DEVELOPMENTS
IN THE
THEORY OF THE ELECTRON

BY

ABRAHAM PAIS

Containing also an abstract of
Field and Charge Measurements in Quantum Theory, a paper by N. Bohr and L. Rosenfeld

1948

INSTITUTE FOR ADVANCED STUDY
AND
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY

DEVELOPMENTS IN THE THEORY OF THE ELECTRON

I. Classical Theory. II. Relativity Theory, (a) Abraham- versus Lorentz electron, (b) Self-energy and self-stress, (c) Approaches to a consistent relativistic theory. III. Non-relativistic Quantum Theory, (a) Zero point energy and quantum self-energy, (b) The correspondence argument in quantum field theory. IV. Relativistic Quantum Theory, (a) General features, the renunciation of the one-body concept, (b) The subtraction problem, general considerations, (c) The subtraction problem, special cases, (d) The self-energy problem in relativistic quantum theory, (e) Approaches to a consistent relativistic quantum theory. V. Concluding Remarks.

This survey aims at reviewing how the theory of the electron, from its original classical form, developed through the advents of the relativity and quantum theories to the present stage.[1] In particular it stresses how this development, like many others, is characterized by two main trends: on the one hand every new phase comprises its predecessor as a limiting case in the description of many phenomena in which electrons are involved (aspect of correspondence); but on the other hand each new step brings novel features with it which force us to a re-evaluation of our picture of the electron itself (aspect of disparity). For example, relativity generally goes over into classical theory in the limit of small velocities. But while in classical theory the electromagnetic mass of the electron can unambiguously be defined (I) this is no longer possible in relativity theory even in the low velocity limit (IIa). In this connection the concepts of self-energy and self-stress are discussed (IIb). These quantities become infinite if a point model for the electron is used. Various attempts have been made to eliminate these infinities so as to attain a consistent classical relativistic theory of a point electron. These attempts are briefly surveyed in section IIc.

Such modifications may of course suggest further lines of progress. However, it must be emphasized that a classical singularity-free theory cannot be considered necessarily to correspond to a similarly consistent quantum theory. For the quantum theory of systems of charged particles

[1] This review of developments in the theory of the electron was originally intended to be an introductory survey to a number of papers on positon theory, already published elsewhere, to be reproduced and compiled in book form. Due to the very recent rapid developments in this field, such a book would not now seem timely. The present survey was closed off in December 1947 and is meant to serve as an orientation of the main trends in electron theory up to that time. I would like to express my deep gratitude to all those whose kind advice has been invaluable in preparing this survey. The numbers in square brackets refer to the bibliography.

and fields again exhibits features of disparity along with an intimate correspondence in its relation to classical theory. This is discussed in section III, where the quantum aspects of the electromagnetic field in interaction with the electron are reviewed from a non-relativistic point of view as far as the electron is concerned. Thus it is recalled in section IIIa how the very process of subjecting a divergence-free theory to quantization leads to the introduction of new divergences due to unavoidable quantum electrodynamic fluctuation phenomena. Therefore the correspondence principle does not provide a guide in the self-energy problems. On the other hand these same electromagnetic fluctuations play an important role in the analysis on correspondence lines of the problems of observability involving electromagnetic fields. This has been shown by Bohr and Rosenfeld (IIIb); their work hence establishes the logical link between quantum electrodynamics and classical theory.

We have finally to make the last step (as things stand today), the transition to relativistic quantum theory. And again, as we have already noted twice, we must distinguish between a domain of correspondence and of disparity, this time with the non-relativistic and non-quantum limiting cases. As a sequel to their previously mentioned paper, Bohr and Rosenfeld analyse the correspondence problem in relativistic quantum theory. Specifically they discuss the influence of the charge-density fluctuations, due to pair formation, on the observability of the electric charge on a body. This work has not yet been published, but an abstract of it is included at the end of this paper.

With regard to the non-correspondence features, it is well known how relativistic quantum theory forces us to renounce the one body picture of an elementary particle (IVa). As a result of this renunciation the self-energy problems appear in an entirely new light (IVd), not only for electrons which satisfy the exclusion principle, but also for particles which satisfy the Bose-Einstein statistics.

Insofar as the various steps in the transition from classical to relativistic quantum theory can be considered irrevocable, it would seem that attempts at further progress which do not admit the concepts of relativistic quantum theory from the outset are highly academic, in view of the typical non-correspondence development in our notions of the elementary particles. This holds in particular for the various tentative schemes, outlined in IIc, for a consistent theory within the classical relativistic framework. The transcription of these ideas into relativistic quantum theory has either failed or not yet been fully developed (IVe).

I. Classical Theory

It is hardly necessary to recall here the great successes of the classical theory of the electron initiated by the pioneer work of H. A. Lorentz [1]. In this pre-relativistic version Lorentz considered the electron to be a small

rigid sphere, the stability of which, rather than being accounted for by the theory, is assumed *ad hoc*. Consequently a detailed analysis of the intrinsic properties of the electron lies beyond the scope of these considerations. Yet there is one feature of that problem which is already inherent in the classical picture, viz. that of the electromagnetic mass. Indeed, closely analogous to the case of a rigid body moving through an incompressible fluid,[2] an electron with mechanical mass m_o and in uniform motion will appear to have a momentum $(m_o + \mu_{cl})v$ and a kinetic energy $(m_o + \mu_{cl})(v^2/2)$. Here it is supposed that v is small compared with the light velocity c and

$$\mu_{cl} = \theta \frac{e^2}{ac^2};$$

also e is the charge of the electron, a its radius, θ a numerical constant depending on what assumptions are made for its (spherical) charge distribution (e.g. for a spherical shell model $\theta = \frac{2}{3}$). Thus, on the classical picture, it would seem that it is the quantity $m_o + \mu_{cl}$ which is observed experimentally, while m_o merely plays the role of a mathematical parameter occurring in the equation which expresses the equality of an inertial force m_o times acceleration and the Lorentz force due to the total electromagnetic field which is present. We will therefore, but only for the moment, identify $m_o + \mu_{cl}$ with the "experimental mass" m of the electron.

The recognition of the existence of a contribution μ_{cl} to the electron mass due to the presence of the electromagnetic field raises the question whether it is possible to formulate the theory of the electron in a "structure independent" way, i.e. to bring it into a form such that all its inferences can be expressed by using e and m rather than e and m_o as electron attributes. Thus for example the Lorentz equation of motion for the electron as well as the electromagnetic field equations can in the non-relativistic case be derived from the Hamiltonian

$$\mathfrak{H} = \frac{\mathbf{p}^2}{2m_o} - \frac{e\mathbf{p} \cdot \mathbf{A}_\perp{}^{(0)}}{m_o c} + \frac{e^2 \mathbf{A}_\perp{}^{(0)2}}{2m_o c^2} + U + \frac{1}{8\pi}\int (\mathbf{E}_\perp{}^2 + \mathbf{H}^2)dv \qquad (2)$$

Here $\mathbf{p}$ is the (mechanical) electron momentum, $\mathbf{A}_\perp{}^0$ the transverse (i.e. divergence-free) vector potential at the position of the electron, $E_\perp$ and H the transverse electric and the magnetic field respectively, while U is the potential energy of the system (including all Coulomb energies[3]). $\mathfrak{H}$ now depends on e and m_o and the question is whether we can (at least for $v/c \ll 1$) bring $\mathfrak{H}$ into a form depending on e and m only. Moreover we shall require

[2] It was by using this hydrodynamical analog that J. J. Thomson was led to the introduction of the concept of electromagnetic mass many years before the discovery of the electron [2].

