihilation. Such a pair, which possesses mass, is capable of

transforming into particles possessing no mass and, conversely, particles with no mass may produce particle-antiparticle pairs having mass.

Quantum phenomena taking place in regions of space with dimensions of the order of 10^{-8} to 10^{-13} cm are beautifully explained by contemporary physics. In smaller regions of space, however, physicists feel less confident of their definitions. Investigation of the phenomena characteristic of such regions has practically only just begun. It constitutes the subject of the *physics of elementary particles*.

In spite of the difficulties of staging laboratory experiments and their prohibitive cost, quite a lot of empirical data have been accumulated in this sphere. A very nice and promising classification of the known elementary particles has been developed which has made possible the theoretical prediction of particles. Many of these predicted particles have now been discovered, and today some thirty or so particles are known.

(A classification of *interactions* of elementary particles has been elaborated. Remarkable internal symmetry properties and new types of conservation laws associated with them have been discovered. Certain types of particle interactions have been found to violate some of the known conservation laws. Thus, in so-called weak interactions parity is not conserved.)

The tremendous faculties of the human brain for speculative reasoning have been demonstrated time and again. Equally staggering is the craftsmanship of contemporary experimentalists. The neutrino, one of the most remarkable of the elementary particles, whose existence was predicted thirty years ago, has no mass, travels with the speed of light and practically does not interact with matter. A neutrino could travel through an iron sphere millions of times larger than the earth's orbit around the sun and experience *only one* interaction. Neutrinos travel freely through the earth, the solar system and whole galaxies. It would seem that there was no way whatsoever of intercepting neutrinos, and yet it has been done. The success of the physical experiment to detect the neutrino represents another triumph of the human mind. And still, many more mysteries lie in store. Already now some won-

derful links are being suggested between the physics of elementary particles and the physics of outer space.

Some elementary particles, called nucleons, are capable of forming stable systems: *atomic nuclei*. There are not many different types of nuclei in nature. The spatial domain of the atomic nucleus is of dimension $\sim 10^{-13}$ cm. Nucleons form into nuclei in conditions existing in the interior of stars, where the "stellar state" of matter prevails.

In ordinary terrestrial conditions matter is made up of atoms and molecules. *Macroscopic bodies* are assemblies of countless numbers of atoms and molecules. All the properties of macroscopic bodies — elasticity and flowability, thermal and electric properties, magnetic, chemical and optical properties, the properties of solid, liquid, and gaseous states — are explained by the properties of atoms and the bonds between them. There can be no doubt that the *phenomena of life*, too, can be explained within the framework of known physical concepts.

It should be clear from what has been said that physical concepts and ideas are the simplest, the most general and the most profound, with them it is possible to study the whole of nature according to a unified method of research.

* * *

The purpose of physics is to establish physical laws, that is, to determine the relationships that exist between the physical quantities characterising a phenomenon.

Physicists use the rules and methods of mathematics to formulate their conclusions. Nevertheless, physics differs radically from mathematics in that it cannot be divorced from experiment. Accordingly, it is essentially made up of two sciences, experimental physics and theoretical physics.

In the course of its evolution physics has developed some more or less general principles of research, e. g., (1) science deals only with observable things; (2) every object displays its properties only in interactions with something external with respect to it (the so-called external classical conditions, which will be discussed later on).

Chapter 1. Classical Mechanics

As pointed out before, every physical law contains space-time relationships in terms of *length* and *time*.

Length

Length is measured by means of measuring rods. A fundamental property of measuring rods is that if any two rods have once been found to coincide when superimposed they will remain equal forever and always, that is to say, they will always coincide whenever they are superimposed. Only rigid bodies can be used as measuring rods.

Time interval

Time intervals are measured by clocks; any system in which a repeatable process takes place can be used for this. A fundamental concept of classical mechanics concerning the dimensions of bodies and intervals of time is their *absolute nature*: a measuring rod, if it is a good one, is always of the same length regardless of whether it is moving with respect to an observer or not; two clocks once set by one another will show the same time, if they have the same rhythm, irrespective of how they may be moving.

Space and time as physical entities

Space and time are physical entities like any other, but their importance by far surpasses that of any other physical entity. To study the properties of space and time one must observe the motion of bodies in them; investigation of the manner in which this motion proceeds offers a clue to the properties of space and time.

Three-dimensional space

The statement of one dimension in space is not enough to locate a point uniquely. From *experience* we know that three dimensions must be stated to define the position of a point in space relative to other points. The

numerical values of these dimensions are called the coordinates of the point. The fundamental property of three-dimensional space is that three coordinates are sufficient to locate precisely a point relative to some rigid body, the *frame of reference*.

**Space
homogeneity
and isotropy**

In further investigating the motion of bodies one finds that the properties of space are identical at different points, and that at each point they are the same in all directions. In other words, space is both *homogeneous* and *isotropic*. These properties of space find expression in the laws of conservation of linear momentum (or simply momentum) and angular momentum (or moment of momentum).

Time homogeneity The same experiments also reveal that different moments of time are equivalent, as far as their physical properties¹ are concerned, i. e., time is *homogeneous*. An expression of this is the law of conservation of energy.

The whole progress of physics has served to confirm that the energy and momentum conservation laws are fundamental laws of nature. This is seen in the fact that the vast number of known physical laws can be developed from a very few general relationships, which include the conservation laws.

Euclidean space Another property of space is its *flatness*. This is known as Euclidean space, because it conforms to Euclidean geometry.¹

The *classical* notions of space and time mentioned here were established experimentally and are valid in a vast domain of physical phenomena.

The particle We shall now examine the properties of motion of the simplest physical entity, a *material point* or *particle*. A material point is a body whose dimensions can be neglected in describing its motion. The location of a material point, as mentioned before, is determined by three spatial dimensions, its coordinates.

¹ The question of the geometry of real space is a purely physical one which can be answered only by experiment.

**Frame
of reference.
Coordinates
of a particle**

Three coordinates uniquely define the position of a point with respect to a body of reference. If a Cartesian coordinate system is attached to this reference body, the location of the particle

in space can be characterised by its radius vector, the components of which parallel to the axes of the coordinate system are equal to the Cartesian coordinates x, y, z of the point.¹

Classical motion According to classical notions of the spatio-temporal properties of motion, a material point occupies a definite place in space at every instant and it possesses definite coordinates. When a material point moves its coordinates change. Accordingly, the radius vector r of a material point can be considered as a function of time: $r = r(t)$.

**Velocity
of a particle**

The *velocity* of a material point is

$$v = \frac{dr}{dt},$$

i. e., the derivative of the radius vector with respect to time.

Acceleration is defined as

Acceleration

$$w = \frac{dv}{dt}.$$

**Degrees
of freedom**

Mechanics studies the motion of systems of material points. The number of independent coordinates required to describe the motion of a mechanical system is called *the number of degrees of freedom* of that system. A material point, evidently, has three degrees of freedom, and a system consisting of n material points has $3n$ degrees of freedom.

**State
of a classical
particle**

If the state of a mechanical system at any given instant is known, its state at any other moment of time can be determined. The state of a system, it is found, is completely defined by stipulating the coordinates and velocities. Acceleration cannot be stated arbitrarily as it is a function of the position coordinates and velocity.

¹ Other types of coordinate systems may also be used. They are usually chosen according to their convenience in describing motion.

The formulas that correlate acceleration, coordinates and velocity are called the *equations of motion* of a system.)

Classical mechanics

The fundamental problem of classical mechanics (that is, mechanics based on the classical notions of space, time and motion) is the study of the motion of any mechanical system by determining its position coordinates as functions of time from stated initial conditions, i. e., from the coordinates and velocities at some initial time.

Inertial frame of reference

As mentioned before, physical phenomena cannot be investigated without reference to a coordinate system. One is free to choose any of a countless number of reference systems moving in all manners with respect to one another. The manifestations of the laws of nature, however, may differ in different systems. If an arbitrary reference system is chosen one may find that the laws governing the simplest phenomena are involved indeed. The problem thus is to find a frame of reference in which all the laws of nature would appear in as simple a form as possible; such a reference system would evidently be best suited for describing physical phenomena.

In seeking such a frame of reference we proceed from the simplest type of motion, that of a material point so distant from all other bodies that any interaction between it and those bodies can be neglected. Such a material point is said to be in *free* motion.

Free motion

Let us take an arbitrary frame of reference and study its properties with the aid of a free material point. We assume that at the initial moment the point was at rest in that system. Generally speaking, then, the next instant the point will no longer be at rest, having begun to move in some direction. In this sense we may state that a frame of reference with arbitrary properties is neither homogeneous nor isotropic.

Two free moving bodies, however, may remain at rest relative to each other indefinitely. Therefore motion can be referred to a coordinate system rigidly attached to some free moving bodies. Such a frame is called an *inertial system*. In an inertial system all directions are physically equivalent and the properties of different points

in space are identical, i. e., it is mechanically homogeneous and isotropic.

The inertia law These properties of an inertial reference system lead to the conclusion that in it the free motion of a material particle takes place at uniform velocity. This statement is known as the *inertia law* (Newton's first law). In particular, as assumed before, if at some initial time the velocity of a material point is zero, it remains at rest indefinitely.

The relativity principle If we consider a reference frame moving with respect to an inertial system in some arbitrary manner, the former will not, in the most general case, be an inertial system. It does not follow from this, however, that there is only one inertial frame of reference in the world. It will be readily observed that there exist an infinite number of such systems, all of them being in uniform rectilinear motion with respect to each other.

The classical relativity principle To inertial frames of reference is applicable the *relativity principle*, which states that their physical properties are equivalent. Coupled with the concept of absolute time, this is known as *Galileo's relativity principle*. From the equivalence of all inertial frames of reference it follows that the equations of motion of any mechanical system do not change in passing from one inertial system to another. }

The Galilean transformations Let \mathbf{r} be a radius vector defining the position of a material point relative to an inertial reference system O at some moment of time t . Let the radius vector and time of the event be respectively \mathbf{r}' and t' in another inertial reference system O' moving with a velocity \mathbf{v}_0 relative to the unprimed system.

According to the classical notions of space and time, the formulas for going over from one set of coordinates and time measurement to another have the form

$$\left. \begin{aligned} \mathbf{r} &= \mathbf{r}' + \mathbf{v}_0 t, \\ t &= t'. \end{aligned} \right\} \quad (1)$$

The first equation expresses the absolute quality of spatial dimensions, the second, the absolute qua-

lity of time. These relations are known as the *Galilean transformations*.

**Composition
of velocities**

Differentiating (1) with respect to time, we obtain

$$\boldsymbol{v} = \boldsymbol{v}' + \boldsymbol{v}_0. \quad (2)$$

This simple formula defines the principle of composition of velocities: velocity \boldsymbol{v} relative to the unprimed reference system is compounded of velocity \boldsymbol{v}' relative to the primed system and the velocity \boldsymbol{v}_0 of system O' relative to system O .

Acceleration

Differentiating (2) with respect to time, and taking into account that \boldsymbol{v}_0 is constant, we obtain

$$\frac{d\boldsymbol{v}}{dt} = \frac{d\boldsymbol{v}'}{dt},$$

i. e., the acceleration is the same in all inertial frames of reference.

**Relative
and absolute**

Insofar as there exists no distinguished "absolute" reference system, the concept of absolute rest is devoid of meaning. If a body is at rest in one inertial system, it is moving with some uniform velocity relative to other systems, and there is no reason why one should be given preference before another. Similarly, the notion of absolute velocity is also meaningless; only the relative velocity of bodies with respect to one another has physical meaning. Absolute acceleration, on the other hand, is physically meaningful since, as we have found out, acceleration is the same in different inertial systems, and the difference between inertial and noninertial systems is of an absolute nature. In future, unless otherwise specified, we shall always be referring to inertial systems.

**The principle
of action
at a distance**

It follows from Galileo's relativity principle that interactions between bodies propagate through space instantaneously, i. e., if the state of one body is altered, it will immediately affect all other bodies interacting with it, no matter how far away they are. In fact, if interaction propagated with a finite velocity then,

from the rule of velocity composition (2), the velocity would be different in different frames of reference and it would be physically possible to distinguish between them, which contradicts the relativity principle. As the velocity of propagation of interactions is assumed to be infinite, classical (Newtonian) mechanics is said to be based on the principle of *action at a distance*.

**Systems
of particles**

Up till now we have been speaking of the properties of motion of a single free moving material point. Now consider a system of material points, assuming it to be so far away from any other bodies that its interactions with them can be neglected (they do not have to be taken into account in investigating the motion of the system). Such systems are termed *closed*.

**Conservation
laws**

In investigating the changes that take place with free moving bodies after they have interacted for some time, we find that, irrespective of the nature of the interaction, certain *conservation laws* hold. In other words, there are certain quantities characterising the state of bodies, the totality of which over all the interacting bodies is not affected by the interaction.

**Conservation
of momentum**

Consider a closed system of particles. Since, by virtue of spatial homogeneity, all configurations occupied by it as a whole in space are equivalent, we may assert that its properties will not change if it is moved parallel to itself to any distance.

A consequence of this circumstance is the conservation of a certain vector quantity characterising the system. Namely, there exists a vector characterising every material point, such that the sum of all the vectors over all the particles of a closed system is independent of time. It is called the vector of *linear momentum*. Momentum is related to velocity by a direct proportionality. The coefficient of proportionality, which is different for different material particles, is called *mass*.¹ *The law of conservation of the momentum of a system* is thus expressed by the

¹ This is its *physical definition*, on which experiments are based.

formula

$$m_1\mathbf{v}_1 + m_2\mathbf{v}_2 + \dots = \sum_{\alpha} m_{\alpha}\mathbf{v}_{\alpha} = \mathbf{P} = \text{const}, \quad (3)$$

where α is the number of particles.

Mass of a particle The law of conservation of momentum suggests a mode for correlating the masses of material points. Thus, for two colliding particles we can rewrite (3) as:

$$m_1\Delta\mathbf{v}_1 = -m_2\Delta\mathbf{v}_2,$$

where $\Delta\mathbf{v}_1$ and $\Delta\mathbf{v}_2$ represent the change in the velocities of the respective particles. Hence,

$$\Delta\mathbf{v}_1 = -\frac{m_2}{m_1}\Delta\mathbf{v}_2.$$

Knowing $\Delta\mathbf{v}_1$ and $\Delta\mathbf{v}_2$, it is possible to correlate the masses of both particles. Obviously the basic mass cannot be measured and it must be taken for unity. It can be chosen arbitrarily: choice carries no deep physical meaning.

Newton's first law The law of conservation of momentum in a closed system can be regarded as a generalisation of the *inertia law* (Newton's first law). In fact, for a free moving particle the momentum $\mathbf{p} = m\mathbf{v}$ remains constant. In the case of a system of material points interacting in any way the momentum of each particle is not constant, but the sum of the momenta of all the particles is conserved, and

$$\mathbf{P} = \sum_{\alpha} \mathbf{p}_{\alpha}.$$

Evidently, the quantity

$$\frac{d\mathbf{p}_{\alpha}}{dt} = \mathbf{F}_{\alpha}, \quad (4)$$

Force which expresses the change in the momentum of a particle in unit time, is a measure of the external action on that particle. This quantity is the *force* which the particles of a system exert on the particle α .

Potential function The interaction of material points is described in classical mechanics in terms of the *potential energy of interaction*

$$U = U(r_1, r_2, \dots, r_{\alpha}, \dots, r_n),$$

which is a function of the coordinates of the interacting particles. The type of potential function is determined in each case by the nature of the interactions. With distances between material points increasing to infinity the potential energy vanishes to zero.

It is apparent that this method of describing interactions presumes their instantaneous propagation, i. e., it is in accord with Galileo's principle of relativity. Actually, the force ¹

$$F_{\alpha} = - \frac{dU}{dr_{\alpha}} \quad (5)$$

exerted on a particle by other particles depends at any instant, in this mode of description, only on the positions of the particles at that instant. A change in the position of one particle affects the other particles at the same instant.

This remark leads us to the conclusion that forces depend only on the mutual positions of material points and not on their motion.

It follows from equations (4) and (5) that

Newton's second law	$\frac{dp_{\alpha}}{dt} = - \frac{dU}{dr_{\alpha}}.$	(6)
--------------------------------	--	-----

This formula gives the most general equation of motion of a system of material points. The equations of mechanics in the form (6) are called the Newtonian equations (Newton's second law).

Summing the equations of motion (6) over a system of particles and taking into account that, from the law of conservation of momentum, the left-hand side of the equation becomes zero, we find that the sum of all the forces in a closed system is zero:

$$F_1 + F_2 + \dots + F_n = 0.$$

In the special case of a system consisting of two material points,

$$F_1 + F_2 = 0.$$

¹ The derivative of a quantity with respect to a vector is a vector the components of which are equal to the derivatives of that quantity with respect to the components of the vector.

This means that the force exerted on the first particle by the second is equal in magnitude and opposite in direction to the force exerted by the second particle on the first. This is known as the *law of action and reaction* (Newton's third law).

**Some properties
of mass**

The concept of momentum can be used to formulate the concepts of rest and velocity of a system as a whole. A system of material points is at rest in that frame of reference in which its momentum is zero. The velocity of a system of material points relative to some reference frame O is defined as the velocity of a frame of reference O_0 in which the system is at rest.

Denoting by \mathbf{v}_α and $\mathbf{v}_{0\alpha}$ the velocities of a material point α relative to reference systems O and O_0 , and by \mathbf{v} the velocity of system O_0 relative to system O , the Galilean transformations yield

$$\mathbf{v}_\alpha = \mathbf{v}_{0\alpha} + \mathbf{v}.$$

Multiplying through by the mass of the corresponding material points and summing over all the points, we obtain

$$\mathbf{P} = \mathbf{P}_0 + \mathbf{v} \sum_{\alpha} m_{\alpha}.$$

The system of material points is at rest relative to the reference frame O_0 , hence, $\mathbf{P}_0 = 0$ and

$$\mathbf{P} = \mathbf{v} \sum_{\alpha} m_{\alpha}.$$

As \mathbf{P} is the momentum and \mathbf{v} the velocity of the system as a whole, the latter relationship expresses the additivity of mass: the mass of a composite body is equal to the sum of the masses of its component parts. In other words, we have the *law of conservation of mass*.

Here are some other properties of mass. Making use of the relationship between the momentum and velocity of a particle, equation (6) can be written for one particle in the form

$$m \frac{d\mathbf{v}}{dt} = - \frac{dU}{dr}.$$

According to Galileo's relativity principle, this equation does not change in passing from one inertial reference system to another. It follows, in particular, that the mass m of a particle is *invariant*, i. e., independent of the choice of reference system and therefore is a true characteristic of the particle.

Another property of mass to be noted is that it cannot be negative,¹ i. e., the directions of the momentum and velocity of a particle coincide.

Law of conservation of angular momentum We have seen that the law of conservation of linear momentum follows from the homogeneity of space with respect to a closed system of particles. Space is also isotropic; and any direction is as good as the other. Hence the properties of a closed system should not change if the whole system is turned through an arbitrary angle about an arbitrary axis. This condition implies the conservation of some vector quantity characterising the system. Namely, every material point of a system is characterised by a vector, such that the sum of all the vectors over all the particles of a given closed system is independent of time. This is the vector of *moment of momentum*, or *angular momentum* as it is more commonly called. It is related in a specific way with the linear momentum vector: the angular momentum of a particle is equal to the product of the radius vector and the linear momentum of the particle. Thus *the law of conservation of angular momentum* takes the form

$$\sum_{\alpha} \mathbf{r}_{\alpha} \times \mathbf{p}_{\alpha} = \mathbf{M} = \text{const.}$$

So far we have been speaking of closed systems only. Let us now consider the motion of a system interacting with another system whose motion is known. In this case the concept of *motion in an external field* is introduced.

We know already that interactions of material points are described by means of the potential function

$$U = U(\mathbf{r}),$$

¹ This follows from the very general physical principle known as the principle of least action.

which is a function of the coordinates of the interacting particles. It describes the interactions of the material points of a system and, if the system is not closed, the interactions between the material points of the system and other bodies.

For a nonclosed system in given external conditions the potential function may be explicitly time-dependent:

$$U = U(\mathbf{r}, t).$$

It represents the *potential energy in an external field*.

And so, motion in external fields or, more precisely, the conservation laws for open systems.

In the most general case, the angular momentum of a system in an external field is not conserved, though it may be in some special cases. Thus, if a system is in a central-symmetrical field (one in which potential energy depends only on the distance from some fixed point, the centre) all directions in space from that centre are equivalent, and the angular momentum of a system with respect to that centre is conserved. With respect to any other point in space, however, it will naturally change.

If a system is not closed but its interactions with surrounding bodies are such that the external conditions do not change in its displacement in some direction l (an axial-symmetrical field), the projection of the linear momentum of the system on that direction is conserved:

$$p_l = \text{const.}$$

Similarly, in an axial-symmetrical field the component of the angular momentum along the axis of symmetry is conserved.

We shall restrict ourselves to the cited cases as they are the most important.

Law of conservation of energy Let us now consider the very important case of motion of an open system in a uniform external field. If a system

is not in a variable external field its properties are not explicitly time-dependent. This follows from the consideration that, in the absence of an external field (or in a uniform external field), all instants of time are equivalent with respect to a given physical system. A conse-

quence of this condition is the conservation of a scalar quantity E characterising the system,

$$\sum_{\alpha} \frac{m_{\alpha} v_{\alpha}^2}{2} + U = E, \quad (7)$$

called *energy*. The expression

$$\sum_{\alpha} \frac{m_{\alpha} v_{\alpha}^2}{2} = T$$

gives the kinetic energy of the system. The function U was mentioned before.

The relationship (7) expresses the *law of conservation of energy of a system*. It will be observed that energy comprises two essentially different components. The kinetic energy T is a quadratic function of velocity, the potential energy U is independent of velocity.

Isotropy of time In developing the law of conservation of energy we spoke of time homogeneity, according to which the physical properties of different moments of time are equivalent. Equation (7) indicates that time is not only homogeneous but isotropic as well, i. e., its properties are the same in both directions: substitution of $-t$ for t leaves equation (7), as well as the equation of motion (6), unchanged. In other words, if a system is capable of some type of motion, a reverse motion (one in which the system goes through the same states in reverse order) is always possible. In this sense, according to the laws of classical mechanics, all motions are *reversible*.

Noninertial frames of reference Up till now we have been considering the motion of a material point or a system of particles relative to inertial frames of reference. Accordingly, the equations of motion of a point had the form

$$m \frac{d\mathbf{v}}{dt} = - \frac{dU}{d\mathbf{r}}. \quad (8)$$

Such equations of motion hold for inertial reference frames, which, as we know, move relative to one another with uniform velocity. But if we go over to a *noninertial*

(accelerated) reference system the equations of motion change.

