CRITIQUE OF
PHYSICS

$$\underbrace{\begin{matrix} e^2 \\ (M) \\ (m) \end{matrix} \qquad\qquad c}_{\begin{matrix} \frac{e^2}{Mc^2} \\ \frac{e^2}{mc^2} \end{matrix}} \qquad \underbrace{\qquad\qquad \begin{matrix} G \\ (M) \end{matrix}}_{\frac{GM}{c^2}}$$

L. L. WHYTE

KEGAN PAUL, TRENCH, TRUBNER & CO., LTD.

CRITIQUE OF
PHYSICS

BY
L. L. WHYTE

LONDON
KEGAN PAUL, TRENCH, TRUBNER & CO., LTD.
BROADWAY HOUSE: 68-74 CARTER LANE, E.C.
1931

PRINTED IN GREAT BRITAIN BY HEADLEY BROTHERS,
18 DEVONSHIRE STREET, LONDON, E.C.2; AND ASHFORD, KENT.

TO

LOTTE

CONTENTS

CHAPTER III

GENERAL FEATURES OF UNITARY COURSE THEORY

PREFACE

THE completion of general relativity theory (1916) and of the general methods of quantum theory (1927) leaves physical theory in a situation of special interest since these theories lack an adequate common foundation and their interrelations are still obscure.

The Critique has been written in the belief that a revision of the fundamental concepts underlying both relativity theory and quantum theory is an essential preliminary to a comprehensive theoretical synthesis and that the empirical facts required for this task are now available. The purpose of the Critique is to make this aim precise and to suggest a method which can be used in the proposed revision. Relativity and quantum theory are therefore examined and reinterpreted in the light of this new aim and method. No theoretical formulation is reached, but the analysis leads to a programme for further research.

The summary given at the start of each section will assist the reader in following the argument. References are starred * in the text and collected on page 165. The notes, referred to by numbers[1], deal with special points where it is undesirable to interrupt the main argument. Two historical appendices have been included on the characteristic features of Einstein's work in relativity theory and of the development of quantum theory. These brief surveys of the two great theoretical achievements of the period 1900-1930 appropriately link the heuristic vista of the Critique to the previous stages of physical thought.

London. December, 1930.

NOTATION

The six primary universal dimensional constants:

c	the velocity of light in vacuo.
G	the gravitational constant.
e	the electronic charge.
M	the mass of the proton.
m	the mass of the electron.
h	Planck's constant of action.

The three pure number ratios:

$\alpha = \dfrac{2\pi e^2}{hc}$	the fine-structure constant.
$\beta = \dfrac{M}{m}$	the ratio of proton to electron mass.
$\gamma = \dfrac{e^2}{GMM}$	the ratio of the electrostatic to the gravitational force between two protons.

" The idea of the measuring rod and the idea of the clock coordinated with it in the theory of relativity do not find their exact correspondence in the real world. It is also clear that the solid body and the clock do not in the conceptual edifice of physics play the part of irreducible elements, but that of composite structures which may not play any independent part in theoretical physics. But it is my conviction that in the present stage of development of theoretical physics these ideas must still be employed as independent ideas, for we are still far from possessing such certain knowledge of theoretical principles as to be able to give exact theoretical constructions of solid bodies and clocks." (Einstein. 1921.)

" The hiatus (between the recognition of the relativistic relation of interval and the assumption of four-coordinate order) probably indicates something more than a temporary weakness of the rigorous deduction. It means that space and time are only approximate conceptions, which must ultimately give way to a more general conception of the ordering of events in nature not expressible in terms of a fourfold coordinate-system." (Eddington. 1922.)

" In addition to the modifications of our ordinary space-time world required by the theory of relativity and characterised by the constant c, and to the inexactitude relations of the quantum theory symbolised by Planck's constant h, still other limitations will appear connected with the universal constants e, m, M. It is not yet possible to see what form these limitations will take." (Heisenberg. 1930.)

CHAPTER I

SPACE-TIME

§1. INTRODUCTORY

A new method and a new formulation of the aim of fundamental physical theory are necessary.

IN the great epoch of classical physics from 1600 to 1900 the space and time of local individual experience were assumed to be universally and exactly valid. Without enquiring how they were to be made universal the classical physicist projected his own space and time throughout the universe by an unconscious act of the mind and within the frame so created was able to trace out the laws of classical physics.

During the last forty years exact experiment has shown that in two respects the classical frame fails to represent phenomena correctly. Space and time measurements have been found to be subject to an inescapable relativity and inexactitude implying limitations to the classical frame which have received expression in the theory of relativity and the quantum theory. In many problems an adequate formulation is possible if attention is paid to only one of these limitations, and the physical laws are then expressed accordingly in terms of *relativistic invariants*, or alternatively of *probabilities*. In other problems and in a general exact theory both limitations should be applied simultaneously, but probabilities in their natural form are not invariants and for this and other reasons a combination of the two restrictions on classical space and time presents difficulties. No general relativistic quantum theory has been found, and it appears likely that here new methods are necessary. This view is supported by the fact that the difficulties of relativistic quantum theory are closely linked with problems such as the duality of matter and radiation, and of electron and proton, which lie in a novel region of physical theory and have been scarcely touched by current methods.[1]

But though relativity and quantum theory leave many problems unsolved they have led to the recognition that the absolute classical frame of space and time was an unconscious assumption, i.e., a pure mental creation unsupported by empirical methods which could be used to establish it. This assumption was unrecognised and therefore left unquestioned until experiment revealed that nature has not provided man with the instruments which would be necessary to extend his local space and time universally and exactly. In place of signals of infinite velocity and fineness we have only light and electrons of finite velocity which cause a finite disturbance in the system they are used to observe. These experimental discoveries led to the realisation that the absolute distinction of extended space and time and the conception of exact position in space and time were unconscious generalisations unsupported by experience. The methods based on the realisation of these two facts may now have reached the limit of their range. If this is so a further advance in physical theory can be founded on the radical recognition that the four-coordinate frame which is still retained in relativity and quantum theory is itself a mental creation which, as we shall see, can neither be justified as empirically necessary nor be usefully applied to certain urgent theoretical problems. Absolute space and time and exact determination in space and time have been eliminated as illegitimate assumptions ; there still remains the assumption of a world describable with four metrical coordinates. It is probable that the deep problems with which physical theory is now confronted arise from uncritical use of a descriptive method which can only provide a clumsy and approximate representation of the essential structure of space-time phenomena.

The Critique is therefore based on the view that the methods of relativistic macroscopic field theory and of general quantum theory have already exhausted the fundamental problems for which they are appropriate—there remain to them many fields of application to special complex systems—and that new methods are necessary which must involve a further severe modification of the classical frame of space and time. The aim is to discover a more general limitation of the classical frame which reduces in different special cases to the relativistic and quantum limitations associated with c and h respectively and at the same time accounts for the appearance of the constant e. (See quotation from Heisenberg, p. xi.) This new method is to be reached by a final elimination from fundamental theory of the assumption of a four-dimensional world describable in terms of four metrical coordinates.

It will be convenient to sketch briefly the kind of method which it is proposed to substitute for the use of four metrical coordinates as being capable in principle of satisfying the above demands. Whether it can in fact satisfy them is not yet known; the method is put forward as worthy of further study and as permitting the analysis of current physical theory from a new angle.

Classical theory used four coordinates which serve as the independent variables of the differential equations of motion. The use of these four coordinates implies a distant simultaneity concept supported by a physical criterion of simultaneity. The new method proposes to take account of the fact that in an exact structural theory time-measurement can be treated as derived from space-measurement, and a dependent time variable is therefore to be used to determine the observed space-time coincidences without extension as a universal time coordinate. Distant simultaneity, time as an independent variable, and the use of four-coordinate laws are eliminated. Their place is taken by a theory of the structure of clocks which yields a dependent or derived time-parameter used only for determining coincidences. Instead of using motions to measure time and then extending this time as a universal coordinate to be applied in the expression of the laws of motion, the derived time parameter is only to be used locally to determine the occurrence of coincidences. The purpose of the Critique is to justify the consideration of so far reaching a modification of classical methods. But it may be seen at once that a method of this kind might well degenerate in special cases to the use of four coordinates with definite restrictions on the independence of the time coordinate, and therefore on the permissible motions, of the kind expressed in the relativistic metric and in the quantum conditions.

The proposed transformation may be restated as follows. The discussions of measurement given by Einstein in the introduction to the general theory of relativity and by Heisenberg in connection with the principle of inexactitude leave two important problems untouched. The first concerns the structure of rods and clocks. These were deliberately treated by Einstein (see quotation on page xi) as primary unanalysable conceptions in order to allow the development of field theory to proceed in the absence of a theory of the structure of matter. Yet we shall see that an exact structural theory of rods and clocks reveals an unequal status in time and space measurements which was provisionally and successfully neglected by relativity theory, and leads to the substitution of a derived time parameter for an independent time coordinate. The second problem concerns the ordering of space-time events implied in the use of a four-coordinate system. The postulation of a metrical relation of interval between events is alone inadequate to justify the use of coordinate systems (see quotation from Eddington, page xi). The question therefore arises whether the conventional assumption of a four-coordinate frame will continue to provide the correct notation for a physical theory which includes a treatment of the structure of rods and clocks. The argument of the Critique will suggest that it does not, and will indicate along what lines an improved method should be sought.

The theory which is to be based on the new method must comprise a theory of structure and reduce in special cases to the concepts and methods of coordinate theory, i.e. relativity theory and quantum theory are to be shown as special cases of a more general limitation on the classical frame. This definition of the aim of the proposed method is not adequately precise to be of heuristic value, and our first task will therefore be to give an exact definition of the degree of unification which should be demanded of a comprehensive theory (§2). The remainder of this chapter is occupied with an analysis of space-time measurement and space-time order from the point of view of a structural theory, and leads to the recognition that there is no objection in principle to the development of a descriptive method which does not use four space-time coordinates (§13). In Chapter II the main results of coordinate theories are discussed on the assumption that the proposed method does in fact provide the correct basis for a unified physical theory. In Chapter III various general problems of scientific concept and method which arise in connection with the proposed modification of the basis of physical theory are examined and the Critique concludes with a programme for further theoretical research.

§2. UNITARY THEORY

The aim of synthetic research is defined.

The aim of fundamental theoretical research is here interpreted as a synthesis which represents gravitation, electromagnetism, and quantum phenomena as aspects of one universal type of process in whose definition no arbitrary numbers appear. A theory providing this radical fusion of phenomena hitherto treated as largely independent will be called a *Unitary Theory.*

The ideal form for a unitary theory is a deductive theory in which the different expressions describing gravitational, radiative, proton and electron systems and their interactions are derived from a minimum of independent assumptions. Moreover these assumptions or axioms cannot be such as to permit the appearance of two (or more) theoretically independent magnitudes of the same kind (e.g. two primary universal lengths) since the ratio of these would constitute an arbitrary number. The prohibition of arbitrary numbers means that all lengths, times or masses appearing in the deductive theory are to be expressible as known functions—not containing arbitrary constants—of *one* primary length, time or mass. This means, for example, that in the structure of nature only one standard of length is involved, whereas in current theory, four independent natural units of length appear in the description of systems of different kinds and no satisfactory derivation has yet been given of the three arbitrary pure numbers which are given by their ratios. We shall see in §7 that the six primary universal dimensional constants yield three pure numbers α $\left(=\frac{2\pi e^2}{hc}\right)$, β $\left(=\frac{M}{m}\right)$, and $\gamma\left(=\frac{e^2}{G.M.M.}\right)$. A unitary theory must therefore provide a theoretical derivation of α, β and γ.

Since a unitary theory covers quantum phenomena it must provide a theory of the structure of matter, including the structure of rods and clocks, i.e. a theory of measurement showing under what conditions and to what accuracy measurement is possible. It may therefore also be defined as follows :—

A UNITARY THEORY IS A DEDUCTIVE THEORY OF MEASUREMENT AND OF THE RELATIONS HOLDING

BETWEEN MEASURED QUANTITIES—BASED ON A THEORY OF THE STRUCTURE OF RODS AND CLOCKS—IN WHICH ALL THEORETICALLY SIGNIFICANT LENGTHS, TIMES OR MASSES APPEAR AS FUNCTIONS—NOT CONTAINING ARBITRARY NUMBERS—OF ONE PRIMARY LENGTH, TIME, OR MASS.

A unitary theory involves a higher degree of unification than has been demanded from unified field theories, such as Einstein's *Einheitliche Feld-theorie* (1928-29), and goes essentially beyond the formal correlation of unmodified gravitational, electromagnetic, and quantum laws. Such correlations while valuable as preliminary steps tend to conceal the possibility of the more profound fusion which is here regarded as the ultimate aim of synthetic physical theory. The definition given to the term "Unitary Theory" represents the radical and precise formulation of the aim of unification which inspires the theoretical physicist. It also satisfies the criterion of a significant synthesis of partial theories, by supplying more information about the partial phenomena than is given by the partial theories themselves.

The possibility of a unitary theory will be assumed, for heuristic purposes, throughout the Critique. We are not concerned with the question—at present unanswerable—whether this assumption is correct, for the heuristic value of an assumption is independent of its truth or falsehood. The precise formulation of a far-going hypothesis often leads to unexpected advances through which the hypothesis itself may even come to be regarded as either false or irrelevant. In the Critique the intention is to reveal the heuristic value of this assumption in suggesting definite lines for research in a new region; if arguments are given showing that the assumption is plausible, or even likely to be correct, no attempt is or can now be made to prove this.

The formulation of the aim of a unitary theory also serves to set a standard for current theory. For example it suggests that in the absence of a physical theory with the power and range of a unitary theory current cosmological theories, though useful, involve a degree of extrapolation which may well be mistaken.

As a result of an analysis of measuring rods, clocks and coordinate systems we shall be led (§14) to extend this heuristic assumption to the form in which it will be used in Chapters II and III :—

The correct representation of the fundamental structure of physical phenomena is given by a unitary theory whose axioms do not involve the use of four extended metrical (space-time) coordinates.

§3. THE STRUCTURE OF MEASURING RODS

Any theory of the structure of matter (or of radiation) must use at least one universal length, or its equivalent.

A unitary theory must provide a theory of the structure of matter and cannot treat rods and clocks as ultimate unanalysable conceptions. It must give a theory of the structure of rods and clocks and of the conditions under which these material systems may be used as instruments of measurement. In this it differs essentially from macroscopic field theory which could proceed without specific hypotheses regarding the structure of rods and clocks. In 1921 Einstein called attention to this fact about relativistic field theory in the passage quoted on page xi. But these words were written ten years ago and it is no longer true that we are far from being able to give an exact theoretical construction of a solid body or a clock. The rapid advance of quantum theory has transformed the situation, and by revealing in this passage the point at which relativistic concepts must ultimately require further analysis Einstein has provided the clue to the correct correlation of relativity and quantum theory. We shall see that the necessary common ground is to be found by an appropriate extension or generalisation of the fundamental concepts of both theories.

Quantum theory has also been able to reach its present advanced stage without a complete discussion of the structure of rods and clocks. The methods of wave mechanics have brought us within sight of an exact theoretical description of the structure of a molecule and of a rod (crystal), but in the attempt to establish a relativistic quantum theory the point has been reached where an exact analysis of what is meant by a time coordinate, and hence by a clock, is an essential preliminary to the correct formulation of any problem. Current quantum-mechanical methods break down at the problem of a relativistic quantum theory. Thus if relativity theory and quantum theory are each contrasted with the ideal comprehensive physical theory, the range of both is seen to be limited by the absence of a complete structural theory of rods and clocks. Einstein deliberately accepted this limitation in order to permit macroscopic field theory to attack the problems of relativity without waiting for a complete theory of structure, while quantum theory in its primary preoccupation with the ordering of new empirical data has continually postponed the fundamental question of the structure of a clock, until in relativistic quantum theory it can no longer be escaped. It seems that the moment anticipated by Einstein has now arrived when further advance must be based on a complete theory of structure. A preliminary discussion of the structure of rods and clocks in relation to current theory (§§3-9) will lead on to the closely related problem of the use of four metrical coordinates.

The treatment of structure is most conveniently begun by an analysis of the structure of measuring rods and of their application in the process of measurement, and for this we may choose matter in the perfect crystalline state. The perfect crystal represents the theoretically simplest type of measuring rod, and the deviation of the instruments in actual use from this theoretical ideal does not affect the present argument.

An atomic or structural theory, whatever its mathematical methods, must treat a crystal as an arrangement of a finite number of atoms (nuclei), in which neighbouring pairs of atoms, at a constant and adequately low temperature, are in stable equilibrium at finite distances from one another. These equilibrium distances are the same for all energetically and chemically similar pairs. A theory of the structure of matter must yield these distances (crystal lattice constants) as functions of the quantities determining the system, i.e. as functions of (1) pure numbers representing the nuclear charges in terms of the electronic charge, and (2) the universal atomic constants. Thus *any* theory of crystalline structure must yield equations of the following or equivalent dimensional form :—

Lattice constant $= N.L$, where N is a pure number, and L is a characteristic length yielded by some combination of the universal atomic constants. WITHOUT A UNIVERSAL STANDARD OF LINEAR SCALE PROVIDED BY UNIVERSAL ATOMIC CONSTANTS IT WOULD BE IMPOSSIBLE FOR A STRUCTURAL THEORY TO REPRESENT THE DEFINITENESS AND CONSTANCY OF SIZE OF A CRYSTAL, OR OF A MEASURING ROD, AND HENCE TO DESCRIBE THE PROCESS OF SPATIAL MEASUREMENT.

The necessity for a standard of linear magnitude in atomic theory has been recognised by individual workers ever since this theory received mathematical expression. Thus Boscovich in 1758 realised that in substituting inverse square fields of force around a point-centre for the earlier hard material particles it was necessary to assume that the forces changed sign at some definite distance, i.e. became repulsive for smaller distances. The first experimental estimate of atomic size was probably that of Young (1805) who deduced from observations on the cohesion of liquids that the individual molecules had a diameter of the order of $10.^{-8}$ cms. These facts received no theoretical treatment, though Larmor* in

1900 called attention to the great importance of the definiteness of molecular size for fundamental dynamical theory.

This neglect by theory of an essential feature of any complete physical theory was due to the fact that the standard of linear magnitude required by a structural theory was alien to the methods of classical physics. We shall immediately examine in detail how the basic classical theories were applicable to systems of any size, and how the fundamental physical concepts were therefore unsuited for the description of structure. The only available method of transition to an atomic theory was to begin by assuming a formal correspondence with macroscopic (classical) laws, even though this meant that the necessary definiteness of size had to be brought in by arbitrary additional postulates. Thus the assumption, made by Lorentz in his electron theory, that the ultimate electrical particles obeyed the same laws as ordinary charged bodies neglected the indispensable unit of linear scale or constant with the dimensions of a length. A consequence of this is that the necessity for a natural unit or standard for determining the scale of the structure of matter, instead of receiving explicit theoretical formulation as a general difference between macroscopic and structural theories, was in fact introduced in an arbitrary and misleading form when empirical facts were collected which happened to require this feature for their description. For example, if the density of black body radiation at a given temperature is to be finite equipartition is impossible, there must be at least one maximum density at a given wave-length, and a theoretical standard of length is necessary to define this wave-length. It is an interesting historical coincidence that Larmor's suggestion that fundamental dynamical theory must be modified so as to supply this feature was made at the same time as Planck's introduction (in connection with the absence of equipartition) of the quantum constant h which, as we shall see, serves this and—in fundamental theory—no other purpose.

But in the state of dynamical theory at that time there was no possibility of the introduction of this new constant being made in a theoretically significant manner. It had therefore to be introduced by postulates, which not only were arbitrary in that they bore no relation to classical theory, but were such as to conceal the essential function of the new constant. The appearance of *h as a constant having the dimensions of action* is an arbitrary historical fact determined by the general form of the physical theories current at the time it was introduced, whereas its *empirical* function, i.e. in connection with empirical data, is in all cases except that of statistical averages to supply in combination with the other constants a universal standard of *length*. This fact is concealed by the use of generalised coordinates in general dynamics, but is of the greatest importance in connection with any revision of physical concepts. Most of the apparent conceptual problems of the quantum theory arise from the fact that this primary necessity of any structural theory (whether of matter or of radiation) had, for historical reasons, to be introduced in an arbitrary and misleading manner.

We shall now trace in detail the relation of classical, relativity, and quantum theory to the demand for the special length, defined by a universal constant having the dimensions of a length, which is essential to any theory of structure.

§4. KINEMATIC SIMILARITY IN CLASSICAL PHYSICS

The kinematic similarity of classical systems permitted (a) the definition of compound dimensional constants, and (b) their description without a universal length.

The general methods of classical theory were developed from the study of macroscopic phenomena in which no unique lengths were involved, i.e. of systems whose laws did not distinguish special systems of specific size as having unique properties. The absence of unique lengths is clearly seen, for example, in Kepler's three Laws and Newton's Law of Gravitation. Kepler's 1st and 3rd Laws assert that planets move in ellipses around the sun as focus, the periods of the planets being proportional to $L^{\frac{3}{2}}$ where L is the semi-major axis of the orbit. (Kepler's 3rd Law is exact if the planets be assumed to have equal or negligible masses.) These laws permit an infinite set of planetary orbits of different sizes geometrically similar to any given ellipse, corresponding times in these orbits being as $L^{\frac{3}{2}}$. No member of this infinite and unlimited set of theoretically possible orbits is kinematically distinguished from the others; for example, there is no finite member of the set for which the orbital velocity degenerates to zero, as for instance it would if there were a position of static equilibrium between sun and planet at some finite distance. The absence of any kinematically unique member in this set of possible orbits is shown analytically in the fact that the laws contain no constant of the dimensions of a length which could define a unique orbit of a special size. The Newtonian law of force has the form $F=\frac{G.\mu\mu'}{r^2}$, where the constant G has not the dimensions of a length.

This property of kinematic similarity is possessed by all systems obeying the Newtonian law, and was well known to Newton. It is instructive, however, to note that he considered an alternative law which provides an interesting analytical problem, and does not permit the same similarity. Newton* showed that if a given (e.g. elliptical) orbit is described under a given central force, $F=f(r)$, then the same orbit with a superimposed rotation (e.g. a rotating ellipse) can be described under the non-homogeneous law of force $F=f(r)-\frac{a}{r^3}$, (a = constant). This law, since (in general) non-homogeneous in r, must contain a constant with the dimensions of a length (or of some power of a length) and hence differs essentially from all classical laws. Orbits of different size under this law are no longer geometrically similar, kinematic similarity disappears, and there is, if c is positive, a position of static equilibrium for the orbital particle at a distance r_1 given by $f(r_1).\ r_1^3 = a$. Moreover any non-homogeneous law of force involving a polynomial in r implies the appearance of one or more constants having the dimensions of a length, permits orbits of specific size having unique kinematic properties, and eliminates kinematic similarity. Though Newton happened to interest himself in a law of this type, such cases never occur in Newtonian or classical dynamics.

In view of the importance of the classical property of kinematic similarity, and its disappearance in relativity and quantum theory, it will be convenient to define it precisely so as to include the classical radiation field :—

An infinite unlimited set of geometrically similar possible orbits (or field distributions) of different sizes in which corresponding times are proportional to a power of a characteristic length defining the linear scale of an orbit (or field distribution), i.e. in which $t \propto l^n$, n being a positive or negative number, will be called a Homogeneous Set, or H-set.

We have seen that no member of an H-set can have unique kinematic properties; hence laws describing H-sets cannot account for the definite equilibrium distances involved in the atomic structure of matter. The existence of H-sets in a given region of physics means the absence of a natural unit of length, and the elimination of H-sets can only come about through the introduction, explicitly or implicitly, of a constant having the dimensions of a length. Many forms of similarity have been discussed by different authors from Galileo onwards[23], but it is important to notice that only H-sets possess what may be called complete kinematic similarity. Thus under the non-homogeneous law just referred to there exists an infinite set of *geometrically* similar orbits (the circular orbits), but the periods of these are not proportional to a simple power of r, and the set is not homogeneous since it contains a kinematically unique member. Every law defining an H-set may be given the following form :—

$$\text{Corresponding Times} = \begin{pmatrix}\text{Dimensional constant}\\ \text{characteristic of H-set}\end{pmatrix} \times \begin{pmatrix}\text{Characteristic length defining}\\ \text{member of H-set}\end{pmatrix}^n.$$

This has the dimensions :

$$T = \begin{pmatrix}\text{Compound dimensional}\\ \text{constant involving}\\ \text{Time and Length}\end{pmatrix} \times L^n$$

THUS THE EXISTENCE OF AN H-SET PERMITS THE DEFINITION OF A COMPOUND DIMENSIONAL CONSTANT, WHOSE DIMENSIONS INCLUDE POWERS OF LENGTH AND TIME. For example, in Kepler's Third Law,

$$T = K.\, L^{\frac{3}{2}}, \text{ where } K = \frac{2\pi}{\sqrt{G.\mu}} \text{ and } \mu = \text{Sun's mass.}$$

In this case the dimensional constant (K) is not a *universal* constant, since it defines solar orbits only. The following are the three classical laws defining H-sets which permitted the definition of the compound universal constants G, c, and e :—

1. *Newton's Law* defining *Gravitational H-sets* with $t \propto l^{\frac{3}{2}}$ and yielding the compound universal constant G of dimensions $M^{+1}.\ L^{+3}.\ T^{-2}$.

2. *Maxwell's Laws*. Any solution of Maxwell's Electromagnetic equations (for vacuum) can be arbitrarily varied in scale, and hence can be taken to represent an *Electromagnetic Field H-set* of geometrically similar field distributions. In these corresponding times are proportional to the distances, i.e. $t \propto l$, and this type of H-set yields the compound universal constant with

dimensions $L^{+1}. T^{-1}$. (Even though c was not first discovered as an electromagnetic field quantity, it is always given empirically in terms of an H-set, i.e. as a velocity independent of the linear scale of the system in which it is measured.)

3. *Coulomb's Law.* This defines *Electronic (and protonic) H-sets* in the case of the open or closed non-radiating classical orbits. These have $t \propto l^{3}$ and yield the compound universal constant e with dimensions $M^{+1}. L^{+3}. T^{-2}$. (The definition of quantity of electric charge and hence the measurement of e depends on the homogeneous inverse square law of force.)

In each case the exponent n in the expression l^{n} determines the powers of length and time which appear in the dimensions of a compound constant defined by an H-set. Only the assumed existence of H-sets allowed the definition of compound dimensional constants of this type. For example, had the Keplerian orbits been found empirically to have periods proportional to $A_{\alpha}.l^{\alpha}+A_{\beta}.l^{\beta}+A_{\gamma}.l^{\gamma}+ \; . \; . \; . \; .$ where α, β, γ represent a series of at least two different indices (as in Newton's example) then gravitational phenomena could not have been summarised in one dimensional constant (e.g. A_{α}) for the constants A_{β}, A_{γ} would also be involved. Only the empirical fact that $A_{\beta}=A_{\gamma}=O$, i.e. the absence of constants $\left(\frac{A_{\beta}}{A_{\alpha}}, \frac{A_{\gamma}}{A_{\alpha}}\right)$ having the dimensions of a length (or of some power of a length), permitted the definition of a single gravitational constant G.

A compound dimensional constant defined in terms of classical conceptions thus asserts a similarity property in systems of a certain type. The other pre-quantum dimensional constants, M and m, can be defined without the assumption of H-sets because they are dimensionally simple, i.e. of the form of ratios of masses (e.g. proton mass/unit mass) and do not like G, c, or e involve more than one of the primary dimensional quantities. We thus reach the interesting observation that universal dimensional constants which either do not define an H-set, or have the effect of eliminating H-sets, must be expressible as dimensionally simple quantities, i.e. as universal constants of length, time, or mass.

The kinematic similarity of the classical H-sets shows that classical theory cannot supply the special lengths required in a structural theory.

§5. THE ABSENCE OF SIMILARITY IN RELATIVITY PHYSICS

The presence of two compound dimensional constants (in addition to M or m) in the description of relativistic systems eliminated similarity and introduced universal lengths into the equations describing orbital motion at high velocities.

We have seen that classical theory assumed the existence of three independent types of H-sets: the H-sets of gravitational orbits, of electromagnetic radiation fields, and of non-radiating electron (or proton) orbits, implicit in the definitions of G, c, and e, respectively. These three types of phenomena were independent in the sense that no unified theory combined them and that complex systems comprising two of them had to be described by special combinations of the separate theories. The radical fusion of any two of these separate regions of theory necessarily modified the H-sets previously permitted in each region. For example, if c represents a maximum velocity for all regions of physics, then the set of possible electronic orbits cannot any longer form an unlimited set with $t \propto l^2$ since the velocity in an orbit being inversely proportional to l increases without limit as l is made smaller and would exceed c for part of the set of orbits. Thus in the special theory of relativity the only H-set which can remain is that of the electromagnetic radiation fields, in which $t \propto l$ or $t = \frac{1}{c} l$. Any fusion of the separate departments of classical physics implied the elimination of two at least of the H-sets which had permitted the original definition of the constants G, c, and e. WHILE THE EXISTENCE OF COMPOUND DIMENSIONAL CONSTANTS IMPLIES A SPECIAL INVARIANCE (TO CHANGES OF SCALE) OF THE PHYSICAL LAWS, THE PRESENCE OF TWO SUCH CONSTANTS OF DIFFERENT DIMENSIONS IN ONE UNIFIED THEORY ELIMINATES THIS INVARIANCE BECAUSE IN CONJUNCTION WITH m OR M THEY PERMIT THE APPEARANCE OF CONSTANTS HAVING THE DIMENSIONS OF A LENGTH $\left(\frac{GM}{c^2}, \frac{e^2}{Mc^2}, \frac{e^2}{mc^2}\right)$.

Any unification of gravitational, radiational, and electronic phenomena must necessarily eliminate the kinematical similarity possessed by the independent classical descriptions of the three regions, as is indicated in the following diagram :

Independent Classical phenomena (without a universal length).

$$\underbrace{\begin{matrix} \text{Electrostatic actions.} & \text{Light.} \\ e^2 & c \\ (M) & \\ (m) & \end{matrix}}_{\frac{e^2}{Mc^2} \quad \frac{e^2}{mc^2}} \qquad \underbrace{\begin{matrix} \text{Gravitational actions.} \\ G \\ (M) \end{matrix}}_{\frac{GM}{c^2}}$$

Combined phenomena (with universal lengths).

We shall now examine briefly how the special and general theories of relativity, in providing notations for such fusions, eliminated the classical similarity by allowing the appearance of specific lengths in the equations of motion. Further detail is given in a note.[3]

A. *Electronic motion in the Special Theory.*

The invariance of the equation describing the H-sets of classical non-radiating electron orbits around a charged nucleus with reference to the transformation $r' = \lambda.r,\ t' = \lambda^{\frac{3}{2}}.t$ is destroyed by the relativistic variation of mass with velocity. The geometrical form of these orbits depends on the ratio of the size of the system to the universal length $(\frac{e^2}{mc^2})$ yielded by the constants e, m, c involved in the description of the system.[3] For example, the advance of the perihelion per revolution of the relativistic ellipses of the Bohr-Sommerfeld theory (neglecting quantisation) is not the same for orbits of equal eccentricity but different size, and is a function of the ratio major axis$/(\frac{e^2}{mc^2})$.

Classical field theory ascribed to the electron a radius equal to $K.\ \frac{e^2}{mc^2}$ where K is a constant to which slightly varying values were given by different theories. In recent theory the length $\frac{e^2}{mc^2}$ appears not as the radius of a spherical electron but as a factor involved in the relativistic description of electron orbits.

Similar arguments apply throughout to the motion of protons in the special theory.

B. *Planetary motion in the General Theory.*

The relativistic equation for a solar orbit with negligible planetary mass is not, like the corresponding equation for a Newtonian orbit, invariant under the transformation $r' = \lambda.r,\ t' = \lambda^{\frac{3}{2}}.t$.[3] In the case of a closed orbit the advance of perihelion depends on the size of the orbit as well as its eccentricity, and for orbits of equal eccentricity is a function of the ratio, major axis/solar constant of length $\frac{G\mu}{c^2}$ (μ =Sun's mass).

In the special case of an open orbit given by the deflection of light in the solar gravitational field the angular deflection depends on the ratio of the shortest distance from the sun to the solar length $\frac{G\mu}{c^2}$.

In the cosmological theory a new constant is introduced which implies the presence of a cosmic length, $\sqrt{\frac{3}{\lambda}}$ (called the radius of curvature of the world), in equations resulting from the modified law $R_{ik} = \lambda . g_{ik}$.

The introduction of c into electronic and gravitational phenomena eliminated the H-sets which had permitted the original definition of G and e and the modification of the meaning of space-time measurements resulting from the theory of relativity provided what were effectively new definitions of these constants. In the special theory this was achieved postulating the constancy of c and the special invariance of the laws and accepting the consequences: the Lorentz transformations for length, time, and mass. The resulting variation of electronic mass with velocity involved the equivalent of a new definition of e, and permitted this constant to be retained though the electronic orbits no longer formed an H-set. In the general theory it was shown that the proportionality of mass and weight, and hence G, could be retained (to an adequate approximation) in a generally covariant theory. But this meant the giving up of the last remaining H-set, and allowing the value of c, as measured in a given coordinate system, to depend on the coordinate system, the gravitational field, and the direction of propagation.

Thus the theory of relativity eliminates the H-sets of classical physics and provides three universal lengths (derived from the constants G, c, e, m, M) which appear explicitly in the relativistic equations describing gravitational, electronic, and protonic orbits of unit masses and charges. (In the case of other masses or charges they appear with numerical factors.) Yet none of these three lengths are appropriate for the description of the static or stationary directly measurable lengths involved in the structure of matter, for they appear only in the relativistic description of motion at high velocities. Thus another universal length is necessary for a structural theory.

§6. THE UNIVERSAL LENGTHS IN QUANTUM PHYSICS

The quantum constant h in combination with the other atomic constants defines the linear scale of all directly measurable structures in matter or radiation.

