PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS

London Office: Bentley House, N.W.1
American Branch: New York

Agents for Canada, India, and Pakistan: Macmillan

First Edition 1948
Reprinted 1949

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)

EDITOR'S PREFACE

AFTER the death of Sir Arthur Eddington in November 1944, Miss Eddington and the Syndics of the Cambridge University Press invited me to supervise the publication of the manuscript of the present volume. I have followed the text without alteration except in correcting obvious slips of the pen, making slight changes to remove obscurities, and supplying references which Eddington had left blank. I have added a few additional notes and have constructed an Index.

By way of general introduction, I may say that the work is complete in itself, and practically replaces all the author's previous writings on his theory of the constants of Nature. Chapters I–V follow closely the treatment adopted in his Dublin lectures of 1943, and Chapters VI–VIII are devoted to the sedenion analysis which had been expounded in his *Relativity Theory of Protons and Electrons* of 1936; but the rest of the book is chiefly new matter, and contains developments of outstanding power and interest. Those who desire a preliminary glance at the results may be advised to turn to:

(i) The table on page 66, which gives the values of the microscopic constants as calculated by Eddington's theory, compared with the observed values.

(ii) The similar table of molar and nuclear constants on page 105.

(iii) The first list of achievements of the theory in nuclear physics given on page 211: the numerical comparisons will be found in the separate sections of Chapter IX, Eddington's intention having been to collect them, together with the discoveries of Chapters X–XII, in a table in the part of the book which he did not live to complete.

(iv) The results for the magnetic moments of the hydrogen atom and the neutron on pages 249 and 251.

For a somewhat fuller introductory account of the theory, reference may be made to an article in the *Mathematical Gazette*, **29** (October 1945), pp. 137–44.

Professor E. T. Copson, of University College, Dundee, in the University of St Andrews, and Professor George Temple, F.R.S., Head of the mathematics department in King's College, London, have most kindly read the proof-sheets with me. I wish also to acknowledge gratefully the help given by the Staff of the Cambridge University Press.

EDMUND T. WHITTAKER

CONTENTS

I. THE UNCERTAINTY OF THE REFERENCE FRAME *page* 1

§ 1. The uncertainty of the origin, *p.* 1. § 2. The physical origin, *p.* 3. § 3. The Bernoulli fluctuation, *p.* 4. § 4. The standard of length, *p.* 7. § 5. Range of nuclear forces and the recession of the galaxies, *p.* 9. § 6. Spherical space, *p.* 10. § 7. Uranoids, *p.* 13. § 8. The extraneous standard, *p.* 14. § 9. Scale-free physics, *p.* 16. § 10. Pseudo-discrete states, *p.* 17. § 11. Stabilisation, *p.* 19.

II. MULTIPLICITY FACTORS *page* 22

§ 12. Complementary fields, *p.* 22. § 13. The rigid-field convention, *p.* 23. § 14. Separation of field and particle energy, *p.* 25. § 15. Application to scale-free systems, *p.* 26. § 16. The 'top particle', *p.* 28. § 17. Standard carriers, *p.* 30. § 18. Mass-ratio of the proton and electron, *p.* 32. § 19. Rigid coordinates, *p.* 35. § 20. The fine-structure constant, *p.* 37. § 21. The inversion of energy, *p.* 39. § 22. Mutual and self energy, *p.* 41. § 23. Comparison particles, *p.* 44.

III. INTERCHANGE *page* 46

§ 24. The phase dimension, *p.* 46. § 25. Interchange of suffixes, *p.* 48. § 26. The two-particle transformation, *p.* 51. § 27. Hydrocules, *p.* 53. § 28. Separation of electrical energy, *p.* 56. § 29. Current masses of the proton and electron, *p.* 58. § 30. Molarly controlled charge, *p.* 60. § 31. Secondary anchors, *p.* 62. § 32. Calculated values of the microscopic constants, *p.* 64. § 33. The Coulomb energy, *p.* 66.

IV. GRAVITATION AND EXCLUSION *page* 68

§ 34. Unsteady states, *p.* 68. § 35. Under-observation, *p.* 70. § 36. Structural and predictive theory, *p.* 72. § 37. Physical and geometrical distribution functions, *p.* 73. § 38. The weight function, *p.* 75. § 39. The genesis of proper mass, *p.* 77. § 40. Absolute determination of m_0, *p.* 79. § 41. Exclusion, *p.* 81. § 42. The negative energy levels, *p.* 83. § 43. Determination of m_0 by exclusion theory, *p.* 85. § 44. Super-dense matter, *p.* 87. § 45. The degeneracy pressure, *p.* 90.

V. THE PLANOID *page* 93

§ 46. Uranoid and planoid, *p.* 93. § 47. Interchange of extracules, *p.* 94. § 48. The special planoid, *p.* 96. § 49. The energy of two protons, *p.* 98. § 50. Non-Coulombian energy, *p.* 101. § 51. The constant of gravitation, *p.* 103. § 52. Molar and nuclear constants, *p.* 105.

VI. THE COMPLETE MOMENTUM VECTOR *page* 106

§ 53. The symbolic frame, *p.* 106. § 54. Miscellaneous properties of E-symbols, *p.* 108. § 55. Equivalence and chirality, *p.* 109. § 56. Rotations, *p.* 112. § 57. Five-dimensional theory, *p.* 114. § 58. Ineffective relativity transformations, *p.* 116. § 59. Strain vectors, *p.* 117. § 60. Real and imaginary E-symbols, *p.* 120. § 61. Reality conditions, *p.* 122. § 62. Distinction between space and time, *p.* 123. § 63. Neutral space-time, *p.* 125. § 64. Congruent spaces, *p.* 127. § 65. Determinants and eigenvalues, *p.* 129.

VII. WAVE VECTORS *page* 131

§ 66. Idempotency, *p.* 131. § 67. Standard form of idempotent vectors, *p.* 132. § 68. Spectral sets, *p.* 134. § 69. Catalogue of symbolic coefficients, *p.* 137. § 70. The wave identities, *p.* 140. § 71. Matrix representation of E-numbers, *p.* 141. § 72. Factorisation of E-numbers, *p.* 143. § 73. Wave tensors of the second rank, *p.* 146. § 74. Wave tensors of the fourth rank, *p.* 149. § 75. Phase space, *p.* 150. § 76. Relative space, *p.* 152. § 77. Vectors in micro space, *p.* 154. § 78. The quantum-classical analogy, *p.* 155.

VIII. DOUBLE FRAMES *page* 159

§ 79. The *EF*-frame, *p.* 159. § 80. Chirality of a double frame, *p.* 161. § 81. The interchange operator, *p.* 162. § 82. Duals, *p.* 164. § 83. The *CD*-frame, *p.* 166. § 84. Double-wave vectors, *p.* 167. § 85. The 136-dimensional phase space, *p.* 169. § 86. Uranoid and aether, *p.* 171. § 87. The Riemann-Christoffel tensor, *p.* 173. § 88. The de Sitter universe, *p.* 175. § 89. The tensor identities, *p.* 177. § 90. The contracted Riemann-Christoffel tensor, *p.* 178. § 91. States and interstates, *p.* 181. § 92. The recalcitrant terms, *p.* 184.

IX. SIMPLE APPLICATIONS *page* 186

§ 93. The metastable states of hydrogen, *p.* 186. § 94. Neutrium and deuterium, *p.* 188. § 95. Mass of the neutron, *p.* 191. § 96. Double intracules, *p.* 193. § 97. Comparison with field theory, *p.* 196. § 98. Mass of the deuterium atom, *p.* 197. § 99. Mass of the helium atom, *p.* 199. § 100. The separation constant of isobaric doublets, *p.* 202. § 101. Isotopic spin, *p.* 205. § 102. Radii of nuclei, *p.* 207. § 103. The nuclear planoid, *p.* 208. § 104. Mass of the mesotron, *p.* 211.

X. THE WAVE EQUATION *page* 215

§ 105. Field momentum, *p.* 215. § 106. The gradient operator, *p.* 217. § 107. Isostatic compensation, *p.* 219. § 108. Wave equation of the hydrogen intracule, *p.* 221. § 109. Solution of the wave equation, *p.* 224. § 110. The interchange momentum, *p.* 227. § 111. The two-frame transformation, *p.* 230. § 112. Electromagnetic potentials, *p.* 233.

XI. THE MOLAR ELECTROMAGNETIC FIELD *page* 236

§ 113. Gauge transformations (molar theory), *p.* 236. § 114. Action invariants, *p.* 239. § 115. Gauge transformations (microscopic theory), *p.* 241. § 116. Indices of wave tensors, *p.* 242. § 117. Magnetic moments, *p.* 245. § 118. Magnetic moment of the hydrogen atom, *p.* 247. § 119. Magnetic moment of the neutron, *p.* 249.

XII. RADIATION *page* 252

§ 121. Radiation by a moving electron, *p.* 252. § 122. Transition probabilities, *p.* 254. § 123. Compton scattering, *p.* 258. § 124. Transverse self energy of a particle, *p.* 261.

EDITORIAL NOTE *page* 264

APPENDIX. THE EVALUATION OF THE COSMICAL NUMBER 265

EDITORIAL NOTE 284

INDEX 285

Chapter I

THE UNCERTAINTY OF THE REFERENCE FRAME

1. The uncertainty of the origin

The quantities occurring in the equations of mathematical physics relate partly to physical objects and events and partly to a mathematical framework introduced for purposes of reference. Both relativity theory and quantum theory stress the distinction between *observables*, i.e. quantities which could be ascertained by a specified observational procedure, and *unobservables*, i.e. quantities depending partly on the auxiliary mathematical frame which cannot be the subject of actual observation. Unobservables are used to systematise description and facilitate calculation, but they are eliminated in the final steps of the calculation of observationally verifiable results.

In classifying quantities as observable or unobservable I shall follow current usage rather than the most literal meaning of the terms. Rightly or wrongly, modern physics is not over-scrupulous in postulating measurements of a highly impracticable kind. But although measurements are often very much idealised, it is recognised that idealisation must stop short of actual self-contradiction. Relativity theory and quantum theory have each laid down conditions of observability which are certainly necessary if self-contradiction is to be avoided; and the term 'observable' will here be understood to mean that the quantity satisfies both tests.[a]

Relativity theory begins with a denial of absolute motion. An observed velocity $d\xi/dt$ of a physical entity is necessarily relative to another physical entity. Likewise an observed coordinate ξ is a relative coordinate of two physical entities.

Quantum theory insists that the connection of a physical entity with the geometrical frame of coordinates is governed by Heisenberg's uncertainty principle. A particle is not exactly locatable at a point in the geometrical frame, or in four dimensions as a world-line. It can only be assigned a probability distribution of position and velocity.

In modern physics these two principles of observability have been applied separately with very far-reaching results; but they have seldom been applied in combination even by those who profess to be developing a relativistic quantum theory. The combined principle is that *a coordinate ξ is observable only if it is a relative coordinate of two entities both of which have uncertainty of position and momentum in the geometrical frame.*

The same considerations apply to momenta and other observables. An observable is always a statistic of a double probability distribution.

The essential point is that an observable coordinate is measured, not from an abstract mathematical point as origin, but from something which is involved physically in the operation which furnishes the measure. Being involved physically it experiences those incalculable reactions which limit the precision of our knowledge in the way laid down by Heisenberg's principle. We must therefore distinguish between the 'physical origin' from which an observable coordinate is measured, and the 'geometrical origin' of the auxiliary mathematical frame. The latter, as already stated, is

[a] Observables and measurables will be fully treated in Chapter XIII, and a logically satisfactory definition will then be given.

eliminated in the final calculation of observationally verifiable results; being therefore aloof from the rough-and-tumble of observational inquisition, it has a sharpness of definition which contrasts with the blurring of all physical landmarks by probability scatter.

Consider a system of particles with coordinates x_r, y_r, z_r ($r = 1, 2, 3, \ldots$) in the geometrical frame. These coordinates are unobservables. To obtain physical, i.e. observable, coordinates we must substitute for the geometrical origin an actual particle or its equivalent, e.g. the centroid of a set of particles. Let the geometrical coordinates of the physical origin be x_0, y_0, z_0; these also are unobservable. But the relative coordinates

$$\xi_r, \eta_r, \zeta_r = x_r - x_0, y_r - y_0, z_r - z_0 \tag{1·1}$$

are observables.

Nominally the exact value of ξ_r could be found by observation, at the expense of infinite uncertainty of the conjugate momentum. Such a measurement would be scientifically useless, since the coordinate would instantly become uncertain again; it is the measurement of a careless experimenter who destroys his specimen by handling it too roughly. Thus our knowledge of ξ_r, η_r, ζ_r at any time is described by a probability distribution. A measurement of ξ_r will give a value taken at random from the pre-existing[a] probability distribution of ξ_r; or equivalently it will give the distance from a random point in the probability distribution of x_0 to a random point in the probability distribution of x_r.

The transformation of coordinates from x_r, y_r, z_r to ξ_r, η_r, ζ_r is a change from an origin fixed in the geometrical frame to an origin with a probability scatter in that frame. It will be necessary later to make a special study of this type of transformation which, of course, is beyond the scope of the ordinary theory of coordinate transformations. In particular, we shall obtain formulae for transforming a probability distribution of physical coordinates into a probability distribution of geometrical coordinates or vice versa, and very much simpler formulae for transforming the probability distribution of the conjugate momenta (§§ 37, 38). But to carry out these transformations it is necessary to know the distribution function $f(x_0, y_0, z_0)$ of the coordinates of the physical origin. This function cannot be found observationally, because x_0, y_0, z_0 are unobservable.

The coordinates postulated in the dynamical equations of wave mechanics must be measured from a physical origin, since they and their conjugate momenta are assumed to be observables, being in fact the typical observables of quantum theory. It will be recalled that the wave-packets, whose propagation and diffusion are studied in wave mechanics, are created by our observational measurements—or more strictly by our becoming aware of the results of measurements and assessing the probability accordingly—so that it is essential to distinguish the variates in which these concentrations of probability can occur.

Thus in some, if not all, of the fundamental equations of quantum theory the coordinates are measured from a physical origin. The urgent question arises: What is this origin, and what distribution function $f(x_0, y_0, z_0)$ has been assumed for it? Writers

[a] After the measurement the information which it furnishes is used to reassess the probability. The probability distribution therefore changes discontinuously at the moment when the observer becomes aware of the result of the measurement. Attention will be paid to this point in § 35, where a very important distinction between 'structural theory' and 'predictive theory' is introduced.

on quantum theory give no hint as to the physical origin they are employing. But their equations can only be valid for some particular origin, since they are not of a form which would be invariant for arbitrary changes of f.

2. The physical origin

The centroid of a large number of particles has the important statistical property that (subject to certain conditions which are ordinarily fulfilled) the form of its probability distribution does not depend on the law of probability distribution of the individual particles. The mean of a large number of uncorrelated variates x_r has a Gaussian distribution whatever, within reason, may be the distribution law of the individual x_r.

Thus if we employ the centroid of a large number of particles as our physical origin, we have the immense advantage of starting with an *a priori* knowledge of the distribution of its geometrical coordinates x_0, y_0, z_0, complete except for the one disposable constant in the Gaussian law. The distribution function of x_0 is then

$$f(x_0) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-x_0^2/2\sigma^2}.$$

If we impose the condition that the distribution of the particles has spherical symmetry, the formula is extended to three dimensions and becomes

$$f(x_0, y_0, z_0) = (2\pi\sigma^2)^{-\frac{3}{2}} e^{-(x_0^2+y_0^2+z_0^2)/2\sigma^2}. \tag{2·1}$$

The standard deviation σ, which is left to be determined later, will be called the *uncertainty constant* of the physical reference frame.

Although the centroid is not directly indicated by a physical landmark, it is admissible as a physical origin. Formally the observational procedure would be to measure the coordinate ξ_{rs} of the rth particle from each of the other particles in turn, and take the mean $\xi_r = \overline{\xi}_{rs}$; the mean can be treated as equivalent to a single measurement from a mean particle at x_0.

Throughout this book we shall employ a physical origin related to the geometrical origin by the Gaussian distribution function (2·1), which is defined observationally as the centroid of a system of n particles with a spherically symmetrical but otherwise unrestricted probability distribution. The number n is always understood to be very large. Sometimes this assemblage will be the principal subject of investigation; but, if not, it is in the background, forming the standard environment (§ 7) of the small object-system that is being particularly studied.

The laws and constants that we derive are valid only on the understanding that the measurements concerned in them are referred to the physical frame defined by (2·1). Since writers on quantum theory leave us to guess what frame they are using, there is no guarantee that our frame will turn out to be the one in which the current quantum equations are valid. We are under no obligation to prove this identity in advance; but as a matter of practical expediency it is very desirable that the frames should agree. By making an early junction with current quantum theory we are saved an overwhelming amount of labour, because we can then take over unchanged all the specialised investigations needed to complete the application of our results to practical experiments; and it is therefore good policy to avoid unnecessary differences of form and definitions. The carrying out of this policy involves a good deal of 'intelligent

anticipation', and steps which determine the form of the theory often have to be justified from this point of view. The reader interested in logical rigour should bear in mind that the development of the theory turns partly on strict deduction and partly on ultimate saving of labour. The former part requires proof, the latter part success.

Fortunately, we can foresee that current quantum theory must be based on a physical origin which is the centroid of a large number of particles. For if it were otherwise, the equations could be of no practical use. Since x_0, y_0, z_0 are unobservable, there is no way of determining $f(x_0, y_0, z_0)$ by observation. It is only when we have theoretical information, such as that furnished by 'the law of large numbers', that we can associate a definite form of f with an observationally defined point. Thus, if quantum theory postulates a non-Gaussian form of f, it is impossible to recognise observationally the measured coordinates to which it applies, and there is no means of connecting its predictions with actual experiment. This is a *reductio ad absurdum*, because there is no doubt as to the general agreement of current quantum theory with experiment.

The foregoing may be described as the problem of 'anchoring' an ideal mathematical frame in the world of observational measurement. Anchoring is made possible by the statistical cancelling of fluctuations in large assemblages; and the small residual fluctuation that remains is necessarily Gaussian.

Starting with an abstract geometrical coordinate frame, we step over from pure geometry into physics by introducing a physical coordinate frame whose origin has the probability distribution (2·1) relative to the geometrical origin. We shall find that the standard deviation σ of this distribution *puts the scale into* the physical frame and everything constructed in the physical frame, whether it be a nucleus, an atom, a crystal or the whole extent of physical space. The main problem in this book is to investigate the way in which the extensions of these various structures are related to σ, and to evaluate the numerical ratios for some of the simpler structures.

In stepping over from the geometrical to the physical frame we appear to have freedom of choice of σ. But the freedom is illusory, because σ can only be measured in terms of the extensions of physical structures whose scale it has itself determined. To double σ would double all linear constants such as the wave-lengths of the hydrogen spectrum; thus the measure of σ in terms of the wave-length of the H_α line as unit would remain unaltered.

3. The Bernoulli fluctuation

Consider a very large number of particles N which all have the same probability distribution of coordinates. Let V_0 be a volume, fixed in the geometrical frame, extensive enough to include a large number of them. Each particle has the same probability p of being within V_0, and the mean or expectation number in V_0 is $n_0 = pN$. Let the actual number in V_0 be n, and set

$$n = n_0 + y.$$

Then, by James Bernoulli's theorem, the fluctuation y has the distribution law

$$f_N(y) = \{2\pi n_0(1 - n_0/N)\}^{-\frac{1}{2}} e^{-y^2/2n_0(1-n_0/N)}. \tag{3·11}$$

If $N/n_0 \to \infty$, this becomes

$$f_\infty(y) = (2\pi n_0)^{-\frac{1}{2}} e^{-y^2/2n_0}. \tag{3·12}$$

Both distributions are Gaussian, and their standard deviations are $(n_0 - n_0^2/N)^{\frac{1}{2}}$ and $n_0^{\frac{1}{2}}$. Hence (3·12) can be obtained by compounding with (3·11) an independent Gaussian fluctuation with standard deviation $(n_0^2/N)^{\frac{1}{2}}$ and distribution law

$$f_e(y) = (2\pi n_0^2/N)^{-\frac{1}{2}} e^{-Ny^2/2n_0^2}. \tag{3·13}$$

Let $\zeta = y/n_0$, so that

$$n = n_0(1+\zeta). \tag{3·2}$$

Then the distribution law of ζ corresponding to $f_e(y)$ is

$$g_e(\zeta) = (2\pi/N)^{-\frac{1}{2}} e^{-\frac{1}{2}N\zeta^2}. \tag{3·3}$$

We thus resolve the Bernoulli fluctuation into two independent Gaussian fluctuations, namely, an 'ordinary fluctuation' (3·12) arising from the finiteness of n_0 and an 'extraordinary fluctuation' (3·3) arising from the finiteness of N. The extraordinary fluctuation is to be combined negatively, so as to give a total fluctuation less than the ordinary fluctuation.

We shall apply this analysis to a system of particles which is in self-equilibrium, so that the probability distribution is steady. According to relativity theory the only distribution of matter which can be in self-equilibrium is a uniform distribution filling a hyperspherical space. This is the well-known 'Einstein universe'. The hyperspherical (or, as it is commonly called, spherical) space has finite volume; so that N/n_0 is finite. The infinite Euclidean space of classical theory corresponds to the limit when $N/n_0 \to \infty$ and the extraordinary fluctuation vanishes. Thus, in passing from classical to relativity theory by taking N finite, two changes are made: the space becomes curved, and an extraordinary fluctuation is introduced. These, however, are not two changes but one. *We are going to show that the space curvature is simply a way of taking into account the extraordinary fluctuation.*

Henceforth we shall deal with the extraordinary fluctuation alone. (The ordinary fluctuation, being common to relativity theory and classical theory, requires no special attention.) Denoting the particle density n/V_0 by s, the fluctuation changes an exact particle density s_0 into a slightly uncertain density

$$s = s_0(1+\zeta). \tag{3·41}$$

Instead of considering an uncertain number of particles n in a fixed volume V_0, we can consider an exact number of particles n_0 and transfer the uncertainty to the containing volume V, where $n/V_0 = n_0/V$. Setting

$$V = V_0/(1+\epsilon)^3, \tag{3·42}$$

the uncertainty is now contained in a linear scale factor $1+\epsilon$.

The distribution function $g_e(\zeta)$ can be transformed into a distribution function of ϵ. If we had to transform a distribution over discrete values of ζ into a distribution over corresponding values of ϵ, the relation would be $(1+\zeta) = (1+\epsilon)^3$. But, in transforming a continuous distribution function, discrete values are replaced by constant ranges, and we have to insert a factor proportional to $d\epsilon/d\zeta$ to transform constant ranges of ϵ into the non-constant ranges of ϵ which correspond to constant ranges of ζ. The relation is therefore

$$(1+\zeta)\,d\zeta = \text{constant} \times (1+\epsilon)^3\,d\epsilon,$$

which gives on integration

$$(1+\zeta)^2 = (1+\epsilon)^4. \tag{3·43}$$

For the distribution function $g_\epsilon(\zeta)$ the values of ζ which have sensible probability are of order $N^{-\frac{1}{2}}$, and are therefore extremely small—actually about 10^{-39}. Hence (3·43) becomes with ample approximation $\zeta = 2\epsilon$. By (3·3), the standard deviation of ζ is $N^{-\frac{1}{2}}$; hence the standard deviation of ϵ is

$$\sigma_\epsilon = 1/2\sqrt{N}. \tag{3·5}$$

The extraordinary fluctuation of the particle density can therefore be represented by a scale fluctuation with the standard deviation (3·5).

The geometrical frame is our standard of fixity when we speak of the uncertainties of physical quantities; and the ideal exact scale $\epsilon = 0$ is the scale on which the geometrical coordinates are measured. In order to take account of the extraordinary fluctuation as a scale uncertainty, we must introduce the uncertain scale $1+\epsilon$ in the system of the physical coordinates ξ, η, ζ. Considering a point distant r from the origin, the difference $x_0, y_0, z_0 = \xi - x, \eta - y, \zeta - z$ between its physical and geometrical coordinates will now consist of

(*a*) a fluctuation with standard deviation σ in all directions, due to the uncertainty of the position of the physical origin, and

(*b*) a fluctuation with standard deviation $\sigma_\epsilon r$ in the radial direction only, due to the uncertainty of the scale of measurement of r.

Remembering that the extraordinary fluctuation represented by (*b*) is to be combined negatively with other sources of fluctuation, the resultant standard deviation is

$$\text{radial } (\sigma^2 - \sigma_\epsilon^2 r^2)^{\frac{1}{2}}, \qquad \text{transverse } \sigma. \tag{3·6}$$

We shall call (3·6) the *local uncertainty* of the physical reference frame. It has been derived as a combination of the uncertainty of a distant origin with the uncertainty of scale; but it can be described more compactly as the uncertainty of a local physical origin relative to a local geometrical origin. We could, by making a local coordinate transformation, introduce 'natural coordinates' such that the local uncertainty in all directions is restored to the original value σ; these coordinates are applicable so long as the distance r from the local origin is small enough for $\sigma_\epsilon r$ to be neglected. Independently of coordinate systems, *the local uncertainty in a given direction defines an extension which might be adopted as the unit for measuring lengths in that direction in that locality.* We shall call this the σ-system of defining lengths, or briefly 'the σ-metric'.

Let ds be the length, reckoned in σ-metric, of a line-element dr, $r\,d\theta$, $r\sin\theta\,d\phi$. By (3·6) the lengths of radial and transverse elements are proportional to $dr/(\sigma^2 - \sigma_\epsilon^2 r^2)^{\frac{1}{2}}$, $r\,d\theta/\sigma$, $r\sin\theta\,d\phi/\sigma$; so that the general formula is

$$ds^2 = \frac{dr^2}{1 - (\sigma_\epsilon^2/\sigma^2)\,r^2} + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2. \tag{3·7}$$

This is the well-known formula for the line-element in spherical space of radius $R_0 = \sigma/\sigma_\epsilon$. Hence by (3·5)

$$\sigma = R_0/2\sqrt{N}. \tag{3·8}$$

It will appear in the next section that the σ-metric is the recognised metric in physical theory; so that we have in this way reached the usual description of space (occupied by a steady distribution of particles) as spherical, and have found the equation (3·8)

which determines the uncertainty constant σ in terms of the cosmological constants R_0, N.

The scale uncertainty is naturally interpreted as the result of measuring with a standard whose actual extension (in the geometrical coordinate frame) is uncertain to this extent. A standard whose uncertainty is no more than 1 part in 10^{39} is very much idealised; but there is no self-contradiction in assigning to a physical system characteristics supposed to have been measured by such a standard. If, on the other hand, the standard is supposed to be exact, it is contradictory to suppose that anything has been measured with it.

If a measured distance r has been determined with a standard which has a fluctuation σ_ϵ, the standard deviation of r cannot be less than $\sigma_\epsilon r$. We have applied the standard deviation $\sigma_\epsilon r$ negatively; this means that we have corrected the whole observed fluctuation of r for the part attributable to uncertainty of the standard. We first represent the observational measures, without any correction, in flat physical space. Recognising that the probability scatter is partly due to fluctuation of the standard employed (it being impossible to make measures at all unless the standard is inexact), we desire to eliminate this part so as to obtain what we should regard as the true distribution corresponding to an exact standard. This elimination changes the σ-metric so that the flat space is transformed into the spherical space (3·7). Thus the cosmical curvature replaces the fluctuation of the standard; and when we use the relativity representation of the universe in spherical space the scale is to be treated as exact.

This elimination of σ_ϵ is statistical, so that the curvature representation is suitable for molar physics, which is concerned with statistical averages of large numbers of particles. But a merely statistical elimination is not good enough for microscopic physics. Consequently in quantum theory we shall not use the curvature representation. We shall revert to flat space, and take account of scale fluctuation in another way (§ 24).

4. The standard of length[a]

In order to make it clear that the σ-metric is the recognised metric in relativity theory, quantum theory and practical metrology, we consider the conditions that must be fulfilled by an ultimate standard of length. The Paris metre is not accepted as an ultimate standard; the mere fact that anxiety is felt as to its constancy shows that physicists have in mind a more trustworthy standard by which it might be judged. The ultimate standard must be available at all times and places. We require a physical structure, not necessarily permanent or transportable, but constructable at any time and place from a recorded specification.

The form of the specification is decided by the condition that the definition of length (and a corresponding definition of time interval) is required at the very beginning of physics; because the definitions of other physical quantities assume that a system of space and time measurement is already in existence. It would therefore be a vicious circle to use any 'dimensional' physical quantities in specifying the standard referred to in the definition of length. The quantitative part of the specification must consist entirely of pure numbers. The specification of physical structure by pure numbers—numbers of elementary particles in configurations or states defined by quantum numbers

[a] This subject is treated at greater length in *The Philosophy of Physical Science*, pp. 70–85.

—is developed in quantum theory. Accordingly, *the standard of length must be a quantum-specified structure.*

The equations of quantum theory determine the various spatial extensions in quantum-specified systems as fixed multiples of a unit $\hbar/m_e c$. Whether or not this unit is supposed to be constant at all times and places, the ratio of two quantum-specified extensions at the same time and place is a fixed constant. Thus all quantum-specified structures give equivalent metrics, differing from one another only by a constant conversion factor.

It remains to show that the unique quantum-specified metric is the same as the σ-metric. This follows at once if we can show that any one quantum-specified extension has a mathematically calculable, and therefore fixed, ratio to the local uncertainty σ. Since the main purpose of this book is to investigate in detail the way in which the extensions of various simple structures are related to σ, ample proof will be furnished in due course. For example, we shall find that the Rydberg constant for hydrogen $\mathfrak{R}$, which is the reciprocal of a length, is given by

$$\mathfrak{R}^{-1} = \frac{16\pi\sqrt{5}}{3} . 136^2 . 137 . \sigma. \tag{4·1}$$

Thus the use of the wave-length of the H_α line as a standard of length available at all times and places is equivalent to using the σ-metric.

For molar measurement the standard is commonly embodied in a rod, which is understood to be calibrated by means of the H_α (or some other quantum-specified) wave-length. Or we may use the extension of a fixed number of lattice spaces in a specified kind of crystal at a temperature specified in some absolute way. The standard of time is likewise defined by periods of light waves or of the vibrations of a crystal. Evidently, in replacing the Paris metre by a wave-length or crystal-lattice standard, and the earth's erratic time-keeping by a quartz clock, the practical metrologist accepts the quantum-specified standard as his ideal, so that there is no difference in the accepted meaning of length and time-interval in theoretical and experimental physics.

The ratio of the wave-length to the period of H_α light is the velocity of H_α light. Thus it follows from the definition of the ultimate standards of length and time that the velocity of light is constant everywhere and everywhen. Alleged experimental evidence for a rather large change of the velocity of light in the last 70 years has been put forward. From the nature of the case there can be no such evidence; if anything is put in doubt by the experimental results, it is the agreement of the standards used by the various observers. More baleful, because it has received more credence, is the speculation of various writers that the velocity of light has changed slowly in the long periods of cosmological time, which has seriously distracted the sane development of cosmological theory. The speculation is nonsensical because a change of the velocity of light is self-contradictory.

It is perhaps not superfluous to add that no question arises as to whether the standard here defined *really* has the same length at all times and places. The question implies that there is a more ultimate standard, invested with 'reality'—whatever that may mean—which would show up the variations, if any, of the quantum-specified standard. The concept of length must be kept free from this kind of metaphysical embroidery.

Length, like other physical quantities, is a term introduced for the purpose of succinct description of observational knowledge; and, if it is defined appropriately for this purpose, no other criticism is relevant.

5. Range of nuclear forces and the recession of the galaxies

The simplest manifestation of the uncertainty of the local physical origin occurs when we consider two particles very close together, as in a nucleus or in the close encounters of two protons in scattering experiments. If ξ_r, ξ_s are physical coordinates of the two particles, their relative position is usually described by the coordinate-difference $\xi_{rs} = \xi_s - \xi_r$. But it is also possible to measure the relative coordinate directly from one particle to the other without the intermediary of an origin. The directly measured relative coordinate will be called ξ'_{rs}. Both ξ_{rs} and ξ'_{rs} are observables, and they have the same mean value; but their probability distributions are different, that of ξ_{rs} having the greater spread. Thus the wave functions associated with them, and the conjugate momenta, are different.

An observation of ξ_r gives the distance from an undetermined point in the probability distribution of the origin to an undetermined point in the probability distribution of the particle. If ξ_s is also measured, the measure has an independent starting point in the probability distribution of the origin. Thus $\xi_s - \xi_r$ will include the coordinate-difference of two random points in the distribution of the origin; this is a quantity having a Gaussian probability distribution with standard deviation $\sigma\sqrt{2}$. By making the measurements directly from one particle to the other we eliminate this source of scatter; hence, in the notation of the theory of errors,

$$\xi_{rs} = \xi'_{rs} \pm \sigma\sqrt{2}. \tag{5·1}$$

This illustrates a principle of wide importance. The description of physical systems by probability distributions requires precautions which are liable to be overlooked because they have no counterpart in the classical conception of physics from which most of our nomenclature is derived. Definitions have to be refined to take account of distinctions unprovided for in classical terminology. This applies even to the distance between two particles, where it is necessary to state explicitly which of two quantities

$$r_{12} = (\xi_{12}^2 + \eta_{12}^2 + \zeta_{12}^2)^{\frac{1}{2}} \quad \text{and} \quad r'_{12} = (\xi_{12}'^2 + \eta_{12}'^2 + \zeta_{12}'^2)^{\frac{1}{2}}$$

is meant. The difference is insignificant unless we are dealing with distances of the order of nuclear dimensions; but in the nucleus it is essential to distinguish r_{12} and r'_{12}. Thus, when a writer uses the term 'range of nuclear forces', we have to ask whether he means range in r_{12} or range in r'_{12}.

Normally the relative coordinates employed in quantum theory are ξ_{12}, η_{12}, ζ_{12}. In particular, the Coulomb energy is e^2/r_{12}. The non-Coulombian energy, however, is a singular energy associated with $r'_{12} = 0$, i.e. with actual coincidence of the particles. The whole electrical energy can therefore be expressed as $e^2/r_{12} + B\delta(r'_{12})$, where δ is Dirac's δ-function (§ 49). By (5·1) the values ξ'_{12}, η'_{12}, $\zeta'_{12} = 0$ correspond to ξ_{12}, η_{12}, $\zeta_{12} = \pm\sigma\sqrt{2}$; so that the point $r'_{12} = 0$ has a Gaussian probability distribution with standard deviation $\sigma\sqrt{2}$ over ξ_{12}, η_{12}, ζ_{12}, and $B\delta(r'_{12})$ is transformed into $Ae^{-r_{12}^2/k^2}$, where k (which is $\sqrt{2}$ times the standard deviation) is equal to 2σ. This is the form in

which the non-Coulombian energy appears in the usual equations. We call k the *range constant* of nuclear forces. By (3·8),

$$k = 2\sigma = R_0/\sqrt{N}. \tag{5·2}$$

The range is simply the effect of the uncertainty of the reference frame, which scatters the singularity $r'_{12} = 0$ into a Gaussian distribution of r_{12}.

Since the range constant has been determined experimentally, chiefly from the scattering of protons by protons, and the cosmological constants R_0, N have been determined by astronomical observation of the recession of the extra-galactic nebulae, we are able, even at this early stage, to apply an observational test to the theory. The well-known formula, first derived by Einstein in 1916, for the mass M of an Einstein universe is

$$\kappa M/c^2 = \tfrac{1}{2}\pi R_0, \tag{5·3}$$

where κ is the constant of gravitation and c the velocity of light. The number of particles (protons and electrons) being N, we have $M = \frac{1}{2}N\mu$, where μ is the mass of a hydrogen atom. Hence

$$R_0/N = \kappa\mu/\pi c^2 = 3{\cdot}95 \times 10^{-53}\,\text{cm}. \tag{5·41}$$

The experimental determination of the range constant from the scattering of protons by protons gives

$$R_0/\sqrt{N} = k = 1{\cdot}9 \times 10^{-13}\,\text{cm}. \tag{5·42}$$

From (5·41) and (5·42) we can obtain N and R_0 separately, and hence find the limiting speed of recession of the galaxies which by Lemaître's formula is $V_0 = c/R_0\sqrt{3}$. The result is $V_0 = 585$ km. per sec. per megaparsec. The actual speed should be rather less than the limiting speed, but the difference is not very important.[a] The observed value, found by Hubble and Humason, is 560 km. per sec. per megaparsec.

The observational determinations of k and V_0 do not claim high accuracy; and an agreement within 10 per cent would have been considered satisfactory. The test is therefore rather rough. But it is of particular interest because it straddles the whole range of physical systems from the nucleus to the cosmos.

Since $k = 2\sigma$, a much more accurate value of k (correct to 8 significant figures if we wish) can be obtained from (4·1). The result is $k = 1{\cdot}921 \times 10^{-13}$. This gives $V_0 = 572{\cdot}4$ km. per sec. per megaparsec.

Reversing the argument, we can deduce from the observational data that the range in r'_{12} is zero; so that non-Coulombian energy is definitely associated with a singularity of r'_{12}. Thus we need not hesitate to reject the 'meson-field' hypothesis altogether. It is in any case quite unnecessary in genuinely relativistic quantum theory. It is not an alternative way of taking into account the uncertainty of the origin, because it gives an energy distribution $Ae^{-\lambda r_{12}}$ instead of $Ae^{-r_{12}^2/k^2}$.[b]

6. Spherical space

The formula for ds in spherical space has alternative forms corresponding to different definitions of the coordinate r. The form (3·7) is obtained when we project the points

[a] From our present knowledge of the average density of matter throughout space, it is estimated that the present radius of the universe is $5R_0$. This will make the actual speed 30 km. per sec. per megaparsec less than the limiting speed V_0. (*Monthly Notices, R.A.S.* **104**, 203.)

[b] It may be expected that the shape of the non-Coulombian potential well will, at a not distant date, be determined experimentally. This will provide a crucial test between the present theory and meson-field theory.

of spherical space *orthogonally* on the tangent flat space at the origin, and take the polar coordinates in the tangent space as r, θ, ϕ. This leads to a simple graphical representation of our results.

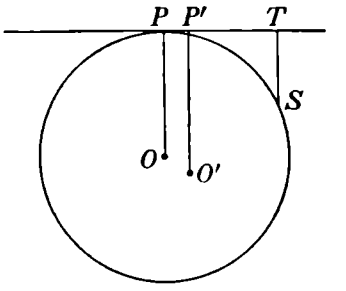
We start with a geometrical origin P and rectangular coordinates x, y, z in flat space. Let the coordinates of a particle at T be x_r, y_r, z_r. When the extraordinary fluctuation is represented by curvature, x_r, y_r, z_r are unaltered, but a fourth coordinate u_r is introduced which displaces T to a point S on the hypersphere. Transferring the origin to the centre O of the hypersphere, the equation of the hypersphere is

$$x^2+y^2+z^2+u^2 = R_0^2. \tag{6·1}$$

For a particle with uniform probability distribution over the hypersphere the mean values are

$$\overline{x_r^2} = \overline{y_r^2} = \overline{z_r^2} = \overline{u_r^2} = \tfrac{1}{4}R_0^2.$$

Thus the standard deviation of a coordinate of a particle (from its mean value 0) is $\frac{1}{2}R_0$; and the standard deviation of a coordinate of the centroid O' of the N particles is $R_0/2\sqrt{N}$, which is equal to σ by (3·8). We denote the components of OO' by x_0, y_0, z_0, u_0. Each has a Gaussian distribution with standard deviation σ.



Let P' be the orthogonal projection of O'. Since PP' is of order 10^{-13} cm., we can regard P' indifferently as a point in the tangent space or in the curved space. Its coordinates x_0, y_0, z_0 have a dispersion σ, agreeing with that of the physical origin. Thus the definition of the physical origin as the centroid of N particles is extended to particles in spherical space by simply ignoring the u coordinate. The fourth component u_0 of OO' represents the scale fluctuation of R_0. Since the radius of the hypersphere determines the linear scale of the whole universe, we naturally associate with the distinction between the geometrical origin P and the physical origin P' a distinction between the geometrical scale OP and the physical scale $O'P'$. We have

$$O'P'/OP = (R_0-u_0)/R_0.$$

Since the standard deviation of u_0/R_0 is $\sigma/R_0 = \sigma_\epsilon$, the scale fluctuation is correctly represented.

What we have here shown is that a rather naïve interpretation of the four-dimensional picture turns out to be correct. This makes four-dimensional theory rather easier than we had a right to expect. I do not think that an alternative proof of the results in § 3 can be obtained in this way. If central or stereographic projection is substituted for orthogonal projection, the standard deviation of the centroid of the projected points is not equal to $R_0/2\sqrt{N}$. It would be difficult to justify the preference for orthogonal projection without reference to the full investigation in § 3.

According to general relativity theory local irregularities of curvature are superposed on the cosmical hypersphere. It might seem that the next step in unified theory would be to derive these local curvatures from statistical fluctuations by some extension of the theory of § 3. But that would be a misunderstanding of the relation between molar and microscopic theory, and of the relation of unified theory to both. Normally the formulae of general relativity theory which covers molar physics and of quantum

theory which covers microscopic physics are not comparable. Being intended for different fields of application, they introduce different kinds of approximation; so that usually when a problem is such that the approximations of general relativity theory are valid, the approximations of quantum theory are invalid, and vice versa. To unite the two theories we have to seek out special conditions in which the approximations of both are satisfied, so that the methods of both are rigorously applicable. A uniform steady distribution, or Einstein universe, provides such a meeting point.

To show how the approximations of the two theories diverge, let us consider the most typical feature of quantum theory. *Quantisation* is a complication which arises from uniformity and symmetry; for in these conditions there is a persistence of certain dynamical integrals (usually integrals of angular momentum) which invalidates the assumption on which the practice of molar averaging is based. A slight non-uniformity is treated in quantum theory as a perturbation, which does not modify the integrals though it reduces the time that they persist. Thus as the non-uniformity increases the importance of quantisation fades away. When it appears that quantal effects are no longer important, the conditions for the usual molar averaging are satisfied; and, by a discontinuous change of method, we pass over to the representation of non-uniformity by irregular curvature.

To show more definitely the incompatibility of method, consider an atom in the slightly non-uniform environment which corresponds to an irreducible gravitational field. The non-uniformity would be treated in quantum theory as a perturbation having no effect on the eigenstates of the atom but inducing transitions between them. The eigenstates are determined by a wave equation which, when expressed in tensor form, contains the tensor $g_{\mu\nu}$. Since the eigenstates are the same with or without the non-uniformity, so also are the wave equation and the coefficients $g_{\mu\nu}$ contained in it. But this directly conflicts with general relativity theory which represents the gravitational field by modifying the $g_{\mu\nu}$.

To take account of an irreducible gravitational field in the wave equation by using the $g_{\mu\nu}$ which represent the gravitational field in molar theory would be, not a refinement, but an error.[a] The principle of equivalence does not apply. Formally this remains true for structures so extensive that the molar $g_{\mu\nu}$ differ considerably from the uniform $g_{\mu\nu}$ used in the wave equation; but, since a wide deviation implies that transitions between the eigenstates are very frequent, the wave analysis ceases to be useful. This is the fading out of quantisation already mentioned, which leaves us free to change our method and redescribe the system in terms of the non-uniform $g_{\mu\nu}$ of molar theory.

The distinction between 'special' and 'general' relativity theory is well known. In considering the connection with quantum theory, it would be useful to distinguish also an 'intermediate' relativity theory. *Special theory* is limited to flat space-time; *intermediate theory* is an extension to curved but uniform space-time; *general theory* is a further extension to non-uniform curvature. It is intermediate theory that links up with quantum theory. Since the formulae of general relativity cover intermediate relativity, they will be used from time to time in our development of quantum theory, but always in their particular application to uniform curvature.

[a] Thus attempts to 'extend Dirac's wave equation to general relativity' are misguided, but probably the intention is only to extend it to generalised coordinates in flat space by putting it into tensor form. This is a purely mathematical transformation in no way dependent on the theory of relativity.

7. Uranoids

For the purpose of investigation we divide the universe into two parts, namely, an *object-system* and its *environment*. The term 'object-system' (object-particle, object-field, etc.) is used to distinguish the part that is being intensively studied. The environment comprises everything not specifically included in the object-system, whether surrounding it or permeating it. It might alternatively be described as the 'background'.

The environment must never be left out of consideration. It would be idle to develop formulae for the behaviour of an atom in conditions which imply that the rest of the matter of the universe has been annihilated. In relativity theory we do not recognise the concept of an atom as a thing complete in itself. We can no more contemplate an atom without a physical universe to put it in than we can contemplate a mountain without a planet to stand it on.

The most elementary formulae of physics relate to very simple object-systems in very simple environments. Just as we have to begin with very simple objects—electrons, two-particle systems, etc.—so we have to begin with very simple environments—uniform, electrically neutral, etc. These simple environments will be called *uranoids*. A uranoid is an ideally simplified universe just as a geoid is an ideally simplified earth, and it is used in an analogous way.

The uranoid adopted as standard environment for our object-systems is naturally taken to be a steady uniform probability distribution of particles. This, as we have seen, constitutes an Einstein universe, and occupies a hyperspherical space. Usually it is further specialised as a 'zero-temperature uranoid' so that the particles are at almost exact rest.[a] The advantage of zero temperature is that the environment then consists of material particles only; whereas if the temperature is not zero it includes radiation. The standard uranoid is also taken to be electrically neutral; so that if a molar electromagnetic field has to be considered, it must be included in the object-system.

The *whole universe*, usually idealised as a standard uranoid, is a partner in every problem. That does not mean that we attribute to the remote environment any greater share in determining local phenomena than is ordinarily admitted in relativity theory. In particular, the most radical change in the distribution of the extra-galactic nebulae only affects small-scale systems to the extent to which it alters $g_{\mu\nu}$ in the locality considered—an effect almost entirely eliminated by a local transformation of co-ordinates. We include the whole environment in order to save the trouble of dividing it. For, if we introduce a boundary, we give ourselves the extra trouble of discovering boundary conditions which shall have the same effect as a continuation of the environment beyond the boundary.

Two lines of approach have led us to consider a system of a very large number of particles in conjunction with the small system that is being intensively studied. In § 2 it was a question of *metric*; the large system determines the uncertainty of the physical reference frame, and hence the scale of the various structures in that frame. Now it is a question of *mechanics*; the environment of the object-system is actually a vast assemblage of particles, and we have to consider the physical interaction. But these two effects are really identical. Einstein's theory, by unifying geometry and mechanics,

[a] For the significance of 'almost exact' see § 10.

unifies the metrical and mechanical effects of the environment; both are included in the description of the influence as a 'field of $g_{\mu\nu}$'. This is at the same time a metrical field and a mechanical (gravitational and inertial) field; and it makes no difference whether we regard it as influencing the measured characteristics of the object-system by physical interaction or by determining the measure system to which they are referred.

In relativity theory the field of $g_{\mu\nu}$ is generally described as the gravitational field. Sometimes, however, as a concession to Newtonian terminology, we describe it as the inertial-gravitational field, and make an artificial separation of the inertial and gravitational parts. When this distinction is made, the standard uranoid provides the inertial part of the field, and the deviation of the actual environment from the standard uranoid provides the gravitational part. 'Gravitational fields' will play a prominent part in our development of relativistic quantum theory. This seems surprising at first, since the effect of a gravitational field in the narrower sense is negligible in systems on an atomic scale. But we use the term in the broader sense which includes the inertial field. It is the inertial part that is important; for without an inertial field there is no inertia, and the masses of particles are unaccounted for.

8. The extraneous standard

We adopt a system of 'natural units' such that

$$c = 1, \quad 8\pi\kappa\hbar^2 = 1, \tag{8·1}$$

where κ is the constant of gravitation, and $\hbar$ is Planck's constant divided by 2π. By taking the velocity of light c to be unity we avoid the obsolete and troublesome distinction between mass and energy; these terms will be regarded as synonymous, and we shall use whichever is the more usual in the particular context. The second condition, as will presently be seen, simplifies the relation between the energy tensor which is the basis of relativity mechanics and the momentum vector which is the basis of quantum mechanics.

The particular relations (8·1) are chosen for convenience; but the imposition of two (and only two) fixed relations between the units of length, time and mass is an essential theoretical requirement. It is needed to remove a redundant fluidity of description of physical systems, occasioned by referrring them to three extraneous standards when one standard is sufficient to fix the scale. In modern physics fluidity of description is provided by a systematic transformation theory (tensor calculus, etc.), which is thrown into confusion if we graft on to it a merely traditional change of units.

The two relations leave one unit at our disposal. It is immaterial whether this is taken to be a length, mass, density or any other combination. If, for example, it is taken to be a length, the corresponding units of mass and time are fixed by (8·1); and every physical quantity has just one *dimension-index* which shows how it varies with the unit of length.

One extraneous standard must be retained. The internal structure of a system can be described wholly by numerical ratios; but to complete the description it is necessary to fix the scale of the system by reference to some standard outside it. If the whole universe were being investigated as one system, an outside standard would be unnecessary. But the analytical method of physics divides the universe into simple

systems of various types which are studied one by one. Each system is supposed to be surveyed from outside, so that the observer and his standard scale are alien to the system. The extraneous standard forms the link that enables us to put together in a definite scale-ratio the fragments into which our analytical method divides the universe.

We have discussed the construction of a standard of length in § 4. As a physical standard, supposed to be used in experimental measurement, it is subject to uncertainty. The linear scale-uncertainty of our adopted physical reference frame is σ_ϵ; and since all observables are understood to be referred to that frame, it is implied that they are measured with an extraneous standard having that uncertainty. That is to say, the standard constructed from the recorded specification must, when brought into use, be assigned a length, not 1 cm., but $1 \pm \sigma_\epsilon$ cm. It follows that any physical quantity of dimensions (length)^y has a scale-uncertainty $1 \pm y\sigma_\epsilon$.

We now see why it is essential to eliminate redundant units. For a mass in natural units $y = -3$; so that masses have a scale-uncertainty $1 \pm 3\sigma_\epsilon$. But if masses in the object-system were compared with a standard gram defined independently of the standard of length, we should have no means of knowing how their scale-uncertainty is related to σ_ϵ. The classical system of definitions involving three independent standards, although tolerable for exact quantities, is not sufficiently precise for the treatment of probability distributions. Unless the definitions indicate the nature of the comparisons by which the quantities are to be ascertained, they leave the kind of ambiguity illustrated in § 5 by ξ_{rs} and ξ'_{rs}. The use of natural units greatly helps to eliminate this ambiguity; for, by abolishing redundant standards, it narrows the choice of procedure of measurement.

The usual equations (with $c = 1$) for the energy tensor $T_{\mu\nu}$ in relativity theory and for the momentum vector p_μ in quantum theory are

$$-8\pi\kappa T_{\mu\nu} = G_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}G, \quad p_\mu = -i\hbar\partial/\partial x_\mu. \tag{8·2}$$

It follows that $8\pi\kappa T_{\mu\nu}$ is a spherical curvature or inverse square of a length, and has therefore the same dimensions as $(p_\mu/\hbar)^2$. Then with $8\pi\kappa\hbar^2 = 1$, $T_{\mu\nu}$ has the same dimensions as p_μ^2, or more appropriately as $p_\mu p_\nu$. Thus:

An energy tensor is, both dimensionally and tensorially, the product of two momentum vectors. (8·31)

In particular, density has the dimensions (mass)^2, so that volume has the dimensions (mass)$^{-1}$.

A particle density (number of particles, or probability of a particle, in unit volume) is a pure number multiplied by a volume-reciprocal, and therefore has the dimensions of a mass. A volume-reciprocal (three-dimensional) is a vector in four dimensions. Thus, both dimensionally and tensorially,

A particle density is a momentum vector. (8·32)

In relativity mechanics it is most convenient to use a density as the extraneous standard, so that the dimension-index of the energy tensor is 1. The dimension-indices of the principal physical quantities referred to a density standard are:

Mass, momentum, energy, $\frac{1}{2}$. Length and time, $-\frac{1}{6}$. Density and pressure, 1.
Angular momentum and action, $\frac{1}{3}$. Electric charge, $\frac{1}{6}$. (8·4)

The results (8·31) and (8·32) enter deeply into the form of the theory that we shall develop; so that in the main series of investigations we commit ourselves to natural units and to the system of dimension-indices in (8·4). Also, since $8\pi\kappa = \hbar^{-2}$, the constant of gravitation enters into the quantum formulae through $\hbar$. It is open to an investigator to choose two other relations in place of (8·1) in order to eliminate the redundant standards; I think he will find the resulting theory less simple, but it will come to the same thing in the end. The point I want to emphasise is that he will have to do a great deal of recasting of the theory to suit his units; for when once we have begun to develop the theory on the basis of (8·31), there is no going back on natural units.

9. Scale-free physics

The standard uranoid has two linear characteristics σ and R_0, which for the present we regard as independent.[a] In a considerable part of physics we are not concerned with the actual values of σ and R_0; this will be called *scale-free physics*. Scale-free physics deals with structures which can be adjusted to any scale. This is an ideal which can only be approximately realised in the actual universe; for clearly the homology must break down if the structure is magnified to a size comparable with the universe or minified to a size comparable with an atom. The exact equations of the system must involve both σ and R_0; but there are many investigations in which we can to a sufficiently high approximation put $\sigma = 0$ and $R_0 = \infty$. Thus physics splits into three branches:

(1) Scale-free physics, involving neither σ nor R_0;
(2) Cosmical physics, involving R_0 but not σ;
(3) Quantal physics, involving σ but not R_0.

There are besides certain unifying investigations which involve both σ and R_0, such as that in § 3; but these are too rare to need formal classification. A more informal nomenclature, which leaves cosmical physics out of account, distinguishes (1) and (3) as *scale-free* and *scale-fixed* physics.

Molar relativity theory is contained in (1) and (2), and classical molar theory in (1). But (1) includes also a considerable part of quantum theory. It should be remembered that Dalton's atomic theory had a long period of fertile application before any knowledge of the actual size of atoms was obtained. In the same way we shall carry out investigations which treat matter as an assemblage of large numbers of particles and involve the mass-ratios of different kinds of particles, but do not introduce any factors which fix the size or mass of an individual particle. This is all part of scale-free physics. The term 'quantal theory' is intended to be more restricted than 'quantum theory'; it refers to the more typical part in which quantisation and discrete eigenstates are introduced. Text-books on quantum theory do not confine themselves to this, and they introduce many developments which belong to scale-free theory. It is in the scale-free part of quantum theory that relativistic treatment introduces the most drastic amendments, and the most fertile advances. We shall therefore devote a great deal of attention to the scale-free part of quantum theory.

[a] They are not actually independent, because the only possible value of N is $\frac{3}{2} \times 136 \times 2^{256}$: cf. *Proc. Camb. Phil. Soc.* **40**, 37, 1943.

The criterion of a scale-free system is formally stated as follows: If we specify the characteristics of a system in terms of an extraneous standard, and consider the series of systems formed by varying the standard but keeping the specification the same, then (for a scale-free system), if one system of the series is physically possible, all are possible.

10. Pseudo-discrete states

Since quantum theory is partly scale-fixed and partly scale-free, two very different types of wave function are employed in it. Those in the scale-fixed or quantal part are discrete and self-normalising; those in the scale-free part are pseudo-discrete and are arbitrarily normalised.

The discrete self-normalised wave functions represent a particle density which rapidly decreases outwards so that the integral over space converges. They are adjusted to give a 'normalised density' $s_n(x, y, z)$ whose integral over the whole domain of x, y, z is 1. This density is said to correspond to 'unit occupation'. But we can associate with the state defined by the wave function an *occupation factor* j, and the corresponding density is then $js_n(x, y, z)$. The occupation factor gives the number of particles in the state or the probability that there is a particle in the state. In more advanced developments the occupation factor is a symbolic operator J, which only reduces to a number (an eigenvalue) when there is definitely an integral number of particles in the state.

The pseudo-discrete or scale-free wave functions are typified by the 'infinite plane waves' of elementary wave mechanics. The infinitude is not to be taken literally, but implies that the distribution extends uniformly to an undefined distance large compared with σ but small compared with R_0—anything from a millimetre to a megaparsec. It would be meaningless to specify the total number of particles in a distribution of undefined extent. We therefore choose an arbitrary normalisation density s_n to be defined as the 'density of unit occupation'; or equivalently we choose an arbitrary normalisation volume V_n, and define unit occupation to be one particle per volume V_n. This does not mean that each particle is distributed over a cell of volume V_n; that would be represented in wave mechanics by a wave packet, not by a simple plane wave. Every one of the particles has uniform probability distribution over the whole extent of the wave. *A particle occupying a pseudo-discrete wave function is an unidentified member of a large assemblage.*

The classification of wave functions as discrete and pseudo-discrete requires explanation because the usual antithesis is between 'discrete' and 'continuous'. A set of distribution functions $f_\alpha(x, y, z)$, distinguished by discrete values of a parameter α, is naturally replaced by a continuous distribution function $f(x, y, z, \alpha)$ when α becomes continuous. But we cannot do this with wave functions. They contain a phase which gives rise to the characteristic wave property of interference when two or more waves exist at the same point. Clearly a representation in which waves at the same 'point' x, y, z, α interfere is not the limiting form (for continuous α) of the representation in which waves with different values of α at the same point x, y, z interfere. Since α cannot be treated as an additional argument of the wave function, we divide the domain of α into arbitrary small ranges $\delta\alpha_r$, associating a wave function ψ_r with each range. These are the pseudo-discrete wave functions.

When the whole occupation is concentrated in a range represented by one pseudo-discrete wave function, the state of the system is said to be *almost exact.* For example, we often have to consider a system in a state of almost exact rest.

The proper mass of a particle is a fixed characteristic, but the proper density can be varied at will by varying the volume over which its probability distribution extends. Thus the density is a scale-free characteristic, and the mass a scale-fixed characteristic. More generally the energy tensor of a particle is scale-free, and the momentum vector is scale-fixed. We have, therefore, the distinction that the particles of scale-free physics are carriers of energy tensors, and the particles of quantal physics are carriers of momentum vectors.

Since relativity mechanics is based on the energy tensor, scale-free particles appear as the first step in passing from molar relativity theory to microscopic theory. By means of pseudo-discrete wave functions we can represent a molar object as a large assemblage of scale-free particles in one or more pseudo-discrete states. The molar energy tensor is thus represented as the sum of contributions $\Delta T_{\mu\nu}$ from individual particles, $\Delta T_{\mu\nu}$ being a characteristic of the pseudo-discrete state.

The individual particles are unidentified members of the whole assemblage; so that, as regards any one particle, we can know only the probability of its being in the various states. In other words the occupation factors j_r of the pseudo-discrete states can be interpreted as probabilities of an individual particle as well as frequencies in the assemblage.[a]

It might be thought that the scope of scale-free theory would be very limited, seeing that the masses of particles are not scale-free. But results usually expressed in terms of mass can often be expressed alternatively in terms of density. Thus the mass-ratio m_p/m_e of a proton and electron can equally well be described as the density-ratio of two constituents (positive and negative) of a molar distribution of hydrogen. As a density-ratio it comes within scale-free theory, and we shall evaluate it by scale-free theory in Chapter II.

It might also be thought that the 'infinite plane waves' representing uniform probability distributions, though useful for a start, would very soon have to be replaced by more complicated waves in order to represent systems of more practical interest. But the more interesting systems are not obtained that way. Atoms, etc., are constructed, not by introducing non-uniformity of distribution of electrons and nuclei, but by *correlating* their coordinates. Consider, for example, a vessel known to contain a proton and an electron, each of them equally likely to be anywhere in the vessel. In due time they will combine into a hydrogen atom, emitting a photon. The electron is still equally likely to be anywhere in the vessel; and so also is the proton. What has happened is that their probability distributions, while remaining uniform, have become correlated. Atomic wave functions, such as 'the wave function of the hydrogen atom' are *correlation wave functions.* These must not be confused with *distribution wave functions.* It is only in a few special problems (deflection of particles in a molar electro-

[a] The definition of probability in physics is always a frequency definition, so that the probability attached to an individual is in any case defined as the frequency in an *ensemble.* But the *ensemble* is commonly an imagined class of individuals or events, e.g. the imagined repetitions of a measurement. Here, however, the *ensemble* is an assemblage supposed to be actually present in the physical universe. The fact that it is not infinite like the ideal *ensemble* would be represented by an extraordinary fluctuation; but this is treated as negligible in the scale-free approximation.

magnetic field) that there can be any occasion to consider other than uniform distribution wave functions.

In later work distribution and correlation wave functions will be distinguished as external and internal wave functions. The external wave functions are pseudo-discrete, scale-free and (to avoid useless complication) uniform; the external particles or 'extracules' represented by them are unidentified members of a large assemblage. Initially the same treatment is applied to the internal particles or 'intracules', so that they also are represented by pseudo-discrete wave functions; but this is preliminary to a more highly developed treatment by quantal theory.

It will be seen that the bifurcation of quantum theory indicated by the two types of wave function has several aspects. We have still to mention the aspect which is perhaps physically the most significant. Quantisation is a distinctively *electrical* phenomenon. This becomes evident when we realise that the quantum constant $\hbar$ is simply the electrical unit e^2/c disguised by including a 'multiplicity factor' 137 (§ 33). It is quantisation that introduces the fixed scale and the discrete type of wave function; so that they appear only in the distinctively electrical part of fundamental theory. We say 'distinctively electrical', because in unified theory electrical and mechanical energy merge into one another like inertial and gravitational energy, and the adopted line of separation is flexible to some extent. Up to a certain point electrical characteristics can be represented by an energy tensor, which can be absorbed into the mechanical energy tensor. Distinctively electrical theory may be said to begin when this treatment proves inadequate and the scale-free characteristic (energy tensor) is replaced by a scale-fixed characteristic (momentum vector) related to the unit e^2/c or $\hbar$. We may therefore broadly describe scale-free theory as mechanical theory and scale-fixed theory as electrical theory—with the rider that the less distinctive parts of electrical theory can be, and commonly are, transferred to scale-free theory.

11. Stabilisation

In theoretical investigations we do not put ourselves quite in the position of an observer confronted with an object of which he has no previous knowledge. The theorist is considering, let us say, an electron with coordinates x, y, z. He recognises that knowledge of x, y, z could only be obtained by observational measurements performed on the electron, and that these measurements will create a conjugate uncertainty in the corresponding momenta. But knowledge of the mass m and charge e is on a different footing. Their values are taken from a list of physical constants. It is true that these values of m and e rest on observation, but not on observation of the object at x, y, z. As applied to that object they are *free information*—not to be paid for by a reciprocal uncertainty of the variates conjugate to them.

The theorist would explain that this free information follows from his knowledge that the object at x, y, z is an electron. But how has he got to know that it is an electron? The only answer seems to be: 'I know it is an electron because (as stated in the title of my paper) that is what the investigation is about.' The conventions of theoretical physics accept this as an admissible source of knowledge. We shall not dispute this. But it must be recognised that in this way a disjunction with purely observational physics is created.

An observer confronted with an unknown particle does not start with knowledge that it is an electron. If he measures the mass and charge he may be able to infer that it is almost certainly an electron; but meanwhile he has by his measurements created reciprocal uncertainty in the conjugate variates. A pure observer cannot obtain free information. On the other hand, an actual observer is a human being, and therefore an inveterate theorist; and he will probably proceed to steal free information by substituting the accurate tabular values of m and e for his own rougher measurements without paying the enhanced price.

We shall call a quantity whose value is given as free information a *stabilised characteristic.*

According to the technical definition of 'observable' a stabilised characteristic is not an observable. An observable is a quantity whose value or whose probability distribution is supposed to have been ascertained wholly by observations of the system that it describes; stabilisation replaces this by a quantity whose value is prescribed as one of the postulated conditions of a theoretical problem.

We have found (§ 5) that it is often necessary to refine the classical definitions intended for exact characteristics before they can be used in connection with probability distributions. Consider the momentum 4-vector p_1, p_2, p_3, p_4 and the proper mass m of a particle. It is generally understood that p_1, p_2, p_3 are directly measured; but it is not clear whether p_4 stands for a quantity supposed to be obtained by a direct measurement of an analogous kind, or whether it is to be computed from other measurements by the formula $p_4^2 = p_1^2+p_2^2+p_3^2+m^2$. Usually the latter procedure seems to be intended; we can indicate this definitely by calling p_4 the *hamiltonian.* The further question then arises: Is m intended to be an observable like p_1, p_2, p_3, or is it a stabilised characteristic of the particle? Both cases are important in physics, and they must be carefully distinguished.

By stabilising m, the probability distribution of the mechanical characteristics of the particle is reduced from a four-dimensional distribution over p_1, p_2, p_3, m (or equivalently over p_1, p_2, p_3, p_4) to a three-dimensional distribution. We shall find that the number of dimensions of the domain of probability distribution (hereafter called the *phase space*) enters as a coefficient into the leading formulae of fundamental theory, and that particles are classified primarily according to this number. A particle whose probability distribution has k dimensions will be called a V_k. Thus a simple (spinless) particle is a V_3 or a V_4 according as its proper mass is or is not stabilised. The number k is the number of degrees of freedom of a system; and stabilisation corresponds to the introduction of constraints which reduce the number of degrees of freedom.

Stabilisation is used lavishly in specifying the environment of the object-system. For there would be no gain in separating for mathematical treatment a simple object-system with only a few degrees of freedom, if we did not at the same time limit the complexity of the environment considered in conjunction with it. Hence we take the environment to be uniform, static, of zero-temperature and electrically neutral. So much restriction is imposed that only two observables R_0 and σ remain; and of these only σ is concerned in the ordinary problems of microscopic theory. When we treat the theory of a hydrogen atom in the standard uranoid, the information that the environment is uniform, static, etc., is on the same footing as the information that the object-system

is a hydrogen atom. Both are free information. We might alternatively treat an atom in a universe found observationally to agree with the standard uranoid within specified limits of uncertainty; but then it would be only logical to take an atom found observationally to be hydrogen within specified limits of uncertainty which do not exclude a chance of its being helium, lithium, etc.

Individual components of a vector or tensor cannot be stabilised without abandoning the tensor transformation properties. Stabilisation can, however, be applied to a tensor in the form of invariant conditions to be satisfied by the tensor as a whole. We can, for example, impose on a tensor of the second rank the stabilising condition that it is antisymmetrical, or that it is the outer product of two vectors, or that it is the outer square of a vector. These conditions are invariant for tensor transformations; and, since they reduce the number of independent variates required to specify the tensor, they reduce the number of dimensions of its probability distribution.

Chapter II

MULTIPLICITY FACTORS

12. Complementary fields

The object-systems studied in microscopic physics are composed primarily of particles. By gravitational and electromagnetic interaction the object-particles disturb the distribution of surrounding particles, so that the environment cannot have the simple specification postulated for the standard uranoid. We treat the disturbed environment as the sum of a uniform environment and a 'disturbance'.

This gives a threefold division of the universe into object-particles, standard uranoid and a disturbance. In theoretical physics the disturbance is tossed about like a shuttlecock between the other two. In general relativity theory it is grouped with the uranoid, and we then have a particle in a rather complex environment. The environment thus becomes the main theme of study, and the theory is described as field theory. The method of wave mechanics requires us to include the disturbance in the object-system. It is then called the *object-field*.

We have said that the environment comprises everything not specifically included in the object-system; but in practice the environment is specified first, as a standard uranoid, and everything else must be included in the object-system. In particular, all fields except the inertial field (which corresponds to the undisturbed uranoid) are to be included as object-fields. We distinguish between *extraneous* and *complementary* object-fields. An extraneous gravitational or electrical field is deliberately introduced in order to study the behaviour of the object-particles in conditions more complex than those provided by the standard environment; we are not concerned with such fields at the present stage. The complementary object-field is an inescapable adjunct of the object-particles, being the form in which the theory takes into account the readjustment of the environment due to their presence. If the object-system has electric charge the complementary electric field represents the induced charge—or rather it is the substitute that we put into the equations to compensate for neglecting induction effects in the environment. Similarly the complementary gravitational field is the compensation for neglecting the gravitational sources induced in the environment by the mass of the object-system.

The complementary field is a quantitative conception. The definition of the 'disturbance' as the residuum left by subtracting the undisturbed from the disturbed environment is made precise, if it is understood that the subtraction refers to quantitative characteristics. Thus the difference of the energy tensors of the disturbed and undisturbed environment is transferred to the object-system as the complementary field energy tensor. It follows that an object-field is specified by the same set of variates —energy, momentum, etc.—as a distribution of particles. (The other mode of description of a field, by potentials, etc., is used only when the disturbance is grouped with the uranoid, as in molar relativity theory.) The field quantities can usually be allotted between the particles, so that each object-particle has a particle energy tensor and a complementary field energy tensor. A close analogy to complementary field energy is

the effective addition to the energy of a cylinder moving through a fluid due to the motion that it sets up in the fluid.

The term 'field' is very much overworked, and it would be difficult to give a definition which does not conflict with some common usage. The sense, which it seems most important to preserve in fundamental theory, associates it with averaged characteristics, in contrast to particle characteristics which are essentially individual even if the individual is an unidentified member of an assemblage. This meaning still applies when we speak of 'the field energy of a particle', because the energy is that of the disturbance created in an environment composed of a very large number of particles whose average distribution is known.[a]

We have to consider the complementary gravitational field and the complementary electromagnetic field. Both are important, but current quantum theory ignores them altogether. Being primarily an empirical theory, it makes up for the omission by empirical factors and terms in its formulae.

Investigation of the complementary electromagnetic field is postponed; but it will be useful to state briefly its effect. Consider a proton as object-system. By electric interaction it polarises surrounding matter. In the theory of electrolytes and in astrophysics the importance of the complementary field energy (Debye-Hückel energy) is well known; but in constructing the hamiltonian of a proton in current wave mechanics, it is ignored. The proton is treated as though it could be superposed on its environment without creating an induced charge. We must conclude that *the current equations of wave mechanics postulate an unpolarisable uranoid.* This makes the field of the proton continue outwards indefinitely instead of being quenched in a short distance by an induced charge, and leads to difficulties of divergence of integrals. In replacing a natural environment by an unpolarisable uranoid, idealisation has been carried too far and essential properties of matter have been discarded. By taking account of the complementary field energy corresponding to the induced charge, we obtain an additional term in the hamiltonian which turns out to be the non-Coulombian energy. We further reformulate the basis of the theory of radiation, and avoid the divergence of the integrals which occurs in the current formulation.

In this chapter we shall deal with the complementary gravitational field. We have a tangled skein to unravel; because in this case the empirical adjustment made in current theory is, not an isolable term, but a set of factors which have been absorbed into the accepted masses of the elementary particles and other fundamental constants, and carried through into every part of microscopic physics.

13. The rigid-field convention

The divergence between the methods of relativity theory and of quantum theory is especially manifested in their treatment of the field. The root of the divergence is that wave mechanics (which is the principal analytical method of quantum theory) is

[a] Owing to the technical difficulty of solving problems relating to a system of more than two or three particles, field treatment is often used as a manageable approximation when the number of particles to be averaged is quite small. In this way we may introduce an intra-atomic field (Hartree's 'self-consistent field'). But in the general theoretical development we set aside intra-atomic fields as merely a practical dodge. Energy of interaction of object-particles on one another is part of the particle energy; it is the interaction of the object-particles with the environment that contributes the field energy.

based on the concept of a rigid field. The metrical (or gravitational) field is unconditionally assumed to be rigid; particles in quantum theory are supposed to be freely displaceable or removable without altering the $g_{\mu\nu}$ of the coordinate frame. If an extraneous electromagnetic field is present, it also is assumed to be rigid, provided that in accordance with the usual practice we exclude the 'transverse' part which is treated separately as a radiation field. In higher approximations the radiation field is not rigid; but this can scarcely be counted as an exception, because it is then treated by particle methods. The fact that wave mechanics, when forced to deal with a non-rigid field, replaces it by photons or oscillators, shows how deeply rigidity enters into its conception of a field.

A rigid field of $g_{\mu\nu}$ and κ_μ (electromagnetic potential) having been specified, wave mechanics proceeds to construct in it a skeleton frame of eigenstates, leaving itself free to decide later to what extent the eigenstates are occupied. This *flexibility of occupation* is essential to the method; for physical change is to be represented by transitions between eigenstates. There is no such flexibility in relativity theory where the $g_{\mu\nu}$ and κ_μ are conditioned by the matter present, and any change of the occupying matter involves a readjustment of the field; the change indeed has no physical significance except through its manifestation by changes of the field. This makes relativity calculations so difficult that even the problem of two bodies has not yet been solved beyond a rather crude approximation. It is no reproach to quantum theory that it parts company with relativity theory in order to find a more practicable way of solving the problems in which it specialises. The rigid-field convention is the basis of its method. Instead of a fluid frame of field variables reflecting every change of the material system investigated, we contemplate a rigid frame of eigenstates unaffected by changes (transitions) of the occupying system.

The method can only be an approximation, and if not used circumspectly it may not even be a valid first approximation. In order that it may be a first approximation, small changes in the degree of occupation of the eigenstates must produce only changes of the second order in the field. Thus:

The field must be stationary for small changes of the occupation factors of the eigenstates. (13·1)

By attending to this condition we secure an infinitesimal flexibility of occupation. Infinitesimal flexibility is not enough for some intended applications, and we shall see later how the flexibility can be extended. But we have to turn to the beginning of a theory—to its first approximations—rather than to its more elaborate adaptations to discover its fundamental concepts and definitions. It is clear that (13·1) must be incorporated in the definition of the 'particles' contemplated in wave mechanics; so that they may possess by definition the freedom of transition that the theory postulates. *Quantum particles*, i.e. the entities supposed to occupy, or to have a probability of occupying, states in a rigid field, must be distinguished from *relativity particles* which are defined as singularities in an essentially non-rigid field. We must avoid the widespread error of applying to quantum particles familiar formulae that have been developed for relativity particles. The fact that relativity particles have Lorentz-invariant properties gives no *prima facie* reason for supposing that quantum particles have Lorentz-invariant properties.

Since (13·1) is embodied in the definition of quantum particles (as occupants of states in a *rigid* field) failure of the stationary condition becomes an inconsistency. As we should usually express it, the failure implies that the field produced by the particles occupying the eigenstates is not the same as the field assumed in calculating the eigenstates. Thus the complementary field is alternatively called the self-consistent field. We have hitherto taken the rather one-sided view that the complementary field is produced by the masses of the object-particles; but the particles have these masses because they occupy states of corresponding energy in the field. In practical calculation the bringing of the field initially assumed into agreement with the field finally calculated usually involves a series of adjustments by trial and error; but happily the self-consistent gravitational field of a scale-free system is found by a very simple calculation.

14. Separation of field and particle energy

We first assume the eigenstates to be discrete. We have then a set of states ψ_r ($r = 1, 2, 3, \ldots$) to each of which is assigned an occupation factor j_r. Since changes of the system are specified by changes of the j_r, the j_r can be regarded as generalised coordinates, or preferably generalised momenta. The total energy of the system including that of the self-consistent field will be a function H^0 $(j_1, j_2, j_3, \ldots)$ of the j_r. Since the self-consistent field has to be recomputed and the eigenstates redetermined after any change of occupation, H^0 will not in general be a linear function. Let

$$E_r = \partial H^0/\partial j_r, \tag{14·1}$$

so that

$$dH^0 = \Sigma E_r dj_r. \tag{14·2}$$

And let

$$E^0 = \Sigma j_r E_r = \Sigma j_r \partial H^0/\partial j_r. \tag{14·3}$$

The energy of a particle in the state ψ_r must be E_r. For, in order that the field may be rigid, the whole change of energy dH^0 must be accounted for as a change of particle energy alone. This is expressed by (14·2), $E_r dj_r$ being the change due to the addition of the fraction dj_r of a particle of energy E_r. Then E^0 is the total particle energy of the system; and the field energy is the difference

$$W^0 = H^0 - E^0 = H^0 - \Sigma j_r \partial H^0/\partial j_r. \tag{14·4}$$

By (14·2) and (14·3)

$$dW^0 = -\Sigma j_r dE_r. \tag{14·5}$$

We note that if H^0 $(j_1, j_2, j_3, \ldots)$ is a homogeneous function of the nth degree,

$$E^0 = nH^0, \quad W^0 = (1-n)H^0. \tag{14·6}$$

For definiteness we have considered energy; but any other characteristic which is conceived as additive can be substituted. Energy-density, momentum, pressure, angular momentum, etc., must be apportioned between the object-particles in their various eigenstates and the object-field by the same formulae. This apportionment is such as to allow small changes of occupation with the E_r and W^0 constant; though for larger changes E_r and W^0 are functions of the occupation.

For molar fields (including therefore the gravitational field) the limitation to small changes dj_r is not so severe a restriction as we might at first suppose. Consider the

transfer of a small element of occupation dj from a state ψ_r to ψ_s. By (14·2) the change of energy is

$$dH^0 = dj(E_s - E_r). \tag{14·7}$$

In order to measure this observationally we must take a large number of systems, say $N = 10^{24}$, of which a smaller but still considerable number n make the transition from ψ_r to ψ_s. The particular systems that change are not identified; so that, as regards any one unselected system, the change is an increase $dj = n/N$ of the probability that it is in state ψ_s with an equal decrease of the probability that it is in state ψ_r. Accordingly the limitation is that N/n must be large.

We can therefore admit large changes including complete transitions of individual systems, on the understanding that

(1) The system is an unidentified member of a large assemblage.

(2) The great majority of the systems in the assemblage remain in the initial state, i.e. the state for which the self-consistent field was calculated.

15. Application to scale-free systems

When the foregoing treatment is applied to scale-free systems a modification is needed because the eigenstates are no longer discrete. A discrete eigenstate is specified by a set of quantum numbers, and there is no ambiguity in identifying the 'same' eigenstate ψ_r after the occupation has been changed and the eigenstates have been recalculated. Quantum numbers being no longer available we have to use some other set of characteristics for classifying the states of a scale-free system.

Let the characteristics used for classification be X_α ($\alpha = 1, 2, \ldots, n$). We shall employ characteristics which all have the same physical dimensions, and adopt an extraneous standard of the same dimensions; so that X_α has the dimension-index 1. The system being scale-free, we obtain another physically possible system by the transformation $X_\alpha \to \lambda X_\alpha$, the corresponding transformation of a characteristic Y of dimension-index l being $Y \to \lambda^l Y$. The transformation must be applied to the whole system, including the self-consistent field.

The X_α are pictured as coordinates of a point in a 'representation space' of n dimensions. An arbitrary point in representation space will not necessarily represent a possible state; indeed the nature of the system contemplated will usually be defined by a set of relations between its characteristics X_α. Thus the possible states will form a k-dimensional locus in the representation space, where usually $k < n$. We call this locus the *phase space* of the system. The number of dimensions k of the phase space will be called the *multiplicity factor*.

We denote the volume of an element of phase space by $d\tau$. This implies that a metric, giving a definite reckoning of volume, has been defined. In the phase spaces that we have occasion to employ, it is possible to develop a systematic metric based on the conception of relativistic equivalence. But for the present purpose any continuous metric will serve provided that it is scale-true; that is to say, provided it is such that the scale transformation $X_\alpha \to \lambda X_\alpha$ transforms a k-dimensional element of volume $d\tau$ into an element of volume $\lambda^k d\tau$.

The discrete occupation factors j_r are now replaced by a continuous occupation function $j(X)$ of the coordinates X_α, such that $j(X)\,d\tau$ is the collective occupation of the states compressed in the range $d\tau$. The consequent changes in the formulae of § 14

are easily found, summations being replaced by integrations, and ordinary differentiation with respect to a variate j_r by Hamiltonian differentiation[a] with respect to a function $j(X)$. Corresponding to (14·1), (14·2), (14·3), (14·5) we have

$$E = \eth H^0/\eth j, \tag{15·1}$$

$$\delta H^0 = \int E\,\delta j\,d\tau, \tag{15·2}$$

$$E^0 = \int Ej\,d\tau, \tag{15·3}$$

$$\delta W^0 = \delta H^0 - \delta E^0 = -\int j\delta(E\,d\tau). \tag{15·4}$$

Let l be the dimension-index of H^0, the dimension-index of X_α being 1 as already stated. The equations show that H^0, E, E^0, W^0 have the same dimensions. Thus $E\,d\tau$ has the dimension-index $l+k$.

Since the system is scale-free, we can take for the variation δ an infinitesimal scale variation $X_\alpha \to (1+\epsilon)X_\alpha$. Then $W^0 \to W^0(1+\epsilon)^l$, and $E\,d\tau \to E\,d\tau(1+\epsilon)^{l+k}$; so that, ϵ being infinitesimal,

$$\delta W^0 = l\epsilon W^0, \qquad \delta(E\,d\tau) = (l+k)\,\epsilon E\,d\tau.$$

Hence, by (15·4) and (15·3),

$$W^0 = -\left(\frac{l+k}{l}\right)E^0, \qquad H^0 = -\frac{k}{l}E^0. \tag{15·51}$$

By arbitrarily dividing phase space into small numbered cells $d\tau_r$, we replace the continuum of states by pseudo-discrete states (§ 10) with occupation factors $j_r = j(X)\,d\tau_r$. Comparing (14·6) and (15·51), we see that

The scale-free condition makes H^0 a homogeneous function of degree $-l/k$ of the pseudo-discrete occupation factors. (15·52)

Usually the characteristic H^0, whose partition is being considered, is itself one of the classifying characteristics X_α. Then $l = 1$; and (15·51) becomes

$$H^0 = -kE^0, \quad W^0 = -(k+1)\,E^0. \tag{15·6}$$

We have seen that the energy tensor is a scale-free characteristic. The simplest kind of scale-free particle will be characterised by an energy tensor and nothing else. The components of the energy tensor are then the classifying characteristics of the state. For such particles (15·6) gives the partition of the total energy tensor $T_{\mu\nu}$ into a particle energy tensor $E_{\mu\nu}$ and a field energy tensor $W_{\mu\nu}$, namely,

$$T_{\mu\nu} = -kE_{\mu\nu}, \quad W_{\mu\nu} = -(k+1)\,E_{\mu\nu}, \tag{15·7}$$

where the multiplicity factor k is the number of independent components of the energy tensor. We have to notice that k may be reduced by applying stabilising conditions to the energy tensor, and the partition of $T_{\mu\nu}$ between the particles and field is then altered.

To apply rigid-field treatment we choose an 'initial state', and partition the initial energy tensor $(T_{\mu\nu})_0$ into $(E_{\mu\nu})_0$ and $(W_{\mu\nu})_0$ by (15·7). For a small change of occupation $\delta T_{\mu\nu} = \delta E_{\mu\nu}$, since the field energy is stationary. Thus

$$\left.\begin{aligned}(T_{\mu\nu})_0 &= -k(E_{\mu\nu})_0 \quad \text{(initial energy)},\\ \delta T_{\mu\nu} &= \delta E_{\mu\nu} \qquad\quad \text{(transition energy)}.\end{aligned}\right\} \tag{15·8}$$

[a] *Mathematical Theory of Relativity*, § 60. If E satisfies (15·2) for small variations δ, it is by definition the Hamiltonian derivative $\eth H^0/\eth j$ of H^0.

It is rather more convenient to use as the classifying characteristics

$$X_{\mu\nu} = -T_{\mu\nu}/k. \tag{15·91}$$

Then for a state $X_{\mu\nu} = (X_{\mu\nu})_0 + \delta X_{\mu\nu}$, we have

$$E_{\mu\nu} = (X_{\mu\nu})_0 - k\delta X_{\mu\nu}. \tag{15·92}$$

Since it is used as the classifying characteristic, we call $X_{\mu\nu}$ the *generic energy tensor*. Like $T_{\mu\nu}$, to which it is constantly related, $X_{\mu\nu}$ has the ordinary Lorentz-invariant properties; that is to say a change of velocity of the system produces in $X_{\mu\nu}$ the same alteration as would be produced by an opposite change of velocity of the coordinate axes. The generic energy tensor therefore gives the *expected* particle energy—which the system would be expected to have by ordinary kinematical calculation disregarding the rigid field. The effect of rigidifying the field is that the transition energy $\delta E_{\mu\nu}$ is $-k$ times the amount expected.

Another way of putting it is that $X_{\mu\nu}$ is the energy tensor of a relativity particle and $E_{\mu\nu}$ the energy of a quantum particle. The method of relativity theory requires us to recalculate the field after any change of occupation; so that for a relativity particle the self-consistent field would be recalculated after a transition, and we should always have $X_{\mu\nu} = -T_{\mu\nu}/k$ in accordance with the formula (15·7) obtained for the state for which the self-consistent field is calculated. The purpose of the quantum method is to provide an approximation (applicable to small transitions) which avoids this continual recalculation; and the energies $E_{\mu\nu}$ of the particles which it introduces are determined accordingly. By (15·92) a relativity particle is formally a quantum particle of multiplicity -1.

16. The 'top particle'

Multiplicity factors will occur very often in these investigations, and we have to learn to handle them familiarly and confidently. I think they are made most easily comprehensible by the *principle of the top particle*.

When H^0 is a classifying characteristic, or has the same dimensions as a classifying characteristic, $l = 1$; and, by (15·52), H^0 is a homogeneous function of the occupation factors of degree $-1/k$. Thus, if we are considering an assemblage of particles in one pseudo-discrete state, H^0 varies as $j^{-1/k}$. In this case j is the whole occupation which (for pseudo-discrete wave functions) is proportional to the particle density s (§ 10); so that

$$H^0 \propto s^{-1/k}, \quad \frac{dH^0}{ds} = -\frac{1}{k}\frac{H^0}{s}. \tag{16·1}$$

The assemblage has indefinite extent; but we consider a unit volume, choosing the unit large enough for s to be very large. If one particle (per unit volume) is removed from the assemblage, the amount of H^0 removed is dH^0/ds; we therefore call dH^0/ds 'the H of the top particle'. But the average H^0 per particle (in unit volume) is H^0/s, which we call 'the H of a mean particle'. Denoting the H of a top particle by $\mathfrak{H}$, and of a mean particle by $\bar{H}$, (16·1) gives

$$\mathfrak{H} = -\bar{H}/k. \tag{16·2}$$

The chief application is to the energy tensor, where $\mathfrak{H}$ and $\bar{H}$ are respectively $E_{\mu\nu}$ and $T_{\mu\nu}$, (16·2) being a form of (15·7). Thus for an assemblage of particles in an initial pseudo-discrete state:

The particle and total energy tensors are respectively the energy tensors of a top particle and a mean particle. (16·3)

If the initial state is the state of almost exact rest, the energy tensor reduces to a single component, the density; and (16·2) gives the relation between the proper densities of top and mean particles. Since the proper masses are in the same ratio as the proper densities, the relation between the mass $\mathfrak{m}$ of a top particle and the mass $\bar{m}$ of a mean particle is

$$\mathfrak{m} = -\bar{m}/k. \tag{16·4}$$

The negative sign in (16·2) is a basal feature of rigid-field theory, which we call the *inversion of energy*. This will be dealt with later (§ 21). Here we consider separately the effect of the multiplicity factor k; so that for the present purpose the minus sign is to be omitted. We have arrived at a picture of the particles of the assemblage arranged in energy levels, which are built up successively as j increases from 0 to its actual value, the energy decreasing as the level rises. *The molar characteristics of assemblage depend on the mean particle; but since we are limited to small changes of j it is always the top particles that take part in quantum transitions.* Thus k appears as a kind of selection factor. The particles of the assemblage are unidentified; and any particle that we select is equally likely to belong to any level. But the rigid-field convention is that our selected particle (object-particle) is always a top particle, and the factor k represents the adaptation of the formulae to this selection.

Since this picture is familiar in quantum theory in other connections, it helps us to place the present series of investigations in relation to more familiar parts of quantum theory. The fact that, when a number of similar particles are present, they cannot all have the same energy but must occupy a series of levels, is usually attributed to the exclusion principle. It will be seen that although we do not explicitly use excluding particles in scale-free theory, the exclusion effect is incorporated in another way. Again, by taking the object-system to be supported, as it were, on a platform of fully packed energy levels, we obtain the rigidity of background which the quantum-theory method of treatment demands.

In this method the gravitational field energy $W_{\mu\nu}$ no longer appears explicitly; in fact *gravitation has been replaced by exclusion.* In this way the treatment of gravitation is assimilated to the general methods developed for other purposes in quantum theory. We shall follow up this treatment of gravitation later. For the present it is more useful to retain $W_{\mu\nu}$; owing to the inversion of energy, it is the sum of the energies of a mean and top particle (both reckoned as positive).

The multiplicity k decides the partition of $T_{\mu\nu}$. If k is changed by partly stabilising the energy tensor, the mean particle is unaffected; but the energy tensor and mass of a top particle are changed in inverse ratio to k. We have called a quantum particle of multiplicity k a V_k (§ 11). Since quantum particles are top particles, the relation between the masses m_1, m_2 of a V_{k_1} and V_{k_2} is

$$\frac{m_1}{m_2} = \frac{k_2}{k_1}. \tag{16·5}$$

The pseudo-discrete assemblage is the form in which the uranoid appears in scale-free physics. An assemblage in a state of almost exact rest corresponds to the zero-temperature uranoid. It is flat and of vague extent, because scale-free treatment depends on the approximations $R_0 = \infty$, $\sigma = 0$. One or more unidentified particles in it are selected to form a small object-system, and these automatically become the top particles. The rest of the assemblage forms the environment of this object-system, and is postulated to remain in the initial state of almost exact rest whilst the object particles have freedom of transition to other states. The particles of the environment or uranoid only need to be treated collectively, so that they can be taken to be mean particles.

Since k is the number of dimensions of the element $d\tau$ defining the pseudo-discrete state, any change of k must normally apply to the whole assemblage. But it makes no difference to the mean particles forming the uranoid, and it affects only the object-particles occupying the top level. As already noted, k may be described as selection factor. Accordingly, (16·5) applies to V_{k_1} and V_{k_2} particles *which have the same environment*; and m_1/m_2 is the genuine mass-ratio as it would be determined by practical measurement.

17. Standard carriers

The ordinary momentum vector and energy tensor have respectively 4 and 10 independent components. But when spin momentum is taken into account, mechanical characteristics are specified by a *complete momentum vector* with 10 independent components or by a *complete energy tensor* with 136 independent components.

These will be investigated fully in Chapters v and vii; meanwhile, the following preliminary explanation may be useful. Linear momentum is represented by a 4-vector and angular momentum by a 6-vector in four dimensions. In our later developments these will appear in combination as a $(4+6)$-vector, together with 6 additional components making 16 in all. The additional components are distinguished by the property that they change sign when a left-handed frame is substituted for a right-handed frame. This property is the criterion by which we distinguish electrical from mechanical characteristics. In the neutral unpolarisable uranoid, which is our standard environment, the electrical characteristics of a particle are dormant, having nothing to act on; and the particle is then fully specified by the complete momentum vector consisting of the 10 mechanical components. The whole set of 16 components constitutes the 'extended momentum vector'.

Corresponding to the extended vector with 16 components, the extended tensor of the second rank has 256 components. The criterion of chirality (right- or left-handedness) again distinguishes the components which are dormant in the standard achiral environment from the active components. The 10×10 combinations of two achiral suffixes and the 6×6 combinations of two chiral suffixes give 136 achiral components. These form the complete energy tensor. The 120 combinations in which one suffix is chiral and the other achiral give the dormant components of the extended energy tensor.

The term 'particle' survives in modern physics, but very little of its classical meaning remains. A particle can now best be defined as *the conceptual carrier of a set of variates*.[a] We shall frequently use the term 'carrier' as an alternative to particle.

[a] It is also conceived as *the occupant of a state* defined by the same set of variates.

The definition includes composite as well as elementary particles. We do not set any limit to the number of variates carried by one particle; but in practice the term will be used only for fairly simple combinations, the more complex carriers being referred to as systems.

We shall freely invent particles to carry the sets of variates which our form of analysis groups together. The provision of a carrier is not so much a necessity of thought as a necessity of language. It might seem desirable to distinguish the 'mathematical fictions' from the 'actual particles'; but it is difficult to find any logical basis for such a distinction. 'Discovering' a particle means observing certain effects which are accepted as proof of its existence; but it seems to be a matter of fashion or convention that one sort of effect rather than another is accepted as critical for this purpose.[a] Experimental test is concerned, not with the conceptual carrier, but with that which it carries; and there is no reason to think the carrier less fictitious when it is a proton or electron than when it is an admittedly fictitious particle. A much more important distinction is between particles resulting from the analysis of actual matter and those resulting from the analysis of matter subjected to artificial mathematical constraints; these may well be distinguished as actual and fictitious particles.

The simplest kind of carrier in scale-free physics is the carrier of a complete energy tensor and nothing more. If the energy tensor is unspecialised, i.e. not subject to stabilising conditions, this will be called a *standard carrier*. The standard carrier is a V_{136}.

An ordinary energy tensor is separated into

$$T^{\mu\nu} = \rho_0 v^\mu v^\nu + S^{\mu\nu}, \tag{17·1}$$

where ρ_0 is the proper density, v^μ the velocity vector of the external motion, and $S^{\mu\nu}$ is an internal stress system. When $S^{\mu\nu} = 0$, $T^{\mu\nu}$ reduces to the outer square of a vector,

$$T^{\mu\nu} = v^\mu \sqrt{\rho_0} \cdot v^\nu \sqrt{\rho_0}. \tag{17·2}$$

When the energy tensor of a particle has the form (17·2), we call $v^\mu \sqrt{\rho_0}$ the *root vector*. That is the strict terminology; but the root vector will more usually be called the *momentum vector*. The recognised momentum vector is not a scale-free characteristic, and does not appear until we pass over to quantum physics; but when the transition is made, the present scale-free particles will be given a fixed scale, and the root vector will then become the recognised momentum vector of the fixed-scale particle.[b] The use of this name for the root vector is therefore only a mild anticipation.

The mechanical characteristics of a classical particle are completely specified by a momentum vector; so that its energy tensor is of the restricted type (17·2). But in

[a] It is generally considered that the companion of Sirius, inferred from the elliptic motion of Sirius itself, remained hypothetical until it was actually seen in 1862; but it is difficult to see why the detection of a radiational effect rather than the detection of a gravitational effect should be regarded as the discovery. The most impressive proof of the existence of certain microscopic particles is the sharpness of their location shown in cloud-chamber tracks; but they are not classical particles, and sharpness of location is not a normal characteristic. Our attitude towards this test seems to be as arbitrary as that of the man who would not believe in the existence of the moon until he saw it eclipse the sun.

[b] This is more clearly seen by considering the reverse transition. Take a fixed-scale particle specified by a momentum vector whose magnitude is defined by reference to σ. The scale-free properties are those which survive when we cut out the connection with σ before making the approximation $\sigma = 0$. Clearly the dropping of the determinate scale will not introduce a new mechanical vector. The root vector must therefore be identified with the momentum vector relieved of its scale.

quantum theory there is no imperative reason for excluding particles (carriers) which have an internal stress-system, more especially as quantum particles admittedly contain an intrinsic angular momentum or spin. Thus we have a choice between particles characterised by an unrestricted energy tensor with 136 components, and particles characterised by a complete root vector or complete momentum vector with 10 components. The former have been called standard carriers or V_{136} particles; the latter will be called vector carriers or V_{10} particles.

Vector carriers are obtained from standard carriers by imposing on the complete energy tensor the stabilising condition that it is the outer square of a complete vector. This condition is invariant for tensor transformations and it reduces k from 136 to 10.

By further stabilisation we can obtain the spinless V_3 and V_4 particles mentioned in § 11; but there is an important difference. A symmetrical tensor of the second rank can always be resolved into a sum of outer squares of vectors; so that by representing matter as composed of V_{10} instead of V_{136} particles we do not lose generality. But by representing it as composed of V_4 or V_3 particles we lose the possibility of representing vorticity; so that the analysis refers to material subjected to a constraint which does not exist in nature. The V_4 and V_3 particles are therefore fictitiously simplified elements of physical structure. The V_{10} particles (subject to adaptations which will be found necessary when electrical theory is developed) are actual particles. So also are the V_{136} particles; but since they include an internal stress system, which in classical theory is a sign of compositeness, they are currently described as composite.

18. Mass-ratio of the proton and electron

The rigid-field treatment is simplified if we analyse the distribution in such a way that initial and transition energies are carried by different particles. In that case we distinguish:

Initial particles, for which no change of state is contemplated. They are generally taken to be in a state of almost exact rest.

Transition particles, which have many possible states. Their initial state is a state of zero energy tensor.

This kind of analysis is not a device invented for the present theory. It is the normal procedure in Newtonian mechanics and astronomy. It was followed in the earlier developments of wave mechanics; but unfortunately later writers abandoned it. The classical procedure is to replace the system by an *external particle* moving with the centre of mass and having the total mass of the system, together with *internal particles* describing orbits relatively to the centre of mass and having suitably 'reduced' masses. Since the total mass (rest energy) is carried by the external particle, the rest energies of the internal particles are zero. Thus, choosing the coordinate frame so that the centre of mass is at rest, the external particle is an initial particle and the internal particles are transition particles.

The special case of a two-particle system is important, not only in itself, but as the basis of the approximate treatment of more complex systems. In celestial mechanics the only practicable line of attack on a system of three or more bodies is by applying perturbations to two-body sub-systems; and a corresponding treatment is applied in quantum mechanics. Our analysis of a microscopic system of two particles will correspond exactly to the treatment of a double star in astronomy. If the masses of

the components are m, m', the system is transformed into an external mass M moving with the centre of gravity and an internal mass μ classifying the relative orbit, where

$$M = m+m', \quad \mu = mm'/(m+m'). \tag{18·1}$$

It is to be noticed that when the internal particle is at rest (i.e. when the relative coordinates are not changing) it adds nothing to the energy of the system; so that its rest energy is not μ but 0. It is therefore a carrier of transition energy only. The external particle carries initial energy only, provided that we do not contemplate changes of the motion of the centre of gravity of the system.

The double-star system as a whole is a carrier of variates, and can be described as a particle. In our ordinary outlook it is a composite particle—a bi-particle—and the carrier is separated into two carriers. This can be done in various ways, and the way that 'strikes the eye' is not the most suitable for celestial mechanics. Mechanical analysis separates it into an external particle carrying initial energy and an internal particle carrying transition energy. A microscopic bi-particle, e.g. a hydrogen atom, is separated in just the same way. If our study is confined to the properties of hydrogen atoms, the other method of splitting the carrier—into a proton and electron—is a confusing irrelevancy. These particles do not enter at all into the theory of the hydrogen spectrum, the magnetic moments of the atom, etc. But the transformation to proton-electron representation is needed when the results for hydrogen are made the basis for the theory of more complex systems.

Consider a standard carrier of mass m_0 in an initial state of almost exact rest, and let it make a transition to a state of momentum p'_1, p'_2, p'_3. At present it is unnecessary to consider large values of p'_α, and the energy computed in the ordinary way without regard to the rigid field is, with sufficient approximation,

$$X = m_0 + \frac{p'^2_1+p'^2_2+p'^2_3}{2m_0}. \tag{18·21}$$

This is the generic energy (§ 15). The effect of the rigid field is to multiply the expected transition energy by $-k$; and by (15·92) the particle energy is

$$E = m_0 - \frac{k(p'^2_1+p'^2_2+p'^2_3)}{2m_0} = m_0 - \frac{p'^2_1+p'^2_2+p'^2_3}{2\mu}, \tag{18·22}$$

where

$$\mu = m_0/k = m_0/136. \tag{18·23}$$

To preserve a formal analogy between rigid-field dynamics and classical dynamics we employ in quantum theory a momentum p_α which is i times the classical momentum so that

$$p_1, p_2, p_3 = ip'_1, ip'_2, ip'_3, \tag{18·31}$$

and

$$E = m_0 + \frac{p^2}{2\mu} \quad (p^2 = p_1^2+p_2^2+p_3^2). \tag{18·32}$$

We have remarked that the standard carrier will appear in our ordinary outlook as composite, since it can contain an internal stress. We now divide it into two carriers, carrying respectively the initial and transition energy, and distinguished as external and internal particles; so that (18·32) is separated into

$$E_e = m_0, \quad E_i = p^2/2\mu. \tag{18·33}$$

Further, since the purpose is to replace the composite particle by particles which conform to the ordinary conception of a simple particle and are therefore not carriers of an internal stress system, we must stabilise the two energy tensors after the separation, reducing their multiplicity from 136 to 10. The standard (V_{136}) carrier is then split into two vector (V_{10}) carriers.

In terms of wave analysis, the double wave function Ψ of the composite particle is factorised into wave functions ψ_e, ψ_i of two simple particles. Not every state of the V_{136} is a combined state of two V_{10}'s, the number of degrees of freedom 20 of the latter combination being too small. We are here considering a transition to a separable state, since otherwise the vector p_1, p_2, p_3 would not exist; but the state is pseudo-discrete, and strictly it is only the mid-point of the 136-dimensional element $d\tau$ that is a separable state by definition. Stabilisation is introduced when we replace $d\tau$ by the product $d\tau_e d\tau_i$ of two 10-dimensional elements defining pseudo-discrete states of the vector particles.

By (16·5) the reduction of k from 136 to 10 multiplies the mass m_0 of the initial particle by 136/10. A change of k does not affect the transition energy, so that E_i is unchanged. Thus the final result is

$$E_e = M, \quad E_i = p^2/2\mu, \tag{18·4}$$

where

$$M = \tfrac{136}{10} m_0, \quad \mu = \tfrac{1}{136} m_0. \tag{18·5}$$

Presumably this combination of an external and an internal particle of the simplest possible kind is realised in the hydrogen atom. Accepting this identification, which is checked in a great many ways by later investigations, the ratio

$$\eta_1 = \frac{M}{\mu} = \frac{136^2}{10} = 1849{\cdot}6 \tag{18·6}$$

is a fundamental constant of the hydrogen atom. If m_p, m_e are the masses of the proton and electron, (18·1) gives $m_p + m_e = M$, $m_p m_e = M\mu$; so that m_p, m_e are the roots of the quadratic equation

$$m^2 - mM + M\mu = 0. \tag{18·7}$$

By (18·5), this becomes

$$10m^2 - 136mm_0 + m_0^2 = 0, \tag{18·8}$$

which is the equation found by the author in 1931.[a] By solving this equation we find the mass-ratio

$$\eta_2 = \frac{m_p}{m_e} = 1847{\cdot}60. \tag{18·9}$$

The constant η_1 is determined more directly from the observational data than η_2, and would really be more suitable for observational comparisons; but since η_2 is the listed constant, we have to use it in order to give the comparison in the customary form.

We call the masses given by (18·8) the *standard masses* of the proton and electron. When we pass over to electrical theory, differently defined masses called *current masses* will be introduced which have a rather different ratio (§ 29). The molar definition of mass obviously does not extend to electrons, and the necessary extension is not by any means self-evident. In any comparison with the so-called observational value of m_p/m_e, it is necessary to scrutinise the experimental procedure and reductions actually carried out in order to ascertain the definition implicitly assumed, and modify our

[a] *Proc. Roy. Soc.* A, **134**, 529.

calculated value so as to adapt it to that definition. The observational test of η is therefore not to be undertaken lightly, and is postponed.

On account of the importance of these results, I add some remarks on points arising in the derivation. Since the molar definition of mass does not apply directly to these microscopic particles, what is the justification for using the term at all in describing them? The answer given in current theory is that a quantum mechanics can be developed which has close formal analogy with molar mechanics, and the terminology arises from this analogy. We have this in mind when we put the rigid-field hamiltonian (18·22) into the form (18·32) so as to make it formally the same as a classical hamiltonian.

It is not necessary to use the method of the top particle in determining the effect of stabilisation. We consider a distribution of s standard particles per unit volume at almost exact rest, and having a total density (including that of the self-consistent field) denoted by T_{44}. Then by (15·7) the density apportioned to the particles is

$$\rho_{136} = E_{44} = -T_{44}/136.$$

The standard particles are then replaced by external and internal particles; but the latter have no share of the initial density and can be ignored at this stage. When the external particles are stabilised, the density apportioned to the particles is changed to $\rho_{10} = E_{44} = -T_{44}/10$; so that $\rho_{10}/\rho_{136} = 136/10$. The standard particles and external vector particles correspond one to one, so that s is unchanged; and the density ratio 136/10 is also the mass-ratio μ/m_0 of the two kinds of particle. We note that only the apportionment of T_{44} is changed; there is no change of the total energy, because we are comparing different modes of description of the same physical distribution—alternative forms of analysis according as double or single wave functions are employed.

We next introduce a small transition energy with density δT_{44}. By (15·8) $\delta T_{44} = \delta E_{44}$; so that all the transition energy is allocated to the particles irrespective of the value of k. We therefore make no change in E_i when stabilisation is introduced.

The whole investigation is scale-free. For convenience we begin by assigning a mass m_0 to the standard particle; but the calculation is concerned throughout with ratios to m_0 which are determined by comparing densities.

19. Rigid coordinates

The rigid-field treatment introduces a field energy which from the ordinary point of view is extravagantly large; for example, the standard particle is associated with a field energy which is -137 times its own mass. Thus the particle is supposed to be in a field of gravitational potential far outside our ordinary experience. If the necessary field were of the irreducible type, like the gravitational field of the earth or sun, results obtained in such fantastic conditions could have no practical application. But gravitational fields can be created by a transformation of coordinates, and the intense fields postulated in rigid-field treatment must evidently be furnished in that way.

The practical physicist will wish to apply the result of the last section, let us say, to a gram of hydrogen in a litre vessel. The theorist, to avoid the complication of a boundary, has made the hydrogen continue with the same density indefinitely. Since the density is about 10^{26} times the average density of matter in the actual universe, the continuation produces a large gravitational potential in the vessel. The systematic

way of removing this potential, and so adapting the theoretical solution to the practical problem, is by a transformation of coordinates. Our procedure of separating the particle energy from the field energy is equivalent; but it is useful to examine the problem also from the point of view of coordinate transformation.

The fundamental requirement, that the field must be stationary for small changes of the occupation factors, can be stated in the more practical form: *the coordinates must be so chosen that* the field is stationary for small changes of the occupation factors. Such coordinates will be called *rigid coordinates*.

We shall determine the rigid coordinates for a uniform distribution of particles in an initial state of almost exact rest. If x, y, z, t are the rigid coordinates and x', y', z', t' the Galilean coordinates, the relation will be shown to be

$$x' = x, \quad y' = y, \quad z' = z, \quad t' = -kt. \tag{19·1}$$

Since the $g'_{\mu\nu}$ have Galilean values, we obtain

$$g_{11} = -1, \quad g_{44} = k^2, \quad \sqrt{(-g)} = -k. \tag{19·2}$$

The negative sign of $\sqrt{(-g)}$ is fixed by the consideration that in changing continuously from the accented to the unaccented system we pass through a branch-point of $(-g)^{\frac{1}{2}}$.

By general relativity theory,[a] the energy and momentum per unit coordinate mesh (which by (19·1) agrees with unit space-volume in rigid as well as in Galilean coordinates) are $\mathfrak{T}^4_\mu = T^4_\mu\sqrt{(-g)}$. Normally a pseudo-tensor-density $\mathfrak{t}^4_\mu$ has to be added, but this vanishes here because the $g_{\mu\nu}$ are constants. The transformation (19·1) gives

$$\mathfrak{T}^4_1 = \mathfrak{T}'^4_1, \quad \mathfrak{T}^4_4 = -k\mathfrak{T}'^4_4. \tag{19·3}$$

Since the space coordinates are unchanged the number of particles per unit coordinate mesh is unchanged, and the momentum and energy of a single particle transform in the same way as $\mathfrak{T}^4_\mu$. This makes it necessary to represent the momentum and energy of a particle by a *covariant* vector p_μ; for the covariant transformation law $p_\mu = p'_\alpha\,\partial x'_\alpha/\partial x_\mu$ gives

$$p_1 = p'_1, \quad p_4 = -kp'_4. \tag{19·4}$$

This is the reason why the momentum vector of a particle in wave mechanics is a covariant expression $p_\mu = -i\hbar\partial/\partial x_\mu$, whereas in relativity theory it is primarily a contravariant expression $p^\mu = m\,dx_\mu/ds$. It is this difference that makes it necessary to identify the velocity of a particle in wave mechanics with the group velocity instead of the wave velocity of the waves.

To create the artificial gravitational field we must, as it were, mistake the transformed coordinates for Galilean coordinates.[b] That is to say, we account for the difference $\mathfrak{T}^4_\mu - \mathfrak{T}'^4_\mu$, not by any peculiarity of the coordinates but by the presence of a field. Thus $\mathfrak{T}^4_\mu - \mathfrak{T}'^4_\mu$ (or $p_\mu - p'_\mu$) is the field portion $\mathfrak{W}^4_\mu$ of the tensor, and the part $\mathfrak{T}'^4_\mu$ which remains when the field is removed is the true particle portion $\mathfrak{E}^4_\mu$. Accordingly

$$\mathfrak{W}^4_\mu = \mathfrak{T}^4_\mu - \mathfrak{T}'^4_\mu, \quad \mathfrak{E}^4_\mu = \mathfrak{T}'^4_\mu. \tag{19·5}$$

Hence, by (19·3),

$$\mathfrak{W}^4_4 = -(k+1)\,\mathfrak{E}^4_4, \quad \mathfrak{T}^4_4 = -k\mathfrak{E}^4_4. \tag{19·6}$$

[a] *Mathematical Theory of Relativity*, § 59. Usually we do not distinguish between upper and lower suffixes or between tensors and tensor-densities; but occasionally (as here) the distinction is introduced.

[b] Ibid. § 16.

Dropping the distinction between upper and lower suffixes and between tensors and tensor densities, (19·6) becomes in our previous notation

$$W_{44} = -(k+1)E_{44}, \quad T_{44} = -kE_{44}.$$

Since we are considering particles at rest, all other components vanish, and (15·7) is satisfied. Thus (19·1) is the required coordinate transformation.

Thus in practice the rigid-field condition becomes:

A time t which is k^{-1} times the Galilean time t' must be employed in wave mechanics. (19·7)

It is understood that the negative sign attached to k has been otherwise dealt with; ordinarily it is got rid of, as in (18·31), by adopting in wave mechanics momenta which are i times the corresponding classical momenta.

After making our calculations by wave mechanics, we have to transform back to Galilean time for the practical application of the results. By suitably modifying the various constants—masses of the particles, magnetic moments, Coulomb energy, etc. —we can adapt the equations to Galilean coordinates, and so take a short cut. The modified constants are, of course, regarded as the 'true constants' by the empiricists; and we have to take the 'adapted wave mechanics' as furnishing the officially recognised definitions. But the fundamental relations are obscured in the adapted version; and to make new advances we must proceed with the undistorted theory. Moreover, by going the long way round, we are able to calculate the modified constants which have hitherto only been determined empirically.

We have now three ways of picturing the difference between total energy and particle energy: (*a*) by recognising the difference explicitly as field energy, (*b*) deriving the particle energy from the total energy by transforming from rigid to Galilean coordinates, (*c*) distinguishing the total energy as that of a mean particle and the particle energy as that of a top particle. The actual calculation (§§ 14, 15) has been based on (*a*). The method (*b*) lends itself to systematic development, and is our stand-by for difficult investigations. The method (*c*) replaces gravitation by exclusion, and goes a long way to meet quantum theory in its specialised form.

20. The fine-structure constant

The standard carrier is the link between molar and microscopic description, because it results from cutting up a molar energy tensor into additive elements without introducing any new properties or descriptions. Thus $k = 136$ is the basal multiplicity. When other multiplicities are introduced by stabilisation, the effect must be allowed for by factors which are usually absorbed into the masses, etc. These complicating factors are to be regarded as the reduction to the basal multiplicity 136 which allows direct passage from microscopic to molar description. The passage is essential if the theory is to have contact with experiment; because experimental data are always molar data obtained with molar measuring apparatus.

The basal relation of particle, total and field energy is then

$$T_{\mu\nu} = -136E_{\mu\nu}, \quad W_{\mu\nu} = -137E_{\mu\nu}. \tag{20·1}$$

The coefficient 137 in the second relation is the *fine-structure constant.*

The fine-structure constant is conveniently described as the ratio of two units or 'atoms' of action. Such natural units are obtained when we multiply a separable element of energy by a time intrinsically associated with it. Corresponding to an atom of particle action there must be an atom of field action 137 times as great. Two well-known atoms of action are found experimentally to be in this ratio. In radiation phenomena a constant unit is obtained by multiplying the energy of a photon by the radian-period of the corresponding light-waves; the product is $\hbar$. In particle phenomena a constant unit is obtained by multiplying the energy e^2/r of an elementary electric doublet (electron and proton or positron) by the time equivalent r/c of the separation; the product is e^2/c. The ratio $\hbar c/e^2$ of these two units is found to be 137 with an experimental accuracy of about 1 part in 10,000 (§ 32). Since the energy of the doublet is negative the ratio is strictly -137, agreeing in sign with (20·1). The negative sign cannot be evaded by considering the energy of a pair of like charges; this is not a separable energy, since the resultant electric field must extend outwards until it is quenched by induced charges in surrounding material.

The theoretical derivation of the relation $\hbar c/e^2 = 137$ will be given in § 33. Meanwhile the experimental proof that $\hbar$ and e^2 (c being as usual set equal to 1) have the required field-particle ratio clearly identifies $\hbar$ as the field unit and e^2 as the particle unit of action. There should also be a unit $136e^2$ or $\frac{136}{137}\hbar$ corresponding to total energy. The factor

$$\beta = \tfrac{137}{136} \tag{20·2}$$

will be called the *Bond factor*.[a] It may seem surprising that the field unit $\hbar$ is more prominent than the total unit $\hbar/\beta$; but this is only a matter of notation, for the factor β often occurs hidden in the empirically determined masses, etc. Quantum theory began with the study of the radiation field (Planck's law), and its early development was very much dominated by spectroscopic applications; so that the field unit became firmly established as its principal unit before the total unit had a look in.

The unit $\hbar$ is formally introduced as the coefficient in the momentum operator

$$p_\mu = -i\hbar\partial/\partial x_\mu. \tag{20·3}$$

The operator therefore refers to field momentum.[b] In our sequence this equation is not derived until § 105, though we shall sometimes use it in anticipation in order to correlate our terminology with that of current theory. The derivation in § 105 will show definitely that the operational form (20·3) is restricted to field momentum, the particle momentum in scale-fixed theory being given by a different form of symbolic operator.

The ratio $W_{\mu\nu}/T_{\mu\nu}$ is $(k+1)/k$, so that as $k \to \infty$ the field tensor and total tensor become equal. The ideal continuous matter postulated in molar relativity theory is a system with an infinite number of degrees of freedom. The conclusion that the field tensor is then the total tensor justifies the common description of molar relativity theory as 'pure field theory'. In passing over to microscopic theory, which is a mixed field and particle theory, the first step is to divide the matter into standard carriers each limited to 136 degrees of freedom; the relation $W_c = T_c$ for continuous matter is then replaced

[a] Attention was first called to it by W. N. Bond, *Nature*, **133**, 327, 1934.

[b] In the initial state of a scale-free distribution the field, total and particle momenta have a fixed ratio, and the scale can be adapted so that (20·3) represents any one of them. Thus it is sometimes unnecessary to insist on this restriction.

by $W = \beta T$. The change is in W, not T; for we contemplate microscopic and molar systems in the same environment, and must be able to replace a sufficiently large collection of particles by continuous matter without changing the total energy. Accordingly the relation is $T = T_c$, $W = \beta W_c$.[a]

In uniform distributions (infinite plane waves) $T_{\mu\nu}$ is independent of the coordinates, and we are able to employ particles which carry no coordinates since they occupy the whole undefined extent of the distribution. This leads us to regard the multiplicity of a particle as decided by the energy tensor or momentum vector alone. But uniform distribution is a special case, and the question arises whether the whole conception will not break down as soon as we admit non-uniformity. Apparently the number of degrees of freedom will then become infinite as in continuous matter. To answer this we must remember that, in counting degrees of freedom, the *potentiality* of non-uniformity is on all fours with the *existence* of non-uniformity. Thus the uniformity of space distribution of our V_{136} or V_{10} particles is a stabilised uniformity. We can in just the same way stabilise any specified non-uniformity, described, for example, by a particular eigenfunction. A distribution constrained to vary in a prescribed way has no more degrees of freedom than a distribution constrained to be uniform. Thus the analysis into V_{136} or V_{10} particles does not break down in a non-uniform distribution. The particles still 'carry no coordinates', but they carry stabilised functions of the coordinates. The variation or non-variation with coordinates is always to be treated as free information; this is evident when we consider the nature of wave analysis, for when we decide to apply Fourier analysis to phenomena we do not thereby make any assumption that an observer is called upon to verify.