[3] Hence U contains also the electrostatic energy of the electron itself, i.e., a term $3\theta e^2/4a$; this term is obviously structure dependent, but can be discarded from a non-relativistic point of view; it is identical with $\mu_{rel}c^2$ defined by (6).

that the transformation achieving this is a canonical one, this especially in view of the transition to quantum theory.

Especially Kramers [3] has stressed that a closer analysis of the question of structure elimination may lead to a deeper insight into the nature of the paradoxes we are faced with at present. In fact, the remarkable situation presents itself that for example the (non-relativistic) quantum theory of a system of electrons in interaction with each other and with radiation fields [4] starts from an expression (2) with the tacit assumption that m_0 and m are identical. It is well known that this theory has to a certain extent led to a satisfactory description of processes like emission and absorption of radiation by atomic systems and of dispersion phenomena. However it might be deemed that the inconsistencies of this same theory might at least partially find their source in the logically unjustified identification of m_0 and m. If we first transform away the electromagnetic mass in the way indicated above and then quantize the new Hamiltonian scheme thus obtained, can we not attain a formalism which is free of the occurrence of the well-known divergences of the customary "structure dependent" theory based on (2)? The answer to this question is negative, as we shall see in discussing below in somewhat more detail the typical quantum features of the interaction of charged particles and electromagnetic fields. Before coming to the introduction of the quantum of action, however, we will first examine how the inclusion of the basic ideas of relativity theory affects the situation. It will appear that the very concept of electromagnetic mass (and thus of experimental mass) in the way we have used it thus far loses its unambiguous meaning in a relativistic theory, even in the limit of very low velocities. Thus relativity theory severely limits the scope of a theory which is structure independent in the sense used hitherto.

II. Relativity Theory

(a) *Abraham- versus Lorentz electron; ambiguity of the electromagnetic mass concept*

Let us first consider, for arbitrary velocities, the electromagnetic energy and momentum E_{elm} and p_{elm} of our rigid electron—which in fact is just the so-called Abraham electron [5]. To fix ideas we take a shell model.[4] Then $\left(\text{with } \beta = \frac{v}{c}\right)$:

$$E_{elm} = \frac{e^2}{2a}\left(\frac{1}{\beta}\ln\frac{1+\beta}{1-\beta} - 1\right) = \frac{e^2}{2a} + \frac{1}{2}\left(\frac{2e^2}{3ac^2}\right)v^2 + \cdots$$

$$p_{elm} = \frac{e^2}{2ac\beta}\left(\frac{1+\beta^2}{2\beta}\ln\frac{1+\beta}{1-\beta} - 1\right) = \frac{2e^2}{3ac^2}v + \cdots$$

[4] All conclusions on the properties of the Abraham relative to the Lorentz electron are independent of the kind of spherical distribution. In particular all expressions for E_{elm} and p_{elm} for the shell models should be multiplied by $\frac{4}{3}$ to get the results for a homogeneous charge distribution.

the second term in the development of E_{elm} representing the electromagnetic kinetic energy E_{elm}^{kin} for small velocities. Hence for small v

$$E_{elm}^{kin} = \frac{p_{elm}^2}{2\mu_{el}}; \qquad \mu_{el} = \frac{2e^2}{3ac^2} \tag{3}$$

as it should be if we wish to consider μ_{el} to be a part of the experimental electron mass.

Considering now the Lorentz shell electron we have (for later purposes we write down the result for the case that there is a mechanical mass m_o):

$$E = \frac{m_o c^2}{\sqrt{1-\beta^2}} + \frac{\mu_{rel} c^2}{\sqrt{1-\beta^2}}\left(1 + \frac{\beta^2}{3}\right), \qquad \mu_{rel} = \frac{e^2}{2ac^2} \tag{4a}$$

$$p = \frac{m_o + \frac{4}{3}\mu_{rel}}{\sqrt{1-\beta^2}} v, \tag{4b}$$

and for $v/c \ll 1$

$$E_{elm}^{kin} = \tfrac{1}{2}(\tfrac{5}{3}\mu_{rel})v^2, \qquad p_{elm} = \tfrac{4}{3}\mu_{rel} v$$

and therefore, even for small velocities, the notion of electromagnetic mass becomes ambiguous, since it is not possible to fulfill (3). It would therefore seem that in analyzing the problem of electromagnetic mass the consideration of the limiting case of classical theory does not provide an adequate basis, since experiment has taught us that we should discard the Abraham electron in favor of the Lorentz model.[5]

(b) *Self-energy and self-stress*

The expressions given for E_{elm} and p_{elm} are special applications of the following general formulae for the energy and momentum of a particle which is the source of certain fields: Let the energy-momentum tensor of the total system of particle (with mechanical mass m_o) and fields be $\mathfrak{T}_{ik}$. The energy of the system then is given by[6] $\int \mathfrak{T}_{44} dv$, ($dv = dx\, dy\, dz$) the momentum by $-\frac{i}{c}\int \mathfrak{T}_{14} dv$ (for a particle moving in the x-direction). Now denoting the $\mathfrak{T}_{ik}$ in the rest system of the particle by $\mathfrak{T}_{ik}(0)$ we have (as $\mathfrak{T}_{14}(0) = \mathfrak{T}_{41}(0) = 0$; $dv = dv_o\sqrt{1-\beta^2}$)

$$E = \int \mathfrak{T}_{44} dv = \frac{\int \mathfrak{T}_{44}(0) dv_o - \beta^2 \int \mathfrak{T}_{11}(0) dv_o}{\sqrt{1-\beta^2}} \tag{5a}$$

[5] It may be recalled that in the well known e/m determinations as a function of the velocity it is only the electromagnetic momentum which enters into the interpretation of the measurements. As this quantity is, for small velocities, the same for the Abraham and the Lorentz electron, it was necessary to go to higher velocities in order to decide experimentally in favor of either model. (It is furthermore interesting to recall that in the numerical interpretation of these experiments it was assumed that the mass is entirely of electromagnetic origin.)

[6] This holds for suitable normalization of the energy; $x_1 = x$, $x_2 = y$, $x_3 = z$, $x_4 = ict$. As regards the mechanical part of $\mathfrak{T}_{ik}$ only the 44-component differs from zero in the rest system.

$$p = -\frac{i}{c}\int \mathfrak{T}_{14}dv = \frac{\int \mathfrak{T}_{44}(0)dv_o - \int \mathfrak{T}_{11}(0)dv_o}{c^2\sqrt{1-\beta^2}}\,v \tag{5b}$$

We now introduce the self-energy $\mu_{rel}c^2$ by

$$(m_o + \mu_{rel})c^2 = \int \mathfrak{T}_{44}(0)dv_o \tag{6}$$

In the electromagnetic case the relation between μ_{cl} and μ_{rel} is

$$\mu_{rel} = \tfrac{3}{4}\mu_{cl}$$

and for the Lorentz electron we find (4) from (5).[7]

The formulae (5) demonstrate that even for a relativistic electron model with finite self-energy—like the Lorentz electron with $a \neq 0$—we still do not have a consistent picture of the electron, as (4) does not give the right relation between energy and momentum of a freely moving particle, however we define the experimental mass. A glance at (5) shows that the cause of the trouble lies in the circumstance that $\int \mathfrak{T}_{11}(0)dv_o$, the "self-stress," is unequal to zero. In physical terms, the electromagnetic forces tend to explode the electron. Whereas it was possible in classical considerations to simply postulate the compensation of the electromagnetic forces acting on the electron so as to have a stable particle, this is no longer admissible when relativity is taken into account. Thus we have as two necessary conditions in relativistic electron theory: (a) that the self-energy be finite, (b) that the self-stress be zero.[8]

In a sense the second condition is more essential than the first. For if the stresses were zero, one would be able to discard in a covariant way the self-energy of a free particle by stating that $m_o + \mu_{rel}$ shall represent the experimental mass of the electron, whether or not μ_{rel} is finite. Thus non-zero stresses prohibit the unambiguous formulation of a structure independent theory. These requirements obviously hold *mutatis mutandis* for any particle coupled with a field.