The primary function of Planck's constant h, in combination with the atomic constants, is to provide physical theory with standards of length in terms of which all directly measurable lengths of theoretical importance can be described. Such directly measurable lengths are the sizes of molecules or crystal-lattices, and the wave-lengths of light, electrons or protons produced under known conditions. Wherever h appears in a fundamental equation (not referring to statistical assemblies), the empirical content of the equation is the correlation of a *measured length* with other quantities. (The special case of the inexactitude relation is treated in §10.) Equations that contain h and do not at first sight appear to determine a measurable length have to be transformed into assertions about measurable lengths (wave lengths, crystal lattice constants, etc.) before they have a direct relation to experiment. This fact is concealed by the original introduction of the constant as a unit of action, but any revision of physical concepts must render explicit the primary function of h:

PLANCK'S CONSTANT (IN APPROPRIATE COMBINATIONS) DETERMINES THE LINEAR SCALE OF THE STRUCTURE OF MATTER AND OF RADIATION, in terms of the selected unit of length, the cm.(4)

This is evidenced by the fact that empirical determinations of h always involve the (direct or indirect) measurement of a length involved in the structure of matter or of radiation. This length need not be a wave length; h can in principle be calculated from the empirical values of the nuclear distances in the hydrogen molecule in a given quantum state, or from any lattice constant as soon as theory is slightly further advanced. (See table below.) The purpose of Planck's constant for fundamental theory would have been more directly represented had a constant having the dimensions of a length been introduced instead. For example instead of $\epsilon=h\nu$, defining h, the exactly equivalent equation $\frac{\epsilon}{mc^2}=\frac{l}{\lambda}$ defining an universal length $l\ (=\frac{h}{mc})$ might have been used if the importance of the energy mc^2 had already been discovered.

Fundamental quantum theory, i.e. the system of equations involving h, may usefully be regarded as the theory of *structures of definite scale* and classical theory as the approximate theory of those processes where the scale of the structure of matter and radiation can be neglected. Planck's constant serves to eliminate finally the similarity properties of classical laws, which still remained approximately valid in relativity theory so long as high velocities were not involved. In addition to this positive function h has also the negative function of limiting the exactitude of coordinate description, which will be dealt with later. The following table gives a summary of all the empirical evidence for Planck's constant —excluding that based on statistical assemblies—showing in each case the particular combination of h with the other constants which is used to provide a constant having the dimensions of a length. (The usual numerical factors are retained.)

Wave-lengths of light.

Characteristic Line Spectra	$\frac{h^3c}{2\pi^2me^4}$	(Rydberg wave length).
Relativistic or spin fine structure	$\frac{h}{mc}$	$(\delta\lambda=f(\alpha^2))$.
Wave lengths in Photo-electric effect	$\frac{h}{mc}$	$(\lambda=\frac{h}{mc}.\frac{2c^2}{v^2},$ for small $v)$.
Shifts in Compton effect	$\frac{h}{mc}$	$(\delta\lambda=\frac{h}{mc}.2\sin^2(\frac{\theta}{2}))$

Wave-lengths of electrons and protons of given momentum.

Electrons	$\frac{h}{mc}$	$(\lambda=\frac{h}{mc}.\frac{c}{v},$ for small $v)$.
Protons	$\frac{h}{Mc}$	$(\lambda=\frac{h}{Mc}.\frac{c}{v},$ for small $v)$.

Sizes of molecules and crystal lattice constants.

$$\left.\begin{array}{l}\text{All molecular sizes}\\ \text{All lattice constants}\end{array}\right\}\frac{h^2}{4\pi^2me^2}$$

(e.g. the nuclear separation in the normal H_2 molecule is obtained by Heitler and London* from an equation in ρ, where $\rho=\frac{R}{r_o}$, with R=nuclear separation, and r_o=Bohr's first radius.)

It is important to note that of the three universal lengths involving h :— $\frac{h}{mc}$, $\frac{h^2}{4\pi^2me^2}$, $\frac{h^3c}{2\pi^2me^4}$ only two are independent, i.e. they can all be expressed as functions of any two of them, or of any one of them and the pure number ratio $\alpha=\frac{2\pi e^2}{hc}$, if numerical factors are neglected.

In classical electron theory the Coulomb Law for the field around a nucleus permitted H-sets of possible electron orbits of any size. Bohr used h to select a discrete series of special orbits, and obtained a formula for the hydrogen line spectra in which the wave lengths were given as a simple numerical factor times one standard length $\left(\frac{h^3c}{2\pi^2me^4}\right)$. In order to describe the fine structure of line spectra it was however necessary to introduce a second standard of length $\frac{h}{mc}$ or its equivalent a pure number (α) whose square represents (but for a numerical factor 2) the ratio of these two standard lengths. The achievement common to all the later theories of the atom (from Sommerfeld to Dirac) which contain either the relativity correction, or the electron spin, or a fusion of both, is that they provide a theoretical system involving two standard lengths. Whatever physical interpretation is given to any of these theories—in terms of particles, waves, or spinning electrons—this formal property remains fundamental because directly related to the empirical facts which support quantum theory. In fact the formal possibility of Dirac's reduction of relativity and spin fine structure to one mathematical form is a result of this feature common to both types of correction: they both introduce a second universal length into the description of the structure of radiation.

§7. THE FOUR UNIVERSAL LENGTHS IN UNITARY THEORY

A unitary theory must reduce the four independent universal lengths obtainable from the six primary constants to one, by providing a derivation of the three arbitrary numbers, α, β, γ.

We have seen that relativity theory eliminated the three types of H-sets of classical physics by fusing phenomena previously treated as independent and thus permitting the appearance of three relativistic universal lengths:

$$\frac{GM}{c^2}, \qquad \frac{e^2}{Mc^2}, \qquad \frac{e^2}{mc^2}.$$

Neither these lengths nor multiples of them are ever directly measured, though they are involved in the relativistic equations governing motion at high velocities.

Quantum theory, on the other hand, eliminates the classical H-sets by introducing a new constant which, with the others, yields the directly measurable lengths found in the structure of matter and of radiation. These lengths are

$$\frac{h}{Mc}, \qquad \frac{h}{mc}, \qquad \frac{h^2}{4\pi^2me^2}, \qquad \frac{h^3c}{2\pi^2me^4}.$$

Thus an inner formal similarity of the two theories which contrasts them with classical theory is that they both involve the appearance of universal lengths in physical law. This fact can be regarded as underlying the discovery of a common mathematical origin for relativistic and spin fine-structure in spectra. Any revision of basic physical concepts will have to provide a reinterpretation of this fact, and a comprehensive physical theory must account for these seven universal lengths.

Of these seven, which all appear in current theory, only four are independent, i.e. four can be chosen so that the other three are expressible as simple functions of them with non-arbitrary numerical factors (2π etc.). For example, we can select the four lengths $\frac{GM}{c^2}$, $\frac{e^2}{Mc^2}$, $\frac{e^2}{mc^2}$, $\frac{h^2}{4\pi^2me^2}$, which are connected by the three non-dimensional ratios, γ, β, and α^2.

A unitary theory was defined (§2) as a comprehensive theory in which no arbitrary numbers occur. This implies that all lengths appearing in the deductive theory are given as functions—not containing arbitrary constants—of one primary length. Thus in a unitary theory these four lengths, though now independent, have to be expressed as functions of one length, i.e. a unitary theory has to provide a derivation of the pure numbers α, β and γ. The definition may therefore be re-written:

> A UNITARY THEORY IS A DEDUCTIVE THEORY OF THE POSSIBILITY OF MEASUREMENT AND THE RELATIONS HOLDING BETWEEN MEASURED QUANTITIES, INCLUDING A DERIVATION OF THE NUMBERS α, β, γ.

A unitary theory may be regarded as based on the elimination of redundant dimensional constants. It was seen in §5 that the presence in one theory of two or more compound dimensional constants in addition to M or m indicated the appearance of a universal constant of length. Therefore the presence of an excessive number of constants may be regarded as resulting from a failure to express the laws more simply in terms of one primary universal length. The possibility of an isolated derivation of α has already been taken up by Eddington. (See §28.) There are advantages in the more general conception of a unitary theory which must provide simultaneously derivations of the three arbitrary numbers which appear in current theory.

Since the success of Hamiltonian methods in the wave theory and matrix theory of quantum mechanics, it has been clear that a comprehensive formal analogy connects the four regions of phenomena: gravitational, protonic, electronic, and radiational. Within certain limitations Hamiltonian and relativistic methods are applicable throughout, and an inverse square law is necessary in the first three regions. It is therefore natural that this formal similarity can be expressed in various mathematical notations[1] none of which adds to our knowledge of the interactions of these different types of phenomena. In contrast to these formal correlations, a unitary theory must supply a radical treatment of the interrelations of the four regions, this being implied in a theoretical derivation of ι, β, and γ.

Having found in quantum theory the special lengths necessary to any theory of the structure of matter, we are now able to continue the consideration of rods, clocks and the process of measurement.

§8. MEASURING RODS (contd.)

In an exact structural theory the conception of length measurement relative to a moving system is unnecessary; the Lorentz transformations may therefore be eliminated from fundamental theory, provided four extended coordinates are no longer used.

We have seen that the universal length necessary in any theory of the structure of matter for the description of molecular sizes and crystal lattice constants is provided by quantum theory in the form $\frac{h^2}{4\pi^2 me^2}$. This quantity is involved in the equations which determine all inter-nuclear equilibrium distances, i.e. the size of all solid material systems of known structure is given by a theoretically determinable numerical factor times this universal length. The *finitude* of this quantity is of course essential to a theory of the structure of matter (e.g. of measuring rods of finite size) and hence to any theory of measurement. For example, if the mathematical formulæ of quantum theory are transformed by allowing h to become vanishingly small, then all crystal lattice constants vanish with h, no molecules or rigid bodies of finite size exist and no measurement would be possible. Thus if the set of classical concepts be understood to include finite measuring rods the limiting form of quantum concepts as $h \longrightarrow 0$ does not exactly correspond to classical concepts. On the contrary the existence of a *finite* constant h, or its equivalent in another form, is implicit in the concept of a measuring rod with structure.

A structural theory of measuring rods has (i) to account for such rods in terms of their structure, and (ii) to describe the conditions under which their lengths (i.e. the lengths of any two rigid bodies), may be compared. One feature necessary to any theory of the structure of rods has just been discussed (§§3, 6); it will be convenient now to consider (ii). The treatment given here is only provisional, since it is subject to the analysis to be undertaken later of certain aspects of the space-time ordering of events.

Direct comparison of two rods is only possible when the rods lie together, parallel and in relative rest. Indirect comparisons of lengths (e.g. the determinations of the ratio of the lattice constants of two crystals) can be carried out in several ways, and quantum theory sets no limit in principle to the accuracy with which such a determination can be made. The Heisenberg inexactitude sets no limit to the possible accuracy of a single measurement such as the comparison of static lengths, and hence in a unitary theory the ratios of the lattice constants must be given *exactly*.

The lattice constants to be given by a unitary theory are the *proper-lengths* of the crystal units, and it is necessary to consider whether unitary theory will permit the indirect comparison of the length of bodies in relative motion, i.e. whether it will need to use the macroscopic relativistic conception of the measurement of a length relative to a moving system. In the special theory of relativity the assumption of the constancy of the velocity of light and of the invariance of the laws of physics under special coordinate transformations led to the result that different observers in different states of motion relative to a body would obtain different results for its length when measuring it indirectly by light experiments. The result of such experiments was called " the length relative to a given system." Moreover in the absence of a theory of the structure of matter such experiments provided the only method available to such observers of inferring the length of a body in relative motion. Therefore in the special theory the conclusion was unavoidable that the length of a body has to be regarded as a function of the coordinate system in which it is measured and is derivable from an invariant interval in the Minkowski four-dimensional space-time continuum, which reduces to the ordinary or proper length for systems in which the body is at rest.

This conclusion is no longer inescapable if a theory of the structure of matter is available to all observers. Each observer can then calculate the lattice constants and hence the proper length of a crystalline rod of known composition and structure. Experiments with light of the kind assumed in the special theory are no longer the only method of inferring the length of a rod in relative motion. For example, if an observer has two rods known to be identical in size and structure by previous examination, and sets them in relative motion, he can choose between two methods of description:

(i) That of the special theory, i.e. a four-dimensional invariant formulation of the lengths obtained by light experiments in different systems. This method deliberately neglects structure since the lengths obtained from such experiments differ from the (proper) lengths known to the observer from the known structure of the rods. Moreover this method is essentially macroscopic and subject to the quantum inexactitude.

(ii) That of a structural theory, which must treat the known proper lengths as primary and adequate. This method eliminates the conception of length measurement relative to systems in relative motion and hence is incompatible with the use of a postulate of the constancy of a *velocity* c in defining distant simultaneity, and with the use of four extended coordinates.

In principle both these methods are possible. They are alternative descriptive methods, which may be useful in different situations. The second method does not *conflict* with the principle of relativity because this principle is necessarily irrelevant to a theory which does not use systems of four coordinates. The question is which is the primary, more exact, or more fundamental method, and which must play a secondary, derived, or approximate rôle in a deductive physical theory? The advantages of the first method are well known. The most important of these are:

(a) It provides a revision of classical space-time following naturally on the absence of a criterion of absolute simultaneity.

(b) It yields a powerful general interpretation of the Michelson-Morley result not dependent on any particular electronic-causal theory of the structure of matter.

(c) More generally, it allows predictions to be made about rods and clocks without any assumptions about their structure, and

(d) It permits a natural generalisation to include gravitation.

On the other hand, viewed as an exact or primary method it has the following disadvantages:

(a) As has just been shown it neglects to use the full possibilities of a structural theory.

(b) It appears unsuitable for application to the structural description and equations of state of a body, since these naturally take their simplest and most significant form in a coordinate system moving with the body. (Cf. Eddington*.)

(c) Its generalisation in the General Theory is unsuitable for an exact theory of measurement based on a theory of the structure of matter. (§10.)

(d) It neglects the inexactitude of simultaneous space-time measurements. (§10.)

(e) It involves the use of a notation which cannot satisfy the demand for an invariant metrical representation of the observed topological facts. (§11.)

(f) It has so far failed to provide an adequate basis for the relativistic representation of quantum phenomena.

These facts render it probable that while the special theory provided the necessary basis for macroscopic field theory in the absence of a gravitational field it is not suitable for a theory in which microscopic structure is regarded as primary. In such a theory the exact microscopic structure of a crystal has to be made to account for such properties as the unique axes of symmetry of the atomic patterns, the values of the lattice constants, and the different characteristic values of macroscopic parameters (conductivities, coefficients of expansion, etc.) along these unique axes. These properties imply that special coordinate systems are given in the material structure in which these unique axes are at rest and in which the laws of structure therefore receive their simplest form.

The alternative method is therefore provisionally selected as fundamental, i.e. as necessary in the primary formations of a unitary theory. This method will be expected to degenerate—in application to the macroscopic study of those relative motions in which structure and quantum inexactitude are neglected and where no unique axes are immediately given in experience—to the macroscopic field laws satisfying the principles of special and general covariance. Thus there is no contradiction between the two methods, their range and accuracy are different, and the assumption of one method may even imply the approximate validity of the other. For a further comparison of the relativistic and the proposed alternative method, see[5].

But if the structural method is treated as primary and, proper lengths are regarded as adequate in basic theory, then the use of the postulate of the constancy of the velocity c in defining relative distant simultaneity and hence permitting the establishment of an extended time-coordinate can no longer play a part in basic theory. This does not mean a conflict with the empirical facts which have hitherto been interpreted as evidence for the constancy of a velocity c, but it implies that the interpretation of light experiments in terms of the concept of velocity and the attempt to use an extended time-coordinate based on the conception of distant relative simultaneity must both be given up. The primary formulations of unitary theory must not involve the conception of distant relative simultaneity which is implied in the use of an extended time-coordinate. Thus the assumption that a unitary structural theory constitutes the correct aim for physical theory suggests that the concept of velocity and the use of four coordinates are mistaken in fundamental theory. In §10 we shall be led to the same conclusion on other grounds.

The status of the Lorentz contraction in classical, relativity, and unitary structural theory may be epitomised as follows: Lorentz sought to describe the contraction causally in terms of electronic forces, Einstein showed that it followed necessarily from the fact that the invariant velocity of light provided the only adequate method of defining an extended time-coordinate, while unitary structural theory in providing an exact theory of the proper lengths of all bodies in terms of their structure renders the indirect measurement of length and hence the Lorentz contraction superfluous for exact theory.

§9. CLOCKS

A structural theory of clocks must yield their periods as a function of their size and structure, and allow time measurement to be treated as dependent on spatial measurement, i.e. provide a dependent time variable in place of the independent variable t of macroscopic theory.

In relativity theory a clock is an ultimate unanalysable conception and is defined as any system undergoing a cyclic process which can be regarded as isolated, i.e. whose environment can be treated as constant. The periods of such a system are defined as equal intervals. This definition covers

i. Rotational inertial motion.

ii. Motion in free or conditioned cyclic paths in gravitational or electrical systems.

iii. Molecular or elastic vibrations.

iv. Light propagation to and fro between fixed mirrors.

Examples of these types are: the earth (though only approximately cyclic owing to secular changes), orbital and pendulum motion, vibrating crystals, and light propagation as in Fitzeau's determination of c.

A unitary theory must provide a structural analysis of these different classes of clocks. The period of a clock of the first class depends on arbitrary initial conditions, but the ratios of the periods of clocks of the other three types must be given by a unitary theory as a function of their size and structure. Thus the proper period of a clock of one of these types can, in principle, be calculated by any observer in relative motion to the clock provided he knows its structure. Such an observer can therefore, in principle, always use proper periods and is not compelled to rely on indirect measurements by light experiments of their periods with reference to his own moving system. Expressed without reference to an observer this means that in a unitary theory the conception of the measurement of the time-period of a clock relative to a frame with reference to which the clock is in motion can be eliminated, since proper-times are given as a function of structure. But this elimination of the Lorentz transformations from exact structural theory implies, as in §8, the discarding of the conception of an extended space-time coordinate system based on the simultaneity criterion implicit in the postulated constancy of c.

Two facts render it necessary in a fundamental *structural* theory to regard time measurement as secondary to, and dependent on, space measurement :

(a) Space measurement cannot be reduced to the measurement of time without the use of rigid bodies or the enumeration of the periods of a spatial process. For example, a space-time coordinate system cannot be established by using the propagation of light alone. In addition a clock consisting of rigid bodies is essential. (Reichenbach*.) Spatial measurements all rest ultimately on the use of rigid bodies as providing a constant standard of length.

(b) On the other hand in the case of the clocks of types (ii.), (iii.), and (iv.) time measurement reduces to space measurement, and a complete structural theory must give the period as a function of the size and structure of the clock. In the case of the clocks of type (i.) time measurement consists in counting the periods of a spatial process. No cyclic or other time-measuring process is known that does not involve motion.

Thus in a strict structural theory of measurement the measuring rod not only can but must take a primary place as the basis of all measurement. IN A STRUCTURAL THEORY SPACE AND TIME MEASUREMENT HAVE DIFFERENT STATUS, SPATIAL MEASUREMENT BEING PRIMARY. This fact is concealed by the mathematical methods of macroscopic relativity theory which expressly exclude considerations of structure. But this conclusion in relation to the *measurement* of space and time does not imply that spatial relations are necessarily to be regarded as primary and temporal relations as secondary in the space-time topological *ordering* of events which underlies the possibility of measurement. (§11.)

§10. SPACE-TIME METRICAL COORDINATES

The finitude of c and h places restrictions on the combination of space and time measurements into an exact four-coordinate description of elementary processes, and the appearance of c and h in these forms may be treated as a consequence of applying four-coordinate description to phenomena better described in another manner.

The fact that space is three dimensional and that time is experienced as a single series has led to the use of four coordinates in the description of space-time observations. If it is assumed that a physical method exists by which the moment *now* (here) can be exactly correlated with single moments defined by physical events everywhere throughout space by a symmetrical criterion of simultaneity (so that if event A is simultaneous with event B, the reverse is also true) then the use of four coordinates is inescapable. Moreover since exact measurements of proper lengths and proper times are in principle possible the four coordinates must be connected with a metric.

This argument amply justifies the use of four metrical coordinates in macroscopic or approximate theory. Yet from the standpoint of a final exact theory serious difficulties arise both in general relativity theory and in quantum theory, which undermine the above argument and render it doubtful whether four metrical coordinates can provide the correct notation for the exact basis of physical theory. An examination of these difficulties will lead to a reconsideration of the supposed necessity for a four-coordinate description. In this section we shall be concerned with metrical coordinates, the non-quantitative topological order being treated in §11.

(a) *Coordinates in the General Theory of Relativity.*

All the empirical applications of the general theory refer to minor corrections to Newtonian theory in fields which may be regarded as static (one body problems). In the more general case of a two body problem it is not known whether exact solutions of the field equations exist. In hypothetical two body problems with non-static intense fields the correction to Newtonian theory is not necessarily small, the velocity of light in a given system varies with the time, the direction, and the system selected, and the non-linearity of the field equations prevents the superposition of solutions. Moreover no physical method exists in fields of this kind by which solutions of the field equations—if they do exist—could be identified with sets of physical events, since a solution of the equations only becomes of empirical significance when expressed in coordinates which have a special character of permanence permitting their identification in nature. In these general fields the *differential ds* of mathematical theory bears no exact relation to possible *finite* measurements with rods or clocks, since the field varies over the volume of the rod and over the space-time volume of the clock process. In intense non-static fields the exact metrical significance of the coordinates is lost for finite regions, and only their topological content remains significant, i.e. the coordinates now do no more than define the *topological ordering*[6] of events as that of $3+1$ dimensional space. (Reichenbach*.)

This lack of exact correlation between the differential *ds* of the geometrical theory and the finite intervals yielded by actual measurements, which only becomes important in hypothetically intense non-static fields or in an exact structural theory, underlies Einstein's remark that the unanalysable relativistic concepts of rods and clocks cannot play an independent part in an ultimate or exact theoretical physics. (See page xi.) Yet in a macroscopic theory of the weak and nearly linear gravitational fields of the known astronomical universe these difficulties do not arise, and hence there can be nearly periodic orbits, and an approximate proportionality between mass and weight. The discovery of these phenomena is from the point of view of the general theory a fortunate chance, since the mass-weight proportionality, for example, disappears or has no meaning in intense non-linear fields.

Though the difficulties which occur in intense fields are of no importance in a macroscopic relativity theory of the known universe, yet they are of considerable heuristic significance. They may suggest the point at which macroscopic relativistic methods have to be modified in order to be appropriate to an exact or microscopic theory. Now the loss of the metrical significance of the coordinates occurs in the general theory as the gravitational fields become very intense, i.e. as velocities in gravitating systems approach c. The disappearance of the empirically significant metric in these hypothetical fields may therefore be regarded as the result of the finitude of c. Thus the finitude of c can be viewed as causing the failure of the descriptive method based on the use of four metrical coordinates. From the standpoint of current theory, if c were not finite these (hypothetical) difficulties in applying a metric in exact theory would not arise. From a different standpoint the same situation may be interpreted as meaning that another and for an exact microscopic theory better method of description exists, and that THE APPEARANCE OF THE VELOCITY c MAY BE REGARDED AS A NECESSARY CONSEQUENCE OF APPLYING A SYSTEM OF FOUR METRICAL COORDINATES TO EVENTS WHOSE ORDERING SHOULD BE DESCRIBED IN ANOTHER MANNER.

(b) *Coordinates in Quantum Theory.*

Heisenberg has shown that the finitude of h implies a definite limit to the accuracy with which two canonically conjugate coordinates of a particle can be simultaneously determined. From this it follows that the path of a given particle cannot be exactly determined and described in terms of the four coordinates x, y, z, t. Thus from the standpoint of current theory the finitude of h is the cause of a second limitation on the exact applicability of four metrical coordinates to the description of physical phenomena. From the point of view of an improved descriptive method THE APPEARANCE OF h AS A CONSTANT OF ACTION MAY BE REGARDED AS THE CONSEQUENCE OF APPLYING FOUR-COORDINATE DESCRIPTION TO FACTS WHICH ARE INADEQUATE. The facts are inadequate in the sense that they fail to supply as much information as is required for exact description in terms of four coordinates. The heuristic importance of this interpretation of the origin of h as a constant of action is not lessened by the fact that the inexactitude may in certain cases be treated as due to the dual pseudo-wave-particle nature of the electrons (protons, etc.) since from the point of view of an improved method of description the wave-properties which are apparent in these cases are themselves partly a consequence of the application of four-coordinate description.

Thus the validity of the description of the motion of a particle in terms of four metrical coordinates is subject to double limitations arising from the existence of the two constants c and h. In interpreting the situation arising from the quantum inexactitude Bohr has emphasised the complementarity or reciprocity of the claims of causality and of the space-time (i.e. four-coordinate) description of events. Though less attention has been paid to the hypothetical difficulties which arise in application of the general theory of relativity to intense fields they are of heuristic importance. The convenience of four-coordinate description in classical theory is due not only to the smallness of h relative to ordinary sense impressions (which allows the inaccuracy of instantaneous position in a given physical frame to be neglected), but also to the smallness of ordinary mechanical velocities relative to c (which permits the establishment of a physical coordinate frame, otherwise unattainable).

Two alternatives were recently proposed: a four-coordinate description with probabilities replacing causation, or an exact mathematical scheme not interpretable in terms of ordinary space-time. These have since proved to be complementary rather than alternative methods. The elementary space-time probability theory of Schrödiger-Born (further developed in the quantum field theory of Heisenberg-Pauli) meets difficulties which are only resolved in the abstract mathematical scheme of Heisenberg-Jordan-Dirac. This latter method is deducible from postulates commonly regarded as having no immediate empirical or physical significance, because they are not capable of four-coordinate representation.

But we have seen that the appearance of the constant h may be regarded as the consequence of using an inappropriate descriptive method, that of fourcoordinates. Since distant simultaneity and four coordinates are not immediately given in elementary physical observations, it is possible that another physically significant method of representing events may be found, in terms of which the abstract axioms of quantum mechanics would acquire immediate physical meaning. In other words, given some new method of representing events it may be possible to deduce the necessity for the abstract algebra and coordinate inexactitude of quantum mechanics. This supports the conclusion reached in §§8, 9, where we saw that a theory of the structure of rods and clocks could in principle dispense with the Lorentz transformations, but that this would imply giving up the interpretation of c as a velocity, and hence also the use of four extended coordinates.

On this view the appearance of both c and h in these forms is to be interpreted as due to, and setting necessary limits to, the application of four-coordinate description to physical observations which are insufficient for this method of representation and should be described otherwise, possibly with fewer independent parameters. Just as the finitude of c involved the relativity of simultaneity, so its variability in the general theory undermines the value in exact theory of the concept of distant relative simultaneity. On the other hand h not only accounts for the existence of finite structures, and hence for the possibility of measurement, but also limits the possible accuracy with which measurements of space and time can be combined into a universal system of four metrical coordinates. Heisenberg* has already called attention to the similar rôles played by c in relation to the principle of relativity and h in relation to the principle of inexactitude. It is possible to develop this similarity from the point of view here put forward. THE APPEARANCE OF c AND h (IN THIS FORM) IS A DIRECT CONSEQUENCE OF APPLYING FOUR-COORDINATE DESCRIPTION TO EVENTS WHOSE ORDERING CAN BE BETTER DESCRIBED IN SOME OTHER MANNER. To this parallel we shall return in §§15 and 23.

In the passage quoted on page xi Heisenberg has suggested that new limitations connected with the constants e, m and M may have to be imposed on the ordinary world of space-time, by which is meant four-coordinate space-time. The task of a unitary theory is to discover these limitations in the form of a unitary or universal type of structure not expressed in terms of four metrical coordinates, which when so expressed gives rise to the necessity for the constants c, h, e, etc. In order to develop methods appropriate to this task, we must now pass from the consideration of metrical coordinates to that of the topological ordering of space-time events.

§11. SPACE-TIME TOPOLOGICAL ORDER

If the demand be made for an invariant metrical description of the observed topological ordering of events (e.g. coincidences) then the empirical quantity c can no longer be interpreted as a velocity.

The use of four metrical coordinates implies, in addition to definite metrical relations between events, that the latter are arranged in the qualitative or topological[(6)] order of the points of a four-dimensional continuum. We shall approach the examination of this assumption through an analysis of the basis of the general theory of relativity.

In introducing the methods of the general theory Einstein justified the adequacy of arbitrary curvilinear coordinates by pointing out that all physical observations can be reduced to the determination of space-time coincidences. Since a coincidence of two events is represented simply by their having the same set of four coordinates, any continuous system of coordinates serves as well as any other. Quantitative physical theory and in particular metrical coordinate systems serve only to simplify the description of the observed totality of space-time coincidences. If it be assumed that physical theory must have the form of a four-coordinate field theory, the absence of unique coordinate systems leads to the principle of general covariance and to the selection of tensor equations as the natural notation for laws which are to be assertions—based on the use of a metric—about space-time coincidences. Einstein has pointed out the various assumptions implicit in this argument. Of these the more important are: that phenomena are reducible to point-events and their relations, that the one definite observable relation between events is the topological property of space-time coincidence, and that rods and clocks exist which enable this description to be given the form of a four-coordinate metrical field theory.

A consideration of these assumptions raises the following questions:

(i) Does the absolute coincidence of point-events provide a convenient basis for physical theory?

(ii) Does the reduction to coincidences imply that the observed *order* need not be retained in the theoretical description of the totality of the coincidences?

(iii) Has the set of observed events, or of hypothetical possible events, the topological order implied in the use of a four-coordinate system?

(iv) Does the use of four-coordinates reveal the essential topological structure of the observable space-time phenomena, or is any alternative method possible and preferable?

These questions will now be considered in turn.

(i) In using the conception of a space-time coincidence Einstein* deliberately neglected the inexactitude involved in the simultaneity and position of two neighbouring but distinguishable events. At that stage the important recognition was that it was not neighbourhood (or coincidence) in space alone or in time alone that could be definitely determined, but the fact of neighbourhood in *space-time*. Since absolute coincidences of different events are not observed, we shall later substitute the term *convergence* where it is necessary to emphasise the fact that what is actually given is the close proximity or neighbourhood of two events. For the present we shall retain the term coincidence.

Physical theory might conceivably be based instead on the ordering of sense-data, or causal world-lines, or electrons. These alternatives are here neglected, as is also the observed colour of events.

(ii) The ordering of observed coincidences is of the utmost importance for physical theory, since a chaotic list of coincidences would have no physical significance. Physical laws have not merely to record numbers of isolated coincidences, they must represent the facts about their ordering which render the coincidences physically significant, e.g. the three dimensional ordering of space, the temporal or causal sequence, the temporal continuity (persistence) of bodies, etc. Thus the task of physical theory is to establish an ordered catalogue of coincidences. In relativity theory the importance of the ordering of the coincidences was not explicitly discussed, but definite types of order were assumed in the use of four coordinates, spherically symmetrical solutions, etc.

(iii) The use of four coordinates implies that the elements (events) represented by the whole range of possible sets of values of the coordinates form " a space of four dimensions " in the strict sense of dimensional theory. The field of possible events, i.e. what is normally understood by space-time, forms a four-dimensional space by definition. It is meaningless to ask whether it can be proved that *actual physical events* form a four-dimensional continuum, because the meaning of " actual events " is not clear [7]. Question (iii) is consequently of no importance. The situation in current theory is that physical laws determine the sequence of actual events against a largely conventional background, which may be called the field of possible events. It is possible to compare the predictions of the laws with the actual events, but not to compare the whole of a finite region of the conventional field with any direct physical observations. The assumption of a four-dimensional field for events has until recently justified itself by allowing the formulation of empirically confirmed coordinate laws.

(iv) The field of possible events defined by four coordinates does not display more of the structure of phenomena than is given in the broad fact that space is three dimensional and time a single series. The responsibility for displaying the rest of the structure of phenomena has been put on the coordinate laws, which have therefore had to describe—as yet without complete success—the facts of discreteness or atomicity, the causal lines of material particles and of light, and the ways in which these interact. The problem of expressing atomicity in physical law has been aggravated by the preliminary assumption of a mathematical continuum as the canvas on which atomicity is to be represented.

Apart from the difficulties in the representation of causal structure which are shown in the pseudo-wave-particle model of quantum mechanics, a profound problem arises in connection with the representation in metrical coordinate theories of the immediate empirical facts (coincidences) which these theories serve to describe. Since the purpose of metrical coordinates is to describe coincidences, it is natural to demand that the metrical invariants of relativistic theory shall be capable of directly representing the observed coincidences. If a four-dimensional metric is fundamental, then the metrical invariants must be capable of representing the observed coincidences. But the coincidence of two events cannot be defined in terms only of the relativistic *interval* between them. If for two points (A, B), $ds_{AB} \longrightarrow O$ or $ds_{AB}=O$, then A, B *may* be neighbouring or coincident events, but they may also be distant events situated on a light-track. Thus the most important topological relation (coincidence) in the theory of physical events, a relation which is fundamental in the theoretical justification of the notation of relativity theory, cannot be directly and unambiguously defined in terms of the metrical invariants alone of the theory. In order to define the relation of close proximity or coincidence either the coordinates or the intervals between at least three events have to be used. THE METRICAL INVARIANTS OF RELATIVISTIC COORDINATE THEORY DO NOT SUFFICE TO DEFINE THE TOPOLOGICAL ORDER OF PHYSICAL OBSERVATIONS.

Another example of the difficulties which arise from the use of an indefinite metric ($ds^2=dx^2+dy^2+dz^2-c^2dt^2$) is that if $ds_{AB}=O$ and $ds_{BC}=O$ then ds_{AC} is not necessarily zero. The property of zero interval between pairs of events is not a transitive relation! This paradox was of little importance so long as physical theory could disregard the topological assumptions involved in the use of coordinates, but if the demand is made that theory shall represent the topology of observed events it requires further attention.

The use of an *invariant velocity* to define an extended time coordinate (Einstein) and the *geometrical representation of this invariant* in a four-dimensional continuum (Minkowski), which is indispensable to the general theory, lead necessarily to the use of an *indefinite metric,* i.e. a metric which cannot express topological invariants directly. Four-coordinate theory cannot give invariant metrical expression to the observed topological invariants. But the metric is introduced according to general relativity theory, precisely in order to describe these topological facts (coincidences). THE INTERPRETATION OF THE INVARIANT c AS A VELOCITY AND THE DEMAND FOR AN INVARIANT METRICAL REPRESENTATION OF THE OBSERVED TOPOLOGICAL FACTS ARE INCOMPATIBLE.