We conclude that the degrees of freedom of a wave function, which determine its k, are not contained in the 'function' part of it. They depend only on the degree of constraint of the 'wave vector' part of it.

21. The inversion of energy

For all values of k the particle energy has opposite sign to the total energy. In examining the significance of this reversal of sign we shall for simplicity consider particles of multiplicity 1; so that by (15·92)

$$\delta E_{\mu\nu} = -\delta X_{\mu\nu}. \tag{21·1}$$

By inserting a particle of mass m at a point P we increase the energy in a small region containing P by the amount m. But this is not the net addition to the energy of the universe. The particle is the source of a field of gravitational potential, and every particle in the environment has a small negative energy in this field. It will not be surprising if the negative energy on integration exceeds the positive addition; so that to obtain a net increase of energy it would be necessary to add a particle of negative mass. For localised classical particles there is a sharp distinction between the immediate and the consequential effects of the addition of the particle at P; but for the particles represented by infinite plane waves, the positive and negative energies are superposed uniform distributions.

[a] The opposite assumption $W = W_c$, $T = T_c/\beta$ has certain attractions. It suggests that the process of dividing matter into standard particles *creates* the initial energy $-W/137$ which we assign to the particles. There is perhaps a legitimate way of developing this view; but it is important to show that it does not cohere with the course of development that we have been following.

Let the mass of a V_1 as ordinarily measured in a Galilean frame be m. In order that rigid-field theory may be applicable we have to add to it a field energy $-2m$, so that its total rest energy is $-m$. Instead of recognising this negative energy explicitly, we reverse (in rigid-field dynamics) the direction of positive reckoning of energy. This involves a reversal of the positive time direction, and eliminates the negative sign in the transformation (19·1). We refer to this as *the inversion of energy*.

If the particle is now given a velocity v in Galilean coordinates, the question arises whether it is the mass m or the mass $m-2m$ that is supposed to be set in motion. Since the field is rigid the field energy $-2m$ is immobile, and the velocity refers only to the particle mass m. Thus the kinetic energy is $\frac{1}{2}mv^2$; but owing to the inversion of energy this will now be reckoned as $-\frac{1}{2}mv^2$. Thus the whole energy in the rigid field is $-m+\frac{1}{2}mv^2$ in direct reckoning and $m-\frac{1}{2}mv^2$ in the adopted inverted reckoning. To avoid dynamical paradox we must write this as $m+\frac{1}{2}m(iv)^2$, thus attributing to the particle a 'velocity' iv. Correspondingly, the particle in the rigid field has a momentum which is i times the classical momentum as noted in (18·31).

This is the source of the mysterious $\sqrt{-1}$ which occurs so widely in the formulae of quantum theory. But it is to be noticed that in the problems solved by wave mechanics it is the quantum momenta that are real, and the corresponding classical momenta are accordingly imaginary.

The inversion of energy is illustrated by considering a system of gravitating particles, e.g. a star cluster. As usual the initial state is taken to be that in which the stars are nearly at rest, and the cluster is therefore widely extended. If it contracts to a new steady state, the kinetic energy K is positive; but there is also a potential energy $V=-2K$, and the whole energy T is $-K$.[a] Regarding the cluster classically, the change of particle energy is $\delta X=K$. Regarding it as a system superposed on a rigid field, the whole change of energy must be assigned to particle energy (there being no other variable energy) in accordance with (15·8); so that $\delta E=\delta T=-K$. Hence $\delta E=-\delta X$ as in (21·1).

We can introduce rigid-field stars (not identifiable individually with classical stars) to carry δE, and set $\delta E=\Sigma\frac{1}{2}mv^2$. But δE is negative, so that either m is negative or v is imaginary in the physically real cluster. This is avoided by inverting the reckoning of energy so that δE becomes positive and δX negative. Then the rigid-field stars have positive m and real v; but if we leave in the cluster one of the original classical stars its kinetic energy $\frac{1}{2}mv^2$ is now negative. Since it has positive mass, it must have imaginary v. But it is a real star with a real classical velocity. Thus the classical velocity is iv.

The analysis into steady states is responsible for this change of view. When a particle is *defined* as the occupant of a steady state, we cannot dissociate from it the energy necessary to the steadiness.[b] The particle energy has therefore to be taken to include V.

We have stated (§ 12, footnote) that energy of interaction of the object-particles with one another is part of the particle energy, the field energy being energy of interaction of the object-particles with the environment. This accords with our conclusion here that, in rigid-field representation, V (which is the *internal* potential energy of the cluster) must be counted as particle energy.

[a] These are well-known conditions for the steady state of a cluster.

[b] An *unsteady* star cluster is treated later. This throws additional light on the relation between the classical and the rigid-field treatment.

The V_1 particles may be called *semi-classical* particles, the true classical particles being formally V_{-1}. We obtain V_1 particles by stabilising the orientation but not the length m of the momentum vector. Then m is the only genuine observable and the orientation of the momentum vector, or equivalently the velocity, must be given as free information. The V_1 particles agree with classical particles in being exempt from the uncertainty principle, but differ in that they are superpositions on a rigid environment whereas classical particles disturb the environment. They are obviously artificial, and are useful only as a stepping stone to the natural quantum particles.

22. Mutual and self energy

When an object-body is observed in conjunction with a reference body, any measurement that we make determines characteristics which belong, neither to one body nor to the other, but to both jointly. It is, however, customary to allot these mutual characteristics to the object-body, or more defensibly to partition them between the two bodies according to some self-consistent scheme. This conceptual transfer, by which self properties are substituted for mutual properties, is a habit of thought which has been elevated into a convention. The language of physics is bound up with this convention, and we can scarcely do otherwise than accept it in principle.

Consider two particles, one of which is the object-particle and the other, used as a reference body, will be called the 'comparison particle'. Let the particles be simple, so that their mechanical characteristics are fully specified by complete momentum vectors p_μ, p'_μ. The mutual energy tensor is necessarily of the form

$$M_{\mu\nu} = \tfrac{1}{2}C(p_\mu p'_\nu + p'_\mu p_\nu), \tag{22·1}$$

where C is a pure number; for there is no other tensor of the second rank, having the dimensions of an energy tensor, depending symmetrically and inseparably on the mechanical characteristics of the two particles. When self energy tensors are substituted, these must for a similar reason be of the form

$$T_{\mu\nu} = Ap_\mu p_\nu, \quad T'_{\mu\nu} = A'p'_\mu p'_\nu. \tag{22·2}$$

For the moment we leave it vague whether these tensors refer to particle or total energy, since the numerical constants C, A, A' can be adapted to either definition.

For particles at almost exact rest, the three energy tensors reduce to densities

$$\rho_m = Cmm', \quad \rho = Am^2, \quad \rho' = A'm'^2, \tag{22·3}$$

where m, m' are the proper masses of the object-particle and comparison particle.

The usual practice is to allot the mutual density wholly to the object-particle. This would give $\rho = \rho_m$, or

$$Am^2 = Cmm'. \tag{22·41}$$

But this is an inconsistent procedure, since it makes no provision for the self density of the comparison particle. We must not shilly-shally with the comparison particle, assuming its presence in order that the measurement of ρ_m may be made and assuming its absence when we interpret what the measured ρ_m stands for. The correct relation $\rho + \rho' = \rho_m$ gives

$$Am^2 + A'm'^2 = Cmm'. \tag{22·42}$$

We have already found that the mass m of a simple V_{10} (proton or electron) satisfies an equation which is of this form. Comparing (18·8) and (22·42), we have

$$m' = m_0, \quad A : A' : C = 10 : 1 : 136. \tag{22·51}$$

The factors 10, 1, 136 are the multiplicity factors of the respective carriers. An unstabilised mutual energy tensor of multiplicity 136 is divided into two portions, one being the self energy tensor of object-particles of multiplicity 10, and the other the self energy tensor of comparison particles of multiplicity 1. By assigning to the comparison particle the multiplicity 1 we express the fact that it is employed as a standard —that it is, in fact, the embodiment of the extraneous standard which takes part in the ideal experiment by which ρ_m is determined. It is therefore idealised as a carrier of only the one characteristic of which it is the standard. This will be examined more fully in § 23.

As in the separation of a standard carrier into two vector carriers, the separation of the carrier of mutual energy into two carriers of self energy is a resolution of a double wave function into two simple wave functions; and it must be carried out in the rigid field postulated by wave mechanics. The energy to be partitioned is therefore total energy; and the constants in (22·42) refer to total energy. To obtain the corresponding particle energies we divide by the multiplicity factors, and the three constants then become equal. The particle energy-densities are simply

$$\rho = A'm^2, \quad \rho' = A'm'^2, \quad \rho_m = A'mm'. \tag{22·52}$$

This method evidently provides an alternative derivation of the fundamental quadratic equation (18·8) for the masses of a proton and electron. It is only necessary to justify the simple rule (22·52) for the particle energies, and then reverse the steps of the argument. We do not stop to elaborate this, since the proof in § 17 suits our sequence of development better.[a]

The same treatment gives the mass μ of an internal particle. Since the internal particle has zero rest energy, $\rho = 0$; so that $\rho' = \rho_m$. Hence, by (22·3) and (22·51),

$$A'm'^2 = C\mu m' = 136A'\mu m', \tag{22·61}$$

so that $\mu = m'/136$. This agrees with (18·5), m' having been identified with m_0.

The external V_{10} particle has a mass $M = 136m_0/10 = Cm'/A$; so that

$$AM^2 = CMm'. \tag{22·62}$$

This agrees with the crude formula (22·41) which makes no provision for the density of the comparison particle. It appears that when a hydrogen atom is analysed into an external and an internal particle the same comparison particle serves for both; but when it is analysed into a proton and electron, each has its own comparison particle. The difference is reasonable, because an actual separation of the external and internal particles is unthinkable—we cannot ionise them apart, as we can a proton and electron.

Another way of looking at it is that, since our external particles are condemned to remain always in the initial state of almost exact rest, there is not much that the experimenter can do with them. If he has knowledge of their density it is not by the standard method of determining the mutual density of an object-particle and comparison particle. He can, of course, surmount the difficulty by measuring separately

[a] A derivation by this method has been given in *Proc. Roy. Soc.* A, **174**, 16, 1940.

the densities of a proton and electron and adding them. That is in fact the way the problem is tackled when M is determined by the mass-spectrograph.

For comparison we collect the formulae showing the partition of a mutual (total) density into self densities. The more definitive notation A_0, m_0 is used instead of A', m':

$$\left.\begin{aligned} 136A_0 mm_0 &= 10A_0 m^2 + A_0 m_0^2 \quad \text{(proton or electron)},\\ 136A_0 \mu m_0 &= \qquad\quad A_0 m_0^2 \quad \text{(internal particle)},\\ 136A_0 Mm_0 &= 10A_0 M^2 \qquad \text{(external particle)}. \end{aligned}\right\} \tag{22·7}$$

It is significant that the particles which have comparison particles (as shown by the term $A_0 m_0^2$) have electrical characteristics, whereas the external particle is neutral.

The root of the matter is that probabilities are multiplicative; so that in an analysis of probability distributions a system appears as the product of its parts rather than as the sum of its parts. But we have to reconcile this with the current representation in which the characteristics of the system are the sums of the characteristics of its parts. When wave mechanics gets fairly started, this is achieved by consigning the additive characteristics to the exponents of wave functions, so that they get added when the wave functions are multiplied. But at the beginning we meet the difficulty that there can be no observable without a double probability distribution; so that, in order to suit both multiplicative and additive representations, the most elementary mechanical characteristic must be equally representable as the product and as the sum of the mechanical characteristics of two parts. It is from this condition that we obtain (22·7), which determines the relative masses of the most elementary parts so that they are consistent both with additive and with multiplicative representation.

There is a rather different way of applying (22·7) which is sometimes more illuminating. Suppose that the mutual energy tensor supplements the two self energy tensors, so that all three co-exist. If m_0 is replaced by $-m_0$, the equations (22·7) express the condition that the resultant density vanishes. The substitution of $-m_0$ turns the comparison particle into a *comparison hole*; and the masses m, m_0 are determined so that we can superpose on the uranoid a system consisting of an object-particle and its comparison hole without changing the density. Since we are considering particles at rest, no other physical characteristic is affected. Thus the change that we are describing is purely formal.

From this point of view the selection of certain particles for intensive treatment as object-particles involves a procedure which we call 'specification'. As object-particles they will have wave functions endowing them with a variety of characteristics, and in order to give them the necessary freedom constraints must be removed. Since we are here calculating rest masses, we are concerned only with the first step—the de-stabilisation, or removal of constraints. This is the process of specification; it means substituting a particle of the required multiplicity for one of the stabilised particles (comparison particles) in the uranoid; or equivalently, adding a particle of the required multiplicity together with a hole cancelling one of the uranoid particles. Equation (22·7) determines the masses so that specification does not affect the total energy tensor. This is a necessary condition, because we must be at liberty to extend or diminish the object-system by transferring particles from or to the environment; for there is no physical distinction between the particles treated intensively by wave functions and those treated collectively as an environment.

23. Comparison particles

A full account of an experimental determination of the mass of an atom would comprise a chain of investigations with the atom at one end and the standard kilogram at the other. For theoretical purposes we shorten the chain and idealise the experiment as one in which representatives of the atom and the standard kilogram directly interact. It would, of course, make nonsense of the experiment if either were unrepresented; and the theorist who forgets to put a representative of the standard into his equations is guilty of the same absent-mindedness as if he had forgotten the atom. The duty of the representative of the standard kilogram is to carry the standard unit of mass. To make the problem manageable we embody the standard in a comparison particle which can interact with the object-particles in the same way that they interact with one another.

As the carrier of the extraneous standard the comparison particle is outside the object-system. But in investigating the theory of the measurement we have temporarily to enlarge the object-system so as to bring it inside. The extended system which includes a comparison particle will be called a *perfect* object-system. The perfect object-system is self-contained in the sense that the measured characteristics attributed to it do not involve anything outside itself. The comparison particle is, however, somewhat modified, because inside the object-system there is only one variate for it to carry, namely the scale variate, whereas outside the system it is an ordinary particle carrying many variates. We stabilise all the characteristics of the internal comparison particle except the one characteristic of which it furnishes the standard.[a]

From the classical point of view it seems paradoxical that the standard of mass should be a body whose only *inexact* characteristic is its mass. But the reason is that, unless a property is inexact, it is un-get-atable by observation. By § 8 the uncertainty of the extraneous standard of mass has a fixed relation to the fundamental scale uncertainty of the physical reference frame. When we embody the scale we embody also the scale uncertainty.

Another difference between the comparison particle inside and outside the object-system, which will lead to important developments later on, may be mentioned here. The external comparison particle is a mean particle; the scale embodied in it is the scale of the physical reference frame which is based on the collective distribution of all the particles of the uranoid (§§ 2–4) and has only the minute fluctuation σ_ϵ. But the comparison particle introduced into the object-system, so as to interact with the object-particles in the same way that they interact with one another, is an individual. This will have a scale fluctuation of order $\sqrt{N}$ times the fluctuation of a mean particle, and therefore of order unity. It would seem that there is gross scale uncertainty of all characteristics of an individual atom. The conclusion does not directly affect external particles, whose masses are not determined by the use of a comparison particle but in a more indirect way (§ 22). The internal energies have in fact large scale uncertainty; but it is a feature of the method of analysis currently employed that the continuous

[a] A measured mass m is the quotient m_1/m_2 of the characteristics m_1, m_2 of the object-body and the standard. But a measured momentum vector p_α is not the quotient of two vector characteristics $(p_\alpha)_1$, $(p_\alpha)_2$ of the object-body and the standard; it is the quotient of $(p_\alpha)_1$ by the scalar characteristic m_2 of the standard. The experiment must therefore be arranged so that characteristics of the second particle other than m_2 are not allowed to affect the measurement—otherwise it is not a 'good' measurement.

probability distribution of scale is replaced by a distribution over discrete eigenscales, so that in each of the resulting eigenstates the scale is exact. The uncertainty of scale remains as an uncertainty as to the state of the system—unless we are told, as free information, that it is in a particular eigenstate.

We cannot at the same time be measuring the mass of a comparison particle and using it as a standard for measuring other masses; but we follow the ordinary convention that its mass when employed as a standard is the same as the mass which we should find by observing it as an object-particle. This convention explains an earlier result which seemed puzzling. By (22·51) the mass m' of a comparison particle is the same as the mass m_0 of a standard carrier, whereas we might have expected it to have the mass $136m_0$ of a V_1, the multiplicity assigned to it being 1. But when we express the mass of a system in terms of the mass of a standard, we refer to masses measured in the same way; so that the unit m' is meant to be the mass found for the comparison particle when it, in its turn, is being treated as an object-system. The identification $m' = m_0$ shows that the comparison particle, before it is introduced into the object-system, is a standard carrier. It is just an unspecialised element of the energy tensor of the uranoid.

In this description of the comparison particle as carrier of the standard of mass we have run ahead of scale-free physics. In the scale-free investigations of § 22 the comparison particles are carriers of the standard of density. We have found in (22·52) that the mutual density is resolved into two densities in the ratio m^2/m'^2. Since these are composed of particles whose masses have the ratio m/m', there is not a one-to-one correspondence of the object-particles and comparison particles in unit volume; that is to say, the two distributions have not the same occupation factor. The conclusion that, in the formula $Ap_\mu p_\nu$ for the energy tensor (particle energy only), A has the same value for all kinds of particles, refers to the particular state of occupation with which the investigation is concerned, and is not to be applied generally. The formula for the energy tensor corresponding to *unit occupation* is evidently

$$E_{\mu\nu} = \frac{1}{V_0}\frac{p_\mu p_\nu}{m}, \tag{23·1}$$

where V_0 is the normalisation volume, since for particles at rest this reduces to $\rho = m/V_0$.

Chapter III

INTERCHANGE

24. The phase dimension

The usual equations of wave mechanics postulate flat space. I do not think that there is anything to be gained by trying to extend wave mechanics to curved space. Curvature and wave functions are alternative ways of representing distributions of energy and momentum; and it is probably bad policy to mix them.

We have introduced the curved space of molar relativity theory as a mode of representation of the extraordinary fluctuation, and have obtained the fundamental relation (3·8) between the microscopic constant σ and the cosmological constants R_0, N. Having got what we want out of it, space curvature no longer interests us; and we return to flat space to pursue the specialised development of microscopic theory. That does not mean that henceforth we neglect curvature; we merely refrain from using the dodge that introduces it. The scale uncertainty, instead of being disguised as curvature, will be taken into account openly; so that there is no loss of rigour.

Accordingly the scale is now treated as an additional variate whose probability distribution is specified along with that of the ordinary momenta and coordinates. The variates of a probability distribution occur in conjugate pairs, and the variate conjugate to the scale will be called the *phase*. Since we have to provide for cases in which the scale reduces to an eigenvalue, the scale is classed as a momentum and the phase as a coordinate. The phase coordinate is represented as a fifth dimension normal to space-time (which is now flat), so that the scale and phase are invariant for the rotations and Lorentz transformations of special relativity theory.[a]

The scale uncertainty is primarily a fluctuation of the extraneous standard. But fluctuations of the standard are reflected in the measured characteristics of the system. The scale momentum is the measure of a characteristic which we may call the *scale-indicator*; it is itself unvarying, but its measure shows these reflected fluctuations. In the ordinary momenta the reflected fluctuations of the standard and the fluctuation of the characteristics themselves are inextricably combined; so that we have to introduce one unvarying characteristic to exhibit the scale fluctuation by itself.

We have employed a comparison particle to embody the extraneous standard, and have 'perfected' the object-system by including the comparison particle within it. The introduction of the scale and phase dimension is an equivalent way of perfecting the object-system; and the scale-indicator is the form taken by the comparison particle when it is brought into the object-system. It is a common practice to use a 6-dimensional space to represent a system of two particles. Here one of the particles is a comparison particle, and we only need to extend the object-space by one dimension. Moreover, since the object-system has always to be considered in conjunction with an extraneous standard, the extra dimension is a permanent feature of its representation.

To represent the extraordinary fluctuation or cosmical curvature the scale momentum must be given a Gaussian probability distribution with standard deviation σ_ϵ.

[a] This is not the same as the fifth dimension introduced by curvature. In § 6 the scale was represented by a distance $O'P'$ in the u direction; but distances normal to space-time now represent phase, the scale being a momentum.

For most purposes this would be a pedantic refinement; and the scale may be regarded as a stabilised characteristic. But now that each particle or small system has its own scale variate, a new field of phenomena is opened to theoretical investigation, which is suppressed in the molar treatment of scale as an averaged characteristic. As remarked in § 23 the comparison particle to be introduced into a microscopic object-system is an individual; and the fluctuation of its energy is of order 1, in contrast to the mean comparison particle whose fluctuations are of order 10^{-39}. We have therefore to distinguish two steps: the substitution of an explicit (5-dimensional) for a concealed (curvature) representation of the mean scale, and in the explicit representation the substitution of individual scales for the mean scale. Since the mean scale is practically a stabilised scale, the second step is described as *the de-stabilisation of scale.*

For some purposes it is convenient to take an angular momentum as extraneous standard, so that the scale momentum is an angular momentum and the corresponding phase coordinate is an angle.[a] This facilitates the stabilisation of scale—or rather it facilitates the de-stabilisation of the fixed scale commonly assumed. The feature of an angular coordinate is that 'infinite uncertainty' corresponds to uniform probability distribution between 0 and 2π. Thus, if J is an angular momentum and θ the corresponding angle, as the uncertainty of J diminishes θ tends to a uniform distribution over the range 2π; and we pass without discontinuity from an almost exact (observed) value to an exact (stabilised) value of J. Conversely, results which assume an exact scale are extended to a slightly fluctuating scale by spreading the distribution uniformly over a thickness 2π in an extra phase dimension. We call 2π the *widening factor.* From the widened distribution we can pass continuously to distributions in which the variation of scale becomes of serious importance.

The widening factor must be taken into account when we compare spherical space (with stabilised scale) and flat space (with fluctuating scale). When the scale is stabilised we have a spherical space whose total volume is $V = 2\pi^2 R_0^3$. Preparatory to de-stabilisation this is to be re-ordered as a volume $V_3 = \pi R_0^3$ of three-dimensional space having a thickness 2π in an extra phase dimension. Comparing it with a flat sphere of radius R_0 and volume $V_4 = \frac{4}{3}\pi R_0^3$, we have

$$V_3 = \tfrac{3}{4}V_4. \tag{24·1}$$

Since V^{-1} in natural units is a mass m, this is a relation of the form

$$m_3 = \tfrac{4}{3}m_4, \tag{24·2}$$

and is an example of the law (16·5) connecting masses of different multiplicity. In V_3 the scale is still exact and the phase necessarily has uniform distribution over the thickness 2π; the representation does not give any extra freedom. In V_4 the scale is de-stabilised and the constraint is relaxed, so that the number of degrees of freedom is raised from $k = 3$ to $k = 4$. Conversely, starting with the volume $\frac{4}{3}\pi R_0^3$ of flat space, we multiply it by a thickness 2π in the phase dimension, then multiply by $\frac{3}{4}$ to stabilise the scale since the stabilisation reduces the number of degrees of freedom from 4 to 3, and so obtain the volume $2\pi^2 R_0^3$ of scale-stabilised space which is three-dimensional but curved.

[a] The complete momentum vector contains both linear and angular momentum, so that there is no incongruity in this choice.

It is noteworthy that, by the stabilisation of scale, a flat sphere is transformed into a hypersphere of the *same* radius R_0. This could scarcely have been foreseen, since ordinarily we do not regard the two 'radii' as comparable characteristics. It is, however, in keeping with the principle of orthogonal projection in § 6.

More usually we take the phase and scale to be a linear coordinate and momentum. The angular width 2π will then be replaced by a linear width $2\pi l$. In the very simplest problems the scale and phase are separable from the other variates of the system and contribute an independent factor to the wave function. By the usual procedure in wave mechanics the distribution of the linear phase x, defined only between the limits $0 < x < 2\pi l$, is represented by a wave function periodic in $2\pi l$, and the eigenfunctions are the Fourier components $e^{inx/l}$, where n is an integer. The corresponding scale momenta are $p = n\hbar/l$. Thus:

When the scale is separable, its eigenvalues form a series proportional to the integers. (24·3)

When compared with other momenta and coordinates, the scale and phase have most affinity with the energy and time. We may, in fact, regard them as a duplicate energy and time referring to the comparison particle, and kept distinct from the energy and time of the object-particle by representation in an extra dimension. In the common representation of a two-particle system we have a similar duplication of the space coordinates and momenta, those relating to the two particles being kept distinct by using six dimensions. There is, however, a significant difference between the scale momentum and the energy, because the scale-indicator is the carrier of reflected fluctuations of the standard rather than of the standard itself. This inversion appears in various ways. It originates in the fact that the extraordinary fluctuation has to be combined negatively. We represented it at the end of § 22 by combining the object-particle with a comparison hole. It is a matter of common sense that a standard is by nature a *divisor*; so that it is the inverse of an actual standard that appears in our multiplicative representation. We shall find that, owing to this inversion, the scale and phase are space-like variates; so that they are distinguished not only in direction but also in character from the time and energy of the object-particle.

25. Interchange of suffixes

We observe only relative positions and relative velocities; consequently an observable coordinate or momentum involves two physical entities. A measurement involves four physical entities, two to furnish the observable that is said to be measured and two to furnish the comparison observable used as standard. For example, in a measurement of distance the extension between two given points is compared with the extension between two graduation marks on a scale. Disregarding the conventional allocation of the measure to two, or even to one, of the four entities concerned in it, what is measured is a characteristic of a set of four entities; we shall call this characteristic a *measurable*. A measurable has a quadruple probability distribution in the geometrical frame.

For a casual measurement the form of the probability distribution is unrestricted. But scientific investigation proceeds by systematic measurement, and the recognised physical quantities are defined by measurements of such a nature that the quadruple

probability distribution degenerates into a double or even a single distribution. One or more of the four entities is taken to be the average of a large number of particles, so that its variance is small and Gaussian; and the effect of the Gaussian fluctuation, if it is not negligible, can be allowed for by appropriate corrections. The physical origin and the mean or stabilised scale are used in this way; with the aid of one or both of them we can generally avoid handling anything more complex than a double probability distribution; but there is always a danger of losing important correlations by premature averaging, and it is occasionally necessary to refer to the primitive quadruple distribution in order to pick them up.

We denote the two object-entities by A_1, A_2 and the comparison entities by A_1', A_2'. We can replace A_1', A_2' by a mean scale; that is the simplification assumed in the method of § 18. Or we can use the physical origin for A_2 and A_2'; that is the simplification assumed in the method of § 22. Either treatment leaves us with a double probability distribution (A_1, A_2 or A_1, A_1'), since the fluctuations of the scale and of the physical origin are known and can be eliminated if necessary.

Let the four entities be a proton and electron A_1, A_2 and their respective comparison particles A_1', A_2'. We denote a measurable of the four entities by $[A_1 A_1' A_2 A_2']$. The transformation

$$[A_1 A_1' A_2 A_2'] \rightarrow [A_1 A_2' A_2 A_1'], \tag{25·1}$$

in which two perfect particles exchange comparison particles, is called *interchange*.

Since there is no recognisable distinction between different comparison particles all observable results are invariant for interchange. Interchange is a relativistic transformation from one comparison system to an equivalent comparison system, and is to be treated on just the same footing as a relativistic rotation of the space-time frame.[a]

In problems concerning two or more particles the exigencies of mathematical method compel us to attach identification suffixes to the particles. These are numbers chalked on the particles by the mathematician for his own convenience; and, since they have no physical connotation, observable results are invariant for their interchange. Invariance for interchange of suffixes is amalgamated with invariance for interchange of comparison particles by making the comparison particles the carriers of the suffixes. Thus in a perfect particle the comparison particle carries the scale-and-phase and the suffix, and the object-particle carries the other physical characteristics.

When a proton and electron are brought together into one system, so that they can be replaced by an external and an internal particle, one comparison particle is eliminated (§ 22). This is true in so far as comparison particles are carriers of the scale, for it would be redundant to have two versions of the extraneous standard in one system. But having now given them the additional duty of carrying the suffixes, we cannot drop a comparison particle so completely. We must preserve a *permutation variate*, whose changes provide the transformation (25·1) which would otherwise be lost. Thus in the fusion of two particles into one system their two scale momenta are replaced by a joint scale momentum and a permutation momentum, and their two phase coordinates by a joint phase coordinate and a permutation coordinate.

We define the *permutation coordinate* to be an angle θ such that the transformation (25·1) corresponds to $\theta \rightarrow \theta + \pi$. Since $\theta \rightarrow \theta + 2\pi$ restores the original measurable the

[a] *Proc. Roy. Soc.* A, **122**, 358, 1928. Interchange energy has since become very familiar, but its relativistic basis is not so widely realised.

uncertainty of θ is limited to a range 2π, and similar conditions apply to the permutation coordinate as to an angular phase. The widening factor 2π belonging to the phase which drops out is bequeathed to the permutation coordinate which comes in in place of it. The coordinate θ is itself unobservable; that is to say it is a cyclic coordinate which can be eliminated by ignoration of coordinates;[a] but the constant angular momentum (permutation momentum) conjugate to it gives a term in the hamiltonian which we call 'interchange energy'. It was pointed out in my first discussion of interchange in 1928 that this is the energy known to us observationally as Coulomb energy. Thus the transformation (25·1) is the beginning of electrical theory in the present sequence of development.

Let $(x, 1; x', 2)$ denote a configuration in which the particle on which we have chalked the number 1 is at x, and the particle with the number 2 is at x'. A passage to the configuration $(x, 2; x', 1)$ can occur either as the result of spatial motion or as the result of the permutation coordinate increasing by π. We limit the term 'interchange' to the latter mode of passage. According to the conceptions of the old quantum theory, this would be a *quantum jump*; but in wave mechanics all such jumps are replaced by a continuous flow of probability from one state to the other. This must be pictured as an extra-spatial circulation involving a dimension normal to space-time. Evidently the distribution of probability will not satisfy the hydrodynamical equation of continuity if we take account only of the spatial flow; and a supplementary term (Coulomb term) representing the interchange flow has to be inserted. In very simple problems, such as the metastable states of the hydrogen atom (§ 93), the combination of spatial flow and extra-spatial flow merely tilts the plane of the resultant motion into a three-dimensional space inclined to ordinary space. But usually the solution of the complete equation of continuity of flow is a difficult problem making heavy demands on our power of analysis.

The importance of a relativistic transformation applicable to the variates of an atomic system or sub-system can only be fully appreciated when we make a detailed study of the procedure of wave mechanics. In a plane of relativistic rotation, i.e. a plane in which all orientations relative to the environment are equivalent, 'symmetrical degeneracy' occurs. The mode of dealing with symmetrical degeneracy in spatial planes is well known, and this must be extended to the degeneracy in the extra-spatial plane of interchange circulation. An important point is that the conditions of quantisation are such that there are no eigenstates with zero angular momentum in a plane of degeneracy. Thus interchange circulation is not merely permissive; it is unavoidable.

The relativistic character of the interchange rotation, which gives it its importance, depends on the indistinguishability of the particles interchanged. It seems to be commonly thought that the difference of mass of the proton and electron, and possibly also the difference of charge, will prevent the application of interchange to the proton-electron system. As the theory is here stated, this objection does not arise since it is the indistinguishable comparison particles that are interchanged. But in any case the objection is ill-founded. We have two positions, two chalked numbers, two masses, two signs of charge; and we have no initial information as to which of the eight combinations is being dealt with. The association of the chalked numbers with the two

[a] In actual application the problem is complicated by the non-commutation of the various angular coordinates of a system.

positions, being relativistic, can be chosen arbitrarily; and we can proceed to develop a dynamics which will enable us to test whether the larger mass must be placed at x_1 or at x_2 in order to satisfy observational data.[a] The supposed non-interchangeability of the proton and electron is based on the mistaken assumption that we begin with *free information* as to which of them is at x_1.

This applies even in classical mechanics. Consider a double star with components so similar that the telescope does not distinguish them. We can observe only the coordinates $x(t)$, $x'(t)$ of two occupied points at various times t_1, t_2, t_3, ... The velocity of a component may be either

$$\frac{x(t_2)-x(t_1)}{t_2-t_1} \quad \text{or} \quad \frac{x'(t_2)-x(t_1)}{t_2-t_1},$$

and the acceleration will be still more ambiguous—leading to a number of possible values of the mass-ratio. It begs the question to say that in order to get the right mass-ratio we ought always to associate the accent with the component of larger mass; we do not know which has the larger mass until we know which is the right mass-ratio. In spite of the indistinguishability the data can be sorted into a consistent orbit and a unique mass-ratio determined. (The task is rather simpler for the double-star observer than for the atom observer, since the former is dealing with exact positions instead of probability distributions.) But mass is never used as a criterion of distinction, since the method is still successful when the mass-ratio turns out to be 1.

The 'indistinguishability' of particles is best understood if we think of them as carriers. It does not apply to the contents of the carriers, and it is to be noted that the contents include the mass and sign of charge as well as less permanent characteristics. It would scarcely have entered our heads to suppose that mass and charge are less amenable to interchange than other variates, were it not that we commonly substitute a stabilised mass and charge instead of the observables actually measured; but the stabilisation should obviously not be allowed to interfere with the interchange. From this point of view the same V_{10} carrier can carry contents characteristic of either a proton or an electron.

Thus no difficulty arises if we treat the permutation as interchange of the proton and electron relative to the comparison system, i.e. as $[A_1 A_1' A_2 A_2'] \to [A_2 A_1' A_1 A_2']$. But, like any other relativistic transformation, it has an alternative aspect as an equivalent transformation of the comparison system (i.e. the generalised reference frame); and that is the form in which it is stated in (25·1).

26. The two-particle transformation

For brevity we call the external and internal particles in two-particle systems *extracules* and *intracules*.

When the hydrogen atom is analysed into an extracule and intracule the theory of interchange is considerably simplified, because the interchange of the proton and electron affects only the intracule. It is therefore desirable at this stage to examine more closely the connection between the two kinds of representation of a binary system. The familiar classical formula

$$M = m+m', \quad \mu = mm'/(m+m'), \tag{26·11}$$

[a] The data will be furnished by the interaction of the two-particle system with other systems.

ought not to be taken over into quantum theory (or into relativity theory) without re-investigation. In wave mechanics it originates as follows.

Let x_α, x'_α ($\alpha = 1, 2, 3, 4$) be the space and time coordinates of two particles with mass-constants m, m'. The coordinates of the corresponding extracule and intracule are

$$X_\alpha = \frac{mx_\alpha + m'x'_\alpha}{m+m'}, \quad \xi_\alpha = x'_\alpha - x_\alpha. \tag{26·12}$$

Hence
$$\frac{\partial}{\partial x_\alpha} = \frac{m}{m+m'}\frac{\partial}{\partial X_\alpha} - \frac{\partial}{\partial \xi_\alpha}, \quad \frac{\partial}{\partial x'_\alpha} = \frac{m'}{m+m'}\frac{\partial}{\partial X_\alpha} + \frac{\partial}{\partial \xi_\alpha}, \tag{26·13}$$

and, defining M and μ by (26·11), we obtain the well-known relation

$$\frac{1}{m}\frac{\partial^2}{\partial x_\alpha \partial x_\beta} + \frac{1}{m'}\frac{\partial^2}{\partial x'_\alpha \partial x'_\beta} = \frac{1}{M}\frac{\partial^2}{\partial X_\alpha \partial X_\beta} + \frac{1}{\mu}\frac{\partial^2}{\partial \xi_\alpha \partial \xi_\beta}. \tag{26·14}$$

Introducing momenta defined by $p_\alpha = -i\hbar\partial/\partial x_\alpha$, etc., this becomes

$$\frac{p_\alpha p_\beta}{m} + \frac{p'_\alpha p'_\beta}{m'} = \frac{P_\alpha P_\beta}{M} + \frac{\varpi_\alpha \varpi_\beta}{\mu}. \tag{26·21}$$

In any volume of hydrogen there is the same number of particles of each of the four kinds, so that the occupation factors are the same for all. Thus their shares of the energy tensor are proportional to their energy tensors for unit occupation, and therefore to $p_\alpha p_\beta/m$, etc., by (23·1). It follows from (26·21) that

$$(T_{\alpha\beta})_m + (T_{\alpha\beta})_{m'} = (T_{\alpha\beta})_M + (T_{\alpha\beta})_\mu. \tag{26·22}$$

That is to say, the protons and electrons taken together yield the same energy tensor for the distribution as the extracules and intracules taken together. Since this is the crucial condition that two equivalent representations have to satisfy, the interpretation of M and μ as mass-constants of the extracule and intracule is verified.

Regarding the two-particle system as a bi-particle with 8 coordinates x_α, x'_α, the foregoing is a transformation of coordinates in the 8-space to another orthogonal system X_α, ξ_α. From (26·12),

$$\frac{\partial(X_\alpha, \xi_\alpha)}{\partial(x_\alpha, x'_\alpha)} = 1, \tag{26·31}$$

so that volumes are unchanged. This applies whether we include or exclude the time coordinates; so that for corresponding volumes of the 6-space

$$\int dx_1 dx_2 dx_3 dx'_1 dx'_2 dx'_3 = \int dX_1 dX_2 dX_3 d\xi_1 d\xi_2 d\xi_3, \tag{26·32}$$

and for the 2-time,
$$\int dt\, dt' = \int dT\, d\tau. \tag{26·33}$$

Also for corresponding ranges of momentum,

$$\int dp_1 dp_2 dp_3 dp'_1 dp'_2 dp'_3 = \int dP_1 dP_2 dP_3 d\varpi_1 d\varpi_2 d\varpi_3. \tag{26·34}$$

We notice also that
$$mm' = M\mu. \tag{26·35}$$

Since volumes in the 6-space are unchanged, the density of the probability distribution of the bi-particle in the 6-space is unchanged. The probability density is expressed as the product of two complex conjugate double-wave functions Ψ, $\Psi^\dagger$, which are accordingly unchanged in value. (The change of their arguments will, of course, give

them a different functional form.) This makes it possible to adopt Ψ as the operand, previously unstated but assumed to be invariant, of the differential operators that have been introduced.

When Ψ is pseudo-discrete, it is necessary to define normalisation volumes. These must be related so as to give a one-to-one correspondence of the four kinds of particles. Three of the particles are located in ordinary x-space, and these must be assigned equal three-dimensional normalisation volumes $V_m = V_{m'} = V_M$. But the coordinates ξ_r of the intracule locate it in a relative space (ξ-space), and its normalisation volume must be delimited in that space. For unit occupation there is one bi-particle per volume $V_m V_{m'}$ of 6-space; and since the probability density is unchanged by the transformation, we have $V_M V_\mu = V_m V_{m'}$. Thus the four normalisation volumes are equal, notwithstanding that one of them is in a different space.

It follows that the density of intracules in ξ-space is equal to the density of extracules in x-space—a result which was by no means self-evident.

It is to be noticed that the energy tensor of the intracule consists of density, pressure, etc., in ξ-space; and there is no immediate justification for adding it to energy tensors in ordinary space. The justification lies in equation (26·22). The transformation has in fact been specially chosen so as (exceptionally) to make the addition legitimate.

Setting $p^2 = p_1^2+p_2^2+p_3^2$, etc., (26·21) gives

$$p^2/m + p'^2/m' = P^2/M + \varpi^2/\mu. \tag{26·41}$$

Introducing hamiltonians

$$\left.\begin{aligned} h &= m + p^2/2m, & H &= M + P^2/2M, \\ h' &= m' + p'^2/2m', & \eta_0 &= \varpi^2/2\mu, \end{aligned}\right\} \tag{26·42}$$

we have, by (26·41) and (26·11),

$$h + h' = H + \eta_0. \tag{26·43}$$

We should like to make the theory symmetrical by giving the intracule a hamiltonian

$$\eta = \mu + \varpi^2/2\mu. \tag{26·44}$$

But $H + \eta \neq h + h'$. We can, however, consider a transformation in which $h + h' \to H + \eta$. That means that when we transform from proton-electron to extracule-intracule representation we change the zero-level of energy reckoning so as to add the required density to the distribution. This transformation (described as 'freeing the intracule') will play an important part in the theory.

27. Hydrocules

In a hamiltonian of the form

$$E = \mu_1 + \frac{p^2}{2\mu_2}, \tag{27·1}$$

we distinguish μ_1 as the rest mass and μ_2 as the mass-constant of the particle. In a pure inertial field the two constants agree. When they differ it is because the particle possesses potential energy in a gravitational or electrical field, and this constitutes the difference $\mu_1 - \mu_2$. The characteristic mass of a particle, which is listed in tables of physical constants, does not include these casual additions, and refers to the mass-constant. When

$\mu_1 = \mu_2$ the particle is said to be *free*; when $\mu_1 = 0$ it is *bound*. In classical theory an intracule is a bound particle.

The principal investigations of quantal (scale-fixed) theory postulate free intracules. In particular, Dirac's wave equation of the hydrogen atom is the equation of a free intracule. We have therefore to connect one important series of investigations in which the intracule is assumed to be bound with another important series in which it is assumed to be free. Evidently in passing from one to the other the zero level, from which the energy of the atom is reckoned, is changed by the amount μ required to free the intracule. Such changes of zero are made light-heartedly in elementary physics; but in relativity theory we have to take into account various consequential changes, because the energy determines the gravitational field and hence the metric.

Since the double-wave function of the V_{136} will have to be re-analysed into simple wave functions of an extracule and a free intracule, the energy must be added in (18·32), before stabilisation is applied; so that (18·33) is replaced by

$$E_e = m_0, \quad E_i = \mu + p^2/2\mu. \tag{27·2}$$

These must continue to be the initial and transition energies, for otherwise the purpose of the analysis would be frustrated. We are limited to small transitions, i.e. the transitions (which considered individually may be large) are limited to a few particles in a large assemblage, the rest of the assemblage remaining in the initial state. Thus in practice 'initial state' means the state of the postulated environment or uranoid; and (27·2) refers to an object-particle in a uranoid consisting of particles of energy m_0. These are not standard particles, because, owing to the change of zero level, the rest energy of a standard particle is now

$$m_0 + \mu = \tfrac{137}{136} m_0 = \beta m_0. \tag{27·3}$$

We shall call these particles *hydrocules*.

We have therefore to recognise two systems of reckoning energy (used in different parts of quantum theory and commonly confused) which will be distinguished as systems A and B. The primary difference is:

System A. Standard particle uranoid; bound intracules.

System B. Hydrocule uranoid; free intracules.

We turn now to the consequential effects of the change. In substituting hydrocules for standard particles we lose $\frac{1}{137}$ of the mass. A change of density of a large assemblage seriously upsets its gravitational equilibrium. According to relativity theory only one density of a steady distribution of particles at rest is compatible with the constant of gravitation κ, namely the density of a pressureless Einstein universe. There are various ways in which the hydrocule uranoid might be adjusted to equilibrium. The particles might be re-spaced; or the constant κ (and consequently $\hbar$) might be changed. But the simplest way is to couple with the change from A to B a change of extraneous standard. The standard is changed by an amount such that

In passing from A to B, measured densities are multiplied by β. (27·4)

We keep the particle density s of the two uranoids the same. Then in substituting hydrocules for standard particles the mass-density is divided by β; but the change of standard restores it to its original value, so that the uranoid is still compatible with the

original constant κ and no other. Since our units are such that $8\pi\kappa\hbar^2 = 1$, $\hbar$ also is unchanged. Accordingly

In passing from A to B, the constants s, κ and $\hbar$ are unchanged. (27·5)

Since the extraneous standard is furnished ultimately by the uranoid, the coupling of a change of standard with a change of uranoid is a natural linkage which, in fact, makes the whole transformation relativistic. The effect of the change of standard on various measured quantities is shown in (8·4). In particular

In passing from A to B, lengths and times are multiplied by $\beta^{-\frac{1}{2}}$. (27·6)

Care must be taken only to apply (27·6) to quantities that are directly measured. The scale-free particles that we have been considering are carriers of density; they are not carriers of mass, and the masses attributed to them cannot be supposed to be directly measured. The quantities m_0, M, μ are densities prematurely converted into masses by a factor left to be decided when we pass over into scale-fixed theory. The factor cannot be subjected to pre-natal transformations—or at least it would be a meaningless complication to introduce them—and it is therefore treated as a fixed constant which, like $\hbar$ and κ, is common to systems A and B. It is for this reason that we keep the particle density s unchanged. The masses m_0, M, μ accordingly transform as densities. In particular, the rest masses m_0/β, m_0 of the hydrocule and standard particle in system A become m_0, $m_0+\mu$ in system B.

The standard particle has been identified with the hydrogen atom, a standard particle at rest being a hydrogen atom at rest both externally and internally. In system B its rest mass is made up of an extracule of mass m_0 and an intracule of mass μ. It is worth while to restate the argument of § 18 in a form applying to this combination. So far as external motion (motion of the centre of mass) is concerned the atom is represented by its extracule. If the extracule has a classical momentum p' the additional energy is, by the usual classical formula, $\delta X = p'^2/2m_0$.[a] If the classical motion is replaced by transition in a rigid field the additional energy is

$$\delta E = -136\delta X = -p'^2/2\mu = p^2/2\mu.$$

The classical motion and transition motion are distinguished by real and imaginary values of p'; the imaginary p' is, of course, not equivalent to a Lorentz transformation of the axes of reference. The formula

$$E_e = m_0 + p'^2/2m_0, \quad E_i = \mu + p^2/2\mu$$

provides for both an external classical motion p' and a transition motion p; but it is understood that these are alternatives, since we have not investigated the cross-terms that might arise if both were present simultaneously.

The point that we have to make clear is that, in adding the energy μ in order to free the intracule, we change the rest mass of the atom to $m_0+\mu$ but we do not change its mass constant m_0 for classical motion, e.g. for motion produced by a change of space-time axes. The added energy is of the nature of potential energy, not set in motion when the atom is set in motion.

[a] The extracule, by definition, cannot change from the initial state; but the momentum p' can be given to it by referring it to moving axes.

28. Separation of electrical energy

We shall now introduce the hydrocule in another way, which throws more light on its nature. The standard carrier or V_{136} carries an energy tensor and nothing more. We have now to consider a particle which carries in addition a permutation variate, so that when it is resolved into two vector particles the particles are distinctively suffixed. The relation of the permutation variate to the 136 components of the complete energy tensor will be investigated later. For our present purpose it is sufficient that the particle has on account of the additional variate a phase space of 137 dimensions, and is therefore a V_{137}. By (16·5) the rest masses m_0, m_0' of a V_{136} and V_{137} are related by

$$m_0' = \tfrac{136}{137} m_0 = m_0/\beta. \tag{28·1}$$

But m_0/β is the mass already found for the hydrocule. The new particle can therefore be identified with the hydrocule.

Both the standard particle and the hydrocule are hydrogen atoms. The difference lies in the amount of observational probing supposed to have been applied. In the standard particle the probing (which necessarily has a disturbing reaction on the atom) is confined to determining, with more or less uncertainty, the two occupied points of space; in the hydrocule it goes deeper and determines, with more or less uncertainty, which particle is at which point. (It is understood that the probing can alternatively be applied to variates conjugate to those mentioned.) Evidently the deeper probing is implied in the picture of the hydrogen atom which the atomic physicist has in mind; so that for him the ideal hydrogen atom is a hydrocule. But for the molar physicist the ideal hydrogen atom is a standard particle, i.e. a carrier of an element of the energy tensor of molar hydrogen. It is only when an additional variate with no molar counterpart is loaded on the carrier that the drop of rest mass from m_0 to m_0' occurs. The change from the standard particle uranoid to the hydrocule uranoid—from system A to system B—is a change from the molar physicist's to the atomic physicist's point of view.

The permutation variate introduces interchange energy, later to be identified with Coulomb energy, so that it furnishes an electrical degree of freedom. Although a unified treatment of electricity and gravitation is sometimes valuable, it is kept in reserve, because practical applications require a more intensive treatment of electrical than of gravitational effects. The purpose of unification is to obtain a clean start for separation—so that early haphazard separation may be replaced by a systematic separation. The unified outlook has given us the fundamental result $m_0' = m_0/\beta$; but we shall now keep the electrical (interchange) degree of freedom separate from the 136 mechanical degrees of freedom, and the self-consistent electric field separate from the self-consistent gravitational field. Thus the hydrocule will be treated as a V_{136}, that being its mechanical multiplicity. The whole theory of Chapter II is equally adapted to the standard particle and the hydrocule, since it deals with the rigid gravitational field, and k is by implication the mechanical multiplicity. The hydrocule will have additional electrical effects represented by correcting terms or factors inserted in the equations of a purely mechanical particle; but these will be found as corrections to the equations of a V_{136} mechanical particle.

The mass m_0 of the hydrocule in system B agrees with the mass m_0 of the standard particle in system A. It is therefore entirely consistent to regard the hydrocule as the

V_{136} of system B and the standard particle as the V_{136} of system A. The difference, as already stated, is the difference of the hydrogen atom as conceived in atomic and in molar physics. Taking the point of view of the atomic physicist, we shall now regard the application of the theory of Chapter II to the hydrocule as its primary application. The results will then refer to system B.

In the two-particle transformation (26·12) the particles are suffixed (the suffix being indicated by an accent). If they were not, ξ_α would have an ambiguous sign and it would be impossible to proceed. *Thus the two-particle transformation can only be legitimately applied to a hydrocule.* Having transformed the proton and electron into an extracule and bound intracule, we usually substitute a free intracule for the bound intracule in further analysis. This converts the system into a standard particle in system B. We may say that the hydrogen atom appears in atomic physics (system B) in two forms: in conception, as a hydrocule consisting of a proton and electron or equivalently an extracule and bound intracule; in analysis as a standard particle consisting of an extracule and a free intracule.

System A is the 'observational system'. We study microscopic entities supposed to be endowed with measurable characteristics; but, as remarked in §1, the terms 'observable' and 'measurable' are not to be understood too literally. All that is guaranteed is that there is no self-contradiction in supposing them to have been observed or measured. Whatever may be the logical status of the measurements supposed to be carried out inside an atom or nucleus, they are not the measurements performed by the practical physicist from which our knowledge of atoms and nuclei is actually derived. The 'hard facts of observation' are molar measurements. Thus when the time comes to connect the results of theoretical analysis with laboratory measurements, the object-system must be put into the environment postulated by the definitions of molar quantities. In passing from molar to microscopic representation the energy tensor of the uranoid must not be changed, either by altering the zero level of energy reckoning or by altering the extraneous standard; all that may be done to adapt it to microscopic theory is to dissect it into additive elements of suitable magnitude. By definition the carriers of these elements of energy tensor are standard particles; so that the uranoid postulated in experimental comparisons is system A.

Thus results obtained in the theoretical system B have to be transformed to system A before they can be compared with observational measurements. Most of the quantities employed in physical description are connected in a roundabout way with observational measurements. It is these intermediary quantities that give the most trouble, directly measured quantities being transformed according to the comprehensive scheme in (8·4). The reduction of the intermediary quantities from B to A can only be found by tracing the detailed procedure by which their so-called 'observational values' have in fact been derived from measurements—since this procedure is an essential part of their definition. Thus the calculation of the β-factors, introduced into the values of the constants by the reduction from B to A, is a continually recurring problem, which crops up afresh whenever a new theoretical constant is to be compared with observation, or rather with the result obtained by the current procedure of reduction of the observations.

We have to accept the convention that any quantity that has been extensively used in the systematisation of observational knowledge has acquired thereby the status

of a *vested interest.* It if does not arise naturally in the theory we have to go out of our way to introduce it in order to avoid talking a different language from everyone else. Unless an internal inconsistency is detected the established procedure of reduction of the measurements—which is the definition of the quantity—must be accepted without amendment.

29. Current masses of the proton and electron

By the quantal theory of the hydrogen atom the Rydberg constant for hydrogen is

$$R = \frac{1}{2}\left(\frac{1}{137}\right)^2 \frac{\mu c}{2\pi\hbar}. \tag{29·1}$$

This is the usual formula simplified by inserting the exact value 137 of $\hbar c/e^2$. (The derivation by fundamental theory leads directly to (29·1) there being no occasion to introduce e^2 except as an abbreviation for $\hbar c/137$.) Since the formula is obtained from the theory of the free intracule, R, μ are the constants R_B, μ_B for system B, $\hbar$, c being the same in both systems. The wave-lengths used in an empirical determination of the Rydberg constant are measures in system A; so that the empirical Rydberg constant $\mathfrak{R}$ is R_A. Since $\mathfrak{R}$ is the reciprocal of a length, $R_B = \beta^{\frac{1}{2}} R_A = \beta^{\frac{1}{2}}\mathfrak{R}$ by (27·6).

We introduce a mass μ_A defined by

$$\mu_B = \beta^{\frac{1}{2}}\mu_A. \tag{29·2}$$

Then

$$\mathfrak{R} = \frac{1}{2}\left(\frac{1}{137}\right)^2 \frac{\mu_A c}{2\pi\hbar}. \tag{29·3}$$

Thus μ_A is the 'current mass' of the intracule. It is the quantity obtained from the measured wave-lengths of hydrogen by the accepted formula (29·1), and quoted everywhere as the 'observed value' of $m_p m_e/(m_p + m_e)$. We have therefore to recognise μ_A as the mass of the intracule in the observational system A. On the other hand the masses of extracules transform like measured densities, and we have

$$M_B = \beta M_A. \tag{29·4}$$

Hence

$$\left(\frac{M}{\mu}\right)_A = \beta^{-\frac{5}{6}}\left(\frac{M}{\mu}\right)_B = \frac{136^2}{10}\beta^{-\frac{5}{6}} \tag{29·5}$$

by (18·6), since the theory of Chapter II is now being applied to hydrocules in system B. Using (29·5), the quadratic equation $m^2 - mM + M\mu = 0$ for the masses of the proton and electron becomes in the observational system,

$$10m^2 - 136mm_0 + \beta^{\frac{5}{6}} m_0^2 = 0, \tag{29·6}$$

which gives the mass-ratio

$$m_p/m_e = 1836{\cdot}34. \tag{29·7}$$

The masses found in this way will be called the *current masses* of the proton and electron to distinguish them from the *standard masses* in (18·8) and (18·9).

Since this is the first of many examples of adaptation of the theory to the observational system, we shall briefly review the principles involved. Hydrocules, in separated electrical theory, are V_{136} particles, and the theory of § 18 applies to them as well as to standard particles; but the hydrocule application implies a hydrocule

uranoid, so that the results refer to system B. In particular it is $(M/\mu)_B$ that is equal to $136^2/10$. The obvious way of comparing this constant with observation is to transform the observational measurements into system B. But it is impracticable to go through the vast accumulation of observational material and re-reduce it. Moreover, we cannot ignore the existing systematisation of observational knowledge, and the vested interest of those quantities that have been widely used in this systematisation. This means that in practice we have to adapt the theory to system A, not the observational data to system B. If, in comparing the theoretical constant $136^2/10$ with observation, the transformation of the observational data from system A to B is overlooked, the comparison will be in error by a factor $\beta^{\frac{5}{6}}$. Adaptation of the theory to system A means that we include this factor in M/μ so as to dispense with the transformation of the measurements. We thus present a distorted version of the theory, which has, however, the merit of being directly applicable to practical measurements. The characteristics of the atom (and the resulting proton and electron) are distorted as compared with the simple theoretical atom; but the essential point is that it is this distorted atom that has been reached by those who proceed from the observational end, following established practice and making no allowance for the transformation.

We proceed to a more general discussion of the observational system. Mass, momentum, charge, etc., are primarily defined as molar measures. In some cases the definition can be extended to microscopic physics by the convention that the mass of a particle is $1/n$ of the molarly measured mass of a large aggregation of n similar particles. The elementary charge e can likewise be defined as $1/n$ of the molarly measured charge of a body from which n electrons have been expelled. Quantities defined in this way will be said to be *molarly controlled*. Molar control can be applied to the mass M of a hydrogen atom and to the masses of all neutral atoms and molecules. But it cannot be applied to the mass m_e of an electron, since electrons cannot be aggregated separately.[a] Nor can it be applied to the mass μ of an intracule, since intracules cannot be aggregated separately from their extracules.

In quantum mechanics a fresh start is made, and the terms mass, momentum, charge, etc., are applied to quantities which are *analogues* of the correspondingly named molar quantities in a well-known analogy between quantum and classical mechanics. We shall often have to emphasise that nomenclature in quantum theory is guided by analogy, not by equivalence. Thus, even when a molarly controlled definition of a quantity exists, quantum theory does not adopt it unless it fits the analogy. We shall see in § 30 that the molarly controlled definition of electric charge has to be rejected because it does not fit the analogy. But—to make a start—quantum theory accepts the molarly controlled definition of the masses of neutral atoms and molecules. This allows us to make a practical determination of the mass M of a hydrogen atom (in the observational system) in the following way.

The steps are: (1) the mass density of a crystal is found by molar measurement; (2) the particle density (number of molecules per cu.cm.) is found by a measurement of the lattice volume, depending on the comparison of the grating formed by the crystal itself with a ruled grating; (3) the quotient of (1) by (2) gives the molarly

[a] In case it is proposed to determine the mass by applying a correction for the electrostatic energy to the molarly measured mass of a large aggregation of electrons, we may point out that if a gram of electrons were put into a vessel of 10 cm. radius, the measured mass (which includes electrostatic energy) would be about 10 million tons.

controlled mass M_c of a crystal molecule; (4) the ratio M_c/M (molecular weight in terms of $H = 1$) is found by chemical density comparisons.

The fact that hydrogen does not form a suitable grating, so that another substance has to be employed as intermediary, may be set aside as an inessential detail. In principle M and other neutral masses are determined by measuring molarly the mass density, and also by the grating method the particle density, of a large assemblage.

There is no molarly controlled definition of μ, and it rests with the quantum physicist to state what he intends it to mean. As usual, the text-books give no explicit definition,[a] and we have to institute a kind of detective inquiry to discover the definition that has been implicitly assumed. Since $\mathfrak{R}$ is much more accurately known than any other physical constant, the relation (29·1) plays a leading part in the accepted observational values; and the observational system is deeply committed to a definition of μ consistent with it. Assuming that $\hbar$ has been duly defined (a matter that will be dealt with later), (29·1) supplies the required definition of μ. The mass defined in this way will be said to be *spectroscopically controlled.* The same system of definition must evidently be applied to atoms of elements other than hydrogen. The accepted convention is accordingly

The masses of extracules are molarly controlled, and the masses of intracules are spectroscopically controlled. (29·8)

The observational determination of μ is completed in the following way. The Faraday constant for hydrogen $\mathfrak{F} = e/Mc$ is found by molar measurement of the charge resulting from electrolysis of a known mass of water. We have

$$\hbar = 137e^2/c = 137\mathfrak{F}^2M^2c. \tag{29·91}$$

Inserting this in (29·3), $$\mu_A = 4\pi \,.\, 137^3\mathfrak{R}\mathfrak{F}^2M^2. \tag{29·92}$$

This, however, requires a small correction. The interpretation of the measured Faraday constant as e/Mc assumes that e is defined by molar control. We shall show in the next section that the system of quantum mechanics requires that the elementary charge e in quantum theory shall differ slightly from the charge e' defined by molar control; and consequently the constant $\mathfrak{F} = e/Mc$ to be used in (29·92) differs from the observed Faraday constant $\mathfrak{F}' = e'/Mc$.

30. Molarly controlled charge

The rules for the change of units in the transformation $A \to B$ are: (a) densities are multiplied by β, (b) lengths and times are multiplied by $\beta^{-\frac{1}{2}}$. Rule (a) is the condition for gravitational equilibrium, and it applies equally to molar and microscopic density; independently of (b) it makes κ invariant. Rule (b) is the result of the relation $8\pi\kappa\hbar^2 = 1$, and it makes $\hbar$ (and therefore e) invariant. But the relation $8\pi\kappa\hbar^2 = 1$ cannot be retained in classical electrodynamics, which is a limiting form obtained by making $\hbar \to 0$; so that rule (b) drops out. We have to consider what rule should be substituted for (b) in transforming quantities defined by, or controlled by, classical electrodynamics.

In a quantum-specified process the action is given as a fixed multiple of the unit $\hbar$ or e^2/c. Since $\hbar$ and e^2/c are invariant, the quantum-specified action is invariant in the

[a] Nor is there any definition of m_e, which would fill the gap equally well.

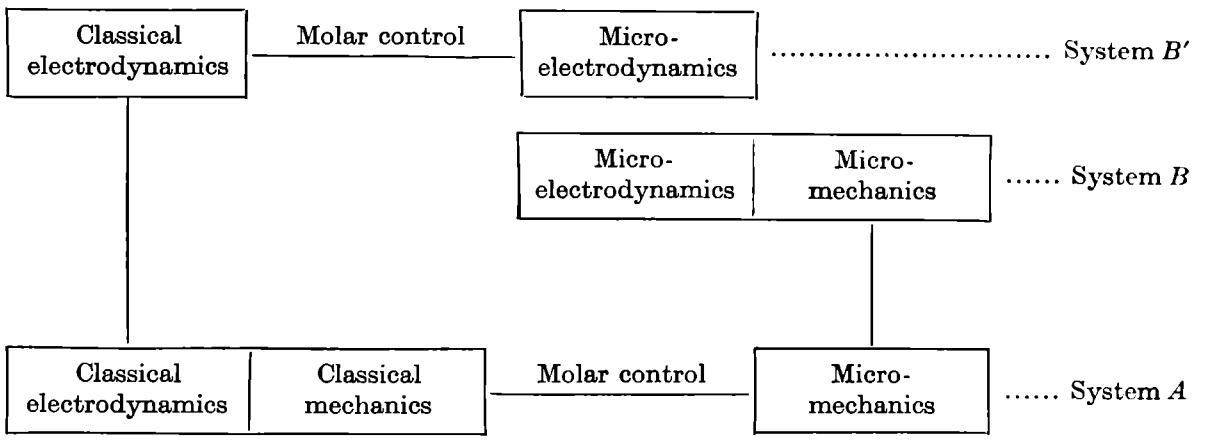
transformation $A \to B$. In classical electrodynamics action is a given function of the continuous variables which specify the process; and the connection between electrodynamics and mechanics is made by adding the electrical to the mechanical action in the usual Hamiltonian equations. Since the basis of quantum nomenclature is the analogy between classical and quantum dynamics, the new rule (b') should be such that, when coupled with (a), the action in a process specified by classical electromagnetic theory is invariant. Since classical action has the dimensions ρL^4 (when $c = 1$), the multiplication of ρ by β involves multiplication of lengths by $\beta^{-\frac{1}{4}}$. Calling the resulting measure-system B', the rule is

In $A \to B'$, lengths and times are multiplied by $\beta^{-\frac{1}{4}}$. (30·1)

Or, combining (27·6) and (30·1),

In $B \to B'$, lengths and times are multiplied by $\beta^{-\frac{1}{11\frac{1}{2}}}$, densities being unchanged. (30·2)

We now have the chain of connection shown in the diagram.



The essential conditions which determine the diagram are:—Classical mechanics and electrodynamics are connected in the observational system A. Microscopic mechanics and electrodynamics are connected by the theory of suffixed particles, and therefore in system B. The changes of measure-system, shown in the diagram, are such that action is invariant. The passage from classical to microscopic theory is made by molar control; that is to say, microscopic elements are treated as simply additive.

The starting point is classical mechanics—the energy tensor $T_{\mu\nu}$ of molar relativity theory. We reach micro-electrodynamics, which contains the action units e^2/c and $\hbar$ by two routes; but the measure-systems disagree. This means that the upper molar control, which is applied to electrical actions and energies, is inconsistent with the lower molar control, which is applied to mechanical energies and actions. Since we have adopted molar control for mechanical quantities (masses of neutral particles), we cannot consistently adopt the molarly controlled definition of e^2/c or $\hbar$.

Consider the effect of the transformation $B \to B'$ on a hydrogen atom in a specified quantum state. By the usual theory of the eigenstates, the various energies, including the Coulomb energy of the state, have fixed ratios to M, and therefore transform like M when the measure-system is changed. But we have seen that M transforms as a

density; and, since densities are unchanged in passing from B to B', M is unchanged. Hence

In $B \to B'$, the Coulomb energy of a specified quantum state is unchanged. (30·3)

Denoting the Coulomb energy by $e^2/\bar{r}$, we have $(e^2)_B/\bar{r}_B = (e^2)_{B'}/\bar{r}_{B'}$. Hence

$$\frac{(e^2)_{B'}}{(e^2)_B} = \frac{\bar{r}_{B'}}{\bar{r}_B} = \beta^{-\frac{1}{12}} \tag{30·4}$$

by (30·2). Setting e, e' for e_B, $e_{B'}$, e is the elementary charge as defined in quantum theory. Primarily it is the measure in system B, but it also applies to the observational system A, since e is invariant for the transformation. The molarly controlled definition gives e'; since it follows the other route from classical mechanics via classical electromagnetic measurement, which gives the measure-system B'. The constants $\hbar$, $\mathfrak{F}$ are derived from e; and we have the equations

$$e' = \beta^{-\frac{1}{24}}e, \quad \hbar' = \beta^{-\frac{1}{12}}\hbar, \quad \mathfrak{F}' = \beta^{-\frac{1}{24}}\mathfrak{F}, \tag{30·5}$$

giving the molarly controlled constants in terms of the actual quantum constants.

We can now proceed to evaluate M/μ from (29·92), remembering that the constant $\mathfrak{F}$ is $\beta^{\frac{1}{24}}$ times (about 1 part in 3300 greater than) the ordinary Faraday constant $\mathfrak{F}'$. The observational data, according to R. T. Birge,[a] are $\mathfrak{R} = 109677{\cdot}58 \pm 0{\cdot}05$ cm.$^{-1}$, $\mathfrak{F}' = 9573{\cdot}56 \pm 1{\cdot}0$ e.m.u.g.$^{-1}$, $M = (1{\cdot}67339 \pm 0{\cdot}00031)\,10^{-24}$ g. With these values (29·92) gives $M/\mu = 1838{\cdot}56 \pm 0{\cdot}51$. The theoretical value (29·5) is 1838·34. Or, since $M/\mu = m_p/m_e + 2 + m_e/m_p$, the comparison can be put in the form

$$m_p/m_e, \quad \text{theoretical } 1836{\cdot}34, \quad \text{observed } 1836{\cdot}56 \pm 0{\cdot}51. \tag{30·6}$$

A rather more accurate observational value, depending on other kinds of measurements, will be given later.

31. Secondary anchors

Theoretical equations must be anchored to observation by giving observational definitions of the quantities involved in them. For each such quantity an exact procedure must be indicated by which its value can be obtained from experimental measurements. In the last two sections we have fulfilled this obligation. It is desirable, though not essential, that the required measurements shall be capable of high accuracy in the present state of experimental physics. This requirement has also been fulfilled.

Having provided a primary anchorage, we may (optionally) throw out secondary anchors—alternative procedures, involving different experimental measurements. These may become very numerous; for example, Birge cites seven essentially different ways of determining e/m. We cannot turn aside from the general development of fundamental theory to enter on a comprehensive study of secondary anchors. On the other hand we cannot ignore them altogether.

I shall here examine briefly three secondary anchors: (1) the spectroscopic determination of m_p/m_e, (2) the deflection determination of $e/m_e c$, (3) the direct determination of h/e. These are measurements which have now attained high accuracy, and have

[a] *Reports on Progress in Physics* (Physical Society), **8**, 90, 1942.

weight in determining the best possible observational values of the physical constants. My examination will be limited to the question whether the usual reductions are correct as regards β-factors. In other respects, I accept the validity of these methods on the authority of current quantum theory, without accepting any responsibility.[a]

(1) *Spectroscopic determination of* m_p/m_e. Applying the formula (26·11) both to hydrogen and deuterium, we have

$$M = m_p + m_e, \quad M_d = m_d + m_e, \quad \mu = m_p m_e/(m_p + m_e), \quad \mu_d = m_d m_e/(m_d + m_e). \quad (31·1)$$

Setting $M_d = \alpha M$, $\mu_d = \gamma\mu$, we obtain

$$\frac{m_p}{m_e} = \frac{\alpha - 1}{\alpha(\gamma - 1)}. \quad (31·2)$$

By (29·1), γ is the ratio of the Rydberg constants for deuterium and hydrogen. Also α, which is the ratio of the atomic weights, can be determined by density comparisons. (On the authority of current theory, we accept a mass-spectrograph determination as equivalent.) Thus (31·2) provides an observational test of the theoretical value of m_p/m_e, independent of that given in (30·6). The result (including values obtained from a similar comparison of helium and hydrogen) is

$$m_p/m_e, \quad \text{theoretical } 1836{\cdot}34, \quad \text{spectroscopic } 1836{\cdot}14 \pm 0{\cdot}22. \quad (31·3)$$

Deuterium will be investigated in § 94. Meanwhile the problem is beyond the stage that we have reached in developing fundamental theory; and the immediate importance of (31·3), is, not so much the additional and closer confirmation of the theoretical value, as its indication of the nature of the extension of the theory to elements other than hydrogen. Our two measure-systems A and B have been based on hydrogen in its two forms, standard particles and hydrocules, but they must serve for all other elements. The ratio γ refers to Rydberg constants in the same measure-system, primarily system B. The energy μ_d required to free the deuterium intracule has no particular relation to the change of energy reckoning between A and B; nevertheless, we are able to apply the quantal theory of the free intracule to deuterium, without introducing measure-systems adapted to the constants of deuterium in the same way that systems A and B are adapted to the constants of hydrogen.

To understand this, we must remember that systems A and B correspond to the two kinds of hydrogen uranoid. There is no question of introducing a deuterium uranoid. Obviously, when we substitute a deuterium atom for a hydrogen atom as our object-system, we do not want to re-construct the whole universe with deuterium in place of hydrogen. Secondly, the difference of measure-system depends on the extra degree of freedom introduced by suffixing the particles. Suffixing is introduced in resolving a hydrogen atom into a proton and electron, but there is no corresponding procedure in resolving a deuterium atom into a deuteron and electron; for the deuteron is introduced, not by analysis, but as a combination of two protons and an electron, and the suffixing is already presupposed when we begin to build it.

[a] Unless some unforeseen obstacle appears, the theory of these methods can be developed rigorously on the basis of the present fundamental theory. But it is impossible for one person to follow up all the detailed applications, and ascertain whether the full investigation confirms the results commonly accepted. I therefore accept the methods as correct, unless I have definite grounds for suspecting error.

The distinction of system B as the theoretical system (for quantal theory) and system A as the observational system, though introduced in connection with hydrogen, covers all kinds of matter. The transformation $A \to B$ separates the 'electrical world' from the 'mechanical world'. These are superposed in the actual universe, and therefore in the observational system; but they are separated in order to apply to the electrical world the intensive treatment developed in quantal theory. By carrying out the separation in hydrogen, we separate them once for all; for all matter is composed of hydrogen.

(2) *Deflection determination of* $e/m_e c$. In this experiment the measured quantities are the curvature of track of the electron and a molar electromagnetic force E or H. The current reduction assumes that the solution is the same as in classical mechanics. The exact theory is particularly difficult, because one partner in the experiment is microscopic and one molar. The molar element in the problem secures that a large number of quanta are involved; so that the action can be treated as continuous and the dynamical equations converge to classical form. But there still remains the question of the correct formula for the Coulomb energy between a molar and a microscopic charge. For two molar charges the formula is $n_1 n_2 e'^2/r_{12}$ and for two microscopic charges it is $n_1 n_2 e^2/r_{12}$. The commonsense conclusion is that in the micro-molar problem it is $n_1 n_2 e' e/r_{12}$; this gives $n_2 e V'$ for a microscopic charge $n_2 e$ in terms of the measured molar potential $V' = n_1 e'/r_{12}$. Thus the charge as given by the current reduction is e (not e'); and the result can be accepted as a correct determination of the constant $e/m_e c$.

I find nothing to suggest that a full investigation will disagree with the 'commonsense' conclusion; but I have not gone far enough to be able to state that it is definitely confirmed.

(3) *Direct determination of* h/e. This depends on the conversion of an energy $e\delta V$ due to the fall of an electron through a molarly measured potential difference δV into an energy of radiation $h\delta\nu$ where $\delta\nu$ is a measured difference of frequency. Here h is evidently the quantum constant (not h'), and e is the same as in the deflection experiment. Assuming the 'commonsense' conclusion, the constant given by the usual reductions is h/e.[a] Experiments are not yet accurate enough to distinguish between h/e and h/e'; but a comparatively small improvement in the data will make this a valuable test.

32. Calculated values of the microscopic constants[b]

All observational data in this section are taken from R. T. Birge, *Reports on Progress in Physics*, **8**, 90, 1942.

Experimental results are expressed in three traditional units which have no relation to any theory. We have therefore to select three measured quantities to be used as conversion constants. They must occur in the part of the theory believed to be exact; and, subject to this, they should be the most accurately ascertained constants available. There can be little doubt that the most suitable choice is the velocity of light (c), the

[a] In *Proc. Phys. Soc.* **54**, 491, 1942, I stated incorrectly that the constant determined in this way was the 'official', i.e. molarly controlled, constant h'/e'.

[b] The corresponding section on molar, cosmical and nuclear constants is § 51. I have given a somewhat similar discussion in *Proc. Phys. Soc.* **54**, 491, 1942. Certain differences appear because in the earlier discussion I accepted (under protest) the molarly controlled value e' as the 'official' value of the elementary charge, and defined the official m_e consistently with it.

Rydberg constant for hydrogen ($\mathfrak{R}$) and the Faraday constant for hydrogen ($\mathfrak{F}'$). The measured values are:

$$\left.\begin{aligned} c &= (2{\cdot}99776 \pm 0{\cdot}00004) \times 10^{10}\,\text{cm.sec.}^{-1}, \\ \mathfrak{R} &= 109677{\cdot}58 \pm 0{\cdot}05\,\text{cm.}^{-1}, \\ \mathfrak{F}' &= 9573{\cdot}56 \pm 1{\cdot}0\,\text{e.m.u.g.}^{-1}. \end{aligned}\right\} \qquad (32{\cdot}1)$$

From (32·1) it would be possible to reconstruct the C.G.S. system if the original standards were lost; and for the present purpose we regard (32·1) as the *definition* of the centimetre, gram and second.

The calculated values of the constants are obtained by combining with (32·1) our theoretical values of the fine-structure constant, the mass-ratio and e'/e. The collected formulae are:

$$\left.\begin{aligned} &(1)\ \frac{hc}{2\pi e^2} = 137, \qquad (2)\ M = \tfrac{137}{10} m_0, \qquad (3)\ \mu = \frac{\beta^{\frac{5}{8}}}{136} m_0, \\ &(4)\ e' = \beta^{-\frac{1}{24}} e, \qquad (5)\ h' = \beta^{-\frac{1}{12}} h, \qquad (6)\ \beta = \tfrac{137}{136}, \\ &(7)\ M = m_p + m_e, \qquad (8)\ \mu = m_p m_e/(m_p + m_e), \\ &(9)\ \mathfrak{R} = \frac{1}{2}\left(\frac{1}{137}\right)^2 \frac{\mu c}{h}, \quad (10)\ \mathfrak{F}' = \frac{e'}{Mc}. \end{aligned}\right\} \qquad (32{\cdot}2)$$

For some of the constants the observed values given by Birge require correction, because e' has not been distinguished from e in the reductions. The constant primarily affected is $e/m_e c$. Birge gives two slightly discordant observational determinations, namely,

deflection	$(1{\cdot}75959 \pm 0{\cdot}00024) \times 10^7$,
spectroscopic	$(1{\cdot}75880 \pm 0{\cdot}00028) \times 10^7$.

The discordance is not in the data but in the reductions. The deflection method gives $e/m_e c$ correctly (§ 31); but the spectroscopic value is obtained from the spectroscopic determination of m_p/m_e by using the Faraday constant $\mathfrak{F}'$, so that it corresponds to $e'/m_e c$. Multiplying it by $\beta^{\frac{1}{24}}$ to obtain $e/m_e c$, the comparison becomes

$e/m_e c$ (deflection)	$1{\cdot}75959 \pm 0{\cdot}00024$,
$e/m_e c$ (spectroscopic)	$1{\cdot}75934 \pm 0{\cdot}00028$,
Mean	$1{\cdot}75947 \pm 0{\cdot}00018$.

The most direct comparison of theory and observation is given by the values of the fine-structure constant and the mass-ratio, since these, being pure numbers, do not involve the conversion constants. In the following comparison we give first the uncorrected observed values given by Birge,[a] and second the corrected values which take account of the distinction between e' and e:

	Uncorrected	Corrected	Theoretical
$hc/2\pi e^2$	$137{\cdot}030 \pm 0{\cdot}016$	137·009	137
m_p/m_e	$1836{\cdot}56 \pm 0{\cdot}56$	1836·27	1836·34

The probable errors of the corrected values are substantially smaller owing to the removal of the discordance between deflection and spectroscopic determinations. They are about 1 part in 10,000 for both constants.

[a] The coincidence of Birge's uncorrected m_p/m_e with the observed value in (30·6) is accidental.

The constants calculated from (32·1) and (32·2) are as follows:

Microscopic constants

Ref. No.	Symbol	Description	Calculated	Observed	P.E.
1	M	Mass of ^{1}H atom	$1{\cdot}67368 \times 10^{-24}$	1·67339	31
2	m_0	Mass of comparison particle	$1{\cdot}23065 \times 10^{-25}$	...	...
3	μ	Mass of intracule	$9{\cdot}10428 \times 10^{-28}$	...	...
4	m_e	Mass of electron	$9{\cdot}10924 \times 10^{-28}$	9·1066	22
5	m_p	Mass of proton	$1{\cdot}67277 \times 10^{-24}$	1·67248	31
6	e	Elementary charge (e.s.u.)	$4{\cdot}80480 \times 10^{-10}$	...	...
7	e'	(Molarly controlled)	$4{\cdot}80333 \times 10^{-10}$	4·8025	10
8	h	Planck constant	$6{\cdot}62908 \times 10^{-27}$	...	...
9	h'	(Molarly controlled)	$6{\cdot}62504 \times 10^{-27}$	6·6242	24
10	h/e	(Direct method)	$1{\cdot}37968 \times 10^{-17}$	1·3800	5
11	$e/m_e c$	...	$1{\cdot}75953 \times 10^{7}$	1·75947	18
12	e/Mc	Faraday constant $\mathfrak{F}$	9576·48	...	...
13	e'/Mc	Faraday constant $\mathfrak{F}'$	...	9573·56	100
14	m_p/m_e	Mass-ratio	1836·34	1836·27	...
15	$hc/2\pi e^2$	Fine-structure constant	137	137·009	...

The observed values of Nos. 11, 14 and 15 have been corrected as explained above. The other observed values are taken from Birge's Report without correction, although in some cases a correction is required. The probable error is given in units of the last decimal place of the observed value.

33. The Coulomb energy

Interchange of the unaccented and accented quantities in (26·12) does not alter x_α, but it reverses the sign of ξ_α; consequently the interchange (Coulomb) energy appears in the equations of the intracule. A full treatment of the Coulomb energy must be deferred until the wave equation of the intracule is investigated in Chapter x; but it is desirable at this stage to explain in general terms the way in which it is introduced.

We shall find (in agreement with current theory) that the internal wave functions of the hydrogen atom are of the Dirac type, which have a half-quantum of spin angular momentum in any plane of rotation. From the physical point of view this is a spurious angular momentum which comes from using a mathematical coordinate frame as reference system; for we cannot help giving a coordinate frame a right- or left-handed character, whereas the physical environment of the atom possesses no such chirality. The spin momentum is therefore to be regarded as a formal addition which our analytical method makes necessary. This applies also to the extra-spatial plane of interchange circulation; but in this case the spin momentum to be added is a whole quantum. This quantum of extra-spatial spin momentum provides the interchange energy. The interchange angular momentum, like the spin angular momentum occurs because we are obliged to give to the mathematical reference system a character which the physical environment does not possess. The mathematical reference system includes a distinctive suffixing of the two particles.

A spatial angular coordinate is measured relatively to the environment, and the conjugate angular momentum properly belongs half to the atom and half to the environment; whereas the permutation coordinate θ is internal to the atom, and the conjugate angular momentum belongs wholly to the atom. Thus the spin angular

momentum to be inserted in the equations is $\frac{1}{2}\hbar$ for a spatial rotation, and $\hbar$ for the interchange rotation. To put it another way: if we give the two-particle system a spatial angular momentum $\frac{1}{2}\hbar$, the rest of the universe must recoil with angular momentum $\frac{1}{2}\hbar$; but if we give it an angular momentum $\frac{1}{2}\hbar$ of interchange of the two particles, there is no corresponding way in which the rest of the universe can recoil. To conserve angular momentum, we must put the recoil angular momentum inside the atom, and picture the interchange as consisting of two equal and opposite circulations each with angular momentum $\frac{1}{2}\hbar$. This is the rigorous representation; and it agrees with our general treatment of continuous transition circulations ('interstates'). But for most purposes, the two interchange circulations can be replaced by one circulation of double weight.

We consider accordingly the addition of a quantum $\hbar$ of interchange angular momentum to the rest state of the two-particle system; and, for the purpose of comparison, we add at the same time a spatial angular momentum $n\hbar$. In accordance with the principles of wave mechanics these additions are to be made in a rigid field; and, since we are here treating the electrical degree of freedom which provides for interchange in just the same way as the mechanical degrees of freedom which provide for spatial rotation, the multiplicity must be taken to be $k = 137$. By (19·7) the rigid time t is k^{-1} times the Galilean time, and the momentum conjugate to it is k times the Galilean momentum. The interchange angular momentum gives the intracule a linear momentum $\hbar/r$ in an extra-spatial dimension normal to r ($r^2 = \xi_1^2+\xi_2^2+\xi_3^2$); if this is the time dimension, the linear momentum reduced to Galilean coordinates becomes $\hbar/kr$. The spatial angular momentum, which we inserted for comparison, gives a spatial linear momentum $n\hbar/r$; and this is unaltered in the transformation from rigid to Galilean coordinates. Thus in the wave equation as ordinarily given, i.e. adapted to Galilean coordinates, the linear momenta to be inserted are respectively

$$\hbar/137r, \quad n\hbar/r. \tag{33·1}$$

The extra-spatial dimension is not the time; but it is the time analogue which plays the part of the time in the analogy between quantum and classical mechanics. To maintain the analogy it must transform in the same way as the time in passing from rigid to Galilean coordinates; so that (33·1) remains valid.

In the linear wave equation, from which the energy levels of hydrogen are derived, the directions of the various momenta are indicated by symbolic coefficients; and we can easily pick out the momentum which has the direction of the time analogue. Since an equation can always be multiplied throughout by a numerical factor, we have introduced a check momentum $n\hbar/r$ in order to test whether the current form of the equation shows the momenta or a multiple thereof. Actually $n\hbar/r$ appears without a numerical factor in the current form, so that the momenta are shown directly. The term which has the symbol of the time analogue is, in the current notation, e^2/cr; and, since this is the term $\hbar/137r$ found in (33·1), we have the identification $e^2 = \hbar c/137$.

In this derivation the fine-structure constant $\hbar c/e^2 = 137$ appears as the k of a V_{137}, instead of the $k+1$ of a V_{136}. That is because we here employ unified theory in which the electromagnetic field is rigidified along with the gravitational theory. In separated electrical theory, only the gravitational field is rigid; and the flexibility of the electric part of the field leads to a re-adjustment of the relations between field energy and total energy.

Chapter IV

GRAVITATION AND EXCLUSION

34. Unsteady states

In molar relativity theory the mass of a system of particles, not necessarily in a steady state, is defined as follows: Taking axes such that the centre of mass is instantaneously at rest, the mass M of the system is equal to the mass of a single particle, having the position and acceleration of the centre of mass, which would produce the same gravitational field, i.e. the same geodesics, at a great distance from the system. If M is changing, the time to which the above-mentioned value refers is earlier than the time of agreement of the geodesics by the amount of the light-time.

The meaning is that, so far as remote gravitational effects are concerned, the system is replaceable by a particle of proper mass M at, and moving with, its centre of mass. Evidently, if the system is in the midst of an empty region large compared with its own dimensions, the system and the equivalent particle make the same net addition to the energy of the universe.

Until recently there was disagreement between the results of different investigators; but the discrepancies have been cleared up and there is now general agreement. For a system of gravitating particles,[a]

$$M = M_0 + 3K + 2V, \tag{34·1}$$

where M_0 is the sum of the rest masses of the particles, K the kinetic energy, and V the internal potential energy $\Sigma\kappa m_s m_t/r_{st}$; K and V are defined as in Newtonian theory, the approximation being only correct to this order. Even this approximation involves very lengthy calculations. Using the Newtonian identity

$$2K + V = \tfrac{1}{2}d^2C/dt^2, \tag{34·2}$$

where C is the moment of inertia of the system about its centre of mass, (34·1) is reduced to the forms

$$M = M_0 + K + V + \tfrac{1}{2}d^2C/dt^2, \tag{34·3}$$

$$M = M_0 - K + d^2C/dt^2.[b] \tag{34·4}$$

By (34·3) the system possesses, in addition to the energy $M_0 + K + V$ usually recognised, an *expansion energy* $\frac{1}{2}d^2C/dt^2$. The expansion energy can be associated with the gravitational waves set up by the changes of the system. We distinguish $M_0 + K + V$ as the *local energy*. The local energy satisfies the law of conservation; and it is the custom to consider this alone. But if we suppose the local energy to be the whole energy, changes of the gravitational potential energy of the remote environment are produced and

[a] Eddington and Clark, *Proc. Roy. Soc.* A, **166**, 469, 1938. I think this is the only paper that gives explicitly the equivalent mass M; but the same general formulae were obtained by Einstein and Infeld and by Robertson. Levi-Civita, who had previously obtained a different result, re-examined his investigation and announced his agreement. There was a mistake in de Sitter's pioneer calculation in 1916.

[b] [*Editorial note*:—A subsequent investigation by Clark has shown that the above definition of the mass of a system (as the mass of an equivalent particle which produces the same field at great distances) is possible only in the case of steady systems, i.e. systems for which d^2C/dt^2 vanishes. Non-steady systems can be included in the discussion if (following Whittaker and Ruse) Gauss' theorem be used as the basis of definition of mass. It is found that the mass can be expressed in terms of two volume-integrals, one of which gives a contribution $M_0 + 3K + 2V$ as above, while the second yields an additional term $-\frac{1}{2}d^2C/dt^2$. Adding these terms we obtain a final result $M = M_0 + K + V$.]

left unbalanced; so that in its wider application the law of conservation breaks down. The expansion energy balances these changes and restores conservation.

Since the laws of quantum mechanics are analogous to those of molar mechanics, we shall apply this result tentatively to a wave packet. Consider a Gaussian wave packet in three dimensions with an initial standard deviation σ in each coordinate, so that $C = 3m\sigma^2$. According to the solution given by current wave mechanics,[a] after a time t the standard deviation σ_t is given by

$$\sigma_t^2 = \sigma^2 + \hbar^2 t^2/4\sigma^2 m^2.$$

Hence the expansion energy of the wave packet is

$$\tfrac{1}{2}d^2C/dt^2 = 3\hbar^2/4\sigma^2 m. \tag{34·51}$$

It is also found that the momentum p_1 has the mean square value $\overline{p_1^2} = \hbar^2/4\sigma^2$; so that the mean kinetic energy $(\overline{p_1^2} + \overline{p_2^2} + \overline{p_3^2})/2m$ is

$$K = 3\hbar^2/8\sigma^2 m. \tag{34·52}$$

Thus $\frac{1}{2}d^2C/dt^2 = 2K$; and it follows from (34·2) that $V = 0$. This was to be expected, because the wave packet represents only one particle, and there is no gravitational or electrical attraction between elements of the same particle.

Setting $V = 0$ in (34·1) the wave packet has a total energy $M_0 + 3K$. A distribution which is not expanding has total energy $M_0 - K$ by (34·4). Thus in order to inhibit expansion we have to incorporate a constraint with energy $V = -4K$. Leaving aside the constant M_0, the energy $T = 3K$ of the free packet is made up of the energy $E = -K$ of the bound packet and the energy $W = 4K$ added by removing the constraint.

Since we are here dealing with a three-dimensional distribution of probability, the multiplicity factor is $k = 3$; and the foregoing is an example of the relation $T = -kE$, $W = -(k+1)E$. Evidently (34·1) is the special case for spinless V_3 particles of a general formula

$$M = M_0 + kK + 2V. \tag{34·6}$$

The term $2V$ calls attention to the fact that the conventional reckoning of M_0 is a double reckoning. For we attribute to the isolated system an energy which is really a mutual energy of the system and the rest of the universe; so that when each part of the universe is treated in turn as an isolated system the energy gets counted twice over. The double reckoning extends to kinetic energy, since the kinetic energy of a particle is derived from its rest energy by a Lorentz transformation. It applies also to the potential energy $-\Sigma m\phi$ or $\Sigma e\Phi$ of the system in an extraneous gravitational or electric field. But it does not extend to the internal potential energy, where the contribution of each pair of particles is reckoned only once. In (34·6), V is multiplied by 2 to transform it to the double reckoning in which all the rest of the energy is expressed.

It will be seen that in a pseudo-discrete assemblage, the total energy is turned into a particle energy $E = -T/k$ by introducing a constraint to inhibit dispersal of the particle. We have to debit the non-dispersing particle with the energy $-W$ of the constraint, or equivalently credit it with an energy W, which we have called the complementary field energy. The relation $\frac{1}{2}d^2C/dt^2 = 2K$ is independent of the size of a wave packet; and the formulae remain true for a particle whose probability distribution occupies the whole extent of the assemblage. Thus the assemblage (which is, by definition, in equilibrium and therefore not dispersing) is analysed into non-dispersing

[a] C. G. Darwin, *Proc. Roy. Soc.* A, **117**, 268, 1927. Darwin's σ is $\sqrt{2}$ times our σ.

particles carrying the particle energy E; and its field energy $W = -(k+1)E$ results from the inhibition of the dispersal. The present analysis shows further that the relation between particle and field energy is not altered by localising the particle in a small wave packet.

35. Under-observation

It is recognised that a system is disturbed by the process of measurement. Measurement in the broadest sense, i.e. the acquiring of information which we formulate as a measure, is not necessarily an artificial interference; the system may spontaneously broadcast information by starting a chain of causation which reaches our sense organs. But whether the information is given spontaneously or extorted by deliberate probing, it has a selective bias; because it refers to the system in a particular state of interaction with the environment. It is impossible that we should have genuine information about states of the system incompatible with the furnishing of information; and we convict ourselves of fabrication if we pretend to describe such states.

We commonly speak of an 'undisturbed state'; but we can have no knowledge of a state undisturbed by interaction with the environment. For consistency the normal state (loosely described as undisturbed) must be defined to be that in which a specified kind and amount of interaction with the environment is occurring, or equivalently *the system is subjected to a conventional amount of observational probing*; and its only characteristics (other than stabilised characteristics) are those revealed by this probing. Sporadic additional information cannot be incorporated, because our possession of it is evidence that the system is not in the normal state that we profess to describe; but it may prompt us to re-define the normal state so as to include the probing necessary to furnish the additional information systematically. Thus it may well happen that in different types of investigation, and at different stages of development of the theory, the normal condition is differently defined. The re-definition creates what is to all intent a new kind of system or particle with different specific characteristics. We have had an example of this in § 28, where the hydrogen atom, first identified with the standard particle, becomes, when probed for additional information, identified with the hydrocule, and its mass is decreased in the ratio β. Similarly, when a V_k is changed to a $V_{k'}$ by stabilisation, the specific character of the particle is altered because the information to be dug out of it is less extensive.

We have insisted on the importance of defining the postulated environment of the object-system. We must equally insist on the importance of defining the amount of interaction or interference from outside. Just as the irregular natural environment is replaced by a standardised uranoid, so the meddlesome activities of the experimenter are replaced by a standardised amount of interference. This supplementation is associated with the transition from molar to microscopic physics; in this transition we greatly elaborate the specification of an object-body, and it is not surprising that some addition to the ordinary molar specification of the environment should be needed. In general theory a system is described by generalised coordinates and momenta x_α, p_α. The standard probing is specified qualitatively by giving a list of the x_α, p_α supposed to be measured. In the simplest kinds of system it is sufficient to specify the total number of such pairs, i.e. the multiplicity. But it is necessary also to specify *the keenness of the probing*; and we shall now consider how this can be done.

In trying to lay down a standard procedure of measurement we encounter the well-known difficulty that observables do not in general commute. Our immediate concern is with the non-commutation of x_α and p_α, which makes it necessary to treat them together as a joint-observable. Our information as to the probability distributions of x_α and p_α corresponds to one type of measurement which we shall describe indifferently as measurement of the coordinate or of the momentum. The simplest way of specifying a standard 'degree of keenness' of this measurement is to require that the combined uncertainty of x_α and p_α shall have the minimum value fixed by Heisenberg's principle. The system is then said to be *fully observed* in this coordinate.

Suppose that the observational probing ceases at a time t_0, and we consider the probability distribution at a later time $t_0+\tau$. The combined uncertainty of x_α and p_α increases with τ, and the system ceases to be fully observed. This is illustrated by the expanding wave packet in § 34, which satisfies the condition of full observation at the initial instant; as t increases the dispersion σ_t of x_α increases, and there is no compensating decrease of the dispersion of p_α. *The time τ, measured from the instant of cessation of full observation, is a 'coefficient of under-observation' of the system.*

Just as we take the zero-temperature uranoid as the standard environment, so we take full observation as the standard probing. When it is necessary to consider a different environment, the disturbance of the uranoid is represented by an object-field which is specified by appropriate variates (potentials). Similarly a difference from the standard probing is represented by a 'field of under-observation', which must be specified by appropriate variates. The general theory of fields of under-observation might be developed as an analytical exercise; but so far as I can see it has no practical interest. I therefore confine attention to the one case of practical importance, namely, the uniform under-observation which results from abrupt cessation of the standard probing. This is specified by a single variate τ equal to the time-interval between the cessation of probing and the instant considered. If this is an observed time-interval it is subject to uncertainty, so that we have a probability distribution over τ; but we may also consider problems in which τ is given as a stabilised characteristic.

Since τ increases with the time $t = t_0+\tau$ that is being considered, there has been a great deal of confusion between τ and the coordinate t. The confusion is increased by taking t_0 as the origin of time, as in the current discussion of the expanding wave packet; so that the same symbol serves in two capacities, as a time-coordinate and as a measure of under-observation.

In statistical mechanics the time is distinguished from the other coordinates at the very outset. 'Probability distribution over time' has no intelligible meaning. A distribution function $f(x,y,z;t)$ gives the distribution *over* x,y,z *at* the time t. We can extend it to $f(x,y,z,\tau;t)$ giving the distribution over x, y, z, τ at the time t; but there is no homology between the measure of under-observation τ and the space coordinates.

If it were desired to treat the four coordinates homologously, the physical origin of time would have to be defined, in the same manner as the physical origin of space coordinates, as the centroid of a large number of events uncorrelated in their time distribution. That leads to mathematical phantasies which have no relation to physical experience. Moreover it would be necessary to postulate a negative σ^2 for the time direction (which could not be provided by any distribution of real events), because the relativistic homology is between space and imaginary time. But we can obtain a

relativistic extension of statistical mechanics by taking as the fourth dimension, not time, but the scale and phase dimension. The essential features of a natural environment are represented by a uranoid of particles with incoherent, i.e. uncorrelated, phase; so that by a straightforward extension of § 2 we obtain a physical origin which has Gaussian uncertainty of its phase coordinate as well as of its space coordinates. The scale fluctuation has to be combined negatively; so that the uncertainty of the origin in this direction is effectively $-\sigma^2$. The scale and phase have therefore a time-like relation to the spatial momenta and coordinates.

In 1928 Dirac made his famous relativistic extension of wave mechanics. The *wave vectors* which he introduced have a very general application, by no means confined to statistical theory; and we shall use them widely in Chapters V–VII for relativistic treatment either of the time coordinate or the phase or both together. But Dirac *wave functions*, which are used to represent probability distributions, are confined to statistical mechanics; and the common idea, that the fourth coordinate concerned in them is the time, is evidently a misapplication. As shown above, it is the phase that is relativistically connected with the space coordinates, like the time in non-statistical theory. The invariance of Dirac's wave equation is not Lorentz-invariance—not invariance for changes of motion of the system or the axes of reference. It is an invariance involving the phase coordinate which has formal analogy with Lorentz-invariance. We shall find later that the phase is the time analogue in the analogy between quantum and classical mechanics.

Thus we have to guard against the triple confusion which occurs in current theory, where the time may be either the coordinate t, or the measure of under-observation τ, or the time analogue (phase).

36. Structural and predictive theory

The problems of wave mechanics fall naturally into two classes, *structural* problems and *predictive* problems. When the theory is used for prediction we must suppose that observation has ceased some time before the epoch to which the prediction applies; and in such problems τ plays an important part. But when it is used to investigate structure—energy levels, magnetic moments, fundamental constants, etc.—the structure sought for is that of the system in the standard fully observed condition, and τ is not involved.[a]

Since concentrations of probability tend to disperse, predictive problems may also be described as 'problems of decay'. The expanding wave packet is the most elementary predictive problem, but it is scarcely typical; as regards method it seems to stand entirely by itself. The most usual practical problem of decay is: given the initial occupation of a set of eigenstates, to determine the change that will occur by transition during a subsequent period τ. Here the variates of the problem are the occupation factors.

We find the same separation in classical statistical mechanics. The theory of thermodynamic equilibrium is the structural part; and the calculation of the rate of decay of deviations from thermodynamic equilibrium is the predictive part. It is emphasised

[a] If 'prediction' is used in the sense of calculating results that can be verified experimentally, structural theory abounds in such predictions. But in the above classification these are excluded; structural theory does not foretell the activities of experimenters, and its results are calculations of what they can do, not predictions of what they will do.

that some of the most significant results, such as the Maxwell-Boltzmann laws, apply only to thermodynamic equilibrium; and I think it is unfortunate that the distinction has not been made equally clear in quantum theory. When comparing the equilibrium wave mechanics of the hydrogen atom with the non-equilibrium wave mechanics of an expanding wave packet, it is the dissimilarity rather than the affinity that should be emphasised.

Our investigations will lie almost wholly in structural theory. Thus the standard conditions of full observation are postulated, and τ is not involved. At a later point predictive theory is linked on. When statistical theory is applied to an isolated system in classical theory it is necessary to 'enclose' the system in order to keep the entropy constant. The enclosing boundary can be regarded as a mathematical constraint which continually restores the negative entropy that leaks away in natural conditions. Here the condition of full observation is the mathematical constraint which keeps the entropy constant. If the observation is not continually renewed, our information becomes stale; and the entropy, which is calculated from the probability relative to existing information, decays.

The standard procedure in wave mechanics is to represent physical change by changes of the occupation factors of steady states. I shall therefore take it as part of the fundamental definition of a wave function that it represents a steady state, the occupation factor alone being variable with time. If we admit unsteady states physical change is represented sometimes by continuous motion and sometimes by transitions; and the treatment becomes unsystematic. The only kinds of motion admitted in our representation are (1) steady circulation of probability within a state, (2) extra-spatial flow of probability between states, which may be steady or unsteady. A Lorentz transformation is contrary to the spirit of wave mechanics, since it introduces a progressive motion not coming into either of these categories.

When the theory is approached in this way there is not much danger of confusing the three kinds of 'time'. The only functions of the time coordinate are occupation factors; the time-like variate in the wave functions is the phase or time analogue. The time τ used as a measure of under-observation is set aside as peculiar to predictive theory.

The common practice is to merge the occupation factor in the wave function as its amplitude $j^{\frac{1}{2}}$. The common wave function $\Psi = \Sigma j_r^{\frac{1}{2}} \psi_r$ is then variable with the coordinate time. I think the introduction of these unsteady wave functions is unnecessary, confusing and, in practice, attended by serious lack of rigour. I would not say that this kind of extension is to be ruled out altogether; but I have no use for it here. Those who contemplate it should remember: (1) that it is perilous to apply the results derived from a study of statistical equilibrium to non-equilibrium conditions, and blind generalisation should not be attempted, (2) that the unsteady state has an expansion energy which must be duly taken into account.

37. Physical and geometrical distribution functions

We return to equation (1·1) to take up a new line of development. If the distribution function $f(x_0)$ of the coordinate x_0 of the physical origin is known, the distribution function $g(x)$ of the geometrical coordinate x of a particle can be converted into a

distribution function $h(\xi)$ of the physical coordinate ξ, or vice versa, by a method familiar in the theory of statistics.

Since x and x_0 are uncorrelated, and $x = x_0 + \xi$,

$$h(\xi) = \int_{-\infty}^{\infty} g(x_0+\xi) f(x_0)\, dx_0. \tag{37·1}$$

Denoting the Fourier integrals of f, g, h by F, G, H, we have the reciprocal relations

$$F(q) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iqx} f(x)\, dx, \tag{37·21}$$

$$f(x) = \int_{-\infty}^{\infty} e^{iqx} F(q)\, dq. \tag{37·22}$$

Then by (37·1)

$$\int_{-\infty}^{\infty} e^{-iq\xi} h(\xi)\, d\xi = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-iq(x_0+\xi)} g(x_0+\xi)\, e^{iqx_0} f(x_0)\, dx_0\, d\xi.$$

The right-hand side is separable, and by (37·21) we obtain

$$H(q) = 2\pi G(q)\, F(-q). \tag{37·3}$$

The physical origin has the Gaussian distribution function $f(x_0) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-x_0^2/2\sigma^2}$. This gives

$$2\pi F(-q) = e^{-\frac{1}{2}\sigma^2 q^2}. \tag{37·4}$$

Therefore

$$H(q) = e^{-\frac{1}{2}\sigma^2 q^2} G(q). \tag{37·5}$$

Thus from $h(\xi)$ we obtain successively $H(q)$, $G(q)$, $g(x)$ by (37·21), (37·5) and (37·22).

Usually we are more interested in momenta, and wish to transform a probability distribution of the physical momenta conjugate to ξ into a probability distribution of the geometrical momenta conjugate to x, or vice versa. For this purpose it is necessary to use wave functions, instead of distribution functions of the coordinates. They are related as follows:

The two real distribution functions of x and p are replaced by one complex wave function, which has reciprocal forms $\mathfrak{f}(x)$, $\mathfrak{F}(p)$ mathematically derivable from one another, the probabilities in the ranges dx, dp being respectively

$$2\pi\hbar\,|\,\mathfrak{f}(x)\,|^2\, dx, \quad 2\pi\hbar\,|\,\mathfrak{F}(p)\,|^2\, dp. \tag{37·6}$$

The forms are reciprocal Fourier integrals[a]

$$\mathfrak{f}(x) = (2\pi\hbar)^{-\frac{1}{2}}\int_{-\infty}^{\infty} e^{ipx/\hbar}\,\mathfrak{F}(p)\, dp, \tag{37·71}$$

$$\mathfrak{F}(p) = (2\pi\hbar)^{-\frac{1}{2}}\int_{-\infty}^{\infty} e^{-ipx/\hbar}\,\mathfrak{f}(x)\, dx. \tag{37·72}$$

Comparing (37·21) and (37·71), q corresponds to $p/\hbar$. By (37·6), $|\,\mathfrak{f}\,|$ is the square root of f; and it suggests itself that if we put

$$q = 2p/\hbar \tag{37·8}$$

[a] Dirac, *Quantum Mechanics*, 2nd ed., p. 103, eq. (41).

the previous analysis for distribution functions will apply to wave functions. Then (37·3) directly converts a distribution of geometrical momenta $G(2p/\hbar)$ into a distribution of physical momenta $H(2p/\hbar)$.

This interpretation of (37·3) proves to be correct. But, owing to the phase variable in the wave functions, the derivation is not straightforward. We shall see in the next section that a certain implicit convention is necessary in order to make the relation between geometrical and physical momenta determinate.

38. The weight function

Consider the case in which the coordinate x of the particle has a Gaussian distribution $g(x)$ with standard deviation s. Then $h(\xi)$ is a Gaussian distribution with standard deviation s', where

$$s'^2 = s^2 + \sigma^2. \tag{38·1}$$

For a Gaussian wave packet with standard deviation s, the distribution of momentum is found in current wave mechanics to be

$$G(p) = (\pi\hbar^2/2s^2)^{\frac{1}{2}}\, e^{-2s^2p^2/\hbar^2}. \tag{38·2}$$

The functions $F(p)$, $H(p)$ correspond similarly to σ, s'; so that by (38·1),

$$H(p) = \text{const.} \times G(p)\, F(-p). \tag{38·3}$$

Since the exponential in $F(-p)$ is $e^{-2\sigma^2p^2/\hbar^2}$, we have, by comparison with (37·4), $q = 2p/\hbar$ as already foreseen. The result (38·3) states:

The probability of a physical momentum p is the combined probability of a geometrical momentum p of the particle and an opposite momentum $-p$ of the physical origin. (38·4)

If we regard the particle and the physical origin as a two-particle system with momenta p, p' replaceable by an external momentum P and an internal momentum ϖ, the physical momentum of the particle is the internal momentum ϖ. The distribution functions of the four momenta must satisfy

$$K(P)\, H(\varpi) = G(p)\, F(p'), \tag{38·5}$$

since in the two-particle transformation $dP\,d\varpi = dp\,dp'$. Let $K(P) = \delta(P)$, where δ is Dirac's δ-function. Then, for all states with non-zero probability, $P = 0$; and by (26·13), $p = -p' = \varpi$. Thus (38·5) reduces to (38·3). This result is independent of the masses assigned to the particle and the physical origin; and (38·3) would evidently not be consistent with any other form of $K(P)$. Accordingly, (38·4) is equivalent to

$$K(P) = \delta(P). \tag{38·6}$$

The coordinate x conjugate to P is an unobservable geometrical coordinate, and it is clear that we cannot have proved (38·6). We have deduced it from the current theory of a wave packet; but that means that we have brought to light an implicit convention of wave mechanics. Since the same convention must be applied consistently throughout wave mechanics, we can remove the original restriction to a Gaussian distribution. Thus (38·3) is the general formula for converting distributions of geometrical momenta into distributions of physical momenta, and vice versa.

The convention is a consequence of the conception of wave mechanics as a system of description and prediction, employing data which consist wholly of observables such

as ξ, ϖ, although in the greater part of the analysis an auxiliary mathematical reference frame is used. Perhaps by a confusion of ideas, p has been left out of the equations as physically meaningless; but it is not meaningless as a variate connecting the auxiliary mathematical reference frame with the observables. The omission does no harm if p is always zero; so that we must accept (38·6) as an implicit convention made in introducing the auxiliary frame, and effectively forming part of its definition. The simplest view of the convention $p = 0$ is that, whenever a particle is given a momentum p a *recoil momentum* $-p$ is imparted to the physical origin. Or, stated in terms of probability distribution, the convention $K(p) = \delta(p)$ means that the probability of occurrence of a momentum p without a recoil momentum of the origin is zero.

We change the notation $F(-p)$ to $w(p)$. Since $F(-p)$ is given by (38·2) with σ in place of s,

$$w(p) = (2\pi\varpi^2)^{-\frac{1}{2}} e^{-p^2/2\varpi^2}, \tag{38·71}$$

where

$$\varpi = \hbar/2\sigma. \tag{38·72}$$

Then

$$H(p)\,dp = \text{const.} \times G(p)\,w(p)\,dp. \tag{38·8}$$

Thus *the distribution of geometrical momenta is turned into a distribution of physical momenta by weighing the ranges dp with the 'weight function' w.* In three dimensions the weight function for an element $dp_1 dp_2 dp_3$ is

$$w(p_1, p_2, p_3) = (2\pi\varpi^2)^{-\frac{3}{2}} e^{-(p_1^2+p_2^2+p_3^2)/2\varpi^2}. \tag{38·9}$$

The weight function is explained dynamically as a factor which makes allowance for the recoil momentum that must be supposed to be imparted to the physical reference frame when we assign a momentum p_1, p_2, p_3 to an object-particle. Later the theory of recoil will be extended to angular momenta; and it will be found to play an important part in the interpretation of the Riemann-Christoffel tensor.

The effect of $w(p)$ is to reduce the frequency of large momenta. The value of ϖ is about $200m_e c^2$; and, since the weight function has hitherto been ignored in practical investigations, there is scope for improving the current formulae for phenomena involving energy transitions of order $200m_e c^2$. It is not suggested that the formulae can be corrected by a crude application of (38·9) which refers to ideal V_3 particles. Attention must be paid to multiplicity factors, and the amendment is a matter for detailed investigation for each type of phenomenon.

Certain integrals, which converge correctly when the weight function is included, diverge when it is overlooked. As H. C. Corben has shown,[a] the supposed 'infinite transverse energy' of an electron has arisen from this oversight. To omit $w(p)$, or equivalently to set $\sigma = 0$, amounts to assuming an infinite universe, and introduces into wave mechanics the same difficulties of infinitude that relativity theory encountered and overcame a quarter of a century ago. There are no divergent integrals in the present theory; and we need not waste time over the difficulties that quantum physicists create for themselves by ignoring the guidance of relativity theory. The infinite universe has long been dead and buried; and those who insist on digging up the corpse must expect to be haunted.

[a] *Proc. Camb. Phil. Soc.* **35**, 195, 1939. Corben's treatment requires some modification, since he does not distinguish sufficiently between distribution functions and wave functions; and he introduces an uncertainty of the time origin, which is illegitimate unless the time analogue is meant.

39. The genesis of proper mass

In the uniform steady distributions hitherto considered we have taken the particles to be at rest. Let us now consider the opposite extreme, namely a distribution in which the geometrical momenta have unlimited uniform probability distribution; that is to say, the number of particles in a range $dp_1dp_2dp_3$ is proportional to $dp_1dp_2dp_3$ for all values of p_1, p_2, p_3 from $-\infty$ to $+\infty$. According to classical statistical mechanics this is the limit which is approached as the temperature tends to infinity. We shall therefore call it the *infinite-temperature uranoid.*

The corresponding distribution of physical momenta is $wdp_1dp_2dp_3$, where w is the weight function (38·9). The mean values of p_1^2, p_2^2, p_3^2 in this distribution are

$$\overline{p_1^2} = \overline{p_2^2} = \overline{p_3^2} = \varpi^2 = \hbar^2/4\sigma^2. \tag{39·1}$$

Since the momenta are large, the accurate energy formula $E^2 = m^2+p_1^2+p_2^2+p_3^2$ must be used. For a reason which will appear presently we consider particles without proper mass, so that $E^2 = p_1^2+p_2^2+p_3^2$. Hence

$$\overline{E^2} = 3\varpi^2 = 3\hbar^2/4\sigma^2. \tag{39·2}$$

By the well-known formulae of molar relativity theory,[a] the pressure P and density ρ of a steady uniform distribution of matter satisfy

$$8\pi\kappa P = -R_0^{-2}+\lambda, \quad 8\pi\kappa\rho = 3R_0^{-2}-\lambda. \tag{39·3}$$

Here R_0 is the radius of space curvature, and λ (which it would be premature to identify with the cosmical constant) is of the nature of a constant of integration. When the total number of particles is fixed R_0 is the fixed constant $2\sigma\sqrt{N}$ by (3·8), but λ is disposable. By varying λ we vary P/ρ, or equivalently the temperature. Let

$$\Theta = \frac{\rho-\rho_0}{\rho_0} = \frac{\text{density of thermal energy}}{\text{density of proper energy}}. \tag{39·41}$$

The proper density ρ_0 is equal to $\rho-3P$. Hence by (39·3),

$$\Theta = \frac{3(\lambda-R_0^{-2})}{6R_0^{-2}-4\lambda}. \tag{39·42}$$

As λ increases from R_0^{-2} to $\frac{3}{2}R_0^{-2}$, Θ increases steadily from 0 to ∞. From the definition of Θ in (39·41) it is appropriate to adopt it as a measure of temperature in relativity theory. In classical theory infinite temperature involves infinite energy, and the temperature is made infinite by taking ρ infinite in (39·41). But in relativity theory an increase of temperature diminishes ρ_0; and the temperature becomes infinite by the vanishing of ρ_0, whilst ρ is still finite. Owing to the vanishing of ρ_0, the particles in the infinite temperature uranoid have zero proper mass—a conclusion that has been anticipated in our calculation of $\overline{E^2}$ in (39·2).

We can now compare the constants of an infinite-temperature uranoid and a zero-temperature uranoid, distinguishing the former by an accent. By (39·3) and (39·42) we obtain

$$\Theta' = \infty, \quad \lambda' = \tfrac{3}{2}R_0^{-2}, \quad 4\pi\kappa\rho' = \tfrac{3}{4}R_0^{-2}, \quad \rho_0' = 0, \tag{39·51}$$

$$\Theta = 0, \quad \lambda = R_0^{-2}, \quad 4\pi\kappa\rho = R_0^{-2}, \quad \rho_0 = \rho. \tag{39·52}$$

Hence

$$\rho_0 = \rho = \tfrac{4}{3}\rho'. \tag{39·6}$$

[a] Tolman, *Relativity, Thermodynamics and Cosmology*, § 139.

Proper mass is a concealed form of energy, and a fundamental theory should show how energy comes to be concealed in this way. We have therefore started with an infinite-temperature uranoid in which there is no concealed energy; and we see that by lowering the temperature a proper density ρ_0 representing concealed energy is created. Let us examine more closely how the change of temperature from ∞ to 0 can give a particle a proper mass m.

The fact that the object-particle itself is reduced to rest by the lowering of the temperature is irrelevant; for the proper mass is an invariant independent of the motion of the particle. It is the temperature of the environment that matters. The effect can be explained roughly by Newtonian theory. The object-particle is in the inertial-gravitational field of the rest of the universe. The greater the mass of the environment, the greater is the strength of the field, and the greater numerically is the negative potential energy of a particle at rest in the field. By decreasing the temperature we decrease the energy or mass of the environment, and decrease the negative potential energy of the object-particle. The subtracted negative energy, regarded as an added positive energy, constitutes the proper mass of the particle in the zero-temperature environment. Substituting Einsteinian for Newtonian theory, the foregoing treatment provides a rigorous calculation of the addition.

Must we then vary the proper mass of a particle according to the temperature of its surroundings? Certainly the change of gravitational potential due to modified energy of the surroundings must, if it is sensible, be taken into account in some way. But the correction is recognised explicitly as gravitational potential energy, and is not included in the proper mass. The proper mass of a particle does not depend on the temperature of its actual environment; it depends on the temperature of the environment postulated as standard. This is the conventional distinction between gravitation and inertia already formulated; the inertial field, which determines the proper mass, corresponds to the standard uranoid, and the gravitational field corresponds to the deviation of the actual environment from the standard uranoid.[a]

The value of λ which corresponds to the zero-temperature uranoid, namely $1/R_0^2$, is called the *cosmical constant*.

The molar energy tensor of the distribution is made up of the self energy tensors of the particles composing it. By (22·2) the self energy tensor of a particle has the form $A p_\mu p_\nu$. Here p_μ is specified by a probability distribution and the mean contribution of a particle to the energy tensor is

$$T_{\mu\nu} = A\overline{p_\mu p_\nu}. \tag{39·71}$$

For a particle in the infinite temperature uranoid this gives $T_{11} = A\varpi^2$, $T_{44} = 3A\varpi^2$ by (39·1) and (39·2), E being now denoted by p_4. Summing for the N particles the total pressure (T_{11}) and density (T_{44}) are

$$P' = NA\varpi^2, \quad \rho' = 3NA\varpi^2. \tag{39·72}$$

[a] This meets a possible criticism of our argument. It might be objected that, since a change of the temperature of the universe from 0 to ∞ would involve a violation of the conservation of energy, no conclusion can be based on it. But we do not tamper with the actual universe; we only consider the effect of adopting an infinite-temperature distribution instead of a zero-temperature distribution as the standard by which gravitation is distinguished from inertia.

In the zero-temperature uranoid the particles are all reduced to rest, so that the momentum vector becomes $(0, 0, 0, m)$; and the total pressure and density, calculated from (39·71), are

$$P = 0, \quad \rho = NAm^2. \tag{39·73}$$

Hence $\rho/\rho' = m^2/3\varpi^2$. But, by (39·6), $\rho/\rho' = \frac{4}{3}$. Hence $m^2 = 4\varpi^2$, and

$$m = 2\varpi = \hbar/\sigma \tag{39·81}$$

by (38·72). Also, by (39·52) and (39·73), $1/4\pi\kappa R_0^2 = \rho = AN\hbar^2/\sigma^2$. Since $R_0 = 2\sigma\sqrt{N}$, this gives

$$A = 1/16\pi\kappa\hbar^2N^2, \tag{39·82}$$

or in natural units, $A = 1/2N^2$.