Consider a frame of reference O' in translatory motion with some velocity $V(t)$ relative to an inertial system O . Denoting by \mathbf{v} the velocity of a material point with respect to system O , and by \mathbf{v}' its velocity with respect to system O' , we get

$$\mathbf{v} = \mathbf{v}' + V(t).$$

Substituting this expression for \mathbf{v} into equation (8), taking into account that $V(t)$ is a given function of time, and introducing the vector $\mathbf{W}(t) = dV(t)/dt$, which is the acceleration of the primed coordinate system, we obtain the equation of motion of a material point in system O' :

$$m \frac{d\mathbf{v}'}{dt} = -\frac{dU}{d\mathbf{r}'} - m\mathbf{W}(t), \quad (9)$$

where \mathbf{r}' is the radius vector of the particle in the primed system.

A comparison of equations (8) and (9) shows that in an accelerated system moving in a straight line there appears an additional uniform field of force,¹ the force acting on the material point being equal to the product of its mass and acceleration $\mathbf{W}(t)$ and directed opposite the acceleration. This force is commonly called the *inertia force*.

If, in addition to translatory motion, the frame of reference is also rotating with an *angular velocity* $\Omega(t)$,² the equation of motion of a material point in that reference system can be obtained by the necessary velocity transformations. It will differ from (9) by three additional terms found in the right-hand side: $m\mathbf{r} \times (d\Omega/dt)$, the inertia force; $m(\Omega \times \mathbf{r}) \times \Omega$, called the *centrifugal inertia force*; and $2m\mathbf{v} \times \Omega$, called the *Coriolis inertia force*. Note that the latter, unlike the forces considered before, depends on velocity.

¹ That is, a field in which the force acting on a particle is the same at all points in space at a given moment.

² It is defined as $d\varphi/dt$, where $d\varphi$ = angle of rotation,

Chapter 2. Relativistic Mechanics

Galileo's relativity principle, we saw, leads to the conclusion that interactions between bodies propagate through space instantaneously. Experience shows, however,

The constant c that there is no such thing as instantaneous interaction in nature. Hence, classical mechanics, which proceeds from the concept of instantaneous propagation of interaction, carries an inherent error. For, if there is a change in one of two interacting bodies, it will tell on the other body only after a certain time interval. Only after this necessary time has elapsed will the processes engendered by the change begin to take place in the second body. Division of the distance between the two bodies by this time interval yields the *velocity of propagation of the interaction*.

From the relativity principle¹ it follows, in particular, that the limiting velocity of propagation of interactions is the same in all inertial frames of reference. In other words, it is a universal constant, denoted by the symbol c . The latest measurements set its value at

$$c = 2.99776 \times 10^{10} \text{ cm/sec.}$$

The tremendous magnitude of this velocity explains why classical mechanics is sufficiently accurate for real

¹ It is appropriate at this juncture to clarify some of the propositions associated with the relativity principle.

Experience shows that the *relativity principle* exists. According to this principle the laws of nature are the same in all inertial frames of reference. That is to say, the equations expressing the laws of nature are *invariant* with respect to coordinate and time transformations from one inertial system to another. This means that an equation describing some law of nature has the same form when it is expressed in terms of space-time coordinates in different inertial reference systems. The relativity principle may also be formulated as the principle of equivalence of all inertial frames of reference.

situations. Most of the velocities with which we have to deal are so small in comparison with c that in practice the proposition concerning the infinity of the latter has no effect on the accuracy of computations.

**Relativistic
relativity
principle**

The relativity principle combined with the postulate of the finality of the propagation speed of interactions gives Einstein's principle of relativity, as opposed to Galileo's which presumed that interactions can propagate with infinite velocity.

The mechanics based on Einstein's relativity principle is called *relativistic*. In the limiting case of small velocities in comparison with c the effects of a finite propagation speed can be neglected, and relativistic mechanics turns into conventional *classical* mechanics based on the assumption of instantaneous propagation of interaction. The limiting transition from relativistic to classical mechanics may be formally expressed as going over to the limit $c \rightarrow \infty$ in the equations of relativistic mechanics.

**Space-time
separation**

We saw before that the limiting propagation speed c is the same in all inertial reference frames. Let us express this mathematically. Consider two coordinate systems O and O' moving with uniform velocity relative to each other. Let the coordinate axes be so directed that axis Ox coincides with axis $O'x'$ and the y and z axes are parallel to the respective primed axes; the time is t and t' in the unprimed and primed systems, respectively.

Consider an event ¹ viewed in system O as taking place at a point x_1, y_1, z_1 at time t_1 in that system. Let the event be the dispatch of a signal ² propagating with velocity c .

¹ An event is specified by the point where, and the time when, it takes place. Thus an event involving some material particle is defined by the three position coordinates of the particle and the time.

It is frequently useful to introduce for more graphic representation an imaginary, *four-dimensional space* with axes for the three spatial coordinates and time. In such a space an event is depicted by a point.

² An interaction propagating from one particle to another is sometimes spoken of as a "signal" or "message" sent from the first particle and "informing" the second of the change suffered by the first. The velocity of propagation of interaction is then called the "velocity of the signal".

Let the second event be the arrival of the signal at a point x_2, y_2, z_2 at time t_2 . The velocity of the signal being c , the distance travelled by it is $c(t_2 - t_1)$. On the other hand, that same distance is equal to

$$[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{1/2}.$$

We can thus write the following relationship between the coordinates of both events in system O :

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = 0. \quad (10)$$

The propagation of the signal and the two events can also be observed from system O' . Let the space-time coordinates of the first event in the primed system be x'_1, y'_1, z'_1, t'_1 and of the second event x'_2, y'_2, z'_2, t'_2 . As the velocity of the signal is the same in both systems, we have, by analogy with (10),

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = 0.$$

The coordinates of any two events being x_1, y_1, z_1, t_1 and x_2, y_2, z_2, t_2 , the quantity

$$S_{1,2} = [c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2]^{1/2}$$

is called the *space-time separation* between the two events.

It follows thus from the constancy of the velocity c that if the separation between the two events is zero in one inertial frame it is zero in any other. For the space-time separation¹ between two infinitely close events we get

$$dS^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2,$$

which is known also as the *squared interval*.

¹ For mathematical convenience the variable τ is sometimes used instead of t . They are connected by the relationship $\tau = ict$. Then

$$S_{1,2}^2 = -[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 + (\tau_2 - \tau_1)^2]$$

and

$$dS^2 = -(dx^2 + dy^2 + dz^2 + d\tau^2).$$

Accordingly, one lays off on the coordinate axes of the imaginary four-dimensional space, not x, y, z, t , but x, y, z, τ . Then $-S_{1,2}^2$ can be interpreted as the square of the distance between points x_1, y_1, z_1, τ_1 and x_2, y_2, z_2, τ_2 , and $-(dS)^2$ as the squared elementary length.

**Invariance
of space-time
separation**

As shown above, if $dS = 0$ in some inertial frame, then $dS' = 0$ in any other. On the other hand, dS and dS' are quantities of the same order of smallness. It follows from these two considerations that dS and dS' must be proportionate:

$$dS = \alpha dS',$$

the coefficient α being dependent only on the absolute value of the relative velocity of both inertial systems. It cannot depend on the space-time coordinates, as otherwise different points of space and moments of time would not be homologous, which contradicts the notion of space and time homogeneity. Neither can it be dependent on the direction of the relative velocity, as this would contradict the notion of space isotropy. Hence, we can write

$$dS' = \alpha dS$$

with the same justification as we wrote $dS = \alpha dS'$, as the velocity of the first system relative to the second is equal to the velocity of the second relative to the first. Substituting $dS = \alpha dS'$ into $dS' = \alpha dS$, we find that $\alpha^2 = 1$, i. e., $\alpha = \pm 1$. For selecting one of these values note that α can be either always $+1$ or always -1 . For if $\alpha(v)$ could be $+1$ for some velocities and -1 for others, for yet others it would have to have values lying between $+1$ and -1 , which is impossible. But if that is so then α must always be equal to $+1$, since a special case of the transformation of $dS' = \alpha dS$ is the identity $dS \equiv dS'$, where $\alpha = +1$. From $dS' = dS$ it follows that for all finite separations $S' = S$.

Thus we arrive at a very important result: the space-time separation between events is the same in all inertial frames of reference, i. e., it is *invariant* in relation to transformations from one coordinate system to another. This invariance is the mathematical expression of the constancy of the velocity c .

**The Lorentz
transformations**

Now let us develop the formulas for going over from one inertial reference frame to another, i. e., the formulas by which, knowing the coordinates x, y, z, t of an event in some reference system O one can find the coordinates

x', y', z', t' of the same event in another inertial reference system O' .

In classical nonrelativistic mechanics this problem is easily solved. As time is absolute, $t = t'$. Then, if the coordinate axes are so chosen that axes x and x' coincide and axes y and z are respectively parallel to axes y' and z' , the motion being along the axes x and x' , the coordinates y and z will, evidently, be equal to the corresponding coordinates y' and z' , while the difference between x and x' will be given by the distance travelled by one system with respect to the other. If the time count has been chosen to start when the origins of the two systems coincided, and the velocity of the primed system relative to the unprimed one is V , the distance travelled is Vt . Thus,

$$x = x' + Vt', \quad y = y', \quad z = z', \quad t = t'. \quad (11)$$

These are the transformation formulas in classical mechanics (the Galilean transformations). It can easily be proved that they do not satisfy *relativity theory*, as could be expected: they do not leave the separation between events invariant.

The relativistic transformation formulas must be sought on the basis of the requirement of separation invariance. Using in the following discourse the more convenient quantity $\tau = ict$, the space-time separation, as we have seen (cf. footnotes on pp. 30, 31) can be regarded as the distance between two corresponding points in a four-dimensional coordinate system. We can say, therefore, that the required transformation must leave unchanged all the lengths x, y, z, τ in the four-dimensional space. But such transformations represent either parallel translations or rotations of coordinate systems. Of these, translation of a coordinate system parallel to itself is of no interest as it simply means a transfer of the origin of the spatial coordinates and a new beginning of the time count. Thus the required transformation must be mathematically expressed as a rotation of a four-dimensional coordinate system x, y, z, τ .

Any rotation in four-dimensional space can be resolved into six rotations, viz., rotations in the planes $xy, zy, xz, \tau x, \tau y, \tau z$ (just as any rotation in ordinary space can

be resolved into three rotations in the planes xy , zy , and zx). The first three of these rotations transform only the spatial coordinates; they correspond to ordinary spatial rotations.

Let us investigate the rotation of plane τx ; the coordinates y and z remain unchanged. If φ is the angle of rotation, the connection between the old and new coordinates is given by the formulas

$$\left. \begin{aligned} x &= x' \cos \varphi - \tau' \sin \varphi, \\ \tau &= x' \sin \varphi + \tau' \cos \varphi. \end{aligned} \right\} \quad (12)$$

We are seeking the formulas for going over from one inertial reference system O to another system O' moving relative to the former with a velocity V along the x axis. It is evident that only the x coordinates and the time t undergo a transformation, hence the transformation must be of the form (12). All that is left is to determine the angle φ , which can be dependent only on the relative velocity V ¹.

Consider the motion of the origin of system O' in system O . In this case $x' = 0$, and the equations (12) take the form

$$x = -\tau' \sin \varphi, \quad \tau = \tau' \cos \varphi,$$

or, dividing the one by the other,

$$x/\tau = -\tan \varphi.$$

But $\tau = ict$, and x/t is, evidently, the velocity V of system O' relative to O . Thus,

$$\tan \varphi = i \frac{V}{c},$$

whence

$$\sin \varphi = \frac{i \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad \cos \varphi = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

¹ Note that throughout this book we are denoting by the symbol V any uniform relative velocity of two inertial frames of reference, and by the symbol v the velocity of a moving particle, which may not necessarily be uniform.

Substituting the expressions for $\sin \varphi$ and $\cos \varphi$ into (12), we find that

$$x = \frac{x' - i \frac{V}{c} \tau'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad y = y', \quad z = z', \quad \tau = \frac{\tau' + i \frac{V}{c} x'}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Substituting, furthermore, $\tau = ict$ and $\tau' = ict'$, we obtain finally

$$x = \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad y = y', \quad z = z', \quad t = \frac{t' + \frac{V}{c^2} x'}{\sqrt{1 - \frac{V^2}{c^2}}}. \quad (13)$$

These expressions are known as the *Lorentz transformations*, and they are of fundamental importance in contemporary physics.

**Corollaries
of the Lorentz
transformations**

It will be readily observed that in the limiting case of classical mechanics ($c \rightarrow \infty$) the Lorentz transformations change into the old Galilean transformations (11).

At $V > c$ the coordinates x and t in equations (13) become imaginary; this corresponds to the statement that motion faster than c is impossible. It is also impossible to use a reference system travelling with the velocity c , as the denominators in the formulas (13) would vanish.

**Length
in relativity
theory**

Now let there be a measuring rod at rest in system O lying parallel to the x axis. Its length measured in the system is $\Delta x = x_2 - x_1$, where x_2 and x_1 are the coordinates of its ends. What is the length of the rod in system O' ? To determine it we must locate the coordinates x'_2 and x'_1 of both ends of the rod in the primed system at precisely the same instant t' . From equations (13) we have

$$x_1 = \frac{x'_1 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad x_2 = \frac{x'_2 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

The length of the rod in the primed system is $\Delta x' = x'_2 - x'_1$. Subtraction of x_1 from x_2 yields

$$\Delta x = \frac{\Delta x'}{\sqrt{1 - \frac{V^2}{c^2}}}. \quad (14)$$

The proper length of the measuring rod is its length in the reference system where it is at rest. Denoting it by the symbol $l_0 = \Delta x$, and its length in some other reference system O' by l , we obtain

$$l = l_0 \sqrt{1 - \frac{V^2}{c^2}}.$$

Thus the rod is longest in the system in which it is at rest. Its length in a reference system relative to which it is moving with velocity V is shorter by the factor $\sqrt{1 - \frac{V^2}{c^2}}$. This consequence of theory is known as the *Lorentz contraction*.

For the limiting case of $c \rightarrow \infty$ (classical mechanics) we obtain from equation (14) $\Delta x = \Delta x'$, which is completely in accord with classical notions concerning the absolute nature of length.

**Time interval
in relativity
theory**

Now consider two events taking place at the same point in space, whose coordinates are x', y', z' in system O' . According to a clock at rest in the system the time between the two events in the primed system is $\Delta t' = t'_2 - t'_1$. Let us find the time interval Δt between the two events as measured in system O .

From (13), we have

$$t_1 = \frac{t'_1 + \frac{V}{c^2} x'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad t_2 = \frac{t'_2 + \frac{V}{c^2} x'}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Subtracting, we obtain

$$t_2 - t_1 = \Delta t = \frac{\Delta t'}{\sqrt{1 - \frac{V^2}{c^2}}},$$

which can be rewritten

$$\Delta t' = \Delta t \sqrt{1 - \frac{V^2}{c^2}}. \quad (15)$$

Time measured by a clock moving together with a given object is called the *proper time* of that object. The above relationship establishes the connection between proper time and time in the reference system relative to which the motion is being considered.

It is evident from equation (15) that the proper time of a moving object is always less than a corresponding time interval in a stationary system. In other words, the rhythm of a travelling clock is slower than that of a stationary clock.

In the limiting case of classical mechanics ($c \rightarrow \infty$) the equality $\Delta t' = \Delta t$ develops, which is in accord with classical notions concerning the absolute nature of time.

Thus relativity theory introduces important changes in the fundamental physical concepts of space and time. Our notions based on daily experience, we find, are approximations due to the fact that in everyday life we deal with velocities far below c .

Another consequence of the Lorentz transformations arises from the formulas which relate the velocity of a material particle in one system to its velocity in another system.

**Velocity
composition
in relativity
theory**

Once again, let system O' be moving with velocity V along the x axis of system O , let $v_x = dx/dt$ be the velocity component of a particle in the unprimed system and $v'_x = dx'/dt'$ its velocity component in the primed system. Differentiating equations (13), we obtain

$$dx = \frac{dx' + V dt'}{\sqrt{1 - \frac{V^2}{c^2}}}, \quad dy = dy',$$

$$dz = dz', \quad dt = \frac{dt' + \frac{V}{c^2} dx'}{\sqrt{1 - \frac{V^2}{c^2}}}.$$

Dividing the first three equations by the fourth, we get

$$\frac{dx}{dt} = \frac{dx' + Vdt'}{dt' + \frac{V}{c^2} dx'}, \quad \frac{dy}{dt} = \frac{dy' \sqrt{1 - \frac{V^2}{c^2}}}{dt' + \frac{V}{c^2} dx'},$$

$$\frac{dz}{dt} = \frac{dz' \sqrt{1 - \frac{V^2}{c^2}}}{dt' + \frac{V}{c^2} dx'}.$$

Dividing further the numerator and denominator of the right-hand sides of these equations by dt' , we obtain

$$v_x = \frac{v'_x + V}{1 + \frac{v'_x V}{c^2}}, \quad v_y = \frac{v'_y \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{v'_x V}{c^2}},$$

$$v_z = \frac{v'_z \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{v'_x V}{c^2}}.$$

These formulas give the law for the change in velocity in going over from one reference system to another and the rule for velocity composition in relativity theory. In the limiting case of $c \rightarrow \infty$ they turn into the formulas

$$v_x = v'_x + V, \quad v_y = v'_y, \quad v_z = v'_z$$

of classical mechanics.

In the special case of a particle moving parallel to the x axis, $v_x = v$, $v_y = v_z = 0$. Then $v'_y = v'_z = 0$ and $v'_x = v'$, and

$$v = \frac{v' + V}{1 + \frac{v'V}{c^2}}.$$

It is readily apparent from this formula that the sum of two velocities less than, or equal to, c cannot be greater than c .

Relativistic mechanics Let us examine some quantities the very existence of which follows from the most general properties of space-time symmetry. We shall begin with a single free particle.

Momentum of a relativistic particle The conservation of the momentum of a free particle, as we know, is a consequence of spatial homogeneity. In relativistic mechanics the momentum of a particle is expressed by the formula

$$\mathbf{p} = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}} \mathbf{v}. \quad (16)$$

When velocity is small ($v \ll c$), or in the limit when $c \rightarrow \infty$, this expression turns into the classical $\mathbf{p} = m\mathbf{v}$.

In classical mechanics every particle is characterised by its mass m . In relativistic mechanics a particle is also characterised by its mass, insofar as in going over from one inertial reference frame to another its value does not change; as they say in relativity theory, the mass m of a particle is a relativistic invariant.

Energy of a relativistic particle The conservation of the energy of a free particle is a consequence of time homogeneity. In relativistic mechanics the energy of a particle is expressed by the formula

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (17)$$

It is seen from this expression that in relativistic mechanics the energy of a particle does not vanish even when the velocity is zero. This "rest energy", i. e., the energy when $v = 0$, is

$$E = mc^2.$$

At low velocities ($v/c \ll 1$) the expression (17) takes the form

$$E \approx mc^2 + \frac{mv^2}{2},$$

i. e., the classical expression for the kinetic energy of a particle less the rest energy.

Formulas (16) and (17) yield the following relationship between the energy and momentum of a free material particle:

$$p = \frac{Ev}{c^2}. \quad (18)$$

At $v = c$ the momentum p (16) and the energy E (17) of a particle become infinite. This means that a particle whose mass is not zero cannot move with the speed c . The momentum of a particle with zero mass written in the form (16) gives, at $v = c$, an indeterminacy of the form $0/0$ and may remain finite. In this case the velocity of the particle must always be c . From equation (18) we have for such particles

$$p = \frac{E}{c}.$$

Thus relativistic mechanics allows for the existence of particles with zero mass moving with the velocity c . Later on we will see that light phenomena can be interpreted in terms of such particles.

**Systems
of relativistic
particles**

The foregoing formulas are equally valid for the motion of a whole composite body consisting of many particles. In this case by mass is always meant the aggregate mass of the body, and by velocity, its velocity as an entity.

Consider a body at rest (as a whole). Then its energy, which we call *internal*, is equal to Mc^2 , where M is the mass of the body. This energy comprises, besides the rest energy of its component particles, their kinetic energy and the energy of their interactions. In other words, Mc^2 is not simply the sum $\sum m_\alpha c^2$, where m_α is the mass of a particle α belonging to the body, and therefore M is not equal to $\sum m_\alpha$.

Thus, in relativistic mechanics the law of conservation of mass is no longer valid; the mass of a composite body is not equal to the sum of the masses of its component

parts. Instead, there is only the law of conservation of energy, which also includes the rest energy of the particles. The difference $\Delta M = M - \sum_{\alpha} m_{\alpha}$ between the mass of a composite body and the net mass of its component parts is called the *mass defect*. The quantity $\Delta M c^2$ is called the *binding energy*.

Consider a body consisting of two parts of masses M_1 and M_2 in a frame of reference in which it is at rest and suppose that it is spontaneously splitting into the two parts, the respective velocities of which are v_1 and v_2 . Then, according to the principle of conservation of energy, we have

$$M c^2 = \frac{M_1 c^2}{\sqrt{1 - \frac{v_1^2}{c^2}}} + \frac{M_2 c^2}{\sqrt{1 - \frac{v_2^2}{c^2}}}.$$

This equation is satisfied only if $M > M_1 + M_2$, i. e., if the mass defect $\Delta M = M - M_1 - M_2$ is positive. Thus, spontaneous decay is possible only if the mass defect of a body is positive with respect to the parts into which it is disintegrating. Conversely, if the mass defect is negative, the body is stable and cannot decay spontaneously. In this case, evidently, disintegration can be achieved only by injecting energy from outside, this energy being at least equal to the binding energy $\Delta M c^2$.

Elementary particles in relativity theory

Let some rigid body¹ be put into motion by an external force applied at one of its points. If the body is absolutely rigid, all its points must start moving at precisely the same instant as the one to which the force is applied. This is quite possible in classical mechanics by virtue of the instantaneous propagation of pressure. In relativity theory, however, this is impossible, as a pressure exerted at one point of a body is transmitted to the other points with a finite velocity; hence all points of the body cannot start moving at once, which means that the body suffers a deforma-

¹ By a rigid body is meant a system of material particles the distances between which are constant.

tion. Thus relativity theory leads to the conclusion that there is no such thing as an absolutely rigid body.

The foregoing leads to certain conclusions regarding elementary particles. By an *elementary particle* is meant a body which participates in all physical phenomena as an entity, i. e., it is meaningless to speak of its parts. In other words, the state of an elementary particle is completely defined by its location and velocity as a whole. Apparently, if an elementary particle possessed finite dimensions it could not be deformable, since deformation is associated with the possibility of independently moving parts of a body. But relativity theory, as we have just found, rejects the concept of absolutely rigid bodies. We thus arrive at a very important conclusion: in relativity theory elementary particles cannot have finite dimensions and must be regarded as geometrical points.