That the conditions (a) and (b) are in principle independent of each other can be seen from the example of the so-called Poincaré-electron [6]: The first equation (4) can be written as follows:

$$E = \frac{(m_o + \tfrac{4}{3}\mu_{rel})c^2}{\sqrt{1-\beta^2}} - PV \tag{7}$$

with

$$P = \frac{\mu_{rel}c^2}{4\pi a^3}, \qquad V = \frac{4\pi a^3}{3}\sqrt{1-\beta^2}$$

[7] The electromagnetic terms are readily found by noting that the trace of the electromagnetic tensor is zero and that in the rest system $\int \mathfrak{T}_{11}dv = \int \mathfrak{T}_{22}dv = \int \mathfrak{T}_{33}dv$, as in that system the charge distribution is spherically symmetrical.

[8] Cf. M. Laue [7]; formulae like (4) and (5) can also be considered to apply in an asymptotic sense to the case of a point electron, ($a = 0$), in as far as the latter is defined as the limiting case of a Lorentz model of finite extension. It should also be noted that the vanishing of the divergence of T_{ik} is, in the presence of singularities, generally not sufficient to make the self-stress zero.

Thus V is the volume of the electron; P has the dimensions of a pressure. If we could eliminate the $-PV$ term, the relations between E, p and v would become correctly relativistic. Poincaré's suggestion now amounts to adding the tensor

$$\begin{array}{rl} & P g_{ik} \rho(\mathbf{x}) \\ \rho(\mathbf{x}) = 1 & \text{inside the electron} \\ = 0 & \text{outside the electron} \end{array}$$

to the electromagnetic tensor. This indeed cancels (in all systems of reference) the $-PV$ term in E and does not give a contribution to p. Thus for $a \neq 0$ both conditions would be fulfilled. Defining a point electron as the limiting case $a \to 0$ we see however that the self-stress would stay zero in the limit, but the self-energy would become infinite. Thus for every model a separate discussion for self-energy and self-stress is necessary. Moreover the Poincaré model $a \neq 0$ is instructive in showing that our two conditions are not sufficient. It was indeed shown by Lorentz [8] that the equilibrium one obtains by introducing the Poincaré tensor is not a stable one with respect to any deformations: Poincaré's cohesion pressure compensates the explosive electromagnetic forces for a given value of the electron radius, but it does not stabilize the electron. The conditions of finite self-energy and zero self-stress are thus closely linked with a relativistic *compensation* of the electromagnetic forces, to be sure, but the actual *stabilization* problem would seem to involve more and deeper lying features which might be connected with such open questions as the very existence of an elementary quantum of charge.

(c) *Approaches to a consistent relativistic theory*

With regard to self-energy and self-stress, the various attempts[9] to formulate a classical relativistic theory of the electron can be tabulated as follows (0 = finite radius, · = point model):

Theory	Model	Self-energy	Self-stress
Lorentz	0	finite	finite, $\neq 0$
Lorentz	·	∞	∞
Poincaré	0	finite	0
Poincaré	·	∞	0
λ-limiting process	·	0	0
Action at a distance	·	0	0
Two-field theories	·	finite	0
Born (non-linear theories)	·	finite	0

[9] This short survey of modifications of the classical Lorentz theory makes no pretense of completeness. It rather aims at indicating along what general lines such modifications have been attempted.

In each of the theories, briefly discussed below, we find represented one of the following lines of thought:

(*a*) One uses the circumstance that the customary Lorentz formulation does not mathematically exhaust the solutions of the Maxwell equations in the presence of point electrons. Thus the definition of a charged point is not unique (λ-process). Moreover, the use of combinations of retarded and advanced potentials leaves some scope for modification (Dirac's classical theory, action at a distance). The static interaction between two electrons is given by Coulomb's law. But the force of an electron on itself is zero, either as a result of suitable limiting processes (λ-process, Dirac's theory) or by definition (action at a distance).

(*b*) One assumes that there are non-electromagnetic forces acting on the electron, described half-phenomenologically (Poincaré) or causally in terms of differential equations, linking these forces with the field concept (two-field theories). While the Maxwell-Lorentz field equations themselves remain unchanged, there is thus a resulting force (differing from Coulomb's law) between electrons such that this total force is zero at the position of the electron.

(*c*) One assumes that the Maxwell-Lorentz equations themselves are changed in the sense that they only represent the state of affairs for weak fields or, equivalently, at large distances from the electron and that they become non-linear at small distances (Born).

We will now give a short characterization of these classical relativistic theories and will come back to the application of these ideas to quantum theory in section IVe, where we will see that the very approach to these problems from a nonquantum mechanical basis would seem inadequate (cf. also p. 2).

λ-*limiting process.* This procedure may easily be contrasted with the Lorentz theory. There the point electron is defined as the limiting case in which a space-like vector, the electron radius, tends to zero. Here a time-like vector is introduced and the actual situation on the world line of the electron is assumed to correspond to the limit where this vector is zero. For example, the electric field $\mathbf{E}(\mathbf{x}_o, t_o)$ at the position $\mathbf{x}_o$ of the electron at the time t_o is defined as $\mathbf{E}(\mathbf{x}_o, t_o) = \lim_{\xi \to 0, \tau \to 0} \mathbf{E}(\mathbf{x}_o + \xi, \mathbf{t}_o + \tau)$ where the transition to the limit is made so that always $|\xi| < c\tau$. As first pointed out by Wentzel [9], one obtains in this way a zero electromagnetic energy and momentum for a freely moving electron. The formalism gives the same results as the Lorentz theory for the situation at any point off the world line of the electron.

The limiting process proposed by M. Riesz, using a method of analytic continuation of the solutions of the inhomogeneous electromagnetic wave equations in the complex plane, has been shown by Ma [10] to be equivalent in results with the λ-process. Also Dirac's classical electron theory [11] is intimately connected to the λ-process: a "proper" and an "external"

field are defined as half the sum and half the difference respectively of the retarded and advanced field, and Dirac's method amounts to subtracting the proper field from the total field to obtain the force on the electron. To perform this subtraction within the framework of a Hamiltonian scheme it is essential to use again the λ-process [12].

The equation of motion of the electron in Lorentz's theory involves terms proportional to positive powers of the electron radius. Such terms are always neglected and the ensuing equation has therefore in principle an approximate character, albeit this approximation is a good one as long as only wavelengths $\gg a$ come into play. Dirac's equation, on the other hand, is a rigorous one. It can be obtained from the Lorentz equation[10] by inserting into the latter a negative mechanical mass so as to make the total mass equal to the observed mass and then letting the electron radius tend to zero. The occurrence of this sink of mechanical energy is closely connected with the appearance of the "non-physical solutions" of the equation of motion of the electron, i.e. solutions corresponding to accelerated motions even in the absence of external fields. The existence of these anomalies was pointed out by Dirac. However, in classical theory these solutions can simply be discarded [13].[11]

Action at a distance.[12] From the point of view of self-energy and self-stress this theory, recently worked out in detail by Wheeler and Feynman [15], is closely related to Dirac's classical theory [11] in that the fields are again invariantly divided into proper and external fields and the assumption is introduced that the proper field of a given electron shall not act on the particle by which it is produced. As a result of this postulate, the concepts of self-energy and self-stress do not appear. The theory is distinct from Dirac's theory in that a formalism is developed, mathematically alternative to Dirac's field theory, in which the electromagnetic field quantities no longer appear explicitly. In this scheme the electromagnetic field thus plays an entirely subordinate role with no degrees of freedom of its own. In particular Wheeler and Feynman show that this theory also allows one to account for the radiative reaction when one considers a universe containing sufficiently many charged particles to absorb completely all the radiation proceeding out of the source. Radiative reaction appears from this point of view as a consequence of statistical mechanics rather than a feature of pure electrodynamics. The elementary interaction between particles is half advanced and half retarded, that is, symmetric in time. Schoenberg

[10] This was pointed out by McManus and kindly communicated to me by Professor Peierls.