In applying the last two formulae to practical calculation we have to remember that the current quantum equations refer to superpositions on a rigid field, and incorporate the multiplicity factors thereby introduced in the adopted values of the constants. For this reason the simple theoretical constant $\hbar$ used in this section will not agree with the practical Planck constant $\hbar$. The relation will be investigated in the next section. In order to avoid a clash of notation we shall denote the constant in the foregoing investigation by γ; so that $A = 1/16\pi\kappa\gamma^2N^2$, and

$$m = \gamma/\sigma = 2\gamma\sqrt{N}/R_0. \tag{39·9}$$

40. Absolute determination of m_0

The particles in the foregoing analysis are V_3 particles; for we have considered only the triple probability distribution of coordinates and momenta. As pointed out in § 20, the inclusion of coordinates as well as momenta does not affect the multiplicity. This is also shown directly in § 38. When the distribution function giving the probability distribution of the coordinates is replaced by the wave function which gives in addition the probability distribution of momenta, the number or degrees of freedom is increased initially from 3 to 6; but the convention $K(p) = \delta(p)$ eliminates 3 degrees of freedom, so that *the probability distribution of momentum is uniquely determined by the probability distribution of coordinates, and vice versa.*

If any other characteristics, besides p_1, p_2, p_3 and the conjugate coordinates, are carried by the particles these must be stabilised. In particular the proper mass, calculated in (39·9), is stabilised. Its computed value is a function of the temperature of the environment; so that in a sense a measurement of the mean temperature of the universe would be an observational determination of m. But the temperature concerned is that of the standard environment which we postulate, not that of the actual environment which we might measure; so that m is a stabilised characteristic, not an observable.

The uncertainty of scale can be taken into account directly or replaced by space curvature. In the cosmological investigation in § 39 it has been replaced by space curvature. Normally in wave mechanics, which is not adapted to curved space, the proper course is to take it into account directly; so that m has a probability distribution arising from the uncertainty of the mass standard (comparison particle) with which it

is measured. The particle is then a V_4. Denoting the masses of V_3 and V_4 particles by m_3, m_4, the mass given by (39·9) is m_3; and, by (16·5), $m_4 = \frac{3}{4}m_3$. Thus

$$m_3 = 2\gamma\sqrt{N}/R_0, \quad m_4 = \tfrac{3}{2}\gamma\sqrt{N}/R_0. \tag{40·1}$$

This is confirmed by the following considerations.

The relation $\rho_0 = \frac{4}{3}\rho'$ in (39·6) implies that when proper mass is released as unconcealed energy, a quarter of the energy disappears. It is easily verified from (39·3) that this continues to hold when only part of the proper mass is released, corresponding to a relatively small change of Θ. It should therefore apply to nuclear transformations. The experimenter tells us that there is no energy loss in a nuclear transformation. But this is not really a contradiction, because he does not measure the changed gravitational energy of the extra-galactic nebulae—which is where the loss occurs. We can interpret the experiment more simply. In order to justify his assertion the experimenter must measure the mass of the particle before and after the transmutation. We cannot even imagine this experiment to be performed on a V_3 particle; because if m is *measured* the particle is *ipso facto* not a V_3. The introduction of a measurable mass turns it into a V_4 particle, and the mass measured is $\frac{3}{4}m_3$. This anticipates the loss of $\frac{1}{4}$ of the energy m_3 in the subsequent transmutation; so that the conservation of energy is found to be experimentally satisfied.

It is only by imposing an unnatural constraint on matter, so as to admit only irrotational motion, that we are able to analyse it into V_3 or V_4 particles. To obtain the constants of the actual universe, we must consider V_{10} particles of mass $\mu = \frac{3}{10}m_3$; so that by (40·1),

$$\mu = \tfrac{3}{5}\gamma\sqrt{N}/R_0. \tag{40·2}$$

To find the relation between γ and $\hbar$, consider a wave function corresponding to a state of small velocity. The momenta $p_\alpha = -i\gamma\partial/\partial x_\alpha$ used in (40·2) have been assumed to be such that $E^2 = m^2 + p_1^2 + p_2^2 + p_3^2$; so that for a V_{10} with small velocity,

$$E = \mu + \frac{p_1^2+p_2^2+p_3^2}{2\mu} = \mu - \frac{\gamma^2\nabla^2}{2\mu}. \tag{40·31}$$

In the rigid-field treatment we introduce separate carriers for initial and transition energy, and by (18·4) the momenta $p_\alpha = -i\hbar\partial/\partial x_\alpha$ are such that

$$E = \mu + \frac{p_1^2+p_2^2+p_3^2}{2\mu} = \mu - \frac{\hbar^2\nabla^2}{2\mu}. \tag{40·32}$$

Comparing the two expressions,

$$\gamma = \hbar\sqrt{(\mu/\mu)} = 136\hbar/\sqrt{10}, \tag{40·4}$$

by (18·6).

A correction is needed, because in separating the carriers of initial and transition energy we double the number of particles. We should therefore start with $\frac{1}{2}N$ particles in § 39 and in (40·2), so as to obtain the N particles of the actual universe. Keeping R_0 fixed and equal to the radius of the actual uranoid, the change $N \to \frac{1}{2}N$ gives $\sigma \to \sigma\sqrt{2}$, and all quantum-specified lengths are thereby multiplied by $\sqrt{2}$. Correspondingly, quantum-specified angular momenta, being of dimensions $(\text{length})^{-2}$, are divided

by 2, so that $\gamma \to \frac{1}{2}\gamma$. We must therefore substitute $\frac{1}{2}N$, $\frac{1}{2}\gamma$ for N, γ in § 39 and (40·2). The latter formula becomes

$$M = \frac{3}{5}\frac{\gamma\sqrt{(\frac{1}{2}N)}}{R_0}. \tag{40·5}$$

Alternatively we can keep γ, N and R_0 the same as in the actual uranoid. Then the value of M given by (40·2) refers to a uranoid with N extracules of mass μ, to which are added N intracules of zero rest mass. The doubling of the number of particles divides σ by $\sqrt{2}$; and since mass has dimensions (length)$^{-3}$, it multiplies M by $2\sqrt{2}$. We have therefore to divide the result in (40·2) by $2\sqrt{2}$, in order to obtain the mass in a uranoid with $\frac{1}{2}N$ extracules and $\frac{1}{2}N$ intracules. This gives (40·5).

By (40·4) and (40·5)
$$M = \frac{136}{10}\frac{3}{4}\frac{\hbar\sqrt{(\frac{4}{5}N)}}{R_0}, \tag{40·6}$$

so that
$$m_0 = \frac{3}{4}\frac{\hbar\sqrt{(\frac{4}{5}N)}}{R_0}. \tag{40·7}$$

It remains to determine the β-factor required to reduce (40·5) from the theoretical to the observational system. We have used the ratio $M/\mu = 136^2/10$, which applies to standard particles in system A or hydrocules in system B. The latter is the definitive application, since it takes account of electrical properties. Thus the radius referred to in (40·6) and (40·7) is $(R_0)_B$, which is $\beta^{-\frac{1}{2}}(R_0)_A$; and R_0 must be replaced by $\beta^{-\frac{1}{2}}R_0$ in the formulae, in order to reduce to the observational system. No other change is required. The mass m_0 determined by the formula is that of the hydrocule in system B; but this is the same as that of the standard particle in system A, and, since it is the standard particles that are directly connected with the molar density or mass, the result is an absolute determination of the mass of the hydrogen atom in the observational system A. The formulae, including β-factors, are accordingly

$$m_0 = \frac{3}{4}\frac{\beta^{\frac{1}{2}}\hbar\sqrt{(\frac{4}{5}N)}}{cR_0}, \quad M = \frac{136}{10}\frac{3}{4}\frac{\beta^{\frac{1}{2}}\hbar\sqrt{(\frac{4}{5}N)}}{cR_0}, \tag{40·8}$$

the velocity of light c being re-inserted so as to remove the restriction on the units.

This is the central formula of unified theory.[a] Since $\hbar/Mc$ is known with great accuracy (§ 32), (40·8) gives an accurate determination of $R_0/\sqrt{N}$ and hence of σ. This replaces the rough determinations of σ from the range of nuclear forces and the recession of the galaxies in § 5. The calculation of natural constants with this formula and comparison with observation will be treated later.

In the next three sections the same formula will be derived by an entirely different method.

41. Exclusion

The earliest version of the exclusion principle refers to the discrete eigenstates of an atom, and asserts that there cannot be more than one particle in a given eigenstate.

[a] My earlier calculation (*Protons and Electrons*, equation (14·71)) was in error by a factor $(\frac{4}{5})^{\frac{1}{2}}$. The correction was announced tentatively by Eddington and Thaxton in *Physica*, **7**, 122, 1940. The factor $\beta^{\frac{1}{2}}$ is a more recent refinement.

We are not concerned with this form of the principle at present, because it is a truism for the hydrogen atom which has only one internal particle. But there is a more general version applicable to a continuum of states, which (as applied to hydrogen) can be expressed as follows:

If the 6-space obtained by taking x, y, z, p_1, p_2, p_3 as coordinates is divided into cells of volume h^3 ($h = 2\pi\hbar$), then in a steady state the maximum number of particles per cell is two electrons and two protons. (41·11)

By applying the two-particle transformation, and using the invariant relations (26·32) and (26·34), we derive a similar exclusion principle for extracules and intracules. The wording is modified because the variates ξ, η, ζ, ϖ_1, ϖ_2, ϖ_3 of the intracule form a distinct (electrical) 6-space. The exclusion principle is:

In a steady state the maximum number of particles is two extracules per cell h^3 of xp-space and two intracules per cell h^3 of $\xi\varpi$-space. (41·12)

This divides the principle into an electrical exclusion principle applying to intracules and a mechanical exclusion principle applying to extracules. The theoretical basis of the exclusion principle will be found at a later stage of the development of the theory. Meanwhile we accept it as a part of quantum theory which it is our task to unify with the rest of physics. Such experimental verification as exists has hitherto been confined to the electrical exclusion principle. Unless the mechanical principle is true also, the mixed form (41·11) cannot be true; but to use this as an argument for the mechanical principle would beg the question whether the current belief in (41·11) is well founded. The following investigations will lead to observational confirmation of the mechanical exclusion principle, and so redress the balance.

We are going to show that the mechanical exclusion principle leads to the formula (40·8) for the masses of particles that we have found by gravitational theory. Thus we can replace our former gravitating (non-excluding) particles by excluding (non-gravitating) particles. *Exclusion is a wave mechanical substitute for gravitation.* It is a wider concept, since it has both a mechanical and electrical application; so that we should perhaps rather say that gravitation (including inertia), i.e. curvature of space, is the form in which molar theory represents the mechanical part of exclusion. But it is also a narrower concept, since the exclusion principle postulates a steady state. This is in accordance with our general conclusion (§ 6) that the methods of molar and of microscopic theory have a very narrow field of overlap.

The application of the exclusion principle to extracules is parallel to the application to intracules in the theory of super-dense (white-dwarf) matter; so that the first few steps are familiar. Super-dense matter is treated in § 44.

Consider a unit volume of three-dimensional space, so that the cell of 6-space corresponds to a cell $dp_1 dp_2 dp_3 = h^3$ of momentum space. Let n be the number of extracules per unit volume of three-dimensional space; and let μ_0 be their mass-constant, so that according to the classical formula the kinetic energy of an extracule is

$$E = p^2/2\mu_0 \quad (p_0^2 = p_1^2 + p_2^2 + p_3^2).$$

For a zero-temperature distribution the energy is a minimum. We have therefore to distribute p_1, p_2, p_3 so that ΣE, and therefore Σp^2 is a minimum subject to the density

not exceeding two particles per cell. The minimum is given by a distribution in momentum space which fills uniformly a sphere of radius $p = \mathfrak{p}$ determined by

$$\tfrac{4}{3}\pi\mathfrak{p}^3 = \tfrac{1}{2}nh^3, \tag{41·21}$$

and the 'top energy' is

$$\mathfrak{E} = \frac{\mathfrak{p}^2}{2\mu_0} = \left(\frac{3n}{8\pi}\right)^{\frac{2}{3}} \frac{h^2}{2\mu_0}. \tag{41·22}$$

Also the mean energy is

$$\bar{E} = \tfrac{3}{5}\mathfrak{E} = \frac{3}{5}\left(\frac{3n}{8\pi}\right)^{\frac{2}{3}} \frac{h^2}{2\mu_0}. \tag{41·23}$$

This is a well-known formula in the theory of white-dwarf matter. Applying it to the $\frac{1}{2}N$ extracules of the zero-temperature uranoid, we have $n = \frac{1}{2}N/2\pi^2R_0^3$. Hence, setting

$$\mu_1 = \mu_0(\tfrac{3}{4}N)^{\frac{1}{3}} \tag{41·3}$$

(41·22) becomes

$$\mathfrak{E} = \frac{3}{4}\frac{N}{2\mu_1}\left(\frac{h}{2\pi R_0}\right)^2. \tag{41·41}$$

This may also be written as

$$2\mu_1\mathfrak{E} = \tfrac{3}{4}\varpi^2, \tag{41·42}$$

where ϖ is the weight constant (38·72), which is equal to $\hbar/2\sigma$ or $h\sqrt{N}/2\pi R_0$.

Consider a top particle. It is one of a distribution of particles at zero temperature, and therefore from the ordinary point of view $\mathfrak{E}$ is its rest mass. $\mathfrak{E}$ is here exhibited as exclusion energy, the particle being boosted up to this energy level by the full occupation of all states of lower energy. But in any case the rest mass is a concealed form of energy, and it is the same energy that we formerly accounted for gravitationally. According to one picture the particles of the uranoid give an object-particle its rest energy by determining an extraordinary fluctuation of scale, which gives curvature to space and consequently provides an inertial-gravitational field; according to the other picture they compete for the lowest energy levels, so that (on the average) an object-particle has a high energy although the temperature is zero. Ultimately the two modes of interaction must be equivalent.

We shall distinguish this exclusion treatment of rest mass as 'sub-threshold theory'. As soon as we pass to ordinary quantum mechanics, the rest mass is accepted as an intrinsic attribute of the particle. In particular the original mass-constant μ_0 is peculiar to sub-threshold theory, and is replaced in quantum mechanics by a mass-constant equal to $\mathfrak{E}$.[a] It will be seen that, although the calculations are precisely similar, there is a difference in the practical interpretation of the present result for extracules and the familiar result for intracules in white-dwarf matter, because the plane of ordinary experience is the top level of the former and the bottom level of the latter.

42. The negative energy levels

Regarding the top level $\mathfrak{E}$ as zero level, the particles of the zero-temperature uranoid completely fill a series of negative energy states existing below the zero level. In principle this agrees with a conception first introduced by Dirac; but it must be emphasised that there are important differences. Here the number of negative energy

[a] Subject to such adjustment (to be considered later) as may be necessary to allow for the fact that not all the particles have the top energy.

states, though large, is finite. The particles filling them are not invented *ad hoc*, but are the rest of the particles of the universe distributed in the way in which the elementary quantum equations postulate them to be, namely so as not to interfere with the behaviour of the object-system that the equations describe. Since the particles are extracules, the theory is symmetrical between positive and negative charge; and the difficulty of the infinite negative charge, which occurs in Dirac's theory, does not arise. Presumably negatrons as well as positrons can exist.

This picture refers to sub-threshold theory; positrons and negatrons belong to super-threshold theory and, although related to the vacant states or 'holes' in the sub-threshold picture, are not directly identified with them. Similarly electrons and protons, which are super-threshold particles, are not directly identified with particles in the sub-threshold picture.

To form a super-threshold object-particle, we must excite a top particle to a higher energy. Then the net addition to the uranoid is a particle at a level above $\mathfrak{E}$ together with a hole at the level $\mathfrak{E}$. This combination can be regarded as a bi-particle carrying the excitation energy. We arrive at once at the theory of the mass of a proton or electron in § 22, in the form given at the end of that section where a comparison hole is used. The bi-particle of multiplicity 136 carries an unspecialised element of the excitation energy tensor added to the zero-temperature uranoid; and this is divided into two parts, one part being stabilised as a V_{10} particle (proton or electron) and the other as a V_1 hole. The hole at the level $\mathfrak{E}$ is accordingly the comparison hole.

A super-threshold positron or negatron is formed similarly by taking a hole at the level $\mathfrak{E}$ and exciting it to a lower level. The net addition is a bi-particle consisting of a hole at the lower level and a comparison particle plugging the original hole. The calculation of the masses is precisely similar; so that the masses of the positron and negatron are equal to those of the electron and proton.

The connection between the particles in sub-threshold theory and those previously studied is therefore:

(1) The top level particles of exclusion theory are the comparison particles of our previous theory.

(2) The stabilisation which yields V_{10} particles is not introduced until energy is added to the uranoid in order to form super-threshold object-systems. Thus the extracules in § 41 are unspecialised elements of energy tensor with multiplicity 136.

Both (1) and (2) identify $\mathfrak{E}$ with m_0.

We have reached the conception of an object-system perched on a firm platform provided by fully packed energy levels; so that it starts with a threshold energy highly boosted up by exclusion. It is essential to employ top particles in forming object-systems; otherwise holes are left in the sub-structure, and it does not constitute the rigid environment which wave mechanics postulates. By (41·23) the top particles have $\frac{5}{3}$ of the mean energy $\bar{E}$. The multiplier $\frac{5}{3}$ may be described as a *selection factor*. By attributing the top mass $m_0 = \mathfrak{E}$ to every particle we magnify the density of the uranoid by $\frac{5}{3}$ in passing from sub-threshold to super-threshold theory. In order that R_0 may be unchanged the increased density must be coupled with a constant of gravitation reduced in the ratio $\frac{3}{5}$. The constant of gravitation applicable to sub-threshold theory is accordingly

$$\kappa_1 = \tfrac{5}{3}\kappa. \qquad (42\cdot1)$$

The name 'constant of gravitation' is inappropriate when gravitation is replaced by exclusion; but the interpretation is given by the relation $8\pi\kappa\hbar^2 = 1$, and we have

$$\hbar_1^2 = \tfrac{3}{5}\hbar^2. \tag{42·2}$$

Thus $\hbar_1$ is the constant to be used in sub-threshold theory, $\hbar$ being the ordinary super-threshold constant.

A slightly different derivation of (42·2) is obtained by considering that $\hbar$ determines the size of a unit cell in exclusion theory. Thus in the analysis of a unit volume in § 41, when $\hbar_1$ is changed to $\hbar$, the volume $dp_1dp_2dp_3$ of the unit cell of momentum space is changed in the ratio $\hbar^3/\hbar_1^3$; so that corresponding values of p are changed in the ratio $\hbar/\hbar_1$ and those of E in the ratio $\hbar^2/\hbar_1^2$. This increase changes $\bar{E}$ to $\mathfrak{E}$. It therefore corresponds to the change of total energy introduced by the selection of top particles.

43. Determination of m_0 by exclusion theory

Let us compare uranoids containing different numbers of particles. We shall make the comparison from the plane of ordinary experience, which is the level $\mathfrak{E}$; so that the uranoids will be viewed from above. That is to say, we take a small object-system to which the usual microscopic constants m_0, $\mathcal{M}$, μ, $\hbar$, σ, ϖ apply, and vary the number $\frac{1}{2}N$ of negative energy levels beneath the threshold. The cosmological constant R_0 will vary as $N^{\frac{1}{2}}$; but otherwise the super-threshold system will be unaffected. Since $m_0 = \mathfrak{E}$, $\mathfrak{E}$ will be unchanged. The constant μ_0 will be changed, but that does not affect the super-threshold system.

By (41·42), μ_1 is unchanged. For further progress it is necessary to determine $\mu_1/\mathfrak{E}$; and, since this ratio is independent of N, it can be found once for all by taking any convenient value of N. The case that can be solved easily is $\frac{1}{2}N = 2$. But we have previously assumed that N is large, and for small values our formulae become ragged. This will not affect the order of magnitude; but to obtain accurate formulae it is necessary to examine more closely how the negative energy levels are distributed.

Owing to the symmetry the eigenstates will correspond to the surface harmonics of the hypersphere of space.[a] By a well-known theorem each single eigenstate in a complete set corresponds to a unit cell. Owing to the symmetry, harmonics of the same order coalesce into a multiple state; so that the cells are arranged in a series of levels, the kth level from the bottom corresponding to harmonics of order $s = k-1$. For surface harmonics on a hypersphere there are $(s+1)^2$ independent harmonics of order s. Thus the kth level contains k^2 cells or $2k^2$ extracules. The bottom level, corresponding to a surface harmonic of zero order, consists of one cell containing two extracules.

We call k the *quantum number* of the level. When k is large, the number of extracules up to the kth level is $\frac{2}{3}k^3$; and the top quantum number $\mathfrak{k}$ is given by $\frac{2}{3}\mathfrak{k}^3 = \frac{1}{2}N$. Thus

$$\mathfrak{k} = (\tfrac{3}{4}N)^{\frac{1}{3}}. \tag{43·1}$$

Hence by (41·3)

$$\mu_1 = \mathfrak{k}\mu_0. \tag{43·2}$$

The top quantum number of the universe is $2{\cdot}607 \times 10^{26}$, and the particles in the top-most complete level would form a planet of mass intermediate between Neptune and Saturn.

[a] The eigenstates are determined by the same conditions as in a symmetrical field of force in atomic physics, except that there is one more dimension. In both cases there is symmetrical degeneracy, due to the impossibility of defining any distinctive orientation of a set of rectangular axes.

The formulae (41·3) and (43·1) are not valid for small numbers; but the raggedness is eliminated in the derived formula (43·2), which remains true down to $\mathfrak{k} = 1$. That is to say, the mechanical characteristics E, p of the level are smoothly related to the quantum number k, although their relation to the maximum occupation becomes irregular at low levels. When the curvature representation is dropped, scale and phase variates are used instead; and in an eigenstate of the system the scale must reduce to an eigenscale. In the uranoid the quantum levels are the (multiple) eigenstates of the particles; each represents a uniform symmetrical distribution and the only difference between the levels is a difference of scale.[a] By (24·3) the eigenvalues of the scale are proportional to the integers. The lowest scale, corresponding to the integer 1, gives the bottom level in the uranoid; and the eigenscale of any other level is then proportional to its quantum number k. Any mechanical characteristic varies as a fixed power of the scale depending on its dimension-index. Thus the simple relations of the mechanical characteristics to k, established for high quantum number, remain valid down to $k = 1$.

For the kth level in a uranoid with top quantum number $\mathfrak{k}$, we now have the exact formulae

$$E/\mathfrak{E} = k^2/\mathfrak{k}^2, \quad p/\mathfrak{p} = k/\mathfrak{k}, \tag{43·31}$$

and the number of extracules in the level is $2k^2$. Comparing uranoids with the same microscopic constants but different $\mathfrak{k}$, the exact formulae are

$$\mathfrak{E} = \text{const.}, \quad \mu_1 = \text{const.}, \quad \mu_0 = \mu_1/\mathfrak{k}, \tag{43·32}$$

and by (41·41),

$$2\mu_1\mathfrak{E} = \mathfrak{k}^3(\hbar_1/R_0)^2, \tag{43·4}$$

remembering that a different constant $\hbar_1$ is to be used in sub-threshold theory.

To justify the treatment in wave mechanics, a quantum particle has been defined as an addition to a rigid environment. It is a feature of the exclusion representation that this condition is automatically fulfilled by taking the quantum particle to be a top particle, which can be inserted or removed without disturbing the fully packed energy levels beneath. Thus the replacement of gravitating by excluding particles replaces the gravitational field by an exclusion field which, unlike the gravitational field, is automatically rigid. This is a merit, and also a limitation; for it prevents the adaptation of exclusion treatment to the majority of molar problems which involve a flexible field.

The rigidity of the exclusion field is not exact, though normally it is an extremely high approximation. When N is not large, we must distinguish between *vertical* and *lateral* exclusion. The part of $\mathfrak{E}$ due to lateral exclusion, i.e. due to the $2\mathfrak{k}^2 - 1$ particles at the same level as the object-particle, is negligible in the actual universe; but it becomes relatively more important as $\mathfrak{k}$ decreases and for $\mathfrak{k} = 1$ the exclusion is wholly lateral. The uranoid then consists of two particles which mutually exclude one another from the zero state $E = 0$. Either particle is a top particle; but it is not in a rigid environment, since its removal would allow the other particle to drop to $E = 0$. When treated as a quantum particle superposed on a rigid environment, its energy must be taken to be $\mathfrak{E} = C + \mathfrak{p}^2/2\mu_0$, where C is the additional energy of the system due to its presence but not included in its own exclusion energy $\mathfrak{p}^2/2\mu_0$. In this simple case C is the exclusion energy $\mathfrak{p}^2/2\mu_0$ of the other particle; so that $\mathfrak{E} = 2C$.

[a] The surface harmonics were introduced only for the purpose of calculating the multiplicity of the uniform distribution which, in symmetrically degenerate conditions, replaces them.

We have hitherto included only kinetic energy in E. Including rest energy, the classical formula is $E = \mu_0 + p^2/2\mu_0$. Since this has to reduce to $\mathfrak{E} = C + p^2/2\mu_0$ at the one-quantum level, we have $C = \mu_0$. Thus in the uranoid $\mathfrak{k} = 1$,

$$\mathfrak{E} = 2\mu_0 = 2\mu_1. \tag{43·5}$$

But $\mathfrak{E}$ and μ_1 are independent of $\mathfrak{k}$, so that $\mathfrak{E} = 2\mu_1$ in any uranoid. Thus (43·4) becomes

$$\mathfrak{E}^2 = \mathfrak{k}^3(\hbar_1/R_0)^2 = \tfrac{3}{4}N(\hbar_1/R_0)^2, \tag{43·6}$$

the latter form being valid in the actual universe where N is very large. By (42·2) this becomes

$$\mathfrak{E}^2 = m_0^2 = \frac{3}{4}\cdot\frac{3}{5}\frac{\hbar^2 N}{R_0^2}, \tag{43·7}$$

which agrees with our previous determination of m_0 in (40·7). The factor $\beta^{\frac{1}{2}}$, reducing the value of m_0 to the observational system, should be inserted in practical calculation as before.

It will be noticed that the classical hamiltonian $E = m + p^2/2m$, which has hitherto been used only for small values of p/m, is in this investigation applied to extremely large values of p/m. It will be shown in § 45 that $m + p^2/2m$ is the correct hamiltonian for the standing waves which represent a system in statistical equilibrium, just as $(m^2 + p^2)^{\frac{1}{2}}$ is the correct hamiltonian for progressive waves. The common description of the hamiltonian $m + p^2/2m$ as 'non-relativistic' is absurd; the argument which shows that $(m^2 + p^2)^{\frac{1}{2}}$ is the correct form of hamiltonian for progressive waves because they *can* be eliminated by a Lorentz transformation of the coordinate frame, equally shows that it is not the correct form of hamiltonian for standing waves because they *cannot* be eliminated by a Lorentz transformation of the coordinate frame. There is therefore no reason to discredit the hamiltonian $m + p^2/2m$ in advance; and the agreement of the derived formula (43·7) with the earlier result obtained by relativity theory (which, we shall find, is verified experimentally to about 1 part in 1000) is a very stringent test of its correctness.

44. Super-dense matter

The mean density of the companion of Sirius is about 50,000 g.cm.$^{-3}$; and other white-dwarf stars with still higher density are known. At such densities the electrons and protons (or nuclei) are in a state of continuous collision; and the problem of separating carriers of electrical and of mechanical energy is radically different from the separation of intracules and extracules in isolated two-particle systems. We shall investigate super-dense matter which is in a steady state, and therefore in statistical equilibrium.

Consider first the molar aspect. The deviation of any physical characteristic from perfect uniformity can be represented by a superposition of waves. There are two wave systems to be considered:

(*a*) Mechanical waves, i.e. sound waves.

(*b*) Electrical waves, i.e. changes of electric polarisation with consequent currents and magnetic effects.

To a first approximation the mechanical and electrical waves are independent; but, if the amplitudes are not treated as infinitesimal, there are cross-terms which will ultimately bring about an equilibrium distribution of energy between them.

At non-zero temperature there is a field of radiation, determined by Planck's law, which induces waves of system (*b*) in the material; these in turn maintain waves of system (*a*) by the above-mentioned slow transfer of energy. But even at zero temperature the wave systems do not vanish, since perfect uniformity is highly improbable statistically. Accordingly at zero temperature we have waves representing fluctuations that result from the law of chance. It is the energy and pressure of these residual waves that we principally wish to determine.[a]

Turning to the microscopic aspect, we consider first the electrical waves. The polarisation can be represented as a distribution of a large number of doublets. We denote the vector moment of a doublet by $e\xi_1$, $e\xi_2$, $e\xi_3$. Considering the n doublets in a unit volume, we count the number with ξ_1, ξ_2, ξ_3 in various ranges, and so arrive at a distribution function of the 'electrical coordinates' ξ_1, ξ_2, ξ_3. In like manner the sound waves are represented microscopically by displacements x_1, x_2, x_3 of a large number of uncharged particles, and a distribution function of the 'mechanical coordinates' x_1, x_2, x_3 can be defined.

As usual the frequency distribution of variates in an assemblage is described as the probability distribution of the variates of a particle which is an unidentified member of the assemblage. Thus the state of the material is described by the probability distributions of the coordinates ξ_1, ξ_2, ξ_3 of an unidentified doublet and of the coordinates x_1, x_2, x_3 of an unidentified neutral particle.

It would not be possible to represent a simple sound wave in this summary way; because the most significant feature is then the correlation of displacement of those particles which are in the same part of the wave. But the method that we are developing is bound up with the theory of statistical equilibrium; and there is no intention of extending it to non-equilibrium phenomena. In equilibrium conditions the fluctuations are of a random character. By Fourier analysis we can represent them by systems of mechanical and electrical waves; but the analysis would be misleading if it causes us to think that correlations of displacement, such as those exhibited in a simple harmonic wave, are involved.

Returning to the molar aspect, there is an alternative treatment which claims attention. If we think of the material as a superposition of positively charged and negatively charged matter, we may consider instead of (*a*) and (*b*):

(*a'*) Waves of displacement of positively charged matter.

(*b'*) Waves of displacement of negatively charged matter.

This is equivalent, since electrical waves are given by oscillations of (*a'*) and (*b'*) in opposite phase, and mechanical waves by oscillations in the same phase. By the same procedure as before, the state of the material is described microscopically by a probability distribution of the displacement x_1, x_2, x_3 of an unidentified positively charged particle, and a probability distribution of the displacement x'_1, x'_2, x'_3 of an unidentified negatively charged particle.

Whereas the wave systems (*a*) and (*b*) are nearly independent, there is very intense interaction between (*a'*) and (*b'*)—so intense that in molar physics we should never think of investigating them separately. Microscopic physics shows the mechanism of

[a] The results will apply to a 'black dwarf', i.e. a super-dense star with zero temperature. In practical application to white dwarfs, the correction due to the internal temperature of the star is separately investigated.

this interaction. Whenever a proton and an electron collide energy is transferred from (a') to (b') or vice versa. In a time in which each electron has had one collision a considerable part of the whole energy has passed between the two systems; so that the time of relaxation is of the order of magnitude of the time between collision. In superdense matter the particles are in continuous collision, and the separate wave systems (a') and (b') can scarcely be said to exist even momentarily.

The eigenstates must therefore conform to the separation of (a) and (b), not (a') and (b'). If the lifetime of a state is long compared with the natural period of its eigenfunction, the state is sharp. As the lifetime decreases the state becomes broadened; and the states of the wave systems (a'), (b') would be broadened out of existence. Broadening of the states means that the energy of the system is not wholly the energy of the particles in the states, but is partly energy of the transitions going on between the states. Practical procedure turns on the fact that the energies of eigenstates are comparatively easy to calculate; whereas the calculation of interstate energy is a difficult problem scarcely to be attempted unless it occurs as a small correction. It is therefore necessary to employ eigenstates between which transitions are known to be rare, so as to ensure that the incalculable interstate energy is negligible; to employ eigenstates subject to frequent transitions misses the point of the method.

It is necessary to emphasise the inadmissibility of analysis based on the separation of (a') and (b'), because it is the initial fallacy in recent investigations of white-dwarf matter; and the resulting pressure formula (the Stoner-Anderson formula) continues to work devastation in astronomy. These investigations take for eigenfunctions a set of Dirac wave functions which represent electrons moving freely under no forces with constant velocity and constant plane of spin. It would be difficult to imagine a more unsuitable choice for the problem attacked. Having chosen eigenstates between which transitions are excessively frequent, the investigator imposes on himself the task of calculating the energy of the transition circulation which has to be added to that of the eigenstates. But this is not attempted and the investigation proves nothing.

Wave analysis is applied in the usual way to the probability distributions of ξ_α and x_α; and the conjugate momenta ϖ_α, p_α are introduced. We obtain in this way wave functions in ξ-space and x-space. (The 'waves' of wave mechanics must not be confused with the original sound waves and electrical waves.) It is to these that the exclusion principle applies, namely, that there are not more than two particles per cell h^3 of $\xi\varpi$-space and not more than two particles per cell h^3 of xp-space.

It is impossible to grasp all at once the implications of this new approach to wave mechanics. We must be content to connect it with the other lines of development by gradual stages. One connection has already been traced: exclusion between the uncharged particles (extracules) in xp-space is a substitute for inertia and gravitation. A sub-threshold investigation was needed to show the equivalence; that is to say, the wave system (a) was taken to include not only the sound waves but the energy concealed in the rest mass of the material. In ordinary (super-threshold) theory the main part of the xp-exclusion is concealed in the rest masses of the extracules; so that normally we have only to deal with the $\xi\varpi$-exclusion.[a]

[a] The replacement of exclusion interaction by rest mass refers to the uranoidal distribution. In compressing the uranoid particles to white-dwarf density, further exclusion energy is added which corresponds to the residual sound waves at zero temperature. This is too small to have any importance; but it can be calculated in the same way as the exclusion energy corresponding to the electric waves.

Another important connection is obtained by considering matter of very low density. For simplicity we postulate a temperature sufficient to ionise the matter. As the density decreases, collisions between positive and negative particles become infrequent; and ultimately the alternative separation (a'), (b') becomes admissible. Thus at low density there will be a transformation connecting the x, ξ variates with the x, x' variates, which must satisfy the condition (26·22) that the two representations yield the same energy tensor. We take this to be the two-particle transformation (26·12), and so make a junction with our previous theory.

It is now possible to make clear some logical points. We began with a general representation of polarisation by doublets, but did not specify the particular distribution of doublets in terms of ordinary constants of the material. This was justified, because super-dense matter only exists in the interior of certain stars, and its constants are by no means 'ordinary'. If it is asked whether the doublet is formed by joining a proton to the nearest electron or to one taken at random from the whole volume considered, we must reply that there are no protons and electrons—or, at least, they are so kept in transition that the question has no meaning. The specialisation, for application to particular physical systems, is therefore held over until we consider the limiting case of very diffuse matter; it then takes the form of the condition that the variates satisfy the two-particle transformation, which provides linkage with our previous theory and hence with observation. The two-particle transformation must not be applied except to diffuse matter, because (26·22) is not satisfied if one of the representations involves energy of transition circulation. The interaction between the intracules (doublets) is here represented entirely by exclusion. As the density decreases, the exclusion effect decreases; and when exclusion becomes negligible we arrive at an assemblage of non-interacting intracules—which is the distribution studied in our earlier theory of intracules.

The last step—the two-particle transformation—assumes the material to be hydrogen. For other elements a more complex transformation would be required, which cannot be treated here. We shall therefore limit the theoretical discussion to super-dense hydrogen. The junction with ordinary theory supplies the important information that the masses of the x and ξ particles are the masses m, μ of the hydrogen extracule and intracule.

45. The degeneracy pressure

Applying the general formula (41·23) to the intracules in super-dense hydrogen the mean exclusion energy per intracule is

$$\bar{E} = \frac{3}{5}\left(\frac{3\sigma}{8\pi}\right)^{\frac{2}{3}}\frac{h^2}{2\mu}, \tag{45·1}$$

where σ is the particle density of the intracules.

For astronomical purposes the pressure is required. This may be calculated independently of the energy as follows. By (23·1) the mean contribution of each intracule in the normalisation volume V_0 is $\overline{\Delta T_{11}} = \overline{p_1^2}/V_0\mu$; and the total contribution of the σV_0 intracules is $P = \sigma\overline{p_1^2}/\mu$. By the spherical form of the momentum distribution $\overline{p_1^2} = \frac{1}{5}\mathfrak{p}^2$; hence by (41·21)

$$P = K\sigma^{\frac{5}{3}}\left(K = \frac{1}{5}\left(\frac{3}{8\pi}\right)^{\frac{2}{3}}\frac{h^2}{\mu}\right). \tag{45·2}$$

This is the 'degeneracy pressure' of white-dwarf matter, i.e. the pressure at zero temperature.

Since
$$\frac{h^2}{\mu}=\frac{h^2}{m_e}+\frac{h^2}{m_p},$$

the pressure can be (rather artificially) split into two parts apparently arising from the electrons and protons respectively, the contributions being inversely proportional to their masses. It is usually assumed that this can be extended to all nuclei or ions. Since the contributions of the positive particles are negligible, (45·2) will then apply to any kind of matter, σ being the particle density of the free electrons. The assumption is probably correct, but it lacks rigorous proof.

There are three white-dwarf stars for which sufficient astronomical data are available. These afford a striking general confirmation of (45·2),[a] though the test is not accurate enough to settle fine points of theory. White-dwarf matter was first investigated by R. H. Fowler, to whom the result (45·2) is due.[b]

The mechanical waves (sound waves) will contribute an additional degeneracy pressure $K_M\sigma^{\frac{4}{3}}$, where K_M is the constant obtained by substituting M for μ in K. The additional pressure is of no practical importance.

The wave functions of the steady states are standing waves. It is not permissible to replace the standing wave by two progressive waves in opposite directions, because that would be a representation of dispersing material; it ignores the collisions which continually reverse the direction of motion and so prevent dispersal. Similarly there are no waves representing circulatory motion, such as occur in an isolated atom; because the continuous collision makes the lifetime of an integral of angular momentum vanishingly small. In a standing wave p_1^2, p_2^2, p_3^2 reduce to eigenvalues; but the momenta p_1, p_2, p_3 are not eigenvalues and their expectation values are zero. This means that the investigation deals with energy tensors, not momentum vectors, and it therefore comes within scale-free physics.

Since the exclusion energy-density and pressure are scale-free characteristics, there is no danger of the formulae breaking down for large values of p^2. In particular the proportionality of P and E to $\sigma^{\frac{4}{3}}$ must hold for all values of σ.[c] We can, if we like, limit the foregoing discussion to relatively small values of p^2 (for which its validity is unquestioned), and then extend the results to large values by a scale transformation. It is instructive to follow up the details of the transformation.

Our analysis separates the electrical (ξ, ϖ) system from the mechanical (x, p) system, applying a similar exclusion principle to each; so that we can apply the scale transformation to either system independently. The masses of the particles have been found by sub-threshold analysis to be concealed exclusion energy of the mechanical system. We wish to examine changes of the electrical exclusion of particles with fixed proper masses, and must therefore apply the scale transformation to the electrical system only. But we shall begin by applying it universally. In natural units P has

[a] In the astronomical calculation the only unknown likely to affect the result is the percentage of hydrogen in the whole mass; and the formula is therefore generally used to give an observational determination of the abundance of hydrogen in the star. The hydrogen effect is not large; and it is a considerable success for the theory that there is a possible hydrogen abundance for each of the stars. Details are given in *Monthly Notices, R.A.S.* **99**, 595, 1939.

[b] *Monthly Notices, R.A.S.* **87**, 114, 1926.

[c] P, E and σ are to be measured in Galilean coordinates at the point considered.

the dimensions ρ, and σ the dimensions $\rho^{\frac{1}{2}}$; so that, in order to satisfy (45·2), K must have the dimensions $\rho^{-\frac{1}{2}}$. This agrees with the form h^2/μ of K, since, by (29·2), μ is multiplied by $\beta^{\frac{1}{2}}$ when the density is multiplied by β. In the actual application of (45·2) K is constant, because the scale transformation is not extended to the mass μ.

Keeping μ constant we can give K the right dimensions by taking h to vary as $\rho^{-\frac{1}{12}}$. Then if ρ is multiplied by β^{-1}, as in the transformation from system B to system A, h is multiplied by $\beta^{\frac{1}{12}}$. By (30·5) this change is eliminated in the constant $h' = \beta^{-\frac{1}{12}}h$ defined by molar control in system A. The definition by molar control depends on the *union* of mechanics and electrodynamics in classical theory, which makes it necessary to apply the same scale to both; the quantum definition depends on the *separation* of microscopic mechanics and electrodynamics by the introduction of extracules and intracules. Thus molar control gives the constant h', which is the unchanged quantity when the scale transformation ($B \to A$) is applied to both systems; and quantum theory gives the constant h, which is the unchanged quantity when it is applied to the electrical system separately. This confirms the formula (30·5).

The question was raised at the end of § 43 whether it is permissible to use

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

as the hamiltonian for standing waves, when p^2 is not small. This is now answered in the affirmative. It occurs, not as an approximation to $(m^2+p^2)^{\frac{1}{2}}$, which on account of its Lorentz invariance is obviously ruled out, but as an exact condition applying to the energy tensor. It may be written $\Delta T_4^{\,4} = -\frac{3}{2}\Delta T_1^{\,1}$, or $\Delta T_4^{\,4} = -\Delta T$; that is to say, $\Delta\rho = \frac{3}{2}\Delta P = -\Delta\rho_0$, ρ_0 being the proper density.

Reference has already been made to the erroneous (Stoner-Anderson) formula, which is currently used. In opposing my criticism of it, Dirac, Peierls and Pryce say:[a]

'Eddington raises objections on similar grounds against the customary treatment of the equation of state of a degenerate gas. Here the situation is considerably simpler *because one neglects the interaction between the particles altogether*.'

That is why I reject altogether the customary treatment.

[a] *Proc. Camb. Phil. Soc.* **38**, 199, 1942. The italics are mine.

Chapter V

THE PLANOID

46. Uranoid and planoid

Having established the formulae (3·8) and (40·7) which connect the cosmological constants N, R_0 with the microscopic constants, we have no further occasion to treat object-systems on a cosmical scale. Being free now to confine attention to systems whose extension is very small compared with R_0 we are able to introduce an important simplification.

When provision has been made for representing the inertial-gravitational field otherwise than by curvature, so that the object-system and its environment are in flat space, it is natural to adopt as standard environment a uniform distribution of particles in flat space. The distribution may be supposed to continue indefinitely so that it is represented by infinite plane wave functions; but the 'environment' of the small object-system is limited to a sphere of radius R_1 about the object-system as centre containing N_1 particles, R_1 and N_1 being chosen so as to give the correct uncertainty constant σ. We call this form of standard environment a *planoid*.

The planoid is not a mathematical transformation of the uranoid, but is a physically distinct distribution. When the scale and phase dimension is used instead of curvature, the uranoid is projected orthogonally into a flat space (§ 6), and its density in the flat space is proportional to $(1-r^2/R_0^2)^{-\frac{1}{2}}$; in the planoid this is replaced by a density independent of r. Regarded as representations of the actual universe, there is nothing to choose between the uranoid and planoid; both differ widely from the actual expanding universe. In using either of them we take advantage of the relativistic principle that extensive changes of the remote environment can be made without affecting the object-system (§ 7). Local phenomena involve only its integrated effect, which is contained in the inertial-gravitational or metrical potentials $g_{\mu\nu}$.

We have to give the planoid a spherical boundary. The boundary would have created difficulty at an earlier stage; but we do not mind introducing a boundary when the theory has got far enough to settle the boundary conditions. Now that the object-system is limited to a small region at the centre, the effect is only to introduce an extraordinary fluctuation by limiting the number of particles to N_1. Whatever wave mechanical substitute for gravitation is employed is a substitute for the extraordinary fluctuation, and therefore implicitly takes account of the boundary condition.

We consider then, as alternative environments for a small object-system,

(a) A zero-temperature uranoid with radius of curvature R_0 containing N particles.

(b) A zero-temperature planoid of radius R_1 containing N_1 particles.

Unless otherwise stated, it is understood that the same units of mass and length are used in (a) and (b), so that all quantum-specified lengths and masses are the same. This requires that σ shall be the same. For uniform distribution over a sphere of radius R_1, the standard deviation of a coordinate is $(\frac{1}{5}R_1^2)^{\frac{1}{2}}$, and for a coordinate of the centroid is $(\frac{1}{5}R_1^2/N_1)^{\frac{1}{2}}$; we have therefore

$$\frac{R_1^2}{5N_1} = \sigma^2 = \frac{R_0^2}{4N}. \tag{46·1}$$

The fundamental formula for m_0 is slightly simplified by using planoidal constants. By (40·7) and (46·1),

$$m_0 = \frac{3}{4}\frac{\hbar\sqrt{N_1}}{R_1}. \tag{46·2}$$

The most illuminating comparison with the uranoid is obtained by writing (40·7) in the form

$$m_0 = \left(\frac{3}{4}\right)^{\frac{1}{2}}\frac{\hbar_1\sqrt{N}}{R_0}, \tag{46·3}$$

where $\hbar_1$ is the sub-threshold constant (42·2).

Usually it is unnecessary to fix N_1 and R_1 separately, since only the combination $1/\sqrt{N_1}$ is required. We shall, however, introduce later a *special planoid* with fixed constants R_1, N_1, which is useful for certain fundamental investigations.

47. Interchange of extracules

We denote the mean value of a quantity averaged over the volume of the planoid by $[\ldots]$. If r is the distance from the centre,

$$[r^{-2}] = 3R_1^{-2}, \quad [r^{-1}] = \tfrac{3}{2}R_1^{-1}. \tag{47·1}$$

These means are not sensibly altered if r is measured from an object-particle reasonably near (e.g. within a billion miles of) the centre. Then, by (46·2),

$$m_0 = \left[\frac{\frac{1}{2}\hbar}{r}\right]\sqrt{N_1}, \tag{47·21}$$

$$m_0^2 = \frac{3}{16}\left[\frac{\hbar^2}{r^2}\right]N_1 = \frac{3}{2}\left[\sum_{s=1}^{s=\frac{1}{2}N_1}\left(\frac{\frac{1}{2}\hbar}{r}\right)^2\right], \tag{47·22}$$

the summation being over the $\frac{1}{2}N_1$ extracules of the planoid. The fluctuation of the summed quantity is relatively insignificant, and we can drop the $[\ldots]$. We have then

$$m_0^2 = \tfrac{3}{2}\Sigma_s p_s^2 \quad (p_s = \tfrac{1}{2}\hbar/r_s). \tag{47·3}$$

Introducing the mass-constant $\mu = m_0/136$ of an intracule, (47·3) gives

$$m_3 = \tfrac{136}{3}m_0 = \Sigma_s \frac{p_s^2}{2\mu}, \tag{47·4}$$

where m_3 is the rest mass of a V_3 extracule by (16·5). We obtain a physical interpretation of p_s by supposing that the V_3 object-particle has a half-quantum of angular momentum about the sth planoid extracule; the corresponding linear momentum is $\frac{1}{2}\hbar/r_s = p_s$.

We shall now take the planoid extracules to be V_3 particles, so that the object-particle to which (47·4) refers is one of an assemblage of similar particles. Rigid-field theory can then be applied. Let the object-V_3 make a transition to a state of momentum p_x, p_y, p_z; its energy is

$$E_3 = m_3 + (p_x^2 + p_y^2 + p_z^2)/2\mu \tag{47·51}$$

by (18·4), since the transition energy is independent of the multiplicity. Hence by (47·4),

$$E_3 = (p_1^2 + p_2^2 + \ldots + p_{\frac{1}{2}N_1}^2 + p_x^2 + p_y^2 + p_z^2)/2\mu. \tag{47·52}$$

This opens up a far-reaching conception. Initial energy is eliminated, and the whole energy is exhibited in the form of transition energy of a system with $\frac{1}{2}N_1 + 3$ degrees

of freedom. The $\frac{1}{2}N_1 + 3$ momenta are formally similar; but only the last three are actual observables. Ideally the others are observables depending on the individual distances r_s of the planoid extracules; but they are treated collectively and replaced by a stabilised average. Not even the average is observed, the planoid being a postulated standard environment.

We recognise that the initial or rest energy m_3 is a concealed energy. In (47·52) it is dragged out of concealment, and exhibited in a comparable form with unconcealed energy. It is seen to arise from the V_3 extracule having a half-quantum of angular momentum about every other extracule in the planoidal assemblage. Since the particle is at rest, the momentum is extra-spatial, and the half-quantum represents interchange circulation. The half-quantum value agrees with § 33, because here the interchange is between the object-particle and its environment; and the doubling, which gives a whole quantum of interchange angular momentum between two parts of an object-system (as in the case of the Coulomb energy) does not apply.

Since the non-observable momenta are treated collectively, we can substitute for their separate momenta p_s a resultant

$$P = \sqrt{(\Sigma p_s^2)}. \tag{47·61}$$

Then $m_3 = P^2/2\mu$, and $$E_3 = (P^2 + p_x^2 + p_y^2 + p_z^2)/2\mu. \tag{47·62}$$

The momenta p_x, p_y, p_z introduced in (47·51) are in quantum reckoning, adapted to the inversion of energy. Using instead the classical momenta, related to them by (18·31), the result is

$$E_3 = \frac{P^2 - p_x'^2 - p_y'^2 - p_z'^2}{2\mu} = \frac{p_u'^2}{2\mu}, \tag{47·63}$$

where p_u' can be interpreted as the fourth component of a vector p_x', p_y', p_z', p_u' of fixed length P. Since we already have a phase coordinate normal to x, y, z, we can take p_u' to be the component in that dimension, i.e. the scale momentum.

We have arrived at the following physical representation, limited at present to V_3 particles. The object-particle has a half-quantum of angular momentum about every particle in the environment; *this is the only momentum contemplated.* But the planes of the angular momenta are unknown, and no assumption is made as to their probability distribution. The corresponding linear momenta will therefore in general have components in space as well as normal to space. These elementary linear momenta combine quadratically, like waves in incoherent phase. Consequently the resultant momenta in the four directions satisfy $p_x'^2 + p_y'^2 + p_z'^2 + p_u'^2 = P^2$, where $P^2 = \Sigma_s(\frac{1}{2}\hbar/r_s)^2$, and is independent of the planes of the angular momenta. The energy E_3 is a quantity $p_u'^2/2\mu$ arising wholly from the extra-spatial component. In order to determine p_u' we have to measure p_x', p_y', p_z', and apply (47·63), P being a stabilised constant.

The space momenta p_x', p_y', p_z' can (owing to the quadratic law of combination) only be zero, if each half-quantum is in a plane through r_s and the extra-spatial axis. Thus in the theory of the rest mass the planes are taken to have this distribution. But it is not a general condition; and, when it is true, it is an inference from the observational fact that p_x', p_y', p_z' have been measured and found to be zero. No other observational information is introduced. Thus the calculated rest mass applies to a particle for which only the three spatial momenta have been measured, i.e. to a V_3 particle. This explains why the mass given by the simple formula (47·4) is m_3.

We have now two alternative substitutes for the molar representation of the rest energy of particles by curvature. The rest energy of an object-particle is either the result of exclusion by the particles of the environment, which compete with it for the levels of lowest energy; or it is the resultant energy of the continual interchange of the object-particle with all the particles of the environment, the interchange being associated with a half-quantum of angular momentum. At present we do not inquire further into the origin of the exclusion principle or the half-quantum principle. It is sufficient to recognise that both principles are familiar in the electrical part of quantum theory; so that in tracing inertial mass to either of these sources we contribute to the unification of theory. The half-quantum principle is the more elementary, the interchange of extracules here treated being immediately related to the interchange of (or rather, within) intracules in Chapter III. The exclusion principle gives a connection with more advanced parts of atomic physics.

To make clear the relation between the three representations of mass, consider a pseudo-discrete assemblage. We began by treating the object-particle as an unidentified member of the assemblage; as such it has its appropriate share of the molar energy tensor which is identified with space-time curvature. If we try to identify the particles by attaching suffixes s to them, continuous interchange occurs between every pair of suffixes. The interchange energy of the identified particle is the same as the mass which we assign to the unidentified particle. Another way of distinguishing the particles is to provide each with a different 'state', the states being carefully chosen (§ 44) so that transitions between them are negligible. There is no longer interchange energy; but the states pile up to a high energy level, and the exclusion energy gives the mass that was assigned to an unidentified particle.

48. The special planoid

In the special planoid $N_1 = N$, $R_1 = R_0$, the quantum is the sub-threshold constant $\hbar_1$, and the particles are V_4 particles.

We have
$$\sigma_1^2 = R_1^2/5N_1 = R_0^2/5N = \tfrac{4}{5}\sigma^2, \tag{48·11}$$
and for the weight constant ϖ in (38·72),
$$\varpi_1 = \frac{\hbar_1}{2\sigma_1} = \left(\frac{3}{5}\Big/\frac{4}{5}\right)^{\frac{1}{2}} \frac{\hbar}{2\sigma} = \left(\frac{3}{4}\right)^{\frac{1}{2}} \varpi. \tag{48·12}$$

Since all quantum-specified lengths and masses are fixed multiples of σ and ϖ, the planoid reckoning of lengths (L) and masses (M) differs from the uranoid reckoning, the transformation formulae being
$$L = (\tfrac{5}{4})^{\frac{1}{2}} L_1, \quad M = (\tfrac{4}{3})^{\frac{1}{2}} M_1. \tag{48·2}$$

Results worked out in the special planoid have to be converted into ordinary measure by these formulae.

It would lead to great complication if we did not keep the same number of particles in the uranoid and the planoid; but the condition $R_1 = R_0$ is not so important; and we can, if we prefer, eliminate the transformation of length. The specification of the planoid is then
$$N_1 = N, \quad R_1 = (\tfrac{5}{4})^{\frac{1}{2}} R_0, \quad \hbar_1 = (\tfrac{3}{4})^{\frac{1}{2}} \hbar, \quad L = L_1, \quad M = (\tfrac{4}{3})^{\frac{1}{2}} M_1. \tag{48·3}$$

The special planoid is to be considered in connection with the exclusion analysis in §§41–43. Exclusion energy is equivalent to inertial-gravitational energy. The latter is conventionally separated into inertial and gravitational energy, the inertial energy being the top exclusion energy. In a large assemblage there is, in addition to the inertial masses of the particles, a negative gravitational energy; this is the negative exclusion energy (measured from $\mathfrak{E}$ as zero), which arises because some of the particles must be below the top level. Since exclusion theory is an equilibrium theory applying only to uniform distributions, we cannot make a detailed comparison with the gravitational energy of aggregations such as stars. But, considering the whole universe, we see that a particle has a negative gravitational energy, due to the field of the rest of the universe, corresponding to $\bar{E} - \mathfrak{E}$, and amounting to $\frac{2}{5}$ of its inertial mass. The factor $\frac{2}{5}$ is a special value for the Einstein universe; in the actual universe it would differ considerably, as also in the planoid. In the special planoid the negative gravitational energy is $\frac{1}{4}$ instead of $\frac{2}{5}$ of the inertial energy, the selection factor being $\frac{4}{3}$ instead of $\frac{5}{3}$. The uncertainty of the factor in the actual universe does not create any practical difficulty, because we always remove the gravitational potential in the region we are considering by a local transformation of coordinates—'local' here being a term wide enough to cover a galaxy.

In the analysis of the uranoid (§42), the selection factor $\frac{5}{3}$, by increasing the density and therefore decreasing κ and increasing $\hbar$, gave $\hbar_1 = (\frac{3}{5})^{\frac{1}{3}}\hbar$; and quantum-specified momenta $\hbar/r$ or energies $\hbar\nu$ would transform in the same ratio. The argument assumes that there is no change of the unit of length, so that the comparable transformation for the planoid is (48·3). This identifies the selection factor for the special planoid as $\frac{4}{3}$. The practical value of the special planoid lies in the fact that it has this particular selection factor.

In §16 we showed how multiplicity factors could be exhibited as selection factors. There is no direct correspondence with the selection factors in exclusion theory; and the connection does not extend beyond a general similarity of method. When the factor is $\frac{4}{3}$, the selection of top particles gives the same change of density as the stabilisation of scale which transforms a distribution of V_4 particles into V_3 particles. When we pass from the flat representation with additional phase coordinate in the planoid to the curved representation in the uranoid, the scale is stabilised; so that V_4 planoid particles become V_3 uranoid particles. We automatically provide for the consequent change of density when we select the top particles.

It is clear that the exclusion principle as enunciated in (41·11) or (41·12) refers to V_3 particles. It contemplates only three classifying characteristics of states, namely the three spatial momenta. An indirect confirmation is given by the result (§47) that the half-quantum interchange, which is equivalent to exclusion, yields the mass m_3 of a V_3 particle. The exclusion calculation of mass does not provide any test of the multiplicity, because the top particle whose mass $\mathfrak{E} = m_0$ is calculated is a comparison particle in super-threshold theory; and in that capacity it behaves like a particle of multiplicity 1, whatever its intrinsic multiplicity. In super-threshold theory m_0 is the mass of a V_{136} object-particle—hydrocule or standard particle according as system B or A is used. This is introduced by a special process explained in §42; but, since there is a one-to-one correspondence, and the masses are the same, we can for practical purposes regard the same particle as having the multiplicity 136 in gravitational theory

and the effective multiplicity 3 in exclusion theory.[a] The only use of the latter multiplicity is to distinguish the V_3 particles in curvature representation and the V_4 particles in scale-and-phase representation; but, since this is more simply expressed as a distinction between top and mean particles in the special planoid, we can eliminate all reference to the exclusion multiplicity. Thus, for practical application, we regard the planoid as composed of V_{136} object-particles, with the usual understanding that only the top particles appear in microscopic observation.

The advantage of the uniform flat environment provided by the planoid is not merely that it simplifies the algebra. It gets over a difficulty of non-integrability of states, which arises in spherical space. When we specify a state by the components of the energy tensor, there is no way of extending the state over the whole of spherical space, because the directions to which the components refer become ambiguous. The result is that we fill spherical space with bits of states which do not join up into whole states. In an early investigation[b] I found that there is at any point of space a certain number N_0 of independent wave functions, and concluded that on the principle that each particle is to be provided with a distinct wave function the total number of particles must be N_0. It has since turned out that the actual number is $N = \frac{3}{4}N_0$. The fallacy lay in supposing that the N_0 bits of states found at every point were equivalent to N_0 states covering the whole hypersphere; but they are not integrable elements of states, and its appears that the equivalent of a quarter of them is lost by the misfit.

49. The energy of two protons

The analysis in Chapter II introduced an elementary two-particle system, which we were able to identify with a proton and electron. This identification was to be expected, for it is evident that a system of two protons or two electrons is not so elementary. The like charges are the source of an extended electric field which induces opposite charges somewhere in the environment. Properly they form, not a two-particle system, but an incompletely separated part of a four-particle system. It is only by introducing certain adaptations that two-particle analysis can be extended to like charges. We shall now investigate this extension, taking for definiteness two protons.

The Coulomb or interchange energy of a proton-electron system has been calculated theoretically; but a corresponding calculation for the proton-proton system would be difficult owing to the non-existence of steady states. However, it is easy to deduce the Coulomb energy of two protons from that of a proton and electron. A charged particle has no Coulomb energy if it is in a neutral environment, the mutual energy being then purely mechanical by definition. Also by definition equal distributions of positively and negatively charged particles constitute neutral matter. It follows that for all values of r the proton-proton electric energy must be equal and opposite to the proton-electron energy $-e^2/r$.

We have therefore to consider two particles with masses m_p and a positive Coulomb energy e^2/r; but this system is not simply superposable on the standard neutral

[a] A multiplicity factor is relative to the adopted system of classificatory characteristics. Normally we adhere to the system provided by components of the energy tensor, but a transformation to another reference system is not forbidden. The treatment of the same particle as V_{136} in gravitational classification and V_3 in exclusion classification can be regarded as an example of the relativity.

[b] *Protons and Electrons*, §§ 16·5–16·7.

environment. A charge distribution, amounting altogether to $-2e$, is induced in the environment. If we treat the two protons formally as an isolated two-particle system, a term must be inserted in the hamiltonian to represent the effect of the suppressed induced charges. In the theory of electrolytes and in astrophysics, this term is known as Debye-Hückel energy. *In the fundamental equations of an object-system in the standard uranoid it is the non-Coulombian energy.*

The following procedure is an adaptation to electrical energy of a treatment that has already been applied to mechanical energy. Instead of analysing matter into relativity particles which disturb the uranoid gravitationally, we have analysed it into quantum particles which are superpositions on an undisturbed uranoid. This will now be extended to electric charge. Instead of relativity charges (classical charges) which induce opposite charges in the environment, we consider quantum charges which are simply superposed on the environment. The protons and electrons in our analysis are carriers of quantum charges.

We shall show that the energy of a pair of quantum charges is

$$\left.\begin{aligned} E &= -e^2/r \quad \text{(unlike charges)},\\ E &= e^2/r + B\delta(r') \quad \text{(like charges)}, \end{aligned}\right\} \tag{49·1}$$

where δ is Dirac's function

$$\int \delta(r)\, dV = 1 \quad (\delta(r) = 0, \text{ if } r \neq 0), \tag{49·2}$$

and B is a constant which will be determined presently. The difference between r and r' has been explained in § 5; r' is measured directly from one particle to the other, and r (which is the more usual measure) corresponds to the coordinate-differences ξ, η, ζ.

Consider a single proton as object-system, and first treat its charge classically. In selecting it as object-system we leave an unpaired electron in the environment, which is the induced charge above-mentioned. Since the uranoid particles have uniform probability distribution, the selection leaves a charge $-e$ distributed uniformly over the uranoid; and the mutual electric energy of the proton and the disturbed uranoid is

$$-\Omega = -[e^2/r] = -e^2[r^{-1}], \tag{49·31}$$

the square bracket denoting mean value.

Now substitute quantum charges, assuming provisionally the energy formula (49·1). The object-proton is now a charge superposed on the undisturbed uranoid, which accordingly contains $\frac{1}{2}N$ electrons and $\frac{1}{2}N$ protons. If V is the volume of the uranoid, each particle has a probability dV/V of being in an element dV; and the mutual energy of the object-proton and a uranoid particle is

$$\left.\begin{aligned} \frac{1}{V}\int E\, dV &= -\Omega \quad \text{for an electron},\\ &= \Omega + B/V \quad \text{for a proton}, \end{aligned}\right\} \tag{49·32}$$

by (49·1). The total mutual energy of the object-proton and the uranoid is therefore $\frac{1}{2}NB/V$. This has to agree with the previous result (49·31) obtained by the classical or relativity calculation; hence

$$\tfrac{1}{2}NB/V = -\Omega = -e^2[r^{-1}]. \tag{49·33}$$

The argument applies to a planoidal as well as a uranoidal environment; and the calculation is made easy by using the special planoid. By (47·1) $\Omega = \frac{3}{2}e^2/R_1 = \frac{3}{2}e^2/R_0$. This, being of the form $\hbar/r$, is a quantum-specified energy, and must be multiplied by $(\frac{4}{3})^{\frac{1}{3}}$ to convert from planoidal to ordinary measure in accordance with (48·2). Hence

$$\tfrac{1}{2}NB/V = -\tfrac{3}{2}(\tfrac{4}{3})^{\frac{1}{3}}e^2/R_0, \tag{49·34}$$

and, since $V = \frac{4}{3}\pi R_0^3$,

$$B = -(\tfrac{4}{3})^{\frac{1}{3}}4\pi e^2 R_0^2/N = -(\tfrac{4}{3})^{\frac{1}{3}}16\pi e^2\sigma^2. \tag{49·4}$$

In order to show the necessity of the form $B\delta(r')$ of the supplementary term, we shall put the calculation in another form. In a geometrical frame x, y, z the probability distribution is continuous; for the concentration of a finite amount of probability in an infinitesimal volume would be contrary to the uncertainty principle. The same is true for a physical frame ξ, η, ζ whose origin is the centroid of a number of particles. But if we employ a frame ξ', η', ζ' measured from a proton as origin, we artificially introduce a singularity. Considering the uranoid particles referred, first to an x-frame or ξ-frame, and secondly to a ξ'-frame, the distinction is that in the latter frame the point $(0, 0, 0)$ is certainly occupied by a proton. This property provides the formal definition of a ξ'-frame or, as we may now call it, a *singular frame*; because, protons being indistinguishable, it would be impossible to say which of the $\frac{1}{2}N$ protons has been chosen as origin. Any one of them has an equal chance of being the occupant of the singular point; so that the probability distribution of a proton consists of a discrete chance $2/N$ of occupying the origin together with a probability $1-2/N$ evenly distributed over space This gives, for the probability in an element dV,

$$\left(\frac{2}{N}\delta(r') + \left(1-\frac{2}{N}\right)\frac{1}{V}\right)dV. \tag{49·51}$$

If the protons are classical protons, each carries independently of its position the electrical energy $-\Omega$ due to the opposite charge of the rest of the uranoid. Multiplying (49·51) by $-\Omega$, the energy distribution is

$$(B\delta(r') - \Omega')\frac{dV}{V}, \tag{49·52}$$

where $\Omega' = (1-2/N)\Omega$, and (as before)

$$B = -2\Omega V/N. \tag{49·53}$$

The distribution of carriers is a matter of conception, the physically significant result being the distribution of that which is carried. We can therefore re-interpret (49·52), by taking dV/V to be the probability distribution of the carrier, and $B\delta(r') - \Omega'$ to be the energy carried. The new carriers are quantum protons. They have uniform distribution in the singular frame. Their chances of occupying the origin are infinitesimal, and so also is their combined chance. The permanent proton at the origin, which we regard as part of the frame, is therefore additional.

The Coulomb energy of the proton is included in Ω' in so far as it arises from protons and electrons with uniform probability distribution uncorrelated to its own; but, of course, the Coulomb energy due to a charge whose distance from it is prescribed (either exactly or as a probability distribution) must be taken into account separately. In particular, the Coulomb energy due to the singular proton is not included in Ω'.

(Strictly, it should be reduced in the ratio Ω'/Ω like the rest of the Coulomb energy of a quantum proton, but the correction is insignificant.) The constant part of the energy Ω' might be included in the mass of the proton, but it is utterly insignificant. Thus we have for the energy of a quantum proton in the field of another proton taken as origin the singular energy $B\delta(r')$ together with the ordinary Coulombian energy.

Since one of the protons is itself the origin of the coordinates ξ', η', ζ', the corresponding r' is the directly measured distance. Relative coordinates and distances defined in this way have not been used in any of our previous investigations. In particular the relative coordinates introduced in the two-particle transformation (26·12) are coordinate-differences. Having obtained the non-Coulombian energy $B\delta(r')$ in the ξ'-frame, we have to transform to the ξ-frame in order to apply it in the ordinary equations. By (5·1), the singular point $r' = 0$ corresponds to a Gaussian probability distribution of ξ, η, ζ with standard deviation $\sigma\sqrt{2}$; so that

$$\delta(r') = (4\pi\sigma^2)^{-\frac{3}{2}} e^{-r^2/4\sigma^2}. \tag{49·6}$$

Combining (49·4) and (49·6),

$$B\delta(r') = -\left(\frac{16}{3\pi}\right)^{\frac{1}{2}} \frac{e^2}{\sigma} e^{-r^2/4\sigma^2}. \tag{49·7}$$

The general reason for the form $B\delta(r')$ can be stated as follows: When a system of two like charges is treated as a superposition on an undisturbed environment, an adjustment of the energy must be made to compensate for the omission of the induced charges. Except when the separation vanishes, no change can be made, owing to the condition that the energy must be equal and opposite to the energy of two corresponding unlike charges, for which the difficulty of superposition does not arise. The adjustment must therefore take the form of a δ-function of the separation. The argument in the second paragraph of this section shows that the vanishing of the separation must be defined in the same way as the coincidence of two unlike charges in a neutral particle. Clearly the latter coincidence is defined by $r' = 0$, the physical origin being irrelevant.

50. Non-Coulombian energy

As usual the elementary result (49·7) is modified by a multiplicity factor when used in current formulae. Before we can calculate this factor it is necessary to decide at what stage in the stabilisation the substitution of quantum protons for classical protons is to be made. The principle of superposition on an undisturbed environment requires that the field shall be rigid gravitationally as well as electrically; otherwise we should have to allow for gravitational disturbance of the environment by the energy which we add to eliminate the electrical disturbance. The Coulomb energy is added in the V_{137} rigid field, and it is primarily in that field that we make the adjustments treated in the last section. In rigid coordinates the Coulomb energy is $\hbar/r$ and the value of B is correspondingly 137 times greater; but the same factor applies also to the mass m_0 of the hydrocule which becomes $137m_0$. Removing it in anticipation of the transformation to Galilean coordinates, we can define the stage at which the energy (49·7) is to be inserted as that in which the mass of the hydrocule is m_0. This means that the adjustment can be made directly in the special planoid where the top particles have the mass m_0.

We have found two ways of introducing protons and electrons. In the first method the hydrocule of mass m_0 is taken as object-particle and split into a proton and electron. This is inappropriate here, because we want to obtain two protons alone. In the second method (§§ 22, 42) the hydrocule of mass m_0 is taken as comparison particle, and protons or electrons are added singly together with comparison holes. Using this method we must first decide the sign of our singular frame—that it is to have a proton, not an electron, as origin. We then regard the comparison particles as an equal mixture of two classes, class (a) having an additional energy $B\delta(r')$, and class (b) having no additional energy. Class (a) comparison particles must be used to form protons, and class (b) to form electrons; or we may put it that, when a proton is introduced, its comparison hole eliminates a particle of class (a). In formula (22·7) determining the rest energy of the proton, m_0 is replaced by $m_0 + B\delta(r')$; and it follows that the correction to the energy m_p of the proton is

$$\frac{m_p}{m_0} B\delta(r'). \tag{50·1}$$

It will be noticed that the factor m_p/m_0 occurs in the non-Coulombian but not in the Coulombian energy. The latter is transition energy unaffected by changes of multiplicity. The non-Coulombian energy is an adjustment of initial energy which allows us to simplify a four-particle system into a two-particle system. It is on the same footing as the rest mass itself, which may be regarded as an adjustment in reducing a $\frac{1}{2}N$-particle system to a one-particle system, since it takes the place of the interchange of the extracule with all the other extracules.

Inserting the factor m_p/m_0 in (49·7), the non-Coulombian energy of two protons is

$$E = -Ae^{-r^2/k^2}, \tag{50·2}$$

where

$$k = 2\sigma = R_0/\sqrt{N}, \tag{50·3}$$

$$A = \left(\frac{16}{3\pi}\right)^{\frac{1}{2}} \frac{m_p}{m_0} \frac{e^2}{\sigma}. \tag{50·4}$$

The value of σ can be found directly from the Rydberg constant. By (40·8), (29·5) and (29·3),

$$M = \frac{136}{10}\frac{3}{4}\frac{\beta^{\frac{1}{6}}\hbar\sqrt{(\frac{4}{5}N)}}{cR_0}, \quad \frac{M}{\mu} = \frac{136^2}{10}\beta^{-\frac{5}{6}}, \quad \mathfrak{R} = \frac{1}{2}\left(\frac{1}{137}\right)^2 \frac{\mu c}{2\pi\hbar}.$$

From these we find

$$\sigma = \frac{R_0}{2\sqrt{N}} = \frac{3}{136^2 \cdot 137 \cdot 16\pi\sqrt{5}} \frac{1}{\mathfrak{R}} = 9{\cdot}604 \times 10^{-14}\,\text{cm}. \tag{50·5}$$

Then, by (50·3) and (50·4),

$$\left.\begin{aligned} k &= 1{\cdot}9208 \times 10^{-13}\,\text{cm.}, \\ A &= 4{\cdot}2572 \times 10^{-5}\,\text{erg} = 52{\cdot}01 m_e c^2. \end{aligned}\right\} \tag{50·6}$$

It remains to consider whether any β-factors are introduced in the transformation to the observational system. Since e is the same in both systems, and σ^{-1} has a calculated ratio to the Rydberg constant, it appears that by using the observational constant $\mathfrak{R}$ in (50·5) the values of A and k are given in the observational system. To make this conclusive, we have to show that the numerical factor (50·5) is obtained as the ratio of σ_B to R_B^{-1} when the whole calculation is carried out in system B. This is

easily confirmed; and, since the ratio is unchanged by the transformation to the observational system, I conclude that (50·6) requires no correction for β-factors. It is understood that A is an optically controlled energy.

Observational values of k and A are found from the scattering of protons by protons. The values given by Thaxton[a] are $k = 1{\cdot}9 \times 10^{-13}$ cm., $A = 51{\cdot}4 m_e c^2$. These determinations are rough. But, although the scattering experiments give weak determinations of A and k separately, they give a good determination of Ak^2. Consequently a good determination of A is obtained by reducing the scattering data with the calculated value $1{\cdot}9208 \times 10^{-13}$ of k. Using the data of Breit, Thaxton and Eisenbud,[b] we find in this way,

$$A/m_e c^2 \quad \text{observed } 52{\cdot}26, \text{ calculated } 52{\cdot}01.$$

The resultant energy (Coulombian and non-Coulombian) of two protons is negative when their separation is between $0{\cdot}03k$ and $2{\cdot}07k$; and the resultant force is attractive when the separation is between $0{\cdot}25k$ and $2{\cdot}69k$. The maximum ratio of the non-Coulombian to the Coulombian energy is 15·2.

For two electrons we have to substitute m_e for m_p in (50·4). The resultant energy is always positive, and the resultant force repulsive. The non-Coulombian energy is never more than $\frac{1}{120}$ of the Coulombian energy, and the non-Coulombian force is never more than $\frac{1}{64}$ of the Coulombian force. Thus in practice the non-Coulombian energy of two electrons is negligible; and, since there is none between unlike charges, the only non-Coulombian energy that need be considered is that between protons.

In finding B by (49·53) we had to calculate the value of $V\Omega$. This calculation was simplified by using the planoid. In an earlier investigation I calculated it in the uranoid, using the relativity equations for the electric potential of a charge in curved space. The result $2\pi R_0^2 e^2$ was the same except that the $(\frac{4}{3})^{\frac{1}{2}}$ factor was absent. The difference arises from the divergence of treatment of energy in relativity theory and quantum theory, which introduces a difference of definition. The constants are adjusted so that the measures agree in a small flat region, but the difference manifests itself when curvature has to be taken into account.

51. The constant of gravitation

We have now three independent ways of determining $R_0/\sqrt{N}$ from laboratory data, two of them from proton scattering experiments using the values of k and A respectively, and the third (which is very much more accurate) from μ, $\hbar$ and c by (40·8). We can determine R_0/N from μ, κ and c by the elementary formula (5·41). We have also a rough determination of R_0 from the observed recession of the nebulae. In Chapter XIII a purely theoretical calculation of N will be made, giving the value $\frac{3}{2} \cdot 136 \cdot 2^{256}$ which should be exact. Leaving out the three weak determinations, our principal formulae are

$$\frac{R_0}{N} = \frac{\kappa M}{\pi c^2}, \tag{51·1}$$

$$\frac{R_0}{\sqrt{N}} = \frac{136}{10}\left(\frac{9}{20}\right)^{\frac{1}{2}} \frac{\beta^{\frac{1}{2}} h}{2\pi c \mu}, \tag{51·2}$$

$$N = \tfrac{3}{2} \cdot 136 \cdot 2^{256}. \tag{51·3}$$

[a] *Physica*, **7**, 122, 1940.

[b] *Physical Review*, **55**, 1018, 1939.

From (51·1) and (51·2), together with $hc/2\pi e^2 = 137$ and $e/mc = \mathfrak{F} = \beta^{\frac{1}{2}\frac{1}{4}}\mathfrak{F}'$, we obtain

$$\frac{\kappa}{\mathfrak{F}'^2c^2} = \frac{136\,.\,137}{10}\left(\frac{9}{20}\right)^{\frac{1}{2}}\frac{\pi\beta^{\frac{1}{2}}}{\sqrt{N}}, \tag{51·4}$$

which gives N in terms of familiar constants. Conversely, using the theoretical value (51·3), we obtain the following comparison:

$$\kappa/\mathfrak{F}'^2c^2 \quad \text{calculated } 8{\cdot}0939\times10^{-37}, \text{ observed } (8{\cdot}098\pm0{\cdot}006)\times10^{-37}. \tag{51·5}$$

Since $\mathfrak{F}'$ is known much more accurately than κ, we use the calculated value of $\kappa/\mathfrak{F}'^2c^2$ to give a determination of κ. The result is

$$\kappa = 6{\cdot}6665\times10^{-8}, \tag{51·6}$$

which should be accurate to about 1 part in 5000. The observed value is

$$(6{\cdot}670\pm0{\cdot}005)\times10^{-8}.$$

The question arises whether (51·6) is necessarily the value of the constant of gravitation that would be determined in actual experimental conditions. In obtaining (51·1) and (51·2) we postulate a universe composed of a steady uniform zero-temperature distribution of protons and electrons. We are, of course, allowed to rearrange the matter of the universe in a way that simplifies the theoretical calculation of $\kappa/\mathfrak{F}'^2c^2$, just as the experimenter is allowed to rearrange the matter in his laboratory in a way that simplifies the experimental determination. But in such rearrangement the experimenter cannot, and the theorist must not, violate the conservation of energy. Formally at least there is an implicit assumption that in the actual universe, the energy of free radiation, cosmic rays, etc., is just sufficient to re-transmute the complex elements into hydrogen, restore the gravitational energy lost in forming condensations, and leave a uniform static distribution at zero temperature. I think that for epistemological reasons this assumption is correct; but it is not really needed for the present purpose. The introduction of a whole uranoid is a mathematical convenience designed to avoid a troublesome investigation of boundary conditions. The uranoid or the boundary conditions fix the scale of measure; but this ceases to have any importance when it is a question of determining a purely numerical ratio such as $\kappa/\mathfrak{F}'^2c^2$. The remote environment has no more effect on the determination of the numerical ratio $\kappa/\mathfrak{F}'^2c^2$ by local measurements than it has on determining the numerical ratio π by local measurements of the circumference and diameter of a circle.

The foregoing question would scarcely have arisen were it not that according to (51·4), $\kappa/\mathfrak{F}'^2c^2$ depends on the total number of particles, so that the remote environment seems to play an essential part. But mathematical dependence is not the same thing as physical causation; and the equation can equally be read as an assertion that the number of particles in the universe is determined by a universal constant $\kappa/\mathfrak{F}'^2c^2$ which we can measure locally. This is the natural physical interpretation. The constant determines the contribution of a particle to the curvature of space, and fixes N by closing the space as soon as the contributions add up to the required amount.

I think therefore that there is no doubt that (51·6) is the constant of gravitation in the actual universe.

The *force-constant* $F = e^2/\kappa m_p m_e$, or ratio of the electrical to the gravitational force between a proton and electron, has a special interest. By (29·6), $m_p m_e = \beta^5 m_0^2/10$. Hence, by (51·4),

$$F = \frac{2}{3\pi\beta^2}\sqrt{(5N)}. \tag{51·7}$$

The limiting speed of recession of the galaxies, given by the Friedmann-Lemaître theory, is $V_0 = c/R_0\sqrt{3}$. The comparison with astronomical observation has been given in § 5. It is of interest to note that

$$kV_0 = c/\sqrt{(3N)}, \tag{51·8}$$

so that the recession-constant can be derived from the range-constant of nuclear forces, or vice versa, with no other observational data except the velocity of light.

The following table of calculated constants is a continuation of the table in § 32. The only observational data used are the three conversion constants in (32·1), and (for certain astronomical constants) the value of the megaparsec in centimetres. Observed values are available for Nos. 16, 22, 24, 26; the comparisons have already been given. A number of additional constants will be determined in Chapters IX sqq. A summary of all the comparisons of theory and observation is given in Chapter XIV.

52. Molar and nuclear constants

Ref. no.	Symbol	Description	Calculated
16	κ	Constant of gravitation	$6{\cdot}6665 \times 10^{-8}$
17	N	Particles in the universe	$2{\cdot}36216 \times 10^{79}$
18	R_0	Einstein radius of space	$9{\cdot}33544 \times 10^{28}$
19	R_0	(In megaparsecs)	302·38
20	M_0	Mass of the universe	$1{\cdot}97675 \times 10^{55}$
21	ρ_0	Density of Einstein universe	$1{\cdot}23088 \times 10^{-27}$
22	V_0	Nebular recession (km.sec.$^{-1}$mp.$^{-1}$)	572·36
23	$e^2/\kappa m_p m_e$	Force constant	$2{\cdot}2714 \times 10^{39}$
24	k	Nuclear range-constant	$1{\cdot}9208 \times 10^{-13}$
25	A	Nuclear energy-constant	$4{\cdot}2572 \times 10^{-5}$
26	$A/m_e c^2$	...	52·01
27	σ	Uncertainty constant	$9{\cdot}604 \times 10^{-14}$

Chapter VI

THE COMPLETE MOMENTUM VECTOR

53. The symbolic frame

We have introduced particles, called V_{10} particles, which have in addition to linear momentum and energy an intrinsic angular momentum or 'spin'. In four dimensions the linear momentum and energy form a 4-vector p_1, p_2, p_3, p_4 and the angular momentum forms a 6-vector $p_{23}, p_{31}, p_{12}, p_{14}, p_{24}, p_{34}$. The mechanical state of the particle is specified by a probability distribution of these 10 variates.

This mechanical specification applies also to a rigid body except that the angular momentum 6-vector is then more specialised, being subject to the condition that when the axes are chosen so that p_1, p_2, p_3 are zero the components p_{14}, p_{24}, p_{34} are zero. This reduces the number of independent components to 7, and the microscopic analogue of a rigid body is a V_7. We can impose rigidity, or coupling of the spin vector to the linear momentum vector, on a microscopic particle as a stabilising condition if occasion arises; but to begin with we treat particles in which all 10 variates are independent. An exact molar analogue of the V_{10} particle is the *mean* of four or more rigid bodies, for the corresponding $\overline{p}_1, \overline{p}_{23}$, etc., are all independent.

For greater symmetry we change the notation of the momentum 4-vector to $p_{15}, p_{25}, p_{35}, p_{45}$. We can then express the whole $(4+6)$-vector as

$$P = \Sigma E_{\mu\nu} p_{\mu\nu} \quad (\mu, \nu = 1, 2, 3, 4, 5; \mu \neq \nu), \tag{53·1}$$

the $E_{\mu\nu}$ being *symbolic coefficients*. The device of expressing a vector as a linear function of its components with symbolic coefficients was first developed in quaternion theory; and the symbolic calculus which will be treated in this chapter is in fact a double quaternion algebra.

The symbols $E_{\mu\nu}$, which will be used very extensively, are the basis of a form of mathematical analysis which I have called *wave tensor calculus*. Like ordinary tensor calculus, this is peculiarly adopted to the concepts of relativity theory. The new calculus is required because ordinary tensor calculus is unable to cope with the distributions of angular momentum which play a leading part in atomic physics. I have given a leisurely development of wave tensor calculus, with full proofs of the formulae in *Relativity Theory of Protons and Electrons*. Here I shall summarise the elementary algebraic properties of the E-symbols very briefly, and pass on as quickly as possible to their application to relativity theory.

Chapters I–V and Chapters VI–VIII form two different lines of approach to fundamental physics which we may distinguish as *statistical theory* and *spin theory*. They have only occasional contact until the end of Chapter VIII, but are brought together in the rest of the book. In the formal sequence of deduction the present approach is supposed to precede our earlier approach; so that various elementary results, previously used by anticipation or referred to as common knowledge, are no longer allowed to be assumed without proof.

The 'complete set' of $E_{\mu\nu}$ consists of 16 symbols. The set of 10 symbols in (53·1) is enlarged to 15 by admitting an extra value 0 for the suffixes μ, ν; the remaining symbol

is $E_{16} = \pm i$. All 16 symbols are square roots of -1; but some pairs commute and some anticommute according to a scheme given below in (53·4). A linear function of the symbols with numerical coefficients (real or complex) is called an *E-number*; the general form of an E-number is accordingly

$$P = E_{16}p_{16} + \Sigma E_{\mu\nu}p_{\mu\nu} \quad (\mu, \nu = 0, 1, 2, 3, 4, 5;\ \mu \neq \nu). \tag{53·2}$$

Since $E_{16} = \pm i$, the term $E_{16}p_{16}$ is algebraic.[a] The algebraic term in an E-number is called the *quarterspur* (abbreviated as 'qs'). Thus

$$\text{qs}\, P = E_{16}p_{16} = \pm ip_{16}. \tag{53·3}$$

We distinguish E-numbers as *spurred* or *spurless* according as $\text{qs}\, P \neq 0$ or $\text{qs}\, P = 0$.

The definition of the symbols in (53·4) will provide that $E_{\nu\mu} = -E_{\mu\nu}$. These are counted as alternative forms of the same symbol; so that only one of them is included in the summation in (53·2). Since the same term may appear either as $E_{\mu\nu}p_{\mu\nu}$ or $E_{\nu\mu}p_{\nu\mu}$, the numerical coefficients also have alternative forms $p_{\nu\mu} = -p_{\mu\nu}$.

The 15 non-algebraic symbols have the following multiplication table: If $\mu, \nu, \sigma, \tau, \lambda, \rho$ is an even permutation of 0, 1, 2, 3, 4, 5,

$$\left.\begin{aligned} E_{\mu\nu}E_{\mu\nu} &= -1, \\ E_{\mu\sigma}E_{\nu\sigma} &= -E_{\nu\sigma}E_{\mu\sigma} = E_{\mu\nu}, \\ E_{\mu\nu}E_{\sigma\tau} &= E_{\sigma\tau}E_{\mu\nu} = E_{16}E_{\lambda\rho}. \end{aligned}\right\} \tag{53·4}$$

The summation convention is not used in connection with these suffixes. It is easily verified that the foregoing multiplication table is self-consistent.

This multiplication table, together with the identification of E_{16} as i or $-i$, is the definition of the E-symbols; and we regard them as having no properties other than those stated in, or derivable from, this definition. The table specifies a 'group-structure', and no more is stated about the $E_{\mu\nu}$ than that they are a set of elements of this group-structure. It is this economy of statement which makes a set of elements defined by a multiplication table the most appropriate framework for expressing the kind of knowledge we can have (or without inconsistency believe ourselves to have) about a world wholly external to the individual mind. All other formulations betray us into stating more than our knowledge justifies; they do not provide a language in which we can say what we want to say without irrelevant embellishments. The group-structure of the E-symbols will later be found to be identical with that of the rotations in the 15 coordinate planes in a Euclidean space of six dimensions; so that we can represent the $E_{\mu\nu}$ graphically by this set of rotations. But it is retrograde to regard the rotations as interpreting the E-symbols; the E-symbols interpret the rotations. For all that we can know as to the nature of rotation, as an element of description applied to a world external to the mind, is comprised in a symbolic specification of the group-structure of the set of possible rotations; and any more vivid conception that our minds associate with rotation is alien to the significance of rotation in physics.

Evidently the sum of two E-numbers is an E-number; and, since E_{16} stands for i or $-i$, the multiplication table shows that the product of two E-numbers is an E-number.

[a] 'Algebraic' will be used throughout in the sense 'elementary-algebraic', i.e. obeying the rules of ordinary commutative algebra. In particular an *algebraic number* is here defined as a symbol which commutes with every other symbol in the calculus.

Thus the E-numbers have a *closed algebra*. It was in order to close the algebra that we extended the number of $E_{\mu\nu}$ from 10 to 15. The 10 symbols included in (53·1) taken by themselves do not give a closed algebra. It will be found in due course that the set of E-symbols defined by the multiplication table (53·4) supplies the appropriate symbolic coefficients for vectors of physics, including the $(4+6)$-momentum vector of a V_{10} as given in (53·1); but at first we confine our attention to their purely algebraic properties. We can pick out three important kinds of sub-set:

(1) *Pentads*. The table (53·4) shows that two symbols commute or anticommute according as they have no suffix or one suffix in common. We therefore obtain a sub-set of mutually anticommuting symbols by fixing one suffix and letting the other vary, e.g. E_{30}, E_{31}, E_{32}, E_{34}, E_{35}. Such a sub-set is called a 'pentad'.

(2) *Anti-tetrads*. The largest sub-set of mutually commuting symbols $E_{\mu\nu}$ is three, e.g. E_{01}, E_{23}, E_{45}. Such a sub-set is called an 'anti-triad'. By adding the algebraic symbol E_{16}, we obtain a set of four mutually commuting symbols called an 'anti-tetrad'.

(3) *Conjugate triads*. Besides the triads of anticommuting symbols which form parts of pentads, there exist cyclic triads $E_{\mu\nu}$, $E_{\nu\sigma}$, $E_{\sigma\mu}$. To any one of these there is a conjugate triad $E_{\tau\lambda}$, $E_{\lambda\rho}$, $E_{\rho\tau}$. We therefore distinguish as a sub-set a 'conjugate triad-pair', e.g. E_{23}, E_{31}, E_{12} and E_{45}, E_{50}, E_{04}. Within each triad the symbols anticommute, but each member of one triad commutes with each member of the other triad.

These structural features—pentads, anti-tetrads, and triad-pairs—form a pattern of interlacing. The description of a structural pattern, without ascribing any particular nature to the elements of the pattern, or even to the operation which we picture as 'weaving', is the essence of the group concept as applied in physics.

54. Miscellaneous properties of E-symbols

We shall often use a one-suffix notation E_μ ($\mu = 1, 2, \ldots, 16$) for the set of E-symbols; so that an E-number is written as

$$P = \sum_1^{16} E_\mu p_\mu. \tag{54·1}$$

The following results are easily established:[a]

(*a*) If E_σ is any one of the symbols, the 16 products $E_\sigma E_\mu$ ($\mu = 1, 2, \ldots, 16$) reproduce the set E_μ in a different order, apart from algebraic factors ± 1 or $\pm i$.

(*b*) The 16 symbols are not connected by any linear algebraic relation; that is to say, the set is not redundant.

(*c*) If an E-number $\Sigma E_\mu p_\mu$ is equal to 0, every component p_μ is equal to 0. If $\Sigma E_\mu p_\mu = \Sigma E_\mu p'_\mu$, then $p_\mu = p'_\mu$.

(*d*) Each symbol, except E_{16}, anticommutes with 8 symbols, namely the remaining members of the two pentads to which it belongs; it commutes with the remaining 8 symbols, which include itself and E_{16}.

(*e*) If P commutes with E_σ, every non-vanishing term of P commutes with E_σ; if P anticommutes with E_σ, every non-vanishing term of P anticommutes with E_σ.

(*f*) Each symbol, except E_{16}, anticommutes with at least one member of a given tetrad. (A tetrad consists of four members of a pentad.) Hence, if an E-number com-

[a] Proofs, where necessary, are given in *Protons and Electrons*, §§ 2·4, 2·5.

mutes with all four members of a tetrad, it is an algebraic number. This is the criterion generally used in practice when we want to prove that a symbolic expression reduces to an algebraic number.

(*g*) The components p_μ of an E-number P are given by

$$p_\mu = \cdot\cdot \operatorname{qs}(E_\mu P). \tag{54·2}$$

For, by (*a*), $E_\mu P$ consists of 16 terms, one of which is algebraic; and we see at once that the algebraic term is $E_\mu . E_\mu p_\mu = -p_\mu$.

(*h*) For two E-numbers P, Q

$$\operatorname{qs}(PQ) = \operatorname{qs}(QP) = -\Sigma_\mu p_\mu q_\mu . \tag{54·3}$$

(*i*) For any E-number P

$$\sum_1^{16} E_\mu P E_\mu = -16 \operatorname{qs} P. \tag{54·4}$$

For, considering a term $E_\sigma p_\sigma$ of P, the sum $\Sigma E_\mu (E_\sigma p_\sigma) E_\mu$ gives $E_\sigma p_\sigma$ eight times and $-E_\sigma p_\sigma$ eight times, except when $\sigma = 16$.

(*j*) If four symbols forming a tetrad, e.g. E_{01}, E_{02}, E_{03}, E_{04}, are given, together with E_{16}, the remaining symbols can be expressed as products of these, so that the whole set is uniquely defined. We therefore regard the complete set as 'generated' by a tetrad.

55. Equivalence and chirality

An E-number P is said to be non-singular if there exists a reciprocal E-number P^{-1} such that $P^{-1}P = 1$. It can easily be shown that the reciprocal, if any, is unique and that it commutes with P; so that the same reciprocal is defined by the condition $PP^{-1} = 1$.

If q is any non-singular E-number, the transformation

$$E'_\mu = qE_\mu q^{-1} \tag{55·1}$$

gives a set of symbols E'_μ which satisfy the multiplication table (53·4) together with $E'_{16} = \pm i$. For example,

$$E'_{\mu\sigma} E'_{\nu\sigma} = qE_{\mu\sigma} q^{-1} qE_{\nu\sigma} q^{-1} = qE_{\mu\sigma} E_{\nu\sigma} q^{-1} = qE_{\mu\nu} q^{-1} = E'_{\mu\nu}.$$

Thus the two sets of symbols E_μ, E'_μ have the same group structure, and are *equivalent sets*.

An E-number $P = \Sigma E_\mu p_\mu$ can also be expressed as an E'-number $\Sigma E'_\mu p'_\mu$. By (54·2) the new coefficients are

$$p'_\mu = -\operatorname{qs}(E'_\mu P) = -\operatorname{qs}(qE_\mu q^{-1} P). \tag{55·2}$$

We regard p_μ and p'_μ ($\mu = 1, 2, \ldots, 16$) as the components of P in two different, but equivalent, reference frames. We shall therefore commonly speak of the set of 16 symbols as a *frame*. The E-frame and E'-frame are relativistically equivalent in the sense in which two Lorentz frames are equivalent. Intrinsically they are indistinguishable, because the symbols have by definition no properties other than that of satisfying the multiplication table. But having arbitrarily chosen a frame to be designated E, we can distinguish other frames E', E'', ..., by giving the components in the initial frame E of the transformation E-numbers q', q'', ... which generate them. Similarly, having

arbitrarily chosen an initial Lorentz frame, we distinguish other frames by giving the velocities, referred to the initial frame, of their origins.

We shall apply E-frames to physics in such a way that the equivalence of Lorentz frames is included in the equivalence of E-frames, and the transformations of special relativity theory are included in the general transformation (55·1).

Consider a physical system S whose characteristics are specified by E-numbers $P, Q, R, \ldots$, where $P = \Sigma E_\mu p_\mu$, $Q = \Sigma E_\mu q_\mu$, ... Let S' be another system correspondingly specified by $P', Q', R', \ldots$, where $P' = \Sigma E'_\mu p_\mu$, $Q' = \Sigma E'_\mu q_\mu$, ... Since the systems are constructed according to the same specification $(p_\mu, q_\mu, \ldots)$ in different but equivalent frames, they have the same intrinsic indistinguishability as the frames, and are 'equivalent systems'.

Regarding the transformation $S \to S'$ as a change undergone by a physical system, we call it a *rotation*. The term 'rotation' is contrasted with 'strain'. We shall use the terms 'rotation' and 'strain' with the widest generality to describe changes of the specification of a system, the distinction being that a strain involves an intrinsic alteration of the system, whereas a rotation merely changes it into another equivalent system. For a rotation we have, by (55·1),

$$P' = \Sigma E'_\mu p_\mu = q(\Sigma E_\mu p_\mu) q^{-1} = qPq^{-1}, \tag{55·3}$$

so that the transformation $q(\ldots)q^{-1}$ applied to any E-number is a rotation, i.e. a rotation of the carrier of the characteristics specified by P. For emphasis we shall sometimes use the expression 'relativity rotation', as a reminder that the system remains relativistically equivalent to its former self—as in the rotations of special relativity theory.

If, in the rotation of a physical system, the reference frame is kept unchanged, (55·3) is re-written as

$$P' = \Sigma E_\mu p'_\mu = qPq^{-1}, \tag{55·4}$$

and the rotation is then specified by the transformation $p_\mu \to p'_\mu$ of the components. The nature of the transformation $p_\mu \to p'_\mu$ will be investigated in § 56.

It is necessary to examine whether all possible rotations are included in $q(\ldots)q^{-1}$. It can be proved[a] that, if E'_μ is any set of E-numbers satisfying the multiplication table, and if $E'_{16} = E_{16}$, there exists a transformation $E'_\mu = qE_\mu q^{-1}$, where

$$q = \alpha \Sigma_\mu E'_\mu E_\sigma E_\mu, \quad q^{-1} = \alpha \Sigma_\mu E_\mu E_\sigma E'_\mu, \tag{55·5}$$

α being an algebraic number, and E_σ being one of the 16 E-symbols. The appropriate E_σ has to be found by trial and error. Usually the product of $\Sigma_\mu E'_\mu E_\sigma E_\mu$ and $\Sigma_\mu E_\mu E_\sigma E'_\mu$ is 0; but there is just one E_σ which makes the product a non-zero algebraic number, and α can then be chosen so that (55·5) satisfies $qq^{-1} = 1$.

The explanation of this rather complicated procedure is as follows. By § 54 (j) the set E'_μ is generated by a tetrad, say E'_{01}, E'_{02}, E'_{03}, E'_{04}. The 16 sets generated by the tetrads $\pm E'_{01}$, $\pm E'_{02}$, $\pm E'_{03}$, $\pm E'_{04}$ are called *images* of one another. These images are sets connected with E_μ by the transformation (55·5), E_σ being different in each case. The image for which $E_\sigma = E_{16}$ is said to 'correspond' to E_μ. Between corresponding sets the transformation (55·5) takes the simpler form

$$q = \alpha \Sigma E'_\mu E_\mu, \quad q^{-1} = \alpha \Sigma E_\mu E'_\mu. \tag{55·6}$$

[a] *Protons and Electrons*, §§ 2·7, 2·8.

Since $E_{16} = \pm i$, it commutes with q; and $qE_{16}q^{-1} = E_{16}qq^{-1} = E_{16}$. Thus there is no transformation $q(\ldots)q^{-1}$ between sets such that $E'_{16} = -E_{16}$.

To sum up: E-frames are divided into two systems according as $E_{16} = i$ or $E_{16} = -i$. In either system every pair of frames is connected by a rotation $q(\ldots)q^{-1}$; but there is no rotation between frames of opposite systems. Evidently the frames of each system form a continuum, and the rotation $q(\ldots)q^{-1}$ can be performed in a series of infinitesimal steps. There is complete discontinuity between the two systems.[a]

Thus the symbolic scheme provides for the representation of structures which, though intrinsically similar, cannot be rotated into one another. This distinction, familiarly illustrated by a right-handed and a left-handed screw, is called *chirality*.[b]

The two kinds of frame will be distinguished as *right-handed* and *left-handed*. Conventionally the frames with $E_{16} = i$ are right-handed. This is non-committal, because it is impossible to define which of the two algebraic square roots of -1 has been designated i.

From the table (53·4) we find that a right-handed frame E_μ and a left-handed frame F_μ may have 10 symbols in common, the remaining 6 being reversed in sign. If, starting with the same generating tetrad E_{15}, E_{25}, E_{35}, E_{45}, we construct right- and left-handed sets, the symbols reversed in sign are

$$E_{01},\ E_{02},\ E_{03},\ E_{04},\ E_{05},\ E_{16}. \tag{55·7}$$

Consider two particles, carriers of variates p_μ, which are specified by E-numbers $P = \Sigma E_\mu p_\mu$, $P^\dagger = \Sigma F_\mu p_\mu$. These, being similarly constructed in chirally opposite frames E_μ, F_μ, have the same kind of opposite chirality as the frames; they are intrinsically similar but cannot be rotated into one another. When both particles are referred to the right-handed frame E_μ, $P^\dagger$ becomes $\Sigma E_\mu p^\dagger_\mu$, where

$$\left.\begin{aligned} p^\dagger_\mu &= -p_\mu, && \text{for the 6 components in (55·7),} \\ p^\dagger_\mu &= p_\mu, && \text{for the other 10 components.} \end{aligned}\right\} \tag{55·8}$$

According to classical electromagnetic theory the distinction between positive and negative electrification consists in opposite chirality, positive and negative charges being sources of chirally opposite rotational strains in the aether. We may therefore expect that P and $P^\dagger$ will describe particles which are electrically or magnetically opposite; and that the 6 components (55·7) which are reversed in sign, will describe distinctively electromagnetic characteristics. The 10 components $p_{\mu\nu}$ with $\mu, \nu = 1, 2, 3, 4, 5$, which are unreversed, will then represent mechanical characteristics. They are the components which we have already provisionally allotted to the linear momentum 4-vector and the angular momentum 6-vector in (53·1).

We have thus an outline of a scheme of physical identification of the components of an E-number. It is not suggested that the foregoing considerations are a proof of the identification; nor, on the other hand, are we indulging in speculation. The position is that we have acquired a new mathematical tool—a symbolic E-frame—and we have to learn the art of using it. We shall not achieve much by prodding about with it aimlessly. It is guidance as to method, not demonstration or hypothesis, that concerns us at this stage.

[a] The recognition of the two systems of frames is due to S. R. Milner.

[b] The term appears to have been first introduced by Kelvin.

Normally when a single particle is contemplated as object-system, the environment is taken to be a neutral unpolarisable uranoid. The distinctively electromagnetic components of P are then dormant; there is nothing for them to interact with, and they have no observational significance. This idealisation of the environment gives the V_{10} particles, whose 'complete momentum vector' consists of the 10 mechanical components. When all 16 components are included we call P the *extended momentum vector.*

56. Rotations

We shall now investigate the transformation $p_\mu \to p'_\mu$ of the components of an E-number corresponding to a rotation $q(\ldots)q^{-1}$. We consider first a 'simple rotation'

$$q = e^{\frac{1}{2}E_{\mu\nu}\theta}. \tag{56·1}$$

An exponential containing a non-algebraic symbol is understood to be defined by the exponential series. From this definition it is easily found that

$$e^{E_{\mu\nu}\theta} = \cos\theta + E_{\mu\nu}\sin\theta, \quad e^{E_{\mu\nu}\theta}e^{-E_{\mu\nu}\theta} = 1, \tag{56·21}$$

just as though $E_{\mu\nu}$ were an algebraic square root of -1. In fact, the usual theory of exponentials applies unless there are non-commuting terms in the exponent; but, for example, $e^{E_{01}\theta_1+E_{02}\theta_2} \neq e^{E_{01}\theta_1}.e^{E_{02}\theta_2}$. For $q = e^{\frac{1}{2}E_{\mu\nu}\theta}$ we have $qX = Xq$ if X commutes with $E_{\mu\nu}$, and $qX = Xq^{-1}$ if X anticommutes with $E_{\mu\nu}$. The rotation $X' = qXq^{-1}$ therefore gives $X' = X$ if X commutes with $E_{\mu\nu}$, and

$$X' = Xq^{-2} = Xe^{-E_{\mu\nu}\theta} = X(\cos\theta - E_{\mu\nu}\sin\theta)$$

if it anticommutes. Hence, if μ, ν, σ, τ are different suffixes,

$$\left.\begin{aligned} E'_{\sigma\tau} &= E_{\sigma\tau}, \\ E'_{\mu\sigma} &= E_{\mu\sigma}(\cos\theta - E_{\mu\nu}\sin\theta) = E_{\mu\sigma}\cos\theta + E_{\nu\sigma}\sin\theta. \end{aligned}\right\} \tag{56·22}$$

The rotation q transforms $P = \Sigma E_\mu p_\mu$ to $P' = \Sigma E'_\mu p_\mu$, which is then referred to the original frame E_μ and re-written $P' = \Sigma E_\mu p'_\mu$; so that

$$P' = \Sigma E'_\mu p_\mu = \Sigma E_\mu p'_\mu. \tag{56·31}$$

Hence, by (54·2), $$p_\mu = -\operatorname{qs}(E'_\mu P'), \quad p'_\mu = -\operatorname{qs}(E_\mu P'). \tag{56·32}$$

Then $$p_{\sigma\tau} = p'_{\sigma\tau}, \tag{56·41}$$

$$\begin{aligned} p_{\mu\sigma} &= -\operatorname{qs}\{(E_{\mu\sigma}\cos\theta + E_{\nu\sigma}\sin\theta)\,P'\} \\ &= -\operatorname{qs}(E_{\mu\sigma}P')\cos\theta - \operatorname{qs}(E_{\nu\sigma}P')\sin\theta \\ &= p'_{\mu\sigma}\cos\theta + p'_{\nu\sigma}\sin\theta. \end{aligned} \tag{56·42}$$

For example, the rotation $q = e^{\frac{1}{2}E_{12}\theta}$ gives

$$p_{01} = p'_{01}\cos\theta + p'_{02}\sin\theta, \quad p_{02} = p'_{02}\cos\theta - p'_{01}\sin\theta, \quad p_{03} = p'_{03}, \quad p_{04} = p'_{04}. \tag{56·5}$$

The whole result of the E_{12} rotation is that 4 pairs of components (p_{01}, p_{02}; p_{31}, p_{32}; p_{41}, p_{42}; p_{51}, p_{52}) are rotated through an angle θ in their respective planes like rectangular components of a vector, and the remaining components are unchanged. By taking $q = e^{\frac{1}{2}E_\mu\theta}$ with $\mu = 1, 2, \ldots, 15$, we obtain 15 independent simple rotations, each

rotating 4 pairs of components and leaving 8 components unchanged. If $\mu = 16$, P is unchanged.

Not every pair of components can be linked in rotation. Two components which can rotate together are said to be *perpendicular*; those which cannot are said to be antiperpendicular. By (56·42), components are perpendicular if their symbols anticommute, and antiperpendicular if their symbols commute. The rotation linking two perpendicular components $p_{\mu\sigma}$, $p_{\nu\sigma}$ is $q = e^{\frac{1}{2}E_{\mu\nu}\theta_{\mu\nu}}$; thus the symbols of the two components and their rotation form a cyclic triad.

Consider four mutually perpendicular components p_{01}, p_{02}, p_{03}, p_{04}. If we represent them geometrically as components of a vector referred to rectangular axes x_1, x_2, x_3, x_4 in 4-space, it follows from (56·5) that the E_{12} rotation is represented as a geometrical rotation in the x_1x_2 plane. Similarly the symbolic rotations with E_μ equal to

$$E_{23},\ E_{31},\ E_{12},\ E_{14},\ E_{24},\ E_{34}, \tag{56·6}$$

are represented by ordinary geometrical rotations in the six coordinate planes of the 4-space.

The transformation which rotates p_{01} and p_{02} also rotates three other pairs of components. This coupling of transformations is commonly provided for by assigning appropriate vector or tensor character to the sets of variables concerned. It is easily shown that the correct coupling is obtained if we represent the components of P by

$$\left.\begin{array}{ll} (a) & \text{a 4-vector } p_{01},\ p_{02},\ p_{03},\ p_{04}, \\ (b) & \text{a 4-vector } p_{15},\ p_{25},\ p_{35},\ p_{45}, \\ (c) & \text{a 6-vector } p_{23},\ p_{31},\ p_{12},\ p_{14},\ p_{24},\ p_{34}, \\ (d) & \text{two invariants } p_{05},\ p_{16}. \end{array}\right\} \tag{56·7}$$

According to (55·7) one 4-vector is mechanical and one electrical; the 6-vector is mechanical, and both invariants are electrical.

We have thus obtained a representation of E-numbers by vectors in a 4-space. Conversely, the ordinary description of physical systems in the 4-space by 4-vectors and 6-vectors can be transformed into a description by E-numbers. That is to say, following the quaternion device of expressing a vector as a linear function of its components, we can now identify the symbolic coefficients with symbols of an E-frame, as suggested in (53·1). Provided that we can distinguish mechanical and electrical 4-vectors, the appropriate symbols can be found at once from (56·7). The essential point is that the ordinary relativistic properties of physical systems for rotations in the 4-space are now contained in the equivalence of E-frames, the relativistic rotations of special relativity theory being a selection (56·6) from the general rotations $q(\ldots)q^{-1}$ which transform an E-frame into an equivalent frame.

The E-frame is a construct of pure mathematics, and we are free to use it in physics in any way that seems likely to be useful. We are now able to define the particular application that we intend to follow up. *The E-frame will be brought into physics by identifying the 4-space (associated with it in the way defined above) with space-time.*

By this step we 'anchor' the E-frame in observational physics. A great many details remain to be elaborated in the course of the development of the theory. Our starting point is that, when physical space-time is rid of the inessentials which embellish

it in our ordinary conception, all that remains to describe is the structure of its relativistic transformations. The E-frame provides a symbolism for formulating this structure.

By way of contrast with rotations, it is useful to consider the transformation $q(\ldots)q$, which we shall call a *pseudo-rotation*.[a] A pseudo-rotation $q = e^{\frac{1}{2}E_{\mu\nu}\theta}$ gives instead of (56·22),

$$\left.\begin{aligned} E'_{\mu\sigma} &= E_{\mu\sigma}, \\ E'_{\sigma\tau} &= E_{\sigma\tau}(\cos\theta + E_{\mu\nu}\sin\theta) = E_{\sigma\tau}\cos\theta + iE_{\lambda\rho}\sin\theta, \end{aligned}\right\} \tag{56·8}$$

where $\mu, \nu, \sigma, \tau, \lambda, \rho$ is an even (odd) permutation of 0, 1, 2, 3, 4, 5 for a right- (left-) handed frame. It follows as in (56·42) that

$$p_{\mu\sigma} = p'_{\mu\sigma}, \quad p_{\sigma\tau} = p'_{\sigma\tau}\cos\theta + ip'_{\lambda\rho}\sin\theta. \tag{56·91}$$

For example, a pseudo-rotation $e^{\frac{1}{2}E_{23}\theta}$ gives in a right-handed frame,

$$p_{01} = p'_{01}\cos\theta + ip'_{45}\sin\theta, \quad ip_{45} = ip'_{45}\cos\theta - p'_{01}\sin\theta. \tag{56·92}$$

Pseudo-rotations are therefore complementary to rotations, each connecting in formal rotation the pairs of components that the other leaves unconnected. A pseudo-rotation rotates antiperpendicular components; the symbols of the two components and the pseudo-rotation connecting them form an anti-triad.

If we make the transformation $E'_\mu = qE_\mu q$, the symbols E'_μ do not satisfy the multiplication table (53·4), and the frame is therefore intrinsically different from the E-frame. Thus pseudo-rotations are not relativistic transformations. That does not mean that they are of minor importance in physics. A pseudo-rotation is a strain; and the representation as a transformation $q(\ldots)q$ provides a basis for the systematic classification of strains.

The most general infinitesimal rotation is given by $q = 1 + \frac{1}{2}d\Theta = e^{\frac{1}{2}d\Theta}$, where $d\Theta$ is any infinitesimal E-number. This is resolvable into simple rotations $e^{\frac{1}{2}E_1 d\theta_1}$, etc. We cannot, in general, resolve a finite rotation quite in this way; because the usual law of combination of exponentials is not valid when the symbols do not commute. The most general finite rotation $q(\ldots)q^{-1}$ can be resolved into a *succession* of simple rotations, the order of their application being indicated. We need not dwell further on this, because the non-commutation of rotations is familiar in elementary geometry; and it is the same property that we here meet with in symbolic form.

57. Five-dimensional theory

The analysis of P into two 4-vectors, a 6-vector and two invariants corresponds to the usual representation of physical systems in space-time. Representation in space-time implicitly excludes rotations which do not transform space-time into itself; for that reason the transformations of the vectors in (56·7) are limited to the 6 rotations (56·6) which correspond to purely internal rotations of space-time. A rotation such as $q = e^{\frac{1}{2}E_{15}\theta}$ would mix up the sets of components (a), (b), (c) which in four-dimensional theory form entirely distinct vectors. To exhibit a more extended range of relativistic

[a] This is called an 'antiperpendicular rotation' in *Protons and Electrons*, but I now prefer a terminology which makes it clear that it is not a rotation.

transformation, we must adopt a five-dimensional representation of physical variates. Then P breaks up into

$$\left.\begin{array}{l}(a)\ \text{a 5-vector, composed of electrical components } p_{01}, p_{02}, p_{03}, p_{04}, p_{05},\\ (b)\ \text{a 10-vector (analogue of a 6-vector in four dimensions), composed of the 10 mechanical components,}\\ (c)\ \text{an invariant (electrical) } p_{16}.\end{array}\right\} \quad (57{\cdot}1)$$

In this representation there are 10 independent relativity rotations in the 10 coordinate planes of the 5-space, the associated symbols being

$$E_{23}, E_{31}, E_{12}, E_{14}, E_{24}, E_{34}, E_{15}, E_{25}, E_{35}, E_{45}. \quad (57{\cdot}2)$$

To exhibit all 15 rotations we require a six-dimensional representation. The components, other than p_{16}, form a 15-vector or antisymmetrical tensor of the second rank in the 6-space. The components are associated with the 15 coordinate planes as are also the independent rotations; so that in applying a rotation to P we are applying a rotation to a rotation (or to something structurally equivalent to a rotation). This identity of operator and operand is the characteristic of a 'group'; so that we may say briefly that the six-dimensional representation is a representation of the group-structure, the group of the E-symbols being the same as the rotation group in 6-space. The ordinary vectors of the 6-space have no counterpart in the E-frame.

Beginning with the six-dimensional representation of the group-structure of the E-frame, the symbols are associated initially with coordinate planes. Dropping one dimension, 5 of the symbols are associated with coordinate axes, the other 10 being associated with coordinate planes as before. This begins to be something like an ordinary space; but it is to be noticed that, when it is compared with physical space there is an inversion of mechanical and electrical characteristics. The symbols of the axes are the electrical symbols $E_{0\mu}$ ($\mu = 1, 2, \ldots, 5$); and the vectors in the 5-space are electrical vectors $\Sigma E_{0\mu} p_{0\mu}$; that is to say, they represent those properties of particles which are dormant in a neutral uranoid. The relation of the 5-space to space-time is not simply the addition of a dimension; it involves a change of the symbolic directions of the axes from

$$E_{15}, E_{25}, E_{35}, E_{45} \quad \text{to} \quad E_{01}, E_{02}, E_{03}, E_{04}, E_{05}. \quad (57{\cdot}3)$$

The formal reason is that each pentad contains at least one electrical symbol (suffix 0); so that we cannot extend the momentum vectors and conjugate position vectors of ordinary mechanics to more than four dimensions.

The relation (57·3) is clarified if we distinguish tensors and tensor-densities in four dimensions. In tensor calculus the measured volume dV and the coordinate volume $d\tau$ of a four-dimensional element are connected by $dV = \sqrt{(-g)}\, d\tau$ or, as it is here more appropriately written $dV = i\sqrt{g}\, d\tau$. In rectangular coordinates $g = -1$, and the factor $\sqrt{-g}$ can be omitted. But in symbolic theory the square root of -1 is not necessarily $\pm i$; it may be an E-symbol. If we introduce a fifth dimension, dV becomes a vector having the symbolic direction E_{05} of the fifth dimension; and this must be taken into account in combining it, by multiplication, with tensors represented in five dimensions. Since $d\tau$ is by definition purely numerical, and is treated as an ordinary number in integration, we have $\sqrt{g} = \pm E_{05}$.

More formally, the four-dimensional volume element dV is the product of four vectors represented symbolically by $E_{15}dx_1$, $E_{25}dx_2$, $E_{35}dx_3$, $E_{45}dx_4$. By the multiplication table $E_{15}E_{25}E_{35}E_{45} = E_{16}E_{05}$. Thus (in a right-handed frame)

$$dV = iE_{05}dx_1dx_2dx_3dx_4. \tag{57·4}$$

Thus a vector P and its four-dimensional density $\mathfrak{P} = P\sqrt{-g}$ are related by

$$\mathfrak{P} = PiE_{05}, \quad P = \mathfrak{P}iE_{05}, \tag{57·5}$$

the second result being derived from the first by multiplying finally by iE_{05}. In particular, we have the correspondence

$$\left.\begin{aligned} i\mathfrak{P} &= E_{01}p_1 + E_{02}p_2 + E_{03}p_3 + E_{04}p_4 + E_{05}p_5, \\ P &= E_{15}p_1 + E_{25}p_2 + E_{35}p_3 + E_{45}p_4 + E_{16}ip_5. \end{aligned}\right\} \tag{57·6}$$

Thus the ordinary momentum 4-vector p_1, p_2, p_3, p_4, which is a vector in the four-dimensional mechanical frame E_{15}, E_{25}, E_{35}, E_{45}, is a vector-density in the five-dimensional frame E_{01}, E_{02}, E_{03}, E_{04}, E_{05}. We notice that, although the complete expression for P includes a component (with symbol E_{05}) perpendicular to p_1, p_2, p_3, p_4, this is not the component that appears as p_5 in the five-dimensional representation.

