Sgríbhinní Institiúid Árd-Léighinn
Bhaile Átha Cliath. Sraith A Uimh. 2
Communications of the Dublin Institute for
Advanced Studies. Series A No. 2

THE COMBINATION OF RELATIVITY THEORY AND QUANTUM THEORY

by

SIR ARTHUR S. EDDINGTON

DUBLIN
THE DUBLIN INSTITUTE FOR ADVANCED STUDIES
64-5 MERRION SQUARE
1943
(Reprinted 1960)

THE COMBINATION OF RELATIVITY THEORY AND QUANTUM THEORY.

By ARTHUR S. EDDINGTON.

Introduction.

Twenty years ago relativity theory had unified molar physics—molar mechanics, molar electrodynamics, and the geometry of molar measurement—into a single formulation, rational and complete. By "complete" I mean that it was as complete as a theory which admittedly covered only one-half of physics had any right to be. Microscopic physics, as represented by quantum theory, was still in its infancy. Since then a highly developed microscopic theory has grown up almost independently of relativity theory, and the problem of unification arises. But if the right inspiration had occurred to those who twenty years ago were casting about for a way of extending relativity theory to microscopic phenomena, our present knowledge might have been reached by a continuation of the systematic development of relativity theory, and there would have been no problem of unifying a theory which had never been divided. How ought we to have proceeded to extend relativity theory to embrace microscopic physics? What avenues were left unexplored? It is easy to be wise after the event; and, knowing what we know now, I think these questions can be answered.

There were two unexplored avenues. We neglected *statistical* relativity theory. We considered relativity transformations of exact quantities, but we did not develop to any serious extent a corresponding relativistic theory of probability distributions. Secondly, we neglected *spin*. The energy tensor $T_{\mu\nu}$ of continuous matter in molar relativity theory represents fluid in irrotational motion, and there is no provision for representing a continuous distribution of vorticity. If these two developments are pursued, all known phenomena are reached.

In 1928 Dirac opened up the spin avenue, and we went down it with a rush. My path bifurcated from the others soon after the start. The others took short cuts to the things that most interested them, and produced a theory adapted to certain parts of physics which are the field of intense experimental activity. But in fourteen years I have never had

the smallest doubt that the direction which I took in 1928 was the one which leads to the unified relativity and quantum theory.

My book *Protons and Electrons* (1936) was based almost wholly on the spin extension of relativity theory. Since then I have been mainly occupied with the statistical extension of relativity theory—the other unexplored avenue—and the present article is devoted to this part of the subject. It includes very little that has previously been published. The article concludes at the point where it would, I think, cease to be profitable to develop the statistical theory alone; that is to say, the natural continuation would be to give the full development of the spin theory, and then pass on to problems in which both are applied concurrently. It has been necessary sometimes to quote results of the spin relativity theory; but only the more elementary results, which are probably by this time familiar in one form or another, have been used. It is an advantage to begin with statistical theory, because it supplies the rigorous definition of the various new concepts which microscopic theory introduces. The spin theory is highly mathematical, and is likely to degenerate into pure algebra unless it is guided by a clear understanding of the concepts to which it is to be applied. I regard the introductory statistical part of the theory as the more difficult, because we have to use our brains all the time. Afterwards we can use mathematics instead.

Although limited to statistical theory (except for a few elementary results of spin theory), this article covers the theoretical calculation of all the fundamental physical constants except the cosmical number. Comparisons with observation are made in §§ 4, 20, 24, 28, the two comparisons in § 20 being much the most stringent.

Part 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 constructed 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 experimental procedure, and "unobservables," i.e. quantities depending on the auxiliary mathematical framework which cannot be the subject of any actual experiment. Unobservables are used to facilitate description and calculation, but are eliminated in the final calculation of observationally verifiable results.

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 the coordinate ξ, of which an observed

velocity is the time-derivative, is a *relative coordinate* of two physical entities.

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

In modern physics these two conditions of observability have been applied separately with very far-reaching results. *In relativistic quantum theory they must be applied in combination.* 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 essential point is that an observable coordinate is measured, not from an abstract geometrical origin, but from something which is involved physically in the experiment which furnishes its observed value. Being involved physically, it experiences those incalculable reactions which limit the precision of our knowledge in the way described 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 mathematical reference frame which is inaccessible to measurement. The mathematical reference frame is eliminated in the final prediction 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 physical landmarks by probability scatter.

There is one simple class of quantum problems, namely, the theory of the eigenstates of a two-particle system, in which there is no need to introduce an extraneous origin, since we can deal directly with the coordinates of one particle relative to the other. But normally we are concerned with a number of particles whose coordinates are measured from a common origin. Consider a system of n particles with coordinates x_r, y_r, z_r in the geometrical frame. These coordinates are unobservable. To obtain physical (observable) coordinates, we must substitute a physical origin which has a probability scatter relative to the geometrical origin. If the (unobservable) geometrical coordinates of the physical origin are x_0, y_0, z_0, the relative coordinates

$$\xi_r, \eta_r, \zeta_r = x_r - x_0, y_r - y_0, z_r - z_0 \qquad (1)$$

are observables. A measurement of ξ_r will give a value taken at random from the pre-existing probability distribution of ξ_r; or equivalently it gives the distance from a random point in the probability distribution of x_0 to a random point in the probability distribution of x_r. Thus (1) is not a simple change of origin, but a change from an origin fixed in the

geometrical frame to an origin with a probability scatter in that frame. It will be necessary (§ 21) to make a special study of this type of transformation which is, of course, outside the ordinary tensor theory of coordinate transformations.

The coordinates postulated in the current dynamical equations of quantum theory must be measured from a physical origin, since they and their conjugate momenta are always assumed to be observables, being in fact the typical observables of wave mechanics.[1] The urgent question arises: How is this physical origin defined, and what probability distribution has been assigned to it? For the current equations are clearly not of a form which would be invariant for arbitrary changes of the probability distribution of the origin.

In treating a system of n similar particles we naturally take their centroid as origin. When n is large, the centroid has the important statistical property that the form of its probability distribution does not depend on the form of the 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 as physical origin, we have the immense advantage of starting with a knowledge of the distribution function $f(x_0, y_0, z_0)$ of its geometrical coordinates, complete except for the one disposable constant in the Gaussian law. It should be remembered that it would be impossible to ascertain $f(x_0, y_0, z_0)$ by observation, since geometrical coordinates are unobservable; so that, unless we have theoretical information, such as that furnished by the "law of large numbers," we can have no knowledge of f. Assuming spherical symmetry, we have

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

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

Throughout our investigations we shall employ a physical reference frame related to the geometrical reference frame by the distribution function (2) of the physical origin. The system of n particles, which provides an observational identification of this origin, will appear in our work sometimes as the principal object of study and sometimes as the "standard environment" of a small system that is being intensively treated. It is essential to use the same physical reference frame with the same σ throughout physics, whether the system treated is on an atomic,

[1] The dynamical equations describe the propagation and dispersal of the concentrations of probability (wave packets) created by our measurements. It is therefore essential to identify correctly the variates in which these concentrations of probability are produced.

molar or cosmical scale; otherwise the laws applicable in the different branches of physics will not grade continuously into one another.

Starting with an abstract geometrical frame of coordinates, we pass from pure geometry into physics by introducing a physical frame whose origin has a Gaussian probability distribution relative to the geometrical origin. The standard deviation σ of this distribution "puts the scale into" the physical frame and all that we construct in it, whether it be a nucleus, an atom, a crystal or the whole extent of physical space. Initially we seem to have freedom of choice of σ; but the freedom is illusory, since σ can only be measured in terms of the extensions of physical structures whose scale it has itself determined. To double σ would be physically meaningless, since it would double the scale of everything in the physical universe, leaving the relative dimensions of things unchanged.

2. 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 particles. Each particle has the same probability p of being within V_0, and the expectation number of particles in V_0 is therefore $n_0 = pN$. Let the actual number in V_0 be n, and let $y = n - n_0$. Then by James Bernoulli's theorem the "fluctuation" y has the distribution function

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

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

$$f_\infty(y) = (2\pi n_0)^{-\frac{1}{2}} e^{-y^2/2n_0}. \qquad (4)$$

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 (4) is obtained by compounding with (3) an independent Gaussian fluctuation with standard deviation $(n_0^2/N)^{\frac{1}{2}}$ and distribution law

$$f_e(y) = (2\pi n_0^2/N)^{-\frac{1}{2}} e^{-Ny^2/2n_0^2}. \qquad (5)$$

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

$$n = n_0(1 + \zeta). \qquad (6)$$

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

$$g_e(\zeta) = (2\pi/N)^{-\frac{1}{2}} e^{-\frac{1}{2}N\zeta^2}. \qquad (7)$$

The whole fluctuation is, therefore, separated into two independent fluctuations (4) and (7), the one depending on the finitude of n_0 and the other on the finitude of N. We distinguish them as the *ordinary* and *extraordinary* fluctuations. The extraordinary fluctuation is to be combined negatively, i.e. removed from the ordinary fluctuation. If, for

example, we are treating in statistical mechanics a unit volume of gas, we generally take into account only the ordinary fluctuation; that is to say, we assume that outside the unit volume the gas extends uniformly without limit in all directions. But an infinite extent of uniform gas is contrary to relativity theory, which introduces a curvature of space determined by the density of the gas, so that space becomes closed. *We shall show that this space-curvature is simply a way of taking the extraordinary fluctuation into account.*

The fluctuation (7) transforms an exact particle density s_0 into an uncertain density

$$s = s_0(1 + \zeta) . \tag{8}$$

Instead of considering an uncertain number of particles n in a fixed volume V_0, we can consider a fixed number of particles n_0 and transfer the uncertainty to the containing volume V. The mathematical analysis for a system of n_0 particles occupying a volume V_0 is adapted to other values of V by a change of linear scale. Thus the uncertainty is now contained in a linear scale-factor $1 + \epsilon$, defined by $V = V_0/(1 + \epsilon)^3$. If we had to transform a distribution of discrete values of ζ into a distribution of discrete values of ϵ, the relation would be $(1 + \zeta) = (1 + \epsilon)^3$. But in transforming a continuous distribution, discrete values of ζ and ϵ must be replaced by constant ranges of ζ and ϵ; and we have to include a factor $d\epsilon/d\zeta$ transforming the constant ranges of ϵ into ranges which correspond to constant ranges of ζ. The relation is accordingly $1 + \zeta = \text{const.} \times (1 + \epsilon)^3 d\epsilon/d\zeta$, which on integration gives

$$(1 + \zeta)^2 = (1 + \epsilon)^4 \tag{9}$$

For the distribution function (7) the values of ζ which have sensible probability are of order not greater than $N^{-\frac{1}{2}}$. Hence (9) becomes with sufficient approximation $\zeta = 2\epsilon$. By (7) the standard deviation of ζ is $N^{-\frac{1}{2}}$; hence the standard deviation of ϵ is

$$\sigma_\epsilon = 1/2\sqrt{N} \tag{10}$$

The geometrical frame is our standard of fixity when we speak of the uncertainties of physical quantities; and the exact scale of measurement of the geometrical coordinates corresponds to $\epsilon = 0$. The uncertain scale-factor $1 + \epsilon$ is introduced in the measure system of the physical coordinates ξ, η, ζ. Considering a point distant r from the origin, the differences $\xi - x, \eta - y, \zeta - z$ will now include, besides the standard deviation σ in all directions due to the uncertainty of the physical origin, a standard deviation $\sigma_\epsilon r$ in the radial direction due to the uncertainty of the physical scale of measure of ξ, η, ζ. Remembering that the extraordinary fluctuation is to be combined negatively, the resultant standard deviation is

$$\text{radial } (\sigma^2 - \sigma_\epsilon^2 r^2)^{\frac{1}{2}}, \quad \text{transverse } \sigma \tag{11}$$

This may be described as the *local uncertainty* of the physical reference frame, since it represents the difference between the physical and geometrical coordinates. We have derived it by combining the uncertainty of a distant origin with the uncertainty of scale; but it can be interpreted more compactly as the uncertainty of a local physical origin relative to a local geometrical origin. The anisotropic distribution could be transformed into an isotropic distribution by introducing an appropriate local coordinate system in place of the coordinates associated with the distant origin. Independently of coordinate systems, the local uncertainty in a given direction defines an extension which might be employed as a 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 of a line-element according to σ-metric. By (11) the lengths of radial and transverse elements are proportional to $dr/(\sigma^2 - \sigma_\epsilon^2 r^2)^{\frac{1}{2}}$ and $rd\theta/\sigma$; so that the general formula is

$$ds^2 = \frac{dr^2}{1 - (\sigma_\epsilon^2/\sigma^2) r^2} + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2 . \qquad (12)$$

This is the well-known formula for a line-element in spherical space of radius $R_0 = \sigma/\sigma_\epsilon$. Hence by (10)

$$\sigma = \sigma_\epsilon R_0 = R_0/2\sqrt{N} . \qquad (13)$$

It will be shown in § 3 that the σ-metric agrees with the recognised definition of length in relativity theory and quantum theory; so that the spherical metric assigned to a steady uniform distribution of particles in relativity theory is the device by which current theory takes into account the effect of the extraordinary fluctuation arising from the finitude of N. It is not *necessary* to employ this device; and for the domestic development of quantum theory it is preferable to keep to flat space and take the scale fluctuation into account explicitly, introducing an extra variate for that purpose. The present investigation exhibits the connection between the two methods of treatment; and it establishes the important formula (13) which determines our fundamental constant σ in terms of the constants R_0, N familiar in cosmological theory.

The coordinates r, θ, ϕ in (12) are the ordinary polar coordinates of the orthogonal projection of the point of spherical space on the tangent flat space at the origin. This leads to a simple geometrical representation of our results. The four rectangular coordinates of a particle on the hypersphere, referred to a geometrical origin O at the centre, satisfy

$$x_r^2 + y_r^2 + z_r^2 + w_r^2 = R_0^2 \qquad (14)$$

so that for uniform probability distribution over the hypersphere the mean values are

$$\overline{x_r^2} = \overline{y_r^2} = \overline{z_r^2} = \overline{w_r^2} = \tfrac{1}{4} R_0^2 \qquad (15)$$

Thus the standard deviation of a coordinate of a particle is $\frac{1}{2} R_0$, and

the standard deviation of a coordinate of the centroid O' (in 4-space) is $\frac{1}{2}R_0/\sqrt{N}$, which is equal to σ. If now we transfer the geometrical origin to a point P on the hypersphere and take the u-axis along $O\,P$, the three-dimensional uncertainty of the physical origin corresponds to the $x,\ y,\ z$ components of $O\,O'$; and the uncertainty of scale corresponds to the use of $O'P$ as standard of length instead of the exact standard $O\,P$.

Besides uncertainty of origin and scale, the physical reference frame will also have uncertainty of orientation. But by the use of tensors we can make our description independent of the orientation of the frame. Current quantum theory is alert to the difficulty of defining orientation, and the problem is treated rigorously in spin theory. We may, therefore, here confine attention to the uncertainty of origin and scale, which has been completely neglected.

The use of a centroid as origin of space coordinates obviously cannot be extended to an origin of time. But we shall deal entirely with steady distributions, and shall not require a time origin.

3. The standard of length.

The definition of length, both in theoretical physics and in practical metrology, agrees with the σ-metric. To make this clear we consider the necessary conditions that must be fulfilled by an ultimate standard of length. It must be available for comparison at all times and at all places. We require a physical structure, not necessarily permanent, but constructable at any place and time from a recorded specification. The form of specification is determined by the condition that the definition of length (and a corresponding definition of time-interval) is required at the very beginning of physics, and must precede the definitions of all other physical quantities. It would, therefore, be a vicious circle to employ any "dimensional" physical quantities in specifying the standard which forms part of the definition of length. The quantitative part of the specification must, therefore, consist entirely of *pure numbers*. It is only in quantum theory that we have developed a method of specifying physical structure by pure numbers—numbers of elementary particles in a configuration specified by quantum numbers. The standard of length must accordingly be a *quantum-specified structure*.[2]

The equations of quantum theory determine the various spatial extensions in quantum-specified systems as fixed multiples of a linear unit $\hbar/m_e c$. Thus, whether the unit is constant or not, the ratio of two such extensions at one time and place is the same as at any other time and place; and they give equivalent systems of reckoning length, differing

[2] For a fuller discussion see Eddington, *The Philosophy of Physical Science*, pp. 70–85.

only by a constant conversion factor. In saying that the metric is quantum-specified we fix it uniquely.

It only remains to show that the unique quantum-specified metric agrees with 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 σ. Abundant calculations of such ratios occur in the course of development of the present theory. For example, the Rydberg constant for hydrogen is found to be[3]

$$R_H = \frac{3}{16\pi\sqrt{5}} \cdot \frac{1}{136^2 \cdot 137} \cdot \frac{1}{\sigma}. \qquad (16)$$

Thus the metric which makes the local uncertainty a constant for every place, time and direction (by adopting it as the standard of length) also makes the wave-lengths of the hydrogen lines constant for every place, time and direction. In other words, the σ-metric is the ordinary metric.

4. Range of nuclear forces and the recession of the nebulae.

The description of physical systems by probability distributions requires certain precautions which are liable to be overlooked, because they have no counterpart in the classical conception of physics, from which most of our nomenclature has been derived. Distinctions have to be made which are unprovided for in the usual terminology.

Consider the coordinates ξ_r, ξ_s of two particles measured from the physical origin. An observational measurement of ξ_r gives the distance from an undetermined point in the distribution of x_0 to an undetermined point in the distribution of x_r. If ξ_s is also measured, the measure will have an independent starting-point in the distribution of x_0. Thus the coordinate difference $\xi'_{rs} = \xi_s - \xi_r$ includes the coordinate difference of two independent points in the probability distribution of x_0; this has a Gaussian distribution with standard deviation $\sigma\sqrt{2}$.

But the relative coordinate $\xi_{rs} = x_s - x_r$ can also be measured directly from the r th to the s th particle without the intermediary of an origin. Both ξ'_{rs} and ξ_{rs} are observables, and they have the same expectation value, but their probability distributions are different, that of ξ'_{rs} having the greater spread. We have in fact (in the notation of the theory of errors)

$$\xi'_{rs} = \xi_{rs} \pm \sigma\sqrt{2} \qquad (17)$$

Evidently the wave functions and the conjugate momenta of ξ'_{rs} and ξ_{rs} will be different.

Since σ is found to be of the order 10^{-13} cm. the difference between ξ'_{rs} and ξ_{rs} appears directly in nuclear problems, and in the

[3] The formulae, from which this result is derived, are given in *Proc. Phys. Soc.*, 54, 491, § 2.

scattering of protons by protons where very close approaches of the particles occur. The non-Coulombian energy of two protons is a singular energy associated with their actual coincidence, i.e. with $\xi_{rs} = 0$ (see § 28). The corresponding values of ξ'_{rs} form a probability distribution with standard deviation $\sigma\sqrt{2}$, so that the non-Coulombian energy has the form

$$A e^{-(\xi'^2_{rs} + \eta'^2_{rs} + \zeta'^2_{rs})/k^2} \qquad (18)$$

where $k = 2\sigma$. We thus get an apparent range of the non-Coulombian forces when the particles are referred to an origin, although the range in ξ_{rs}, η_{rs}, ζ_{rs} is actually zero.