[11] It was shown by Eliezer [14] that there are no physical solutions for the system of an electron bound to a proton, i.e. no solution describing in this case an electron spiralling inwards around the proton. This interesting result has in itself of course no further physical implications as a classical system of this kind does not correspond to anything existing in nature.

[12] I am indebted to Professor Wheeler for discussions on this subject.

has considered a formulation [16] in which one has—in addition to such half-advanced half-retarded forces describable purely in terms of action at a distance—also a field with degrees of freedom of its own.

Two-field theories. By assuming that the electron is the point source not only of the electromagnetic field, but furthermore of a second field—which necessarily has to have a short range of influence—one tries to obtain relativistic compensation of the electromagnetic forces by the forces due to this new field. This balancing can be achieved by assuming the short range field to be of the scalar type [17]. Another possibility is to take this field to be of the vector type provided one subtracts rather than adds the corresponding energy momentum tensor from the electromagnetic one [18], thus running into well-known difficulties of non positive definiteness.

Non-linear electromagnetic theories. In this approach, initiated by Born [19], one starts from a relativistic and gauge invariant action function of such a form that the field equations derived from it are no longer linear. The action function is of course so chosen that in the limit of weak fields the Maxwell-Lorentz theory shall hold. From an investigation of the spherically symmetric electrostatic problem of a point source it can be shown that for a point electron both conditions of finite self-energy and zero self-stress can be satisfied. The latter conditions combined with those of invariance are not sufficient to determine uniquely the action function, several forms of which have been investigated [20]. For more complicated systems the solution of the field equations becomes exceedingly difficult. Even for the case of two point charges at rest no exact solution is known, although it has been shown that for an arbitrary number of point charges present the fields are uniquely determined by the charges [21]. For the case of one magnetic dipole it can be shown that no solution with finite energy exists [22], at least when one uses Born's original "square root Lagrangian."

Born links up the non-linear theory with a unitary program, by which here is meant that m_o shall be zero, *i.e.* that energy and momentum of the electron shall be entirely of electromagnetic origin.[13] From the point of view of the requirement of relativistic compensation, the unitarity is of course not a necessary condition.

[13] This is a weaker concept of unitarity than that of the program of Mie. Mie tried to develop a covariant (but not gauge invariant) formalism in which not only the electromagnetic energy and momentum are finite but also the fields are regular everywhere with the ultimate aim to derive the equations of motion of the electron from the field equations alone [23]. In the case of an electric point charge the D-field of Born is still singular. It was kindly pointed out to me by Professor Pauli that such field singularities occur in every gauge invariant non-linear electromagnetic field theory, in contrast to Mie's theory with absolute potentials, and that the fact that the place of the point charge is still an exception of the law of nature div $\mathbf{D} = 0$, which has to be replaced by div $\mathbf{D} = -4\pi_o\delta(\mathbf{x} - \mathbf{x}_0)$, makes it impossible to derive the law of motion for the charged particles from the field equations alone.

has considered a formulation [16] in which one has—in addition to such half-advanced half-retarded forces describable purely in terms of action at a distance—also a field with degrees of freedom of its own.

Two-field theories. By assuming that the electron is the point source not only of the electromagnetic field, but furthermore of a second field—which necessarily has to have a short range of influence—one tries to obtain relativistic compensation of the electromagnetic forces by the forces due to this new field. This balancing can be achieved by assuming the short range field to be of the scalar type [17]. Another possibility is to take this field to be of the vector type provided one subtracts rather than adds the corresponding energy momentum tensor from the electromagnetic one [18], thus running into well-known difficulties of non positive definiteness.

Non-linear electromagnetic theories. In this approach, initiated by Born [19], one starts from a relativistic and gauge invariant action function of such a form that the field equations derived from it are no longer linear. The action function is of course so chosen that in the limit of weak fields the Maxwell-Lorentz theory shall hold. From an investigation of the spherically symmetric electrostatic problem of a point source it can be shown that for a point electron both conditions of finite self-energy and zero self-stress can be satisfied. The latter conditions combined with those of invariance are not sufficient to determine uniquely the action function, several forms of which have been investigated [20]. For more complicated systems the solution of the field equations becomes exceedingly difficult. Even for the case of two point charges at rest no exact solution is known, although it has been shown that for an arbitrary number of point charges present the fields are uniquely determined by the charges [21]. For the case of one magnetic dipole it can be shown that no solution with finite energy exists [22], at least when one uses Born's original "square root Lagrangian."

Born links up the non-linear theory with a unitary program, by which here is meant that m_o shall be zero, *i.e.* that energy and momentum of the electron shall be entirely of electromagnetic origin.[13] From the point of view of the requirement of relativistic compensation, the unitarity is of course not a necessary condition.

[13] This is a weaker concept of unitarity than that of the program of Mie. Mie tried to develop a covariant (but not gauge invariant) formalism in which not only the electromagnetic energy and momentum are finite but also the fields are regular everywhere with the ultimate aim to derive the equations of motion of the electron from the field equations alone [23]. In the case of an electric point charge the D-field of Born is still singular. It was kindly pointed out to me by Professor Pauli that such field singularities occur in every gauge invariant non-linear electromagnetic field theory, in contrast to Mie's theory with absolute potentials, and that the fact that the place of the point charge is still an exception of the law of nature div $\mathbf{D} = 0$, which has to be replaced by div $\mathbf{D} = -4\pi_e\delta(\mathbf{x} - \mathbf{x}_0)$, makes it impossible to derive the law of motion for the charge particles from the field equations alone.

III. Non-relativistic Quantum Theory

(a) *Zero point energy and quantum self-energy*

While from the foregoing it has become apparent that a classical approach cannot fully bring out the nature of the fundamental problems of elementary particle physics since they are essentially of a relativistic character, equal caution is needed in judging the scope of attempts in which the quantum of action has not been taken into account from the outset. This inadequacy of the correspondence principle approach will become especially clear in considering relativistic quantum theory. As a first orientation it may be useful, however, to consider in a non-relativistic approximation the new features of a quantum treatment of elementary particle problems as compared to the classical situation. Thus we start from the Hamiltonian (2), as Dirac did in his first formulation of a quantum theory of dispersion [4b], quantizing the electromagnetic field in the now well-known way. The non-relativistic aspects of the theory in this form refer only to the treatment of the electrons, not to that of the fields.