Since the relation (57·5) between P and $\mathfrak{P}$ is reciprocal, it would be possible to interchange P and $\mathfrak{P}$ in (57·6). That would give a different anchorage of the symbolic frame in observational physics. We adopt (57·6) as it stands, because then the typical 4-vectors of physics (momentum vectors and position vectors) remain vectors in four-dimensional symbolic theory; it would cause confusion of nomenclature if they became vector-densities when put into symbolic form. But it may be remarked that the alternative anchorage is more logical if we wish to present the theory in purely deductive form, beginning with an E-frame and proceeding to find physical identifications of the quantities derived from it. In that case we should begin with a pentad of five axes $E_{0\mu}$, and introduce 5-vectors $P = \Sigma_\mu E_{0\mu}p_\mu$. We should then find that the momentum and position 4-vectors (ordinarily so-called) are actually vector-densities in a particular sub-space of four dimensions, the true vectors in that sub-space being electrical.

58. Ineffective relativity transformations

It is important to understand clearly the function of relativity transformations. The language of physics is such that two different descriptions may specify the same object or condition. This is very natural, because we do in fact encounter different aspects of the same object and describe it from those aspects. But, through generalisation and systematisation, multiplicity of description has been extended much beyond the original intention. It is regulated by a more or less systematic transformation theory, and descriptions are introduced for the sake of completeness, although they could scarcely arise from any imaginable aspect. It would not be possible to systematise observational knowledge without this multiplicity of description; but it needs an antidote, because much mystification results if we do not easily recognise that two descriptions are equivalent. The antidote is a knowledge of relativity transformations, these being the transformations which relate equivalent descriptions.

We shall call a relativity transformation *effective* if it relates two modes of description in current use. If the transformation is *ineffective* it only means that the language of

physics is not so redundant as it might have been. When we investigate relativity transformations systematically, as in the theory of the rotations of the E-frame, effective and ineffective transformations are mixed together indiscriminately. This may seem an invitation to extend the terminology of physics so as to provide expression for the results of transformations which, relative to the present terminology, are ineffective. But that is a mathematical, rather than a physical, attitude. The physicist would scarcely be well advised to extend the multiplicity of his descriptions for no other purpose than to give more opportunity to the mathematician to remove the resulting mystification.

The common view is that there are six effective relativity rotations (three spatial rotations and three Lorentz transformations), but this is not a hard and fast rule applying to all branches of physics. The distinctive property of these rotations is that they transform flat space-time into itself; so that the descriptions, identified as equivalent by these transformations, are descriptions of objects and conditions in the same flat space-time. But already our descriptions have been forced to transcend flat space-time in two ways. The extraordinary fluctuation has introduced (as an alternative to the more radical complication of curvature) an extra phase dimension. The suffixing of particles has introduced extra-spatial interchange circulation. We have therefore to recognise equivalences of descriptions over a wider field than that comprised in space-time representation, and this gives scope for others of the 15 relativity rotations of the frame to become effective. At the same time the three Lorentz transformations become ineffective in statistical physics; for we have seen that the time coordinate is differentiated from the space coordinates at the outset of the study of probability distributions. The curious insistence on introducing Lorentz transformations, which appears so often in the literature of quantum theory, is an example of the introduction of mystification which can have no other purpose than to give the mathematician an opportunity of removing it.

Our attitude is that there is no general principle by which we can decide once for all which of the rotations are effective. I may, however, say in anticipation that normally in microscopic theory the new effective rotations compensate for the loss of the Lorentz transformations; so that there are six effective rotations which have close analogy with, but are not identical with, the six rotations of molar space-time. One of the uses of the E-frame is to show how the molar and microscopic systems of relativity transformation are related.

No apology is needed for the ineffectiveness of a rotation; it could be made effective by extending our terminology, but it is judged not worth while to do so. It may seem rather a fiasco that we should introduce an E-frame providing a considerable increase in the number of relativity rotations only to find that most of the increase is ineffective. But that is a misunderstanding. It is not the relativity rotations, but the strains that we are after. The recognition of equivalence of description is a clearing of the ground. It is when we come to the strains that positive construction begins.

59. Strain vectors

When P represents a quantity distributed over a region of space, we usually have to consider the density of P rather than P itself. The density of a vector in the three-dimensional space x_1, x_2, x_3 will be called a *strain vector*. A distinctive name is required

because the term vector-density in tensor calculus refers to the density in a four-dimensional element.

The reciprocal V^{-1} of a three-dimensional volume element is a 4-vector which, for a volume of the space x_1, x_2, x_3, is in the x_4 direction. Moreover, in natural units, it is a momentum vector, and it has entered into our theory as such. It has therefore the same symbolic coefficient E_{45} as the p_4 component of a momentum vector, and we can write

$$V^{-1} = \pm E_{45} v^{-1},$$

where v is a number. The density of P, when distributed over V, is accordingly $\pm PE_{45}v^{-1}$. Thus a strain vector has the form PE_{45}, where E_{45} is the symbol associated with the time axis.

The strain vector S corresponding to a vector P is defined by

$$S = -PE_{45}. \tag{59·1}$$

Both the sign and the order of the factors P, E_{45} are conventional. Multiplying (59·1) finally by E_{45}, we have

$$P = SE_{45}. \tag{59·2}$$

Then P is the quantity in a unit volume which corresponds to the density S.

Setting $P = \Sigma E_\mu p_\mu$, $S = \Sigma E_\mu s_\mu$, (59·1) gives the following relation between the components:[a]

Momentum 4-vector:

$$p_{15}, p_{25}, p_{35}, p_{45} = -s_{14}, -s_{24}, -s_{34}, is_{16},$$

Spin 6-vector:

$$p_{23}, p_{31}, p_{12}, p_{14}, p_{24}, p_{34} = is_{01}, is_{02}, is_{03}, s_{15}, s_{25}, s_{35}, \tag{59·3}$$

Electrical components:

$$p_{01}, p_{02}, p_{03}, p_{04}, p_{05}, p_{16} = is_{23}, is_{31}, is_{12}, s_{05}, -s_{04}, is_{45}.$$

In molar theory there is an alternative definition of a three-dimensional volume which may be called the *ordered volume*. The ordered volume contained by three orthogonal linear elements d^1x_μ, d^2x_μ, d^3x_μ is[b]

$$\boldsymbol{V}_\tau = \epsilon_{\mu\nu\sigma\tau} d^1x_\mu d^2x_\nu d^3x_\sigma, \tag{59·41}$$

where $\epsilon_{\mu\nu\sigma\tau}$ is the well-known alternating operator. If the elements have the directions of the three space axes, they are represented by $E_{15}d^1x_1$, $E_{25}d^2x_2$, $E_{35}d^3x_3$, and (59·41) reduces to

$$\boldsymbol{V}_4 = \epsilon_{1234} E_{15}E_{25}E_{35} d^1x_1 d^2x_2 d^3x_3 = E_{16}E_{04}v. \tag{59·42}$$

Thus the ordered volume has a symbolic coefficient $\pm iE_{04}$ instead of the coefficient $\pm E_{45}$ found for V. The difference is explained by reference to ordinary tensor calculus where it is shown that the ordered volume is strictly, not a vector, but a 'vector-volume' $\boldsymbol{V}_\tau/\sqrt{-g}$.[c] By comparison we obtain $iE_{04}\sqrt{-g} = \pm E_{45}$; so that $\sqrt{-g} = \pm iE_{05}$, agreeing with § 57.

This treatment exhibits the connection between the chirality of the E-frame and the chirality of a set of rectangular axes. It is assumed in (59·42) that the order of multiplication of the elements is the order of the axes along which they are directed. If this

[a] We have written i for E_{16}. In future all formulae containing i are given for a right-handed frame, unless otherwise stated.

[b] We use the summation convention temporarily.

[c] *Mathematical Theory of Relativity*, 2nd ed. p. 244.

is changed from a right-handed to a left-handed sequence, the sign of V_4 is reversed. The sign of the right-hand side of (59·42) is reversed by changing E_{16} from i to $-i$, i.e. by reversing the chirality of the frame. Thus reversing the chirality of the E-frame is equivalent to reversing the chirality of the axes associated with it.

The special importance of strain vectors is that they form an intermediate step between the purely abstract 'problem of one particle' and the realistic but complicated problem of a system of particles. A single particle in a geometrical or algebraic frame affords nothing observable; we have to deal with it as an auxiliary conception, but we want to pass on as quickly as possible to something that has the semblance of an observable system.

Irrespective of the constitution of the rest of the system, a radical change of conception occurs when a particle is considered as part of a system. A system introduces the concept of *simultaneity*. Each particle has its own space-coordinates; but there is only one time coordinate, which is in the direction of the resultant momentum vector of the system. Anticipating that in any observational problem the particle will be part of a system, in which a time axis and a corresponding simultaneous space will be defined independently of the particle itself, we consider P in conjunction with this independent time direction which we represent by the symbol E_{45}. The strain vector $S = -PE_{45}$ arises out of this association. To understand the relation of S to P, consider the motion of the particle in the x_1 direction which is specified by the momentum component $E_{15}p_{15}$. This could be eliminated by a Lorentz transformation in the plane E_{14}. But, since our time axis has already been fixed by other considerations, this rotation is inhibited. The 'system', composed of the particle together with the source of the fixed time direction, differs intrinsically from—is strained in comparison with—a system in which $p_{15} = 0$. The nature of the strain is appropriately specified by the inhibited rotation which would have removed it, and has therefore the symbolic coefficient E_{14} of the inhibited rotation. This is the strain vector mode of description, in which a strain component $E_{14}s_{14}$ takes the place of a momentum component $E_{15}p_{15}$. By (59·3) we have $is_{14} = p_{15}$, the i indicating that the inhibited rotation is hyperbolic.

The passage from vector to strain vector description corresponds to a changed conception of motion. Instead of being an unobservable relation of the particle to a mathematical reference frame, it is an observable deformation of a physical system in which the particle is contained.

When we describe the momentum strain vector S as the three-dimensional density of the momentum vector P, we refer to tensor character, not to dimensions. By (59·1), S and P have the same dimensions. Normally S is the quantity directly defined, so that S and P both have the dimensions of a momentum-density or mass-density; so that in this way we introduce 'momentum vectors' which have the dimensions of energy tensor and 'masses' which have the dimensions of densities. It is, of course, the dimensions, not the tensor character, that determine the transformation due to a change of extraneous standard. This helps us to understand the anomaly in § 26, etc., where the mass $m_{0,\,M}$ of scale-free particles transforms as densities. In a pseudo-discrete assemblage the unidentified individual particle must from the start be considered as part of a system; so that strain vector description is prior to vector description. If P is introduced, it is derived by (59·1) from the strain vector which represents the density, and retains the dimensions of density.

60. Real and imaginary E-symbols

The multiplication table (53·4) shows at once that some of the E-symbols must be real and some imaginary. Let us examine the most general way in which real or imaginary character can be assigned to them consistently with the multiplication table. Let E_{12} be one of the real symbols. It can be proved by *reductio ad absurdum* that not all the 8 symbols which anticommute with E_{12} are imaginary. Accordingly let E_{12}, E_{13} be real symbols; then their product E_{23} is also real. Turning to the conjugate triad, two cases arise: (a) all 3 symbols are real, or (b) one symbol, say E_{45}, is real, and E_{50}, E_{04} are imaginary.

In case (a) we find that the remaining 10 symbols are imaginary. Since each of the 6 pentads contains 2 members of the real triad-pair E_{22}, E_{31}, E_{12}, E_{45}, E_{50}, E_{04}, a pentad always contains 2 real and 3 imaginary symbols. In case (b) we find, by multiplying the triad-pair by the real symbol E_{45}, that E_{01}, E_{02}, E_{03}, E_{04}, E_{05}, E_{16} are imaginary; the other symbols are real. There are therefore two possible systems of assignment:

System (a), 6 real and 10 imaginary symbols. The real symbols form a pair of conjugate triads. Each pentad has 2 real and 3 imaginary symbols.

System (b), 10 real and 6 imaginary symbols. The imaginary symbols form a pentad together with E_{16}.

In saying that the E-frame consists of real and imaginary symbols with a character-distribution given either by (a) or (b), we do not go beyond the original definition of the E-symbols as a set of elements satisfying a certain multiplication table. But in practical application it is scarcely possible to leave the choice between (a) and (b) undecided. If we choose one or other system, it is no longer true to say that we regard the E-symbols as having no properties other than those stated in, or derivable from, their multiplication table. The choice introduces an additional property of the symbols, which can only be justified fundamentally by deepening our foundations. We must give the E-frame a *pre-history*, which shows how the property originates. For the most part we regard the pre-history of the E-frame as an epistemological prologue to physics, which is segregated from the main development of fundamental theory and dealt with in Chapter XIII. But here it is necessary to go one step further back.

Consider a pair of conjugate triads $E_{\mu\nu}$, $E_{\nu\sigma}$, $E_{\sigma\mu}$, $E_{\tau\lambda}$, $E_{\lambda\rho}$, $E_{\rho\tau}$, and let

$$\zeta_1, \zeta_2, \zeta_3 = E_{\mu\nu}, E_{\nu\sigma}, E_{\sigma\mu}, \quad \theta_1, \theta_2, \theta_3 = E_{\tau\lambda}, E_{\lambda\rho}, E_{\rho\tau}. \tag{60·1}$$

It will be found that the complete set of E_μ is given by

$$\zeta_\alpha, \quad \theta_\alpha, \quad i\zeta_\alpha\theta_\beta, \quad i. \tag{60·2}$$

The symbols ζ_1, ζ_2, ζ_3 are anticommuting square roots of -1, each of which is the product of the other two. Introducing an algebraic symbol $\zeta_4 = 1$, we call a linear function $a_1\zeta_1+a_2\zeta_2+a_3\zeta_3+a_4\zeta_4$ with numerical coefficients (real or complex) a ζ-number. The ζ-numbers form a closed algebra. Similarly we have a closed algebra of θ-numbers, which is homomorphic with and commutes with the ζ-algebra. The expression (60·2) for the complete set of E_μ shows that the algebra of E-numbers is the 'direct product' (outer product) of the ζ- and θ-algebras; or, since the latter algebras are identical, it is the direct square of the ζ-algebra.

The ζ-algebra (which is Hamilton's quaternion algebra) is the simplest possible non-commuting algebra. For, if there exist two anticommuting elements ζ_1, ζ_2, their product ζ_3 necessarily anticommutes with both and provides the third linearly independent element needed to complete the algebra. The E-algebra, being its direct square, is the next simplest non-commuting algebra.

A ζ-number with real coefficients a_1, a_2, a_3, a_4 will be called a *real ζ-number*. The real ζ-numbers form a closed algebra, called the real ζ-algebra. The direct square of the real ζ-algebra gives a closed algebra of *real E-numbers*

$$\Sigma E_\mu p_\mu = \Sigma a_{\alpha\beta} \zeta_\alpha \theta_\beta \quad (\alpha, \beta = 1, 2, 3, 4), \tag{60·3}$$

with real coefficient $a_{\alpha\beta}$. By (60·2), p_μ is real for the 6 symbols of the triad-pair, and imaginary for the other 10 symbols. This implies that $E_{\mu\nu}, E_{\nu\sigma}, E_{\sigma\mu}, E_{\tau\lambda}, E_{\lambda\rho}, E_{\rho\tau}$ are to be counted as real symbols, the other 10 symbols being imaginary. This is the distribution which we have called system (a).

Imaginary numbers are unknown in experimental physics. If they abound in theoretical physics, it is because the theorist has at some stage in the collation and systematisation of observational data made a definite departure which introduces them. If not already present, they are introduced by the definition of the E-symbols in (60·2); and the imaginary character of certain of the E-symbols is due to the factor i which appears explicitly in their definition.

By this deepening of the foundation our theory now rests on a ζ-algebra which is not so much 'real' as 'characterless'. Its reality is unsophisticated, because it precedes the use of complex numbers in the subject. The resulting reality system of the E-frame is system (a). I do not think there is any alternative foundation which would give system (b). At any rate we shall adopt system (a), and commit ourselves to the foregoing pre-history of the E-frame employed in our investigations.

System (a) has another property of considerable theoretical importance. There exists a representation of E-numbers by fourfold matrices (§ 71). The representation is said to be a 'true' representation, if real E-numbers are represented by wholly real matrices, and imaginary E-numbers by wholly imaginary matrices. It can be proved that a pentad composed of matrices which are wholly real or wholly imaginary contains two real and three imaginary matrices. Thus true representation is possible if we adopt system (a), but not if we adopt system (b).

Adopting system (a), we arrange the notation of the suffixes so that the scheme is

$$\left.\begin{array}{ll} \text{real symbols,} & E_{23}, E_{31}, E_{12}, E_{04}, E_{05}, E_{45}, \\ \text{imaginary symbols,} & E_{01}, E_{02}, E_{03}, E_{14}, E_{24}, E_{34}, E_{15}, E_{25}, E_{35}, E_{16}. \end{array}\right\} \tag{60·4}$$

Equation (60·1) then becomes

$$\zeta_1, \zeta_2, \zeta_3 = E_{23}, E_{31}, E_{12}, \quad \theta_1, \theta_2, \theta_3 = E_{45}, E_{50}, E_{04}. \tag{60·5}$$

We have found that a rotation $q(\ldots)q^{-1}$ gives an E'-frame equivalent to the E-frame. But an additional property has now been given to the symbols, and for *full equivalence* of the frames it is necessary that E'_μ should have the same real or imaginary character as E_μ. If a rotation $q = e^{\frac{1}{2}E_{\nu\sigma}\theta}$ is applied to $E_{\mu\sigma}$, we have $E'_{\mu\sigma} = E_{\mu\sigma}q^{-2}$; so that the

character is not preserved unless q^2 is real. Except for the special value $\theta = 90°$, q cannot be a pure imaginary. Thus

For the group of rotations which give full equivalence, q is real. (60·6)

If, in the application of the E-frame to physics, any physical interpretation is given to the real or imaginary character of the E_μ, relativistic equivalence must be identified with full equivalence. It will therefore be a condition for relativity rotation that q is real.

61. Reality conditions

The representation of physically real characteristics by imaginary or complex numbers is a common device in elementary physics. When the quantity so represented has a probability distribution, certain restrictions must be imposed. Consider, for example, the complete momentum vector. To define its 10 components p_μ in such a way that only real numbers are used would be a harassing regulation. On the other hand, if they are unrestrictedly complex, the domain of probability distribution is extended from 10 to 20 dimensions—which is not at all our intention. To avoid this the use of complex numbers is controlled by 'reality conditions'. Reality conditions are formally similar to stabilising conditions, which reduce the domain of probability distribution to a locus of fewer dimensions than the mathematical representation space. Usually, but not necessarily, they have the simple form that certain specified components are real and other specified components are imaginary.

When we speak of 'reality conditions' the term refers to physical reality which, as we have just seen, is not usually congruent with mathematical reality of the representative symbols. The basis of our determination of reality conditions will be that physical reality is invariant for relativity rotations; since there can be no intrinsic equivalence between structures which are physically real and those which are not. This is, of course, subject to the proviso that the rotations are themselves physically real. Thus our first step must be to determine the reality conditions for rotations. Evidently the physically real rotations must form a closed group—a subgroup of the group of mathematical rotations.

Whatever the reality condition for rotations may be, it can be stated in the form that the physically real simple rotations are $q = e^{\frac{1}{2}E_{\mu\nu}\alpha_{\mu\nu}\theta}$, where θ is real, and $\alpha_{\mu\nu}$ is a certain array of complex numbers of modulus 1. The $\alpha_{\mu\nu}$ are independent of θ, since a large rotation can be divided into a succession of small rotations.

If θ is a small quantity whose cube is neglected, we obtain

$$e^{\frac{1}{2}E_{\mu\sigma}\alpha_{\mu\sigma}\theta}\, e^{\frac{1}{2}E_{\nu\sigma}\alpha_{\nu\sigma}\theta}\, e^{-\frac{1}{2}E_{\mu\sigma}\alpha_{\mu\sigma}\theta}\, e^{-\frac{1}{2}E_{\nu\sigma}\alpha_{\nu\sigma}\theta} = e^{\frac{1}{2}E_{\mu\nu}\alpha_{\mu\sigma}\alpha_{\nu\sigma}\theta^2}. \tag{61·1}$$

The left-hand side is a succession of real rotations; therefore $e^{\frac{1}{2}E_{\mu\nu}\alpha_{\mu\sigma}\alpha_{\nu\sigma}\theta^2}$ is a real rotation and, by the reality condition, $\alpha_{\mu\sigma}\alpha_{\nu\sigma} = \pm\alpha_{\mu\nu}$. By repeated application of this result,

$$\alpha_{\mu\tau}\alpha_{\nu\sigma}^2 = \pm\alpha_{\mu\sigma}\alpha_{\nu\sigma}\alpha_{\tau\sigma}\alpha_{\nu\sigma} = \pm\alpha_{\mu\nu}\alpha_{\tau\nu} = \pm\alpha_{\mu\tau}, \tag{61·2}$$

so that $\alpha_{\nu\sigma}^2 = \pm 1$ and $\alpha_{\nu\sigma} = \pm 1$ or $\pm i$. The sign has no significance, and the physically real rotations are accordingly determined by a set of characters $\alpha_{\nu\sigma} = 1$ or i associated with the $E_{\nu\sigma}$.

The real or imaginary character of symbols themselves may be similarly described by a set of characters $\beta_{\nu\sigma} = 1$ or i. Then $E_{\nu\sigma}\beta_{\nu\sigma}$ is always real; and by (60·6) the rotations which give full equivalence are $q = e^{\frac{1}{2}E_{\mu\nu}\beta_{\mu\nu}\theta}$. It is evident that the rotations which give full equivalence must be identified with the physically real rotations, so that $\beta_{\mu\nu} = \alpha_{\mu\nu}$. This is a matter of definition rather than of deduction, and does not admit of direct proof. Previously the symbolic frame has been anchored in a generalised universe in which 'measurements' are not restricted to real numbers; and we are now introducing a tighter anchorage in a physically real world. All that we can be expected to show is that the anchorage condition $\beta_{\mu\nu} = \alpha_{\mu\nu}$ is necessary to prevent the physics and the mathematics getting to cross-purposes. By rigorous deduction the choice has been narrowed so that in a given plane there are only two rotations with characters 1 and i to be considered. It is only reasonable to choose the mathematical representation so that full mathematical equivalence corresponds to real physical equivalence, rather than the direct opposite.

Accordingly (60·6) applies to the physically real rotations:

$$\textit{A rotation } q(\ldots)q^{-1} \textit{ is physically real if } q \textit{ is real.} \tag{61·3}$$

62. Distinction between space and time

An expression which is wholly real or wholly imaginary will be called *monothetic*. Two expressions will be called *homothetic* if both are real or both imaginary, and *antithetic* if one is real and the other imaginary.

When the angle of a rotation is imaginary we shall denote it by iu; so that by (61·3) the real relativity rotations are $q = e^{\frac{1}{2}E_\mu\theta}$ or $e^{\frac{1}{2}E_\mu iu}$ according as E_μ is real or imaginary. Consider two perpendicular components $p_{\mu\sigma}$, $p_{\nu\sigma}$. If $E_{\mu\sigma}$, $E_{\nu\sigma}$ are homothetic, their product $E_{\mu\nu}$ is real; and the rotation $e^{\frac{1}{2}E_{\mu\nu}\theta}$ gives a circular rotation of $p_{\mu\sigma}$ and $p_{\nu\sigma}$ as shown in (56·42). If $E_{\mu\sigma}$, $E_{\nu\sigma}$ are antithetic, $E_{\mu\nu}$ is imaginary, and the rotation $e^{\frac{1}{2}E_{\mu\nu}iu}$ gives $p_{\mu\sigma} = p'_{\mu\sigma}\cos(iu) + p'_{\nu\sigma}\sin(iu)$, $p_{\nu\sigma} = -p'_{\mu\sigma}\sin(iu) + p'_{\nu\sigma}\cos(iu)$, or

$$p_{\mu\sigma} = p'_{\mu\sigma}\cosh u + ip'_{\nu\sigma}\sinh u, \quad ip_{\nu\sigma} = p'_{\mu\sigma}\sinh u + ip'_{\nu\sigma}\cosh u, \tag{62·1}$$

which is a hyperbolic rotation between $p_{\mu\sigma}$ and $ip_{\nu\sigma}$ (or between $ip_{\mu\sigma}$ and $-p_{\nu\sigma}$).

Thus the real E_μ give circular relativity rotations, and the imaginary E_μ give hyperbolic relativity rotations.

We can now determine the reality conditions for the components p_μ of a vector by applying the condition that physical reality is invariant for relativity rotations.[a] Thus $p_{\mu\sigma}$ and $p'_{\mu\sigma}$ must satisfy the same reality condition. For the circular rotation (56·42) this requires that $p_{\mu\sigma}$ and $p_{\nu\sigma}$ shall be homothetic; for the hyperbolic rotation (62·1) it requires that $p_{\mu\sigma}$ and $p_{\nu\sigma}$ shall be antithetic. Both cases are covered by the condition that $E_{\mu\sigma}p_{\mu\sigma}$ and $E_{\nu\sigma}p_{\nu\sigma}$ are homothetic. This gives the general reality condition for a vector, namely, that all its symbolic terms $E_\mu p_\mu$ are homothetic with the possible exception of $E_{16}p_{16}$. It is here assumed that all the components are connected (directly or indirectly) by effective relativity rotations. The modification when some of the rotations are ineffective will be examined later.

Since a pentad consists of three imaginary and two real symbols, not more than three mutually perpendicular components can have homothetic symbols; so that there

[a] 'Relativity rotation' will henceforth mean physically real rotation. Rotations which are not physically real may be described as *formal* relativity rotations.

cannot be more than three mutually perpendicular directions connected by circular rotations. 'Space', defined as a domain in which the relativity rotations are circular, is accordingly restricted to three dimensions. If we extend this domain to four dimensions, the fourth axis must be connected with the other three by hyperbolic rotations (Lorentz transformations); in other words, it is a time axis.

The three-dimensionality of space, and the time-like character of the fourth dimension are thus deduced directly from the properties of the E-frame. To what extent this amounts to an *a priori* proof that the space-time of physical experience must be of this kind, depends on our inquiry into the ultimate origin of the E-frame in Chapter XIII. Here we would emphasise that, independently of epistemological considerations, the result is a substantial contribution to the unification of physics; for the E-symbols were originally introduced into physics as an instrument of quantum theory.

We accordingly associate three imaginary symbols E_{15}, E_{25}, E_{35} with the space axes and a real symbol E_{45} with the time axis. (The suffix 0 is avoided, because it has been reserved in (55·7) for distinctively electrical characteristics.) The symbols E_{23}, E_{31}, E_{12} giving the circular spatial rotations are real, and E_{14}, E_{24}, E_{34} giving the Lorentz transformations are imaginary, as the foregoing analysis requires. The notation agrees with (53·1), and justifies the use of E-symbols as coefficients in that formula.

The position 4-vector and momentum 4-vector are denoted by

$$\left.\begin{aligned} X &= E_{15}ix_1 + E_{25}ix_2 + E_{35}ix_3 + E_{45}t, \\ P &= E_{15}ip_1 + E_{25}ip_2 + E_{35}ip_3 + E_{45}\epsilon, \end{aligned}\right\} \tag{62·2}$$

where x_1, x_2, x_3, t and p_1, p_2, p_3, ϵ are real numbers. As here defined X and P are real. This is a convention, because the reality conditions only state that X and P are monothetic. The opposite convention, which takes X and P to be imaginary, is probably more familiar. But it is clearly more fitting that the real numbers which have absolute distinction of sign should be used to represent time components which have distinctive directions towards future and past; that the imaginary numbers whose sign is reversible should represent space components which have no one-way property.

Having marked out in the actual world a Lorentz frame of three rectangular axes and a corresponding time, we label the positive directions of these axes $E_{15}i$, $E_{25}i$, $E_{35}i$, E_{45}. These directions will be called a 'realisation' of the four symbols attached to them. The angular directions of relativity rotation of the Lorentz frame likewise supply a realisation of the symbols E_{23}, E_{31}, E_{12}, $E_{14}i$, $E_{24}i$, $E_{34}i$. We have then a physical interpretation of the ten components forming the mechanical part P_m of a vector P, namely, as magnitudes associated with linear or angular directions marked out in the actual world. There is no separate realisation of the symbol i; it occurs only as part of a realised symbol. Thus physical interpretation is only provided for real magnitudes associated with the realised symbols; e.g. $E_{15}ix$, $E_{23}\theta$ can only be interpreted if x and θ are real. It follows immediately that in order to have a physical interpretation P_m must be wholly real. In other words the reality condition for P_m is that P_m must be real; for reality conditions are conditions necessary to secure that an expression has a physical interpretation.

The electrical 4-vector must be monothetic, since its components are connected by effective relativity rotations; but we cannot immediately decide whether it is homothetic with P_m or antithetic. This very important question will be treated in the next

section. The complication is that in standard conditions electrical components are dormant. A dormant vector is not expected to have a physical interpretation; and we cannot apply to it the ordinary reality condition which is intended to secure that a vector shall have a physical interpretation!

63. Neutral space-time

The E-frame provides a fifth direction perpendicular to the axes x_1, x_2, x_3, t; and the position vector (62·2) can be extended to

$$X = E_{15}ix_1 + E_{25}ix_2 + E_{35}ix_3 + E_{45}t + E_{05}t_0, \tag{63·1}$$

where according to the reality conditions t_0 should be real. In molar relativity theory space-time, i.e. the domain in which the particles of the uranoid are located, is a four-dimensional locus in a Euclidean representation space of five dimensions. It is tempting to identify t_0 with the fifth coordinate in the representation space; but the interpretation turns out to be not so simple.

Take as origin a point P in the spherical space, and let W be the tangent flat space at P containing the axes x_1, x_2, x_3. Then if the suggested identification is correct, the centre of curvature is on the axis t_0; and, for all points in the neighbourhood of P, t_0 has the same sign, say positive. In particular, the position vector of any uranoid particle in the neighbourhood of P will have a positive component t_0.

Now construct a second uranoid according to the same specification in a chirally opposite frame; so that the two uranoids are similar but chirally opposite. The position vector of the corresponding particle in the second uranoid is obtained by reversing the sign of the term $E_{05}t_0$ in (63·1). Thus, when both uranoids are referred to the same frame E_μ, t_0 is changed to $-t_0$. Thus the chirally opposite uranoids are on opposite sides of the tangent plane W. But this is impossible because they are precisely the same uranoid; a uranoid, being electrically neutral, is its own chiral opposite. By definition it would be unaffected by interchanging positive and negative charge.

The neutral uranoid (or the space which it occupies) cannot be curved either in the positive or in the negative direction of the t_0 axis, because such curvature represents a condition biased as between positive and negative charges. If it has a property which can be legitimately represented as curvature in a direction perpendicular to the space-time axes, this direction must be represented symbolically by $E_{05}E_{16}$; for $E_{05}E_{16}$, being the product of two chiral symbols, is not altered by reversing the chirality, and it is perpendicular to (anticommutes with) E_{15}, etc. Denoting the coordinate in the $E_{05}E_{16}$ direction by x_0, the fifth term in the position vector is $E_{05}E_{16}x_0$. Or, adopting as usual a right-handed frame, the position vector is

$$X = E_{15}ix_1 + E_{25}ix_2 + E_{35}ix_3 + E_{45}t + E_{05}ix_0, \tag{63·2}$$

with x_1, x_2, x_3, t, x_0 real.

Two important consequences arise. Since E_{05}, like E_{45}, is real, its rotations with the space axes are hyperbolic, and the corresponding coordinate t_0 is time-like. But in (63·2) we have substituted ix_0 for t_0; so that x_0 is formally an imaginary time, or equivalently a real space-like coordinate. The radius of curvature R_0 of the neutral uranoid is accordingly space-like, as assumed in molar relativity theory and confirmed by the observational comparisons of the constants into which it enters. Thus the E-frame

predicts correctly the character of the curvature as well as the (3 + 1)-dimensional structure of space-time.

Secondly, since x_1 and x_0 are both space-like the rotation between them is circular, and is given by $q = e^{\frac{1}{2}E_{01}\theta}$ with θ real. But, since E_{01} is imaginary, this is a formal, not a real, relativity rotation; and the equivalence between lengths in and normal to space is formal, not physical. The fifth dimension used to represent curvature is a formal construct, and there is no physical extension of space in that direction. Thus a distinction obvious to common experience is correctly shown in the symbolic representation.

As a theoretical exercise we might consider a *monochiral uranoid* composed of elementary particles all of one sign. This would naturally be the simplest system to construct in a chiral E-frame; and the originally suggested position vector (63·1) applies to it. If it could be confined to a four-dimensional locus, it would have a time-like radius of curvature in the real E_{05} direction, which changes from positive to negative when the sign of the charges is reversed. But the fifth dimension has real relativity rotations with the other dimensions, so that the monochiral uranoid occupies a three-dimensional space and a two-dimensional time.

Anything like the conditions of a monochiral uranoid is utterly beyond experience, and we need not be surprised at the scarcely imaginable results, e.g. the two-dimensional time. Molar matter, even when said to be 'highly charged' is electrically neutral to an excessively high approximation as judged by the natural criterion, namely, the ratio of the number of protons to the number of electrons contained in it; a disparity of 1 in 10^{10} is outside reasonable possibility. Thus only faint vestiges of chirality penetrate into molar physics; and our ideas of space and time measurement are based on neutral (Riemannian) space-time. In the elementary particles of microscopic physics we meet for the first time chirality which is not excessively diluted.

It is easily calculated that an inequality of 1 in 10^{39} in the number of protons and electrons in the uranoid would give a potential everywhere of 550,000 volts. In a region where this potential actually occurs the conditions can be regarded as 10^{-39} of the way from a neutral to a monochiral uranoid; and the radius of curvature is tilted through an angle 10^{-39} out of the standard direction corresponding to potential zero. The tilt, being a pseudo-rotation (since there is no rotation $q(\ldots)q^{-1}$ between $E_{05}E_{16}$ and E_{05}), is interpreted as a strain; and the potential measures this particular strain in the environment. Our usual procedure is to treat an actual environment as a standard uranoid plus a disturbance; and the two components of the curvature then become separated. The electrical part P_e of the momentum vector has no physical interpretation in the standard uranoid; but it has a physical interpretation in the added chiral disturbance. When we speak of a characteristic as physically significant, we mean that it can be detected by its interaction with an appropriate test-body. For the isolated carrier of a momentum vector the test-body is the environment. The foregoing result agrees with our common experience that different test-bodies are required for P_m and P_e.

It is now clear that the 'reality conditions' for P_e are the conditions that it shall have physical significance in a chiral environment but not in an achiral environment, and the reality conditions for P_m are the conditions that it shall have physical significance in an achiral, but not in a chiral environment. Using the component $E_{05}x_{05}$ of the position vector as a test, we see by (63·1) and (63·2) that the conditions for physical

significance in the two environments are antithetic. Thus the reality condition for a vector P in the standard environment is

The electrical part P_e is antithetic to the mechanical part P_m. (63·3)

By our convention P_m is real and P_e imaginary. An alternative, and more formal derivation of this condition will be given in the next section.

The strain vector $S = -PE_{45}$ will be physically real if P is real. This also can be divided into mechanical and electrical parts $S_m = -P_m E_{45}$, $S_e = -P_e E_{45}$; and the reality condition requires S_m and S_e to be antithetic. This condition can be put in a simpler form. The components s_μ of S_m and S_e are shown separately in (59·3); and we see that the electrical part corresponds to the six real E_μ and the mechanical part to the ten imaginary E_μ. Thus the antithesis of S_m and S_e is provided for in the symbolic coefficients, and the components s_μ are all homothetic. The convention makes them imaginary:

The reality condition for a strain vector is that all the components s_μ are imaginary. (63·4)

The antithesis in (63·3) has a simple physical explanation. In ordinary units the dimensions of electric charge e are irrational; we cannot express e in terms of the standard M, L, T but we can express e^2. The result of any actual measurement must be a real number; but formally we may describe by a symbol the 'ratio' of two quantities which are not experimentally comparable. Thus the ratio of an electric charge to our mechanical standard may be expressed by a symbol, but not by a real number. But the square of the symbol must be a real number since e^2 is experimentally comparable with the standard. This is satisfied by taking the charge to be an imaginary number in terms of the mechanical standard. We have therefore two standards whose ratio is represented by the symbol i or E_{16}; and the antithesis of P_m and P_e means that the one part in order to be physically real must be experimentally comparable with the standard 1 and the other part with the standard i. The dormancy of P_e is due to the fact that the standard environment defines only the former standard.

64. Congruent spaces

According to the previous section, reality conditions depend on the environment; so that, for example, we postulate a standard neutral uranoid when we formulate the conditions (63·3) and (63·4). This is reasonable, because a physical structure which is possible in one environment may be impossible in another. Reality conditions are, in fact, disguised boundary conditions, which have a simple form because we consider only simple environments. The disguise is that they ostensibly relate the physically possible vectors to a symbolic or geometrical frame instead of to the environment; but, by pre-arrangement, the frame forms a compendium of the properties of the environment, so that it comes to the same thing. The treatment therefore depends on a congruence of the properties (symmetry and chirality) of the frame and environment, which makes the frame, in effect, a symbolisation of the environment and the environment a realisation of the frame.

The complication is that the uranoid which completely realises an E-frame is monochiral—a distribution extravagantly remote from experience. The neutral uranoid

realises only the mechanical part of the E-frame. In the actual universe the most we can do is to introduce an extremely dilute chirality into the environment; but this is treated (as usual) as a superposition on the standard neutral environment—an electromagnetic object-field—and does not modify the relation of the frame and standard environment. The chiral part of an electron inserted in the neutral space-time determined by an equal mixture of electrons and protons is not related to the frame in the same way as the chiral part of an electron inserted in a universe composed wholly of electrons. It is easy to determine the latter relation, but it is the former that is required in physics.

This is elucidated by reference to § 57. We begin with a 5-space (3-space + 2-time) with rectangular axes which are realisations of E_{01}, E_{02}, E_{03}, E_{04}, E_{05}.[a] This corresponds to a monochiral uranoid. The axes give realisation of the chiral symbols and the angular directions in the coordinate planes give realisations of the mechanical symbols. We have next to reduce this to a neutral space-time. Conceptually we shall regard this as the same space, except for the dropping of a dimension; so that the geometrical axes remain fixed, but are now realisations of E_{15}, E_{25}, E_{35}, E_{45}. More precisely, we regard the new space as a superposition of two opposite chiral spaces; for it would be self-contradictory to regard a neutral space and a chiral space as the same. When the two opposite 5-spaces are superposed, the realisations of their E_{04} axes have opposite directions; and we cannot identify the realisation of E_{45} in the new space with both. We must identify it with the realisation of $E_{04}E_{16}$, which does not change sign on reversing chirality. Similarly the realisation of $E_{15}i$ in neutral space must be identified with the realisation of $E_{01}E_{16}i$ in the chiral spaces. There is no other way, because we have not provided any separate realisation of E_{16} or i. The result is that the axes in the 4-space and in the 5-space, which are conceived as geometrically congruent, have antithetic symbols. In the right-handed 5-space the axes are realisations of E_{01}, E_{02}, E_{04}, $E_{04}i$, $E_{05}i$, the first four being congruent with the realisations of $E_{15}i$, $E_{25}i$, $E_{35}i$, E_{45} in a Lorentz frame.

Primarily this implies that in passing from four to five dimensions we have to uproot and re-anchor the symbolic frame with respect to the geometrical axes; but we have found a simpler way of making the change by identifying the vectors $P^{(5)}$ in five dimensions with imaginary vector-densities $i\mathfrak{P}^{(4)}$ in four dimensions. Thus by considering vectors and vector-densities together we combine the two representations in one. This was shown in (57·6). The new point is that the position 5-vectors (and correspondingly the momentum 5-vectors) are imaginary, in contrast to the position and momentum 4-vectors in space-time which are real. These 5-vectors are introduced into space-time as vector-densities $i\mathfrak{P}$; since P is homothetic with $i\mathfrak{P}$, the corresponding vectors P are imaginary. Thus the vectors introduced into space-time from five dimensions are antithetic to the native vectors.

Returning to the 5-space, the rotation between axes $E_{01}, E_{02}, \ldots$ is the same as between $E_{01}E_{16}, E_{02}E_{16}, \ldots$; so that the foregoing anomaly affects only the 5-vector and not the other ten components. The 10-vector in 5-space is real, and remains real when introduced into space-time.

There is no confusion between native and imported vectors. The ten mechanical components are real, and are represented (in different ways) either in 5-space or in

[a] Subject to the insertion of appropriate i factors, to be determined later.

neutral 4-space. There are no native electrical components, since such components would be meaningless in a neutral environment. Thus, when the mechanical momentum vector (complete momentum vector) is extended to include electrical components, these have to be imported; and it is the imported addition that is antithetic to the rest.

The essential step in the foregoing deduction is the setting up of a geometrical congruence of the axes $E_{15}i$ and $E_{01}E_{16}i$. This gets over the difficulty of representing neutral and monochiral environments in the same geometrical space—a representation which is required when we treat an electric field as a superposition on a neutral uranoid. The monochiral addition is a misfit in the neutral space-time, which it has had no share in determining. But we have made congruent with the neutral space a space into which it can be fitted by the recognised device of using imaginary as well as real axes.

65. Determinants and eigenvalues

We summarise for future reference some further definitions and results relating to E-numbers. They are mostly simple generalisations of definitions and results relating to matrices, which are a special case of E-numbers (§ 71). Proofs are given in *Protons and Electrons*, Chapters II, III.

The *determinant* of an E-number is a certain homogeneous quartic function of its components, namely

$$\det P = \Sigma p_\mu^4 \pm 2\Sigma p_\mu^2 p_\nu^2 + 8\Sigma p_{\mu\sigma}p_{\mu\tau}p_{\nu\sigma}p_{\nu\tau} \pm 8\Sigma p_{\mu\nu}p_{\sigma\tau}p_{\lambda\rho}p_{16}. \tag{65·1}$$

In the second term the sign is positive if $E_\mu E_\nu$ anticommute and negative if they commute; in the fourth term the sign is positive or negative according as $\mu, \nu, \sigma, \tau, \lambda, \rho$ form an even or odd permutation of 0, 1, 2, 3, 4, 5. It is understood that terms which have two or more equivalent forms (e.g. $p_{12}p_{13}p_{42}p_{43}$ and $p_{21}p_{24}p_{31}p_{34}$) are included once only.

The determinant has the property

$$\det(PQ) = \det P \det Q. \tag{65·2}$$

Also the condition that P is singular is

$$\det P = 0. \tag{65·3}$$

It follows from (65·2) that a product of E-numbers is singular if any one of them is singular.

By (65·1), $\det 1 = \det E_{16}(-i) = 1$. Hence $\det q^{-1} \det q = 1$; so that

$$\det P' = \det(qPq^{-1}) = \det q \det P \det q^{-1} = \det P. \tag{65·41}$$

Thus the determinant is invariant for rotations. Also by (65·1),

$$\det E_\mu = 1, \tag{65·42}$$

$$\det(e^{\frac{1}{2}E_\mu\theta}) = 1, \quad \text{if} \quad \mu \neq 16. \tag{65·43}$$

Thus the determinant is also invariant for pseudo-rotations, except the pseudo-rotation $q = e^{\frac{1}{2}E_{16}\theta}$.

If λ is an algebraic number, $\det(P-\lambda)$ is formed by substituting $p_{16}+i\lambda$ for p_{16} in (65·1). Thus the equation

$$\det(P-\lambda) = 0 \tag{65·51}$$

is a quartic equation in λ, which may be written as

$$f(\lambda) \equiv (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)(\lambda-\lambda_4) = 0. \tag{65·52}$$

This is called the *characteristic equation* of P. It is an ordinary algebraic equation, but there is a remarkable theorem (called the Hamilton-Cayley theorem in matrix theory) that P satisfies its own characteristic equation; so that we have also a symbolic equation

$$f(P) \equiv (P-\lambda_1)(P-\lambda_2)(P-\lambda_3)(P-\lambda_4) = 0. \tag{65·53}$$

If X is a symbol of any kind, and ϕ is another symbol such that

$$X\phi = \alpha\phi \quad (\phi \neq 0), \tag{65·6}$$

where α is an algebraic number, ϕ is called an *eigensymbol* of X, and α the corresponding *eigenvalue* of X. It is easily shown that if X satisfies a polynomial equation $f(X) = 0$, the only possible eigenvalues of X are the roots of the equation $f(\alpha) = 0$. It follows from (65·53) that the possible eigenvalues of P are λ_1, λ_2, λ_3, λ_4. Since $E_\mu^2 = -1$, the eigenvalues of E_μ are $\pm i$.

A symbol which has an eigenvalue 0 is singular. For if, in (65·6), X has a reciprocal R, we have

$$\phi = RX\phi = R\alpha\phi,$$

which is impossible if $\alpha = 0$. It follows that, if α is any eigenvalue of X, $X-\alpha$ is singular. Applying this to an E-number, if λ is an eigenvalue of P, $P-\lambda$ is singular so that $\det(P-\lambda) = 0$; and the possible eigenvalues are the roots of the characteristic equation. This also follows directly from the fact that P itself satisfies the characteristic equation.

In elementary theory we separate the *magnitude* and *direction* of a vector X, the magnitude being specified by a linear invariant l called the length, and the direction by a normalised or unit vector X/l. It is often useful to make a similar separation in an extended vector P. We have a choice of two linear invariants of P, namely $\mathrm{qs}\, P$ and $(\det P)^{\frac{1}{4}}$, either of which can be regarded as the 'magnitude'. Correspondingly the 'direction' of P is represented by

$$(a) \quad P/\mathrm{qs}\, P, \qquad (b) \quad P/(\det P)^{\frac{1}{4}}. \tag{65·7}$$

These are alternative, but not equivalent, definitions of a normalised or unit vector. Generally the normalisation (b) is preferable because it is invariant for pseudo-rotations as well as rotations, and it also preserves the correspondence of vectors, vector-densities and strain vectors. But it is inapplicable if P is singular. Some of the most fundamental vectors in our theory (idempotent vectors) are singular; and for these we employ normalisation (a).

Chapter VII

WAVE VECTORS

66. Idempotency

A quantity X is *idempotent* if

$$X^2 = X. \tag{66·1}$$

Then all positive integral powers of X are equal to X. Negative powers are non-existent; for (66·1) shows that the eigenvalues are 1 and 0, so that X is singular.

Consider a probability distribution of a variate X over a set of values X_r ($r = 1, 2, 3, \ldots$). Let f_r be the probability of a value X_r, the total probability Σf_r being 1. The mean or expectation value of X is $\overline{X} = \Sigma f_r X_r$. In addition we are likely to be concerned with the expectation value of X^2, namely $\overline{X^2} = \Sigma f_r X_r^2$, and possibly other functional means such as $\overline{e^{iX}} = \Sigma f_r e^{iX_r}$.

The means $\overline{X}$, $\overline{X^2}$, $\overline{X^3}$, ... are independent moments of the probability distribution, and are not related like the powers $X, X^2, X^3, \ldots$ of individual quantities. If $\overline{X^2} = \alpha(\overline{X})^2$, α is called an *averaging factor*. If, as is common in rough approximations, we neglect averaging factors the procedure is equivalent to treating means as though they were individuals. A less obvious error is to incorporate the averaging factors which correspond to a commonly occurring standard probability distribution in the empirical constants of the equations; so that what are really means are treated as pseudo-individuals.

There is immense simplification if X is idempotent, so that $X = X^2 = X^3$ and $\overline{X} = \overline{X^2} = \overline{X^3}$. The treatment of means as pseudo-individuals then becomes fully justified.

It will not be surprising if in our gropings into the structure of things a legend of individuality has attached itself to the carrier of an idempotent variate. In statistically grounded theory it is the closest counterpart of the obsolete classical particle. We now know that matter cannot be analysed into elements having the individual distinctness that classical particles were supposed to have; but in the carriers of idempotent variates we reach elements which, though not less statistical than other carriers, do not betray their statistical character in the ordinary calculations of dynamics.

The association of pseudo-individuality with idempotency has profound importance in the logical structure of physical science. We observe only molar phenomena, and predict only molar phenomena. (It need scarcely be said that the water-drop tracks in a Wilson chamber are molar phenomena.) Underlying the molar world we are accustomed to picture a microscopic world populated by individuals which yield the averages concerned in molar phenomena; and it is further supposed that protons and electrons are such individuals. Even those who more cautiously or more speculatively —the two terms are often synonymous—regard protons and electrons as not necessarily the ultimate elements of structure, would probably concede that there is some limit at which statistical analysis can go no further because we are no longer dealing with means but with individuals. But it now appears that the only kind of individuality known in physics is the pseudo-individuality conferred by idempotency. There is no

objective limit to the fine-grainedness of statistical analysis. We begin with averages and there is no reason to suppose that our procedure of analysis ever leads to anything but averages. The accepted stopping-point is decided by the consideration that the procedure of analysis has achieved its aim when it has removed the indeterminacy of the relations between linear and non-linear characteristics. The runner stops when the race is won, not because further motion is impossible; it is for the same reason that statistical analysis ends in carriers of idempotent vectors or tensors.

When we realise that the partition of characteristics of molar objects among a large number of carriers is an analysis directed towards a mathematically defined, not an objectively defined, goal, the calculation of the total number of these carriers by purely theoretical reasoning will not seem too audacious.

Fundamental importance must accordingly be attached to momentum vectors and momentum strain vectors which are idempotent; and we shall investigate the conditions of idempotency. Idempotency is invariant for rotations; for if $P^2 = P$, and $P' = qPq^{-1}$, we have $P'^2 = qPq^{-1}qPq^{-1} = qP^2q^{-1} = qPq^{-1} = P'$.

If P is idempotent, $1-P$ is also idempotent. We call $1-P$ the *image* of P.

We first notice certain trivial solutions of $P^2 = P$, namely,

(*a*) $P = 0, P = 1$,

(*b*) $P = \frac{1}{2}(1+iE_{\mu\nu})$, and vectors obtained by applying rotations to it.

These are distinguished invariantly by the fact that their quarterspurs are 0, 1, or $\frac{1}{2}$. We define *a non-trivial idempotent vector* to be one whose quarterspur has not these values. We shall find that the quarterspur is either $\frac{1}{4}$ or $\frac{3}{4}$.

67. Standard form of idempotent vectors

It is easily verified that if $\mu, \nu, \sigma, \tau, \lambda, \rho$ is an even permutation of 0, 1, 2, 3, 4, 5, the vector

$$P = -\tfrac{1}{4}E_{16}(E_{\mu\nu}+E_{\sigma\tau}+E_{\lambda\rho}+E_{16}) \tag{67·1}$$

is idempotent. We shall prove that *every non-trivial idempotent vector can be reduced to the form* (67·1), *or its image, by a rotation.*

Taking $P = \Sigma E_\mu p_\mu$, multiplying out P^2, and equating coefficients of E_μ on both sides of $P^2 = P$, we obtain 15 conditions of the form

$$2i(p_{01}p_{16}+p_{23}p_{45}+p_{24}p_{53}+p_{25}p_{34}) = p_{01}, \tag{67·21}$$

together with

$$-\sum_1^{16} p_\mu^2 = ip_{16}. \tag{67·22}$$

The frame is here taken to be right-handed. Multiplying (67·21) by p_{01}, we obtain

$$p_{01}^2(1-2ip_{16}) = 2i(p_{01}p_{23}p_{45}+p_{01}p_{24}p_{53}+p_{01}p_{25}p_{34}). \tag{67·23}$$

There are 15 triple products of the form $p_{\mu\nu}p_{\sigma\tau}p_{\lambda\rho}$; so that each occurs three times in the set of 15 equations typified by (67·23). We see from (67·23) that the order of the suffixes is in each case an even permutation; so that the products always occur with the same sign. The sum of the 15 products will be denoted by

$$v = \Sigma p_{\mu\nu}p_{\sigma\tau}p_{\lambda\rho}. \tag{67·24}$$

By summing the 15 equations,

$$6iv = (1-2ip_{16})\sum_1^{15} p_\mu^2 = -(1-2ip_{16})(p_{16}^2+ip_{16}) \tag{67·25}$$

by (67·22).

A *pentadic part of* P is a set of five terms whose symbols form a pentad. There are six pentadic parts of P, denoted by ϖ_α ($\alpha = 0, 1, 2, 3, 4, 5$), namely,

$$\varpi_\alpha = E_{\alpha 0}p_{\alpha 0}+E_{\alpha 1}p_{\alpha 1}+E_{\alpha 2}p_{\alpha 2}+E_{\alpha 3}p_{\alpha 3}+E_{\alpha 4}p_{\alpha 4}+E_{\alpha 5}p_{\alpha 5}. \tag{67·31}$$

One of the terms is meaningless, and is to be ignored. In $\Sigma_\alpha \varpi_\alpha$ each term of P except the quarterspur occurs twice; so that

$$\tfrac{1}{2}\Sigma_\alpha \varpi_\alpha = P - ip_{16}. \tag{67·32}$$

The square of ϖ_α is algebraic; for example,

$$-\varpi_0^2 = p_{01}^2+p_{02}^2+p_{03}^2+p_{04}^2+p_{05}^2. \tag{67·33}$$

Also

$$-\tfrac{1}{2}\Sigma_\alpha \varpi_\alpha^2 = \sum_1^{16} p_\mu^2 - p_{16}^2. \tag{67·34}$$

For an idempotent P this reduces by (67·22) to

$$\tfrac{1}{2}\Sigma_\alpha \varpi_\alpha^2 = p_{16}^2+ip_{16}. \tag{67·35}$$

By summing the five equations of the form (67·23) which correspond to the terms on the right-hand side of (67·33), we obtain

$$-\varpi_0^2(1-2ip_{16}) = 2iv, \tag{67·41}$$

and by symmetry this applies to any of the pentads. Since trivial idempotency is excluded, the quarterspur ip_{16} is not equal to 0, 1, or $\frac{1}{2}$; so that, by (67·25), v is not zero. Thus the squares of all the pentads have the same non-zero value

$$\varpi_\alpha^2 = -\frac{2iv}{1-2ip_{16}} = \tfrac{1}{3}(ip_{16}+p_{16}^2) \tag{67·42}$$

by (67·25).

Having established that the pentads are not null vectors, we can choose axes in five dimensions so that ϖ_0 reduces to a single component $E_{01}p_{01}$. Then

$$p_{02}, p_{03}, p_{04}, p_{05} = 0. \tag{67·51}$$

From the conditions (67·21) we then obtain immediately[a]

$$p_{12}, p_{13}, p_{14}, p_{15} = 0. \tag{67·52}$$

This reduces ϖ_2 to $E_{23}p_{23}+E_{24}p_{24}+E_{35}p_{35}$; and, since ϖ_2 is not a null vector, it can be reduced to a single component $E_{23}p_{23}$ by rotations E_{34}, E_{35}. These rotations leave p_{01} unaltered, and do not disturb the results (67·51) and (67·52). We have now

$$p_{24}, p_{25} = 0. \tag{67·53}$$

Since the only surviving $p_{0\mu}$ is p_{01}, and the only surviving $p_{2\mu}$ is p_{23}, the only surviving triple product is $p_{01}p_{23}p_{45}$. This occurs in the equations (67·23) for p_{01}^2, p_{23}^2, p_{45}^2; and

[a] For example, $p_{12}(1-2ip_{16}) = 2i(p_{03}p_{45}+p_{04}p_{53}+p_{05}p_{34})$, and $ip_{16} \neq \frac{1}{2}$.

these accordingly are the only surviving components besides the quarterspur. Thus (67·42) becomes

$$-p_{01}^2 = -p_{23}^2 = -p_{45}^2 = -\frac{2ip_{01}p_{23}p_{45}}{1-2ip_{16}} = \tfrac{1}{3}(ip_{16}+p_{16}^2). \tag{67·61}$$

Hence $2ip_{01} = \pm(1-2ip_{16})$, and

$$\tfrac{1}{4}(1-2ip_{16})^2 = \tfrac{1}{3}(ip_{16}+p_{16}^2). \tag{67·62}$$

The roots are $ip_{16} = \frac{1}{4}, \frac{3}{4}$. It is sufficient to consider the root $\frac{1}{4}$, since $\frac{3}{4}$ will then correspond to the image vector. The value $ip_{15} = \frac{1}{4}$ gives $p_{01} = \pm\frac{1}{4}i$; and P is reduced to

$$P = -\tfrac{1}{4}i(\pm E_{01} \pm E_{23} \pm E_{45} + E_{16}). \tag{67·7}$$

In order to satisfy (67·61) the number of negative signs in the bracket must be 0 or 2. Subject to this rule of signs, the idempotency of (67·7) is easily verified by direct multiplication.

The reversal of two signs is provided for in the enunciation of (67·1), since

$$E_{\mu\nu} - E_{\sigma\tau} - E_{\lambda\rho} = E_{\mu\nu} + E_{\tau\sigma} + E_{\rho\lambda},$$

and the order of the suffixes is still an even permutation. D. E. Littlewood has pointed out that when $\mu, \nu, \sigma, \tau, \lambda, \rho$ is an *odd* permutation the square of

$$\tfrac{1}{2}(E_{\mu\nu} + E_{\sigma\tau} + E_{\lambda\rho} + E_{16})$$

is -1. The existence of these anti-tetradic roots shows that it would be erroneous to jump to the conclusion that all symbolic square roots of -1 are equivalent. One would expect the anti-tetradic roots to have an important physical application; but thus far I have not found any.

68. Spectral sets

A momentum vector P which satisfies the reality condition cannot be strictly idempotent; for the reality condition makes the quarterspur (an electrical component) imaginary, and the idempotent condition makes it real ($\frac{1}{4}$ or $\frac{3}{4}$). We consider instead a vector with 'imaginary idempotency' which satisfies $P^2 = iP$, or $(-iP)^2 = -iP$.

The conclusion that the strictly idempotent vector must be $-iP$, not P, might have been avoided by adopting the opposite reality convention in § 60; and, in view of the great importance attached to idempotency, this might seem a strong reason for preferring the opposite convention. We shall find, however, that the i in $P^2 = iP$ is not the sign of a badly chosen convention, but a premonition of the transformation from classical to quantum momentum vectors. It will be remembered that the real momenta in quantum theory are i times the classical momenta. We have not yet considered how this result, applying to the momentum 3-vector, is to be generalised; but it already suggests that the idempotency which bestows pseudo-individuality on quantum particles is an idempotency of $\pm i$ times the classical momentum vector P.

The vector
$$P = \tfrac{1}{4}(E_{01} + E_{23} + E_{45} + E_{16}) \tag{68·1}$$

satisfies the reality conditions and $P^2 = iP$. The suffixes 1, 2, 3 can, of course, be interchanged cyclically; but otherwise the distribution of the suffixes is unique. Any other

combination, e.g. $\frac{1}{4}(E_{05}+E_{14}+E_{23}+E_{16})$ would violate the reality conditions. This enables us to infer quite definitely the constitution of a particle carrying a momentum vector which satisfies $P^2 = iP$.

In the reduction to the standard form (68·1) the axes have been chosen so that the particle is at rest; this is shown by the absence of momentum components E_{15}, E_{25}, E_{35}. There are two mechanical components of P, namely an energy or mass $E_{45}\frac{1}{4}$ and an angular momentum $E_{23}\frac{1}{4}$. The component $E_{16}\frac{1}{4}$, which is an invariant having an electrical character, evidently represents a charge. The remaining component $E_{01}\frac{1}{4}$ is an electromagnetic characteristic directed along the x_1-axis. A charged particle spinning in the E_{23} plane will have a magnetic moment directed along the x_1-axis; accordingly $E_{01}\frac{1}{4}$ is interpreted as the magnetic moment.

The corresponding strain vector is

$$S = -PE_{45} = -\tfrac{1}{4}i(E_{23}+E_{01}+E_{16}+E_{45}). \tag{68·2}$$

This is strictly idempotent by (67·7). The components are three-dimensional densities of mass $(-\frac{1}{4}iE_{16})$, charge $(-\frac{1}{4}iE_{45})$, spin momentum $(-\frac{1}{4}iE_{01})$ and magnetic moment $(-\frac{1}{4}iE_{23})$. In the strictly idempotent expression we can reverse the signs of any two components other than E_{16}; we thus obtain four associated strain vectors S_a, S_b, S_c, S_d given below. We also give the corresponding vectors P_a, P_b, P_c, P_d; the physical meaning of the association is more easily recognised in these.

$$\left.\begin{aligned}
S_a &= -\tfrac{1}{4}i(E_{23}+E_{01}+E_{16}+E_{45}) && \text{charge }+,\ \text{spin }+,\\
S_b &= -\tfrac{1}{4}i(-E_{23}-E_{01}+E_{16}+E_{45}) && \text{charge }+,\ \text{spin }-,\\
S_c &= -\tfrac{1}{4}i(-E_{23}+E_{01}+E_{16}-E_{45}) && \text{charge }-,\ \text{spin }+,\\
S_d &= -\tfrac{1}{4}i(E_{23}-E_{01}+E_{16}-E_{45}) && \text{charge }-,\ \text{spin }-;
\end{aligned}\right\} \tag{68·3}$$

$$\left.\begin{aligned}
P_a &= \tfrac{1}{4}(E_{01}+E_{23}+E_{45}+E_{16}) && \text{charge }+,\ \text{spin }+,\\
P_b &= \tfrac{1}{4}(-E_{01}-E_{23}+E_{45}+E_{16}) && \text{charge }+,\ \text{spin }-,\\
P_c &= \tfrac{1}{4}(-E_{01}+E_{23}+E_{45}-E_{16}) && \text{charge }-,\ \text{spin }+,\\
P_d &= \tfrac{1}{4}(E_{01}-E_{23}+E_{45}-E_{16}) && \text{charge }-,\ \text{spin }-.
\end{aligned}\right\} \tag{68·4}$$

The condition that the number of negative signs is 0 or 2 causes automatic reversal of the magnetic moment if either the spin or the charge is reversed, in agreement with elementary theory. For all four combinations the mass component of P is $\frac{1}{4}E_{45}$, indicating a positive mass $\frac{1}{4}$. Thus the condition that the strain vector is idempotent distinguishes a direction along the E_{45}-axis as the positive direction or, as it is usually called, the *future*. This is not affected if the frame is left-handed; qs S must still be $+\frac{1}{4}$, and the corresponding term in $P = SE_{45}$, is $+\frac{1}{4}E_{45}$. To put it another way: if we make it a condition of individuality that the strain vector shall be idempotent, individual particles will have momentum vectors which all point the same way along the time-axis, whether their charges are positive or negative.

By direct multiplication we find that the product of any two of the four S's or of the four P's is zero. Thus the strain vectors in (68·3) satisfy

$$S_a^2 = S_a, \quad S_aS_b = 0, \quad S_a+S_b+S_c+S_d = 1. \tag{68·5}$$

These are the equations which define a *spectral set*.[a] They show that the S's behave as selection operators which separate an operand into constituents which are pure, non-

[a] G. Temple, *Proc. Roy. Soc.* A, **338**, 479, 1932.

overlapping and exhaustive. Since $S_a S_a = S_a$ it makes no difference how often the operation of selection is repeated; this is characteristic of pure selection as contrasted with selective operations such as 'fractionation'. Since $S_a S_b = 0$, that which has been selected by the operation S_b yields nothing to the selection made by S_a; that is to say, the constituents selected by S_a and S_b are mutually exclusive. The third equation shows that the separated constituents, when added together, make up the original mixture, so that nothing has been left out. Spectral analysis of light is an operation of this kind (with an infinite number of operators), yielding constituents which are pure, mutually exclusive and together make up the light that is analysed.

When the S-symbols stand alone, the operand on which the spectral selection is performed is the strain vector 1 or $E_{16}(-i)$. This consists of one component $S_{16} = -i$ representing energy density or mass density; so that the corresponding matter is neutral, spinless and at rest, as in the standard uranoid. Spectral analysis exhibits this as a mixture of four kinds of pure particles in equal proportions. The balanced combination of the four pure particles will be called a *scalar particle*. In the coordinate system here employed its momentum vector consists of an energy component only, $p_{45} = 1$; but in other coordinate systems it becomes a general 4-vector, $p_{15}, p_{25}, p_{35}, p_{45}$. Thus the scalar particle is a V_3 or V_4 particle according as the proper mass is treated as exact or almost exact.

Idempotent E-numbers with quarterspurs $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ are said to be of *rank* 1, 2, 3.[a] Distinguishing idempotent strain vectors of ranks 1 and 2 by S_1 and S_2, we have the following connection. By adding any two of the S_1's in (68·3) we obtain six S_2's, namely,

$$-\tfrac{1}{2}i(E_{16} \pm E_{23}), \quad -\tfrac{1}{2}i(E_{16} \pm E_{01}), \quad -\tfrac{1}{2}i(E_{16} \pm E_{45}). \tag{68·6}$$

By multiplying any two of these S_2's (excluding pairs which differ only in sign and give a product 0) we obtain the original S_1's; e.g.

$$\{-\tfrac{1}{2}i(E_{16}+E_{23})\}\{-\tfrac{1}{2}i(E_{16}+E_{45})\} = -\tfrac{1}{4}i(E_{16}+E_{23}+E_{45}+E_{01}). \tag{68·7}$$

Thus *an S_2 is the sum of two S_1's, and an S_1 is the product of two S_2's*. In additive theory S_1 appears to be the more elementary form; but in multiplicative theory S_2 is the more elementary.

The image $S_3 = 1 - S_1$ of an idempotent E-number of rank 1 is idempotent of rank 3. Assuming additive representation, S_3 represents either two electrons and a proton or two protons and an electron. The latter combination occurs naturally as a deuteron. An electron S_c and a deuteron $S'_c = 1 - S_c$ satisfy the conditions for a spectral set

$$S_c^2 = S_c, \quad S_c'^2 = S'_c, \quad S'_c S_c = 0, \quad S'_c + S_c = 1. \tag{68·8}$$

Thus there is a rational foundation for an analysis of matter into deuterons and electrons. In particular, the deuteron may have the pseudo-individuality conferred by idempotency. Things do not work out quite like that, because the primary mode of combination in quantum theory is multiplicative, not additive. But the deuteron (as the most elementary particle which has zero isotopic spin) is certainly the key-particle in nuclear theory; and it is interesting to note that the symbolic theory has already begun to hint at its importance.

[a] This agrees with the definition of rank in matrix theory.

For the present we shall be concerned only with idempotency of rank 1; and the corresponding particles will be called *pure particles*. The pure particle at rest is defined either by $S^2 = S$ or $P^2 = iP$. The latter condition is (as already pointed out) invariant for rotation; but strain vectors are not transformed by rotation in the same way as vectors, and the condition $S^2 = S$ is not invariant. Thus the general definition is

A pure particle is the carrier of a momentum vector which satisfies $P^2 = iP$, *and* $\mathrm{qs}P = \frac{1}{4}$. (68·9)

The pure particles are the 'germs' of protons and electrons. The complete properties of protons and electrons cannot appear until an environment has been provided in which these properties can manifest themselves; and at present we are studying them in a very abstract way in the environment represented by a simple E-frame. But it is natural to call these embryonic particles by the names that they will earn later, when properties arising out of multiplicity factors, interchange, etc., are added.

The introduction of idempotent vectors is important in restoring the unity of the E-frame, which seemed about to break down through the ineffectiveness of so many of the relativity rotations. The grandiose scheme of relativistic equivalence with which we started is whittled down by the need for a neutral, not a monochiral, uranoid. It might begin to be doubted whether there is any object in stringing together the odd collection of vectors and invariants listed in (56·7) and usually treated as altogether distinct vectors. This criticism would, I think, be valid if field quantities alone were concerned. But the fixed-scale particle theory which we are now developing introduces another kind of linkage not concerned with rotation—a cross-pinning of antiperpendicular components. Relativity rotation gives a comparison of space and time measures, and shows the fundamental equivalence which exists between 300,000 km. and 1 sec.; but it can do little more. The pure particle is a four-armed comparator determining fundamental equivalences between a wider variety of physical characteristics—mass, angular momentum, charge, magnetic moment.

69. Catalogue of symbolic coefficients

In terms of mathematically real variates p_μ or $\varpi_\mu = -ip_\mu$, the electrical part of a momentum vector is

$$P_e = E_{01}p_{01} + E_{02}p_{02} + E_{03}p_{03} + E_{04}i\varpi_{04} + E_{05}i\varpi_{05} + E_{16}p_{16}. \tag{69·1}$$

In the last section p_{01} has been identified with a magnetic moment. Thus the 3-vector p_{01}, p_{02}, p_{03} is the ordinary magnetic moment, and ϖ_{04} is the time component introduced when the magnet has a velocity in the reference frame. Since these quantities form part of a momentum vector, it is more natural in the present context to call the magnetic moment 'magnetic momentum' and the component ϖ_{04} 'magnetic energy'. The two 4-vectors in P are then distinguished as the mechanical and magnetic momentum vectors

$$P^{(4)} = E_{15}ip_1 + E_{25}ip_2 + E_{35}ip_3 + E_{45}\epsilon, \quad M^{(4)} = E_{01}\mu_1 + E_{02}\mu_2 + E_{03}\mu_3 + E_{04}i\mu_4. \tag{69·2}$$

The components ϖ_{05}, p_{16} are invariants in space-time. The latter has been identified with electric charge; so it suggests itself that ϖ_{05} is the magnetic charge or pole-strength. This is verified because, multiplying a magnetic pole-strength $E_{05}i\mu$ by a displacement

$E_{15}i\delta x_1$, we obtain $\pm E_{01}(\mu\delta x_1)$ which has the correct symbolic coefficient for a magnetic moment $\mu_1 = \mu\delta x_1$.

We can now find without difficulty the symbolic coefficients of the principal physical quantities in much the same way as their dimensions or tensor characters are determined, using the equations which define their interconnections. In the following table we give the coefficient to be attached to the mathematically real measure. No attention is paid to sign, since this generally depends on casual conventions.

This table completes our task of anchoring the symbolic frame in molar physics. *It refers wholly to molar (classical) quantities.* We have remarked that quantities in quantum theory are often analogues, not direct representatives of the correspondingly named molar quantities. The anchored frame provides a symbolic nomenclature which will enable us to define precisely the new quantities introduced in extending our analysis to microscopic physics.

Symbolic coefficients

Electric charge	E_{16}
Magnetic pole-strength	$E_{05}i$
Energy, mass, time displacement	E_{45}
Momentum, spatial displacement	$E_{15}i$, $E_{25}i$, $E_{35}i$
Angular momentum, angular displacement	E_{23}, E_{31}, E_{12}
Magnetic moment (um)	E_{01}, E_{02}, E_{03}.
Magnetic energy	$E_{04}i$
Electric moment	E_{15}, E_{25}, E_{35}
Velocity, mechanical force	$E_{14}i$, $E_{24}i$, $E_{34}i$
Angular velocity	$E_{01}i$, $E_{02}i$, $E_{03}i$
Volume, particle density	E_{45}
Energy density, mass density	$E_{16}i$
Charge density, electrostatic potential	$E_{45}i$
Vector potential	E_{15}, E_{25}, E_{35}
Electric force	E_{14}, E_{24}, E_{34}
Magnetic force	$E_{23}i$, $E_{31}i$, $E_{12}i$
Momentum density	$E_{14}i$, $E_{24}i$, $E_{34}i$
Angular momentum density	$E_{01}i$, $E_{02}i$, $E_{03}i$
Radius of space curvature	$E_{05}i$

A few special points may be noticed. In accordance with the molar outlook, a momentum and its conjugate coordinate are considered to have the same direction; they have therefore the same symbolic coefficient and are homothetic. The volume (symbol E_{45}), used in forming densities is the 'unordered volume', which is a vector; the ordered product of three linear elements has the symbol $E_{15}i\,.\,E_{25}i\,.\,E_{35}i = -E_{04}$ and is a vector-density. The symbolic coefficient of electrostatic potential ϕ is obtained from the condition that $e\phi$ is a mechanical energy, or alternatively from the condition that the operator ∇^2 is real and algebraic. The symbols for velocity can be obtained from the consideration that velocity (in terms of the velocity of light) is the imaginary angle of a Lorentz transformation.

The most conspicuous absentee in the foregoing table is *pressure*. Our symbolic nomenclature does not yet provide for symmetrical tensors of the second rank; they require the double frame treated in Chapter VIII. Mass density and momentum density are also components of a second rank tensor in a double frame; but they can be treated alternatively as components of a strain vector in a simple frame. There is no such alternative treatment for pressure.

The symbolic coefficient describes the nature of the quantity, 'nature' being defined by reference to such concepts as relativistic equivalence, chirality, idempotency, all of which are comprised in the general concept of group structure. We call these *structural concepts*,[a] and the corresponding nature of the physical quantity is its structural nature.

In Chapters I–V physical quantities have been measured in natural units, with one extraneous standard left disposable. The treatment begun in Chapter VI involves different units altogether; for the components of P are not of homogeneous dimensions in any possible reckoning. The only consistent course is to assign to them zero dimensions, so that they are pure numbers independent of extraneous standards. This is clinched by the introduction of idempotent vectors, which are necessarily pure numbers. We have, in fact, begun to develop the system of description of physical structure by pure numbers, which logically precedes the introduction of dimensional quantities, since it is required in the definition of a reproducible standard of length.

For each kind of quantity occurring in the momentum vector there is a 'fixed-scale unit', and the coefficients p_μ are pure numbers representing ratios to the units. If we can recognise observationally a pure particle, we can determine the various fixed-scale units in C.G.S. measure; we know that the mass, angular momentum, charge and magnetic moment in fixed-scale measure are each $\frac{1}{4}$, so that it is only necessary to measure the same characteristics observationally in C.G.S. measure. The pure particle, having the idempotent property which is the statistical substitute for individuality, is the goal of our analysis; we cannot go beyond it, and we cannot rest satisfied till we reach it. And, in identifying it observationally, we identify also the fixed-scale units, or rather quarter-units, which it carries. In identifying the fixed-scale units we have to remember that certain factors (multiplicity factors, β-factors) are concealed in the practical definitions of physical quantities and should be eliminated. In these theoretical chapters (VI–VIII) the formulae are not as a rule given in a form which immediately corresponds to the definitions adopted for practical purposes; and it would therefore be misleading to express the units which they postulate in practical measure. The adaptation of the formulae, by insertion of multiplicity and β-factors, is one of the principal matters to be dealt with in each kind of practical application that we make.

The mystical conception of dimensions dies hard; and I suppose there will be some readers who find it hard to swallow a description of a certain angular momentum and a certain energy as each $\frac{1}{4}$. The feeling is that a certain difference of intrinsic nature is being lost sight of, which ought to have been indicated by different dimensions.[b] But it is not true that we describe an angular momentum and energy as each $\frac{1}{4}$; we describe them as $\frac{1}{4}E_{23}$, $\frac{1}{4}E_{45}$. Equality of measure is clearly distinguished from equality of the characteristics themselves.

Although the measures in fixed-scale theory are pure numbers, there is still a rudiment of dimensionality in the formulae; because numbers may be real or imaginary—they may have the dimensions i or i^2. By the table of symbolic coefficients the standard i (or E_{16}) by itself corresponds to charge and the standard i^2 (or $E_{16}i$) to density; and it will be found that in all cases where the same symbols occur with or without an

[a] *Philosophy of Physical Science*, p. 144.

[b] But dimensions never have been an indication of intrinsic nature. Surely there is no resemblance of intrinsic nature between a right angle and the population of a city; but they have the same dimensions. There is no difference of intrinsic nature between energy and mass; but (currently) they have different dimensions.

i-factor, the interpretation is in one case electrical and in the other case mechanical. Thus an equation which had incorrect i-dimensions would equate electrical to mechanical characteristics.

For a pure particle at rest, the relations $P = SE_{45}$, $S^2 = S$, give

$$-S = P^2. \tag{69·3}$$

We use S to describe mass density and momentum density which are components of the energy tensor; so that (69·3) satisfies the familiar condition that the energy tensor has the dimensions of the square of the momentum vector. This may be regarded as an accidental coincidence, because we are not using the natural units in which the dimensional relation was established, and in defining three-dimensional densities by $-PE_{45}$ we paid no attention to dimensions. Whether accidental or not, it has the important result of easing the transition from natural units to fixed-scale units. The separation of an energy tensor into the product of two momentum vectors is dimension-true in both parts of the theory.

70. The wave identities

The standard form (67·1) represents an idempotent vector in a specially chosen frame. We now seek to express the conditions of idempotency covariantly so as to apply in any frame.

We shall show that for an idempotent E-number P of rank 1,

$$(\varpi_\alpha - ip_{16})P = 0 \quad (\alpha = 0, 1, 2, 3, 4, 5), \tag{70·1}$$

or, in words: P is an eigensymbol of every pentadic part of P, and the corresponding eigenvalue is the quarterspur.

By straightforward multiplication we obtain

$$(\varpi_0 - ip_{16})P = E_{01}a_{01} + \ldots + E_{05}a_{05} + E_{23}b_{23} + \ldots + E_{45}b_{45} + ic_0, \tag{70·2}$$

where

$$a_{01} = p_{02}p_{12} + p_{03}p_{13} + p_{04}p_{14} + p_{05}p_{15}, \tag{70·31}$$

$$b_{12} = i(p_{03}p_{45} + p_{04}p_{53} + p_{05}p_{34} - p_{12}p_{16}), \tag{70·32}$$

$$ic_0 = p_{16}^2 + \varpi_0^2. \tag{70·33}$$

By (67·21)
$$2i(p_{12}p_{16} + p_{03}p_{45} + p_{04}p_{53} + p_{05}p_{34}) = p_{12}.$$

Hence
$$b_{12} = p_{12}(\tfrac{1}{2} - 2ip_{16}) = 0,$$

since the quarterspur is $\frac{1}{4}$. The other $b_{\mu\nu}$ vanish similarly. Also (67·42) gives

$$\varpi_0^2 = \tfrac{1}{3}(\tfrac{1}{4} - \tfrac{1}{16}) = \tfrac{1}{16} = -p_{16}^2;$$

so that $c_0 = 0$. The result is therefore

$$(\varpi_0 - ip_{16})P = \Sigma E_{0\mu}a_{0\mu}. \tag{70·4}$$

Multiplying initially by $\varpi_0 + ip_{16}$, the left-hand side vanishes since $\varpi_0^2 + p_{16}^2 = 0$. Hence

$$\varpi_0 \Sigma E_{0\mu}a_{0\mu} = ip_{16}\Sigma E_{0\mu}a_{0\mu}.$$

On multiplying out, the two sides are found to have no components in common, so that both sides must vanish. Hence $a_{0\mu} = 0$; and the result (70·1) is established.

It is useful to notice that the vanishing of the $a_{0\mu}$ gives

$$p_{02}p_{12}+p_{03}p_{13}+p_{04}p_{14}+p_{05}p_{15}=0, \tag{70·5}$$

and 14 similar equations, showing that the components form 15 pairs of orthogonal 4-vectors.

The conditions (70·1) are necessary conditions. We shall next show that sufficient conditions for $P^2=P$ are

$$\left.\begin{array}{ll}(a) & \text{any three of the equations (70·1),}\\ (b) & \text{qs}\,P=\frac{1}{4}.\end{array}\right\} \tag{70·6}$$

By (70·2), the equation (70·1) with $\alpha=0$ secures that the $b_{\mu\nu}$ with neither suffix equal to 0 vanishes; and the equation with $\alpha=1$ secures that the $b_{\mu\nu}$ with neither suffix equal to 1 vanishes. There remains only b_{01} which is made to vanish by any one of the four remaining equations. Thus the condition (a) is sufficient to give $b_{\mu\nu}=0$. Using $ip_{16}=\frac{1}{4}$, we immediately derive the original conditions (67·21) for idempotency. Transforming these into (67·23), and summing in groups of five, we obtain $-\varpi_\alpha^2(1-2ip_{16})=2iv$ for all values of α. Thus ϖ_α^2 is the same for all pentads, and the remaining condition required for idempotency (67·22) is then found to be satisfied.

It can be proved similarly that $P(\varpi_\alpha-ip_{16})=0$; and that any three of these equations together with $\text{qs}\,P=\frac{1}{4}$ form sufficient conditions. The two sets of equations

$$(\varpi_\alpha-ip_{16})P=0, \quad P(\varpi_\alpha-ip_{16})=0, \tag{70·7}$$

will be called the *wave identities*. The reason for the name will appear later.

Leaving aside condition (b), the wave identities are satisfied by any multiple of an idempotent E-number of rank 1; e.g. by the momentum vector P which is $-i$ times an idempotent E-number. They do not apply directly to a strain vector, since that only satisfies $S^2=S$ in a special coordinate system. But, since $P=SE_{45}$, we have

$$(\varpi_\alpha-ip_{16})SE_{45}=0, \quad SE_{45}(\varpi_\alpha-ip_{16})=0.$$

Multiplying the first of these equations initially and finally by E_{45}, the wave identities for a strain vector are

$$\{E_{45}(\varpi_\alpha-ip_{16})\}S=0, \quad S\{E_{45}(\varpi_\alpha-ip_{16})\}=0. \tag{70·8}$$

71. Matrix representation of E-numbers

We can find a set of 16 matrices of 4 rows and columns which satisfy the definition of the E-symbols, on the understanding that the multiplication in the multiplication table (53·4) is matrix multiplication and that the unit matrix plays the part of the algebraic number 1. From the nature of the E-symbols the representation is not unique; for the set of matrices used to represent a frame E_μ might equally well have been used to represent an equivalent frame E'_μ. But in any one representation there is a one-to-one correspondence of E-numbers and matrices; so that every E-number is represented by a fourfold matrix, and every fourfold matrix represents an E-number. The theory of matrix representation of E-numbers is given in *Protons and Electrons*, §§ 3·1–3·6.

The following is one of the simplest representations of the E_μ:

$$E_{15}=\begin{matrix}0&0&i&0\\0&0&0&-i\\i&0&0&0\\0&-i&0&0\end{matrix}\qquad E_{25}=\begin{matrix}0&0&0&-i\\0&0&-i&0\\0&-i&0&0\\-i&0&0&0\end{matrix}\qquad E_{35}=\begin{matrix}i&0&0&0\\0&i&0&0\\0&0&-i&0\\0&0&0&-i\end{matrix}\qquad E_{45}=\begin{matrix}0&0&-1&0\\0&0&0&-1\\1&0&0&0\\0&1&0&0\end{matrix}$$

$$E_{01}=\begin{matrix}0&i&0&0\\i&0&0&0\\0&0&0&i\\0&0&i&0\end{matrix}\qquad E_{02}=\begin{matrix}i&0&0&0\\0&-i&0&0\\0&0&i&0\\0&0&0&-i\end{matrix}\qquad E_{03}=\begin{matrix}0&0&0&i\\0&0&-i&0\\0&-i&0&0\\i&0&0&0\end{matrix}\qquad E_{04}=\begin{matrix}0&-1&0&0\\1&0&0&0\\0&0&0&1\\0&0&-1&0\end{matrix}$$

$$E_{23}=\begin{matrix}0&0&0&-1\\0&0&-1&0\\0&1&0&0\\1&0&0&0\end{matrix}\qquad E_{31}=\begin{matrix}0&0&-1&0\\0&0&0&1\\1&0&0&0\\0&-1&0&0\end{matrix}\qquad E_{12}=\begin{matrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{matrix}\qquad E_{14}=\begin{matrix}i&0&0&0\\0&-i&0&0\\0&0&-i&0\\0&0&0&i\end{matrix}$$

$$E_{24}=\begin{matrix}0&-i&0&0\\-i&0&0&0\\0&0&0&i\\0&0&i&0\end{matrix}\qquad E_{34}=\begin{matrix}0&0&-i&0\\0&0&0&-i\\-i&0&0&0\\0&-i&0&0\end{matrix}\qquad E_{05}=\begin{matrix}0&0&0&-1\\0&0&1&0\\0&-1&0&0\\1&0&0&0\end{matrix}\qquad E_{16}=\begin{matrix}i&0&0&0\\0&i&0&0\\0&0&i&0\\0&0&0&i\end{matrix}$$

An E-number $T=\Sigma E_\mu t_\mu$ is then represented by the matrix

$$\begin{matrix}u_{16}+u_{35}+u_{02}+u_{14}, & u_{01}-u_{24}+t_{12}-t_{04}, & -u_{34}-t_{45}+u_{15}+t_{13}, & -u_{25}-t_{23}-t_{05}+u_{03},\\ u_{01}-u_{24}-t_{12}+t_{04}, & u_{16}+u_{35}-u_{02}-u_{14}, & -u_{25}-t_{23}+t_{05}-u_{03}, & -u_{34}-t_{45}-u_{15}-t_{13},\\ -u_{34}+t_{45}+u_{15}-t_{13}, & -u_{25}+t_{23}-t_{05}-u_{03}, & u_{16}-u_{35}+u_{02}-u_{14}, & u_{01}+u_{24}+t_{12}+t_{04},\\ -u_{25}+t_{23}+t_{05}-u_{03}, & -u_{34}+t_{45}-u_{15}+t_{13}, & u_{01}+u_{24}-t_{12}-t_{04}, & u_{16}-u_{35}-u_{02}+u_{14},\end{matrix}\qquad(71{\cdot}1)$$

where $u_{\mu\nu}=it_{\mu\nu}$. In this representation the E_μ are *four-point matrices*, i.e. matrices with four elements ± 1 or $\pm i$ and the other elements zero. This does not apply to representations generally.

The representation has been chosen so that real E-symbols correspond to real matrices and imaginary symbols to imaginary matrices. Thus the terms monothetic, homothetic, antithetic apply equally to the symbol E-numbers and their representative matrices. This will be called a *true representation.*[a] There is a general theorem (*Protons and Electrons*, § 3·5) that, if complex matrices are excluded, a pentad of fourfold matrices must contain three imaginary and two real matrices. Thus true representation is possible only when the characters of the symbols are assigned according to system (*a*) in § 60.

The foregoing set of matrices satisfies the multiplication table, and must always represent a set of E-symbols, though it may not be the particular set that we have chosen to designate E_μ. Thus in a general representation of the E_μ, there will always be an equivalent frame $E'_\mu=qE_\mu q^{-1}$ to which the representation (71·1) applies.

[a] Dirac's matrix system (*Quantum Mechanics*, 2nd ed. p. 255) is not a 'true representation'.

The elements of a matrix T are denoted by $T_{\alpha\beta}$ ($\alpha, \beta = 1, 2, 3, 4$) the first suffix giving the number of the row and the second the number of the column. *We use the summation convention for the row-and-column suffixes of matrices.* Unless expressly stated the summation convention is not used in this book for any other suffixes, except in quoting well-known formulae from molar relativity theory.

The diagonal sum or *spur* of a matrix T is $T_{\alpha\alpha}$. By (71·1)

$$\text{spur } T = T_{\alpha\alpha} = 4u_{16} = 4it_{16} = 4 \text{ qs } T, \tag{71·2}$$

so that the spur is four times the quarterspur—a not undesigned coincidence. This is proved for the frame E_μ which has the representation (71·1); but both $T_{\alpha\alpha}$ and qs T are invariant for rotations, so that applies also in any other frame. To show that the spur is invariant, we recall that the rule of matrix multiplication is $(pq)_{\alpha\beta} = p_{\alpha\gamma} q_{\gamma\beta}$; hence

$$T'_{\alpha\alpha} = q_{\alpha\beta} T_{\beta\gamma} q^{-1}_{\gamma\alpha} = (q^{-1}_{\gamma\alpha} q_{\alpha\beta}) T_{\beta\gamma} = (1)_{\gamma\beta} T_{\beta\gamma} = T_{\gamma\gamma}.$$

We have in all representations,

$$(E_\mu)_{\alpha\alpha} = 0 \quad \text{if} \quad \mu \neq 16, \qquad (E_{16})_{\alpha\alpha} = 4i. \tag{71·3}$$

In the representation (71·1) the imaginary E_μ are symmetrical matrices and the real E_μ are antisymmetrical matrices. This continues to hold in all true representations. If $\overline{E_\mu}$ is the transpose of E_μ, so that

$$(\overline{E_\mu})_{\alpha\beta} = (E_\mu)_{\beta\alpha}, \tag{71·4}$$

we have

$$\left.\begin{array}{ll} \text{For imaginary matrices} & \overline{E_\mu} = E_\mu, \\ \text{For real matrices} & \overline{E_\mu} = -E_\mu. \end{array}\right\} \tag{71·5}$$

By (63·4) the coefficients S_μ of a momentum strain vector are all imaginary. It follows that the real part of S is a symmetrical matrix and the imaginary part an antisymmetrical matrix; in other words,

The momentum strain vector is a hermitic matrix. (71·6)

Denoting the complex conjugate by a dagger,

$$\overline{S} = S^\dagger. \tag{71·7}$$

The determinant of an E-number defined by (65·1) is the determinant of the representative matrix (71·1). The agreement extends to all representations, because the determinant of a matrix satisfies (65·2) and is accordingly invariant for transformations $q(\ldots)q^{-1}$.

72. Factorisation of E-numbers

A matrix $T_{\alpha\beta}$ may or may not be an outer product of two four-valued factors ψ_α, χ_β. In any case it can be expressed as the sum of such products (the required number of products being not more than four). If $T = \psi_\alpha \chi_\beta$, we have for its square, formed according to the rules of matrix multiplication,

$$\begin{aligned} (T^2)_{\alpha\beta} = T_{\alpha\gamma} T_{\gamma\beta} &= \psi_\alpha \chi_\gamma \psi_\gamma \chi_\beta = \psi_\gamma \chi_\gamma \psi_\alpha \chi_\beta \\ &= T_{\gamma\gamma} T_{\alpha\beta} = 4it_{16} T_{\alpha\beta} \end{aligned}$$

by (71·2). If T is normalised so that its quarterspur is $\frac{1}{4}$, this becomes $T^2 = T$. Hence

A factorisable matrix is idempotent when normalised to quarterspur $\frac{1}{4}$. (72·1)

There is also a converse theorem:

If an E-number with quarterspur $\frac{1}{4}$ *is idempotent, the matrix representing it is factorisable.* (72·2)

To prove this, let a matrix T, satisfying $T^2 = T$ and spur $T = 1$, be expressed as a sum of products

$$T_{\alpha\beta} = \psi^a_\alpha \chi^a_\beta + \psi^b_\alpha \chi^b_\beta + \psi^c_\alpha \chi^c_\beta + \ldots. \tag{72·31}$$

Let the number of terms on the right be the fewest possible. Then there can be no linear relation between the ψ's or between the χ's; for if there were a relation

$$\psi^a = c_1 \psi^b + c_2 \psi^c + \ldots,$$

we should have $$T_{\alpha\beta} = \psi^b_\alpha (\chi^b + c_1 \chi^a)_\beta + \psi^c_\alpha (\chi^c + c_2 \chi^a)_\beta + \ldots,$$

and the number of terms would be reduced by one. We write A_{ab} for the spur $(\psi^a \chi^b)_{\gamma\gamma}$, so that a product $\psi^a_\alpha \chi^a_\gamma \psi^b_\gamma \chi^b_\beta$ reduces to $A_{ba} \psi^a_\alpha \chi^b_\beta$. The idempotent condition gives

$$0 = T^2 - T = (A_{aa} - 1)\, \psi^a_\alpha \chi^a_\beta + (A_{bb} - 1)\, \psi^b_\alpha \chi^b_\beta + \ldots + A_{ba} \psi^a_\alpha \chi^b_\beta + A_{ab} \psi^b_\alpha \chi^a_\beta + \ldots. \tag{72·32}$$

Setting $\beta = 1, 2, 3, 4$ this gives four linear relations between the ψ's. Thus (72·31) is not irreducible unless the coefficients of the ψ's vanish in all four equations; that is, unless

$$(A_{aa} - 1)\, \chi^a + A_{ba} \chi^b + A_{ca} \chi^c + \ldots = 0, \tag{72·33}$$

and similar equations. But linear relations between the χ's are also excluded; so that the coefficients on (72·33) must vanish. Thus $A_{aa} = 1$, $A_{bb} = 1$, etc. Then by (72·31)

$$\begin{aligned} \text{spur } T &= \text{spur } (\psi^a \chi^a) + \text{spur } (\psi^b \chi^b) + \text{spur } (\psi^c \chi^c) + \ldots \\ &= A_{aa} + A_{bb} + A_{cc} + \ldots \\ &= 1 + 1 + 1 + \ldots. \end{aligned}$$

Since spur $T = 1$, there is only one term on the right; so that T is a single product $\psi^a \chi^a$.

It follows that:

The momentum vectors and momentum strain vectors of pure particles are factorisable. (72·4)

Conversely,

Any factorisable momentum vector, unless it is spurless, is a numerical multiple of the momentum vector of a pure particle. (72·5)

For, by dividing by a numerical factor, the factorisable vector is normalised to have quarterspur $\frac{1}{4}$, and (72·1) applies.

Owing to the concealed multiplicity factors, we commonly use momentum vectors which are numerical multiples of the true vectors; so that the momentum vectors of pure particles appear in unnormalised form. It is therefore convenient to express the purity condition in a form applicable to unnormalised vectors; in particular in a form which does not make it necessary to investigate whether the quarterspur of the true

vector is $\frac{1}{4}$ or $\frac{3}{4}$. Effectively, *factorisability is the condition for purity*. This is the condition that we apply in practice.[a]

The factors (if any) of a matrix are unique, except that they can be multiplied by reciprocal algebraic numbers a, a^{-1} real or complex.

The four-valued quantities ψ, χ, which we introduce as factors of the matrices, representing E-numbers are called *wave vectors*. When there is danger of confusion, we shall distinguish the ordinary vectors of physics, including extended vectors, as *space vectors*.

In order to distinguish inner and outer multiplication without recourse to suffixes, we adopt the following notation: a starred vector ψ^* has inner multiplication with a following matrix or vector, and therefore outer multiplication with a preceding factor; an unstarred vector ψ has inner multiplication with a preceding factor and outer multiplication with a following factor.[b] Starred and unstarred vectors will be called *initial* and *final* vectors. An important example of this notation is

$$\chi^*\psi = \text{inner product}, \quad \psi\chi^* = \text{outer product}, \tag{72·61}$$

so that

$$\chi^*\psi = \text{spur}\,(\psi\chi^*) = 4\,\text{qs}\,(\psi\chi^*). \tag{72·62}$$

We note that χ^*E_μ and $E_\mu\psi$ are wave vectors ($(\chi^*E_\mu)_\alpha = \chi^*_\beta(E_\nu)_{\beta\alpha}$), but $E_\mu\chi^*$ and ψE_μ are wave tensors of the third rank.

A hermetic matrix, if factorisable, is the product of complex conjugate factors ψ, $\psi^\dagger$. Hence, by (71·6), we have for a pure particle,

$$S = \psi\psi^\dagger \tag{72·71}$$

(the dagger being counted as a star). Then

$$P = \psi\chi^* \quad (\chi^* = \psi^\dagger E_{45}). \tag{72·72}$$

Thus S and P are fully determined if one factor ψ is specified.

In the wave identities (70·7), which are the conditions for purity and therefore for factorisability, we can write $P = \psi\chi^*$. Then $(\varpi_\alpha - ip_{16})P$ is the outer product of two vectors $(\varpi_\alpha - ip_{16})\psi$ and χ; and it will not vanish unless one of the two vectors is zero. Since $\chi \neq 0$, the wave identity becomes $(\varpi_\alpha - ip_{16})\psi = 0$. Treating the second equation similarly, (70·7) reduces to

$$(\varpi_\alpha - ip_{16})\psi = 0, \quad \chi^*(\varpi_\alpha - ip_{16}) = 0. \tag{72·81}$$

Similarly (70·8) becomes

$$\{E_{45}(\varpi_\alpha - ip_{16})\}\psi = 0, \quad \psi^\dagger\{E_{45}(\varpi_\alpha - ip_{16})\} = 0. \tag{72·82}$$

These equations can be used to determine the factors ψ, χ, $\psi^\dagger$ when P is given, or in certain cases when part of P is given together with the datum that it is factorisable. It is, of course, unnecessary to solve more than one of the four sets of equations, since $\psi^\dagger$ and χ can be found directly from ψ. We usually employ the set $(\varpi_\alpha - ip_{16})\psi = 0$.

Thus the equations originally derived as conditions for the existence of factors are converted into equations for determining the factors. A short direct proof of (72·81), not depending on the conditions for idempotency, is given in § 89.

[a] It would not exclude particles whose momentum vector is a numerical multiple of that of a pure particle. But in fixed-scale theory there are usually other conditions which exclude such particles.

[b] The asterisk is not permanently attached to the vector. The same vector may occur with or without the asterisk according to the mode of multiplication intended. It is dropped when the multiplication is indicated by suffixes.

The equation given by the pentad ϖ_5 has a special importance, since it includes the components of the momentum 4-vector. We give it in full, setting $p_{16} = m$, and omitting the term $E_{05}p_{05}$ which in practice is taken to be zero because it represents magnetic charge

$$(E_{15}ip_1 + E_{25}ip_2 + E_{35}ip_3 + E_{45}\epsilon - im)\psi = 0, \tag{72·83}$$

or, in the equivalent strain vector form (72·82),

$$(E_{14}ip_1 + E_{24}ip_2 + E_{34}ip_3 + \epsilon + imE_{45})\psi = 0. \tag{72·84}$$

Multiplying (72·83) initially by $E_{15}ip_1 + E_{25}ip_2 + E_{35}ip_3 + E_{45}\epsilon + im$, we obtain

$$p_1^2 + p_2^2 + p_3^2 - \epsilon^2 + m^2 = 0, \tag{72·85}$$

showing that the proper mass is $\pm m$. When p_1, p_2, p_3 are replaced by the usual differential operators, (72·84) is Dirac's linear wave equation.[a]

At this stage of the theory, the linear wave equation (as applied to practical problems such as the energy levels of hydrogen) is still a long way off. A great deal of intermediate theory is essential before we can deal with it. But a comparison at this stage helps us to find our bearings. In order to put our equations in a comparable form we have written m for p_{16}; but it must be pointed out that the notation is misleading. The component p_{16} is the charge. By the peculiar properties of a pure particle, the proper mass (as also the resultant spin and magnetic moment) is equal to $\pm p_{16}$; so that for the most part the misreading of m as mass does no particular harm. But it must not be supposed that the mass is made negative by reversing the sign of m; it is the charge that is reversed.

The question sometimes arises: given one pentad of P, say ϖ_0, what is the most general solution of the equation $(\varpi_0 - ip_{16})\psi = 0$ satisfied by a factor ψ of P? Since $\varpi_0^2 = -p_{16}^2$, the equation is satisfied by $\psi = (\varpi_0 + ip_{16})U$, where U is any wave vector U_α; or, introducing suffixes,

$$\psi_\alpha = (\varpi_0 + ip_{16})_{\alpha 1}U_1 + (\varpi_0 + ip_{16})_{\alpha 2}U_2 + (\varpi_0 + ip_{16})_{\alpha 3}U_3 + (\varpi_0 + ip_{16})_{\alpha 4}U_4, \tag{72·9}$$

where U_1, U_2, U_3, U_4 are arbitrary constants. If the four elementary solutions were independent, no restriction at all would be imposed on ψ_α. But $\varpi_0 + ip_{16}$ has eigenvalues $2ip_{16}$, 0, and is therefore singular. Its determinant vanishes, and there is a linear relation between the four columns of the matrix. Thus one of the four terms in (72·9) is redundant, and the solution for ψ_α contains only three arbitrary constants.

73. Wave tensors of the second rank

The components of a factorisable E-number $P = \psi\chi^* = \Sigma E_\mu p_\mu$ are given by

$$p_\mu = -\tfrac{1}{4}\chi^* E_\mu \psi. \tag{73·1}$$

This follows from (54·2). For, when suffixes are inserted, (73·1) is

$$p_\mu = -\tfrac{1}{4}\chi_\alpha (E_\mu)_{\alpha\beta}\psi_\beta = -\tfrac{1}{4}(E_\mu)_{\alpha\beta}(\psi\chi^*)_{\beta\alpha} = -\tfrac{1}{4}\,\mathrm{spur}\,(E_\mu P) = -\mathrm{qs}\,(E_\mu p).$$

By (73·1)
$$ip_\mu/p_{16} = \chi^* E_\mu \psi \div \chi^*\psi. \tag{73·21}$$

If a numerical characteristic a and an operator $\boldsymbol{a}$ are so related that

$$a = \chi^* \boldsymbol{a}\psi \div \chi^*\psi, \tag{73·22}$$

[a] *Quantum Mechanics*, 2nd ed. p. 255, equation (8).

we call $\boldsymbol{a}$ the *operational form* of the characteristic and a its *expectation value* (or average value) in the system which carries ψ, χ. In the special case in which ψ or χ^* is an eigensymbol of $\boldsymbol{a}$, the expectation value agrees with the eigenvalue.

It follows from (73·21) that

$$\boldsymbol{p}_\mu = -ip_{16}E_\mu, \tag{73·23}$$

so that, apart from a normalisation factor $-ip_{16}$, *the symbolic frame is the operational form of the momentum vector*. The momentum vector referred to here, and throughout Chapters VI–VIII, will later be distinguished as the 'particle momentum vector'; since in Chapter X we shall introduce a 'field momentum vector' which has a different operational form $-i\hbar\partial/\partial x_\mu$.

When the wave vectors are functions of position the relation between $\boldsymbol{a}$ and $\boldsymbol{a}$ is defined by

$$\boldsymbol{a} = \iiint \psi^\dagger \boldsymbol{a} \psi \, dx_1 dx_2 dx_3 \div \iiint \psi^\dagger \psi \, dx_1 dx_2 dx_3, \tag{73·24}$$

the strain vector $\psi\psi^\dagger$, which is a three-dimensional density, being used when an integration is necessary. In this case we have instead of (73·23),

$$\boldsymbol{s}_\mu = -is_{16}E_\mu, \tag{73·25}$$

where s_μ is the component of the strain vector.

When a space vector $P = \psi\chi^*$ is transformed by rotation into $P' = q\psi\chi^*q^{-1}$, P' is the product of two wave vector factors

$$\psi' = q\psi, \quad \chi'^* = \chi^* q^{-1}. \tag{73·3}$$

We regard (73·3) as the transformation law of wave vectors.

The rotation $q = e^{\frac{1}{2}E_{16}\theta} = e^{\frac{1}{2}i\theta}$ has hitherto been disregarded because it makes no difference to P; but it transforms ψ, χ into $a\psi$, $a^{-1}\chi$, where a is algebraic. The factors of a matrix are unique except for this purely algebraic transformation; thus, by including E_{16} rotations, (73·3) covers the transformation of any two factors of P into any two factors of P'. It is therefore the most general transformation law for ψ and χ consistent with the condition that their product $\psi\chi^*$ transforms as a space vector.

Wave vectors which, under a rotation q, transform like ψ will be called *covariant*, and wave vectors which transform like χ will be called *contravariant*.

The distinction between covariant and contravariant vectors is independent of the distinction between initial and final vectors. An alternative form of (73·3) is

$$\psi'^* = \psi^*\bar{q}, \quad \chi' = \bar{q}^{-1}\chi, \tag{73·4}$$

where $\bar{q}$ is the transpose of q. For (73·3) gives $\psi'_\alpha = q_{\alpha\beta}\psi_\beta = \psi_\beta \bar{q}_{\beta\alpha} = (\psi^*\bar{q})_\alpha$.

The covariant and contravariant transformations become the same if $\bar{q} = q^{-1}$; for (73·4) then gives $\psi'^* = \psi^* q^{-1}$, agreeing with the transformation of the contravariant vector χ^* in (73·3). This coalescence occurs if $q = e^{\frac{1}{2}E_\mu\theta}$ and E_μ is an antisymmetrical matrix. Then E_μ is real, and the rotation is circular. When E_μ is symmetrical (imaginary), the transformations are inverse to one another; that is to say, a contravariant vector under a rotation q transforms like a covariant vector under the inverse rotation q^{-1}. Thus:

Covariant and contravariant wave vectors behave in the same way in circular rotations, and oppositely in hyperbolic rotations. (73·5)

There are four kinds of wave tensor of the second rank. Let ψ, ϕ be covariant wave vectors and χ, ω contravariant wave vectors. The possible combinations, and their transformation laws given by (73·3) and (73·4) are:

$$\left.\begin{array}{ll} \text{(1) Covariant wave tensor} & \psi\phi^* \to q(\psi\phi^*)\bar{q}, \\ \text{(2) Mixed wave tensor} & \psi\chi^* \to q(\psi\chi^*)q^{-1}, \\ \text{(3) Mixed wave tensor} & \chi\psi^* \to \bar{q}^{-1}(\chi\psi^*)\bar{q}, \\ \text{(4) Contravariant wave tensor} & \chi\omega^* \to \bar{q}^{-1}(\chi\omega^*)q^{-1}. \end{array}\right\} \tag{73·6}$$

No. 2 is immediately identified with a space vector by its transformation law $q(\ldots)q^{-1}$. *A space vector is a mixed wave tensor of the second rank.* This is the primary link between wave tensor calculus and ordinary tensor calculus.

Setting $\bar{q}^{-1} = q'$, the transformation law of no. 3 is $q'(\ldots)q'^{-1}$; so that it also is a space vector. The space vectors (2) and (3) rotate differently under the same frame transformation—a behaviour provided for in ordinary tensor calculus by recognising two classes of space vectors, covariant and contravariant. We have $q' = q$ for the circular rotations (real antisymmetrical E_μ) and $q' = q^{-1}$ for the hyperbolic rotations; thus (2) and (3) behave in the same way under spatial rotations and oppositely under Lorentz transformations—agreeing with the distinction between covariant and contravariant vectors in elementary vector theory. We accordingly identify the wave tensors (2) with covariant space vectors, and (3) with contravariant space vectors.[a]

To interpret the transformation of (1), we first apply it to E_{45}. The transformation $e^{\frac{1}{2}E_\mu\theta}E_{45}e^{\frac{1}{2}\overline{E_\mu}\theta}$ leaves E_{45} unchanged if

(*a*) E_μ is symmetric and anticommutes with E_{45}, or

(*b*) E_μ is antisymmetric and commutes with E_{45}.

We find that the only rotations which change E_{45} are the six electrical rotations $E_\mu = E_{01}, E_{02}, E_{03}, E_{04}, E_{05}, E_{16}$. Restricting q to mechanical rotations we next apply the transformation $q(\ldots)\bar{q}$ to a strain vector $S = -PE_{45}$; it gives

$$S' = qS\bar{q} = -qPq^{-1}qE_{45}\bar{q} = -P'E_{45}.$$

Thus if the transformation $q(\ldots)q^{-1}$ is applied to P, and the transformation $q(\ldots)\bar{q}$ is applied to S, they still preserve the correspondence $S' = -P'E_{45}$ which expresses that S' is the three-dimensional density of P'. Thus (1) is identified with a strain vector. More precisely it is a covariant strain vector, i.e. the three-dimensional density of a covariant space vector.

Since the transformations (3) and (4) can be written as $q'(\ldots)q'^{-1}$, $q'(\ldots)\bar{q}'$, no. (4) is the three-dimensional density of a contravariant space vector and is identified as a contravariant strain vector.

In arriving at this identification we have excluded electrical rotations; but we shall now accept the transformation law $q(\ldots)\bar{q}$ as the fundamental definition of a strain

[a] The opposite identification might seem more natural, since (2) rotates in the same way as the coordinate axes which are generally depicted as contravariant vectors. But the essential criterion is that the momentum vector is of type (2). In molar theory this is a contravariant vector, but in microscopic theory it is a covariant vector. Since our main applications are in microscopic theory, we have adopted the nomenclature best suited to it.

vector—treating the original definition (as a three-dimensional density) as a special interpretation, which is legitimate so long as electrical rotations are not contemplated. The special interpretation implies that we are able to define a real three-dimensional volume by position vectors in space-time. Since electrical rotations mix the two antithetic parts of the extended space vector, they upset the reality conditions, and their exclusion is understandable.[a]

We have considered factorisable wave tensors. By expressing a non-factorisable wave tensor as the sum of factorisable wave tensors, the transformation formulae (73·6) and consequent identifications with space vectors or strain vectors are immediately extended to all wave tensors. Our general result is:

Covariant and contravariant space vectors are mixed wave tensors of the second rank.

Covariant and contravariant strain vectors are covariant and contravariant tensors of the second rank.

74. Wave tensors of the fourth rank

The most fundamental tensors in physics, such as $g_{\mu\nu}$ and $T_{\mu\nu}$ are symmetrical tensors of the second rank. To form these we require outer products of two space vectors $\psi\chi^*$ and $\phi\omega^*$, or sums of such products; so that they are wave tensors of the fourth rank. Since the asterisk notation becomes unmanageable, we use suffixes. Contravariant character will be indicated either by upper position of the suffixes or by underlining it; and the four vectors are denoted by $\psi_\alpha, \phi_\alpha, \chi^\alpha, \omega^\alpha$, or by $\psi_\alpha, \phi_\alpha, \chi_{\underline{\alpha}}, \omega_{\underline{\alpha}}$. Thus $P_{\alpha\underline{\beta}} = \psi_\alpha \chi^\beta$ is a covariant space vector; $P_{\underline{\alpha}\beta} = \chi^\alpha \psi_\beta$ is a contravariant space vector. The space tensors require two covariant and two contravariant wave vectors; and with these we can also form mixed strain tensors of the second rank. It is customary to use the notation $T_{\mu\nu}$, which strictly refers to one component, for the tensor as a whole; but this is confusing in equations between wave tensors and space tensors which apply only to the tensor as a whole. We therefore use the notation T_{00}, $T_0{}^0$, T^{00} for the covariant, mixed and contravariant space tensors.

We obtain the following associated group of tensors:

$$\left.\begin{aligned}
T_1 &= T_{\alpha\underline{\beta}\gamma\underline{\delta}} = \psi_\alpha \chi^\beta \phi_\gamma \omega^\delta = -T_{00} && \text{covariant space tensor,}\\
T_2 &= T_{\alpha\underline{\beta}\underline{\gamma}\delta} = \psi_\alpha \chi^\beta \omega^\gamma \phi_\delta = \ \ T_0{}^0 && \text{mixed space tensor,}\\
T_3 &= T_{\underline{\alpha}\beta\gamma\underline{\delta}} = \chi^\alpha \psi_\beta \phi_\gamma \omega^\delta = \ \ T^0{}_0 && \text{mixed space tensor,}\\
T_4 &= T_{\underline{\alpha}\beta\underline{\gamma}\delta} = \chi^\alpha \psi_\beta \omega^\gamma \phi_\delta = -T^{00} && \text{contravariant space tensor,}\\
T_5 &= T_{\alpha\underline{\beta}\gamma\underline{\delta}} = \psi_\alpha \phi_\beta \chi^\gamma \omega^\delta = -Z_0{}^0 && \text{mixed strain tensor,}\\
T_6 &= T_{\underline{\alpha}\beta\gamma\delta} = \chi^\alpha \omega^\beta \psi_\gamma \phi_\delta = -Z^0{}_0 && \text{mixed strain tensor,}\\
T_7 &= T'_{\alpha\beta\underline{\gamma}\underline{\delta}} = \psi_\alpha \phi_\beta \omega^\gamma \chi^\delta = \ \ Z'^0_0 && \text{mixed strain tensor,}\\
T_8 &= T'_{\underline{\alpha}\underline{\beta}\gamma\delta} = \omega^\alpha \chi^\beta \psi_\gamma \phi_\delta = \ \ Z'^0{}_0 && \text{mixed strain tensor.}
\end{aligned}\right\} \qquad (74\cdot1)$$

In all there are 24 permutations giving 24 associated tensors; but the foregoing 8 exhibit the various kinds of relation which can occur. Covariant and contravariant

[a] We may regard neutral space as a superposition of two opposite chiral spaces and correspondingly divide the volume V into two chiral volumes $V_+ + V_-$. A mechanical rotation has no chiral bias, and V_+, V_- rotate together. An electrical rotation rotates V_+ and V_- oppositely, so that it disrupts the neutral volume; and three-dimensional density is no longer definable.

strain tensors do not appear, because these require four wave vectors of the same kind which cannot be formed into space tensors.

In Galilean coordinates, raising or lowering a suffix reverses the sign of the space components which are represented by symmetrical matrices. Permuting the suffixes of the two wave vector factors transposes the matrix and reverses the signs of the antisymmetrical matrices. Thus the transformation $-\psi_\alpha\chi^\beta\phi_\gamma\omega^\delta \to \psi_\alpha\chi^\beta\phi_\delta\omega^\gamma$ raises the second suffix, and changes T_{00} to $T_0{}^0$, in agreement with (74·1). Thus the association of the first four tensors in (74·1) agrees with the usual association of T_{00}, $T_0{}^0$, $T^0{}_0$ and T^{00} in ordinary tensor calculus.[a] The permutation transformation can be expressed in the form

$$(-T_{00})_{\alpha\beta\gamma\delta} = (T_0{}^0)_{\alpha\beta\delta\gamma}. \tag{74·2}$$

We have introduced two different strain tensors Z, Z' given by the permutations

$$(Z_0{}^0)_{\alpha\gamma\beta\delta} = (-T_{00})_{\alpha\beta\gamma\delta}, \quad (Z'^0_0) = (-T_{00})_{\alpha\delta\beta\gamma}, \tag{74·3}$$

or

$$(Z_0{}^0)_{\alpha\delta\beta\gamma} = (T_0{}^0)_{\alpha\beta\gamma\delta}, \quad (Z'^0_0) = (T_0{}^0)_{\alpha\delta\gamma\beta}. \tag{74·4}$$

We shall call $Z_0{}^0$ the *cross dual* and Z'^0_0 the *straight dual* of $T_0{}^0$. The study of this associated group of space tensors and strain tensors forms a very fundamental part of our theory, and will be taken up after double frames have been introduced.

The outer product $S_1 \times S_2$ of the strain vectors S_1, S_2, which correspond (by the formula $S = -PE_{45}$) to the space vectors P_1, P_2, is a strain tensor which corresponds to the space tensor $P_1 \times P_2$. This 'correspondence' must be distinguished from the 'association' of strain tensors and space tensors defined by (74·3) or (74·4).

In double-suffix notation $T_{\mu\nu}$ becomes $T_{\mu\epsilon,\nu\sigma}$; so that T includes, in addition to the ordinary space tensor of the second rank, third and fourth rank tensors restricted by conditions of antisymmetry. In particular the Riemann-Christoffel tensor has the requisite antisymmetry, and can be included in the form $T_{\mu\epsilon,\nu\sigma}$. Thus fourth rank wave tensors are sufficient to cover all the tensors ordinarily used in molar physics; and we shall not need to consider combinations of more than four wave vectors. The ordinary definition of 'rank' is evidently unsuitable when we consider extended space vectors and tensors. We shall therefore re-define it by counting two antisymmetrical suffixes as one unit. Then space tensors of the first rank (space vectors) are wave tensors of the second rank; and space tensors of the second rank are wave tensors of the fourth rank.

75. Phase space

In a neutral environment electrical components are dormant, and instead of the extended momentum vector we consider the *complete momentum vector* consisting of mechanical components only and carried by a V_{10} particle. Correspondingly we have to consider a complete strain vector. By (59·3) the purely mechanical part of a strain vector is

$$S_m = \Sigma' E_\mu s_\mu \quad (\mu = 1, 2, \ldots, 10), \tag{75·1}$$

the accent indicating that the summation is over the imaginary (symmetrical) E_μ. The s_μ are interpreted as strains in a physical system containing the V_{10} particle; and each value of S_m represents a different state of strain of the system. As we have to consider

[a] Association by raising or lowering suffixes is usually regarded as identity—they are covariant, mixed and contravariant forms of the same tensor T. Thus the space tensors in (74·1) are forms of the *same* tensor. The group of 24 permutations includes other space tensors which are forms of related tensors.

probability distributions over the possible states, it is convenient to represent the states by points in a phase space. By the reality condition (63·4), the s_μ are all imaginary; we therefore set $S_m = iX$, where

$$X = \Sigma' E_\mu x_\mu, \tag{75·2}$$

and represent the real coefficients x_μ as coordinates of a point in phase space.