Chapter 3. Electromagnetic Field Theory

Up till now we have been discussing the properties of particles. We have found that interactions between particles can be described by means of the

Fields

concept of field of force. Instead of stating that a particle is acting on another we can say that a particle creates a field around itself. Any other particle inside this field is subjected to the action of some force. In classical mechanics field is but a means of describing particle interaction. In relativity theory, however, the postulate of the finality of the propagation speed of interactions changes this state of affairs substantially. The forces acting on particles at a given instant are not determined by the positions of the particles at that instant. Any change in the relative position of one particle affects other particles only after a certain interval of time. This means that *field* is a physical reality *per se*. It is wrong to speak of direct interaction of particles at a distance. Interaction at any instant is possible only between neighbouring points in space (close-quarter or contact interactions). Hence, what actually must be considered is interaction between a particle and a field followed by interaction between the field and another particle.

Let us see what is meant by the physical entity called *field*. We shall consider two types of field, gravitational and electromagnetic. First, the *electromagnetic field*.

Classical electrodynamiccs A field displays its properties only in interaction with particles. It follows, then, that a field is characterised by the effect it has on the motion of particles in it. Thus, two fields are physically identical if at any time they exert the same action on a particle at any point in space.

**Electric charge
of a particle**

The interaction of a given electromagnetic field with a particle is determined by a single quantity characterising that particle, its charge.¹ Charge may be positive, negative or zero. Thus we distinguish between charged and uncharged, or neutral particles.

**Field of a
particle at rest**

We shall begin with the simplest case of a uniform electrical, or electrostatic, as it is commonly called, field created by charged particles at rest. In such a field, acting on a particle having a charge e_1 is a force

$$F = \frac{ee_1}{r^2}, \quad (19)$$

where e is the charge responsible for the field and r is the distance between charges e and e_1 . This is known as *Coulomb's law*.

It is now easy to determine the energy of interaction of the two charges. For, taking into account the well-known relationship $F = -dU/dr$, equation (19) yields

$$U = \frac{ee_1}{r}. \quad (20)$$

The formulas for force and potential energy would not change if the sign of both charges were reversed, thus suggesting that the name of a charge—positive or negative—is purely arbitrary; what matters is the difference in signs.

From equation (20) it also follows that potential energy may have different signs: it is positive when both charges have the same sign and negative when they have different signs. This corresponds to the observed fact of repulsion and attraction of charges.²

¹ By definition, the charge of a particle is an invariant, i. e., it does not depend on the choice of reference system.

² Insofar as the potential energy of interacting particles is always chosen so that it becomes zero when the distance between them is infinite, in which case the interaction can be neglected, the sign of potential energy is determined by the type of interaction: whether it is an attraction or a repulsion. It is positive for repulsion and negative for attraction.

The quantities characterising an electric field as such are *intensity*, or *field strength*

$$E = \frac{F}{e_1}, \quad (21)$$

which is force referred to a unit charge; and *potential*

$$\varphi = \frac{U}{e_1}, \quad (22)$$

which is energy referred to a unit charge.

Using equations (19) and (20), the formulas (21) and (22) can be written for point charges in the form

$$E = \frac{e}{r^2}$$

and

$$\varphi = \frac{e}{r}.$$

These formulas give the intensity and potential of a field at a point located at a distance r from the charge e that produces the field.

Field of a system of charges

If we have a system of charges, we must proceed from the following very important property of the electromagnetic field to establish the form of the field. Experience shows that electromagnetic fields obey the so-called *superposition principle*. This means that if a charge produces a field and another charge produces another field, the field due to both charges together presents a simple addition of the two fields. This, in turn, means that the strength of the resultant field at any point is equal to the vector sum of the component field strengths at that point.

Consider a system of several charged particles. Let us determine the field produced by this system at a great distance from it. From the superposition principle, it is equal to the sum of the fields produced by each charge. So we have

$$E = \frac{e_1}{r^2} + \frac{e_2}{r^2} + \frac{e_3}{r^2} + \dots$$

We do not write a vector sum as we assume that the particles are contained in a very small volume and the

direction of the radius vectors of each charge is the same. Hence, the foregoing equation can be rewritten as follows:

$$E = \frac{e_1 + e_2 + e_3 + \dots}{r^2}.$$

We see that the only difference between the field of a simple and a compound particle is that, instead of the charge e , we have a sum of charges $e_1 + e_2 + e_3 + \dots$. This relationship expresses the *law of conservation of charge*. It shows that the charge of a system of charges equals the sum of the component charges. Consider a system of point charges. Its energy can be determined from the expression (20), which gives the interaction energy of two charges. For mathematical convenience, namely, that the equations be of a symmetrical form, let us write this expression in another notation:

Energy of a system of charges

$$U = \frac{e_1 e_2}{r}. \quad (23)$$

Introducing the potentials of the fields produced by charges e_1 and e_2 , $\varphi_1 = e_2/r$ and $\varphi_2 = e_1/r$, equation (23) can be written in the form

$$U = \frac{1}{2} (e_1 \varphi_1 + e_2 \varphi_2).$$

Generalised for an arbitrary number of charges, the equation yields

$$U = \frac{1}{2} \sum_{\alpha} e_{\alpha} \varphi_{\alpha},$$

where φ_{α} is the potential of the field produced by all the charges at the point of location of the charge e_{α} .

"Self"-energy of a particle

Applying this formula to a single charged elementary particle (an electron, for example) and the field produced by it, we come to the conclusion that the charge must possess a potential "self"-energy, equal to $e\varphi/2$, where φ is the potential of the field made by the charge at the location of the charge. But in relativity theory, we know,

an elementary particle must be regarded as a point. On the other hand, the potential $\varphi = e/r$ of its field becomes infinite at point $r = 0$. Thus, according to electrodynamics, an electron should possess an infinite self-energy and, consequently, infinite mass (which is the quotient of energy divided by c^2). The physical absurdity of this result shows that the fundamental principles of electrodynamics of necessity lead to the conclusion that its sphere of application is restricted.

It could be noted that the infinite self-energy and mass yielded by electrodynamic theory make it impossible to pose within its framework the question whether the whole mass of an electron is electromagnetic (i. e., whether it is connected with the electromagnetic self-energy of the particle).

**Limits
of application
of classical
electrodynamics**

Insofar as the appearance of the physically absurd concept of infinite self-energy of an elementary particle is due to the need for considering such a particle as a point, the conclusion can be drawn that electrodynamics as a logically isolated physical theory becomes intrinsically contradictory when very small distances are involved. Of what order of magnitude are these distances, it may be asked. The answer will be forthcoming if it is noted that the electromagnetic self-energy of an electron should be of the order of the rest energy mc^2 .

If, on the other hand, an electron is regarded as having a dimension r_0 , its potential self-energy would be of the order of e^2/r_0 . From the requirement that both quantities must be of the same order, i. e., $e^2/r_0 \sim mc^2$, we get ¹

$$r_0 \sim e^2/mc^2.$$

This formula defines the limits of application of electrodynamics which follow from its basic propositions. It should be noted, however, that actually the limits of application of "classical" electrodynamics outlined here are even higher due to *quantum phenomena*.²

¹ The quantity $r_0 = e^2/mc^2$ is called the "classical electron radius" and is equal to about 10^{-13} cm.

² Quantum effects come into play at distances of the order of h/mc , i. e., 10^{-10} cm, where h is Planck's constant.

Maxwell's equations We have examined the basic concepts of the electromagnetic field theory, or *classical electrodynamics* as it is called, on the simplest example of a uniform field. Now let us investigate the case of an arbitrary electromagnetic field and derive the equations describing it.

Any solution of a field equation must describe a field that can be produced in actual conditions. According to the superposition principle, the sum of any such fields should represent a field also capable of being reproduced physically, i. e., it must satisfy the field equations. One of the properties of linear differential equations is that the sum of any solutions of an equation is also a solution. Consequently, a field equation must be a linear differential equation. The set of simultaneous equations describing an electromagnetic field is known as Maxwell's equations, and they are the basic equations of electrodynamics.¹

Invariance of field equations

As a field is a relativistic entity, Maxwell's equations must evidently be invariant with respect to the Lorentz transformations. The field transformation laws reveal that electric and magnetic field parameters, like most physical quantities, are relative, i. e., they are different in different frames of reference. In particular, an electric or magnetic field may be zero in one reference system and exist in another.

Laws of field transformation

Since, by virtue of the finality of the velocity of propagation of interactions, a field must be regarded as an independent system with its own "degrees of freedom", it apparently must possess energy, linear momentum and angular momentum. It is also apparent that all the corresponding conservation laws must hold both for a free field and for a field with charges in it.

The conservation laws in electrodynamics

Fields in vacuum

Maxwell's equations possess one very important property: their solution in vacuum need not be zero. This means that an electromagnetic field may exist even in the absence of any charges.

¹ Note that Maxwell's equations commonly assume knowledge of the distribution and motion of charges in space, and the field produced by them is found as a solution of the equations.

Electromagnetic waves

Electromagnetic fields in vacuum without charges are called *electromagnetic waves*. In the absence of charges such a field must of needs be variable, which follows from the very form of the field equations.

Plane waves

A *plane wave* is a special case of electromagnetic waves in which the field depends on only one coordinate, say, x (and on time). The solutions of the field equations in this case take the form

$$f = f_1(x - ct) + f_2(x + ct),$$

where f is any vector component characterising the field and f_1 and f_2 are arbitrary functions.

Take, for example, $f_2 = 0$, whence $f = f_1(x - ct)$. Let us investigate the meaning of this solution. The field changes with time in every plane $x = \text{const}$; at any given instant it is different for different values of x . The field, evidently, has the same value for the coordinates x and times t satisfying the relationship $x - ct = \text{const}$, i. e.,

$$x = \text{const} + ct.$$

This means that if at some time $t = 0$ the field strength at a point x in space had some specific value, then after a time interval t it will have the same value at a distance ct along the x axis from the initial point. We can say that all values of the electromagnetic field travel through space along the x axis with the velocity c .

Thus, $f_1(x - ct)$ represents a plane wave travelling in the positive direction of x . It will readily be observed that $f_2(x + ct)$ is a wave travelling in the opposite, negative, direction of x . Obviously, these waves propagate independently of one another, i. e., without interacting.

The solutions of Maxwell's equations for empty space (in the absence of charges) thus take the form of travelling waves. The velocity of propagation of electromagnetic waves, including light, in vacuum is equal to the universal constant c ¹.

¹ This is why c is commonly referred to as the velocity of light,

**Monochromatic
wave**

It is interesting to note that the ratio between the energy and momentum of an electromagnetic wave is the same as for particles travelling with the velocity c . An important special case are electromagnetic waves whose fields are simple harmonic functions of time. These are known as *monochromatic waves*. The time dependence of all quantities (the components of the field-strength vectors) for a monochromatic wave is determined by a factor of the form $\cos \omega t$, where ω is the wave frequency. The period of a wave is equal to $2\pi/\omega$.

For a plane wave travelling in the positive direction of x the field is a function of $x - ct$ only. Therefore, if a plane wave is monochromatic its field is a simple harmonic function of $x - ct$.

Let f be any component of the vectors characterising a field; then for a monochromatic wave f is most conveniently expressed as the real part of the complex expression¹

$$f = a \exp \left[-i\omega \left(t - \frac{x}{c} \right) \right], \quad (24)$$

where a is a constant called the *amplitude*.

The quantity $\lambda = 2\pi c/\omega$ is called the *wavelength*, and it gives the period of change of a field at coordinate x at a given time t .

Introducing a unit vector n in the direction of wave propagation, equation (24) can be rewritten as follows:

$$f = a \exp \left[-i \left(\omega t - \frac{\omega}{c} nr \right) \right].$$

The vector

$$k = \frac{\omega}{c} n = \frac{2\pi}{\lambda} n$$

is called the wave vector. Consequently, a plane monochromatic wave is described by the expression

$$f = a \exp [i(kr - \omega t)]. \quad (25)$$

¹ Complex expressions are most conveniently used in view of the linearity of Maxwell's equations. This enables all operations to be performed not with trigonometric expressions but with simpler exponential expressions, then going over to their real part. In future we shall be using the complex notation, always presuming the real part of the corresponding expression.

The expression in the parentheses is called the *phase* of the wave.

**Interference
of waves**

Let two waves f_1 and f_2 be propagating in the same direction and let their vibrations be also in the same direction.

They will yield one wave, whose function is

$$f = f_1 + f_2.$$

The mean intensity of the wave (which is determined by the square of the field) has the form

$$\bar{f}^2 = \bar{f}_1^2 + \bar{f}_2^2 + 2\bar{f}_1\bar{f}_2. \quad (26)$$

Thus, the mean intensity of the resultant wave is not, generally, the sum of the mean intensities of the component waves. This phenomenon is known as *interference*. If equation (26) does not contain the product of the two waves (i. e., if $\bar{f}_1\bar{f}_2 = 0$), no interference takes place. The latter equality means, in turn, that both waves are physically independent (they come from different sources). For if two waves are independent of one another the mean of the product $\overline{f_1 f_2}$ equals the product $\bar{f}_1 \bar{f}_2$ of the mean values of \bar{f}_1 and \bar{f}_2 and, insofar as both are zero [the mean of oscillating multipliers of the form $\exp(\pm i\omega t)$ is zero], $\overline{f_1 f_2} = 0$.

**Geometrical
optics**

One of the properties of a plane wave is that its direction of propagation and amplitude are everywhere the same.

Arbitrary electromagnetic waves do not possess this property. Although in the most general case this is not true of electromagnetic waves, within small regions of space they can be treated as plane waves. For this, apparently, the amplitude and direction of the wave should not change perceptibly over a distance of the order of the wavelength.

If this condition is satisfied the concept of a *wave front* as a surface at all points of which a wave has the same phase at a given instant may be introduced. The wave front of a plane wave is, evidently, a plane normal to the direction of propagation. In a small region of space any wave may be assumed to be travelling normal to the wave front. A fundamental concept in geometrical

optics is that of a ray of light the tangent of which at any point is in the direction of propagation of the wave.

Geometrical optics treats of electromagnetic propagation, light in particular, as of rays, ignoring its undulatory character, that is, the wavelength is allowed to tend to zero.

The laws of geometrical optics are thus strictly valid only in the ideal case when a wavelength can be assumed to be infinitesimal.¹ The greater the departure from this condition, the greater the deviations from the laws of geometrical optics.

Wave diffraction The phenomenon which is observed as a result of these deviations is called *diffraction*. Diffraction can be observed when an obstacle is placed in the path of an electromagnetic wave or when a wave is made to pass through an aperture in an opaque screen. Geometrical optics predicts that the shadow beyond a screen must be sharply defined. Diffraction, however, produces a fairly complex intensity pattern on the fringes of the illuminated part, the phenomenon being the more pronounced, the smaller the screen or aperture and the longer the wave.

Diffraction phenomena are directly related to interference. This is especially apparent in the following experiment. Let a screen with two narrow slits be in the path of an electromagnetic wave. The beam passing through one of the slits when the other is closed produces a certain intensity pattern on a screen behind the slit. A similar pattern is observed when the beam is made to pass through the second slit only. But when it passes through both slits at the same time the pattern beyond is not that of a simple superimposition of the original two. Due to interference an alternation of intensities is observed.

Wave radiation We began our acquaintance with the basic concepts of electromagnetic field theory with a uniform field produced by charges at rest; then we considered variable charge-free fields. Now let

¹ Geometrical optics is applicable to real solutions by virtue of the smallness of the wavelengths concerned in comparison with the characteristic parameters of the problem.

us examine some properties of variable fields with charges in arbitrary motion.

Consider a field created by a system of moving charges at a great distance from that system (great in comparison with the size of the system). If, furthermore, the distances are also great in comparison with the length of the electromagnetic waves radiated by the system, in small sections of the field the waves may be treated as plane.

Insofar as we are considering a system of particles moving in a localised region of space, the charges are moving with acceleration. It follows from this that only accelerated charges are capable of radiation. Charges moving uniformly in a straight line do not radiate. This, of course, follows also from the relativity principle, since a uniformly moving charge may be investigated in an inertial system in which it is at rest, and stationary charges do not radiate. Furthermore, as fields possess energy, electromagnetic radiation involves energy transfer.

Chapter 4. Gravitational Field Theory

Another type of field that exists in nature is the so-called *gravitational field*. All general considerations concerning fields as relativistic entities presented in connection with the electromagnetic field are valid, apparently, for the gravitational field as well.¹

Gravitational fields

The basic property of gravitational fields is that all bodies, irrespective of mass or charge, move in them in the same way (provided, of course, that the initial conditions are the same). Thus, for instance, the laws of free fall in the earth's gravity field are the same for all bodies irrespective of mass: they all receive the same acceleration.²

The principle of equivalence

This property of the gravitational field establishes a very important analogy between the motion of a body under gravity and a body outside of any external field considered from the point of view of a noninertial frame of reference. In an inertial reference system free motion is always uniform and in a straight line, and if two velocities were the same at some initial instant they will be the same all the time. Therefore if investigating free motion in a noninertial system, all bodies will also be found to be moving in the same way with respect to it.

¹ Which is why we are proceeding at once with the relativistic approach.

² This, it should be noted, is not the case with electromagnetic fields, where acceleration depends on the ratio of the charge of a body to its mass, which is, of course, different for different bodies.

The conclusion is that the properties of motion in a non-inertial frame of reference are the same as in an inertial system with a gravitational field or, in other words, a noninertial reference system is equivalent to a gravitational field. This is known as the *principle of equivalence*.

Consider motion in a uniformly accelerated frame of reference. Bodies of any mass freely moving in such a system will, evidently, all have the same uniform acceleration, equal in magnitude and oppositely directed to the acceleration of the reference system.¹

This is the type of motion observed in a homogeneous, uniform gravitational field such as the earth's (or rather in small portions of it, where it can be regarded as homogeneous). Thus, a uniformly accelerated reference system is equivalent to a uniform, homogeneous external field. A somewhat more general case is that of a reference system moving parallel to itself in a straight line with non-uniform acceleration: it is evidently equivalent to a homogeneous, but variable, gravitational field.

Some special features It should be noted, however, that fields equivalent to noninertial reference systems are not absolutely identical to "real" gravity fields which exist also in noninertial systems. Namely, there exists a very important difference with regard to their properties in infinity. At an infinite distance from the body responsible for a field a "real" gravity field always tends to zero. On the other hand, fields equivalent to noninertial reference systems increase indefinitely in infinity, or at least they remain finite. For example, the centrifugal forces developing in a rotating frame of reference increase indefinitely with the distance from the axis of rotation; a field equivalent to an accelerated reference system moving in a straight line is the same throughout the whole of space into infinity.

Fields equivalent to noninertial reference systems disappear as soon as one goes over to an inertial system. "Real" gravitational fields, however, which exist also in inertial reference frames, cannot be excluded by any choice of coordinate system. This follows from the mentioned distinction between infinity conditions for "real"

¹ See equation (9), page 28.

fields of gravity and fields equivalent to noninertial frames of reference. Since the latter do not tend to zero in infinity it is obvious that no choice of reference system can exclude the "real" fields, which vanish in infinity.

All that can be achieved by a judicious choice of reference system is to exclude a field of gravity in a region of space sufficiently small to regard the field in it as homogeneous (over a small interval of time). This can be achieved by choosing a system moving with an acceleration equal to that obtained by a particle placed in the given region of the field.

The foregoing remarks indicate that the principle of equivalence, as applied to "real" gravity fields, is generally valid only when small regions of space are considered, where the field may be regarded with sufficient accuracy as homogeneous (over a small interval of time).

Curved coordinate systems To derive further corollaries from the foregoing propositions, the principle of equivalence must be investigated in action. Consider a noninertial frame of reference O' revolving with uniform velocity relative to an inertial system O about their common axis z . Let us write the separation formula for both systems and compare the two. Space-time separation in an inertial Cartesian coordinate system is given, it will be recalled, by the expression

$$dS^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \quad (27)$$

We also know that the separation remains the same in going over to any other inertial frame of reference, i. e., with respect to the Lorentz transformations. When a noninertial frame of reference is involved, however, the form of dS^2 changes. Thus, in our case of a uniformly rotating coordinate system,

$$\begin{aligned} x &= x' \cos \Omega t - y' \sin \Omega t, \\ y &= x' \sin \Omega t + y' \cos \Omega t, \\ z &= z', \end{aligned}$$

where Ω is the angular velocity of rotation, the space-time separation takes the form

$$\begin{aligned} dS^2 &= [c^2 - \Omega^2(x'^2 + y'^2)] dt^2 - dx'^2 - dy'^2 - dz'^2 + \\ &\quad + 2\Omega y' dx' dt - 2\Omega x' dy' dt. \end{aligned} \quad (28)$$

Whatever the time transformation law, this expression for dS^2 cannot be reduced to the form (27).

Thus, in a noninertial frame of reference the squared interval is quadratic in the general form of the differentials of the coordinates. This result means, in particular, that a four-dimensional coordinate system (with one time coordinate) is *curved* with respect to noninertial frames of reference.

The separation expression determines all geometric properties of space-time in every given curved coordinate system or, as they say, establishes the metric of the space-time continuum.

Non-Euclidean properties of space-time. Nature of gravitational field	To get back to our rotating reference system, a comparison of the expressions (27) and (28) reveals that the geometric properties of space suffer a change in going over from an inertial to a noninertial frame of reference.
--	--

Geometry becomes *non-Euclidean*, as distinct from conventional Euclidean geometry.

And finally, applying the principle of equivalence to our case (any noninertial frame of reference is equivalent to an inertial system with a corresponding gravitational field) we may draw the conclusion that a gravitational field represents nothing more than a change in the geometrical properties of the space-time continuum (i. e., a change in its metric).

The foregoing can be clarified by considering the simple example of a rotating frame of reference, specifically, a circle drawn in the $x'y'$ plane of a system O' with the centre on the axis of rotation. If there were no rotation, the ratio of the circumference l to the diameter d of the circle would be π . In the case of rotation relative to the inertial system O , all the length elements along the circumference experience a Lorentz contraction with respect to the unprimed system, the elements along the diameter (perpendicular to the velocity) remaining unchanged. As a result, the ratio l/d will no longer be π , and we find that geometrical relations in the noninertial frame of reference are non-Euclidean, thus differing from inertial systems. Furthermore, if there are two identical clocks rotating in the primed system, one on the circumference

and the other at the centre, an observer in the unprimed system will find that the former runs slower than the latter. The same, of course, is observed inside the primed system.¹ Thus the properties of time also change in going over to a noninertial frame of reference.

**Length
and time intervals
in general
relativity theory**

We have established that time intervals depend on the gravitational field. It is readily observed from the foregoing that distances between bodies are also affected by gravity. The presence of a gravitational field means a change in the space-time metric, such, in particular, that the space metric changes, being in the general case dependent on time. This leads to the conclusion that the component bodies of a system can never be at rest with respect to each other. The inevitability of such deformations is especially apparent from the fact that, as we have just seen, in non-Euclidean space the ratio of the circumference of a circle to its radius is not 2π and, generally speaking, it changes with time. Therefore, even if the distances between bodies along the radius of a circle remain unchanged, distances along the circumference must change, and vice versa. Hence, the mutual position of the component bodies of a system cannot be regarded as unchangeable.