Let us now consider a system consisting solely of one electron with momentum zero. Thus in particular all occupation numbers of the (transverse) photon states are zero. The energy of this system contains in the first place the electrostatic energy $\mu_{rel}c^2$ as comprised in U (cf. footnote 3). But this is not all: indeed following the usual approximation procedure we consider the coupling terms between electron and radiation field (second and third term in (2)) as a perturbation and first solve the eigenvalue problem without these couplings. For the state in question, this procedure yields $\mu_{rel}c^2 + \frac{1}{2}\Sigma h\gamma$ for the energy, the second term being the zero point energy of the radiation field. A simple calculation (neglecting the recoil of the electron) then shows that up to the second order in e, there is a contribution arising from the $\mathbf{A}^{\circ}{}_{\perp}{}^2$-term which, considering only those Fourier components of the radiation field whose wavelengths are $\geq a$, is $e^2\hbar/\pi m_o c a^2$. Hence to the approximation indicated the energy of the system is

$$\mu_{rel}c^2 + \frac{e^2\hbar}{\pi m_o c a^2} + \frac{1}{2}\sum h\gamma = \mu_{qu}{}^{NR}c^2 + \frac{1}{2}\sum h\gamma,$$

$$\mu_{qu}{}^{NR} = \mu_{rel} + \frac{e^2\hbar}{\pi m_o c^3 a^2} \tag{8}$$

Here $\mu_{qu}{}^{NR}$ is the electromagnetic self-energy on non-relativistic quantum theory if we define the self-energy as the difference in energy of the states "vacuum plus one electron" and "vacuum," as a consequence of which the zero point energy of the field drops out. Quite apart from the meaning one wishes to attach to the possibility of formally transforming away the zero point energy of the radiation field in a relativistically invariant way [24, 25],

this would not affect the self-energy in as far as it is the zero point amplitude rather than the zero point energy of the electromagetic field, which causes the quantum contribution to μ_{qu}^{NR}.

Thus the revision of the vacuum concept in quantum as compared to classical theory introduces new features in the electron problem which lie beyond the grasp of arguments of correspondence with classical physics.[14] In this connection it may be noted that, owing to the occurrence of the quantum terms in (8), the notion of the so-called "classical electron radius": e^2/mc^2 would seem to lose its original meaning in so far as it is linked up with an identification of μ_{cl} or μ_{rel} with the experimental mass m. That such an identification cannot even have an approximate meaning follows from the circumstance that the second term in (8) cannot be considered as a correction to the first one. In fact, the former is $\sim \hbar c/e^2$ times the latter (cf. (4a)). We will see later that in the relativistic quantum theory the situation differs considerably from that in the present non-relativistic version, but it does remain true also in the relativistic theories that no connection exists between m and quantities like μ_{rel} (cf. p. 28). More generally we may conclude that it has little meaning to define the structure of the electron in a classical way; the problem of the intrinsic properties of the electron is fundamentally of a quantum nature.

Of course the second term in (8) does not constitute the entire quantum contribution to the self-energy, as there will also occur terms $\sim e^4, e^6 \ldots$ if the higher approximations are taken into account. In this connection it may be recalled that the use of the approximation procedure of a power series development in $e^2/\hbar c$, as applied to the interaction of electric particles and radiation fields in the low energy region, is often justified by a correspondence reasoning. The argument is that from classical theory we are allowed to conclude that the reaction of radiation on the system by which it is emitted is relatively small if low frequencies only are involved. As has been pointed out above, however, correspondence would precisely seem to fail as regards the self-energy, where in fact no clear cut distinction can be made between ponderomotive force and radiation reaction. Thus in the present stage the power series development has not been proved to be an adequate mathematical tool for the approach of questions of the self-energy type.[15]

[14] This is for example clearly demonstrated in the treatment by Pauli and Fierz [26] of what is sometimes misleadingly called the infra-red catastrophe. Starting from a formalism in which the "classical structure" is eliminated, these authors find an expression for the cross section of the radiative scattering of an electron in the region of small energy loss which nonetheless still depends on the electron radius a. This reappearance of a is quite analogous to the fact that if one eliminates μ_{rel} one will still have according to (8) a self-energy involving a due to the quantization of the electromagnetic field. For a further discussion of the problem of radiative scattering cf. pp. 34 and 35.

[15] In this connection an instructive result was obtained by Kramers [27], who by taking into account radiation reaction in a semi-classical way obtains an expression

(*b*) *The correspondence argument in quantum field theory* [28]

Thus the synthesis of the field concept and quantum theoretical considerations in as far as elementary particle problems are concerned has yielded results which have no counterpart in the classical domain. On the other hand one can deal strictly on correspondence lines with the application of the quantum electrodynamical formalism to the problems of observability of electromagnetic field quantities, under the important restriction that one may look apart from the atomic structure of the measuring instruments. This has been shown in a penetrating analysis by Bohr and Rosenfeld [28], starting from the following considerations.

The components of electric and magnetic field strengths do not always commute, depending on what pair of components is chosen and on the relative position of the space-time points to which these components refer. Such non-commutativity generally gives rise to a complementary indeterminateness of the space-time point functions concerned. However, the only observables in the true sense are space-time averages of field components, for in observing field strengths one makes use of test bodies of finite extension and the actual measurement takes a finite time interval (e.g. in measuring an electric field one observes the momentum of a testbody at the beginning and the end of such interval). It is then shown that the two space averages of field components of the same kind (i.e. both electric or both magnetic) always commute if the time intervals coincide, while the time averages of two components of different kind always commute if the space domains coincide. From this result it follows in particular that all averages of components over the same space-time domains are independently observable with arbitrary sharpness. If the space time intervals do not overlap, there occur complementary uncertainties in these mean values. Then the analysis of idealized measurements leads to reciprocal limitations which are exactly those derived from the commutation relations.

In considering how a given measurement actually can be made with given latitude, e.g. how the complementary indeterminacy of position and momentum of the testbody itself affects the situation, it is important to specify the nature of this body. Thus Bohr and Rosenfeld show that if the test body is taken to be adequately heavy and homogeneously charged with adequately large charge density, quantum electrodynamics does not impose limitations additional to those of classical theory on the definability and observability of the mean values of field quantities over any finite space

---

$\frac{\hbar c}{a}\tan^{-1}\frac{2e^2}{3mc^2a}$ for the quantum self-energy. For example, developing the $\tan^{-1}$ as a power series for $a \gg \frac{e^2}{mc^2}$ gives as a first term the quantum term given in (8) and provides a simple instance of fallacious conclusions which may be drawn from illegitimate expansions.

time region. In particular it turns out that the radiative properties of accelerated charged test bodies do not (as was sometimes believed [29]) affect the uncertainty relations involving the momentum and position of that body.

IV. Relativistic Quantum Theory

(a) *General features; the renunciation of the one-body concept*

Thus far we have endeavored to penetrate into the elementary particle problems either along the lines of relativistic considerations, leaving out all quantum phenomena, or the other way round. Proceeding along either direction features presented themselves which had no counterpart in the classical picture from which we started. Now we shall have to examine the situation from a point of view in which both relativity and quantum theory are taken into account. As we shall see, this synthesis has not yet been achieved in a satisfactory way, but already at the present stage of development it brings into evidence new features of the elementary particle problems which, again, have no correspondence with the notions we have developed so far.

One of the main features of relativistic quantum theory is indeed that the very concept of "one" elementary particle loses its original meaning. Thus in the first attempt at constructing a relativistic invariant wave equation, the Schrödinger-Gordon-Klein equation, difficulties were encountered in trying to define suitably a probability density satisfying the requirements of positive definiteness, normalizability and relativistic invariance [30]. Some time later Pauli and Weisskopf [31] showed: (a) that starting from this equation a theory of spinless charged particles could be developed (on the lines of Bose-Einstein field quantization) provided one gives up the notion of probability density in configuration space as used in non-relativistic quantum theory, (b) that due to the coupling of these particles with the electromagnetic field the total number of such particles is no longer an observable due to creation and annihilation processes.