We can make an observational test of the results thus far obtained. According to the latest experimental data available,[4] $k = 1{\cdot}9 \cdot 10^{-13}$ cm.; so that

$$\frac{R_0}{\sqrt{N}} = 2\sigma = k = 1{\cdot}9 \cdot 10^{-13} \text{ cm.} \qquad (19)$$

From the well-known formula $\kappa M_0/c^2 = \frac{1}{2}\pi R_0$, connecting the mass M_0 and radius R_0 of the Einstein universe, we obtain

$$\frac{N}{R_0} = \frac{\pi c^2}{\kappa M} \qquad (20)$$

where M is the mass of a hydrogen atom and κ is the constant of gravitation. From (19) and (20), we obtain N and R_0 separately; and we can then calculate the limiting speed of recession of the galaxies $V_0 = c/R_0\sqrt{3}$. The result is $V_0 = 585$ km. per sec. per megaparsec. The observed value given by Hubble and Humason is 560 km. per sec. per megaparsec.

Thus the two most direct methods of determining σ—from the range of nuclear forces and from the recession of the galaxies—are in good agreement. Owing to the inaccuracy of the observational data, the test is rather rough. Much more accurate (but indirect) methods of determining σ are found in the later developments of the theory. The definitive value is

$$\sigma = 9{\cdot}604 \cdot 10^{-14} \text{ cm.} \qquad (21)$$

From the point of view of the ordinary theory of statistics the confusion of ξ'_{rs} with ξ_{rs} is an error of the most elementary kind. Its consequences appear straightforwardly in an apparent (but really non-existent) range of the non-Coulombian interaction of two particles; and this agrees with the value of σ found by extra-nuclear observation.

---

[4] Eddington and Thaxton, *Physica*, 7, 122 (1940).

This was pointed out in 1937.[5] The hypothesis of a "meson-field," which has since been so prominent in nuclear physics, is clearly a retrograde step. It is, of course, possible that the formal mathematics of meson-field theory may constitute a valid treatment of the $\sigma\sqrt{2}$ dispersion, since a mathematical scheme admits of many alternative physical interpretations; but, if so, the "mesons" of the meson-field are totally unconnected with the mesotrons produced by cosmic rays.

5. 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 system that is being intensively studied. The environment comprises everything not specifically included in the object-system, whether it surrounds it or permeates it.

The most elementary formulae relate to the most simple object-systems in the most simple environments. Just as we have to begin with simple objects—electrons, two-particle systems, etc.—so we have to begin with simple environments. 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. By general relativity theory such a distribution occupies a spherical space, and constitutes an "Einstein universe." Usually it is further specialised as a "zero-temperature uranoid"; the pressure is then zero, and there is no radiation present.

Two lines of thought have led us to contemplate a system of a large number of particles as an environment or background for the objects intensively studied. In § 1 it was a question of *metric*; the system of particles determines the uncertainty of the reference frame, and hence the scale of the various structures related to that frame. Now it is a question of *mechanical interaction*; a vast assemblage of particles is present in the actual universe, and we have to take account of its influence (e.g. gravitational influence) on the object-system under consideration. But the two aspects are not really distinct. General relativity, by unifying geometry and mechanics, amalgamates the metrical and mechanical influences of the environment; both are included in the description of the influence as a "field" of $g_{\mu\nu}$. This field is at the same time a metrical field and an inertial-gravitational field; and it is indifferent whether we regard it as influencing the characteristics of the object-system by mechanical agency or by determining the measure-system in which they are expressed.

[5] *Proc. Roy. Soc.*, A, 162, 155.

In the present series of investigations we shall generally take into account the *complete* uranoid. This does not imply that the remote environment plays a greater part in determining local phenomena than is ordinarily admitted in relativity theory. In particular, the most radical changes in the distribution of the extragalactic nebulae will have no observable effect on small scale systems, provided that there is no sensible alteration of the field of $g_{\mu\nu}$ in the region occupied by the system; and the observable results for small scale systems will be the same in the actual irregular expanding universe as in the uniform static uranoid used for calculation. We take into account the whole uranoid, because mathematically it is easier to treat a whole universe than part of one. If we do not include the whole, we give ourselves the extra trouble of discovering boundary conditions which shall have the same effect as the part that has been left out. On the other hand, it should be realised that the remote environment only affects the problem through integral properties which *could* be expressed as boundary conditions.

Einstein's theory removed the absolute distinction between gravitation and inertia; but in practice we continue to separate them conventionally, and, for example, we distinguish the gravitational energy $-m_e\phi$ of an electron from its inertial energy $m_e c^2$. In the present nomenclature, the standard environment provides the inertial part of the $g_{\mu\nu}$-field, and the deviation between the actual and the standard environment provides the gravitational part.

We shall confine attention almost exclusively to uniform probability distributions of particles. It might be thought that it would soon become necessary to pass on to non-uniform distributions, so as to obtain the kind of system studied in quantum theory. But atoms, etc., are constructed, not by introducing non-uniformity of distribution of nuclei and electrons, but by introducing *correlations*. Consider, for example, a vessel known to contain an electron and proton. If they are uncombined, they are each equally likely to be anywhere in the vessel. If they combine into a hydrogen atom, the electron is still equally likely to be anywhere in the vessel, and so is the proton. What has happened is that the coordinates of the electron and proton have become correlated, though they each continue to have uniform probability distribution. Atomic wave functions, such as "the wave function of the hydrogen atom," are *correlation wave functions*. These must be carefully distinguished from the *distribution wave functions* of protons, electrons, etc. Except in a few special problems (deflection in a molar electromagnetic field), there is no occasion to consider non-uniform distribution wave functions. Thus the distribution wave functions are generally limited to the "infinite plane waves" introduced in elementary wave mechanics, and all the complication is reserved for the correlation wave functions.

6. The extraneous standard and scale-free physics.

We adopt a system of "natural units" such that

$$c = 1 \qquad 8\pi\kappa h^2 = 1 \quad . \qquad (22)$$

These particular relations are conventional; but the imposition of two fixed relations between the three units of length, time and mass is essential to the theory. It removes a redundant fluidity of description of physical systems, which is occasioned by referring them to three extraneous standards when one standard is all that is required to fix the scale. In modern physics fluidity of description is provided for by a systematic transformation theory (tensor calculus, etc.); and we lose the benefit of systematic treatment if we graft on to it a haphazard traditional transformation of units.

The relations (22) leave one unit at our disposal. It is immaterial whether this is taken to be a mass, length, time or any combination of them. If, for example, a mass is adopted, the corresponding units of length and time are fixed by (22); and every physical quantity has just one *dimension-index* showing how it varies with the unit of mass.

We have seen (§ 2) that in the physical frame the linear scale has an uncertainty σ_ϵ. This may be regarded as an inherent uncertainty of the standard of length. By expressing our results in natural units, all characteristics of the system are described in terms of the standard of length; and a characteristic of dimensions (length)^y has an uncertainty $y\sigma_\epsilon$ arising from the uncertainty of the standard, as well as the scatter contained in its own geometrical probability distribution. For a mass we find that $y = -3$. If the mass were measured by comparison with an independent standard (the gram), the uncertainty would have no relation to σ_ϵ. The definitions of physical quantities must state explicitly the nature of the comparisons by which they are understood to be measured; and the traditional system of definitions which involves three extraneous standards, whilst adequate for exact quantities, is not precise enough for the treatment of probability distributions, and leads to the same kind of ambiguity as the confusion of ξ'_{rs} and ξ_{rs} (§ 4). We remove such ambiguity by introducing the system of natural units with only one extraneous standard.

One extraneous standard must be retained. We can describe the internal structure of a physical system wholly by numerical ratios; but to complete the description it is necessary to fix the scale by reference to some outside standard. If the whole universe were investigated as one system, no outside standard would be needed. But the analytical method of physics divides the universe into simple systems of various types, which are investigated one by one. Each system is supposed to be surveyed from outside. The single extraneous standard is the "loose end," which

enables us to put together in their proper scale relation the fragments into which our analytical method divides the universe.

The standard uranoid has two linear characteristics σ and R_0 which we can regard as independent.[6] Our reproducible standard of length has (apart from its small uncertainty σ_ϵ) everywhere a fixed ratio to σ and to R_0; but in a considerable part of physics, the actual value of the ratio is not involved. We shall call this *scale-free* physics. Scale-free physics deals with structures which can be adapted to any scale. The criterion 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, then if one system of the series is a physically possible system all are physically possible. This abstract ideal can only be approximately realised in the actual world; for the homology must break down when the system is enlarged to a size comparable with the universe or diminished to a size comparable with an atom. The exact equations must always involve σ and R_0; but there are many investigations in which to a very high approximation we can set $\sigma = 0$, or $R_0 = \infty$, or both. We therefore distinguish three branches of physics:—

(1) Scale-free physics: problems involving neither σ nor R_0.

(2) Cosmical physics. problems involving R_0.

(3) Quantal physics: problems involving σ.

The term quantal physics is not intended to be synonymous with quantum theory; because text-books on quantum theory include a great deal of scale-free physics as well as quantal physics, and wave analysis is applied to both. In general, wave functions which are not self-normalising, and are arbitrarily normalised to represent a certain number of particles per unit volume, are scale-free.

The scale-free approximation neglects the "transverse" part of the field. When we include the transverse part of the electric field we pass over into quantal physics. When we include the transverse part of the gravitational field we pass over into cosmical physics.

7. Stabilised characteristics.

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 observation, and is therefore limited by the

[6] Actually they are not independent, because it appears in later developments that N necessarily has the value $\frac{3}{2} \times 136 \times 2^{256}$.

uncertainty principle; but knowledge of the mass m and the charge e is treated on a different footing. The values of m and e are taken from a table of constants. It is true that they rest on observation, but not on observation of the object at x, y, z. They are treated as *free information*, not to be paid for by a reciprocal uncertainty of the conjugate variates.

The theorist would explain that this free information follows from his knowledge that the object is an electron. But how has he got to know that it is an electron? The 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; and we shall not dispute it. But it must be recognised that it creates a disjunction with observational physics.

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

Consider, for example, a simple spinless particle characterised mechanically by a momentum vector p_1, p_2, p_3 and a proper mass m. It is generally understood that p_1, p_2, p_3 are observables; but m may be either an observable or a stabilised characteristic.[7] By stabilising m, the probability distribution of momentum 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, where $p_4 = (m^2 + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}}$) to a three-dimensional distribution. We shall find that the number of dimensions of the probability distribution enters as a coefficient into many leading formulae, and that particles must be classified primarily according to this number. A particle whose probability distribution has k dimensions will be called a V_k particle. Thus a spinless particle is a V_3 or a V_4 particle according as its proper mass is or is not stabilised. The number k may also be described as the number of degrees of freedom of the 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 a system. 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, electrically neutral, zero-temperature. So much restriction is imposed that only two observables σ and R_0 remain; and of these only σ is concerned in the ordinary problems of quantum theory. The restriction is given as free information. When we treat the theory of a hydrogen

[7] The term "observable" is used ungrammatically in quantum theory with the meaning "supposed to have been observed"; a stabilised characteristic might have been observed, but has not been observed.

atom in the standard environment, 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. We might alternatively consider an atom in a universe found observationally to agree with the standard uranoid within certain limits of uncertainty; but then it would be only logical to treat the atom as having been found observationally to be hydrogen within 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 in the form of invariant conditions imposed on the vector or 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 quantities required to specify the tensor, they reduce the number of dimensions of its probability distribution.

8. Pseudo-discrete states.

Two kinds of wave functions are employed in quantum theory. The more typical kind represents concentrated distributions of probability whose density falls off at great distances, so that the integral converges. The "normalised" density distribution $\rho_n(x, y, z)$ is such that the total mass is the mass of one particle; the state is then said to be *in unit occupation*. But we may associate with the state an occupation factor j different from unity, and the density is then $j\,\rho_n(x, y, z)$. The occupation factor represents the probability that there is a particle in the state, or the number of particles in the state. Strictly speaking, it is a symbolic operator J which only reduces to an eigenvalue when there is definitely an integral number of particles in the state. But in elementary theory we take it to be a number j which is the expectation value of J.

The foregoing "self-normalising" wave functions occur only as correlation wave functions. The distribution wave functions are of the other kind, typified by the "infinite plane waves" of elementary quantum theory. The infinitude is not to be taken literally, but implies that the distribution extends uniformly to a distance large compared with σ but small compared with R_0—anything from a thousandth of a millimetre to a million parsecs. The degree of occupation of this type of wave function cannot be specified by the number of particles in the whole undefined extent of the distribution. We accordingly select an arbitrary density ρ_n as the density of unit occupation; or equivalently we select an arbitrary normalisation volume V_n, and define unit

occupation as "one particle per volume V_n." This does not mean that there is one particle in a volume V_n. (That would be represented by a wave-packet, not a simple plane wave function.) Each particle has an even probability distribution over the whole extent of the wave function. Any particle occupying a wave function of this kind must be regarded as *an unidentified member of a large assemblage.*

Wave functions of the second kind are *pseudo-discrete.* It is important to realise that pseudo-discrete wave functions are not a limiting case of continuous wave functions. A set of discrete eigenfunctions has the general form $\psi(x, y, z, a)$, where a stands for a set of parameters defining constant characteristics of the eigenstate. When the eigenstates are continuous we have to consider the occupation of a continuous range of values of a. This may be described by a distribution function $j(a)$ or a wave function $\chi(a)$. In the latter case the wave functions ψ, χ can be amalgamated into a *continuous wave function* $\psi(x, y, z, a)$ over the coordinates and parameters. In practice, however, we do not extend wave treatment to the distribution over a, and are content to use the distribution function $j(a)$. A pseudo-discrete state is formed by concentrating a finite amount of occupation j into a range of a small enough to be treated as infinitesimal. When the whole probability of the system is concentrated in one pseudo-discrete state, its state is said to be *almost exact.*

Relativity mechanics is based on the energy tensor $T_{\mu\nu}$, which includes the density (energy-density or mass-density) as a component T_{44}. The energy tensor is a scale-free characteristic, since the density contributed by any kind of particle can be varied at will by varying the volume over which its probability distribution extends. On the other hand, the momentum vector, which includes the mass as a component, is not scale-free. We have, therefore, the distinction that the particles of scale-free physics are characterised by energy tensors, and the particles of quantal physics by momentum vectors. Evidently the pseudo-discrete wave functions are the wave functions of scale-free physics, and the self-normalising wave functions are the wave functions of quantal physics.

Pseudo-discrete wave functions appear as the first step in the passage from molar relativity theory to microscopic theory. They represent the molar object as a large assemblage of particles in one or more pseudo-discrete states. The molar energy tensor is represented as the sum of contributions $\Delta T_{\mu\nu}$ of the particles, $\Delta T_{\mu\nu}$ being a characteristic of the state of the particle. The particles are unidentified; so that, as regards any one particle, we know only the probability of its being in the various states; in other words, the occupation factors j_r of the pseudo-discrete states may be interpreted either as frequencies in the assemblage or as probabilities of an individual particle. Since it is usually

unnecessary to consider non-uniform distributions, the molar object subjected to this analysis may be taken to be a uniform gas.

It might be thought that the scope of scale-free theory in microscopic physics would be very limited, since the masses of particles are not scale-free. But results usually expressed in terms of mass can often be equally well expressed in terms of density. Our next aim will be to determine theoretically the mass-ratio m_p/m_e of the proton and electron. This can alternatively be described as the density-ratio of two constituents (positive and negative) of molar hydrogen. As a density-ratio it is a problem of scale-free theory, and we shall treat it in that way. It may be recalled that Dalton's atomic theory was scale-free, and enjoyed a long fruitful career before anything was discovered as to actual scale of the atom.

In the course of development of the theory, it is found that the self-normalising wave functions represent electric energy, whereas the pseudo-discrete wave functions represent mechanical energy. By "electric" we here mean *distinctively* electric. (Naturally in unified theory electric and mechanical energy merge into one another, like inertial and gravitational energy, and the exact line of separation must be conventional.) The ratio of the electrical to the gravitational force between two particles is of the order $\sqrt{N}$, and the counterpart in mechanical theory of the quantum of electrical angular momentum $\hbar$ is $\hbar\sqrt{N}$. Thus σ and $\frac{1}{2}R_0$, which are in the ratio $\sqrt{N}$, are corresponding linear units of structure. If the same principles which provide a definite scale of structure for electrical energy are applied to mechanical energy, the structure of the latter is on the scale R_0. This is the quantum aspect of the cosmical curvature of space. Except in cosmical physics R_0 is treated as infinite ($R_0/\sqrt{N}$ remaining finite), and the mechanical waves in wave mechanics are structureless plane waves.

So long as we deal only with protons and electrons, the argument ξ, η, ζ of a self-normalising electrical wave function is geometrically related to the argument x, y, z of a mechanical wave function. But physically ξ, η, ζ is an electrical characteristic, namely, polarisation (separation of an electric doublet) or an analogous quantity (separation of two like charges). Existing quantum theory may be said to have been primarily designed for the intensive study of electrical polarisation. Although it is ideally possible to develop unified theory from an arbitrarily chosen starting point, it is not surprising that we should prefer a less specialised point of entry. Accordingly, the typical quantum wave functions referring to polarisation only begin to appear in the theory towards the end of the present investigation.

PART II.—MULTIPLICITY FACTORS.

9. The rigid-field treatment.

Quantum theory and relativity theory from the very outset employ opposite methods of treating the field. The root of the difference is that wave mechanics, which is the principal analytical method used in quantum theory, depends on the concept of a rigid field. The metrical (or gravitational) field $g_{\mu\nu}$ is unconditionally assumed to be rigid. The molar electromagnetic field κ_μ is also assumed to be rigid, provided that (following the usual practice in quantum theory) we exclude the "transverse" part, which is treated separately as radiation. Even radiation is not a genuine exception, because it appears in wave mechanics, not as a field, but as an assemblage of particles (photons).

A rigid field of $g_{\mu\nu}$ and κ_μ having been specified, we construct in it a frame of eigenstates, leaving ourselves free to decide later to what extent each eigenstate is occupied. This *flexibility of occupation* is essential to the methods and conceptions of wave mechanics; for the main subject of study is the transition of particles between eigenstates. There is no such flexibility in relativity theory, where the $g_{\mu\nu}$ are conditioned by the matter present.

The rigid field treatment is at best an approximation, and, if it is not used circumspectly, it is liable to be altogether fallacious. In order that it may be legitimate as an approximation, small changes of occupation of the eigenstates must produce only changes of the second order in the field:—

The field must be stationary for small changes of the occupation factors of the eigenstates.

This condition must be regarded as incorporated in the definition of the particles contemplated in quantum theory, so that they may have the characteristic freedom of transition which wave mechanics postulates. That is to say, we define a *quantum particle* as the occupant of a state in a rigid field, in contrast to a *relativity particle* which is a singularity of an essentially non-rigid field. We must avoid the widespread error of applying to quantum particles familiar formulae which have been developed for relativity particles.

It would be inconsistent with this definition to contemplate a system of quantum particles in a non-rigid field. We, therefore, call the field which satisfies the stationary condition the *self-consistent field* for the state of occupation contemplated. As we should ordinarily express it, failure of the stationary condition implies that the field assumed in calculating the eigenstates is not the field produced by the particles occupying the eigenstates.