It is possible that the difficulties of four-coordinate theory may be overcome by the use of a new type of metric as the basis of physical theory and that the demand for the direct representation of topological properties may be of value in suggesting this improved metric. (Metric is here and in future used in the sense of any law defining space and time measurements.) Moreover the demand for topological representation may lead to the correct fusion or revision of the methods of relativity and quantum theory since this demand necessarily implies the discarding of the interpretation of c as a *velocity* and the concept of velocity lies at the root of the difficulties of a relativistic quantum theory. We shall therefore use the demand for the representation of the topology of events—as the principle of covariance was used—as a heuristic method to suggest new forms of physical law.

It is interesting to note that the demand for a direct representation of the fact that forces act primarily at small distances is made by Einstein* the basis of an objection to the use of higher-dimensional spaces in quantum theory, since in such spaces the fact of space-time neighbourhood does not receive such a simple expression as in ordinary space-time coordinates. This is a special case of the general demand for the representation of topological facts which is here put forward.

§12. METRIC AND ORDER IN PHYSICAL THEORY

Physics is the study of the ordering of events, and metric describes a special type of regularity in this ordering i.e. the congruence of adjacent event-patterns.

In the general theory physics is treated as the study of space-time coincidences, and we have seen that this implies that the aim of any physical theory must be to give an *ordered* catalogue of coincidences. It will be convenient to express this more generally by saying that physics is the study of the ordering of events. The ordering of events may be taken to mean all the relations between events which are comprised in the terms *between, coincident* (or *convergent*), *before and after.* These words describe non-quantitative relations between events from which the observed order of events can be constructed. They may be called the *topological relations*(6) of physics, by using this term in a wider sense than that of abstract mathematical topology which does not include the conception of temporal succession. An exact discussion of these relations would be of no value at this stage, and it will be sufficient here to assume that the properties of causal continuity, of discreteness, and of the dimensionality of space and time can be reduced to these primary relations between events.

But if physics is the study of the topological ordering of events, what part is played by metric? The relation of metrical and topological properties in pure mathematics is a subject of much interest, but we are here concerned with the description of physical situations and must not accept uncritically analogies from pure mathematics. A field of ordered physical events is given; the problem is to find the status of metric in the description of this field. First of all it is not possible to reduce the order to purely metrical relations, since—as we shall see—order is from our present point of view the more general concept and topological relations may persist where metrical relations cannot be determined. Moreover a metric can only be applied to a field of events whose topological form (e.g. dimensionality) is already given, and we have seen that the invariant metrical relations of relativity theory are inadequate to define the observed facts of ordering, for example coincidences. Further the temporal order (succession) of two events is a descriptive or qualitative relation which can never be given pure metrical expression since it involves a criterion or definition of the non-metrical relation of before and after. Metric is a geometrical conception and does not comprise any asymmetrical relation between two points capable of representing the fact of succession, e.g. the necessary succession in an irreversible physical sequence. These are special instances of the fact that the concept of order is from a certain point of view more general than that of metric, though of course the assumption of a metric implies relations, such as the equality of two separated lengths, which are not included in the relations of order. But the *observable* metrical relations, such as the equality of two adjacent lengths, depend on relations of order; the congruence of two adjacent event-patterns is in fact recognised by observing the *order* of the events. Thus the class of relations between events comprised in the term order is more general than and includes the *directly observable* relations comprised in the term metric. (This refers only to order in the physical sense, the corresponding propositions in pure mathematics being false.)

It is therefore necessary in a deductive theory to regard metric as a mathematical method which is useful for describing a special aspect of the ordering of events; metric describes the uniformities in the ordering of events revealed by the congruence of adjacent event-patterns. This suggests that other mathematical methods may be necessary to describe other kinds of uniformity in the ordering of events which do not take the form of congruences, and to this we shall return later. (§35.)

But the selection of order as the primary subject-matter of physics and the interpretation of metric as a description of one part only of the uniformities or regularities shown in this order, appears at first sight to be a reversal of the fertile tradition in physical thought of the reduction of qualitative to quantitative description. This requires a further analysis because the term qualitative is here ambiguous. The reduction of "quality" to quantity has been necessary because numbers representing the results of measurements appeared to provide the only unambiguous agreed basis on which an objective science could be built. But wherever measurements are beyond dispute, so also is the observed order; only because observed order does not admit of disagreement, do different observers agree on measurements. Order is the ultimate objective fact, quantities are a shorthand for observed congruences in this order, and sooner or later as exact science widens its range it must explicitly recognise this situation. The development of exact science necessarily passes from subjective (here more accurately indefinable or disputable) *quality* to *quantity*, and from *quantity* finally to *order*. Order is a descriptive, non-quantitative property, but it is not qualitative, in the sense of subjective or undefinable. The generalisation of quantity to order is the rational conclusion, not the reversal, of the tendency towards objective precision displayed in the history of physical thought.

An advantage of regarding the topological facts as primary and metric as a method of describing certain symmetries (congruences) in the observed topological order is that it raises the question, when is metric applicable? What kinds of observations justify or require the use of a metric, and what form of metric is necessary? Now physical symmetries (congruences) which are important can only be directly observed when the pattern of events persists either by static permanence (rod) or by periodic repetition (clock). Symmetries in patterns of events can only form the basis of a physically significant metric when they are given as the congruence of finite proper lengths of two rods, or the proper periods of two clocks. The extension of these congruences to a four-dimensional metric resting on an extended time coordinate is here regarded as the origin of the difficulties of current theory. This extension is a legitimate experiment but is justified theoretically only while it provides a basis for the successful development of physical theory. It is therefore possible that a revision of current physical theory might be based on the restriction of metric in exact fundamental theory to the measurement of finite proper lengths and times. This suggestion is supported by the fact that these measurements are the only ones permitted with absolute accuracy in quantum theory.

The proposed view of physical theory may be summarised thus:

Physical Nature presents a discernible ordering of events.

Physical Theory is a conceptual and symbolic system describing and predicting this order.

Metrical Laws describe those uniformities in the observed order which are based on symmetrical congruences. Pure metrical laws are based on congruences only and do not express any asymmetrical two-termed relation between events.

Unitary Theory must be based on a new form of metric capable of providing a description of the topological facts, and not going beyond the given types of congruence.

Non-metrical Laws will also be necessary to describe those uniformities in the observed order which are beyond the range of metric since they involve the asymmetrical relation of succession. (§35.) Such laws will be called descriptive laws.

§13. COURSE CALCULUS

For physics nature is a mesh of lines linking coincidences, here called closed and open courses, and a mathematical method of describing the mesh without an extended time coordinate is called a Course Calculus.

Since the reduction of space-time phenomena to a random collection of isolated coincidences involves the neglect of essential characteristics of the ordering of events, it is necessary to discover the primary element to which phenomena may be reduced without the loss of features which are essential to physical theory.

It is clear that physical theory is concerned with the linkage of coincidences by causal lines. In any physical experiment at least two observations are necessary, an earlier and a later observation connected by the fact that they refer to the same system. Now since all observations reduce to the determination space-time coincidences, the elementary type of physical experiment—on which all others are based—is the determination of a pair of coincidences, the earlier pair being linked to the later by two causal lines. All experiments can therefore be reduced to the observation of a pair of linked coincidences. Thus the primary element to which the ordering of observable space-time phenomena can be reduced is a pair of causally linked coincidences. Since these causal lines are no longer to be regarded as describable in terms of four coordinates, they will be given a new name. A pair of causal tracks leading from a coincidence will be called a *course*. If they meet again they form a *closed course*, if not, an *open course*.

The following heuristic hypothesis is proposed: THE SIGNIFICANT STRUCTURE OF THE SPACE-TIME ORDERING OF EVENTS IS THAT OF THE CAUSAL LINKAGE OF PAIRS OF COINCIDENCES, i.e. CLOSED OR OPEN COURSES.

This implies that four events, arranged as causally linked pairs of coincidences, are the minimum for which a structural relationship can be established, which, in conjunction with the three-dimensional form of space, is adequate to serve as a basis for physical theory. The linkage may be complex, and the absence of linkage in an open course, i.e. the failure of two tracks to converge, is itself a significant, though negative, structural relation. The coincidences are as a rule directly observed, while the linkage is at least in part hypothetical. The postulate implies that the whole of physics (apart from the study of static systems) is concerned with the closing (or not) of a given course, i.e. *whether in a given system whose initial conditions are known in the form of one or more coincidences a specific later coincidence will or will not occur.* Nature is to be regarded as a mesh of closed and open courses. The junctions of the mesh are the facts about which physics can make definite assertions.

From this point of view a metric is a method found convenient in describing the properties of the mesh. For example, the particular kind of metric used in a four-coordinate metrical field theory has provided a powerful method of establishing an ordered catalogue of coincidences by embedding the mesh in a metrical continuum of four dimensions. In doing so coordinate theory made possible many more assertions about the mesh than are actually given in the immediate facts of observation, i.e. more than are necessary to describe the actual linkage of coincidences. In fact metrical coordinate theory implies more information than observation and experiment can yield, as has been proved in the quantum theory of coordinate inexactitude. Our heuristic hypothesis suggests that every physical law must describe a closed course and not, like coordinate laws, begin with arbitrary initial conditions expressed in terms of the assumption of an extended system of coordinates, without implying how these coordinates are to be established, i.e. without giving the initial conditions a precise empirical meaning.

For the various reasons already discussed a four-coordinate theory may not provide the correct basis for a comprehensive and exact physical theory. But if the use of four coordinates is discarded, then physical laws can no longer define actual events (in a system defined by known initial conditions) as a selection from a four-dimensional field of conventionally possible events. The new form of physical law must define the actual ordering of events in some other manner, and a new type of law will be adequate if it can predict the linkage of observable coincidences in open and closed courses. It will be convenient to give a general name to any such method which does not involve the use of four coordinates: *A mathematical method determining the ordering of events in terms of open and closed courses without using four extended coordinates will be called a Course Calculus, and a physical theory based on such a method will be called a Course Theory.*

A course calculus does not permit arbitrary initial conditions since it reduces all phenomena to the primary type of a closed or open course, the initial conditions in either case being given as one or more coincidences. A course calculus is thus a method of predicting if a subsequent coincidence will occur along two causal lines leading from one (or more) earlier coincidences without the use of a relative or absolute simultaneity concept for distant events.

Though no course calculus is at present available the assumption that it provides the correct notation for the basis of a unitary theory will prove of considerable value. A course calculus differs from four-coordinate methods in three ways:

(i) It does not extend the concept of local time to a universally extended coordinate. Coordinate theories use motions to measure local time, and then generalise this local time to a universal independent variable for the expression of laws of motion. Course theory denies the necessity and empirical validity of this generalisation, and asserts that the local time parameter is only required along world lines, and here only to determine that coincidences do or do not occur.

(ii) Since time is no longer required as an independent variable defining a four-dimensional field of possible events, course calculus can make use of the fact that all proper time measurements are dependent on spatial determinations, i.e. all time measurement can be regarded as secondary to spatial measurement. (§9.)

(iii) With the elimination of the conventional field of possible events, the arbitrariness of possible initial conditions also disappears, and in a course calculus all initial conditions are given as coincidences.

Any physical structure providing a resolution of the pseudo-wave-particle duality and a fusion of quantum and relativity theory must necessarily rest on novel and therefore difficult conceptions. The concept of distant simultaneity is the weakest point in the chain of concepts which forms the basis of current theory, because it is the furthest from direct observation. It is possible to retain an elementary or commonsense view of immediate observations—comparisons of rods and clocks, the determination of coincidences, the three-dimensionality of space—without the idea of extended simultaneity, since simultaneity is given in experience only as a local relation. An examination of the situation in general relativity and in quantum theory shows that exact physical meaning has already disappeared from the concept of distant simultaneity except in special cases. Course theory accepts this situation, and seeks to eliminate from fundamental exact theory (e.g. from wave or matrix mechanics) the use of an extended time-coordinate.

The formulation of the aim of a course theory represents a natural extension of observations made by many thinkers on the nature of temporal order, time measurement, and the use of coordinates. (Mach*, Poincaré*, Einstein*, Eddington*, Robb*, Russell*, Bridgman*, Broad*, Reichenbach*.)

§14. UNITARY COURSE THEORY

It is assumed that a course calculus provides the correct notation for the basic principles of a unitary theory.

In §2 the formulation of a unitary theory was proposed as the aim of synthetic theoretical research. The analysis in §§3-12 of the conditions of measurement and of certain difficulties attending the use of four coordinates led to the recognition that there is, in principle, no bar to the development of an alternative mathematical method of representing space-time phenomena. In §13 any alternative method satisfying certain conditions was named a course calculus. The aim of a unitary theory is now combined with the general type of method of a course calculus :

THE HEURISTIC HYPOTHESIS IS MADE THAT A UNITARY THEORY IS POSSIBLE AND MUST BE BASED ON A COURSE CALCULUS. Such a theory will be called a *Unitary Course Theory*, and is the aim of the programme of research which will be developed in Chapters II and III.

This hypothesis will be shown to provide a stimulus to research in a new region of theoretical physics. Without any further knowledge as to the specific form which a course calculus may take it is possible to use this hypothesis to open up a number of problems which have hitherto lain beyond the reach of any general theoretical method. For example, the hypothesis can be used to divide the characteristics of coordinate theories into

(i) those which require four-coordinate description, and hence can be regarded as partly the consequence of the introduction of coordinates, and

(ii) those which can be expressed without the use of four coordinates, and may be interpreted as representing deeper-lying structural features to be retained in a course theory.

The preliminary outlines of this analysis are given in Chapter II. For this purpose it is necessary to summarise the general characteristics which a unitary course theory must have. It is obvious that any such summary must be tentative and will be subject to revision at a later stage. After the analysis of Chapter II has been completed various necessary features of a unitary course theory will be evident which cannot be anticipated at this stage. (§26.) The following characteristics have however already been discussed:

1. A unitary course theory must represent the ordering of events, e.g. coincidences (or convergences) and their causal or other linkage.

2. It must use a modified form of metric providing a basis for a theory of the structure of matter and of radiation, and for a theory of measurement in which only direct comparisons of lengths and of times can be made exactly.

3. It will modify the use of quantitative time in two ways, (a) by eliminating the distant simultaneity concept, (b) by treating (in a theory of the structure of clocks) local time measurement as always derived from, or secondary to, spatial measurements. Taken together these modifications imply the elimination of time as an independent variable. Course theory may thus provisionally be regarded as equivalent to the use of a discrete coordinate system following the lines of world structure and providing a derived time-parameter only for the purpose of determining whether or not two lines will cross.

4. It will define a unitary type of space-time structure without arbitrary initial conditions (possibly a closed course of some definite type) to which all physical processes may be reduced.

5. It must show that the description of this unitary structure in terms of four coordinates implies the appearance of (i) the plurality of types of space-time process called gravitational, protonic, electronic, radiational, and of (ii) the six dimensional constants, by providing derivations of α, β, γ.

Proper lengths and proper times will be retained, but a unitary course theory cannot use in its fundamental formulations any concept dependent on an extended time-coordinate, e.g. the instantaneous distance between particles in relative motion, velocity, momentum, mass, acceleration, energy, field quantities, differential laws, space-time metrical geometry, or wave motion. The calculus must show under what conditions and with what accuracy these secondary and sometimes inappropriate and inexact concepts may be used.

CHAPTER II

ANALYSIS OF COORDINATE THEORIES FROM THE STANDPOINT OF A UNITARY COURSE THEORY

THE following analysis of the characteristics of four-coordinate theories is based on the assumption that a unitary course theory is the correct form for a comprehensive physical theory. This assumption will enable a new interpretation to be given to certain aspects of current theory, which will be supported by a classification of the methods and concepts of coordinate theory into those which can and those which cannot be retained in a course theory. Suggestions will also be made regarding the methods which are available in a course theory to take the place of those which cannot be carried over from coordinate theory.

The analysis starts with the fundamental concepts of coordinate theory. This leads to the consideration of the general form of coordinate laws, the principles of covariance, and the general principles and special problems of quantum theory. The results of this examination are then epitomised in a study of the six universal dimensional constants. A survey follows of the general symmetry properties of coordinate theory, and of the problems left unsolved in coordinate theory. Finally the new features of unitary course theory made clear by this analysis are summarised.

§15. THE BASIC CONCEPTS OF COORDINATE THEORY

Of the quantities of coordinate theory only proper lengths, proper times, and the ratio distance ÷ time over a finite closed course can be retained in a course theory, though this ratio cannot be called a velocity. The relativity corrections and the quantum conditions may be regarded as compensations for the neglect of the closed course property involved in the use of the concept velocity.

The majority of the concepts of four-coordinate physics involve the use of an extended time-coordinate, and hence are not available in the basic formulations of course theory. A list of these forbidden concepts, most of which depend on differential coefficients with respect to an extended time-coordinate, has already been given. (§14.) These concepts must appear in course theory as the result of applying four-coordinate description to a unitary structure defined in some other manner. Thus certain of the primary concepts of classical physics, such as inertia and instantaneous velocity, which have retained their importance in relativity theory and in a severely restricted form in quantum theory, must in course theory appear as complex derived concepts possibly only of approximate validity.

The only quantitative concepts based on space-time measurement which can be carried over into course theory are *proper lengths* and *proper* or *local time periods.* These two quantities can however in certain cases be combined as follows to form an exact equivalent to certain types of velocity: a uniform to-and-fro or cyclic velocity (as in the to-and-fro propagation of light, or in a circular electron orbit) may be retained in a course theory, if reinterpreted as a finite distance ÷ a finite local time taken, i.e. as the distance travelled in a closed course (to-and-fro or cyclic) ÷ the time as measured by a resting clock at the start-and-end point. This ratio cannot be called a *velocity* in course theory, for it is no longer an instantaneous differential coefficient but refers to a finite process and does not depend on an extended time-coordinate involving a distant simultaneity criterion. It has the dimensions $\frac{L}{T}$, but is not physically a velocity, though it is the equivalent of a velocity in the special case of a velocity over a finite closed course. This ratio is a new type of physical quantity which can be defined in terms of special combinations of classical quantities and is the only primary kinematical quantity which can be carried over from four-coordinate into course theory. *If our heuristic hypothesis is correct we shall expect to find that quantities of this kind are more important than can be explained on the basis of coordinate theory.* To this we shall return presently.

The non-metrical concepts of coordinate theory can be retained in course theory, though the particular form given to them may be modified. For example, the process of the *counting* of similar objects remains unchanged. The *topological ordering of events* becomes, as we have seen, of increased importance. Hitherto largely implicit, in course theory the topological assumptions have to be made explicit and used to justify the application of a new form of metric.

The arrangement of events in *finite closed series* (cyclic order) is the topological form definitely given in every positive physical observation. (A conventional element is of course present in the assumption of causal or particle continuity.) This form is therefore to be taken as the significant basis of space-time structure.

If the essential structure of the ordering of events cannot be conveniently described by the representation of events as points in a four-dimensional metrical continuum, then four-coordinate description must somewhere break down, and *the manner of its failure may indicate what other topological form corresponds better to the significant ordering of events.* But we have seen that the constants c and h determine limitations on the validity of exact four-coordinate (metrical) description, and it is therefore possible to regard the appearance of these constants as the result of the assumption—in the use of four coordinates—of a background or physical frame of unsuitable topology, or in other words of the neglect of the essential structure of observable phenomena. Now if the conditions of the measurement of c and h be analysed, the elementary topological form of space-time structure to which these measurements (like every physical experiment) are reducible is that of a closed course. This is seen easily in the to-and-fro measurements of c, and in the determination of h by wave-length observations made possible by the diffraction and interference of coherent rays which form closed courses from the source to the point of interference. In other cases the topological form of a closed course is less obvious only because more complex. The part played by c and h in setting limits to the validity of four-coordinate description suggests therefore that their appearance may be a consequence of the neglect of the closed course structure of phenomena.

This argument may be restated thus. Though all *complete* physical experiments involve the determination of closed courses, yet not all physical quantities or dimensional constants refer directly to complete closed courses. For example the constants e, m, M, G, do not refer to complete experiments, i.e. to closed courses, but imply the use of an already established extended coordinate system and thus refer directly to a part only of a closed course. On the other hand precisely the two constants which set limits to the possible accuracy of four-coordinate description are defined and measured in terms of complete closed courses. c can only be measured either as a to-and-fro velocity or over a more complex closed course on the assumption that the velocity is the same in all directions, and h as well as being originally applied theoretically to define cyclic integrals is actually measured in terms of diffraction patterns of coherent rays, i.e. closed courses. *The limits to coordinate description* *are set by constants which refer directly to complete closed courses.*

An analogy may serve to clarify this situation. Just as in general relativity fields of force are evidence of the assumption of an incorrect, or more accurately, inappropriate *metrical geometry*, so in unitary course theory the appearance of certain of the dimensional constants will be treated as consequences of the use of inappropriate *topological forms* as the background or basis of physical theory. This analogy may be carried further. In general relativity the *proportionality* of mass and weight alone permitted their being ascribed, in the Principle of Equivalence, to a common origin. In course theory the *topological similarity* (homeomorphy) of the space-time structures involved in the determination of c and h alone permits the appearance of these two constants in coordinate theory to be treated as a dual consequence of the neglect of one universal type of space-time topological structure. The numerical proportionality of mass and weight permitted their fusion in the general theory; the dependence on the topological form of a closed course of the measurements of c and h leaves open the possibility of their being traced to a common origin. In fact, when expressed in the forms c, $\frac{e^2}{h}$, they both are kinematic quantities of the only type permissible in course theory: ratios of finite distances over closed courses to finite local times, for light and electrons respectively. Thus the heuristic hypothesis has received an initial justification, and has indicated the first step in the development of a new type of metric. The only type of kinematical quantity which the heuristic hypothesis permits is actually more important in current theory than could be understood on the basis of coordinate theory.

Coordinate theories have so far failed to give any account of the appearance of constants such as c and h. These constants *underlie* the use of coordinates, c being used to define distant simultaneity, and both c and h serving to limit the exactitude of four-coordinate metrical description. The heuristic hypothesis of the possibility of a unitary course theory transforms this situation; THE FACT THAT c AND $\frac{e^2}{h}$ ARE MEASURED OVER CLOSED COURSES MAKES IT POSSIBLE TO REGARD BOTH THE RELATIVITY CORRECTIONS AND THE QUANTUM CONDITIONS AS NECESSARY COMPENSATIONS FOR THE NEGLECT OF THE CLOSED COURSE PROPERTY, THIS NEGLECT BEING INVOLVED IN THE USE OF THE CONCEPTS "DISTANT SIMULTANEITY" AND "INSTANTANEOUS VELOCITY."

Since topological order underlies metrical geometry a mistaken or non-significant topology may reveal its unsuitability by compelling the appearance of more quantities than are necessary in a perfected descriptive method (unitary course theory). Unitary course theory thus seeks not only to render explicit hitherto neglected topological assumptions, but to apply an improved metric based on a topological order representing more closely the significant structure of the data of observation to show that what are called gravitational, electromagnetic, protonic and electronic processes are partial aspects of one universal form of topological and metrical structure.

§16. THE GENERAL FORM OF COORDINATE LAWS

In unitary course theory elementary unitary laws without either arbitrary units or arbitrary initial conditions must replace the general partial laws of coordinate theory.

The general forms of coordinate laws discussed here are primarily those of classical theory, the consideration of the modifications introduced by relativity and quantum theory being treated in detail in subsequent sections.

General Partial Laws.—The fundamental laws of classical and macroscopic relativity physics have the form of equations holding *generally*, i.e. for all systems of a given type whatever their magnitude or structure. Thus the Newtonian law of gravitation asserts a relation between masses, distances, and accelerations which applies to masses of any magnitude. Similarly, the conservation principles, the variational principles of classical mechanics and wave theory, Maxwell's equations, and Einstein's gravitational field equations express general relations supposed to hold between certain quantities whatever their magnitude and whatever the structure of any material system involved. Neither specific structure nor quantities of special magnitude are involved in the definition of the concepts used in the formulation of these laws. The description of particular systems is obtained from the general laws by inserting in the law (or in the general solution of the differential equations) the arbitrary *special values of initial conditions* (masses, charges, fields, distances, velocities, etc.) determining the particular system. Each general law refers only to systems of a given type, and it has been necessary to establish a group of apparently independent general laws describing the different types of systems composed of gravitational or electromagnetic fields or of charged bodies. From the standpoint of a unitary theory these laws may be called *general partial laws*, since though at present independent they are regarded as representing partial aspects of one underlying form of space-time structure. The macroscopic laws of coordinate physics are all of this type.

A unitary course theory providing an exact theory of structure must be based on fundamental laws of a different type. *Firstly*, the fundamental laws must be unitary and not partial, i.e. represent a single universal type of process, which under different conditions appears under the different partial forms (gravitational, electromagnetic, etc.) and *secondly*, the basic laws instead of being general must be structural, i.e. refer to particular and elementary types of structure, from which larger or more complex structures can be built up. Thus a structural theory cannot rest on general laws of the classical type in which the initial values of all the variable quantities are completely arbitrary. A theory of structure must treat as fundamental particular or *elementary* laws referring to certain standard or elementary systems (for example, the hydrogen atom), in place of general laws referring to any system of one of the partial types. But while limited to this elementary system the fundamental law must include within one formulation the gravitational, radiational, and electromagnetic actions between such elementary systems. Thus in place of *General Partial Laws,* unitary course theory must use *Elementary Unitary Laws* without arbitrarily variable initial conditions.

A suggestion of the way in which different partial types of interaction may be derived from one mathematical formulation is given by the derivation of both the chemical and the Van der Waals forces between atoms from one quantum-mechanical notation. (London*.)

To replace the generality of the classical laws elementary unitary laws must be supplemented by a principle of combination enabling a description of complex systems to be built up from the particular laws of the elementary systems.

Elementary though not unitary laws are used in quantum theory, where the basic equations refer specifically to the single electron, the hydrogen atom, etc. The elementary nature of the quantum laws (i.e. their reference to particular elementary systems) is evidenced by the presence of the micro-structural universal constants e, m, M, h, combinations of which refer only to systems of known micro-structure, while the general laws of classical physics contain only the constants G and c which do not refer to micro-structure. The elementary partial laws of quantum theory represent a transition stage between the general partial laws of classical theory and the elementary unitary laws of unitary theory.

Dimensional Homogeneity.—The general laws of macroscopic physics dealt with systems capable of existing in an infinite range of spatial magnitudes (H-sets, §4) and thus not providing any universally available physical unit of length. Hence the dimensional homogeneity of the laws was necessary to compensate the arbitrariness of the selected units of length, time and mass. The particular or elementary laws of quantum theory (and those of unitary theory) describe systems which do provide natural units of length and time (mass is irrelevant to a theory not using four-coordinates) and hence these elementary laws when expressed in arbitrary units must involve the appearance of universal lengths, such as the length $\frac{h}{mc}$ in Dirac's relativistic equation for the electron. The dimensional quantity $\frac{h}{mc}$ has no *intrinsic* importance in the description of physical structures, since it merely expresses the ratio of a natural length to some arbitrary unit of length. The laws of a structural physics cannot be expected to contain a reference to the historical fact that it was convenient in classical physics to be able to use arbitrary units. In other words these laws must receive their most natural expression—in so far as lengths are concerned—as a direct formulation of the ratios of the lengths of different natural structures, without reference to arbitrary units. Laws of this kind imply the elimination of the dimensional constants from fundamental physical theory.

Differential Equations provided the natural formulation of general coordinate laws, the arbitrary integration constants permitting the application of the law to an infinite range of initial conditions, but for the same reason differential equations of the classical type are unsuited for the description of micro-structures of definite size. Quantum theory overcomes this difficulty by substituting for the classical differential equation with its continuous range of solutions a matrix or complex pseudo-wave equation with its eigenwerte or proper values. In certain important cases the solutions of the wave equation only satisfy appropriate boundary and continuity conditions provided the initial conditions are restricted by the energy being limited to a discrete series of special values (discrete eigenwerte). Course theory must discard partial differential equations involving a time-coordinate.

Field Laws were developed as a consequence of the view that all actions are reducible to the contact actions of neighbouring elements and require the use of differentials, which cannot play a part in course theory. Nevertheless the emphasis on topological order in course theory preserves the fundamental importance of space-time neighbourhood.

Conservation Principles.—The principles of the conservation of mass, momentum, energy and electricity (and the corresponding relativistic theorems) involve the use of an extended time-coordinate and cannot be applied in course theory, which must deduce them as secondary principles resulting from the introduction of four coordinates. Conservation principles are very closely related to the measurement of time; every process defined by conservation principles can be used as a clock, and every process selected as a clock permits the definition of a conserved energy function. There is therefore a conventional element in conservation principles which arises from the fact the "independent variable" t is not truly independent but is obtained by using the processes which it is applied to describe. But apart from this conventional aspect conservation principles represent important empirical facts; for example, the root of the principle of the conservation of energy is that for an isolated system an energy-function of the coordinates exists which is independent of the path by which a given state is reached. (Planck*.) This must be accounted for in course theory as a necessary consequence of the introduction of coordinates.

Variational Principles.—In pre-quantum coordinate theory variational principles are a general expression of the existence of differential equations of motion describing uniquely determined continuous reversible paths in a four-coordinate space-time (if the system is isolated, i.e. the field constant) or irreversible paths (if the field is a known function of the time). Variational principles are condensed mathematical expressions of the existence of laws of motion of this kind. The different formulations of the principles express this property for systems of varying degrees of generality, and the usual forms contain conventional elements chosen so as to render them convenient for treating mechanical problems. (Voss*.)

In view of the fact that the existence of uniquely determined motions in four-coordinate space-time is denied in quantum theory, variational principles no longer have the same physical significance. Their confirmation in classical and relativity theory is evidence only of

(i) an approximate determination of continuous paths, and
(ii) particular laws of force or of potential distribution.

Though variational principles cannot be used in course theory, the fact that they refer to *finite integrals* may render them useful in deducing coordinate laws from the fundamental principles of course theory.

§17. SPECIAL AND GENERAL COVARIANCE

The principles of covariance are relevant only to a non-structural four-coordinate field theory; they must be deduced in course theory as necessary consequences of the (approximate) introduction of four coordinates.

Since the Principles of Special and General Covariance can play no part in the basis of a theory which does not use four coordinates, course theory must show that it leads to results which, when expressed in terms of four coordinates, are equivalent to those which follow from the principles of covariance. We shall therefore examine briefly how the methods of relativity theory may be regarded from the standpoint of a course theory.

(a) The special theory is based on the interpretation of the Michelson-Morley experiment as establishing the invariance of a physical *velocity* c. This view takes no account of the fact that c must be measured either in to-and-fro observations or on the assumption that c is the same in all directions, this fact being supposed not to have any special physical significance. The emphasis is on metrical representation, the topology of the situation being neglected. Course theory proposes another interpretation of the measurements based on to-and-fro experiments. The empirical invariant c is not to be regarded as a velocityinvolving a time-coordinate, but as the ratio of a finite distance over a closed course to the to-and-fro time measured at the start-and-end point. The emphasis is here on topological representation, the metric being chosen to conform to topological facts. These two alternative interpretations of the empirical invariant c lead respectively to the theory of relativity (which is occupied in eliminating reference to any particular one of the coordinate systems implied in the interpretation of c as a velocity) and to course theory (which wholly discards four coordinate systems).

It is to be noted that *all* determinations of c, not only those involving direct to-and-fro propagation, involve the observation of closed courses. This is merely a special case of the fact that all complete physical experiments take this form.

(b) In relativity theory the existence of a symmetrical algebraic form for the interval if the units of length and time are chosen such that $c=1$ implies a correlation between space and time measurements. In course theory this correlation is retained, time measurement being defined in terms of space measurement.

(c) Though the Lorentz transformations have no place in course theory they must appear as necessary accompaniments of the deduction of coordinate field physics from course theory.

(d) The failure of the metrical invariants of relativity theory to represent topological invariants has been discussed in §11.

(e) The principles of covariance assert that the laws of nature are to be expressed without reference to any special coordinate system, and their heuristic value lay in the search for *simple* covariant laws. The theory of relativity does not deny the existence of unique local coordinate axes (e.g. non-rotating axes) but only asserts that the differential equations expressing field laws have covariant form. Covariance is a property of differential equations and is irrelevant to the finite quantities involved in a structural theory, which correspond to special integrals of the equations.

(f) The symmetry of certain aspects of space and time measurements in relativity theory may be regarded as due to the absence of any necessity to distinguish in the mathematical notation between real and imaginary quantities or between definite and indefinite metrics. In quantum mechanics the distinction between real and imaginary quantities is important. These facts combine with other difficulties to prevent a satisfactory formulation of relativistic quantum mechanics.

(g) The equivalence of mass and energy in relativity theory is of special importance. Course theory must show this to be a necessary consequence of the application of coordinate systems to the phenomena described by its own principles.

(h) Gravitational actions can be regarded by unitary course theory as a limiting degenerate form of a general type of action primarily represented in current theory by the more powerful quantum and electromagnetic actions. Gravity and quantum-electromagnetism need not be given equal status in a unitary theory; gravity being a much weaker and mathematically simpler type of action may be treated as a special residual effect. Moreover gravitational theory is only macroscopic and approximate. Quantum phenomena are therefore of more importance than gravitation for the transformation from coordinate to course theory.

(i) In the absence of an exact relativistic quantum theory the interactions of radiation and matter are not known adequately to justify reliance on Einstein's cosmological theory (1917) or in the new theories which have been based on it. (Lemaitre*, Tolman*, Eddington*.) These theories involve a degree of extrapolation from direct observations which, in the absence of a unified physical theory, may easily be at fault.

§18. QUANTUM THEORY

The general methods of quantum theory have been given a final form, though the fundamental mathematical axioms have not yet received physical interpretation.