**Frame of reference
in general
relativity theory**

One remark concerning the concept of frame of reference in the presence of a gravitational field is called for in connection with what has been said. As long as the gravitational field did not enter the scene we could use for a reference system any assembly of bodies at constant distances, that is, at rest with respect to one another. The gravitational field makes this impossible insofar, as we have seen, that the notion of bodies being stationary relative to one another becomes meaningless. Moreover, the very concept of a definite relative velocity of bodies becomes meaningless.

Accordingly, to locate bodies accurately in space with a gravity field it would be necessary, strictly speaking,

¹ Using the principle of equivalence, this result can be interpreted as meaning a change in the rhythm of a clock caused by a gravitational field. If one of two identical clocks has been placed in a gravitational field, it will soon be found to be slower.

to have a system comprising an infinite number of bodies filling the whole of space. Such a system of bodies, together with clocks having arbitrary rhythms connected with each body, represents a frame of reference in general relativity theory.

Space and matter Now is the time to see how the contents of the basic physical concepts of *length* and *time interval* change with the expansion of the scope of the phenomena considered.

In nonrelativistic mechanics lengths (or distances) and time intervals were the same in all frames of reference; they were absolute, in other words. In special relativity theory they depend on the choice of reference system.

With the introduction of the gravitational field (general relativity) space-time relationships—distances and time intervals—undergo a radical change. Space becomes *curved* or *non-Euclidean*, as distinct from *flat* or *Euclidean* space in the absence of gravitational fields. Time intervals vary not only in different reference systems but even at different points of space in one and the same system.

Furthermore, since a gravitational field can be produced by any physical object whatsoever, and the field itself is nothing but a change in the space-time metric, we arrive at the fundamental and remarkable conclusion that the geometrical properties of the space-time continuum are determined by physical phenomena and are not intrinsic properties of space and time themselves.

Let us investigate the motion of a material particle in non-Euclidean space or, in other words, its behaviour in a gravitational field.¹ Generally a grav-

**Free particle
in non-Euclidean
space**

itational field varies at different points of space and is time-dependent at every point. Furthermore, as length

and time interval depend on the field values and therefore differ at different points in space (within the same

¹ It should be remembered that the presence of a field at a given point means a change in the geometrical properties of space, which becomes non-Euclidean, at that point.

frame of reference), the concept that unites the two notions of space and time—velocity—must also have different values at different points of space. As velocity is a vector quantity, and in the most general case a change in velocity involves a change in both speed and direction, we conclude that *free motion of a particle in real three-dimensional non-Euclidean space is neither rectilinear nor uniform*. This, incidentally, explains the statement that in general relativity the notion of a specific relative velocity of free moving bodies is meaningless.

Gravitational field equations It should be borne in mind in investigating gravitational field theory that a quantitative physical theory is complete if the “equations of motion” have been developed for the phenomena involved (the field equations in our case).

To establish the form of the field equations we proceed from the fundamental physical property of the gravitational field: the principle of equivalence according to which the gravitational field is equivalent to a noninertial frame of reference. Since, furthermore, gravitational fields may be “arbitrary”, so may reference systems be arbitrary. Thus, in general relativity the choice of reference systems connected by reciprocally unique and continuous transformations is in no way restricted. In other words, any quantitative relationships describing a field, including the field equations, must be of a form that is independent of the choice of a reference system.

Covariance of equations It is appropriate to note at this juncture that the greater the scope of the phenomena involved the higher the class of transformations with respect to which the laws of physics are invariant. Thus, the equations of classical mechanics are invariant with respect to the Galilean transformations. This means that any equation describing a classical-mechanical phenomenon in terms of position coordinates and time has the same form in different inertial frames of reference. The equations of the more general relativistic mechanics and electromagnetic field theory are invariant with respect to a broader class of space-time transformations—the Lorentz transforma-

tions—though only inertial reference frames are still considered.

When gravity is taken into account, the laws of physics become so comprehensive and “exhaustive” that they are invariant with respect to any completely arbitrary space-time transformations (a property known as *covariance of equations*); in other words, physical laws are at last independent of the choice of frame of reference,¹ i. e., of the whim and will of the observer.

Mathematical form of equations The mathematical apparatus capable of satisfying the covariance requirements of general relativity equations is known as *tensor analysis*, and it studies four-dimensional geometry in arbitrary curved coordinates.² It follows that the gravitational field equations in general relativity theory must be of tensor form.

In establishing this feature of the field equations we proceed from the equivalence principle which, however, applies only for small regions of the space-time continuum. The conclusion, therefore, is that the field equations must be of differential form as, incidentally, the “equations of motion” in all other physical spheres.

Thus, the basic formulas of general relativity—the gravitational field equations—are differential tensor equations. They were first derived by Einstein and, as we shall see, their content is extremely broad.

Every branch of physical phenomena is characterised by the region of space-time in which they are observed. Thus, the laws of particle motion in classical and relativ-

¹ Now it becomes more or less clear why the gravitational field theory based on relativity theory is called *general relativity*.

² Note the following in this connection: in generalising known relationships for the case of an arbitrary *given* gravitational field these relationships are written in tensor form, i. e., an arbitrary four-dimensional curved coordinate system is introduced. This, in particular, is how the equations of motion of a particle in a given gravitational field are derived; this too is what is done in generalising the electrodynamic equations for the case of an arbitrary given gravitational field.

As distinct from the case when known phenomena are studied in the presence of a gravitational field, we are concerned here with the laws governing the change of the field itself, i. e., with developing the field equations.

istic mechanics are valid in any but the very smallest regions of space-time. The same holds for the laws governing electromagnetic and gravitational fields, and we have observed before that electrodynamic laws are no longer valid in small spatial regions.

Cosmic space-time relationships, we shall see later, are described by general relativity. In going over to the world of extremely small dimensions we must note, first of all, that it lies outside the realm of general relativity, in the domain of the so-called *laws of quantum*. We shall examine phenomena in very small regions of space-time after reviewing general relativity theory. At this stage note the characteristic dimensions of the spatial domain of quantum phenomena:

(i) quantum properties of particles (atomic phenomena) are manifest in regions with dimensions of the order of $\lesssim 10^{-8}$ cm;

(ii) quantum properties of electromagnetic fields are manifest at distances of the order of $\sim 10^{-10}$ cm;

(iii) specific nuclear and decay processes involving "elementary" particles are characterised by spatial domains of 10^{-13} cm and smaller and time intervals of 10^{-23} to 10^{-8} sec.

Running ahead, it will be noted that in the quantum domain at energies of the order of the "rest energy" of particles the boundaries between the concepts of particle and field begin to disappear and the concept of "quantised field" appears. Moreover, there are observed processes of reciprocal transformations of quanta of different fields. Higher energies enable the production of new particles to be observed and they help to penetrate into regions of even smaller distances between particles.¹

The physics of very small space-time regions is progressing extremely rapidly. It is natural to expect that in their development the ideas of "microphysics" must merge at some point with the ideas of "macrophysics", otherwise there will be a lack of harmony in the whole edifice of physical science.

¹ This explains the tremendous efforts devoted to particle acceleration techniques.

**Features
and properties
of gravitational
fields**

Getting back to the features and properties of the gravitational field equations, we must first note that they are *nonlinear*. Therefore the *principle of superposition* does not apply to gravitational fields, which sets them apart from electromagnetic fields in special relativity theory.

Furthermore, the structure of the gravitational field equations is such that they may contain characteristics describing any physical systems. It can be concluded from the very form of the field equations that gravitational fields emanate from all physical entities. Thus, gravitational fields are produced by particles as well as by electromagnetic fields and electromagnetic waves; more, a gravitational field may be produced by another gravitational field and also by gravitational waves, i. e., fields in space devoid of matter (particles or electromagnetic fields).

Thus, unlike electromagnetic fields, gravitational fields emanate from all physical entities, and these include particles and fields. Particles are characterised by mass, and fields, by energy. Common to both is energy, which is to say that any kind of energy is responsible for a gravitational field. It should be borne in mind, however, that the energy density of natural free radiation is very small in comparison with the energy density of material bodies, which possess rest energy. That is why the investigation of gravitational fields emanating from electromagnetic fields in the absence of mass presents no particular interest. Gravitational interaction becomes important only when bodies of appreciable mass are involved (due to the smallness of the gravitational constant¹). Therefore the gravitational field must perforce be investigated in connection with macroscopic bodies.

The gravity equations possess yet another remarkable feature: through a number of identical transformations they can be turned into the "equations of motion" of the physical system responsible for the gravitational field concerned. Thus, the gravitational field equations incorporate the equations of the matter (material particles or

¹ See footnote on p. 67.

electromagnetic field) which creates the field. In contradistinction, the electromagnetic field (Maxwell's equations) incorporate only the equation of conservation of total charge and not the equations of motion of the charges that produce the field.

Hence, in the case of an electromagnetic field the locations and motions of the charges are arbitrary as long as the total charge does not change. A statement of charge distribution (in terms of Maxwell's equations) in this case defines the field produced by the charges. In a gravitational field the configuration and motion of the matter responsible for the field cannot be stated arbitrarily; on the contrary, they must be determined together with the field.

It should be noted, however, that the gravitational field equations do not determine completely the distribution and motion of matter. Namely, they do not incorporate the equations of state, i. e., the equations connecting pressure and energy density, which must be stated alongside the field equations.¹

**The conservation
laws
in general
relativity theory**

Every physical process, as we know, takes place in space and time. Insofar as space and time possess certain properties, these properties, evidently, impose certain restrictions on possible processes involving physical entities. Namely, by virtue of the symmetry properties of space and time, processes involving any physical things must always fulfil the requirements of conservation of energy, momentum and angular momentum. In fact, energy, momentum and angular momentum are quantities which owe their existence to the most general properties of space and time. In other words, they are general physical concepts characterising things of any physical nature whatsoever.

By virtue of the finality of the velocity of propagation of interactions, field must be regarded as an independent system with its own "degrees of freedom". In other words, field of any kind is an independent physical entity. There-

¹ It is evident from this that the ideas and methods of general relativity theory are inadequate for cognition of such an entity as the universe; this requires further knowledge of the physics of "stellar and interstellar" states of matter.

fore all fields, gravitational included, are characterised by energy, momentum and angular momentum.

A few general remarks should be made before discussing the respective conservation laws for the gravitational field and their peculiarities.

Note that the equations of motion for any system of phenomena are obtained from a most general physical principle called the *principle of least action*. Its formulation includes a function, bearing the name of Lagrange, which can be specialised according to the most characteristic traits and features of a given sphere of phenomena. Note that the principle of least action itself, irrespective of the concrete form of the Lagrangian function, imposes a restriction on the totality of possible "motions" of a system. Other restrictions follow from the known symmetry properties of space and time and find expression in the respective conservation laws. The formal derivation of the latter involves both the corresponding properties of space-time and the "equations of motion". Thus, in effect, the "equations of motion" are included in the conservation laws.

This holds for physical laws in the absence of gravity. The presence of a gravitational field introduces certain "closing" features. Any attempt at adequately formulating the conservation laws for systems including gravitational fields leads to a peculiar feature: degeneration of the conservation laws into identities. On the other hand, as we have seen, the gravitational field equations embody the conservation laws for the matter (particles or electromagnetic field) responsible for the field.

The conclusion thus is that, unlike other physical entities, the gravitational field cannot be incorporated in a closed system or, in other words, it must be regarded as an "external condition" with respect to the system.

Thus, when gravitational field is taken into account, the structure of physical laws becomes such that essentially one can speak only of the field equations and not of the conservation laws. Nevertheless, in some cases application of the conservation laws yields some very interesting results. One of them is the equality of so-called *gravitational* and *inertial* masses. Gravitational mass is the mass which determines the gravity field produced by

a body, being the mass which enters into Newton's law of gravity. Inertial mass, on the other hand, determines the relationship between the momentum and velocity of a body; in particular, the rest energy of a body equals the inertial mass times c^2 .

**Consequences
of the field
equations**

The gravitational field equations in vacuum, i. e., in the absence of matter (represented by particles or electromagnetic field), have a nonzero solution, which, as in the case of electromagnetic fields, has the form of a running wave.¹ The velocity of propagation of *gravitational waves* is c .

Gravitational waves possess energy. In the investigated case of a weak gravitational field, which can exist in a restricted region of space, the energy can be determined uniquely. In this, and only this case, can one speak of the energy of gravitational waves.

A system of masses moving in a restricted region *radiates* gravitational waves. This is analogous to what was observed in discussing the radiation of electromagnetic waves. It should be noted that the numerical value of the loss of energy due to gravitational radiation is so small, even for objects of astronomical dimensions, that its effect on motion, even on a cosmic time scale, is negligible (the expression for the loss of energy through electromagnetic radiation contains the term $1/c^3$, whereas the expression for gravitational radiation contains $1/c^5$).

Let us investigate the limiting transition to nonrelativistic mechanics in the gravitational field equations. The assumption concerning the low velocities of all particles requires the gravitational field to be assumed weak, otherwise a particle in it would acquire a high velocity.

Note also that for gravitational fields the field equations are linear in the first approximation and consequently, the principle of superposition holds.

In this nonrelativistic approximation acting on a particle of mass m_1 in a gravitational field is a force

$$F = -\frac{Gmm_1}{r^2}, \quad (29)$$

¹ This result is obtained for a weak field approximation.

where m is the mass of the particle responsible for the field, r is the distance between particles m and m_1 , and G is a constant.¹

Equation (29) expresses Newton's famous *law of gravitation*. The minus is a reflection of the fact that gravitational forces cause only *attraction* between bodies. That the right-hand side of equation (29) is negative is seen from the consideration that the quantities in it are essentially positive, for mass can only be positive, and G is positive.

Insofar as Newton's nonrelativistic gravity theory deals essentially with weak fields and low velocities, its sphere of application is, evidently, not so great. In considering very large regions of the universe, in which fields cannot be assumed weak, Einstein's relativistic gravity theory comes into play.

Applications of general relativity

In speaking of the applications of general relativity theory and the effects predicted by it,² the phenomena observed when light travels in a gravitational field present the most vivid confirmation of the concepts of non-Euclidean space-time and hence, of the general theory of relativity.

Consider the path followed by a beam of light in a gravitational field. The stronger the field, the greater the "non-Euclidity" of space. As a result, a beam of light passing by bodies responsible for gravitational fields must be curved, for in non-Euclidean space, it will be recalled, a free particle (in this case a photon, a particle with zero mass) follows a curved path.

Consider further the change in the frequency of light propagating in a gravitational field. We have seen that time intervals depend on the field, viz., the stronger the field, the slower the rhythm of a clock placed in it, and vice versa. In other words, the passage of time is differ-

¹ G is called the *gravitational constant*. It is a universal constant numerically equal to $G = 6.67 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ sec}^{-2}$.

² An interesting aside: Einstein developed the general theory of relativity by purely deductive reasoning; only subsequently was it confirmed by astronomical observations.

ent at different points of space. Evidently, then, the frequency is the number of oscillations of a wave per unit time. When a wave recedes from a body producing a gravitational field the frequency decreases, when a wave approaches such a body the frequency increases.

General relativity has also been used to predict planetary motions. Insofar as the speed of a planet is small in comparison with c , relativistic gravity theory introduces very minor corrections to planetary orbits computed according to Newtonian theory. Essentially the difference lies in that, whereas in the Newtonian two-body problem the orbits are ellipses fixed in space, Einstein's theory leads to the conclusion that they undergo a very slow precession in their own planes.

**The universe
and general
relativity**

As the above-mentioned phenomena are investigated from the viewpoint of general relativity in comparatively small regions of space-time, in which fields are not very strong, their effects are small. In investigating the structure and evolution of the universe on a grand scale (the cosmological problem) general relativity plays a dominant part.

The very first applications of Einstein's theory to this problem yielded very interesting and important results. Investigation of the field equations of general relativity has led to the conclusion that, on a large scale, the geometrical properties of the universe must be changing with time. As predicted by theory and confirmed by observations, this change is of an "expanding" nature.

Further progress in this sphere is encountering difficulties which are due both to insufficient astronomical data and to formidable mathematical obstacles in the general investigation of Einstein's gravitational field equations.

To continue our discussion of space-time properties of the world as a whole, we must make a few more remarks based on applications of the conservation laws.

Space homogeneity, a consequence of which is the *law of conservation of linear momentum*, means that the properties of a physical system do not depend on its location in space. In other words, the properties of isolated phys-

ical systems (e. g., atoms, elementary particles) are the same here on earth and on distant celestial bodies.¹

Space isotropy, a consequence of which is the *law of conservation of angular momentum*, means that the properties of a physical system do not depend on the direction in which that system is moved. In other words, any direction in the universe is as good as the other.

Time homogeneity, a consequence of which is the *law of conservation of energy*, means that the properties of isolated physical systems do not depend on the moment of time at which they are investigated. In other words, the properties of isolated physical systems are the same today as they were millions of years ago and they will be the same millions of years hence.

These propositions, however, are to some extent conditioned by the fact that in investigating large regions of the universe gravitational fields come increasingly into play and this, as we have seen, means a change in the properties of space-time. This, evidently, in various ways affects the conservation laws, which are expressions of certain properties of space-time. Specifically, as we know, a gravitational field cannot be included in a closed system and must be regarded as "something external" with respect to it.

On the other hand, we have found that the solution of the gravitational field equations in the framework of general relativity for vast spatial scales leads to the conclusion that gravitational fields are time-dependent.

Thus, in general relativity the world as a whole must be treated not as a closed system but as a system in a variable gravitational field.

Time isotropy (or reversibility of time) which is considered in nonquantum physics, has no place, we shall find later, in quantum mechanics. This nonequivalence

¹ For example, the line spectra emitted by atoms (and we judge of the properties of atoms according to their spectra) on the sun are precisely like the spectra emitted by the same kinds of atoms on earth. This is not to be confused with the observed emission spectra of atoms on the sun, when account must be taken of the effects of the sun's gravity field on the propagation of light: the so-called "red shift", which is a displacement of the spectrum lines as compared with the lines of the same emission spectrum on earth.

of the two directions in time is manifest in connection with the fundamental process of quantum mechanics, that of interaction between a quantum system and a system governed, with sufficient accuracy, by classical mechanics.

Absence of isotropic symmetry of time indicates the direction of processes in time. This direction is such that the entropy (concerning which see Chapter 8) of the world as a whole is increasing steadily without, however, tending to any limit, which follows from the concept of the nonstationary universe.

Chapter 5. Quantum Mechanics (Nonrelativistic Theory)

Our study of the properties of motion now takes us from macrospace to microspace. In the infinitesimal confines of the microcosmos only *electromagnetic interactions* are of any consequence. In the world of elementary particles the ratio of the intensity of gravitational to electromagnetic interaction, Gm^2/e^2 , is of the fantastically small order of 10^{-40} .

We will first restrict our investigation to *particles* of very small, but not zero, mass. Such particles exist in the nonrelativistic domain of motion, when $v \ll c$, which is the one we shall deal with in this chapter.

Electromagnetic fields (and the particles of zero mass to which they are correlated) are, of course, relativistic entities and from the outset they demand relativistic investigation.

Atomic phenomena When classical mechanics and electrodynamics are used to explain *atomic phenomena*, which involve particles of very small mass and occupy very small regions of space, the results are found to clash with experience. This is demonstrated most vividly by the contradiction inherent in the conventional electrodynamic model of the atom with electrons circling the nucleus on "classical" orbits. Since the electrons represent charges moving with acceleration, one would expect them to emit a steady stream of electromagnetic waves. This in turn would lead to a dissipation of energy which would ultimately result in their falling on the nucleus. Thus the classical electrodynamic atomic

model turns out to be an unstable arrangement. This, of course, is not the case.

Such a contradiction between theory and experiment is an indication that a workable theory applicable to atomic phenomena requires a fundamental revision of basic classical concepts and laws.

Diffraction of particles

A convenient point of departure for such a revision is the experimentally observed phenomenon of *diffraction of particles*. It was found that when a homogeneous beam of particles is passed through a crystal the emerging beam contains an alternation of maxima and minima of intensities similar to the diffraction patterns of electromagnetic waves. Thus, under suitable conditions material particles may behave like waves.

How greatly this contradicts conventional notions of motion is seen from the following. Imagine a beam of particles falling on an opaque screen with two slits cut in it. The density of the particles in the beam is low enough to neglect interactions between them. When we close one slit we observe an intensity pattern on a screen behind the slit. A similar pattern is observed when the beam is passed through the other slit. Now on the basis of conventional notions we should expect the pattern produced by the beam passing through both slits at once to be a simple superimposition of the original two: every particle moving in its path passes through one of the slits without in any way affecting the particles passing through the other. Actually, though, a diffraction pattern is observed which, due to interference, is not merely the sum of the patterns produced through each slit separately. Obviously, this result cannot be reconciled with the concept of particles moving in a path. On the atomic scale the idea of path or trajectory does not exist.

Wave-corpuscule duality

In fact, in investigating the motion of subatomic particles we observe the phenomenon of *wave-corpuscule duality*.

It is hardly surprising that entirely new principles govern the properties of motion on the atomic level. Science deals only with observable things. But a thing can be observed only if it is made to interact with something external with respect to it.

**Observation
disturbance.
The quantum
of action**

This means that any observation involves a disturbance, or perturbation, of the observed thing. An observable is treated as a *classical-mechanical* system when the disturbance caused by the observation can be neglected and its behaviour can be described by classical mechanics. If, on the other hand, there exists a certain "limiting perturbation", a certain "quantum of action"¹ which cannot be neglected, we have a *quantum-mechanical* system, and a new theory is needed to describe it.

**Statistical
probability**

Thus, a quantum-mechanical system cannot be observed without causing major perturbations in it and, consequently, one cannot expect complete determinism of the results of observations: uncertainty and statistical relations come into play. Quantum theory, therefore, must allow for a mathematical prediction, not of the results of observations, but of the *probability* of obtaining one or another result of measurement.

Let us elaborate this. If an observation is carried out on a quantum-mechanical system in a given state the results are indeterminate; in other words, one and the same experiment repeated several times in identical conditions may lead to several different results. If, however, the experiment is repeated a sufficient number of times, every calculated result will be obtained in a certain ratio to the total number of tests, i. e., there exists a certain probability that a specific result will be obtained. Theory enables this probability to be calculated. A unique experimental result is possible only in the special case when the probability of a result is unity.

Measurement

It follows from the foregoing that in order to study the properties of quantum-mechanical objects they must be brought into interaction with classical-mechanical objects since, as we have seen, only classical objects can be said with certainty to be, at any given instant, in some known state with

¹ Represented by *Planck's constant*, which has the value $h = 6.62 \times 10^{-27}$ erg-sec.

specific values of the physical quantities characterising it. With quantum-mechanical objects this is impossible.

If, now, a quantum-mechanical object is made to interact with a classical object, the latter's state will change. The character and degree of the change depend on the state of the quantum-mechanical object, and can therefore serve as a quantitative characteristic of the latter.

Accordingly, the classical object is the *instrument*, and the process of its interaction with the quantum-mechanical object is the *measurement*.