We shall first take the eigenstates to be discrete.[8] We have then a set of states represented by ψ_r ($r = 1, 2, 3, \ldots$) to each of which is attached an occupation factor j_r. Since changes of the system are specified by changes of the j_r, the j_r are regarded as generalised coordinates (or momenta). The total energy of the system, including that of the self-consistent field, can then be expressed as a function $H^\circ(j_1, j_2, j_3, \ldots)$. It will not in general be a linear function of the j_r. Let

$$E_r = \partial H^\circ / \partial j_r \tag{23}$$

so that

$$dH^\circ = \Sigma E_r dj_r \tag{24}$$

And let

$$E^\circ = \Sigma j_r E_r = \Sigma j_r \partial H^\circ / \partial j_r \tag{25}$$

Then the energy of a particle in the state ψ_r must be identified with E_r. For, in order that the field may be rigid, the whole change of energy dH° must be accounted for as a change of particle energy alone. This condition is expressed by (24), $E_r dj_r$ being the change due to the addition of the fraction dj_r of a particle of energy E_r. The difference

$$W^\circ = H^\circ - E^\circ = H^\circ - \Sigma j_r \partial H^\circ / \partial j_r \tag{26}$$

is the energy of the field. By (24) and (25)

$$dW^\circ = -\Sigma j_r dE_r . \tag{27}$$

We note that if $H^\circ(j_1, j_2, j_3, \ldots)$ is a homogeneous function of the n th degree, (25) and (26) give

$$E^\circ = nH^\circ \qquad W^\circ = (1 - n)H^\circ \tag{28}$$

We have considered the energy for definiteness, but the same theory of partition applies to any other characteristic conceived as additive, e.g. energy-density, momentum, angular momentum, pressure.

By (23) the energy E_r of a particle varies with the initial state of occupation. But, having chosen an initial state, we can then make small changes of occupation keeping E_r and W° fixed.

10. Rigid fields in scale-free physics.

We shall now apply the rigid-field condition to scale-free systems. Some modification is necessary, because the eigenstates become continuous.

---

[8] There is then a one-to-one correspondence of eigenstates when the occupation factors and the corresponding self-consistent field are varied; so that we can identify unambiguously the *same* eigenstate $\psi_r(j)$, $\psi_r(j')$ in two different states of occupation of the system.

Discrete states are identified by quantum numbers, and there is no ambiguity in identifying the "same" state in a system after the occupation factors have been changed. Quantum numbers being no longer available, we have to introduce some other set of classifying characteristics in order to specify the states.

Let the adopted set of classifying characteristics be

$$X_\alpha \ (\alpha = 1, 2, \ldots n).$$

We shall use characteristics of the same physical dimensions (usually components of a tensor) and adopt an extraneous standard of the same dimensions, so that the X_α, have 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 correspond to a possible state of the system. In general, the possible states will be restricted to a k-dimensional locus. We shall call this locus the *phase space* of the system. The volume of an element of phase space will be denoted by $d\tau$. This implies that a metric, giving a definite reckoning of volume, has been defined. For the phase spaces which 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,* i.e. it must be such that the 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 $j(X)$, such that $j(X)\,d\tau$ is the probability associated with the states in the range $d\tau$. The consequent changes in the formulae of § 9 are easily found, summations being replaced by integrations and ordinary differentiation with respect to a variable j_r by Hamiltonian differentiation[9] with respect to a function $j(X)$. Corresponding to (23), (24), (25), (27), we have

$$E = \eth H^\circ / \eth j \tag{29}$$

$$\delta H^\circ = \int E \delta j \, d\tau \tag{30}$$

$$E^\circ = \int E j \, d\tau \tag{31}$$

$$\delta W^\circ = \delta H^\circ - \delta E^\circ = - \int j \delta (E \, d\tau) \tag{32}$$

Let l be the dimension-index of H°, the dimension-index of X_α being 1 as already stated. The equations show that $E(X)$, E°, W°

[9] Eddington, *Mathematical Theory of Relativity,* § 60.

also have the index l. Thus $E\,d\tau$ has the index $l + k$. Since the system is scale-free, we may take as a possible variation an infinitesimal change of scale $X_\alpha \to (1 + \epsilon)\,X_\alpha$. Then $W^\circ \to W^\circ\,(1 + \epsilon)^l$, and $Ed\tau \to Ed\tau\,(1 + \epsilon)^{l+k}$; so that, ϵ being infinitesimal,

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

Hence, by (32) and (31),

$$W^\circ = -\frac{l+k}{l}\,E^\circ \qquad H^\circ = -\frac{k}{l}\,E^\circ . \tag{33}$$

By arbitrarily dividing phase space into small numbered cells $(d\tau)_r$, $(r = 1, 2, 3, \ldots)$ we can replace the continuum of states by pseudo-discrete states (§ 8) with occupation factors $j_r = (j\,(X)\,d\tau)_r$. Comparing (33) and (28) we see that—

The scale-free condition makes H° a homogeneous function of degree $-l/k$ of the pseudo-discrete occupation factors.

In particular, if the whole probability is concentrated in one pseudo-discrete state, e.g. if the system is in a state of almost exact rest, we have

$$H^\circ \propto j^{-l/k} .$$

This shows the importance of distinguishing between observables and stabilised characteristics. For, if stabilising conditions are imposed on the classifying characteristics X_α, the number of dimensions k of the phase space is reduced.

If the additive characteristic H°, whose partition is being considered, is itself one of the classifying characteristics, we have $l = 1$; and (33) becomes

$$H^\circ = -\,k\,E^\circ \qquad W^\circ = -\,(k + 1)\,E^\circ \tag{34}$$

11. 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* (consisting of a linear momentum 4-vector and an angular momentum 6-vector) with 10 independent components. The corresponding *complete energy tensor* has 136 independent components.[10]

There are in addition 6 dormant components of the momentum vector and 120 dormant components of the energy tensor (making the totals 16 and 256), which are distinguished because they change sign when left-handed axes are substituted for right-handed. These are suppressed because the standard physical reference frame (unlike a geometrical

[10] *Protons and Electrons*, §§ 5·3, 6·4, 10·6.

coordinate frame) is without chirality. Chiral characteristics are electromagnetic, and the standard uranoid is stabilised as electrically neutral.

The term "particle" survives in quantum theory, but very little of its classical meaning is retained. It is now best described as *the conceptual carrier of a set of variates.* The simplest kind of particle in scale-free physics is the carrier of an element of complete energy tensor and nothing more. If the energy tensor is unspecialised, i.e. not limited by stabilising conditions, the particle will be called a *standard carrier* or V_{136} particle.

An ordinary energy tensor is separated into

$$T'^{\mu\nu} = \rho_0 v^\mu v^\nu + s^{\mu\nu} \qquad (35)$$

where ρ_0 is the proper density and v^μ the velocity vector of the molar 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

$$T'^{\mu\nu} = v^\mu \sqrt{\rho_0} \cdot v^\nu \sqrt{\rho_0} \qquad (36)$$

We shall call $v_\mu \sqrt{\rho_0}$ the *root vector.* In natural units the root vector has the dimensions of a momentum vector, and it can be made to agree with the recognised momentum vector by suitable choice of normalisation volume. We shall, therefore, generally refer to it as the momentum vector.[11]

The mechanical characteristics of a classical particle are completely specified by a momentum vector, so that its energy tensor is of the restricted type (36). But in quantum theory there is no imperative reason for excluding particles containing an internal stress-system, more especially as quantum particles (electrons and protons) admittedly contain an internal angular momentum or spin. Thus we have a choice between particles characterised by an unrestricted complete 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; the latter will be called vector particles (V_{10} particles). *Vector particles are obtained 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 obtain the V_4 and V_3 particles (spinless particles) mentioned in § 7. But there is an important difference. A general energy tensor can always be represented as the sum of the outer squares of a number of vectors; so that by substituting V_{10} for

[11] There are important distinctions between the root vector and the recognised momentum vector, e.g. the root vector has ambiguous sign, but they do not concern us here.

V_{136} particles we do not lose generality. They correspond to different modes of microscopic analysis of the same molar energy tensor. But if we substitute spinless V_4 or V_3 particles, we lose the possibility of representing *vorticity*; so that the analysis refers to molar material subjected to a constraint which does not exist naturally. The V_3 and V_4 particles are, therefore, recognised to be *fictitious particles.* The V_{10} particles are *actual* particles.[12] So also are the V_{136} particles; but, since they include an internal stress system, they are from the ordinary classical outlook regarded as composite. The V_{136} particles will be identified later with hydrogen atoms.

The components $T_{\mu\nu}$ of the energy tensor must be taken as the classifying characteristics of the states of these particles, since no other characteristics exist. Thus the total energy tensor is partitioned into a particle energy tensor $E_{\mu\nu}$ and a field energy tensor $W_{\mu\nu}$ by the formula (34). We have

$$T_{\mu\nu} = -\,k\,E_{\mu\nu} \qquad W_{\mu\nu} = -\,(k+1)\,E_{\mu\nu} \tag{37}$$

The factor k, which represents the number of independent components of the energy tensor, will be called the *multiplicity factor.*

To apply the 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 (37). Since $W_{\mu\nu}$ is unaltered by transitions, we have

$$\left.\begin{aligned} (T_{\mu\nu})_0 &= -\,k\,(E_{\mu\nu})_0 && \text{(initial energy)} \\ \delta T_{\mu\nu} &= \delta E_{\mu\nu} && \text{(transition energy)} \end{aligned}\right\} \tag{38}$$

It is understood that the particle is an unidentified member of a large assemblage of which all but a small proportion remain in the initial state.

The classification of states was determined by $T_{\mu\nu}$, but it is now more convenient to use the equivalent classifying characteristics

$$X_{\mu\nu} = -\,T_{\mu\nu}/k \tag{39}$$

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

$$E_{\mu\nu} = (X_{\mu\nu})_0 - k\,\delta X_{\mu\nu} \tag{40}$$

We call $X_{\mu\nu}$ the *generic energy tensor.* Like $T_{\mu\nu}$ (to which it is constantly related), it has the ordinary Lorentz-invariant properties; that is to say, a change of velocity of the particle produces in $X_{\mu\nu}$ the same alteration as an opposite change of velocity of the reference frame. The generic energy tensor is therefore the energy which we should expect the particle to have by the ordinary rules of kinematical calculation, disregarding the presence of a rigid field. Lorentz-invariance does not apply to $E_{\mu\nu}$, and $W_{\mu\nu}$, the latter being unaffected by a change of velocity

[12] Subject to adaptations which may be found necessary when electric charge is taken into account.

(transition) of the particle, but undergoing the usual tensor transformation if the velocity of the frame of reference is changed. Thus the meaning of (40) is that, owing to the rigid field, the transition energy is $-k$ times as much as we should have expected.

The treatment is simplified if we analyse systems in such a way that the initial energy and transition energy are associated with different particles (carriers). We then 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 being 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 classical mechanics and astronomy for systems of more than one particle; and it was followed in the earlier developments of quantum theory, though later writers have unwisely abandoned it. The system is replaced by an *external particle* moving with the centre of mass, and having the total mass of the system, together with internal particles describing relative orbits and having suitable "reduced masses." The rest energy of an internal particle is zero. The external and internal particles correspond respectively to initial and transition particles.

12. Mass-ratio of the proton and electron.

Consider a standard carrier of mass m_0 in an initial state of rest. If it makes a transition to a state of momentum p_1', p_2', p_3', the generic energy (i.e. the expected energy) is, to the ordinary approximation,[13]

$$X = m_0 + \frac{p_1'^2 + p_2'^2 + p_3'^2}{2m_0} \tag{41}$$

so that by (40) the particle energy is

$$E = m_0 - \frac{p_1'^2 + p_2'^2 + p_3'^2}{2\mu} \tag{42}$$

where

$$\mu = m_0/k = m_0/136 . \tag{43}$$

To preserve a formal analogy between rigid-field dynamics and classical dynamics, we adopt in quantum theory a momentum which is i times the classical momentum[14]

$$p_1, p_2, p_3 = ip_1', ip_2', ip_3' \tag{44}$$

[13] We have no occasion to consider large values of p_1', p_2', p_3'. It would not be a better approximation to adopt $X = (m_0^2 + p_1'^2 + p_2'^2 + p_3'^2)^{\frac{1}{2}}$, since the Lorentz formula is not valid in a gravitational field.

[14] Further explanation of this step is given in § 13.

so that by (42)

$$E = m_0 + (p_1^2 + p_2^2 + p_3^3)/2\mu \ . \qquad (45)$$

We have already remarked that the standard carrier will from the ordinary point of view be regarded as composite, since it can contain an internal stress. The deviation from the classical conception of a particle is further indicated by the fact that its *rest mass* m_0 differs from its *mass-constant* μ. We therefore divide it into two carriers, an external particle carrying the initial energy and an internal particle carrying the transition energy, if any. Thus (45) is separated into

$$E_e = m_0 \qquad E_i = (p_1^2 + p_2^2 + p_3^2)/2\mu \qquad (46)$$

Further, since the ordinary outlook admits only vector particles, we must, along with the separation, introduce stabilisation reducing the multiplicity from 136 to 10.

Consider a distribution of n particles per unit volume at rest, and let the total density T_{44} be given. Then by (37) the density apportioned to the particles is $\rho = E_{44} = -T_{44}/k$. If k is changed by stabilisation from k_1 to k_2, the corresponding densities are in the ratio $\rho_1/\rho_2 = k_2/k_1$. If the masses are m_1, m_2, we have $\rho_1 = n m_1$; $\rho_2 = n m_2$; so that

$$m_1/m_2 = k_2/k_1 \ . \qquad (47)$$

The formula (47), being based on (37), applies only to the external particles, which represent the initial part of T_{44}. The transition part of T_{44} is by (38) wholly particle energy, so that the transition part of E_{44} is independent of changes of k. Hence when we apply stabilisation to the whole energy of the standard carrier, E_e is changed in inverse ratio to k, and E_i is unchanged; so that (46) becomes

$$E_e = \frac{136}{10} m_0 \qquad E_i = \frac{p_1^2 + p_2^2 + p_3^2}{2\mu} \qquad (48)$$

Denoting the mass (rest energy) of an external V_{10} particle by M, we have

$$M = \frac{136}{10} m_0 \qquad \mu = \frac{1}{136} m_0 \qquad (49)$$

The internal V_{10} particle has no rest energy; but its mass (mass-constant) is defined as the coefficient μ in the expression (48) for the energy, just as in classical mechanics.

Presumably this combination of an external and an internal particle of the simplest kind (V_{10} particles) is realised in a hydrogen atom.

Accepting this identification, which is checked in a great many ways in subsequent developments of the theory, the ratio

$$\eta_1 = \frac{M}{\mu} = \frac{136^2}{10} = 1849{\cdot}6 \qquad (50)$$

is a fundamental constant of the hydrogen atom which can be determined observationally. It is here found primarily as a density-ratio in an assemblage of standard carriers; but, since the division of the standard carriers gives one internal particle to each external particle, it can also be interpreted as a mass-ratio for individual atoms.

If the masses of a proton and electron are m_p, m_e the masses of the corresponding external and internal particles are given by

$$M = m_p + m_e \qquad \mu = m_p m_e/(m_p + m_e) \qquad (51)$$

so that m_p, m_e are the roots of

$$m^2 - M m + M\mu = 0 \qquad (52)$$

or by (49)

$$10\, m^2 - 136\, m\, m_0 + m_0^2 = 0 \qquad (53)$$

which is the equation found by the author in 1931.[15] From this we obtain the mass-ratio

$$\eta_2 = \frac{m_p}{m_e} = 1847{\cdot}60 \qquad (54)$$

1836.2

We call m_p, m_e defined in this way the *standard masses* of the proton and electron. When we pass over to electrical theory (§ 17) modifications of the definition will suggest themselves, and the ratio may be slightly changed. But there is no very obvious definition of the mass of an electron; and in any comparison with an observational determination of m_p/m_e the question arises, What precisely has the observer determined? Having ascertained this, we have to re-reduce his results so as to correspond to the definition adopted here or in later theoretical formulae. The full comparison is given in § 20, and the agreement is found to be perfect.

13. The inversion of quantum energy.

In molar relativity theory, as in classical theory, matter is supposed to be composed of particles to which the uncertainty principle does not apply. Such particles can be included in the present theory as a special case $k = 1$. This implies that the velocity (represented by a unit vector in four dimensions) is treated as free information, not involving any reciprocal uncertainty of position; and the only mechanical

[15] *Proc. Roy Soc.*, A, 134, 532.

characteristic which is a genuine observable is the length m of the momentum vector. This is consistent with the usual elementary relativistic outlook, in which m appears as an absolute quantity, whereas the orientation of the momentum vector is relative to an unobservable space-time frame. V_1 particles, therefore, resemble classical particles in being exempt from the usual uncertainty conditions; but, as quantum particles, they are superpositions on a rigid environment, unlike classical particles, whose essential characteristic is that they carry fields of force which disturb the environment. We, therefore, describe the V_1 particles as *semi-classical.* By the definition of generic energy, $X_{\mu\nu}$ is the particle energy of a classical particle; if $E_{\mu\nu}$ is the particle energy of a V_1 quantum particle; we have in a transition

$$\delta E_{\mu\nu} = - \delta X_{\mu\nu} \qquad (55)$$

by (40). The change of sign of the transition energy will be referred to as the *inversion of energy.*

The inversion of energy can be understood by considering a system of gravitating particles in a steady state, e.g. a star cluster. The initial state is taken to be that in which the stars are nearly at rest, and the cluster is, therefore, widely extended. If the cluster contracts to a new steady state, the kinetic energy K is increased, but the whole energy $T = K + V$ is diminished, V being the potential energy. The condition for a steady state is $T = - K$. Treating the cluster as a classical system, the gain of particle energy is $\delta X = K$; but, treating it as a system of quantum particles superposed on a rigid field, the whole change of energy must be represented as a change of particle energy, so that the gain of particle energy is $\delta E = T = - K$.

The mass or rest energy of a body, as defined in molar physics, is not a net addition to the energy of the universe. It is an addition made in a particular region; but elsewhere there is a decrease, which is recognised as the negative potential energy of surrounding objects due to the gravitational field emanating from the body. This distinction, between the direct change of energy due to the presence of a particle and the indirect change due to its field, can only be made if the particle is localised. Quantum particles are not localised, and only the total energy change in any region can be considered. The non-localisation is accentuated by the fact that in mechanical theory the particles are represented by "infinite" plane waves.

We generally confine attention to steady states, because a quantum particle is defined as the occupant of an eigenstate; but a brief reference may be made to the relativistic theory of unsteady states. In general relativity theory the mass M of a *system* is defined as follows. Taking axes such that the centre of mass is instantaneously at rest, M 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.[16] It is found that for a system of gravitating particles[17]

$$M = M_0 + 3K + 2V \qquad (56)$$

where M_0 is the sum of the rest masses, K the kinetic energy and V the internal potential energy. Using the identity[18]

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

where C is the moment of inertia of the system about the centre of mass, (56) is reduced to either of the two forms

$$M = M_0 + K + V + \tfrac{1}{2} d^2C/dt^2 \qquad (57\,a)$$

$$M = M_0 - K + d^2C/dt^2 \ . \qquad (57\,b)$$

If the system is steady, (57b) gives $M = M_0 - K$, the negative sign indicating the inversion of energy. For an unsteady system there is also an expansion energy $\frac{d^2C}{dt^2}$. In wave mechanics the expansion would be represented as a series of transitions between steady states; and the expansion energy is then the perturbing term in the Hamiltonian, which induces the transitions.