Setting aside the proposed physical application, we shall first investigate the pure geometry of the space (75·2). It is transformed into itself by geometrical rotations in any of its 45 coordinate planes; these are true rotations qXq^{-1} if the axes in the plane are perpendicular, and pseudo-rotations qXq if the axes are antiperpendicular (§ 56).

Consider first two perpendicular terms $E_{\mu\sigma}x_{\mu\sigma}$, $E_{\nu\sigma}x_{\nu\sigma}$; then the geometrical rotation between them is a true rotation qXq^{-1}, and the matrix in q is the product $E_{\mu\sigma}E_{\nu\sigma} = E_{\mu\nu}$. Since $E_{\mu\sigma}$, $E_{\nu\sigma}$ are imaginary, $E_{\mu\nu}$ is real and antisymmetrical; so that $\bar{q} = q^{-1}$. Considering next two antiperpendicular terms $E_{\mu\nu}$, $E_{\sigma\tau}$, the geometrical rotation is a pseudo-rotation qXq, and the matrix in q is $E_{\lambda\rho} = iE_{\mu\nu}E_{\sigma\tau}$. Since $E_{\mu\nu}$, $E_{\sigma\tau}$ are imaginary, $E_{\lambda\rho}$ is imaginary and symmetrical; so that $\bar{q} = q$. In either case the geometrical rotation is $qX\bar{q}$, which is the transformation of a strain vector. Thus, when the position vector in phase space represents a strain vector, the tensor transformations of the strain vector are the internal rotations of phase space.

To understand the significance of this result, it must first be realised that, when we use the complete momentum vector instead of the extended vector, it is not assumed that the electrical components vanish. For example, an electron is a V_{10} particle, notwithstanding its electrical characteristics. We must be allowed to investigate the probability distribution of a set of mechanical variates without necessarily assuming that no other variates exist. But in relativity theory one of our chief instruments of investigation is rotation and the tensor transformations connected with it; and rotation of the most general kind is liable to mix electrical and mechanical components. If we separate P_e from P_m the two parts do not transform independently under electrical rotations. The foregoing result shows that, if we separate S_e from S_m, S_m transforms independently *under all rotations*. The change of S due to the rotation is, so far as the mechanical components are concerned, a purely internal transformation of phase space which transforms mechanical components into mechanical components.

The 15 ordinary rotations q each give geometrical rotation in three coordinate planes of phase space; the six real E_μ give circular rotation and the nine imaginary E_μ give hyperbolic rotation. By (65·43) these are all *unitary* transformations, whether they are true rotations or pseudo-rotations; that is to say they do not alter $\det S_m$. In addition there is an E_{16} pseudo-rotation which is not unitary; it multiplies S_m by a real algebraic factor, and may be regarded as a scale transformation. We can eliminate the scale and the E_{16} transformation by considering the normalised strain vector

$$\Sigma_m = S_m/(\det S_m)^{\frac{1}{4}},$$

as in (65·7). This defines a curved 9-dimensional locus, which may be called 'orientation phase space'. It exhibits all the 45 geometrical rotations. It is not a closed space.[a] For example, the strain vector $S_m = -\frac{1}{2}i(E_{16} + E_{15})$ satisfies the reality conditions, and is idempotent and therefore singular. For a singular S_m, Σ_m is infinite.

[a] In *Protons and Electrons*, § 7·4, phase space was wrongly stated to be closed.

A volume element of orientation phase space is a solid angle $d\omega$ of directions of the position vector of phase space. We regard solid angles which are transformed into one another by geometrical rotation (i.e. by strain vector transformation) as equal. The volume element of phase space is then $r^9 d\omega dr$, where r is proportional to $(\det S_m)^{\frac{1}{2}}$. In this way a systematic measure of volume is defined in phase space, as was stated in § 15. This metric, however, does not seem to be required in any practical development. Owing to the non-closure of orientation phase space, we cannot contemplate a uniform probability distribution of orientation either as a naturally occurring distribution or for comparison purposes. Thus far our use of phase space has been limited to cases in which the probability is concentrated in a small region (almost exact states).

76. Relative space

The two-particle transformation (§ 26),

$$x_\alpha = \frac{mx_\alpha + m'x'_\alpha}{m+m'}, \quad \xi_\alpha = x'_\alpha - x_\alpha \quad (\alpha = 1, 2, 3, 4, 0), \tag{76·1}$$

is extended to five dimensions by including phase coordinates x_0, x'_0, whose conjugate momenta p_0, p'_0 are the scale momenta (§ 24). The transformation replaces the particles m, m' by an extracule and intracule, and their separate scales by a combined scale and an interchange momentum (§ 25). Since interchange affects only the relative coordinates ξ_α, the interchange momentum must be identified with the momentum ϖ_0 conjugate to ξ_0; this leaves P_0 to be identified with the combined scale—as we should naturally expect. In short, the interchange variates are the scale and phase of the relative space.

The separation of intracules and extracules is a point of bifurcation of theory. Extracules are normally dealt with by molar theory, or by simple extensions of it to provide for the analysis of the energy tensor into scale-free particles. Intracules are dealt with more intensively by quantal theory. To begin with the extracules and intracules have five coordinates x_α, ξ_α locating them respectively in an x-space and a ξ-space of five dimensions, the ξ-space being a relative space. After the separation the two spaces are each reduced, but in different ways, to four dimensions. The x-space is reduced by stabilising the scale, so that the coordinate x_0 drops out; it then becomes ordinary molar space-time. In ξ-space the spatial coordinates ξ_1, ξ_2, ξ_3 and the interchange coordinate ξ_0 are clearly indispensable; but there is no obvious use for a differential time coordinate ξ_4. We therefore reduce the ξ-space to four dimensions by stabilising ϖ_4, so that the relative-time dimension ξ_4 drops out.

We have therefore four-dimensional x- and ξ-spaces which have three dimensions in common, in the sense that the directions of the axes ξ_1, ξ_2, ξ_3 are geometrically congruent with the directions of x_1, x_2, x_3. Instead of being extended to five dimensions, the two-particle transformation is effectively reduced to three. For x_0 and ξ_4 no longer exist as coordinates; at the most they are formal variates invented to satisfy the transformation.

The four-dimensional ξ-space and x-space are structurally similar; and there is analogy between the properties of intracules referred to ξ-space and of extracules referred to x-space. We shall show presently how to put the vectors in the two spaces into correspondence so as to exhibit the analogy. It is natural to make this corre-

spondence the basis of nomenclature in fixed-scale quantum theory. For example, the non-spatial component of the momentum 4-vector in x-space is called energy, and by analogy the non-spatial component of the momentum 4-vector in ξ-space is called energy. But it is important to realise that this 'designation by analogy' is not an identification.

Thus in the very extensive specialised developments of quantal theory the nomenclature follows a *quantum-classical analogy*. This is very useful for assisting progress. But it creates confusion when the actual quantities and their quantum analogues occur together in the borderline investigations which unify physical theory. We must distinguish carefully between designation by analogy and designation by identity.

The absence of a time dimension in ξ-space needs closer investigation. Our first impression is that the omission means that we consider only steady states. But 'steady' means independent of x_4, which in any case is not a coordinate of the intracule. A variation with the differential time ξ_4 would not make the system unsteady. The real meaning of the omission of ξ_4 is that the particles m, m' are being contemplated together as *one system*; for it is implicit in the conception of a system that its parts are simultaneous. A proton to-day and an electron yesterday do not constitute the system which we call a hydrogen atom; the earth now and the moon $0^{s}{\cdot}01$ ago do not form the earth-moon system whose orbit is studied in celestial mechanics. Analysis of a system means resolving it into a set of simultaneous constituents which are together equivalent to the system at the same instant. This principle has been applied in introducing strain vectors (§ 59), when it was pointed out that the difference between a particle contemplated in isolation and the same particle contemplated as part of a system is that in the latter case it is associated with planes of simultaneity furnished by the system as a whole. When a set of particles is regarded as a system, the characteristics of the system at any instant are (by definition) a synthesis of the characteristics of the particles at that instant.

At first sight the condition $\xi_4 = 0$, resulting from the definition of a system, is a mental constraint. It does not restrict the freedom of the particles in the way that a condition such as $\xi_1 = 0$ would do. But we have seen in § 35 that 'time', when it forms a dimension of a domain of probability distribution, is a coefficient of under-observation, and means the time elapsed from the cessation of full observation. Thus the condition $x_4 = x_4'$ means that the two particles, if not fully observed, are under-observed to the same extent; in other words, $\xi_4 = 0$ is the condition for *uniform observation*. Uniform observation, like uniform environment, is naturally accepted as part of the standard conditions in which an object-system is supposed to be studied.[a]

Elementary relativity theory has made familiar—much too familiar—the fact that the term 'simultaneity' was used in classical Newtonian theory in circumstances in which no observational definition could be supplied. But there was no ambiguity in the definition of simultaneity in a two-particle system, e.g. a double star; nor is there any ambiguity in a hydrogen atom. The planes of simultaneity of a system are those of a Lorentz frame in which the system as a whole is at rest, so that the momenta P_1, P_2, P_3 of the extracule are 0. It is to this frame that the condition $\xi_4 = 0$ refers.

[a] The study of non-uniform observation is scarcely likely to be practically important; but, just as in treating non-uniform environment, the procedure would be to treat the non-uniformity as a disturbance superposed on the standard conditions of uniform observation, and incorporate it in the object-system as an extraneous object-field of a kind hitherto uninvestigated.

Thus, when the general mathematical transformation (76·1) is applied to a physical system, it is subject to the restriction

$$P_1,\ P_2,\ P_3 = 0. \tag{76·2}$$

The incorporation of (76·2) makes the transformation no longer Lorentz-invariant. The terms referring to the extracule in (26·21), (26·41), etc., are thereby very much simplified. The purpose of the special choice of axes is, however, not to simplify the extracule, but to obtain a straightforward description of the relative 4-space by four independent coordinates. If the frame is not chosen in accordance with (76·2), there are five coordinates ξ_α connected by a linear relation expressing the condition of uniform observation.

The quantum-classical analogy begins as an analogy of relative ξ-space and ordinary x-space; but it delivers a relation between two 4-spaces which has a much wider application. The simple relative space described by ξ-coordinates occurs only in the two-particle problem; but the substitution of phase for time as the fourth coordinate is a general feature of quantal theory. A more general nomenclature is therefore needed. We distinguish *molar space* and *micro space*. Micro space is derived from molar space (i.e. ordinary space-time) by taking as the fourth coordinate the phase instead of the time; or (in terms of momentum) by de-stabilising the scale and stabilising the energy.

77. Vectors in micro space

We shall now investigate in detail the conversion of molar space into micro space by de-stabilising the scale and stabilising the energy. By (57·5) the terms[a]

$$P = (E_{15}p_{15}) + E_{45}p_{45} + E_{05}p_{05} + E_{16}p_{16} \tag{77·11}$$

of a momentum vector in molar space correspond to a vector density $\mathfrak{P}$, where

$$i\mathfrak{P} = (E_{01}p_{15}) + E_{04}p_{45} - E_{05}ip_{16} - E_{16}ip_{05}. \tag{77·12}$$

From its very nature the scale cannot be zero; so that, whether unstabilised or stabilised it must be included among the components of the momentum vector. The stabilised scale of molar physics is a real algebraic constant unassociated with direction, and is accordingly identified with the component p_{16} of P. When it is de-stabilised, we introduce a fifth dimension x_0 in which to represent its probability distribution and that of its conjugate phase. Thus in de-stabilisation the scale is turned into a new direction in which it has a symbolic coefficient E_{05} instead of E_{16}.[b] A change of scale, numerical or symbolic, reflects a change of the extraneous standard used as unit for the measurement of all components of P. Thus the factor $-iE_{05}$, which converts the scale $E_{16}p_{16}$ into $E_{05}p_{16}$, must be applied to all the components; and P is converted into $-iPE_{05}$, which is equal to $-\mathfrak{P}$ by (57·5). The sign is a matter of convention, and can be made positive by measuring the de-stabilised scale in the negative direction of E_{05}. We have

[a] A bracket is used to denote the sum of three similar spatial terms.

[b] It is unimportant whether the coefficient is E_{05} or $E_{05}i$, because the factor i would be eliminated in the final result (77·3); but E_{05} can be shown to be correct. From the expression (68·4) for a pure particle, it appears that, corresponding to a real p_{16}, the energy p_{45} is real. Thus p_{16} is a time-like unit—a unit of energy rather than of linear momentum; and when represented in a direction normal to space-time, it must have the coefficient E_{05} characteristic of a time-like radius of curvature. The space-like radius of curvature of actual space-time has the symbol $E_{05}i$ (§ 63).

accordingly the result that, as a preliminary to the de-stabilisation of scale, P is changed to $\mathfrak{P}$.

In $\mathfrak{P}$ the five axes are associated with the symbols E_{01}, E_{02}, E_{03}, E_{04}, E_{05}, and we therefore designate them x_1, x_2, x_3, x_4, x_5. This involves a change of the designation of the phase coordinate from x_0 to x_5. The stabilisation of the x_5 momentum (scale) changes $\mathfrak{P}$ to P, where

$$\mathfrak{P} = PiE_{05}. \tag{77·21}$$

Similarly the stabilisation of the x_4 momentum (energy) changes $\mathfrak{P}$ to P_q where

$$\mathfrak{P} = P_q iE_{04}. \tag{77·22}$$

A simple explanation of (77·21), (77·22) is that in order to eliminate a dimension of a domain of probability distribution we must integrate over the coordinate that is to be dropped. In an x_5 integration the symbolic coefficient E_{05} of the differential dx_5 is incorporated in the integrated vector.

From (77·21) and (77·22)

$$P_q = PE_{45} = -S \tag{77·3}$$

by (59·1). Since the transformation from P to P_q, via $\mathfrak{P}$, consists in de-stabilising the scale and then stabilising the energy, P_q is the representation in micro space of the vector P in molar space.

The vectors of micro space are the strain vectors of molar space. (77·4)

This is the basal quantum-classical analogy. Micro space is used in quantum theory; and the vectors of quantal theory are the vectors P_q of micro space. *These are analogues of the classical vectors P, but are actually the classical strain vectors S.* We express this by saying that P_q is a vector in analogous designation and a strain vector in identical designation.

In particular, the quantum momentum vector of an intracule is classically a strain vector. This is easily understood if we remember that the atom to which the intracule belongs has an extended probability distribution in x-space; consequently the momentum vector P_q of the intracule is spread out as a density in molar space. The strain vector S representing this density is, as it stands, a vector P_q in micro space, and so becomes the momentum vector in fixed-scale theory.

An extended probability distribution is often described as a wave. Considering the wave in x-space, the molar point of view is that the wave itself is the intracule, which is accordingly described by a strain vector. The micro point of view is that the intracule is a compact particle at some unknown point in the wave, and it is accordingly described by a momentum vector. Equation (77·3) reconciles the wave and particle aspects, the same expression being a strain vector in molar space describing a wave and a momentum vector in micro space describing a compact particle.

78. The quantum-classical analogy

In order to display the quantum-classical analogy we introduce in micro space a frame E'_μ formed from E_μ by the transposition of suffixes

$$4, 5, 0 \rightarrow 5, 0, 4. \tag{78·1}$$

The real or imaginary character of the symbols is not changed, the three suffixes being equivalent in this respect; also, since (78·1) is an even permutation, E_{16} is unchanged. The vector $P_q = PE_{45} = \Sigma E_\mu E_{45} p_\mu$, referred to the E'-frame, is found to consist of

$$\left.\begin{aligned} &(E'_{15}p_{15}) + E'_{45}p_{05} && \text{(momentum 4-vector)},\\ &(E'_{23}ip_{01}) - (E'_{14}ip_{23}) && \text{(spin 6-vector)},\\ &(E'_{01}p_{14}) + E'_{04}p_{04} - E'_{05}ip_{16} + E'_{16}ip_{45} && \text{(dormant components).}\end{aligned}\right\} \qquad (78\cdot2)$$

This gives the 'dictionary'

$$\left.\begin{aligned} (p'_{15}),\ p'_{45} &= (p_{15}),\ p_{05},\\ (p'_{23}),\ (p'_{14}) &= (ip_{01}),\ (-ip_{23}),\\ (p'_{01}),\ p'_{04},\ p'_{05},\ p'_{16} &= (p_{14}),\ p_{04},\ -ip_{16},\ ip_{45},\end{aligned}\right\} \qquad (78\cdot3)$$

by which we translate quantum designation p'_μ into classical designation p_μ.

The governing consideration is that micro space agrees with molar space as regards the spatial dimensions but the phase coordinate (which for an intracule is the interchange coordinate) takes the place of the time. Representing the phase coordinate as a molar fifth dimension x_0, the momentum 4-vector $p'_{15}, p'_{25}, p'_{35}, p'_{45}$ of micro space must consist of the components $p_{15}, p_{25}, p_{35}, p_{05}$ in molar space. This identification is shown in (78·3); and it is sufficient to fix the permutation (78·1).

It will be seen that only four components $p_{15}, p_{25}, p_{35}, p_{04}$ have the same designation by analogy and by identity. All other components are more or less disguised by the analogy.

The analogy will not be complete unless it extends to reality conditions. Reality conditions are boundary conditions furnished by the postulated environment; so that a general treatment would resolve itself into the formal problem of determining a micro environment which is the analogue of the neutral uranoid in molar space. It is more useful to treat the special problem of finding the reality conditions for an intracule in its normal environment, namely in a hydrogen atom in a neutral uranoid. We shall show that the *relative* reality conditions—those which decide whether a term $E'_\mu p'_\mu$ is homothetic with or antithetic to a term $E'_\nu p'_\nu$—are formally the same as the molar reality conditions. The *absolute* reality condition is opposite, e.g. $E'_{16}p'_{16}$ for an intracule is antithetic to $E_{16}p_{16}$ for a molar particle or an extracule.

The proof consists mainly in recapitulating the derivation of the molar reality conditions, showing that the analogy holds at each stage. The only step presenting any difficulty occurs at the outset. To prove that the 4-vector $(E'_{15}p'_{15}) + E'_{45}p'_{45}$ is monothetic, we have to show that there is a mathematically real rotation between p'_{45} and a spatial component; or in classical designation between p_{05} and a spatial momentum. The ordinary relativity rotations of the E-frame which affect p_{05} are inhibited by the postulated neutrality of the environment; but for an intracule there exists an additional interchange rotation, which is relativistic (§ 25). This is in a plane through the radius vector ξ_1, ξ_2, ξ_3 and the extra-spatial axis x_0; and it rotates p_{05} with the component of (p_{15}, p_{25}, p_{35}) in the radial direction. The proof therefore turns on whether interchange is a mathematically real rotation. By means of a double frame we shall find in § 81 a symbolic expression for the interchange rotation which shows that it is mathematically real. Accordingly the momentum 4-vector of the intracule is monothetic.

If a particle at a real point of space has a physically real linear momentum its angular momentum about the origin is physically real. The angular momentum is the vector product of the position vector and the linear momentum. The congruence of the three-dimensional x- and ξ-spaces is a congruence of real points so that the position vectors $(E_{15}x_1)$ and $(E'_{15}\xi_1)$ for real points in the two spaces are homothetic. It does not follow that the momentum 3-vectors are homothetic; but it does follow that the angular momentum vectors are homothetic or antithetic according as the linear momentum vectors are homothetic or antithetic. Since the whole mechanical $(4+6)$-vector in molar space is monothetic, it follows that the mechanical $(4+6)$-vector[a] of the intracule is monothetic.

It remains to consider the dormant terms. For ξ-space the postulated uniformity of observation is the analogue of the postulated neutrality of the uranoid in x-space; the one postulate makes the standard conditions such as to afford no criterion for distinguishing past and future directions of relative time, and the other makes them such as to afford no criterion for distinguishing positive and negative sign of charge. The argument of §§ 63, 64 applies *mutatis mutandis* to ξ-space; and we conclude that the dormant terms are antithetic to the active terms. Thus:

The relative reality conditions for an intracule correspond by analogy (but not by identity) to the molar reality conditions. (78·4)

The momentum 3-vector is the same in both designations; but we have found (§§ 18, 21) that the momentum 3-vector of an intracule is an imaginary classical momentum. Thus p_{15} (or p'_{15}) for an intracule is antithetic to p_{15} for a classical particle. By (78·4) the antithesis must be extended to all pairs of corresponding components; that is to say, p'_μ for an intracule is antithetic to p_μ for a classical particle.

In quantum designation, the active part of the momentum vector of an intracule is imaginary, and the dormant part is real. In the strain vector $(S = -P_q E'_{45})$ *the coefficients* s_μ *are all real.* (78·5)

The same result is obtained more fundamentally from the consideration that the standard in molar physics is a quantum-specified standard; so that the scale which occurs as a component of the momentum vector of the intracule also serves as the molar scale. It must therefore satisfy the molar reality conditions as well as those of an intracule. In quantum designation the scale momentum p_{05} has become the energy p'_{45}. The molar reality condition is that p_{05} is imaginary; hence p'_{45} for an intracule is imaginary. This is antithetic to the condition for the classical energy p_{45} of a classical particle.

As would be expected, it is the fixity and discreteness of the quantum energy p'_{45} of the states of an intracule that makes it serviceable as a standard for molar physics. The so-called eigen-energies of the states are actually (i.e. classically) eigen-scales. The quantum-classical analogy changes our point of view, so that p'_{45} is regarded as a self characteristic of the intracule. But classically the scale p_{05} is the representative of the extraneous standard, added to the object-system in perfecting it, and its changes reflect changes of the extraneous standard. Thus when we use a quantum-specified standard in molar physics, we literally incorporate the p_{05} of a certain state of the

[a] I.e. the vector designated mechanical by analogy.

intracule[a] as the scale momentum of every molar object. Compared with this p_{05}, molar energies and momenta are enormous; so that in the ordinary approximation p_{05} is put equal to 0. This is the scale-free approximation. We then use a standard which has a large but unspecified ratio to the microscopic standard. This cannot be represented by p_{05} which is zero; but it is introduced as a stabilised scale, represented by p_{16} in accordance with § 77.

In the scale-free approximation $p_{05} = 0$ *for all molar systems.* (78·6)

It is useful to give explicitly the reality conditions for the various components of a momentum vector for both designations and for both kinds of particle.

Classical designation:

A 1. Molar particle: real (p_{23}), (p_{01}), p_{45}, p_{16}; imaginary (p_{15}), (p_{14}), p_{04}, p_{05};

A 2. Intracule: real (p_{15}), (p_{01}), p_{45}, p_{04}; imaginary (p_{23}), (p_{14}), p_{16}, p_{05}.

Quantum designation:

B 1. Molar particle: all imaginary;

B 2. Intracule: real (p'_{15}), (p'_{14}), p'_{04}, p'_{05}; imaginary (p'_{23}), (p'_{01}), p'_{45}, p'_{16}. (78·7)

The conditions for a molar particle apply also to an extracule.

Components which satisfy the same reality condition in an intracule and in a molar particle will be called *persistent*. The persistent components are

$$(p_{01}),\ (p_{14}),\ p_{45},\ p_{05} \quad \text{or} \quad (p'_{23}),\ (p'_{01}),\ p'_{16},\ p'_{45}. \tag{78·8}$$

The importance of the persistence of p_{05} has already been noticed. It is also important that the magnetic moment (p_{01}) is persistent; this means that the intracule has a real classical magnetic moment, which will give it a real energy in a molar magnetic field.

In quantum designation, the idempotent vector

$$-\tfrac{1}{4}i(E'_{01}+E'_{23}+E'_{45}+E'_{16}) \tag{78·9}$$

is persistent, and satisfies the reality conditions both for an intracule and an extracule.

[a] Primitively p_{05} would be the p'_{45} of the intracule in a state in which it has an idempotent momentum vector. But multiplicity factors are incorporated in the accepted procedure of transferring the scale to molar physics. It should be remembered that quantum energies are not definable by molar control; and the procedure of connecting μ with a molarly controlled mass M (§§ 29, 30) is very roundabout. This agrees with the fact that p'_{45} and p_{45} are antithetic, so that there is no direct experimental comparison.

Chapter VIII

DOUBLE FRAMES

79. The *EF*-frame

For the treatment of space tensors of the second rank, such as the energy tensor, we employ an EF-algebra which is the direct square of the E-algebra, and therefore the direct fourth power of the ζ- or quaternion algebra.

A double frame consists of 256 symbols $E_\mu F_\nu$ ($\mu, \nu = 1, 2, \ldots, 16$). The E_μ and F_μ form two equivalent simple frames, and separately obey the multiplication table (53·4) together with $E_{16} = \pm i$, $F_{16} = \pm i$. In contrast to the equivalent frames E_μ, E'_μ previously treated, the F_μ are not E-numbers. They are new symbols altogether, and every F_μ commutes with every E_μ. The multiplication table therefore includes the additional rule

$$E_\mu F_\nu = F_\nu E_\mu. \tag{79·1}$$

A linear function of the $E_\mu F_\nu$ with algebraic coefficients is called an EF-number. The product of two EF-symbols reduces to an EF-symbol multiplied by ± 1 or $\pm i$. Thus sums and products of EF-numbers are EF-numbers, and the algebra is closed.

Alternatively we use a single-suffix notation EF_μ ($\mu = 1, 2, \ldots, 256$) for the EF-symbols. Thus an EF-number T is written in the alternative forms

$$T = \sum_1^{16}\sum_1^{16} E_\mu F_\nu t_{\mu\nu} = \sum_1^{256} EF_\mu t_\mu. \tag{79·2}$$

An equivalent double frame is obtained by the transformation

$$EF'_\mu = qEF_\mu q^{-1}, \tag{79·3}$$

where q is any non-singular EF-number. The theory of relativistic equivalence applies to double frames in the same way as to simple frames (§ 55). If

$$T = \Sigma EF_\mu t_\mu, \quad T' = \Sigma EF'_\mu t_\mu,$$

so that T' is constructed in the accented frame in the same way that T is constructed in the unaccented frame, T' and T have the same equivalence as the frames; and the change $T \to T'$ is described as a rotation, i.e. an alteration without strain, of the physical system represented by T. The rotations of a physical system are accordingly given by $T' = qTq^{-1}$; and by re-resolving T' in the unaccented frame, so that it is expressed as $T' = \Sigma EF_\mu t'_\mu$, we obtain the transformations $t_\mu \to t'_\mu$ which represent rotations of a physical system in a fixed frame.

When a standard environment is postulated and account is taken of the boundary conditions which it imposes, rotations which would violate the boundary conditions are ineffective. In molar theory only the six rotations which transform the domain of molar space-time into itself are effective. As in the single frame, the physically real rotations correspond to real values of q. It must be remembered that the EF_μ are square roots of $+1$, so that in comparing EF-rotations with E-rotations, EF_μ corresponds

to $E_\mu i$ or $E_\mu F_{16}$. Circular and hyperbolic rotations are respectively $q = e^{\frac{1}{2}EF_\mu i\theta}$, and $q = e^{\frac{1}{2}EF_\mu u}$, with θ and u real. The six relativity rotations in molar space-time are

$$\left.\begin{aligned} q &= e^{\frac{1}{2}(E_\mu F_{16}+F_\mu E_{16})i\theta} \quad \text{with } \mu = 23, 31, 12,\\ q &= e^{\frac{1}{2}(E_\mu F_{16}+F_\mu E_{16})u} \quad \text{with } \mu = 14, 24, 34.\end{aligned}\right\} \tag{79·4}$$

To obtain a symbolic representation of the outer product of two space vectors, we associate one of them with the E-frame and the other with the F-frame. Thus considering two momentum 4-vectors $p_{\mu 5}$, $p'_{\mu 5}$, their symbolic forms are taken to be

$$P = (E_{15}p_{15}) + E_{45}p_{45}, \quad P' = (F_{15}p'_{15}) + F_{45}p'_{45}, \tag{79·5}$$

and their outer product has the symbolic form

$$P \times P' = \Sigma E_{\mu 5}F_{\nu 5}p_{\mu 5}p'_{\nu 5} \quad (\mu, \nu = 1, 2, 3, 4). \tag{79·61}$$

Correspondingly the ordinary (unextended) energy tensor $T_{\mu\nu}$ has the symbolic form

$$T = \Sigma E_{\mu 5}F_{\nu 5}\alpha_{\mu\nu}T_{\mu\nu}, \tag{79·62}$$

where $\alpha_{\mu\nu} = \pm 1$ or $\pm i$. For the real momentum vector ordinarily denoted by p_1, p_2, p_3, p_4, we have p_{15}, p_{25}, p_{35}, $p_{45} = ip_1$, ip_2, ip_3, p_4; correspondingly for the real energy tensor ordinarily denoted by $T_{\mu\nu}$, we have $\alpha_{\mu\nu} = -1$ for pressure terms, i for momentum terms, and $+1$ for the density terms. If the mixed tensor $T_\mu{}^\nu$ is substituted, $\alpha_{\mu\nu}$ becomes $+1$ for the pressure and density terms and $\pm i$ for the momentum terms. As the pressure and density terms are the most important, we adopt the notation

$$T_{00} = \Sigma E_{\mu 5}F_{\nu 5}t_{\mu 5,\nu 5} = \Sigma E_{\mu 5}F_{\nu 5}T_\mu{}^\nu, \tag{79·71}$$

where $T_\mu{}^\nu$ differs from the ordinary notation only in the momentum components which are modified by a factor $\pm i$. It should be understood that, notwithstanding the mixed tensor on the right introduced to connect the symbolic with the ordinary notation, (79·71) is a covariant tensor—the product of two cogredient vectors, or sum of such products. In a symbolic rotation $q(\ldots)q^{-1}$, a vector $\Sigma E_\mu p_\mu$ necessarily rotates in the same way as the coordinate axes to which the symbolic frame is anchored, and we cannot make two vectors rotate contragrediently unless we use transposes $\bar{E}$, $\bar{F}$ to relate them to the symbolic frame. The associated mixed and contravariant tensors are

$$-T_0{}^0 = \Sigma E_{\mu 5}\bar{F}_{\nu 5}T_\mu{}^\nu, \quad T^{00} = \Sigma \bar{E}_{\mu 5}\bar{F}_{\nu 5}T_\mu{}^\nu, \tag{79·72}$$

because raising or lowering a suffix reverses the sign of a space-like component, and transposing a matrix reverses the sign of a time-like component.

We here assume, as also in (79·4), that the coordinate axes are anchored in the frames E_μ and F_μ in the same way, so that their respective rotations are 'directly linked'. Occasionally we introduce a contragrediently linked EF-frame, in which the effective rotations (79·4) are modified so as to make vectors in the two frames rotate contragrediently; there is a certain gain of mathematical symmetry, but contragredient linkage is confusing for ordinary purposes.

Just as P is part of an extended momentum vector with 16 components, so T is part of an extended energy tensor with 256 components. These reduce to 136 independent components if T_{00} is constrained to be a symmetrical tensor. For uniformity we generalise (79·71) to

$$T_{00} = \Sigma E_\mu F_\nu t_{\mu\nu} = \Sigma E_\mu F_\nu T_\mu{}^\nu, \tag{79·73}$$

where T_μ^ν is loosely described as the 'ordinary energy tensor', although it is only for a small selection of the components that any ordinary notation pre-exists.

Taking the F-frame to be right-handed, $-iF_{16} = 1$. Thus we can express an E-number $P = \Sigma E_\mu p_\mu$ as an EF-number $P = \Sigma E_\mu F_{16}(-ip_\mu)$. A vector given in the E-frame can be transferred into the EF-frame in this way. It is helpful to regard an EF-number as an E-number in which the coefficients instead of being taken from the field of algebraic numbers are taken from the more extensive field of F-numbers. Conversely, an E-number is a degenerate EF-number involving only one F-symbol which, since it commutes with all the other symbols employed, is indistinguishable from i.

If a strain vector S is expressed as an EF-number

$$S = \Sigma_\mu E_\mu F_{16} \sigma_\mu, \tag{79·8}$$

all the coefficients σ_μ are real.

80. Chirality of a double frame

A double frame may be *homochiral* ($F_{16} = E_{16}$) or *antichiral* ($F_{16} = -E_{16}$); but, unless expressly stated, we shall consider only homochiral frames. The properties of an antichiral frame differ considerably; in particular it has no interchange operator (§ 81).

A homochiral frame may be right-handed ($E_{16} = F_{16} = i$) or left-handed ($E_{16} = F_{16} = -i$). The $E_\mu F_\nu$ are divided into 136 active and 120 dormant symbols, the former being unchanged and the latter reversed in sign if the chirality of the frame is reversed. Distinguishing the 6 chiral E_μ from the 10 achiral E_μ as in (55·7), the 136 active (achiral) double symbols are made up of the 10×10 products of achiral E_μ and achiral F_ν and the 6×6 products of chiral E_μ and chiral F_ν. The dormant components are so called because in an environment entirely without chirality, such as a neutral uranoid, they have no physical manifestation.

The i referred to in defining chirality belongs to an algebraic frame $(1, i)$ which supplies the field of complex numbers used as coefficients in the EF-numbers. But just as i is replaced by F_{16} in (79·8), so in a general EF-number it can be replaced by the algebraic symbol G_{16} of a third frame; and the EF-numbers are then degenerate EFG-numbers. This view eliminates the concept of absolute chirality (right- or left-handedness), leaving only relative chirality (homochirality or antichirality) of pairs of frames, symbolic or algebraic.

It would be easy to extend the construction of a double frame to triple or more highly multiple frames; and the description of a universe of 10^{79} particles provides scope for highly multiple frames if we care to tackle it that way. But we must not let multiple frame building run riot. The double frame provides for the representation of symmetrical tensors of the second rank, which in our classification includes tensors of the Riemann-Christoffel type. An EFG-frame would provide for symmetrical tensors of the third rank; but neither third nor higher rank tensors form any part of our current apparatus of description of the physical universe. The method of analysis has apparently been designed to avoid multiple frames. When a system is too complex to be represented by a second rank tensor, it is treated as a combination of simple systems represented by second rank tensors together with interactions that can also be

represented by second rank tensors. The epistemological reason for this pre-eminence of double frames, or equivalently of second rank tensors, will appear later.

This brings about a practical distinction between dormancy in an E-frame and dormancy in an EF-frame. It would be premature to drop dormant components in the E-frame; because they can combine with dormant components in the F-frame to form active components in the double frame. But the double frame is final; its dormant components get no chance of combination, and may as well be dropped at once. This is a result of method, not of necessity. They would contribute to active components in a triple frame; that is to say, in a slightly chiral environment the dormant components of a second rank tensor have physical manifestations which could be represented by third rank tensors. But physical method does not use these third rank tensors; and, although it is bound to account for the physical manifestations it does so in a way which does not connect them with the dormant components.[a] This evasion of third rank tensors is connected with the insolubility of the problem of three bodies. Broadly speaking, there is no three-particle theory in physics; and triple systems are treated by adaptations of two-particle theory, using approximate perturbation methods.

Thus in the normal application of symbolic frames to fundamental physics we do not contemplate any extension beyond double frames. The 120 chiral components of tensors are not only *dormant* but *dead*; and there is no distinction between right- and left-handed homochiral frames.

One reservation must be made, although it will not concern us until Chapter XIII. There is a quadruple $EFGH$-frame in the background, though it does not figure in the analysis. The double probability distributions that we deal with arise from degeneration by averaging of quadruple distributions (§ 25); and similarly the double frames arise from the degeneration of quadruple frames. In a measurement resulting from a casual comparison of an object-system with a comparison system the object-system would be represented by EF-numbers and the comparison system by GH-numbers; but in systematic measurement the comparison system is simplified and standardised, so that its only characteristic is an invariant scale with symbolic coefficient $G_{16}H_{16} = 1$. There is one special investigation, namely, the evaluation of the cosmical number in Chapter XIII, which stands apart from the rest of physics because of the much higher approximation involved; and in it the non-degenerate quadruple frame is used. The dormant components of the double frame must then be resurrected, since they are given opportunity of combination. The full number of components, 256, appears explicitly in the formulae of Chapter XIII. In all other problems the dormant EF_μ are to be considered non-existent.

81. The interchange operator

Let
$$I = \tfrac{1}{4}\Sigma E_\mu F_\mu \quad (\mu = 1, 2, \ldots, 16). \tag{81·1}$$

Consider the product $E_\nu I F_\nu$. It has 16 terms of the form $\frac{1}{4}E_\nu E_\mu . F_\mu F_\nu$. If E_μ, E_ν anticommute, F_μ, F_ν anticommute, and the term reduces to the form $\frac{1}{4}E_\sigma(-F_\sigma)$. If E_μ, E_ν

[a] There is a similar evasion in the simple frame. An electron has dormant momenta which will have a physical manifestation if the environment includes a proton. But the interaction of the electron and proton is described by interchange rotation in the intracule, instead of by building a second rank tensor out of the dormant momenta of the two particles.

commute, F_μ, F_ν commute, and the term reduces to the form $\frac{1}{4}E_{16}E_\sigma . F_{16}F_\sigma$. The frame being homochiral, $E_{16}F_{16} = -1$; so that in either case the term has the form $-\frac{1}{4}E_\sigma F_\sigma$, and the 16 terms are the terms of $-I$ in a different order. Thus

$$E_\nu I F_\nu = -I. \tag{81·21}$$

Again, $IE_\nu F_\nu$ consists of the terms of $-I$, except that (for $\nu \neq 16$) the eight terms which anticommute with E_ν are reversed in sign. In the sum $\Sigma_\nu IE_\nu F_\nu$ each term of $-I$ occurs eight times with correct sign and eight times with reversed sign, except that $-\frac{1}{4}E_{16}F_{16}$ occurs 16 times with correct sign. Thus $\Sigma_\nu IE_\nu F_\nu = 16(-\frac{1}{4}E_{16}F_{16}) = 4$. Hence

$$I^2 = \tfrac{1}{4}I\Sigma E_\nu F_\nu = 1. \tag{81·22}$$

Multiplying (81·21) first by initial IE_ν and secondly by final $F_\nu I$, we obtain

$$F_\nu = IE_\nu I, \quad E_\nu = IF_\nu I. \tag{81·23}$$

Hence
$$E_\mu F_\nu = IF_\mu IIE_\nu I = I(F_\mu E_\nu)\, I. \tag{81·24}$$

Thus the operation $I(\ldots)I$ interchanges the E- and F-frames. Since $I^2 = 1$, the operation can also be written as $I(\ldots)I^{-1}$; it is therefore one of the relativity rotations $q(\ldots)q^{-1}$ of the double frame. We call I the *interchange operator*.

The distinction indicated by the letters E, F might alternatively have been indicated by suffixes. In considering interchange, it is useful to start with an unlabelled frame Γ_μ, and regard the 1 or 2 strokes which turn Γ into F or E as suffixes 1 or 2. Then the momentum vectors p_μ, p'_μ of two unsuffixed particles are represented symbolically by $\Sigma\Gamma_\mu p_\mu$, $\Sigma\Gamma_\mu p'_\mu$; when the particles are suffixed these become $\Sigma F_\mu p_\mu$, $\Sigma E_\mu p'_\mu$, and interchange of suffixes is effected by the operation $I(\ldots)I$. Thus the present interchange operation is the same as in § 25, symbolic frames taking the place of comparison particles as carriers of the suffixes.

In § 25 interchange was regarded as continuous. Continuous interchange is here given by the rotation

$$q = e^{\frac{1}{2}Iu}. \tag{81·3}$$

By (81·1), I is real; so that for a physically real rotation u must be real, and the rotation is hyperbolic (§ 79). Complete interchange $q = I$ is not included in (81·3); in exponential form it is $q = e^{\frac{1}{2}(I+1)i\theta}$ with $\theta = \pi$. Except for the one value $\theta = \pi$ this is not a physically real rotation.

Since there is no circular interchange rotation, there cannot be a steady state in which the only rotation is a circulation of suffixes. This could have been foreseen. For interchange momentum has been identified with Coulomb energy; and two particles with Coulomb energy cannot form a steady system unless there is in addition a spatial rotation. Steady interchange angular momentum occurs in the same conditions as the angular momentum components p_{14}, p_{24}, p_{34} which also correspond to hyperbolic rotation; by themselves they would involve a hyperbolic rotation which cannot persist, but they can occur as components of a steady circular angular momentum in a plane tilted with respect to the adopted reference frame.

The interchange operator can be factorised. We have $I = -I_1 I_2$, where

$$\left.\begin{aligned} I_1 &= \tfrac{1}{2}(E_{23}F_{23} + E_{31}F_{31} + E_{12}F_{12} - 1), \\ I_2 &= \tfrac{1}{2}(E_{45}F_{45} + E_{50}F_{50} + E_{04}F_{04} - 1). \end{aligned}\right\} \tag{81·41}$$

This is easily verified by straightforward multiplication. Or, in the notation of (60·5),

$$\left.\begin{aligned} I_1 &= \tfrac{1}{2}(\zeta_1\zeta_1' + \zeta_2\zeta_2' + \zeta_3\zeta_3' - 1), \\ I_2 &= \tfrac{1}{2}(\theta_1\theta_1' + \theta_2\theta_2' + \theta_3\theta_3' - 1). \end{aligned}\right\} \tag{81·42}$$

In current quantum theory the spin components σ_x, σ_y, σ_z are defined as quantities whose square is $+1$, so that $\sigma_x = i\zeta_1$, etc. Treating the θ's analogously,

$$\sigma_x, \sigma_y, \sigma_z = i(\zeta_1, \zeta_2, \zeta_3), \quad \tau_u, \tau_v, \tau_w = i(\theta_1, \theta_2, \theta_3). \tag{81·5}$$

In this notation
$$I_1 = -\tfrac{1}{2}\{1 + (\boldsymbol{\sigma}, \boldsymbol{\sigma}')\}, \quad I_2 = -\tfrac{1}{2}\{1 + (\boldsymbol{\tau}, \boldsymbol{\tau}')\}, \tag{81·6}$$

$$I = -\tfrac{1}{4}\{1 + (\boldsymbol{\sigma}, \boldsymbol{\sigma}')\}\{1 + (\boldsymbol{\tau}, \boldsymbol{\tau}')\}, \tag{81·7}$$

The factor I_1 is Dirac's interchange operator,[a] which is used in a great many investigations.

We shall call $\boldsymbol{\sigma}$ the *spin*, and $\boldsymbol{\tau}$ the *co-spin*.

82. Duals

When matrix representation is employed, the commutative product $E_\mu F_\nu$ is represented by the outer product of the matrices E_μ, F_ν. This is an array of 4^4 elements which we denote by $(E_\mu F_\nu)_{\alpha\beta\gamma\delta}$. The array is called a *double matrix*. Every EF-number is represented by a double matrix, and every four-deep double matrix represents an EF-number. It is convenient to give names, *columns*, *rows*, *piles*, *tiers* to the four-dimensions of the array; so that the first suffix gives the number in the column, the second the number in the row, the third in the pile series and the fourth in the tier series.

The same system of matrix representation, e.g. that given in § 71, is used for E_μ and F_μ; for it would be misleading to disguise the equivalence of the frames by gratuitously introducing a difference in their representation. As in § 81, we denote the abstract matrix by Γ_μ. By changing the notation to E_μ we associate it with columns and rows in the array, and by changing the notation to F_μ we associate it with piles and tiers. But the notation Γ_μ is retained if the association is already indicated by the suffixes $\alpha, \beta, \gamma, \delta$. Thus

$$(E_\mu F_\nu)_{\alpha\beta\gamma\delta} = (\Gamma_\mu)_{\alpha\beta}(\Gamma_\nu)_{\gamma\delta}. \tag{82·1}$$

It follows that
$$(E_\nu F_\mu)_{\gamma\delta\alpha\beta} = (\Gamma_\nu)_{\gamma\delta}(\Gamma_\mu)_{\alpha\beta} = (E_\mu F_\nu)_{\alpha\beta\gamma\delta}, \tag{82·21}$$

and (81·24) can be expressed symbolically by

$$(\gamma\delta\alpha\beta) = I(\alpha\beta\gamma\delta)\, I. \tag{82·22}$$

The transpose $\overline{P}$ of a simple matrix P is obtained by interchanging rows and columns. Correspondingly the double matrix obtained by any interchange of columns, rows, piles and tiers of a double matrix T is a transpose of T. The associated space tensors and strain tensors in (74·1) are all transposes of one another.

We call the transpose obtained by the permutation $(\alpha\delta\gamma\beta)$ the dual of T, and denote it by $\tilde{T}$. Thus

$$\tilde{T}_{\alpha\beta\gamma\delta} = T_{\alpha\delta\gamma\beta}. \tag{82·31}$$

[a] *Quantum Mechanics*, 2nd ed. p. 226. It was first given by Dirac in *Proc. Roy. Soc.* A, **123**, 724, 1929. The complete interchange operator was introduced by Temple, *Proc. Roy. Soc.* A, **127**, 342, 1930. It was factorised, and the relation to Dirac's operator pointed out, by Eddington, *Proc. Roy. Soc.* A, **138**, 39, 1932. The complete operator is now often used in nuclear theory.

We denote the dual of $E_\mu F_\nu$ by $C_\mu D_\nu$. Thus

$$(E_\mu F_\nu)_{\alpha\beta\gamma\delta} = (\Gamma_\mu)_{\alpha\beta}(\Gamma_\nu)_{\gamma\delta}, \quad (C_\mu D_\nu)_{\alpha\beta\gamma\delta} = (\Gamma_\mu)_{\alpha\delta}(\Gamma_\nu)_{\gamma\beta}. \tag{82·32}$$

Then if $T = \Sigma E_\mu F_\nu t_{\mu\nu}$, we have $\tilde{T} = \Sigma C_\mu D_\nu t_{\mu\nu}$. We can also express $\tilde{T}$ directly as an E-number $\Sigma E_\mu F_\nu \tilde{t}_{\mu\nu}$; so that there are equivalent forms

$$\left.\begin{aligned} \tilde{T} &= \Sigma E_\mu F_\nu \tilde{t}_{\mu\nu} = \Sigma C_\mu D_\nu t_{\mu\nu}, \\ T &= \Sigma E_\mu F_\nu t_{\mu\nu} = \Sigma C_\mu D_\nu \tilde{t}_{\mu\nu}. \end{aligned}\right\} \tag{82·33}$$

Since $(C_\mu D_\nu)_{\alpha\beta\gamma\delta}$ is a 4^4 array it can be expressed as an EF-number $\Sigma_{\sigma,\tau} E_\sigma F_\tau v_{\mu\nu,\sigma\tau}$; or it can be expressed in the equally general form

$$C_\mu D_\nu = E_\mu F_\nu \Sigma_{\sigma\tau} E_\sigma F_\tau w_{\mu\nu,\sigma\tau}, \tag{82·4}$$

which is found to be more convenient. We shall determine the coefficients $w_{\mu\nu,\sigma\tau}$. Inserting suffixes (82·4) becomes

$$(\Gamma_\mu)_{\alpha\delta}(\Gamma_\nu)_{\gamma\beta} = \Sigma_{\sigma,\tau}(\Gamma_\mu\Gamma_\sigma)_{\alpha\beta}(\Gamma_\nu\Gamma_\tau)_{\gamma\delta} w_{\mu\nu,\sigma\tau}.$$

Multiplying both sides by $(\Gamma_\lambda\Gamma_\mu)_{\beta\alpha}(\Gamma_\rho\Gamma_\nu)_{\delta\gamma}$,

the left-hand side reduces to $(\Gamma_\lambda\Gamma_\mu.\Gamma_\mu.\Gamma_\rho\Gamma_\nu.\Gamma_\nu)_{\beta\beta} = (\Gamma_\lambda\Gamma_\rho)_{\beta\beta} = \text{spur}\,(\Gamma_\lambda\Gamma_\rho)$. The right-hand side reduces to $\Sigma_{\sigma,\tau} w_{\mu\nu,\sigma\tau}\,\text{spur}\,(\Gamma_\lambda\Gamma_\mu\Gamma_\mu\Gamma_\sigma)\,\text{spur}\,(\Gamma_\rho\Gamma_\nu\Gamma_\nu\Gamma_\tau)$, or $\Sigma_{\sigma,\tau} w_{\mu\nu,\sigma\tau}\,\text{spur}\,(\Gamma_\lambda\Gamma_\sigma)\,\text{spur}\,(\Gamma_\rho\Gamma_\tau)$. The only non-vanishing term is that given by $\sigma = \lambda$, $\tau = \rho$; so that the right-hand side is $16 w_{\mu\nu,\lambda\rho}$, and we have

$$\begin{aligned} 16 w_{\mu\nu,\lambda\rho} = \text{spur}\,(\Gamma_\lambda\Gamma_\rho) &= -4, \quad \text{if} \quad \lambda = \rho, \\ &= \ \ 0, \quad \text{if} \quad \lambda \neq \rho. \end{aligned}$$

Inserting this in (82·4),

$$C_\mu D_\nu = E_\mu F_\nu(-\tfrac{1}{4}\Sigma_\sigma E_\sigma F_\sigma) = -E_\mu F_\nu I. \tag{82·5}$$

Hence, by (82·33),

$$\tilde{T} = -TI. \tag{82·6}$$

We can now express the following permutations in terms of I operations

$$(\alpha\delta\gamma\beta) = -(\ldots)I, \quad (\gamma\beta\alpha\delta) = -I(\ldots), \quad (\gamma\delta\alpha\beta) = I(\ldots)I. \tag{82·7}$$

An EF-number T_s is *symmetrical*, i.e. symmetrical with respect to the E- and F-frames, if $T_s = IT_sI$. Multiplying both sides by I, this gives $IT_s = T_sI$. Similarly an antisymmetrical EF-number T_a satisfies $IT_a = -T_aI$. Any EF-number can be divided into symmetrical and antisymmetrical parts satisfying

$$IT_s = T_sI, \quad IT_a = -T_aI. \tag{82·81}$$

Setting

$$\gamma_{\mu\nu} = \tfrac{1}{2}(E_\mu F_\nu + E_\nu F_\mu), \quad \zeta_{\mu\nu} = \tfrac{1}{2}(E_\mu F_\nu - E_\nu F_\mu), \tag{82·82}$$

and introducing single-suffix notation γ_μ, ζ_μ, we can write

$$T_s = \sum_1^{136} \gamma_\mu (t_s)_\mu, \quad T_a = \sum_1^{120} \zeta_\mu (t_a)_\mu. \tag{82·83}$$

A symmetrical rotation $q_s = e^{\gamma_\mu \theta}$ does not alter I; but an antisymmetrical rotation $q_a = e^{\zeta_\mu \theta}$ gives $I' = q_a I q_a^{-1} = I q_a^{-2}$. Hence the duals $\tilde{T} = -TI$, $\tilde{T}' = -T'I$ of T and T' do not satisfy $\tilde{T}' = q\tilde{T}q^{-1}$ unless q is a symmetrical rotation. The relation of duality is invariant only for symmetrical rotations.

The algebraic component $E_{16}F_{16}t_{16,16}$ of T will be denoted by qs T and called the quarterspur, though the name is no longer strictly appropriate. For any component,

$$t_{\mu\nu} = \text{qs}\,(E_\mu F_\nu T), \tag{82·91}$$

or in matrix notation

$$t_{\mu\nu} = \tfrac{1}{16}(\Gamma_\mu)_{\beta\alpha}(\Gamma_\nu)_{\delta\gamma}T_{\alpha\beta\gamma\delta}. \tag{82·92}$$

83. The *CD*-frame

The symbols $E_\mu F_\nu$ are converted into their duals $C_\mu D_\nu$ by interchanging rows and tiers in the 4^4 array. In a certain sense the $C_\mu D_\nu$ form a frame equivalent to the *EF*-frame; but the equivalence belongs to a different order of ideas from anything hitherto contemplated. We can apply the same multiplication table to the *CD*-symbols as to the *EF*-symbols; *but the operation of multiplication is different.* Multiplication of two 4^4 arrays gives different results according as we express them as *EF*-numbers and use the *EF*-multiplication table or as *CD* numbers and use the *CD*-multiplication table.

In particular the number 1, identified in the two frames with $-E_{16}F_{16}$ and $-C_{16}D_{16}$ has different meanings. In the *EF*-frame it is the unit double matrix $\delta_\alpha{}^\beta\delta_\gamma{}^\delta$. Applying the permutation $(\alpha\delta\gamma\beta)$, the number 1 in the *CD*-frame is the double matrix $\delta_\alpha{}^\delta\delta_\gamma{}^\beta$. Correspondingly numbers which are algebraic in the *EF*-frame are not algebraic in the *CD*-frame.

By (82·6)
$$-\tilde{1} = 1\,.\,I = I, \quad \tilde{I} = -I\,.\,I = -1, \tag{83·1}$$

so that -1 (or $E_{16}F_{16}$) and I are duals. By (82·5), $C_{16}D_{16} = I$. Also, if $I_{CD} = \frac{1}{4}\Sigma_\mu C_\mu D_\mu$, we have $I_{CD} = \tilde{I} = -1$; so that

$$C_{16}D_{16} = I_{EF}, \quad E_{16}F_{16} = I_{CD}. \tag{83·2}$$

But I_{CD} is not the interchange operator for the *CD*-symbols unless we interpret the multiplications in the operation $I_{CD}(\ldots)I_{CD}$ as *CD*-multiplications.

To admit two kinds of multiplication would lead to a great deal of ambiguity; and we shall therefore not recognise *CD*-multiplication formally. The symbols $C_\mu D_\nu$ will be used as a notation for the duals of the $E_\mu F_\nu$, and not as elements of a frame. It is, however, worth while to keep in mind that there is a parallel view which would regard the dual world as fundamental and the *EF*-world as derivative. If we do not employ it, it is because a continually shifting view-point, though sometimes illuminating, has practical disadvantages.

Let $P = \Sigma E_\mu p_\mu$, $P' = \Sigma E_\mu p'_\mu$ be two vectors in an *E*-frame. To form their outer product we transfer P' to an *F*-frame by the transformation $IP'I = \Sigma F_\mu p'_\mu$. Then

$$P \times P' = P\,.\,IP'I = PI\,.\,P'I = \tilde{P}\tilde{P}', \tag{83·3}$$

by (82·6). *The outer product of two vectors is the straight product of their duals.* The converse—that the straight product of two vectors is the outer product of their duals—is not true. The dual $(-PI)$ of a space vector P is a space tensor, and the outer product would be a tensor of the fourth rank.

The transpose given by the permutation $(\alpha\gamma\beta\delta)$ is very important physically, though mathematically it is not so simple as the dual. It will be called the *cross-dual* and denoted by $T^\times$. We have

$$(C_\mu\bar{D}_\nu)_{\alpha\beta\gamma\delta} = (C_\mu D_\nu)_{\alpha\beta\delta\gamma} = (E_\mu F_\nu)_{\alpha\gamma\delta\beta} = (E_\mu\bar{F}_\nu)_{\alpha\gamma\beta\delta},$$

so that
$$C_\mu\bar{D}_\nu = (E_\mu\bar{F}_\nu)^\times. \tag{83·4}$$

If $T = \Sigma E_\mu F_\nu t_{\mu\nu}$, $T^\times = \Sigma C'_\mu D'_\nu t_{\mu\nu}$, so that $C'_\mu D'_\nu$ is the cross-dual of $E_\mu F_\nu$; we find

$$C_\mu \bar{D}'_\nu = -E_\mu \bar{I} \bar{F}_\nu \quad (\bar{I} = \tfrac{1}{4}\Sigma E_\mu \bar{F}_\mu). \tag{83·5}$$

This is proved by the same method as (82·5).

By (74·1), we find

$$Z'^0_0 = \tilde{T}^0_{\ 0}, \quad Z^0_0 = T^\times_{00}. \tag{83·6}$$

Thus the dual and cross-dual relations are used to derive strain tensors from space tensors.

Although vectors and tensors are primarily defined by transformation properties, we continue to array physical characteristics as momentum vectors and energy tensors when there is no intention of applying transformations. The utility of the complete energy tensor as a standard arrangement of 136 different characteristics of a particle is not lessened if, as usually happens, our formulae postulate a special time axis which it would be illegitimate to change. When there is no question of transformation, we may regard 'tensor' as a courtesy title. The value of this courtesy nomenclature extends to the correspondence of strain vectors and space vectors, and the association of groups of tensors in (74·1). It is convenient to have a simple way of indicating reversal of sign of certain components of T^0_0 by lowering a suffix, although we may be quite uninterested in the fact (which is also expressed by the notation) that the resulting tensor T_{00} is covariant.

By (74·1),

$$(T_{00})_{\alpha\beta\gamma\delta} = -(T^0_0)_{\alpha\beta\delta\gamma}, \quad (Z^0_0)_{\alpha\beta\gamma\delta} = -(Z'^0_0)_{\alpha\beta\delta\gamma}. \tag{83·7}$$

Thus there is a certain justification for describing Z^0_0 as Z'_{00}, which gives (83·6) the more symmetrical form

$$Z'^0_0 = \tilde{T}^0_{\ 0}, \quad Z'_{00} = T^\times_{00}. \tag{83·8}$$

This is legitimate if Z'_{00} is a courtesy tensor. The lowering of the second suffix in Z'^ν_μ here indicates just the same reversal of sign of components with space-like ν, as the lowering of the second suffix in T^ν_μ. But this sign reversal has not the same effect on tensor character in a strain tensor as in a space tensor; and Z'_{00} remains a mixed tensor. Thus when transformations are contemplated, we have to give it the more suitable notation Z^0_0.

84. Double-wave vectors

In an EF-number $T_{\alpha\beta\gamma\delta}$ the first two suffixes refer to the E-frame and the second two suffixes refer to the F-frame. These form separate chains in matrix multiplication; so that the full notation for a rotation $T' = qTq^{-1}$ is

$$T'_{\alpha\beta\gamma\delta} = q_{\alpha\epsilon\gamma\zeta}\, T_{\epsilon\zeta\eta\theta} q^{-1}_{\eta\beta\theta\delta}. \tag{84·1}$$

If T is resolvable into factors which transform separately, the factors are double-wave vectors Ψ, X with transformation laws

$$\Psi'_{\alpha\gamma} = q_{\alpha\epsilon\gamma\zeta}\, \Psi_{\epsilon\zeta}, \quad X'_{\beta\delta} = X_{\eta\theta} q^{-1}_{\eta\beta\theta\delta}. \tag{84·2}$$

By adapting the asterisk notation, the transformation laws can be put in a form similar to the laws (73·3), (73·4) for simple wave vectors

$$\Psi' = q\Psi, \quad \Psi'^* = \Psi^* \bar{q}, \quad X'^* = X^* q^{-1}, \quad X' = \bar{q}^{-1} X, \tag{84·31}$$

where

$$\bar{q}_{\alpha\beta\gamma\delta} = q_{\beta\alpha\delta\gamma}, \tag{84·32}$$

$\bar{q}$ being obtained from q by changing E_μ, F_μ to $\bar{E}_\mu$, $\bar{F}_\mu$. By analogy Ψ, X are called respectively covariant and contravariant double-wave vectors. If Φ is another covariant double-wave vector, the outer products

$$T = \Psi X^*, \quad Z = \Psi \Phi^* \tag{84·41}$$

are mixed and covariant double-wave tensors with transformation laws

$$T' = qTq^{-1}, \quad Z' = qZ\bar{q}. \tag{84·42}$$

Mixed and covariant simple wave tensors are identified with space vectors and strain vectors; and similarly mixed and covariant double-wave vectors are identified with space tensors and strain tensors. But when defined as double-wave vectors the space (strain) tensors have a more extended range of transformation than when defined in the ordinary way as the outer products or sums of outer products of space (strain) vectors. It is not uncommon for the range of transformation of a tensor to be more limited in some applications than in others, the circumstances being such as to make the more general transformations ineffective; and we shall continue to use the names 'space tensor' and 'strain tensor' for T and Z although the transformations have been generalised. But the classification of a space tensor or strain tensor as covariant, mixed or contravariant introduces a limitation of q to such forms as give a linked rotation of the E- and F-frames. The forms are

$$q_{\alpha\beta\gamma\delta} = q_{\alpha\beta}q_{\gamma\delta} \quad \text{or} \quad q_{\alpha\beta}\bar{q}^{-1}_{\gamma\delta}. \tag{84·5}$$

Since the suffixes α, β refer to the E-frame and γ, δ to the F-frame, the rotation $q_{\alpha\beta}q_{\gamma\delta}$ gives the same rotation to the two frames. We call this *cogredient linkage*. The rotation $q_{\alpha\beta}\bar{q}^{-1}_{\gamma\delta}$ makes covariant space vectors in the F-frame rotate the same way as contravariant space vectors in the E-frame. We call this *contragredient linkage*. In contragredient linkage the frames are anchored so that E_{45} and F_{45} are opposite directions along the time axis; and the frames rotate so that they remain opposite.

Contragredient linkage is an intolerable complication for ordinary purposes; and, for example, it does not harmonise with the point of view (§ 81) that E and F are derived from an unlabelled frame Γ merely by attaching suffixes. It is therefore only employed exceptionally. But it makes the general relations of space tensors and strain tensors more symmetrical. The reason is that a two-particle system must consist of particles of opposite charge if it is to be separable from its environment. This would fit more symmetrically into an antichiral double frame; but, since we only use homochiral frames, the same symmetrisation is effected by anchoring the E- and F-frames with their chiral axes E_{05}, F_{05} in opposite directions. This involves anchoring E_{45}, F_{45} also in opposite directions, an even number of reflections being necessary if the chirality is unchanged.

Consider first contragredient linkage. If $\Psi_{\alpha\gamma} = \psi_\alpha\omega_\gamma$, $X_{\beta\delta} = \chi_\beta\phi_\delta$, the transformation (84·2) with $q = q_{\alpha\beta}\bar{q}^{-1}_{\gamma\delta}$ breaks up into

$$\psi'_\alpha = q_{\alpha\epsilon}\psi_\epsilon, \quad \omega'_\gamma = \bar{q}^{-1}_{\gamma\zeta}\omega_\zeta, \quad \chi'_\beta = \chi_\eta q^{-1}_{\eta\beta}, \quad \phi'_\delta = \phi_\theta\bar{q}_{\theta\delta}. \tag{84·6}$$

Hence by (73·3), (73·4), ψ, ϕ are covariant and ω, χ are contravariant wave vectors. Then $T_{\alpha\beta\gamma\delta} = \Psi_{\alpha\gamma}X_{\beta\delta} = \psi_\alpha\chi^\beta\omega^\gamma\phi_\delta$, which is a mixed space tensor T^0_0 by (74·1).

If, on the other hand, $\psi_\alpha\omega_\gamma$ and $\chi_\beta\phi_\delta$ are both covariant double-wave vectors Ψ, Φ, we have ψ, χ covariant and ω, ϕ contravariant. Then $Z_{\alpha\beta\gamma\delta} = \Psi_{\alpha\gamma}\Phi_{\beta\delta} = \psi_\alpha\chi_\beta\omega^\gamma\phi^\delta$, which is a mixed strain tensor $Z_0{}^0$.

Next consider cogredient linkage. This will reverse the characters of the two wave vectors in the F-frame, which have suffixes γ, δ; so that

$$T_{\alpha\beta\gamma\delta} = \psi_\alpha\chi^\beta\omega_\gamma\phi^\delta, \quad Z_{\alpha\beta\gamma\delta} = \psi_\alpha\chi_\beta\omega_\gamma\phi_\delta, \tag{84·7}$$

showing that T and Z are covariant space and strain tensors T_{00}, Z_{00}.

Since the strain tensors included in the group of tensors associated by transposition of suffixes are mixed strain tensors, the association of a strain tensor and space tensor can only be invariant for the rotations $q_{\alpha\beta}\bar{q}^{-1}_{\gamma\delta}$ which represent contragredient linkage. But commonly we are not interested in the invariance of the association, the tensor appellations being 'courtesy titles'; and it is then unnecessary to introduce the complication of a contragredient anchorage of the frames.

A space tensor T_{00} which is the product of four wave vectors may either be factorised into two space vectors U, V or into two double-wave vectors Ψ, Φ the relation being

$$(T_{00})_{\alpha\beta\gamma\delta} = U_{\alpha\beta}V_{\gamma\delta} = \Psi_{\alpha\gamma}\Phi_{\beta\delta}. \tag{84·81}$$

Performing the transposition $(\alpha\gamma\beta\delta)$ we obtain the cross-dual

$$(Z_0{}^0)_{\alpha\beta\gamma\delta} = \Psi_{\alpha\beta}\Phi_{\gamma\delta} = U_{\alpha\gamma}V_{\beta\delta}. \tag{84·82}$$

Denoting the strain-vector factors of the cross-dual by $U^\times$, $V^\times$, and the double-wave vector factors by $\Psi^\times, \Phi^\times$, we have the relation

$$U^\times = \Psi, \quad V^\times = \Phi, \quad \Psi^\times = U, \quad \Phi^\times = V. \tag{84·9}$$

Since the cross-dual relation is not invariant, (84·9) is not a tensor equation. Thus although Ψ and Φ are the strain-vector factors of $Z_0{}^0$ they do not transform as strain vectors.

85. The 136-dimensional phase space

Of the 256 symbols $E_\mu F_\nu$, 136 are real and 120 imaginary. The real symbols are the 100 products of the 10 imaginary E_μ and 10 imaginary F_μ and the 36 products of the 6 real E_μ and 6 real F_μ.

The 136 real symbols are obtained by multiplying the 136 achiral symbols by $E_{45}F_{45}$.[a] The correspondence $S = -PE_{45}$ of strain vectors and space vectors gives a correspondence

$$Z = TE_{45}F_{45} \tag{85·1}$$

of strain tensors and space tensors. Thus the separation of achiral and chiral (active and dormant) components of a space tensor corresponds to the separation of real and imaginary symbolic coefficients in the corresponding strain tensor.

Corresponding to the 10-dimensional phase space associated with a simple frame, we have a 136-dimensional phase space associated with a double frame, the coordinate axes being associated with the real symbols; so that the position vector in phase space is

$$X = \Sigma' EF_\mu x_\mu, \tag{85·2}$$

[a] The ten imaginary E_μ are obtained by multiplying the ten achiral (mechanical) E_μ by E_{45}.

the accent indicating summation over the real EF_μ. As in § 75, we seek the transformations which transform phase space into itself.

Phase space is transformed into itself by geometrical rotations in any of its 9180 coordinate planes. These are either true rotations qXq^{-1} or pseudo-rotations qXq, according as the axes in the plane have anticommuting or commuting symbols. It is easily found that if two real EF-symbols anticommute the product symbol giving the plane of rotation is imaginary, and if they commute it is real. Thus both kinds of rotation are included in $qX\bar{q}$, if the bar is defined so that

$$\overline{EF}_\mu = -EF_\mu \ (EF_\mu \text{ imaginary}), \quad \overline{EF}_\mu = EF_\mu \ (EF_\mu \text{ real}). \tag{85·3}$$

We find at once from (85·3) that $\overline{E_\mu F_\nu} = \bar{E}_\mu \bar{F}_\nu$, where $\bar{E}_\mu$, $\bar{F}_\nu$ have the usual meaning; so that

$$(\overline{EF}_\mu)_{\alpha\beta\gamma\delta} = (EF_\mu)_{\beta\alpha\delta\gamma}, \quad \bar{q}_{\alpha\beta\gamma\delta} = q_{\beta\alpha\delta\gamma}. \tag{85·4}$$

This agrees with the definition of $\bar{q}$ in (84·32); and accordingly the transformation $q(\ldots)\bar{q}$, which gives purely internal rotation of phase space, is that of a covariant double-wave vector or strain tensor Z, as shown in (84·42). Thus the phase tensor, i.e. the physical tensor which is represented or partly represented by a point in phase space, is a strain tensor. For a simple frame, the corresponding result (§ 75) is that the phase vector is a strain vector.

The general theory of 136-dimensional phase space follows by analogy with 10-dimensional phase space. The $E_{16}F_{16}$ transformation is a scale transformation, and the other 255 transformations are unitary. We can introduce an 'orientation phase space' of 135 dimensions; but it contains singular points or loci, and is not closed. The phase tensor may have the full number of 256 components; it is not supposed that the unrepresented components vanish, but that we are uninterested in their probability distribution. The phase tensor being identified with a strain tensor, the unrepresented part is identified with the dormant part by (85·1); and for that reason it is not concerned in the strains occurring under standard conditions.

We see therefore that the 136-dimensional phase space represents the possible strains produced by a V_{136} particle in the system to which it belongs in the same way that the 10-dimensional phase space represents the possible strains produced by a V_{10} particle; and that the limitation to 136 or 10 dimensions is due to the omission of components which are dormant in the conditions contemplated. The system providing the planes of simultaneity is characterised by the energy tensor $E_{45}F_{45}$ or the momentum vector E_{45}; either description signifies that it is at rest relatively to the adopted time axis.

The phase tensor Z is mixed or covariant according as we adopt contragredient or cogredient linkage. Adopting contragredient linkage, we can set $Z_0{}^0 = \psi_\alpha \phi_\beta \omega^\gamma \chi^\delta$. Then the cross-dual permutation gives a space tensor $T_{00} = \psi_\alpha \omega^\beta \phi_\gamma \chi^\delta$. The active phase tensor, whose symbolic coefficients satisfy $\overline{EF}_\mu = EF_\mu$, satisfies

$$(Z_0{}^0)_{\alpha\beta\gamma\delta} = (Z_0{}^0)_{\beta\alpha\delta\gamma}. \tag{85·51}$$

Hence (interchanging the second and third suffixes on both sides) the corresponding part of T_{00} satisfies

$$(T_{00})_{\alpha\gamma\beta\delta} = (T_{00})_{\beta\delta\alpha\gamma}. \tag{85·52}$$

By (82·22) this is the condition $T_{00} = IT_{00}I$ for a symmetrical tensor.

In molar relativity theory, symmetry is an essential property of an energy tensor. In microscopic analysis there is no objection to introducing antisymmetrical terms in the individual elements which cancel in the molar average;[a] but in the standard particle, which is defined simply as the carrier of an element of molar energy tensor, the symmetry condition must apply. Its energy tensor (in classical designation) has therefore the form $T_s = \sum_1^{136} \gamma_\mu (t_s)_\mu$, found in (82·83); and the number of degrees of freedom is accordingly 136. The foregoing result shows that, in order that all possible values of T_s may be represented by points in 136-dimensional phase space, the phase tensor must be the cross-dual of T_s. We have therefore the fundamental relation

$$Z_0{}^0 = (T_{00})^\times \tag{85·6}$$

between the phase tensor and the classically designated energy tensor.

If we introduce other types of microscopic particle, not so directly connected as the standard particle with the molar energy tensor, the dormant components of the phase tensor correspond to antisymmetrical components of the classical energy tensor. It is therefore possible to give an interpretation to antisymmetrical terms in a classical energy tensor if chirality is not excluded from the environment. In molar relativity theory it is well known that the symmetry of the energy tensor, and more generally the symmetry of the Riemann-Christoffel tensor, is the result of postulating Riemannian, i.e. neutral, space-time; for when a generalised R.C. tensor ${}^*B_{\mu\epsilon\nu\sigma}$, in which $\mu\epsilon$ is not symmetrical with $\nu\sigma$, is admitted we pass over to Weyl's geometry or to the author's affine extension of it, which embodies electromagnetic potentials as well as a metrical tensor in the geometrical frame.[b]

86. Uranoid and aether

From an energy tensor $T_0{}^0$ we can derive strain tensors by two fundamentally different processes. 'Correspondence' gives a strain tensor $TE_{45}F_{45}$, which will be called Z_u; 'association' gives a strain tensor $\tilde{T} = -TI$, which will be called Z_a. Thus

$$Z_u = TE_{45}F_{45}, \quad Z_a = -TI. \tag{86·1}$$

If the strain tensor represents strains produced by the particle carrying T in a system of which it forms part, we must interpret Z_u and Z_a as the strains produced by the particle in two different types of system U and A. The feature of the system U is that its mechanical characteristics define a time axis with corresponding planes of simultaneity; if it has any other properties these are assumed not to affect the strains Z_u. The system U is therefore fully represented by a momentum vector $E_{45}m$, or in a double frame by an energy tensor $E_{45}F_{45}\rho$; that is to say, it is a particle or assemblage of particles at rest. Since the standard uranoid is such an assemblage, Z_u is most simply defined as the tensor describing the strains produced by the particle in the standard uranoid U.

Since the energy tensor of the system U is a multiple of $E_{45}F_{45}$, we tentatively assume that the energy tensor of the system A is a multiple of $-I$. To determine the nature

[a] For example, the bi-particle composed of a V_{10} object-particle and a comparison particle carries an element of energy tensor which is usually unsymmetrical.

[b] *Mathematical Theory of Relativity*, § 92.

of A, consider the part of the energy tensor $-I$ which is recognised in molar relativity theory, namely

$$-(E_{15}F_{15}\tfrac{1}{4}+E_{25}F_{25}\tfrac{1}{4}+E_{35}F_{35}\tfrac{1}{4}+E_{45}F_{45}\tfrac{1}{4}). \tag{86·2}$$

By (79·71) this is the energy tensor ordinarily denoted by

$$T_{\mu}^{\nu}=-\tfrac{1}{4}\delta_{\mu}^{\nu}. \tag{86·3}$$

It is invariant for Lorentz transformations; so that the mechanical characteristics of system A are indifferent to orientation in space-time. As we should ordinarily say, the velocity (a momentum) is entirely uncertain. The terms of the energy tensor are the mean square values of the uncertain momenta (the linear mean values being zero). In the full expression $I=\frac{1}{4}\Sigma E_{\mu}F_{\mu}$, this applies to all 16 components of the momentum vector; so that system A can be described as a distribution whose extended momentum vector is entirely uncertain.

The systems adjoint to Z_u and Z_a are therefore extreme opposites, the first having an exact momentum vector (along the time axis) and the second an entirely uncertain momentum vector. By Heisenberg's principle, the first has entirely uncertain position, and the second (if fully observed) has exact position.

The strain vector was originally introduced in order to associate with the object-particle an extraneous time direction independent of its own momentum. In Z_a this purpose is nullified, because the particle is associated with a system which, like empty space, does not discriminate any particular time direction. The energy tensor (86·3) of system A has just the same form as the cosmical energy tensor $T_{\mu}^{\nu}=-(\lambda/8\pi\kappa)\,\delta_{\mu}^{\nu}$; so that Z_a describes the strains which would be produced by the particle in the carrier of the cosmical energy tensor. The carrier of the cosmical energy tensor is commonly called empty space; but I do not think the modern fashion of calling the aether 'empty space' is conducive to clearness, and I shall here keep the classical name. System U is the uranoid, and system A is the aether. Or, in another familiar terminology, system U is an Einstein universe and system A is a de Sitter universe.

It is worth while emphasising that the best known and most crucial difference between matter and aether (which are both carriers of energy tensors, strains, etc.), namely that the former provides a reference system for velocities and the latter does not, is exhibited directly in the energy tensors $E_{45}F_{45}$ and $-I$, since $E_{45}F_{45}$ defines a time axis, with corresponding definitions of rest and simultaneity, and I is completely indifferent to axial orientation. Indeed if we had set ourselves the problem of describing mechanically a medium which does not provide any reference system for the measurement of velocity and angular velocity, we should have found directly that such a medium is specified by an energy tensor proportional to I.

We can see in a general way that Z_u is the appropriate strain tensor when T is the energy tensor of a quantum particle. For the quantum particle is a superposition on the standard uranoid, and in strain representation is the embodiment of a set of strains in U added to the undisturbed U. Equally Z_a is the appropriate strain tensor when T is the energy tensor of a relativity particle. For in relativity theory each particle independently is the source of a 'field' or state of strain of the aether; and the action of one element of matter on another is transmitted through the aether. The two ways of introducing a strain tensor are therefore closely connected with the transition from the classical to the quantum point of view. In further developments we have to con-

sider this in conjunction with the change of designation introduced by the quantum-classical analogy.

The treatment in this section is introductory. A fuller understanding of the subject will be reached later in this chapter.

87. The Riemann-Christoffel tensor

In ordinary tensor calculus, a vector A^μ taken by parallel displacement round the perimeter of a surface element $dS^{\nu\sigma}$ receives an increment

$$dA^\mu = \tfrac{1}{2}\Sigma_{\epsilon,\nu,\sigma} A_\epsilon dS^{\nu\sigma} B^{\mu\epsilon}{}_{\nu\sigma}. \tag{87·1}$$

This may be taken as the mathematical definition of the Riemann-Christoffel tensor $B^{\mu\epsilon}{}_{\nu\sigma}$. In order that it may be a physical definition, parallel displacement must be defined physically. This implies that there is at every point a physical reference frame, the criterion of parallel displacement being that the components of the vector referred to this frame remain unchanged. In general the physical frame which determines parallel displacement is not uniformly related to the geometrical reference frame to which the tensors in (87·1) are referred. It consists of local frames changing from point to point, because parallelism (in tensor calculus) is defined only for infinitesimal distances, and is not an integrable relation.

We shall employ rectangular coordinates in (87·1). The Riemann-Christoffel tensor implies curvature; but we shall only need to consider a small region in which the effect of curvature on the coordinate system is negligible.

Let the surface element be a small circle of radius r in the coordinate plane $x_\gamma x_\delta$ having its centre at the origin, so that $dS^{\nu\sigma}$ reduces to two components

$$dS^{\gamma\delta} = -dS^{\delta\gamma} = \pi r^2.$$

Let the vector be initially along the axis x_α, so that it reduces to one component A^α. By receiving an increment dA^β $(\beta \neq \alpha)$ the vector is rotated in the plane $x_\alpha x_\beta$ through an angle $d\theta'^{\alpha\beta}$ given by

$$A_\alpha d\theta'^{\alpha\beta} = dA^\beta = \pi r^2 A_\alpha B^{\beta\alpha}{}_{\gamma\delta},$$

so that

$$d\theta'^{\alpha\beta} = -\pi r^2 B^{\alpha\beta}{}_{\gamma\delta}, \tag{87·2}$$

the R.C. tensor being antisymmetrical in the suffixes α, β.

Thus if the carrier of A^μ describes the circle with angular velocity $\omega^{\gamma\delta}$, the vector itself will rotate with angular velocity $\omega'^{\alpha\beta}$, where

$$\frac{\omega'^{\alpha\beta}}{\omega^{\gamma\delta}} = -\tfrac{1}{2}r^2 B^{\alpha\beta}{}_{\gamma\delta}. \tag{87·3}$$

In general the whole angular velocity ω' will consist of a number of components $\omega'^{\alpha\beta}$ in different planes $x_\alpha x_\beta$.

In this motion the components of A relative to the physical frame undergo no change. It follows that ω' is also the local angular velocity of the physical frame. The R.C. tensor determines a linkage between angular displacements of the object-particle carrying A^μ and angular displacements of the physical frame. As we should ordinarily say, ω' is the recoil of the physical frame due to the motion ω of the object-particle. The angular velocities ω, ω' are relative to the geometrical reference frame.

As an example, consider rotations in the plane x_2x_3. If m is the mass of the object-particle, its angular momentum is $\Omega^{23} = mr^2\omega^{23}$; hence, by (87·3),

$$\omega'^{23} = -(B^{23}{}_{23}/2m)\Omega^{23}. \tag{87·41}$$

The physical frame recoils as though it had a moment of inertia $2m/B^{23}{}_{23}$. In an Einstein universe the value of $B^{23}{}_{23}$ is R_0^{-2}; so that the effective moment of inertia of the frame is then

$$C = 2mR_0^2. \tag{87·42}$$

This is four times the moment of inertia of a mass m spread uniformly over the hypersphere.

In this interpretation of the R.C. tensor we make no hypothesis. We have only translated the common interpretation in terms of curvature or non-integrability into an interpretation in terms of recoil. The existence of two equivalent interpretations, one geometrical and one mechanical, is a consequence of the unification of geometry and mechanics in molar relativity theory. The geometrical interpretation is followed up in general relativity theory, where it leads to far-reaching developments. Here we follow up the mechanical interpretation. We need a definition of the R.C. tensor which will survive when the curvature representation is replaced by a scale momentum and phase coordinate in a fifth dimension. By the mechanical definition the R.C. tensor becomes an array of coefficients giving the relation between angular displacements of an object-system in various planes and the resulting recoil displacements of the physical frame. In elementary dynamics these relations would be fixed by the moments of inertia of the two reacting systems. From the experimental point of view, our object-particle describing a small circle is a test body which enables us to apply a known amount of angular momentum[a] and study the recoil. This is the method which would be used in practice to investigate the coefficients of inertia of an object. Thus the R.C. tensor is to be regarded as a description of the inertial characteristics of the physical reference frame.

We have allowed for only one mobile carrier of object-vectors; so that it is implicitly assumed that each object-particle has its own physical comparison frame recoiling independently. The individual frames correspond to the comparison particles. A comparison particle in the standard environment has a uniform probability distribution throughout space. In an irregular environment the distribution will be non-uniform and the coefficients of inertia will be modified; this modification is shown by the deviation of the value of $B^{\mu\epsilon}{}_{\nu\sigma}$ from the value in the standard Einstein uranoid. We call the whole aggregation of comparison particles the 'comparison fluid'. The mean recoil of the comparison fluid due to any one object-particle would be ω'/N. Thus if a mean physical reference frame defined by the comparison fluid is employed, the recoil is exceedingly small, as we should expect. But in this case the cause of the recoil is that somewhere in the universe there is a particle which has been given an angular momentum Ω. The chance that this is *the* particle, carrying the vector A^μ whose change we have been studying, is $1/N$. We can therefore equivalently regard (87·3) as the relation between the rotation ω'/N of the whole comparison fluid and the expectation rotation ω/N (i.e. $1/N$ chance of a rotation ω) of the particle under consideration.

[a] More strictly a known amount of the nameless quantity of which angular momentum is the time derivative, which like angular momentum is conserved. We are handicapped by poverty of nomenclature in mechanics, and have had to use ω instead of $\delta\theta$, not because the time differentiation is relevant, but in order to obtain quantities with familiar names.

According to our formula any particle when given the rotation ω produces the same recoil of the physical frame; so that by this criterion all particles have the same mass. This is in accordance with our view that the currently recognised differences of mass are due to multiplicity factors improperly absorbed into the mass in adapting fundamental dynamics to experimental problems. By (87·42) the moment of inertia of a comparison particle in the standard uranoid is four times that which would be obtained by spreading its mass (assumed to be the same as that of an object-particle) uniformly over space. The factor 4 comes from the duplication of the matter of the universe in our analysis, each particle being counted as an object-particle and also as part of the recoiling environment of other particles.

When $B^{\mu\varepsilon}{}_{\nu\sigma}$ contains non-diagonal terms the recoil is not in the same plane as the object-rotation—as happens in elementary dynamics when the momental ellipsoid has unequal axes. A diagonal term may be negative, indicating a 'recoil' in the same direction as the object-rotation. Linear displacement in space or time can be regarded as angular displacement about the centre of curvature of space. Since the symbols associated with these displacements are E_{15}, E_{25}, E_{35}, E_{45}, the terms in the R.C. tensor which determine the corresponding recoils are $B^{15}{}_{15}$, $B^{25}{}_{25}$, $B^{35}{}_{35}$, $B^{45}{}_{45}$; so that for time displacement,

$$\omega'^{45}/\omega^{45} = -\tfrac{1}{2}R_0^2 B^{45}{}_{45}. \tag{87·51}$$

It would be nonsensical to admit a time interval between the object and its reference system. We must therefore make it a condition that the recoil drives the frame forward in time, keeping pace with the object-particle. Thus $\omega'^{45} = \omega^{35}$, and

$$B^{45}{}_{45} = -2/R_0^2. \tag{87·52}$$

This result, applying to the Einstein universe, cannot be verified directly, since the component $B^{45}{}_{45}$ is not included in the usual four-dimensional R.C. tensor; but it will be useful in interpreting the extended R.C. tensor.

88. The de Sitter universe

For the sake of uniformity with previous formulae we shall use $B_{\mu\varepsilon}{}^{\nu\sigma}$ rather than $B^{\mu\varepsilon}{}_{\nu\sigma}$; $B_{\mu\varepsilon}{}^{\nu\sigma}$ relates covariant angular displacements $\theta_{\nu\sigma}$, $\theta'_{\mu\varepsilon}$ in the same way that $B^{\mu\varepsilon}{}_{\nu\sigma}$ relates contravariant angular displacements.

In symbolic form the R.C. tensor is

$$B_{00} = \Sigma E_{\mu\varepsilon} F_{\nu\sigma} B_{\mu\varepsilon}{}^{\nu\sigma}, \tag{88·1}$$

where, as in (79·71), $B_{\mu\varepsilon}{}^{\nu\sigma}$ agrees with the ordinary notation as regards pressure and density terms, but differs by a factor i as regards momentum terms.[a] The form (88·1) exhibits the ordinary R.C. tensor as part of an extended tensor with a great many additional components which cannot be defined in terms of curvature but are defined in terms of recoil.

If B_{00} consists of diagonal components only, rotation in any coordinate plane gives recoil in the same plane. If further the diagonal components are all equal, we have symmetrical recoil, ω'/ω being independent of the plane of rotation. For symmetrical recoil B_{00} is proportional to $\Sigma E_\mu F_\mu$; so that

$$B_{00} = \alpha I. \tag{88·2}$$

[a] The generalised definition is: a term $E_\mu F_\nu t_{\mu\nu}$ is a density term if E_μ, F_ν are real, a pressure term if they are imaginary, and a momentum term if they are antithetic.

Symmetrical recoil corresponds to a de Sitter universe. This is a hyperboloidal space-time, circular in its space dimensions and hyperbolic in its time-dimension. Its ordinary R.C. tensor is

$$B_{\mu\epsilon}{}^{\nu\sigma} = (\delta_\mu{}^\nu\delta_\epsilon{}^\sigma - \delta_\mu{}^\sigma\delta_\epsilon{}^\nu)/R^2, \tag{88·3}$$

which reduces to $1/R^2$ for a diagonal component and 0 for a non-diagonal component. The component corresponding to displacement in time is positive like the other components. This appears to conflict with the condition at the end of § 87, that the object-particle and physical reference frame move forward in time together. But the circumstances in a de Sitter world are unique. It is well known that the de Sitter universe contains no matter; so that when a particle is selected as test body the recoiling environment is—what is left when you take away a particle that is not there. Calling this an antiparticle,[a] its displacements in space and time are such that it always coincides with and cancels the test particle. Thus we obtain the de Sitter R.C. tensor by first imposing the condition $\omega' = \omega$ for all displacements, and then reversing the sign to change the particle into an antiparticle. This agrees with the result stated.

It has here been assumed that the E- and F-frames have cogredient anchorage. It would be confusing to use contragredient anchorage in discussing the direction of recoil. But contragredient linkage of the frames is mathematically simpler. The only difference is that the covariant tensor B_{00} becomes a mixed tensor $B_0{}^0$; so that the de Sitter R.C. tensor in contragredient E- and F-frames is

$$B_0{}^0 = \alpha I. \tag{88·4}$$

The dual of this is the strain tensor

$$Z_0{}^0 = \tilde{B}_0{}^0 = \alpha\tilde{I} = \alpha E_{16}F_{16}, \tag{88·5}$$

by (83·1).

We shall find (§ 90) that the extended R.C. tensor and the extended energy tensor are one and the same; so that (88·4) is also known as the energy tensor of the de Sitter universe. The dual $Z_0{}^0$ is the energy tensor of a distribution of particles completely at rest as in the standard uranoid or planoid. Strictly (88·5) represents the planoid, not the uranoid; because in the Einstein uranoid there exist components

$$B_{2323},\ B_{3131},\ B_{1212} = R_0^{-2}$$

of the R.C. tensor, which form part of the extended energy tensor, but do not appear in (88·5).

Thus the dual transformation transforms the energy tensor (space tensor) of the de Sitter world into the energy tensor (strain tensor) of the planoid. Equivalently it transforms the energy tensor of the aether into the energy tensor of the planoid; for the de Sitter world contains no matter or radiation. This is a somewhat more precise formulation of the ideas introduced in § 86.

The transformation is important as connecting the classical (or molar relativity) outlook and the quantum outlook. For classical particles are insertions in the aether whereas quantum particles are superpositions on the planoid. This relation will be further developed in § 91.

[a] It is not a 'hole'; a hole is made by taking away something that (originally) is there.

89. The tensor identities

It can be shown that for any wave vector ψ,

$$\sum_{\sigma=0}^{\sigma=5}(E_{\mu\sigma}\psi)_\alpha(E_{\mu\sigma}\psi)_\beta-(E_{16}\psi)_\alpha(E_{16}\psi)_\beta=0. \tag{89·1}$$

This is proved by straightforward multiplication using the matrix representation in § 71. Let $P=\Sigma E_\mu p_\mu=\psi\chi^*$; then multiplying (89·1) by χ_α and using $p_\mu=-\frac{1}{4}\chi^*E_\mu\psi$, we obtain

$$(\Sigma_\sigma p_{\mu\sigma}E_{\mu\sigma}\psi-p_{16}E_{16}\psi)_\beta=0.$$

This is the wave identity $(\varpi_\mu-ip_{16})\psi=0$, found by another method in (72·81).

We shall investigate a corresponding identity for double-wave vectors. For any two wave vectors ψ, ϕ, we can show similarly that

$$\sum_{\sigma=0}^{\sigma=5}\{(E_{\mu\sigma}\psi)_\alpha(E_{\mu\sigma}\phi)_\beta+(E_{\mu\sigma}\phi)_\alpha(E_{\mu\sigma}\psi)_\beta\}-\{(E_{16}\psi)_\alpha(E_{16}\phi)_\beta+(E_{16}\phi)_\alpha(E_{16}\psi)_\beta\}=0. \tag{89·2}$$

We have

$$(E_{\mu\sigma}\psi)_\alpha(E_{\mu\sigma}\phi)_\beta+(E_{\mu\sigma}\phi)_\alpha(E_{\mu\sigma}\psi)_\beta=(E_{\mu\sigma})_{\alpha\gamma}(E_{\mu\sigma})_{\beta\delta}(\psi_\gamma\phi_\delta+\phi_\gamma\psi_\delta).$$

If $\Psi_{\alpha\beta}=\psi_\alpha\phi_\beta+\phi_\alpha\psi_\beta$, (89·2) gives

$$\{\Sigma_\sigma(E_{\mu\sigma})_{\alpha\gamma}(E_{\mu\sigma})_{\beta\delta}-(E_{16})_{\alpha\gamma}(E_{16})_{\beta\delta}\}\Psi_{\gamma\delta}=0. \tag{89·31}$$

Since any symmetrical double vector can be expressed as the sum of double vectors of the form $\psi_\alpha\phi_\beta+\phi_\alpha\psi_\beta$, (89·31) is an identity satisfied by any symmetrical double vector Ψ.

Let X be another symmetrical double vector, and let $T=\Psi X^*$. Then

$$\Psi_{\gamma\delta}X_{\alpha\beta}=T_{\gamma\alpha\delta\beta}. \tag{89·32}$$

Multiplying (89·31) by $X_{\alpha\beta}$,

$$\{\Sigma_\sigma(E_{\mu\sigma})_{\alpha\gamma}(E_{\mu\sigma})_{\beta\delta}-(E_{16})_{\alpha\gamma}(E_{16})_{\beta\delta}\}T_{\gamma\alpha\delta\beta}=0. \tag{89·33}$$

Also changing β to ζ in (89·31), and then multiplying by $X_{\alpha\beta}(E_{\mu\nu})_{\beta\zeta}$,

$$\{\Sigma_\sigma(E_{\mu\sigma})_{\alpha\gamma}(E_{\nu\sigma})_{\beta\delta}\}T_{\gamma\alpha\delta\beta}=0\quad(\mu\neq\nu). \tag{89·34}$$

The last step depends on the multiplication rule $(E_{\mu\nu})_{\beta\zeta}(E_{\mu\sigma})_{\zeta\delta}=(E_{\nu\sigma})_{\beta\delta}$ when $\sigma\neq\nu$. The term for which $\sigma=\nu$ is $(E_{\mu\nu})_{\alpha\gamma}(-1)_{\beta\delta}$, which cancels the final term $-(i)_{\alpha\gamma}(iE_{\mu\nu})_{\beta\delta}$ because of the symmetry of T for interchange of α,γ with β,δ. Equations (89·33) and (89·34) are combined in

$$\{\Sigma_\sigma(E_{\mu\sigma})_{\alpha\gamma}(E_{\nu\sigma})_{\beta\delta}-\delta_\mu{}^\nu(E_{16})_{\alpha\gamma}(E_{16})_{\beta\delta}\}T_{\gamma\alpha\delta\beta}=0. \tag{89·35}$$

Hence, by (82·92),

$$\Sigma_\sigma t_{\mu\sigma,\nu\sigma}-\delta_\mu{}^\nu t_{16,16}=0. \tag{89·4}$$

Since $\Psi_{\alpha\rho}X_{\gamma\delta}=T_{\alpha\gamma\beta\delta}=T^\times{}_{\alpha\beta\gamma\delta}$, $T^\times$ is made up of terms for which E_μ and F_ν are both symmetrical (imaginary). We have classified such terms as 'pressure terms'. Thus (89·4) is subject to the condition that the cross-dual consists of pressure terms only.

We consider next the case in which $T^\times$ consists of density terms only; so that T is the product of antisymmetrical double-wave vectors Ψ, X. The antisymmetrical identity corresponding to (89·2) is found to be[a]

$$\sum_{\sigma=0}^{\sigma=5}\{(E_{\mu\sigma}\psi)_\alpha(E_{\mu\sigma}\phi)_\beta-(E_{\mu\sigma}\phi)_\alpha(E_{\mu\sigma}\psi)_\beta\}+3\{(E_{16}\psi)_\alpha(E_{16}\phi)_\beta-(E_{16}\phi)_\alpha(E_{16}\psi)_\beta\}$$
$$=4(\phi^*E_{mn}\psi).(E_{mn})_{\alpha\beta}, \tag{89·5}$$

[a] Some steps in the calculation are given in *Protons and Electrons*, § 11·5.

where E_{mn} is the product of the two real symbols in the μ pentad. Setting

$$\Psi_{\alpha\beta} = \psi_\alpha \phi_\beta - \phi_\alpha \psi_\beta,$$

this gives

$$\{\Sigma_\sigma (E_{\mu\sigma})_{\alpha\gamma} (E_{\nu\sigma})_{\beta\delta} + 3(E_{16})_{\alpha\gamma} (E_{16})_{\beta\delta}\} \Psi_{\gamma\delta} = 2(E_{mn})_{\delta\gamma} \Psi_{\gamma\delta} (E_{mn})_{\alpha\beta}. \quad (89{\cdot}61)$$

If X is another antisymmetrical double-wave vector, and $T = \Psi X^*$, we obtain on multiplying by $X_{\alpha\beta}$,

$$\{\Sigma_\sigma (E_{\mu\sigma})_{\alpha\gamma} (E_{\nu\sigma})_{\beta\delta} + 3(E_{16})_{\alpha\gamma} (E_{16})_{\beta\delta}\} T_{\gamma\alpha\delta\beta} = -2(E_{mn})_{\delta\gamma} (E_{mn})_{\beta\alpha} T^\times{}_{\gamma\delta\alpha\beta}. \quad (89{\cdot}62)$$

Hence by (82·92)

$$\Sigma_\sigma t_{\mu\sigma,\mu\sigma} + 3t_{16,16} = -2t^\times{}_{mn,mn}. \quad (89{\cdot}7)$$

For $\mu = 0, 1, 2, 3, 4, 5$ the corresponding values of mn are 23, 31, 12, 05, 04, 45.

The corresponding formula for $\Sigma_\sigma t_{\mu\sigma,\nu\sigma}$ $(\mu \neq \nu)$ is complicated, and does not seem to be of much interest.

90. The contracted Riemann-Christoffel tensor

The summation $\Sigma_\sigma t_{\mu\sigma,\nu\sigma}$ in (89·4) resembles the operation of contracting a tensor $T_{\mu\epsilon}{}^{\nu\sigma}$. We shall apply it to the Riemann-Christoffel tensor (88·1),

$$T = B_{00} = \Sigma E_{\mu\epsilon} F_{\nu\sigma} B_{\mu\epsilon}{}^{\nu\sigma}. \quad (90{\cdot}1)$$

The Einstein tensor $G_{\mu\nu}$ and invariant G are defined by

$$G_\mu{}^\nu = B_{\mu 1}{}^{\nu 1} + B_{\mu 2}{}^{\nu 2} + B_{\mu 3}{}^{\nu 3} + B_{\mu 4}{}^{\nu 4}, \quad (90{\cdot}21)$$

$$G = G_1{}^1 + G_2{}^2 + G_3{}^3 + G_4{}^4. \quad (90{\cdot}22)$$

The sums are limited to four terms so as to agree with the recognised definitions; but (90·21) defines additional components of $G_\mu{}^\nu$ with suffixes 5 and 0. The momentum terms of $B_{\mu\epsilon}{}^{\nu\sigma}$ differ from the ordinary notation by a factor i. These yield the terms $G_1{}^4, G_2{}^4, G_3{}^4, G_4{}^1, G_4{}^2, G_4{}^3$, which will differ similarly. The reader will easily verify that this difference of notation does not affect the results which follow.

If $T^\times$ consists of pressure components only, (89·4) is applicable and gives

$$\Sigma_\sigma B_{\mu\sigma}{}^{\nu\sigma} - \delta_\mu{}^\nu B_{16}{}^{16} = 0. \quad (90{\cdot}23)$$

Hence, by (90·21),

$$G_\mu{}^\mu = B_{16}{}^{16} - B_{5\mu}{}^{5\mu} - B_{0\mu}{}^{0\mu}. \quad (90{\cdot}24)$$

We have transposed the suffixes in the last two terms, using their antisymmetrical property. Summing (90·24) for $\mu = 1, 2, 3, 4$,

$$G = 4B_{16}{}^{16} - G_5{}^5 - G_0{}^0$$
$$= 2B_{16}{}^{16} + 2B_{05}{}^{05},$$

by (90·24). Hence

$$B_{16}{}^{16} = \tfrac{1}{2}(G - 2\lambda) \quad (\lambda = B_{05}{}^{05}). \quad (90{\cdot}3)$$

By (90·23) and (90·21),

$$B_{\mu 5}{}^{\nu 5} + B_{\mu 0}{}^{\nu 0} = -G_\mu{}^\nu + \delta_\mu{}^\nu B_{16}{}^{16}$$
$$= -\{G_\mu{}^\nu - \tfrac{1}{2}\delta_\mu{}^\nu (G - 2\lambda)\}$$
$$= 8\pi\kappa T_\mu{}^\nu, \quad (90{\cdot}4)$$

by the usual formula[a] in molar relativity theory for the energy tensor $T_\mu{}^\nu$ $(\mu, \nu = 1, 2, 3, 4)$. Thus the ordinary energy tensor consists of two parts which we shall distinguish as

$$(8\pi\kappa T_\mu{}^\nu)_{\text{mech}} = B_{\mu 5}{}^{\nu 5}, \quad (8\pi\kappa T_\mu{}^\nu)_{\text{mag}} = B_{\mu 0}{}^{\nu 0}. \quad (90{\cdot}5)$$

[a] *Mathematical Theory of Relativity*, equation (54·71).

The first part is composed of the components of B whose symbolic directions agree with those of a mechanical energy tensor or outer product of two mechanical momentum 4-vectors $p_{\mu 5}$, $p'_{\mu 5}$. The second part is similar except that the magnetic momentum 4-vectors $p_{\mu 0}$, $p'_{\mu 0}$ take the place of mechanical momenta. Thus the components of the extended R.C. tensor B in directions appropriate for an energy tensor form the actual energy tensor. In other words

The extended energy tensor and the extended Riemann-Christoffel tensor are the same tensor. (90·6)

The distinctive property of the energy tensor defined by

$$-8\pi\kappa T_\mu^{\ \nu} = G_\mu^{\ \nu} - \tfrac{1}{2}g_\mu^{\ \nu}(G - 2\lambda)$$

is that it satisfies the conservation of energy and momentum *identically*. It is proved in general relativity theory that no other metrical tensor of the second rank, not containing derivatives of the $g_{\mu\nu}$ beyond the second, has this property; so that its physical identification with the energy tensor is clearly indicated. We shall distinguish it as the *conserved energy tensor*. Our developments have had little to do with conservation, which controls changes of occupation rather than the structure of states. So long as we deal with steady states energy and momentum are, of course, conserved—and conservation of pressure is thrown in gratuitously. Conservation is, however, introduced when we bring in the R.C. tensor as a measure of recoil; and (90·4) shows it in the familiar form.

It is noteworthy that magnetic momenta contribute to the conserved energy tensor in just the same way as mechanical momenta. They were called magnetic momenta (rather than magnetic moments) because they are components of the extended momentum vector; but we had not hitherto found any dynamical resemblance. The present result shows that they are momenta in fact as well as in name. Since the reality conditions for mechanical and magnetic momenta are antithetic, the magnetic momenta contribute to the energy tensor a negative density and pressure.

In (90·3) the component $B_{05}{}^{05}$ is identified with the cosmical constant λ. All other components are treated as variable characteristics of the system; so that at first sight it is incongruous that $B_{05}{}^{05}$ should be identified with a universal constant. But, since E_{05} represents the scale dimension, $B_{05}{}^{05}$ is the pure scale component. We have seen that a 'perfect' system carries a scale-indicator whose changes, if any, are reflections of the extraneous standard, so that it is invariant for any changes of the system itself; and the symbolic frame has been anchored, so that the duty of scale-indicator is associated with the suffix-pair 05. It is therefore entirely fitting that $B_{05}{}^{05}$ should be a constant unalterable by any physical change of the distribution specified by B, and alterable only by change of the extraneous unit of density. The identification of this constant with λ agrees with molar theory, where λ performs the corresponding duty of a scale or, as it is more usually called in that theory a 'gauge'. In particular, the 'gauging equation' in empty space is $G_{\mu\nu} = \lambda g_{\mu\nu}$. The density $\lambda/8\pi\kappa$ will be called the *cosmical standard*.

The observational standard, by which lengths and times are defined in molar as well as microscopic physics, is a quantum-specified standard depending on the properties

of the intracule. This is a mass standard, and the corresponding scale momentum is (in classical designation) the component p_{16} of the momentum vector (§ 77). Densities in molar physics are normally regarded as scale-free. But in cosmical theory they are not strictly scale-free; and the breakdown occurs because the cosmical constant λ is no longer negligible. Scale-free physics lies between cosmical and quantal physics; and we can import into it either a density standard from cosmical physics or a mass standard from quantal physics, the corresponding scale indicators being $B_{05}{}^{05}$ and p_{16}.

We take the opportunity to make clear the relation between the quantum-specified standard and the cosmical standard. We begin with the principle that for local measurements the only possible standard is quantum-specified (§ 4); so that molar theory must borrow its scale from quantum theory, and cannot be self-contained. The quantum-specified standard is a σ-standard. Later we introduce a multiple of the σ-standard, namely, $R_0 = 2\sigma\sqrt{N}$, which is shown to be the radius of curvature of the uranoid. By geodetic measurements on the earth's surface we can determine the radius of the earth, or equivalently we can express lengths in terms of the earth's radius as unit; similarly by triangulations in the uranoid we might ideally determine lengths in terms of its radius R_0 as standard. As the required measurements are molar, this would relieve molar theory from the necessity of borrowing its standard from quantum physics. But this breaks down because the uranoid is a postulated, not an actual, environment. A more plausible way is to use the aether instead of the uranoid, introducing the cosmical standard by the equation $G_{\mu\nu} = \lambda g_{\mu\nu}$ satisfied in space empty of matter. There is no space empty enough for this purpose within our galaxy; but regions where $G_{\mu\nu} = \lambda g_{\mu\nu}$ is satisfied to a rough approximation must exist between the galaxies. It is not sufficient that such regions should exist; we must be able to identify them. We ought at least to make sure that there is not more than one atom per cubic metre. Evidently the test of emptiness to this degree will have to be supplied by microscopic theory; so that we have not rendered molar theory self-contained, but have only changed its mode of dependence on microscopic theory.[a]

Attempts have been made to generalise the gauging equation so as to provide a cosmical standard independent (at any rate to the first order) of the density of the matter present, and having everywhere a constant ratio to the practical quantum-specified standard. Thus Weyl[b] introduced the condition $G - 6(\kappa^\alpha)_\alpha + 6\kappa_\alpha\kappa^\alpha = 4\lambda$ as an appropriate generalisation of the condition $G = 4\lambda$ which holds in empty space. Here κ_α is the electromagnetic potential vector, which, unlike the electromagnetic force, is not determined uniquely by molar measurement. The resulting standard is therefore not an observational standard. The same applies to many subsequent proposals for a unified geometry of gravitational and electromagnetic fields.

We conclude that, as an observable standard, the cosmical standard is not a rival to the quantum-specified standard. It cannot be introduced without stabilisation; certain conditions—uranoid distribution, empty space, values of an unobservable potential, etc.—have to be postulated without reference to observation. But its introduction as a supposititious molar standard is a convenient device for sealing off molar

[a] The extreme delicacy of the test is not strictly relevant. We need a test of emptiness, other than that which molar theory itself can supply, in order to get out of a vicious circle—defining our metric by $G_{\mu\nu} = \lambda g_{\mu\nu}$ in empty space, and defining empty space (molarly) as a region where the energy tensor (measured according to that metric) vanishes, the condition that the energy tensor vanishes being $G_{\mu\nu} = \lambda g_{\mu\nu}$.

[b] *Mathematical Theory of Relativity*, equation (89·2).

physics from the rest of physics. In our formal extension B_{00} of the molar Riemann-Christoffel tensor, we keep within the sealed-off domain; and the scale naturally refers to the cosmical standard.

91. States and interstates

We have found in (90·6) that the energy tensor and R.C. tensor are included in the same extended tensor; so that the same extended tensor has been treated under two names and with two notations T_{00}, B_{00}. (Allowing for the difference of units, $B_{00} = 8\pi\kappa T_{00}$.) In future we shall use the name energy tensor and the notation T_{00}. By (85·6) its cross-dual is the phase tensor $Z_0{}^0$. The result (90·4), which is now written as

$$T_{\mu 5}{}^{\nu 5} + T_{\mu 0}{}^{\nu 0} = T_{\mu}{}^{\nu} \quad (\mu, \nu = 1, 2, 3, 4), \tag{91·1}$$

connecting it with the conserved energy tensor $T_{\mu}{}^{\nu}$, is subject to the condition that $Z_0{}^0$ contains pressure components only. The momentum components of $Z_0{}^0$ are the 120 dormant components not represented in phase space; these correspond to anti-symmetrical terms in T_{00}. Thus the inhibition of momentum components is the ordinary condition of symmetry, regarded as part of the definition of an energy tensor except in theories which have recourse to non-Riemannian geometry. The exclusion of density components is more surprising. But it is evident from (89·7) that, when Z contains density components, the conserved energy tensor is not contained wholly in T_{00}; it includes also terms taken from $T^{\times}$, i.e. from $Z_0{}^0$ itself.[a] The density part of Z consists of 36 components $Z_{\mu}{}^{\nu}$ with $\mu, \nu = 23, 31, 12, 04, 05, 45$. The exclusion makes no obvious difference to the form of T_{00}; but it relates the components in such a way that the number of degrees of freedom is reduced from 136 to 100.

Systems for which Z consists of pressure terms only provide a field of overlap of relativity mechanics and symbolic mechanics. When density terms are present, the theories bifurcate. Relativity mechanics continues to define the energy tensor by the criterion of conservation, and symbolic mechanics by the criterion of appropriateness of direction; and, since the criteria are no longer equivalent, their formulae cease to be comparable. Ultimately the recalcitrant part of Z—the density part—will be the centre of interest; because the phenomena associated with this part are outside the scope of relativity mechanics. But we must first explore the field of overlap.

A steady distribution of probability is called a *state*. In limiting cases this may be a rest state, but usually it is a state of steady circulatory flow of probability. The state itself is an element of the framework of our analysis; the physical element is an occupied or partially occupied state, obtained by attaching an occupation factor j. The factor j is also interpreted as a probability; so that we have to consider probability distributions over a set of states, as well as the probability distributions which form the states. Physical change is represented by a change of occupation factors, the analytical framework of states being fixed. The simplest element of physical change is a transition or partial transition between two states, i.e. the transfer of a quantity of occupation δj from one to the other. When the transfer occurs continuously we get a flow of probability which may in special cases be a steady circulation. A steady transition circulation involves no physical change, and may form part of the description of a steady system.

[a] Explicit expressions for the diagonal components of $G_{\mu}{}^{\nu} - \frac{1}{2}\delta_{\mu}{}^{\nu}(G - 2\lambda)$ for this case are given in *Protons and Electrons*, formulae (11·58).

A steady circulation of probability between two states will be called an *interstate*. Like the state, the interstate is part of our analytical framework; the physical element is an occupied or partially occupied interstate obtained by attaching an occupation factor. The occupant of an interstate will be called an *oscillator*. In general a steady system will consist of a number of particles occupying states and oscillators occupying interstates; and both particles and oscillators will make contributions to the energy tensor of the system.

The interstate flow of probability must be represented in dimensions outside those of the two states; but since the double frame provides 136 symbolic directions, we may reasonably expect it to provide for this representation. We define a 'simple pure oscillator' to be the carrier of a factorisable momentum vector $P = \psi\chi^*$. (Temporarily the terms vector and tensor, when referring to oscillators, will be courtesy titles having no reference to transformation properties.) We assume provisionally that a simple pure oscillator is physically possible. By definition, the interstate is associated with two states. These must be the pure states specified by the wave vectors ψ and χ; for the factors of P are unique, and there is no other pair of states with which it is possible to connect P. By (72·71) the momentum strain vectors of the pure states are $U = \psi\psi^\dagger$, $V = \chi\chi^\dagger$.

If χ is changed continuously to $\psi^\dagger$, P is changed to U; so that the circulation in U may be regarded formally as an interstate circulation between conjugate states specified by ψ and $\psi^\dagger$, and correspondingly by strain vectors U, $U^\dagger$. State circulation is a special case of interstate circulation, the states being conjugate. The outer product $UU^\dagger$ is an energy tensor characteristic of the state circulation; and, since it is a strain tensor, we may adopt it as the phase tensor Z. The interstate circulation gives correspondingly a characteristic energy tensor $T = PP^\dagger$. This stands for $T_{\alpha\beta\gamma\delta} = \psi_\alpha\chi_\beta\psi^\dagger{}_\gamma\chi^\dagger{}_\delta$; and its cross-dual is $T^\times{}_{\alpha\beta\gamma\delta} = \psi_\alpha\psi^\dagger{}_\beta\chi_\gamma\chi^\dagger{}_\delta$, or $T^\times = UV$.

The notation is more symmetrical if we take $P = \psi\chi^\dagger$, the strain vectors of the states being $U = \psi\psi^\dagger$, $V = \chi\chi^\dagger$ as before. Then the state circulations are described by energy tensors $Z = UU^\dagger$, $Z = VV^\dagger$; and the interstate circulation is described *in a corresponding way* by an energy tensor T whose cross-dual is $Z = UV^\dagger$. It follows that T is a space tensor, and P is a space vector.

The transition flow corresponds to the ordinary classical conception of momentum, which is first introduced as a measure of physical changes we observe. We have here considered the special case of a steady transition flow which gives no physical change; but the unsteady flow of probability from state to state, which represents physical change, will follow the same route and be represented by momenta with the same symbolic directions. It is therefore the interstate flow that is directly represented by a space vector P or space tensor T in the EF-frame anchored in molar physics.

In the energy tensor the transition flow represented by P is coupled with a recoil flow represented by $P^\dagger$. The reason for this will be understood by reference to reality conditions. The vector $\psi\chi^*$ does not in general satisfy the reality conditions (63·3) for a momentum vector in neutral space-time; these limit a factorisable vector to the form $\psi\psi^\dagger E_{45}$ found in (72·72). This means that P cannot be simply superposed on a neutral environment, but will induce a disturbance representable by another vector $Q^\dagger$ superposed on the undisturbed environment. Since the flow represented by P cannot exist independently of the induced or recoil flow represented by $Q^\dagger$, it is sufficient that

the energy tensor $PQ^\dagger$ of the combination shall satisfy the reality conditions. The reality conditions for a tensor must be consistent with those for its cross-dual. Since $U = \psi\psi^\dagger$, $V = \chi\chi^\dagger$ satisfy the reality conditions for strain vectors, $PP^\dagger$ satisfies the reality condition for a space tensor; and it is easily seen that $PQ^\dagger$ will not satisfy the condition unless $Q^\dagger = P^\dagger$. Thus $P^\dagger$ is identified as the momentum vector of the induced effect in the neutral uranoid.

The double frame accordingly pictures the state flow and transition flow as follows. The transition flow is represented by a classical momentum vector P in the E-frame, and there is a recoil flow represented by $P^\dagger$ in the F-frame. The state flows are correspondingly represented by momentum vectors U and V in the C'- and D'-frames, respectively (§ 83). This is one of the cases in which the recognition of two parallel views of world structure corresponding respectively to the EF- and $C'D'$-frames is important. For if the interstate momentum vectors are transformed to the $C'D'$-frame they become general energy tensors; and, although the effects of transition can be duly calculated, we lose the means of picturing it as a circulation similar to that occurring in the states. Naturally in quantum theory where the primary emphasis is on states, not on transition circulation, we mentally interchange the EF- and $C'D'$-frames, so that we think of the frame in which the state circulation is represented as the primary or EF-frame; this, however, is not the frame that has been anchored in molar physics. The picture having served its purpose, we express the relation between states and interstates by means of the cross-dual instead of introducing a $C'D'$-frame explicitly.

We have seen that the quantum momentum vector is classically a strain vector (§ 77); so that the quantum energy tensor is classically a strain tensor. It is therefore to be identified with $Z_0{}^0$. We thus reach a more fundamental view of the quantum classical analogy:

The quantum energy tensor is the cross-dual of the classical energy tensor. (91·2)

The mathematical basis of the quantum-classical analogy is the analogy of the EF- and $C'D'$-frames. We shall now commonly distinguish T_{00} and $Z_0{}^0$ as the classical energy tensor and the quantum energy tensor.

The set of states used in analysis, generally chosen as eigenstates of particular characteristics, may be varied according to the problem that is being considered. In quantum theory the aim is to choose a set of states between which transitions are rare; so that to the first approximation the interstates are unoccupied, and the state energy tensor is the whole energy tensor of the system. This has the form $Z = \Sigma jUU^\dagger$. In the second approximation, an interstate energy $Z = \Sigma jUV^\dagger$, due to small occupation factors j of the interstates, is taken into account as a perturbation energy; that is to say, it is regarded as a mutual energy of the object-system and a perturbing system which induces transitional circulation between the states of the object-system.

The interstates are not wholly unoccupied even in an isolated two-particle system—a hydrogen atom—unless it is in the ground state. In this case the 'perturbing system' must be the aether. For uniformity, perturbations due to extraneous particle systems should be represented as effects transmitted through the aether; so that the aether is always the immediate perturbing agent. The interstate energy tensor is then exhibited as a mutual energy tensor of the strains produced by the object-system in the aether and the rest of the strain system of the aether. If there are no additional strains due

to extraneous particles or radiation, the strain system of the aether corresponds to its own energy tensor, namely the cosmical energy tensor $\lambda g_{\mu\nu}$.

When we adopt the quantum, instead of the classical point of view, the parts played by T and Z become interchanged. In analogous designation Z is a space tensor T', and T is a strain tensor Z'. More precisely $T = Z'_a$, where Z'_a is the tensor representing the strains in the aether; there is also a strain tensor $Z'_u = T'E_{45}F_{45}$ representing strains in the uranoid (§ 86). The identification of the classical tensor T with the aetherial strain tensor accords with the conclusion in the preceding paragraph. Since quantum particles are defined as superpositions on the uranoid, Z'_u is the strain tensor for the ordinary purposes of quantum theory.

92. The recalcitrant terms

The quantum energy tensor Z (also described as the phase tensor) consists of pressure terms Z_p and density terms Z_d. The corresponding parts of the classical energy tensor are $T_p = Z_p{}^\times$, $T_d = Z_d{}^\times$. The suffixes p, d refer to classification in Z; both T_p and T_d include pressure, density and momentum components of T. The important physical difference is that T_p is a combination of recognised tensors of molar physics—the conserved energy tensor and the Riemann-Christoffel tensor—whereas T_d has no such connection with familiar quantities.