The general methods of quantum theory are less known than those of classical and relativity physics and it will be convenient to give a preliminary analysis of the structure of the theory before going on to consider its relation to course theory. (§§19-22.) We are here concerned with quantum theory in the standard form which was reached in 1927 [8]. In this definitive theory the pioneer work of Planck, Bohr, Einstein, Sommerfeld, Heisenberg, de Broglie, Schrödinger, Pauli and many others was crowned on the physical side by Born's Probability Interpretation (1926) and Heisenberg's Principle of Inexactitude (1927), and on the mathematical side by the Dirac-Jordan Transformation Theory (1927). The only important addition made since 1927 is Dirac's relativistic theory of the electron (1928). The quantum theory of 1927, supplemented by this new feature, is able to account correctly for a very wide range of atomic phenomena, and consolidates within one general theory all the ground won since 1900, though various important problems are still left in an unsatisfactory state. There is little doubt that the general methods of the transformation theory represent a final stage of general quantum theory, i.e. that so long as four coordinates are used these methods provide the indispensable basis for any general theory. (If the use of four coordinates is discarded the name quantum theory must be given up.) The history of the theory is briefly sketched in Appendix B ; here we shall examine the structure of the theory in its final form. In subsequent sections different aspects of the theory will be taken up in greater detail and discussed from the standpoint of a course theory.

Various methods of presenting the 1927 quantum theory are possible. It may, for example, be approached through the complementary nature of the only physical concepts now available for describing experiments : waves and particles. The mathematical theory is then reached by a process of induction from the experimental facts which imply the breakdown of these concepts when they are applied separately. This method is followed by Heisenberg*. The other extreme is the strict axiomatic method, which seeks to deduce the observed phenomena from a series of highly abstract postulates. (Jordan*, Neumann*, Born*.) An alternative to these is the compromise which accepts the deductive method but attempts to keep closer to the physical significance of the mathematics by avoiding the use of a set of independent axioms. (Dirac*.) All these methods have their advantages and disadvantages, and they should be regarded as supplementary to one another. Yet another approach lays the emphasis on the empirical importance of the wave concept in elementary problems, and on the analogy of the new mathematical notation with that of classical wave theory. This aspect of quantum theory is of great convenience and practical importance, but easily leads to misinterpretation through neglect of the fact that classical waves as well as classical particles are essentially inadequate as representations of quantum phenomena.

In the following analysis the logical structure of the theory is our main concern, and the deductive method will therefore be most convenient. It also has the advantage of laying emphasis on the mathematical expressions which correctly represent experience rather than on the classical physical concepts which are now known to be inadequate. It must be remembered however that the object of this analysis is to help towards the formulation of a new physical concept capable of providing immediate physical meaning to the abstract basis of quantum theory.

The Structure of the 1927 *Quantum Theory.*—The mathematical structure of the theory rests on the use of a *system of symbolic algebra* involving a *non-commutative rule of multiplication.* Certain primary symbols may be taken to represent a class of higher numbers (q-numbers), or matrices, or operators in a Hilbert (higher-dimensional complex) space. This symbolic system is supplemented by *interpretative rules* which yield to the symbolism the power to predict empirical results. Thus a definite combination of the symbols is defined to represent the average value of a given observable quantity in the state of a given physical system represented by another symbol. This interpretation, together with certain axioms about measurement, permits the symbolism to define, in general, the possible results of any experiment and the *probability* of any of these results when a previous experiment on the same system has given a definite known result. The function which transforms the symbolic expression of a known observation into the probability for some other subsequent empirical observation, is called the *transformation-function,* or sometimes the probability-amplitude because it must be squared to yield the observed probabilities. A special case of the transformation-function is the *Schrödinger wave-function*, normally in a higher-dimensional space. The possible values of an observable quantity can be represented by the *Eigenwerte* of a *Heisenberg Matrix*, and in the case of the matrix representing the energy these are also the eigenwerte of the Schrödinger *wave-equation*. The standard expression of the transformation theory is not in relativistic form, the normal formulation of a probability in any actual problem not being a relativistic invariant. (Neumann*.)

This completes the general mathematical notation of the quantum theory, i.e. includes all the general physical laws which do not apply only to special systems. To obtain definite results from this notation the symbols must be given special properties which render them appropriate to special physical systems. These special properties are given by the *equations of motion* (Hamiltonian) and the *quantum conditions* (commutation rules) of the special quantum-mechanical system which it is desired to describe. The Hamiltonian and the commutation rules constitute the specification of the physical system, which when inserted into the general notation yields statistical predictions about all possible measurements which in some cases reduce to exact causal predictions. The specification of a particular system (electron, proton, photon, or systems of many of these, with or without interactions) has to be taken direct from experience, though various methods are available which suggest appropriate functions. Thus the quantum Hamiltonian may according to the *Correspondence Principle* often be taken direct from classical theory, though this sometimes leaves an ambiguity resulting from the new importance of the order of the non-commuting symbols. The quantum conditions constitute a generalisation of classical theory, the non-commutative multiplication permitting the introduction of the characteristic *quantum constant h*, and *h* need only be introduced at this one point. The appropriate quantum conditions (commutation rules) for systems of electrons lead to the Pauli *Exclusion principle* and the Fermi-Dirac statistics, and for photons to the Bose-Einstein statistics.

This theory when applied to quantum-mechanical matter and quantum-mechanical radiation shows that within definite limitations matter and radiation can both behave in elementary situations as classical *waves*. Thus certain elementary quantum-mechanical systems necessarily display the properties of particle motion and of wave diffraction and interference, within definite limitations. These limitations define an *inexactitude* in the possibility of the four-coordinate description of space-time processes, which underlies the use of the concepts of waves and particles. This inexactitude on the one hand can be regarded as a necessary consequence of a *wave-particle dualistic model*, and on the other hand corresponds to the fact that the postulated interpretation of the abstract algebra of quantum mechanics in general only yields relative probabilities where classical theory yielded exactly determined values. Finally a *relativistic Hamiltonian for the electron*, though inappropriate to the general notation of the theory and presenting various difficulties, predicts correctly the fine-structure of the hydrogen spectrum, and suggests that in addition to its wave-particle properties the electron will behave, to a definite approximation, as though it had a *spin* of the empirically found magnitude, and will therefore require *four quantum numbers* for its specification.

The theory so completed provides an adequate basis for a comprehensive description of the ordinary atomic, physical, and chemical properties of matter, including the statistical theory of the general electrical and thermal properties of gases, liquids and simple crystalline substances, the structure of the periodic table, chemical valency, etc. Difficulties arise in an exact relativistic (electrodynamic) theory, e.g. of nuclear structure and radioactivity, and in the calculation of the properties of complex atoms and molecules. No attempt is made to correlate the different Hamiltonians, or to provide a space-time interpretation of the algebraic axioms of the theory. Leaving aside these and other unsolved problems which appear to lie in a new region of synthetic theory, the 1927 quantum theory has already proved that it contains most if not all the general principles necessary to account for atomic phenomena.

Before passing to the discussion of the different aspects of this theory in relation to course theory it will be convenient to summarise some of the most important of the axioms which are necessary in a strict deductive formulation of the theory. These are:

1. *Principle of Interference of Measurements.*

   Every measurement alters the measured object, and two measurements consequently interfere unless they can be replaced by one measurement. But the alteration caused by a measurement is such that this measurement remains valid, i.e. the same result is got by an immediate repetition of the measurement.

2. *Representation of Physical Quantities by Operators (or Matrices).*

   $\phi_r\, a\, \psi_r$ represents the average value of the complex operator a representing a particular observable for the state r of a system represented by the conjugate imaginary symbols ϕ_r, ψ_r. It follows that the eigenwerte a of the equation $a\psi_r = a\psi_r$ represent the possible values of the observable a, and that $|\phi_s\psi_r|^2$ represents the probability of a given observation on ϕ_r agreeing with the same observation on ψ_r. (Dirac's notation.)

3. *Definition of Operators describing special systems.*

   This is given by

   (i) The classical Hamiltonian, in certain cases suitably modified (Correspondence Principle).

(ii) The quantum conditions, which lead in certain cases to the substitution of $-\frac{ih}{2\pi}.\frac{d}{dq_r}$ for p_r, and of $\frac{i.h}{2\pi}.\frac{d}{dt}$ for H. $(i=\sqrt{-1})$.

Of these axioms the first two lead to the general statistical theory of quantum mechanics, while the third defines its application to special problems.

§19. GENERAL PRINCIPLES OF QUANTUM THEORY

The quantum conditions are restrictions which have to be imposed on the dynamical variables in order that they may have the special values essential to any structural theory.

We now proceed to a more detailed discussion of the relation to course theory of the following features of the general methods of quantum theory:

(i) The abstract basis of the mathematical notation.
(ii) The interpretative rules (Probabilities, Inexactitude)
(iii) The correspondence principle.
(iv) The quantum conditions.

(i) *The Abstract Algebra.*

The mathematical notation of quantum mechanics, though yielding definite predictions about measurements, is based on the use of symbols which do not represent ordinary numbers, or ordinary physical quantities. The symbols may for example represent complex numbers, and in certain cases obey a non-commutative law of multiplication. This notation has been reached as the culmination of a process which began in 1900: the development of an algebraic (symbolic) description of phenomena which was left uninterpreted whenever current physical concepts proved inadequate. The first uninterpreted algebraic rule was that introduced by Planck, which he successfully incorporated in his theory of radiation though he was unable to give it any general theoretical basis. Algebraic rules of this kind have continually been discovered, combined, and generalised till in the quantum theory of 1927 they dominate the whole of atomic theory. As the rules have been gradually generalised, they have conflicted with classical theory at successively more and more fundamental points, and the attempts at physical re-interpretation have never made up on the rapid growth of this comprehensive abstract notation. Planck's theory of radiators with discrete energy changes, Einstein's light-quanta, Bohr's stationary orbits, Schrödinger's waves, Born's probability interpretation, and Heisenberg's wave-particle inexactitude, have each expressed new physical ideas modifying the conceptual structure of classical theory at successively deeper levels. Heisenberg has pointed out that his inexactitude principle was evolved subsequently to and in order to explain the statistical interpretation of quantum variables. (Compare Appendix B.) Yet none of these new physical suggestions has proved adequate to provide a comprehensive and unified physical interpretation of the algebraic system which has grown up in analogy to but considerably generalising the mathematics of classical physics. The last stage of these physical interpretations, the wave-particle-probability-inexactitude, on the one hand has shown the inadequacy of any physical model now known, and on the other hand is unable to provide a physical justification for the special form of the axioms of the quantum-mechanical notation.

This can only be because none of these suggestions has yet modified classical theory at its root. The assumption of the necessity of a four-coordinate description of phenomena has been left intact, even in the statistical theory. COURSE THEORY THEREFORE SEEKS TO PROVIDE NEW PHYSICAL CONCEPTIONS, NOT BASED ON FOUR COORDINATES, WHICH

DESCRIBE THE GIVEN STRUCTURE OF SPACE-TIME OBSERVATIONS, AND BY DOING SO YIELD A PHYSICAL INTERPRETATION TO THE ABSTRACT ALGEBRA OF QUANTUM MECHANICS.

A general reason for the appropriateness of the notation used in quantum theory is that it leads to the formulation of any physical problem as an eigenwert problem. These latter are the only known class of mathematical problems which can yield as solutions either a discrete set of special values, or a continuous range of values, or a combination of these. Quantum mechanics, since it has to describe (e.g. in spectra) combinations of special discrete values with continuous ranges, necessarily came to be formulated in terms of an eigenwert problem.

(ii) *The Interpretative Rules.*

The interpretation of dynamical variables as representing probabilities leads to a new type of quantum superposition (London*, Dirac*) differing essentially from classical superposition. A classical assertion of the presence of one system *as well as* another system, is to be translated in quantum theory into a statistical assertion of the presence *either* of the one system *or* of the other. This is a consequence of accepting the finite lower limit to the accuracy with which the position and momentum of a particle can be simultaneously determined. But in spite of the fact that in general only probabilities are yielded by the theory, yet there are always some quantities for which exact values can be predicted subsequently to an exact measurement of initial conditions. Thus a strict causal determination governs certain quantities ; so far as is at present known it is only the demand for a complete four-coordinate description that brings about an inevitable inexactitude.

But if, as course theory suggests, four coordinates imply more metrical relations than are given in experience, then the expression in terms of four coordinates of a law, which in some non-coordinate notation is definite and exact, must necessarily lead to inexactitude. From the point of view of course theory the inexactitude in coordinate description arises because it is impossible to combine (for example) two measurements of position and momentum to form a complete coordinate description in such a way as to allow for the mutual interference of the two measurements. The aim of a course calculus is to eliminate the necessity of combining simultaneous measurements by adopting a non-coordinate description of the causal structure of physical processes. The probability law of the distribution of measurements about a mean will in course theory be regarded as the expression of the arbitrariness involved in selecting one out of many possible ways in which coordinates may be introduced into the description of a closed course.

Heisenberg* has suggested that the kernel of the quantum theory is expressed in the necessity for a different description of a whole process when an observation is made half-way through the process. Now an observation corresponds to the closing of a course, i.e. the noting of a coincidence. This feature of quantum statistical descriptions may therefore be expressed by saying that theory must not neglect the individual smaller closed courses that make up the larger closed course of a complex experiment.

(iii) *Correspondence Principle.*

The final form taken by Bohr's correspondence principle asserts that the classical equations of motion (Hamiltonians) may be taken over into quantum theory, such special modifications being made as are necessary. These Hamiltonians specify the particular type of system considered, in terms of the dimensional constants and law of force involved in the system. But the correspondence principle even in this form is merely a heuristic guide to the general kind of function required since if regarded as a precise principle it breaks down in various ways: (i) there is no classical analogy to certain quantum-mechanical systems (see §20),

(ii) the classical Hamiltonian must in certain cases be expressed in special coordinate systems if it is to provide the correct form for the corresponding quantum system, and (iii) the arbitrariness of the order of variables in classical Hamiltonians gives rise to ambiguities. Moreover the formal analogy underlying the correspondence principle conceals the profound physical difference implied by the appearance of h and i in the quantum conditions.

The aim of a unitary course theory is to provide a unified derivation of the required quantum Hamiltonians to take the place of this heuristic and provisional analogy with classical systems.

(iv) *The Quantum Conditions.*

The quantum conditions, expressed in the form of commutation rules, represent the essential feature of the notation of quantum theory, as opposed to classical and relativity theory. Planck's constant h need be introduced into the theory only through these rules, and all phenomena in which h are included may be regarded as resulting from them. Thus the finite limit to inexactitude and the determination of special lengths are two of the most important consequences of the commutation rules. This fact may be expressed in another way, (i) compatible observations, i.e. non-interfering simultaneous measurements, can only be made of quantities which commute, and (ii) the definition in terms of h of discrete special values for observables is a consequence of the eigenwert-problem which necessarily results from the fact that the non-commutating variables have to be interpreted as operators or as matrices. Course theory accepts the necessity of defining special values, but seeks to eliminate the inexactitude by eliminating the use of four coordinates.

We now pass from these formal properties of the commutation rules to their dynamical meaning as quantum conditions imposed on a quantum-mechanical system.

A classical dynamical system is described by equations of motion which determine the subsequent values of all the variables, when their initial values are given. In quantum mechanics this description is inadequate, a quantum-mechanical system being adequately described only when the quantum conditions are known in addition. These are extra relations connecting the variables at any one time, and may thus be regarded as *restrictions put on the initial values of the variables*, which are such that they persist throughout the subsequent history of the system. The quantum conditions can be viewed as initial restrictions, which limit the arbitrariness of the initial conditions, but taken alone do not determine them. These new dynamical relations called the quantum conditions are introduced into the mathematical scheme in the form of commutation rules involving h. The commutation rules therefore provide restrictions on the possible values of the dynamical variables which hold good throughout the history of the system.

But we have seen (§3) that any structural theory must place restrictions on the range of values open to classical variables allowing the determination of special values, and that the constant h performs just this function. (§6.) The conclusion is therefore reached that whatever particular part is played by the commutation rules in relation to the special mathematical calculus of transformation theory, their primary or fundamental function is to restrict the dynamical variables so as to allow the determination of the special values necessary to any structural theory.

Side by side with this general interpretation of the commutation rules it is necessary to put their other function in quantum theory, which is to define which measurements interfere and cannot be made simultaneously with exactitude, and which do not interfere. As has already been mentioned it follows from the abstract algebra and the interpretative rules of quantum mechanics that compatible observables (whose measurements do not interfere) commute, and conversely that simultaneous exact measurements cannot be made of observables which do not commute.

As the result of examining a further type of quantum condition applying to many body systems we shall be led in the next section to generalise as follows the interpretation already given:

THE COMMUTATION RULES OR QUANTUM CONDITIONS SERVE TO RESTRICT IN THE MANNER ESSENTIAL TO ANY STRUCTURAL THEORY (i) THE POSSIBLE VALUES OF THE DYNAMICAL VARIABLES OF A SIMPLE SYSTEM, AND (ii) THE WAYS IN WHICH SIMPLE SYSTEMS MAY BE COMBINED TO FORM COMPLEX SYSTEMS. These restrictions are precisely of the kind to be expected in the transition from non-structural field theory to a structural theory requiring variables with special values. Important examples of the effect of these restrictions are (a) the determination in terms of h of the special lengths involved in the structure of matter and radiation, and (b) the organisation of the elements into a systematic periodic table on the basis of Pauli's exclusion principle.

It will be convenient to give a special name to dynamical elements (e.g. particles) whose initial conditions are arbitrary, and to those for which the initial conditions have to be subjected to special restrictions. Dynamical elements (particles) will be called *analytically independent* in the former and *analytically dependent* in the latter case. For example in classical theory particles were analytically independent, whereas in the quantum theory of atomic structure or of assemblies of electrons or protons they are analytically dependent. In quantum theory the persistence of these restrictions throughout the history of a system implies that in addition to the conservation of energy there is a conservation or persistence of form in the function representing the complex system. Though the particles can be arbitrarily interchanged this aspect of the form of the wave-function is definite, and is preserved throughout the history of the system it represents.

The " analytical dependence of quantum particles " may be taken as expressing an important consequence of the existence of the quantum conditions, which is that the classical conception of the number of degrees of freedom of a dynamical system is not applicable to quantum-mechanical systems, since—for example—the system of electrons in an atom can in certain cases be described equally well by one parameter (the energy) as by the complete set of coordinates of each electron. (Neumann* ; cf. §33.) The existence of restrictions on the initial conditions of a system necessarily means that the number of degrees of freedom of a system no longer has a definite meaning.

Finally, another consequence of the form of the quantum conditions and the fact that they contain i is that a quantum variable is representable as a set of Fourier terms periodic in the time. (Dirac*.) To this we shall return in §21.

§20. SPECIAL QUANTUM-MECHANICAL PROBLEMS

The difficulties arising from special quantum problems lie in the new region of research implied in the aim of a unitary course theory.

(i) Non-relativistic systems.
(ii) Resonance.
(iii) Systems of similar particles.
(iv) Relativistic Quantum Theory.
(v) Dirac's Relativistic Electron.
(vi) Electron and Proton.
(vii) Matter and Radiation.
(viii) Difficulties.

(i) *Non-relativistic systems.*

In certain cases the correspondence principle yields no direct assistance in finding even a non-relativistic approximation to the correct Hamiltonian for a quantum-mechanical system. For example, there exists no classical system corresponding to the situation presented by a photon interacting with an atom. In other cases ambiguity arises owing to the importance of the order of two multiplied quantities in quantum theory. Moreover the selection of the different commutation rules (quantum conditions) for many-body systems of electrons, protons, and photons, has not yet been justified by any general theory. Thus many arbitrary features remain in non-relativistic quantum theory.

(ii) *Resonance.*

The important phenomenon of quantum-mechanical resonance between two systems is of special interest in relation to course theory. For example, the quantum-mechanical description of two identical atoms near one another indicates what would classically be regarded as a non-stationary oscillation of energy from one to the other. This may however be treated as due to the superposition of two primary stationary oscillations of slightly differing frequency. In quantum mechanics such " superposition " means that there is a certain probability of finding one or the other of the two primary stationary states. (For certain atoms the Pauli exclusion principle prohibits one of these states.) The splitting of an initial stationary state into two distinct states owing to the perturbation caused by the presence of a second atom, is thus the consequence of the reduction in quantum theory of *a single non-stationary state to a group of alternative states* each associated with a definite probability. This reduction of non-stationary states to stationary states in which the relative phases are arbitrary is the equivalent of the elimination of progressive motion from problems whose classical representation involves the movement of particles or the propagation of waves. Another example is the stationary wave description of the Bohr atomic states whose classical description involves the progressive motion of a particle round an orbit. In these examples the elimination of non-stationary, i.e. *moving* systems by wave-theory may be regarded as representing a step towards the complete elimination of the use of a time-coordinate in descriptions of quantum systems.

(iii) *Systems of similar particles.*

Just as an elementary quantum-mechanical system is not adequately specified without the quantum conditions, so a complex system is not completely defined by the quantum descriptions of the simple systems which compose it but requires in addition a rule determining how the description of the complex system is to be built up out of the descriptions of the simple systems. In other words a rule is required restricting the initial arrangements, and hence also the subsequent arrangements, of the elements composing the complex systems in terms of their coordinates and momenta. This restriction, as in the case of simple systems, can be expressed in the form of a commutation rule. In this case the rule determines the commutation properties of the wave functions obtained by a permutation of the particles (electrons or photons), and the two special rules which have to be assumed for electrons and photons lead to the use of wave functions respectively antisymmetrical and symmetrical in the coordinates of the particles. Since the commutation rules are such that if they hold initially they persist throughout the history of the system, the wave function if once symmetrical (or antisymmetrical) remains so. From the functions so selected the Pauli exclusion principle and the Fermi-Dirac statistics for electrons, and the Bose-Einstein statistics for photons, follow immediately.

In this quantum description of systems of similar particles the individual particles lose their identity and analytical independence. The emphasis is put on the persistence of a certain type of order or form in the complex system, though the elements which define the arrangement have no individual identity. Electron tracks, for instance, correspond not to the paths of large bodies but rather to the contours on a hill, whose form is defined by a law of arrangement not expressible in terms of separate individual particles with arbitrary initial conditions.

(iv) *Relativistic Quantum Theory.*

No systematic relativistic treatment of quantum systems has yet been developed and it is uncertain whether the abstract algebra of quantum mechanics is in its present form appropriate for relativistic problems. A direct attack on the problem leads to difficulties, and the attempt made to establish a relativistic quantum field theory[1] has not produced encouraging results.

From the standpoint of course theory both the relativity corrections and the quantum conditions represent corrections which are necessary under certain conditions to compensate for the use of four-coordinate description, and the neglect—implicit in the use of the concept velocity—of the closed course structure of phenomena.

It has already been suggested that the systems for which both the quantum conditions and the relativity corrections would be simultaneously necessary may not be capable of satisfactory description in terms of four coordinates. This view is supported by the fact that in a relativistic quantum theory an exact definition should be given of the way in which time-coordinates are obtained, i.e. a structural theory of clocks should be given. But relativity theory rests on the idea of a clock as an unanalysable or primary concept, and a relativistic quantum theory is therefore likely to give rise to contradictions. The alternative is a course theory based on a study of the structure of clocks.

The fact that a relativistic theory must lead to difficulties can also be seen as follows. In classical mechanics t was an independent variable and not an observable quantity. In relativity mechanics t, in addition to being an independent variable, had to be included in the observable parameters necessary to describe the state of a system. This caused no difficulty because relativity mechanics restricted itself primarily to simple systems. But in a relativistic quantum mechanics the double use of t as observable and as independent variable gives rise to difficulties because some of the most important quantum problems are many body problems. Only in very simple cases can time be successfully used both as independent variable and as an observable; in complex problems it necessarily fails because it implies the superposition of methods based on the assumption of unanalysable clocks on methods designed to give an analysis of clocks.

(v) *Dirac's Relativistic Electron.*

Though a general treatment of relativistic quantum theory for complex systems has not yet been given, Dirac has developed a first order relativistic wave equation for the free electron. This leads to a term in the angular momentum integrals which corresponds to the mechanical moment of a spinning electron, and to terms in the energy which correspond *to a definite approximation* with the classical energy due to the equivalent magnetic moment. Thus within certain restrictions the spinning electron model is justified by this equation. The restrictions on the physical validity of the model imply that the spin moment can never be directly measured, and that in certain cases no spin term appears in the expression for the energy owing to the direction of spin being undefined. It is important to notice that here as in the relativity theory of electronic and gravitational orbits the relativistic correction takes the form of a superimposed rotation, in this case in the form of an electron spin.

Dirac's equation represents a step beyond the normal forms of wave equation suggested by the correspondence principle. The equation raises difficulties which are intimately connected with the asymmetry of electron and proton, and though invariant in content is not covariant in form. Moreover the equation is incompatible with the general methods of quantum theory since if combined with them it implies that the electron always has the velocity c. (Breit*, Heisenberg*.)

Course theory regards these problems like all those of exact relativistic quantum theory as insoluble in terms of four-coordinate theory, since relativity theory treats clocks as unanalysable ultimate conceptions, while an exact quantum theory must give complete structural descriptions.

(vi) *Electron and Proton.*

The electron and proton differ in mass, and also in the part they play in nuclear and atomic structure and in radioactivity. Quantum theory has to assume the appropriate Hamiltonians and quantum conditions for electrons and protons, and offers no general interpretation of these. A relativistic quantum theory, if it were possible, would necessarily throw light on this asymmetry. Unitary course theory seeks to represent the electron and proton as two aspects of the coordinate description of a unitary non-coordinate process.

(vii) *Matter and Radiation.*

Matter and radiation are essentially different though in quantum theory capable of description by similar laws. The differences of matter and radiation are expressed in quantum theory by the absence of proper-mass and charge in the photon, by the polarisation properties of the photon which are different from those of spinning material particles, by the different commutation rules obeyed by complex systems of photons and systems of electrons or protons, and by the lack of conservation in the number of photons in a system. In spite of these differences, which quantum theory cannot explain but has to assume, both matter and radiation are alike in that they display some of the features of particles and of waves. This similarity is to be regarded in course theory as a consequence of applying coordinate description to a unitary course phenomenon. (See §21.)

(viii) *Difficulties.*

The main difficulties or arbitrary features in quantum theory are

(a) The inadequacy of the correspondence principle in defining Hamiltonians.
(b) The problems of relativistic quantum theory.
(c) The lack of a general theory justifying the quantum conditions (commutation rules) for complex systems.
(d) Nuclear theory (connected with (b)).
(e) Inadequate mathematical technique for complex systems.

All these problems, excluding (e), lead into a new region of physical theory. This new field may be defined as the justification of the use of the different dimensional constants, Hamiltonians, and commutation rules, within one comprehensive unitary theory.

§21. PARTICLES AND WAVES

The quantum pseudo-particles and pseudo-waves are partial descriptions of a unitary physical structure which underlies the axioms of the mathematical theory.

Until recently physical theories were based either on the conception of "particles," or on that of "waves." In 1905 Einstein used particle theory in order to modify the wave theory of radiation, and in 1925 de Broglie applied wave theory in order to improve the particle theory of electrons. These successful modifications of the earlier independent particle and wave theories showed that each concept represented phenomena correctly only when appropriately modified by the application of the other. Both matter and radiation display some of the features of both models, though the symmetry is limited by the different use of the dual model for representing matter and radiation. This situation found formal representation in the transformation theory, which, combined with the inexactitude principle, renders precise the limitations which each model places on the other.

But this theory also shows that in the exact description of complex systems theparticle and wave concepts, while still indispensable in the preliminary identification of a physical situation, do not represent the essential features of a quantum-mechanical problem. The analogy which quantum phenomena display to the expressions of particle and wave theory and which is used in the identification of a problem is fundamentally misleading owing to the use in quantum theory of imaginary and complex quantities in higher dimensional spaces. The classical wave and particle analogies cannot be adequate, for whereas the classical wave and particle have incompatible properties (in spite of the partial Hamiltonian analogy) the quantum pseudo-wave and pseudo-particle are fused in one unified mathematical theory.

The union of the two concepts implies that neither is valid in its classical form, and before going on to consider what new mathematical form or physical model is to take their place, it will be convenient to summarise the properties of (i) the ideal classical particle and the quantum pseudo-particle, and (ii) the ideal classical wave and the quantum pseudo-wave.

(i) *The ideal classical particle* is a permanent centre of a spherically symmetrical field, exactly localised in a four-coordinate system, with dimensions so small that the distinction between its parts can be neglected in the interactions of particles. In addition the particle of classical theory obeyed the laws of classical mechanics, usually exerted inverse square forces, possessed inertial mass, and contained no element capable of defining a specific length. (The rigid electron of radius $k.\frac{e^2}{mc^2}$ is a departure from the ideal particle.) The coordinates and momenta of a system of particles may be given arbitrary values, i.e. the particles are analytically independent.

The quantum pseudo-particle (electron, proton) of elementary mechanical problems has axial symmetry, exerts non-central magnetic forces, displays a wave-length in diffraction phenomena, and undergoes motions not exactly definable in terms of four observable coordinates, but determined by a probability function. The initial conditions of a set of electrons (or protons) are subject to restrictions, i.e. the particles are analytically dependent. These restrictions introduce dimensional constants capable of defining special lengths. The relevance of the particle concept to atomic structure is limited by the fact that the assignment of definite electron configurations to the energy levels in complex spectra is an approximate procedure which has no exact general meaning. (Condon*.)

The light-quantum (photon) has different properties which, again, are not those of the classical particle. In extreme cases certain aspects of the behaviour of electrons, protons and photons degenerate to that of the ideal particle.

(ii) *The ideal classical wave* is an extended system of field quantities, propagated in accordance with a linear "wave" equation connecting the differential coefficients of the wave-function with respect to the four coordinates and defining a velocity of phase propagation. An oscillatory wave solution of the equation has a definite period, and the velocity of propagation (and wave-length) may be treated as dependent on the medium at each point. The wave-length is observed by means of the interference properties of coherent waves from one source, owing to the property of the linear superposition of amplitudes.

The quantum pseudo-wave representing electrons or protons is a scheme of complex quantities (related to the relative probabilities of observable events) propagated in general in a higher dimensional space with a velocity which is a function of position relative to one or more fixed centres (nuclei). The complex quantities are not in general separable into two real parts as in ordinary classical waves. The wave-equation contains special dimensional constants capable of defining specific lengths. In special cases the quantum pseudo-wave reduces to a classical four-dimensional real wave-function.

The quantum pseudo-wave representing radiation (photons) approaches more closely the classical wave but cannot be exactly localised, and is compatible with the particle properties of radiation.

It has been stated that in certain simple cases the quantum pseudo-wave and pseudo-particle degenerate to the ordinary classical forms, and it is this fact which renders the terms valuable in the provisional description of atomic systems, and in dealing intuitively with simple problems. Yet even in these cases there are limitations (on the validity of the classical wave and particle descriptions) which have been interpreted as implying the elimination of the "physical reality" which localised particle tracks and propagated oscillatory waves enjoyed in classical theory. These limitations are:

(i) the exact mathematical representation of the concept of an electron path cannot be used, e.g. in describing the process of electron diffraction at a grating, unless a definite experiment is carried out to determine the path, while this experiment necessarily modifies the diffraction process, (Heisenberg*), and

(ii) the exact representation of the concept of a light wave, or other solution of Maxwell's equations representing the radiation from an atom, cannot be used unless an experiment is made to determine the coordinates of the nucleus, while this experiment necessarily modifies the radiation. (Heisenberg.*)

This situation has been described as the loss of reality or objectivity in the electron path and the light wave. The significance of this assertion is that the elementary process cannot be given exact description in terms of a descriptive method (four coordinates), the use of which already implies the prior application of other physical processes (application of clocks and rods). In other words, the process of establishing metrical coordinates should be included in, rather than assumed by, the description of elementary processes. From the standpoint of a course theory, in reference to the above examples, WHAT IS CALLED THE LOSS OF OBJECTIVITY OF THE PARTICLE MODEL OF MATTER, AND OF THE WAVE MODEL OF RADIATION, IS MERELY ONE ASPECT OF THE FAILURE OF FOUR-COORDINATE DESCRIPTION. In the literature of coordinate physics " objectivity " and " reality " usually mean describability in terms of four coordinates. It is important however to remember that the cases of the isolated oscillatory radiation field and the single localised electron path, though of great interest, are highly specialised situations, and that in more general problems such as the interactions of electrons and radiation or quantum-mechanical resonance the classical models break down completely.

The brief summaries given above of the properties of the quantum pseudo-particles and pseudo-waves serve to indicate the failure of the classical concepts rather than to describe the essential nature of quantum-mechanical problems. Moreover the quantum pseudo-particles and pseudo-waves have different properties when they are used to describe different kinds of material or radiational systems. It will therefore be necessary to discard entirely the terms waves and particles in fundamental synthetic theory. Waves and particles are conceptions derived from macroscopic phenomena which provide partial and misleading descriptions of two complementary aspects of one microscopic phenomenon. " Light and matter are unitary physical phenomena, their apparent duality arises from the essential inadequacy of our language." (Heisenberg*.) " The waves and particles should be regarded as two abstractions which are useful in describing the same physical reality." (Dirac*.)

The question immediately arises: why are these conceptions as useful as they are, if they do not correctly represent the actual phenomenon? What underlying unitary structure or process gives rise to these complementary dual aspects? Quantum mechanics regards waves and particles as attempts to describe the consequences of a set of abstract rules which determine phenomena. It must therefore be possible to find in these abstract rules the features which permit or justify the use of the particle and wave concepts in different situations. An analysis of these rules leads to the following results:

The rules

(i) determine an "eigenwert" problem, i.e. a mathematical problem leading to special values as well as continuous ranges for the variables.

(ii) involve i [8].

(iii) imply a special type of probability super position.

The Wave Theory interpretation of certain aspects of these rules is possible because

(i) Eigenwerte problems can be interpreted as representing systems of waves.

(ii) The appearance of i permits interpretation in terms of a periodic function.

(iii) A wave theory implies a limit to the accuracy of localisation of a diffracting particle, this inexactitude being compatible with the above superposition of probabilities.

The Particle Theory interpretation of other aspects is possible because the rules in certain cases imply an *approximate* localisation of phenomena (matter, energy) into small regions of space, or limited tracks in space-time.