We have defined an instrument as being a physical object which is subject to classical mechanics to a sufficient degree of accuracy. Such, for instance, is a body of sufficient mass. It should not be assumed, however, that macroscopic qualities are absolutely essential for an instrument. In some conditions a purely microscopic object can be used as an instrument, insofar as the notion "with sufficient accuracy" depends on the problem under consideration.

Thus, the flight of an electron through a Wilson cloud chamber is observed by a track of water vapour much thicker than atomic dimensions; with such a degree of accuracy in determining its path, the electron can be treated as a classical-mechanical object.

Relationship between quantum and classical mechanics

The above reasoning points, first of all, to the special nature of the relationships between quantum theory and classical mechanics. Usually a general theory can be formulated in a logically consistent manner independent of the "not-so-general" theory representing its limiting case. Thus, relativistic mechanics can be developed out of its fundamental principles without any reference to Newtonian mechanics. Quantum theory, however, cannot in principle be substantiated without falling back upon classical mechanics. It thus occupies a unique position among physical theories: it includes classical mechanics as a *limiting case* and at the same time it depends on this limiting case for *its own substantiation*.

A few remarks The foregoing shows that the mechanics of atomic phenomena—quantum mechanics, that is—must be based on concepts of motion which differ in principle from the concepts of classical mechanics. The paradoxical properties of motion of subatomic particles mentioned here point not merely to the inaccuracy of the classical laws of motion but to the unsuitability of the basic concepts of classical mechanics for describing atomic phenomena.

The characteristic features of the new theory mentioned before must serve as the basis for developing its mathematical apparatus.

Quantum mechanics, as we shall see later on, generally *restricts*, in comparison with classical mechanics, the set of values which can be assigned to various physical quantities (such as energy), i. e., values that can be detected by measuring the quantity concerned. The methods of quantum mechanics must provide the means of determining those “allowed” values.

The wave function Furthermore, in quantum mechanics we find that some physical quantities cannot be measured *simultaneously*, i. e., they cannot possess specific values at the same time. The special features of atomic phenomena are reflected in the mathematical formalism of quantum mechanics. It is based on the consideration that every state of a quantum-mechanical system can be described at a given instant by a specific (generally complex) coordinate function Ψ , whose square of the modulus, $|\Psi|^2$, gives the probability distribution of the coordinates of the system. It is called the *wave function* of the system.

When we have a compound system consisting of many particles the wave function depends on all the coordinates of the system; it is the function of a point, not in real physical space, but in a *multidimensional configuration space*. It follows from these properties of the wave function that it cannot be interpreted as a kind of field pervading space like electromagnetic and other fields.

Although the function cannot be measured experimentally, it can be constructed *only* as a result of measurement. The value $|\Psi|^2$ or, what is the same thing, $\Psi\Psi^*$, has definite physical meaning.

The principle of superposition

Fundamental in the description of quantised states is the *principle of superposition*, which postulates a number of properties of the wave function. Suppose that a measurement made in a state characterised by the wave function Ψ_1 yields with certainty result 1, and a measurement made in state Ψ_2 yields result 2. It is accepted then that any linear combination of Ψ_1 and Ψ_2 , i. e., any Ψ function of the form $c_1\Psi_1 + c_2\Psi_2$ (where c_1 and c_2 are constants) gives a state in which the same measurement yields *either* result 1 *or* result 2.

The principle of superposition of states is the *fundamental principle of quantum mechanics* insofar as essentially it embodies all the specific features of atomic phenomena mentioned before. In particular, it follows from the superposition principle that in quantum mechanics the sum of two probabilities is not equal to the common value of the two, but there exists an interference of probabilities:

$$|\Psi|^2 = |c_1|^2 |\Psi_1|^2 + |c_2|^2 |\Psi_2|^2 + (c_1^* c_2 \Psi_1^* \Psi_2 + c_1 c_2^* \Psi_1 \Psi_2^*).$$

The principle of superposition thus takes into account both the wave-corpuscule properties of quantum-mechanical objects and the statistical probability of the results of observations.

Linearity of equations

It follows directly from the principle of superposition that all equations which are satisfied by wave functions must be linear ¹ with respect to Ψ . These equations will be derived further on, but some remarks are called for here.

Remarks

Knowledge of the wave function Ψ makes it possible in principle to calculate the probability of results for any other measurements (not only coordinate measurements). All the probabilities will be determined by expressions ² in which Ψ enters multiplied by Ψ^* .

¹ Generally speaking, the state of a quantum-mechanical system, and with it the wave function, will change with time. In this sense the wave function may be regarded as being also a function of time.

² Such an expression should, of course, include a function depending on the type of measurement and its result.

When a quantity having a specific physical meaning is calculated by means of the wave function it always includes the product of Ψ multiplied by Ψ^* . It follows from this that the normalised wave function¹ is determinate only to the accuracy of a constant "phase" multiple of the form $e^{i\alpha}$ (where α is any real number) the modulus of which is unity. This indeterminacy is intrinsic and cannot be eliminated but, on the other hand, it is of no consequence as it has no bearing whatsoever on physical results.

Operators In quantum mechanics any observable is represented by a mathematical object called an *operator*. The values which a given observable L can assume in conditions compatible with the observation are called the *eigenvalues* ("self" or "proper" values) of the quantum-mechanical operator \hat{L} . The *eigenfunction* Ψ_l of this operator corresponding to an eigenvalue l describes the state of the system in which the given quantity has a definite value l .

Range of eigenvalues The values which a given physical quantity may assume are called its eigenvalues. The sum total of possible eigenvalues of a quantity constitutes the *range of eigenvalues*. In classical mechanics, all physical quantities can generally pass through a continuous range of values. In quantum mechanics, too, there are quantities which may pass through a continuous range of eigenvalues (coordinates, for example). In this case we speak of a *continuous range*, as distinct from a discrete, or point, range when the eigenvalues make up a discontinuous series.

Simultaneous measurement It is interesting to note that there exists a formal similarity between the basic relationships of quantum mechanics in operator form and the corresponding relationships of classical mechanics. There is, however, an essential difference between the algebra of operators and the algebra of numbers: the former is, generally, characterised by noncommutativeness of multiplication.

¹ The condition for normalisation of wave functions is that the sum of the probabilities of all possible values of the coordinates of a system must be unity. This condition is a statement of fact that the system is present in space.

A physical expression of the latter statement is the impossibility of measuring various quantum-mechanical quantities *simultaneously*, i. e., they cannot have definite values at the same instant. If the specific values of two quantities L and M are known at a certain instant, this means, according to the mathematical formalism of quantum mechanics, that a given state is described by a wave function which is the *common eigenfunction* of the two operators \hat{L} and \hat{M} corresponding to these quantities. This in turn means that the operators \hat{L} and \hat{M} *commute*, i. e., their successive application (multiplication) yields a result that does not depend on the order in which they were applied: $\hat{L}\hat{M} = \hat{M}\hat{L}$. This symbolic equality may also be written in the form

$$\hat{L}\hat{M} - \hat{M}\hat{L} = 0.$$

Thus we arrive at an important conclusion: if two quantities, L and M , can simultaneously take definite values, their operators commute.

The uncertainty principle

If, on the other hand, $\hat{L}\hat{M} - \hat{M}\hat{L} \neq 0$ there exists no state in which both quantities L and M are simultaneously determinate. In this case we say that the quantities L and M are characterised by uncertainty relations, which express an important feature of quantum-mechanical objects. Let us consider, in particular, the commutation rules for momentum and coordinate operators. They have the form:

$$\begin{aligned} \hat{p}_x \hat{x} - \hat{x} \hat{p}_x &= -i\hbar, \\ \hat{p}_x \hat{y} - \hat{y} \hat{p}_x &= 0, \\ \hat{p}_x \hat{z} - \hat{z} \hat{p}_x &= 0. \end{aligned} \tag{30}$$

Similar relationships hold for other components of the momentum and coordinate operators.

The equations (30) show that the coordinate of a particle along one of the axes may have a certain value simultaneously with the momentum components along the two other axes; the coordinate and momentum component along the same axis, however, cannot be specified simul-

taneously. In particular, a particle cannot be located at a specific point in space and at the same time have a specific momentum \mathbf{p} .

A corollary of the commutation rules are the relationships

$$\begin{aligned}\Delta p_x \Delta x &\sim h, \\ \Delta p_y \Delta y &\sim h, \\ \Delta p_z \Delta z &\sim h,\end{aligned}\tag{31}$$

which express the *uncertainty principle*.

It is seen from equations (31) that the more exactly we know the position coordinate of a particle (i. e., the smaller Δx), the higher the uncertainty in the value of the momentum component Δp_x along the axis, and vice versa. In particular, if a particle is localised at some precisely specified point in space ($\Delta x = \Delta y = \Delta z = 0$), then $\Delta p_x = \Delta p_y = \Delta p_z = \infty$. This means that all momentum values are equally probable. Conversely, if a particle has a precisely specified momentum \mathbf{p} , any position in space is equally probable.

It follows from this, in particular, that the state of a free particle (its wave function) is completely given by either its momentum vector \mathbf{p} *alone* or its radius vector \mathbf{r} *alone*. In classical mechanics, as we know, the state of a particle is given by the simultaneous statement of its momentum \mathbf{p} and position coordinate \mathbf{r} .

Uncertainty relations exist for other quantities as well. In particular, for energy and time (in a system with infinitely weak interactions with external objects) we have

$$\Delta(E - E') \sim \frac{h}{\Delta t}.\tag{32}$$

The expression (32) is the *uncertainty relationship for energy*. Its physical meaning, however, is quite unlike that of the uncertainty relationship $\Delta p \Delta x \sim h$ for a position coordinate and momentum. In the latter, Δp and Δx are uncertainties in the precise values of the momentum and coordinate at the same instant; they indicate that the two cannot be precisely defined at the same time. In contrast, the energies E and E' of a given system can be measured at a given time with any accuracy. The quantity

$E - E'$ in (32) is the difference between two precisely measured values at two different moments of time; it is not an uncertainty in the energy value at some given instant; Δt is the time lapse between the two measurements.

The wave equation As mentioned before, the wave function Ψ completely defines the state of a physical system in quantum mechanics. This means that statement of the function at any instant not only describes the properties of the system at the instant but also predicts its behaviour at any subsequent moment of time, to the extent, of course, which is generally achievable in quantum mechanics. In mathematical formalism this means that the quantity $\frac{\partial \Psi}{\partial t}$ must be determined at any given instant by the Ψ function itself at that instant and, according to the superposition principle, the relationship must be linear. In the most general case we may write

$$i \frac{\partial \Psi}{\partial t} = \hat{L} \Psi,$$

where \hat{L} is a linear operator and i is a multiple introduced for mathematical convenience.

Investigation of the properties of the operator \hat{L} reveals that, in the absence of time-dependent fields, it is an *energy operator*. This operator, which we shall denote by the symbol \hat{H} , is called the *Hamiltonian of the system*. The relationship between $\frac{\partial \Psi}{\partial t}$ and Ψ takes the form

$$i h \frac{\partial \Psi}{\partial t} = \hat{H} \Psi. \quad (33)$$

If the form of the Hamiltonian is known, the equation (33) gives the wave functions of a given physical system. This basic equation of quantum mechanics is called the *wave equation*.

The Hamiltonian Let us now determine the form of the Hamiltonian, a problem of primary importance, since it determines the form of the wave equation. We shall begin with considering a free particle, i. e., one that is not in any external field. By virtue of space homogeneity, the Hamiltonian of the particle can-

not contain the coordinates explicitly and can depend only on the momentum operator. Furthermore, as the total energy and momentum of a free particle remain unchanged both quantities can, apparently, exist simultaneously. Inasmuch as the momentum vector completely defines the state of a particle, the eigenvalues of the energy E must be expressed as functions of the momentum in the same state. E being a function only of the absolute value of the momentum, and not of its direction. This follows from the isotropy of space with respect to a free particle, i. e., the absence of any privileged direction in it. The very form of the function $E(p)$ is completely determined by the requirements of the Galilean relativity principle, which must be fulfilled in nonrelativistic quantum mechanics in just the same way as in classical (nonrelativistic) mechanics. In classical mechanics this requirement leads to a quadratic function of energy versus momentum:¹

$$E = p^2/2m.$$

For this relationship to appear for all eigenvalues of the energy and momentum, a similar relationship must hold for their operators:

$$\hat{H} = \frac{1}{2m} (\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2). \quad (34)$$

This is the Hamiltonian function of the free moving particle.

The Hamiltonian of one particle located in an external field is

$$\hat{H} = \frac{p^2}{2m} + U(x, y, z), \quad (35)$$

where $U(x, y, z)$ is the potential energy of the particle in the external field. The first term can be regarded as the kinetic energy operator, the second, as the potential energy operator, the latter being reduced to a simple multiplication by the function U . Note that the eigenvalues of the kinetic energy operator are always positive,

¹ This formula is easily arrived at by applying the free-particle expressions for energy, $mv^2/2$, and velocity, $v = p/m$.

as the operator is equal to the sum of the squares of the operators of the momentum components taken with positive factors.

Schrödinger's equation

Substitution of the Hamiltonian expressions (34) and (35) into the general equation (33) yields the wave equations for the corresponding systems. These equations were derived by Schrödinger and bear his name.

Properties of solutions

Let us examine the basic properties of Schrödinger's equations. First of all we should note that the conditions which satisfy their solutions are of a most general nature: the wave function must be unique, continuous and finite over all space.

Free particle

For a *free particle* the equation has finite solutions in all space at any *positive* value of the energy E (including zero). The energy range of a free moving particle is thus *continuous* from zero to $+\infty$.

For the solutions of Schrödinger's equation for a free particle there can be taken the general eigenfunctions of the operators of the three momentum components. The wave functions for *stationary states*, i. e., states in which energy has specific values, will then have the form

$$\Psi = \text{const} \times e^{i\left(\frac{p}{\hbar}r - \frac{E}{\hbar}t\right)}, \quad (36)$$

where $E = \frac{p^2}{2m}$. Every such function describes a state in which the particle possesses a specific energy E and momentum p . A purely formal comparison of the function (36) with equation (25) reveals that it is a plane wave of frequency E/\hbar and wavelength $2\pi\hbar/p$ propagating in the p direction. It is called the *de Broglie wavelength of a particle*. Thus we have the following basic particle relationships in quantum mechanics:

$$p = \hbar k; \quad E = \hbar\omega, \quad \lambda \cong \frac{\hbar}{p}. \quad (37)$$

Schrödinger's equation for a *constrained particle*, i. e., one which cannot move to infinity from the centres of at-

Constrained particle traction, has solutions satisfying the above conditions of finality, but only with specific negative energy values. The energy range of a constrained particle is a discontinuous one.

“Zero” energy The least kinetic energy eigenvalue in a discrete range is never zero. Thus, in quantum mechanics the energy of motion of a constrained particle cannot be zero (even at absolute zero temperature).

The “tunnel” effect It also follows from the solutions of Schrödinger’s equation that in quantum mechanics a constrained particle may be located in regions of space where the potential energy U is greater than the total energy E of the particle. The probability $|\Psi|^2$ of a particle penetrating into such a region, though it rapidly tends to zero with the depth of penetration into the region increasing, is not zero at finite distances.¹ In this respect there is a difference of principle from classical mechanics, in which a particle can never penetrate into a region where $U > E$. The reason is that at $E < U$ the kinetic energy would be negative, hence the velocity would be imaginary, which is absurd. In quantum mechanics the kinetic energy eigenvalues are also positive. Nevertheless there is no contradiction since, if in the process of measurement a particle is localised at some point in space, the observation disturbs the state of the particle so that it ceases to possess definite kinetic energy.

Limiting transition to classical mechanics Since quantum mechanics includes classical mechanics as a limiting case the question arises of the way in which the limiting transition is effected. In quantum mechanics a particle is described by a wave function giving the various values of the position coordinates; this function represents a solution of a linear differential equation (Schrödinger’s equation). In classical mechanics a particle is regarded as moving in a trajectory which is fully determined by the equations of motion. A relationship *in some respects*

¹ If such a region of space is very small there is always the probability of the particle “tunnelling through” it.

analogous to that existing between quantum and classical mechanics is found between wave and geometrical optics in electrodynamics. In wave optics electromagnetic waves are described by electric and magnetic field vectors satisfying a definite set of simultaneous linear differential equations (Maxwell's equations). Geometrical optics, on the other hand, investigates the propagation of light along trajectories or rays. The analogy leads to the conclusion that the limiting transition from quantum to classical mechanics is similar to the transition from wave to geometrical optics. In the latter case the transition corresponds to going over to a zero wavelength limit $\lambda \rightarrow 0$. In all quantum-mechanical phenomena, however, there is always present the quantum constant h . It follows, then, that the transition to classical mechanics can be formally described by going over to the limit $h \rightarrow 0$.

Investigation of the limiting case of Schrödinger's equation reveals that in the most general case motion described by a wave function does not turn into motion on a definite trajectory. The link with classical motion lies in that if, at some initial time, the wave function, and with it the position coordinate probabilities, are given, the probabilities will undergo a "displacement" according to the laws of classical mechanics.

And finally, quantum-mechanical operators reduce in the limit to a plain multiplication of Ψ by the corresponding physical quantity. This follows from the fact that quantum-mechanical relationships in operator form look just like the corresponding classical relationships.

Measurement again Speaking of Schrödinger's equation, it is necessary to remember the fundamental role played in quantum mechanics by the process of interaction between a quantum-mechanical and classical-mechanical object which is conventionally called *measurement*. It should be stressed, however, that this process does not presume the existence of an observer. In the present case such a connection between the laws of nature and the properties of the observer is not valid. Let us examine the situation in quantum mechanics.

Consider a system consisting of quantum-mechanical and classical objects. Suppose that before an interaction

the quantum-mechanical object was described by some arbitrary normalised wave function. As a result of an interaction with the classical object the latter goes over from its initial to some other state, from which we judge of the state of the quantum-mechanical object.

The properties of interaction between a quantum-mechanical object and classical conditions¹ are such that after the interaction the quantum-mechanical object assumes a state described by a specific wave function. Only such states of quantum-mechanical objects enable us to judge of their properties. Thus the process of interaction between a quantum-mechanical object and a system subject, to a sufficient degree of accuracy, to classical mechanics is *fundamental* to quantum mechanics since only in consequence of it can the properties of quantum-mechanical objects be described.

Causality One general remark is called for in connection with Schrödinger's equation. A prerequisite for the concepts of *cause* and *effect* is the existence of the notions *before* and *after*. The "equations of motion" for various systems of physical phenomena provide the temporal relationships of the states of physical entities. If the "external conditions" ("cause") are stated and the initial state of a thing is known, the "equations of motion" uniquely define the state of the thing at any subsequent moment of time ("effect"). In classical, relativistic and quantum mechanics the external conditions were defined by fields; in field theories (electromagnetic and gravitational) the external conditions are stated in terms of corresponding "charge configurations" (subject to the specific features of the gravitational field). The state of an observable at a given instant is determined in classical and relativistic mechanics by its position coordinates and momentum; in electromagnetic field theory it is defined by the values (over all space) of electric- and magnetic-field vectors; in gravitational field theory, by the value of the metric tensor; in quantum mechanics, by the wave function.

¹ Irrespective of whether these conditions are created by the observer or they are natural conditions existing besides and independently of any observer.

The "equation of motion" of quantum mechanics—Schrödinger's equation—uniquely defines the dependence of the wave function of state on time. If we know the initial state $\Psi(x)$ of a quantum-mechanical system created by a measurement, the state of that system $\Psi(x, t)$ at any instant t will be found uniquely from the wave equation. If we decide to measure some quantity in the state $\Psi(x, t)$, the result will be indeterminate, the probability of the values being "smeared out". At the same time, all the probabilities can be precomputed, and they are determined by expressions including the function $\Psi(x, t)$ and a function depending on the nature and result of the measurement.

In physics two independent principles can be said to hold: the *principle of causality*, which is expressed by the "equations of motion", and the principle according to which any physical entity displays its properties *only* in interaction with external classical conditions. For, as we have seen, though the "equations of motion" presume a knowledge of the initial state of a system, the manner in which the information concerning that state was obtained, the manner in which it was described, is quite another and unrelated question.

To determine the state of a physical system it must interact with external classical things.¹ Such interaction is accompanied by disturbances of the system itself: in nonquantum physics they are neglected, in quantum physics they must be taken into account. As a consequence of this the interaction of a system with external classical objects is uniquely described in nonquantum physics and is indeterminate in quantum physics. The outcome of the investigation is the determination of the state of a system. The subsequent fate of the created condition is linked with the "equations of motion".

¹ This was the case in classical and relativistic mechanics and in classical electrodynamics, where a field is characterised by its action on charged "test bodies". This was the case in gravitational field theory and quantum mechanics. This is the case, as we shall see later on, in quantum electrodynamics, where measurement of field components requires a determination of the momentum of a charged test body (with finite classical charges and current distributions acting as the test bodies).

Irreversibility of time

All equations of nonquantum physics are known to be invariant with respect to any substitution of past for future.

Neither is Schrödinger's equation altered by a substitution of $-t$ for t in it;¹ nevertheless, quantum mechanics essentially includes the nonequivalence of the two directions in time. It is manifest in connection with the fundamental process of interaction between a quantum-mechanical object and a system subject, with sufficient accuracy, to classical mechanics.

This process, we know, always affects the quantum-mechanical object, changing its state and, accordingly, its wave function, thereby destroying, as it were, the whole of the past (prior to the interaction). We have seen that if a quantum-mechanical object is subjected to two successive interactions A and B , the effects of process A can be used to determine the probability of the effects of process B , regardless of what was happening to the object before process A began. This demonstrates the dual nature of measurement in quantum mechanics: its role is different with respect to the past and future of an object. Thus quantum mechanics does not display complete physical symmetry in both directions of time.

The conservation laws in quantum mechanics

We have seen that the most general symmetry properties of space and time lead to the conclusion that all physical entities, quantum-mechanical systems

included, must be characterised by such quantities as energy, momentum and angular momentum. Moreover, in closed systems these quantities must be conserved. In quantum mechanics the laws of conservation are expressed in operator form, with all the consequences therefrom.

From time homogeneity with respect to a system arises the energy conservation principle. Its quantum-mechanical peculiarity is that it can be verified by means of two measurements only to an accuracy of the order of $h/\Delta t$, where Δt is the time interval between the two measurements [see equation (32)]. The energy ranges of free and constrained particles were discussed before.

¹ Provided the wave function Ψ is at the same time replaced by Ψ^* .

From space homogeneity with respect to a system arises the momentum conservation principle. Naturally, the range of possible eigenvalues is considered only for free particles.

Insofar as quantum mechanics considers phenomena restricted to very small regions of space, and a free moving particle may travel into infinity, it is natural to expect that the range of energy and momentum eigenvalues of a free moving particle should not differ from the possible values in classical mechanics. In fact, we found that the energy range of a free particle is continuous from 0 to $+\infty$ (negative values, which are discrete, correspond to constrained particle states). It is now evident that the range of eigenvalues of the momentum components is also continuous and extends from $-\infty$ to $+\infty$.