We have seen that, in order to compensate the inversion of sign of transition energy, a quantum momentum p_α ($\alpha = 1, 2, 3$) is introduced which is i times the classical momentum p_α'. This is the origin of the $\sqrt{-1}$, which appears so mysteriously in quantum formulae. The classical momentum operator would be $p_\alpha' = -\hbar\,\partial/\partial x_\alpha$.

Both in classical and quantum theory we deal with momenta which are real; so that effectively there is no overlapping of classical and quantum systems, except when the momentum is zero. This is illustrated in equations (41) to (45), where we proceed as though we were going to give the carrier m_0 an *external* velocity p_1'/m_0, p_2'/m_0, p_3'/m_0; but it turns out that our analysis, as applied to a *real* rigid-field problem, refers to a system in which the carrier remains externally at rest and has a steady *internal* motion with momenta p_1, p_2, p_3. The fact is that the mechanical quantities in quantum theory are analogues, not necessarily direct representatives, of the correspondingly named classical quantities; and in some cases the direct representatives are imaginary.

14. Rigid coordinates.

In order to validate the rigid-field treatment adopted in wave mechanics we have to put the system into an extravagantly strong gravitational field.

[16] If M is varying, the time to which M refers is earlier than the time of comparison of geodesics by the amount of the light-time.

[17] Eddington and Clark, *Proc. Roy. Soc.*, A, 166, 469.

[18] *Mon. Not. R. Astr. Soc.*, 76, 525.

The standard particle, for example, is put into a field which gives it a potential energy – 137 times its own mass. If the gravitational field were "transverse," i.e. irreducible like the field of the sun, the results obtained in these fantastic conditions would have no practical application. We are, therefore, limited to initial distributions for which the self-consistent gravitational field is "longitudinal," i.e. creatable and removable by a coordinate transformation.

The basis of our investigation has been that wave mechanics, which introduces particles as occupants of states in a rigid field, is only valid if the characteristics of the field are stationary for small changes of the occupation factors. We can now elucidate the proviso by making it read, "*if the coordinates are chosen* so that the characteristics of the field are stationary . . ." Such coordinates will be called *rigid coordinates.*

We can determine the rigid coordinates for a uniform distribution of particles at rest, which is the initial state usually adopted. The required transformation from Galilean coordinates x', y', z', t' to rigid coordinates x, y, z, t is

$$x' = x \qquad y' = y \qquad z' = z \qquad t' = -kt \tag{58}$$

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

$$g_{44} = k^2 \qquad g^{44} = k^{-2} \qquad \sqrt{-g} = -k.$$

By general relativity theory[19] the energy and momentum per unit coordinate mesh (which by (58) agrees with the space volume in rigid as well as Galilean coordinates) are $\mathfrak{T}_\mu{}^4 = T_\mu{}^4\sqrt{-g}$. Normally a pseudo-tensor density $t_\mu{}^4$ is added, but this vanishes here because the $g_{\mu\nu}$ are constants. The transformation (58) gives

$$\mathfrak{T}_1{}^4 = \mathfrak{T}'_1{}^4 \qquad \mathfrak{T}_4{}^4 = -k\mathfrak{T}'_4{}^4 . \tag{59}$$

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 are transformed in the same way as $\mathfrak{T}_\mu{}^4$. This makes it necessary to represent the momentum and energy of a particle by a covariant vector p_μ; for the transformation (58) gives $p_1 = p_1'$, $p_4 = -kp_4'$. This agrees with the quantum definition of the momentum vector by the covariant operator $p_\mu = -i\hbar\partial/\partial x_\mu$.

To create a field by coordinate transformation we must, as it were, mistake the transformed coordinates for Galilean coordinates; so that the change of $\mathfrak{T}_\mu{}^4$ and p_μ is attributed to additional field energy and momentum. Thus, adapting our former notation,

$$\mathfrak{W}_\mu^\nu = \mathfrak{T}_\mu^\nu - \mathfrak{T}_\mu^{\prime\nu}; \qquad \mathfrak{E}_\mu^\nu = \mathfrak{T}_\mu^{\prime\nu} .$$

---

[19] Eddington, *Mathematical Theory of Relativity*, § 59.

Since we are considering particles at rest, the only non-zero component is $\mathfrak{T}_4^4$; and by (59) the condition

$$\mathfrak{T}_\mu^\nu = -k\mathfrak{E}_\mu^\nu \qquad \mathfrak{W}_\mu^\nu = -(k+1)\mathfrak{E}_\mu^\nu$$

is satisfied. This (in rather more elaborate notation) is the condition (37).

The rigid-field condition, therefore, in practice reduces to this:—*in wave mechanics we must employ a time* t *which is* 1/k *times the true time* t'.[20]

This is, of course, covered up in current quantum theory by empirical adjustments of the constants. But our aim here is to develop the theory without empirical constants.

15. Mutual and self energy.

By observing an object-body in conjunction with a reference body, we determine 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 as self characteristics, or more defensibly to partition them between the two bodies according to some logical plan. This conceptual transfer, by which self properties are substituted for mutual properties, is a habit of thought which has been elevated into a convention. Since the language of physics is bound up with this convention, 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, is called a *comparison particle.* Let the particles be of the simple type for which the mechanical characteristics are completely specified by a momentum vector. If the momentum vectors are p_μ, p_μ' the mutual energy tensor is necessarily of the form

$$M_{\mu\nu} = \tfrac{1}{2}C(p_\mu p_\nu' + p_\nu p_\mu') \qquad (60)$$

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

$$T_{\mu\nu} = Ap_\mu p_\nu \qquad T_{\mu\nu}' = A'p_\mu' p_\nu'. \qquad (61)$$

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

$$\rho_m = Cmm' \qquad \rho = Am^2 \qquad \rho' = A'm'^2. \qquad (62)$$

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

$$Am^2 = Cmm'. \qquad (63)$$

[20] This is additional to the inversion of energy, which gets rid of the negative sign.

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 when the measurement is made, but assuming its absence when the measurement is interpreted in physical terms. The correct relation $\rho + \rho' = \rho_m$ gives

$$A m^2 + A' m'^2 = C m m'. \qquad (64)$$

We have already found that the masses m of the elementary V_{10} particles (protons and electrons) fulfil a relation (53) which is of this form. By comparison of (64) and (53), we obtain

$$m' = m_0 \qquad A : A' : C = 10 : 1 : 136. \qquad (65)$$

The factors 10, 1, 136 are the multiplicity factors of the respective carriers. The distribution of unspecialised energy $M_{\mu\nu}$ of multiplicity 136 is replaced by two distributions of specialised energy, one being the energy of object-particles of multiplicity 10 and the other the energy of comparison particles. By assigning to a comparison particle the multiplicity 1, we express the fact that it is employed as a standard, and, therefore, idealised as a carrier only of that characteristic of which it is the standard.

We have, therefore, the following result:—For *particle* energy the three coefficients are equal, and the densities are simply

$$A' m^2, \qquad A' m'^2, \qquad A' m m'. \qquad (66)$$

The relation $\rho + \rho' = \rho_m$ applies to the *total* energies; and the quadratic equation giving the mass of the proton or electron expresses the partition of the total mutual density $136\, A'm\,m'$ into total self densities $10\, A'm^2$, $A'm'^2$.

The root of the problem is that *probabilities are multiplicative.* In wave mechanics we deal, not with particles, but with probabilities of particles (occupation factors). Thus a system appears primarily as the product, rather than the sum, of its parts. When wave mechanics gets fairly started, this condition is met by consigning the ordinary additive characteristics to the exponents of wave functions; so that they get added when the wave functions are multiplied. But we are here occupied by this condition as it affects the beginning of wave mechanics. There can be no observable without a double probability distribution; so that the most elementary mechanical observable is a product of two parts m, m' or more generally p_μ, p'_μ. We have to reconcile this with the common conception of the composition of mechanical characteristics by addition of parts.

The same method gives the mass μ of an internal particle. In this case $\rho = 0$, since the internal particle has zero rest energy. Hence $\rho' = \rho_m$, or

$$A' m'^2 = C \mu m' = 136\, A' \mu m'. \qquad (67\,a)$$

Hence $\mu = m'/136$, which agrees with (49), m' having been identified with m_0 by (65).

The value $M = 136\,m_0/10$ of the mass of the external particle gives $10\,A'M^2 = 136\,A'Mm'$, or

$$AM^2 = CMm'. \qquad (67\,b)$$

This corresponds to the crude formula (63), which makes no provision for the density of the comparison particle. It appears then that, when a hydrogen atom is analysed into external and internal particles, the comparison particle associated with the internal particle serves for both; but when it is analysed into a proton and electron two comparison particles are present. Presumably this is connected with the fact that the proton and electron can be taken apart, but the external and internal particle cannot be dissociated from one another. We can think of a proton without an electron, but not of an inside without an outside. The "decoupling" of the proton and electron will be treated in § 28.

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

$$\left.\begin{aligned} 10\,A_0m^2 + A_0m_0^2 &= 136\,A_0mm_0 && \text{(proton or electron)}\\ A_0m_0^2 &= 136\,A_0\mu m_0 && \text{(internal particle)}\\ 10\,A_0M^2 &= 136\,A_0Mm_0 && \text{(external particle).}\end{aligned}\right\} \quad (68)$$

It is significant that the particles which have comparison particles attached to them (proton, electron, internal) have electrical characteristics, whereas the external particle is neutral.

PART III.—ELECTRICAL THEORY.

16. Interchange of comparison particles.

A measurement is an operation in which four entities are physically involved; for the observable to be measured is a relation between two entities A_1, A_2, and the measure is its ratio to a comparison observable which is the corresponding relation between another pair of entities A'_1, A'_2. Accordingly—

A measure is a statistic of a quadruple probability distribution.

In a casual measurement the form of the quadruple distribution is unrestricted; but the foundation of scientific investigation is the substitution of systematic for casual measurement. In particular, the definitions of physical quantities are based on an ideal systematic plan of measurement. In systematic measurement one or more of the four entities is an average of a large number of particles, so that its variance is standardised. The quadruple distribution has then a simplified form, the simplification being unsymmetrical as regards A_1, A_2, A'_1, A'_2.

Normally two entities A'_1, A'_2 are selected as the comparison system and standardised by averaging. Their sole function is to provide the comparison standard (i.e. the extraneous standard materialised in a form in which it can be imagined to take part in the physical process of measurement). Thus the comparison system is so much simplified by averaging that it retains only one dimension of probability distribution and is the carrier of only one variate, namely, the scale.[21] In short, it reduces to a V_1 particle.

The entities A_1, A_2 are treated differently, according as we are studying distribution or correlation. In distribution theory A_1 is selected as object-particle and A_2, standardised by averaging, is the reference particle or physical origin. In correlation theory A_1, A_2 form the object-system, and the quantity measured is admittedly a relation between A_1 and A_2 in which both play a symmetrical (or anti-symmetrical) part. In that case we introduce an internal particle as the carrier of the relative coordinates and momenta; and the scale is then correspondingly associated with the internal particle of the system A_1', A_2' standardised by averaging.

The ordinary outlook regards a single particle as a carrier of mass. This, as we have seen, involves a conceptual partition of mutual properties into self properties. The self mass of A_1 may be arrived at by two routes. Either we measure the mutual energy of A_1, A_2 by comparison with the mutual energy of A'_1, A'_2, and partition the result between A_1 and A_2; or we partition the mutual energy of A_1, A_2, between A_1 and A_2, and measure the part allotted to A_1 by comparison with the correspondingly obtained energy of A'_1. These correspond to the two methods by which we have obtained and interpreted the equation (53) for the masses of protons and electrons, since the method of § 12 was based on treatment of the system $A_1\ A_2$, and the method of § 15 on treatment of the system $A_1\ A'_1$. Both presuppose a preliminary simplification or stabilisation which has reduced the quadruple to a double probability distribution. We have now to examine this preliminary step, so as to rescue what has been lost in assuming it.

A system such as A_1, A'_1 consisting of an object-particle and a comparison particle will be called a *perfect particle*. We consider a measurable, denoted by $[A_1\ A_1'\ A_2\ A_2']$, furnished by a perfect electron and a perfect proton. The transformation

$$[A_1 A_1' A_2 A_2'] \rightarrow [A_1 A_2' A_2 A_1'] \qquad (69)$$

in which the two perfect particles exchange comparison particles, will be called *interchange*.

---

[21] Regarding the scale as a momentum, it has a conjugate coordinate which we call the *phase coordinate*.

Since comparison particles are alike, interchange has no observable effect. All observable results must be invariant for interchange. Interchange is, therefore, a relativistic transformation of the system of reference, which is to be treated in the same way as relativistic rotations of the space-time frame.

In treating a double probability distribution, it is necessary to discriminate between two kinds of distribution function or wave function—

(*a*) Expressing the probability that two points x, x' are occupied.

(*b*) Expressing the probability that particle No. 1 is at x and No. 2 at x'.

So far as I am aware, no mathematical theory of wave functions of type (*a*) exists, and treatment is always based on functions of type (*b*). This involves the assignment to the particles of identification suffixes which have no observable counterpart, the phenomena being invariant for interchanges of the suffixes. We may assimilate this invariance for interchange of suffixes to the afore-mentioned invariance for interchange of comparison particles by making the comparison particles the carriers of the identification suffixes.

Thus in a perfect particle the object-particle is the carrier of the coordinates and their conjugate momenta, and the comparison particle is the carrier of the scale and the suffix. We have found (§ 15) that when a proton and electron are brought together into one system, so that they are replaced by an external and internal particle, one of the comparison particles is eliminated. This is true in so far as the comparison particles are carriers only of the scale, since it would be redundant to have two versions of the extraneous standard associated with one system. But since they carry the suffixes, it is necessary to retain a permutation variate whose changes represent the transformation (69), which would otherwise be lost. Either we replace the two comparison particles by an external comparison particle carrying the scale and an internal comparison particle carrying the permutation variate, or we combine them into a V_2 comparison particle with two degrees of freedom corresponding to the scale and the permutation variate.

The permutation variate is most conveniently taken to be an angular coordinate θ, such that the transformation (69) corresponds to $\theta \rightarrow \theta + \pi$. Since the coordinate itself is unobservable, it is a cyclic coordinate which can be eliminated by ignoration of coordinates; but the angular momentum conjugate to it constitutes an interchange energy which must be taken into account in the Hamiltonian of the system. I pointed out in 1928[22] that this interchange energy is the Coulomb energy

[22] *Proc. Roy. Soc.*, A, 122, 358.

of the proton and electron. It will be shown in §§ 19, 20 that the exact value of the Coulomb energy can be found in this way.

It seems to be commonly thought that the difference of mass of a proton and electron makes them "distinguishable," so that the theory of interchange does not apply to them. But this is a rudimentary fallacy. Clearly we must distinguish the individual particles *before* the masses can be measured, so that the mass cannot be used as a criterion of distinction. This applies even in classical mechanics. If $x(t)$, $x'(t)$ are the coordinates of two occupied points observed at two times t_1, t_2, the velocity of the particle originally occupying $x(t_1)$ may be either

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

Since the velocity is observationally ambiguous, the whole dynamics is ambiguous; and the mass of the particle is undefined. The only way to proceed is to assume identification numbers (at both times) as free information. Then the velocity, mass, etc., can be deduced. We may then, if we like, employ the distinctive masses to identify the particles; but this can only reproduce the distinction already assumed as free information, on the basis of which the masses were determined.

Let $(x, 1; x', 2)$ denote the state in which particle No. 1 is at x and No. 2 at x'. A passage to the state $(x, 2; x' 1)$ may occur as the result of motion through space, i.e. by a coordinate interchange; or it may occur as the result of an increase π of the permutation coordinate giving the transformation (69), i.e. by suffix interchange. The term "interchange," unless otherwise stated, refers to suffix interchange. The important point is that the equation of continuity of flow of probability is not satisfied if we take into account only the flow through space; and the mechanical equations must include the extra-spatial flow representing suffix interchange.

17. Electric energy.

The standard carrier is a V_{136} particle, which carries an energy tensor and nothing more. We have now to consider a particle which carries a permutation variate in addition; so that when it is divided into two particle distributions the particles are distinctively suffixed. This will be called a *bi-particle*. There are two points of view. The additional degree of freedom can be provided by extending the complete energy tensor so as to include one of the "dormant" electric components (§ 11). The bi-particle then appears as a V_{137} particle; and by (47) the relation between the masses m_0, m_0' of the standard carrier and bi-particle is

$$m_0 = \frac{137}{136} m_0' = m_0' + \mu' \qquad (70)$$

where

$$\mu' = m_0' / 136. \qquad (71)$$

But, although completely unified treatment of gravitational and electrical energy is sometimes useful, it is ordinarily held in reserve, since our subject requires more intensive treatment of electrical than of gravitational phenomena. We shall, therefore, not admit the electric or permutation degree of freedom on the same footing as the rest. Both standard particles and bi-particles will be treated as V_{136} particles, the electrical properties of the latter being taken account of by inserting special correcting terms in the Hamiltonian. The principal correction is the Coulomb interchange energy. But, before this is inserted, we have to remove an averaged electric energy, which is present in the adopted initial state in our scale-free investigations, and included in m_0 as though it were mechanical energy.

There is an obvious reason for the difference of mass in (70). In the initial distribution of standard particles at rest the occupied points are at rest, but there is a continual interchange circulation of the suffixed particles which occupy them. Thus standard particles at rest are equivalent to bi-particles in (extra-spatial) motion; and by (70) it follows that μ' is the energy of this motion. Thus in passing from the standard particle to the bi-particle representation of the same distribution, there is a reduction μ' in the initial energy and an increase μ' in the transition energy; so that (46) becomes

$$E_e = m_0 - \mu' = \frac{136}{137} m_0 \qquad E_i = \mu' + (p_1^2 + p_2^2 + p_3^2)/2\mu. \quad (72)$$

In a Hamiltonian of the form E_i we distinguish μ' as the rest mass and μ as the mass-constant. When they agree the particle can be described as *free*. It is a feature of the new form (72) that the internal particle is very nearly free. The reason why it is not quite free is that the gravitational equations are non-linear; so that the separation of E into two additive terms does not coincide with the physical separation into two systems which we can think of as existing separately. In the next section we shall follow the physical separation by treating E_e and E_i as at first existing in separate worlds; it will then appear that the difference $\mu - \mu' = \mu/137$ is introduced in combining the two worlds.

When the rest mass and the mass-constant differ, the ordinary interpretation is that the mass-constant is the true inertia, and the rest mass includes in addition either gravitational or electrical potential energy. "Longitudinal" electric energy can be treated indifferently as electrical or mechanical. If we count it as mechanical, the longitudinal electric energy in the rest mass is counted as true inertia, and is, therefore, included in the mass-constant. If we count it as electrical, it is not included in the true inertia; the rest mass is unchanged but the mass constant is reduced. Naturally before introducing electrical theory we include as much as possible of the electrical energy in the mechanical

energy. Consequently the standard particle (an element of the mechanical energy tensor) includes electric energy in the inertia or mass-constant, and m_0 is the mass-constant as well as rest mass, as assumed in (41). When we introduce electrical theory, the mechanical energy is more narrowly defined. The energy m_0 is split into the pure mechanical energy of a bi-particle at rest both spatially and extra-spatially, together with the electrical energy of its extra-spatial motion in the state in which it is equivalent to a standard particle. The inertia or mass-constant is then the pure mechanical energy m'_0. Let

$$\beta = 137/136. \tag{73}$$

Then for a bi-particle in the initial state in which it is equivalent to a standard particle at rest

$$\text{mass-constant } (m_0') = \text{rest mass } (m_0) \div \beta. \tag{74}$$

Both standard particles and bi-particles are being treated as V_{136} particles, and the stabilisation which reduces k from 136 to 10 applies to both. Thus for a hydrogen atom, when the electrical energy is distinguished from the pure mechanical energy.

$$\text{mass-constant } (M') = \text{rest mass } (M) \div \beta. \tag{75}$$

"Transverse" electric energy cannot be included in the mechanical energy; but it only becomes important when the momenta are large. For that reason we have hitherto considered only small values of p_1, p_2, p_3, deferring the higher approximation until bi-particles are reached and the whole electric energy can be treated together.