This analysis implies a radical reversal of what was at first the normal method of interpretation. Until it was realised that waves and particles are only provisional and incorrect terms, it was natural to consider that quantum mechanics deals with eigenwert problems because there are waves in physical systems, that the probability-inexactitude aspect of quantum mechanics is a consequence of the wave-particle duality, and that i appears in the mathematics because it is a representation of oscillatory waves. Now that this attitude is no longer possible, it is necessary to reverse the interpretation and to accept the rules of quantum mechanics as the *primary facts at present known* about the structure of physical nature. Thus, for example, it is the presence of i in these rules which permits their interpretation —in certain cases—as the representation of a periodic function[8]. The fact that the quantum observables can be treated as periodic in the time is a direct deduction from and consequence of the presence of i in the commutation rules. This is an inversion of the original physical application of i in the representation of periodic phenomena, since i was then introduced in order to simplify the mathematical description of an oscillatory phenomenon, whereas in quantum mechanics its presence in the axioms permits a phenomenon (which is not known to be always or necessarily oscillatory) to be treated in terms of the mathematics of oscillations (though no oscillations in an atom have been directly observed).

The oscillatory wave and particle concepts are on this view attempts to describe a unitary structure, which succeed —as far as they do—on account of special characteristics of the axioms of quantum mechanics, these latter expressing as much as is yet known about the unitary structure. The various features of these axioms: the probability-superposition principle, the appearance of eigenwert problems, the presence of certain dimensional constants and of i, must express the properties of the unitary structure which is inaccurately described by the concepts of particles and waves and by the mathematics of classical particles and wave theory. All that can now be said of the unitary structure is that it must lead to the results of quantum mechanics, i.e. it must provide

(i) equilibrium structures of special sizes;

(ii) changing structures obeying an exact non-coordinate law, and probability coordinate laws.

(iii) a substitute for wave theory in terms of the closed courses involved in the interference of coherent rays.

(iv) a justification of the quantum probability superposition principle.

(v) an interpretation of the presence of i in the axioms of quantum mechanics.

(vi) an interpretation of the differences in the mode of application of the dual-wave particle model to radiation and to matter.

A unitary model or structure with these properties serving as the basis of a unitary theory must provide a derivation of the numbers α and β and this implies the treatment of radiation and matter as aspects of one underlying type of process. Thus the UNITARY STRUCTURE MUST DISPLAY, WHEN EXPRESSED IN TERMS OF FOUR COORDINATES, A PAIR OF DUAL ASPECTS: THE ASPECTS CALLED MATTER AND RADIATION, AND CUTTING ACROSS THESE THE ASPECTS CALLED PARTICLES AND WAVES. A unitary theory will therefore, in accounting for the partial validity of wave and particle models, also throw light on the transformation of matter into radiation.

§22. CLASSICAL, RELATIVITY, QUANTUM, AND UNITARY COURSE THEORY

The methods of relativity and quantum physics are shown as transition stages between those of classical physics and the proposed methods of unitary course theory.

The differences of classical, relativity, quantum and unitary course theory, which have been referred to in previous sections, may be summarised as follows:

Classical Theory is the approximate description of motions in terms of three spatial coordinates and a time-coordinate which, in the absence of a theory of the structure of rods and clocks, is regarded as independent of the spatial coordinates. The four coordinates define a background in which differential equations with arbitrary initial conditions determine the motions. The dynamical variables have a continuous range of possible values, and may be regarded as representing the interactions of analytically independent particles.

Relativity Theory is the approximate relativistic description of motions in terms of the metrical invariants of the space-time continuum which are related to the velocity of light and the proper-lengths and periods of unanalysed rods and clocks. The time-coordinate is still an independent variable, no general theory of the structure of clocks or of the origin of coordinates being given. Dynamical variables have a continuous range of possible values though velocities are restricted by an upper limit. Initial conditions are arbitrary, and motions are given by solutions to field equations.

Quantum Theory is the most exact description of motions and of the structure of matter and of radiation which is possible in terms of the four coordinates of classical theory. Dynamical variables have ranges of special values as well as continuous ranges, and physical systems are therefore described in terms of the solutions of eigenwert problems. The restriction on initial conditions implies the presence of analytically dependent elements (quantum particles) and eliminates the classical concept of the degrees of freedom of a system. In certain problems the substitution of an operator (defined as a function of the space coordinates) for the

Comparative Analysis

Classical.	*Relativity.*
3 Space Coordinates. 1 Time Coordinate.	4 Space-Time Coordinates.
Special coordinate systems. Non-invariants.	Arbitrary coordinate systems. Metrical invariants.
"Continuity," or more exactly no special values of dynamical variables, e.g., no special lengths.	"Continuity," though velocities limited in range, and special lengths involved at high velocities.
Kinematic similarity, permitting definition of Compound Dimensional Constants (c, G, e.).	No similarity at high velocities. Constants reinterpreted.
Arbitrary units, and dimensional theory.	Nearly arbitrary units, and dimensional theory.
Differential Laws.	Differential Metrical Laws.
Initial Conditions arbitrary. Solutions with given initial conditions.	Initial conditions nearly arbitrary. Special spherically symmetrical solutions.
	Classical analogy leads to use of ic in metrical form.
Periodic orbits as special case.	No exact solutions to two-body problem.
Laws imply more data than is available in many cases. Various problems therefore left undefined.	Laws imply less data than classical laws. "Ether-drift" eliminated and problems of relative motion, being adequately defined, are soluble.

OF PHYSICAL THEORIES

Quantum.	*Unitary Course.*
4 Space-Time Coordinates.	3 Space Coordinates. 1 Derived local time parameter.
(Uncertain.) (Nearly invariants.)	No four-coordinate systems, hence invariance meaningless.
" Discontinuity," i.e. special values of variables. Special lengths.	" Discontinuities " (e^2, h) reduced to common origin.
No similarity. Simple dimensional constants appear (masses and lengths: m, M, $\frac{h}{mc}$, etc.)	Kinematic similarity meaningless. Stationary structures have characteristic relative sizes. Dimensional constants eliminated by (i) derivation of α, β, γ, (ii) discarding use of arbitrary units.
Arbitrary units retained hence simple dimensional constants in all formulæ.	The relative sizes of elementary structures described without reference to arbitrary units. (Dimensional theory relevant only after introduction of four-coordinate system.)
In general no partial differential coefficients of dynamical variables. Operator or Matrix equations. Initial conditions restricted by quantum conditions. Eigenwerte in terms of dimensional constants.	No differential coefficients. Initial conditions given by coincidences.
Classical analogy leads to use of ih in quantum conditions.	ic and ih to be shown as consequences of using local derived time of course theory as extended coordinate and independent variable.
Analogy with classical periodic orbits fundamental.	Periodicity of rotation important. Periodicity of wave excluded since no extended time-coordinate.
Laws imply less data than classical laws. Exact space-time position and hence also phase of atomic resonators being eliminated, atomic interaction problem is now soluble.	Problem of interactions of electrons protons, radiation, and gravitation, (which is inadequately defined in earlier theories) to become soluble, i.e. derivation of α, β, γ, possible, at cost of laws containing less information than is implied in use of four coordinates.

momentum (defined classically in terms of a time-coordinate) implies the elimination of time as an independent coordinate. A strict relativistic quantum theory leads to contradictions because the use of relativistic methods implies the absence of any structural analysis of a clock, while quantum methods must supply this.

Course Theory is the exact description of direct measurements and space-time coincidences in terms of a theory of the structure of matter and of clocks, which does not assume a universal time-coordinate but derives a time parameter which is, for a process of given structure, a known function of the spatial coordinates. Course theory regards all coordinate processes as aspects of one underlying process and extends to all processes the use of a derived time parameter such as was defined in relativity only in terms of light propagation. Course theory however applies this parameter to defining coincidences without extending it as a universal time-coordinate. The problems of course theory will include (i) the derivation without the use of four coordinates of directly observed ratios of wave-lengths and molecular sizes (crystal lattice constants) and (ii) the derivation of the particular dimensional constants which are necessary if the unitary non-coordinate process is to be described in terms of inverse square laws in a four-coordinate description.

This survey of the three existing theories and the proposed aim of a unitary course theory is developed in the table on pages 98 and 99.

§23. THE SIX UNIVERSAL DIMENSIONAL CONSTANTS

The necessity in coordinate theory for six primary dimensional constants and the numerical values of the ratios α, β, γ, must be derived in course theory as the consequence of applying the methods of coordinate theory to a unitary structure.

Of the dimensional constants used in fundamental physical theory six (G, c, M, m, e, h) may be treated as primary since the remainder refer only to statistical assemblies. Of these constants the first five are used in the classical, relativity, and quantum energy functions (or Hamiltonians) for physical systems of different types and h is used in the formulation of the quantum conditions of a quantum-mechanical system. The six constants[9] may be taken as epitomising the whole of gravitational, electromagnetic, and atomic theory and an examination of them from the standpoint of a course theory will provide an opportunity of summarising from a fresh angle the conclusions already reached.

This examination will fall into two parts:

(i) the elimination from the six constants of trivial information resulting from the use of arbitrary units of length, time and mass.

(ii) the consideration from the standpoint of a course theory of the remaining information about physical structure provided by the existence of the constants.

(i) *The elimination of trivial information.*

The six empirical constants constitute six quantities which in coordinate theory are regarded as independent. Of these six independent quantities, three (in the form of combinations of the constants) are determined by the arbitrarily selected units of length, time and mass, while three dimensional invariants are left. These are the pure numbers α, β, γ, or combinations of these. Thus THERE REMAINS AS THE

SIGNIFICANT INFORMATION PROVIDED BY THE APPEARANCE OF THE SIX CONSTANTS THE QUALITATIVE FACT OF THE NECESSITY OF SIX CONSTANTS OF THIS TYPE IN COORDINATE THEORY, AND THE NUMERICAL VALUES OF THE THREE PURE NUMBERS.

The same information may be expressed in another manner:

The existence of three natural measures (physical units) of length, time and mass and of three universal constants which may be expressed either as non-dimensional numbers or as quantities of any desired dimensions.

In view of

(a) the appearance of universal *lengths* in the fundamental equations of current physics.

(b) the attempt made in physical theory to represent phenomena *geometrically*, and

(c) the priority of *space* measurement over time measurement (§9), it will sometimes prove convenient to give these three constants the form of additional lengths, making in all the four universal lengths discussed in §7. It is natural for a geometrical representation of physical theory to interpret the pure numbers α, β, γ, as ratios of lengths.

(ii) *Analysis of this information from the standpoint of a course theory.*

Firstly: The qualitative fact of the necessity for six constants with the particular dimensions possessed by G, c, M, m, e, h.

c when interpreted as a velocity, permits time to be measured in terms of space, and is used in establishing an extended time-coordinate. Thus c underlies the use of coordinates. In four-coordinate theory the (special) invariance of c compels the representation of space-time phenomena in a continuum of indefinite metric, which fails to provide invariant metrical representation of the experienced topological ordering of events. If the demand for topological representation is accepted, the appearance of c compels explicit recognition of the topological fact that it is measured, not as a classical velocity, but as a ratio of finite quantities over a closed course. Thus reinterpreted c can be retained in course theory. All formulæ containing c in relativistic correction terms represent compensations for the neglect (involved in the conception of the velocity of light, i.e. in the use of four coordinates) of the closed course property of light observations. Such corrections involve $\frac{v^2}{c^2}$. In relativistic formulæ bearing exact formal correspondence with classical formulæ (e.g. in the metrical form corresponding to the Euclidean form) c appears as $i.c.$ This fact must be used in the transition to course theory. (Compare under h.)

h must *in some form* underlie the use of coordinates, since combinations with h not only define the scale of the rods used in establishing coordinates, but also set a limit to the accuracy with which they can be measured. But as a constant of action h implies coordinate description, and cannot be retained in course theory. It can only be retained (i) in combinations yielding directly measurable lengths (see under m), and (ii) in the form $\frac{e^2}{h}$ which is given as a closed course distance ÷ time in the theory of electron orbits. All formulæ containing h represent either special lengths found in physical structures, or compensations for the neglect (involved in the conception of electron velocity, i.e. in the use of four coordinates) of the closed course property of electron observations, e.g. in the Heisenberg inexactitude formulæ. In all the quantum conditions required to supplement the expressions obtained by the use of the correspondence principle from classical formulæ h appears as $i.h.$ This parallel with c must be used in the transition to course theory.

G is macroscopic, is not used in establishing four coordinates, therefore does not underlie coordinates, cannot be retained in course theory, and must be interpreted as the result of the introduction of coordinates. From the standpoint of unitary course theory it represents the residual interaction of "neutral" bodies at macroscopic distances, and must be derived as an approximate expression in coordinate theory of a degenerate case of atomic (electromagnetic) actions.

e does not underlie the use of coordinates, cannot be retained *in this form* in course theory, and expresses in terms of coordinates a uniformity in electronic and nuclear interactions. The combination $\frac{e^2}{h}$ may however be retained in course theory.

M does not underlie the use of coordinates, cannot be retained in course theory, and expresses the coordinate property of the existence of identical finite inertial units. This is to be as an expression in terms of coordinates of a symmetry in the inter-relatedness of course phenomena. Special combinations of M representing special wave-lengths (e.g. $\frac{h}{Mc}$) can be retained in course theory.

m does not underlie the use of coordinates, cannot be retained in this form in course theory, and is to be derived as the result of the application of coordinates to electron courses. Combinations such as $\frac{h}{mc}, \frac{h^2}{4\pi^2me^2}$, etc., may be retained in course theory since they represent directly measurable lengths.

Summary of analysis of the qualitative fact of the necessity of six constants of this type.

(a) One constant, G, is macroscopic, approximate, represents a limiting case of material interactions, and does not appear at all in exact course theory.

Four constants, M, m, e, h, can be retained in course theory, in other combinations. One constant c can be retained in the same form after reinterpretation.

(b) c and $\frac{e^2}{h}$ represent ratios of finite lengths/finite times over closed courses for light and for electrons respectively, c being directly measured as such and $\frac{e^2}{h}$ inferred principally from the interference of coherent rays forming closed courses. These quantities can be retained in course theory and are dimensionally and topologically similar. The deep correlation of these quantities is emphasised by the fact that in formulæ bearing a formal analogy to classical expressions, both c and h appear with the factor i.

(c) In addition to these, the lengths $\frac{h}{mc}$, $\frac{h^2}{4\pi^2 me^2}$, $\frac{h^3 c}{2\pi^2 me^4}$ (and combinations involving $\frac{M}{m}$), can be retained in course theory.

Secondly: The Numerical Values of the Ratios.

$\alpha = \frac{2\pi e^2}{hc}$ From the standpoint of unitary course theory there is one universal unitary type of space-time structure capable of description without four coordinates. When coordinates are introduced, this structure displays various aspects. One aspect is called light, and permits the measurement of c. Another aspect is called electronic motion, and permits (in relation to radiation) the inference of a quantity $\frac{e^2}{h}$. The ratio $\frac{2\pi e^2}{hc}$ is therefore to be derived in course theory by consideration of the way in which physical theory has distinguished the two aspects called "light" and "electrons." In spectral fine structure the directly measured wave-lengths are functions of $\frac{h}{mc}$ and $\frac{h^3c}{me^4}$, and α^2 is the empirically given quantity. Bohr has pointed out that the error in the quantum-mechanical neglect of radiation reaction is of the order of α^2. Course theory regards the inaccuracy of treating electrons and light as independent processes, instead of as aspects of one process, as of this same order.

$\beta=\frac{M}{m}$. A numerical coefficient involved in the derivation of the phenomena of inertia in coordinate theory from the fundamental principles of course theory. The partial mathematical similarity of electrons and protons may conceal a far-going asymmetry in their relation to the underlying structure.

$\gamma=\frac{e^2}{GMM}$. To be derived in the deduction of coordinate physics from course theory as the ratio of the primary interaction represented by e^2, to a limiting (degenerate) form of interaction. Large numbers of this order can easily appear in view of the exponential forms used in quantum mechanics.

§24. SYMMETRIES IN COORDINATE LAWS

The special form of coordinate laws raises questions regarding their symmetry which may be of importance in the transition to course theory.

The symmetry of physical laws with reference to the four coordinates presents several important problems, which will now be discussed.

Reversibility. In classical theory isolated processes are reversible, i.e. their representation in a non-rotating coordinate system involves only even powers of t. In a rotating coordinate system processes are irreversible, i.e. they involve an odd power of t. Reversibility is therefore not an invariant property. Electronic motions in a magnetic field are only reversible if the field also be reversed.

In relativity theory local coordinate systems can always be found for which a process at a point is reversible. Non-local or general reversibility cannot be discussed in relativity theory since (i) the property is not invariant and (ii) exact solutions of the two body problem are not available.

In quantum theory the wave function for an isolated system (in N-dimensional complex space) obeys a reversible law, provided the spin terms are also reversed. But the physical importance of this is limited by the following facts:

(i) No physical process is isolated.

(ii) No relativistic wave equation for complex systems has yet been given.

(iii) The statistical nature of quantum laws modifies the relation of the reversibility of laws to the reversibility of actual physical processes. This point will now be examined in detail, starting with classical statistical laws.

Reversibility in Statistical Law.

(a) The classical statistical description of assemblies.

Since an exact coordinate description of an assembly of molecules was in classical theory unattainable for practical reasons, a statistical description was given instead in terms of the average, or the most probable, value of the macroscopic parameters. Irreversible heat processes were treated as the change of these parameters from less probable to more probable values. This theory is of interest for two connected reasons, (i) it involves the use of an asymmetrical relation of temporal succession which is alien to pure metrical physics (cf. §§34, 35) and (ii) though the theory appears to start from axioms devoid of the relation of succession, it obtains the result that systems tend most frequently to move from less probable to more probable states.

The first point will be taken up later. The second shows that the idea of earlier and later has somewhere been introduced into the mathematical theory. The solution to this paradox is that though the primary classical concepts do not involve the relation of succession, yet the statistical concept of the " probability of a process, given the statistical description of the initial state " implies the use of this relation (10).

In equilibrium states statistically *inverse* processes are equally probable (Principle of Detailed Balancing) ; these are distinct from the kinematically *reversed* processes. (Dirac*, Jordan*.)

(b) The quantum statistical description of individual processes.

Since the exact coordinate description of an individual elementary process is in quantum theory unattainable for theoretical reasons, a statistical description is given in terms of the probabilities of different observations if the elementary process be repeated. The probability function (Schrödinger function in higher space) obeys a reversible equation in the case of isolated systems, but this does not mean that in the special solutions which represent actual experiments the probability function is necessarily reversible. The nature of quantum-mechanical observation and description is such that the four-coordinate description necessarily takes the form of the relative probabilities ZZ of different observations *subsequent to* an initial observation Y, e.g. in the spreading of wave packets. It is not possible, in general, to give a physical meaning to the reversal of this special solution, i.e. it will in general be empirically valueless to say that given a system with initial relative probabilities ZZ a definite observation Y can be made subsequently. This assertion, though mathematically legitimate, does not refer to an empirically identifiable situation, since the empirical initial conditions of an individual process cannot be known in the form of relative probabilities, except in very special cases.

The general character of the quantum theory prevents any simple and unambiguous meaning to be given to the terms "reversible" and "irreversible," since isolated processes are reversible, but any observation of them alters them irreversibly. The simple reversibility of classical processes disappears because it is empirically meaningless in quantum theory to consider an isolated process. In neither classical nor quantum theory can non-isolable systems be reversed.

Ambiguities due to alternative signs.

The quadratic expressions underlying classical, relativity and quantum theory permit ambiguities in their solutions owing to the possibility of taking positive and negative square roots. This leads to no difficulty in the case of ordinary reversible processes, where the ambiguity $\pm t$ is permissible. But in quantum theory it results in allowing positive and negative energies and positive and negative charges with symmetrical properties which are not required empirically.

A similar but less serious ambiguity is that of the sign of integration constant in the spherically symmetrical solution of the field equations of general relativity. This integration constant can have arbitrary positive or negative values, but its empirical interpretation as the gravitational mass of a particle at the centre requires only positive values to be permitted. A negative value would result in a form of geodesic implying a repulsive gravitational action.

These ambiguities are due to the use of differential equations and quadratic forms and can be eliminated only by the use of linear expressions not permitting arbitrary constants.

Enantiomorphic Forms.

The coordinate representation of an axial vector quantity changes sign when a left-handed coordinate system is substituted for a right-handed system. Thus Maxwell's equations in their usual form do not remain invariant under this substitution, but require reformulation with changed signs (cf. Einstein*). It is improbable that either the right-handed or the left-handed screw need have priority in physical theory, and the principle of covariance suggests that physical laws should be formulated in a form which is indifferent to the selection of right- or left-handed coordinate systems. A deductive derivation of coordinate laws from the principles of course theory would be expected to yield these laws in invariant form.

Cartesian and Polar Coordinates.

A pervasive duality runs through coordinate theory corresponding to the properties which receive simple symmetrical expression in Cartesian and in polar coordinates respectively. Translational motions obeying the conservation of linear momentum, and rotational motions under central forces implying the conservation of angular momentum display on the one hand formal similarities, and on the other hand essential differences. The similarities result in the great value of the Lagrange notation of generalised dynamical coordinates; the differences arise chiefly from the fact that translational motion must be reversed to return to the starting point, whereas rotational motion is essentially periodic. This asymmetrical duality in the four-coordinate descriptions of motions leads finally in quantum theory to the two constants c and h. The reduction of c and $\frac{e^2}{h}$ to one common origin must be based on the view that in applying the macroscopic coordinate principles of the conservation of linear and angular momentum to a microscopic unitary course phenomenon, the latter has necessarily been split into two aspects (light and electricity) described in simple cases by a characteristic to-and-fro velocity (c), and a characteristic circular orbital velocity $(\frac{e^2}{h})$, respectively.

Beneath these analogies there lies the deep distinction between rotational and translational quantities arising from the fact that the length of an arc can only be defined as the limit of a sum of finite chords. The primary meaning of length for empirical theory rests on the direct comparison of finite structures, and has no meaning for curved lines. Thus both for empirical and for pure mathematical theory π is definable only as the limit of a sum of series of finite terms. This may be of importance for a theory which deals only with the ratios of finite quantities, and does not admit differentials in the formulation of primary physical law. For example the recognition that quantum theory is based entirely on the measurement of special lengths along straight lines (e.g. wave lengths, sizes of molecules) suggests the observation that the angular momentum and angular periodicity properties of microscopic quantum-mechanical systems are merely inferred from the pseudo-classical form of the mathematics and are not directly based on empirical fact.

§25. UNSOLVED PROBLEMS IN COORDINATE THEORY

The completion of general relativity theory and of the general methods of quantum theory leaves many problems unsolved.

The following are the most important problems at the present time left in an unsatisfactory state by coordinate theory:

1. The examination of the relation of time measurement to space measurement on the basis of a structural theory of matter, i.e. of rods and clocks.
2. The fusion of gravitational, electromagnetic, and quantum phenomena.
3. The absence of exact solutions of two or more body problems in the general theory of relativity and of an adequate analysis of the empirical meaning (if any) of the differential metric in intense non-static fields.
4. The necessity in general relativity of assuming the Newtonian attractive law as a first approximation in order to exclude the "repulsive" geodesics permitted by the mathematical notation.
5. The absence of an exact relativistic quantum theory.
6. The peculiar nature of the dualities, (a) electron and proton, (b) matter and radiation, (c) particle and wave.
7. The inadequacy of the correspondence principle.
8. The absence of a mathematical technique for the exact solution of the many-body problems involved in complex atomic spectra, chemical actions, molecular structure, crystal structure, etc.
9. The interpretation of the abstract algebraic basis of the transformation theory.
10. An exact quantitative theory of nuclear structure, atomic weights, and radioactivity.

§26. SUMMARY

The form of a unitary course theory is closely defined by the conditions it must satisfy.

The analysis of four-coordinate theories from the standpoint of a unitary course theory has revealed the following characteristics of a unitary course theory in addition to those already summarised in §14.

(i) Unitary course theory must rest on elementary unitary laws representing the whole interaction (radiational, electrostatic and magnetic, gravitational) of particular elementary systems.

(ii) Course theory may use as a fundamental kinematic quantity the ratio of proper length to proper time over a closed course.

(iii) Unitary course theory must represent the relativity corrections and quantum conditions as expressions of the restriction of course laws to this special kinematic quantity, i.e. as compensations for the neglect of the closed course property resulting from the use of the concept velocity.

(iv) It must supply a unitary course structure with special topological and metrical properties to supplant the pseudo-wave-particle model of quantum coordinate theory.

(v) The unitary course structure cannot have arbitrary initial conditions, but must be applicable only to determine coincidences following on earlier initial coincidences, i.e. the laws of motion and the initial conditions of classical theory must be combined in the descriptive method of the unitary course theory.

(vi) Unitary course laws must describe the ratios of the sizes and periods of natural structures of different complexities without reference to arbitrary units.

(vii) Finally, unitary course theory since definable as a deductive theory of measurement and of the relations between measured quantities, must provide solutions for the unsolved quantitative problems of coordinate theory collected in §25.

Chapter III

THE GENERAL FEATURES OF UNITARY COURSE THEORY

§27 Introductory

The fundamental revision of physical theory which is implied in the aim of unitary course theory is justifiable only if it can restore simplicity and clarity to basic physical concepts.

In Chapter I an examination of space-time phenomena led to the hypothesis that a unitary course theory provides the correct form for fundamental physical theory. In Chapter II the special characteristics of coordinate theories were examined from the standpoint of this hypothesis and further features of a unitary course theory were deduced. Chapter III is concerned with certain general properties of unitary course theory on its own ground. We are now no longer using the results of coordinate physics to suggest what a unitary course theory must be, but are studying the relation of the conception of a unitary course theory reached in the previous chapters to general problems of physical concept and scientific method. The normal procedure in connection with the reformulation of a special region of physical theory is to postpone the consideration of general problems of concept and method until a precise mathematical expression has been established which can control and limit the range of these wider problems. The reverse procedure is required where the aim is a comprehensive synthesis based on a revision of the fundamental conceptions underlying the whole of physical theory. In this case the recognition of new general principles must precede mathematical theory. A fusion of this kind can only be attained by a deep analysis of the meaning of quantitative physical theory necessarily involving a revision of current conceptions of physical method. In face of the mass of detailed fact accumulated during over two centuries and the prestige of the methods which have rendered this accumulation possible, no far-going revision can be accepted even as worthy of study unless it can be shown in advance that the precision and clarity of scientific method will not be thereby damaged. The complexity of the fundamental concepts of quantum theory and the many unsolved problems which they raise render them unsuitable to serve as the basis of scientific theory. In the following sections it is suggested that if a unitary course theory can be established it holds the possibility of a clarity, precision and synthetic range adequate to justify a radical modification of fundamental concepts.

§28. THE THREE NUMBERS α, β, γ OF COORDINATE THEORY

A theoretical derivation of α, β, γ by a method going behind the use of coordinates provides the natural culmination of the movement towards unification in the history of physics.

We have seen (§§2, 7) that the definition of a unitary theory implies that it must provide a theoretical derivation of the three numbers α, β, γ appearing in coordinate theory. It will be convenient to examine by what general types of method such derivations could be obtained.

A theoretical derivation of one of these numbers implies the *identification* and *equation* of two quantities previously treated as of independent origin. Thus if α is to be derived, a theory must assert that $\frac{2\pi e^2}{h}$ means the same thing as i.e. is identical with $a.c$, where a is a special numerical factor given by the theory. THE ESSENCE OF A DERIVATION OF α, β, OR γ IS THE POSTULATE THAT TWO PHYSICAL QUANTITIES HITHERTO TREATED AS INDEPENDENT HAVE A COMMON ORIGIN AND HAVE IDENTICAL MEANING IN REFERENCE TO ULTIMATE PHYSICAL STRUCTURE PROVIDED AN APPROPRIATE NUMERICAL FACTOR IS INTRODUCED. We may now ask, what general forms can such an identification take ?

Within the normal range of the concepts of four-coordinate classical field physics it is improbable that any significant identification could be made. For example various trivial attempts have been made to derive one or more of these numbers using elementary coordinate theory, but these cannot be regarded as significant, since they are based on *ad hoc* hypotheses which throw no new light on physical theory as a whole. Outside classical field theory, however, significant identifications of this kind become possible. Thus in the general theory of relativity inertia and gravitation are treated as aspects of a non-euclidean physical geometry, and so could be identified. But the identification of these concepts, though explaining the proportionality of mass and weight, led to no new numerical relationship. Another extremely important case is provided by the identification of the relativity effects in the fine-structure of spectral lines with the consequences of electron spin. These phenomena have been shown by Dirac to arise as consequences of an appropriately selected relativistic wave equation for the electron. But this identification of spin and relativity effects, or more exactly their reduction to a common mathematical basis, did not supply a new numerical relationship since the electron spin had not been measured independently of the fine-structure of spectral lines.

On the other hand an interesting example of the possibility, in principle, of the derivation of α has been given by Eddington [11] starting from the Fermi-Dirac statistics and Dirac's relativity equation for the electron. In these theories two forms of interaction between electrons are assumed; (a) the Coulomb interaction with a potential $\frac{e^2}{r}$, and (b) the Fermi-Dirac formulation of the Pauli exclusion principle. In the interests of simplicity of description, it is assumed that these two forms of interaction are not independent, and that the more general form of interaction (b) shall include (a) as one of its consequences. The most natural expression of the exclusion principle is found to lead to a term in the wave equation for two electrons proportional to $\frac{1}{r}$, i.e. having the same form $(\frac{1}{136} \cdot \frac{h}{2\pi} \cdot \frac{1}{r})$ as the term of the Coulomb potential $(\frac{1}{c} \cdot \frac{e^2}{r})$. The equation of these terms yields the integer 136 (later modified to 137) for $\frac{1}{a}$. The equated terms represent interactions between two electrons expressed as probability factors, i.e. differences in the probable expections in relation to one electron owing to the presence of the other. Essentially, *probabilities* are equated, and the integral value for $\frac{1}{a}$ arises from an enumeration of the number of degrees of freedom assigned to the system. For various reasons [11] this derivation is not generally accepted as correct, but it provides a clear example of the general characteristics which any derivation of a, β, or γ must display.

It is notable that Eddington's method goes beyond the concepts of ordinary four-coordinate field theory and is compelled to use the higher spaces of the probability theory. This exemplifies the need, in any significant identification of quantities of four-coordinate theory to go beyond its normal range. But instead of passing to the higher space in which quantum probabilities are calculable the same end might be achieved by the elimination of coordinates. Course theory seeks to derive the three numerical constants a, β, γ, by identifying three pairs of coordinate quantities. This identification is to be justified by interpreting each pair as a dual description (in terms of coordinate theory) of one essential type of space-time structure definable without the use of four coordinates. Eddington correlates h and $\frac{e^2}{c}$ by identifying two forms of interaction expressed as *probabilities;* unitary course theory seeks to correlate all three pairs of quantities by identifying each pair as dual aspects of a unitary course structure.

The possibility must be kept open that further research may prove that the six-dimensional constants are not exact constants. For example, the value of e might be found to vary under different conditions of measurement or the effects of electron spin might conceivably be such as to prohibit the measurement of e beyond a certain accuracy. If this situation arose empirically then a unitary theory would be called upon only to provide a derivation of the limiting values reached by a, β, γ when the six constants are measured under given ideal conditions.

The conception of a theoretical derivation of a, β, γ, based on the identification of quantities now treated as of different origin is the radical expression in the present situation of the ideal of simplification which has inspired the greatest advances in physical theory. The Newtonian Theory of Gravitation rested on the numerical identification (1682) of the attraction of the earth on the moon with the terrestial gravitational attraction ; Maxwell's Electromagnetic Theory of Light was based on the numerical agreement (1864) of the ratio of the two electrical units with the velocity of light, and modern Electron Theory is founded on the empirically discovered equality of the electronic and protonic charge (1897). The importance of G and c as *universal* constants rests on identifications which once seemed as arbitrary—from the point of view of older theories—as does the proposal of a triple derivation of α, β, γ from the standpoint of coordinate theory.

The history of the dimensional constants reveals in a concise form the ceaseless movement towards unification which marks scientific thought. In the epoch from 1676 (Römer) to 1899 (Planck) the gradually collected knowledge of fundamental physical processes was being summarised in *six independent universal constants.* The years 1900 to 1930 have been spent in combining the regions symbolised by these constants, i.e. in developing physical laws which involve *combinations of the constants.* This must be followed by a period which will be occupied in attempting to provide a unified *derivation of the three pure numbers* which epitomise the total fundamental knowledge won during the first two periods. The hypothesis of a unitary course theory is put forward in order to focus attention on the new problems of this third period.

§29. LOGIC, PHYSICAL CONCEPT, AND EMPIRICAL FACT

The logical, conceptual, and empirical aspects of physical science find their common root in the idea of order.

We now pass from the special characteristics of unitary course theory to a consideration of its relation to the problems of scientific method. Every physical theory presents the three aspects: logical development, assumption as to the nature of physical reality expressed in the selection of fundamental physical concept, and empirical fact based on sense-perception. (Planck*.) These aspects will serve as useful lines of approach to the general questions of scientific method in connection with a unitary course theory.

The *logical form* of a physical theory is expressed with greatest precision in its mathematical development, which usually aims at the deduction of the whole theoretical system from a minimum set of axioms. When expressed in this axiomatic form[12], the unity, simplicity, and self-consistency of a conceptual or mathematical system are most easily recognised. The *physical concepts* of the system necessarily imply certain metaphysical assumptions, i.e. a conception of the nature or form of physical reality which usually goes beyond the given facts and provides a physical model or other leading idea enabling the variety of immediate fact to be organised in a synthesis. The selection of what is to be treated as *empirical fact* out of a group of vaguely defined sense-perceptions provides the theory with its crucial basis in experience, but involves a preliminary choice of the way in which sense-perceptions are to be arranged. This choice depends primarily upon the concepts which are to be used. Every physical theory displays these three aspects; it must be developed logically, it must use concepts defining to what the theory refers, and it must correlate empirical facts based on sense-perceptions.

An exaggerated emphasis on each of these complementary aspects leads to the treatment of physical theory as an *exercise in logic*, as a philosophic or conceptual *interpretation* of experience, or merely as the *simplest description* of phenomena. No one of these views of physical science is adequate; the axiomatic or logical view in stressing the rationality of theory neglects the fact that no complete science yet exists and that all discovery rests on the synthetic non-rational intuition that there is a persisting uniformity and simplicity in nature; the interpretative view collapses because in a deductive theory the primary concepts are necessarily indefinable; while pure empiricism cannot stand alone because it lacks a precise criterion of what is observable, i.e. of what is indeed fact. These partial aspects and interpretations of science are complementary, and represent not independent elements arbitrarily combined in science, but the inevitable features of the process of systematic conceptual thought about experience.