From space isotropy arises the angular momentum conservation principle. Its quantum-mechanical peculiarities consist in that the three angular momentum components do not commute and hence they cannot have definite values at the same time (except when they are all simultaneously zero). In this respect angular momentum differs essentially from linear momentum, the three components of which may have specific values at the same time. Thus a quantum-mechanical particle does not possess an angular-momentum vector. Specific values can be simultaneously assigned to the absolute value of the angular momentum and to one of its components.

Note that, as a rule, angular momentum is considered only in connection with constrained systems. The possible eigenvalues of angular momentum are discontinuous (both for absolute value and the component in any direction). Neither does angular momentum commute with linear momentum, i. e., the two quantities cannot be assigned specific values at the same time.

Space homogeneity and isotropy lead, in turn, to complete *space symmetry with respect to left and right*, a corollary of which is the invariance of the laws of nature with regard to mirror reflection, or inversion. In non-quantum physics invariance with respect to inversion does not lead to any conservation laws. In quantum physics, however, it leads to the *law of conservation of parity*.

Its substance is as follows. An inversion of a wave function results in the function either not changing at all or in a reversal of its sign. In the former case the wave function (and the corresponding state) is *even*, in the latter it is *odd*. In other words, if the state of a closed system has parity (i. e., if it is even or odd), its parity does not change with time.

Together with the other conservation laws, the parity conservation principle regulates the possibility of the decay or formation of compound systems. Thus, if internal forces break up a closed system into several parts its angular momentum and parity do not change. This may make it impossible for a system to disintegrate even if the energy conditions for disintegration are there.

The two-particle problem Consider the motion of a system of two interacting particles. To begin with, note that statement of the Hamiltonian defines the properties of a quantum-mechanical system, insofar as the Hamiltonian determines the form of the wave equation from which the wave functions of the system are found.¹

In our problem the Hamiltonian is time-independent and coincides with the energy operator. The latter commutes with the momentum, angular momentum and inversion operators, which means that they may have common wave functions. This imposes certain restrictions on solutions of the wave equation, and the wave function of the system is completely defined by statement of the energy, momentum and angular momentum (the parity is found from the angular momentum of the system).

Furthermore, since the motion of the system can be represented as the sum of two independent motions (a translation of the system as a whole and the relative motions of the component particles) the wave function of the system will consist of two independent components: the wave function of free motion and the wave function of constrained motion. The former, we know, is complete-

¹ It is precisely the Hamiltonian (for closed systems) that displays invariance with regard to the three mentioned transformations of a coordinate system: parallel translation, rotation and inversion.

ly defined by the momentum vector, hence the latter is completely defined by a simultaneous statement of energy and angular momentum (the absolute value of the angular momentum and its component in an arbitrary direction).

“Internal” degrees of freedom We see that the maximum number of physical quantities completely defining the state of our quantum-mechanical system and simultaneously having specific values is equal to the number of degrees of freedom of the system. Our quantum-mechanical system also possesses such characteristics as parity and so-called *spin*. Both are typically quantum-mechanical parameters which vanish in going over to classical mechanics. Thus, quantum-mechanical objects possess specific “internal” degrees of freedom.

Spin Elementary particles, it was found, must be endowed with intrinsic angular momentum which is in no way associated with its motion through space. Known as *spin*, it is to be distinguished from the angular momentum associated with motion in space, the *orbital angular momentum*.¹

Since an elementary particle must be treated as a point it is quite meaningless to picture its intrinsic angular momentum as being produced by a rotation about its axis. This goes to show that spin is a purely quantum-mechanical parameter which does not allow of any “classical” interpretation.

In the foregoing discourse we have always assumed the wave function of a particle to be a function of its coordinates. With the introduction of the notion of spin the wave function must be a function of both the coordinates and the spin variable, which gives the value of the projection of the spin on a chosen direction in space and passes through a limited number of discrete values. These values may be assigned to wave functions as indices. Thus the wave function of a particle whose spin is not zero actually represents not one but a set of several coordinate functions with different spin indices.

¹ We shall see later on that particles are also characterised by an intrinsic parity. The latter is taken into account in the production and annihilation of particles.

Systems of like particles We shall see later on that quantised systems consisting of like particles possess certain features that are without analogy in classical systems. In classical mechanics, identical particles do not lose their individuality despite the likeness of their physical properties. One can imagine that the particles comprising a given physical system have been numbered at some instant and the motion of each can be traced along their respective paths, thus enabling identification at any instant.

In quantum mechanics the uncertainty principle introduces an entirely new state of affairs. In fact, the concept of path or trajectory becomes meaningless. For velocity, like momentum, cannot be assigned a definite value simultaneously with the position coordinates of a particle. But velocity multiplied by a differential time element Δt determines the displacement of the particle in the time Δt . Hence, the absence of velocity simultaneously with the coordinates means that, if a particle is precisely located at a given instant, the next instant its coordinates will have no definite value at all. Which means that having localised and numbered our particles at some moment of time we shall still be unable to identify them at any subsequent instant; having localised a particle at some point of space at a later instant, we can never say which of the particles has in fact arrived at that point.

Thus, in quantum mechanics, it is in principle impossible to track any of a number of identical particles, and hence to distinguish them. Identity of physical properties leads in the case of quantum-mechanical systems to the impossibility of defining any individual particle.

This so-called *principle of indistinguishability of identical particles* plays a fundamental part in the quantum-mechanical investigation of systems of like particles.

Consider a system of two particles. Owing to their identity, the states resulting from a simple interchange of the particles should be physically equivalent. That is to say, the wave function of the system may change by no more than a negligible phase multiple. Let $\Psi(\xi_1, \xi_2)$ be the wave function of the system, ξ_1, ξ_2 denoting the

totality of the three position coordinates and the spin projections of each particle.¹ Then we should have

$$\psi(\xi_1, \xi_2) = e^{i\alpha} \psi(\xi_2, \xi_1),$$

where α is a real constant. A second interchange brings us back to the initial state, but the Ψ function is now multiplied by $e^{2i\alpha}$. It follows, then, that $e^{2i\alpha} = 1$, or $e^{i\alpha} = \pm 1$. Thus,

$$\psi(\xi_1, \xi_2) = \pm \psi(\xi_2, \xi_1).$$

We come to the conclusion that there are only two possibilities: the wave function is either *symmetrical* (it does not change when the particles change places) or *antisymmetrical* (it changes its sign when the particles interchange). It is evident that the wave functions of all states of the same system must have the same kind of symmetry, otherwise a wave function, which represents a superposition of different symmetry states, would turn out to be neither symmetrical nor antisymmetrical. This result can be directly generalised for a system of an arbitrary number of like particles. Thus we find that certain restrictions are imposed on the possible states of systems of identical particles.

The properties of assemblies of particles of the same kind are described by either symmetrical or antisymmetrical wave functions, depending on the species of particles. Particle systems described by antisymmetrical functions are said to be governed by Fermi-Dirac (or simply Fermi's) statistics, systems described by symmetrical functions obey Bose-Einstein (or Bose's) statistics (see Chapter 9).

Relativistic quantum theory demonstrates (see next chapter) that the statistics governing a species of particles is uniquely related to their spin: particles having half integral spin obey Fermi's statistics, those having integral spin obey Bose's statistics.

Some important consequences follow from the foregoing discourse: no two particles in any system of identical particles obeying Fermi's statistics may occupy the same

¹ A simple interchange of the coordinates of two particles does not yet correspond to a physical interchange, or "permutation".

quantum state at the same instant (*Pauli's exclusion principle*). In a system of like particles subject to Bose's statistics any number of particles can occupy the same quantum state at the same time. It will thus be observed that even when there are no direct interactions in a quantum system of like particles, there exists a peculiar kind of reciprocal influence between them—the *exchange effect*.

The atom

Let us apply the above principles of quantum mechanics to the atom. An *atom* is a system of electrically interacting electrons moving in the Coulomb field of the nucleus. As we know, the energy, orbital angular momentum and parity of a system of particles in a central-symmetrical external field are conserved. Therefore every stationary state of an atom is characterised by certain angular momentum and parity. Furthermore, stationary states of systems of identical particles possess a certain exchange symmetry with which the spin of a system of electrons is associated.

Insofar as an atom is a system of constrained particles, the values of the above quantities comprise a discrete set. They are found from solutions of Schrödinger's equations for the respective atomic systems.

A system of electrons obeys Fermi's statistics (an electron's spin is half integral) and is therefore subject to Pauli's exclusion principle. This leads to certain laws governing the manner in which electrons fill atomic quantum states which provide a beautiful explanation of the nature of the periodic change in properties displayed by the elements arranged according to their atomic numbers (Mendeleyev's periodic system).

Conclusion

Nonrelativistic quantum mechanics, outlined in this chapter, represents a complete and logically self-contained physical theory brilliantly confirmed in its field of application by innumerable experiments. It owes its existence to Max Born, Werner Heisenberg, Erwin Schrödinger, Paul Dirac and other great physicists. The discovery of quantum mechanics represents one of the greatest triumphs of the human brain. In fact, it has shown that man is capable of rising above his own imagination and cognising things that he is incapable of imagining.

Chapter 6. Relativistic Quantum Theory

In the preceding chapter we discussed the properties of *particles* within small regions of space and found that their behaviour obeys the laws of quantum. The properties of field as a physical entity also undergo changes in small regions of space.

Systems with variable numbers of particles

We have seen that on the atomic scale the only type of interactions that matter are electromagnetic. As noted before, the properties of an electromagnetic radiation field are "analogous" to the properties of particles having zero mass, so-called photons, travelling with the velocity c . This analogy enables an electromagnetic field to be regarded *as an assembly of photons*. It is evident that the corpuscular properties of the field (the momenta of the photons) may be displayed fully only when the energy of the photons is of the same order as the energy of the particles interacting with them (electrons). The relationship $h\omega = mc^2$ offers an estimate of the order of magnitude of photon frequencies ($\omega \sim 10^{20} \text{ sec}^{-1}$) and wavelengths ($\lambda \sim 10^{-10} \text{ cm}$). Thus, at distances of the order of $\sim 10^{-10} \text{ cm}$ the quantum nature of the field becomes especially prominent.

Furthermore, as an electromagnetic field interacting with matter exchanges energy and momentum with it, a photon field must, evidently, be a system with a variable number of particles, photons being absorbed and emitted by matter. Besides, as the field and the corresponding photons are essentially relativistic entities, we conclude

that real relativistic quantum systems are systems with *variable numbers* of particles.

If we seek to generalise nonrelativistic quantum mechanics of particles having nonzero mass over the relativistic domain, we also arrive at the concept of systems with a variable number of particles. Actually, the equations of relativistic quantum mechanics (Dirac's equation) allow for a transformation, by virtue of so-called *charge conjugation*, in which a particle is replaced by an antiparticle.¹ The antiparticle of the *electron* is the *positron*. A characteristic feature of a particle-antiparticle pair is their annihilation or production as a result of interaction. Thus, electron-positron systems are systems with a variable number of particles; on this basis such systems may be interpreted as a kind of *quantised field*.

The equations of relativistic quantum mechanics of particles have the characteristics of field equations and they describe systems with variable numbers of particles, but not the properties of individual particles. Only in certain extremely simplified conditions can the behaviour of a relativistic electron be described in terms of an individual particle.

Quantised field In the relativistic quantum realm the notions of particle and field are replaced by the new concept of *quantised field* (electron-positron field, photon field).

In the most general case these fields cannot be regarded as free: they always interact. This is seen in the transformation of an electron-positron pair into a photon and the reciprocal transformation of photons into electron-positron pairs. This is also seen in the fact that an electron (like a positron) always interacts with the electromagnetic field (the photon field) produced by it. An electron emits and absorbs photons. Photon emission is extremely short-lived (virtual emission, they call it) and an emitted photon is immediately absorbed back again.

¹ The property of the equations mentioned here is that, if all particles of a system are replaced by their respective antiparticles and the sign of the electromagnetic field is reversed, the equations will not change. When we speak of a particle-antiparticle pair the question of which of the two is the "particle" and which is the "antiparticle" is one of formalism and has no physical meaning.

Hence, from the relationship $\Delta E \Delta t \sim \hbar$ the expression for photon self-energy is ambiguous and a photon may produce an electron-positron pair which is subsequently annihilated with the emission of photons.

Electrons and positrons interact through the medium of a photon field. Photons, on the other hand, interact through the medium of an electron-positron field. Thus, an electron-positron field and a photon field comprise a *single quantised electron-positron-photon field*. This field is investigated in quantum electrodynamics (the quantum theory of electromagnetic field).

Quantitative description

The equations of the electron-positron-photon field must constitute a set of Dirac and Maxwell equations in operator form. Naturally enough, the mathematical apparatus and the concepts of relativistic quantum theory must be unlike those of nonrelativistic quantum mechanics. The nonrelativistic quantum-mechanical description theory does not investigate explicitly the processes of particle production and annihilation (or, more correctly, interchange). In quantum mechanics the square of the wave function gives the probability of a particle occurring in a given state. In relativistic quantum theory the wave functions of quantised fields describe assemblies of particles contributing to the fields concerned and the interchange processes, which are explicitly contained in the theory. Accordingly, the wave functions of quantised fields acquire operator form and separate into operators of particle production and annihilation between which "commutative" relationships are established. *Operator wave functions* are determined by the field equations and commutative relationships.

Thus, the field functions cease to be functions in the conventional sense and become operators of a wave function common to all fields, called the *amplitude of state*. In relativistic quantum theory the probabilities of state are given by quadratic forms of the amplitude of state.

The variety of processes of photon radiation and absorption by electrons and positrons, photon scattering, production of electron-positron pairs and their annihilation with photon emission and many others are considered

as transitions of the electron-positron-photon field from one state to another.

There is also possible a field state (the lowest, basic state) in which the particles considered as its quanta vanish. This state is called *vacuum*. When a field is in an excited state it contains *excited quanta* — particles, in fact.

Property of vacuum Due to the weakness of interactions between electron-positron and photon fields they can, to some approximation, be considered separately. It should be remembered, however, that actually they *always* interact (thanks to the existence of vacuum states) and the notion of localised free fields is in principle wrong.

The state of vacuum of a photon field is one in which the number of photons is zero. It possesses a certain *zero energy*. The effect of so-called *zero-point oscillations of a photon field* is observed in a number of phenomena and experimentally.

The vacuum state of an electron-positron field is one in which the number of electrons and positrons is zero. The mean value of the electric charge of this state is zero. However, thanks to zero-point fluctuations of the vacuum state, charges of varying sign are observed to appear and vanish in different sections of space. This effect of *electron-positron vacuum polarisation* is observed experimentally.

Remarks We saw in general relativity theory that the geometrical properties of space-time are conditioned by physical phenomena. In relativistic quantum theory we find that vacuum—empty space—is endowed with physical properties.

Speaking of the photon field, note that its excitation quanta (photons) possess integral spin, hence they obey Bose's statistics: an infinite number of photons may be in every state of the photon field. Thus a limiting transition to the conceptions of classical electrodynamics is possible.

Speaking of the electron-positron field, note that its excitation quanta (electrons and positrons) can appear and disappear only in pairs, in accordance with the law of conservation of charge. Electrons and positrons have

half-integral spins and they obey Fermi's statistics: no more than one particle can occur in each quantised state of an electron-positron field. Thus the question of any limiting transitions (to classical theory) is meaningless and an electron-positron field is essentially a quantum-relativistic entity.¹

An important achievement of relativistic quantum theory is that it has established the connection between spin and statistics.

Point interaction It was pointed out before that the interaction between a photon and an electron-positron field is very weak. Owing to this, the basic equations of relativistic quantum theory are solved by the method of successive approximations. An interaction is regarded as a small increment (a small perturbation), all quantities being expressed in the form of series in powers p of the interaction constant (which is small in comparison with unity).

As we know, in relativistic theory interaction can take place only when two particles are simultaneously at the same point in space (the concept of close-quarter or, more correctly, *point interaction*), and a photon and electron interact at a point. But we found in classical electrodynamics that point interaction leads in computations to results with meaningless infinite expressions. Thus, in computing the self-energy (mass) of a point electron we obtain infinity.

These difficulties extend to quantum theory as well. The interaction of an electron with the zero-point oscillations of a photon field yields an infinite value for its mass; interactions between an electron and zero-point oscillations of an electron-positron field yield an infinite charge.

In classical electrodynamics the difficulty is surmountable insofar as it does not deal with small spatial scales. Not so in quantum electrodynamics. If in solving for some effect we succeed in computing the first approximation, subsequent approximations of theory lead to diverg-

¹ In this connection note that the effect of interaction of photons through an electron-positron field (*a nonlinear effect*), does not exist in classical theory.

ing results and cannot be solved. For some phenomena the very first approximation yields infinities. Nevertheless, the divergencies of quantum theory are much less pronounced than in the classical theory.

Renormalisation The infinities appearing in relativistic quantum theory are all of the same type and are reduced to infinite variations of charges and masses. The difficulties were overcome in the following manner. It is experimentally impossible to observe the mass of a "free" electron, m_0 , since electrons *always* interact with vacuum. We can only measure the "observable" mass

$$m = m_0 + \delta m,$$

where δm is the divergent insertion to mass m_0 due to the electron's interaction with zero-point oscillations of the photon field.

At this juncture we must make reference to the physical principle according to which science deals only with "observables"; in other words, a physical theory must be formulated in terms of quantities that are in principle observable. According to this concept it is the observable m , not the isolated quantities m_0 or δm , that is investigated by theory. This is known in relativistic quantum theory as the *mass renormalisation principle*.

A similar situation applies to the electron charge. Since any external charge *always* polarises vacuum (i. e., an interaction takes place between the charge and zero-point oscillations of the electron-positron field) theory must allow for *charge renormalisation*

Note the common features of mass and charge renormalisation. In both cases the divergencies are associated with unobservable phenomena. Thus, in renormalising charge (mass) the divergence δe (δm) changes only the "bare" charge e_0 ("bare" mass, m_0) of the electron. The "bare" charge or mass need not necessarily be finite. All that matters is that the resultant (observable) physical charges and masses be finite. Insofar as the observable quantities are e and m , the divergencies in the theory disappear if the results are written in terms of the observed charges and masses. It follows from what has been said

that even if there were no infinities theory would have to be renormalised.

Note again that the infinities appearing in theory are all of the same type and are reduced to infinite changes in charge and mass. When the charges and masses are renormalised a theory free from infinities can be developed in which it is possible to compute not only the first approximation but any further approximations as well.

Latest developments of theory Relativistic quantum theory describes electromagnetic interactions, or in other words, interacting electrons, positrons and photons. Its description embraces all extra-nuclear phenomena: the structure of atomic electron shells, radiation, absorption and scattering of light quanta by atomic systems, collisions of electrons among themselves and with atomic nuclei, positron production, etc.

The situation is entirely different when strong particle interactions are investigated.¹ Attempts to apply ordinary quantum-electrodynamical formalism to them have failed.

The most promising school in the theory of strong interactions is probably the school which extends the most general concepts of relativistic quantum theory to the domain of strong interactions. Account is taken of the fact that it is impossible, within the framework of relativistic quantum theory, to measure any quantities relating to interacting particles. Exact measurement of the coordinates of one particle would presume the production of a tremendous number of pairs, while exact measurement of momentum in a short time interval contradicts energy-time uncertainty relationships.

In the framework of relativistic quantum theory only the momenta and polarisation of free particles can be measured, since in measuring them we are not restricted by time (thanks to the conservation of momentum). Therefore, in the theory only relationships between quantities describing free particles have physical meaning.

¹ Strong (nuclear) interactions are interactions of an entirely different nature. Their intensity is several thousand times higher than that of electromagnetic interactions (see p. 102).

Investigation of the processes of particle interaction in this theory is impossible in principle. The theory investigates only the behaviour of the particles *before* and *after* the interaction, not the interaction itself. A consequence of such conceptions, as applied to the problem of strong interactions of elementary particles, is a theoretical formalism which excludes Ψ -operators, as they contain quantities that are unobservable in principle. A Hamiltonian, on the other hand, can be constructed only from Ψ -operators, hence the theory of strong interactions cannot be based on Hamiltonian functions.

The new theoretical formalism is based on so-called unitary relationships and the principle of point interaction, which manifests itself in the analytical properties of the basic quantities of the theory, for example in various kinds of dispersion relationships.

Since the concept of compound particles is linked with the investigation of particle interaction, the conclusion may be drawn that it is a meaningless concept in the framework of relativistic quantum theory. In this theory all particles must be equal in the sense that they are all equally elementary and compound. Thus, such an approach to theory makes the old question of the "elementarity" of particles quite meaningless.

Chapter 7. Physics of Elementary Particles

The sum total of experimental data has enabled scientists to formulate a number of general propositions and principles concerning the *physics of elementary particles*.

Classification of elementary particles and their interactions First of all, *elementary particles and their interactions* must be classified. To begin with, every particle is associated with an antiparticle. In some cases the particle and antiparticle may be identical, as in the case of photons and π^0 -mesons. The main feature of particle-antiparticle pairs is their capacity to annihilate.

All the known elementary particles are classified into four families:

1. Baryons (nucleons, N ; hyperons, Λ , Σ , Ξ) and antibaryons.
2. Mesons (π -mesons [pions]; K -mesons) and antimosons.
3. Leptons (electrons, e ; neutrinos, ν ; μ -mesons [muons]) and antileptons.
4. Photons, γ .

The interactions in which elementary particles can participate are also naturally classified into the following three types:

1. *Strong interactions* (which occur between baryons, antibaryons and mesons). They are responsible for nuclear forces (interactions between nucleons) and the production of mesons and hyperons in high-energy nuclear collisions.

2. *Electromagnetic interactions* (which bind photons with all charged particles). Leptons, with the exception of the neutrino, are characterised by these interactions.

3. *Weak interactions*, which are responsible for slow particle decay.¹

The intensity of interactions is given by the following figures:

1. Strong interactions, ~ 1 ;
2. Electromagnetic interactions, $\sim 10^{-3}$;
3. Weak interactions, $\sim 10^{-14}$, and, to complete the picture,
4. Gravitational interactions,² $\sim 10^{-40}$.

The state of affairs is best described by means of successive approximations. In the *first approximation*, interactions of types (2) and (3) are not taken into account. This means that leptons and photons cannot interact with baryons or mesons or among themselves. Baryons, anti-baryons and mesons take part in reactions and transmutations according to the specific laws of strong interactions; decay with emission of leptons or photons does not take place.

The *second approximation* includes particle charges, which means that interactions of types (1) and (2), but not (3), are possible. Processes involving baryons, anti-baryons and mesons are not affected by electromagnetic phenomena, but decay with emission of leptons and photons is permitted.

In the *next approximation*, which probably presents the fullest description of matter (excluding gravitation), weak interactions enter.

Interactions differ substantially in their *time characteristics*. Thus, processes which can take place in the first approximation are characterised by times of the order of

¹ The neutrino is the only particle characterised by *weak interactions alone*. That is why collisions between neutrinos and particles of matter are so rare. In matter of normal density the distance between two collisions of a neutrino with other particles is measured by the astronomical figure of the order of 10^{17} km. This means, for example, that the earth, whose diameter is $\sim 12 \times 10^3$ km is quite "transparent" to a stream of neutrinos. Nevertheless, experimental physicists today are able to detect neutrinos directly.