The formulae

$$M = m + m' \qquad \mu = mm'/(m + m') \tag{76}$$

refer to mass-constants.[23] On the other hand, in the formula $M/\mu = 136^2/10$, obtained in the treatment of standard particles in § 12, M refers to the rest mass. It would be troublesome to change the notation in (76); so we shall denote the mass-constant and the rest mass by M and $M\beta$. Then $M = 136\,m'_0/10$, $M\beta/\mu = 136^2/10$; and the quadratic equation for the masses m, m' of the suffixed particles is

$$10\,m^2 - 136\,mm'_0 + \beta m'^2_0 = 0. \tag{77}$$

This gives a mass-ratio $m_p/m_e = 1834{\cdot}1$. The difference from the former result $1847{\cdot}6$, which applies to unsuffixed particles, is that the masses of the proton and electron are now freed from electric energy; formerly they included the electrical energy of the simple (but highly

[23] This appears formally in (79) below. But in any case it is clear that the formula (76) cannot refer to rest masses; since in its ordinary use in classical mechanics the rest mass of the internal particle is zero.

unnatural) distribution adopted as initial state.[24] Having cleared away initial electric energy, we can make a clean start in calculating and inserting the Coulomb interchange energy of the particular configurations we wish to treat.

The result (77) was obtained in *Protons and Electrons,* equation (15 . 84). But we shall see in the next section that it is not yet final. The trouble is that in removing the electrical energy from the system, so as to obtain purely mechanical masses of the proton and electron, we upset the self-consistent gravitational field which was adjusted when the electrical energy was present. We have to devise a treatment which adjusts the field after the electric energy has been removed. The fact that the internal particle in (72) is not completely freed is a symptom of this trouble. We shall, therefore, first complete the freeing of the internal particle.

The familiar classical formula (76) ought not to be taken over into relativity theory or quantum theory without re-derivation. In wave mechanics it originates as follows:—Let x_α, x_α', X_α, ξ_α ($\alpha = 1, 2, 3, 4$) be the coordinates and times of the four particles; then

$$X_\alpha = (m x_\alpha + m' x_\alpha') / (m + m') \qquad \xi_\alpha = x_\alpha' - x_\alpha , \tag{78}$$

and, defining M and μ by (76), we immediately derive the identity

$$\frac{\Box_x}{m} + \frac{\Box_{x'}}{m'} = \frac{\Box_X}{M} + \frac{\Box_\xi}{\mu} \tag{79}$$

where $\Box$ is the usual wave operator $\partial^2/\partial t^2 - \nabla^2$ for the set of coordinates indicated by its suffix.

For a free particle we shall now use the relativistic Hamiltonian, valid for all values of the momentum, namely,

$$p_4 = (m^2 + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}} , \tag{80}$$

and the wave equation of a free particle is then

$$-\hbar^2 \Box = m^2 . \tag{81}$$

It is not possible for all four particles to be free, because (79) would then give $m + m' = M + \mu$, which is untrue. We take the proton and electron, to be free, because it is their free masses which we wish to determine. We further decide to make the internal particle a free particle; so that only the external particle is abnormal.

The reason for the last decision is that the internal particles (correlation wave functions) have to bear the brunt of all the important

[24] It is unnatural because in an actual uniform assemblage of protons and electrons the total electrical energy would be zero. By treating it as an assemblage of two-particle systems not interacting with one another, we omit the positive energy which in fact cancels the internal negative energy of the proton-electron pairs.

practical developments of wave mechanics, which are highly complicated; so that it is essential to have the simplest possible concept as a basis. By freeing the internal particle and giving it the normal Hamiltonian of relativistic mechanics, we are able to represent it straightforwardly by Dirac wave functions. External particles (distribution wave functions) are only passengers in the theory. They can be left to be added at the end, although, as we shall see, the addition is not quite so simple as is usually assumed.

It will be seen that we here make a slight change in the separation shown in (72), by letting the correction $\mu' - \mu$ for non-linearity of the combination fall on E_e instead of on E_i.

18. The β-factors.

Thus far we have been content with a mathematical separation of mechanical and electrical energy co-existing in the same space. The detailed development of quantum theory demands a more drastic separation, which virtually puts them into separate worlds. The internal particles, now completely freed and representable by Dirac wave functions from the "electrical world"; and it is in this world that the whole development of quantal physics takes place. Quantal theory—the typical part of quantum theory—is concerned exclusively with electrical energy, for we have seen (§ 8) that the corresponding quantum of mechanical energy has the much higher order of magnitude $\hbar \sqrt{N}$. The mechanical world, so long as it is kept separate from the electrical world, is scale free, except in regard to cosmical scale.

This separation of mechanical and electrical worlds may be regarded as a "dis-unification" of our theory; but, of course, such a step is necessary if we are to correlate unified theory with the existing non-unified development of theoretical physics, which contains an immense amount of technical development that is of permanent value. We have to exhibit non-unified physics as included in unified physics, just as we have to exhibit Newtonian gravitational theory as included in Einstein's theory. But, having seen the genesis of the separate mechanical and electrical worlds in unified theory, we shall be able to combine them rigorously instead of by guesswork. The combination is not linear, but introduces certain β-factors which will be investigated in this section. The β-factors must be taken into account in any comparison with observation, since the world which we observe is the combined world.

The internal particle being now recognised as purely electric, the hydrogen atom is a combination of a mechanical (neutral) external particle and an electric doublet. The coordinate ξ of the internal particle is the separation of the doublet, and it is primarily significant as a measure of electric polarisation. Picturing it, however, as an ordinary coordinate, it locates the electric particle in a ξ-space, which is distinct

from the ordinary x-space in which the proton, electron and external particle (occupying the centre of mass of the electron and proton) are situated. The ξ-space is the space of the electric world. There is a fourth coordinate σ, which formally corresponds to imaginary time, in the electric world; in physically real systems the probability extends over real values of σ—unlike the mechanical world, where the values of the time are required to be real.[25]

We here deal with the electrical world only so far as it is introduced in an assemblage of hydrogen atoms; but a brief reference may be made to more complex systems. In systems of more than two particles the basis of the separation into mechanical and electrical worlds is the separation of symmetrical and antisymmetrical wave functions. This reduces to the separation of external and internal particles in the simple case of a two-particle system. It is commonly said that the wave functions of systems of simple particles are antisymmetrical; but that is merely because current wave mechanics selects the antisymmetric (electric) characteristics as its special field of study. The general theory of the separation has many ramifications, which we do not enter into here. It is sufficient to call attention to a practical point. In practice the external energies of complex particles are compared with one another by mechanical methods (atomic weight determinations), and the internal energies are compared with one another by optical methods (wave-length determinations of the emitted radiation). Thus the observational data follow the theoretical separation into two worlds; and the corrections (β-factors) which we apply in combining the two worlds appear in practice as corrections to be applied in comparing energies measured on the atomic-weight scale with energies measured on the wave-length scale. By considering the hydrogen atom, for which the precise theory can be worked out, we determine once for all the correction necessary to adjust the two scales.

Consider two steady distributions both containing n particles per unit volume spatially at rest, the particles being—

A, standard particles, regarded as bi-particles with electric energy.

B, bi-particles without electric energy.

By relativity theory, only one density of a steady distribution of particles at rest, namely, the density of an Einstein universe, is compatible with the actual constants of nature. All our previous work has been based on *A* as initial state, so that we have implicitly assumed natural constants compatible with it. When the electric energy is removed to obtain *B*, the density is reduced from $n m_0$ to $n m'_0$, and the distribution is no longer in equilibrium with the same natural constants. It would be troublesome to introduce a new set of physical constants

[25] See § 19.

associated with B; and it is, therefore, simpler to couple with the removal of the electric energy a change of extraneous standard such as to multiply densities by the factor β. This restores the density to $n m_0$; so that it is compatible with the original physical constants (the same numerically, though the number refers to a different unit). In natural units length has dimensions (density)$^{-\frac{1}{6}}$ so that the change of standard multiplies lengths by $\beta^{-\frac{1}{6}}$. Thus the effect of removing the electric energy may be stated as follows:—

In passing from A to B there is no change of density, but lengths and times are multiplied in the ratio $\beta^{-\frac{1}{6}}$ (82)

Now A is the initial state in our unified theory which describes the combined world. By the analysis in (78) and (79) we separate off the free electric particles, and consign them to a separate world. The system B which is left forms the purely mechanical world, in which longitudinal electric energy is no longer counted as mechanical energy.

Let x_α as before be coordinates in the combined world, and $\bar{x}^\alpha$ be coordinates in the mechanical world. Let $\xi_\alpha = x_\alpha' - x_\alpha$, $\bar{\xi}_\alpha = \bar{x}_\alpha' - \bar{x}_\alpha$. By (82), $\bar{x}_\alpha = \beta^{-\frac{1}{6}} x_\alpha$, so that $\bar{\xi}_\alpha = \beta^{-\frac{1}{6}} \xi_\alpha$. Thus if we calculate the mass of an internal particle from the wave equation $-\hbar \Box \bar{\xi} = \bar{\mu}^2$, instead of from $-\hbar^2 \Box \xi = \mu^2$, we obtain

$$\bar{\mu} = \beta^{\frac{1}{6}} \mu . \tag{83 a}$$

There is no change in density or in n, so that the rest mass of the external particle is unchanged. But, as pointed out in § 17, the M which appears in (76) is the mass-constant in system A; and the rest mass in either system is

$$\bar{M} = \beta M . \tag{83 b}$$

The analysis of the energy tensor $T_{\mu\nu}$ in §§ 9–12 applies to the combined world or the mechanical world according as we do or do not include longitudinal electric energy in $T_{\mu\nu}$. We are now excluding the electric energy, so that the result (50) is $\bar{M}/\bar{\mu} = 136^2/10$. This applies to system B. Passing to system A (combined world) we then have

$$\frac{M}{\mu} = \beta^{-\frac{5}{6}} \frac{\bar{M}}{\bar{\mu}} = \frac{136^2}{10} \beta^{-\frac{5}{6}} . \tag{84}$$

This transformation from B to A does not mean that we cease to exclude longitudinal electrical energy from the mechanical tensor $T_{\mu\nu}$. (If we re-admitted it, we should get back to standard particles, and obtain $M/\mu = 136^2/10$ as in our original discussion.) The transformation is a re-measurement of the purely mechanical energy; the same mechanical energy, originally measured by the constants $\bar{M}, \bar{\mu}$ in a world containing no electric energy, is re-measured in a world where there exists one free electric particle for each mechanical particle. The re-measure-

ment consists in a change of extraneous standard, which introduces the factors $\beta^{\frac{1}{6}}$, β in (83 a) and (83 b). The masses resulting from (84) represent purely mechanical energy in an environment which is not purely mechanical—which is the condition in which the current dis-unified theory requires us to treat it.

Using (84), the quadratic equation for the masses m_p, m_e, becomes

$$10\,m^2 - 136\,mm_0 + \beta^{\frac{5}{6}} m_0^2 = 0\,, \qquad (85)$$

m_0 being correspondingly re-defined. The resulting mass-ratio is

$$m_p / m_e = 1836{\cdot}34.$$

These masses may be distinguished as the *dis-unified masses* of the proton and electron. We shall adopt this definition of the masses in § 20, where comparison is made with observation.

We shall briefly recapitulate the ideas leading to (84) and (85). The analysis of a system into two suffixed particles, or equivalently into a mechanical external particle and an electrical internal particle or doublet, is necessarily made in the combined world. In quantal theory the electrical part is treated quite separately (which is possible, since the analysis in (79) is arranged to give free electric particles), quantisation of angular momentum is introduced, and for the first time a definite scale of structure is fixed by the discrete unit of angular momentum $\hbar$ which then appears. This scale of structure is imported into the other half of the combined world by the condition of one-to-one correspondence of external and internal particles which the analysis imposes. It is adopted as the normal scale for all purposes, and the constants of nature are referred to it. We turn now to the mechanical half of the combined world. Owing to the omission of the rest masses of the free electrical particles its density is less in the ratio $1/\beta$, and it could not be in self-equilibrium consistently with the scale and physical constants already fixed. But we wish to consider it by itself, because our plan is to treat the mechanical and electrical worlds separately and then put them together—a plan not only conforming to current theoretical procedure but also to observational practice. Having first separated it mathematically, we change the extraneous standard so as to multiply densities by β and lengths by $\beta^{-\frac{1}{6}}$ keeping $\hbar$ nominally unchanged. That puts it into self equilibrium, and we can proceed with investigations according to the steady state theory already developed. In particular, the result $\bar{M}/\bar{\mu} = 136^2/10$ applies. Reverting to the normal scale, $\bar{M}$, which is proportional to the density, is replaced by $M = \bar{M}/\beta$; and $\bar{\mu}$, which is a mass-constant unassociated with rest-mass in the mechanical theory, is replaced by $\mu = \bar{\mu}/\beta^{\frac{1}{6}}$. The result (84) then follows.

Mass, momentum, charge, etc., are primarily defined by molar measure-

ment. In some cases the definition can be extended to microscopic quantities by the convention that the mass of a particle is $1/n$ of the molarly measured mass of an aggregation of n similar particles. Similarly, the elementary electric charge can 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 are said to be *molarly controlled*. Molar control applies to the mass of a hydrogen atom and other neutral atoms, but it is obviously inapplicable to the masses of electrons or other charged particles. Similarly, there is no molar control of μ.

In quantum mechanics a fresh start is made, and the terms mass, momentum, charge, etc., are applied to quantities which appear as *analogues* of the molar quantities in a well-known analogy between classical and quantum mechanics. Even when molar control exists, the quantum quantity is not necessarily equal to the molarly controlled quantity. We fix the quantum units so that the masses of neutral atoms agree; but it then turns out that the quantum charge e, which is analogous to molar charge, is not equal to the molarly controlled charge e. That is to say, the Coulomb term $z_1 z_2 e^2/r$, which appears in the quantum Hamiltonian of two particles with z_1 and z_2 elementary charges, is not identical with the Coulomb term $Z_1 Z_2 e'^2/r$, which appears in the classical Hamiltonian of two molar bodies with Z_1 and Z_2 elementary charges. The reason for this difference is that *action* is not invariant in the transformation from A to B.

To find the relation between e and e', we compare the microscopic and molar methods of extending pure mechanics to include electrical energy and momentum: (1) the quantum method of combining external and internal (or more generally symmetrical and antisymmetrical) wave functions; (2) the classical method of adding an electrical action to the mechanical action, and determining the pondermotive electric force, etc., by applying the variation principle to the combined action. In (1) we had to lay stress on securing a steady initial state of equilibrium compatible with the natural constants employed; but this is irrelevant in (2), since the classical equations apply equally to steady and unsteady motion. In (2) we have still to multiply densities by β in passing from A to B; but this is only to satisfy molar control of neutral masses, not to restore equilibrium, and it is unnecessary to couple with it multiplication of lengths and times by $\beta^{-\frac{1}{6}}$. Instead, we impose another condition, namely, *molar control of action*; for, since we want to express the electric action in A in such a way that it can be simply added to the mechanical action in B, it must be unaltered in the change of units between A and B. In order that action may be unchanged when the density is multiplied by β, lengths and times must be multiplied by $\beta^{-\frac{1}{4}}$. Calling the resulting system B', we pass from B' to B by multiplying lengths and times by $\beta^{-\frac{1}{6}+\frac{1}{4}} = \beta^{\frac{1}{12}}$. But the particles

in B, B' are neutral particles with the same molarly controlled masses. Their actions will, therefore, differ in the same ratio as the times in B and B'; so that for an action $\hbar$, we have

$$\hbar = \beta^{\frac{1}{1+2}} \hbar' . \tag{86}$$

For example, if in quantum mechanics n neutral particles of mass M each receive additional action $\hbar$ as the result of taking into account electrical characteristics, the corresponding description in molar physics will be that a molar body of mass $n M$ has, besides its mechanical action, an electromagnetic action $n \hbar'$.

In § 19 the elementary charge is identified as

$$e = (\hbar c / 137)^{\frac{1}{2}} . \tag{87}$$

Hence by (86) the corresponding charge in classical electrodynamics is

$$e' = \beta^{-\frac{1}{2}\frac{1}{4}} e. \tag{88}$$

It may be remarked that masses are multiplied in the ratio $\beta^{\frac{1}{2}}$ in the transformation $A \rightarrow B$, and by $\beta^{\frac{1}{4}}$ in the transformation $A \rightarrow B'$. But neither result applies to the *mass of a particle*. The reason is that the number of particles is not kept constant in the transformation; the constant quantity is the number of particles n per unit volume. It is to be remembered that there is no natural subdivision of the scale-free mechanical system into particles; and the density-ratios, which we actually determine, are interpreted as mass-ratios in cases in which there is a one-to-one correspondence of particles. In transformations we must remember that the masses of external particles are really densities and transform them as such; or equivalently we must always consider the same number of external particles per unit volume whatever unit of length is being used.

19. The fine-structure constant.

The factor β can be considered from another point of view. The continuous matter and æther treated in molar theory may be regarded as systems with an infinite number of degrees of freedom. For such systems $k = \infty$, and by (37)

$$\frac{T_{\mu\nu}}{W_{\mu\nu}} = \frac{k}{k+1} = 1 . \tag{89}$$

The total energy is then pure field energy; and, in fact, molar relativity theory is generally described as a "pure field theory." To pass over to particle theory we divide the energy tensor into a large number of elements borne by standard carriers, each limited to 136 degrees of freedom. This step introduces a particle energy $E_{\mu\nu} = - W_{\mu\nu}/137$ not previously recognised, and the total energy is reduced to $W_{\mu\nu}/\beta$. The factors are such that it is legitimate to rigidify the field and treat

small changes of energy as changes of particle energy alone. Thus the phenomena are represented as an activity of microscopic particles in a fixed metrical field, instead of by changes of field variables.

Thus at the very first step we divide the molar energy tensor by β. Since this does not refer to any change in the physical system, but only to our mode of regarding it, it is most suitably described as a change of extraneous standard. When we change from standard particle to bi-particle representation (system A to system B) we again change the extraneous standard so as to multiply densities by β. This brings us back to the original molar standard. It will be seen that the introduction of suffixed particles (unlike unsuffixed particles) does not directly add anything to the energy; but indirectly it adds the energy arising from the interchangeability of the suffixes. Further additions may be made at a later stage by stabilisation.