Logical or mathematical development and empirical fact are obviously both necessary to any systematic thinking about physical experience. It is not equally clear that concepts implying metaphysical assumptions are also necessary, and one school of thought has suggested that such assumptions can and should be avoided by the progressive elimination of all that is not immediately observable. We shall therefore examine whether a physical model or some other constructive physical conception is necessary.

A physical model in the sense of the English tradition of the last century can play no part in the final formulation of a strict deductive theory. A model of this kind, for example a system of mechanical, wave, or vortex motions, implies inexact or incomplete argument by analogy from a macroscopic to a *different* microscopic situation. Analogy rests on the selection of common features in different situations; by its meaning argument by analogy cannot be complete or exact. For example, a mechanical model of the atom (i.e. a model using only the conceptions of macroscopic mechanics such as the motions of spherical particles, etc.) could never explain the rules governing the combination of atoms into complex chemical systems, because macroscopic mechanics contains no analogous rules. Moreover we have seen that macroscopic analogies necessarily lack the definiteness of linear scale indispensable to a micro-structural theory. (§§3, 6.)

Argument by analogy is the supreme heuristic weapon; but when discovery is to be succeeded by precise formulation it must be laid aside since it is incompatible with the exact processes of a deductive theory. The axioms of such a theory must represent elementary forms of structure or order underlying all the more complex patterns into which they may be combined. It is logically as absurd to reduce mechanical phenomena and waves in real physical media to atoms, and then atoms to mechanical particles and physical waves, as to interpret earth, fire and water as the interaction of atoms and then atoms as themselves built up of earth, fire and water. Yet this is what is attempted if the pseudo-waves and pseudo-particles of quantum mechanics are regarded as anything more than severely modified mathematical forms which bear a few traces of their ancestry.

Though models cannot play a part in a final deductive formulation, a constructive physical conception of a different type is indispensable. It is necessary to know to what the assertions of a scientific theory refer, and a physical concept is therefore required which defines the subject matter of the science. But science is possible, and is only possible, where facts can be classified or arranged, for science is nothing but the classification of fact. The most general form such classification can take is in terms of a given type of relation-structure or *relation-number* [18], i.e. a pattern of relationships of definite logical types (symmetrical, or asymmetrical, transitive, etc.). Science, as the classification of fact, necessarily seeks to formulate the structure of the relations involved in the particular classification of fact found in experience. Now it happens that space-time order may here be substituted for "relation-structure" without loss of generality since every logical type of relation required in physics can be represented in terms of the symmetrical and asymmetrical relations of physical events. Physical science may therefore be defined as the study of observed order, as was done in §12. NO METAPHYSICAL ASSUMPTION OTHER THAN THE CONDITION THAT NATURE IS TO BE THE SUBJECT OF SYSTEMATIC REPRESENTATION IS INVOLVED IN THE SELECTION OF ORDER AS THE PRIMARY PHYSICAL CONCEPT. This is an immediate consequence of the fact that by order is meant a pattern of relationships general enough to include any possible classification of facts. Order is, in fact, the most general scientific concept, and all other scientific concepts must define special forms of order. Thought about physical nature is possible because order is both a logical property (i.e. a combination of relations of abstract logical types) and a physical property (i.e. space-time observations provide systematic examples of relations of definite logical types).

Logic, physical concept and fact meet in the conception of order. Order is the essence of given fact, the most general physical concept, and a projection of the logical structure of thought. This permits a fusion of the three partial interpretations of science which arise from emphasising each of the aspects of scientific theory: logic, concept and fact. Science, defined as the study of order, is more than axiomatics, for it emphasises the fact that only because a systematic order is found is reason applicable to nature; more than interpretative philosophy, for it reveals all concepts as merely names for special types of order; more than pure empiricism, for it defines what is observable. Science uses and goes beyond mathematics, philosophic interpretation, and observation, for it is the precise, i.e. symbolic, formulation of observed order.

The key to the proposed revision of physical theory must thus be provided by a formal intuition which is at once an induction from fact, the construction of a new type of space-time order, and the formulation of an abstract relation-structure.

We shall now proceed to the discussion of the logical and conceptual structure of a unitary course theory (§30) and its properties of a simple description of fact (§31).

§30. THE STRUCTURE OF UNITARY COURSE THEORY

The deductive sequence of a unitary course theory must reverse the historical order of the discovery of physical principles.

A deductive theory starts with a set of undefined concepts and unproved propositions (axioms) connecting them. The axioms of unitary course theory have to define a universal type of ordering of events in space and time. This may be regarded as an elementary form of space-time structure with both topological and metrical properties, all other structures or processes being regarded as special combinations or special aspects of this universal type of structure. In §21 this universal or elementary type of space-time ordering underlying the phenomena now partially described by the terms "particle" and "wave" was called the *unitary structure*. The axioms of unitary course theory define the unitary structure.

These axioms must include one defining an asymmetrical transitive two-termed relation since a relation of this type is essential in defining temporal succession and cannot be derived from other types of relation. The axioms must also include, or permit the derivation of, a symmetrical relation in a pattern of four-events, such as the symmetry (congruence) of two neighbouring physical systems. A persisting physical symmetry will then serve as a definition of congruence, and permit the definition of the process of spatial measurement, and of the secondary derivation of a time parameter in terms of spatial processes.

The further development of the theory will be guided towards the complex conceptual and mathematical structure of current physics by the forced or logically arbitrary adoption of physical representation in terms of a special structural form (a four-coordinate system) which being applied in a manner incompatible with the axioms produces logical inconsistencies (appearing as the paradoxes and physical falsity of relativistic quantum theory) and numerical inexactitude (physical indeterminateness).

The adoption of a descriptive method which is inconsistent with the axioms since it treats t as an independent variable though only given in the axioms as a dependent variable is an act of logical violence which produces remarkable results. Prior to the act we have a self-consistent set of axioms describing a single comprehensive type of order in nature, and rendering possible the separate measurement of space and of time up to any accuracy. Let us regard this group of axioms as a stencil of single harmonic pattern through which physical nature may be viewed, and being so viewed reveals its ultimate structure. Subsequent to the act, which we may interpret as the superposition on the first stencil of a second stencil of different pattern from the first, our view of phenomena is complex and obscure. We find a pattern certainly, but it is both complicated and inexact. It is no longer a complete or logical pattern, for it contains inconsistencies and inexactitudes ; the stencils have to be warped a little to fit. Different vistas cross this new pattern ; we call them electrons, protons, radiation, gravitation. Thus we find four patterns where previously there was only one, and these four are of different scales determined by four universal lengths. The ratios of these four lengths constitute three pure numbers which are the result of using the second stencil and play no part in the ultimate structure of nature. Yet for isolated regions of phenomena these bastard sub-patterns are useful, for where we can neglect their inexactitude and their interactions they obey well known and convenient mathematical laws. As instruments for the specialist they have been and will remain invaluable. Yet they are based on false assumptions and are only made possible by forcing an arbitrary descriptive method on to the primitive pattern of nature.

In the following diagram the historical development of coordinate theory from its classical foundations to the abstract principles of quantum mechanics is contrasted with the proposed deductive derivation of the same material from the physically significant axioms of a unitary course theory.

The Structure of Coordinate and Unitary Course Theory

Coordinate Theory.	*Unitary Course Theory as proposed.*
Tendency of historical development (approximate).	Deductive sequence reversing historical order.
A. In four-coordinate space-time frame.	
Inertia concept. Momentum conservation principle. Energy concept. Energy conservation principle. Inverse Square Laws of Gravitation and Electrostatics (G,e). Wave concept. Wave amplitude superposition principle. Electromagnetic field Laws (c).	Classical particle and wave theories provide representations of the unitary structure, valid within definite limitations.
Covariance concept. Special and general covariance, and equivalence principles. Field law of gravitation.	Covariant theory provides high approximation to consequences of course theory when expressed in terms of four coordinates.
Quantum constant h. Probability-inexactitude and Eigenwerte principles. Correspondence principle. Various special quantum conditions.	Quantum theory provides the most exact description of the unitary course structure possible in terms of four coordinates. The numbers α, β, γ result from introduction of coordinates.
(Transition from coordinate space-time to abstract notation.)	(Introduction of four-coordinate description and use of arbitrary units causes transition from unitary course theory to statistical, inexact coordinate theory.)
B. Abstract, or without immediate physical significance, because not in coordinate space-time.	*C*. Physically significant, though not in terms of four coordinates.

(Abstract notation to be given physical basis in terms of a non-coordinate theory.)

Matrix and Operator concepts. Probability transformation theory. Non-commutative multiplication rules.	$\longrightarrow$	Closed course concept. Derived local time parameter. Unknown new type of metrical law.

§31. SIMPLICITY IN PHYSICAL THEORY

A microscopic structural theory can combine simple concepts and simple basic laws.

Science is necessary because nature is complex, and is possible because the complexity of nature is limited. Thus science is the discovery and formulation of the fact that the complexity of nature is limited by an underlying order, and scientific theory must therefore in one respect be simple and in another respect mirror the complexity of nature. Therefore in the deductive development of a theory from fundamental concept and law to the description of sense perceptions there must be aspects where it is simple and other aspects where it is complex. For example the laws, if they are to serve their purpose, must be simple, while the descriptions of perceptions must be capable of a high degree of variety. This has suggested the view that "advances in scientific knowledge must bring about the result that an increase of formal simplicity can only be won at the cost of an increased distance or gap between the fundamental hypothesis of theory on the one hand, and the directly observed facts on the other." (Einstein*.) "Nature is such that simple things do not enjoy simple laws, so that in order to simplify laws we must complicate the meaning of their terms." (Nicod*.) Here "simple things" are those in terms of which common sense describes the world. According to this view there are two kinds of simplicity which may be present in a physical theory: the *intrinsic* simplicity of the laws obeyed by its primitive concepts, and the *extrinsic* simplicity of the interpretation of these primary concepts in terms of sense-perceptions. The intrinsic simplicity of the laws of mathematical physics has been won by using abstract (i.e. extrinsically complex) concepts. It is therefore suggested that intrinsic and extrinsic simplicity are incompatible, and that physical theory necessarily seeks and wins intrinsic simplicity at the cost of an increasing extrinsic complexity.

Cogent evidence supports this argument. It is natural to assume that the stupendous variety of perception-patterns can only be reduced to order by the use of concepts based on a process of successive abstraction from the data of perception. Moreover this has been the actual tendency of theoretical physics from Newton to Dirac; repeated attempts to remain close to the facts have never been able to arrest the increasing distance from perception in the historical series of concepts: gravitational acceleration, differential laws, potentials, covariant metrical field laws in a continuum of indefinite metric, wave functions in a complex higher space, higher numbers with non-commutative multiplication. This movement has not been appreciably checked by the elimination of unobservable elements by Einstein (absolute simultaneity) or by Heisenberg (exact electron orbits), because neither of these workers was able to propose new fundamental concepts which represented immediate features of sense perception rather than the extrinsically complex symbolic forms (tensor or matrix equations) which those features obeyed. Hence in both cases the elimination of an unobservable was accompanied by an increased extrinsic complexity in the basic concepts of theory.

Yet this tendency is not inescapable. It is in principle possible to have a theory with extrinsically simple concepts obeying intrinsically simple laws, which is none the less capable of describing the variety of phenomena. For example, a theory based on the conception of a certain type of space-time ordering of events given in observation, e.g. a closed course, would be extrinsically simple since having immediate commonsense meaning. Microscopic examples of this ordering could be described by simple laws, while macroscopic situations might have to be described by complicated deductions from these primary microscopic laws. Thus if the variety and complexity of nature is macroscopic only, then intrinsic and extrinsic simplicity can be combined in a microscopic theory, the variety of nature now being mirrored only in the complicated deductions necessary in the description of a macroscopic situation.

This is in fact the aim of course theory, as may be seen by examination of the relation of the ordering of sense perceptions to the axioms of physical geometry.

In an extrinsically complex theory the axioms of physics do not directly represent the observed ordering of sense-data. Hence the necessity arises of connecting these axioms and the ordering of sense data by building from the sense-data logical constructions obeying the axioms of physical geometry (Nicod, see Russell*), or by relating the assumptions of physics more closely to the ordering of the sense-data. The latter method implies a movement towards extrinsic simplicity, and is followed by course theory. A course calculus representing the topology of events offers an appropriate basis for a theory combining extrinsic and intrinsic simplicity (while macroscopically complex) since it rests on a microscopic formulation of relations given directly to perception in every experiment (the ordering of events implied in a closed course), and does not seek to establish a hierarchy of concepts by successive abstraction. The search for a unitary course theory is an attempt to combine extrinsic and intrinsic simplicity in an essentially microscopic theory in which the coordinate laws valid for systems of all sizes (conservation and variational principles) will appear as complex deductions, valid just in so far as the introduction of four coordinates is empirically valid.

§32. THE LIMITED COMPLEXITY OF NATURE

The atomic hypothesis implies a limit to the micro-structure of nature.

The argument of the last section rested on the assumption, implicit in the methods of analytical science, that beneath the macroscopic complexity of nature there is a microscopic simplicity. Atomic theory postulates that a finite number of microscopic laws can represent phenomena, which means that there exist a finite number of types of primary structural elements. This atomic postulate does not of course involve an assertion that the ultimate units of structure have actually been reached in the atom, or the proton, or the electron, but implies that such ultimate units exist, and that only a finite number are present in any given material system. In some sense this must be correct since otherwise it would be difficult to account for the remarkable success of the atomic hypothesis. Moreover it has recently been suggested (Keynes*, see also Nicod*) that a Principle of Limited Variety bearing a close resemblance to this postulate is indispensable if the process of scientific induction is to be given rational justification.

But if the atomic postulate be accepted, then atomic science, in seeking to describe the macro-complex in terms of the micro-simple revealed by its progressive analysis, must sooner or later reach a limit to micro-analysis, i.e. a point beyond which experimental micro-analysis becomes impossible, and theoretical micro-analysis infertile. No indication is given in the postulate as to the point at which this will occur, but when this critical stage is reached certain things may be expected to happen :

(i) Experimental micro-analysis will appear to be arrested not by technical difficulties but by an absolute theoretical barrier.

(ii) Theoretical micro-analysis after passing the valid limit to micro-structure will yield spurious ultimate units of increasing complexity, in place of the simple ultimate units required by an ideal theory.

(iii) When the final micro-structure of nature has been reached, any simplification or revision of the fundamental laws should result for the first time in the correct theoretical justification of the laws describing the simplest *combinations* of the ultimate units.

These occurrences must be expected, though possibly in a disguised form, and interpreted as hints that further micro-analysis may prove impossible. But apart from these negative inferences, such occurrences will draw attention, when this critical moment arrives, to the necessity of a new descriptive method capable of restoring simplicity to the ultimate structural units of physical theory.

§33. THE EXHAUSTION OF MICRO-ANALYSIS

The evidence suggests that a limit to the micro-structure of nature has been reached, and that the opportunity for a fundamental conceptual revision has therefore arrived.

A situation has just been outlined which has to be expected at some point in the history of micro-analytical science. A considerable body of evidence points to the conclusion that this moment has arrived and that a limit to the micro-structure of nature has been reached. By " nature " is here meant all that can ever be accessible to physical theory and experiment as these terms are at present understood. The evidence is of four kinds, of which the first three correspond to the three phenomena described in the previous section.

(i) *Strict limits are now set to the possible accuracy of the simultaneous measurement of four coordinates.* The successful reinterpretation of physical theory with probabilities replacing exact four-coordinate causation suggests that these limits are not merely temporary, but represent essential barriers to further experimental micro-analysis.

(ii) *Further theoretical micro-analysis has recently resulted in an increased complexity of the primary structural elements of physical theory, which has been compensated by special* *simplifying principles involving an effective reduction in the number of degrees of freedom of atomic systems.*

The spherically symmetrical electron has proved inadequate and it is necessary to regard the electron as an aspect of a complex system of wave-functions which imply that it has an axis of spin in addition to its wave and particle properties. The ultimate structural element of current theory has thus lost the simplicity without which it cannot provide an appropriate basis for theory. The mathematical representation of an electron is now too complex to be used in the exact calculation of complex phenomena and most of the applications of quantum mechanics have therefore to rely on approximate methods. The view that this complexity arises because an inappropriate method of mathematical representation is being applied to nature is supported by the fact that the complicated electron model has to be supplemented by two principles which may be regarded as simplifying the model by retracing steps of successive analysis :

Pauli's Exclusion Principle. This represents a rule of ordering defining the ways in which different electronic motions may co-exist in one system. This principle is so powerful that certain systems containing large numbers of electrons may by its application be described by the same mathematical expression as issued to describe a one-electron system. Moreover under this principle other highly complex systems of definite kinds necessarily have the property of spherical symmetry. The exclusion principle may therefore be regarded as a rule of synthesis, by which the number of degrees of freedom in a system which has been excessively analysed can in certain cases be radically reduced.

Dirac's Linear Wave Equation for the electron shows that the " spin " and the " relativity corrections " which were previously regarded as independent properties of the electron are in fact special aspects of one fundamental characteristic.

Thus Dirac's equation is the first step towards a simplification of the relativistic spin electron; the mathematical complexity of the equation must be regarded as a consequence of the particular mathematical technique used.

(iii) *The theory of the atom as supplemented by these simplifying principles has been able to account adequately for chemical properties hitherto untouched by fundamental physical theory.*

The application of the exclusion principle to the spin electron led immediately to a qualitative quantum-mechanical interpretation of the phenomena of chemical valency, while the exact quantitative development of chemical theory appears only to be held up by mathematical difficulties. The extension of quantum mechanics to the whole field of atomic and molecular physics including the calculation of all the material and chemical constants awaits an improvement in mathematical methods rather than further experimental analysis of the electron and proton. Experimental research is required on combinations of atoms, on nuclear structure and on radioactivity, rather than on the electron and proton in isolation.

(iv) *Evidence arising from the simplicity, stability and completeness of the crystalline forms of matter and from the simplicity of the periodic table.*

The regularity and symmetry of the atomic patterns in the perfect crystalline state suggests that a crystal is built up of ultimate simple units and that these have been reached in the atom or nucleus. If there were further complex micro-structure within the atom, it would be at least improbable that such complexities could remain wholly latent and fail to disturb the regularity and symmetry of the atomic patterns. According to physical theory prior to quantum mechanics any motion within an atom would have complex interactions on the similar movements in its neighbours leading inevitably to instability. The quantum conditions are precisely necessary to prohibit such interactions, i.e. to neutralise or prohibit the (empirically absent) interactions which result from the assumption of a dynamical structure within the atom. Thus the stability and simplicity of the perfect crystalline state provides *prima facie* evidence that the correct theory must not rest on the application of micro-analytical methods beyond the simplest of the atomic units (i.e. hydrogen nuclei) which determine the lattice pattern and crystalline type.

The simple regularity of the atomic numbers and of the periodic table also suggests that further analysis beyond the proton and electron is unnecessary.

This evidence may be summarised by saying that the most novel features of quantum theory, i.e. those having least likeness to classical physics, may be interpreted as resulting from the excessive analysis which is implied in the assumption of analytically independent elements where this independence is not found in nature. Whether or not a limit, either final or provisional, to the micro-structure of nature has in fact been reached, the evidence brought forward shows that the assumption that a final limit has been reached may prove of considerable heuristic value.

Now if this assumption be granted, physical theory has reached a point where a fundamental revision of its methods may not only bring about a great simplification but also provide the basis for a final theoretical system of concepts and methods appropriate to all space-time phenomena. The concepts and methods used in describing the ultimate structure of space-time phenomena are necessarily of importance for all the departments of natural science. Therefore in anticipation of such a possibility a fundamental study of the general conditions which must be satisfied by any comprehensive theory of natural science becomes of importance, whereas at other times such studies may tend to lead physical thought away from its legitimate attention to empirical detail. The assumption that a unitary course theory can provide a comprehensive synthesis implies that finest experimental detail may be reinterpreted in terms of new conceptions developed from arguments of the greatest generality.

§34. GENERAL CONCEPTS OF NATURAL SCIENCE

A revision of physical theory leading to a unitary theory based on the concept of order is of importance for the general conceptual system of natural science.

One of the consequences of a successful revision of physical theory would be the development of more powerful mathematical methods permitting an extension of the application of fundamental physical theory to complex physical and chemical structures. The question therefore arises whether there is a limit to the field of application of fundamental physical theory.

Physics has here been interpreted as the study of the space-time ordering of phenomena, and so defined physics covers in principle the whole field of natural science. If order is the fundamental physical concept, then none of the natural sciences can require an independent set of concepts, i.e. concepts not derivable as forms of order.

It would of course be inappropriate to give the title "physics" to a comprehensive natural science. The important issue is not the naming of the sciences, but the way in which their concepts are inter-related. The possibility of a single comprehensive system of concepts for natural science has often been considered, but attempts to develop such a system have been rightly postponed. So long as fundamental physical theory claimed to restrict itself to metrical relations, or quantities, independent concepts were indispensable to all those sciences which required the wider class of observable relations comprised in topological order.

For example, *irreversible heat process* in statistical physics, *katabolism* and *protoplasmic gradient* in physiology, *function* and *organism* in biology, are concepts which in addition to metrical aspects, involve also the asymmetrical relation of temporal succession, i.e. these concepts are aspects of order which are more general than those of a purely metrical reversible physics. The irreversible heat process was in classical physics described in terms of the probable behaviour of a statistical assembly and in the definition of the probability of a process the non-metrical asymmetrical relation of before and after is involved[10]. This shows that the attempted restriction of physical theory to metrical relations was mistaken.

The situation between physics and biology was therefore as follows. Physical theory was misunderstood, and interpreted as purely metrical, while biological theory was clearly a study of order, including metrical aspects. So long as metric was mistakenly believed to be the only basis for exact unambiguous (i.e. symbolic) description, certain biological concepts appeared indefinite because non-metrical. But in fact both metrical physics and descriptive biology have their origin in the study of order and both need the general relations of order as well as the special relations of metric, though in differing degrees. The apparent disjunction of the concepts of physics and biology was a result of two errors: (i) the belief that physics was purely metrical, and (ii) the view that the precise objective clarity of symbolic mathematical expression could only be applied to elementary numbers.

The proposed reinterpretation of the relations of the fundamental concepts of physics and biology, though of theoretical value, has no relevance to the question whether organisms can be built up out of inanimate structures. Even though organisms are a type of space-time order, and physics must supply the basis conceptual technique for all studies of space-time phenomena, yet it is possible that a type of space-time order exists in organisms which cannot be reduced to those special types of order that physics finds in its elementary units. The available information is still inadequate to permit an answer to this question; the reinterpretation of basic theory here proposed is merely one part of the theoretical technique which will be necessary in the analysis of the observational material when it finally becomes available.

Natural science must on this view take the form of a hierarchy of correlated sciences describing systems of different complexity and based on the general concepts of space-time order defined by physical theory. Lack of clarity as to the basis of physics has hitherto permitted the view that physics is an exact science dealing only with quantity while biology is a descriptive science dealing with forms and functions incapable of exact definition in terms of physical concepts. This distinction is seen to be relative not absolute. The four-coordinate description of physical measurements is necessarily inexact, measurement itself rests on elementary observations of order, and the classification of organic forms on the observation of more complex forms of order. Certain aspects of both inanimate and animate systems can be given exact quantitative treatment, while other aspects of *both* classes of systems cannot be described in terms of quantity alone.

§35. GENERAL PRINCIPLES OF NATURAL SCIENCE

The description of the regularities of space-time order involves, in addition to pure metrical laws describing congruences, descriptive laws describing irreversible sequences.

A complete description of the uniformities of the observed space-time order involves the use of descriptive principles representing irreversible sequences in addition to the purely metrical principles representing congruences. The known metrical principles have already been discussed; the descriptive principles, which must include the comprehensive generalisations of theoretical biology, are still to be discovered, or at least correctly formulated.

When these principles are fully known, natural science must have the following general structure. Its basis must be given by physical principles—probably not expressed in terms of four coordinates—determining the ultimate topological and metrical structure which underlies all space-time phenomena. From this foundation the development of the principles of natural science will proceed in two complementary directions:

(i) *Towards four-coordinate metrical physics.* The underlying structure permits the existence of static rods and reversible cyclic clocks and the approximate establishment of a four-coordinate system. The whole of coordinate metrical theory follows as a consequence.

(ii) *Towards a general descriptive morphology*, or descriptive science of space-time forms. The underlying structure gives rise not only to static rods and clocks but also to regularities of temporal succession, i.e. to irreversible physical sequences which can be described in terms of topological concepts (including the relation of succession) supplemented by the general conception of spatial symmetry. The determination of the necessary temporal order of two states A, B, in an irreversible process may be given by a general descriptive rule neither requiring nor capable of pure quantitative expression. Such rules would form the basis of a descriptive morphology, this term being used in the physical sense of a science of the time sequences of forms having special types of spatial symmetry.

These two classes of principles are closely inter-related, and yet their separation is logically necessary and may prove of considerable value in the general theory of natural science.

It is possible that comprehensive descriptive generalisations exist which have as wide an application as the great quantitative laws of metrical physics. Just as the principle of the conservation of energy may be applied to systems of any type, so certain general descriptive principles may apply to irreversible processes whether elementary or statistical, physical or biological. These descriptive principles may form a systematic theory capable of deduction from the axioms describing the underlying unitary structure and correlated at every point with the metrical laws.

Since these principles would constitute an essential part of the theoretical system resulting from the establishment of a unitary course theory the following examples are given of the general form which they can take.

(i) *Symmetry tends to be established;* more symmetrical states follow on less symmetrical states. (Mayer, Mach*.)

Determined changes in any system imply an initial lack of symmetry, for if there is complete symmetry no factor exists to determine the direction of a possible change. Only differences can initiate determined changes. Now the asymmetry which determines a change must either increase or decrease. If it always increases, symmetrical forms would not only be unstable but would never be produced. But symmetrical forms (e.g. crystals, organic forms) are produced, and in some cases the symmetrical patterns are certainly formed by the disappearance of asymmetries (e.g. the decrease of molecular irregularity in crystallisation). Thus if any general descriptive rule of this type exists, then it must have the form: symmetry tends to be established, or more symmetrical follow on less symmetrical states. This rule has applications to the formation of molecules and crystals, to statistical (heat) processes, and to organic developmental processes.

The definition of the class of systems to which this rule is applicable and its correlation with the reversible processes of classical theory must be postponed.

(ii) The essential features of elementary protoplasmic organisation have to be described largely in topological terms: environment, surface, interior. Other essential concepts are axes of symmetry, and gradients. Thus any rules governing this organisation will take the form of descriptive principles as defined above.

(iii) The combination of general rules of the types (i) and (ii) when applied to special biological problems must lead to deductions describing observed facts, e.g. that in organic developmental processes systems with a given type of symmetry necessarily follow and never precede systems with another given type of symmetry.

A systematic descriptive formulation of the non-metrical features of inanimate and animate processes can have the same logical consistency, lack of ambiguity, and objective validity as metrical laws. Descriptive and metrical methods are both indispensable for every department of natural science. The metrical laws provide the systematic and precise formulation of the detailed facts resulting from the measurement of inanimate and animate nature. The descriptive laws are of the nature of synthetic generalisations which supplement metrical theory at those significant points where its method of approach necessarily fails.

The formulation of a unitary course theory would be of the utmost value in providing an adequately permanent system of universal physical concepts and laws as the basis for a new attack on the problems of biological theory. The evidence provided by the quantum theory that a limit to the micro-structure of nature has been reached, and the present concentration of research on biological problems have prepared the ground for an advance in biological theory. In the epoch of classical physics it was conceivable that physical laws did not provide any clue to the correct form for the biological laws. This situation would be radically altered by the establishment of a unitary theory. A theoretical system covering the whole field of fundamental physical law from gravitation to radiation and chemical action must be expected to bear some relation to the laws of life. If the laws of life were independent of the physical laws life could neither exist within the physical universe nor discover its laws[14].

§36. LINES FOR RESEARCH

The argument leads to a programme of research.

The analysis of coordinate theories and of the general features of a unitary course theory has now been carried as far as is possible in the present state of physical science. It has however revealed a number of lines along which theoretical research may advance towards the aim of a unitary course theory. The possibility of using some of these methods has long been recognised and it is probable that many of the points raised are well known to individual workers. In critical epochs in the history of science speculative anticipations of a new method often arise independently in many minds. Yet the possibility of real advance is only given when current thought has developed far enough to permit such speculations to be converted into a precise programme. There are many examples in the history of mathematics and of physical theory where the formulation of a definite problem was of crucial importance in focussing the endeavours of a school of thought into a sustained and successful attack. The general outlines of the heuristic method here proposed could have been recognised at any time in the last twenty years, but only since the completion of the general notation of quantum mechanics have the fundamental problems which these methods are designed to attack become urgent. Since 1927 the chief need in physical theory has been a re-examination of the fundamental concepts in order to provide a proper correlation of the methods of relativity and of quantum theory. The explicit formulation of a new programme for fundamental synthetic research is moreover indispensable if the ever growing range of detailed fact is to be adequately controlled by well founded general principles capable of being understood by the ordinary scientific worker. Just today when physical thought is fascinated by the great successes of relativity and quantum theory is it necessary to avoid excessive preoccupation with methods whose productive period in connection with fundamental problems is probably over and to attract attention to new fields of research.

A number of lines for research which bear on the development of a unitary course theory are collected in the following list with references to the section in which they are discussed. The different methods are closely related to one another, but their separate formulation opens up in each case a precise problem which can be developed by known mathematical or symbolic methods. It will be convenient later to refer to this scheme for research as the 1930 Programme. It is proposed to publish, after an interval of an adequate number of years, a survey of any contributions which have been made to physical theory along the lines of this programme.

1. *Topological demand.* The transformation of the four-dimensional metric to a new type of metric permitting the representation of the primary topological relation of coincidence. (§11.)

2. *Elimination of time as an independent variable and of arbitrary initial conditions.* The development of a form of quantitative law which

   (i) does not use t either as an extended coordinate or as an independently measurable local time, but instead is based on a structural theory of clocks, and

(ii) does not permit arbitrary initial conditions, but traces every system back to a standard initial state such as a coincidence. (§§9, 13.)

3. *Unitary physical structure underlying the dual "wave-particle" and abstract quantum algebra.* The search for a unitary structure appropriate to a law of the kind outlined in 2. and such that when described in terms of four coordinates it will yield a physical interpretation of the abstract algebra of quantum mechanics and appear in simple cases as a dual wave-particle model. (§§19, 21.)

4. *Elimination of arbitrary units.* The development of a quantitative law not dependent on the use of arbitrary units of length and time, but instead expressing the direct comparison of elementary physical structures of definite scale, e.g. by combining in pairs the quantum mechanical laws referring to systems of different types. (§16.)

5. *Origin of imaginary quantities.* The analysis of the parts played by linear and quadratic forms and real and imaginary quantities in coordinate theory, with a view to the interpretation of the appearance of $i.c$ and $i.h$ as the analytical consequence of substituting an independent variable t for the derived time parameter given in the underlying unitary structure.

6. *Physical re-interpretation of electrons, protons, radiation, gravitation.* The exact analysis of the space-time structures involved in all the methods by which the six universal dimensional constants are determined, and hence of the possible forms which the derivation of α, β, γ can take in unitary course theory, on different physical assumptions as to the unitary structure

underlying electrons, protons, radiation, gravitation. (§§22, 29.)

7. *General mathematical and logical analysis of natural science.* A rigorous analysis by the methods of mathematics and symbolic logic of the quantities and types of relation-structure required in the deduction of the quantitative laws of physics and the descriptive laws of natural science from unitary axioms of space-time order. (§§28, 29, 34, 35.)

§37. CONCLUSION

The Critique to which we have submitted current physical theory has led to the following conclusions:

(i) That physical science is ripe for the formulation of a comprehensive unitary theory.

(ii) That this is so primarily because experimental and theoretical research have reached a limit to the micro-structure of nature.

(iii) That a unitary theory must rest on a new type of metric implying definite restrictions on the validity of a four-coordinate metric.

(iv) That electrons, protons, radiation, and gravitation are aspects of the four-coordinate description of a unitary phenomenon describable in the new metric.

(v) That the new metric is to be discovered by a closer attention to topological facts, and will eliminate the use of dimensionally independent quantities.

(vi) That this emphasis on topological fact can be used to provide a physical interpretation of the abstract algebraic axioms of quantum theory.

(vii) That a physical theory based on the concept of order can supply the foundations of a conceptual system for a comprehensive natural science.

And finally,

(viii) That the publication of a definite programme for research towards a unitary course theory may stimulate the development of these methods.

NOTES

Note. 1. *Methods of fundamental research already in use.* (§1.)

The following are the chief methods which have been applied towards the formulation of a wider theoretical synthesis :—

(*a*) The development of new geometrical field laws in the relativistic space-time continuum comprising electromagnetism and gravitation (Eddington. Weyl. Einstein 1928-9).

(*b*) The five-dimensional representation of electromagnetism and gravitation (Kaluza, Klein).

These two methods start from the methods of relativistic macroscopic field theory and do not appear to take adequate account of quantum phenomena. Einstein believes that in spite of this a macroscopic field synthesis should come first.

(*c*) The development of relativistic quantum theory, either as a quantised electrodynamic field theory (Heisenberg-Pauli) or for special systems as a relativistic wave-mechanics (Dirac). The first method is of great difficulty, while the second does not conform to the general notation of the transformation theory of quantum mechanics. Dirac's successful expression for the electron has however proved the value of this approach.

(*d*) Special attempts to extend quantum mechanics so as to obtain theoretical derivations of one or more of the pure number ratios. (Eddington, α, 1928; proposal by Dirac for β, 1929.) This line of attack has not yet achieved any generally accepted success.

(*e*) Methods aiming at a formal generalisation of quantum mechanics capable of comprising gravitation. (Weyl and others.) No methods of this kind have yet been developed providing any constructive advance beyond the present independent theories of quantum mechanics and relativistic gravitation.

All these methods rest ultimately on the four-coordinate description of space-time phenomena. Though the abstract axioms of quantum mechanics are not representable as relations between ordinary coordinate quantities, yet the empirical consequences of this theory are expressed in terms of the four space-time coordinates.

A preliminary outline of the methods proposed in the Critique has been published in Zeit. f. Phys. **56**, 809, 1929; **61**, 274, 1930; and a third paper in 1931.