² Note that the upper limit of the spatial region in which strong and weak interactions are observable lies at about 10^{-13} cm; the radius of electromagnetic and gravitational interactions is virtually unlimited, which explains why only the latter two are observed in classical situations.

10^{-23} sec ("fast" processes); those which occur in the second approximation, but not in the first, have a duration of $\sim 10^{-15}$ sec (*electromagnetic* processes); finally, those occurring *only* in the third approximation are characterised by times of the order of 10^{-8} sec ("slow" processes).

The conservation laws in the physics of elementary particles

The general relationships governing the processes of particle production, absorption and decay are the *conservation laws*. In addition to the known laws of conservation of energy, momentum, angular momentum (including the "intrinsic" angular momentum of particles, spin) and parity (including the "intrinsic" parity of particles), in the physics of elementary particles there exist a number of specific conservation laws.

They are most conveniently examined according to the above scheme of successive approximations. Thus, in the first approximation—strong interactions—we observe *conservation of "baryon" charge* (conservation of the difference between the number of baryons and antibaryons), *isotopic spin* (an intrinsic degree of freedom of baryons, antibaryons and mesons characterising their charge properties) and "*strangeness*" (a new discrete characteristic of particles).

The strangeness for nucleons, antinucleons, and pions (the "fundamental" particles) is zero; for their "excited" states—hyperons and *K*-mesons—the "strangeness" has integral numbers. The law of conservation of strangeness provides an explanation for some features of processes involving "strange" particles; thus, in fast processes they are never produced singly and they have a long lifetime (slow decay).

Introduction of electromagnetic interactions (the second approximation) eliminates degeneracy in the isotopic spin of particles, which results in violation of *isotopic invariance* (violation of the conservation of isotopic spin). Thus, for example, such particles as nucleons, *N*, and pions decay into a number of "charge" (electrical) states: differences develop between the proton, *N*⁺, and the neutron *N*⁰; the pion has three charge states.

In the second approximation the conservation of strangeness holds. The concepts of strangeness and isotopic

spin made possible a very convenient grouping of strong-interacting particles into *charge multiplets* (an example of which we have just seen). The concepts of strangeness and isotopic spin are not applicable to leptons.

In the "electromagnetic" approximation mentioned above the "baryon" charge is conserved. Naturally, in this approximation the *law of conservation of electrical charge* appears.

With the introduction of weak interactions (the third approximation) the latter two conservation laws are valid; strangeness *is not* conserved; nor, naturally, isotopic spin either.

The laws of conservation of energy, momentum and angular momentum are valid for *all interactions*, i. e., they are universal.¹

Strong and electromagnetic interactions are invariant with respect to particle-antiparticle interchange. In weak interactions the invariance is violated. Parity is conserved in strong and electromagnetic interactions but not in weak interactions.

All interactions are invariant with respect to the totality of particle-antiparticle interchange and space inversion (*combined parity*).

The law of conservation of combined parity is thus a universal one, and it leads to some interesting conclusions. In the case of a neutrino, for instance, we arrive at the concept of a *two-component longitudinally polarised neutrino* and an associated new discrete particle characteristic called *helicity*.

The concept of combined parity leaves empty space completely symmetric, asymmetry being transferred to particles. Our world turns out to be asymmetrical with respect to left and right.

Remarks It was pointed out before that the logical integrity of physical concepts gives hope of a linkup between the ideas of physics of elementary particles and general relativity (cosmology). In

¹ In gravitational interactions these general conservation laws become in effect meaningless. In the physics of elementary particles, however, gravitational interactions are negligible.

this respect three facts, which may or may not be inter-related, are of special interest.

Fact number one. There is an apparent preference in the observable universe for nucleons and electrons, whereas the concepts of particle physics require complete symmetry in density distribution of particles and antiparticles.

Fact number two. The concepts of particle physics lead to the conclusion that our world is not mirror-symmetrical.

Fact number three. Cosmological concepts lead to the conclusion of the expanding universe.

Are these three facts in any way related? At present physicists are unprepared even to formulate the question adequately.

Chapter 8. Nuclear Physics

In the previous chapter we considered the properties of processes involving *individual* elementary particles. This brief chapter deals with the properties of *systems of elementary particles*. The spatial dimensions of such systems are of the order of 10^{-13} cm. In practice physicists deal mainly with systems of nucleons. Hyperons, which can be regarded as “excited states” of nucleons, decay into nucleons and pions. Assemblies of bound states of nucleons comprise stable formations of systems of elementary particles called *atomic nuclei*.

Nucleons in nuclei are subject to all types of interactions: strong, electromagnetic and weak. From this it would seem that a complete theory of bound states of nucleonic assemblies—the *atomic nucleus theory*—can be developed only after a theory of all types of interactions of elementary particles.

Data concerning nuclear properties are obtained mainly from laboratory experiments. Utilisation of the most general principles of quantum mechanics enables only an “analogue” interpretation of certain nuclear properties to be carried out, different families of nuclear phenomena requiring different analogues.

A comprehensive nuclear theory will be able to answer the question concerning the limits of the periodic system of elements.

Chapter 9. Statistical Physics

Up till now we have been examining the properties of isolated micro-entities and small assemblies of them. Now we shall investigate the properties of physical systems comprising vast numbers of such micro-entities as atoms and molecules. These are *macroscopic* systems or bodies.

Macroscopic bodies Macroscopic bodies obey laws which are generally not concerned with whether the motions of the component particles of a body are described by classical or quantum mechanics. Their substantiation, however, requires a different course of reasoning in either case. For the sake of convenience we shall proceed with our investigation on the basis of classical-mechanical assumptions.

In principle, if one can write as many equations of motion for a mechanical system as it has degrees of freedom, and solve (integrate) them, one can obtain exhaustive information concerning the motion of the system. However, if one must deal with a system which, though obeying the laws of classical mechanics, possesses very many degrees of freedom, one would have to solve so many differential equations as to make one's task virtually impossible. Moreover, even should one be able to integrate the equations in general form, one would be unable to substitute the initial velocities and position coordinates of the particles into the general solution.

It would thus seem at first glance that with the number of particles increasing the complexity and confusion of the properties of a mechanical system must increase so much as to utterly bury under the natural laws governing its behaviour. This, of course, is not the

case, and when the number of particles is very great new features and laws appear.

Statistical laws These are known as *statistical laws*, and they are conditioned precisely by the great number of particles comprising a body. They can in no way be reduced to purely mechanical laws. Their special features consist in that they lose meaning if applied to mechanical systems with few degrees of freedom. Thus, although the motions of systems possessing very many degrees of freedom are subject to the same mechanical laws as systems with few degrees of freedom, the very existence of so many degrees of freedom results in *qualitatively new laws*, which comprise the subject of what is called *statistical physics* or simply *statistics*.

Phase space Before formulating the basic tasks of classical statistics let us introduce the concept of *phase space*. Suppose a given macroscopic mechanical system has s degrees of freedom. In other words, its points are localised in space by s coordinates q_i , where the subscript i passes through the values $1, 2, \dots, s$. Then the state of the system at a given instant is determined by the simultaneous values of the s coordinates q_i , and the corresponding s momenta p_i . The different states of the system may be mathematically represented by points in so-called *phase space* (which is, of course, a purely mathematical concept); the coordinates and momenta of the given system are laid off on coordinate axes in the phase space. Every system has its own phase space, the number of dimensions of which is equal to twice the number of degrees of freedom of the system. Every point of the phase space, which corresponds to certain values of the coordinates q_i and momenta p_i of the system, represents a certain state of the system. The state of the system changes with time, and correspondingly the point of phase space describing the state traces a line called the *phase path*.

Subsystems Consider a macroscopic body or closed system of bodies, which has no interactions with other bodies. Mentally isolate a part of the system, very small in comparison with the whole, but still macroscopic; obviously, if the system consists of a sufficient number of particles, even a small part of it may

still have very many particles. We shall call such relatively small, but macroscopic, bodies *subsystems*. A subsystem is a mechanical system that is no longer closed and interacts with other parts of the system. Due to the tremendous number of degrees of freedom of those other parts the interactions are of a very complex and involved nature and the state of the subsystem changes with time in a very complex and involved way.

An exact solution of the problem of the behaviour of the subsystem is possible only by solving the mechanical problem for the whole closed system, i. e., by developing and solving all the differential equations of motion for the given initial conditions. This, as mentioned before, is a virtually hopeless task. Fortunately enough, however, the very complexity of the mode in which the state of a subsystem changes, which makes methods of mechanics inapplicable, provides an approach from another side.

Statistical distribution

Owing to the extreme complexity of the external forces exerted by the other parts of the system, our isolated subsystem, given sufficient time, will pass through all its states many times. This may be formulated more precisely in the following way. Denote by the symbol $\Delta p \Delta q$ any small element of "volume" of the phase space of the subsystem corresponding to its coordinates q_i and momenta p_i lying within some small intervals Δq_i and Δp_i . We may claim that, given a sufficient time interval T , the meandering phase path will pass many times through every such element of the phase space. Let Δt be the part of the total time T in which the subsystem was in a given phase space element $\Delta p \Delta q$.¹ When the total time T is increased to infinity the ratio $\Delta t/T$ tends to a limit

$$\omega = \lim_{T \rightarrow \infty} \frac{\Delta t}{T}. \quad (38)$$

This quantity can be regarded as the *probability* of observing the subsystem in the given phase space element $\Delta p \Delta q$ at an arbitrary time.

Thus, there must exist a certain function $\rho(p, q)$ which represents the "probability density" in phase space.

¹ By the latter expression we mean that the system is in the states denoted by the phase points in that element.

Knowledge of this function enables the expression (38) to be rewritten as follows:

$$\Delta w = \rho(p, q) \Delta_p \Delta_q.$$

The function ρ is called the *statistical distribution function*, or simply the distribution function, for a given body.

The following consideration is of great importance in statistics. The statistical distribution of a given subsystem does not depend on the initial state of some other small part of the same system, as the effects of that initial state are completely superseded over a sufficiently long period of time by the effects of other, much larger parts of the system. Neither do they depend on the initial state of the isolated small part, insofar as in time it passes through all possible states, each of which may be taken for the initial state. Therefore statistical distribution can be found for small parts of a system without solving any mechanical problems taking into account the initial conditions.

Determination of the statistical distribution for any subsystem constitutes the basic task of statistics. Speaking of "small parts" of a closed system, it should be borne in mind that the macroscopic bodies, with which we are concerned, are usually themselves small parts of a larger closed system comprising those bodies and the medium in which they are immersed.

If the problem is solved and the statistical distribution for a given subsystem is known, we can calculate the probability of the different values of any physical quantities that are dependent on the state of the subsystem (i. e., on its coordinates q and momenta p). It is also possible to compute the mean value of any such quantity by multiplying all possible values by the corresponding probabilities and summing (integrating) over all states.

Averaging by means of the distribution function (statistical averaging, as it is called) makes it unnecessary for us to follow the changes of the actual values of the physical quantity with time to determine the mean value. At the same time, it is evident that, by virtue of the very definition of probability, it follows from equation (38) that statistical averaging is equivalent to averaging over time.

Macroscopic quantities

It is evident from what has been said that statistical conclusions and predictions concerning the behaviour of macroscopic bodies are *indeterminate*. This is the difference between statistical and classical mechanics, the conclusions of which are determinate. It should be stressed, however, that the indeterminacy of statistics is due not to the nature of the entities it deals with but only to the fact that statistical data are obtained on the basis of much less information than is required for a complete mechanical description (it is not necessary to know the initial values of all the coordinates and momenta).

In practice, however, the indeterminacy of statistical descriptions of macroscopic bodies is usually in no way *observable*. For in observing a macroscopic body in steady-state (i. e., time-independent) conditions for a sufficiently long time we find that all physical quantities¹ characterising the body are virtually constant (equal to their mean values), with appreciable deviations occurring very rarely.² This fundamental rule is the more valid, the more complex and bigger the body under consideration.

Thus, by providing a method of calculating mean values of quantities characterising macroscopic bodies, statistics enables predictions to be made which are true to a very high degree of accuracy for practically any time interval; so high, in fact, as to completely eliminate the effects of the initial state of the body. In this sense the predictions of statistics are, for all practical purposes, determinate.

Statistical equilibrium

A closed macroscopic system is said to be in *statistical equilibrium* if the *macroscopic* physical quantities describing any macroscopic part of it are equal to their mean values to a high degree of relative accuracy. It can be seen from

¹ *Macroscopic* quantities, of course, which characterise the body as a whole or separate macroscopic parts of it, but not individual particles.

² The following example offers graphic illustration of the great accuracy of this principle. If a volume containing 0.01 gram-molecule is isolated in a gas, the mean relative deviation of energy in this amount of matter from the mean value is $\sim 10^{-11}$. At the same time, the probability of detecting (from one observation) a relative deviation of the order of, say, 10^{-6} is given by the incredibly small number $\sim 10^{-3 \times 10^{18}}$.

the foregoing that if a closed macroscopic system is observed for a sufficiently long time it will be in statistical equilibrium for most of that time. If at any initial instant a closed macroscopic system was not in statistical equilibrium (for instance, it was deliberately disturbed by an external force which was later withdrawn, i. e., it became a closed system again), it will inevitably attain equilibrium. The time for the necessary transition to statistical equilibrium is called the *relaxation time*. When we spoke above of "sufficiently long" time intervals we meant intervals longer than the relaxation time.

Liouville's theorem To continue with the properties of the statistical distribution function, suppose now that we are observing a subsystem for a very long time. At every instant the system is denoted by a point in its phase space. The distribution density of all points in the phase space is at any given place in the limit proportional to the distribution function $\rho(p, q)$ which, according to its meaning, determines the probabilities of different states of the subsystem.

It is obvious that, just as at the initial time, at any other instant of time all these points will be distributed in the phase space according to the distribution function $\rho(p, q)$. In other words, the distribution density of the moving phase points is uniform for any given place and proportional to the corresponding value of ρ . We thus arrive at the important conclusion that the distribution function is constant along the phase paths of a subsystem (*Liouville's theorem*).

Remember that the subsystems are not themselves closed. On the contrary, they are continually subject to forces from other parts of the system. But as these parts, though small in comparison with the whole system, are nevertheless macroscopic bodies, we may assume that over time intervals that are not too long they behave approximately like closed systems. For participating in the interactions of a subsystem with surrounding parts are mainly those particles that are located near its surface. The relative number of such particles as compared with the total number of particles in the subsystem decreases rapidly with the size of the subsystem increasing, and for a subsystem of sufficient size the energy of its interaction with neigh-

bouring parts is small as compared with its internal energy. We can thus say that subsystems are *quasi-closed*, always remembering that this holds for time intervals that are not too long. If the time is sufficiently long, effects of interactions between subsystems will be manifest no matter how weak they are. More, it is just such comparatively weak interactions that lead ultimately to statistical equilibrium.

Insofar as we are speaking of quasi-closed subsystems, the foregoing result (Liouville's theorem) is valid only for time intervals that are not too long, during which a subsystem behaves like a closed system to a sufficient accuracy.

**The conservation
laws
in statistical
physics**

It follows from Liouville's theorem that the distribution function must be expressed only in terms of such combinations of the variables p and q which remain constant when a subsystem moves as a closed entity. We know from mechanics that these constant combinations of the variables p and q are energy, momentum and angular momentum. Thus we arrive at an important conclusion. The conserved quantities—energy, momentum and angular momentum—completely determine the statistical properties of a closed system, i. e., the statistical distributions of any subsystems within it, and consequently, the mean values of any physical observables pertaining to them. These conserved quantities take the place of the immeasurable amount of data (initial conditions) which would be needed for a mechanical approach.

The linear and angular momenta of a closed system are associated with its motion as a whole—uniform translatory motion and uniform rotation. It can be said, therefore, that the *statistical state of a system* undergoing a given motion depends *only on its energy*. Thus, in statistics energy is of special importance.

**Gibbs'
distribution**

The considerations set forth here enable a simple distribution function to be developed for a closed system. This is known as *Gibbs' distribution*, and it provides a description of the equilibrium statistical properties of *any* macroscopic bodies.

**Special features
of quantum
statistics**

Going over to the peculiarities of quantum statistical physics, note first of all that in quantum mechanics a purely mechanical approach to the problem of determining the behaviour of a macroscopic body is as hopeless as in classical mechanics. Such an approach would require Schrödinger's equation to be solved for a system comprising all the particles of a body, a problem even more hopeless, if possible, than the integration of the classical equations of motion. Even should it prove possible in some specific case to develop a general solution of Schrödinger's equation, it would be hopeless to select and write the partial solution satisfying the specific conditions of the problem, as it is characterised by the specific values of a tremendous amount of different quantum numbers. More, we shall find later on that for a macroscopic body the concept of stationary states is, in a sense, conditional—and this is a point of fundamental importance.

**Energy ranges
for bodies**

Let us determine some of the features characterising macroscopic bodies from the purely quantum-mechanical point of view as compared with systems consisting of a relatively small number of particles. First, there is the extreme distribution density ("almost a continuity") of levels in the range of self-energy values of a macroscopic body. The reason for such density is readily understood if one notes that, thanks to the enormous number of particles in a body, any energy may, roughly speaking, be "distributed" among the different particles in a tremendous variety of ways. Note also that at the beginning of a macroscopic body's energy range the distances between the first energy levels are not at all small and may even be independent of the dimensions of the body (the number of particles).

**Impossibility
of stationary
states**

Owing to the frequency of levels a macroscopic body is actually unable to remain in a strictly stationary state. In any case, the energy value will be "spread out" by a quantity of the order of the energy of interaction of the system with other bodies. But the latter is infinitely greater than the distances between the energy

levels, and this refers not only to quasi-closed subsystems but also to systems which could be regarded as closed from any point of view. In nature, of course, there are no absolutely closed systems whose interactions with other bodies would be precisely zero; but any actual interaction, even though it be so small as to have no effect whatsoever on any other properties of a system, will still be very great in comparison with the vanishingly small intervals in its energy range.

Statistical matrix

The state of a macroscopic body cannot be described by a wave function since the possible reserve of data concerning the state of such a body is far smaller than the total amount needed to develop the wave function of state. A quantum-mechanical description based on incomplete data about a system is effected by means of a so-called *density matrix*. With a density matrix one can calculate the mean value of any quantity characterising a system, as well as the probabilities of different values (in statistics a density matrix is called a *statistical matrix*).

In quantum statistics the statistical matrix takes the place of the distribution function of classical statistics. Whereas the distribution function $\rho(p, q)$ gives the distribution of probabilities of different coordinates and momenta of the particles of a body, the statistical matrix gives the probability of a body occurring in this or that quantum state.

All the most general considerations outlined above for classical statistics, in particular, that about the practically determinate nature of statistical predictions, refer completely to quantum statistics as well.

Entropy. The law of increasing entropy

It was mentioned before that if a closed system is not in statistical equilibrium, its macroscopic state changes with time until equilibrium is achieved. In characterising a macroscopic system by the distribution of energy among its subsystems one may say that the series of successive states through which the system passes corresponds to an increasingly probable energy distribution. This probability increase is, generally speaking, considerable, insofar as it is of an exponential nature. Namely, the probability is given by the expression e^θ , the exponent

being the so-called *entropy of the system*.¹ It can be said, therefore, that processes in a closed system not in equilibrium proceed in such a way that the system passes steadily to states with higher and higher entropy, until complete statistical equilibrium is achieved. This is known as the *law of increasing entropy*.

Thus, in natural closed systems entropy never decreases: it increases or at best remains constant. According to these two possibilities all processes involving macroscopic bodies are conventionally classified as *irreversible* and *reversible*. To the former belong processes in which the entropy of a closed system increases; they cannot be repeated in reversed order as this would involve a decrease in entropy. Reversible processes are those in which the entropy of a closed system does not change and which, therefore, can proceed in both directions. A strictly reversible process is, naturally, an idealised limiting case. Real natural processes may be reversible only to a higher or lower degree of accuracy.

Application of the law of increasing entropy to the world as a whole *does not necessitate* statistical equilibrium. Actually, the statement that, given a sufficient period of time, a closed system must necessarily attain equilibrium refers to systems in steady-state conditions. But we have seen (see pages 65 and 69) that within the framework of general relativity the world as a whole is not a closed system, it is a system in a variable gravitational field. This arises from the fact that a gravitational field cannot be included as a component of a closed system as otherwise the conservation laws, which constitute the foundation of statistics, would turn into identities. Gravitational field itself is, generally speaking, a function not only of the position coordinates but of time as well, which means that the "external conditions" are far from stationary.

¹ *Entropy* is a very important statistical concept. It characterises the "degree of smearing" of a macroscopic state of a system over its microscopic states. In other words, the entropy of a system is determined by the number of microscopic ways in which a given macroscopic state of a system can be effected (or more precisely, the logarithm of that number).

Thermodynamic relationships

The physical observables characterising macroscopic states of bodies are called *thermodynamic quantities*. Some of them have, besides a thermodynamical, also a purely mechanical meaning, such as energy or volume. There is, however, a category of quantities appearing only in connection with purely statistical relationships which are quite meaningless if applied to nonmacroscopic systems; to this category belongs entropy. Some relationships between thermodynamic quantities are valid irrespective of the specific bodies to which they apply. These are known as *thermodynamic relationships*.

Temperature The concept of entropy leads us to the important thermodynamic quantity called *temperature*. Insofar as the entropy of a system is a function of its energy, $\sigma(E)$, there can exist a quantity

$$\frac{d\sigma}{dE} = \frac{1}{\Theta}, \quad (39)$$

which has the same value for a whole system in statistical equilibrium. The quantity Θ is called *temperature*. Like entropy, temperature is a purely statistical quantity which makes sense only for macroscopic bodies.

Entropy σ is a dimensionless quantity. It follows, therefore, from (39) that temperature Θ has the dimension of *energy*, which means that it can be measured in units of energy, ergs, for instance. In practice, however, special units called *degrees* are conventionally used. If Θ is temperature measured in ergs and T is the same temperature in degrees, the relationship between the two values is

$$\Theta = kT,$$

where k is a coefficient of proportionality, called *Boltzmann's constant*, equal to 1.380×10^{-16} erg/degree.

Pressure A macroscopic body evidently acts with some force on the boundaries of its volume. The absolute magnitude of the force per unit area of that boundary is called *pressure*, it is another thermodynamic quantity.

Equation of state Macroscopic (statistical) systems are characterised by a relationship between pressure, volume and temperature, called the *equation of state* of the given system. A comprehensive determination

of the state of a system, however, requires knowledge of its energy as well.

The equation of state can be derived if the microscopic characteristics of the system (its energy range) are known. For, on the one hand, between all thermodynamic quantities there exists a relationship which is manifest in the general thermodynamic formulas. On the other hand, such a thermodynamic quantity as entropy is closely linked with the microorganisation of macrosystems; thus, if the microstructure of a body is known, its thermodynamic parameters can be calculated.