A more direct expression of the distinction between T_p and T_d is given in § 89; T_p corresponds to symmetrical and T_d to antisymmetrical double-wave vectors Ψ, X. Allowing for a possible generalisation of wave vectors into wave functions, we have the rule:

T_p corresponds to symmetrical wave functions, and T_d to antisymmetrical wave functions. (92·1)

The result (92·1) is an important junction with current quantum theory; since the need for separating symmetrical and antisymmetrical wave functions is insisted on by all authorities.

If Z_d and Z_p are separated, the states of Z_d are specified by strain vectors of the form

$$U_d = E_{23}S_{23} + E_{31}S_{31} + E_{12}S_{12} + E_{04}S_{04} + E_{05}S_{05} + E_{45}S_{45}, \qquad (92\cdot2)$$

the state energy tensor being $Z_d = \Sigma q U_d U_d{}^\dagger$. Only dormant components of the strain vector occur in U_d. Similarly the strain vectors U_p of the states of Z_p consist of active components only. We can have also interstate energy tensors

$$Z_d = \Sigma q U_d V_d{}^\dagger, \quad Z_p = \Sigma q U_p V_p{}^\dagger$$

representing transitions between the states of Z_d, or between the states of Z_p. But there is no interstate energy tensor of the form $U_d V_p{}^\dagger$, since this is composed of momentum terms of Z, or equivalently of antisymmetrical terms of T. Thus

There are no transitions between states corresponding to symmetrical wave functions, and states corresponding to antisymmetrical wave functions. (92·3)

This also agrees with current theory. It results from the symmetry imposed on the energy tensor by Riemannian geometry. It is not an absolute physical rule, but is an expression of the fact that the agencies (third rank tensors) required to bring about the transitions are too complicated to be represented in a double frame (§ 80). The transi-

tions are 'forbidden', in the usual sense which does not exclude their occurrence in sufficiently complex conditions.

We describe the antisymmetry contained in symbolic factors Ψ, X as *spin antisymmetry*. There is another kind of antisymmetry of wave functions contained in the arguments of the functions, which we describe as *functional antisymmetry*. The coordinates ξ_1, ξ_2, ξ_3 in relative space are antisymmetrical for interchange of the particles m, m'; so that we must regard relative space as inherently antisymmetrical relative to x-space. When interchange is taken into consideration, this leads to the condition $f(-\xi_1, -\xi_2, -\xi_3) = -f(\xi_1, \xi_2, \xi_3)$ for the wave functions of the intracule. In this case the separation of the symmetrical and antisymmetrical wave functions is so complete that they are represented in different spaces and carried by different particles; transitions between them are unthinkable.

We are content here to point out this junction with an important branch of current theory, leaving it to future investigation to decide how much of the current development of the theory of symmetry and antisymmetry can suitably be annexed to fundamental theory.[a] Here we shall use only the two most elementary separations of symmetrical and antisymmetrical states, namely the separation of intracules from extracules and the separation of Z_d from Z_p. In both cases the separation provides a line of demarcation between mechanical and electrical phenomena, but the line is drawn in different ways.

Any kind of division of unified physics is necessarily conventional; and the line of division between the electrical world and the mechanical world is liable to shift according to our point of view. For example, it is a matter of terminology whether non-Coulombian energy is to be counted as electrical or mechanical, or whether a field of electromagnetic potential whose curl vanishes can properly be considered an electrical characteristic. In the most elementary theory mechanical effects are determined by a symmetrical tensor $g_{\mu\nu}$ and electrical effects are determined by an antisymmetrical tensor $F_{\mu\nu}$; and the guiding principle of our terminology is to describe effects associated with antisymmetry as electrical and effects associated with symmetry as mechanical.

Unless there is antisymmetry there is no advantage in employing quantum methods. When there are no intracules, the distribution is a scale-free aggregation of extracules, affording no problems other than those which are within the scope of general relativity theory. When there are no recalcitrant terms Z_d, we can use the classical energy tensor T_p, which has an immediate interpretation in terms of the tensors of general relativity theory. Thus the approach to quantum theory consists in adding a microscopic system consisting purely of intracules or of Z_d tensors to a mechanical system of ideally simple structure. We use a planoid consisting of extracules at rest, or equivalently specified by the energy strain tensor $\alpha E_{16} F_{16}$ (§ 88) which is of the form Z_p. So far as the addition of intracules is concerned, this is the approach used in our earlier chapters. But we have now an alternative kind of addition to the planoid, represented by a recalcitrant energy tensor Z_d, which will be found to lead to important developments, both theoretical and practical.

[a] I think that the most important part that must be annexed or adapted, is the application to 'n-legged intracules', i.e. intracules whose positive ends coalesce in a nucleus. I have not dealt with this problem. In other respects I think this book goes at least as far as the current development, though the results are generally obtained in a different form. The distinction between Fermi-Dirac and Einstein-Bose particles is treated in § 126.

Chapter IX

SIMPLE APPLICATIONS

93. The metastable states of hydrogen

The energies of the ground state and metastable states of hydrogen can be found by a rather simple construction.

The symbolic frame provides two axes normal to space, namely a time axis x_4 and a phase axis x_0. We have found (§ 63) that, when the space-time containing x_1, x_2, x_3, x_4 is neutral, the axis x_0 is space-like, being associated with the achiral symbol $E_{05}E_{16}$. Accordingly the geometrical rotation between x_4 and x_0 is hyperbolic. By applying such a rotation we obtain new axes x_4', x_0', the space axes x_1, x_2, x_3 being unchanged. This rotation will be called a *time-tilt.*

In the usual representation of a system by vectors in space-time, the component p_{05} is stabilised. To obtain an equivalent time-tilted system the stabilised component must be changed to p_{05}'. We must first remove all discriminatory treatment of the x_0 axis, then apply the time-tilt rotation, and then re-introduce discriminatory treatment of the new axis x_0'. By § 57 this means that the time-tilt rotation is to be applied to the vector-density $\mathfrak{P}$, and not directly to P. In $\mathfrak{P}$ the time and phase directions have symbols E_{04}, E_{05}; and the time-tilt rotation is

$$q = e^{\frac{1}{2}E_{45}iu}. \tag{93·1}$$

It is not a relativity rotation; the real relativity rotation in that plane, which is circular, is inhibited by the postulated neutrality of space-time. But the time-tilt gives a new system formally similar to the original system.

Consider an intracule with a momentum vector $E_{45}p_{45} + E_{01}p_{01}$, consisting of an energy $p_{45} = \epsilon$ and a magnetic moment $p_{01} = g\hbar$. The designation is classical, and by (78·7, A 2) ϵ and g are real. Before adding it to an extracule we give it a time-tilt, so that its momentum vector is $P = E_{45}'\epsilon + E_{01}'g\hbar$ in a time-tilted frame E_μ'. The vector-density is given by

$$i\mathfrak{P} = -PE_{05}' = E_{04}'\epsilon - E_{15}'g\hbar. \tag{93·21}$$

Transformed to the untilted frame, this becomes

$$i\mathfrak{P} = E_{04}\epsilon\cosh u + E_{05}i\epsilon\sinh u - E_{15}g\hbar\cosh u + E_{14}ig\hbar\sinh u, \tag{93·22}$$

and the corresponding vector is

$$P = E_{45}\epsilon\cosh u - E_{16}\epsilon\sinh u + E_{01}g\hbar\cosh u - E_{23}g\hbar\sinh u. \tag{93·23}$$

By (78·3) this becomes, in quantum designation,

$$P = E_{16}i\epsilon\cosh u + E_{05}i\epsilon\sinh u + E_{23}ig\hbar\cosh u + E_{14}ig\hbar\sinh u. \tag{93·24}$$

Rewriting (93·24) in the notation

$$P = E_{16}i\mu + E_{05}if + E_{23}ij\hbar + E_{14}i\alpha\hbar, \tag{93·3}$$

we have the relations

$$\epsilon^2 = \mu^2 - f^2, \quad g^2 = j^2 - \alpha^2, \tag{93·41}$$

$$\epsilon/g = \mu/j = f/\alpha. \tag{93·42}$$

Hence $\epsilon/f = g/\alpha$, or
$$\frac{\epsilon}{(\mu^2-\epsilon^2)^{\frac{1}{2}}} = \frac{(j^2-\alpha^2)^{\frac{1}{2}}}{\alpha}. \tag{93·5}$$

According to familiar principles, the tilt may be expected to have discrete eigenvalues determined by quantisation. The terms in (93·24) which admit of quantisation are the two angular momentum components $E_{23}ij\hbar$, $E_{14}i\alpha\hbar$. In order to apply the elementary principles of quantisation it is necessary to refer P to rigid coordinates.

Before adding a free intracule the extracules are placed in a rigid field; and it is in the corresponding rigid coordinates that the whole wave mechanical theory of representation by wave functions and quantisation of angular momentum is developed. By (19·7) the Galilean time is k times the rigid time; so that if u and u_0 are the hyperbolic angles of the time-tilt measured in Galilean and rigid coordinates we have
$$\tanh u_0 = k \tanh u. \tag{93·61}$$

The momentum vector in rigid coordinates corresponding to (93·24) will have the form
$$P_0 = E_{16}i\epsilon_0 \cosh u_0 + E_{05}i\epsilon_0 \sinh u_0 + E_{23}ig_0\hbar \cosh u_0 + E_{14}ig_0\hbar \sinh u_0. \tag{93·62}$$

The constants ϵ_0, g_0, u_0 refer to the undistorted representation of the free intracule, which must be used in investigating the theoretical principles of its structure; but the distorted representation in Galilean coordinates, described by the constants ϵ, g, u, conforms to the observational system and is the officially recognised intracule.

We assume that the quantisation conditions for (93·62) are that the E_{23} component of angular momentum is an integral number of quanta $j\hbar$ and the E_{14} component is one quantum $\hbar$. Thus
$$\tanh u_0 = 1/j \quad (j = 1, 2, 3, \ldots), \tag{93·63}$$
so that $\tanh u = 1/jk$. The E_{23} component, being in a spatial plane,[a] is unaffected by the reduction to Galilean coordinates; but the E_{14} component is divided by k. The rigid field in which the intracule is free corresponds to hydrocules at rest, so that $k = 137$. Thus, for the Galilean vector P, the foregoing quantisation conditions give
$$j = \text{integer}, \quad \alpha = \tfrac{1}{137}. \tag{93·7}$$

Equation (93·5) coupled with the conditions (93·7) is the well-known Sommerfeld formula for the energies ϵ of the metastable states of hydrogen (including the ground state $j = 1$). As the atom approaches ionisation, $j \to \infty$ and $\epsilon \to \mu$; thus μ is identified with the rest mass of the free intracule. Since there is no doubt as to the correctness of Sommerfeld's formula, we conclude that the metastable intracule is actually a time-tilted intracule of the simple kind here discussed.

By (93·63) the ground state corresponds to $u_0 = \infty$. The ground state is a limiting state in which the intracule is time-tilted as far as it will go. The ordinary view of the ground state as the first of a series of similar quantised states, scarcely does justice to its extreme character; one gets the impression that it is a usurper, the rightful bottom state being the 0-quantum state which has failed to materialise. But the state

[a] In comparing this with § 19, we must note that $\mathfrak{T}_\mu{}^4$ is a strain vector, and the covariant momentum vector p_μ which transforms in the same way is the quantum momentum vector—not the classical vector which is contravariant. The permutation of suffixes (78·1) has no effect on E_{23}. It is therefore right to use the quantum designation here.

$u_0 = \infty$ is rock bottom independently of any consideration of quantisation; and the fact that it is also the 1-quantum state is of secondary interest.

The Sommerfeld formula for the unstable eigenstates of hydrogen is

$$\frac{\epsilon}{(\mu^2-\epsilon^2)^{\frac{1}{2}}} = \frac{n+(j^2-\alpha^2)^{\frac{1}{2}}}{\alpha}, \tag{93·81}$$

where n is an integer, or

$$\epsilon/f = (n+g)/\alpha. \tag{93·82}$$

There appears to be no way of generalising the foregoing construction to cover these states. It is not surprising that they elude elementary treatment, because we shall find later that the unstable states, unlike the metastable states, have non-algebraic wave functions—a complication which it is scarcely possible to embody in a geometrical picture.

It will be seen that in this treatment we obtain a representation of the intracule as an integrated whole, in contrast to the later wave-equation treatment which represents it by a density distribution over ξ-space. In this integrated representation ξ-space is not introduced. It will be seen that the momentum vector (93·3) satisfies the reality conditions B 1 for a molar particle in (78·7), not the conditions B 2 given for an intracule. The discussion of reality conditions in § 78 was not intended to apply to integrated representation; in particular, the principle of uniformity of observation, used in deriving the conditions, ceases to be relevant when the intracule is not represented as a distribution over relative space.

94. Neutrium and deuterium

In a high quantum state the time-tilt of the intracule tends to zero, and the momentum vector is $E_{45}\mu + E_{01}j\hbar$ in classical designation or $E_{16}i\mu + E_{23}ij\hbar$ in quantum designation. The determination of the eigenstates of hydrogen is a special problem depending on the theory of the free intracule, so that the energy μ has to be taken equal to the mass-constant of the intracule; but this is a matter of adjustment of the zero-level of energy reckoning, and the intracule as it first appears in the two-particle problem is a bound intracule with energy 0 in a high quantum state. Thus in general theory the only essential part of the intracule is the quantum angular momentum (classical magnetic moment) $E_{23}ij\hbar$, the term $E_{16}i\mu$ being an arbitrary addition made by varying to the adopted zero of energy reckoning. We shall indicate its arbitrary character by a square bracket; so that the quantum momentum vector is $(E_{16}p_{16}) + E_{23}p_{23}$, or more generally

$$P_1 = [E_{16}p_{16}] + E_{23}p_{23} + E_{31}p_{31} + E_{12}p_{12}. \tag{94·1}$$

By (92·2) the momentum vector of a state of the antisymmetrical (recalcitrant) part Z_d of the quantum energy tensor is[a]

$$U_d = [E_{16}p_{16}] + E_{23}p_{23} + E_{31}p_{31} + E_{12}p_{12} + E_{45}p_{45} + E_{50}p_{50} + E_{04}p_{04}, \tag{94·2}$$

where a similar optional term has been inserted, so that the occupant may be treated as bound or free according to the choice of zero of energy reckoning. One half of U_d is

[a] The classical strain vector given in (92·2) is equivalently the quantum space vector. The quantum designation further introduces the suffix permutation 4, 5, 0 → 5, 0, 4 but this does not alter the form of U_d.

representable by a hydrogen intracule. It suggests itself that in general theory another type of intracule will be required to represent the other half, having a momentum vector

$$P_2 = [E_{16}p_{16}] + E_{45}p_{45} + E_{50}p_{50} + E_{04}p_{04}. \tag{94·3}$$

Following the nomenclature in § 81, we distinguish P_1 and P_2 as *spin* and *co-spin intracules*.

In classical designation P_2 becomes $[E_{45}p_{45}] + E_{05}p_{05} + E_{16}p_{16} + E_{04}p_{04}$, and the corresponding vector-density is $\mathfrak{P}_2 = [E_{04}\mathfrak{p}_{04}] + E_{16}\mathfrak{p}_{16} + E_{05}\mathfrak{p}_{05} + E_{45}\mathfrak{p}_{45}$. A time-tilt only changes the values of $\mathfrak{p}_{04}$, $\mathfrak{p}_{05}$ without altering the form of $\mathfrak{P}_2$; and, in fact, no rotation which leaves three-dimensional space invariant can alter the form. Thus when a co-spin intracule is added to an initial distribution of extracules, there is nothing corresponding to the series of eigenstates which were obtained by time-tilting the spin intracule. A simple co-spin intracule analogous to the simple spin intracule in § 93 has only one state. This state may be given different orientations in a co-space whose coordinate planes have symbols E_{45}, E_{50}, E_{04}, just as any one of the metastable states of the spin intracule may be given different orientations in ordinary space; but in neither case does the orientation affect the energy, which here has the symbolic direction E_{16} and is not preferentially related to any orientation either in space or co-space.

The ζ- and θ-algebras are sometimes useful in this connection. By (60·5) we can write

$$P_1 = [a_0 i] + a_1\zeta_1 + a_2\zeta_2 + a_3\zeta_3, \tag{94·41}$$

$$P_2 = [b_0 i] + b_1\theta_1 + b_2\theta_2 + b_3\theta_3. \tag{94·42}$$

The momentum vectors of spin and co-spin intracules are respectively ζ-numbers and θ-numbers.

The combination of a co-spin intracule and an extracule should be a system of the same elementary status as the hydrogen atom which is a combination of a spin intracule and an extracule. This points to a physical identification with the neutron. We shall adopt this identification, which will be confirmed in due course by calculating the mass and magnetic moment (§§ 95, 119).

Since the angular momentum $j\hbar$ (components p_{23}, p_{31}, p_{12}) is by definition zero in a co-spin intracule, the neutron may be regarded as supplying the missing state $j = 0$ of the hydrogen atom. 'A hydrogen atom in the 0-quantum state' is a convenient description of the neutron. When we use it, we must keep in mind that the sequence of quantised states terminates at $j = 1$ as a limit; so that $j = 0$ is a singular state, extra to the sequence. In the equivalent representation by a proton and electron, the distinction is described as a difference of *binding* of the electron. We accordingly recognise two kinds of binding. Nuclear (or neutron) binding corresponds to a co-spin intracule and atomic binding to a spin intracule.

It is recognised that a neutron can transform into a hydrogen atom, and therefore into a proton and electron. To deny that a neutron is composed of a proton and electron is as pedantic as to deny that water is composed of oxygen and hydrogen. In either case the dissociation involves emission of energy and other components of the complete energy tensor. We may, however, distinguish between an emission whose energy tensor consists wholly of Z_p terms and an emission which includes Z_d terms. When quantum theory is approached in the way described at the end of § 92, so that its application is

directed to Z_d systems superposed on a Z_p environment, the former emission merges at once into the Z_p of the environment, whereas the Z_d emission has recalcitrant properties which can only disappear by forbidden transitions between antisymmetrical and symmetrical states. It is clear that the change of momentum vector in the conversion of a co-spin into a spin intracule corresponds to a Z_d energy tensor. There is therefore a justification for taking note of the corresponding emission, which has been called a neutrino.

Two systems which differ only in the Z_p part of the energy tensor will be said to be *modularly equivalent*. The idea is that the conversion of one into the other would involve changes of energy, momentum, angular momentum, etc., which we are accustomed to accept without comment as the ordinary accompaniment of transitions, the balance being provided by or taken up by the Z_p environment. Correspondingly, for systems described by momentum vectors, modular equivalence means that they differ only in the symmetrical part U_p; for in the energy tensor cross-terms of the form U_pV_d are dormant. We have allowed for a very limited form of modular equivalence by the addition of a bracketed energy term in the formulae in this section. The general modular equivalent would be obtained by substituting $[U_p]$, where U_p is an arbitrary symmetrical matrix.

By transferring the centre of interest from the mechanical to the electrical world, we invert our previous ideas of activity and dormancy. What was formerly looked upon as the dormant part of P is now the most active part, and the former active part has become dormant. Thus an alteration of U_p (which has now become the dormant part) is treated as having no classificatory significance; its effects are not in the half of the world that we are investigating. In § 63 we pointed out that for the experimental measurement of P_e (which has become U_d in quantum designation) it is necessary to have an extraneous electrical standard represented by i, the mechanical standard being taken as 1. The electrical standard now becomes the primary standard; and indeed a separate mechanical standard is unnecessary, mechanical and electrical quantities being distinguished as having even and odd dimension index in terms of the electrical standard. When the electrical standard is substituted for the mechanical standard the vector P is changed to $-iP$. It would be confusing to have different notation, nomenclature, standards, etc., for the mechanical and electrical worlds, since ultimately we have to consider them in conjunction. So we must be content to make clear the change in our point of view, without attempting to embody it formally in the mathematics.

The E-frame is the outer product of a ζ-frame and θ-frame; and similarly the F-frame is the outer product of a ζ'-frame and a θ'-frame. The ζ- and ζ'-frames are anchored similarly with respect to the geometrical axes x_1, x_2, x_3. The θ- and θ'-frames are also anchored either similarly or contragrediently (according as the E- and F-frames are cogredient or contragredient) in molar physics. The combination of two particles represented by momentum vectors is multiplicative, and the combined system is characterised by the outer product of the momentum vectors. Thus the combination of two spin intracules is represented by a product of ζ- and ζ'-numbers, which is an EF-number; and a combination of two co-spin intracules is represented by a $\theta\theta'$-number which is an EF-number. But we have a choice of two representations of a combination of a spin and a co-spin intracule; it is represented either by a $\zeta\theta'$-number which is an

EF-number, or by a $\zeta\theta$-number which is a pure E-number. By choosing the latter representation we enter on a radically new development.

Consider the two free intracules (94·41), (94·42), and let their momentum vectors be normalised so that $a_0, b_0 = 1$. For the outer product $P_1 \times P_2$, which in this case coincides with the ordinary product by multiplication of E-symbols, we have

$$P_1 \times P_2 = i(P_1 + P_2) + [Q], \tag{94·5}$$

where Q consists of symmetrical matrices only. Thus multiplicative combination is modularly equivalent to additive combination. The intrusion of the factor i is avoided by writing (94·5) in the form $(P_1/i) \times (P_2/i) = (P_1 + P_2)/i - [Q]$; that is to say, the momentum vector P/i expressed in electrical units should be used in this connection, as foreseen above. Not only is a combination of a spin and co-spin intracule represented by an E-number, but the E-number represents them additively.

The combination of a spin intracule and a co-spin intracule will be called a double intracule (abbreviated as D.I.). The momentum vector of a bound D.I. has the general form

$$U_d = E_{23}p_{23} + E_{31}p_{31} + E_{12}p_{12} + E_{45}p_{45} + E_{50}p_{50} + E_{04}p_{04}, \tag{94·6}$$

so that the occupants of the antisymmetrical states which contribute the recalcitrant part Z_d of the energy tensor are D.I.'s. The D.I. has therefore fundamental importance in our analysis. If symmetrical terms Q are added to make it resolvable (multiplicatively) into two free intracules, the resulting D.I. may be described as 'free'. But such terms are, from the present point of view, dormant; and they have no classificatory significance in defining the states of the D.I.

The double intracule must be added to a double extracule; but since extracules are scale-free, the change from a single to a double extracule is merely a matter of doubling the occupation factor. The resulting system is equivalently composed of two protons and two electrons, one electron having atomic binding and the other nuclear (or neutron) binding. We recognise this system as a deuterium atom.

A deuterium atom is a combination of a double intracule and a double extracule. (94·7)

If the two protons and two electrons are arranged as a hydrogen molecule, only spin intracules are introduced; but there are *four* of them, since there are four sets of relative coordinates ξ_α joining a proton and electron. The same difficulty would arise if both electrons had nuclear binding. The complexity of these systems agrees with our analytical result that the two intracules must be of opposite kinds in order to form a simple linear combination. This result is the starting point of nuclear theory.

95. Mass of the neutron

We shall determine the mass of the neutron by the scale-free method developed in Chapter II. In this method the density of a uniform distribution of neutrons at almost exact rest is compared with the density of a corresponding distribution of hydrogen atoms. The calculation depends on a knowledge of the multiplicity factors of the particles. We are concerned particularly with the multiplicity of the hydrogen (spin) intracule and the neutron (co-spin) intracule.

The intracule was introduced as a carrier of transition energy and its standard mass μ was determined accordingly. Unlike the other masses introduced, which were rest

masses, μ is a mass-constant; as such it is independent of the multiplicity of the intracule. This is shown explicitly in (22·7), where the formulae for the masses of the proton, electron and extracule involve their multiplicity factor 10, but there is no term introducing the multiplicity of the intracule. When the intracule is freed it is given an initial energy equal to μ, and it is then that μ becomes associated with a definite multiplicity.

In the two-particle transformation it is postulated that the intracule has the same multiplicity 10 as the other particles. For the two sides of the equation correspond to two ways of resolving a double-wave function into two simple wave functions, and the intracule is therefore represented as of the same formal type as the other particles. The argument is made more precise by considering phase spaces. The probability density of the bi-particle in 6-space is specified in § 26 by a double-wave function Ψ which is invariant for the two-particle transformation. Since the distributions are pseudo-discrete, an individual Ψ corresponds to the states comprised in a small element of the combined phase space of the proton and electron, and after the transformation to a small element of the combined phase space of the extracule and intracule. It is therefore postulated that the two double phase spaces are divisible into elements which can be put into one-to-one correspondence. This means that they must have the same number of dimensions. Since each of the other particles contributes 10 dimensions, the intracule contributes 10 dimensions.

Accordingly, considering hydrogen atoms in Galilean coordinates, the free intracule is a particle of rest mass μ and multiplicity 10; and μ then forms part of the initial energy of the distribution. The intracule is represented as a particle distributed over ξ-space; and its antisymmetrical character is contained in its coordinates ξ_α which are reversed in sign by interchange of the proton and electron. We have distinguished this as functional antisymmetry (§ 92). Another representation of the hydrogen intracule has been found with spin antisymmetry instead of functional antisymmetry. This is necessarily an integrated representation (integrated over ξ-space), since the two kinds of antisymmetry would neutralise one another if the coordinates were not eliminated. The momentum vector of a state then has the form (94·1), and has 4 components (in a time-tilted frame); and if the intracule were constrained to be in this state its multiplicity would be 4. The actual multiplicity 10 is due to the fact that the hydrogen intracule has freedom of distribution over the whole series of metastable states, and in addition over unstable states not yet treated.

We are going to apply a formula which is only valid for scale-free distributions, and must therefore confine our attention to states of very high quantum number for which the fixed-scale effects of quantisation are evanescent. The time-tilt is then negligible, and the energy is the initial energy μ. Considering one such state, we shall calculate the energy of a particle constrained to occupy this state. The energy is increased by the constraint which inhibits transition to other states, or, since the high quantum states merge into a continuum, inhibits 6 of the 10 degrees of freedom. The constraint is a stabilisation applied to the momentum vector, reducing the multiplicity from 10 to 4; and, since μ is initial energy, we can apply the usual formula (16·5). The energy μ_4 of the constrained intracule is

$$\mu_4 = \tfrac{10}{4}\mu. \tag{95·1}$$

If such a constrained intracule exists naturally, it will be recognised as a particle which remains in one state of its own accord. The co-spin intracule is such a particle,

since it has only one state. Moreover, it is scale-free since the momentum vector is not controlled by quantisation. The difference of the momentum vectors of the co-spin and spin intracules can be regarded as the momentum of a constraint incorporated in the constitution of the former which nullifies 6 degrees of freedom. By the scale-free theory the energy of a stabilising constraint is determined solely by its effect on the multiplicity; and (95·1) applies equally to a V_4 intracule created artificially by constraining a V_{10} intracule to one metastable spin-state, or a natural V_4 co-spin intracule self-constrained to one state because it has no other. The only limitation is that the energy of the system must not (in the states considered) be dependent on quantisation, since (16·5) is expressly derived from the scale-free condition. For a co-spin intracule the removal of the constraint, so as to allow transition to spin-states, is the process commonly called absorption of a neutrino.

We can accordingly identify the mass μ_n of the co-spin intracule with μ_4; and the difference of the masses n', H' of a neutron and hydrogen atom is[a]

$$n' - H' = \mu_n - \mu = 1{\cdot}5\mu. \tag{95·2}$$

For complete accuracy it is necessary to consider β-factors. In (29·5) the reduction of the ratio M/μ from system B to system A introduced a factor $\beta^{-\frac{5}{6}}$ resulting from the different transformation law of the optically controlled mass μ and the molarly controlled mass M. This factor is now irrelevant since the mass of the neutron is defined in the same way as that of the hydrogen atom by molar control; and the energy (95·2) transforms in the same way as M. We therefore employ here (and throughout this chapter) the standard mass

$$\mu = 10M/136^2. \tag{95·3}$$

In atomic weight units the mass M of a hydrogen atom is 1·008130; this gives

$$\mu = 0{\cdot}00054505 \text{ a.w.u.} \quad \text{(standard mass)}. \tag{95·4}$$

Having removed the irrelevant factor we have still to consider whether (95·2) requires further correction. It will be found in § 98 that the additional energy has to be multiplied by β; so that the final value is

$$n' - H' = 1{\cdot}5\beta\mu = 0{\cdot}0008236. \tag{95·5}$$

The present observational value $0{\cdot}00082 \pm 0{\cdot}00003$ is too inaccurate to give more than a rough check.

96. Double intracules

A factorisable double-wave vector $\Psi_{\alpha\beta} = \psi_\alpha \chi_\beta$ specifies a combined state of two pure particles whose states are specified separately by ψ, χ. The double-wave vectors $\Psi^s = \frac{1}{2}(\psi_\alpha \chi_\beta + \chi_\alpha \psi_\beta)$, $\Psi^a = \frac{1}{2}(\psi_\alpha \chi_\beta - \frac{1}{2}\chi_\alpha \psi_\beta)$ are not factorisable; they specify two sets of states, but the occupants of these states are not resolvable into two independent particles. We commonly regard unit occupation of Ψ as equivalent to unit occupation

[a] This result was first obtained in *Proc. Roy. Soc.* A, **174**, 42, 1940. But I owe to H. O. W. Richardson the suggestion that μ_n/μ was likely to be the factor $\frac{10}{4}$ (which had already occurred in my theory in connection with magnetic moments, *Protons and Electrons*, § 12·8). Richardson himself had partially developed this suggestion in a paper accepted for publication by the Royal Society, which he afterwards withdrew.

of each of the states Ψ^s, Ψ^a. It is true that wave vectors are not in general additive; and if we form from them energy tensors which *are* additive,

$$(\Psi^s + \Psi^a)(\Phi^s + \Phi^a) \neq \Psi^s\Phi^s + \Psi^a\Phi^a.$$

But the difference $\Psi^s\Phi^a + \Psi^a\Phi^s$ is composed entirely of momentum terms of Z, and these are dormant in standard conditions. Thus, unless there are present perturbing conditions of the complicated kind necessary to induce forbidden transitions between symmetrical and antisymmetrical states, the combination of two particles in the states ψ, χ is equivalent to the combination of two particles in the states Ψ^s, Ψ^a.

The particles occupying Ψ^s and Ψ^a are not independent; for it is only when the occupation factors are equal that the combination represents a distribution of simple particles. If in a pseudo-discrete assemblage the occupation factor of Ψ is j, the occupation factors j_1, j_2 of ψ, χ may be any numbers such that $j_1 j_2 = j$; but the occupation factors of Ψ^s, Ψ^a are necessarily $\sqrt{j}$. Calling, provisionally, the occupants of Ψ^s, Ψ^a extracules and intracules, occupation of an intracule implies equal occupation of a corresponding extracule. The rule that there must be one extracule for each intracule, obvious in the elementary case of hydrogen extracules and intracules, is thus extended to extracules and intracules separated by spin antisymmetry. The new extracules and intracules occupy double-wave vectors and are therefore bi-particles—double extracules and double intracules. By §§ 89, 92, the Z_p part of the energy tensor comes from the Ψ^s double vectors, and the Z_d part from Ψ^a; so that Z_p is the energy tensor of the double extracules, and Z_d of the double intracules.

We have now two fundamental modes of analysis of an extended energy tensor:

(*a*) The element of energy tensor corresponds to a standard particle, consisting of a simple extracule and intracule, or equivalently a proton and electron.

(*b*) The element is divided into Z_p and Z_d, corresponding respectively to a double extracule and double intracule, or equivalently two protons and two electrons.

In (*b*) the separation of Z_p and Z_d is stabilised. The absence of Z_d components in the energy tensor of a double extracule is a matter of definition, not observation; for if such components were observed we should detach them from the extracule and a double intracule would be introduced to carry them. Thus the double extracule is a V_{100} bi-particle, and the double extracule is a V_{36} bi-particle. Except when modified by dormancy, the multiplicity of a bi-particle is the square of the multiplicity of a corresponding simple particle. Thus in (*b*) the simple extracule and intracule are V_{10} and V_6 particles; they correspond respectively to the 10 mechanical and 6 electrical components of the momentum vector. The simple extracules in (*a*) and (*b*) are identical; but the intracules introduced in (*a*) to represent the electrical part of the world are entirely unlike the V_6 intracules in (*b*). They correspond one-to-one, since in either analysis there is one intracule per extracule.

The primary, but not the only, difference is the multiplicity factor. We have seen (§ 95) that the (*a*) intracules of mass μ are V_{10} particles; if by stabilisation these are transformed to V_6 intracules, the mass becomes

$$\mu_6 = \tfrac{10}{6}\mu, \tag{96·1}$$

so that the density of the electrical part of the world is increased in the ratio $\frac{5}{3}$ relatively to that of the mechanical part of the world which is unchanged. But since it is the

intracules that supply our quantum-specified standards, it would be more logical to say that the density of the mechanical part of the world is decreased in the ratio $\frac{3}{5}$ relatively to that of the electrical part of the world which remains constant.

When we passed from sub-threshold to super-threshold theory we multiplied the density (then wholly mechanical) by the selection factor $\frac{5}{3}$ before introducing any electrical characteristics (§ 42). The transformation to V_6 intracules just removes this selection factor. In § 16 we showed that the multiplicity factors in gravitational theory are in fact top-particle selection factors; and if the relative multiplicity factor $\frac{10}{6}$ is exhibited in this way, it is seen to be not merely equal to but identical with the selection factor in sub-threshold analysis. The relation between analyses (*a*) and (*b*) is now apparent; (*a*) is super-threshold theory, (*b*) is sub-threshold theory. Equivalently, (*a*) is gravitational theory, (*b*) is exclusion theory.

We see at once a number of details that fit in with this conclusion. The exclusion treatment introduces a unit cell occupied by two protons and two electrons; this corresponds to an element of energy tensor (*b*). When exclusion supersedes gravitation, so that there is no gravitational interaction between the energy tensors of the particles, energy tensors become *physically* additive; that is to say, the mathematical additivity and separability of energy tensors (which is imposed by our method of analysis) corresponds to physical juxtaposition and separation. This is presupposed when we divide Z into $Z_p + Z_d$, and interpret Z_p, Z_d as energies of a mechanical and an electrical system without reserving any part of Z to represent energy of interaction between them. The whole theory of Z, based on its cross-dual relation to T, identifies it with the aether-strain tensor, not the uranoid-strain tensor, so that it applies to a world built up from nothing as in sub-threshold theory, and not added to a uranoid. Moreover, in recognising a 'recalcitrant' part Z_d of the energy tensor, we definitely broke away from gravitational theory in which the definitions are such that the law of conservation is satisfied identically; so that inevitably Z_d takes us back to the pre-gravitational sub-threshold theory. Finally, the close association of exclusion with the separation of symmetrical and antisymmetrical wave functions is familiar in current theory.[a]

In passing from analysis (*a*) to sub-threshold theory, the change of the multiplicity of the simple intracule to 6, and of the double intracule to 36, is automatic. It is the double intracule that is physically important. Mathematically it is divided into twin V_6 intracules; but these are not to be taken apart. For if one of them is put by itself in the standard environment, its momentum vector U_d becomes dormant. We already know that the physical division, when it occurs, is unsymmetrical, the constituents being a spin and a co-spin intracule. This rejection of the V_6 intracule as a separable particle is not inconsistent with the use that we made of the U_d momentum vector in § 94. That was a vector connected with the D.I. by modular equivalence, and used to exhibit its *composition*; it did not profess to represent its energy and momentum.

As usual we have left to the end the investigation of β-factors. Such a factor arises, because in analysis (*a*) we introduce an extra degree of freedom of the element of energy tensor in order to provide for electrical properties. This is not required in

[a] But it is misstated. There is exclusion between symmetrical wave functions precisely like the exclusion between antisymmetrical wave functions; but as shown in Chapter IV it is currently replaced by inertial-gravitational interaction.

analysis (*b*) where the electrical properties are contained in the recalcitrant part of the energy tensor. Thus the division of densities, and of the masses which transform as densities, by β, which causes so much trouble in analysis (*a*), does not occur in (*b*). That leaves the mass unit for simple particles β times greater in (*b*) than in (*a*); and for double particles the factor becomes β^2. In particular we have to multiply the mass of the D.I. by β^2 in order to express it in the mass unit of analysis (*a*) which has been connected with observational measurement.

We have found that the mass of the standard particle calculated by super-threshold theory in (40·7) and by sub-threshold theory in (43·7) is the same. This verifies that there is no initial difference of mass-unit between the two analyses, such as would compensate the foregoing β^2-factor. The fact is that the difference of rest mass of the standard particle and the hydrocule should have been taken into account in calculating the various fundamental constants from observational data. Since it has not been taken into account we are saddled with a system of constants, 'right' by custom, but wrong from the theoretical standpoint; so that each new development introduces fresh β-factors, which have no theoretical right to be there, but are necessary to cover up the initial mistake in the system of constants.

97. Comparison with field theory

The analyses (*a*) and (*b*) may suitably be distinguished as *linear* and *quadratic* analysis. In order to employ them in the routine of practical calculation, we need a simplified working conception of the way they are related. Linear and quadratic analysis occur also in field theory; and it is helpful to examine the field treatment, since their connection there is more easily grasped.

In field theory mechanical characteristics are specified by a symmetrical tensor $G_{\mu\nu}$ with 10 components and electrical characteristics by an antisymmetrical tensor $F_{\mu\nu}$ with 6 components. This has fairly close correspondence with the representation in particle theory of mechanical characteristics by the 10 symmetrical components and electrical characteristics by the 6 antisymmetrical components of the quantum momentum vector. The combination[a]

$$H_{\mu\nu} = G_{\mu\nu} + iF_{\mu\nu} \tag{97·1}$$

corresponds in this way to the extended momentum vector. It occurs naturally and fundamentally in affine field theory. Linear and quadratic analysis depend respectively on two action-invariants

$$K_1 = H = H_{\mu\nu} g^{\mu\nu} = G, \tag{97·21}$$

$$K_2 = H_{\mu\nu} H^{\mu\nu} = (G_{\mu\nu} + iF_{\mu\nu})(G^{\mu\nu} + iF^{\mu\nu}) = G_{\mu\nu} G^{\mu\nu} - F_{\mu\nu} F^{\mu\nu}. \tag{97·22}$$

The first is the usual action G, employed in mechanical theory. The fact that $F_{\mu\nu}$ contributes nothing to G expresses the dormancy of electrical components in linear analysis. The second equation is more appropriately written

$$G_{\mu\nu} G^{\mu\nu} = H_{\mu\nu} H^{\mu\nu} + F_{\mu\nu} F^{\mu\nu}. \tag{97·23}$$

For the energy tensor derived from G satisfies the law of conservation identically; so that, when electrical energy is introduced, it must be included in $G_{\mu\nu}$. Equation (97·23)

[a] The factor i is required if $G_{\mu\nu}$ and $F_{\mu\nu}$ are real.

corresponds to $Z = Z_p + Z_d$ in particle theory. The absence of cross-terms between $H_{\mu\nu}$ and $F_{\mu\nu}$ corresponds to the absence of momentum components of Z.

The usual electromagnetic action $F_{\mu\nu}F^{\mu\nu}$ appears in (97·23) but it is coupled with $H_{\mu\nu}H^{\mu\nu}$ instead of with the usual mechanical action $H = G$. We have especially to consider the relation of $H_{\mu\nu}H^{\mu\nu}$ to H. The element in quadratic analysis is a double element which we think of as a combination of two simple elements represented by the two factors of K_2. One of these taken by itself should form the simple element of linear analysis. 'By itself' means that aether is substituted for the second element. For the aether $G^{\mu\nu} = \lambda g^{\mu\nu}$, $F^{\mu\nu} = 0$, so that $\lambda g^{\mu\nu}$ must be substituted for the second factor $H^{\mu\nu}$ in K_2. The result is λH. Thus in order to use the two analyses concurrently without an adjusting factor when we pass from one to the other, we must adopt units such that the cosmical constant $\lambda = 1$. Then, so far as mechanical characteristics are concerned, the element of quadratic analysis is simply a duplication of the element of linear analysis; but for the investigation of electrical characteristics the double element has to be kept intact.

All these results have their counterpart in particle theory. The double extracule splits up into two simple extracules; but the double intracule is inseparable. The aether-strain tensor Z is used, so that the simple and double elements are added to the aether, not the uranoid. But the next step is peculiar to field theory which is scale-free, and it would not be permissible to apply it to the corresponding particle analysis. The electrical action $F_{\mu\nu}F^{\mu\nu}$ of the double element is assigned half to each simple element; so that we have a linear analysis with action invariant

$$G = H + \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu}. \tag{97·3}$$

These formulae are in units such that $\lambda = 1$. Consider a change of unit such that $\lambda = \beta^{-1}$. This is a gauge transformation in general relativity theory. The fundamental tensor $G_{\mu\nu} + iF_{\mu\nu}$ derived from affine field theory is by its nature invariant for uniform gauge transformations so that $G_{\mu\nu}$, $F_{\mu\nu}$, $H_{\mu\nu}$ are unaltered. The transformation is therefore $g_{\mu\nu} \to \beta^{-1}g_{\mu\nu}$; so that $g^{\mu\nu} \to \beta g^{\mu\nu}$. The invariants (97·23) and (97·3) are then transformed to

$$\beta^2(H_{\mu\nu}H^{\mu\nu} + F_{\mu\nu}F^{\mu\nu}), \tag{97·41}$$

$$\beta(H + \tfrac{1}{2}\beta F_{\mu\nu}F^{\mu\nu}). \tag{97·42}$$

In field theory this is simply the removal of the restriction to special units, and β is an arbitrary constant. But in particle theory this transformation of the unit, with $\beta = 137/136$, is introduced as explained at the end of the last section. It is useful to exhibit the transformation in two stages as in (97·42). One factor β is engaged when we replace the D.I. (corresponding to $F_{\mu\nu}F^{\mu\nu}$) by two simple intracules (spin and co-spin) additive to the two extracules which correspond to H. The other factor β represents the change of unit of the whole linearised energy.

98. Mass of the deuterium atom

We define the 'upper state' of the deuterium atom to be that in which it is disintegrated into two protons and two electrons at rest, and the 'lower state' to be that in which it is a deuteron and electron at rest. The difference of energy of the two states is the mass-defect of deuterium.

In the upper state a pseudo-discrete distribution of the deuterium atoms is indistinguishable from a distribution of hydrogen atoms, and the same quantum energy tensor Z applies to both. The classical energy tensor for the rest state is $E_{45}F_{45}\rho$; its cross-dual Z contains both Z_p and Z_d components. If the distribution is regarded as hydrogen, Z is left unseparated; it is then the phase tensor of one standard particle, and the resolution into an extracule and intracule, or into a proton and electron, proceeds by linear analysis. If the distribution is regarded as deuterium, Z is separated into Z_p and Z_d, and the separation is stabilised; the two parts are the phase tensors of a double intracule and double extracule in quadratic analysis. The double extracule is equivalent to single extracule with occupation factor 2 in linear analysis.

We have to consider transitions from these limiting states. If in quadratic analysis an arbitrary transition energy δZ is added this has to be divided into symmetrical and antisymmetrical parts δZ_p, δZ_d added to Z_p, Z_d respectively. But this destroys all connection with the separation of intracules and extracules in linear analysis where the extracules are incapable of carrying transition energy. The representation of transition energy in this way is doubtless admissible, but it leads nowhere. This is not surprising because quadratic analysis is sub-threshold analysis, and it is inappropriate to introduce transitions at the sub-threshold stage. We accordingly begin the treatment of transitions in another way.

We take the simple extracule of deuterium and provide for its transitions, precisely as we did in the case of hydrogen, by adding an intracule whose antisymmetry is functional. Everything is the same as in § 18, including the masses m_0, μ; the change from the hydrogen to the deuterium extracule is a change of the occupation factor of the standard particle, and does not affect the specification of its states. The functional antisymmetry of the intracule adds an interchange degree of freedom, changing the standard particle to a hydrocule, and liberating the energy required to free the intracule. We have accordingly an assemblage of doubly occupied hydrocules each with a free intracule of initial energy μ. Both particles at this stage have multiplicity 136; and the intracule can take up as transition energy a complete energy tensor of an unrestricted kind. A suitable addition of transition energy must bring it to the upper state, where it will have an equivalent representation as the double intracule of quadratic (sub-threshold) analysis.

Fortunately there is no need to inquire into the relation between the two representations. The D.I. is a V_{36} particle, so that a stabilising constraint must be applied reducing the number of degrees of freedom from 136 to 36. The transition energy required to be added is the energy of this stabilising constraint. This energy then becomes rest energy of the D.I. which incorporates the constraint in its structure. The energy of the V_{36} particle is μ_2, where μ_2 is given by the usual formula

$$\mu_2/\mu = 136/36. \tag{98·1}$$

It will be seen that the method is precisely similar to that by which we found the mass of the neutron.

The energy $\mu_2 - \mu$ is lost when the atom makes a transition from the upper state in which the intracule is a D.I. to the state in which it is a simple intracule at rest. The latter is evidently the lower state defined above, in which transition energy is carried

by a simple intracule association with the relative coordinates of a deuteron and electron. Leaving out β-factors for the moment, the mass-defect of deuterium is

$$2H - D = \mu_2 - \mu = 100\mu/36 = 0{\cdot}001514 \text{ atomic weight units.} \qquad (98{\cdot}2)$$

As pointed out at the end of § 96, the energy of the double intracule has to be multiplied by β^2 to reduce to the recognised system of constants. The precise mode of insertion of β-factors in (98·2) is found at once by using (97·42) as a guide. The corrected formula is accordingly

$$2H' - D' = \beta(\beta\mu_2 - \mu) = \tfrac{137}{136}\tfrac{101}{36}\mu = 0{\cdot}0015404 \text{ a.w.u.} \qquad (98{\cdot}3)$$

The observational value (Mattauch) is $0{\cdot}001539 \pm 0{\cdot}000002$.

A question arises whether we ought not to use in (98·3) the mass μ_d of the deuterium intracule rather than the mass μ of the hydrogen intracule. This would increase the foregoing result to 0·0015408. It should be remembered, however, that μ_d is a mass-constant; and, although spectroscopic investigations introduce a free intracule having μ_d as rest mass, there does not appear to be any reason for introducing a rest mass μ_d in the present investigation. I therefore incline to the view that (98·3) is correct; but I am unable to answer the question definitely.

The β-factor, which was inserted in anticipation in (95·5) in determining the mass-excess of the neutron, has the same origin. Put in a form corresponding to (98·3) the formula is

$$n' - H' = \beta(\mu_n - \mu). \qquad (98{\cdot}4)$$

Remembering that μ_2 refers to a double intracule and μ_n, μ to simple intracules, we see that the β-factors in (98·3) and (98·4) are introduced consistently.

99. Mass of the helium atom

An ideal system composed of two deuterium atoms rigidly coupled so that their spins and co-spins are anti-parallel will be called a 'balanced atom'. The rigid coupling is a stabilising constraint which reduces the number of degrees of freedom. When it is imposed as an arbitrary constraint, the energy of the system is changed by the energy of the constraint. We shall find that this change can be calculated without difficulty. If there exists a natural state of two deuterium atoms in which the forces of interaction themselves lock the spins in this balanced arrangement with almost perfect rigidity, the energy of the interaction will be the same as the energy of the stabilising constraint. The method therefore gives a short cut to the calculation of the interaction energy of such a state. We have reason to think that the helium atom is a state of the eight-particle system fulfilling this condition to a very high approximation, subject to a correction for the obvious non-rigidity of the distribution of the two satellite electrons.

We shall calculate the mass of a helium atom by this short cut. The long way round is to form quantum equations for the eight particles, including both Coulombian and non-Coulombian interactions, and solve them to determine the steady states that can exist; the energy of the lowest state is the quantity we are seeking. The short cut assumes that the non-rigidity of the spin-coupling in this state is negligible. We shall, however, be able to find a limit to the possible correction for non-rigidity, and show that the theoretical balanced atom is a very close approximation to the natural helium atom.

Consider an assemblage of balanced particles in a pseudo-discrete state. Superficially,[a] it appears to be a deuterium distribution whose probability is equally divided between two pseudo-discrete states with opposite spin and co-spin; but mechanically it differs, because the constraint forbids variation of the occupation of one state without an equal change of occupation of the other. The fact that they have only a joint-occupation factor makes them parts of one state. The spin and co-spin parts U, $-U$ of the momentum vectors of two deuterium atoms form equal energy tensors $Z_d = UU^\dagger$, which are additive. But in the balanced particle U and $-U$ are included in one state and systematically cancel one another in the resultant momentum. The constrained balance of spins and co-spins is expressed by omitting the Z_d components of the energy tensor, the corresponding degrees of freedom being inhibited.

In wave mechanics the distinction between two states with independent occupation factors and two parts of a state with only a joint-occupation factor is represented by incoherent and coherent phase of the corresponding wave functions. The waves of the two deuterium atoms have incoherent phase and do not interfere; but, when coupled into a balanced particle, the waves are made coherent and interference occurs.

We have accordingly an upper state of the assemblage in which the eight-particle system is two deuterium atoms (in their upper state, i.e. disintegrated into protons and electrons) or equivalently four standard particles; and a lower state in which coupling of spins and co-spins has eliminated the electrical or intracule part Z_d of the energy tensor. The result is obtained by quadratic (sub-threshold) analysis. Changing over to linear (super-threshold) analysis, the difference between the two states still consists in the elimination of intracule energy; so that the four standard particles become four hydrocules (or more formally a standard particle state with occupation factor 4 is transformed into a hydrocule state of occupation factor 4). Accordingly the transition from the upper to the lower state causes a loss of $\frac{1}{137}$ of the upper energy. As the upper state is that in which the atom is disintegrated into 4 protons and 4 electrons its energy in the observational system is $4M$; and the mass-defect of a balanced atom is

$$4M/137. \qquad (99{\cdot}1)$$

In comparing this investigation with §§ 28, 29, it is to be noticed that here the electrical energy is *extinguished* by balancing; previously it was *reserved* for intensive treatment. We cannot get rid of the electrical energy of a distribution of protons and electrons merely by dropping the intracules, because equality of the number of extracules and intracules is an essential feature of the representation. In Chapter III we separated the electrical energy for specialised treatment but compensated the removal by a change of unit which multiplied the remaining energy by β; but here the electrical energy is made zero by a balancing of the intracules, and there is no compensating factor β. Thus the energy is β^{-1} times that of the four hydrogen atoms. The intracules containing the spin and co-spin momenta (corresponding to $F_{\mu\nu}$ in field theory) are double; so that balancing is only possible when there are at least eight simple particles

[a] Each proton and electron has even probability distribution over the whole extent of the assemblage; and we shall not notice correlations of position unless the search for correlations of coordinates forms part of our method of investigation. In the short cut, coordinates are not introduced; and, surveying the distribution from this point of view, there is no particular reason to notice the aggregation of groups of particles into compact nuclei.

in the system. With helium there begins a new possibility, unprovided for in any less complicated system.

The satellite electrons of helium provide a 'two-legged intracule' to carry transition energy. This is a doublet with one positive charge $2e$ and the negative charge split into two parts $-e$. This is intermediate between one and two intracules of the hydrogen type, and its mass will be intermediate between μ and 2μ. It is fairly evident that the mass is $\mu\sqrt{2}$; and this is verified by reference to the corresponding theory of the interchange of extracules in § 47. We there found that, when one particle is interchanging simultaneously with a number of particles, the squares of the momenta of the separate interchange circulations are additive; by calculating the resultant momentum according to this quadratic law we obtained the same masses of the extracules that had been found by other methods. Here the two legs of the intracule correspond to two interchange circulations, and the resultant momentum vector has therefore a scale $\sqrt{2}$ times that of a simple intracule.

To obtain the helium atom we must not eliminate the electric energy wholly, as in the balanced particle, but must leave enough to provide a free two-legged intracule. The calculated mass-defect of helium is accordingly

$$4H - \mathrm{He} = \frac{4M}{137} - \mu\sqrt{2} = 0{\cdot}02867 \text{ a.w.u.} \tag{99.2}$$

In terms of μ, the calculated mass-defect (99·2) is

$$d = \left(\frac{4\,.\,136^2}{10\,.\,137} - \sqrt{2}\right)\mu = 52{\cdot}589\mu. \tag{99·3}$$

This is nearly the same as the nuclear energy-constant A. In (50·4) and (50·6), A is the energy defined by optical control. To make it comparable with (99·3) we must convert it into molarly controlled measure. The analysis being linear, this multiplies the value of A by $\beta^{\frac{1}{2}}$.[a] The result is

$$A = 52{\cdot}416\mu, \tag{99·4}$$

where μ is the standard mass both in (99·3) and (99·4). The difference between (99·3) and (99·4) is not due to the omission of a minor correction. Mathematically the close agreement is just a coincidence; but the coincidence has important physical consequences. It means that the electric energy of the eight particles (less the amount retained in the two-legged intracule) is just about the amount of non-Coulombian energy released by welding two protons into one, i.e. by constraining them to have the same coordinates; for $-A$ is by definition the non-Coulombian energy of two protons whose coordinate differences ξ, η, ζ are 0.

Consider the following ideal change in the eight-particle system. We remove all the internal Coulombian energy of the six particles forming the nucleus; that is to say, we inhibit interchange circulation between them. But we retain the charges of two of the protons for external purposes; so that these two protons interchange with the two satellite electrons or, if the atom is ionised, with the charges that they induce in the environment. They therefore retain their mutual non-Coulombian energy, which remains as calculated in § 50. The other four particles, being deprived of their external

[a] The irrelevant factor $\beta^{\frac{5}{6}}$ is omitted and the relevant factor β is inserted, as in the case of the additional energy of the neutron (§ 95).

fields, have no non-Coulombian energy. Assuming in advance that the two active protons will be found to be rigidly coincident, they constitute with the two satellites a two-legged intracule. It will be seen that the result of our tampering with the system —removal of all electric energy except that represented by a two-legged intracule— amounts to the addition of the energy $-d$ that has been calculated. In order that this change may come about naturally and not by artificial constraint, this energy must be supplied by internal adjustment in the nucleus. We have left only one adjustable source of energy, namely the non-Coulombian energy of the two active protons. These must draw nearer together until they reach a distance r at which the energy $Ae^{-k^2r^2}$ is $-d$.

If $d < A$, the elimination of the energy d leaves a non-rigid nucleus having an internal coordinate r representable as the separation of two protons. If $d > A$, the minimum non-Coulombian energy is not low enough to enable the atom to reach the ideal state without artificial constraint. The fortunate near coincidence of d and A allows us to adopt either of them as a close approximation to the mass-defect of the actual helium atom. Evidently d and A are respectively upper and lower limits to the mass-defect; so that we have

$$4H - \mathrm{He} = (52{\cdot}502 \pm 0{\cdot}086)\,\mu$$
$$= 0{\cdot}02862 \pm 0{\cdot}00004 \text{ a.w.u.}, \qquad (99{\cdot}5)$$

the error indicated being the extreme error. The observational value is 0·02866.

100. The separation constant of isobaric doublets

A nucleus of mass number n and atomic number N_p is commonly said to contain N_p protons and $N_n = n - N_p$ neutrons. I shall accept this as a current notation, but without endorsing the hypothesis that neutrons exist individually in the nucleus. In my own formulation the atom consists of n protons and n electrons, N_n of the latter having co-spin (nuclear) binding and N_p having spin (atomic) binding. Neither the spin nor the co-spin electrons are bound to individual protons; and the neutron has no relevance except as affording a simple example of co-spin binding.

If for analytical purposes the particles are grouped in twos or fours, allowance must be made for continuous interchange of membership of the groups. There is a great advantage in using the deuterium group of four particles as a unit of structure, because it constitutes a unit cell in exclusion theory. Then, instead of calculating directly the energy due to particles of one group interchanging with particles of another group, we can calculate the exclusion energy which was found in § 47 to be another representation of the same interaction.

Consider a complex atom composed of $\frac{1}{2}n$ deuterium atoms, so that $N_n = N_p = \frac{1}{2}n$. The extracules are passengers in the theory, and we can confine attention to the $\frac{1}{2}n$ double intracules. To calculate their interaction by exclusion theory we proceed just as in § 41; but the application is simpler, because the known data to which the formulae must be fitted are at the lowest level instead of at the top. By (43·1) the top quantum number is

$$\mathfrak{k} = (\tfrac{3}{4}n)^{\frac{1}{3}}. \qquad (100{\cdot}1)$$

Also, by (43·31), the energy at the kth quantum level varies as k^2; so that for different values of n, the top energy is

$$\mathfrak{E} = A\mathfrak{k}^2 \quad (A = \text{constant}). \qquad (100{\cdot}2)$$

It has been pointed out that, whereas (100·1) is only accurate for large values of n, (100·2) is accurate for all integral values of $\mathfrak{k}$.

The case $\mathfrak{k} = 1$ corresponds to a single deuterium atom, and the energy μ_2 of its D.I. has been found in § 98. It was shown in § 43 that at $\mathfrak{k} = 1$ the energy is half rest mass and half exclusion energy; thus the constant A in (100·2) is $\frac{1}{2}\mu_2$ and

$$\mathfrak{E} = \tfrac{1}{2}\mu_2 \mathfrak{k}^2 = \tfrac{1}{2}\mu_2 (\tfrac{3}{4}n)^{\frac{1}{2}}. \tag{100·3}$$

The second form is accurate only for large values of n. The important quantity for our purpose is

$$\mathfrak{E}_0 = \mathfrak{E} - \tfrac{1}{2}\mu_2 = \tfrac{1}{2}\mu_2 \{(\tfrac{3}{4}n)^{\frac{1}{2}} - 1\}. \tag{100·4}$$

This is the exclusion energy of the D.I. as ordinarily reckoned, i.e. the excess above the energy of the ground state $\mathfrak{k} = 1$. To correspond with the analysis of extracules in § 43, we have to picture the unit cell as occupied by two particles which in this case are the two V_6 intracules whose combination forms the D.I.; these mutually exclude one another from the level $k = 0$, so that even in the lowest state the D.I. contains exclusion energy $\frac{1}{2}\mu_2$. But, except in this comparison, the D.I. is treated as an indivisible particle; and 'exclusion energy' refers to the exclusion energy $\mathfrak{E}_0$ between D.I.'s.

We can usefully employ (100·4) in certain differential comparisons of atomic masses. These depend on the fact that the top particle is more or less loose, and can be modified without much affecting the rest of the system. If n is so large that lateral exclusion is negligible, the top particle stands on a rigid platform formed by fully packed energy levels beneath, just as in § 42; and the conditions are the same as those of an isolated particle superposed on a rigid environment. We shall assume these conditions; but, since n is not very large in practical applications, there will be some inaccuracy due to the neglect of lateral exclusion.

The modification of the top particle that will be considered is the substitution of two spin bound electrons for one spin bound and one co-spin bound electron. We set

$$n = N_n + N_p, \quad T_w = \tfrac{1}{2}(N_n - N_p), \tag{100·5}$$

so that the composition of the atom is specified by two characteristics (n, T_w). The change considered—conversion of a nuclear into a satellite electron—is $(n, 0) \to (n, -1)$. We have treated in § 98 the transformation $(2, 0) \to (2, -1)$ of a deuterium atom into a hydrogen molecule. The present transformation is similar except that it must be applied to a top particle; for if a lower particle were changed it would disturb the exclusion energies of all particles above it. The change of $\mathfrak{E}_0$ is proportional to the change of μ_2 in the deuterium transformation; so that, if d is the mass-defect of deuterium,

$$\Delta \mathfrak{E}_0 = \frac{\mathfrak{E}_0}{\mu_2} d. \tag{100·61}$$

If this is the only change of energy, the difference of mass of the two atoms is

$$M(n, -1) - M(n, 0) = \tfrac{1}{2} d \{(\tfrac{3}{4}n)^{\frac{1}{2}} - 1\}, \tag{100·62}$$

where $M(n, T_w)$ denotes the mass of the atom (n, T_w).

It is customary to express mass-differences in thousandths of an atomic weight unit (mMu.). By (98·3), $d = 1{\cdot}540$ mMu.; and (100·62) becomes

$$M(n, -1) - M(n, 0) = \tfrac{1}{2} . 1{\cdot}272 n^{\frac{1}{2}} - 0{\cdot}770. \tag{100·63}$$

An essential point in the theory is that the field giving the particle the energy $\mathfrak{E}_0$ is a rigid exclusion field, unchanged by the insertion of the particle. As n becomes smaller and lateral exclusion becomes relatively more important, this approximation becomes insufficient; a correction for non-rigidity will then be required. When there is only one particle, and therefore no exclusion field, the whole mass μ_2 is 'a correction for non-rigidity'; in fact, as n decreases, the top particle changes gradually from a quantum particle superposed on a rigid field to a relativity particle characterised entirely by the disturbance it creates in the field. In the exclusion calculation the ground-state mass μ_2 is not a constant mass, possessed by the particle in addition to exclusion energy, and carried up with the particle as it is pushed up to higher levels; it is a correction peculiar to the ground level, to be replaced at higher levels by a similar, but much reduced, correction for the lateral exclusion remaining at the higher level. It is for this reason that we use $\mathfrak{E}_0$, not $\mathfrak{E}$ or $\mathfrak{E}_0+\mu$, in obtaining (100·62).

In practical calculations n is rather small, and it is worth while to introduce a refinement that improves the approximation. This eliminates the systematic effect of the residual lateral exclusion, though it cannot eliminate the raggedness associated with small values of n. In the exclusion calculation we use a uniform distribution of momentum over a sphere in momentum space; and the unit cells are successive spherical shells. The energy $\mathfrak{E}$ of the bounding sphere is taken as the energy of the top particle. This obviously exaggerates the exclusion energy, since the particle occupies uniformly the cell next the boundary. One way of correcting this is to give the top particle the mean energy of the top level and of the level next below; then in (100·63) $n^{\frac{3}{4}}$ is changed to $\frac{1}{2}\{n^{\frac{3}{4}}+(n-2)^{\frac{3}{4}}\}$, or approximately $(n-\frac{2}{3})/n^{\frac{1}{4}}$. But the following treatment is perhaps more rigorous.

On adding the top particle the bounding sphere rises from the $\frac{1}{2}n-1$ to the $\frac{1}{2}n$ level. Assuming rigidity of the lower particles this must be regarded as the effect of self-exclusion of the top particle, i.e. exclusion between different parts of it (the particle in the smoothed representation being infinitely divisible). Self-exclusion is lateral; and it would be inconsistent to take it into account without also taking into account the small lateral exclusion between the top particle and other particles. It is desirable to separate the two aspects of a particle, namely as excludors and excludees. As excludors the particles have no separate individuality, and each contributes an equal share to the general exclusion field; as excludees they occupy different cells and so acquire a domiciliary individuality. Regarding the addition of a particle as the addition of an excludor and an excludee, the excludee necessarily goes to the top cell, and the excludor intensifies the exclusion field in the ratio $\frac{1}{2}n-1$ to $\frac{1}{2}n$. We wish to include in $\mathfrak{E}_0$ the self-exclusion of the top particle.[a] We have therefore to adopt the average exclusion field during the gradual addition of the top particle. This is less than the bounding sphere value in the ratio $\frac{1}{2}n-\frac{1}{2}$ to $\frac{1}{2}n$. We have therefore to correct (100·63) by multiplying $n^{\frac{3}{4}}$ by $(n-1)/n$, so that it becomes

$$(n-1)/n^{\frac{1}{4}}. \qquad (100\cdot64)$$

The error of approximation ought to be considerably reduced if the main calculation is made for the middle of the range instead of for one end of it. If n is odd, the atoms $(n, \frac{1}{2})$, $(n, -\frac{1}{2})$ have integral N_p and N_n; and $(n, -\frac{1}{2})$ is obtained from $(n, \frac{1}{2})$ by changing

[a] If it were not included as exclusion energy it would have to be added as rest energy additional to the exclusion energy derived from other particles.

a nuclear into a satellite electron. The atom $(n, 0)$, for which $\mathfrak{E}_0$ is calculated then comes into the middle of the range. It contains a half D.I., and there is no physical atom answering this description; but this does not affect the exclusion calculation, which uses a smoothed representation that can be varied continuously. Making this change as well as (100·64) in (100·63), we have

$$M(n, -\tfrac{1}{2}) - M(n, \tfrac{1}{2}) = \tfrac{1}{2} \cdot 1{\cdot}272 \frac{n-1}{n^{\frac{1}{2}}} - 0{\cdot}770. \qquad (100{\cdot}7)$$

There remain errors of raggedness. For $n < 10$ the smooth formula is very precarious, and we cannot expect more than the roughest agreement; but onwards from $n = 11$, where the third quantum level begins, (100·7) should be a fair approximation.

The pairs of atoms $(n, \frac{1}{2})$, $(n, -\frac{1}{2})$ are called *isobaric doublets*. The experimental data for 18 isobaric doublets ranging from $n = 3$ to $n = 41$ have been investigated by E. P. Wigner, who gives the formula[a]

$$M(n, -\tfrac{1}{2}) - M(n, \tfrac{1}{2}) = \tfrac{1}{2} \cdot 1{\cdot}27 \frac{n-1}{n^{\frac{1}{2}}} - 0{\cdot}79. \qquad (100{\cdot}8)$$

His separation constant 1·27 is purely empirical, being chosen so as to give the best general representation of the data. Thus (100·7) is observationally confirmed.

101. Isotopic spin

Considerable progress has been made in nuclear theory by extending to co-spin angular momentum the analysis used for spin angular momentum in the theory of atomic spectra. The leading principle is that both the resultant angular momentum and the component in a certain control plane are quantised. The control plane is determined by the environment of the particle or group of particles considered; it is a plane of symmetry of that characteristic of the environment which interacts most strongly with the spin of the object-system. For ordinary spin angular momentum the control may be provided either by an extraneous magnetic field or by non-isotropic characteristics of the atom to which the group of particles belongs.

For co-spin angular momentum the control plane must be a particular plane in the 3-space formed by the directions E_{45}, E_{50}, E_{04}. Alternatively we denote these symbols (in undetermined order) by θ_u, θ_v, θ_w. In the standard environment each of these axes has been given a distinctive physical interpretation; so that any of the three planes might be a controlling plane. Let θ_w be the axis exercising the strongest control. This means that there is strong opposition to the θ_u, θ_v rotations of the object-system, which would break its coupling with the θ_w axis, and relatively little opposition to the θ_w rotation which is in the vw-plane. It is evident that the weakly opposed rotation is the time-tilt introduced in § 93; because the perturbation introduced by it is represented by the slightly imperfect stability (metastability) of time-tilted states. The plane of time-tilt is E_{45} in five-dimensional representation; so that it corresponds to the E_{45} component of a vector-density $\mathfrak{P}$. This becomes the E_{04} component of a classical momentum vector P; and by (78·3) the quantum designation is the same:

The control plane for quantisation of co-spin momentum is the E_{04} plane; and the quantised component is the magnetic energy (time-component of magnetic moment) p_{04}. (101·1)

[a] *Reports on Progress in Physics* (Physical Society), **8**, 279, 1941. The constant 0·79 in (100·8) was intended to be the mass of the neutron.

This identification is confirmed by the following considerations. The spin and co-spin terms form the electrical part of a momentum vector; and their symbols ζ_1, ζ_2, ζ_3, θ_u, θ_v, θ_w introduce two 3-spaces useful for the representation of electrical characteristics. Ordinary space, created and occupied by a planoid, is a 3-space of a different type; and its axes (E_{15}, E_{25}, E_{35}) correspond to the symbols $i\theta_w(\zeta_1, \zeta_2, \zeta_3)$ with $\theta_w = E_{04}$. Thus the first effect of associating the electrical system with the standard environment is to single out in θ-space a particular axis θ_w to play the part of E_{04}. This is clearly the principal discrimination of directions in θ-space exercised by the environment; and the further discrimination of two particular axes in the uv-plane must be very subsidiary. Accordingly for a complex atom (considered as a whole) in the standard environment the controlling plane for quantisation of co-spin is the plane E_{04} of the symbolic frame anchored in the standard environment. The controlling force is, in fact, the inertial field.

If we are treating part of an atom, e.g. a top particle or a small group of top particles analogous to the group of valency electrons in an atom, the environment includes the rest of the atom. It is possible that the core of the atom may exercise a control stronger than that of the inertial field, and in these circumstances the quantised component might have a direction differing from E_{04}. I am unable to decide whether this actually happens. The important point is that in no circumstances does the co-spin momentum lack a strong controlling field.

Ordinary angular momentum consists of two parts, usually called spin and orbital momentum, but distinguished in our theory as particle and field angular momentum. There is nothing to correspond to field angular momentum in co-space; or, at any rate, it has not been made use of in theory up to the present. We therefore consider only the particle or spin angular momentum in developing an analogous treatment of co-spin angular momentum.[a]

Denote the resultant spin momentum of the atom, measured in the unit $\hbar$, by S; and let its components be S_x, S_y, S_z, the controlled component being S_z. The established rule is that S is an integer or half integer, and that there are $2S+1$ states with $S_z = S$, $S-1$, ..., $-S$. It is assumed similarly that the resultant co-spin momentum T is an integer or half integer, and that there are $2T+1$ states (constituting different isobaric atoms) with the controlled component $T_w = T$, $T-1$, ..., $-T$. The controlled component is called the *isotopic spin* and is identified with the T_w defined in (100·5). Isobaric doublets correspond to $T = \frac{1}{2}$, $T_w = \frac{1}{2}$, $-\frac{1}{2}$. We can also have isobaric multiplets with 3, 4, 5, ... components, corresponding to $T = 1, \frac{3}{2}, 2, \ldots$.

We have to examine the connection between this treatment and the exclusion treatment adopted in § 100. If $T = \frac{1}{2}n$, the $n+1$ eigenvalues of T_w give isobaric atoms of mass number n and atomic numbers ranging from 0 to n. In the known elements and isotopes T_w is much less than $\frac{1}{2}n$; and, although this does not necessarily imply that T is less than $\frac{1}{2}n$, there is indirect evidence that in the lighter elements (say, up to $n = 50$) T is quite small. The same happens with S; theoretically the resultant spin of Z satellite electrons might be as much as $\frac{1}{2}Z$; but in spectroscopic theory S is always much smaller, being determined by the valency electrons. The reason is that the spins

[a] When not overpowered by other interactions, the field angular momentum provides the control plane for the spin angular momentum. Co-spin angular momentum has not this control to fall back on, but it has the inertial-field control instead.

of the intracules forming the core of the atom are approximately balanced, and can be left out of consideration. The analogous treatment of T evidently assumes that the nucleus similarly consists of a core with approximately balanced co-spins; and T is the resultant co-spin of a small number of co-valency intracules, not included in the core and more or less loose from it. These co-valency intracules correspond to the occupants of the topmost levels in the exclusion treatment.

The isobaric doublets with $T = \frac{1}{2}$ are monocovalent. Their co-valency intracule is the top particle in § 100. The interpolation, which we use when n is odd, makes it in this case half a D.I., which becomes a co-spin intracule in the state $T_w = \frac{1}{2}$ and a spin intracule in the state $T_w = -\frac{1}{2}$. Ideally the top particle interacts with the core only to the extent that the core gives it a constant exclusion energy. This ideal detachment is also assumed in the spin theory, since quantisation of T and T_w is only effective if the external perturbations are not so strong as to cause very frequent transitions between the quantised states. The exclusion treatment is easily extended to triplets, quadruplets, etc., by detaching in the same way the top strata containing the necessary number of D.I.'s or half D.I.'s to represent the required co-valency.

The two methods supplement one another usefully. The smoothed exclusion representation in nuclear theory is analogous to the Fermi-Thomas treatment of the satellite electrons in a complex atom. In the satellite electron system, however, the smoothed representation fails to take account of the closed shells which make the valency a periodic function of the number of electrons. Apparently the nucleus is more amorphous and therefore better suited to the smoothed treatment. There is no sign of periodicity of co-valency; and all odd nuclei up to $n = 35$ have monocovalent isobars.

102. Radii of nuclei

The methods used in this chapter are such that there has been no occasion to consider the coordinates of the particles.[a] Officially we have been unaware that the nucleus is an unusually close agglomeration of particles. It can, however, be inferred from the very large intracule energies that have been found that, whether they are Coulombian or non-Coulombian energies, some particles must be in close proximity. Without further inquiry we shall accept the observational conclusion that the protons and co-spin electrons form a compact nucleus of small radius. Presumably the boundary is ill-defined; and, if an exact 'radius' is attributed to the nucleus, it must be defined conventionally. We follow the convention in Wigner's theory by which the actual nucleus is compared with a simple model. In the model each particle has uniform probability distribution over a sphere of radius r; and r is adjusted to give the best agreement with the observed energies.

The electrostatic energy of the net charge $N_p e$ uniformly distributed over the sphere of radius r is $\frac{3}{5}N_p^2 e^2/r$. This requires correction on account of the aggregation of the charge into finite units e, since it includes a spurious energy corresponding to the non-existent attraction between elements of the same unit e. Wigner assumes that there are N_p charges e, and the corrected energy is accordingly $\frac{3}{5}N_p(N_p - 1)\, e^2/r$. This is unsatisfactory since it depends on the unfounded hypothesis that neutrons exist individually in the nucleus. But we can substitute a more general treatment, which

[a] Except in the second part of the calculation of the atomic mass of helium.

does not raise this question. If we derive the Coulomb energy from the formula (100·7) which has been corrected for the finite divisibility of the particle distribution, or from Wigner's empirical formula (100·8) which must include the correction, we must use an expression for the Coulomb energy which has been correspondingly corrected. But our original formula (100·63) refers to a smoothed distribution in which the particles are infinitely divisible, and this will transform directly into the correspondingly smooth Coulomb energy which treats the carriers of charge as infinitely divisible. We therefore use the smoothed continuous distribution throughout the calculation, which gives the exclusion energy (100·63) and the electrostatic energy $\frac{3}{5}N_p^2e^2/r$.

Exclusion energy can equivalently be represented by interchange energy, which in turn has been identified with Coulomb energy; so that broadly speaking our exclusion treatment is a method of calculating the Coulomb energy of the atom. The different points of view adopted in the three kinds of treatment lead to somewhat different definitions and distinctions; and it would be more accurate to say that the treatments overlap, but are not identical in their scope. In (100·63) we readily identify the term involving n with the electrostatic energy as ordinarily defined. The constant term (which would ordinarily be classed as non-Coulombian energy) is an intrinsic difference of energy between spin binding and co-spin binding. Wigner, assuming that the electron is bound to a particular proton, takes the constant term to be the mass excess of the neutron. We take it to be equally bound to all the protons in the nucleus. Owing to the smoothed treatment, our value 0·77 of the constant term is strictly the limiting value when the number of protons is very large. When the electron is co-spin bound to one proton as in the neutron the constant term is 0·82. It is noteworthy that it makes so little difference to the energy whether the binding is to one proton or many protons.

Since $N_p = \frac{1}{2}n - T_w$, the difference of the electrostatic energy $\frac{3}{5}N_p^2e^2/r$ in the transformation of T_w from $\frac{1}{2}$ to $-\frac{1}{2}$ is $\frac{3}{5}ne^2/r$. (In conformity with the 'looseness' of the top particle (co-valency particle) it is assumed that the general distribution of the particles which determines r is unaltered by small changes of T_w.) This difference corresponds to the term $\frac{1}{2}.1{\cdot}272.n^{\frac{2}{3}}$ in (100·63). Hence

$$e^2/r = 1{\cdot}060n^{-\frac{1}{3}} \text{ mMu.}$$

which gives
$$r = 1{\cdot}460 \times 10^{-13}.n^{\frac{1}{3}} \text{ cm.} \qquad (102{\cdot}1)$$

The particle density of protons in the nucleus is

$$\sigma_p = n/\tfrac{4}{3}\pi r^3 = 7{\cdot}674 \times 10^{37} \text{ cm.}^{-3} \qquad (102{\cdot}2)$$

This is independent of n; but, owing to the approximation employed, it is restricted to atoms in which T_w is small.

103. The nuclear planoid

The investigation of the satellite electron system of a complex atom is naturally divided into peripheral problems (concerned with valency electrons) and core problems. Nuclear theory may similarly be divided into peripheral theory and core theory. The existing peripheral theory, as developed especially by Wigner, forms a natural specialised continuation of our fundamental theory (§ 101); and, with only minor changes and re-interpretations, we can annex it *en bloc*. Several attempts to investigate

the core have achieved substantial progress, and there must be some measure of truth in the corresponding formulations; but the progress has not gone far enough to make contact with fundamental theory. The present gap is too wide to be bridged without a great deal of specialised investigation which I am unable to undertake. The following development does not close the gap, but is intended to narrow it.

The knowledge of the nature and magnitude of the Coulombian and non-Coulombian forces between particles derived by fundamental theory is presumably adequate for the solution of all nuclear problems; the difficulty is to find a method of applying it to highly complex systems, which is not altogether intractable. Nuclear physicists have proceeded on the assumption that the quantum method developed in extra-nuclear physics is suitable also for the nucleus; and it is clear that the assumption is justified, partially at any rate, by the results. We have found that the quantum method breaks away from the classical or relativity method by introducing quantum particles as superpositions on a rigid environment; and it is a fair conclusion that those who apply quantum methods to nuclear structure are dealing with particles of this kind. Probably unconsciously, they have introduced a nuclear uranoid or planoid, and are pursuing the super-threshold theory that results.

The uniform distribution found in the last section, which has a density independent of the mass-number n of the atom, offers itself as a nuclear planoid. The proton density σ_p, found in (102·2), is $1{\cdot}044 \times 10^{41}$ times the proton density in the standard uranoid. This ratio has been obtained by purely theoretical calculation.[a] Besides having greater density, the nuclear planoid differs from the uranoid in having one electron per two protons instead of one electron per proton.

The non-Coulombian energy of a proton with another proton at a distance r is $-Ae^{-r^2/k^2}$. Thus, when it is in a distribution of protons with particle density σ_p, its non-Coulombian energy is $-E$, where

$$E = \tfrac{1}{2}\int Ae^{-r^2/k^2}.\sigma_p dV = \tfrac{1}{2}\pi^{\frac{3}{2}}k^3A\sigma_p, \qquad (103{\cdot}1)$$

the factor $\frac{1}{2}$ being inserted to avoid reckoning the total non-Coulombian energy of the nucleus twice over. This, however, refers to the standard uranoid in which only spin binding of the charges occurs. Here the charge $-e$ induced by a proton in its environment is half spin bound and half co-spin bound. From the extra-nuclear point of view we are concerned only with the spin bound induction. By the calculation of non-Coulombian energy in § 49, the constant A for a nuclear proton is half the constant for a free proton, because its field is screened by the co-spin bound charge $-\frac{1}{2}e$ which it induces in the nucleus. The intra-nuclear point of view gives the same result, because we are then only concerned with the induction in the nuclear planoid which is the co-spin bound charge $-\frac{1}{2}e$.

Accordingly, halving (103·1), and inserting the values of A, k, σ_p already found, the non-Coulombian energy per proton is

$$E = \tfrac{1}{4}\pi^{\frac{3}{2}}k^3A\sigma_p = 21{\cdot}631 \text{ mMu}. \qquad (103{\cdot}2)$$

The Coulombian energy is $\frac{3}{20}ne^2/r = 0{\cdot}159n^{\frac{5}{3}}$; so that the total energy is

$$-21{\cdot}631 + 0{\cdot}159n^{\frac{5}{3}} \text{ mMu. per proton.} \qquad (103{\cdot}3)$$

The particles will of course have large kinetic energy in this 'potential well' since they exclude one another from the bottom level of the well. If there are a reasonably large

[a] Omitting β-factors, the ratio is $\frac{1}{2}\pi(\frac{5}{3})^3(\frac{5}{4})^{\frac{9}{2}}N^{\frac{1}{3}}$.

number of protons, the usual result that the mean energy $\bar{E}$ is $\frac{3}{5}$ of the top energy $\mathfrak{E}$ will apply.

The foregoing is the sub-threshold analysis of the nuclear planoid. The particles which figure in ordinary quantum analysis are super-threshold particles, carriers of energy superposed on that of the rigid planoid. In order that these may be identified with the protons and electrons of extra-nuclear theory, which are carriers of energy superposed on that of the rigid uranoid, the top level of the nuclear planoid must coincide with the top level of the standard uranoid. This adjustment is necessary in order that there may be no discontinuity of energy reckoning when particles pass from or to the nuclei in atomic transmutations. We have determined r so that the Coulomb energy $0{\cdot}159n^{\frac{5}{3}}$ of the nucleus agrees with the ordinary reckoning $\Sigma e^2/r_{st}$ outside the nucleus; it will therefore not affect the adjustment of top levels. Thus the constant term $-21{\cdot}631$ in (103·3) is the depth of the bottom level of the nuclear planoid below the common top level of the nuclear planoid and uranoid; that is to say,

$$\mathfrak{E} = 21{\cdot}631 \text{ mMu.} \tag{103·4}$$

We may attempt to calculate the exclusion energy $\mathfrak{E}$ directly from the density of the particle distribution in the nucleus, as in the theory of white dwarf matter. The cell formulation of the exclusion principle does not apply to protons unless the full complement of electrons is present, and we have to revert to extracules and intracules. The new energy is exclusion energy of the extracules, which are packed more closely in the nuclear planoid than in the uranoid; the exclusion energy of the double intracules has already been taken into account (§ 100). Adapting (45·1) the energy per extracule is

$$\frac{3}{5}\left(\frac{3\sigma_p}{8\pi}\right)^{\frac{2}{3}}\frac{h^2}{2M} = 23{\cdot}116 \text{ mMu.} \tag{103·5}$$

In (45·1) this is the mean exclusion energy $\bar{E}$, because we there dealt with a super-threshold application of the exclusion principle. Here we are dealing with a sub-threshold application, and the constant h^2 should be replaced by $\frac{3}{5}h^2$ in accordance with (42·2). Or, what comes to the same thing, the result (103·5) is the top energy $\mathfrak{E}$.

This second calculation of $\mathfrak{E}$ is not rigorous; although the nuclear sphere of radius r contains n protons, it is not at all clear that it contains n extracules. Half the extracules are associated with spin intracules and the corresponding electrons are far outside the sphere. It is therefore not surprising that (103·5) exaggerates the exclusion energy a little. The second method is not really a suitable way of attacking the problem; but the rather close agreement of the two results confirms our interpretation of $\mathfrak{E}$ as the depth (from top to bottom level) of the nuclear planoid.

Since the mean exclusion energy of a proton in the nuclear planoid is $\frac{3}{5}\mathfrak{E}$, there is a deficit $\frac{2}{5}\mathfrak{E}$ per proton of energy in the nuclear planoid as compared with the ordinary rest energies of the same particles in the standard uranoid. This means that (in so far as the planoid is an adequate approximation) a nucleus of mass number n will have a mass-defect

$$\left.\begin{aligned}\tfrac{2}{5}\mathfrak{E}n &= 8{\cdot}65n \text{ mMu., compared with hydrogen}\\ &= 0{\cdot}52n \text{ mMu., compared with oxygen.}^{a}\end{aligned}\right\} \tag{103·6}$$

I think that this is about the closest possible linear representation of the observed mass-defects from, say, $n = 30$ to 150. For $n > 150$ the mass-defect is smaller. It is

[a] Corresponding to a 'packing fraction' 5·2.

also smaller for $n < 30$; but this is accounted for by edge-effect, which reduces our value (103·2) of the non-Coulombian energy calculated for a uniform proton distribution of indefinite extent.

If we think of $\frac{2}{5}\mathfrak{E}$ only as mass-defect, the close agreement with observed mass remains an unexplained curiosity. The nucleus contains Coulombian energy large compared with the errors of the formula (103·6); and this must be counterbalanced by negative energy superposed on the planoid. The key to the practical validity of (103·6), which is a slightly improved form of the 'whole number rule' is that the energy $\frac{2}{5}\mathfrak{E}$ is a *margin of safety*. We shall always have particles with an energy 8·65 mMu. greater than the mean energy. Presumably the most stable state is that in which a nucleus can just retain its top particle, and cannot admit another. The mass-defect (103·6) then follows automatically. To the usual approximation, which neglects lateral exclusion, the planoid and the system superposed on the planoid are independent; and the presence of the latter will not affect the foregoing condition.

Evidently it is a boundary condition for the superposed system that its total energy is approximately zero. As n increases from 40 to 238 the Coulomb energy (superposed on the planoid) increases from 1·86 mMu. per proton to 6·10. No doubt it is the difficulty of balancing this energy that causes the nucleus to draw in electrons in excess of the planoidal quota. From $n = 238$, this excess of electrons reduces the Coulomb energy from 6·10 to 3·60; so that the direct effect is considerable. But we forbear to pursue the technical ramifications of nuclear physics; and the theory of the nuclear planoid is a suitable halting point.

It may be convenient to summarise the present contributions of fundamental theory to nuclear theory. These are:

(1) Determination of the law and the constants A, k of the non-Coulombian energy of two protons.

(2) Determination of the mass-defect of deuterium.

(3) Determination of the mass-defect of helium. This is possible by a favourable combination of circumstances not applying to any other atom.

(4) Determination of the separation constant of isobaric doublets, and hence of the radii of nuclei.

(5) Determination of the binding energy of a co-spin electron (*a*) in a neutron, (*b*) in a nucleus. The former is the mass-excess of a neutron.

(6) Suggested use of a nuclear planoid, leading to a possible explanation of the whole number rule.

(7) Unification of the isotopic spin theory with the rest of fundamental theory.

(8) Rejection of meson-field theory, and of the constitution of the nucleus out of protons and neutrons. The latter hypothesis, however, if used merely as an *approximation*, is comparatively harmless, and may at times be a convenient simplification. No empirical constants are used in the calculations; and the results in all cases agree with the experimental values.

104. Mass of the mesotron

The so-called reality condition which forbids momentum components of the quantum energy tensor Z is a boundary condition furnished by the standard environment. It refers to a simple system, fully specified by Z and therefore containing not more

than one D.I., in isolation or in a non-interacting assemblage. A complex atom is formed by uniting several simple systems with energy tensors Z^r ($r = 1, 2, 3, \ldots$). Any one of the simple systems is then in an environment containing electrical particles, and the standard reality condition does not apply to it. Alternatively, defining Z^r to be the energy tensor which the simple system would have if it were isolated, the whole energy tensor Z of the complex atom contains in addition to ΣZ^r an interaction energy for which the reality condition has not been determined. There is therefore nothing to prevent atoms more complex than helium from containing a certain amount of forbidden energy Z_m. We exclude helium because its constitution has been determined in § 99, and it is clear that there is no appreciable amount of forbidden energy.

The carrier of an energy tensor Z_m consisting of momentum components will be called a *forbidden oscillator*. Normally a complex system is analysed into symmetrical and antisymmetrical states chosen so that the part of Z_p and Z_d not included in the state energy tensor is as small as possible. A small residue of Z_d will be an interstate tensor representing transition circulation between the antisymmetrical states, and a small residue of Z_p will represent transition circulation between the symmetrical states; Z_m will be wholly an interstate tensor representing transition circulation between symmetrical and antisymmetrical states.

Physically we distinguish Z_p and Z_d as mechanical and electrical energy. Our conclusion is that in the heavier nuclei the sharp distinction between 'mechanical' and 'electrical' breaks down, and we can have energy in a mechanical-electrical interstate. It is natural that a distinction, first introduced and defined in very simple conditions, should under increasing complication become muzzy. The forbidden components fall into two groups:

$$Z_{m1} = \Sigma E^s_\mu F^a_\nu p_{\mu\nu}, \quad Z_{m2} = \Sigma E^a_\mu F^s_\nu p_{\mu\nu}, \tag{104·1}$$

where s and a indicate symmetrical and antisymmetrical matrices. We therefore recognise two types of forbidden oscillator. It appears that one of them, say Z_{m1}, transfers occupation from a mechanical to an electrical state, and the other from an electrical to a mechanical state. The argument is that the terms $E_\mu F_\nu$ and $E_\nu F_\mu$ in Z must be distinguished in some way because in the state energy tensor they count as independent degrees of freedom. When they occur in an interstat energy tensor they are interchanged by interchanging the two states. I do not think there is any other way of distinguishing two kinds of transition flow.[a] Since Z_{m1}, Z_{m2} each contain 60 components the forbidden oscillator is a V_{60}.

To facilitate the calculation of masses, etc., we consider uniform assemblages of particles in a steady state, just as the practical physicist arranges his material in a way which simplifies his research. Like him, we afterwards apply our results in different conditions, e.g. to isolated particles in unsteady conditions. This practice has hitherto worked smoothly because we have treated particles obeying the standard reality conditions and therefore transferable into any neutral environment. The forbidden V_{60} does not obey the standard reality conditions, so that when transferred to the standard neutral environment it can only exist transiently. If by motion or otherwise

[a] Steady transition circulation will imply equal occupation of $E_\mu F_\nu$ and $E_\nu F_\mu$; but this does not affect their independence, since the steadiness is not compulsory. There is no need for the coefficients of $E_\mu F_\nu$, $E_\nu F_\mu$ to be equal in a state energy tensor.

the highly complex conditions in which the oscillator was born become replaced by a standard environment, Z_m will presumably not be able to disappear instantaneously; we may therefore expect occasionally to observe forbidden oscillators as free but short-lived particles.

The free Z_{m1} particle, which transfers occupation from a mechanical to an electrical state, will spend itself in transforming mechanical energy of the uranoid into electrical energy. The Z_{m2} will similarly spend itself in transforming electrical energy of the uranoid into mechanical energy. The fact that the uranoid contains no electrical energy is a trivial obstacle; after the Z_{m2} particle has done its work there will be a minus quantity of electrical energy. Thus the ultimate products of the two kinds of forbidden oscillator are related in the same way as the electron and positron.

It is clear that the forbidden oscillator is electrically charged. Z_{m1} and Z_{m2} are reversed in sign by reversing the chirality of the double frame; so that they react with an extraneous electromagnetic field.

The process of decay can be studied more precisely. Consider a uniform assemblage of $Z_{m1}^{\bullet}$ particles. The energy tensor has to change from the Z_m state which does not satisfy the reality conditions into the usual $Z_p + Z_d$ state which does satisfy them. Thus the V_{60} particles transform themselves by transition into V_{136} particles. If $\mathfrak{m}$ is the mass of a Z_{m1} particle, the mass of the resulting V_{136} is $60\mathfrak{m}/136$; and the difference

$$76\mathfrak{m}/136 \qquad (104{\cdot}2)$$

is radiated in the transition. Radiation cannot carry away the charge;[a] and the V_{136} will be charged—as we should expect since it is formed by conversion of mechanical into electrical energy. The V_{136} is therefore not a standard particle, but the charged bi-particle treated in § 22 which consists of an electron (or proton) and a comparison particle. The comparison particle merges in the uranoid so that, according to the usual description the Z_{m1} particle changes into an electron (or proton). The Z_{m2} particle similarly changes into a positron (or negatron).

The reason for the genesis of a comparison particle is easily understood. We have to remember that Z refers to quadratic, i.e. sub-threshold, analysis. We have found that two protons and two electrons in super-threshold analysis correspond to a unit cell in sub-threshold analysis. This refers to the state energy tensor. The interstate energy tensor consists of interaction or perturbation of this system of particles; and its conceptual carriers, being supernumerary to the system of particles, have been distinguished as oscillators. We cannot introduce in super-threshold theory a new particle without antecedents in sub-threshold theory. The object-particles in super-threshold theory are obtained, as described in § 42, by exciting a comparison particle. Thus, if an oscillator is to be turned into a particle it has to provide the energy m_0 necessary to constitute an additional particle in the sub-threshold as well as the excitation energy corresponding to the recognised mass m_e or m_p of the resulting object-particle. In short, Z_m is a sub-threshold energy tensor reckoned from the bottom level in exclusion theory, which is m_0 below the zero level of energy reckoning in super-threshold theory.

We recognise observationally transient particles, called mesotrons, which decay by transition into electrons or positrons. It can scarcely be doubted that these are the forbidden oscillators here discussed. Accepting this identification we can calculate the

[a] We are dealing with a large assemblage so that the molar law of conservation of charge applies.

mass of the mesotron. The mass of the charged V_{136} bi-particle (comparison particle + electron) is m_0+m_e. Hence $\mathfrak{m} = \frac{136}{60}(m_0+m_e)$, and the energy radiated in the transition is

$$\mathfrak{m}-(m_0+m_e) = \tfrac{76}{60}(m_0+m_e). \tag{104·3}$$

The mass $\mathfrak{m}_0$ of the mesotron is by this much greater than the mass m_e of the electron into which it decays. Thus[a]

$$\mathfrak{m}_0 = \tfrac{76}{60}(m_0+m_e)+m_e. \tag{104·4}$$

If we take account of β-factors the quadratic energy $\mathfrak{m}$ is multiplied relatively to the linear energy m_0+m_e by β. The result becomes

$$\mathfrak{m}_0 = \tfrac{77}{60}(m_0+m_e)+m_e, \tag{104·5}$$

where m_0 is as usual $10M/136$. (The effect of β-factors on m_e is trivial and has been ignored.) The result is

$$\mathfrak{m}_0 = 173{\cdot}98 \tag{104·6}$$

in terms of the current electron mass m_e.

There seems to be no reason why there should not also exist *heavy mesotrons* which decay into protons and negatrons. Their mass is obtained by substituting m_p for m_e in (104·5). The result is $2{\cdot}38m_p$.

The principal, if not the only source, of mesotrons must be high-energy stimulation of nuclei containing forbidden oscillators. Since they are produced by extra-terrestrial cosmic rays falling on the earth's atmosphere, we infer that nitrogen and oxygen (either or both) contain forbidden oscillators. The occupation of these oscillators is, of course, very small in the normal atom, and an oscillator filled with unit occupation must break away from the atom. It may be rather important that hydrogen and helium contain no forbidden oscillators, and are therefore transparent to cosmic rays; since these elements have very high cosmical abundance, and constitute perhaps 99 per cent of the mass of interstellar matter.

[a] This formula was first given in *Proc. Roy. Soc.* A, **174**, 46, 1940. The multiplicity factor 60 was then assigned tentatively as an interpretation of current views as to the spin properties of the mesotron. The theory here given is much more cogent.

Chapter X

THE WAVE EQUATION

105. Field momentum

Let a covariant wave vector ψ be a function of the coordinates x_1, x_2, x_3, x_4; and let X be the position vector $E_{15}x_1+E_{25}x_2+E_{35}x_3+E_{45}x_4$. A rotation q in the 4-space gives $\psi' = q\psi$, $X' = qXq^{-1}$. Setting

$$\psi = f(x_1, x_2, x_3, x_4), \quad \psi' = f'(x_1', x_2', x_3', x_4'), \tag{105·11}$$

we wish to determine the transformation $f \to f'$ of the functional operator f.

Consider an infinitesimal rotation $q = e^{\frac{1}{2}E_{12}d\theta_{12}}$. By § 56,

$$x_1' = x_1 - x_2 d\theta_{12}, \quad x_2' = x_2 + x_1 d\theta_{12}, \quad x_3' = x_3, \quad x_4' = x_4. \tag{105·12}$$

Hence $$\psi' = f'(x_1 - x_2 d\theta_{12}, x_2 + x_1 d\theta_{12}, x_3, x_4) = e^{l_{12}d\theta_{12}} f'(x_1, x_2, x_3, x_4),$$

where $$\boldsymbol{l}_{12} = x_1 \partial/\partial x_2 - x_2 \partial/\partial x_1. \tag{105·21}$$

Therefore $$f'(x_1, x_2, x_3, x_4) = e^{-l_{12}d\theta_{12}}\psi' = e^{-l_{12}d\theta_{12}+\frac{1}{2}E_{12}d\theta_{12}}\psi$$
$$= e^{(-l_{12}+\frac{1}{2}E_{12})d\theta_{12}} f(x_1, x_2, x_3, x_4).$$

We set $$\boldsymbol{L}_{\mu\nu} = -\boldsymbol{l}_{\mu\nu} + \tfrac{1}{2}E_{\mu\nu} = -(x_\mu \partial/\partial x_\nu - x_\nu \partial/\partial x_\mu - \tfrac{1}{2}E_{\mu\nu}). \tag{105·22}$$

Then for a rotation $d\theta_{\mu\nu}$ in any of the six coordinate planes

$$f' = e^{\boldsymbol{L}_{\mu\nu}d\theta_{\mu\nu}} f. \tag{105·23}$$

Let $\boldsymbol{q}_\mu(\theta)$ denote the operation of rotating an operand through an angle θ in the E_μ plane, i.e. changing it to the equivalently constructed operand in a frame rotated through this angle but retaining the original E-frame as reference frame. When θ is infinitesimal, we have

$$\boldsymbol{q}_\mu(\theta)\,\psi = e^{\frac{1}{2}E_\mu\theta}.\psi, \quad \boldsymbol{q}_\mu(\theta) f = e^{\boldsymbol{L}_\mu\theta}.f. \tag{105·24}$$

For an operand ψ, $$\left(\frac{\partial \boldsymbol{q}_\mu}{\partial\theta}\right)_{\theta=0} = \tfrac{1}{2}E_\mu = \frac{i}{2p_{16}}\boldsymbol{p}_\mu$$

by (73·23), $\boldsymbol{p}_\mu$ being the operational form of the momentum component p_μ. Thus

$$\boldsymbol{p}_\mu = -2ip_{16}\frac{\partial \boldsymbol{q}_\mu}{\partial\theta}. \tag{105·3}$$

We shall adopt (105·3) *as the general definition of the momentum operator.*

For a physical system specified by a wave function f, we have

$$\boldsymbol{p}_\mu = -2ip_{16}\boldsymbol{L}_\mu, \tag{105·41}$$

since (105·24) gives $\partial \boldsymbol{q}_\mu/\partial\theta = \boldsymbol{L}_\mu$ in this case. Since the wave function is defined over a 4-space, only six components of the momentum vector are defined by (105·41), namely, the components of the angular momentum 6-vector. To agree with the usual notation we set

$$p_{16} = \tfrac{1}{2}\hbar. \tag{105·42}$$

Then, by (105·41) and (105·22),

$$\boldsymbol{p}_{\mu\nu} = -i\hbar(x_\mu \partial/\partial x_\nu - x_\nu \partial/\partial x_\mu - \tfrac{1}{2}E_{\mu\nu}), \tag{105·5}$$

which is the well-known formula for angular momentum in current theory. It consists of two parts, which we shall distinguish as

$$\left.\begin{array}{ll} \text{field angular momentum} & -i\hbar(x_\mu \partial/\partial x_\nu - x_\nu \partial/\partial x_\mu), \\ \text{particle angular momentum} & -i\hbar(-\tfrac{1}{2}E_{\mu\nu}). \end{array}\right\} \tag{105·6}$$

The rotation (105·12) is about the origin, and the field angular momentum is correspondingly about the origin. If $\boldsymbol{p}_\mu$ is a linear momentum at x_μ the corresponding angular momentum about the origin is $x_\mu \boldsymbol{p}_\nu - x_\nu \boldsymbol{p}_\mu$. Thus the field angular momentum is the result of a field linear momentum

$$\boldsymbol{p}_\mu = -i\hbar \partial/\partial x_\mu. \tag{105·7}$$

In the early chapters (105·7) has sometimes been used in anticipation. But this is the first time field momentum has appeared in the systematic development which began in Chapter VI. Throughout the last four chapters we have treated wave vectors and tensors which are either isolated or constant in space; so that the momentum has always been of the kind which we now describe more particularly as particle momentum. Thus, in the logical sequence of the theory, (105·7) is a new result now met with for the first time. It provides the accepted definition of the constant $\hbar$; and, since it is a field momentum, $\hbar$ is essentially a field unit of action; and we see immediately the reason for $\hbar c/e^2$ being 137 rather than 136 (§ 20).[a]

There are conspicuous differences between the particle and field momentum vectors. The latter is limited to a linear momentum 4-vector; for, although we employ also an angular momentum 6-vector, this is merely a different way of using the linear 4-vector, not an additional characteristic. In the particle momentum vector all the components represent independent characteristics. The commutation properties of the field and particle vectors are, of course, quite different.

The definition (105·3) gives insight into the nature of momentum. Writing $\psi(\theta)$ for $\boldsymbol{q}_\mu(\theta)\psi$, the expression $\chi^* \dfrac{\partial}{\partial\theta}\psi(\theta)$ is a measure of the differential effect on $\chi^*\psi$ of rotating one factor relatively to the other—a process which may be regarded as creating a strain in the binding of ψ and χ. The momenta p_μ (expectation values) are therefore coefficients of elasticity, defined in the usual way by considering small virtual displacements. In order that the strain in the binding may be separated from intrinsic strains in the factors themselves, the virtual displacement must be such as to produce no intrinsic strain in ψ; and accordingly it is limited to a relativity rotation. Extending the same interpretation to the product of two wave functions f, g, the angular momenta (105·5) are coefficients of rotational elasticity of the binding of f and g. As before the virtual displacement must be such as to produce no intrinsic strain in f; and the condition for this is that it is the relativity rotation $f \to f'$ determined in (105·23).

Although quantum theory ostensibly adopts the non-relativistic outlook which associates observable momenta with entities rather than with relations between entities,

[a] This is contingent on the present distinction between particle and field characteristics being found to agree with the distinction in Chapter II. This will be verified later.

its practice is better than its profession; since even the simplest 'entity' is represented by the product of two factors as though it were composed of two parts, and the observable momenta are contained in the binding of the conceptual counterparts of the two factors.

It will be seen that the field angular and linear momentum is here introduced as a relativistic correction to the particle angular momentum, the correction being required when wave vectors are replaced by wave functions. This is just the opposite to the procedure in current text-books which assume the momentum (105·7) and deduce the 'spin', i.e. particle, angular momentum as a correction.

The investigation brings out a point, not always fully appreciated, that a vector wave function is by no means the same thing as a wave vector which varies with position. The wave function is an *operator* f which, when acting on the coordinates, yields the variable wave vector ψ. The recognition of the operational nature of a wave function is in keeping with the trend of modern physics which, by substituting operations for algebraic numbers, reduces the subject to a formulation of group structure. Unfortunately we cannot very well show the distinction in the notation, because it is the general practice—in pure mathematics as well as in physics—to use the same symbol for the function and the functional operator. We shall in future use ψ to denote the wave function as well as the variable wave vector.

In (105·3) the virtual rotation is applied to the final factor. Since the strain in the binding is produced by relative rotation, the sign is reversed if h denotes rotation of the initial factor. For an initial contravariant vector $\chi = g(x_1, x_2, x_3, x_4)$, the result corresponding to (105·23) is

$$g' = ge^{(-l_{12}-\frac{1}{2}E_{12})\,d\theta_{12}}, \tag{105·81}$$

where $\boldsymbol{l}_{12}$ now operates to the left. Allowing for the reversal of sign, the angular momentum is $i\hbar(\boldsymbol{l}_{12}+\frac{1}{2}E_{12})$. Using $\delta/\delta x_\mu$ to denote the differential operator $\partial/\partial x_\mu$ written after its operand, the formulae (105·6) are adapted to an initial contravariant wave function by setting

$$\frac{\partial}{\partial x} = -\frac{\delta}{\delta x}. \tag{105·82}$$

In particular the linear momentum operator becomes

$$\boldsymbol{p}_\mu = i\hbar\frac{\delta}{\delta x_\mu}. \tag{105·9}$$

106. The gradient operator

If the gradient $\partial/\partial x_\mu$ is treated in the same way as an ordinary space vector, it has the symbolic form

$$\mathbf{grad} = E_{15}\partial/\partial x_1 + E_{25}\partial/\partial x_2 + E_{35}\partial/\partial x_3 + E_{45}\partial/\partial x_4, \tag{106·11}$$

where the real time coordinate is ix_4. The same operator operating to the left is denoted by

$$\mathbf{darg} = E_{15}\delta/\delta x_1 + E_{25}\delta/\delta x_2 + E_{35}\delta/\delta x_3 + E_{45}\delta/\delta x_4. \tag{106·12}$$

Let ψ, χ be vector wave functions satisfying the differential equations

$$(\mathbf{grad} - M)\,\psi = 0, \quad \chi^*(-\mathbf{darg} - M) = 0, \tag{106·21}$$

where M may be a function of the coordinates, possibly containing E symbols. Multiplying the first equation initially by χ^* and the second finally by ψ, and subtracting,

$$\chi^*(\mathbf{grad}+\mathbf{darg})\,\psi = 0, \tag{106·22}$$

or

$$\Sigma\frac{\partial}{\partial x_\mu}(\chi^* E_{\mu 5}\psi) = 0 \quad (\mu = 1,2,3,4).$$

Using (73·1), in which p_μ denotes *particle* momentum only, this gives

$$\Sigma\frac{\partial p_{\mu 5}}{\partial x_\mu} = 0, \tag{106·23}$$

so that the equations (106·21) express the condition that the particle momentum 4-vector is solenoidal.

Since the wave functions represent a continuous distribution, it is preferable to use the strain vector which gives the three-dimensional density of momentum. We set

$$\mathbf{grad}_5 = -E_{45}\,\mathbf{grad} = E_{14}\partial/\partial x_1 + E_{24}\partial/\partial x_2 + E_{34}\partial/\partial x_3 + \partial/\partial x_4, \tag{106·24}$$

and $\mathbf{darg}_5 = -E_{45}\,\mathbf{darg}$. By (72·72), $\chi^* = \psi^\dagger E_{45}$; and (106·2) can be written alternatively as

$$(\mathbf{grad}_5 + E_{45}M)\,\psi = 0, \quad \psi^\dagger(-\mathbf{darg}_5 + E_{45}M) = 0. \tag{106·3}$$

By the same procedure as before we obtain

$$\partial s_{14}/\partial x_1 + \partial s_{24}/\partial x_2 + \partial s_{34}/\partial x_3 + \partial s_{16}/\partial t = 0, \tag{106·4}$$

where $t = ix_4$. This is the equation of conservation of mass for a fluid of mass density s_{16} and momentum density s_{14}, s_{24}, s_{34}.

Since the expectation value of an operator $\boldsymbol{a}$, which in special cases reduces to its eigenvalue, is defined in (72·24) by

$$a = \iiint \psi^\dagger \boldsymbol{a}\psi\, dx_1 dx_2 dx_3 \div \iiint \psi^\dagger\psi\, dx_1 dx_2 dx_3,$$

it is usual to normalise wave functions so that

$$\iiint \psi^\dagger\psi\, dx_1 dx_2 dx_3 = 1. \tag{106·5}$$

Then $\psi\psi^\dagger$ is not strictly the momentum strain vector, which would have given

$$\iiint \psi^\dagger\psi\, dx_1 dx_2 dx_3 = 4i\iiint s_{16}\, dx_1 dx_2 dx_3 = 4i\epsilon,$$

ϵ being the energy of the integrated distribution. When necessary we distinguish the vector normalised by (106·5) as the *stream vector*. The stream vector is the particle momentum vector divided by $4i\epsilon$. When it is used, we can interpret $4is_{16}$ as the probability density and $4i(s_{14}, s_{24}, s_{34})$ as the probability flux of the carrier of ϵ; and (106·4) is the equation of continuity of the flow of probability.

We secure that the flow of probability satisfies the equation of continuity in three-dimensional space, by making the wave functions satisfy differential equations of the form (106·3). But if interchange circulation occurs, so that there is an extra-spatial transfer of probability, the spatial flow will not satisfy the equation of continuity. We then have to correct (106·3) in such a way as to make allowance for the extra-spatial transfer. It is important to understand that extra-spatial flow is treated, not by ex-

tending the analysis to include the extra dimension, but by inserting correcting terms representing the modification of the conditions to be satisfied by the spatial flow.

The developments that are of practical importance begin with the introduction of interchange; but we must first examine more closely the significance of the elementary equations without interchange. We have

$$-i\hbar\,\mathbf{grad} = i\hbar\,\mathbf{darg} = \Sigma E_{\mu 5}\mathbf{p}'_{\mu 5}, \tag{106·6}$$

where the accent is used to distinguish field momentum from particle momentum (unaccented). Consider the case in which M is an algebraic function or constant. Then (106·21) becomes

$$(\Sigma_\mu E_{\mu 5}\mathbf{p}'_{\mu 5} + i\hbar M)\,\psi = 0, \quad \chi^*(\Sigma_\mu E_{\mu 5}\mathbf{p}'_{\mu 5} + i\hbar M) = 0. \tag{106·7}$$

These equations have the same form as the wave identities (72·81) satisfied by the particle momentum vector. Similarly (106·3) reduces to the same form as the wave identities (72·82) satisfied by the particle momentum strain vector. This suggests that the field momentum vector is a simple multiple of the particle momentum vector. To prove this, multiply the first equation of (106·7) initially by $\chi^* E_{\nu 5}$ and the second finally by $E_{\nu 5}\psi$, and add; we obtain

$$2\chi^*(-\mathbf{p}'_{\nu 5} + i\hbar M E_{\nu 5})\,\psi = 0 \quad (\nu = 1, 2, 3, 4, 0). \tag{106·81}$$

Using (73·24), $\quad p_{\nu 5} = \chi^*(-ip_{16}E_{\nu 5})\,\psi \div \chi^*\psi, \quad p'_{\nu 5} = \chi^*\mathbf{p}'_{\nu 5}\psi \div \chi^*\psi.$

Thus (106·81) reduces to

$$p'_{\nu 5} + \frac{\hbar M}{p_{16}} p_{\nu 5} = 0. \tag{106·82}$$

We have therefore the following result: the wave equation

$$(-i\hbar\,\mathbf{grad} - m')\,\psi = 0 \tag{106·91}$$

gives wave functions containing a field momentum vector which is n times the particle momentum 4-vector, where

$$n = m'/ip_{16}. \tag{106·92}$$

Since the field vector has no component p'_{05}, the component p_{05} of the particle momentum vector vanishes by (106·82). This was to be expected since the equations (106·21) or (106·3) are such as to satisfy the equation of continuity in three-dimensional space, and therefore exclude an extra-spatial flow of probability which would be represented by p_{05}.

In deriving (106·81), it is assumed that M commutes with $E_{\nu 5}$ ($\nu = 1, 2, 3, 4$), so that, by § 54 (f), M must be algebraic, as stated earlier, and m' is therefore algebraic. The effect of this limitation is to exclude electromagnetic fields. In a molar field with electromagnetic potential vector κ_μ, M includes terms of the form $\Sigma E_{\mu 5}\kappa_\mu$. To extend the wave equation to electromagnetic fields, we have to introduce interchange circulation. This will be treated in due course.

107. Isostatic compensation

The use of wave functions instead of wave vectors enables us to relax a hampering restriction. Hitherto we have only been able to deal with transitions that do not alter the field energy and momentum. In practice this means that the particles undergoing

transition must be unidentified members of a large assemblage, the other members of the assemblage being in an initial state specified by scale-free characteristics. But a wave function represents an addition both to particle momentum and field momentum; and it may be possible to adjust these additions so that the field remains self-consistent. The equation which imposes the condition that the additional particle and field momentum are so related that the field remains self-consistent (with sufficient accuracy for a first approximation) is called the *wave equation.*

In scale-free physics the field and particle momentum of the initial state have a fixed ratio determined by the multiplicity factor. By taking n in (106·92) equal to this ratio, we obtain wave functions adjusted to fulfil the initial condition. 'Particles' represented by such wave functions can be added in any number to the initial distribution without violating the self-consistency of the field. Or the initial distribution, field and all, can be represented as composed of such particles. By (106·91) the required wave equation is

$$(-i\hbar\Sigma E_{\mu 5}\partial/\partial x_\mu - m')\psi = 0, \tag{107·1}$$

where the field mass m' is n times the particle mass. Since m' is now a constant, the solution has the well-known form

$$\psi = \psi_0 e^{i\theta}, \quad \chi = \chi_0 e^{-i\theta}, \quad \theta = (p'_{15}x_1 + p'_{25}x_2 + p'_{35}x_3 + p'_{45}x_4)/\hbar. \tag{107·2}$$

The term 'uranoid' was derived from analogy with the geoid. Carrying the analogy a little farther, the field momentum and energy is the isostatic compensation of the mountains (object-particles) added to the uranoid. We have not hitherto used this analogy because geodesy is not accustomed to deal with mountains which are 'unidentified members of a large assemblage'. But we are now preparing, formally at least, to deal with individual mountains; and an important step is as far as possible to attach to each individual its own isostatic compensation, instead of the collective compensation that suffices in scale-free physics. In so far as the 'mountain' is an excrescence it is represented by a wave vector; the whole mountain with its underlying isostatic compensation is represented by a wave function. The latter is a self-contained unit which can be dumped anywhere on the uranoid without upsetting the equilibrium of the uranoid particles.

The particular case leading to the wave equation (107·1) is too elementary to be more than a formalism. All that has been accomplished is that, instead of separate descriptions of the particle and field constituents of the energy tensor, we can now give a combined description contained in one pair of wave functions. It will be seen from (107·2) that this is done by making use of the E_{16} rotation, which is idle in the theory of wave vectors, but finds application in the theory of wave functions when θ_{16} is taken to be a function of the coordinates. But the wave functions (107·2), which are infinite plane waves, represent pseudo-discrete states. Thus we are still dealing with large assemblages of particles, and have not yet got down to the problem of individual compensation. The importance of this elementary case is its linkage on the one hand to the scale-free theory and on the other hand to the investigation of angular momentum in § 105, the fertile developments being in systems with integrals of angular momentum.

Since the field momentum contained in a wave function is fully specified by a 4-vector, there is not sufficient flexibility to provide full isostatic compensation for all the

components of the particle momentum vector. The incomplete compensation implies imperfect superposability, or imperfect adjustment to the boundary conditions given by a neutral zero-temperature uranoid; and in general the maladjustment will constitute a perturbation causing the state to decay more or less rapidly. This is not a blemish in the method; if it were confined to perfectly compensated systems, it would overlook metastable states, and states of still less stability, which nevertheless last long enough to be physically important.

We can now state formally the purpose of the wave equation. We want to add to the standard uranoid microscopic systems that satisfy the conditions:

(*a*) The system is itself in equilibrium.

(*b*) Its addition does not disturb the existing equilibrium of the uranoid.

To satisfy (*a*) the flow of probability represented by the stream vector must be a steady circulation, and it must satisfy the equation of continuity[a] as modified to take account of the extra-spatial transfer of probability by interchange. To satisfy (*b*) the added system must include its own isostatic compensation. The conditions (*a*) and (*b*) may therefore be called the conditions of continuity and isostasy. The form (106·21) or (106·3) is determined primarily by the condition of continuity; and the same considerations determine the correcting term (recognised as Coulomb energy) which has to be introduced in order to take account of interchange. The condition of continuity is concerned with particle momentum only. But the wave functions thus determined introduce field energy; and we are able to adjust M so that this supplies the most essential part of the isostatic compensation. The wave equation, as it were, kills two birds with one stone.

108. Wave equation of the hydrogen intracule

For the internal wave equation of the hydrogen atom the coordinates are the relative coordinates $\xi_\alpha = x'_\alpha - x_\alpha$ of the electron and proton; and the domain of the wave function is the ξ-space of § 76. It is four-dimensional, the phase coordinate taking the place of the time coordinate. The coordinate frame is chosen so that the external momenta P_1, P_2, P_3 are zero; that is to say, the atom as a whole is at almost exact rest. Lorentz transformations are inapplicable to relative coordinates; but there is a transformation of ξ-space analogous to the Lorentz transformation of x-space (phase taking the place of time). The intracule is treated as free; this means that the investigation is carried out in system B (§ 27); the transformation of the results to the observational system A has been allowed for in the experimental comparisons in Chapter III.

Although we are dealing with ξ-space, this will not appear explicitly in our notation and terminology, since (conformally with current practice) we shall use the quantum-classical analogy. An occasional reminder will be given that our statements are to be interpreted with the aid of the 'dictionary' in (78·3).

We couple with the wave equation the condition that the field energy $-i\hbar\partial/\partial x_4$ has an exact value (eigenvalue) independent of the coordinates. This condition is very familiar in current theory, but it here appears in a different light. The quantum energy

[a] Steadiness involves the additional condition that s_{16} in (106·4) is independent of t, if the system considered is in x-space. The systems of practical interest are in ξ-space, and time is not concerned. We still have to couple with the wave equation the condition that s_{16} is constant, but for a different reason; s_{16} (or p_{45}) is then the scale and there are obvious reasons for requiring that the scale shall be uniform throughout the distribution.

is only an analogue of the classical energy and is actually a classical scale; moreover it is the quantum specified scale taken over into, and used in, molar physics which has no definable scale of its own. The ultimate object of any quantitative description of the characteristics of an atom must be to express them in terms of the scale of molar experimental measurement; a quantitative description associated with a scale indeterminately or non-uniformly related to the molar scale does not contribute to that end. Thus an eigenscale is essential if we are looking for immediate practical results.