Note 2. *Similarity Principles.* (§4.)

Many writers from Galileo and Newton onwards have developed and applied principles of similarity. The most important references are Galileo, *Conversations* (at the beginning of the First, and in the middle of the Second Day); Newton, *Principia*, Book II, Section 7. Prop. 32; Bernouilli (Joh), Opera 1, p. 514; Bertrand, *Comptes Rendus*, 1847, **25**, p. 163; Mach, *Mechanics*, p. 165; Stäckel, *Journal f. Math.*, 1891, **107**, p. 328. An interesting example of the application of a principle of dynamic similarity is given by Larmor's derivation of the Lorentz transformations (*Æther and Matter*, p. 176). Out of many modern papers on the subject the best analytical treatment is that of London *Phys. Zeit.* **23**, pp. 262, 289, 1922, though the crucial assumption (contained in equation (9), p. 292) is invalid for the reasons discussed in the text. Interesting discussions of the wider problem of general physical similarity are given by Campbell, *Physics, The Elements*, p. 409, and in Bridgman's *Dimensional Analysis*. These treatments have now only historical and practical interest as they are inconsistent with the facts underlying relativity and quantum theory. Byk (*Handbuch der Physik*, vol. IX, p. 330) has pointed out in a discussion of general physical similarity that the quantum constant eliminates classical similarity.

Compare also Weyl, *Philosophie der Mathematik und Naturwissenschaft*, 1927, p. 102.

Note 3. *Relativistic Orbits.* (§5.)

The electron orbits of the special theory and the planetary orbits of the general theory (in the case of negligible planetary mass) are both rotating ellipses. The geometrical form of the orbits of both types is different for orbits of different size, i.e., the perihelion advance in revolutions per revolution (δ) is a function of the size of the orbit. Since δ is a pure number, it must be a function of the ratio of some constant length to the size of the orbit. Neglecting terms in δ^2, and taking nearly spherical orbits, then

for electron orbits

$$\delta = \tfrac{1}{2}.\ (\textit{Universal length } \frac{e^2}{mc^2}) \div (\textit{radius of orbit})$$

while for planetary gravitational orbits

$$\delta = 3.\ (\textit{Solar length } \frac{G\mu}{c^2}) \div (\textit{radius of orbit}).$$

Compare: Eddington, *Math. Th. of Rel.*, p. 88.
Sommerfeld, *Atomic Structure*, 1923, p. 468.

This analogy between electron and gravitational orbits has no immediate meaning since the perhelion rotation has a different origin in the two cases and is six times as great in the general theory as the corresponding value in the special theory.

It is easy to see that classical kinematic similarity must disappear in the general theory or in any covariant theory of a geometrical field. Covariant field equations in which the field quantities are functions of the $g_{\mu\nu}$ and their derivatives cannot describe H-sets since every spherically or axially symmetrical solution involves at least one integration constant having the dimensions of a length. Viewed analytically, the $g_{\mu\nu}$, being dimensionless functions of r, must have the form $f\left(\frac{m'}{r}\right)$ where m' is a length. The same may be seen from geometrical considerations since no similar figures can exist in a non-Euclidean space, the geometry of a figure depending on the ratio of its size to the components of curvature at each point.

Note 4. *Planck's constant h.* (§6.)

In view of the importance of h in current theory it may be convenient to restate—at the cost of some repetition—the argument of the text.

The failure of fundamental dynamical theory to take account of the fact that any structural theory must contain dimensional constants defining the linear scale of the structure described, resulted—as was seen in §3—in Planck's constant being regarded as more remarkable than the other constants and Planck's and Bohr's postulates as peculiarly arbitrary. The problem of the meaning of h was created by a mental inertia which continued to regard classical laws as fundamental. But there was nothing strange in the appearance of h, a new constant being necessary to any theory of the structure of matter or of radiation. Larmor (*Æther and Matter*, p. 189) came near to predicting the appearance of such a constant, for he realised that the equations of gravitational and æther theory could not be adequate for a structural theory. The peculiar feature is not the appearance of h but the success of classical dynamics and field theory in spite of their neglect of that definiteness of scale which is now seen to be essential to any complete physical theory.

The fact that the new constant was given the dimensions of *action* resulted from the historical circumstances of its discovery; had the history of physics been slightly different the new constant might well have been discovered in a manner resulting in it being given the dimensions of $\frac{h}{m}$, i.e., *Area* ÷ *Time*. A constant of this form appears in certain fundamental equations, and would have served equally well. Or again, had Larmor's argument been developed further a universal length might have been introduced into physical theory before the constant h had been found necessary. These are the alternatives that were open to the chances of history; the arbitrary fact that one particular alternative occurred must not conceal the necessity common to them all: that a structural theory dealing with patterns of definite scale required a new constant capable with the others of yielding one or more universal lengths appropriate to stationary or static systems, the particular dimensions of the constant being indifferent provided this condition was satisfied.

Planck noted the fact that h permitted the determination of natural units of length, time and mass (*Berliner Sitzungsber*, p. 465, 1899). But he expressed these in forms of no theoretical value since not required in describing natural structures, and he later discarded them.

In introducing his theory of the hydrogen line spectrum, Bohr also emphasised the fact that a length appropriate to the description of the size of atoms does not appear amongst the quantities characterising the Rutherford atom (e,m,M) whereas when Planck's constant is introduced it becomes possible to determine a length of the order of magnitude required. (*Phil. Mag.*, 26. p. 2, 1913.) It is important to notice that a length nearly of the right order of magnitude was already available ($\frac{e^2}{mc^2}$) but that this was not the one required in describing the principal lines of the hydrogen spectrum. Any theoretical derivation of the fine-structure constant has precisely to account for this duality: the existence of *two* characteristic electronic lengths for example $\frac{e^2}{mc^2}$ and $\frac{h^2}{4\pi^2me^2}$.

Note 5. *Comparative analysis of the methods of relativity theory and the proposed methods of course theory.* (§8.)

The points of divergence of the proposed methods of course theory from those of relativity theory are shown in the following analysis of Einstein's two classic papers introducing the special and general theories: (i) *Annalen der Physik*, 17. 1905 (special theory) and (ii) *Annalen der Physik*, 49. 1916 (general theory). The references and extracts are from the English translation in *The Principle of Relativity* (Methuen, 1923). C. stands for the proposals of the Critique.

(i) *Special Theory.*

p. 38. "The theory to be developed is based on the kinematics of the rigid body."

C. A "rigid body" can no longer remain a primary concept, since a theory of its structure has now to be given.

p. 39. "It might appear possible to overcome all the difficulties attending the definition of 'time' by substituting 'the position of the small hand of my watch' for 'time.' And in fact such a definition is satisfactory when we are concerned with defining a time exclusively for the place where the watch is located; but it is no longer satisfactory when we have to connect in time series events occuring at different places, or—what comes to the same thing—to evaluate the times of events occuring at places remote from the watch."

C. The time correlation by means of a simultaneity concept of events at distant places is asserted to be unnecessary and the text down to "located" expresses the proposed method. Simultaneity is to be restricted to (nearly) coincident events.

pp. 41-42. (Arguments on the relativity of lengths and times.)

C. The concept of velocity is no longer permissible, and hence the conception "the length of the rod in the moving system" as defined in the text cannot be used.

(ii) *General Theory.*

p. 111 onwards. (Arguments re coordinate systems.)
C. Systems of four coordinates are not to be assumed at the start, it being necessary to consider how the complex structures called rods and clocks are used to determine coordinates.

p. 117. "The general laws of nature are to be expressed by generally covariant equations."
C. The demand for covariance is not relevant to a theory which does not use four coordinates.

p. 117. "The introduction of a system of reference serves no other purpose than to facilitate the description of the totality of space-time coincidences."
C. Hence another descriptive method may be used if it can fulfil this purpose.

p. 118. "As all our physical experience can be reduced to such coincidences—— "
C. Provided the ordering of the coincidences is retained, and attention is paid to the question of the exactitude of the coincidences, and (possibly later in the theory) to their observed colour.

p. 118. "The path we have entered upon is psychologically the natural one."
C. Or was until quantum theory provided structural descriptions of rods, and revealed that the methods of relativistic field theory could only be approximate.

p. 119. "Two infinitely proximate point-events."
C. Infinitely proximate point events cannot be given invariant metrical definition since $ds = O$ may define also a light ray of any length.

p. 159. "Equations (67) and (68) together are equivalent to Newton's Law of Gravitation."
C. Those two equations from relativity theory are only equivalent to Newton's Law provided it is *assumed* that an arbitrary numerical factor k (proportional to the gravitational constant) is positive. Relativity theory does not exclude repulsive gravitation.

The prestige of the theory of relativity rests on its empirical success, logical consistency, and elimination of empirically meaningless classical problems. It is well known that it does not represent the only possible kind of space-time theory. (Compare quotation from Einstein on page xi.)

There can be no conflict between an exact microscopic structural theory (e.g., course theory) and an approximate macroscopic field theory neglecting structure such as relativity theory. Course theory will have to show that in so far as four coordinates can be legitimately used the theory of relativity and its consequences are necessarily correct, but that a set of problems exists for which the relativistic concepts and calculus are inadequate. The assumption is made here that further observations will continue to confirm the predictions of the general theory.

Note 6. *Physical Topology and Order.* (§10.)

It is necessary to define and explain the meaning for physical theory of the mathematical terms " topology " and " order."

Abstract Topology is the extension of geometry which studies those properties of geometrical structures which remain unchanged under continuous transformations, i.e., the study of pure position and connection relationships apart from metrical relations. Important concepts of pure topology are *neighbourhoods, continuity, connectivity, ordered sets, dimensionality*. The subject matter of *Analysis situs*, and of the *Theory of Sets* overlaps that of topology, though the different names imply different methods of handling similar material. Important problems of topology are the definition of the dimensions of a space, the properties of closed, open and interlacing curves, the connectivity of surfaces, etc.

Physical Topology deals with the qualitative, i.e., the non-quantitative, ordering of physical phenomena, and cannot be defined in a parallel manner to abstract topology since the latter does not contain the physically indispensable relation of *before* or *after*. Physical topology is at once a generalisation and application of abstract topology, and the *topological ordering* of a set of physical events is most conveniently defined as the group of relations between the events describable in terms of *near, between*, and *before* (or *after*). These terms, in their physical sense, are to apply to the finite patterns of events which form the immediate empirical basis of physical theory and do not refer to hypothetical infinitesimal patterns of points. The definitions given here are provisional, and serve merely to allow a preliminary approach to physical topology.

Physical or Applied Topology has only recently begun to be recognised as a necessary branch of physical theory. The three-dimensionality of space, the single series of temporal succession, the finite unlimited world of Einstein's cosmological theory, are examples of topological forms, but no comprehensive theory of physical topology has yet been developed in connection with these problems. This is because the dimensionality of space and of time appear to be ultimate facts resisting further analysis. Physical topology has so far been approached from two rather different standpoints: through the theory of coordinates, and through the causal structure of phenomena. Eddington (*Math. Th. of Rel.*, p. 225, quoted on page xi) refers to the topological assumptions regarding the ordering of events implicit in the use of coordinates and not theoretically justifiable by the postulation of the relation of interval. Russell (*Analysis of Matter*) emphasises the importance of topology and proposes a physical theory dispensing with coordinates. Reichenbach (*Philosophie der Raum-Zeit-Lehre*) discusses physical topology in relation to the general theory of relativity and the causal structure of phenomena. But no constructive results have been reached and the subject awaits a thorough analysis.

The term *order* will also be used here in a special physical sense. In pure mathematics order is an abstract (logical) property of series generated by any asymmetrical transitive and connected two-termed relation. In physical theory order is most conveniently interpreted rather differently. Physical order is constituted by any topological relations between physical elements (usually events), i.e., it includes spatial and temporal order, and the exact properties of physical order are a matter for study, and not for mathematical or logical definition.

In the physical sense here used the directly observable relations between events comprised in the concept of order are a more general class than the observable relations underlying the concept of metric. In this special sense order is more general than metric.

It is interesting to note that topological ideas have had great influence on the development of physical theory. For example Faraday's mode of viewing electro-magnetic phenomena was based largely on topological pictures of interlacing lines and tubes of force, and Maxwell's achievement was to provide a quantitative reconciliation of these pictures with the abstract mathematical methods then current on the continent. The greater part of mathematical physics is concerned with the quantitative treatment of problems which are defined in terms of some special topological structure, such as a closed circuit or a closed surface.

Note 7. *The conventional element in Physical Topology and Geometry.* (§11.)

It is not possible to give a definite answer to the question "Is the topology of physical events correctly represented by the use of four coordinates?" because the term "physical events" is ambiguous. It can mean the set of actually observed events or the field of all possible events. Only in the latter sense is the question of interest, and in this case it is to be answered affirmatively by definition, for the field of possible events is in all current physical theory *defined* as a four-dimensional continuum. If the current convention is retained that physical law is to select actual events from a four-dimensional field of possible events, the use of four coordinates follows necessarily.

The question, " Is physical space-time non-Euclidean ? " has the same character, for the term " physical space-time " is ambiguous until a relation between its metric and the laws of physical bodies is given. Physical geometry plus physical law determines experience ; a geometry may be chosen arbitrarily (within certain limits) provided the laws are adjusted appropriately. This has been made clear by Poincaré (*Monist*, 1898 and elsewhere), Einstein (*Geometry and Experience*, in *Sidelights on Relativity*, p. 35) and Russell (*Analysis of Matter*, pp. 78-80). At certain stages in physical thought, as for instance in the general theory of relativity, it proves convenient to separate geometry and physical law by a suitable convention which then permits a definite answer to the above question.

This analogy can be used to clarify the problems of physical topology. It is only the assumed physical topology plus the general form of physical laws which determines the order of observed phenomena. Coordinate theory starts with the useful convention of a four-dimensional background, and this is permissible as long as coordinate laws are adequate. Another type of physical theory may make some other topological convention, but then it must show how the use of four coordinates has provided an adequate basis for approximate theory.

Note 8. $i=\sqrt{-1}$. (§21.)

In view of the appearance of i in the fundamental axioms both of relativity and of quantum theory when these are expressed in forms bearing the closest analogy to classical expressions, it may be useful to give a few facts about its history.

In pure mathematics. The use of imaginary and complex quantities was developed by Cardan, de Moivre, Cotes, Euler, Argand, Gauss, Cauchy and others. In 1714 Cotes derived the relation $i.x=log\ (cos\ x\ +i.sin\ x)$ (in another notation) and Euler subsequently introduced the symbol i and the use of the imaginary exponental e^{ix}.

In applied mathematics. (i) Periodic phenomena. In 1850 Kelvin used e^{ix} to simplify the treatment of Fourier series describing physical phenomena, and Rayleigh employed it in 1877 in his " Theory of Sound." In 1893 Steinmetz and Kennelly transformed the theory of electrical engineering by applying e^{it} to alternating current problems. From 1919 onwards it has provided the natural notation for the periodic phenomena generally assumed to underly quantum processes.

(ii) Non-periodic phenomena. Hamilton generalised the imaginary unit and discovered the quaternion units in 1843. This notation can be applied to various empirical problems but is inconvenient. (The Pauli-Dirac spin matrices are closely related to the quaternion units.) The imaginary time-coordinate in the form *ict*, was applied in relativity theory by Poincaré in 1905 and by Minkowski in 1907 to permit a symmetrical formulation of the four-dimensional metric.

(iii) It is not yet certain whether the use of i in Heisenberg's multiplication rule for non-cummutating variables (1925) is to be classed under (i) or (ii), i.e. whether the presence of i in the axioms of quantum theory means that *all* physical phenomena are periodic, or whether its presence is due as in the Minkowski relativity notation to a non-periodic situation, and merely permits this situation under certain circumstances to be represented by the concepts and mathematics of wave or other periodic motions.

Note 9. *The Six Primary Universal Dimensional Constants.* (§23.)

	First Determination.		Best value (1930).		
c	Römer ..	1676	2.99796	±.002	$.10^{+10} cm.sec^{-1}$
G	Newton (estimate) Maskelyne, Cavendish	1684	6.664	±.002	$.10^{-8} cm^3 sec^{-2} gm^{-1}$
M	Estimated from	1860	1.6608	±.0017	$.10^{-24} gm$
e	Johnstone-Stoney Townsend	1874 1897	4.770	±.005	$.10^{-10} (cm^3 sec^{-2} gm)^{\frac{1}{2}}$
m	Thomson, Wiechert	1897	9.035	±.010	$.10^{-28} gm$
h	Planck ..	1899	6.547	±.010	$.10^{-27} cm^2 sec^{-1}$gm

These values yield

$$\alpha = 7.283\ (\pm .006)\ .10^{-3} \text{ or } \frac{1}{137.29}$$

$$\beta = 1838.\ (\pm 1.)$$

$$\gamma = 1.24\quad .10^{36}.$$

(Sources: Birge, *Phys. Rev.* Suppl., July, 1929; Millikan, *Phys. Rev.*, May, 1930. The earlier deflection values for e/m

have been excluded in view of Perry and Chaffee, *Phys. Rev.*, 36, p. 904, 1930; and Kirchner, *Phys. Zeit.*, p. 1073, 1930.)

These six constants may be taken as epitomising the whole of fundamental physical theory. G, c, e, and h refer directly to Newton's Law, Maxwell's Equations and the Theory of Relativity, Coulomb's Law, and the Quantum Conditions respectively, while m and M represent the inertial units involved in all material systems. An explicit statement of the meaning of the six constants would comprise all the empirically indispensable parts of fundamental physical theory.

Note 10. *Irreversibility in Statistical Theory.* (§24.)

The place of irreversible heat processes in classical theory is of special interest when viewed from the point of view of logical form, or more exactly relation-number (see Note 13). If the classical laws are reversible, depend only on t squared, and contain no asymmetrical transitive relationship which can provide a description of *succession*, how can the result be deduced that systems tend to pass from less probable to more probable states? At what point in the argument is the asymmetrical relation of succession introduced unto the argument, and hence into physical theory?

The answer, which is valid both for the classical and for the quantum-mechanical derivation of the entropy theorem, is that the distinction between earlier and later moments is implicit in the conception of the "*probability of a process, given the statistical description of the initial state.*" The collision (or transition) probabilities define the probability of one state changing to a *subsequent* different state, and the formulation of the meaning of "the probability of a process" in the notation of symbolic logic involves the use of a symbol meaning "is later than." An abstract intelligence robbed of the knowledge of succession and confronted with a record of physical phenomena not involving the relation of earlier and later could not use the ordinary concept of the probability of a process. The "probability of a process" is not a pure metrical concept; it rests on the wider class of relations of physical order, since it involves succession as well as metric.

This important fact about process-probabilities is discussed by Weyl, *Gruppentheorie und Quantenmechanik*, p. 95. Cf. also Jordan, *Die Lichtquantenhypothese*, in *Ergebnisse der exakten Naturwissenschaften*, No. 7, 1928, p. 168.

It is often interesting to discover at what point in a mathematical treatment of statistical theory a complete symbolism would involve the introduction of a symbol with the meaning "is later than," since the t^2 of classical mechanics need not be associated with this meaning. In Fowler's *Statistical Mechanics* the time differential $\frac{df_1}{dt}$, equation (1183), page 424, introduces the asymmetrical relation of succession into the argument by which the entropy theorem is derived from quantum statistics. A strict treatment should use a symbolism which makes clear the point at which new mathematico-logical relations of this kind are introduced into a deductive calculation.

Note 11. *Eddington's derivation of* α. (§28.)

The best experimental values for the constants involved in α have been analysed by Millikan (*Phys. Rev.*, May, 1930) who obtains the value 137.29 for $\frac{1}{\alpha}$, and rejects Eddington's proposed value 137. Eddington's argument has not been generally accepted as theoretically legitimate, apart from the question of the correctness of the predicted value for α. The principal reasons for this are the difficulty of the argument, the fact that the treatment is not invariant but assumes stationary electrons, and the failure to throw any light on the interactions of electrons and radiation, which might reasonably be expected of a significant derivation of α. These objections do not however lessen the interest of the general principle involved in this pioneer attempt at a derivation of α by the identification of two forms of electron interaction, the Coulomb law and the exclusion principle.

Note 12. *The Axiomatic Method.* (§29.)

The strict deductive, or *axiomatic*, formulation of a physical theory is an ideal which has seldom been attained. But attempts at an axiomatic formulation lead to the clarification of ideas and the simplification or generalisation of mathematical methods, and are of the greatest value in studying the interrelations of different theories.

The emphasis on unification and precision which characterises the axiomatic or mathematical mind is the necessary complement to the experimental worker's respect for fact. Both are indispensable elements in an exact science, though they lead to somewhat differing interpretations of physical law. The axiomatic mind sees in physical laws concealed definitions by which one quantity is reduced to identity with quantities which are only apparently different, while the experimenter regards laws as discovered relations between quantities which are considered different because measured independently.

A branch of theory which has reached a definite pause in its development tends naturally to approach the ideal deductive form. Yet the axiomatic form does not necessarily represent the ultimate ideal for a complete science, since the axioms of one science may in each case be the complex deductions of a neighbouring science. The highest possible aim for science is the formulation of a self-consistent closed chain of concepts and principles, permitting deductive argument at one direction at every point of the chain. (See Note 14.)

Note 13. *The Structure of Relations, or Relation-number.* (§29.)

The conception of the formal structure of a system of relations, called its *relation-number*, is of great importance and was introduced by Bertrand Russell in *Principia Mathematica*, Vol. II. Relations have certain formal or logical characteristics, e.g., they are:

(i) two-termed, three-termed, or many-termed; if two-termed, then

(ii) symmetrical, asymmetrical, or non-symmetrical;

(iii) transitive, or intransitive; etc.

Similar relations, or complex patterns of relations, are those having the same formal characteristics, i.e., having the same logical structure. The structure, or *relation-number*, of a group of similar relations (or systems of relations) is the formal logical structure common to them all. For example, the structure or relation number of the following group of similar relations:

(*a*) Event A is earlier than event B,

(*b*) a is an ancestor of b,

(*c*) α is larger than β,

is that of a two-termed asymmetrical transitive relation. Relation-number is the abstract logical pattern implicit in a given concrete relation or system of relations, and whenever two relations have the same relation-number, all their logical and mathematical properties are identical.

The importance of the structure of relations is that "it is only structure that we can validly infer from perceptions, and structure is what can be expressed by mathematical logic, which includes mathematics." (Russell, *Analysis of Matter*, p. 254.) The valid content of a scientific proposition, i.e., its legitimate meaning, is given by its structure. Structure is the real essence of scientific knowledge; it is, in fact, the meaning of all exact language. To different types of relations, or systems of relations, we give the names space, time, between, permanence, matter, and so on. The actual names given are of no importance, so long as we are agreed about the formal properties of the situations they represent. Scientific concepts are labels for different types of experienced relation-structure. For example, space and time are simply names for experienced relationships of certain definite logical types, and waves and particles names for more complex types of relations. These relations are all aspects of the (space-time) order given in experience. (The above discussion of the structure of relations is based directly on Russell.)

Science seeks to discover and formulate thestructure (relation-number) of experienced order. This implies a search for simple formal properties within the apparent complexity of phenomena. But these simple formal properties, i.e., the structure of physical law, need not be abstract in the sense of without immediate physical meaning. In fact two of the most general and simple types of experienced relationships are precisely what is meant by "space" and "time." Thus structure may be abstract in the sense of completely expressible in a formal symbolism, and yet the opposite of abstract in another sense, since having immediate physical significance.

An example will make this clear. The higher algebra of quantum mechanics is now abstract in both senses, it is expressed symbolically and it has no direct physical significance. What is desired is not the substitution of words for the symbolism, but the transformation of the equations so that they have direct physical meaning, i.e. so that they express types of symmetry, asymmetry, etc. corresponding to those given in elementary physical perceptions. The abstract algebra will then be seen to have physical significance. Now just as ordinary numbers express metrical properties, so higher numbers (complex numbers, matrices, etc.) of the kind which appear in this quantum algebra may express non-metrical topological relations as well as metrical relations. Elementary and higher algebra are the appropriate methods for expressing the properties respectively of the ordinary numbers and the relation-numbers involved in metrical geometry and in topology.

The discovery of the correct interpretation of the algebra of quantum mechanics in terms of physical topology is delayed chiefly by the plurality of meanings of the terms of space, time, etc. Only when the metrical and topological relations which are combined in these terms have been separated, can the immediate physical significance, if any, of quantum algebra be discovered.

An exact logical analysis of the structures implied on the one hand in four-coordinate space-time and on the other hand in quantum algebra, is the only available method of testing the undiscussed implicit assumptions of coordinate physics. This analysis can only escape the use of concepts which combine several meanings if based on the conception of relation-number. It is probable that a development of the Russellian concept of relation-number will provide for the topological basis of physics, and also for descriptive natural science (see §35), what the Newtonian differential calculus was for metrical physics, a notation adequate to stimulate the discovery and permit the exact formulation of new laws. Relation-number must supplement number in the expression of the uniformity of nature.

The most important references on Structure and Relation-number are:

Russell, *Principles of Mathematics*, p. 262.

Russell and Whitehead. *Principia Mathematica*, Vol. II, p. 303.

Russell, *Introduction to Mathematical Philosophy*, chaps. 5 and 6.

Russell, *Analysis of Matter*, chap. 24. (This book provides a very thorough analysis of the foundations of physical theory.)

A concise survey of mathematical or symbolic logic with some applications to topology, etc., is given by Carnap, *Abriss der Logistik*. (Springer 1929.)

Note 14. *The Circle of Structure.* (§35.)

In §34 the concepts of natural science are regarded as representing different types of space-time order. The fundamental or elementary types of order are defined by physics, while the primary concepts of the other natural sciences represent more complex types of order. On this view the conceptual structure of natural science is that of *an open series of concepts*, passing from the elementary concepts of physical theory (dimensionality, succession, congruence, etc.) to the complex concepts of biology (organisms, etc.). The possibility of arranging all the concepts of natural science in one main series with side branches, i.e. in a hierarchy, does not necessarily imply that the more complex (organic) forms of order are reducible to combinations of the special types of order found in atomic theory; it only implies that an organism is one type of space-time order, i.e. it will fall within the general scheme of concepts of a theory dealing with order.

It is the task of a more comprehensive science to link the two ends of the main series of concepts, i.e. to bring into relation with one another the primary concept: *order*, and the most complex: *man*. This can be done by showing that order, with which natural science has to begin, is a representation of the structure of thought. Thought is the organisation of the delayed reactions of *homo sapiens* to the influences of the environment, and this organisation is based on similarities in the structure of experienced situations. Order is a general term for all situations with definite recognisable structure. Thus the open series of concepts can be closed; order, while primary to natural science, is reached again in human biology as the general form of man's organisation of experience. The closed series of concepts may be called the *circle of structure*, since it links the structure of "matter" with the structure of "thought." The closing of the circle is possible because the structure of matter is necessarily also a structure which can be thought about, and because the kind of structure which can be thought about is determined by the structure of the physical observing system. Symbolic logic and mathematics provide the notation for the circle of structure, its subject matter being the structure or relation-number of experience.

The circle of structure may be divided into several links. Each link represents a science with its own concepts and principles. If we select physics, biology, and psychology we have three sciences each studying structure. The structure is often regarded in these three sciences as the structure of something, e.g. the structure of matter, the structure of life, and the structure of thought or experience. These are convenient names, for certain purposes, but science is concerned solely with correlations of relation-structure, strictly of relation-number (13). Science can never say that either consciousness, or matter, or life, has priority, or that any of the sciences can rightly dominate the others. When emphasis is placed on different links in the circle different important vistas are opened up and we view existence in terms of the laws of matter, life, or mind. Moreover the selection of these links is arbitrary, the circle can, for example, be divided into two links: the structure of space-time phenomena and the structure of mind.

The circle of structure suggests what must be regarded as the ultimate aim of science, i.e. the establishment of a circle of *principles*, so that the axioms of each science follow from the conclusions of the preceding science. The laws of physics would then be the necessary consequences of the manner in which man had set about systematising his experience, the laws of biology would follow from those of physics, and the laws of thought from those of biology. At present it is not known (i) if the axioms of physics are such that they could not have been thought about by man if they were otherwise, (ii) if organic processes are special cases of a more general class of physical process, and (iii) if the structure of mind can be regarded as a consequence of the structure of brain. These are still matters for research. But the establishment of a circle of structural principles represents the highest possible unification of experience and therefore provides a standard to which every comprehensive scientific theory can be compared.

If this aim were ever reached, the honours of success would be equally with reason and intuition. The circle of structure can only be built by intuition using reason as its instrument. Intuition is the recognition of the similarity of relation-structure in two situations, reason the tracing of the consequences of this similarity. Intuition and reason operate on relation-numbers; the use of concepts is only a provisional aid which ultimately tends to confuse the operation (i) of the non-conceptual or pure intuition in recognising the structure of situations, and (ii) of the non-conceptual or pure reason in tracing logical consequences. But the use of reason in science itself rests on a rationally unjustifiable principle of induction, i.e. on an intuitive assumption of a persisting uniformity in nature. By the closing of the circle of structure this intuition is seen as a necessary form governing the reactions of man to his environment. The intellectual problem presented by the actual success of the inductive method then loses its peculiar urgency and becomes part of the inescapable mystery of the whole interrelated circle of experience being as it is.

REFERENCES

Page xi	Einstein.	*Geometrie und Erfahrung* (1921); translated in *Sidelights on Relativity* (1922), p. 35
	Eddington.	*Mathematical Theory of Relativity*. 2nd ed., p. 225
	Heisenberg.	*Physical Principles of the Quantum Theory* (1930), p. 104. Quotation translated from German edition
§3	Larmor.	*Æther and Matter* (1900), p. 189
§4	Newton.	*Principia*, vol. I, §9. Cf. Whittaker, *Analytical Dynamics*. (Theorem of revolving orbits), p. 83
§6	Heitler and London.	*Zeit. f. Physik*, **44**, 455 (1928)
§8	Eddington.	*Math. Th. of Rel.*, p. 34
§9	Reichenbach.	*Zeit. f. Physik*, **34**, 34 (1925); and *Philosophie der Raum Zeit-Lehre*, p. 202
§10	Reichenbach.	*P. der Raum-Zeit-Lehre*, §42
	Heisenberg.	*Zeit. f. Physik*, **43**, 179 (1927)
§11	Einstein.	*Annalen d. Physik*, **17**, 893 (1905); **49**, 774 (1916). (On higher spaces). Solvay Congress (1927), *Reports*, p. 256
§13	Mach.	*Science of Mechanics*, pp. 224, 225, 541
	Poincaré.	*Dernières Pensées; L'Espace et le Temps*, p. 45
	Einstein.	Quotation on page xi. *Ann. d. Physik*, **49** (1916), 3
	Eddington.	Quotation on page xi. *Math. Th. of Rel.*, pp. 34, 225; *Space, Time, and Gravitation*, p. 149
	Robb.	*Theory of Time and Space*, p. 6
	Russell.	*Analysis of Matter*, chaps. 7 and 38

	Bridgman.	*Logic of Modern Physics*, pp. 67-79, 113-4, 150-166, 176
	Broad.	*Scientific Thought*, pp. 361-386
	Reichenbach.	*Phil. der Raum-Zeit-Lehre*, p. 202, §42
§16	London.	*Eisenschitz and London. Zeit. f. Physik*, 60, 491 (1930)
	Planck.	*Das Prinzip der Erhaltung der Energie*, p. 110
	Voss.	*Münchener Berichte* (1901), p. 167
§17	Lemaitre.	*Annales de la Soc. Sci. de Bruxelles*, **47**A, 49 (1927)
	Eddington.	*Monthly Notices.* R.A.S., May, 1930
	Tolman.	*Proc. Nat. Acad. Sci.* (1930), p. 582
§18	Heisenberg.	*Phys. Princ. of the Quantum Theory* (1930)
	Jordan.	*Zeit. f. Physik*, **40**, 809 (1926-7)
	Neumann.	*Göttingen Nachrichten; Math-Phys. Kl.* (1927), pp. 1, 245. (On the non-relativistic nature of the Transformation theory), *Zeit. f. Physik*, **48**, 868 (1928)
	Dirac.	*Principles of Quantum Mechanics* (1930)
	Born and Jordan.	*Elementare Quantenmechanik*, p. 288
§19	London.	In *Quantentheorie und Chemie; Leipziger Vorträge* 1928, p. 65
	Dirac.	*Princ. of Qu. Mech.*, p. 15. (Deduction of periodicity), p. 114
	Heisenberg.	*Phys. Princ. of the Qu. Th.*, pp. 61-2
	Neumann.	*Gött. Nachr.* (1927), p. 248
§20	Breit.	*Proc. Nat. Acad. Sci.*, **14**, p. 553 (1928)
	Heisenberg.	*Phys. Pinc. of the Qu. Th.*, p. 85. *Zeit. f. Physik*, **65**, 4 (1930)
§21	Condon.	*Phys. Rev.*, p. 1121 (1930)
	Heisenberg.	Quotation from German edition of *Phys. Princ. of the Qu. Th.* Cf. English ed., p. 10. (On representation of waves and particles), pp. 78, 92
	Dirac.	*Princ. of Qu. Mech.*, p. 2

§24	Dirac.	*Proc. Roy. Soc.*, **106** (1924), p. 581
	Jordan.	*Naturwissenschaften* (1927), p. 792
	Einstein.	*Berl. Ber.* (1916), p. 184
§29	Planck.	*Das Weltbild der neuen Physik* (1930), pp. 11-13
§31	Einstein.	*The Times*, Feb. 5th, 1929
	Nicod.	*Foundations of Geometry and Induction*, p. 12
	Russell.	In *Introduction* to Nicod's *Foundations of Geometry*
§32	Keynes.	*Probability*, p. 258
	Nicod.	*Foundations of Geom. and Induction*, p. 277
§35	Mach.	*Populär-wissenschaftliche Vorlesungen*, p. 499

TWO HISTORICAL APPENDICES

TWO HISTORICAL APPENDICES

An understanding of the two great theories of this century cannot be reached without a study of their development. Though one must believe that physical science advances towards an objective or generally acceptable formulation of unchanging natural law, yet the form of the particular theories developed during short periods such as thirty years depends not only on the somewhat arbitrary classes of facts made available in these years by experimental research, but also on the mental temperaments of individual workers and the dominant traditions and outstanding problems at the time of their intellectual development. The following essays are merely an introduction to the study of an exceptionally interesting period; no more has been attempted than to emphasise the special methods and tendencies which gave rise to the characteristic features of the two theories, without going into further detail than is necessary to pay honour to great names. Though the historical approach has various disadvantages, it has also the ample justification that it encourages an escape from the fashion of the moment which always over-emphasises one particular mathematical method and tends to neglect deeper lying and more permanent features of physical theory.