One is led to a similar renunciation of unambiguously defining a one particle problem by considering a different approach to the relativistic quantum problem which was developed in the meantime; in 1928 Dirac published his famous wave equation,

$$(c(\boldsymbol{\alpha} \cdot \mathbf{p}) + \beta m_o c^2)\psi = i\hbar \frac{\partial \psi}{\partial t}, \tag{9}$$

surely one of the most basic equations in our present picture of the physical world. $\boldsymbol{\alpha}$ is the velocity vector matrix whose components α_x, α_y, α_z anti-commute with each other and with β while $\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1$ [33]. The equation (9), obtained by a linearization of the relativistic quadratic relation between energy, momentum, and mechanical mass m_o of a particle,

is itself properly covariant. Moreover a probability-density can be defined in accordance with the requirements enunciated above. It is well known how the existence of the electron spin follows from (9). Furthermore the analysis of the electromagnetic interactions gives the correct value for the magnetic moment of the electron while also a satisfactory theory of the atomic fine structure could be given along these lines.[16] The application of the Dirac theory to Compton scattering leads to the well known Klein-Nishina formula [34] which has been checked thus far up to photon energies $\sim 10 MeV$ and at least in this domain is in fair agreement with experiment. Likewise the treatment of deceleration radiation [34] has confirmed the basic ideas of the Dirac theory.

These consequences of the theory are to a great extent independent of a fundamental novelty, viz. the occurrence of negative energy eigenvalues of (9) which cannot generally be discarded in quantum theory, as was possible classically in the equation $E = \pm c\sqrt{p^2 + m_0^2 c^2}$. In order to prevent all electrons in the world from falling eventually into the "sea" of negative energy states, Dirac later introduced what amounts to a redefinition of the vacuum [35]. In the absence of external agents all negative energy states are assumed to be filled up in accordance with the Pauli exclusion principle. The density of energy, momentum, and charge of this state is taken to be zero. Hence, roughly speaking, one has to define the energy (or any other observable) of any given dynamical system as the difference of the energy of the "vacuum + our system" and the vacuum energy itself. The precise procedure and meaning of such subtractions is the subject of various papers to be briefly discussed below. This new interpretation does not affect the successful conclusions of the Dirac theory mentioned above.[17] However, the notion of probability density in the one body problem, while making sense in the primitive form of the Dirac theory (in the following referred to as "one particle theory"), again—as in the case of spinless particles mentioned before—loses its meaning in the reinterpreted theory. This feature is intimately connected with the possibility of pair creation: a "hole" in the sea of negative energy electrons has to be interpreted as an anti-particle, i.e. a particle with charge, energy, and momentum of sign opposite to that of the missing electron. On this picture the transition of a negative energy electron to a positive energy state is equivalent to the creation of a pair of oppositely charged particles.

Originally Dirac believed that the anti-particle was the proton. However, it was pointed out soon afterwards by Oppenheimer [37] and by

[16] We will see later (cf. p. 33) that both the interpretation of the magnetic interaction as well as the fine structure needs further refinement.

[17] With regard to the Klein-Nishina formula and the deceleration radiation cross section, this holds strictly speaking only for the first non-vanishing approximation in which they are usually computed [36]. Also the refinements in the treatment of the fine structure and the magnetic moment depend on whether one uses hole theory or one-particle theory (cf. p. 33).

Tamm [38] that this interpretation would lead to a two-quanta annihilation of hydrogen atoms at a rate which is inconsistent with the known stability properties of matter. Oppenheimer remarked [37] moreover that this difficulty vanishes if one gives up the particle-anti-particle relation between electron and proton, but rather considers them to be independent. The necessity of this view was also stressed by Weyl [39] who on symmetry arguments came to the conclusion that particle and anti-particle should have the same mass. Following these suggestions Dirac then stated [40] "a hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron." Two years later Anderson [41] announced the discovery of the positon,[18] found in Wilson chamber pictures of cosmic radiation, a finding soon confirmed by Blackett and Occhialini [42] from similar experiments. Thus Dirac's ingenious idea of the theory of holes found its experimental support.[19]

It should be noted that it often makes a good deal of difference whether we consider the positon as the anti-particle of the negaton or as an independent elementary particle whose behavior is described by the Dirac equation. Thus it was pointed out by Bhabha [60] that the cross section for a collision between a negaton and a positon is very sensitive to exchange effects. It turns out, in fact, that if the positon is considered as an unoccupied negaton state of negative energy, one obtains a result essentially different from that we should get if we were to consider the positon as an independent positively charged particle in a state of positive energy. Experiments by Mrs. Ho-Zah-Wei [61] seem to give preliminary support to the theoretical result obtained by Bhabha on the hole theory picture.

Hole theory thus gives us not only qualitative predictions of such exchange effects and other new features like pair formation and annihilation

[18] In the following we will understand by negaton what was up till now called electron, i.e. the negatively charged particle (charge e), by positon the corresponding anti-particle. Henceforward "electron" will be used as a collective word for negaton and positon. This convention is in accordance with the "Resolution on names for elementary particles" adopted by the Cosmic Ray Commission of the International Union of Physics in its meeting at Cracow on October 9, 1947. From this resolution we quote: "The . . . Commission . . . recognizes the general use of the term *electron* to denote both positive and negative particles of electronic mass . . . looks with favor upon the terms *positon* and *negaton* as means to distinguish between the two signs of charge."

[19] In order to avoid the conceptual difficulties connected with the introduction of infinite seas, alternative treatments have been given in which negaton and positon occur in a more obviously symmetric way. Such a formulation was first proposed by Fock [43]. Heisenberg [52] uses a similar formalism. The connection between charge symmetry and the representations of the Dirac matrices is discussed by Majorana [44], Racah [45] and Kramers [46] and surveyed in an article by Pauli [47]. Of course in such formalisms too, one has to introduce subtraction prescriptions. Incidentally it may be remarked that Heisenberg's treatment [48] of not completely filled shells of atomic electrons provides an instructive example for the use of (a finite number of) "holes" as endowed with particle properties.

but moreover we can make qualitative inferences as to the cross section for such phenomena by following the prescription of considering only the contribution to the relevant transition probabilities in the lowest non-vanishing order of approximation in powers of $e^2/\hbar c$. We find by this prescription predictions which within the present experimental accuracy agree with experience, just as we did for the effects approximately describable by the one particle theory (deceleration radiation, etc.). It need hardly be recalled here that in the present stage the justification for the mode of calculation just mentioned lies in the fact that "it works" rather than in its logic. Indeed all higher order contributions to the matrix elements involved are infinite, and it will be one of the important requirements of future theories to elucidate how it is possible that our present formalism, while clearly inconsistent, yet allows us to codify such a considerable number of experimental data by using computational methods which at present we can only consider as arbitrary. In this connection it should be noted that in relativistic quantum theory also we meet the same still arbitrary rule that m_o in (9) shall represent (at least to a good approximation) the experimental mass of the electron (cf. also pp. 28–31 below). In fact our present quantum theoretical picture is of a thoroughly dualistic nature.

(*b*) *The subtraction problem: general considerations* **[49, 50, 51, 52, 54, 55]**

Let us now consider in more detail the problems of subtraction which are encountered in hole theory. Suppose we have in space, apart from the vacuum electrons, a system of electrically charged particles, generally consisting of electrons and particles of other kind. The total charge and current distributions in space-time are assumed to generate the electromagnetic field in accordance with equations, such as div $\mathbf{E} = 4\pi\rho$, etc. (ρ = charge density, E = electric field). We now wish to subtract from ρ an amount ρ_o such that $\rho - \rho_o$ is finite and is linked up with the electric field by

$$\text{div } E = 4\pi(\rho - \rho_o(F)) = 4\pi\rho_{eff}. \tag{10}$$

The amount ρ_o to be subtracted will in general depend on the external electromagnetic field itself, as symbolically indicated by writing $\rho_o(F)$, and we have more precisely as first requirement:

(a) $\rho - \rho_o(F)$ and $\mathbf{s} - \mathbf{s}_o(F)$, the effective current density, shall be finite for any F and such that if $F = 0$, $\rho - \rho_o(F) = 0$, $\mathbf{s} - \mathbf{s}_o(F) = 0$. It may be directly noted here that in all treatments of the subtraction problem given thus far one defines the external field such that it also comprises the electromagnetic field produced by the sea distribution itself, i.e. one always starts from a zeroth approximation in which these electrons are considered as free. Equations like (10) imply of course the polarizability of the vacuum distribution, and it is quite essential that this polarizability be field dependent. It was in fact investigated by Peierls [49] whether finite effective densities could be obtained by assuming that ρ_o, s_o were simply the densities

corresponding to the field free situation,[20] and this turned out not to be the case.