For a standard particle in the initial state

$$W_{\mu\nu} = -137 E_{\mu\nu}. \tag{90}$$

The coefficient 137 is *the fine-structure constant.*

The fine-structure constant may be thought of as the ratio of two separable elements or "atoms" of action. Such an atom is formed when we multiply a separable element of energy by a time intrinsically associated with it. Two atoms of action are known. Multiplying the energy $-e^2/r$ of an elementary doublet by the time-equivalent r/c of the separation, we obtain the constant $-e^2/c$. Also $\hbar$ is a discrete element of angular momentum or action. The ratio $-\hbar c/e^2$ is -137, the value being verified experimentally to 1 part in 10,000 (§ 20). Thus the atoms are corresponding quantities of particle action and field action.[26]

By considering field energy we obtain a direct interpretation of the constant 137. But usually it arises in the form 136β. That is to say, it comes from the total energy of a standard particle, together with a β factor introduced in the transformation from standard particle to bi-particle representation.[27]

We shall now explain briefly the calculation of the Coulomb interchange energy. We have seen that the internal particle of a hydrogen atom is a free electric particle represented by a Dirac wave function. It is well-known that such a particle has a half-quantum of spin momentum in every spatial plane of rotation. This is an analytical result obtained immediately from relativity theory,[28] and consequently applies to extra-

[26] It will be noted that the sign, as well as the value, agrees. Two like charges do not provide a *separable* element of energy; since they give an electric field extending outwards until it merges into the fields of other particles.

[27] In other words it is the k of a V_{137} particle rather than the $k + 1$ of a V_{136} particle.

[28] *Protons and Electrons,* § 8·3.

spatial rotation (interchange) in so far as the conditions are comparable. It is found, however, that there is, exceptionally, a whole quantum of interchange rotation. This can be shown analytically, but we shall be content here to give a physical explanation. When we introduce spin angular momentum in a single particle, it is implied that an equal recoil angular momentum is given to the rest of the universe; but since the standard form of analysis applies to an assemblage of particles with equal probability of spin in opposite directions the recoil momentum cancels out and is ignored. In the case of interchange, there is no way of putting a recoil interchange into the rest of the universe; and we, therefore, have to modify the wave function, so that the recoil interchange is contained in the particle itself—where it simply doubles the circulation which interchanges the suffixes.[29]

The quantum of angular momentum $\hbar$ gives the internal particle a linear momentum $\hbar/r$ normal to r ($r^2 = \xi^2 + \eta^2 + \zeta^2$). For extra-spatial interchange this is in the direction of a coordinate σ normal to space, which we call the interchange coordinate. This coordinate is, of course, associated with the relative ξ-space in which the internal particle is located, and plays the part of a relative time $\tau = t_2 - t_1$ of the two particles; but the actual identification is $\sigma = i\tau$, so that it rotates circularly (not hyperbolically) with the coordinates ξ, η, ζ. Since the physical distribution is over real values of σ, the connection with $t_2 - t_1$ is only formal. The intrusion of i is due to the inversion of energy-density of the internal particle, which makes it necessary to employ quantum momenta i times the classical momenta in order that the Hamiltonian $p^2/2m$ may give the right sign to the energy. The component p_4 of a quantum momentum vector would naturally represent a "quantum energy" which is i times the classical energy p_4'. But the interchange energy is a real classical energy, and therefore corresponds to imaginary time in the ξ-space.

The interchange momentum $\hbar/r$ has to be inserted in the bi-particle before it is separated and stabilised as two vector particles; so that the system is then in the rigid field of the bi-particle. Using the method of § 14, the time coordinate is then k^{-1} times the true Galilean time; this applies also to the differential time τ and to σ. Correspondingly the momenta $-i\hbar\partial/\partial t$ and $-i\hbar\partial/\partial\sigma$ are k times the true momenta. Thus the true (Galilean) interchange momentum is $\hbar/kr$. It is to be noted that this reduction does not apply to spatial momenta, the space coordinates being unchanged in the transformation (58). Thus

[29] Another point of view is that the mass of a particle is a mutual energy which really belongs half to it and half to the rest of the universe, but the interchange energy of two particles is a mutual energy belonging solely to the two particles. Since our standard reckoning of energy gives a double reckoning of mechanical energy, we have to double the electrical energy to correspond.

the half-quantum of spatial momentum $\frac{1}{2}\hbar/r$ (which is inserted at the same time) appears unmodified in the wave equation, or in the associated integrals of angular momentum, after they have been reduced as usual to Galilean coordinates.

Treating the bi-particle as a V_{137} particle, we obtain directly $\hbar/137\,r$ as the true interchange momentum. It is, however, more in keeping with our previous methods to regard it as $\hbar/136\,r$, modified by a β-factor through the change of extraneous standard in reducing the standard carrier to a bi-particle.

Thus in the internal wave equation of a hydrogen atom we have to insert an interchange momentum

$$\hbar/137\,r. \qquad (91)$$

with the same symbolic coefficient as the internal energy. Current practice inserts such a term under the name of Coulomb energy, and denotes it by $e^2/c\,r$. We have therefore

$$\hbar c/e^2 = 137. \qquad (92)$$

20. Comparison with observation.

The theory determines two important constants $m_p/m = 1836{\cdot}34$ and $\hbar c/e^2 = 137$ (equations (85) and (92)). We shall compare these theoretical values with the latest observational determinations given by R. T. Birge.[30] His results require correction because (like other writers) he has not distinguished between e' and e in obtaining the constant $e/m_e c$. His deflection value corresponds to $e/m_e c$, and his spectroscopic value to $e'/m_e c$. His data are

deflection	...	$e/m_e c = 1{\cdot}75959 \pm {\cdot}00024$
spectroscopic	...	$e'/m_e c = 1{\cdot}75880 \pm {\cdot}00028$

whence we obtain by (88)

deflection	...	$e'/m_e c = 1{\cdot}75905 \pm {\cdot}00024$

The correction removes the discrepancy between the two determinations. The resulting mean value $1{\cdot}758925$ of $e'/m_e c$ is 1 part in 6400 less than Birge's adopted value $1{\cdot}7592$. This constant enters directly into Birge's m_p/m_e which must accordingly be reduced in the same proportion. The results are then—

Birge, uncorrected 1836·56, corrected 1836·27; theory 1836·34.

The probable error is given by Birge as $\pm {\cdot}56$, but this is considerably reduced by the removal of the internal discrepancy of the spectroscopic and deflection data.

---

[30] Physical Society "Reports on the Progress of Physics," 8, 90 (1941). Fuller details of the comparison with theory are given by the author in *Proc. Phys. Soc.*, 54, 491 (1942).

Denoting $\hbar c/e^2$ by A, the observational value of A^{-3} contains the product of the foregoing constant $e'/m_e c$ and the Faraday constant e'/Mc. In both cases the numerator should be e. Thus Birge's A^{-3} requires reduction by 1 part in 6400, and then multiplication by β^{11} to reduce e'^2 to e^2. The resulting values of A are then—

Birge, uncorrected 137·030, corrected 137·009; theory 137.

The probable error is given as $\pm$ ·016, but is smaller in the corrected value.

Both constants accordingly agree with observation, the tests being accurate to about 1 part in 10,000.

PART IV.—GRAVITATION, EXCLUSION AND INTERCHANGE.

21. Physical and geometrical momenta.

We go back to § 2 to take up another line of development. When the distribution function $f(x_0)$ of the physical origin is known, we can convert the distribution function $g(x)$ of the geometrical coordinate x of a particle into the distribution function $h(\xi)$ of the physical coordinate ξ of a particle, or *vice versa*, by a method familiar in the theory of statistics.

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

$$h(\xi) = \int_{-\infty}^{\infty} g(\xi + x_0) f(x_0)\, dx_0 : \qquad (93)$$

Denoting Fourier integrals of the functions f, g, h by corresponding capital letters, we have the reciprocal relations

$$F(q) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iqx} f(x)\, dx . \qquad (94)$$

$$f(x) = \int_{-\infty}^{\infty} e^{iqx} F(q)\, dq . \qquad (95)$$

Then by (93)

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

The right-hand side is separable, and by (94) we obtain

$$H(q) = 2\pi\, G(q)\, F(-q) . \qquad (96)$$

For a Gaussian distribution of the physical origin we have $f(x_0) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-x_0^2/2\sigma^2}$, which gives $2\pi F(-q) = e^{-\frac{1}{2}\sigma^2 q}$. Hence

$$H(q) = e^{-\frac{1}{2}\sigma^2 q^2} G(q) . \qquad (97)$$

Thus from $h(\xi)$ we can derive successively $H(q)$, $G(q)$, $g(x)$ by (94) (97) and (95).

In most quantum problems we are more concerned with the distribution of momenta than of coordinates; and we wish therefore to be able to transform a distribution of physical momenta (conjugate to ξ) into a distribution of geometrical momenta (conjugate to x).[31]

We shall first consider the case in which $g(x)$ is a Gaussian wave packet $(2\pi s^2)^{-\frac{1}{2}} e^{-x^2/2s^2}$; so that $h(\xi) = (2\pi s'^2)^{-\frac{1}{2}} e^{-\xi^2/2s'^2}$, where

$$s'^2 = s^2 + \sigma^2. \tag{98}$$

The distribution function $G(p)$ of momentum corresponding to a Gaussian distribution function $g(x)$ of position is well known,[32] namely

$$G(p) = (\pi\hbar^2/2s^2)^{-\frac{1}{2}} e^{-2s^2p^2/\hbar^2}. \tag{99}$$

We have similar expressions for the momentum distributions $H(p)$, $F(p)$ corresponding to the position distributions $h(\xi)$, $f(x_0)$. Hence by (98)

$$H(p) = \text{const.} \times G(p)\, F(-p). \tag{100}$$

This is of the same form as (96). Comparing the form $2\pi F(-q) = e^{-\frac{1}{2}\sigma^2 q^2}$ with (99), we see that our previous variate q can be identified as

$$q = 2p/\hbar. \tag{101}$$

The factor 2 appears because the distribution function is effectively the square of the wave function. The result (100) states that—

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. (102)

In other words the transformation of a distribution of physical momentum into a distribution of geometrical momentum is made subject to the condition that the resultant geometrical momentum $P = p + p_0$ is zero.

Proceeding to the general case in which $g(x)$ is non-Gaussian, we have to notice that the calculation of $h(\xi)$ from $g(x)$ and $f(x_0)$ would have been impossible without the knowledge (provided by the definition of the physical origin) that x and x_0 are uncorrelated. Similarly, if the corresponding momenta are denoted by p_ξ, p, p_0, the calculation of $H(p_\xi)$ from $G(p)$ and $F(p_0)$ is impossible without information as to the correlation of p and p_0. Since p and p_0 are unobservables this information must ultimately be a matter of definition. We can determine the correlation either by scrutinising the definitions of wave mechanics or by applying the formulae of wave

[31] Roughly speaking, (96) is a formula for converting a physical distribution of momentum $H(p)$ into a geometrical distribution $G(p)$, as may be seen by comparing it with (100).

[32] Darwin, *Proc. Roy. Soc.* A, 117, 268, 1927.

mechanics which depend on the definitions. When the wave mechanical solution is known, it is easier to adopt the latter course; and accordingly, for a Gaussian wave packet, we have determined the correlation of p and p_0 and found that it is complete, namely $p_0 = -p$. But, since the calculation does no more than reveal a truism contained in the definitions of wave mechanics, the result is general and must apply also to non-Gaussian wave packets.

The result $p_0 = -p$ is a consequence of the fundamental conception of wave mechanics as a system of description and prediction employing only observable coordinates and momenta ξ, p_ξ. Such a system makes no provision for the momentum P, and is therefore bound up with the convention that $P = 0$. The convention may be stated in the form that, whenever a particle receives a momentum p, it is to be understood that a recoil momentum $-p$ is imparted to the physical reference frame. From the transformation (78) we obtain

$$p = mP/(m+m_0) + p_\xi \qquad p_0 = m_0 P/(m+m_0) - p_\xi , \quad (103)$$

so that, when $P = 0$, we have independently of the assumed masses m, m_0

$$p_\xi = p = -p_0 \quad (104)$$

and the result (102) follows.

It is convenient to regard $F(-p)$ as a "weight function" $w(p)$, and to write (100) as

$$H(p)\,dp = G(p).\,w dp \quad (105)$$

Then we have the result—

The distribution of physical momenta is derived from the distribution of geometrical momenta by weighing the ranges dp with the weight function $w(p)$ (106)

Evidently the weight function is explained dynamically as a factor which makes allowance for the recoil momentum imparted to the physical reference frame. Corresponding to (99) we have

$$w(p) = F(-p) = (\pi\hbar^2/2\sigma^2)^{-\frac{1}{2}} e^{-2\sigma^2 p^2/\hbar^2}. \quad (107)$$

Introducing the standard deviation ϖ of the weight function, this becomes

$$w(p) = (2\pi\varpi^2)^{-\frac{1}{2}} e^{-p^2/2\varpi^2} \quad (108)$$

where

$$\varpi = \hbar/2\sigma . \quad (109)$$

In three dimensions the weight function becomes

$$w(p_1, p_2, p_3) = (2\pi\varpi^2)^{-\frac{3}{2}} e^{-(p_1^2+p_2^2+p_3^2)/2\varpi^2}. \quad (110)$$

22. The creation of proper mass.

The weight function has the effect of reducing the probability of very high energies. The value of ϖ is about $200\, m_e$. Since the weight function is always ignored in current investigations, there is considerable scope for improving the existing theory of problems involving energy transitions of order $200\, m_e$. The application of the weight function in such problems will, of course, require detailed investigation (taking account of multiplicity factors) in each case; and it is not suggested that there is any simple way of introducing the correction.

Naturally the omission of the weight function has led to the divergence of certain integrals. It has been pointed out by H. C. Corben[34] that the supposed infinite transverse self energy of an electron is due to this omission. These divergences, about which so much is still being written, are merely the difficulties of infinitude which relativity theory encountered and overcame in its infancy. To omit the weight function is to set $\varpi = \infty$ and $\sigma = 0$, and brings us back to the infinite universe. The "infinite universe" was dead and buried 25 years ago; and those who insist on digging up the corpse must expect to be haunted.

In the uniform steady distributions hitherto considered we have generally taken the particles to be at rest. Let us now consider a distribution in which *the geometrical momenta have unlimited uniform probability distribution*; that is to say, the number of particles in a range $dp_1\, dp_2\, dp_3$ is simply proportional to $dp_1\, dp_2\, dp_3$ for all values of p_1, p_2, p_3 from $-\infty$ to ∞. According to classical statistical mechanics this is the distribution in a gas at infinite temperature; and we shall therefore call it an "infinite-temperature uranoid." Introducing the weight function, the distribution of physical momenta is $w\, dp_1\, dp_2\, dp_3$, where w is given by (110). Then the mean values of p_1^2, p_2^2, p_3^2 in the physical distribution are

$$\overline{p_1^2} = \overline{p_2^2} = \overline{p_3^2} = \varpi^2 = \hbar^2/4\sigma^2 . \qquad (111)$$

Since the momenta are very large, the accurate energy formula $E^2 = m^2 + p_1^2 + p_2^2 + p_3^2$ must be used. For a reason that will appear later we consider particles with no proper mass, so that $E^2 = p_1^2 + p_2^2 + p_3^2$. Hence

$$\overline{E^2} = 3\varpi^2 . \qquad (112)$$

By the well-known formulae of molar relativity theory,[35] the pressure P and density ρ of a uniform static distribution of matter satisfy

$$\left.\begin{aligned} 8\pi\kappa P &= -R_0^{-2} + \lambda , \\ 8\pi\kappa\rho &= 3R_0^{-2} - \lambda . \end{aligned}\right\} \qquad (113)$$

Here R_0 is the radius of space curvature, and λ is of the nature of

[34] *Proc. Camb. Phil. Soc.*, **35**, 195, 1939.

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

a constant of integration. When the total number of particles N is fixed, R_0 is a fixed constant $2\sigma\sqrt{N}$ by (13), but λ is at our disposal. By varying λ we vary P/ρ and consequently the temperature. Let

$$T = \frac{\rho - \rho_0}{\rho_0} = \frac{\text{density of thermal energy}}{\text{density of proper energy}}. \qquad (114)$$

The proper density $\rho_0 = \rho - 3P$. Hence by (113)

$$T = 3(\lambda - R_0^{-2})/(6R_0^{-2} - 4\lambda). \qquad (115)$$

As λ increases from R_0^{-2} to $\frac{3}{2}R_0^{-2}$, T increases steadily from 0 to ∞. From the definition of T in (114) it is reasonable to adopt it as a measure of the temperature in relativity theory. In classical theory infinite temperature involves infinite energy, and T would be made infinite by taking ρ infinite in (114). But in relativity theory an increase of temperature diminishes ρ_0, and T is made infinite by the vanishing of ρ_0 while ρ is still finite. It is for this reason that we have assigned no proper mass to the particles in the infinite-temperature uranoid.

Distinguishing values which refer to the infinite-temperature uranoid by an accent, we have

$$\lambda' = \tfrac{3}{2}R_0^{-2} \qquad 4\pi\kappa\rho' = \tfrac{3}{4}R_0^{-2} \qquad \rho_0' = 0. \qquad (116)$$

In the zero-temperature uranoid

$$\lambda = R_0^{-2} \qquad 4\pi\kappa\rho = R_0^{-2} \qquad \rho_0 = \rho. \qquad (117)$$

Hence

$$\rho_0 = \rho_1 = \tfrac{4}{3}\rho'. \qquad (118)$$

The transformation from (116) to (117) shows how proper mass originates. The particles initially without proper mass acquire proper mass by the lowering of the temperature. The proper mass of a particle is created by the change of temperature of the environment; for it is invariant for changes of velocity of the particle itself. This effect can be explained roughly by Newtonian theory. The object-particle is in the gravitational field of the rest of the matter of the universe, and has a negative potential energy in that field. If the field is reduced by reducing the mass (energy) of the rest of the universe, the particle undergoes a decrease of negative energy, i.e. an increase of energy. The proper mass is accounted for as energy added in this way.

It should further be noted that the proper mass of an object-particle does not depend on the temperature of its actual environment, but on the temperature of its standard environment. For the deviation of the actual from the standard environment is represented by a gravitational field; and the potential energy due to it is kept distinct from the inertial mass (mass-constant). The usual convention is that the proper mass or mass-constant is the mass of the particle when at rest in the zero-

temperature uranoid; and we have only relaxed the convention to show how this mass originates, and thereby obtain a means of calculating it.

Regarding the energy tensors of the two uranoids as made up of the self energy tensors of the particles composing them, we set the energy tensor of a particle equal to $Ap_\mu p_\nu$ as in (61). When p_μ is specified by a probability distribution, the mean contribution to the energy tensor is

$$T_{\mu\nu} = A\overline{p_\mu p_\nu}. \tag{119}$$

For a particle in the infinite-temperature uranoid this gives

$$T_{11} = A\varpi^2 \qquad T_{44} = 3A\varpi^2$$

by (111) and (112), E being now denoted by p_4. Summing for the N particles, the total pressure (T_{11}) and density (T_{44}) are

$$P' = NA\varpi^2 \qquad \rho' = 3NA\varpi^2. \tag{120}$$

Now reduce the temperature to zero. The particles are all reduced to rest, so that the momentum vector becomes $(0, 0, 0, m)$; and the total pressure and energy are

$$P = 0 \qquad \rho = NAm^2. \tag{121}$$

Since $\rho = \frac{4}{3}\rho'$, we have $m^2 = 4\varpi^2$. Hence

$$m = 2\varpi = 2\hbar\sqrt{N}/R_0. \tag{122}$$

by (13) and (109).