The first essay deals with the unique characteristics of Einstein's work in the theory of relativity, and provides an opportunity for an appreciation of his inspired persistence. The second traces the development of the quantum theory and shows the contrast to that of relativity theory. The two essays taken together are intended to suggest some of the historical causes which have contributed to the peculiar nature of the present situation in physical theory. The relation of the two theories is better understood when it is considered that physical theory might well have developed differently during this period even if necessarily converging on the same comprehensive theory ahead.

A. EINSTEIN AND THE THEORY OF RELATIVITY

It is now twenty-five years since Albert Einstein, in his first paper on the theory of relativity, began a life work in many ways unique in the history of scientific thought. The chief steps in this work already lie some way way back in the past, and it is possible to trace its development from a broad perspective. We shall attempt to show how Einstein's thought is distinguished from that of his contemporaries, to make a forecast of the considered judgment of future generations on his theory and to reveal something of the personal motives which inspire his work.

Einstein's intellectual development took place at a time when many efforts were being made to explain the result of the Michelson-Morley experiment, which had failed to show any influence of the earth's orbital motion on the measurement of the velocity of light. During the last decade of the nineteenth century various physicists, including Lorentz, Fitzgerald and Larmor, had been considering whether the apparatus used in this experiment could not be regarded as having changed in size owing to the forces between the atoms being influenced by their motion through the ether.

Provisional explanations of this kind had already been given, but Einstein, a younger man and more ready to attack fundamentals, set out to discover what could be done by assuming that the problem—of explaining why the experiment was not affected by the earth's motion—had only arisen through a wrong method of approach. His first paper, that of 1905, uses the method of theoretical research which is characteristic of all his work. He regards problems or difficulties, not as representing a situation inherent in nature and requiring complicated explanation, but as arising from the use of wrong conceptions in approaching an essentially simple situation. Thus instead of attempting a detailed atomic explanation of the experimental result, he sought a more general method of approach which would eliminate rather than solve the problem, and yield results without assuming any special theory of the atomic structure of matter. Experiment indicated that physical laws and the velocity of light were the same for all observers in uniform relative motion, so he took these two facts and formulated them as fundamental postulates.

This is the crucial step in the Special Theory: the expression as postulates of facts given directly by experience, and the ruthless acceptance of their logical consequences, which in this case meant the giving up of the deep-rooted conception of the simultaneity of distant events. Thus the atomic problem of what caused the alteration in size was transformed into a demand for the revision of a fundamental concept. This revision led to the substitution of relative for absolute simultaneity, which in turn gave as a natural consequence an alteration in the sizes of bodies and the rates of clocks in relative motion. This alteration was no longer surprising, for Einstein's postulates gave a new meaning to space-time measurements. Other consequences were that electric and magnetic fields have to be regarded as aspects of one essential type of electromagnetic field, and the result that energy always possesses mass.

In connection with the 1905 paper it is interesting to note two facts regarding the relation of Einstein's work to that of his contemporaries. In the Special Theory he is not tackling a new problem not yet noticed by others, but developing new and more fundamental methods for an old problem. From 1890 onward this problem had been recognised and in 1904 and 1905 Lorentz and Poincaré published papers containing results very similar to Einstein's. But they were still using earlier ideas which had neither the logical simplicity nor the generality of those of Einstein's

Special Theory. On the other hand, Einstein was not the first to use a space-time continuum with an imaginary time-coordinate as a fourth spatial dimension. This powerful mathematical method was first applied by Poincaré in 1905, then further developed in 1907 by Minkowski, and finally became of great importance through the fact that it provides the natural mathematical basis for the whole theory of relativity, though Einstein did not use it in his first paper.

These contrasted facts show that Einstein's greatness is as an imaginative physical thinker rather than as an applied mathematician. It is not the virtuosity of a mathematician, but the conviction that the uniformity of nature is best expressed in simple mathematical law which leads him to use mathematics. He finds himself compelled to express his new ideas or physical intuitions in an algebraic formula, for this alone can guarantee that "highest purity, clarity, and certainty" which he has described as the ideal of the theoretical physicist.

The postulates of the Special Theory had that clarity, but they could not be final, for their range was limited by the exclusion of gravitational phenomena and accelerated co-ordinated systems. In 1906 two lines of advance were open. The obvious one for those who believed in the Special Theory arose from the fact that Newton's law of gravitation assumed instantaneous action at a distance, whereas in the Special Theory of Relativity no influence could be propagated with a velocity greater than that of light.

Several workers therefore tackled the problem of modifying the Newtonian law to conform to Einstein's Special Theory. But, unlike these followers, Einstein was not himself satisfied that this theory provided an adequately general or secure basis for a complete physical theory. He therefore selected another much harder but more fundamental method of advance, which led him after nearly ten years' search to the General Theory of Relativity.

In 1907 he was already able to outline the new method.

This was to be based on a generalisation of the first postulate of the Special Theory ; the laws of physics were now to hold in all coordinate systems. (Principle of General Covariance.) Its companion, the Principle of Equivalence, postulated that all the effects of a gravitational field were equivalent to the use of an accelerated coordinate system. The value of this principle was that it permitted Einstein to deduce facts about the behaviour of clocks in a gravitational field without knowing anything about the structure of the clocks.

The velocity of light which provides a kind of clock, and the theory showed that the velocity of light was variable and that light rays did not travel in straight lines in a gravitational field. This enabled Einstein in 1911 to predict a value for the deflection of stellar light passing near the surface of the sun ; but almost immediately afterwards he began to develop another line of thought which indicated that this result might need revision.

If the velocity of light is variable the definition of relative simultaneity provided by the Special Theory, which postulated that this was constant, can no longer be used. Several workers were tackling this new situation by arbitrary and special methods when Einstein showed, in 1913, how the equivalence of a gravitational field to an accelerated coordinate system could be expressed by a fusion of gravitation and geometry.

This advance in thought involved a similar transformation to that of the Special Theory. According to Einstein, the problem of gravitation arises only because we have assumed the truth of Euclid's geometry, instead of a more general space-time geometry, which permits gravitational and inertial motions to be treated as essentially the same. The general idea of the fusion was clear in 1913, but the definite quantitative formulation which would permit the prediction of gravitational motions was still lacking. The task was to find a geometry (i.e. a law of gravitation) which would hold in any coordinate system.

In 1914 he gave up this aim, believing that he had found reasons why such a law could not exist, but in the next year he realised this was a mistake and was able to publish in his General Theory field equations of gravitation satisfying the necessary condition. Moreover these equations, or others very similar, provided the only gravitational law which could be expressed independently of a specially selected coordinate system, and gave the Newtonian law as a first approximation. Einstein then showed that his law of gravitation led to the three well-known predictions (including a revised value for the deflection of light by the sun) which have been confirmed as far as the accuracy of the evidence permits.

It is interesting to note that Einstein's law alone, without the assumption of Newton's law as a first approximation, does not say whether gravitation is an attractive or repulsive action, and indeed permits both forms.

One difficulty still remained in gravitational theory. The General Theory predicted a rotation of the orbit of the planet Mercury, i.e. a rotation relative to the inertial frame called by Newton " absolute space " and distinguished in relativity theory by special geometrical properties. But a complete relativistic theory, instead of leaving this frame unrelated to anything else, must show that it depends on all the masses in the stellar universe. An argument along these lines led Einstein in 1917 to modify his 1915 law and to propose a Cosmological Theory based on the view that space is finite. No final agreement has yet been reached as regards the experimental evidence for this theory.

During the following ten years, and, indeed, right up to the present time, Einstein has been occupied in developing a theory comprising electromagnetism as well as gravitation. His view is that it should be possible to complete a large scale or macroscopic field theory for both types of action, and that quantum phenomena should be considered subsequently.

He has therefore sought for a unified expression of his own General Theory and Maxwell's field equations of electro-magnetism. In 1928-29 he was able to formulate his "Unitary Field Theory" in which this fusion is achieved by using a new kind of space-time geometry designed for the purpose. It is too early to form a proper judgment on this development, but many physicists fail to find in its present formulation that self-consistency and inevitability which marked the earlier theories. In attempting to combine gravitation and electromagnetism Einstein may for the first time have entered a region where the unification sought for must be reached, not by the kind of field theory he uses, but by new conceptions built on atomic or quantum theories.

Yet even if new ideas arising from the quantum theory do prove to be required the problem of a unified physical theory can be solved only by building on work done by Einstein, for during the years 1905-17 he was able, while developing the theory of relativity, also to make very important contributions to the quantum theory of light. But the fact that Einstein's penetrating imagination failed to establish any substantial connection between quantum and gravitational theory during those twelve years when he was himself making fundamental additions to both, shows how difficult or how new the necessary conceptions must be.

We have reviewed the most important stages in the development of the theory of relativity and shall now attempt a forecast of the valuation that will be put upon it in the future.

1. It will probably be regarded as showing a progressive development by one man of new basic conceptions unique in the history of physics.

2. Einstein's own estimate will surely be approved: that "it is distinguished by the degree of formal speculation, the slender empirical basis, the boldness of theoretical construction, and finally the fundamental reliance on the uniformity of the secrets of nature and their accessibility to the speculative intellect."

3. Valued strictly within their legitimate ground as large scale field theories neglecting microscopic structure, the Special and General Theories must be considered to leave nothing to be desired as regards logical consistency, or degree of unification won by the elimination of unnecessary hypotheses and concepts. As regards empirical validity no forecast can be made, but within the accuracy of present data these theories are fully confirmed.

4. Its neglect of structure and use of a four-dimensional continuum with an imaginary time-coordinate (in place of the space and time of physical experience) shows that the theory of relativity is not a general or final theory of physical space and time. Future researches may well emphasise the fact that no conclusions about space and time can be drawn with any certainty from the theory outside its proper range.

No great thinker regards his work as final, and Einstein has pointed out that Newton knew the weak points of his theory better than the generations of scholars who followed him. It is likely that in the future, when the limitations of the relativity theory are known more definitely, it will be seen that the same can be said of Einstein. Some of the points at which he has expressly asserted the approximate or provisional nature of his theory can be made clear in a few lines, though they have been neglected by most of his exponents. Throughout the theory "measuring rods" and "clocks" are treated as simple conceptions, neither requiring nor capable of analysis, but Einstein has himself pointed out that this is convenient only until an atomic theory of the structure of matter is available. This point of contact between Einstein's theory and atomic physics may prove important for future research. Moreover the theory is admittedly only approximate, since it neglects the inexactitude of all space-time measurements.

These and other points where relativity theory fails to provide what must be demanded of a final theory of space and time will probably only be cleared up when gravitational and atomic theory have been brought into a closer relationship. Some exponents of the theory of relativity have suggested that it has reduced space and time to the same level but Einstein has emphasised the fact that time enters the fundamental formulæ of the theory differently from space and that a distinction is thus maintained.

One fact stands out from our survey: the development and cumulative effect of the successive achievements, as though they were guided by an aim seen clearly from the start. Einstein's life work is distinguished from that of many others who were simultaneously interested in the same problems by this mark of a personal conviction enabling him patiently to work out new fundamental methods without discouragement at their novelty or difficulty.

Such persistence is probably always the sign of a thinker whose ideas come earlier than their justification, and this has in fact been Einstein's experience. It is likely that often in the years between 1907 and 1915 he felt about the possibility of carrying out his scheme for a general relativity theory just as Newton felt when he wrote in 1676 about a geometrical theorem: "It is plain to me by the fountain I draw it from, though I will not undertake to prove it to others." Yet the development of new methods in order to justify the intuition within a logically consistent theory proved to be of extreme difficulty, and, as we have seen, there was at least one moment when he was temporarily led astray. He attempted things he could not achieve, and gave up as impossible tasks to which he later returned successfully.

Theoretical physics has three aspects: the basis of experimental fact, the physical concepts implying a particular view of physical reality, and the mathematical development of theory. All scientists accept empirical fact as the final test of validity; mathematicians are chiefly interested in the mathematical processes, while the theoretical physicist is concerned with the study and revision of physical concepts. During the period before Einstein's student days Mach had attacked the view of reality implied in the ideas of mechanics and his criticism of physical concepts exerted much influence on Einstein. Yet while Einstein was helped by Mach's views to a scepticism of current concepts, his own faith in an ultimate mathematical harmony was necessary to stimulate and guide him towards new constructive ideas. For Einstein the concept, or idea, is of chief importance. The concept is physically significant if it permits the representation of the uniformity of nature in simple mathematical form. Thus an algebraic equation is significant for him because the symbols represent numbers, which in turn express the uniformity and harmony of nature. In this Kepler, Newton and Maxwell have been his great leaders, and his name will be remembered with theirs.

There is nothing finer in human life than a clear vision followed in patience to its ultimate fulfilment. Such devotion, tested and proved fertile, yields a dignity to life without which its perpetual rhythm would sink into worthlessness. For this the world owes Albert Einstein as much as it can owe to any man. We cannot help asking what is the inspiration of such a character. Einstein's own words on Planck throw light also on himself and can be accepted as a true statement of the motive which gave rise to the theory of relativity:

"The longing to see this pre-established harmony is the source of the inexhaustible patience and persistence which we see in Planck's devotion to the most general problems of our science, undeflected by easier or more thankful tasks. I have often heard that colleagues sought to trace this characteristic to an extraordinary will-power and discipline; but I believe this to be wholly wrong. The emotional condition which renders possible such achievements is like that of the religious devotee or the lover; the daily striving is dictated by no principle or programme, but arises from an immediate personal need."

B. THE DEVELOPMENT OF THE QUANTUM THEORY, 1899-1928.

RELATIVITY theory and quantum theory represent different but overlapping methods of approach to the subject matter of physics, and though leaving many problems still unsolved they include together all the fundamental theoretical advances that have been made in the last thirty years. From 1905 onwards relativistic methods have deeply influenced the development of the quantum theory, and yet this theory guided by its own range of fact has continually developed its own methods, and no radical fusion of the two theories is yet in sight. Light can be thrown on this situation by the recognition that in spite of their simultaneous development (relativity theory 1905-1917; quantum theory 1899-1928) a remarkable contrast exists in the manner in which the two theories evolved. Certain aspects of the history of relativity theory have been described in Appendix A., we are now concerned with the contrast shown by the history of quantum theory. In this brief survey it will not be possible to refer to experimental workers, but it is important to emphasise that detailed experimental work played a more intimate and immediate part in the growth of the theory than ever before in the history of physics.

In order to bring out the main outlines of the development of the quantum theory it will be convenient to state in advance three aspects in which it is in strong contrast to relativity theory. First, the theory of relativity is effectively the work of one man, and its unity and self-consistency throughout successive stages of development result from this fact. On the other hand quantum theory rests on the important theoretical contributions of some fifteen or more workers. The possibility of the final convergence of their contributions in one comprehensive theory, the transformation theory of 1927, is to be accounted for partly by the fact that many of these workers were consciously building on each other's work, or were deeply influenced by personal contact, but also in important cases as a result of the intrinsic situation in physics. Independent workers were frequently compelled by the facts to develop complementary, equivalent, or even identical theories.

The second contrast is in the approach of the two theories to fundamental physical law. Einstein starts from the conceptual problem of simultaneity, and is led to new conceptual methods which suggest the use of special mathematical notations. At each stage in relativity theory a new fundamental physical idea inspires and unifies the development of the mathematical theory. In quantum theory on the contrary no new *primary* concepts are introduced adequate to allow a complete physical interpretation of the mathematics; the important new ideas of quantum theory are secondary, in that they have only received their definition in terms of classical ideas and do not justify the axioms of the theory. The absence of adequate new concepts is indicated by the necessity for a general principle governing the analogies with classical concepts, and the abstract character of the symbolic algebra of quantum theory may be regarded as a natural result of the absence of appropriate new fundamental concepts.

Yet a third antithesis is in the relation of the theories to experimental fact. Relativity theory grew from conceptual problems arising from a narrow range of fact, and the confirmation of the general theory is limited to three highly special kinds of observation, of which probably only one (the orbit of Mercury) can be considered as absolutely unquestionable. The quantum theory, however, is based on a vast range of detailed atomic facts, both descriptive and quantitative, which have always been arranged under widely valid and purely empirical rules before being incorporated in the wider scheme of rules called the quantum theory.

These contrasting aspects of the two theories may be summarised thus; the theory of relativity is the mathematical expression of a conceptual revision of physical theory rendered necessary by a few important facts and accomplished by one man, while the quantum theory is the formulation in abstract symbolic algebra of an interrelated system of rules covering a mass of empirical detail collected and arranged by many workers and partially supported by the concepts of classical and relativity physics. These features will now be traced in a rapid outline of the history of quantum theory.

The character of the quantum theory as a system of algebraic rules for describing empirical facts is shown from the start in the manner in which Planck originally discovered the constant of action (h) necessary in the empirical formula for the density of black body radiation. In 1899 two radiation laws were current, one of which, Wien's Law, was valid for low temperatures and short waves, while the other, the Rayleigh-Jeans Law, was correct for high temperatures and long waves. Planck has described how he hit on the correct expression by regarding these two laws as limits reached in extreme cases, and obtaining a more general law by simple combination of the two limiting laws. In his own words the new law was at this stage "a fortunate guess at an interpolation formula." But after a few weeks of intense work he obtained his well-known theoretical interpretation of the new law based on the postulate that the energy of the classical resonators of a given frequency can only vary by a constant amount proportional to their frequency. The arbitrary nature of this postulate in relation to classical theory shows that Planck's was only a partial theory. A complete theory, i.e. a theoretical derivation from a unified system of concepts and laws was, and still is, impossible as long as classical concepts are given a primary position.

The situation in which Planck found himself has been that of many quantum physicists since, and his work is typical of the whole development of the theory. A group of experimental facts is brought under one rule by elementary arithmetical methods, and the rule is then—subsequently to its discovery—given such theoretical interpretation as is possible by partial revision or limitation of classical concepts. But the revisions which were attempted have always been found inadequate and hence the development of quantum theory, forced at each stage by new empirical rules, passed necessarily from Planck's original attempt at a physical theory, through the similar brilliant attempts of Bohr, Einstein, and Schrödinger to the abstract formalism of Heisenberg-Jordan-Dirac. Planck's pioneer step has from this point of view exactly the same status as all later contributions; neither he nor any of his successors modified classical ideas adequately to produce a comprehensive and unified physical theory of the different empirical rules. Each in turn either extended the range of old rules, or improved their mathematical expression, using as guide either partial physical analogies from classical relativity physics or pure mathematical analysis, but never basing the mathematics on new concepts as fundamental as those of relativity theory.

Einstein was the first (1905-1909) to extend Planck's ideas to a wider range of phenomena, by limiting not only the classical resonators but also the classical radiation field to discrete quanta of energy, here called light-quanta. This hypothesis though unsupported by any classical theory of how this limitation came about, was able to correlate three empirical rules (Planck's Law, the photo-electric equation, and the variation of specific heats) as well as a theory of the fluctuations of radiation based on Boltzmann's probability-entropy relationship. The quantum theory of specific heats was improved by Debye and the light-quantum hypothesis, as subsequently further developed by Einstein and others, gave rise to the branch of quantum theory which led to the work of de Broglie and Schrödinger. This line of thought was guided largely by *physical* analogies based on classical models, and some of the workers using this method even believed that quantum phenomena could be regarded as special cases of the processes already studied in classical theory.

The other parallel branch of the theory was founded by Bohr (1913) who used Planck's constant in a set of theoretically unexplained postulates about the possible electron orbits and modes of radiation of the hydrogen atom. On this basis a partial theoretical derivation was given of Balmer's series and the Rydberg constant, and in 1915 Bohr drew attention to the fact that the fine-structure of the hydrogen spectral lines might be ascribed to the relativistic modifications of the classical electron orbits assumed in his earlier theory. This was taken up by Sommerfeld and made the basis of an exact theory of relativistic fine-structure. During the following years Bohr began to stress the view of the quantum postulates as a simple set of rules for predicting spectra which were not to be regarded as inviting classical interpretation. The essence of his Correspondence Principle (1919-23) was that classical analogies were to be used not to provide a physical theory of quantum phenomena, but to suggest mathematical expressions for describing them. The necessity for a heuristic principle of this kind arose because the Bohr-Sommerfeld theory of the atom left many atomic properties undetermined. With a genial certainty of the inadequacy of classical physical models Bohr developed a school of thought in which it gradually came to be accepted, that, for the time being at any rate, classical models were simply to be used as an heuristic aid to the discovery of the correct algebraic rules for describing phenomena. This revolutionary outlook, which inspired the development of the second branch of the quantum theory, was followed up keenly by a group of young workers in Copenhagen and Göttingen.

The pioneer who transformed these tentative methods into a strict mathematical discipline was Heisenberg (1925). He sought to establish a mathematical scheme connecting the directly observable quantities without the support of the concepts of classical theory. The new scheme was to be reached by analogy and generalisation of the mathematical forms of classical theory; these forms were to be substantially retained, and only the method of deducing experimental results modified to suit quantum phenomena. With the assistance of Born and Jordan an abstract symbolic scheme (matrix theory) was developed (1925-26) which, though difficult because entirely novel to physicists, was able in principle to answer many of the questions which were beyond the range of the earlier theory. Various alternative forms of the same methods were soon evolved of which the most important are the operator theory of Born and Weiner (1926) and the "*q* number" theory of Dirac (1926). Yet the matrix theory of quantum mechanics left many problems still untouched, and was extremely unsuited to the solution of actual problems. These difficulties were only overcome by the new methods which had been developed by others in connection with Einstein's light-quantum hypothesis.

For the last century the atomic or *particle theory of matter* and the *wave theory of radiation* had held the field. The success of his own application of the particle conception to radiation had suggested to Einstein the complimentary idea of applying wave theory to material particles, but only in 1924 was he able to do this in extending Bose's new statistical method to the molecular theory of gases. The same idea had occurred to de Broglie who in 1922 using classical and relativistic methods derived theoretically the correct value for the wave-length of a particle with given momentum. It is interesting to note that though it rested on arbitrary assumptions this magnificent step not only completed the symmetry of the wave-particle theories of matter and radiation but also linked the quantum constant with the methods of relativity theory. Yet this symmetry and linkage could only be superficial in view of the essential inadequacy of classical ideas, and the physical form given to de Broglie's discovery soon proved to be misleading from the point of view of a general theory.

De Broglie's application of wave theory to particle phenomena was given a more general and precise formulation by Schrödinger (1926), in a pseudo-classical theory which allowed the application of a powerful mathematical technique to those problems which could not be conveniently handled by matrix theory. Yet the great success of the de Broglie-Schrödinger theory in predicting electron and light wave-lengths and in treating other atomic problems tended to conceal the fact that just as waves had proved inadequate for describing the properties of light, so must they be inadequate for material particles. This was shown, for example, by the failure of a relativistic Schrödinger wave equation to yield the correct fine-structure formula as already given by the earlier Sommerfeld theory. Moreover in complex problems the Schrödinger "waves" have to be represented not in three-dimensional space but in a higher space of many dimensions, the simpler physical interpretation leading in such cases to wrong results. Precisely the inestimable advantage of Schrödinger's theory, the provision of a pseudo-classical model with correspondingly powerful and well known mathematical methods, by concealing the essential inadequacy of classical methods proved a disadvantage from the point of view of further theoretical advance. This situation would have delayed further progress had not the Heisenberg and Schrödinger methods been discovered to be equivalent. In the Heisenberg matrices and the Schrödinger waves the two lines of advance had reached mathematically equivalent formulations, as was immediately shown by Schrödinger and others (1926). The remarkable fact that methods so completely different in inspiration and mathematical technique should lead to the same results shows that the empirically given facts of a region of physics may be arranged in apparently different but actually equivalent ways, and that the valid content of a theory may therefore often be less than its apparent content. Thus the real content of physical theory at this point was that all the mathematical problems of atomic theory could be expressed in terms of what is called an "eigenwert" problem; wave theorists had often incorrectly tried to imply in addition the validity of classical wave models.

It was natural therefore that the subsequent advance came from the side of the more abstract school. The mathematical equivalence of the two methods had been proved, but the physical inter-relations of the two were still obscure. The first link and the clue to further advance was supplied by Born (1926), who showed that in certain cases a function of the Schrödinger wave quantity expressed the probability of a particle being found at a given point at a given moment. This idea was extended by Pauli and Jordan, and led to the final fusion of wave and matrix mechanics and of the pseudo-wave and particle models in the Probability Transformation Theory of Dirac and Jordan. This theory represents a natural and probably final generalisation of the earlier partial quantum theories, it was developed independently by Dirac and Jordan in 1927, and has not been essentially modified since. The transformation theory is a set of general and abstract algebraic rules which when applied to any given system can be interpreted so as to permit the calculation (i) of all the possible results of a given measurement, and (ii) of "transformation functions" by which the numbers representing the given initial conditions are transformed into an expression of the probability of any given result being obtained in a measurement made on the system. Just enough physical interpretation is given to the abstract rules to allow the determination of the results of experiments on systems of known type (protons, electrons, radiation, and their interactions), but no general physical meaning is given to the rules themselves. An example of the former unification achieved by the theory is that Planck's constant is introduced in one postulate, and not in two as in Bohr's original theory.

Another essential physical link in the interpretation of the wave, matrix and transformation theories was provided by Heisenberg (1927) in his Principle of Inexactitude, subsequently amplified by Bohr. This principle leads to the recognition that wave and particle models are to be regarded as provisional analogies which in general break down unless each is used to correct the implications of the other, though in special cases one model may alone be adequate. Heisenberg has pointed out that the physical interpretation of the fact that quantum laws determine probabilities based on the necessary inexactitude of the wave-particle model was discovered subsequently to the mathematical formulation of the probability transformation theory. This is another example of the passage from mathematical rule to partial physical interpretation which marks the whole development of quantum theory.

Yet another striking case is given by the electron spin phenomena. In 1924 Landé discovered a general empirical rule permitting the qualitative classification of the spectral lines in complex types of Zeeman effect, and Pauli, in an attempt to interpret this, was led in 1925 to the axiom that every electron requires four quantum numbers. This was without theoretical justification until later in 1925 Goudsmit and Uhlenbeck proposed the pseudo-classical model of a spinning electron which gave an immediate interpretation to the necessity for four quantum numbers. But no pseudo-classical model could remain adequate within the general mathematical scheme of quantum theory and hence Dirac was driven to find (1928) an abstract mathematical expression satisfying the mathematical principles of quantum theory which would—at least to an adequate approximation—lead to the same results as the relativistic-spin-electron model.

In the Dirac-Jordan transformation theory the two branches of quantum mechanics are fused and each receives its justification. The abstract mathematical methods of the correspondence principle and matrix theory are shown to have greater generality and theoretical importance, while wave mechanics has the enormous practical advantage of supplying a physical picture adequate to simple cases and permitting the solution of many problems which in matrix theory become inconvenient or intractable. For example, Dirac's important relativistic treatment of the electron (1928) is formulated as a wave-equation, the more general symbolic method proving in this case unsuitable.

From 1927 onwards many workers have applied wave and matrix mechanics to a wide range of physical phenomena such as chemical action, radio-activity, thermionic emission, conduction in metals, the crystalline state, molecular spectra, etc. In most cases the new quantum mechanics gives evidence of providing a correct theoretical approach to these problems, though the mathematical technique available does not often allow an exact treatment. Various approximate treatments have been developed but these rest principally on a pseudo-classical approach and the greatest need remains a simple and more powerful mathematical calculus for complex problems. This lack may be connected with the absence of a satisfactory relativistic theory, since the present form leads to inconsistencies. Dirac's view is that the abstract rules and general physical principles which form the basis of quantum mechanics are adequate for the description of the ordinary physical and chemical properties of matter. It is not, however, yet certain that they provide a suitable basis for a relativistic quantum theory and here the future alone can decide.

It is generally agreed that with the transformation theory (1927), the principle of inexactitude (1927), and Dirac's relativistic electron (1928) a definite stage was reached in the development of quantum theory. After a rapid period of development, starting in 1923 and reaching its greatest intensity in 1926, the chief problems which have so far occupied theoreticians were solved in 1927-28, and since then no generally accepted fundamental improvements have been made. In this situation an examination of the status of the transformation theory is of great importance, and various views have already been expressed. Dirac, treating it as a final culmination of quantum theory, suggests that nature's "fundamental laws (the abstract algebraic axioms of quantum mechanics) do not govern the world as it appears in our own mental picture . . . but instead control a substratum of which we cannot form a mental picture without introducing irrelevancies." This extreme standpoint is based on the view that the symbolic method goes most deeply into the nature of things, and is a natural, though perhaps daring, generalisation from the developments of the last decade. Heisenberg takes a rather different view. To him the breakdown of current physical concepts (waves, particles, ordinary space-time) is a result of their inadequacy to describe experience, and does not necessarily mean that the ultimate laws must be purely symbolic. If this view is correct the symbolic basis of quantum mechanics may receive physical interpretation in terms of new modifications of our idea of space-time demanded by the primary conditions limiting physical observations. This standpoint is consistent with the general tendency we have traced in the history of the theory, though the physical interpretation of these general abstract rules may take years where Planck's special interpolation formula required only a few weeks.

Thus in the interpretation of the transformation theory we reach the point where accepted formulation passes into speculation about the future development of physical theory.

It will, therefore, be useful to look back from the established mathematical scheme of the transformation theory and examine once more in the light of this scheme the methods that led to it. Here we can only summarise the conclusions that arise from this analysis.

(i) The use of classical models, though hitherto of the utmost importance as a guide to theoretical and empirical research, has now exhausted its importance as a theoretical guide. These models are valid only in highly special situations, and approximate or partial concepts of this kind cannot usefully be made the basis of further theoretical research. In this sense, i.e. *for further research*, the work of de Broglie and Schrödinger is less important than that of Heisenberg, Born, Jordan, Pauli, and Dirac, whose methods must contain either the starting points for new advances or the points where any new theory, possibly built on a more general basis, must seek its contact with quantum mechanics.

(ii) The independent growth of different yet equivalent mathematical methods, such as those of Schrödinger and Heisenberg, emphasises a fact of great importance for physical theory. The forms given in physical experience may be described in many different mathematical notations, with the aid of different physical assumptions. It is therefore always of the utmost importance to reduce mathematical methods to their simplest form to relate these methods to the physical postulates and the physical postulates to the data given in immediate observation.

(iii) In the absence of any primary guiding concept, the physical situation has itself compelled workers of very different kinds to work along convergent lines. Thus the different groups of workers who approach new problems from the old physical models (Planck, Einstein, de Broglie, Schrödinger), from the empirically given observations or rules (Heisenberg, Born, Pauli) or by general symbolic methods (Jordan, Dirac) have each contributed to the final result.

This rapid survey of the growth of quantum theory makes clear the contrast with relativity theory which has already been summarised. But even if we do not find an equivalent figure to Einstein, Bohr stands out with a unique physical instinct for the kind of method required, and not his least contribution has been the inspiration and encouragement of younger men. Of these Heisenberg is supreme; in him a clear mathematical mind is tempered with physical intuition. Einstein, de Broglie, Pauli, Born, Schrödinger, Bose, Fermi, have made indispensable contributions, and countless others have assisted by accumulating and arranging empirical facts. Jordan and Dirac have contributed the formal intelligence necessary in a mathematical fusion of different methods. The history of quantum theory, particularly since 1923, is the story of an intricate flux of ideas developed in almost daily interchange. Never in this history of thought has the growth of new methods been so rapid or so collective as in 1925-27 between the younger group, Heisenberg, Pauli and Jordan, assisted and guided by Bohr and Born. Outside this great school, the work of de Broglie and Schrödinger and of Dirac grew relatively independently. If we pass from men to methods, the contrast with relativity theory is still as marked. There is no equivalent in quantum theory to Einstein's new concepts proposed in advance of mathematical theory; the correspondence and inexactitude principles express the way in which classical concepts may be used but supply no substitute. Heisenberg's principle of inexactitude and Born's statistical interpretation of quantum quantities represent the chief quantum principles limiting the validity of ordinary space-time conceptions, but these have not yet been used to provide any physical justification for the form given to the abstract axioms of the theory. In fact there was inadequate time during the rapid development of 1923-28 for a systematic study of the common fundamentals of quantum and relativity theory such as might lead to new physical concepts adequate to provide a basis for a comprehensive theory. The vast range of spectroscopic facts offered sufficient problems to the ordering mind; most were preoccupied with the numerical rules obeyed by these facts and found no time for the fundamental theoretical problems which link relativity and quantum theory.

In conclusion emphasis must be laid on one aspect which relativity theory and quantum theory have in common. They are both the mature product of a long period of research, and represent the definitive expression of physical science on certain limited problems using certain definite methods of approach. They are not, however, self-enclosed logical structures leaving neither points of contact with each other nor unexplained features for subsequent study. The great textbooks of Eddington and Pauli on relativity theory and the brilliant treatments of quantum theory by Dirac and Heisenberg are full of indications of unsolved problems; even those whose task it is to expound what has already been achieved cannot avoid giving hints of all there is still to do. Both theories leave unexplored a very important borderland which links the two regions. The examination of this neglected region brings with it the study of the theory of measurement and of coordinates, and must open up a new stage in physical theory which will now have to concern itself not so much to improve these two great theories, as to explain in terms of some wider vista how they came about and why they leave some problems still unsolved. Neither in philosophic discussion nor in teaching can relativity or quantum theory be correctly represented as offering a general theory of space-time phenomena. Yet both theories are within their own regions so nearly complete that the form which will be given to their correlation is a matter of the utmost importance for the whole of natural science.

The present moment in the development of physical theory is thus one of extraordinary interest. Technology is one of the gifts of science to man; another less easily misused is the display of ideas in the process of growth. It is in the interests of human culture that this process be more fully studied, and nowhere could richer material be found than in the history of physical theory from 1900-1930. The next steps may be slow, but when achieved they are certain to provide an enlightening commentary on the work of this period.