Of course the prescription (a) does not uniquely determine ρ_{eff} and s_{eff} as the subtraction of one infinity from another is quite an ambiguous affair. Thus Peierls [49] showed that in a semi-classical picture (describing the vacuum distribution as a Thomas-Fermi gas) it is even possible to get no polarizability at all; but, as he pointed out, this conclusion by no means precludes polarizations due to quantum effects. Another possibility amounts to first determining the contribution to ρ and s due to electrons moving with a certain momentum and then subtracting the singular part of that contribution. This procedure was considered by Peierls and by Furry and Oppenheimer [50] and was recognized by these authors to be in general inadmissible, as it is not gauge invariant. This particular difficulty can be overcome, as was done by Peierls, by replacing the momentum in this procedure by $\mathbf{p} - \frac{e}{c}\mathbf{A}$ where $\mathbf{A}$ is the vector potential. Whatever the difficulties in detail, it is clear in principle that invariance arguments provide a guide in how to subtract. Thus:

(b) The subtraction shall be gauge invaraint; i.e. ρ_o, $\mathbf{s}_o$ must satisfy conservation law.

(c) The subtraction shall be covariant, i.e. $\rho - \rho_o$, $\mathbf{s} - \mathbf{s}_o$ must form four vector. Hence ρ_o, $\mathbf{s}_o$ must form a four vector.

Nevertheless, these three conditions do not lead to a unique solution. In the last analysis arguments of simplicity have been used in making a choice of expressions for charge and current. Nor is any exact method of tackling the problem known which due to its great intricacy has been treated only by approximation methods.[21]

The method first proposed by Dirac [51] is that of the Hartree approximation. Here it is supposed that each electron has its own individual wave function and that it moves in a definite electromagnetic field which is the same for each particle. A density matrix $(\mathbf{x}', t'|R|\mathbf{x}'', t'')$ is defined such that (always within the Hartree approximation) our $\rho(\mathbf{x}, t)$ mentioned above is equal to $e\ lim_{\mathbf{x}'=\mathbf{x}'', t'=t''}(\mathbf{x}', t'|R|\mathbf{x}'', t'')$. Dirac considered first the field free case where R can be computed exactly and is free of singularities for $\mathbf{x}' \neq \mathbf{x}''$, $t' \neq t''$, but contains certain terms which become singular on the light cone. Then the case of the presence of fields is considered; it is assumed that R has the same types of singularities as in the field free case with coefficients depending on the external field parameters. These singularities shall then constitute the subtractive terms; by performing all calculations with non-zero "off-diagonal distance"[22] one can give a precise meaning to every term, and by introducing the prescription that the off-diagonal distance

[20] *Loc. cit.*, equation (7′).

[21] This holds for the treatments in [49, 50] as well as those to be discussed below.

[22] This quantity is defined as $\sqrt{(\mathbf{x}' - \mathbf{x}'')^2 - c^2(t' - t'')^2}$.

shall tend to zero only after the subtractions have been performed, one obtains an unambiguous subtraction prescription. Dirac gives the equations which the coefficients of the singularities must satisfy and discusses the general aspects of their solutions.

This treatment is applied by Heisenberg [52] to the case in which one admits in the coefficients in question only terms quadratic in the field strengths and linear in their first derivatives.[23] This author also shows that this subtraction is consistent with the law of a conservation law of electric charge. Furthermore the covariance of the scheme is discussed.[24]

Heisenberg also treats the same problems by using the Jordan-Wigner representation [53] of quantized electron amplitudes. This method is also used by Furry and Oppenheimer [50]. It may be noted that the use of this so-called second quantization does not imply any new physical assumptions. One might indeed just as well use many electron wave functions in configuration space, provided they are properly anti-symmetrized. However, such a procedure is very cumbersome, especially for the discussion of effects of higher order than the second in e. The great practical advantage of the second quantization method is precisely that it takes account of the exclusion principle in a simple and elegant manner.

Having performed the subtractions in the density matrix, one can find not only the expressions for the effective charge and current density, but also the effective energy-momentum tensor density, cf. loc. cit. [52] equ. (16). The 44-component of the latter is then the new Hamiltonian. After subtraction, Heisenberg's Hamiltonian has the form $\sum_{n=0}^{4} H^{(n)}$, where $H^{(n)} \sim e^n$; cf. loc. cit. equ. (57)–(61). In as far as $H^{(0)}$ and $H^{(1)}$ are concerned, Heisenberg's results are identical with those of Fock [43], Peierls [49], and Furry and Oppenheimer [50].

It has been shown by Weisskopf [54] that the mathematical methods of Dirac and Heisenberg are from a more directly physical point of view consistent with the following prescriptions, although not thereby uniquely fixed:

1. The vacuum energy in field free space shall be omitted.
2. The charge and current density in field free space shall be omitted.
3. There shall be no correction terms in the energy density which are of the second order in the field strengths, i.e. there shall be no constant field independent electric and magnetic polarizability of the vacuum.

In the treatments of the subtraction problems discussed thus far, the

[23] This assumption, made for reasons of simplicity, is of course not the most general one.

[24] The existence of conservation laws for energy and momentum is only proved, though, for that part of the energy momentum tensor corresponding with $H^{(0)}$ and $H^{(1)}$ defined below.

electromagnetic field quantities are essentially considered to be c-numbers. Wherever use is made of quantum electrodynamics (as e.g. in Heisenberg's treatment [52] of the self-energy of a photon) it is assumed without proof that the customary commutation relations for the electromagnetic field may be used. It has to be verified however whether these commutation relations and the subtractions which have been performed are consistent with each other. And this brings us to a new source of difficulties.

We have indeed to inquire whether we have found in (*a*), (*b*), and (*c*) all the conditions which must be fulfilled to perform in a consistent way the subtractions discussed hitherto. Now, especially in view of a transition to quantum-electrodynamics, it is to be required that the Hamiltonian system from which we started shall still be Hamiltonian after the subtractions have been performed:

(d) The subtraction must be a canonical transformation.

Heisenberg's formalism does not satisfy (d), as has been pointed out by Serber [55]. Let us call the initial Hamiltonian H_I and the quantity into which it is transformed by the subtractions H_{II}. Thus H_{II} is Heisenberg's Hamiltonian. Now by means of H_I we can define the canonically conjugate variables and quantize the electromagnetic field in the usual way. Of course the commutation rules are consistent with the equations of motion, the question is whether this consistency still holds after subtraction. For example, the values of the commutators may not change with time; but Serber shows that, as a consequence of the subtractions, the time derivative of the commutator of the vector potential **A** and the electric field **E** is not only different from zero, but even singular. It may be shown that this inconsistency would disappear if we would consider the electromagnetic quantities occurring in the terms $H^{(2)}$, $H^{(3)}$ and $H^{(4)}$ of H_{II} as c-numbers. But the meaning of such an "approximate quantization" is obscure since there is no domain (of frequencies, say) in which the terms $H^{(2)}$, $H^{(3)}$, $H^{(4)}$ in H_{II} may be neglected in comparison with $H^{(0)}$ and $H^{(1)}$. In fact, as we shall see in discussing the scattering of light by light (p. 24) we would be in trouble already in the long wave length region if we would make the arbitrary cut in the formalism of taking seriously the terms $H^{(0)}$ and $H^{(1)}$, but not the others. As pointed out by Serber, no unambigous meaning can under these circumstances be attached to Heisenberg's computation [52] of the self-energy of a photon—a particular challenge to any theory, as the photon should, on our present picture, have no inertia.[25]

While thus Heisenberg's formalism satisfies the conditions (a), (b), (c) but not (d), one can also conceive a scheme which is in accordance with (a), (d) but not (b), (c), i.e. in which one has a canonical but not an invariant

[25] This self-energy is not only due to the virtual creation and annihilation of pairs of electrons but also to virtual formation of other particles which can actually be created in pairs by electromagnetic radiation.

theory. Such a scheme appears to be possible [55] by giving the off diagonal distance a fixed but arbitrary value different from zero.