In the elementary wave functions (107·2) the components $p'_{15}, p'_{25}, p'_{35}$ of field momentum are also eigenvalues. Such wave functions are useless in ξ-space, since a steady relative momentum means that the electron separates itself from the proton with which it has been partnered, and loses itself in the environment. For a steady internal state it is an essential condition that $\boldsymbol{p}'_{15}, \boldsymbol{p}'_{25}, \boldsymbol{p}'_{35}$ shall *not* reduce to eigenvalues, and we have to find some other way of characterising particular solutions of the wave equation.

The origin of the x-coordinates is an arbitrary point, and we can freely transform to another origin; but it would be nonsensical to 'change the origin' of relative coordinates. To put it formally: the transformation $x_\alpha \to (x_\alpha + \text{const.})$ is relativistic, but $\xi_\alpha \to (\xi_\alpha + \text{const.})$ is non-relativistic. This is a vital distinction between relative and ordinary space. Since the origin is a distinctive point of ξ-space, angular momentum about the origin has a physical significance which is quite lacking in x-space. In place of the eigenvalue conditions imposed on the linear momenta in x-space in (107·2), we shall impose eigenvalue conditions on the angular momenta in ξ-space.

The wave equation (106·91) will be written as

$$(\boldsymbol{W} - \mu)\,\psi = 0 \quad (\boldsymbol{W} = -i\hbar\,\mathbf{grad}), \tag{108·1}$$

where μ is the mass of the intracule. According to our previous discussion, the algebraic term should be $(k+1)\mu$ in order that the field momentum in the wave functions may isostatically compensate the particle momentum. But in current theory multiplicity factors are absorbed into the constants; and, since we have dealt elsewhere with the problem of disentangling them, we here follow the current practice of ignoring them. It would, in fact, be a vicious circle to insert a multiplicity factor in (108·1), because it is as it stands the equation from which the current (optically controlled) constant μ is obtained.[a]

By (78·7, B 2), p_{15}, p_{25}, p_{35} are real and p_{45}, p_{16} are imaginary. The eigenvalue of the field component p'_{45} will be denoted by $i\epsilon$, so that ϵ is a real energy (classically a real scale momentum). Since $E_{16}p'_{16}$ is here represented by μ, the reality condition is fulfilled by a real constant μ.

The equation (108·1) is to be amended by inserting in $\boldsymbol{W}$ a term which makes allowance for the effect of interchange circulation on the equation of continuity of flow of probability. This term—the Coulomb energy—has been partially discussed in § 33; it depends on quantisation of the interchange angular momentum. It would be

[a] The solution (109·6) of the wave equation determines ϵ/μ; so that if μ is multiplied by $k+1$, ϵ is multiplied by $k+1$. In the reduction from rigid to Galilean coordinates ϵ is divided by k. Thus the factor that has been absorbed is $(k+1)/k = \beta$. When we determine μ by another method, e.g. from the mass-defect of deuterium, compensating β-factors are introduced in the formula employed, as is seen in (98·3).

awkward to develop the theory of quantisation of extra-spatial circulation before treating the quantisation of spatial circulation. We therefore postpone the actual derivation of the Coulomb term to § 110, but use it now in anticipation. The amended operator is

$$\boldsymbol{W} = -i(E_{15}\partial/\partial x_1 + E_{25}\partial/\partial x_2 + E_{35}\partial/\partial x_3 + E_{45}\partial/\partial x_4 + E_{45}\alpha/r), \qquad (108{\cdot}2)$$

where $\alpha = 1/137$. To simplify the formulae we adopt units such that $\hbar = 1$.

From this point up to the end of § 109, the treatment follows familiar lines, and is substantially the same as that first given by Dirac.[a] The purely symbolic treatment (using E-symbols instead of matrix representations) is due to Temple.[b] The symbolic treatment is, I think, the easiest analytically; but our main purpose in giving the investigation here is to complete the *liaison* between atomic quantum theory and the rest of fundamental theory.

The procedure depends on finding four operators $\boldsymbol{W}$, $\boldsymbol{U}_1$, $\boldsymbol{U}_2$, $\boldsymbol{U}_3$, which have a common eigensymbol, so that they reduce to eigenvalues simultaneously. In practice this is taken to be equivalent to the condition that $\boldsymbol{W}$, $\boldsymbol{U}_1$, $\boldsymbol{U}_2$, $\boldsymbol{U}_3$ mutually commute. The latter condition is rather less general, but it is unlikely that any possible solutions have been lost by substituting it. Two operators which have to reduce to eigenvalues are already given, namely (108·2) and

$$\boldsymbol{U}_3 = \partial/\partial x_4 = i\partial/\partial t, \qquad (108{\cdot}3)$$

t being the real phase coordinate. By straightforward test we can verify that two operators commuting with these and with one another are

$$\boldsymbol{U}_1 = -iE_{45}\{E_{23}(x_2\partial/\partial x_3 - x_3\partial/\partial x_2) + E_{31}(x_3\partial/\partial x_1 - x_1\partial/\partial x_3) + E_{12}(x_1\partial/\partial x_2 - x_2\partial/\partial x_1) + 1\}, \qquad (108{\cdot}4)$$

$$\boldsymbol{U}_2 = -i\{x_2\partial/\partial x_3 - x_3\partial/\partial x_2 - \tfrac{1}{2}E_{23}\}. \qquad (108{\cdot}5)$$

The eigenvalues will be denoted by

$$\boldsymbol{W} = \mu, \quad \boldsymbol{U}_1 = j, \quad \boldsymbol{U}_2 = u, \quad \boldsymbol{U}_3 = \epsilon. \qquad (108{\cdot}6)$$

Of these, μ is given as a natural constant; the others are to be evaluated in the course of the investigation. The eigenvalues j, u are pure numbers.

There is a rigorous theorem that commuting matrices have a common eigensymbol. Also the maximum set of commuting E-symbols is an antitetrad, so that we cannot have more than four mutually commuting E-numbers. We are now proceeding as though these theorems extended to symbols containing differential operators. I do not think that any proof has been given. The proof would be only of minor interest; since it matters little whether there is rigorous foundation for our expectation that the procedure will be successful so long as it actually does succeed. Another justification for taking four operators is that by varying μ, j, u, ϵ (not necessarily continuously) we obtain a quadruply infinite system of classification, which suitably replaces the elementary classification by four coordinates or by four momenta.

[a] In current text-books this analysis appears in a theoretical setting which is fundamentally unsound. See, in particular, the author's reply to Dirac, Peierls and Pryce, *Proc. Camb. Phil. Soc.* **38**, 206–8, 1942.

[b] *Proc. Roy. Soc.* A, **127**, 349, 1930.

109. Solution of the wave equation

By direct multiplication,

$$(E_{15}x_1+E_{25}x_2+E_{35}x_3)\left(E_{15}\frac{\partial}{\partial x_1}+E_{25}\frac{\partial}{\partial x_2}+E_{35}\frac{\partial}{\partial x_3}\right)$$
$$=E_{23}\left(x_2\frac{\partial}{\partial x_3}-x_3\frac{\partial}{\partial x_2}\right)+E_{31}\left(x_3\frac{\partial}{\partial x_1}-x_1\frac{\partial}{\partial x_3}\right)+E_{12}\left(x_1\frac{\partial}{\partial x_2}-x_2\frac{\partial}{\partial x_1}\right)$$
$$-\left(x_1\frac{\partial}{\partial x_1}+x_2\frac{\partial}{\partial x_2}+x_3\frac{\partial}{\partial x_3}\right)$$
$$=-iE_{45}\boldsymbol{U}_1-1-r\frac{\partial}{\partial r}. \tag{109·11}$$

We set
$$E_r=(E_{15}x_1+E_{25}x_2+E_{35}x_3)/r, \tag{109·12}$$
so that $E_r^2=-1$, and E_r anticommutes with E_{45}. Then, by (108·2),

$$iE_r\boldsymbol{W}=-\frac{iE_{45}\boldsymbol{U}_1}{r}-\frac{1}{r}-\frac{\partial}{\partial r}+E_rE_{45}\left(\boldsymbol{U}_3+\frac{\alpha}{r}\right). \tag{109·13}$$

Multiplying this finally by the common eigensymbol ψ, $\boldsymbol{W}$, $\boldsymbol{U}_1$, $\boldsymbol{U}_3$ reduce to their eigenvalues μ, j, ϵ, and we have

$$\left(\frac{\partial}{\partial r}+\frac{1}{r}+\frac{iE_{45}j}{r}-E_rE_{45}\left(\epsilon+\frac{\alpha}{r}\right)+iE_r\mu\right)\psi=0. \tag{109·14}$$

Setting $\psi=r^{-1}\phi$, this gives

$$\left(\frac{\partial}{\partial r}+\frac{iE_{45}j}{r}-\frac{E_rE_{45}\alpha}{r}-E_rE_{45}\epsilon+iE_r\mu\right)\phi=0. \tag{109·15}$$

Let
$$F=-E_rE_{45}\epsilon+iE_r\mu,\quad G=iE_{45}j-E_rE_{45}\alpha. \tag{109·21}$$

Then
$$FG+GF=-2\alpha\epsilon,\quad F^2=f^2,\quad G^2=g^2, \tag{109·22}$$

where
$$f^2=\mu^2-\epsilon^2,\quad g^2=j^2-\alpha^2 \tag{109·23}$$

as in (93·41). The wave equation (109·15) becomes

$$\frac{\partial\phi}{\partial r}+\left(F+\frac{G}{r}\right)\phi=0. \tag{109·3}$$

Consider first the case in which the wave function ψ is 'separably algebraic', i.e. having the form $\psi_0 f(x_1,x_2,x_3,x_4)$, where f is an algebraic function and ψ_0 a constant wave vector. Then ϕ is separably algebraic; and the constant wave vector ϕ_0 must be an eigensymbol of $F+G/r$ for all values of r, and therefore an eigensymbol of F and G separately. Then, since F and G have a common eigensymbol, $FG-GF$ is either singular or zero. By (109·21),

$$\tfrac{1}{2}(FG-GF)=iE_r\epsilon j-E_rE_{45}\mu j+iE_{45}\mu\alpha. \tag{109·41}$$

The three terms on the right anticommute; hence, squaring,

$$\tfrac{1}{4}(FG-GF)^2=\epsilon^2j^2-\mu^2j^2+\mu^2\alpha^2. \tag{109·42}$$

Since a singular E-number has no reciprocal, its square cannot be a non-vanishing algebraic number. Hence $\epsilon^2 j^2 - \mu^2 j^2 + \mu^2 \alpha^2 = 0$, or

$$\frac{\epsilon^2}{\mu^2 - \epsilon^2} = \frac{j^2 - \alpha^2}{\alpha^2}, \tag{109·43}$$

which agrees with (93·5). Thus the states which correspond to separably algebraic wave functions are those already found by the method of time-tilt in § 93, namely the metastable states and ground state.

Returning to the general case, we introduce new variables

$$y = 2fr, \quad \chi = e^{\frac{1}{2}y}\phi, \tag{109·51}$$

This is the change of variables used in solving an algebraic equation of the form (109·3), the transformed equation being $\partial\chi/\partial y + G\chi/y = 0$. Here the transformed equation is

$$\frac{\partial\chi}{\partial y} + \frac{G}{y}\chi + \frac{1}{2}\left(\frac{F}{f} - 1\right)\chi = 0, \tag{109·52}$$

the coefficient of the third term being singular instead of zero. Assume a solution in series

$$\chi = \sum_{s=0}^{s=n} C_s y^{s+p}, \tag{109·53}$$

where the C_s may be non-algebraic. We assume a terminating series, because it is found that a non-terminating series diverges. In order that the probability in a region enclosing the origin may be finite, $\int \psi^\dagger \psi r^2 dr$ must be finite; thus ψ must not be more divergent than $r^{-\frac{1}{2}}$, and ϕ and χ must not be more divergent than $r^{-\frac{1}{2}}$ or $y^{-\frac{1}{2}}$. Hence

$$p > -\tfrac{1}{2}. \tag{109·54}$$

Substituting (109·53) in (109·52), we obtain the recurrence relation

$$(s + p + G)\, C_s = -\tfrac{1}{2}(F/f - 1)\, C_{s-1}. \tag{109·55}$$

Putting $s = 0$, $(p + G)\, C_0 = 0$; so that C_0 is an eigensymbol of G and $-p_0$ is an eigenvalue. By (109·22) the only eigenvalues of G are $\pm g$. It will be found later that the lowest value of j is 1; so that, by (109·23), g (taken as positive) is greater than $\frac{1}{2}$. Hence $p = g$, the alternative $p = -g$ being excluded by (109·54).

Putting $s = n + 1$ in (109·55), we have $C_{n+1} = 0$, so that $(F/f - 1)\, C_n = 0$. Hence C_n is an eigensymbol of F, and the corresponding eigenvalue is f.

Multiplying (109·55) by $(F/f + 1)$, and putting $s = n$,

$$(F/f + 1)\,(n + g + G)\, C_n = -\tfrac{1}{2}(F^2/f^2 - 1)\, C_{n-1} = 0 \tag{109·56}$$

by (109·22). Then, since $FG = 2\alpha\epsilon - GF$,

$$\{(n + g)\,(F/f + 1) + G(1 - F/f) + 2\alpha\epsilon/f\}\, C_n = 0. \tag{109·57}$$

Since $FC_n = fC_n$, this gives

$$n + g + \alpha\epsilon/f = 0, \tag{109·58}$$

or, by (109·23),

$$\frac{\epsilon}{(\mu^2 - \epsilon^2)^{\frac{1}{2}}} = n + \frac{(j^2 - \alpha^2)^{\frac{1}{2}}}{\alpha}. \tag{109·6}$$

This is Sommerfeld's general formula covering all states of the hydrogen atom. Here n is a positive integer or zero; but it remains to determine the possible values of j.

Take spherical polar coordinates r, θ, ϕ, so that ϕ is the azimuthal angle in the plane x_2x_3. Then

$$\boldsymbol{U}_2 = -i\frac{\partial}{\partial\phi} + \tfrac{1}{2}iE_{23}, \tag{109·71}$$

and, by (108·6),

$$(\boldsymbol{U}_2 - u)\,\psi = -i(\partial/\partial\phi - \tfrac{1}{2}E_{23} - iu)\,\psi = 0. \tag{109·72}$$

Let $\zeta = e^{-(\frac{1}{2}E_{23}+iu)\phi}\psi$; then, by (109·72), $\partial\zeta/\partial\phi = 0$, so that ζ is independent of ϕ. In particular $\zeta(0) = \zeta(2\pi)$, so that

$$\psi(0) = e^{-(\frac{1}{2}E_{23}+iu)2\pi}\psi(2\pi) = e^{-i(\frac{1}{2}+u)2\pi}\psi(2\pi). \tag{109·73}$$

But wave functions are understood to be single-valued, so that $\psi(2\pi) = \psi(0)$. Hence $e^{-i(\frac{1}{2}+u)2\pi} = 1$. Thus

$$u + \tfrac{1}{2} = \text{integer (positive or negative), or 0.} \tag{109·74}$$

We write $-i(\omega_1, \omega_2, \omega_3)$ for the angular momentum operators, so that

$$\omega_1 = x_2\partial/\partial x_3 - x_3\partial/\partial x_2 - \tfrac{1}{2}E_{23}. \tag{109·81}$$

It is easily verified that
$$\omega_2\omega_3 - \omega_3\omega_2 = -\omega_1. \tag{109·82}$$

By (108·4) and (108·5),

$$iE_{45}\,\boldsymbol{U}_1 = E_{23}\omega_1 + E_{31}\omega_2 + E_{12}\omega_3 - \tfrac{1}{2}, \quad \boldsymbol{U}_2 = -i\omega_1. \tag{109·83}$$

Squaring the first expression we obtain[a]

$$\boldsymbol{U}_1^2 = -\omega_1^2 - \omega_2^2 - \omega_3^2 + \tfrac{1}{4}. \tag{109·84}$$

Hence
$$(\omega_2^2 + \omega_3^2)\,\psi = (-\boldsymbol{U}_1^2 + \boldsymbol{U}_2^2 + \tfrac{1}{4})\,\psi = (-j^2 + u^2 + \tfrac{1}{4})\,\psi. \tag{109·85}$$

Let $\psi_1 = (\omega_2 - i\omega_3)\,\psi$. By (109·82),

$$-i\omega_1(\omega_2 - i\omega_3) = (\omega_2 - i\omega_3)\,(-i\omega_1 - 1),$$

so that
$$\begin{aligned}\boldsymbol{U}_2\psi_1 &= (\omega_2 - i\omega_3)\,(\boldsymbol{U}_2 - 1)\,\psi = (\omega_2 - i\omega_3)\,(u-1)\,\psi \\ &= (u-1)\,\psi_1.\end{aligned} \tag{109·86}$$

Since ω_2 and ω_3 (like $\boldsymbol{U}_2$) commute with $\boldsymbol{U}_1$, $\boldsymbol{U}_3$, $\boldsymbol{W}$, ψ_1 can be substituted for ψ without affecting their eigenvalues. Similarly, if $\psi_{-1} = (\omega_2 + i\omega_3)\,\psi$, ψ_{-1} is an eigensymbol of $\boldsymbol{U}_2$ with eigenvalue $u+1$. Repeating the process, the functions

$$\psi_r = (\omega_2 - i\omega_3)^r\,\psi, \quad \psi_{-r} = (\omega_2 + i\omega_3)^r\,\psi, \tag{109·87}$$

if they do not vanish, are eigensymbols with $\boldsymbol{U}_2 = u - r$ and $u + r$, respectively, the eigenvalues of $\boldsymbol{U}_1$, $\boldsymbol{U}_3$, $\boldsymbol{W}$ being unchanged.

Since $\omega_1^2 = -\boldsymbol{U}_2^2$ the eigenvalues of ω_1^2 are all negative; and by symmetry the eigenvalues of ω_2^2 and ω_3^2 are all negative. It can be deduced that the eigenvalues of $\omega_2^2 + \omega_3^2$ are all negative. Then, by (109·85), $u^2 < j^2 - \frac{1}{4}$; so that for a fixed value of j, the series of eigenfunctions ψ_r, ψ_{-r} must terminate in both directions. The limiting functions ψ_k, $\psi_{-k'}$ are such that

$$(\omega_2 - i\omega_3)\,\psi_k = 0, \quad (\omega_2 + i\omega_3)\,\psi_{-k'} = 0. \tag{109·91}$$

[a] In the corresponding investigation in *Protons and Electrons* an error occurred at this point (between equations (9·441) and (9·442)).

Multiplying these initially by $\omega_2+i\omega_3$ and $\omega_2-i\omega_3$,

$$(\omega_2^2+\omega_3^2+\boldsymbol{U}_2)\,\psi_k=0, \quad (\omega_2^2+\omega_3^2-\boldsymbol{U}_2)\,\psi_{-k'}=0. \tag{109·92}$$

Hence, by (109·85),

$$-j^2+u_k^2+\tfrac{1}{4}+u_k=0, \quad -j^2+u_{-k'}^2+\tfrac{1}{4}-u_{-k'}=0,$$

so that

$$u_k+\tfrac{1}{2}=|\,j\,|, \quad u_{-k'}-\tfrac{1}{2}=-|\,j\,|. \tag{109·93}$$

Since $u+\frac{1}{2}$ is an integer, the possible values of $|\,j\,|$ are integers (excluding 0). The possible values of u range from $|\,j\,|-\frac{1}{2}$ to $-(|\,j\,|-\frac{1}{2})$. This completes the derivation of the Sommerfeld formula.

The inference (used above) that, since the eigenvalues of ω_2^2 and ω_3^2 are all negative, the eigenvalues of $\omega_2^2+\omega_3^2$ are all negative, may be roughly justified as follows. Expectation values, being mean values, range between the greatest and least values; so that the expectation values of ω_2^2 and ω_3^2 are always negative and therefore the expectation values of $\omega_2^2+\omega_3^2$ are always negative. Since an eigenvalue is a particular case of an expectation value, the eigenvalues of $\omega_2^2+\omega_3^2$ are always negative.

110. The interchange momentum

We return to the wave equation (108·1) without the Coulomb term, so that

$$\boldsymbol{W}=-i\hbar(E_{15}\partial/\partial x_1+E_{25}\partial/\partial x_2+E_{35}\partial/\partial x_3+E_{45}\partial/\partial x_4). \tag{110·1}$$

We have to couple with the wave equation $(\boldsymbol{W}-\mu)\,\psi=0$, the eigenscale condition that $\partial/\partial x_4$ reduces to an eigenvalue. We have seen (§ 108) that the momentum component of an intracule in any direction in the 3-space normal to x_4 cannot reduce to an eigenvalue. Thus, for a given ψ, there is only one direction in the 4-space in which the gradient is an eigengradient; and our notation adopts this as the x_4 axis.

A difficulty arises because the 4-space in which the intracule is represented (ξ-space) has relativistic properties analogous to those of space-time, and $\boldsymbol{W}$ is correspondingly invariant for transformations analogous to Lorentz transformations. Just as there is no means of defining an absolute time direction in x-space, there is no means of defining an absolute phase direction in ξ-space. Until we have defined the direction to be denoted by x_4, it is premature to impose the condition that the gradient in the x_4 direction, and that direction only, is an eigengradient.

Since we can only define relative directions, the phase direction for a particular eigenstate (which may be called the 'proper phase', by analogy with proper time) must be defined relatively to the directions of other characteristics of the eigenstate. The only other characteristic helpful in defining orientation is the angular momentum. The angular momentum (expectation value) is a 6-vector; and a 6-vector can always be reduced to two antiperpendicular components $E_{23}\Omega_{23}+E_{14}\Omega_{14}$, defining a pair of antiperpendicular planes which may be taken as the coordinate planes x_2x_3 and x_1x_4. The two planes are distinguished from one another by the fact that their relativity rotations are respectively circular and hyperbolic. We can then define the proper phase direction x_4 to be a time-like axis in the fixed plane x_1x_4; but this leaves its particular direction in the plane undefined; and the system contains no characteristic relative to which it could be defined.

Thus the eigenscale condition is not that a particular component $-i\hbar\partial/\partial x_4$ reduces to an eigenvalue, but that the component in some indeterminate direction in a

particular plane x_1x_4 reduces to an eigenvalue. This is the phenomenon of *symmetrical degeneracy*, well known in other contexts. It is treated by a procedure analogous to the ignoration of coordinates in classical dynamics. If x_4' is an arbitrary fixed axis in the plane x_1x_4, the angle θ between x_4 and x_4' is indeterminate. Since θ cannot be determined, the substates corresponding to different assumed values of θ are run together into one state. Then θ, instead of being the label of a substate, becomes an internal coordinate of the state; and the angular momentum conjugate to it appears in the hamiltonian of the state.

The angular momentum conjugate to θ has the following significance. If the substates corresponding to different values of θ were distinguished, transitions between them would occur. It is well known that as the difference of the characteristics of two states decreases the frequency of transition increases. Here we have an extreme case in which the difference disappears entirely, the characteristic θ which nominally distinguishes the substates being unobservable. Owing to the high frequency of transition, the energy of the transition circulation (interstate energy) will form a large part of the total energy of the intracule; and the purpose of the analysis into states is stultified. The remedy is to introduce the combined state in which θ is an internal coordinate; so that what was a transition circulation between substates becomes an internal circulation in the combined state, and its energy is included in the state energy determined by the wave functions of the combined state.

If ψ_0 is a wave vector referred to the fixed frame $x_1'x_2x_3x_4'$, it becomes

$$\psi = e^{-\frac{1}{2}E_{14}\theta}\psi_0 \tag{110·2}$$

when referred to the frame $x_1x_2x_3x_4$ rotated through the angle θ (in the sense $x_1 \to x_4$) relatively to the fixed frame. By (105·5) the angular momentum in the plane x_1x_4 is $-i\hbar(\partial/\partial\theta - \frac{1}{2}E_{14})$, which when applied to (110·2) gives $i\hbar E_{14}$. Thus a wave function $\psi(x_1, x_2, x_3, x_4)$ referred to an indeterminate frame contains an angular momentum $i\hbar E_{14}$ resulting from the indeterminacy. By taking this momentum into account explicitly we correct for the indeterminacy, so that the direction x_4 is thenceforth to be treated as a determinate direction. Accordingly our procedure is to introduce a modified operator $\boldsymbol{W}_0$, derived from $\boldsymbol{W}$ by including a term corresponding to the angular momentum $E_{14}i\hbar$. The indeterminacy of x_4 having been eliminated, we can now apply the condition that $\partial/\partial x_4$ reduces to an eigenvalue $i\epsilon$. This eliminates the coordinate x_4 in the wave equation $(\boldsymbol{W}_0 - \mu)\psi = 0$; so that its solutions give the usual wave functions $\psi(x_1, x_2, x_3)$ representing distribution over a 3-space. This 3-space can be taken to correspond to $x_4 = 0$.

At a point $(x_1, x_2, x_3, 0)$ the angular momentum $E_{14}i\hbar$ in the plane x_1x_4 gives a linear momentum $E_{45}i\hbar/x_1$ in the x_4 direction. Before combining this with $\boldsymbol{W}$, we have to make an adjustment on account of the change of volume element introduced by ignoration of θ.

Using polar coordinates, the usual wave functions are such that $\psi^\dagger(r^2 dr d\omega)\psi$ is the probability in an element $dr d\omega$. When the probability distribution is spread over four dimensions by the indeterminacy of orientation of the 3-space, the sphere of radius r is spread into a hypersphere.[a] Treating the section $x_4 = 0$ as the equatorial plane of

[a] Strictly a hyperhyperboloid.

the hypersphere, we denote the latitude by λ. The ordinary wave functions must then be interpreted as giving the probability distribution in an equatorial zone of infinitesimal width $d\lambda$; so that $\psi^\dagger(r^2 dr d\omega d\lambda)\psi$ is the probability in an element $dr d\omega d\lambda$ at $\lambda = 0$, multiplied by a constant factor due to the renormalisation. But in (110·2) the wave function ψ_0 is extended to four dimensions in a different way, the distribution over $x_1 x_2 x_3$ being treated as occupying a *lune* of the hypersphere of infinitesimal width $d\theta$. We have

$$dx_4 = r d\lambda = x_1 d\theta. \tag{110·3}$$

The linear momentum in the element $dr d\omega d\theta$ must be multiplied by $d\lambda/d\theta = x_1/r$, in order to give the momentum in the usual element $dr d\omega d\lambda$. The new momentum to be combined with $\boldsymbol{W}$ is accordingly

$$E_{45}\frac{i\hbar}{x_1}\frac{x_1}{r} = E_{45}\frac{i\hbar}{r}. \tag{110·4}$$

The momentum (110·4) is to be subtracted from $\boldsymbol{W}$. The momentum $-i\hbar\partial/\partial x_4$ which is put equal to $-i\epsilon$ is the whole momentum in the E_{45} direction including the momentum $i\hbar/r$ found in (110·4). Thus $-i\hbar\partial/\partial x_4 - i\hbar/r$ is the momentum which remains when the part contributed by transition circulation between the substates is excluded. This remainder corresponds to the internal flow in a substate which satisfies the elementary equation of continuity on which the wave equation $(\boldsymbol{W}-\mu)\psi = 0$ was based. Thus by substituting

$$-i\hbar\frac{\partial}{\partial x_4} - \frac{i\hbar}{r} \quad \text{for} \quad -i\hbar\frac{\partial}{\partial x_4} \tag{110·5}$$

in $\boldsymbol{W}$, we obtain the modified operator $\boldsymbol{W}_0$ in which the momentum due to the indeterminacy of the frame is eliminated; so that $\boldsymbol{W}_0$ satisfies the elementary equation $(\boldsymbol{W}_0 - \mu)\psi = 0$ derived for a fixed frame.

As shown in § 33, the momentum $-i\hbar/r$ is divided by 137 in the reduction from rigid coordinates to the observational system A; so that in the current wave equation, the substitution is

$$-i\hbar\left(\frac{\partial}{\partial x_4} + \frac{\alpha}{r}\right) \quad \text{for} \quad -i\hbar\frac{\partial}{\partial x_4} \quad (\alpha = \tfrac{1}{137}). \tag{110·6}$$

This gives the result (108·2) which we used in anticipation in the theory of the hydrogen intracule.

Attention should be called to a new point that we were not aware of in the discussion in § 33. The momentum $-i\hbar/r$ was there divided by 137 on account of the transformation of the time coordinate. But here x_4 is the phase, not the time, direction. But since we have freely employed the quantum-classical analogy in which the phase is the time-analogue, it is implied that the phase is to be transformed in the same way as the time in order to preserve the analogy. Thus the theory of the factor α in § 33 is not invalidated by the substitution of the time-analogue for the actual time. The other part, $-i\hbar\partial/\partial x_4$ of the momentum in the x_4 direction, is also divided by 137 in the transformation from rigid coordinates to the observational system; but this occurs automatically, without any alteration of form, through the substitution of the coordinate x_4 in the observational system for the rigid coordinate x_4. In any case the transformation scarcely concerns us, since $-i\hbar\partial/\partial x_4$ is immediately replaced by its eigenvalue $-i\epsilon$, and ϵ is defined as the energy in the observational system.

The transition circulation between the substates, which provides the Coulomb term α/r in the wave equation is the interchange circulation treated from a new aspect. Its plane E_{14} agrees with that given in (93·3) in the investigation of metastable states by the method of time-tilt.

111. The two-frame transformation

Let I be the interchange operator (81·1) of the E- and F-symbols; and let

$$G_{\mu 5} = \frac{I(E_{\mu 5} + iF_{\mu 5})}{1+i}, \quad H_{\mu 5} = \frac{(E_{\mu 5} + iF_{\mu 5})I}{1+i} \quad (\mu = 1,2,3,4,0). \tag{111·11}$$

We can easily derive the inverse formula

$$E_{\mu 5} = \frac{I(G_{\mu 5} - iH_{\mu 5})}{1-i}, \quad F_{\mu 5} = \frac{(G_{\mu 5} - iH_{\mu 5})I}{1-i}. \tag{111·12}$$

We shall first show that $G_{\mu 5}$, $H_{\mu 5}$ have the properties of pentads. The complete sets G_μ, H_μ generated by these pentads commute with one another, and have the same interchange operator I as E_μ, F_μ. The G- and H-frames have opposite chirality to the E- and F-frames. We have

$$\begin{aligned} G_{\mu 5}^2 &= -\tfrac{1}{2}iI(E_{\mu 5} + iF_{\mu 5})\, I(E_{\mu 5} + iF_{\mu 5}) \\ &= -\tfrac{1}{2}i(F_{\mu 5} + iE_{\mu 5})\,(E_{\mu 5} + iF_{\mu 5}) \\ &= -1, \end{aligned} \tag{111·21}$$

$$\begin{aligned} G_{\mu 5}G_{\nu 5} + G_{\nu 5}G_{\mu 5} &= -\tfrac{1}{2}i\{(F_{\mu 5} + iE_{\mu 5})\,(E_{\nu 5} + iF_{\nu 5}) + (F_{\nu 5} + iE_{\nu 5})\,(E_{\mu 5} + iF_{\mu 5})\} \\ &= 0 \quad (\mu \neq \nu). \end{aligned} \tag{111·22}$$

We set $E_{16} = F_{16} = i'$, where i' may or may not denote the same square root of -1 as i. Then $i'E_{05} = E_{15}E_{25}E_{35}E_{45}$, $i'F_{05} = F_{15}F_{25}F_{35}F_{45}$. Let

$$i'G'_{05} = G_{15}G_{25}G_{35}G_{45}. \tag{111·31}$$

Then $G'^2_{05} = -1$, and

$$i'G'_{05} = -\tfrac{1}{4}(F_{15} + iE_{15})\,(E_{25} + iF_{25})\,(F_{35} + iE_{35})\,(E_{45} + iF_{45}). \tag{111·32}$$

This gives

$$E_{05}G'_{05} = -\tfrac{1}{4}(1 - iE_{15}F_{15})\,(1 - iE_{25}F_{25})\,(1 - iE_{35}F_{35})\,(1 - iE_{45}F_{45}),$$

$$iF_{05}G'_{05} = -\tfrac{1}{4}i(1 + iE_{15}F_{15})\,(1 + iE_{25}F_{25})\,(1 + iE_{35}F_{35})\,(1 + iE_{45}F_{45}).$$

When the right-hand sides are multiplied out, both expressions consist of 16 terms containing different symbols $E_\sigma F_\sigma$. The sum is found to be $(1+i)I$; hence

$$(E_{05} + iF_{05})\,G'_{05} = (1+i)\,I. \tag{111·33}$$

Multiplying initially by I and finally by G'_{05},

$$-I(E_{05} + iF_{05}) = (1+i)\,G'_{05}.$$

Hence, by (111·11),

$$G_{05} = -G'_{05} = i'G_{15}G_{25}G_{35}G_{45}. \tag{111·41}$$

The results (111·21), (111·22) and (111·41) show that $G_{\mu 5}$ is a pentad, and $G_{16} = -i'$. Similar results are found for $H_{\mu 5}$. It is easily verified that the symbols $G_{\mu 5}$ commute with the symbols $H_{\mu 5}$.

The two pentads generate complete sets G_μ, H_μ; but the equations (111·11), (111·12) apply only to the generating pentads. We see at once from (111·11) that $H_{\mu 5} = IG_{\mu 5}I$; and this relation extends to all the symbols, e.g. $H_{12} = H_{15}H_{25} = IG_{15}IIG_{25}I = IG_{12}I$. Thus I is the interchange operator of the G- and H-frames; so that

$$\tfrac{1}{4}\Sigma G_\mu H_\mu = I = \tfrac{1}{4}\Sigma E_\mu F_\mu. \tag{111·42}$$

Consider the 5-vectors ($\mu = 15, 25, 35, 45, 05$),

$$\boldsymbol{p} = \Sigma E_\mu p_\mu, \quad \boldsymbol{p}' = \Sigma F_\mu p'_\mu, \quad \boldsymbol{P} = \Sigma G_\mu P_\mu, \quad \boldsymbol{\varpi} = \Sigma H_\mu \varpi_\mu, \tag{111·51}$$

where

$$P_\mu = \frac{1}{\sqrt{2}}(p_\mu + p'_\mu), \quad \varpi_\mu = \frac{1}{\sqrt{2}}(p'_\mu - p_\mu). \tag{111·52}$$

Then

$$I(E_\mu p_\mu + iF_\mu p'_\mu) = \tfrac{1}{2}I\{(E_\mu + iF_\mu)(p_\mu + p'_\mu) + i(F_\mu + iE_\mu)(p'_\mu - p_\mu)\}$$

$$= \frac{1}{\sqrt{2}}(1+i)(G_\mu P_\mu + iH_\mu \varpi_\mu), \tag{111·53}$$

so that

$$i^{\frac{1}{2}}(\boldsymbol{P} + i\boldsymbol{\varpi}) = I(\boldsymbol{p} + i\boldsymbol{p}'). \tag{111·54}$$

The relation can also be written as

$$i^{\frac{1}{2}}\boldsymbol{P} + i^{\frac{1}{2}}\boldsymbol{\varpi} = \tilde{\boldsymbol{p}} + i\tilde{\boldsymbol{p}}' \tag{111·55}$$

by (82·6).[a]

It appears from (111·52) that the transformation of the double frame from EF to GH accompanies a transformation of two ordinary particles into external and internal particles. It is subject to two limitations. First, it applies to two particles of equal mass, e.g. two satellite electrons in an atom or two protons in a nucleus. Secondly, the simple transformation (111·54) applies only to the linear momentum vector. This suggests that its proper application is to field momentum. Since i has not been identified with i', the sign of i can be reversed in (111·54); so that it is equivalent to two real equations determining $\boldsymbol{P}$ and $\boldsymbol{\varpi}$ separately.

For practical application we need a corresponding relation between strain vectors. As in (72·84), the strain vector form of the wave equation is $(H - \epsilon)\psi = 0$, where, if there is no interchange energy,

$$i'H = E_{14}p_1 + E_{24}p_2 + E_{34}p_3 + E_{45}m \quad (p_\alpha = -i\hbar\partial/\partial x_\alpha). \tag{111·61}$$

Consider two non-interacting particles with wave equations

$$(H_x - \epsilon)\psi = 0, \quad (H_{x'} - \epsilon')\psi' = 0$$

in the E-frame and F-frame respectively. The two equations can be combined as

$$(H_x \pm iH_{x'} - \epsilon_0)\Psi = 0, \tag{111·62}$$

where Ψ is the double wave function $\psi\psi'$. We wish to transform (111·62) into a double wave equation for the extracule and intracule, having a corresponding form

$$(H_X \pm iH_\xi - \eta_0)\Phi = 0. \tag{111·63}$$

We are only concerned with the transformation of the 'hamiltonian' $H_x \pm iH_{x'}$ into $H_X \pm iH_\xi$, because the eigenfunctions Φ and eigenvalues η_0 will be determined by solving the equation (111·63), not by transforming Ψ and ϵ_0. Eigenstates of the

[a] In (111·54) $\boldsymbol{p}$ is associated with the F-frame and $\boldsymbol{p}'$ with the E-frame.

extracule and intracule do not in general correspond to eigenstates of the original particles; so that there is no question of transforming the quantities Ψ, ϵ which are characteristic of an eigenstate.

To suit (111·61) we must use the pentad 4 instead of 5 as generating pentad of the GH-frame; so that the suffix $\mu 4$ is substituted for $\mu 5$ in (111·11). We drop the factor $\sqrt{2}$ in (111·52), so that

$$P_\mu = p_\mu + p'_\mu, \quad \varpi_\mu = p'_\mu - p_\mu. \tag{111·64}$$

In the ordinary two-particle transformation (26·12) for particles of equal mass $P_\mu = p_\mu + p'_\mu$, but $\varpi_\mu = \frac{1}{2}(p'_\mu - p_\mu)$. Thus H_ξ as here defined is twice the usual hamiltonian of the intracule.

Associating H_X with the G-frame and H_ξ with the H-frame we obtain, by (111·53),

$$I(H_x + iH_{x'}) = \tfrac{1}{2}(1+i)(H_X + iH_\xi), \tag{111·71}$$

with

$$\left.\begin{aligned} i'H_X &= G_{14}P_1 + G_{24}P_2 + G_{34}P_3 + G_{45}2m, \\ i'H_\xi &= H_{14}\varpi_1 + H_{24}\varpi_2 + H_{34}\varpi_3. \end{aligned}\right\} \tag{111·72}$$

The H_{45} term vanishes by (111·64), the masses m of the two particles being equal. It will be remembered that in the two-particle transformation the intracule is bound; so that it is correct that there should be no rest-mass term in H_ξ.

Thus far the analysis has only formal interest. But we can now insert the Coulomb term in the hamiltonian H_ξ of the intracule, and use (111·71) to transform to the x, x' representation. By § 110 the effect of the Coulomb term is to change ϵ to $\epsilon - 2\alpha\hbar/r$ in the wave equation $(H_\xi - \epsilon)\psi = 0$. The term is here doubled, because H_ξ is twice the usual hamiltonian and ϵ is twice the actual energy. It has also been reversed in sign, because we are now considering like charges.[a] The Coulomb correction is to be included in H_ξ, and gives an addition

$$\delta H_\xi = 2\alpha\hbar/r = 2e^2/cr. \tag{111·81}$$

Hence by (111·71), the increments of H_x, $H_{x'}$ due to the Coulomb correction are given by

$$I(\delta H_x + i\delta H_{x'}) = (1+i)ie^2/cr. \tag{111·82}$$

The equation with $-i$ substituted for i must also be satisfied. Hence

$$\delta H_x = -Ie^2/cr, \quad \delta H_{x'} = Ie^2/cr. \tag{111·83}$$

The term δH_x depends partly on the coordinates of the accented particle, and will therefore be regarded as a perturbing energy in the wave equation $(H_x - \epsilon)\psi = 0$ of the accented particle. We have therefore the result that the interchange energy e^2/cr found in the relative hamiltonian of two electrons (or two protons) appears also as a perturbing energy in the hamiltonian of either electron (or proton); but whereas in the relative hamiltonian it is an algebraic term, as a perturbing energy it has the symbolic coefficient $\pm I$.

The ambiguity does not mean that the sign is a matter of indifference, but that the convention necessary to define the sign has not yet been introduced. If the coordinate of one particle is not measured from the other particle it is measured from the physical origin. In (111·64) the sign of ϖ is such that the unaccented particle is treated as origin

[a] The Coulomb energy of like charges was deduced from that of unlike charges in § 49.

for the accented particle—not vice versa. It is therefore the coordinate x' that is comparable with ξ; so that, by (111·83), the positive sign should be taken. Accordingly the perturbing Coulomb energy is

$$Ie^2/cr. \tag{111·9}$$

Evidently the same rule applies to non-Coulombian energy. In the relative wave equation of two protons (used in the calculation of scattering of protons by protons) it is the algebraic quantity Ae^{-r^2/k^2}, if the equation is in strain-vector form. When introduced as a perturbing energy in the wave equation of one of the protons (also in strain-vector form), it is $IAe^{-k^2r^2}$.

As shown in § 81, I is resolvable into two factors which may be distinguished as the spin and co-spin factors. In extra-nuclear physics only the spin factor (Dirac's interchange operator) is used; presumably there is some justification for assuming that in these conditions the co-spin factor reduces to 1. It is quite common to use the complete operator I, given by fundamental theory, for the interaction between protons in a nucleus.

112. Electromagnetic potentials

According to the last section, the hamiltonian of an object-electron perturbed by other electrons is

$$H = H_0 + \Sigma_s \delta H_s, \tag{112·11}$$

where

$$iH_0 = -i\hbar\left(E_{14}\frac{\partial}{\partial x_1} + E_{24}\frac{\partial}{\partial x_2} + E_{34}\frac{\partial}{\partial x_3}\right) + E_{45}m, \tag{112·12}$$

$$\delta H_s = I_s e^2/cr_s, \quad I_s = \tfrac{1}{4}\Sigma_\mu E_\mu (F_\mu)_s, \tag{112·13}$$

and r_s is the distance of the sth electron. The object-electron is in the frame E_μ; and the frames $(F_\mu)_s$ of the perturbing electrons commute with E_μ and with one another.

The unperturbed hamiltonian H_0 postulates a neutral environment. Since we cannot create negative charges without counterbalancing positive charges, perturbations due to positive charges should be included in (112·11). The analysis of § 111 applies to particles of equal mass; so that positrons must be employed as the carriers of the positive charges. There, however, need not be the unattached positrons observed in a Wilson chamber. Normally the object-electron will be in an environment of protons and electrons; but, for the purpose of perturbation theory, we represent the proton as the sum of a positron and a neutral particle (not a neutron). The neutral particle is part of the standard environment postulated for the unperturbed hamiltonian; so that only the positron contributes to $\Sigma\delta H_s$.

We shall investigate the perturbation due to a very large number of electrons and positrons, none of which is near enough for its individual effect to be significant. The calculation of $\Sigma\delta H_s$ is then simplified by statistical averaging. By (73·25) $-is_{16}F_\mu$ is the operational form of the particle momentum strain vector of the particle in the F-frame. The averaging consists in replacing F_μ by its expectation value is_μ/s_{16}, neglecting the fluctuation which will become insignificant when the Σ_s summation is applied.[a] As in (106·4), s_{16} is the density and s_{14}, s_{24}, s_{34} is the flux; so that $(s_{14}, s_{24}, s_{34})/s_{16}$

[a] The procedure also neglects correlation of E_μ and F_μ. Reference to this will be made later.

is the velocity u_1, u_2, u_3 of the carrier of s_μ. Denoting by (I) the part of I corresponding to $\mu = 14, 24, 34, 16$, we have, after averaging,

$$(I) = \tfrac{1}{4}i(E_{14}u_1 + E_{24}u_2 + E_{34}u_3 + E_{16}). \tag{112·2}$$

By (112·13), this gives a perturbation energy

$$(\Sigma_s \delta H_s) = \tfrac{1}{4}i(E_{14}e_0\kappa_1 + E_{24}e_0\kappa_2 + E_{34}e_0\kappa_3 + ie_0\kappa_4), \tag{112·3}$$

where e_0 is the charge $(-e)$ of the object-electron, and

$$\kappa_1, \kappa_2, \kappa_3 = \Sigma \frac{e_s(u_1, u_2, u_3)_s}{cr_s}, \quad \kappa_4 = \Sigma \frac{e_s}{cr_s}. \tag{112·4}$$

This is the classical formula for the electromagnetic potential vector κ_μ, disregarding retardation.

Including (112·3) the wave equation $(H - \epsilon)\psi = 0$ becomes

$$\left\{E_{14}\left(-i\hbar\frac{\partial}{\partial x_1} - \tfrac{1}{4}e_0\kappa_1\right) + E_{24}\left(-i\hbar\frac{\partial}{\partial x_2} - \tfrac{1}{4}e_0\kappa_2\right)\right.$$
$$\left. + E_{34}\left(-i\hbar\frac{\partial}{\partial x_3} - \tfrac{1}{4}e_0\kappa_3\right) - i(\epsilon - \tfrac{1}{4}e_0\kappa_4) + E_{45}m\right\}\psi = 0. \tag{112·5}$$

Thus the wave equation for an electron in a molar electromagnetic field of (unretarded) potential κ_μ is obtained by substituting

$$-i\hbar\frac{\partial}{\partial x_\mu} - \tfrac{1}{4}e_0\kappa_\mu \quad \text{for} \quad -i\hbar\frac{\partial}{\partial x_\mu} \quad (\mu = 1, 2, 3, 4) \tag{112·6}$$

in the elementary wave equation.

We here reach an important junction with current theory; since, apart from the factor $\frac{1}{4}$, this is the result currently accepted. The factor $\frac{1}{4}$ is very important, but it must be considered in conjunction with the other numerical factors (multiplicity and β-factors) which current wave mechanics ignores. The applications of (112·6) are treated later.

Equation (112·5) does not pretend to be the complete solution of the problem of finding the resultant perturbation due to a large number of distant charges. For example, it does not include the perturbation responsible for the Compton effect. The part of the perturbing field taken into account in (112·5) is generally called the *longitudinal field*; there remains a *transverse* or *radiation field* which is investigated by other methods. The distinctive feature of the part of the perturbation given by (112·3) is that it depends only on characteristics of the perturbing system familiar in molar measurement, namely the distribution of charge density and current density. If we introduce averaged values of other components of F_μ, little or no progress is made; because there still remains the theoretical problem of connecting these averaged values with quantities that could be measured observationally.

There is another ground for discriminating between the terms here treated and the rest of the perturbation. The function of the wave equation is to determine a field momentum 4-vector isostatically compensating the corresponding part of the particle momentum vector; and it is this part of the particle momentum vector that we have

taken into account in (ΣH_s). When we deal with the uncompensated components new considerations arise. The lack of compensation involves instability of the states, with a consequent interstate energy and an associated radiation field. Thus the determination of the part of the perturbation not included in (112·5) is bound up with the quantum theory of radiation.

By averaging F_μ without reference to its co-factor E_μ in I, we effectively assume that the average value of $E_\mu F_\mu$ is the product of the average values of E_μ and F_μ, or that the state of the perturbing system is uninfluenced by the state of the object-electron. Actually the object-electron would polarise the perturbing system. If therefore there is sensible energy due to interaction between the object-particle and the charge and current which it itself induces in the environment (Debye-Hückel energy), this has been neglected in (112·5). In fact, the current wave equation postulates a rigid electromagnetic as well as a rigid gravitational field.

An object-proton can be treated in a similar way; and its wave equation is modified in accordance with (112·6).

Chapter XI

THE MOLAR ELECTROMAGNETIC FIELD

113. Gauge transformations (molar theory)

A molar electromagnetic field is specified by a 4-vector potential κ_μ, whose curl $F_{\mu\nu}$ gives the magnetic and electric force. In 1918 H. Weyl[a] introduced a unified geometry of gravitational and electromagnetic fields, in which κ_μ was identified with a gauge vector concerned in the definition of the local standard of length. It has since appeared that Weyl's gauge vector should have been identified with $i\kappa_\mu$; so that real electromagnetic potentials correspond to complex gauge. With this amendment Weyl's representation is an important auxiliary in quantum theory.

In 1921 I put forward a generalisation of Weyl's theory.[b] This has been further developed by Einstein and others. Probably the most far-reaching extension is Schrödinger's recent theory.[c] This kind of development has a rightful place in theoretical physics; but I do not think it is the road to quantum theory, and it will not much concern us here. Just as quantum mechanics links up with intermediate rather than with general relativity, so quantum electrodynamics links up with Weyl's elementary gauge theory rather than with my generalisation of it. Investigators have had the idea that by developing geometries of the widest generality imaginable they may hit on the complication responsible for the phenomenon of atomicity. But we should rather look to the other end—to specially simplified conditions. For, in statistical physics, simplicity is paradoxically the most potent source of complication since it gives rise to degeneracy.

I shall first summarise the molar theory of gauge transformation. The coordinates are unaltered but the measure ds of an interval is changed to $ds' = \gamma ds$, where γ is a function of the coordinates. We then have

$$x'_\mu = x_\mu, \quad g'_{\mu\nu} = \gamma^2 g_{\mu\nu}, \quad g'^{\mu\nu} = \gamma^{-2} g^{\mu\nu}, \quad \sqrt{-g'} = \gamma^4 \sqrt{-g}. \tag{113·1}$$

As in § 87 we introduce 'parallel displacement' of vectors. The relation described as parallelism is defined only for small distances and is not integrable. It sets up a one-to-one correspondence of vectors at neighbouring points. Take a line element at a point x_μ, and transfer it by parallel displacement to $x_\mu + dx_\mu$. Let its initial length measured with the local standard at x_μ be l, and its final length measured with the local standard at $x_\mu + dx_\mu$ be $l + dl$. Evidently dl must be proportional to l; and since dl/l is a function of dx_1, dx_2, dx_3, dx_4, we can set

$$dl/l = \sigma_1 dx_1 + \sigma_2 dx_2 + \sigma_3 dx_3 + \sigma_4 dx_4, \tag{113·21}$$

where the σ_μ are functions of the coordinates. It is here assumed that dl/l is independent of the orientation of the line element; this is a simplifying condition in Weyl's theory, which is discarded in the generalised theories mentioned above. Using the summation convention, we have

$$l = l_0 e^\theta, \quad \theta = \int \sigma_\mu dx_\mu. \tag{113·22}$$

[a] *Berlin Sitzungsberichte*, 30 May 1918.

[b] *Proc. Roy. Soc.* A, **99**, 104, 1921.

[c] *Proc. Roy. Irish Acad.* Sect. A, **49**, 53, 1943; 225, 1944; 237, 1944; 275, 1944; **50**, 143, 1945.

Since $\sigma_\mu dx_\mu$ is not necessarily a perfect differential, l at a distant point is not determinate unless a route of transfer is specified.

Applying a gauge transformation, l becomes $l' = \gamma l$; so that

$$\sigma'_\mu dx_\mu = dl'/l = dl/l + d\gamma/\gamma. \tag{113·23}$$

Hence
$$\sigma'_\mu = \sigma_\mu + \partial(\log\gamma)/dx_\mu \tag{113·24}$$

and
$$\operatorname{curl}\sigma'_\mu = \operatorname{curl}\sigma_\mu. \tag{113·25}$$

In classical electromagnetic theory the potential vector κ_μ is indeterminate to the extent of an arbitrary additive gradient. By (113·24) σ_μ is likewise indeterminate to the extent of an arbitrary additive gradient if arbitrary change of gauge system is admissible. The principle of Weyl's theory is an identification of κ_μ with σ_μ, so that the characteristic indeterminacy of the potential 4-vector is represented by an indeterminacy of gauge system. No hypothesis is involved in this identification; it is simply a question of employing a geometrical representation of fundamental electromagnetic quantities which suitably exhibits their characteristic properties. But it will not advance the unification of physics unless it harmonises with a representation of the mechanico-metrical quantities which exhibits their characteristic properties Here lies a difficulty, because in the mechanico-metrical part of physics there is no indeterminacy of gauge system. Lengths and times are defined everywhere by a quantum-specified standard in the locality.

The difficulty is overcome by identifying σ_μ with $i\kappa_\mu$. Then (113·24) gives

$$\kappa'_\mu = \kappa_\mu + \partial\phi/\partial x_\mu \quad (\gamma = e^{i\phi}), \tag{113·31}$$

and the indeterminacy of measure associated with the indeterminacy of electromagnetic potential κ_μ is expressed by gauge transformations $e^{i\phi}$, where ϕ is a real function of the coordinates. In the transformations $ds' = e^{i\phi}ds$,

$$|\,ds\,| \text{ is invariant.} \tag{113·32}$$

Comparison with the quantum-specified standard furnishes a measure which is necessarily real. It determines the invariant modulus, but leaves the complex argument of ds indeterminate.

Hitherto we have fixed the complex argument of ds, and of other physical quantities by 'reality conditions'; thus ds is imaginary for a length, and real for a time. Reality conditions inhibit the gauge transformations $ds \to e^{i\phi}ds$, on which unified theory is based. But we have seen that reality conditions are boundary conditions furnished by the neutral uranoid. If we retain the same conditions when an electromagnetic field is present, it is implied that the field is included in the object-system, so that the environment of the object-system is still a neutral uranoid. This is the normal procedure. But to place κ_μ in the object-system, and $g_{\mu\nu}$ (or at least the inertial part of $g_{\mu\nu}$) in the environment, is the reverse of unified treatment. In unified theory the environment is the whole non-neutral distribution of matter responsible for the combined electromagnetic and inertial-gravitational field; and the standard reality conditions are then inapplicable. Gauge transformations are no longer inhibited.

It will be seen that (with the amended identification of σ) a gauge system is a 'reality system'. In a neutral environment one particular reality system stands out as the

least complicated way of connecting the mathematical variates with observational measures; it gives the standard reality conditions. In non-neutral environment no reality system is distinguished as being simpler than the others. Our inability to define a preferential reality system has to be met in the same way as our inability to define a preferential time axis. When we are unable to give a physical definition of the particular time-axis we are using, we can only employ equations which are invariant for transformations of the time axis, so that they are true however the axis is identified; similarly, when we are unable to define the reality we are using, we can only employ equations which are invariant for transformations of the reality system. Thus a non-neutral environment forces on us the study of gauge transformations or transformations of reality system; and particular use is made of tensors and equations which are gauge-invariant.

Since the electromagnetic force $F_{\mu\nu}$ is curl κ_μ, it follows from (113·31) that

$$F'_{\mu\nu} = F_{\mu\nu}. \tag{113·41}$$

Thus $F_{\mu\nu}$ is a gauge-invariant tensor. On the other hand, $F'^{\mu\nu} = \gamma^{-4}F^{\mu\nu}$. Both κ_μ and $F_{\mu\nu}$ are real in the gauge system in which the σ_μ that have been identified with $i\kappa_\mu$ were first defined. Hence, by (113·31) and (113·41), they are real in all gauge systems. But $F^{\mu\nu}$ is not in general real.

The next step is to find gauge-invariant tensors containing $g_{\mu\nu}$.[a] We introduce a modified 3-index symbol

$$^*\{\mu\nu, \alpha\} = \{\mu\nu, \alpha\} + i(g_{\mu\nu}\kappa^\alpha - g_\mu{}^\alpha\kappa_\nu - g_\nu{}^\alpha\kappa_\mu), \tag{113·42}$$

which is easily shown to be gauge-invariant. The gauge-invariant R.C. tensor, obtained by substituting $^*\{\ldots\}$ for $\{\ldots\}$ in the ordinary R.C. tensor, is a rather complicated expression which gives on contraction the gauge-invariant Einstein tensor

$$^*G_{\mu\nu} = G_{\mu\nu} + 2\kappa_\mu\kappa_\nu - 2g_{\mu\nu}\kappa_\alpha\kappa^\alpha - i(2F_{\mu\nu} + \kappa_{\mu\nu} + \kappa_{\nu\mu} + g_{\mu\nu}\kappa^\alpha{}_\alpha). \tag{113·43}$$

(Where κ has two suffixes, the second suffix denotes covariant differentiation.) Multiplying by $g^{\mu\nu}$, we obtain the gauge-covariant invariant

$$^*G = G - 6\kappa_\alpha\kappa^\alpha - 6i\kappa^\alpha{}_\alpha, \tag{113·44}$$

which satisfies

$$^*G' = \gamma^{-2}(^*G). \tag{113·45}$$

We can go on to construct certain integrals over four-dimensional regions which are in-invariant, i.e. invariant both for coordinate transformations and gauge transformations. The simplest in-invariants are

$$\int(^*G)^2\sqrt{-g}\,d\tau, \quad \int F_{\mu\nu}F^{\mu\nu}\sqrt{-g}\,d\tau, \quad \int {}^*G_{\mu\nu}{}^*G^{\mu\nu}\sqrt{-g}\,d\tau, \quad \int\sqrt{\{-|\,{}^*G_{\mu\nu}\,|\}}\,d\tau, \tag{113·5}$$

where $|\ldots|$ denotes the determinant of the components.

The most compact way of formulating molar mechanics is by an action principle. Accepting this formulation, the problem of unified molar theory is to find the appropriate generalisation of the mechanical action integral (which is known to be $\int G\sqrt{-g}\,d\tau$) that will cover also electromagnetic phenomena. The mechanical action is real; and in order to play the same role the generalised action should be real in the gauge system (i.e. reality system) which we employ. But as soon as we introduce electromagnetic

[a] *Mathematical Theory of Relativity*, §§ 86, 87.

fields, we lose the criterion by which the original reality system was defined. Being unable to specify a particular system, we have to make the generalised action real in all gauge systems. It must therefore be an in-invariant—presumably one of the simple in-invariants in (113·5) or a combination of them.

The integrals (113·5) are invariant, not only for the imaginary gauge transformations in which we are primarily interested, but for real gauge transformations. They are therefore scale-free characteristics. We can be quite sure that the quantum is not lurking in any of them. It would seem that this line of investigation is directed more towards setting up a molar theory logically independent of quantal (fixed-scale) physics than towards unifying molar and microscopic physics. However, a considerable part of wave mechanics is scale-free; and for this gauge-invariance is appropriate.

114. Action invariants

To illustrate the main principles we take for the action density the elementary, but rather artificial combination,

$$\mathfrak{A} = ({}^*G^2 - \alpha F_{\mu\nu}F^{\mu\nu})\sqrt{-g}, \tag{114·1}$$

where α is a numerical constant. It may be assumed that α has order of magnitude unity; because ${}^*G_{\mu\nu} + 2iF_{\mu\nu}$ occurs as a natural combination in (113·43), and ${}^*G^2$ is of the same order of magnitude as ${}^*G_{\mu\nu}{}^*G^{\mu\nu}$.

If $\mathfrak{A}$ is to play the part of a generalisation of mechanical action density it will have to be real—real in all gauge systems because it is in-invariant. The κ_μ are unconditionally real, but the character of the $g_{\mu\nu}$ depends on the gauge system. We can always choose the gauge system, so that

$$\kappa^\alpha{}_\alpha = 0. \tag{114·2}$$

It is a common practice to impose the condition (114·2) on κ_α, thus restricting (but not wholly removing) its indeterminacy; in particular the κ_μ calculated as the retarded potential due to charge-current sources J_μ by the wave equation $\Box\kappa_\mu = J_\mu$ has this restricted definition.[a] Then, by (113·44), $\mathfrak{A}$ will be real if $g_{\mu\nu}$ is real. Thus, with the action-density (114·1), there is at any rate no inconsistency in assuming the existence of a gauge system in which the $g_{\mu\nu}$ as well as the κ_μ are real; and this will naturally be adopted as the standard gauge system. In this system the action becomes

$$A = G^2 - \alpha F_{\mu\nu}F^{\mu\nu} - 12G\kappa_\alpha\kappa^\alpha + 36(\kappa_\alpha\kappa^\alpha)^2. \tag{114·3}$$

In the absence of electromagnetic potential $A = G^2$. Thus the action here employed is the square of the ordinary mechanical action, which is proportional to G. The squaring does not essentially alter the derived mechanical equations, which depend on infinitesimal variations of G or G^2; but the magnitude of the action is altogether different unless we employ a unit of length such that (for the system that is being treated) G is of order unity. Since $G = 6/R^2$, where R is the radius of space-time curvature, this unit is enormous;[b] so that in any actual electric field potential gradients measured in this unit are very large compared with the potentials themselves. Thus $\kappa_\alpha\kappa^\alpha$ is very small compared with $F_{\mu\nu}F^{\mu\nu}$. Thus, α and G being of order unity, the third term in

[a] *Mathematical Theory of Relativity*, § 74·1.

[b] For density 1 g.cm.$^{-3}$, R is of order 10^8 km.

(114·3) is in all practical conditions very small compared with the second, and the fourth term is utterly insignificant. For practical purposes (114·3) becomes

$$\mathfrak{A} = (G^2 - \alpha F_{\mu\nu} F^{\mu\nu})\sqrt{-g}, \tag{114·4}$$

the minute cross-terms having only academic significance.

The mechanical action is all-comprehensive, since the derived energy tensor satisfies the law of conservation identically. It would be idle to add electrical action $F_{\mu\nu}F^{\mu\nu}$ to an invariant which already contains it. Thus if (114·4) has physical significance it must be as representing the difference—the non-electrical part of the mechanical action—and the constant α is accordingly positive.

We are left with the question, How is the mechanical action density, here given by $G^2\sqrt{-g}$, to be reduced to the action density $G\sqrt{-g}$ of ordinary mechanics? In Weyl's original investigation the energy tensor found by Hamiltonian differentiation of $G^2\sqrt{-g}$ was found to be

$$2\lambda\{G_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}(G - 2\lambda)\}, \tag{114·5}$$

where, after the variation, $\frac{1}{4}G$ was put equal to λ. If λ is identified with the cosmical constant, this is the usual energy tensor. But the identification is inadmissible, since $\frac{1}{4}G$ is neither constant nor cosmical. The gulf between $G^2\sqrt{-g}$ and $G\sqrt{-g}$ is much wider than Weyl supposed; it is the gulf between quadratic and linear analysis—between sub-threshold and super-threshold theory.

Whereas the sources of the electrical potentials κ_μ are usually comprised in a small region which can be treated in isolation from the rest of the universe, all parts of the universe contribute to the inertial-gravitational potentials $g_{\mu\nu}$. The *localised* mechanical action $G\sqrt{-g}$ used in ordinary mechanics, which like $F_{\mu\nu}F^{\mu\nu}\sqrt{-g}$ is treated as isolable, is an adaptation of the more fundamental action $G^2\sqrt{-g}$. It involves wholesale re-definitions. Mutual energies of pairs of particles are replaced by self energies localised in the particles. A convention, such as that adopted in rigid field theory, is required so that changes of a local system can be studied without regard to the consequent changes of gravitational energy of the remote environment. Thus, in molar unified theory, the final step required to connect the quadratic action $G^2\sqrt{-g}$ with the linear action $G\sqrt{-g}$, far from being the simple step commonly supposed, raises many of the problems which have been treated in detail earlier in this book. There is not much inducement to us to pursue farther a line of approach which raises these problems over again. I shall therefore refer only briefly to the other invariants in (113·5), and then proceed to the application of gauge transformations to the theory developed in this book.

Attempts have been made to find an action in-invariant which avoids the artificiality of (114·1) but leads to substantially the same result, differences being confined to the cross-terms which in any case are too small to have practical significance. Consider first the action density $\mathfrak{A}_1 = {}^*G_{\mu\nu}{}^*G^{\mu\nu}\sqrt{-g}$. Proceeding as before the condition that the $g_{\mu\nu}$ as well as $\mathfrak{A}_1$ are real is

$$g_{\mu\nu}\kappa^\alpha{}_\alpha + \kappa_{\mu\nu} + \kappa_{\nu\mu} = 0. \tag{114·6}$$

So far as I can make out, this condition cannot be imposed on the κ_μ without restricting the generality of the electromagnetic field F_μ. Disregarding this difficulty, the action reduces to $G_{\mu\nu}G^{\mu\nu} - 4F_{\mu\nu}F^{\mu\nu}$ together with negligible terms in κ_α. The chief point of

interest is that the negative sign appears automatically. The original identification of κ_μ with a real gauge vector gave $G_{\mu\nu}G^{\mu\nu} + 4F_{\mu\nu}F^{\mu\nu}$ which is not a physically significant combination.

The action density $\mathfrak{A}_2 = \sqrt{\{-|{}^*G_{\mu\nu}|\}}$ is now generally accepted as the most fundamental of the in-invariants.[a] This is because of its importance in the generalised theory mentioned at the beginning of § 113. Starting with a geometry defined by an affine connection, we can, before introducing a metric, obtain an affine R.C. tensor ${}^*B_{\mu}{}^{\epsilon}{}_{\nu\sigma}$. Contracting this, we obtain the tensor ${}^*G_{\mu\nu}$. Still without introducing any metric, we can form the determinant $|{}^*G_{\mu\nu}|$ and hence the in-invariant density $\mathfrak{A}_2$. The other integrals in (113·5) all involve $\sqrt{-g}$, and cannot be introduced until a metric has been defined. If the key to world-structure lay in molar physics, we should certainly expect to find it in $\mathfrak{A}_2$.

115. Gauge transformations (microscopic theory)

In this section we shall consider *uniform* gauge transformations $\gamma = e^{i\phi}$, where ϕ is a constant.

The usual formulae for the mechanical and electrical energy tensors are

$$T_\mu{}^\nu = -(8\pi\kappa)^{-1}\{G_\mu{}^\nu - \tfrac{1}{2}g_\mu{}^\nu G\}, \tag{115·11}$$

$$E_\mu{}^\nu = -F_{\mu\alpha}F^{\nu\alpha} + \tfrac{1}{4}g_\mu{}^\nu F_{\alpha\beta}F^{\alpha\beta}. \tag{115·12}$$

In a uniform gauge transformation $G_{\mu\nu}$ and $F_{\mu\nu}$ are unchanged; and, since $g^{\mu\nu}$ varies as γ^{-2},

$$T_\mu{}^\nu \propto \gamma^{-2}, \quad E_\mu{}^\nu \propto \gamma^{-4}. \tag{115·2}$$

The difference of 'gauge dimensions' forbids any simple association of $T_\mu{}^\nu$ and $E_\mu{}^\nu$. This shows at once why, in attempts to construct a unified molar theory, $E_\mu{}^\nu$ is associated with a mechanical tensor derived from G^2 instead of with $T_\mu{}^\nu$ which is derived from G. Introducing the tensor density $\mathfrak{E}_\mu{}^\nu = E_\mu{}^\nu\sqrt{-g}$, the forms unchanged by gauge transformation are

$$T_{\mu\nu}, \quad \mathfrak{E}_\mu{}^\nu. \tag{115·3}$$

This dissimilarity, which is an obstacle to molar unification, is just what is required by the microscopic unification of theory in §§ 89–91. It means that E is a state energy tensor, and therefore combines not with the classical or interstate tensor T but with its cross-dual $Z_0{}^0$. In our previous work we have not distinguished between tensors and tensor densities; but, when attention is paid to the $\sqrt{-g}$ factor, the correct relation between the interstate and state energy tensors is

$$T^{\times}{}_{00} = Z_0{}^0. \tag{115·41}$$

This is proved as follows:

The analysis in §§ 89–91 deals with wave vectors, not wave functions, and therefore with pseudo-discrete distributions. To avoid unnecessary complication the adopted normalisation corresponds to one particle per unit volume, so that the energy density of a particle is the same as the energy. The region containing one particle is marked out invariantly in the coordinate system, so that when the gauge is changed the coordinate volume $\int dx_1 dx_2 dx_3$ is invariant, although the metrical volume is changed. Densities per unit coordinate volume are represented by tensor-densities. In Galilean coordinates a three-dimensional energy or momentum density is represented by

[a] *Mathematical Theory of Relativity*, equation (88·4) and § 101.

components either of a space tensor of a strain vector; or in general coordinates (when the density is per unit coordinate volume) by a space tensor-density or a strain vector-density. Thus, setting $Z_0{}^0 = S_0 S'^0$ where S, S' are the strain vector factors of the strain tensor, the equation $T^{\times}{}_{00} = Z_0{}^0 = S_0 S'^0$ found in Galilean coordinates becomes more generally

$$(T_{00}\sqrt{-g})^{\times} = S_0\sqrt{-g}\,.\,S'^0\sqrt{-g}, \tag{115·42}$$

so that $T^{\times}{}_{00} = Z_0{}^0\sqrt{-g} = \mathfrak{Z}_0{}^0$, as stated.

Physically, $T^{\times}{}_{00} = \mathfrak{Z}_0{}^0$ connects the density (for unit occupation) of the interstate with the product of the densities (for unit occupation) of the two states concerned in it. As unit occupation is defined with reference to a normalisation volume fixed in the coordinate system, and having therefore an invariant coordinate volume, the factor $\sqrt{-g}$ is introduced linearly on one side and quadratically on the other side of the equation, as shown in (115·42).

Our earlier difficulty was that $\mathfrak{E}_\mu{}^\nu$ cannot be included in $\mathfrak{T}_\mu{}^\nu$ because of the difference of gauge dimensions. But it can be included in $\mathfrak{Z}_\mu{}^\nu$. For the matrix $\mathfrak{Z}_\mu{}^\nu$ is simply a transpose of the matrix $T_{\mu\nu}$, which by (115·3) is gauge-invariant like $\mathfrak{E}_\mu{}^\nu$.

The recognition that $\mathfrak{E}_\mu{}^\nu$ is part of the state energy tensor $\mathfrak{Z}_\mu{}^\nu$ of quantum theory, means that the electrical parts of the quantum energy tensors of a large number of systems can build up directly a molar electromagnetic field energy tensor $\mathfrak{E}_\mu{}^\nu$; whereas the mechanical energy tensor can only be derived by a cross-dual transformation. The description of microscopic systems in quantum theory by a tensor directly connected with the electromagnetic field energy tensor but only indirectly with the mechanical field tensor is a natural consequence of the historical development. The quantum first made its appearance in the theory of radiation—electromagnetic waves; and the formative stages of development of the theory were dominated by spectroscopy.

116. Indices of wave tensors

A quantity which in gauge transformations varies as γ^r will be said to have index r.

When matter is analysed into particles, the classical energy tensor $T_{\mu\nu}$ is divided into elements $\Delta T_{\mu\nu} = p_\mu p'_\nu$. This is the tensor generalisation of $\Delta\rho = mV^{-1}$, where m is the mass of the particle and V is the volume over which it is distributed. The representation of a volume-reciprocal as a 4-vector in the corresponding time direction depends on the invariance of $\sqrt{-g}\,dx_1dx_2dx_3dx_4$. If $U = \int\sqrt{-g}\,dx_1dx_2dx_3$, U^{-1} has the same character as dx_4, and is therefore a contravariant vector $(U^{-1})^\nu$. Thus the identification of the second factor in $T_{\mu\nu}$ is $p'^\nu = (U^{-1})^\nu$. In a gauge transformation $dx_1dx_2dx_3$ is invariant, so that U^{-1} varies as $(\sqrt{-g})^{-1}$ or γ^{-4}. Thus p'^ν varies as γ^{-4}, and p'_ν varies as γ^{-2}. By (115·3) $T_{\mu\nu}$ is gauge-invariant; hence p_μ varies as γ^2:

The momentum vector p_μ of an object-particle has index 2. The energy tensor $T_{\mu\nu}$ is the product of vectors p_μ, p'_ν of indices 2 and -2. (116·1)

Since p_μ has index 2, p^μ has index 0. Thus the contravariant momentum vector is gauge-invariant and remains real in gauge transformations. This was to be expected because the foregoing analysis refers to the classical tensor $T_{\mu\nu}$; and in classical physics the momentum vector, whose components directly correspond to the observed energy and momentum, is contravariant.

Introducing wave vectors, the energy tensor T_{00} is the outer product of the symbolic vectors

$$P = \psi\chi^*, \quad P' = \phi\omega^*. \tag{116·21}$$

Since P has index 2 and P' has index -2, we take ψ, χ to have index 1 and ψ, ω to have index -1. The strain tensor $Z_0{}^0$ is obtained by the cross-dual permutation, and is the outer product of the strain vectors

$$S = \psi\phi^*, \quad S' = \chi\omega^*. \tag{116·22}$$

The strain vectors are gauge-invariant, each having one factor of index 1 and one factor of index -1.[a]

We have seen that $Z_0{}^0$ is the quantum energy tensor, and its factors S, S' are quantum momentum vectors. Although classically they are strain vectors they are the space vectors of micro space. Thus, in quantum designation, a momentum vector $P = \psi\chi^*$ is the product of wave vectors of indices 1 and -1; and the gauge transformation is

$$\psi = \psi_0 e^{i\phi}, \quad \chi = \chi_0 e^{-i\phi}, \tag{116·31}$$

where ϕ is real. The corresponding strain vector $S = \psi\psi^\dagger$ is derived by putting $\chi^* = \psi^\dagger E_{45}$; so the $\psi^\dagger$ has index -1, and the gauge transformation of S is

$$\psi = \psi_0 e^{i\phi}, \quad \psi^\dagger = \psi_0{}^\dagger e^{-i\phi}. \tag{116·32}$$

It will be seen that the complex conjugate relation of ψ and $\psi^\dagger$ is not disturbed by a gauge transformation.

We might alternatively take ψ to have index -1, and χ or $\psi^\dagger$ to have index 1. This is equivalent to reversing the sign of i (without altering ψ, χ); so that, if initially the frame is right-handed, i is changed from E_{16} to $-E_{16}$ and the chirality of the frame is reversed. This will reverse the sign of the charge represented by the wave vectors. Taking (116·31) and (116·32) to refer to a positively charged particle, the gauge transformation for a negatively charged particle is

$$\psi = \psi_0 e^{-i\phi}, \quad \psi^\dagger = \psi_0{}^\dagger e^{i\phi}. \tag{116·33}$$

Setting $\kappa = 0$ in (113·31), we see that a field of electromagnetic potential $\kappa'_\mu = \partial\phi/\partial x_\mu$ can be created artificially by a gauge transformation, just as a field of gravitational potential can be created artificially by a coordinate transformation. In both cases the artificial field is more restricted than a natural field; so that it is not possible to transform away a natural electromagnetic or gravitational field. A gravitational field is irreducible if $B_{\mu\varepsilon\nu\sigma} \neq 0$, and an electromagnetic field is irreducible if $F_{\mu\nu} \neq 0$. Formally, we can regard an irreducible field as having been created by a *non-integrable* coordinate or gauge transformation applied to an initially field-free system.

For the artificial field created by the gauge transformation (116·31) of a constant wave vector ψ_0,

$$\hbar\kappa_\mu\psi = \hbar\frac{\partial\phi}{\partial x_\mu}\psi = \left(-i\hbar\frac{\partial}{\partial x_\mu}\right)\psi,$$

so that $\hbar\kappa_\mu$ is the field momentum vector. Let

$$p^0_\mu = -i\hbar\frac{\partial}{\partial x_\mu} - \hbar\kappa_\mu. \tag{116·4}$$

[a] We have seen that $T^{\times}{}_{00}$ is strictly $\mathfrak{Z}_0{}^0$. But when the indices of the tensors are specified, so that their gauge transformations are explicitly stated, it is redundant to continue to distinguish tensors and tensor-densities; and we can revert to the earlier notation $T^{\times}{}_{00} = Z_0{}^0$. The unitary factor $\sqrt{-g}$ only manifests itself in the gauge transformation.

In this special case $p_\mu^0 = 0$. In general a wave function has not the special form $\psi_0 e^{i\phi}$ and is therefore not reducible to a constant wave vector by a change of gauge system. Then p_μ^0 is a gauge-invariant operator; that is to say, its expectation value is gauge-invariant. For in the transformation $\psi' = e^{i\phi}\psi$, p_μ^0 becomes

$$(p_\mu^0)' = -i\hbar\frac{\partial}{\partial x_\mu} - \hbar\left(\kappa_\mu + \frac{\partial\phi}{\partial x_\mu}\right), \tag{116·51}$$

so that

$$(p_\mu^0)'\,\psi' = e^{i\phi}p_\mu^0\psi. \tag{116·52}$$

Hence, multiplying by $\psi'^\dagger = e^{-i\phi}\psi^\dagger$,

$$(\psi^\dagger p_\mu^0\psi)' = \psi^\dagger p_\mu^0\psi. \tag{116·53}$$

In gravitational theory the elementary equations in Galilean coordinates are put into tensor form by substituting covariant derivatives for ordinary derivatives. The tensor equations cover artificial gravitational fields, and by the principle of equivalence they are assumed to apply also to natural fields. Similarly we put the elementary wave equation into gauge-invariant form by substituting (116·4) for $-i\hbar\partial/\partial x_\mu$. It then covers artificial electromagnetic fields, and by the principle of equivalence it is assumed to apply also to natural fields. The assumption of equivalence is none the less a hypothesis because it is dignified by the name of a principle. Here, however, the conclusion is immediately justified, because we have already determined the wave equation in an electromagnetic field by a direct method. We found in (112·6) that, for a particle in x-space, the required substitution is

$$p_\mu^0 = -i\hbar\frac{\partial}{\partial x_\mu} - \tfrac{1}{4}e_0\kappa_\mu. \tag{116·6}$$

Here κ_μ is in ordinary units. We have been using a theoretical unit defined by the identification $\sigma_\mu = i\kappa_\mu$ in § 113. Comparing (116·4) and (116·6), the identification is

$$i\kappa_\mu = 4.137e\sigma_\mu, \tag{116·7}$$

when κ_μ is in ordinary units. The operator (116·4) refers to a positive particle with $e_0 = e$; for a negative particle ψ has index -1, and the sign of the second term is reversed. The reversal is provided for automatically in (116·6) by the sign of e_0.

A uniform gauge transformation is a (complex) scale transformation, and the multiplicity factor k enters into gauge theory in the same way as in the elementary theory of scale-free systems. Considering a pseudo-discrete distribution, H^0 varies as $s^{-1/k}$ by (16·1), where s is the particle density and H^0 is any component of the total energy tensor; in particular ρ varies as $s^{-1/k}$. Primarily we define wave vectors so that $\psi\chi^*$ is the momentum vector in the normalisation volume; but when wave vectors are replaced by wave functions we follow the usual practice of normalising them to represent the stream vector (§ 106). Distinguishing the stream vector by $\psi_s\chi_s^*$, the mass-density ρ is a component of the momentum strain vector and the particle density or probability density s is a component of the stream strain vector. The rule $\rho \propto s^{-1/k}$ shows that in a scale transformation, if $\psi\chi^*$ is multiplied by γ^2, $\psi_s\chi_s^*$ is multiplied by γ^{-2k}; and this will apply to the complex transformation $\gamma = e^{i\phi}$. We have found that the momentum vector (classical) has index 2, so that the stream vector has index $-2k$. The negative sign, which applies to every kind of particle, has been provided for by the inversion

of energy; and we are only concerned with the multiplication of the index of the stream vector by k when the multiplicity is changed from 1 to k. When the cross-dual transformation is applied as before to obtain the quantum momentum and stream vectors, the stream vector will be the product of wave functions ψ_s, χ_s of indices k and $-k$. These are the wave functions currently employed.

Thus, when the multiplicity factor is taken into account, the operator (116·6) becomes

$$p_\mu^0 = -i\hbar\frac{\partial}{\partial x_\mu} - \tfrac{1}{4}ke_0\kappa_\mu. \tag{116·8}$$

It is useful to compare (116·8) with the corresponding operator in the wave equation of the intracule. By (110·6), p_4^0 for the intracule is $-i\hbar\partial/\partial x_4 - ie^2/r$; or if we take x_4 to be the real coordinate,

$$p_4^0 = -i\hbar\frac{\partial}{\partial x_4} + \frac{e^2}{r}. \tag{116·91}$$

The other components of field momentum are unmodified. This will agree with (116·8) if

$$\kappa_\mu = (0, 0, 0, e/r), \quad e_0 = -e, \quad k = 4. \tag{116·92}$$

Thus the wave equation of the intracule is the same as that of an electron in a molar electrostatic field of potential e/r, if the electron is assigned a multiplicity 4.

The foregoing is a mathematical comparison only. The actual multiplicity of a electron is not 4 but 10; and the accepted value of its mass is determined on that basis. A molar electrostatic potential κ_4 is associated with the actual time direction, whereas (116·91) is associated with the time analogue. Current text-books take the electron moving about a fixed centre of molar force to be a literal representation of the conditions in a hydrogen atom. This is quite indefensible. Perhaps the most serious result of this error is that the magnetic moment of the intracule is mistaken for the magnetic moment of the electron. Since, in (116·8), k is 10 for an electron its magnetic moment is 2·5 times larger than that of an intracule. Magnetic moments will be investigated in detail in the next three sections.

117. Magnetic moments

We have found that the wave equation in a molar electromagnetic field is obtained by substituting (116·8) for $-i\hbar\partial/\partial x_\mu$. The equation is accordingly $(\boldsymbol{W} - m)\psi = 0$, where

$$\boldsymbol{W} = \sum_1^4 E_{\mu 5} p_\mu^0, \quad p_\mu^0 = -i\hbar\partial/\partial x_\mu - \tfrac{1}{4}ke_0\kappa_\mu. \tag{117·1}$$

Here x_4, κ_4 are the imaginary coordinate and potential. Multiplying the equation by $\boldsymbol{W} + m$, we obtain

$$(\boldsymbol{W}^2 - m^2)\psi = 0. \tag{117·2}$$

which is called the *second order wave equation*. Squaring $\boldsymbol{W}$, we obtain

$$\boldsymbol{W}^2 - m^2 = -\{(p_1^0)^2 + (p_2^0)^2 + (p_3^0)^2 + (p_4^0)^2 + m^2\} + L, \tag{117·31}$$

$$L = -\tfrac{1}{4}ike_0\hbar\Sigma E_{\mu\nu}(\partial\kappa_\mu/\partial x_\nu - \partial\kappa_\nu/\partial x_\mu) = -\tfrac{1}{4}ike_0\hbar\Sigma E_{\mu\nu}F_{\mu\nu}, \tag{117·32}$$

where $F_{\mu\nu}$ is the electromagnetic force (6-vector).

If the field κ_μ is artificial $L = 0$, and (117·2) gives

$$(ip_4^0)^2 = m^2 + (p_1^0)^2 + (p_2^0)^2 + (p_3^0)^2, \tag{117·33}$$

so that p^0_μ formally resembles a momentum vector; it is, in fact, the momentum vector generalised to the arbitrary gauge defined by κ_μ. For the first order wave equation the principle (i.e. hypothesis) of equivalence is valid, and the equation obtained for an artificial field applies unchanged to an irreducible field; but we see that the principle is not valid for (117·33). An irreducible field introduces an extra term L, which might be regarded as due to an additional component p^0_5 where $L = -(p^0_5)^2$.

If p^0_1, p^0_2, p^0_3 are not too large, we can replace (117·33) by the approximate formula

$$\epsilon^0 = m + \frac{(p^0_1)^2 + (p^0_2)^2 + (p^0_3)^2}{2m}, \tag{117·34}$$

ϵ^0being the real energy ip^0_4. Then an irreducible field adds the term $-L/2m$ on the right-hand side of (117·34). The term $-L/2m$ is gauge invariant.

If the field consists of a magnetic force $F_{23} = H$ along the x_1 axis, the additional energy is

$$-\frac{L}{2m} = \frac{ke_0\hbar H}{8m} E_{23}i = \mathfrak{M}HE_{23}i, \tag{117·4}$$

where

$$\mathfrak{M} = ke_0\hbar/8m. \tag{117·5}$$

Thus the additional energy is the product of the magnetic force H and the quantity $\mathfrak{M}$, which must accordingly be a magnetic moment. The symbolic coefficient $E_{23}i$ agrees with this interpretation. By the table in § 69 a real classical energy has the coefficient E_{45}; hence, setting $\mathfrak{M}HE_{23}i = E_{45}\delta\epsilon$, we obtain

$$\delta\epsilon = H \,.\, E_{01}\mathfrak{M}. \tag{117·6}$$

By the same table E_{01} is the correct symbolic coefficient for the x_1 component of a real classical magnetic moment. Regarded as a vector, the magnetic moment is

$$E_{01}\mathfrak{M}, \quad E_{02}\mathfrak{M}, \quad E_{03}\mathfrak{M}. \tag{117·71}$$

This is the operational form; in application, E_{01}, E_{02}, E_{03} (which in quantum designation is the spin $E_{23}i$, $E_{31}i$, $E_{12}i$) will be replaced by eigenvalues or expectation values.

If the field consists of an electric force iF_{14}, we find in the same way an electric moment

$$E_{15}\mathfrak{M}, \quad E_{25}\mathfrak{M}, \quad E_{35}\mathfrak{M}. \tag{117·72}$$

By the table in § 69 this is a real electric moment. But there is an important contrast between the magnetic and electric moments. The extended momentum vector in a neutral uranoid contains a magnetic but not an electric moment. The standard reality conditions make the linear momentum $E_{15}i$ the real operator (i.e. the operator having a real expectation value) and forbid the electric moment E_{15}.[a] Thus the electric moment can only occur as a second order effect due to the modification of the reality conditions by the presence of the field. The contrast is recognised in elementary physics; permanent magnetisation is possible, but not permanent electric polarisation.

[a] Mathematical notation does not distinguish between an $E_\mu i$ component, and i times an E_μ component. Similarly it does not distinguish between the derivative with respect to a coordinate ix, and $-i$ times the derivative with respect to a coordinate x. This sometimes leads to physically imaginary quantities appearing in the equations which should in fact be zero. Thus in the standard reality conditions the electric moment is not $-i$ times the real quantity $\mathfrak{M}(E_{15}i, E_{25}i, E_{35}i)$, but simply zero; and there is no question of trying to provide a physical interpretation of an imaginary energy resulting from its interaction with the electric force.

An applied magnetic field makes apparent the magnetic moments already existing in particles in neutral space; an applied electric field only makes apparent the electric moments which it has itself created by polarising the particles.

It is instructive to compare the magnitudes of the terms in κ_μ and in $F_{\mu\nu}$ in the second order wave equation. By a gauge transformation we can make $\kappa_\mu = 0$ at the centroid of the probability distribution; then elsewhere κ_μ is of order $lF_{\mu\nu}$ where l is the linear extension of the distribution. The ratio of $\mathfrak{M}/e$ is roughly $\hbar/m$, or in practical units $\hbar/mc$, which is at most 10^{-10} cm. Thus normally $e\kappa_\mu$ will be larger than $\mathfrak{M}F_{\mu\nu}$; but its effect is unsystematic, positive and negative values of κ_μ largely compensating one another. On the other hand $F_{\mu\nu}$ (being a molar force) is constant over a microscopic system; and conditions of quantisation commonly arise in which its co-factor $E_{\mu\nu}$ has a constant eigenvalue over the system.

The foregoing theory relates to particles in x-space; but with a little adaptation it can be applied to intracules. The electric moment then drops out of consideration altogether; because the dimension x_4 corresponding to the components F_{14}, F_{24}, F_{34} does not exist in ξ-space, being replaced by the phase dimension. The axes x_1, x_2, x_3, with which the magnetic force is associated are co-gredient in ξ-space and x-space. Since the field is molar, we can treat $\partial\kappa_\mu/\partial x_\nu$ as constant over a microscopic system. Then the sum of the terms $e_0\kappa_\mu$ for the proton and electron is

$$e\kappa_\mu(x_\nu) - e\kappa_\mu(x'_\nu) = -e\frac{\partial\kappa_\mu}{\partial x_\nu}(x'_\nu - x_\nu) = -e\xi_\nu\frac{\partial\kappa_\mu}{\partial x_\nu}.$$

Since this is a function of ξ_ν only it is carried by the intracule. Denoting it by $e_0\kappa_\mu(\xi_\nu)$, we have $e_0 = -e$, $\partial\kappa_\mu/\partial\xi_\nu = \partial\kappa_\mu/\partial x_\nu$. Thus, if the intracule is treated as a particle with charge $-e$ (with the convention that ξ is measured from the positive to the negative charge), the equivalent magnetic field in ξ-space is the same as in x-space. We have seen that in applying (116·8) to the intracule k must be taken to be 4. Hence, by (117·8) the magnetic moment of the intracule is $-\frac{1}{2}e\hbar/\mu$, or in practical units

$$\mathfrak{M} = -\frac{e\hbar}{2\mu c} \quad \text{(intracule)}. \tag{117·8}$$

This is the value currently accepted, only it is wrongly attributed to the electron.

118. Magnetic moment of the hydrogen atom

Hitherto the intracules have received the largest share of our attention. To avoid unnecessary complication, they have been attached to extracules in a state of almost exact rest, whose momentum vector accordingly reduces (almost exactly) to a single component $E_{45}p_{45}$. In general, however, the extracule has, like other simple particles, an extended momentum vector with 16 components. In neutral space 6 components are dormant, but some of these become active when the extracule is in a molar magnetic field. Evidently the activated components are those which have been found to represent magnetic moment, namely p_{01}, p_{02}, p_{03}. This would appear to change the extracule from a V_{10} to a V_{13} particle. But, as explained in § 28, we do not normally treat electromagnetic and mechanical degrees of freedom on the same footing. Unified theory is confined to certain fundamental investigations, and the results are translated in terms of separated theory. The effects of the additional degrees of freedom are accordingly

put into the form of correcting terms or factors to be inserted in the mechanical equations of a V_{10} extracule.

Since the operational form $\boldsymbol{p}_\mu$ of p_μ is proportional to $E_{\mu\nu}$, we set

$$\boldsymbol{p}_{01}, \boldsymbol{p}_{02}, \boldsymbol{p}_{03} = \mathfrak{M}(E_{01}, E_{02}, E_{03}). \tag{118·1}$$

By (73·23), $\mathfrak{M} = -ip_{16}$; but this identification is uninformative. In $\Sigma E_\mu p_\mu$ particular fixed scale units are used in order that mass, charge, angular momentum and magnetic moment may be dealt with homogeneously. Owing to the practice of absorbing multiplicity factors into the constants, the fixed scale units may differ for different kinds of particle. Thus the identification of $\mathfrak{M}$ with $-ip_{16}$ does not tell us its value in practical units of magnetic moment, even if we happen to know the value of $-ip_{16}$ in practical units of mass or of angular momentum. We have, however, the formula (117·5) giving $\mathfrak{M}$ in practical units; inserting the constant c (which has, as usual, been put equal to 1 in the theoretical investigation),

$$\mathfrak{M} = ke_0\hbar/8mc. \tag{118·2}$$

The term 'magnetic moment' is used both for the constant $\mathfrak{M}$ and the vector (118·1) or its expectation value. Normally in a distribution of hydrogen atoms the magnetic moment of the extracules, i.e. the vector, is zero. This, of course, means that the pseudo-discrete wave functions representing the assemblage of extracules are such that the expectation values of E_{01}, E_{02}, E_{03} are zero; it does not refer to $\mathfrak{M}$, which is a natural constant for hydrogen extracules.

Since $k = 10$ and $m = \mu$ for the hydrogen extracule, (118·2) gives

$$\mathfrak{M} = \frac{5}{2}\frac{e\hbar}{2\mu c}. \tag{118·3}$$

This will be called the *standard* magnetic moment.

The only point that presents any difficulty is whether we are justified in setting $e_0 = e$ for an extracule. It might indeed be suggested that in view of the neutrality of the extracule $e_0 = 0$. But that is clearly impossible, because $\mathfrak{M}$, which is equal to $-ip_{16}$, cannot vanish; and in any case, we know that a system can have a magnetic moment although its charge may be zero. The effect of neutrality of the extracule is that the observation definition of the directions of the E_{01}, E_{02}, E_{03} axes conventionally adopted as positive breaks down; so that $\mathfrak{M}$ must be regarded as a signless quantity.

For the purpose of microscopic theory it would be much simpler if the electromagnetic potential were defined as $e\kappa_\mu$, so as to avoid the complication of introducing the square root of $\hbar c/137$. The charges of a proton and electron would then be $+1$ and -1, and the magnetic moment of the extracule would be $\mathfrak{M}/e = \frac{5}{2}\hbar/2\mu c$. It would then be evident that the sole function of the charge is to fix the sign; and the only effect of absence of charge is to leave the sign indeterminate for lack of the usual reference standard.

There exist several observational methods of determining the magnetic moment of the hydrogen extracule. In comparing the theory with the observational results, it cannot be assumed that the observational data have been reduced in such a way as to give the standard magnetic moment. As in the case of the masses of the elementary particles, we have to distinguish the standard and the current value. By far the most accurate determination is obtained by the gyromagnetic method; and we shall now

proceed to calculate the *current magnetic moment*, i.e. the quantity directly comparable with the published gyromagnetic values.

The experiment is performed in an alternating magnetic field having a fixed axis which we take to be x_1. The result is that the component p_{01} is de-stabilised, and set in forced oscillation (ultimately adjusted to resonance) by the field. The extracule executing the oscillation is a V_{11}, and accordingly has a rest energy $M' = \frac{10}{11} M$, the mass M being that of the stabilised V_{10} extracule. This reduction of rest mass represents the removal of the stabilised energy associated with p_{01} preparatory to the insertion of the specifically calculated energy $\mathfrak{M}HE_{01}$ due to this degree of freedom. Thereafter the extracule is treated as a mechanical V_{10} particle perturbed by the energy $\mathfrak{M}HE_{01}$; and accordingly $\mathfrak{M}$ must be calculated as for a V_{10} particle of mass M'. This gives

$$\mathfrak{M} = \frac{5}{2} \cdot \frac{11}{10} \frac{e\hbar}{2Mc}. \tag{118·4}$$

We have further to determine the β-factors. The relevant theory is in §§ 96, 97, and has been rather stringently checked by the calculation of the mass-defect of deuterium. Additional energies, brought into the theory after the primary adjustment of the constants to the observational system, have to be multiplied by β if the analysis is linear and by β^2 if the analysis is quadratic. Here the analysis is of a third kind, the magnetic energy being introduced in the second order wave equation; but it is easily seen that the factor is β^2, as in quadratic analysis. The magnetic energy first appears as the term L in (117·31), and is added to m^2 or ϵ^2. A new theoretical term added to ϵ has to be multiplied by β, so as to reduce it to the observational scale to which ϵ already conforms; correspondingly the new term L added to ϵ^2 has to be multiplied by β^2. Alternatively, we see directly from (97·42) that additional energy arising from an electromagnetic field $F_{\mu\nu}$ must be multiplied by β^2. Our final result is accordingly

$$\mathfrak{M} = \frac{5}{2} \cdot \frac{11}{10} \beta^2 \frac{e\hbar}{2Mc} = 2{\cdot}7906 \frac{e\hbar}{2Mc}. \tag{118·5}$$

The observational results are given in 'nuclear magnetons'. The nuclear magneton is properly $e'\hbar'/2m_p c$, where e', $\hbar'$ are the molarly controlled values; but actually $e'\hbar/2m_p c$ is used. This is because the measured quantity is not the energy $\mathfrak{M}H$, but the corresponding frequency $\mathfrak{M}H/\hbar$; so the constant directly determined is $\mathfrak{M}/\hbar$. Since the distinction between $\hbar$ and $\hbar'$ is not recognised, the factor $\beta^{\frac{1}{11}}$ required to reduce $\hbar$ to $\hbar'$ has been omitted in the reductions. Our result (118·5) gives

$$\mathfrak{M} = 2{\cdot}7899 e'\hbar/2m_p c. \tag{118·6}$$

The observed value[a] is $(2{\cdot}7896 \pm 0{\cdot}0008)\, e'\hbar/2m_p c$.

119. Magnetic moment of the neutron

The hydrogen atom contains two magnetic moment vectors carried respectively by the intracule and extracule. In a strong magnetic field the weak coupling between them is broken, and each precesses independently about the axis of the field. Thus in the gyromagnetic experiment the resonance frequency for the extracule is found without interference from the intracule. In order that its magnetic moment may be quite separate from that of the extracule, the intracule must be free; and consequently the extracule is a hydrocule extracule. This is the separation we have become accustomed

[a] Millman and Kusch, *Phys. Rev.* **60**, 91, 1941.

to, and it has been taken for granted in the preceding theoretical calculation of the moment of the extracule. It is worth noticing for future reference that this separation is responsible for one factor β in (118·5); in conformity with the observational system, we assign the molarly controlled mass M to the hydrocule instead of to the standard particle, but we rectify it when we insert β-factors, so that the true mass M/β of the hydrocule is ultimately used. The other factor β rectifies the field strength $H = F_{23}$, as required by (97·41).

In the neutron there is only one magnetic moment vector, that of the co-spin intracule being zero. This follows because what is here called (classically) the magnetic moment vector is the spin vector in quantum designation; and the co-spin state of the intracule is defined as a state of zero spin. Thus in the conditions of the gyromagnetic experiment the neutron behaves as a whole—not as an extracule and intracule resonating independently.

It is only to a limited extent that the extracule and intracule can be treated as independent particles. Generally such independence as there is is one-sided; transitions of the intracule do not affect the extracule; but any change of the extracule from the initial state affects the intracule. In the two-particle transformation the frame has to be such that the extracule is at rest; so that motion of the extracule involves a Lorentz transformation of the characteristics of the intracule. Other complications will occur if the extracule is given an angular momentum or magnetic moment. These complications are described by the general term 'coupling'; their chief importance is in spectroscopy. In the neutron even the one-sided independence ceases to exist, because there are no transitions, the co-spin intracule having only one state. The magnetic field does not cause any separation; and it appears that (except in nuclear conditions) the neutron behaves as an indivisible particle.

Let us assume that for the purpose of applying magnetic perturbation energy the neutron is to be treated as a simple particle. We have not hitherto met with this kind of problem, though it must occur frequently in the passage from multiplicative to additive analysis. Previously when two particles have been considered together they have been treated as a bi-particle represented by double wave functions. Our theory starts from the standard bi-particle; so that the whole process is that a bi-particle is separated multiplicatively into an extracule and intracule, and these are joined additively into a particle treated as simple. The important difference for our purpose is that the original occupation factor j of the bi-particle is split into occupation factors $j^{\frac{1}{2}}$ of the extracule and intracule; so that $j^{\frac{1}{2}}$ is the occupation factor of the neutron. Thus the gauge-index of the occupation factor, and therefore of the particle density and stream vector, is halved. This is due to the fact that the particle density is now a density in 3-space instead of in 6-space. Since the momentum vector is unaffected, the relation $\rho \propto s^{-1/k}$ in § 116 is replaced by $\rho \propto s^{-2/k}$; and $\frac{1}{2}k$ must be used instead of k in the formula for the magnetic energy.

The dependence $\rho \propto s^{-2/k}$ is a case of the general dependence $\rho \propto s^{-l/k}$ given by (15·52). The usual simplification $l = 1$, depends on the classifying characteristics having the same dimensions as the characteristic ρ that is being investigated. By adopting the characteristics of the extracule and intracule, instead of their products, as classifying characteristics, of the neutron, we halve the dimensions of the classifying characteristics, so that $l = 2$.

We have found (§ 95) that the multiplicity of the co-spin intracule is 4—a result which has been checked by the observed mass of the neutron. The multiplicity of the extracule is 10. Thus, regarded as a simple particle, the neutron has 14 degrees of freedom; and the factor $\frac{1}{2}k$ to be used in place of k in the magnetic moment is 7. In place of (118·4) we have

$$\mathfrak{M}_n = \frac{7}{4}\cdot\frac{11}{10}\frac{e\hbar}{2m_n c},$$

where m_n is the mass of the neutron. The factor $\frac{11}{10}$, due to the removal of the stabilised magnetic energy contained in the ordinary rest mass prior to insertion of the specifically calculated energy, affects the neutron in the same way as the hydrogen extracule.

The β-factor which rectifies the measurement of $F_{\mu\nu}$ has to be inserted; but the other β-factor, associated with the freeing of the intracule, is absent in the neutron, since the intracule is included in the resonating system. In other words, the mass m_n of the neutron takes the place of the mass μ/β of the hydrocule extracule. The final formula is accordingly

$$\mathfrak{M}_n = \frac{7}{4}\cdot\frac{11}{10}\beta\frac{e\hbar}{2m_n c}. \tag{119·1}$$

Expressed in nuclear magnetons, this gives

$$\mathfrak{M}_n = 1{\cdot}9371\frac{e'\hbar}{2m_p c}. \tag{119·2}$$

The observed coefficient (Alvarez and Bloch) is $1{\cdot}935 \pm 0{\cdot}020$.

A much more accurate observational value can be obtained if we may assume that the magnetic moment of the neutron is the difference of the moments of deuterium and hydrogen. The observed difference is

$$\mathfrak{M}_H - \mathfrak{M}_D = 1{\cdot}9331\frac{e'\hbar}{2m_p c}. \tag{119·3}$$

The difference between $\mathfrak{M}_n$ and $\mathfrak{M}_H - \mathfrak{M}_D$ is about 12 times the probable error. It is in the right direction; for the assumption that in deuterium the hydrogen and neutron moments are rigidly antiparallel gives a lower limit to the moment of the neutron. The discrepancy there presumably measures the looseness of coupling. As a determination of this quantity may be useful in nuclear physics, it is worth while to find it as accurately as possible. Combining the theoretical moments $\mathfrak{M}_H$ and $\mathfrak{M}_n$ with the observed ratio $\mu_H/\mu_D = 3{\cdot}257 \pm 0{\cdot}001$, we have

$$\mathfrak{M}_H - \mathfrak{M}_n - \mathfrak{M}_D = -0{\cdot}0038 \pm 0{\cdot}0003 \text{ nuclear magnetrons.} \tag{119·4}$$

Chapter XII

RADIATION

121. Radiation by a moving electron

[*Editorial note. There was no* § 120 *in the manuscript.*]

The quantum theory of radiation is usually prefaced by a classical calculation of the force on an electron in non-uniform motion. The customary treatment, which divides the elementary charge into infinitesimal elements and assumes that these act on one another according to Maxwellian laws, is made obsolete by the interchange theory of Coulomb energy; electromagnetic interaction first arises when we consider two charges whose carriers can undergo interchange, and there is no question of electrical forces between elements of the same elementary charge. A further objection to the treatment is that the result contains an infinite term which has to be ignored. Also the analysis is unnecessarily complicated. We shall give an independent derivation of the formula for the radiation of momentum, appropriate to the present theory.

The retarded electrostatic potential Φ and vector potential A_1, A_2, A_3 due to a point charge e can be expanded in the series[a]

$$\left.\begin{aligned} \Phi &= e\left(\frac{1}{r}+\frac{1}{2}\frac{d^2r}{dt^2}-\frac{1}{6}\frac{d^3r^2}{dt^3}+\frac{1}{24}\frac{d^4r^3}{dt^4}-\dots\right), \\ A_\alpha &= e\left(\frac{u_\alpha}{r}-\frac{du_\alpha}{dt}+\frac{1}{2}\frac{d^2(ru_\alpha)}{dt^2}-\frac{1}{6}\frac{d^3(r^2u_\alpha)}{dt^3}+\dots\right), \end{aligned}\right\} \tag{121·1}$$

where the velocity u_α of the charge and the distance r from the charge to the point considered are *unretarded*; that is to say, they refer to the same instant as Φ and A_α. The electric force is

$$X_\alpha = -\frac{\partial\Phi}{\partial x_\alpha}-\frac{\partial A_\alpha}{\partial t}. \tag{121·2}$$

We consider first the third term in Φ and the second term in A_α. Denoting the components of r by x_α, we have $\partial r^2/\partial x_\alpha = 2x_\alpha$ and $dx_\alpha/dt = -u_\alpha$. The two terms together give

$$X_\alpha = e\left(-\frac{1}{3}\frac{d^2u_\alpha}{dt^2}+\frac{d^2u_\alpha}{dt^2}\right) = \tfrac{2}{3}e\ddot{u}_\alpha. \tag{121·3}$$

Thus the charge e exerts on any other charge e' (not necessarily concentrated at a point) a mechanical force $\frac{2}{3}ee'\ddot{u}_\alpha$ which is independent of the distribution of e'. In particular a charge $\pm e$ exerts on the opposite charge which it induces in its environment a mechanical force $P_\alpha = -\frac{2}{3}e^2\ddot{u}_\alpha$. Inserting a factor c^{-3}, as required by the dimensions, the force in practical units is

$$P_\alpha = -\frac{2}{3}\frac{e^2}{c^3}\ddot{u}_\alpha. \tag{121·4}$$

The force derived from the other terms of (121·1) is not independent of r, and cannot be calculated unless we know the distribution of the induced charge. This would require a detailed specification of the conductivity and dielectric constant of the sur-

[a] *Mathematical Theory of Relativity*, 2nd ed. p. 253.

rounding material. We therefore confine attention to an environment in which these terms are negligible. Since the earlier terms become large as $r \to 0$ and the later terms become large as $r \to \infty$, the induced charge must be kept within limits r_1, r_2, where r_1 is not unduly small and r_2 not unduly large. This means that the point-charge must be isolated with no other particles within a distance r_1, and that there must be sufficient polarisable matter between r_1 and r_2 to quench its field.

The physical reason for these conditions is apparent, if we regard the momentum communicated to the induced charge as transmitted by radiation. The distance r at which the radiated momentum is regarded as finally detached from the system in which it originates is necessarily conventional, and the limit r_1 is taken as a criterion that the momentum has got clean away. The limit r_2 is introduced in order that the momentum may be quickly absorbed and passed on to the particles in the environment. As r_2 increases the proportion of radiant momentum in transit increases, and our formula which gives the momentum taken up by the induced charge, i.e. the particles of the environment, begins to deviate from the total momentum radiated.

It is clear that, so far as the theory of radiation is concerned nothing would be gained by investigating the terms omitted in (121·3). The condition that they are negligible is in fact the condition that the investigation applies to radiation.

Besides the induced charge and currents the environment contains the charges responsible for the molar electromagnetic field which causes the changes of u_α. This field is specified as a datum of the problem, and the momentum necessary to make it vary with time in the prescribed way is furnished by unstated sources. It is therefore to be treated as rigid. Our calculation of P_α has therefore to be confined to the induced charge and current which constitute the non-rigid part of the environment. In radiation theory the formal distinction is that the interaction between the initial charge distribution which provides the extraneous field and the object-system is represented by an inflowing wave of radiation and the interaction between the object-system and the induced charge distribution is represented by an outflowing wave.

The magnetic force derived from (121·1) contains a term

$$H_1 = \tfrac{1}{2}e\frac{d^2}{dt^2}(l_3u_2 - l_2u_3), \tag{121·5}$$

and corresponding components H_2, H_3, where l_1, l_2, l_3 are the direction cosines of r. This term is of the same order as (121·3), and its action on the induced currents should be added to P_α. The justification for neglecting it appears to be that, when u_α is a rapidly alternating velocity, the phase of the induced current will vary rapidly with r, so that the integrated force is small. (A steady velocity u_α would, of course, be eliminated by a Lorentz transformation before applying (121·4).)

In current theory the classical formula (121·4) is made the basis of the quantum calculation, a quantum dynamics being derived from classical dynamics by semi-empirical rules. That is not our method; and we are concerned with the classical formula only as a check, the quantum formula being required to converge to the classical formula under certain limiting conditions. For either purpose it is desirable to relate the foregoing investigation rather more closely to the quantum point of view. Taking a rigid environment which contains an electromagnetic field, we superpose on it an intracule, i.e. an elementary electric doublet. For small separations the doublet as a whole is

taken as object-system, and the perturbation of its quantised states by the extraneous field (Zeeman and Stark effects) is calculated. For large separations the carrier of the negative charge is treated separately as object-system, and the positive charge then becomes the induced charge added to the environment. At the junction of the two treatments, the momentum transmitted from the negative to the positive charge, which was at first a result of internal interaction and formed part of the momentum balance of the steady quantised state, becomes re-classified as momentum radiated from the object-system to the environment.

122. Transition probabilities

Consider a system represented by wave vectors ψ_0, χ_0 of indices 1, -1. If a field of uniform electrostatic potential κ_4 is introduced, the resulting perturbation can be found in the following way. We eliminate the potential by a gauge transformation $\gamma = e^{\frac{1}{2}i\omega t}$, where ω/κ_4 is a constant depending on the multiplicity. This turns the wave vectors into wave functions

$$\psi = \psi_0 e^{\frac{1}{2}i\omega t}, \quad \chi = \chi_0 e^{-\frac{1}{2}i\omega t}. \tag{122·11}$$

The momentum vector $P = \psi\chi^*$ remains constant in time but this constancy is relative to the modified gauge. By (116·1) a classical momentum vector has index 2; so that the ordinary unit-standard, constant in the initial (Galilean) gauge, becomes $e^{i\omega t}$ in the modified gauge. Thus in the initial gauge the momentum vector of the system is $P' = Pe^{-i\omega t}$. Accordingly the perturbation caused by the field κ_4 consists in changing the constant vector P into a periodic vector P'. The effect is due to the fact that wave vectors of indices 1, -1 give the quantum analogue of momentum, actual (classical) momentum being given by wave vectors of indices 1, 1. The measures are made to agree in the initial gauge system; but when κ_4 is introduced, or the equivalent change of gauge is made, a periodic difference appears.

We apply this to the oscillator occupying the interstate between two states Q_1, Q_2. There is a conjugate interstate with momentum vector $P'^\dagger$, which in steady conditions is constrained to have the same occupation (§ 91); so that, although the momentum vectors of the two oscillators are complex, their sum is real. The wave functions ψ, χ represent the transition circulation in the interstate. Since P' (the momentum in ordinary measure) is reversed in direction at intervals π/ω, this is evidently the half-period of the circulation. If at $t = 0$ there is one unit of occupation in each of the two states, there will again be one unit in each state at $t = \pi/\omega$, but the two units will be interchanged. Thus π/ω is the average time in which a particle makes a transition from Q_1 to Q_2 or from Q_2 to Q_1. The reciprocal

$$w = \omega/\pi \tag{122·12}$$

is the transition probability per unit time.

It may be suggested that the result should be doubled because two oscillators are engaged in the transfer; but (122·12) is right as it stands. The state energy tensor $Z_0{}^0$ corresponds to unit occupation of two states; and when it is transformed into an interstate energy tensor T_{00}, the result corresponds to unit occupation of two interstates. Thus the two particles in continual transition with half-period π/ω correspond to two oscillators.

It is fairly evident that the two oscillators with complex time factors $e^{i\omega t}$, $e^{-i\omega t}$ can be replaced by two oscillators with real time factors $\cos\omega t$, $\sin\omega t$, or more generally (since the origin of t is arbitrary) $\cos(\omega t+\alpha)$, $\sin(\omega t+\alpha)$. By giving α uniform probability distribution we obtain a smooth instead of a lumpy circulation.

For Q_1, Q_2 we shall take two orthogonal eigenstates of an intracule represented by wave functions $\psi_1 e^{i\epsilon_1 t_0/\hbar}$, $\psi_1^\dagger e^{-i\epsilon_1 t_0/\hbar}$ and $\psi_2 e^{i\epsilon_2 t_0/\hbar}$, $\psi_2^\dagger e^{-i\epsilon_2 t_0/\hbar}$. The quantum energies of the two states are ϵ_1, ϵ_2; and the phase coordinate, or time analogue, is here denoted by t_0 to distinguish it from the actual time t. (The coordinate referred to in the preceding calculation of transition probability must, of course, be the actual time.) The wave functions are as usual normalised to represent probability density and flux. Thus $\psi_1\psi^\dagger$ and $\psi_2\psi^\dagger$ are strictly stream strain vectors; but we shall not discriminate between stream vectors and momentum vectors, until the distinction becomes pertinent.

Making the cross-dual permutation, the momentum space vector of the interstate is the product

$$P_0 = \psi_1\psi_2^\dagger e^{i(\epsilon_1-\epsilon_2)t_0/\hbar} = \psi_1\psi_2 e^{i\omega_0 t_0}, \tag{122·2}$$

where

$$\hbar\omega_0 = \epsilon_1 - \epsilon_2. \tag{122·3}$$

Since P_0 is a factor of the classical energy tensor, it has index 2 by (116·1). We have therefore to assign the indices so that ψ_1, $\psi_2^\dagger$ have index 1 and $\psi_1^\dagger$, ψ_2 have index -1. Setting $P = \psi_1\psi_2^\dagger$, we derive P_0 from P by the gauge transformation $\gamma = e^{\frac{1}{2}i\omega_0 t_0}$, which also transforms ψ_1, $\psi_2^\dagger$ into

$$\psi_1 e^{\frac{1}{2}i\omega_0 t_0}, \quad \psi_2^\dagger e^{\frac{1}{2}i\omega_0 t_0}. \tag{122·41}$$

We have seen (§ 116) that, although the covariant momentum vector P_0 varies with gauge, the contravariant vector P^0 is gauge-invariant. In order that the treatment of interstates may be formally assimilated to the treatment of state we must use the gauge-invariant vector in both cases, namely the covariant momentum vector for the state and the contravariant momentum vector for the interstate.

The momentum vector $P = \psi_1\psi_2^\dagger$ in the initial gauge can indifferently be taken to be covariant or contravariant.[a] But if we take it to be contravariant the index of one of the factors ψ_1, $\psi_2^\dagger$ must be reversed. Thus P^0 is the product of the wave functions

$$\psi_1 e^{\frac{1}{2}i\omega_0 t_0}, \quad \psi_2^\dagger e^{-\frac{1}{2}i\omega_0 t_0}, \tag{122·42}$$

which is gauge-invariant, as required.

Quantum theory differs from classical relativity theory in using a covariant instead of a contravariant momentum vector. It will be seen that, if we continue to use wave functions with opposite exponentials as in (122·42), the change over from covariant to contravariant representation occurs automatically in the transformation from the quantum or state energy tensor to the classical or interstate energy tensor; it is required to counteract the bringing together of two wave functions of the same index by the cross-dual permutation.

[a] The wave functions ψ_1, $\psi_2^\dagger$ in the E- and F-frames give, on multiplication, a space vector in the crossed frame (C-frame). Since they are functions of the space coordinates, the space axes in the three frames must be taken as congruent; but there is nothing to connect directions of the time and phase axes in the C-frame with those in the E- and F-frames, and covariance in the E-frame has to be defined independently of covariance in the C-frame. Since the C-frame contains classical vectors, covariance in the C-frame has to agree with the classical definition; but in the E-frame it is decided by an arbitrary convention which can be reversed. The combination of E and F wave vectors chosen to give a covariant C vector will, when the convention is reversed, give a contravariant C vector.

In general P^0 will contain non-vanishing components p_{15}, p_{25}, p_{35}, which require isostatic compensation in order that the steady state may be maintained. As a first approximation we shall use mean values (taken in a way to be defined later) so that p_{15}, p_{25}, p_{35} are constants independent of x_1, x_2, x_3. The field momentum required for compensation is provided by introducing the exponential factors

$$e^{\pm i(p_{15}x_1+p_{25}x_2+p_{35}x_3)/\hbar} \tag{122·51}$$

in the wave functions. It is well known that in order to maintain steady occupation of two quantum states, and therefore a steady circulation between them, a field of radiation must be present. It is the momentum of this equilibrating field that is introduced in (122·51). From another point of view, the momentum is introduced because the energy tensor $Z_0{}^0$ of the quantum states Q_1, Q_2 contains recalcitrant terms, and therefore the classical tensor T_{00} into which it transforms is not self-conserved.

The complete isostatic compensation includes also compensation for the component in the phase dimension, so that (122·51) is extended to[a]

$$e^{\pm i(p_{15}x_1+p_{25}x_2+p_{35}x_3+p_{45}t_0)/\hbar}. \tag{122·52}$$

The fourth term must agree with the term already present in the wave functions (122·42), so that

$$p_{45}/\hbar = \tfrac{1}{2}\omega_0. \tag{122·53}$$

This conclusion requires justification; for we have added momentum p_{15}, p_{25}, p_{35} to equilibrate the system, and it seems arbitrary to assume that no p_{45} momentum will need to be added. Why should we assume that the radiation field maintaining the steady occupation contains no p_{45} momentum? The reason is that the p_{45} momentum is the scale; and to make a change in it would mean that we changed our scale in passing from the states to the interstates. The interstate energy tensor, which is classical, is intrinsically scale-free, since classical physics provides no standard of its own; but a definite scale is given to it by importing a quantum-specified standard. When new field momentum is introduced to equilibrate the interstate circulation, there must be no change of scale momentum, *because it is just at this point that the fixed scale provided by the quantisation of the states Q_1, Q_2 of the intracule is being transferred to the classical description by interstate tensors.* Equation (122·53) is the condition that our final result (transition probability per unit time) shall be referred to the quantum-specified standard of time.