23. Determination of m_0 and M.

The particles in the foregoing analysis are V_3 particles. For we have represented them by wave functions describing a triple probability distribution. If they carry any other characteristics besides p_1, p_2, p_3 and their conjugate coordinates, these must be stabilised; for no provision has been made for incorporating a probability distribution of the proper mass m or of spin momenta, since we have not investigated the uncertainty of the "origin" or zero-point from which the coordinates conjugate to these momenta are measured. In particular the m which is carried by the V_3 particle in the zero-temperature uranoid is stabilised. It has been computed theoretically from the temperature and number of particles in a standard environment which we postulate, not in the actual environment which we observe.

We have seen (§ 2) that the uncertainty of scale may either be taken into account directly or may be replaced by space curvature. In the preceding investigation it is taken into account as space curvature. Normally in quantum theory, which is not adapted to curved space, it must be taken into account directly, so that m has an uncertainty arising from the uncertainty of the mass-standard or comparison particle with which it is compared. The probability distribution is then four-

dimensional and the particle is a V_4 particle. Denoting the masses of V_3 and V_4 particles by m_3, m_4, equation (122) gives m_3; and by (47) $m_4 = \frac{3}{4} m_3$. Thus

$$m_3 = 2\hbar\sqrt{N}/R_0 \qquad m_4 = \tfrac{3}{2}\hbar\sqrt{N}/R_0 . \qquad (123)$$

This is confirmed by the following considerations:—

The relation $\rho_0 = \frac{4}{3}\rho'$ in (118) implies that, when proper mass is released as kinetic or radiant energy, a quarter of the energy disappears. It is easily verified from (113) that this remains true when only part of the proper mass is released in this way, so that it should apply to nuclear transmutations. The experimenter tells us that there is no change of energy in nuclear transmutations; but that is not really a contradiction. We cannot even imagine the experiment to be performed on a V_3 particle without inconsistency; for if a change of mass is measured experimentally, the mass is not a stabilised quantity and the particle is *ipso facto* not a V_3 particle. Substituting a measurable mass makes it a V_4 particle, and the mass measured is, not m_3, but $m_4 = \frac{3}{4} m_3$. This anticipates the loss of $\frac{1}{4}$ of the energy in the transmutation of a V_3 particle, so that the conservation of energy in transmutation is experimentally satisfied.

The constant in (123) is not the definitive $\hbar$ accepted as the well-known natural constant, because in deriving (123) we have not paid any regard to the rigid-field condition postulated in the wave-mechanical theory in which $\hbar$ is officially defined. We shall therefore change the symbol $\hbar$ in the preceding analysis to γ, and investigate the relation of γ to the true $\hbar$.

The V_3 and V_4 particles are fictitious particles, and it is only by imposing an artificial constraint on matter, so as to prevent irrotational motion, that we are able to analyse it into these particles (§ 11). To obtain the actual natural constants we must consider V_{10} particles of mass $M = \frac{3}{10} m_3$; so that by (123) in the new notation

$$M = \tfrac{3}{5}\gamma\sqrt{N}/R_0 . \qquad (124)$$

Consider a wave function corresponding to a small velocity. The momenta $p_\alpha = -i\gamma\partial/\partial x_\alpha$, which we have been using, 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} particle with small velocity

$$E = M + (p_1^2 + p_2^2 + p_3^2)/2M = M - \gamma^2\nabla^2/2M \qquad (125)$$

In the rigid-field treatment we introduce separate carriers for the initial and transition energy, and by (48) the momenta $p_\alpha = -i\hbar\partial/\partial x_\alpha$ are such that

$$E = M + (p_1^2 + p_2^2 + p_3^2)/2\mu = M - \hbar^2\nabla^2/2\mu . \qquad (126)$$

Comparing (125) and (126)

$$\gamma = \hbar\sqrt{(M/\mu)} = 136\,\hbar/\sqrt{10} \qquad (127)$$

An adjustment is needed, because in separating the carriers of initial and transition energy we double the number of particles. Starting with an actual V_{10} uranoid, the N particles are half external and half internal. Eliminating the separation of initial and transition energy, these fuse into $\frac{1}{2}N$ particles with the constant γ. Keeping R_0 fixed and equal to radius of the actual uranoid of N protons and electrons, when $N \to \frac{1}{2}N$, $\sigma \to \sigma\sqrt{2}$; and the extensions of all quantum-specified systems, which have a fixed relation to σ, are multiplied by $\sqrt{2}$. With our natural units, angular momentum has dimensions (length)$^{-2}$, so that when $\sigma \to \sigma\sqrt{2}$, $\gamma \to \frac{1}{2}\gamma$. Thus when N and R_0 are the constants of the actual uranoid, (124) is changed to

$$M = \frac{3}{5}\frac{\gamma}{2}\frac{\sqrt{\frac{1}{2}N}}{R_0}. \tag{128}$$

Hence by (127)

$$M = \frac{136}{10}\frac{3}{4}\frac{\hbar\sqrt{\frac{1}{5}N}}{R_0} \tag{129}$$

so that

$$m_0 = \frac{3}{4}\frac{\hbar\sqrt{\frac{4}{5}N}}{R_0}. \tag{130}$$

This is the central formula of unified theory. My earlier calculation (*Protons and Electrons*, equation (14.71)) was in error by a factor $(\frac{4}{3})^{\frac{1}{2}}$. The correction was announced tentatively by Eddington and Thaxton, *Physica*, 7, 122, 1940. Although the theoretical calculation of N cannot be dealt with here, it may be added that a correction has also been made in it which makes N equal to $\frac{3}{4}N_0$, N_0 being the cosmical number $2 \times 136 \times 2^{256}$ which represents the number of independent quadruple wave functions at any point of space. The two changes compensate one another in the calculated value of the constant of gravitation (§ 24), but not in the calculation of σ and of the nebular recession. The corrected value of N is

$$N = \tfrac{3}{2} \times 136 \times 2^{256}. \tag{131}$$

To take account of electric charge, we have to make the transformation from standard particle to bi-particle representation; which leaves M unchanged but multiplies lengths by $\beta^{-\frac{1}{6}}$. This changes R_0 to $R_0\beta^{-\frac{1}{6}}$, and we have finally

$$M = \frac{136}{10} \cdot \frac{3}{4}\frac{\beta^{\frac{1}{6}}\hbar\sqrt{\frac{4}{5}N}}{cR_0} \tag{132}$$

the constant c being inserted so as to remove the limitation to natural units.

24. The constant of gravitation.

Since $\hbar/Mc$ is determined observationally with high accuracy, (132) gives a very accurate determination of $\sigma\ (= R_0/2\sqrt{N})$. Previously we have had only rough determinations of σ from the range of nuclear forces and from the recession of the galaxies (§ 4).

By combining (132) with the ordinary molar equation (20) we obtain

$$\kappa = \frac{136 \cdot 137}{10} \frac{3\pi}{2} \frac{\beta^{\frac{1}{3}} F^2 c^2}{\sqrt{5N}} \qquad (133)$$

where $F = e/Mc$. For practical purposes it is better to write $\beta^{\frac{1}{3}} F'^2 c^2$ in place of $\beta^{\frac{1}{3}} F^2 c^2$, where $F' = e'/Mc$ by (88). Then F' is the Faraday constant for hydrogen determined by molar experiment. This provides an experimental determination of the number N of electrons and protons in the universe. Alternatively we can use the value (131) of N, and so obtain a purely theoretical determination of the numerical constant $\kappa/F'^2 c^2$. Since $F'c$ is known much more accurately than κ, we can use it to obtain a value of κ, which should be accurate to 1 part in 5000. The result is $6{\cdot}6665 \cdot 10^{-8}$. The direct observational value of κ is $(6{\cdot}670 \pm {\cdot}005) \cdot 10^{-8}$.

In deriving (132) we have postulated an environment consisting of a steady uniform zero-temperature distribution of protons and electrons only. We are, of course, allowed to re-arrange the matter of the universe in a way that simplifies the theoretical calculation of $\kappa/F'^2 c^2$ just as the experimenter is allowed to re-arrange the matter in his laboratory in a way that simplifies the experimental determination; only we must be careful not to violate the conservation of energy in our re-arrangement. There is therefore an implicit assumption that the energy of free radiation, cosmic rays, etc., is just sufficient to transmute the complex elements into hydrogen and leave a uniform distribution at zero temperature. I think that one of the results of unified theory is to show that this is *necessarily* true, as the result of the way in which energy is defined; but this is not yet fully proved. The practical question—whether the formula (133) gives the ratio $\kappa/F'^2 c^2$ in the actual universe, or only in a somewhat simplified model—can be settled in another way.

Our theory supposes κ to be measured for a system in a standard environment, and it also supposes $F'c$ to be measured in a standard environment. To most simple phenomena the principle of equivalence applies, and the effect of changes in the remote environment is automatically compensated by changes of the local reference system; indeed the remote environment only appears in the theory because it is mathematically simpler to include it than to introduce boundary conditions. But even if κ and $F'c$ are affected by changes of the remote environment, the effect should be eliminated in the purely numerical ratio $\kappa/F'^2 c^2$. This would scarcely be questioned, were it not that by (133) the ratio

involves a constant N apparently furnished by the remote environment. But it is equally legitimate to regard N as a constant furnished *to* the remote environment; for the natural constants κ, F', etc., determine how far the uniform distribution extends and therefore the number of particles it will include. That is actually the way I have proceeded in determining N. I introduce first the cosmical number $N_0 = 2 \times 136 \times 2^{256}$, which is the total number of independent quadruple wave functions at any point. These wave functions are not integrable in spherical space, and the corresponding number of particles is $N = \frac{3}{4} N_0$, a quarter of a particle being, as it were, lost in joining up the wave functions in different regions. If we substitute $\frac{3}{4} N_0$ for N in (133), we have a formula *which contains no reference to the remote environment.* The detachment is indeed emphasised, because the N_0 wave functions to which it refers are non-integrable and cannot be continued far outwards.

I therefore conclude that the value $6{\cdot}6665 \,.\, 10^{-8}$ is the result to be expected from accurate measurement of κ in actual conditions.

25. Exclusion.

The exclusion principle can be formulated as follows:—If the 6-space obtained by taking x, y, z, p_1, p_2, p_3 as coordinates is divided into unit cells of volume h^3 $(h = 2\pi\hbar)$, then in a steady state the maximum number of particles per cell is two protons and two electrons, or equivalently two external and two internal particles of hydrogen.

The exclusion condition for internal particles concerns electrical theory, and the consequences in quantum theory are familiar. The condition for external particles concerns mechanical theory; and we are going to show that it leads to the same formula (130) for the masses of the particles as that obtained by gravitational theory. *Thus exclusion is a wave-mechanical substitute for gravitation.* We can replace our former gravitating (non-excluding) particles by excluding (non-gravitating) particles.

Consider a unit volume of ordinary space, so that a unit cell of momentum space is $dp_1 \, dp_2 \, dp_3 = h^3$. Let n be the number of external particles per unit volume. Let μ_0 be their mass-constant, so that the kinetic energy (according to the classical formula) is $E = p^2 / 2\mu_0$, where $p^2 = p_1^2 + p_2^2 + p_3^2$.

For a zero-temperature distribution the energy is a minimum, and we have to distribute p_1, p_2, p_3 so that ΣE is a minimum, subject to the density not exceeding two particles per cell. The momentum distribution then fills a sphere of radius $p = \mathfrak{p}$ determined by

$$\tfrac{1}{2} n h^3 = \tfrac{4}{3} \pi \mathfrak{p}^3$$

and the "top energy" is

$$\mathfrak{E} = \frac{\mathfrak{p}^2}{2\mu_0} = \left(\frac{3n}{8\pi}\right)^{\frac{2}{3}} \frac{h^2}{2\mu_0}. \qquad (134)$$

The mean energy is

$$\overline{E} = \tfrac{3}{5}\mathfrak{E} = \tfrac{3}{5}\left(\frac{3n}{8\pi}\right)^{\frac{2}{3}}\frac{h^2}{2\mu_0}. \tag{135}$$

This is a well-known result in the theory of white-dwarf matter. Applying (134) to the $\frac{1}{2}N$ external particles of the zero-temperature uranoid, we have $n = \frac{1}{2}N/2\pi^2 R_0^3$; so that

$$\mathfrak{E} = \tfrac{3}{4}\frac{N}{2\mu_1}\left(\frac{h}{2\pi R_0}\right)^2 \tag{136}$$

where

$$\mu_1 = \mu_0(\tfrac{3}{4}N)^{\frac{1}{3}}. \tag{137}$$

We can also write (136) as

$$2\mu_1\mathfrak{E} = \tfrac{3}{4}\left(\frac{\hbar}{2\sigma}\right)^2 = \tfrac{3}{4}\varpi^2 \tag{138}$$

by (109).

The eigenstates, representing stationary oscillations of the uranoid, correspond to the surface harmonics of the hypersphere. By a well-known theorem each discrete eigenstate corresponds to a unit cell. Owing to the symmetry the harmonics of the same order coalesce into a multiple state, and the particles are arranged in a series of levels. There are $(s+1)^2$ independent harmonics of order s; so that the k th level, corresponding to harmonics of order $k-1$, contains k^2 cells or $2k^2$ external particles. We call k the quantum number of the level. If k is large the number of particles up to the k th level is $\frac{2}{3}k^3$. The top quantum number $\mathfrak{k}$ is given by $\frac{2}{3}\mathfrak{k}^3 = \frac{1}{2}N$; so that

$$\mathfrak{k} = (\tfrac{3}{4}N)^{\frac{1}{3}} \tag{139}$$

and by (137)

$$\mu_0 = \mu_1/\mathfrak{k}. \tag{140}$$

The top quantum number of the universe is $2{\cdot}607\,.\,10^{26}$, and the particles in the top level would form a planet of mass intermediate between Neptune and Saturn.

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; but in any case rest mass is some kind of concealed energy, and it makes no difference in mechanics whether it is accounted for exclusionally or gravitationally or (as is more usual) left unexplained. We may distinguish the present analysis of the source of the 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 the sub-threshold theory, and is replaced in quantum mechanics by a mass-constant equal to, or at least associated with, $\mathfrak{E}$. It will be

seen that, although the calculations are precisely similar, there is a difference in the practical interpretations of these results for external particles and the results for internal particles in white-dwarf matter; because the plane of ordinary experience is the top level of the former and the bottom level of the latter.

Let us compare uranoids with different numbers of particles. We keep the rest mass $\mathfrak{E}$ of a particle and the uncertainty constant σ unaltered, so that masses and lengths of quantum-specified systems are unaltered, and hence $\hbar$ is unaltered. Then by (138) μ_1 is unaltered; and by (140) μ_0 varies inversely as $\mathfrak{f}$.

This holds rigorously down to $\mathfrak{f} = 1$. The foregoing proof is limited to large values of N; and indeed the relation of N to $\mathfrak{f}$ becomes ragged for small values. But the relation of μ_0 to $\mathfrak{f}$ must remain uniform. For the uranoids corresponding to different values of $\mathfrak{f}$ are evidently replicas of the same uniform distribution on a series of discrete scales. This is not contradicted by their association with surface harmonics of different orders; each level forms one unsplit state, fully occupied when $\mathfrak{f}$ is an integer. (The surface harmonics were only introduced for the purpose of determining the multiplicity.) It is fairly obvious that the eigenscales will form a regular series. By formal investigation of the scale and its conjugate phase variate, we can show that if the extraneous standard is an angular momentum the eigenvalues of the scale are proportional to the integers, the integer being the quantum number $\mathfrak{f}$. Any mechanical characteristic varies as a power of the scale; and hence the relation of μ_0 to $\mathfrak{f}$, which has been shown to hold for large values, will remain exact down to $\mathfrak{f} = 1$.

To justify the treatment adopted in wave mechanics a quantum particle has been defined as an addition to a rigid environment. This condition is automatically fulfilled by a top particle: it can be inserted or removed without disturbing the fully packed energy-levels below. Thus the replacement of gravitating by excluding particles replaces the gravitational field by an exclusion field, which, unlike the gravitational field, is automatically rigid.[36] We must, however, distinguish between *vertical* and *lateral* exclusion. The part of $\mathfrak{E}$ due to lateral exclusion, i.e. due to the particles at the same level as the object-particle, is usually insignificant; but for $\mathfrak{f} = 1$ the exclusion is wholly lateral. The lowest level is a single cell consisting of two particles which mutually exclude one another from the 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 the state $E = 0$. Regarded as a quantum particle superposed

[36] For this reason exclusion theory is not an alternative to gravitational theory in ordinary gravitational problems, any more than gravitational theory is an alternative to exclusion theory in ordinary quantum problems. In saying that exclusion is the wave mechanical substitute for gravitation, we mean for the gravitation which enters into wave mechanical problems.

on a rigid environment its energy must be taken to be $\mathfrak{E} = C + \mathfrak{p}^2/2\mu_0$, where C (the "sub-threshold rest mass") is the change of energy of the actual environment, i.e. the other particle, due to its presence. Thus $C = \mathfrak{p}^2/2\mu_0$, and $\mathfrak{E} = 2C$.

The classical formula for the energy of a particle with rest mass is $E = \mu_0 + p^2/2\mu_0$. Since this has to become $\mathfrak{E} = C + \mathfrak{p}^2/2\mu_0$ at the one-quantum level, we have $C = \mu_0$; and $\mathfrak{E} = 2\mu_0$ at that level. For $\mathfrak{k} = 1$, $\mu_0 = \mu_1$; so that $\mathfrak{E} = 2\mu_1$ at the one-quantum level. But this last result is independent of $\mathfrak{k}$. $\mathfrak{E}$ and μ_1 being constants as shown above. Inserting it in (136), we have for all values of N

$$|\mathfrak{E}^2 = \tfrac{3}{4} N \left(\frac{h}{2\pi R_0}\right)^2. \qquad (141)$$

Hence $\mathfrak{E} = \mathfrak{k}^{\frac{3}{2}} \hbar / R_0$.

It will be noticed that the classical expression $p^2/2\mu_0$ for the energy (hitherto used only for small values of p) has here been used for extremely large values of p. I have shown elsewhere[37] that it is the exact relativistic hamiltonian for standing waves. Here I need only call attention to the extraordinary persistence of the error of applying the hamiltonian $(\mu_0^2 + p^2)^{\frac{1}{2}}$ to standing waves, especially in the theory of white dwarf matter. The argument which shows that this must be 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. The only information derivable from special relativity theory is that, whatever the correct hamiltonian for standing waves may be, it is not $(\mu_0^2 + p^2)^{\frac{1}{2}}$. General relativity theory (gravitational theory) is more informative. But the supreme test of the correct "relativistic hamiltonian" must be that it is such as to give results which accord with relativity theory; and we are now engaged in applying that test to the hamiltonian $p^2/2\mu_0$ by showing that it gives the same masses of the particles as the gravitational theory did.

26. The negative energy levels.

Regarding the top level $\mathfrak{E}$ as zero level, the particles of the standard environment 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 is necessary to emphasise that there are important differences. Here the number of negative energy levels is finite; the particles filling them are not invented *ad hoc*, but are the rest of the

[37] *Monthly Notices*, R.A.S., 100, 582, 1940.

particles of the universe in the state in which the usual quantum equations postulate them to be, i.e. not interfering with the object-system that is being studied. The theory is symmetrical between protons and electrons, so that presumably negatrons as well as positrons can exist; and the difficulty of the infinite negative charge in Dirac's theory does not arise. Also, as will be seen below, it is essentially a sub-threshold picture; and the particles and holes in it are related to, but not directly equivalent to, electrons and positrons.