States of matter As is known, at more or less conventional temperatures and pressures matter is either solid, liquid or gaseous. Extremely high temperatures and pressures lead to various kinds of "stellar states" of matter.

The ideal gas The simplest physical macrosystem is the *ideal gas*, which is a gas in which practically no interactions take place between its molecules.¹

The distribution of the gas particles among the various quantum states is given by Gibbs' distribution (as, incidentally, is the description of equilibrium statistical properties of all macroscopic systems). If the gas constitutes a system of like particles then, as mentioned in Chapter 5, even in the absence of direct force interactions in the system there exists a specific kind of interaction between the particles (exchange effects) which governs the rules according to which the particles of the system fill the quantum states. As we know, there exist only two such rules, according to which there are two kinds of statistical distribution functions for systems of like particles: two quantum statistics, Bose's and Fermi's.

It was noted before that the energy range of macroscopic systems appears as almost continuous, except for the very beginning of the range. It is clear, then, that the low-excitation states of a system (very low temperatures) are described essentially by quantum statistics while highly excited states (high temperatures) are described by classical statistics.

¹ Gas at atmospheric pressure approaches this state.

Since the interactions of particles in ideal gases can be neglected, it is evidently possible to compute the thermodynamic functions for them in general form (i. e., a form suitable for describing all kinds of gaseous systems regardless of the structures of the component molecules). The general equation of state for ideal gases is called *Clapeyron's equation*.

Solid state Characteristic of the *solid state of matter* is that the atoms oscillate slightly around equilibrium positions. The configuration of the equilibrium positions corresponding to the statistical equilibrium of the body is a preferred one among other possible distributions. In other words a solid body in statistical equilibrium must be *crystalline*.

But there also exist amorphous solids in which the atoms oscillate around randomly located points. From the above point of view such bodies are not in equilibrium and should crystallise with time. Actually, though, the relaxation times are so long that amorphous bodies behave as stable bodies for practically infinite time.

Note that for a body to be solid a relatively low temperature is required. The quantity kT , in any case, must be small in comparison with the interaction energy of the atoms (at higher temperatures all solids melt or disintegrate). This is a consequence of the fact that the oscillations of the atoms of a solid around their equilibrium centres are very small.

The latter consideration, in turn, accounts for the simple type of heat motion in solids, making it possible to calculate the thermodynamic parameters of solids in general form.

Liquids Unlike gases or solids, *liquids* do not allow for a calculation of thermodynamic parameters in general form. This is because interactions between the molecules are strong, while at the same time the simple heat motions observed in solids are absent. Due to the high intensity of molecular interactions thermodynamic parameters must be calculated on the basis of specific interaction laws, which differ for different liquids.

The only thing that can be undertaken in general form is an investigation of the properties of liquids at temper-

atures close to absolute zero, although helium is the only known substance that remains liquid near absolute zero. According to classical mechanics, at *absolute zero* the atoms are motionless and their equilibrium potential energy of interaction must be minimum. Therefore, at sufficiently low temperatures the atoms should be expected to perform small oscillations, that is, all bodies should be solid. Actually, though, quantum effects may condition exceptions from the rule. Such an exception is helium. All other substances solidify long before quantum effects become of importance.¹ Thus, thanks to the weakness of interactions between the atoms of helium, it remains liquid down to temperatures where quantum effects begin to matter (a *quantised liquid*), after which solidification becomes impossible.

**“Stellar states”
of matter**

Another problem of interest is the properties of substances when the temperature is increasing. First, we know, from its condensed states (solid and liquid) matter becomes gaseous. Ordinary gases are molecular. As the temperature rises to around 10^3 °K thermal dissociation begins (molecules decompose into atoms) and gases become atomary.

At temperatures of about 10^4 °K the atoms ionise. Ionised matter—*plasma*—consists mainly of ions and electrons ($T \sim 10^6$ °K).

At around 10^7 °K ionisation is complete and the plasma consists of “bare” nuclei and free electrons. As the temperature rises higher nuclear transformations begin ($\sim 10^8$ °K).

Above 10^9 °K the nuclei disintegrate and matter consists of protons and electrons ($T \sim 10^{11}$ °K).² Finally, above 10^{13} °K general transmutations of elementary particles are possible. Thus, to produce a nucleon-antinucleon pair

¹ Quantum effects come into play when the de Broglie wavelength, which corresponds to the thermal motion of atoms, becomes comparable with intra-atomic distances. In liquid helium this happens at 2.3 °K.

² Neutrons are unstable particles which decay into protons with the emission of an electron and a neutrino; the latter does not interact with matter and leaves the system.

energy of the order of Mc^2 is required, where M is the mass of a nucleon. From the relationship $Mc^2 \sim kT$ we obtain the above temperature.

Also of considerable interest is the investigation of the properties of matter at very high densities (in the above listing of property changes of matter with temperature, pressure was assumed normal). Let us see how the properties of matter change with the density increasing, assuming that the temperature remains relatively low.

When the pressure reaches about 10^8 atm the electron shells of atoms are deformed and their internal energy increases markedly. The electric fields of individual nuclei overlap more and more, as a result of which electrons in atomic shells become less and less associated with specific atoms. Free motion of the external electrons is possible. When matter is compressed to about 10^{12} atm the interactions of the electrons and nuclei cease to be of importance and the substance can be regarded as a kind of electron gas of great density (known as *degenerated* gas, and it is described by Fermi's statistics).

When the density and pressure of the gas become of the order of 10^6 g/cm³ and 10^{18} atm respectively¹ the electron gas becomes relativistic, i. e., the average energy of the electrons is comparable with mc^2 .

A further increase in density leads to states in which the thermodynamically advantageous processes are nuclear reactions in which nuclei capture electrons with the emission of neutrinos. The charge of the nucleus decreases, its mass remaining the same, which results in a decrease in the binding energy and, correspondingly, in the mass defect.

At still higher densities and pressures nuclei continue to capture electrons with a further reduction of their charges. Finally, nuclei containing too many neutrons become unstable and disintegrate. At densities of $\sim 10^{11}$ g/cm³ and pressures of $\sim 10^{24}$ atm neutrons begin to predominate in number over electrons, and at densities of $\sim 10^{12}$ g/cm³

¹ Osmium, the densest substance on earth, has a density of 2.2×10 g/cm³ at normal temperature and atmospheric pressure.

they begin to predominate also in the pressure produced by them. This is the domain in which matter may be regarded basically as a neutron Fermi gas (of course, a number of protons and electrons produced by neutron decay will always be present).

At $\sim 10^{27}$ atm the density of the neutron gas is equal to nuclear density, i. e., $\sim 10^{14}$ g/cm³. This state of matter may be maintained up to temperatures of $\sim 10^{120}$ K.

The states of matter corresponding to very high temperatures and densities exist naturally in stars. Here are some figures. Temperatures in the central regions of the sun are in the neighbourhood of 10^7 °K, at a pressure of about 10^{11} atm. In these conditions matter is in a state of completely ionised plasma, i. e., it consists of bare nuclei and electrons. Inside giant stars the temperature is $\sim 10^8$ °K. The highest pressures are found in so-called white dwarfs— 10^{19} atm¹—and matter there is in a state of relativistic degenerated electron gas.

Stars are natural physical entities characterised mainly by electron-nuclear states of matter. Neutron states of matter, in which practically all electrons have been captured by protons and matter may be regarded as a degenerated neutron gas, are in all probability not to be found naturally in any celestial bodies. However, given sufficient mass, the neutron state ought to be thermodynamically the most advantageous as the overall energy of a system in neutron state is less than in electron-nuclear state. For, although transformation of nuclei and electrons into free neutrons involves considerable expenditure of energy, if the mass of a system is large enough this expenditure is more than made up for by the liberation of gravitational energy due to the decrease in size and increase in density of the system.

If consideration is given to the “expanding universe” effect,² visible manifestations of which are the “fleeing”

¹ The density of these stars is $\sim 10^6$ g/cm³, i. e., one cubic centimetre of their matter would weigh several tons in terrestrial conditions.

² This effect, it will be recalled, is a corollary of general relativity theory.

galaxies, it seems highly probable that neutron states of matter may have existed in the past.

In connection with the problem of the origin and distribution of chemical elements it could be noted that conditions for nuclear production are most likely inherent in the conditions for the origination and evolution of the "stellar state" of matter.

Chapter 10. Physical Kinetics

Up till now we have been assuming that the macroscopic systems under investigation were in statistical equilibrium. In other words, the time intervals over which we considered them were long in comparison with their relaxation times.

Commonly, however, it is necessary to examine a system over a time comparable with, or even less than, the relaxation time. For large systems this is possible owing to the existence, in addition to the complete statistical equilibrium of a closed system as a whole, of so-called partial equilibria.

Partial equilibrium The relaxation time increases with the dimensions of a system. Accordingly, individual small parts of a system come into equilibrium much sooner than equilibrium is established among all of them.

Another type of partial equilibrium is due to differences in the speeds of the various processes going on in a system as a whole. A graphic example of this is the partial equilibrium in a mixture of several chemically reacting substances. Thanks to the comparative slowness of chemical reactions, equilibrium in regard to molecular motion sets in much sooner than equilibrium in regard to the reciprocal transformations of the molecules, i. e., to the composition of the mixture. Thus partial equilibria in mixtures can be regarded as equilibria for given chemical compositions (which are actually not in equilibrium).

Macroscopic description again Existence of such incomplete equilibrium makes it possible to introduce the concept of *macroscopic states* of a system (already widely used before in this book). Unlike mechanical microscopic description (i. e., statement of the coordinates and momenta of all particles

of a system if the classical description of the particles is adequate), in *macroscopic description* the mean values of the physical quantities determining this or that state of partial equilibrium are given. These can be, for example, the mean values of quantities characterising individually very small, but macroscopic, parts of the system, each of which can be regarded as being in a particular equilibrium.

Kinetic equations Consider such a sufficiently small subsystem. Due to interactions with other small subsystems, it may not be in equilibrium. Consider the behaviour of its nonequilibrium distribution function, which gives the distribution of the particles of the subsystem according to the latter's quantum states. The change in this function with time is determined both by the spontaneous microprocesses of transition from one state to another inside our quasi-isolated system and the external forces acting on the subsystem. This relationship, which determines the time-dependence of the distribution, is called a *kinetic equation*.

The *kinetic properties* of a body will, evidently, be wholly determined by their specific microstructure. It follows from this that the specific form of the kinetic equations must differ for different bodies.

The kinetic equations are the basic equations of the theory of physical processes involving macroscopic systems, which is known as *physical kinetics*.

Kinetic properties of bodies We have seen that in the theory of statistical equilibrium there exists a universal distribution function—Gibbs' distribution—which provides a description of the equilibrium statistical properties of *any* macrosystems. This means that there exists a large number of relationships among macroscopic quantities which are valid irrespective of the specific bodies to which the quantities refer (thermodynamic relationships).

In kinetics there is no universal distribution function; the change of nonequilibrium states with time is wholly dependent on the specific microstructure of a body (but, of course, as the state approaches equilibrium the distribution functions gradually turn into Gibb's equilibrium function). As a result, there are no general kinetic

equations equally applicable to all bodies, with the exception of a few limiting cases (e. g., ideal gases). Accordingly, too, barring a *few exceptions*, there exist no general quantitative macroscopic connections between kinetic quantities.

**Mechanics and
electrodynamics
of continuous
media**

Phenomenological kinetics, which is one of the methods of investigating kinetic phenomena, considers macroscopic systems as continuous media, completely abstracting from their microstructures and operating exclusively in macroscopic quantities. Here belongs the vast domain of mechanics of continuous media: the *theory of fluid motion (fluid mechanics)* and the *theory of elasticity*. Here also belongs the *electrodynamics of continuous media*, i. e., the theory of electromagnetic fields in material media and the theory of the macroscopic electrical and magnetic properties of matter. The basic relationships of phenomenological kinetics find substantiation in statistical kinetics.

Conclusion

In this discourse we have dealt with properties of various physical entities, from elementary particles, nuclei and atoms to macroscopic bodies, stars, galaxies and the universe as a whole. In studying the properties of these entities we treated them as *given*, without going into their origin. This is not accidental, for within the framework of existing concepts such questions cannot be posed.

It seems likely that questions of the *origin* of all physical entities, from elementary particles to galaxies, can be reasonably posed (or will arise) only after a complete system of ideas has been developed in the physics of elementary particles and after these ideas are *correlated* with the ideas of general relativity.

INDEX

- Absolute zero**, 121
Acceleration, 17, 20
Action and reaction, law of, 24
Action at a distance, 20
Action quantum, 73
Amplitude of state, 96
Amplitude, wave, 50
Angular momentum, 25, orbital, 90
Annihilation, 102
Antibaryon, 102
Antilepton, 102
Antimeson, 102
Antiparticles, 102
Atom, 93; nucleus, 107; phenomena, 71

Baryon, 102; charge, conservation of, 104
Binding energy, 41
Boltzmann's constant, 118
Born, Max, 93
Bose-Einstein (Bose's) statistics, 92

Cartesian coordinates, 17
Causality, 85
Charge, baryon, 104; -conjugation, 95; conservation of, 46; electric, 44; -multiplet, 105; -renormalisation, 99
Clapeyron's equation, 120
Classical electrodynamics, 43, 47
Classical electron radius, 47 fn
Classical mechanics, 18
Classical motion, 17

Classical relativity principle 19
Clock, 15
Closed systems, 21
Combined parity, 105
Composition of velocities, 20, 37
Configuration space, 75
Conservation laws, 16, 21, 25, 48, 64, 87, 104, 114; angular momentum, 25, baryon charge, 104; charge, 46; energy, 26; isotopic spin, 104, mass, 24; momentum, 21; parity, 88; strangeness, 104
Constant, Boltzmann's, 118; gravitational, 67 fn; Planck's, 73 fn
Constrained particle, 83
Coriols inertia force, 28
Coulomb's law, 44
Coordinate system, 16, 17; Cartesian, 17, curved, 57
Curved space, 59

De Broglie wavelength, 82
Degenerated gas, 122
Degrees of freedom, 17, internal, 90
Density matrix, 116
Diffraction, of particles, 72; of waves, 52
Dirac, Paul, 93
Dirac's equation, 95

Eigenfunction, 77
Eigenvalue, 77
Einstein, Albert, 61
Einstein's principle of relativity, 30

- Electric charge, 44
- Electric field, 44
- Electrodynamics, classical, 43, 47; conservation laws in, 48
- Electromagnetic field, 43; quantum theory of, 96
- Electromagnetic waves, 49
- Electron, 102; classical radius, 47 in; -positron field, 97; -positron polarisation, 97
- Electrostatic field, 44
- Elementary particles, 41, 102
- Energy, binding, 41; conservation of, 26; internal, 40; -operator, 80; potential, 26; relativistic, 39; rest, 39; "self"-, 46; zero, 83
- Entropy, 116
- Equation of state, 118
- Equations of motion, 18, 23, 86
- Equilibrium, partial, 125; statistical, 112
- Equivalence, principle of, 54
- Euclidean space, 16
- Event, 30
- Exchange effect, 93
- Exclusion principle, Pauli's, 93

- Fast processes, 104
- Fermi-Dirac (Fermi's) statistics, 92
- Field, 25, 43; electric, 44; electromagnetic, 43; electromagnetic, quantum theory of, 96; electron-positron, 95; electrostatic, 44; gravitational, 54, 60, 63; -intensity, 45; of force, 28, 43; photon, 95; potential of, 45; quantised, 62, 95; -strength, 45
- Force, 22; centrifugal inertia, 28, Coriolis inertia, 28; field of, 28, 43; inertia, 28
- Four-dimensional space, 30 in
- Frame of reference, 16, 17, 58; inertial, 18, noninertial, 27
- Free motion, 18, 60
- Free particle, 82
- Frequency, 50
- Function, antisymmetrical wave, 92; Lagrangian, 65; potential, 22, statistical distribution, 111; symmetrical wave, 92; wave, 75

- Galilean transformations, 19
- Galileo's relativity principle, 19
- Gas, degenerated, 122; ideal, 119; neutron Fermi, 123
- General relativity, 67; conservation laws in, 64; length in, 58; time interval in, 58
- Geometrical optics, 51
- Gibbs' distribution, 114
- Gravitation, Newton's law, 67
- Gravitational constant, 67 in
- Gravitational field, 54, 60, 63
- Gravitational interactions, 103
- Gravitational mass, 65
- Gravitational waves, 66

- Hamiltonian, 80
- Heisenberg, Werner*, 93
- Helicity, 105
- Homogeneity, of space, 16, 68; of time, 16, 69
- Hyperon, 102

- Ideal gas, 119
- Indistinguishability of identical particles, 91
- Inertia force, 28
- Inertia law, 19, 22
- Inertial frame of reference, 18
- Inertial mass, 66
- Instrument, 74
- Interaction, 30; at a distance, 43; electromagnetic, 102, gravitational, 103; particle, 102, point, 98; velocity of, 29; strong, 102; weak, 103
- Interference, of probabilities, 76; of waves, 51
- Internal degrees of freedom, 90
- Internal energy, 40
- Irreversibility of time, 87
- Irreversible processes, 117
- Isotopic invariance, 104
- Isotopic spin, 104
- Isotropy, space, 16, 69; time, 27, 69

- Kinetic equations, 126

- Lagrangian function, 65
- Law of action and reaction, 24
- Law of increasing entropy, 116
- Least action, principle of, 65
- Length, 15, 35; in general relativity, 58; Lorentz contraction of, 36; proper, 36
- Lepton, 102
- Light ray, 52
- Limiting velocity of propagation, 29
- Linear momentum, 21
- Liouville's theorem, 113
- Liquid, 120; quantised, 121
- Lorentz contraction, 36
- Lorentz transformations, 32

- Macroscopic description, 126
- Macroscopic quantities, 112
- Macroscopic systems, 108
- Mass, 21, 22, 24; conservation of, 24; -defect, 41; gravitational, 65; inertial, 66; -renormalisation principle, 99
- Material point, 16
- Matter, states of, 119-121
- Maxwell's equations, 48
- Measurement, 74, 77, 84
- Measuring rod, 15
- Mechanics, classical, 18; quantum, 74; relativistic, 30, 39
- Meson, 102
- Mirror symmetry, 88
- Momentum, 21; conservation of, 21; moment of, 25; relativistic, 39
- Monochromatic wave, 50
- Motion, classical, 17; equations of, 18, 23, 86; free, 18, 60
- Multidimensional configuration space, 75
- Muon, 102

- Non-Euclidean geometry, 57
- Non-Euclidean space, 59
- Noninertial frame of reference, 27
- Neutrino, 102, 103 fn; two-component longitudinally polarised, 105
- Neutron Fermi gas, 123
- Newton's first law, 19, 22
- Newton's law of gravitation, 67
- Newton's second law, 23

- Newton's third law, 22
- Nuclear physics, 107
- Nucleon, 102
- Nucleus, atomic, 107

- Observables, 73, 99
- Orbital angular momentum, 90
- Operator, 77; energy, 80

- Partial equilibrium, 125
- Parity, combined, 105; conservation of, 88
- Particle, 16; -annihilation, 102; -antiparticle pair, 102; constrained, 83; -diffraction, 72; elementary, 41, 102; free, 82; identical, 91; -interaction, 102; with zero mass, 40
- Pauli's exclusion principle, 93
- Period, wave, 50
- Phase path, 109
- Phase space, 109
- Phase, wave, 51
- Photon, 102
- Physical kinetics, 125
- Pion, 102
- Planck's constant, 73 fn
- Plane wave, 49
- Plasma, 121
- Point interaction, 98
- Potential energy, 26
- Potential function, 22
- Pressure, 118
- Principle of equivalence, 54; of indistinguishability of identical particles, 91; of least action, 65; of superposition, 45, 76
- Probability, statistical, 73; interference of, 76
- Propagation limiting velocity of, 29
- Proper length, 36
- Proper time, 37

- Quantised field, 62, 95
- Quantised liquid, 121
- Quantum laws, 62
- Quantum-mechanical systems, 73
- Quantum mechanics, 74; conservation laws in, 87; relativistic, 94
- Quasi-closed subsystems, 114

Radiation, wave, 52
Radius vector, 17
Ray, light, 52
Red shift, 69 fn
Relativistic energy, 39
Relativistic mechanics, 30, 39
Relativistic momentum, 39
Relativistic quantum theory, 94
Relativity theory, 67; conservation laws, 64; length contraction, 58; time dilation, 58; velocity composition, 37
Relativity principle, Galileo's (classical), 19; Einstein's (relativistic), 30
Relaxation time, 113
Renormalisation, 99
Rest energy, 39
Reversible processes, 117
Rigid body, 41

Schrödinger, Erwin, 93

Schrödinger's equation, 82

Self-energy, 46

Signal, 30

Simultaneous measurement, 77

Slow processes, 104

Solid state of matter, 120

Space, 15; curved, 59; Euclidean, 16; four-dimensional, 30 fn; homogeneity of, 16, 68; isotropy of, 16, 69; non-Euclidean, 59; three-dimensional, 15

Space-time metric, 57

Space-time separation, 30

Spin, 90; isotopic, 104

Squared interval, 31

State, amplitude of, 96, equation of, 118; of matter, 119; of a system, 17; solid, 120; stellar, 121

Statistical distribution, 110

Statistical equilibrium, 112

Statistical laws, 109

Statistical matrix, 116

Statistical physics, 108; conservation laws in, 114

Statistical probability, 73

Statistics, Bose-Einstein (Bose's), 92; Fermi-Dirac (Fermi's), 92
Stellar state of matter, 121
Strangeness, 104
Strong interactions, 102
Subsystem, 109; quasi-closed, 114
Superposition principle, 45, 76
Symmetrical wave function, 92
Systems of like particles, 91

Temperature, 118; absolute zero, 121

Tensor analysis, 61

Thermodynamic quantities, 118

Three-dimensional space, 15

Time, 15, 36; dilation of, 37; homogeneity of, 16, 69; irreversibility of, 87; isotropy of, 27, 69; proper, 37

Tunnel effect, 83

Two-particle problem, 89

Uncertainty principle, 78

Vacuum, 97

Velocity, 17; composition of, 20; limiting, 29

Wave, -amplitude, 50; -corpuscle duality, 72; -diffraction, 52; electromagnetic, 49; -equation, 80; -frequency, 50; -front, 51; -function, 75, 92; gravitational, 66; -interference, 51; -length, 50; de Broglie -length, 82; monochromatic, 50; -period, 50; -phase, 51; plane, 49; -radiation 52; -vector, 50

Weak interactions, 103

Zero energy, 83

Zero mass, 40

Zero-point oscillations, 97

TO THE READER

Peace Publishers would be grateful for your comments on the content, translation and design of the book. We would also be pleased to receive any other suggestions you may wish to make.

*Our address is:
2, Pervy Rizhsky Pereulok,
Moscow, U.S.S.R.*

Printed in the Union of Soviet Socialist Republics