We have often remarked that the quantum method of treating an object-system as a superposition on a rigid environment leads to a double reckoning of energy; by taking each particle in turn as object-particle, and adding the energies, we get twice the whole energy of the universe. This is relative to the single reckoning in molar relativity theory, where the environment is not treated as rigid. Thus in transferring the quantum scale to the classical tensors it is the practice to halve the reckoning or scale of energy and momentum—an operation more usually viewed in the converse way as doubling the reckoning of energy when we pass from the relativity to the quantum treatment. To allow for this (122·53) must be replaced by

$$p_{45}/\hbar = \tfrac{1}{4}\omega_0. \tag{122·54}$$

[a] Isostatic compensation is limited to the 4-vector corresponding to the dimensions of the domain of probability distribution. If a time coordinate is introduced, this is not extended to a 5-vector, because probability is not distributed *over* time. The system is, as it were, left to tumble over in the time direction; or, to put it more formally, the unbalanced momentum in the time direction is the cause of transitions.

We have been proceeding step by step from the quantum system of representation to the classical. The coordinates x_1, x_2, x_3 are still the relative coordinates, used in the description of the states Q_1, Q_2 of the intracule, and p_{15}, p_{25}, p_{35} are relative momenta conjugate to them. These have now become incongruous since the classical tensor is exhibited in x-space, not ξ-space; and our next step is designed to eliminate them.

Since we are using the contravariant vector P^0, we can write $p_{\mu 5}/p_{45} = v_\mu$, where (v_1, v_2, v_3) has the usual meaning of velocity (of flow of probability). Following an element of the probability in its motion,

$$\delta(p_{15}x_1 + p_{25}x_2 + p_{35}x_3) = p_{45}v^2\delta t = \tfrac{1}{4}\hbar\omega_0 v^2\delta t \qquad (122{\cdot}61)$$

by (122·54). The theory refers to wave functions superposed on a rigid field, so that $v^2\delta t$ is here reckoned in rigid coordinates. In passing to Galilean coordinates $v^2\delta t$ transforms as x^2/t; so that, by (19·7), it is divided by 136. As in the corresponding reduction of the Coulomb energy term, the divisor is changed to 137 in the final reduction to the observational system. Thus, in the observational system,

$$\delta(p_{15}x_1 + p_{25}x_2 + p_{35}x_3) = \tfrac{1}{4}\hbar\omega_0 v^2\delta t/137;$$

and the exponential (122·51) becomes

$$e^{\pm\frac{1}{2}i\omega t} \quad (\omega = \tfrac{1}{2}\omega_0 v^2/137). \qquad (122{\cdot}62)$$

By (122·12), this gives a transition probability per unit time,

$$w = \frac{1}{2\pi}\cdot\frac{1}{137}v^2\omega_0. \qquad (122{\cdot}63)$$

Reverting to rigid coordinates, the complete exponential obtained by combining (122·42) and (122·51) is equal to $e^{\pm(\frac{1}{2}i\omega_0 t_0 + ip_{45}v^2t/\hbar)}$, which by (122·54) can be written $e^{\pm 2i(mc^2t_0 + \frac{1}{2}mv^2t)/\hbar}$, p_{45} being put equal to m or (since $c = 1$) mc^2. Thus the kinetic energy $\frac{1}{2}mv^2$ is associated with t in the same way that the rest energy mc^2 is associated with t_0. If we had not taken into account the change from double to single reckoning, and had used (122·53), we should have had mv^2 and mc^2 associated with t and t_0—an obvious incongruity.

The wave functions are normalised, according to the usual practice, to give directly the probability density and flux corresponding to one unit of probability in each state and therefore one unit in each interstate; so that the expectation value of v is given by $v_\mu = \int\psi_2^\dagger E_{\mu 5}\psi_1 dV$. Or, if E_{r5} is the direction of the resultant velocity in the interstate,

$$v = v_r = \int\psi_2^\dagger E_{r5}\psi_1 dV. \qquad (122{\cdot}7)$$

The momentum vector of an oscillator does not satisfy the standard reality conditions; so that a physically real momentum $|\,p_{15}\,|$ will be represented symbolically by $E_{15}e^{i\alpha}\,|\,p_{15}\,|$, where α is no longer necessarily 0 or $\frac{1}{2}\pi$. The expectation value of the momentum along the physically real axis x_1 will be given by the operational form $E_{15}e^{i\alpha}$. The factor $e^{i\alpha}$ has been omitted in (122·7), but the omission is rectified by taking the modulus of the result, since the momentum along a real axis is necessarily real. We define v by (122·7), and must therefore substitute $|\,v\,|^2$ for v^2 in (122·63).

For a linear oscillator the use of the average velocity is a good first approximation, but it is unsuitable when the flow has varying directions in three dimensions. We

therefore separate the averaging in magnitude from the averaging in direction, and treat the three-dimensional oscillator as a superposition of linear oscillators in all directions. This is the more necessary because the discrete states Q_1, Q_2 are defined as having a definite orientation, whereas in a natural assemblage the atoms will have varying orientation, and the representative atom which we study will have a probability distribution of orientation of the states concerned in transition. Accordingly, the whole transition probability is divided into elements $w_r d\Omega$, where $d\Omega$ is an element of solid angle in the direction E_{r5}; and our formula becomes finally

$$w_r d\Omega = \frac{1}{137} \cdot \frac{|v_r|^2}{2\pi c^2} \omega_0 d\Omega, \tag{122·8}$$

where v_r and ω_0 are given by (122·7) and (122·3).

This agrees with the current formula for spontaneous dipole transition probability.

123. Compton scattering

The velocity vectors v_α, v'_α of the two interstates, referred to physically real axes, are equal and opposite. This is evident from general considerations.[a] The contravariant vector P^0 from which v_α is derived is constant with respect both to t_0 and t. But the probability cannot continue indefinitely to flow in one direction in the CD-frame; and the other interstate must provide for its return. Thus the analysis exhibits an element of probability as passing from Q_1 to Q_2 with velocity v_α in one interstate and returning to Q_1 with velocity $-v_\alpha$ in the other interstate.

It follows that the half-period of the circulation, besides being the mean time of transition of a particle from one state to the other, is also the mean time of transition of an oscillator from one interstate to the other. Thus the formula (122·8) is given a new field of application as determining the transition probability of an oscillator.

Consider an electron with steady momentum p_1 along the x_1 axis. This is not a 'state', since the distribution (relative to the physical origin) is progressively changing. We can treat it as an interstate provided that we couple with it a conjugate interstate in which the electron has momentum $-p_1$; to satisfy the reality conditions the two interstates must have equal occupation.

For steady circulation the particle momentum vector p_1 of the oscillator must be isostatically compensated by a field momentum $-p_1$. The nature of this field momentum will be more fully investigated later; meanwhile we use the well-known result that the field momentum associated with transitions consists of photons of frequency $\nu = p/\hbar$, or radian frequency $\omega = p_1/\hbar$, moving in the direction of their momentum p.

Accordingly, the two interstates comprise (a) an electron of momentum p_1 and a photon of momentum $-p_1$, and (b) an electron of momentum $-p_1$ and a photon of momentum p_1. When the transition from (b) to (a) occurs we say that the electron has scattered an incident photon p_1 backwards in the $-x_1$ direction. This case, in which the motion of the photon is directly reversed, will be called *collinear scattering*. The part of the theory of scattering which requires wave mechanics is limited to collinear scattering, since the formulae for oblique scattering can be derived from it by a

[a] We have changed the notation so that v_α is the physically real velocity previously denoted by $\pm|v_\alpha|$. The substitution of $P^\dagger$ for P will not alter $|v_\alpha|$, as determined by (122·7). But the analysis leaves unsettled the ambiguity of sign; and we have recourse to other considerations in fixing it.

Lorentz transformation. It will be seen that in collinear scattering the momentum and energy of the system photon + electron are conserved; and since the conservation law is invariant for Lorentz transformations, this applies also to oblique scattering.

The wave function of an interstate is

$$\psi_0 e^{i(p_1 x_1 + m l_0)/\hbar}, \tag{123·1}$$

where ψ_0 is constant and m is the proper mass m_e of an electron. Since this is pseudo-discrete, it represents an assemblage of similar systems. By taking the occupation to be such that there are two electrons, one in each interstate, to a wave-length $l = 2\pi\hbar/p_1$, we obtain a very simple representation of the transitions. By (122·63), disregarding the factor 137 introduced in the later reduction to the observational system, the time between transitions is $\tau = 1/w = 2\pi/v^2\omega_0 = \pi\hbar/v^2 m$, since $m/\hbar$ in (123·1) corresponds to $\frac{1}{2}\omega_0$ in (122·42). The free path of the electron is $\tau v = \pi\hbar/mv = \pi\hbar/p_1 = \frac{1}{2}l$. Thus the transition or reversal of motion occurs at each half wave-length. We picture two opposite streams of electrons with an exchange (or collision) occurring at each half wave-length, so that an individual electron oscillates to and fro over a range $\frac{1}{2}l$ and the system remains steady. This regular representation refers, of course, to the probability, and not strictly to the particles themselves.

The foregoing result is exact. We have used the formula $p = mv$, which is apparently non-relativistic; but it is to be remembered that a rigid field is present so that Lorentz formulae do not apply, and further both here and in (122·63) v is the velocity of flow of probability which may differ from the velocity of the particles by the difference between wave velocity and group velocity. The relation $p = mv$ makes $\frac{1}{2}mv^2$ (which is identified with the kinetic energy in § 122) equal to $p^2/2m$ which has been shown (§ 45) to be the correct relativistic hamiltonian for standing waves.

In the physical interpretation of our calculations, p_1 is taken to be the real *classical* momentum of the scattering electron; but in (123·1) p_1 is a real *quantum* momentum $-i\hbar\partial/\partial x_1$. The explanation is that (123·1) refers to a steady distribution in which the moving electron is incorporated; and field energy, requisite to maintain the steadiness, has been added. As in our example of the star cluster (§ 21), the conversion of a dispersing into a steady system inverts the energy, or equivalently multiplies the spatial momenta by i, thereby changing real classical into real quantum momenta. The condition for steadiness yields important information. The isostatic compensation of the assemblage (in so far as it has not already been taken into account in the rest masses of the electrons) is provided by the electromagnetic radiation field, which we have divided into photons of momentum $\pm p_1$. In particular, the pressure T_{11} of the particles must be compensated by the pressure of the radiation; or, to put it another way, at the boundary of the assemblage the electrons must be kept from dispersing by the pressure of the incoming stream of photons. Let σ_e, σ_p be the number of electrons and number of photons in unit volume. For one-dimensional motion the pressure of the electrons is $\sigma_e p_1^2/m$. For radiation the pressure is the momentum $\sigma_p p_1$ per unit volume multiplied by the velocity of light c'. We must note that c' is specifically the velocity of light, i.e. of the photons, and not the universal constant c that has been set equal to 1; because in the rigid coordinates that we are using, the velocity of light is $137c$ in accordance with (19·7). Equating the two pressures, we have

$$\sigma_p/\sigma_e = p_1/137m. \tag{123·2}$$

As a step towards three-dimensional theory we extend the cellular division to the x_2 and x_3 axes; so that the space is divided into cubes of side l_1, each cube containing two electrons, one in each interstate, and therefore σ_p/σ_e photons in each interstate. This implies that the electron oscillates simultaneously and independently with velocity $\pm v_1 = \pm v_2 = \pm v_3$ along each axis—an unrealisable condition which we shall rectify later. Consider a beam of radiation of frequency corresponding to p_1 and of unit cross-section travelling along the x_1 axis. The probability that a photon is scattered by a particular electron lying in the track of the beam is the 'effective cross-section of the electron' for scattering. This probability is the product of

(1) the probability l^2 that the track of the photon goes through the cell containing the electron;

(2) the time l/c (measured in the observational system, $c = 1$) that the photon takes to go through the cell;

(3) the probability per unit time (measured in the observational system) that the electron makes a transition. Since the electron travels $\frac{1}{2}l$ between transitions the probability is

$$\frac{1}{2}\cdot\frac{1}{137}\cdot\frac{v}{\frac{1}{2}l}, \tag{123·3}$$

the reduction factors 2 (on account of current double reckoning) and 137 being the same as in § 122;

(4) a 'dilution factor' σ_p/σ_e.

The dilution factor is required because (123·3) has been calculated for an interstate or assemblage of interstates containing one photon per electron. This concentration of photons would give much too high a radiation pressure for equilibrium, and we find that the steady distribution contains σ_p/σ_e of a photon per electron. Thus $1-\sigma_p/\sigma_e$ of the transitions in (123·3) are due to photons non-existent in our system; and the transition probability of the electron must be reduced accordingly.

The four factors give a cross-section

$$A_0 = l^2 . l . \frac{v}{137l}\cdot\frac{p_1}{137m}. \tag{123·41}$$

Since $v = p_1/m$ and $l = 2\pi\hbar/p_1$, the result is

$$A_0 = \frac{4\pi^2}{137^2}\frac{\hbar^2}{m^2} = 4\pi^2\frac{e^4}{m^2}. \tag{123·42}$$

We have been representing the electron as a combination of three independent oscillators with wave functions $e^{i(p_1x_1+mt_0)/\hbar}$, $e^{i(p_2x_2+mt_0)/\hbar}$, $e^{i(p_3x_3+mt_0)/\hbar}$. The true three-dimensional oscillator has a wave function $e^{i(p_1x_1+p_2x_2+p_3x_3+mt_0)/\hbar}$. The distinction is that the three independent phases t_0 are in the three-dimensional oscillator replaced by a single phase t_0. In the combination oscillator the components p_1, p_2, p_3 are compared separately with the scale, so that there are three independent scale fluctuations, and three phase dimensions are required; in the three-dimensional oscillator the resultant of p_1, p_2, p_3 is compared with the scale, so that there is only one scale fluctuation and one phase dimension. To introduce three phase dimensions is, of course, a departure from the convention on which all our previous analysis has been based. It involves a widening factor $(2\pi)^3$ instead of the factor 2π allowed for in the change over

from curvature to flat representation. Thus when the natural electron in three dimensions is substituted for the 'unrealisable condition' temporarily used to represent it, there is a redundant factor $4\pi^2$ to be removed from the volume measure. Of the four factors in (123·41) the last two are found independently of the cellular division; in the first two l^3 is to be replaced by $l^3/4\pi^2$. The final value of the cross-section for collinear scattering is

$$A_0 = e^4/m_e^2. \tag{123·5}$$

This agrees with current theory.

If this result is stated in the form that A_0 is the cross-section for collinear scattering of radiation of frequency ν by an electron with momentum $-\hbar\nu$ in the direction of the radiation, all reference to quantum theory is eliminated, and the further development of the formula for practical application is a straightforward problem of special relativity theory. Consider radiation of frequency ν in a fixed direction scattered by an electron with arbitrary momentum vector p_i within a solid angle $d\Omega$ in the direction θ. By the conservation of energy and momentum we can determine the final momentum vector p_f of the electron. We now apply the Lorentz transformation which reduces the resultant $p_i + p_f$ to zero. In the new frame (distinguished by an accent) $p_i' = -p_f'$. The momentum of the radiation is not necessarily in the direction of p_f'; but the component in the direction of p_f' satisfies the conditions for collinear scattering and the transverse component is unscattered. We can therefore find the cross-section A_θ' corresponding to the solid angle $d\Omega'$. Re-transforming to the original frame we obtain the element of cross-section $A_\theta d\Omega$ associated with the element of solid angle $d\Omega$ in any direction of scattering θ.

It is unnecessary to pursue the calculation here. Having obtained by our theory the result (123·5) which is the special case of the Klein-Nishina formula for collinear scattering, agreement in the general case follows automatically if the Klein-Nishina formula is Lorentz-invariant.

The formula is inaccurate for high energies because the uncertainty of the physical origin, from which the coordinates of the electron are measured, has been neglected.

124. Transverse self energy of a particle

The rest mass of an object-particle is a mutual energy of interaction of the particle with all the other particles in the environment. In Chapters III to V we have found three equivalent representations of the interaction, by inertia-gravitation, by exclusion, and by interchange. Following the ideas of the present chapter, we obtain a more detailed development of the interchange representation as a steady transition circulation between the state of the object-particle and the various states of the environment particles. Whichever representation we use, the mutual energies of the object-particle paired with each environment particle in turn must add up to the mass of the object-particle, or to twice the mass if single reckoning is used. That is how we have determined the mass m_0, and hence the masses of the elementary particles; so that we are secure against any failure of this condition. But in current theory the rest masses of particles are introduced empirically as self properties; and the first suggestion of a representation of mass as a mutual property arises in the calculation of the 'transverse self energy' of a particle. This name is given to the energy calculated by adding the energies of transition circulation between the rest state of the particle and all other possible states.

Thus in current theory the agreement between the transverse self energy and the rest mass becomes an important test of consistency. The test is not fulfilled.

It is not a long step from the results of §§ 122, 123, where we have found agreement between current radiation theory and fundamental theory, to the calculation of the transverse self energy; so that the fallacy of the current calculation must occur near the end. The expression for the transverse self energy is finally reduced to the integral[a]

$$W = \frac{1}{137\pi m}\int_0^\infty k\,dk, \tag{124·1}$$

k being the energy of the photon emitted or absorbed in the transition between the rest state and another state, or equivalently the energy-difference of the states. The divergence is a consequence of assuming an infinite universe. Since W should be equal to m, (124·1) is actually an equation for m^2 giving $m = \infty$—which it would be, according to our formulae, in a universe with N infinite. (Put $N = \infty$ in (40·8).)

To rectify (124·1) we first recall that the object-particle is always a top particle; and the other particles, which form the uranoid or planoid, pack the energy levels beneath it. We have previously calculated the top energy by exclusion theory; but the interchange energy, i.e. energy of transition circulation of the top particle with all the particles beneath, should amount to the same thing. Since m is the difference of energy between the top state and the bottom state, it is the upper limit of k. We have therefore to amend (124·1) by putting m instead of infinity for the upper limit. This gives

$$W = \frac{1}{2\pi}\cdot\frac{1}{137}m. \tag{124·2}$$

It may be noticed that the divisor $2\pi\,.\,137$ also appears in (122·63). It remains to trace the origin of this factor.

The usual practice of normalising wave functions so as to correspond to one particle per unit volume in x-space, makes the density of a particle equal to its mass. The calculation leading to (124·1) is based on the interstate energy tensor, involving the product of two momentum vectors, and W is strictly a density; m, on the other hand, is used strictly as the mass in the same calculation. The discrepancy in (124·2) is therefore a discrepancy between a density and a mass which were expected to agree. It is due to the current practice of incorporating neglected factors in the empirical constants. Writing (124·2) in the form

$$(2\pi\,.\,137)^2\,W = (2\pi\,.\,137)\,m, \tag{124·3}$$

we see that the time component of the momentum vector has been reckoned too small by the factor $2\pi\,.\,137$ since it occurs linearly on the right and quadratically on the left. These factors do not betray themselves in purely quantal investigations where we compare masses with masses, or in purely scale-free investigations where we compare densities with densities; so that there is little opportunity of detecting the discrepancy until the present part of the theory is reached.

As the discrepancy does not arise in our treatment which takes account of multiplicity and other factors at each state, it is scarcely our business to explain the factor in (124·3). However, its origin is fairly obvious. The factor 137 is supplied when we make the transformation from time in the observational system to the rigid time

[a] Heitler, *The Quantum Theory of Radiation*, p. 183, equation (23).

coordinate required by the theoretical equations. The factor 2π is the widening factor applied to the volume when we introduce a phase dimension. When the independent phases of the two state wave functions are combined into a single phase of the interstate wave functions, one phase dimension is eliminated, and it is the corresponding widening factor that appears in (124·2). The theory is similar to that of the factor $4\pi^2$ resulting from the elimination of two phase dimensions in § 123.

The equation (124·1) was intended to apply only to an electron; but in the amended form it applies to all particles, being a general relation between the interchange and exclusion representations of rest mass.[a]

It is instructive to examine another way of removing the divergence of (124·1), which connects the interchange representation with the gravitational representation of energy. We have seen (§ 38) that the current assumption of an infinite universe leads to neglect of the uncertainty of the origin and of the consequent weight function. The weight function reduces the number of states of large k and must therefore be inserted in (124·1) which is an integral over all possible states. Thus

$$W = \frac{1}{137\pi m}\int_0^\infty e^{-k^2/2\varpi^2}\, k\,dk = \frac{\varpi^2}{137\pi m}.$$

The value of ϖ is given in (38·72) as $\hbar/2\sigma = \hbar\sqrt{N}/R_0$. But in (40·4) $\hbar$ is changed to γ, and in terms of the practical constant $\hbar$ we have

$$\varpi^2 = \frac{136^2\hbar^2 N}{10}\frac{}{R_0^2} = \frac{5}{3}\cdot\frac{4}{3}\,136Mm_0$$

by (40·7). Factors $\frac{3}{5}$, $\frac{3}{4}$, 2π, $\frac{137}{136}$ are all familiar in connection with planoidal treatment, and we may conclude that these (together with a factor $\frac{1}{2}$ which might arise in many ways) have been omitted in the current theory.

[a] The detailed calculation of exclusion energy in § 43 applies to the particle of mass m_0; but the more rudimentary introduction of exclusion in the form of the 'principle of the top particle' in § 16 applies to any particle, and is sufficient for our purpose here.

Editorial Note

The manuscript ends at this point. The following note, probably written on the last day of Eddington's working life, was found with it, and indicates what the continuation would have been, had he been able to complete his plan. Mr N. B. Slater and Mr G. L. Clark have examined Eddington's papers in the hope of finding first drafts of the missing chapters, but the search has been unsuccessful.

Chapter XII

§ 125. Symbolic occupation.

§ 126. Einstein-Bose particles.

§ 127. Photons.

§ 128. Life-time of the mesotron.

Chapter XIII. Epistemological Theory

(as in *Proc. Camb. Phil. Soc.*, **40**, 37, expanded)[a]

Chapter XIV. Summary

§ 137. The principles of fundamental theory.

(Begin with scale, and the various forms in which scale fluctuation has been taken into account in different parts of the book. Then take the relation between T_{00} and $Z_0{}^0$ as the most fundamental starting-point. Survey of multiplicity factors. End with list of constants whose values are calculated in the book.)

[a] The title of the article in *Proc. Camb. Phil. Soc.* was 'The Evaluation of the Cosmical Number', printed as an Appendix on the opposite page.

APPENDIX

THE EVALUATION OF THE COSMICAL NUMBER[a]

1. The cosmical number $N = \frac{3}{2} \cdot 136 \cdot 2^{256}$ is most picturesquely described as 'the number of protons and electrons in the universe'. This in itself would be a matter of idle curiosity. But N has a more general significance as a fundamental constant which enters into many physical formulae; it determines the ratio of the electrical to the gravitational forces between particles, the range and magnitude of the non-Coulombian forces in atomic nuclei, and the cosmical repulsion manifested in the recession of the nebulae. Its special interpretation as the number of particles in the universe arises in the following way. If we consider a distribution of hydrogen in equilibrium at zero temperature, the presence of the matter produces a curvature of space, and the curvature causes the space to close when the number of particles contained in it reaches a certain total; this total is N. We cannot say with the same confidence that the number of particles in the actual universe is precisely N, because the admission of radiation, complex nuclei, and unsteady conditions takes the problem outside the range of rigorously developed theory; but to the best of our belief these complications do not affect the total number of protons and electrons composing the matter of the universe.

Before calculating their number, it would not be unreasonable to demand a definition of 'proton' and 'electron', and adopt for this purpose the mathematical specification that has been deduced from their observational properties. But the present investigation is more ambitious. It seeks to determine N directly from the principles of measurement. The proposition is that, as soon as we become obsessed with the idea that the right way to find out about the universe is to measure things, we are committed to an analytical conception which implicitly divided the universe into $\frac{3}{2} \cdot 136 \cdot 2^{256}$ particles. Naturally, in the course of counting the particles, we shall arrive at a mathematical specification of that which is being counted. From this specification we can determine the observational properties of the particles, and identify them with protons and electrons.

We have to show, not that there are N particles in the universe, but that anyone who accepts certain elementary principles of measurement must, if he is consistent, think there are. A logically complete demonstration, if it is possible, would be extremely prolix; and it is not the kind of problem I could myself attempt. But I shall try to show that at each stage the investigation is being driven by its own momentum—that the moves leading to a universe of N particles are forced. Or at least there is so much pressure behind the moves that, when we find the physicist actually does think there are N particles, there can be no doubt that it is the result of this pressure and not because of any peculiarity in the external world.

The whole calculation of N is an essay in the representation of conceptions by symbolic algebra. *It is the conceptions that matter.* We have to express in mathematical symbolism what we think we are doing when we measure things; for if we had no conception of what we were doing, the results of the measurements would not persuade us

[a] *Proc. Camb. Phil. Soc.* **40**, 37, 1944.

to believe anything in particular. All our results are derived from the condition that the conceptual interpretation which we place upon the results of measurement must be consistent with our conceptual interpretation of the process of measurement. Having examined critically our conceptual interpretation of the process of measurement, we have to define symbols with properties that correspond precisely to the conceptions introduced.

The investigation, although homogeneous in method, falls into two parts; §§ 2–4 might be described as a prologue to physics, and §§ 6–11 as an epilogue. The main body of physics lies up a side-branch indicated in § 5.

A preliminary investigation on these lines has been given in *P. & E.*, Chapter XVI.[a] The result there found ($2.136.2^{256}$) is now modified by a factor $\frac{3}{4}$, whose source will appear in equation (21) of this paper. The present paper is the first attempt at a full investigation.

2. We observe only relative positions and relative velocities; consequently an observable coordinate or momentum always involves two physical entities. A measurement involves four physical entities, two to furnish the observable that is said to be measured and two to furnish the comparison observable used as standard. For example, in a measurement of distance, the extension between two given objects is compared with the extension between two graduation marks on a standard scale. The usual allocation of the measure to two (or even to one) of the four entities concerned in it should be disregarded for the present; it prejudges the application that will be made of the measure in theory—a prejudgement that is often mistaken as, for example, when a supposed measurement of right ascension of a star is used as a measurement of the error of the clock. What is measured is something associated with four entities; we shall call this a *measurable*. Entities, observables involving two entities, and measurables involving two observables or four entities, form the rudiments of the structure that we are going to develop.

The most primitive entity contemplated resembles a geometrical point in that it has 'no parts and no magnitude'; but instead of 'position only'—a far from primitive attribute—it has existence (or non-existence) only. It has no magnitude, because magnitude is an attribute of a measurable. That it has no parts means that its existence is an unanalysable concept, not resolvable into existences of several parts, any of which can be conceived to exist without the others. We are not here referring to any metaphysical or absolute significance of existence—whether the entity exists in the sense in which *I* exist. All the concepts which we shall employ are structural concepts (*P.P.S.* pp. 144, 162); and the question is only whether the entity exists (or, more simply, is) in the structure contemplated. The symbol which we employ to represent the existence-attribute of an entity must therefore be such that its meaning only becomes definite when it is conjoined with another symbol representing the structure contemplated. This reduction of a symbol to an interpretable form by association with another symbol is expressed mathematically as a reduction to an eigenvalue. The kind of symbol

[a] References to the author's other publications are: *P. & E., Relativity Theory of Protons and Electrons* (Cambridge, 1936); *P.P.S., The Philosophy of Physical Science* (Cambridge, 1939); *D., The Combination of Relativity Theory and Quantum Theory*, a memoir published by the Dublin Institute of Advanced Studies, *Communications*, Series A, No. 2. [Editorial note: For the convenience of those readers of the present volume who do not possess *D.*, a note regarding it is inserted at the end of this Appendix.]

needed to represent the existence-attribute can be described as a 'yes-no symbol'. Distinguishing entities by suffixes, the existence symbol J_r of the rth entity has two, and only two, eigenvalues, namely a yes-eigenvalue a_{r1} and a no-eigenvalue a_{r0}. Its characteristic equation is therefore

$$(J_r - a_{r1})(J_r - a_{r0}) = 0. \tag{1}$$

It follows from (1) that $J_r - a_{r0}$ is an eigensymbol of J_r corresponding to the yes-eigenvalue a_{r1}; so that the entity exists in a structure represented by $J_r - a_{r0}$. Since this is formally the simplest structure in which the entity exists, we adopt $J_r - a_{r0}$ as a symbolic representation of the entity itself.

A fuller justification of this symbolism is as follows. Our mode of acquaintance through sense organs with a world, common to many human minds and therefore not contained within any individual human mind, is such that knowledge of it is necessarily limited to knowledge of structure. Our only means of describing abstract structure (that which possesses the structure being unknowable) is provided by the methods of symbolic algebra; these define a structural pattern by the interrelations of a closed group of symbolic operations (*P.P.S.* Chapter IX). As elements of such a structure, entities and their existence-attributes must be represented by symbols. The self properties of any symbol—those definable without reference to other symbols of the structure—are contained in its characteristic equation. But in this connexion the field of numbers, which supplies the coefficients in the characteristic equation, must be counted as symbols of the structure; because it is part of the symbolic method to replace symbols by ordinary numbers when they behave as ordinary numbers with respect to all the symbolic operations contemplated, just as multiples of the unit matrix may be, and commonly are, replaced by ordinary numbers in matrix calculus. Thus the only genuine self property of J_r postulated in (1) is the degree 2 of its characteristic equation, or equivalently the number 2 of its eigenvalues. It was necessary to choose for the existence-attribute a symbol with self properties which correspond to our conception of the attribute as a choice of two alternatives, existence or non-existence. The only possible correspondence is that the characteristic equation shall be a quadratic offering a choice of two roots.

The symbol J_r by itself represents the offer of a choice. It is a feature of modern theoretical physics that we are in no hurry to decide the choice. If the entity is one of those concerned in a measurement that has actually been made, J_r must, of course, reduce to the yes-eigenvalue. But in the theoretical systematisation of physical knowledge we often introduce intermediary measurables, which are combined[a] or transformed into other measurables before any question of actual measurement arises. For intermediary measurables it is not necessary that J shall reduce to an eigenvalue in the structure contemplated; the entity is then said to have *partial existence* in the structure.

Partial existence is usually explained as 'probability of existence'; but whether the explanation clarifies or obscures its nature may be questioned. In science, 'probability' most commonly occurs in connexion with statistical description, and is then an ordinary number between 0 and 1. Numerical probability applies to a statistical individual, or unidentified member of a large assemblage. But as applied to a single individual or event, which affords no scope for statistical interpretation, all we can say about it is

[a] The most common kind of combination is 'averaging'.

that it is a non-numerical concept with two numerical limits 1 and 0 which have definite meaning. This comes to the same thing as saying that it is a yes-no symbol with eigenvalues 1 and 0.

We have referred to intermediary measurables which are not supposed to be actually measured. It has often seemed illogical that, although physicists are continually stressing the importance of distinguishing between observables and unobservables, they are not over-scrupulous in postulating observations of a highly impracticable kind. The inconsistency is removed if we recognise that the terms 'observable' and 'measurable' refer to structure. We appeal to actual observation and measurement to determine the structure that an observable or measurable must have; but thereafter anything having the defined structure is classed as observable or measurable, without implying that appliances for observing or measuring it could be designed. The unpardonable inconsistency is to treat as a measurable something that has not the structure of a measurable.

3. The most primitive measurable is provided by four entities whose existence-attributes are independent. It will not exist unless all four entities exist. It has therefore an existence symbol of the form $M = J_r J_s J_l J_u$, the multiplication being understood to be commutative (outer or 'direct' multiplication). This has 16 eigenvalues, of which $a_{r1}a_{s1}a_{l1}a_{u1}$ is the yes-eigenvalue and the other 15 are no-eigenvalues. Thus the characteristic feature of a measurable is that it has 15 different ways of not existing.

The four entities will provide a variety of things to measure, and the one existence symbol M applies to them all collectively. The measures are therefore to be regarded as belonging to one measurable whose existence-attribute is represented by M. We shall call this a *tensor measurable*; and the array of measures constitutes its *tensor measure*. We shall be concerned throughout with tensor measurables and tensor measures.

We have found that the entity exists in $J_r - a_{r0}$. In the same way we find that the measurable exists in

$$M' = (J_r - a_{r0})(J_s - a_{s0})(J_l - a_{l0})(J_u - a_{u0}), \tag{2}$$

and, since this is analytically the simplest expression in which the measurable exists, we adopt M' as the symbolic representation of the measurable itself. M', like M, has 16 eigenvalues. The measurable is thus represented as the outer product of the four entities involved in it. A similar rule applies to any system composed of conceptually independent parts, so that its existence is contingent on the exercise of all its parts; it is represented symbolically as the product of its parts. This has two very general consequences. First, the conception of a system as the product (not the sum) of its parts lends itself to a probability interpretation, since probabilities are multiplicative. Secondly, all those characteristics, whose additive character has led to the familiar conception of a system as the sum of its parts, must be represented as exponents of exponentials, so that they may combine by addition when the exponentials are combined by multiplication. This is a well-known principle of wave mechanics.

Physical science may be defined as 'the systematisation of knowledge obtained by measurement'. It is a convention that this knowledge shall be formulated as a description of a world—called the 'physical universe'. We now apply the results that have been obtained to the physical universe. Since the data are measures, the measurable

forms the natural starting-point for the theory. Having satisfied ourselves that a measurable is represented by a symbol with 16 eigenvalues, we now drop the preliminary scaffolding; and any entities introduced in the physical scheme will be defined in terms of measurables, instead of vice versa. The structure repeats itself indefinitely; that is to say, four non-primitive entities will furnish non-primitive measurables, which in turn may be used to define further entities. This cyclic return of structure into itself is characteristic of the mathematical definition of structure by closed groups and algebras. The measurables in successive cycles are always symbols with 16 eigenvalues; the only difference between them is in the undefined significance of the unit of eigenvalue.

We have assumed that the measure depends only on the four entities concerned in it. This is by no means self-evident. The four entities do not shout their measure at us. We shall later examine the practice of measurement more closely to see whether it postulates the existence of something, which we may call Z, besides the four entities (§ 9). It is almost always assumed, and to a very high approximation it is true, that whatever else is required to make the measurement possible exists unconditionally in the actual universe; its non-existence could only come about by one of those fantastically improbable coincidences which are formally possible in a statistical assemblage, but for all practical purposes are ignored. The non-existence of Z is therefore something that the physicist refuses to conceive; so that, although Z is necessary to the measure, this does not increase the number of ways (fifteen) in which the measurable might not exist. It happens that the calculation of N is the one investigation for which this approximation is inadequate. This places us in a difficulty as regards nomenclature. To mark the special character of the calculation of N, I have described it in § 1 as an epilogue to physics. I shall therefore use the term 'measurable', in accordance with the basal approximation of physics, for the concept associated solely with the four entities; but in the epilogue it will be necessary to drop the approximation and introduce the 'true measurable'—a combination of the measurable and Z—to which the measure properly belongs. This will have more than 16 eigenvalues; or we may put it that the higher magnification used in the epilogue shows a fine structure in each of the original 16 eigenvalues.

The data of physics are measures; but we can make nothing of a mere collection of measures without any note of the objects and circumstances to which they refer. The crux of the problem is to supply 'connectivity' to the measures; *so that in the theoretical treatment there may be an equivalent for that part of the procedure of measurement which consists in noting the objects and circumstances to which the measures relate.*

In the mathematical specification we indicate distinctions between different measurables and between different measures by tagging suffixes to them; so that we deal with a set of measurables M_p and a set of measures X_q. But the numbering of the different components of a structure is useless unless it is accompanied by a key-plan. Thus in using M_p, X_q there is an implied promise to supply a key-plan—to define how the distinction referred to in the suffixes is made.[a] This sets before us a definite problem, namely, to discover a structure of measures and measurables which is such that this

[a] The difficulty of distinguishing measures is less obvious than the difficulty of distinguishing measurables. In a scalar measure there is no difficulty, because the measure-number is the distinction. But a tensor measure is an array of measure-numbers; and we have to define the distinction between different permutations of the numbers in the array, which are, of course, entirely different measures.

promise can be fulfilled. This will be called 'the structure problem'. It appears to have a unique solution, bearing in mind the condition that the measurables must have just 16 eigenvalues.

The position is that we cannot make a beginning in structural theory without running into definitionary debt. We may not run up a debt greater than we can pay off. After discharging the initial debt, we pay as we go. So long as we are in debt our moves are forced (§ 1) by the necessity of extricating ourselves. At the end of the structure problem we stand all square with everything structurally defined; and thereafter we should be careful to define new quantities or distinctions as we introduce them by relating them to the quantities which the structure problem has defined.

4. A measurable has a unique measure, but any number of measurables may have the same measure X_q. There are two ways of regarding the relation of measure and measurable: (1) the measure is 'carried' by the measurable; (2) the measurable 'occupies' the measure. We have hitherto taken the first view. Following now the second view, we can, without particularising the distinction between different measurables, state how many measurables occupy X_q. We introduce therefore an occupation symbol K_q, such that the eigenvalue of K_q is the number of measurables occupying X_q. Just as the existence symbol M_p has in the structure a correlative M'_p representing that which exists, so the occupation symbol K_q has a correlative K'_q, called the *occupation operand*, representing that which occupies. Initially, K'_q gives a purely quantitative representation of the occupation, namely, $1, 2, 3, \ldots$ measurables; or, if K'_q is not an eigensymbol of K_q, it corresponds to a symbolic quantity of occupation.

By the definition of K_q its eigenvalues are the positive integers (including 0). In the study of small systems artificially separated from the rest of the universe, an enumeration of the number of carriers of each particular measure is appropriate; but this application of K_q comes much later in physical theory, after we have developed analytical machinery for dividing the universe into semi-independent systems. At present we can only treat the universe as a whole; and an enumeration of the particles throughout the universe which have a measure X_q is scarcely the right way to approach 'the systematisation of knowledge obtained by measurement'. What more nearly concerns our knowledge is the change in the number, brought about by the phenomena we investigate. We therefore modify the definition of K_q so that its eigenvalue is the *excess of occupation* above a large, but at present unstated, integer $\frac{1}{2}n$. The eigenvalues of K_q are then the positive and negative integers, or equivalently the roots of $\sin \pi k = 0$. The characteristic equation of the symbols K is accordingly

$$\sin \pi K = 0. \tag{3}$$

More strictly, this is the limiting form of the characteristic equation when the possible eigenvalues range from $-\frac{1}{2}n$ to $\frac{1}{2}n$, and $n \to \infty$.

According to the rule that a structure is represented as the product of its parts, if K'_1 represents occupation by one excess measurable, $(K'_1)^m$ will represent occupation by m excess measurables. Since

$$m(K'_1)^m = \frac{\partial}{\partial(\log K'_1)} (K'_1)^m, \tag{4}$$

$\partial/\partial(\log K'_1)$ is a symbol which, in conjunction with an excess occupation operand, reduces to an eigenvalue m equal to the number of measurables forming the excess

occupation. But this is the definition of the symbol K; so that $K = \partial/\partial(\log K_1')$. Hence, putting

$$K_1' = e^{i\theta}, \tag{5}$$

we have

$$K = -i\partial/\partial\theta. \tag{6}$$

Of course θ, like K_1', is a symbol (not a number); and we shall presently find that it has a property which would be impossible for a number.

For any function $f(\theta)$,

$$e^{\pi\partial/\partial\theta}f(\theta) - e^{-\pi\partial/\partial\theta}f(\theta) = 2i\sin\pi K f(\theta) = 0,$$

by (3); so that, by Taylor's theorem,

$$f(\theta+\pi) - f(\theta-\pi) = 0. \tag{7}$$

Thus θ is a symbol of such a nature that $\theta+\pi$ is indistinguishable from $\theta-\pi$. Such symbols have a familiar representation as geometrical angles. This is the first suggestion of a spatial picture of any part of the structure of the physical universe. By adopting it we make no ultimate difference, because geometrical concepts, like other concepts, are defined by symbolic algebra; but we obtain a ready-made vocabulary, and so get a respite from the strain of inventing our language as we proceed. By representing θ more particularly as a rotation angle we are able to apply the theory of rotation groups.

By (5) the occupation operand corresponding to $K = 1$ is $e^{i\theta}$. Since this is a symbolic representation of whatever is occupying the measure (in this case just one measurable) $e^{i\theta}$ is the representation of an excess measurable. But since we have been considering only the quantitative aspect of the occupation, $e^{i\theta}$ does not as it stands show any distinction between different measurables. If, however, we replace it by $e^{i\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ is a symbol representing a rotation angle θ *in a particular plane*, we introduce a qualitative distinction. This distinction can be structurally defined by means of a rotation group. The group is unique, because $e^{i\boldsymbol{\theta}}$, being a measurable, has 16 eigenvalues.

We have therefore to pick out from the known rotation groups the one in which the rotations are symbols with 16 eigenvalues. This is the group which defines the double frame, or EF-frame, treated in *P. & E.* Chapter x. The matrices with 16 rows and columns form the same group. Thus $e^{i\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$ are EF-numbers, or equivalently 16-fold matrices.

It may be remarked that, if we *assume* that measurables can be represented by matrices, it follows directly from (2) that the matrices are 16-fold. For the J_r, having 2 eigenvalues, will be represented by 2-fold matrices, and the measurable is an outer product of four of them. But this short cut, besides involving an assumption, misses points that will be needed in the subsequent developments.

The full representation $e^{i\boldsymbol{\theta}}$ of a measurable is reduced to the merely quantitative representation $e^{i\theta}$ by associating an eigensymbol with it. In general the same eigensymbol will reduce a number of EF-symbols to eigenvalues simultaneously. It is therefore not necessary that $\boldsymbol{\theta}$ should be a simple rotation in one plane of the EF-frame; it may be the sum of rotations in different planes, provided that the plane-symbols have a common eigensymbol. The general form of $\boldsymbol{\theta}$ is

$$\boldsymbol{\theta} = \Sigma EF_\mu\theta_\mu \tag{8}$$

summed over a subset of commuting EF-symbols; for it is always possible to find a common eigensymbol of a set of commuting matrices. Planes whose symbols commute

are said to be antiperpendicular; so that $\boldsymbol{\theta}$ is composed of components in antiperpendicular planes.

The eigenvalues of the EF-symbols are ± 1; so that, in conjunction with an eigensymbol, (8) reduces to

$$\theta = \Sigma(\pm\theta_\mu). \tag{9}$$

But there are relations between the simultaneous eigenvalues of the EF_μ, so that the combination of signs in (9) is not entirely arbitrary.

5. A measure X_q is an array of numbers $x_{q1}, x_{q2}, \ldots$. We commonly represent such an array as a symbolic sum $X_q = \Sigma A_r x_{qr}$, where the symbols A_r distinguish the places in the array. But unless the A_r are structurally defined, i.e. connected with symbols whose group relations form a defined structure, we lack the 'key-plan' for interpreting the notation. At the present stage the most general number array that is structurally defined is an EF-number. We must therefore take the X_q to be EF-numbers.

Measures, unlike measurables, are not safeguarded against extreme complexity by a limitation of the number of their eigenvalues. (A measure is not to be thought of as existent or non-existent; the worst that can happen to it is to be totally unoccupied.) Thus the measures concerned in the structure problem, which are EF-numbers, may be a selection from a larger class of measures—a selection forced on us by the conditions of the structure problem. The point is that, when in § 4 we enumerated the measurables occupying X_q without explaining how X_q is distinguished from other measures, we contracted a definitionary debt which makes the whole investigation meaningless, unless it takes the form of a cycle supplying the missing definition; so that our results apply to the occupation of measures selected by the condition that their distinction can be defined by those results.[a]

It is worth while to consider this selection of measures from the physical point of view. Every comparison of two comparable observables is a measure, and has a right to representation in the physical universe which embodies all knowledge obtained by measurement. But in scientific investigation we do not go about casually comparing any two observables we come across. We have gradually discovered which of many possible measurements provide a basis for systematisation and unification; and it is these basal measures that have received recognised names. It is therefore not a blemish on the present investigation that it does not immediately confront us with the whole disordered mass of possible measures, but selects a set of basal measures. We shall see immediately that physical science, in its search for systematisation, had already arrived at just the same selection.

The EF-numbers are space tensors of the second rank (*P. & E.* § 10·2).[b] It is well known that, in the systematisation of molar physics provided by relativity theory, the fundamental measures $(g_{\mu\nu}, T_{\mu\nu})$ are tensors of the second rank. The EF-frame (with its characteristic number 136) is also the basis of the author's development of relativistic quantum theory in *P. & E.* and *D.*; the particle theory begins with the standard carrier, which is the carrier of an unspecialised element of energy tensor and nothing

[a] The first round of a structural cycle is always, so to speak, a trial run. In later rounds all quantities are defined cyclically in terms of one another, so that the cycle can be continued forwards or backwards indefinitely.

[b] In the classification of rank here adopted, two antisymmetrical suffixes count as one; so that a 6-vector (like a 4-vector) is of the first rank, and the Riemann-Christoffel tensor (like the energy tensor) is of the second rank.

more (*D.* § 11). The 'standard carrier' and its 'complete energy tensor' are in fact the names under which the structural concepts of a measurable and its second-rank tensor measure pass into the main parts of physics.

Thus when the structure theory reaches the *EF*-frame it enters territory already widely explored, and we pass straight on to relativity theory and relativistic quantum theory. It is shown in *D.* that all the fundamental physical constants *except the cosmical number* can be derived, and that the values agree with observation. There has been great improvement in the experimental data in the last few years, and the observational test is now very stringent. For the cosmical number we must return to the structure problem. The *EF*-frame enables us to distinguish measures, but we have yet to show explicitly how measurables with the same measure are distinguished from one another.

It may be asked, why should we trouble to distinguish measurables (now identified as standard particles) from one another, seeing that it is a fundamental tenet of quantum theory that elementary particles have no distinctive identity? But, as we pointed out earlier, it is the conception that matters—not whether particles are distinguishable, but whether they are conceived as distinct. Clearly, the physicist conceives electrons as distinct, otherwise he would not talk about exchanging them; if they are continually exchanging identity, they must be conceived as having an identity to exchange. We are not concerned with any metaphysical conception of identity. Whatever it is that is exchanged—whether it is called 'identity' or merely a 'suffix' —has to have a structural equivalent.

But the fundamental reason for distinguishing measurables is that which has been referred to in § 3, namely to provide connectivity. An actual measure without note of the object and circumstances is useless scientifically. It has to be treated as a numbered part to be fitted into a numbered place in the key-plan of a complicated structure; and we must continue the same system of attachment of measures to particular measurables in the analytical theory. The structure of the universe has therefore to be formulated in such a way that this association of measures with structurally identifiable measures is latent in it. We say 'latent', because fundamental physics rarely brings it to light. That is because fundamental physics is concerned with generalisations, which are true independently of particular identifications; but in that respect fundamental physics is not typical of physical science as a whole.

6. We take advantage of the connexion of the structure theory with the ordinary developments of physics (made through the *EF*-frame) to introduce less abstract terminology. The basal measure, which forms an *EF*-number, is the energy tensor. More precisely, it is the 'complete energy tensor' with 136 independent components, which includes the ordinary energy tensor as well as spin components and other consequential variates. These components are all classed as generalised momenta. It is in keeping with the uncertainty principle that no coordinates are included; for if the momenta are exact, the coordinates are entirely uncertain, i.e. unmeasurable. Before coordinates can be introduced we require a conception not yet provided for, namely, a measurable with uncertain measure. This conception will arise presently; but it is not necessary for our present purpose to pursue the developments to which it leads.

A 'conceptual carrier of variates' is generally called a particle. 'Particle' is a wider term than 'measurable'; but, since a measurable is one kind of particle, we shall often refer to it as a particle.

The measure of a measurable is a definite characteristic of it; so that the complete specification $e^{i\boldsymbol{\theta}}$ of a measurable must contain all the data necessary for determining the array of numbers which forms its measure. We have therefore to consider how the measure is to be extracted from the symbolic specification of the measurable. Since the measure is an EF-number, the natural suggestion is to identify it either with $\boldsymbol{\theta}$ or $e^{i\boldsymbol{\theta}}$ which are EF-numbers. But this implies that the measurable is completely defined by the measure which it occupies, and is contrary to the principle that there must be a latent distinction between the various measurables which may happen to have the same measure. Thus the measure X must be extracted from $\boldsymbol{\theta}$, but there must be something left over which is included in $\boldsymbol{\theta}$ but not in X. An important clue to the nature of that which is left over is provided by the consideration that the distinction between the measurables occupying X must be *discrete*. Two energy tensors may be nearly the same, but we cannot speak of measuring the energy tensors of two particles which are 'nearly the same'; conceptually, if two particles are different at all, they are altogether different—there are no degrees of difference of identity. In the structure that we have developed the only basis for a discrete distinction is provided by the sign ambiguities in (9); and this indicates definitely the direction in which our analysis must proceed.

Before treating the EF-frame it is convenient to examine the corresponding problem in a simple E-frame, using analogues of measures and measurables. The 'measurable' is then an E-number and can be represented by a fourfold matrix. By suitable orientation of the axes in the E-frame, any E-number whose four eigenvalues are all different can be reduced to the antitetradic form

$$i\boldsymbol{\theta} = E_{\mu\nu}\theta_1 + E_{\sigma\tau}\theta_2 + E_{\lambda\rho}\theta_3 + E_{16}\theta_4, \tag{10}$$

$\mu, \nu, \sigma, \tau, \lambda, \rho$ being different suffixes.[a] The four terms in (10) commute; so that $\boldsymbol{\theta}$ has components in four antiperpendicular planes.

The measure X is also an E-number or fourfold matrix. Denoting the matrix by $X_{\alpha\beta}$, we distinguish especially the measures for which $X_{\alpha\beta}$ is the product of two vectors (wave vectors) ψ_α, χ_β; these will be called *pure measures*. Purity is invariant for tensor transformations; and it can be described independently of matrix representation, the condition being that, if the E-number is normalised so as to make its quarterspur (i.e. its algebraic component $E_{16}x_{16}$) equal to $\frac{1}{4}$, it becomes idempotent. It has been shown (*P. & E.* § 5·5) that by suitable orientation of axes a normalised pure measure can be reduced to the antitetradic form

$$X = -\tfrac{1}{4}i(E_{\mu\nu} + E_{\sigma\tau} + E_{\lambda\rho} + E_{16}), \tag{11}$$

and further that the wave vector factors ψ, χ are right and left eigensymbols of $E_{\mu\nu}$, $E_{\sigma\tau}$, $E_{\lambda\rho}$, E_{16}.

[a] The corresponding theorem in matrix algebra is that any matrix of which the characteristic roots are distinct can be transformed into a diagonal matrix. The limitation to distinct eigenvalues can be disregarded in the physical application, since exact coalescence is a limiting conception which could be approached but never actually reached in natural conditions.

Provisionally we shall suppose that the basal measure is a pure measure, and that its planes agree with those of the measurable which carries it. This assumption will be justified later. It can then be defined (apart from a possible numerical factor) as the outer product of the right and left eigenvectors of $\boldsymbol{\theta}$. In conjunction with ψ or χ or both, (10) reduces to

$$\theta = \pm\theta_1 \pm \theta_2 \pm \theta_3 \pm \theta_4, \tag{12}$$

the eigenvalues of the E-symbols being $\pm i$.

By (6), θ is (according to the usual nomenclature) a generalised coordinate conjugate to the generalised momentum K. We therefore call θ the *phase coordinate*, or simply the phase. (K can be identified as the *scale momentum*.) We call θ_1, θ_2, θ_3, θ_4 the phase subcoordinates. Adopting the positive signs in (12), the wave front or locus of constant phase is a three-dimensional flat in the subcoordinate system, namely $\theta_1+\theta_2+\theta_3+\theta_4 =$ constant. This gives the familiar picture of $e^{i\theta}$ as a system of plane periodic waves ($\theta+\pi$ being indistinguishable from $\theta-\pi$). The waves, however, are in a non-Pythagorean space, the axes being antiperpendicular (commuting) instead of perpendicular (anticommuting).

A symbol $E_{\mu\nu}$ represents not only a plane but a positive direction of rotation in the plane; and there is an alternative symbol $E_{\nu\mu}$ ($= -E_{\mu\nu}$) to represent the plane with an opposite direction of rotation. (For E_{16} the alternatives are $\pm i$.) Thus the same set of four antiperpendicular planes has 16 different symbolic representations; these will be called *reflexions*. Different reflexions correspond to different choices of the positive directions of the four subcoordinate axes. We have written down (10) and (11) as the relation between measure and measurable; but, if the relation is simply that the planes agree, this is too precise, for the plane-symbols in (10) may be any reflexion of the four plane-symbols in (11). Thus when the components of the measure are given, the measurable may be represented by any one of 16 wave systems resulting from the different combinations of sign in (12).[a]

The set of 16 reflexions will be called the *grid*, the number 16 being the *grid constant*. Reflexions will be distinguished by a grid-number g, which may conveniently be taken as the number on the binary scale formed by writing 1 or 0 for + and − in the sign combination.

Apart from a numerical factor, a pure measure X_1 can be transformed into any other pure measure X_2 by a relativity rotation in the E-frame. To preserve agreement of the planes, the subcoordinate space, carrying the grid with it, must be rotated along with the measure. Thus it is possible to identify corresponding grid-numbers in different measures. *The grid-number of a measurable therefore forms an identifying characteristic which is independent of the measure.* This enables us to introduce the concept (mentioned earlier) of a measurable with uncertain measure. An existent measurable, distinguished by a grid-number, can be supposed to occupy partially two or more measures, the occupation factors being symbols whose sum is unity.

[a] This does not follow directly from (12), because in (12) as originally derived the sign ambiguities are not all independent, and only 8 combinations are admissible. The 16-fold ambiguity arises at an earlier stage, as explained above. In the simple E-frame a complication occurs, because the reality conditions are such that the real planes of the measure correspond to imaginary planes of the measurable, and vice versa. But since this complication does not occur in the corresponding analysis of the EF-frame (to which the analysis of the E-frame is merely a preliminary) it does not really concern us—except as an indication that the simple frame is inadequate for representing true measurables.

We now proceed to justify our provisional assumption that the basal measure is pure. An impure measure is reducible to the form $X = E_{\mu\nu}x_1 + E_{\sigma\tau}x_2 + E_{\lambda\rho}x_3 + E_{16}x_4$. Every constituent of this expression which is not an absolute constant must be determined by the measurable $e^{i\theta}$. The four planes of X and θ must agree, since θ does not define any other set of four antiperpendicular planes. The coefficients x_α, if they are not equal to an absolute constant, must be determined by the components θ_α; but, on the other hand, the connexion must not be such that the θ_α are completely determined by the x_α. Evidently there is no relativistically invariant way of satisfying this; and accordingly the four x_α must be equal to the same absolute constant. This condition makes X pure. We notice that the scalar magnitude of the basal measure has to be an absolute constant, differences of measure being limited to differences of orientation in the E-frame. In ordinary language the discrete measurables which occupy the basal measures are all similar particles. These similar particles are the 'standard carriers', which form the starting-point of the relativistic quantum theory developed in D.[a]

To summarise the argument: we have to connect two E-numbers, a measurable and a measure, in such a way that the measurable completely determines the measure, but the measure does not completely determine the measurable; and further that, when the measure is given, the indeterminacy of the measurable (or part of the indeterminacy) consists of a discrete set of alternatives. These conditions are sufficient to show that the measure is a pure E-number, and that the set of alternatives is a grid associated with the antiperpendicular planes of the pure E-number.

7. The foregoing treatment applies with very little modification to the double frame. The general EF-number, or 16-fold matrix, is (if the eigenvalues are distinct) reducible by suitable choice of axes in the EF-frame to the form

$$\theta = \Sigma E_\alpha F_\beta \theta_{\alpha\beta} \quad (\alpha, \beta = 1, 2, 3, 4), \tag{13}$$

where E_1, E_2, E_3, E_4 and F_1, F_2, F_3, F_4 are antitetrads. The 16 symbols in (13) commute with one another, and define 16 antiperpendicular planes. They have common right- and left eigensymbols (double wave vectors) whose outer product gives a pure EF-number X. When X is normalised so as to be idempotent, it becomes $X = -\frac{1}{16}\Sigma E_\alpha F_\beta$. By our previous argument this must be the measure of $e^{i\theta}$. The coefficient i (which, if attended to, gives trouble in the simple frame) does not occur in the double frame. The only other difference is that there are now 16 phase subcoordinates, and the grid consists of 2^{16} reflexions.

For later use we require the corresponding results for a quadruple frame or $EFGH$-frame. There are then 256 antiperpendicular planes represented by $E_\alpha F_\beta G_\gamma H_\delta$ $(\alpha, \beta, \gamma, \delta = 1, 2, 3, 4)$, formed by combining sets of four commuting symbols in each of the four constituent frames. There are accordingly 256 phase subcoordinates, and the grid consists of 2^{256} reflexions.

We shall treat the grid constant as extremely large. This is scarcely true of the constant 2^{16} of the double grid; but ultimately it is the quadruple grid, with constant 2^{256}, that has a physical application.

8. The next step is dictated by the fact that our results from § 4 onwards refer to excess measurables; in particular the measurables that possess a distinctive grid number are excess measurables. From the observational point of view this is the right starting-

[a] This refers to the corresponding particles in the double frame.

point; for any small system that we measure is contemplated as an excess superposed on a constant environment which constitutes the rest of the universe. I have called the idealised universe, which forms the standard environment of any microscopic system, the *uranoid* (by analogy with the geoid in geodesy). Of course, no mathematical formula ever professes to determine what a system will actually do; it states what the system *would do* under certain conditions, some of which may be stated explicitly and others are assumed implicitly. The implicit conditions include the specification of the standard environment or uranoid on which the system is superposed. The measurables of the standard environment contribute the standard occupation of the measure X; and the measurables that we are particularly interested in—those of the object-system inserted in the environment—accordingly appear as excess measurables additional to the standard occupation.

Accordingly the excess particle is the practical starting-point. But one cannot build a universe in which every particle simultaneously is an excess particle. On the other hand, the rest of the universe is supposed to be formed of particles similar to those which form the excess object-system. Our problem now is to introduce particles which, although not necessarily associated with excess occupation, have characteristics (phase and grid-number) similar to those of excess particles. These are called *Einstein-Bose* (E.B.) particles.

The wave system $e^{i\theta}$ represents one excess measurable. Taking another eigenvalue k of K, we obtain a wave system $e^{ik\theta}$ representing k excess measurables. The whole excess is represented by one wave system in which k measurables are fused together, so that one phase θ and one grid-number g apply to them collectively. They are therefore to be regarded as forming one measurable of multiplicity, or weight, k. The point to be noticed is that there is never more than one phase associated with the same grid-number, multiple occupation being indicated by the wave-length and not by the coexistence of several phases. The same condition must be supposed to apply to E.B. particles, since they are in all respects similar to excess particles. If this condition were not imposed, we should have two different ways of representing multiple occupation, and a great deal of confusion would ensue.

We are concerned with three kinds of distinction, q specifying the measure, k the quantitative occupation, and g the qualitative occupation. All three are combined in the 'state' S_{qgk}. The term 'state' is used for any combination of distinctions or variates, and the term 'particle' for a carrier of the combination or an occupant of the state. This gives a fluidity of description, which is needed in transferring the properties of excess particles to E.B. particles. The latter are described as occupants of states S_{qgk} or equivalently as k-occupants of states S_{qg}. We introduce an occupation symbol J_{qgk} whose eigenvalue is the number of k-occupants of S_{qg}. The only eigenvalues of J_{qgk} are 1 and 0; for we have seen that two k-occupants with the same g would fuse into a $2k$-occupant. The total occupation of the measure X_q is then

$$\Sigma_g k J_{qgk}, \tag{14}$$

and K_q is the excess of this above a certain standard occupation.

For definiteness we take a uranoid composed of particles of weight 1, and therefore drop the suffix k.[a] Then, if n is the grid constant, and nx the standard occupation,

$$K_q = \Sigma_g J_{qg} - nx = \Sigma_g (J_{qg} - x). \tag{15}$$

[a] This means in effect that we treat a universe composed of hydrogen (as remarked in § 1).

If $x = \frac{1}{2}$, the eigenvalues of $J_{qg} - x$ are $\pm\frac{1}{2}$, and the eigenvalues of K_q are the integers from $-\frac{1}{2}n$ to $\frac{1}{2}n$. The characteristic equation of K_q is therefore

$$K \prod_{r=1}^{r=\frac{1}{2}n} \left(1 - \frac{K^2}{r^2}\right) = 0, \tag{16}$$

which converges to $\sin \pi K = 0$ when n is large. Since this does not hold if x differs appreciably from $\frac{1}{2}$, we see that, in order that (15) may be consistent with (3), $x \to \frac{1}{2}$ as $n \to \infty$. In our application n is extremely large; and we therefore have

$$K_q = \Sigma_g Y_{qg}, \quad Y = J - \tfrac{1}{2}. \tag{17}$$

We distinguish Y as the E.B. occupation factor, and J as the F.D. (Fermi-Dirac) occupation factor. The E.B. occupation of S_{qg} is measured from an initial occupation $\frac{1}{2}$, and the F.D. occupation is measured from an initial occupation zero. We can regard q as a major characteristic and g as a minor characteristic of the state S_{qg}. If minor characteristics are ignored, the separate states S_{qg} combine into a state S_q (identical with the measure X_q) of multiplicity n, whose occupation is $K_q = \Sigma_g Y_{qg}$. It will be seen that, when E.B. occupation and corresponding E.B. particles are used, the concept of 'excess' no longer appears explicitly. An interesting point is that, when the occupation factor K_q of the multiple state reduces to an eigenvalue, so that the occupation is represented by $e^{ik\theta}$, it has to choose a reflexion for its wave system and so acquires a grid-number g'. According to the usual outlook, the distinction g indicates what we loosely call the 'individuality' of the particle. We create a particle of composite individuality by combining other particles; but this immediately acquires a simple individuality g' of its own, homologous with, but not derived from, the individualities of the particles from which it was formed. In fact it may well happen that the original particle with grid-number g' was non-existent ($J_{qg'} = 0$).

The distinction between particles has been based on the three characteristics q, g, k. There is a fourth characteristic θ carried by the particle, but this does not provide any additional means of distinction; because, although it may have different values, there cannot exist any other particle with the same g discriminated by a different value of θ. There is no way of defining structurally a common zero-point from which the phases belonging to different reflexions are reckoned; as we ordinarily say, the phases are *incoherent*. From the point of view of wave mechanics this is the most practical consequence of the grid distinction; since it is a fundamental principle in current theory that wave functions associated with different particles have incoherent phase; whereas wave functions associated with the same particle (as in a wave-packet) have coherent phase.

Since there is one phase associated with each grid-number, an enumeration of the particles as distinguished in current theory by incoherent phase, should agree with our enumeration of the particles distinguished by grid-numbers.

9. We have treated a measure as a number, associated with two observables, that we somehow become aware of. But the interpretation of measurement introduces an additional conception, namely *graduation*. If we play about with a pin and a metre standard, we do not become aware of any number in particular unless the standard has been graduated, or unless we use a process of displacement which we interpret as adding pin-extensions. Since length (a measure) is not assigned to the pin till the procedure is completed, we have to use a different term 'extension' in describing the processes which lead up to the determination of length. Moreover, the length is a

property of the measurable formed by the pin and standard, whereas the property of extension is supposed to reside in the pin alone or standard alone. Extension may be called an *observe*, because it is related to an observable in the same way that a measure is related to a measurable. The length that we determine is interpreted conceptually as the ratio of the extension of the pin to the extension of the standard.

Thus the scientific interpretation of measurement introduces two new postulates: (1) the measure is the ratio of two observes, and (2) there is an additivity of observes which is the basis of graduation. But if observes can be added and divided, they form a field homologous with numbers. Thus, although we have carefully distinguished observes from measures (extensions from lengths), we find ourselves attributing to observes properties scarcely distinguishable from those of measures. This often leads to confusion, as when we catch ourselves saying 'the length of an object is the ratio of its length to that of the standard'.

In accordance with our principle that it is the conception that matters, we emphasise that these postulates refer to the conceptual interpretation of measurement. Scientific measurement is a process which has evolved, and is still evolving; and its evolution is guided by our conceptual interpretation of it. It is not a practical process hit on by accident whose meaning we have to discover; we first make up our minds what we think a measurement ought to mean, and then design a procedure which will yield a 'good' measurement, i.e. a measurement to which we can without inconsistency assign that meaning.

The fact is that, although measurement is primarily a process involving four entities, the conceptual interpretation of measurement postulates in addition the existence (in the structure contemplated) of something referred to as Z in § 3. The existence of Z is the condition that makes graduation an exact concept. Following Whitehead, we might call Z a 'basis of uniformity'; but it is more familiar in modern physics as a 'metrical field'. It is by conjunction with Z that observables acquire numerical characteristics which seem more properly to belong to measurables. In the actual universe there exists a basis which is uniform to a very high approximation, and this serves for almost all purposes; but to the much higher approximation required in the calculation of N the ideal exact basis of uniformity does not exist. The finitude of N is, in fact, the cause of the failure of the approximation. In the present investigation we must take account of the failure of the approximation; but we cannot throw over the two postulates which have been based on the approximation, since these have become indissolubly associated with the interpretation of measurement. I may again remind the reader that we are not determining the number of particles in the universe, but the number which a physicist who accepts the elementary principles of measurement will (if he is consistent) *think* there are—a number which he will consequently deduce from his measurements interpreted in the way in which he *thinks* they should be interpreted.

10. We distinguish between *casual measures* and *standard measures*. In a standard measure one observable, called the object-observable, is arbitrary, but the other is a special observable called the standard. Since the measures employed in systematic investigation are almost always standard measures, it is necessary to give a precise definition of the ultimate standard adopted. This is treated in *D*. §§ 2–6. It commonly happens that concepts, which in classical physics are assumed to be exact, are found in wave mechanics to have an irreducible uncertainty; and the standard is no exception.

The physical universe does not provide an exact standard of length, mass, etc. But it is shown in *D.* § 2 that a universal standard can be defined whose variance is only $1/\sqrt{N}$, which is of order 1 part in 10^{39}. There are three ways of dealing with the small fluctuation of scale which results:

(1) In molar physics it is neglected entirely.

(2) In microscopic physics, as developed in *D.*, it is taken into account statistically.

(3) In the present fundamental structure theory it is taken into account rigorously by using a non-numerical symbol to represent the scale of a standard measure.

The approximation (1) corresponds to the assumption that the symbolic scale reduces to an eigenscale; and the approximation (2) corresponds to the assumption that it can be replaced by a probability distribution over eigenscales.

The approximation (1) is responsible for the practice of treating a standard measure as though it were a characteristic of the object-observable only; for it is only if the standard is exact and universal that we can afford to let it drop out of sign in this way. If length, or mass, belonged only to the object-observable it would be an observe, not a measure. Thus the approximation leads to a *displaced representation* in which standard measures are represented as observes; so that the observes have now the structure of measures and are symbols with 16 eigenvalues. Thus, having obtained a measure 0·025 (a pure number) which belongs conjointly to a pin and a standard rod, we displace it into an observe 0·025 m. (not a pure number) which belongs only to the pin. To a certain extent we guard against confusion by recognising that the observe or displaced measure is a dimensional quantity or tensor, whereas the true measure is a purely numerical tensor.

The fact that all *dimensional* quantities in physics arise from the displacement of measures shows how deeply the system of physics has become committed to displaced representation. Even when we find that it rests on an inadequate approximation, we cannot undo the displacement; we must take account of the failure of the approximation in some other way. The observe must still be identified with an exact-scale measure, i.e. a measure made with an exact and universal standard; but it is now only an intermediary concept, since the ideal exact standard does not exist. In accordance with the principle of measurement in § 9, the standard measure, i.e. the measure made with the existent inexact standard, is interpreted as the ratio of two observes, namely, the exact-scale measure of the object-observable and the exact-scale measure of the standard.

Thus for practical application we must adapt the structure theory that has been developed, so that *standard measures* and *exact-scale measures* take the place of *measures* and *observes*.

The measure concerned in the structure problem (the *EF*-number) is an energy tensor, of which the density is a typical component; so we consider more particularly the measurement of density. In displaced representation the exact-scale density becomes an object-observe—agreeing with our usual conception of density as a dimensional quantity residing wholly in the object-system. Allowing for inexactness of the standard, the standard density, or density-measure, is the ratio ρ/ρ' of two density-observes, ρ being the exact-scale density of the object-system and ρ' the exact-scale density of the existent standard of density. Changes of the measure ρ/ρ' may be true changes $\delta\rho$ of the object-system, or scale fluctuations $\delta(1/\rho')$ due to inexactitude of the standard. These changes cannot be observationally separated, because the exact scale

necessary to measure ρ and ρ' does not exist. The observes are conceptual attributes of the object-system and standard, inconvertible into measures except when the structure furnishes a scale that reduces to an eigenvalue; but their ratio is a measure.

Extending this to the energy tensor as a whole, the tensor measure is the ratio of two tensor observes $T_{\mu\nu}$, $T'_{\mu\nu}$; that is to say, it is the fourth-rank tensor $U^{\sigma\tau}_{\mu\nu}$ defined by

$$T_{\mu\nu} = U^{\sigma\tau}_{\mu\nu} T'_{\sigma\tau}. \tag{18}$$

The standard measure is accordingly a fourth-rank tensor or $EFGH$-number; and the frame to be used in our final discussion is the $EFGH$-frame with grid number 2^{256}.[a]

The corresponding measurable is an $EFGH$-number and has therefore 256 eigenvalues. In comparing this with our earlier conclusions, it must be remembered that we are now working to what would ordinarily be considered a fantastically high approximation. For ordinary purposes a measurable has 16 eigenvalues; but under the supermagnification here employed (which is not content to overlook 1 part in 10^{39}) there is a fine structure which splits each of them into 16 components.

In conjunction with § 8, this result shows that the excess particle or particles forming a microscopic object-system are added to a standard occupation of $\frac{1}{2}n = \frac{1}{2}.2^{256}$ particles. These belong to the standard environment or uranoid. We have now broken the back of the problem, since $\frac{1}{2}n$ is roughly of the order of magnitude of the number of particles in the universe. We must expect N to be greater than $\frac{1}{2}n$, because $\frac{1}{2}n$ refers to the occupation of a single measure. We can avoid treating more than one measure by considering a universe in which all the particles are at rest and accordingly have the same energy tensor; but in that case we shall have to introduce a factor (multiplicity factor) to compensate for the constraint which is thereby imposed. The evaluation of these final factors follows the same general theory (*D.* part II) that is used in my determinations of the other natural constants.

11. Consider a uniform distribution of particles at rest in an arbitrary coordinate system. Let σ be the particle density (number of particles per unit volume) and ρ the mass density or energy density. By general relativity theory,

$$8\pi\kappa\rho = 8\pi\kappa T = G,$$

where κ (the constant of gravitation) is a fixed natural constant, and the velocity of light is 1. To represent a fluctuation of the standard of length, we put

$$ds^2 = g_{\mu\nu}dx_\mu dx_\nu = \lambda^{-1}g'_{\mu\nu}dx_\mu dx_\nu,$$

where the $g'_{\mu\nu}$ are fixed numerical functions of the coordinates, and λ has a small fluctuation about the value 1. Then $g^{\mu\nu} = \lambda g'^{\mu\nu}$. The Christoffel brackets $\{\mu\nu, \sigma\}$ are independent of λ, so that $G_{\mu\nu}$ is independent of λ. Hence G, and therefore ρ, varies as λ. A volume varies as ds^3 or $\lambda^{-\frac{3}{2}}$; so that σ varies as $\lambda^{\frac{3}{2}}$, and therefore as $\rho^{\frac{3}{2}}$. Hence

$$\frac{d\rho}{\rho_0} = \frac{2}{3}\frac{d\sigma}{\sigma_0}. \tag{19}$$

It should be noted that λ, and consequently $d\sigma$, $d\rho$ are symbolic,[b] the fluctuation being expressed by the fact that they are not replaceable by definite numbers. If λ reduces

[a] The analysis in §§ 4–8 was not limited to standard (as opposed to casual) measures; but it is evident that the useful physical application is to standard measures. For there would be no interest in enumerating the total number of casual measurables occupying a given measure.

[b] We pass straight from the treatment (1) to the treatment (3) in § 10, the intermediate approximation (2) being foreign to our present methods. Thus we do not contemplate a probability distribution over eigenscales, which would be the common way of describing a fluctuation.

to an eigenvalue, the fluctuation disappears; and we have a fixed eigenscale which, however, will generally differ from the scale $\lambda = 1$ chosen as norm.

The change of particle density constitutes an apparent change of occupation. Since there is no absolute scale, it is impossible to distinguish between 'real' changes of occupation due to excess particles and changes due to scale fluctuation. We can therefore apply to scale fluctuation our previous results for excess occupation by putting

$$d\sigma/\sigma_0 = K/\tfrac{1}{2}n, \tag{20}$$

K being the (apparent) excess occupation above the standard occupation $\frac{1}{2}n$.

We now transform to the usual displaced representation. This assumes that the scale is an eigenscale, so that $d\sigma$ and K reduce to numbers. Then ρ is interpreted as a density-observe. If the particles are such that their densities (density-observes) are simply additive, their number is proportional to the total density ρ; and the proportionate change of occupation is

$$d\rho/\rho_0 = K/\tfrac{3}{4}n, \tag{21}$$

by (19) and (20). Thus, taking $K = 1$, the addition of one particle increases the number present in the ratio $1 + (\frac{3}{4}n)^{-1}$; in other words, the initial occupation of the density-observe is $\frac{3}{4}n$.

This change of the initial occupation from $\frac{1}{2}n$ to $\frac{3}{4}n$ is the result of transforming from particles defined as the occupants of density-measures to particles defined as the occupants of density-observes; or equivalently, from the occupants of standard measures (which in the present high approximation are fourth-rank tensors) to the occupants of exact-scale measures (second-rank tensors). The physical reason for such a modification is easily seen. Suppose that we have a distribution of j_0 particles together with the possibility of an additional particle whose presence or absence is left unsettled. By relativity theory the extra particle will, if it is present, change the metric and hence the masses and densities of the j_0 particles. In rigorous theory one unsettled element in the universe unsettles everything; nothing can be settled until everything is settled. The current approximation, taking advantage of the fact that *in the actual universe* the final scale is unlikely to deviate more than 1 part in 10^{39} from a certain exact scale, settles the scale first, and considers changes of energy tensor referred to this fixed scale. Then the unsettled point is whether a certain net addition to the energy tensor is present or not; and a new type of particle—a net, or quantum, particle—is introduced as carrier of this addition. The net particle is the occupant of a density-observe. It has a different nature from the relativity particle, since it includes in its own energy tensor those changes of energy tensor which relativity theory would assign to the j_0 particles whose metric it affects. The $\frac{1}{2}n$ relativity particles have to be replaced by $\frac{3}{4}n$ net particles to provide the same initial energy tensor, assuming that the energy tensors are additive.

But the energy tensors are not simply additive. If we add a large number of particles so as to increase σ and ρ substantially, λ is altered; and the scale, even if it is an eigenscale, is not the normal scale corresponding to $\lambda = 1$. If we repeat the procedure in the last paragraph with $j_0 + y$ particles certainly present and one unsettled particle, the energy tensor of the net particle will differ from that of the net particle added to the j_0 distribution. The general formula connecting j and ρ is therefore non-linear; though it is, of course, sufficiently nearly linear in the very small range of scale comprised in the scale fluctuation of the actual universe. The determination of the law of variation

of ρ with j is a fundamental (though elementary) investigation in relativistic quantum theory; and the result is used in many other investigations besides the calculation of N. It is treated in *D.* §§ 9, 10. The result is that, when (as here) all the particles are in a state of almost exact rest, or more generally when they are all in one pseudo-discrete state, ρ varies as $j^{-1/k}$, where k is the number of dimensions of the phase space, i.e. the number of independent components of the measure capable of continuous variation. In the present problem, k is the number of independent components of the complete energy tensor, namely 136. Thus $d\rho/\rho_0 = -dj/kj_0 = -dj/j_0'$, where

$$j_0' = kj_0. \tag{22}$$

We can therefore compensate for the non-linearity by changing the initial occupation from j_0 to j_0', i.e. from $\frac{3}{4}n$ to $\frac{3}{4}nk$.[a]

The argument requires care, otherwise we are liable to insert the factor k the wrong way up. If X is the density whose initial occupation is j_0, both referring to an exact scale within the limits of fluctuation of the scale of the actual universe, we have $d\rho = Xdj$, the linear law being valid for the addition of a very few particles dj. It follows that $\rho_0 = Xj_0' = Xkj_0$. Thus either we must assign the measure Xk to the densities of the j_0 initial particles (which conflicts with the understanding that we are discussing the total occupation of a measure X); or we must take the j_0 initial particles to be measurables $e^{ik\theta}$ of weight k, the dj excess particles being of weight 1. We cannot admit any dissimilarity between the excess particles and initial particles, i.e. between the object-system and the rest of the universe. The $\frac{3}{4}n$ particles of weight k are therefore to be replaced by their equivalent in particles of weight 1, namely, $\frac{3}{4}nk$.

The essential point is that the interpretation of measurement accepted in physics saddles us with a displaced representation, based on an approximation which, if it were exact, would involve a linear relation of ρ and j. It is therefore better to compensate for non-linearity by a factor k before applying displaced representation, rather than attempt to represent non-linearity directly (by weighting the measurables) in displaced representation. The factor k can alternatively be regarded as the compensation for treating the tensor measure with k components as though it were a single scalar measure (density). It is shown in *D.* § 10 that any continuous distribution of the energy tensor over k degrees of freedom constitutes a multiple state equivalent to k discrete states.

Our final result is that the standard occupation, to which an excess particle is added, consists of $\frac{3}{4}nk$ similar particles. Each particle, considered in turn as an excess particle, is inserted in an environment of $\frac{3}{4}nk$ particles; in other words, the universe consists of $\frac{3}{4}nk$ of these particles or their equivalent. The particles here enumerated are occupants of, or carriers of, a basal measure (displaced into an observe) which is an unrestricted complete energy tensor. These are the 'standard carriers' whose properties are developed in the author's relativistic quantum theory. It is shown in *D.* §§ 11, 12, that the standard carriers are hydrogen atoms. (The hydrogen atom appears first in the theory, and is analysed subsequently into a proton and electron.) The total number of protons and electrons in the universe is therefore

$$N = \tfrac{3}{2}nk = \tfrac{3}{2} \cdot 2^{256} \cdot 136. \tag{23}$$

[a] The negative sign attached to dj/j_0', which is important in some developments, does not affect the present calculation. It involves a change of K into $-K$, and therefore a merely nominal change of the measurable from $e^{ik\theta}$ to $e^{-ik\theta}$.

EDITORIAL NOTE

In view of the references to the Dublin lectures in the foregoing Appendix, the following table of correspondences between the Dublin lectures and the present volume may be found useful.

D.	F.T.	D.	F.T.
§ 1	§ 1	§ 15	§ 22
§ 2	§ 3	§ 16	§ 25
§ 3	§ 4	§ 17	§ 28
§ 4	§ 5	§ 18	§§ 27, 30
§ 5	§ 7	§ 19	§ 33
§ 6	§§ 8, 9	§ 20	§ 32
§ 7	§ 11	§ 21	§ 37
§ 8	§ 10	§ 22	§ 39
§ 9	§ 13	§ 23	§ 40
§ 10	§ 15	§ 24	§ 51
§ 11	§ 17	§ 25	§ 41
§ 12	§ 18	§ 26	§ 42
§ 13	§ 21	§ 27	§ 47
§ 14	§ 19	§ 28	§§ 49, 50

INDEX

(The numbers refer to pages.)

A, the 'observational' system, 54, 57
absolute and relative chirality, 161
action, units of, 38, 216; integral, 238; invariants, 197, 239
active and dormant symbols, 161; inversion of ideas of, 190
actual particles, 31
additive composition of symbols, 43
aether, and matter, 172; energy tensor of, 176; strain system of, 183; stress in, 172
algebraic, 107
almost exact state, 18
analogue of time in quantum mechanics, 67
analogy, designation by, 153; in quantum mechanics, 59
analysis, linear and quadratic, 196
anchoring of frame, 113, 138
anchors, primary and secondary, 62
angular momentum, 66; as coefficient of rotational elasticity, 216; co-spin, 205
antichiral double frame, 161
antiparticle, 176
antiperpendicular planes and components, 113, 272; rotations, 114
antisymmetrical, wave functions, 184; suffixes, 150
antisymmetry, spin and functional, 185, 192
anti-tetrads, 108
antithetic, 123
association of strain tensors, and space tensors, 150; and energy tensors, 171
atom, balanced, 199; complex, formation of, 212
average value of a characteristic, 147
averaging factor, 131

B (system of reckoning), 54, 57; the transformation $A \rightarrow B$ separates the electrical from the mechanical world, 64
B' (system of reckoning), 61
balanced atom, 199
basal multiplicity, 37
Bernoulli fluctuation, 4
β, Bond factor, 38
bi-particle, 33, 194
Bond factor, 38
bound particle, 54

carrier, standard, 30, 276
casual measures, 279
catalogue of symbolic coefficients, 138
CD-frame, 166
$C'D'$-frame, 167, 183
characteristic equation, of an E-number, 130; of the symbols K, 270
characteristic, stabilized, 20
characteristics, additive, represented by exponents, 268
chirality, 111; in molar physics, 126; of a double frame, 161
circular and hyperbolic rotations, 123
circulation, extra-spatial, 50; of probability within a state, steady, 73; transition, in interstate, 254
classical outlook and quantum outlook, 176
classical particles as V_{-1}, 41
closed algebra, 108
cogredient linkage, 168
coherent phase, 200, 278
collinear (Compton) scattering, 258
combined system characterised by outer product, 190
commutation rules for E-numbers, 107
comparison, fluid, 174; hole, 43; particle, 41, 44
complementary field, 22, 23
complete, energy tensor, 30; momentum vector, 30, 112, 150; strain vector, 150
Compton scattering, 258
confusion between τ (measure of under-observation) and t (time), 71
congruent spaces, 127
conjugate triads, 108
connectivity of measures, 269
conservation of mass for a fluid, 218
conserved energy tensor, 179
constants, the microscopic, 66; molar and nuclear, 105
constraint, affecting energy, 69
continuity and isostasy, conditions of, 221
contracted Riemann-Christoffel tensor, 178
contragredient linkage, 168
contravariant space, strain, and wave vectors, 147, 148
control, molar and spectroscopic, 60; plane for quantisation of co-spin momentum, 205
conversion constants, 64
Corben, H. C., 76
correlation, of coordinates, 18; wave functions, 19
correspondence, of strain tensors and energy tensors, 171; of strain tensors and space tensors, 150
corresponding images, 110
cosmic rays, 214
cosmical constant, 78; identified with B_{05}^{05}, 178, 179
cosmical number N, 4, 105, 265
cosmical standard, 179, 180
co-spin, 164; intracules, 189; angular momentum, 205, 206; binding, 208; electrons in nucleus not bound to protons individually, 202; magnetic moment of co-spin intracule is zero, 250
Coulomb energy, 50, 66; between a molar and a microscopic charge, 64; makes allowance in the wave equation for the effect of interchange circulation on the equation of continuity of flow of probability, 222, 230; term derived, 229
coupling of extracule and intracule, 249
covariant space, strain, and wave vectors, 147–9
cross-dual, 150, 166–9; of the classical energy tensor T_{00} is the quantum or phase energy tensor Z_0^0, 181, 183
crossed frame (C-frame), 255
cross-section for scattering of radiation by an electron, 260
current, magnetic moment of hydrogen extracule, 249; masses of the intracule and extracule, 58; masses of the proton and electron, 34, 58, 62

curvature of space, 5; is a form in which molar theory represents the mechanical part of exclusion, 82

dagger defined, 143
darg defined, 217
Debye-Hückel energy, 23, 99, 235
decay, problems of, 72
deflection determination of $e/m_e c$, 64
degeneracy, is source of complication, 236; pressure, 90; symmetrical, 50, 228
delta-function in the energy of two like charges, 9, 101
designation by analogy, 153
de Sitter universe, 175
density, molar, is scale-free, 18, 180; of top and mean particles, 29; measurement of, 280, 281; of mass, 138; of interstate connected with the product of the densities of the two states concerned, 242; of unit occupations, 17; is a momentum vector, 15; term, condition for, 175
de-stabilisation, 43, 47
determinant of an E-number, 129
determination, (deflection) of $e/m_e c$, 64; (direct) of h/e, 64
deuterium, 63; as combination of double intracule and double extracule, 191; mass of, 197
deuteron, 136
dilution factor, 260
dimension index, 14; of physical quantities, 139
dipole transition probability, 258
Dirac's interchange operator, 164
Dirac's wave equation, 146, 223; its invariance involves the phase coordinate, 72
direction of an extended vector, 130
directly linked frames, 160
discrete wave functions, 19
dispersing system, conversion into a steady system, 259
displaced representation, 280
distinct particles, conception of, 273
distribution, functions, 73; wave functions, 18
divergence of integrals in current wave mechanics, 23, 76, 262
dormancy and activity, inversion of ideas of, 190
dormant, characteristics, 30; components, 112; symbols, 161
double, frame, 159, 276; transformation, 230; intracule, 191; matrix, 164; reckoning, 69; wave function, 34, 52, 54; wave tensors, 168; wave vectors, 167
doublets, isobaric, 202, 206
dual and cross-dual relations, 167, 176
duals, 164

e', charge of electron defined by molar control, 60; value of, 66
e^2 is a particle unit of action, 38
E-frame brought into physics, 113
E-number, 106; as a degenerate EF-number, 161; matrix representation of, 141; representation by quaternions, 121; representation by vectors in 4-space, 113
EF-frame, 159; its group of rotations, 271
EF-numbers are space tensors of the second rank, 272
effective relativity transformations, 117
eigen-energies of states are classically eigen-scales, 157
eigenscale, 222; condition that $\partial/\partial x_4$ reduces to an eigenvalue, 227
eigenstates of an intracule, transitions between, 254
eigensymbol, of yes-no existence symbol, 267; of a symbol, 130
eigenvalue, of angular momentum in ξ-space, 222; of the field energy $-i\hbar\,\partial/\partial x_4$, 221
eight simple particles are necessary for balancing, 200
Einstein-Bose, particles defined, 277; occupation factor Y, 278
Einstein universe, 5, 13; mass of, 10
electric, force between elements of the same elementary charge does not exist, 252; moment, 246; potential, retarded, 252
electrical, characteristics, 30; and mechanical phenomena, separation between, 185; (interchange) degree of freedom, 56; part of a momentum vector, 138; part P_e is antithetic to the mechanical part P_m, 127; theory, 19; charge, symbolic coefficient for, 138
electromagnetic, characteristics, 111; field, irreducible, created by non-integrable gauge transformation, 243; field, molar, 236; potentials, 234
electron, in electromagnetic field, wave equation of, 234; magnetic moment of, 245, 247; moving, radiation by, 252; multiplicity of, 245; origin of its mass, 84; perturbed, hamiltonian of, 233
electrostatic potential, symbol for, 138
energy, of a quantum particle, 28; of a wave packet, 69; of interaction of object-particles with one another is part of the particle energy, but interaction with the environment is field energy, 40; non-Coulombian, 9, 23, 101; quantum, is a classical scale, 222; stabilised in micro space, 154; zero level of, 53
energy tensor, 15; conserved, 179; corresponding to unit occupation, 45; extended, of 256 components, 160; relation to phase tensor Z_0^0, 171; of a relativity particle, 28; of the aether, 176; (strain tensor) of the planoid, 176; symmetry and antisymmetry in, 171
entity, most primitive, 266
environment, 13; and the frame, 127
equivalence, the principle of, 244, 246
equivalent, double frames, 159; sets of E-numbers, 109
exact, properties are unobservable, 44; scale measures take the place of observes, 280; state, almost, 18
excess, measurables, 276; of occupation, 270
excludees and excludors, 204
exclusion, energy, equivalent to gravitational, 82, 97; holds between symmetrical as well as between antisymmetrical wave functions, 195; lateral, 204; principle, 29, 81; treatment and spin treatment, connection between, 206
existence represented by J, 267
existence symbol, 267
expansion energy, 69, 73
expectation value, 147, 218
expected particle energy, 28
extended, energy tensor, 30, 160; energy tensor, analysis of, 194; momentum vector, 30, 112

external, particle, 32, 43; comparison particle is a mean particle, 44; and internal particles, transformation of equal masses into, 231; wave functions, 19
extinguishing of electrical energy by balancing of intracules, 200
extracule, 19, 51, 247; and intracule, treatment as independent particles, 250
extraneous object-field, 22
extraordinary fluctuation, 5
extra-spatial circulation, 50, 66, 73, 218, 221

factorisable E-numbers and matrices, 144
factorisability is condition for purity, 145
factors, determination of, 145
Faraday constant for hydrogen, 60, 65
Fermi-Dirac occupation factor, 278
Fermi-Thomas treatment of satellite electrons, 207
fictitious particles, 31
field, 23; and particle angular momentum are called orbital and spin momentum in current theory, 206; angular momentum, 216, 217; energy, 23, 25, 27, 69; gravitational and inertial, 14; linear momentum operator, 216; momentum vector, 147; momentum vector is a multiple of the particle momentum vector, 219; must be stationary for small changes of the occupation factors of the eigenstates, 24; theory, 22, 196
fifth dimension, representation of the phase coordinate, 46
final vector ψ_f, 145
fine-structure constant, 37, 66, 67, 229; why it is 137 rather than 136, 216
five-dimensional theory of physical variates, 114
fixed-scale, unit, 139; provided by quantisation transferred to classical description by interstate tensors, 256
flexibility of occupation, 24
flow of probability, equation of continuity of, 218
fluctuation, Bernoulli, 4; ordinary and extraordinary, 5
forbidden oscillator, 212; is a mesotron, 213
force-constant (ratio of electrical to gravitational force between a proton and electron) calculated, 105
forces, range of nuclear, 9
formal relativity rotations, 123
four dimensions, passage to five dimensions, 128
fourth-rank wave tensors, 149
frame, 109, 127; crossed (C-frame), 255; double and quadruple, 276; singular, 100; anchoring in molar physics, 138; is the operational form of the momentum vector, 147; of opposite system, 111
free, information, 19; intracule, 53; double intracule, 191; particle, 54
full equivalence of frames, 121
fully observed system, 71, 73
functional antisymmetry, 185, 192
future, the, 135

galaxies, recession of, 10
Galilean time, 67
gauge, identified with imaginary electromagnetic potential, 237; invariant vector is covariant momentum vector for state and contravariant momentum vector for interstate, 256; invariant tensors, 238; transformations, 236, 241; transformation, uniform, is a scale transformation, 244
generation of E-numbers by a tetrad, 109
generic energy tensor, 28, 33
geometrical angle introduced in development, 271
grad defined, 217
graduation, 278
gravitational, and inertial energy distinguished, 97; energy, negative, 97; field, 14; created by non-integrable coordinate transformation, 243; field replaced by exclusion field, 29, 86; the constant calculated, 104
grid, 275–8
ground state of hydrogen, 187
group, of rotations with 16 eigenvalues is the group of rotations of an EF-frame, 271; structure of E-numbers, 107
gyromagnetic experiment, 249

$\hbar$, field unit of action, 38, 216
h, Planck constant, calculated value of, 66
Hamiltonian differentiation, 27
hamiltonians, for standing and progressive waves, 87, 91; of an electron perturbed by other electrons, 233
harmonics, surface, of hypersphere, 85
helium atom, mass of, 199
hole, comparison, 43
homochiral double frame, 161
homothetic, 123
hydrocule, defined, 54; is a V_{137} and is a hydrogen atom, 56; is the V_{136} of system B, 57
hydrogen, all matter is composed of, 64; atom is a combination of a spin intracule and an extracule, 189; magnetic moment of, 247; mass of, 66; separation into two carriers, 33; Sommerfeld's formula for the states of, 225; intracule, wave equation of, 222; metastable states of, 50, 186; ground state of, 187; unstable states of, 188; solution of the wave equation of, 224; universe composed of, 277
hyperbolic and circular rotations, 123
hypersphere of space, surface harmonics of, 85; volume of, 47

i factor in symbols, 139
idempotency, 131; imaginary, 134
identities, the tensor, 177
image of an idempotent quantity, 132
images, 110
imaginary, and real E-symbols, 120; idempotency, 134; in quantum theory, its source, 40
incoherent phases of wave functions, 200, 278
indices of wave tensors, 242
indistinguishability of particles best understood if we think of them as carriers, 49, 51
individuality of a particle, 131, 278
ineffective rotations, 116, 159
inertial, field, 14; and gravitational energy distinguished, 97
infinite, negative charge of Dirac's theory, 84; temperature uranoid, 77; transverse energy of an electron fallacious, 76
information, free, 19
in-invariants, 240
initial, energy, elimination of, 94; particles, 32; state, 27; vector ψ^*, 145

integrals invariant for both coordinate transformations and gauge transformations, 238
interchange, 49; and exclusion representations of rest mass, relation between, 263; applicable to proton-electron system, 50; circulation, extra-spatial plane of, 66; energy, 50, 56; energy as energy of transition circulation, 262; energy appears as perturbing energy in the hamiltonian of an electron, 232; momentum, 227; operator, 162; operator, Dirac's, 164; representation, connection with gravitational representation of energy, 263; variates are the scale and phase of the relative space, 152
intermediate relativity, 12, 236
internal, particle, 32, 34; wave functions, 19
interstate, and state energy tensors, relation $T^{\times}{}_{00}=Z_0{}^0$ between, 241; circulation, described by an energy tensor, 182; conjugate, 254; defined as a steady circulation of probability between two states, 182; density of, connected with the product of the densities of the two states concerned in it, 242; energy, 89; wave function of, 259; (continuous transition circulation), 67
intracule, 19, 51; and extracule, treatment as independent particles, 250; double, is the combination of a spin intracule and a co-spin intracule, 191, 193; essential part is the quantum angular momentum (classical magnetic moment), 188; its quantum momentum vector is essentially a strain vector, 155; magnetic moment of, 247; multiplicity of, 191, 194; wave equation of, 245; transitions between eigenstates of, 255; two-legged, carrying transition energy, 201; n-legged, 185; spin and co-spin, 189
inversion of energy, 29, 39
isobaric, atoms, 207; doublets, separation constant of, 205
isostasy and continuity, conditions of, 221
isostatic compensation, 219; as regards phase and time derivatives, 256; provided by radiation field, 259
isotopic spin, 136; is the controlled component of co-spin angular momentum, 206

J, symbol of existence, 267

keenness of probing, specification of, 70
Klein-Nishina formula for scattering, 261

lateral exclusion, 86, 203
left-handed frames, 111
length, standard of, 7, 15
light velocity, is constant, 8; its value, 65; in rigid coordinates is $137c$, 259
like charges, energy of two, 99, 101
linear (super-threshold) and quadratic (sub-threshold) analysis, 196
linkage, cogredient and contragredient, 168
lists of constants, microscopic, 66; molar and nuclear, 105
local, energy, 68; uncertainty of a physical reference frame, 6
longitudinal field, 234
Lorentz, frames, equivalence of, is included in equivalence of E-frames, 110; transformation is contrary to the spirit of wave mechanics, 73; transformations ineffective in statistical physics, 117

m_e, mass of electron, 34, 66
m_0, mass of standard carrier, 33, 42, 66; absolute determination of, 81; determination by exclusion theory, 85
m_p, mass of proton, 34, 66
M, mass of hydrogen atom, 33, 51; absolute determination of, 81
magnetic, energy, 138, 205; moment, 135, 138, 245; momenta are components of the extended momentum vector, 179; moment of the hydrogen atom, 247; moment of the neutron, 251
magnetisation, permanent, 246
magneton, nuclear, 249
magnitude of an extended vector, 130
mass, and momentum vector are scale-fixed characteristics, 18; as a mutual property, 261; constant of a particle, 53; defects of nuclei of mass number from 30 to 150, 210; differences of, are due to multiplicity factors improperly absorbed, 175; of scale-free particles transforming as densities, 119; relativity definition of, 68; of deuterium atom, 197; of Einstein universe, 10; of electron, 34, 66; of extracule is molarly controlled, 60; of intracule is spectroscopically controlled, 60; of helium atom, 199; of mesotron, 211; of proton, 34, 66; of standard particle, 81, 85; origin of, 77, 83, 95
matrix, double, 164; representation of E-numbers, 141
mean particle, energy tensor of, 28, 37
measurable, defined, 48, 266; in EF-frame are wave systems in spaces with antiperpendicular axes, 275; primitive, symbolic representation of, 268; set of alternatives of, when the measure is given, is a grid, 273, 275, 276; identified as standard particles, 273; tensor, 268; true, a combination of the measurable and Z, 269
measure, is the ratio of two observes, 279; tensor, 268; X_q defined, 272; casual and standard, 279
measures, connectivity of, 269; exact-scale, take the place of observes, 280; pure, 274
mechanical, characteristics, 111; components of a vector, physical interpretation, 124; momentum vector, 137; theory, 19
meson-field hypothesis, 10, 211
mesotron (cosmic ray), nature and mass of, 213; heavy, 214
metastable states of hydrogen, 50, 186
metrical field Z, 279
micro, point of view, that the intracule is a compact particle, 155; space, 154
microscopic constants, list of, 66
modularly equivalent systems, 190
molar, and nuclear constants, list of, 105; characteristics of an assemblage depend on the mean particle, 29; definition of mass, 35; electromagnetic field, 236; point of view that the wave is the intracule, 155; relativity as field theory, 38; space, 154
molarly controlled quantities, 59
momenta, as coefficients of elasticity, 216; are relations between entities, 217; directions of, indicated by symbolic coefficients, 67; to be inserted in wave equation, 67
momentum, and its conjugate coordinate are homothetic, 138; and position vectors as E-numbers, 124; and position vectors as vector-densities in a sub-space, 116; density, 138; in wave mechanics

is i times the corresponding classical momentum, 33, 37; operator defined, 216; strain vector is a hermitic matrix, 143; term, condition for, 175; vector, 15, 36, 106; vector, field and particle, 147
monochiral uranoid, 126
monothetic, 123
motion, conception of, 119
μ, mass of intracule, 33, 42, 51, 66
multiple occupation indicated by wave-length and not by coexistence of phases, 277
multiplets, isobaric, 206
multiplication, different for CD-symbols and EF-symbols, 166; table of E-numbers, 107
multiplicative and additive composition of systems, 43, 136
multiplicity factor, 26; decides the partition of the energy tensor, 29; decided by the energy tensor or momentum vector alone, 39; as selection factor, 29, 195; relative to the adopted system of classificatory characteristics, 98
multiplicity, of the intracule, 191; of comparison particle, 42
mutual, energy carrier, separation into two carriers of self energy, 42; properties and self properties, 41

N, cosmical number, 10, 105, 265, 283
natural units, 14
nebular recession, 10, 105
negative energy levels, 83
negatron, origin of, 84
neutral space-time, 125
neutrino, emitted when a co-spin is changed into a spin intracule, 190; absorption of, 193
neutron, as indivisible particle, 250; is a combination of a co-spin intracule and an extracule, 189; magnetic moment of, 251; mass of, 191
n-legged intracule, 185
non-Coulombian energy, 9, 23, 101; is an adjustment which allows us to simplify a 4-particle system into a 2-particle system, 102
non-singular E-numbers, 109
non-trivial idempotent vector, 132
normal state, defined, 70
normalisation of wave functions, 218
normalised vector, 130
nuclear, and molar constants, list of, 105; forces, range of, 9; magneton, 249; planoid, 208; theory, starting-point of, 191; theory, list of contributions of fundamental theory to, 211; transformations, energy loss in, 80
nucleus, constitution of, 202; co-spins in the, 206; radius of, 207
number of particles in the universe, 10, 105, 265, 283

object-field, 22
object-particle is always a top particle in quantum transitions, 29
object-system, 13; perfect, 44
observable, 1; definition of, 266
observe, 279
occupation, excess of, 270; multiple, indicated by wave-length, 277
occupation factor, 17; the only function of the time coordinate, 73; continuous, 26; Einstein-Bose, 277; Fermi-Dirac, 278
occupation operand, 270; qualitative and quantitative, 277; symbol, 270, 277
operational form of a characteristic, 147
operator, expectation value of, 218
opposite systems of frames, 111
ordered volume, 118; symbol for, 138
ordinary fluctuation, 5
orientation phase space, 151, 170
origin, geometrical and physical, 1, 3
orthogonal projection, 11
oscillator, 3-dimensional, 258; forbidden, 212; forbidden, is a mesotron, 213; occupying interstate between two states, 182, 254; simple pure, defined, 182
outer product of two vectors, 160; is the straight product of their duals, 166

parallel displacement of vectors, 236
partial existence, 267
particle, actual and fictitious, 31; angular momentum, 216; and field angular momentum are called spin and orbital momentum in current theory, 206; classical, is insertion in the aether, 176; defined, 30, 274, 277, 282; density is a momentum vector, 15; Einstein-Bose, 277; energy and total energy, 25, 37, 39, 69; energy, expected, 28; energy tensor, 23, 27; individuality of, 278; momentum in scale-fixed physics, 38; momentum vector, 147; momentum vector is solenoidal, 218; occupying a pseudo-discrete wave function is an unidentified member of a large assemblage, 17; of quantal physics is carrier of momentum vector, 18; of scale-free physics is carrier of energy tensor, 18; pure, 137; quantum, 24; relativity, 24; scalar, 136; simple (spinless), 20; theory begins with the standard carrier, 272; transformation of two of equal masses into external and internal particles, 231; transverse self energy of, 261; which carries no coordinates, 39; which has comparison particle has electrical characteristics, 43
pentad, 108; part of idempotent vector, 133
perfect object-system, 44, 179
permutation coordinate, 49, 66; introduces interchange energy, 56
permuting suffixes of wave vector factors, effect of, 150
perpendicular components, 113
persistent components, 158
phase, defined, 275; and scale, 46, 72; is the time analogue in the analogy between quantum and classical mechanics, 72, 154, 227; space, 20, 26, 150, 169; sub-coordinates, 275; tensor, is a strain tensor, 170; tensor, is the quantum energy tensor, 184; tensor, is the cross-dual of the classical energy tensor, 171, 181
physical interpretation of the 10 mechanical components of a vector, 124
physics, nature of, 266–9
piles, 164
plane, of interchange circulation extra-spatial, 66; of relativistic rotation, 50
planoid, defined, 93; energy tensor (strain tensor) of, 176; the nuclear, 208
position and momentum vectors as E-numbers, 124
positive and negative charges correspond to P and $P^\dagger$, 111
positron, origin of, 84

potential, electrostatic, symbol for, 138; electromagnetic, 234; retarded electric, 252
predictive theory, 2; and structural problems, 72
pre-history of the E-frame, 120
pressure, 77, 138; degeneracy, 90; term, condition for, 175
probability, distribution of momentum is determined by probability distribution of coordinates, 79; distribution over a set of states, 181; extra-spatial flow of, 50, 66, 73, 218, 221; equation of continuity of flow of, 218; of existence, 267; transition, 254
probing, specification of keenness of, 71
product, outer, of two vectors is the straight product of their duals, 166
progressive and standing waves, 87, 91
projection, orthogonal, 11
proper mass, explanation of, 54, 78, 96
proper phase, analogous to proper time, 227
protons, energy of two, 10, 99, 101
pseudo-discrete states, 17
pseudo-individuality, 131
pseudo-rotations, 114, 170
pure, measures, 274; oscillator, simple, 182; particle, 137; particles, momentum vectors and momentum strain vectors of, are factorisable, 144; particle, its proper mass, spin, and magnetic moment, 146

$\bar{q}$, definition of, 167, 170
quadratic (sub-threshold) and linear (super-threshold) analysis, 196, 240
quadruple frame, 162, 276; probability distribution, 48
quantal theory, 16
quantisation, 19
quantum-classical analogy, 153, 155, 183
quantum, energy is a classical scale, 222; energy tensor is classically a strain tensor and is the cross-dual of the classical energy tensor, 183; jump, 50; momentum vector is classically a strain vector, 183; number of a level, 85; outlook connected with classical outlook, 176; particles, 24; specified standard, 8, 180; theory uses a covariant momentum vector, 255
quarterspur, 107; of an EF-number, 166
quaternion algebra, the algebra of E-numbers is the square of, 121

R, Einstein radius of space, 6, 10, 81, 105
radiation, necessary to maintain steady occupation of two quantum states, 256; of momentum by a moving electron, 252; quantum theory of, 234
range constant of nuclear forces, 9, 102, 105
rank of extended space vectors and tensors, 150; of idempotent E-numbers, 136
ratio of masses of proton and electron, 34, 66; alternative derivation, 42; spectroscopic determination, 62
reality conditions, 120–4; for strain vector, 127; in connection with gauge transformations, 237; relative and absolute, 156
recalcitrant, terms, 184, 256; energy tensor Z_d, 185
recession of the galaxies, 10, 105
recoil, interpretation of the Reimann-Christoffel tensor in terms of, 173; momentum of the physical origin, 75; symmetrical, corresponds to a de Sitter universe, 176
reduction of the 5-spaces of extracules and intracules to 4-spaces, 152
reflection, 275
relative, and absolute chirality, 161; and absolute reality conditions, 156; space, 152; and ordinary space, distinction between, 222
relativity, intermediate, 236; particles, 24; rotations, 50, 110; rotations in molar space-time, 160
representation, displaced, 280; space, 26
resolution of rotations into simple rotations, 114
rest-mass of a particle, 54, 78; represented as energy of interchange with particles of environment, 96
retarded electric potential, 252
Riemann-Christoffel tensor, 173; symbolic form of, 175; extended, is the same as the extended energy tensor, 176, 179; contracted, 178
right-handed frames, $(E_{16}=i)$, 111
rigid, coordinates, 35; field, 23; field, a particle has a momentum which is i times the classical momentum, 40; time, 67
root vector, 31
rotations, 110, 112, 159; ineffective, 159; in the 15 coordinate planes in Euclidean space of 6 dimensions, 107, 115; plane of relativistic, 50; transforming the 136-dimensional phase space into itself, 170; which give full equivalence, 121
Rydberg constant for hydrogen, 8, 58, 65

scalar particle, 136
scale, and phase, 47–8; and phase, as the fourth dimension, 72; as a momentum, 46; fixed physics, 16, 19; free physics, 16, 19, 26; free physics lies between cosmical and quantal physics, 180; indicator, 46; momentum, 47, 275; of a standard measure represented by a non-numerical symbol, 280; true, 26; uncertainty, 15, 44
scattering, collinear (Compton), 258
second order wave equation, 245
secondary anchors, 62
selection factor, 84; for planoid, 97
self, energy, transverse, of a particle, 261; properties and mutual properties, 41
semi-classical particles are V_1 particles, 41
set, spectral, 135
σ, uncertainty constant, 3, 67, 81, 102
σ_ϵ, standard deviation of ϵ, 6; is the standard deviation of the scale momentum, 46
simple pure oscillator, 182
simultaneity, 119; implied in conception of a system, 153
singular, E-numbers, 109; frame, 100
Sirius, companion of, 31
six-dimensional representation of the group-structure of the E-frame, 115
solenoidal property of particle momentum, 218
solution of the wave equation of hydrogen, 224
Sommerfeld's formula for the metastable states of hydrogen, 187; more general formula for the states, 225
space, congruent, 127; is restricted to 3 dimensions, 124; spherical, 6, 10, 98; tensors, 168; vector, 145, 168; vector is a mixed wave tensor of the second rank, 148; vectors, outer product of, 160

special relativity, rotations as rotations of E-frame, 113, 123; transformations of, are included in transformations of E-numbers, 110; time axis, 167
specification, 43
spectral set, 135
spectroscopic determination of m_p/m_e, 63
spectroscopically controlled quantities, 60
spherical space, 6, 10; non-integrability of states in, 98
spin, 164; and co-spin intracules, 189; angular momentum, 66; antisymmetry, 185, 192; factor of I (Dirac's interchange operator), 233; isotopic, 136; isotopic, is the controlled component of co-spin angular momentum, 206; theory, 106; treatment and exclusion treatment, 206
spinless particle, a V_3 or V_4 particle, 20; a fictitiously simplified element of physical structure, 32
spur, 143
spurred and spurless E-numbers, 107
stabilisation, 19, 34
standard carrier or V_{136} particle, 31, 37, 276; currently described as composite, 32; division into two carriers, 34; an unspecialised element of the energy tensor of the uranoid, 45; particle theory begins with, 272
standard, environment, 30; form of a non-trivial idempotent vector, 132; magnetic moment of hydrogen extracule, 248; masses of the proton and electron, 34, 58; measure, 279; measure, is a fourth-rank tensor, 281; of length, 15; particle, identified with hydrogen atom, 55; particle, is the V_{136} of system A, 57; quantum-specified and cosmical, relation of, 180
standing and progressive waves, 87, 91
star cluster, 40
Stark effect, 254
starred vector ψ^*, rule for multiplication, 145
state, circulation, 182; defined as a steady distribution of probability, 181; defined as any combination of distinctions or variates, 277; and interstate energy tensors, relation $T^{\times}{}_{00}=Z_0{}^0$ between, 241; metastable, of hydrogen, 186; unstable, 188
statistical theory, 106
steady system, of particles occupying states and oscillators occupying interstates, 182
Stoner-Anderson formula criticised, 89, 92
straight dual defined, 150
strain system of the aether, 184
strain tensors, 168; and space tensors, 150; derivation from energy tensors by 'correspondence' and by 'association', 171; derived from space tensors by dual and cross-dual relations, 167; in quantum theory and relativity, 172
strain vector, defined, 117, 148; importance of, 119; expressed as an EF-number, 161; reality condition for, 127
strains, 110, 114, 117
stream vector, 218, 244
structural, and predictive problems, 72; concepts, 139; theory, 2
structure, problem, 270; quantum-specified, 8
sub-coordinates, phase, 275; two hundred and fifty-six, 276
sub-threshold theory, 83, 84; is exclusion theory and pre-gravitational, 195
suffixes of wave vector factors, effect of permuting, 150
summation convention, 143
super-dense matter, 87
super-threshold, particles, 84; theory is gravitational theory, 195
symbolic, coefficients, catalogue of, 138; frame, 106; frame, anchoring in molar physics, 138; frame is the operational form of the momentum vector, 147; representation of the most primitive measurable, 268
symmetrical, degeneracy, 50, 228; EF-numbers, 165; recoil corresponds to a de Sitter universe, 176; wave functions correspond to pressure components of T, 184
symmetry and antisymmetry in energy tensor, 171
system, conception of, implies simultaneous constituents, 153; is represented symbolically as the product of its parts, 268
systems A and B of reckoning energy, 54; modularly equivalent, 190

T_{00}, $T_0{}^0$, $T^0{}_0$, T^{00}, notation for covariant, mixed, and contravariant space tensors, 149
$T^{\times}$, cross-dual, defined, 166
T_p (pressure components of T) corresponds to symmetrical wave functions, 184
T_d (density components of T) corresponds to antisymmetrical wave functions, 184
$\tilde{T}$, dual, defined, 164
τ, the measure of under-observation, 71; peculiar to predictive theory, 73
temperature in relativity theory, 77
Temple's symbolic treatment of Dirac's wave equation, 223
tensor, densities in four dimensions, 115; gauge-invariant, 238; identities, 177; measurable and measure, 268
tetrad, 108
third-rank tensors, evasion of, 161
three-dimensional oscillator, 258
three-dimensionality of space, 124
threshold energy, 84
tiers, 164
tilt, of plane of angular momentum with respect to the adopted reference frame, 163; of the radius of curvature of a uranoid, 126; time, 186; of the plane of a motion, 50
time, analogue of, in quantum mechanics, 67; axis, special, 167; -like variable in wave functions is the phase, 73; rigid and Galilean, 67; in wave mechanics, 37; -tilt, 186; three kinds of, 72
top particle, 28; takes part in quantum transitions, 29
transformation, of a double frame from EF to GH, 230; of probability distribution of physical momenta into probability distribution of geometrical momenta, 74; of two particles of equal masses into external and internal particles, 231
transition, between eigenstates of an intracule, 254; circulation between substates provides the Coulomb term, 230; circulation in interstates, 254; energy, 28, 33; energy, whole energy exhibited as, 94; flow corresponds to classical conception of momentum, 182; flow is coupled with a recoil flow, 182; forbidden between states corresponding to symmetrical wave functions and states corresponding to antisymmetrical wave functions, 184; probabilities, 254; probabilities, formula for spontaneous dipole, 258

transpose of a matrix, 164
transverse, or radiation field, 234; self energy of a particle, 261
trivial idempotent quantities, 132
true, measurable, a combination of the measurable and Z, 269; representation, 142; representation of E-numbers by matrices, 121; rotations transforming the 136-dimensional phase space into itself, 170
two-frame transformation, 230
two-legged intracule carrying transition energy, 201
two-particle system, 32; transformation can be applied legitimately only to a hydrocule, 57
two protons, energy of, 10, 99, 101

uncertainty, constant σ, 3; local, of a physical reference frame, 6; scale, 6, 15, 44
under-observation, 71
uniform, distribution (infinite plane waves), 39; gauge transformations, 241; observation, 153
unit, natural, 14; vector, 130
unitary transformations, 151
universe, Einstein, 5, 13; mass of, 10
unobservables, 1
unstable states of hydrogen, 188
unsteady wave functions are unnecessary, 73
uranoid, 13; as a pseudo-discrete assemblage, 30; infinite temperature, 77; monochiral, 126, 128; stresses produced by particle in, 171

V_k, a quantum particle of multiplicity k, 20
V_{k_1} and V_{k_2}, relation between masses of, 29
V_{137} particle carries an energy tensor and a permutation variate, 56; is a hydrocule, 56
V_3 particles, 79, 97
V_4 particles, 80
V_{10} particle or vector carrier, 32
vector, in micro space is strain vector of molar space, 155; physical interpretation of its ten mechanical components, 124; wave function is not the same as a wave vector, 217
velocity, of light is constant, 8; of light in rigid coordinates ($137c$), 259; of a particle in wave mechanics is group velocity, 36; reference system for, provided by matter but not by aether, 172; symbols for, 138
vertical and lateral exclusion, 86
volume, element in three dimensions, reciprocal of, is a 4-vector with symbolic coefficient E_{45}, 118; element in phase space, 152; of uranoid, 47

wave, an extended probability distribution is often described as, 155; and particle aspects reconciled, 155; identities, 140
wave equation, defined, 220; for an electron in a molar electromagnetic field, 234; its function is to determine a field momentum 4-vector isostatically compensating the corresponding part of the particle momentum vector, 234; of hydrogen, solution of, 224; second order, 245
wave function, and wave vector, contrast between, 217; confined to statistical mechanics, 72; correlation and distribution, 18; degrees of freedom of, 39; discrete, external, and internal, 19; of an interstate, 259; normalisation of, 218, 244
wave mechanics, based on concept of rigid field, 23–4; new approach to, 89
wave tensor, calculus, 106; indices of, 242; of the second rank, 146; of the fourth rank, 149; double, 168
wave vector, 72; defined, 145; transformation law, 147; double, 167; and wave function, contrast between, 217; with its underlying isostatic compensation is represented by a wave function, 220; normalisation of, 244
weight function for converting distribution of geometrical momenta into a distribution of physical momenta, 76, 263
Weyl, attempt to provide a cosmical standard, 180; gauge vector, 236
widening factor 2π, 47
Wigner, theory of the nucleus, 205

x_0, fifth coordinate in position vector, 125

yes-no category of symbols, 267

Z, quantum energy tensor or phase tensor, 171, 184, 195; systems for which Z consists of pressure terms only, provides overlap of relativity mechanics and symbolic mechanics, 181
Z_0^0, Z^0_0, mixed strain tensor, 149; is cross-dual of energy tensor T_{00}, 181; factors of, are quantum momentum vectors, classically strain vectors, but space vectors of micro space, 243
Z_a, strain produced by particle in aether, 171; is strain tensor when T is the energy tensor of a relativity particle, 172
Z_μ, strain produced by particle in uranoid, 171; is strain tensor when T is the energy tensor of a quantum particle, 172
Z_m, momentum terms of Z, inhibited by condition of symmetry, 181; carrier of, is called a *forbidden oscillator*, 212
Z_p, pressure terms of Z, 184; is energy tensor of double extracule, 194; is mechanical energy, 195, 212
Z_d, density terms of Z, 184; is recalcitrant part of energy tensor, 185, 195; corresponds to double intracule, 194; is electrical energy, 195, 212
Z, metrical field, 269, 279
Zeeman effect, 254
zero level of energy, 53