An object-particle is formed by exciting a top-level particle. But the object-particle (being defined as an addition to the fixed uranoid) is to be identified, not with the excited particle, but with something which together with the unexcited particle is equivalent to the excited particle. This analysis of the excited particle into an object-particle plus an unexcited particle with fixed rest mass corresponds to the analysis in § 15. We can therefore connect the particles in the exclusion theory with the particles in our previous theory as follows:—

(1) The stabilisation which provides V_{10} particles is introduced when we add an object-system to the zero-temperature uranoid, and does not concern the sub-threshold theory. The particles treated in § 25 are accordingly unspecialised elements of energy tensor.

(2) The top-level particles are the comparison particles in our previous theory.

Both (1) and (2) identify $\mathfrak{E}$ with m_0.

In order to create a positron we must take a hole at the top level and excite it (negatively) to a lower level. Here again the excited hole is, not the positron, but the resultant of the positron and the comparison hole. The treatment in § 15 can be applied equally to holes and particles, and gives the same quadratic equation for the masses of negatrons and positrons as for protons and electrons.

We thus reach the conception of an object-system perched on the firm platform provided by fully-packed energy levels, so that it begins at a threshold energy highly boosted up by exclusion. It is essential to employ only top particles in forming object-systems, otherwise the rigid environment postulated in wave mechanics is not supplied. If an excessively large object-system were considered, we might have to excavate to levels substantially below $\mathfrak{E}$ in order to obtain enough particles, and the energy of the object-system would be correspondingly reduced. This reduction is the negative gravitational potential energy of large bodies, now appearing as negative exclusion energy. But confining attention to microscopic systems, we have to notice that wave mechanics always selects top particles which, by (135), have an energy $\frac{5}{3}$ of the energy of an average particle. Wave mechanics therefore adds a selection energy $\frac{2}{5}\mathfrak{E}$ to the true energy $\frac{3}{5}\mathfrak{E}$ of unselected particles.

Accepting the principle that every particle, when its turn comes for microscopic treatment, is a top particle with rest mass $\mathfrak{E}$ or m_0, we

must couple with the mass a constant of gravitation κ which is $\frac{3}{5}$ of the "true" constant of gravitation κ' in order to compensate the selection effect. Since $8\pi\kappa\hbar^2 = 1$, we have correspondingly

$$\hbar^2 = \tfrac{5}{3}\hbar'^2 . \qquad (142)$$

The true constant $\hbar'$ must be used in the sub-threshold investigation by which m_0 is calculated; $\hbar$ is the ordinary constant in quantum theory, in which m_0 is assigned as the mass to every standard particle or comparison particle in turn. We must therefore change h^2 to $\frac{3}{5}h^2$ in (141), obtaining finally

$$\mathfrak{E}^2 = m_0^2 = \tfrac{3}{4}\tfrac{3}{5}N\left(\frac{h}{2\pi R_0}\right)^2 . \qquad (143)$$

This agrees with the result (130) obtained by gravitational theory.

It should be remembered that we have dealt with the exclusion of external particles only, i.e. the part neglected in current quantum theory. We thus leave the exclusion of internal particles to be taken into account in the usual way. Formally the exclusion principle[38] also applies directly to electrons and protons; but the application would involve such a mixture of theories that it is scarcely likely to have a useful outcome. The application to internal particles is commonly mistaken for application to electrons.

27. Interchange of external particles.

When provision is made for representing the inertial-gravitational field otherwise than by curvature of space, it is usually simplest to adopt as standard environment a uniform distribution of particles in flat space. The *distribution* may be supposed to extend indefinitely, so that it is represented by a pseudo-discrete plane wave-function; but the *environment* of any microscopic object-system is limited to a sphere of radius R_1 about the object-system as centre, R_1 being chosen so as to give the right uncertainty constant σ. Apart from determining σ, the only effect of the boundary is to introduce an extraordinary fluctuation by limiting the total number of particles N_1 in the environment (§ 2). Whatever wave-mechanical substitute for curvature we employ is a substitute for this extraordinary fluctuation. We shall call this environment a *planoid*. For a planoid we easily find $\sigma^2 = R_1^2 / 5N_1$.

We have then, as alternative environments for a microscopic object-system,

(*a*) a zero-temperature uranoid of radius of curvature R_0 containing N particles, and

(*b*) a zero-temperature planoid of radius R_1 containing N_1 particles.

[38] I refer to the "cell" formulation of the principle at the beginning of § 25. The special Pauli principle applying to electrons in an atom requires no comment.

It is understood that the same units are employed in (*a*) and (*b*), so that lengths and masses of quantum-specified structures are the same. In particular the constants σ, $\hbar$ m_0, M are unchanged. Since σ is unaltered, we have

$$\frac{R_1^2}{5N_1} = \sigma^2 = \frac{R_0^2}{4N}. \tag{144}$$

The fundamental formula (130) for m_0 is slightly simplified by using the planoidal constants N_1, R_1; it becomes

$$m_0 = \tfrac{3}{4}\hbar\sqrt{N_1}/R_1. \tag{145}$$

Denoting by $[\ \ldots\ ]$ the mean value of a quantity averaged over the planoid, we have for the distance r from the centre

$$[r^{-2}] = 3R_1^{-2} \qquad [r^{-1}] = \tfrac{3}{2}R_1^{-1},$$

so that

$$m_0 = \left[\frac{\frac{1}{2}\hbar}{r}\right]\sqrt{N_1}, \tag{146}$$

$$m_0^2 = \frac{3}{16}\left[\frac{\hbar^2}{r^2}\right]N_1 = \frac{3}{2}\Sigma_s\left(\frac{\frac{1}{2}\hbar}{r_s}\right)^2, \tag{147}$$

the summation being over the $\frac{1}{2}N_1$ external particles.

Suppose that the object-particle has a half-quantum of angular momentum about the s th planoid particle. The corresponding linear momentum is

$$p_s = \tfrac{1}{2}\hbar/r_s. \tag{148}$$

Thus

$$m_0^2 = \tfrac{3}{2}\Sigma_s p_s^2. \tag{149}$$

Introducing the mass $m_3 = \frac{136}{3}m_0$ of an external V_3 particle, and the usual transition mass-constant $\mu = m_0/136$, (149) gives

$$m_3 = \Sigma_s p_s^2/2\mu. \tag{150}$$

We have thus expressed the rest mass of a V_3 particle as the sum of $\frac{1}{2}N_1$ elements $E_s = p_s^2/2\mu$, each element being the energy associated with a half-quantum of angular momentum about an external particle in the planoid.

The significance of this result is more easily appreciated if we give the V_3 particle a transition energy corresponding to a state of momentum p_x, p_y, p_z. By (46) the transition energy, which is independent of the multiplicity of the particle, is $(p_x^2 + p_y^2 + p_z^2)/2\mu$. The whole energy is then

$$E = \frac{p_1^2 + p_2^2 + \ldots + p_{\frac{1}{2}N_1}^2 + p_x^2 + p_y^2 + p_z^2}{2\mu}. \tag{151}$$

This opens out a far-reaching conception. Initial energy is eliminated, and E appears as the transition energy of a particle with $\frac{1}{2}N + 3$

coordinates. The coordinates are dynamically similar, but are physically distinguished by the fact that only the last three are observable; the momenta conjugate to the others are stabilised quantities furnished by the half-quantum rule. To determine an individual p_s it would be necessary to determine r_s by observation; but in postulating a planoid as standard environment, we prescribe the average number of particles in each range of r_s as a stabilised datum. An observational identification of the particular particle to which each term in (151) refers would be irrelevant to the dynamics.

Since the particles in the planoidal environment are treated collectively we substitute for their separate contributions p_s a resultant $P = \sqrt{(\Sigma p_s^2)}$. Then $m_3 = P^2/2\mu$, and

$$E = (P^2 + p_x^2 + p_y^2 + p_z^2)/2\mu . \qquad (152)$$

Evidently P must be represented in a direction orthogonal to p_x, p_y, p_z, and therefore in an extra-spatial direction.

Why is it that the mass obtained in this simple way is that of a V_3 particle? It is appropriate that the particle should have a multiplicity corresponding to the number of dimensions of the planoid from which its mass is derived; but it is instructive to obtain a more precise answer.

The momenta in (152) are quantum momenta which allow for the inversion of energy. If we use instead the classical momenta, we have

$$E = (P^2 - p_x'^2 - p_y'^2 - p_z'^2)/2\mu = p_0'^2/2\mu , \qquad (153)$$

where p_0' forms the fourth component of a vector of invariant length P. The interpretation is as follows:—The object-particle has a half-quantum of angular momentum about every external particle in the planoid, *and no other momentum is ever 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 in general have components in space as well as in the direction normal to space. Owing to the quadratic law of combination of elementary momenta, the resultant linear momenta in the four directions will satisfy

$$p_x'^2 + p_y'^2 + p_z'^2 + p_0'^2 = P^2, \quad \text{where} \quad P^2 = \Sigma_s (\tfrac{1}{2}\hbar/r_s)^2 ,$$

and is independent of the planes of the angular momenta. By (153) the energy is a quantity $E = p_0'^2/2\mu$ arising exclusively from the momentum p_0' in the extra-spatial direction. There is no way of calculating p_0' by pure theory; but we can determine its value by *measuring observationally* p_x', p_y', p_z' and using (153).

The space momenta p_x', p_y', p_z', can only be zero if each element of angular momentum is in a plane through r_s and the extra-spatial axis. Our earlier assumption that the elements are in these planes, so as to give only extra-spatial momenta $p_0' = p_s$ was correct for the special

problem considered, namely determination of the rest mass. But it is not a general law; and, when it is true, it is an inference from the fact that $p_{x'}$, $p_{y'}$, $p_{z'}$ have been measured and found to be zero. No other observational information was employed, because P is a constant depending only on the postulated environment. Thus the result which we obtained applies to a particle for which $p_{x'}$, $p_{y'}$, $p_{z'}$ and nothing else have been measured, i.e. to a V_3 particle.

We have seen that a half-quantum of interchange angular momentum is introduced by suffixing the two particles of a hydrogen atom, and that this constitutes the Coulomb energy of the charges. The question naturally arises, why do we not get a similar interchange angular momentum and energy by suffixing the external particles of two hydrogen atoms. The formal answer is that the external particle has no comparison particle to carry a suffix (§ 15) so that the machinery of interchange is lacking. We have already noted the significant fact that the particles which have comparison particles associated with them are charged particles (§ 15). But now, by deeper analysis, we see that in appropriate conditions the external particles do have a half-quantum of interchange angular momentum; and the interchange energy with all the other external particles in the universe constitutes the mechanical energy—both rest energy and kinetic energy. In short, when the particles are defined quantum-mechanically as superpositions on a rigid environment, all energy is primarily interchange energy.

28. Non-Coulombian energy.

The most elementary two-particle system has been found to consist of a proton and electron. The opposite sign of the two charges is the result of the assumption that the system possesses steady states. We shall now consider two like charges—for definiteness, two protons.

If ξ, η, ζ are the relative coordinates and $\rho = (\xi^2 + \eta^2 + \zeta^2)^{\frac{1}{2}}$, the electric energy of two protons is

$$E = e^2/\rho \qquad (\rho \neq 0) \qquad (154)$$

The formal proof that it must be equal and opposite to the electric energy of a proton and electron calculated earlier, is obtained by considering that neutral matter consists of equal distributions of positive and negative charges, and (by definition) the energy of a particle resulting from a neutral environment is wholly inertial-gravitational. The Coulomb energies must therefore cancel.

If we select a proton as sole object-particle we leave an unpaired electron in the environment. Since the environment particles have uniform probability distribution, there is a charge $-e$ uniformly distributed over the uranoid. If ϕ is the electric potential due to the proton, and $\bar{\phi}$ its mean value over the uranoid, the mutual electric

energy of the proton and its environment is

$$\Omega = -e\bar{\varphi}. \tag{155}$$

To eliminate this energy (so as to obtain a proton which is superposed on the standard neutral environment) we take the energy of a pair of particles to be

$$\left.\begin{aligned} E &= -e^2/\rho && \text{(unlike charges)} \\ E &= e^2/\rho + B\delta(\rho) && \text{(like charges)} \end{aligned}\right\}. \tag{156}$$

where δ is the δ-function, and B a constant. The form of the second term is fixed because (154) is established at all points except $\rho = 0$, but there is nothing to forbid the introduction of a singular energy at $\rho = 0$. To fix ideas, we take one particle as origin, and consider a second particle whose relative coordinates have even probability distribution over the volume V of the uranoid. Then by (156) the mutual energy is

$$\int E\frac{dV}{V} = \Omega \quad \text{or} \quad -\Omega + \frac{B}{V} \tag{157}$$

according as the charges are unlike or like. Hence, if the first particle is in the standard environment with *equal* numbers of protons and electrons, the mutual energy of the particle and its environment is $\frac{1}{2}NB/V$. This will agree with (155) if

$$B = 2\Omega V/N. \tag{158}$$

We thus transform the "classical charge", which induces an equal and opposite charge in the environment, into a "quantum charge", which is simply superposed on a neutral environment, by taking the law of energy to be (156). The term $B\delta(\rho)$ will be called the *decoupling energy*.

The foregoing transformation can be regarded more physically. In a uniform probability distribution of protons and electrons, we choose a proton at random and take it as the origin of the relative coordinates ξ, η, ζ. Consider now the original distribution referred to the frame ξ, η, ζ. The frame is similar to an x, y, z frame except for the peculiarity that the point $(0, 0, 0)$ is always occupied by a proton. Each proton has an equal chance of being the occupant of this point; its probability distribution therefore consists of a singular point with a chance $2/N$, and an even distribution of the rest of the probability. The energy at the singular point must be determined by the condition that the total energy agrees with that calculated in an ordinary frame x, y, z without a singular point. This gives the result $B\delta(\rho)$.

As usual the elementary result is complicated by multiplicity factors when we compute the corresponding energy term to be inserted in the wave equation. Our starting point is an object-system containing a coupled proton and electron. We have to take two such systems, decouple the

protons and combine them into another system. It will be simplest if at the same time we combine the decoupled electrons into a two-electron system. Clearly, this reshuffling must be performed on unspecialised elements of energy tensor; and it is not until the two-proton and two-electron carriers of energy tensor have been formed that we can stabilise the particles as V_{10} particles. In this stabilisation the initial energy $2m_0$ of the two standard carriers is magnified into an initial energy $2m_p$ of the two protons, and $2m_e$ of the two electrons. It is not difficult to see that the decoupling energy, being an initial energy, is included in this magnification. Thus the coefficient for protons is

$$B_p = \frac{2\Omega V}{N} \frac{m_p}{m_0}. \tag{159}$$

It may seem odd that the decoupling energy is initial energy, whereas the ordinary Coulomb energy e^2/ρ is transition energy and is therefore unaffected by the subsequent stabilisation. But the decoupling energy corresponds to the initial Coulomb energy of the proton in its negatively charged environment, and is comparable with the initial Coulomb energy which we remove from a standard particle distribution in order to form a bi-particle distribution.

The value of $\bar{\phi}$ for a spherical universe can be calculated by general relativity theory.[39] The result is $\bar{\phi} V = 2\pi R_0^2 e$. Hence

$$-B_p = 4\pi R_0^2 e^2 m_p / N m_0 = 16\pi\sigma^2 e^2 m_p / m_0 .$$

A rather subtle correction is required to take account of the fact that the calculation refers to a three-dimensional distribution over ξ, η, ζ, whereas the wave equation of the relative distribution of the two protons refers to a four-dimensional distribution including the interchange coordinate $\sigma = i\tau$.[40] It would be difficult to discuss this without a lengthy digression; but the result gives a factor $(\frac{4}{3})^{\frac{1}{2}}$, so that

$$-B_p = \left(\frac{4}{3}\right)^{\frac{1}{2}} \frac{m_p}{m_0} 16\pi\sigma^2 e^2 . \tag{160}$$

We have seen (§ 4) that the relative coordinate of two particles can be defined either as the difference ξ_{12}' of their physical coordinates, or as a coordinate ξ_{12} directly measured from one to the other. In the foregoing discussion ξ evidently refers to ξ_{12}, so that the decoupling term is $B_p \delta(\rho_{12})$. But the internal wave equation of the two-proton system is obtained by the transformation (78), and the

[39] *Proc. Roy. Soc.*, A, 162, 159, 1937. I have since found that the easiest procedure is to develop the theory of the planoid rather more fully than in § 27, and then carry out the whole calculation in the planoid instead of the uranoid. The result (160), including the factor $(\frac{4}{3})^{\frac{1}{2}}$, is then obtained.

[40] Commonly mistaken for the time t (which is an independent variable having no connection with the probability distribution).

coordinate ξ occurring in it is ξ_{12}'. Since $\xi_{12} = 0$ corresponds to a Gaussian probability distribution of ξ_{12}' with standard deviation $\sigma\sqrt{2}$,

$$\delta(\rho_{12}) = (4\pi\sigma^2)^{-\frac{3}{2}} e^{-\rho_{12}'^2/4\sigma^2}. \qquad (161)$$

Thus the decoupling term in the wave-equation is

$$Ae^{-\rho^2/4\sigma^2} \qquad (162)$$

where

$$A = (4\pi\sigma^2)^{-\frac{3}{2}} B_p = -\left(\frac{16}{3\pi}\right)^{\frac{1}{2}} \frac{m_p}{m_0} \frac{e^2}{\sigma}. \qquad (163)$$

This is the non-Coulombian energy of two protons. The value of A calculated from (163) is $52{\cdot}01\ m_e c^2$; the observational value is $51{\cdot}4\ m_e c^2$.[41] The comparison of the range given by (162) with observation has been discussed in § 4.

The non-Coulombian energy of two electrons is much smaller, m_p being replaced by m_e in (163). There is no non-Coulombian energy in the proton-electron system. The internal wave equation for hydrogen is therefore exact, subject to modifications arising from the interaction of the internal particle with the radiation field.

The result (163), without the $(\frac{4}{3})^{\frac{1}{2}}$ factor, was given in an earlier paper.[42] The present deduction enables us to go a little further. Taking a proton as origin, let another proton and an electron have the same coordinates ξ, η, ζ. Then the sum of the two energies associated with a point ξ, η, ζ is $B_p \delta(\rho)$ by (156). Thus the non-Coulombian energy of a proton and neutron is the same as that of two protons. The energy of two neutrons is nearly the same, since the mutual non-Coulombian energy of the two electrons contained in them is very small compared with that of the two protons.

---

[41] I have investigated the β-factors, and find that their effect is to multiply the value of A given in (163) by $\beta^{\frac{1}{6}}$, and multiply the range-constant 2σ by $\beta^{-\frac{1}{6}}$. It may be added that the scattering experiments give only a rough determination of A and σ separately, but give a good determination of $A\sigma^2$; we therefore obtain a much more delicate test of (163) by reducing the observations with the known value (21) of σ. The observational result is then found to be $A = 52{\cdot}26\ m_e c^2$.

[42] *Proc. Roy Soc.*, A, 162, 155, 1937.