Serber [55] also investigates an alternative possibility of interpretation: one can consider H_I as a kind of pre-Hamiltonian and H_{II} as the true Hamiltonian from which the real equations of motion and the actual commutation relations should be derived. However, in performing the transition from H_I to H_{II} one makes, for example, use of the relation $\mathbf{E} = -\dot{\mathbf{A}}$ while the use of H_{II} as Hamiltonian would yield $\mathbf{E} = -\dot{\mathbf{A}}$ + additional terms, incompatible with the previous equation; moreover the additional terms are singular. Thus it seems that when one eliminates the infinities at one place, they turn up again somewhere else. No solution of the subtraction problem satisfying (*a*) to (*d*) has thus far been given, nor has its existence been proved. It should furthermore be pointed out that the difficulties just discussed are in principle independent of the self-energy difficulties, although a remedy for the one set of problems may of course shed new light on the other.

Added in proof. Recent work by Schwinger has brought important progress in formulating positon theory in accordance with the conditions (*a*)—(*d*). In Schwinger's canonical formalism the covariance and gauge invariance is exhibited at every stage.

(*c*) *The subtraction problem: special cases* [51, 52, 54, 57, 58, 59, 64, 65, 65a, 66, 70, 71, 72]

It is with these limitations in mind that one must examine those consequences of the subtractions which give rise to in principle observable deviations from classical electromagnetic theory. These deviations can conveniently be grouped into two classes.

First, for fields which are weak, but vary in space or time or both, the field equations are no longer the customary second order differential equations but contain higher order space-time derivatives. The modified equations are still linear, however. The second class comprises the effects due to the non-linear aspects of the theory, defined more precisely below. In dealing with either group one is often interested in steady state problems in which the fields are sufficiently weak and weakly varying to prohibit the actual creation of pairs, i.e. fields for which the potentials vary less than mc^2 over distances of the order of $\hbar/mc$ [56]. In terms of field strengths this means a limitation to values less than E_{crit} where

$$E_{crit} \sim \frac{m^2c^3}{\hbar e} \tag{11}$$

We will start with a survey of the linear modifications in Maxwell's equations. An interesting case here is that of the steady electrostatic field. (It was in fact in connection with an analysis of this problem by Dirac [51] that the subtraction problem was for the first time raised in its entirety). The treatment is relatively simple when the electric charges creating the

electrostatic disturbance are themselves not electrons. For then one may look apart from phenomena of exchange between external agent and vacuum distribution and may approximate as follows: Consider the academic situation of a static charge distribution $\rho_{ext}(\mathbf{x})$ in space in the absence of the vacuum distribution. The corresponding electrostatic potential Φ_{ext} is then given by the equation

$$\Delta\Phi_{ext} = -4\pi\rho_{ext}$$

The problem now is to find how this equation is modified due to the presence of the vacuum negatons. We then shall have an equation of the form

$$\Delta\Phi = -4\pi(\rho_{ext} + \rho')$$

where we have to perform subtractions to make ρ' finite. Of course ρ' depends on ρ_{ext} itself, as ρ_{ext} will polarize the sea. In order that this problem shall be truly electrostatic, it is necessary that Φ_{ext} shall not be too strongly varying in space, as otherwise pairs might be formed, and that Φ_{ext} shall be brought adiabatically to its actual value. Thus the field strengths are limited by (11).

This problem of the polarization of the vacuum has been treated by several authors. The essential points were first given by Dirac in his Solvay report [51]. A more detailed treatment is given by Heisenberg [52], who also briefly discusses the case of fields varying with time. A more complete treatment of varying fields is given by Serber [57], while a simplified mathematical derivation of Serber's results has been given by Pauli and Rose [58]. In all these papers the Hartree approximation method is used and the induced charge and (for varying fields) current distributions are computed up to the order e^2. To this same order the polarization has also been derived by Weisskopf [54] by using the representation of q-number electron wave functions and then applying ordinary perturbation calculus. This procedure leads to the same results as the previously mentioned method. For the purely static case one obtains for ρ' in this approximation [54]:

$$\rho' = \rho^{(0)} + \frac{e^2}{\hbar c}\sum_{n=0}^{\infty}\lambda_n\left(\frac{\hbar^2}{m^2c^2}\Delta\right)^n \cdot \Delta\Phi_{ext} \tag{12a}$$

Here $\rho^{(0)}$ is the charge density of the vacuum electrons in field free space, which is of course infinite and is subtracted, (Weisskopf's condition 2, see above). In the series development the λ_n are numerical constants; λ_0 is infinite, but the term with $n = 0$ must also be subtracted, according to Weisskopf's condition 3. As this term is directly proportional to ρ_{ext}, it may be considered with some optimism as renormalizing ρ_{ext} itself [54]. The λ_n with $n > 1$ are finite. These terms correspond to effects which, according to hole theory, should in principle be observable consequences of the presence of the vacuum electrons.

(9) can now alternatively be written as:

$$\Delta\Phi_{ext} - \frac{e^2}{\hbar c}\sum_{n=1}^{\infty}\lambda_n\left(\frac{\hbar^2}{m^2c^2}\Delta\right)^n\Delta\Phi_{ext} = -4\pi\rho_{ext} \qquad (12b)$$

From this form it is manifest that, as pointed out by Furry and Oppenheimer [50], hole theory leads to a deviation from Coulomb's law. The deviation from Coulomb's law and its influence on the Rutherford scattering and the displacement of atomic spectral lines has been discussed by Uehling [59] who shows that the correction to the Coulomb energy has the same sign as the Coulomb energy itself. For large distances between the charges the correction can be neglected; for too small distances (12b) loses its validity due to the limitations imposed on E_{ext} mentioned above. The nS-levels in a hydrogen-like atom are shifted downwards by an amount $\sim n^{-3}Z^2\alpha^3$ Rydbergs (Z = nuclear charge number, $\alpha = 1/137$). $P_2, D_2 \cdots$ levels are much less affected. Thus in particular the doublet separation would increase. The small deviation which has to be expected for proton-proton scattering can in principle hardly be disentangled from the scattering anomalies due to the nuclear forces.

It should be noted that in Uehling's work no account is taken of exchange effects. These calculations therefore hold only if neither of the two electric charges whose interaction is considered are electrons. Thus the evaluation of the deviation from Coulomb's law in the hydrogen atom, as performed by Uehling, needs refinement, for here one of the particles involved is an electron. Also the proton-proton interaction would have to be reconsidered in view of exchange effects, if we assume the existence of a sea of negative energy protons.

Thus far we have mainly considered such modifications of the Maxwell equations which give rise to effects $\sim e^2$ leading to equations of the type (12) which are still linear. If effects of higher order in e are taken into account, this linearity is no longer conserved. This implies that the superposition principle, characteristic of the classical electromagnetic theory as well as of non-relativistic quantum theory, no longer holds. Thus it is a consequence of positon theory that two beams of light interact with each other. As a first example of this interaction we have the possibility of pair creation by two photons of sufficient energy. The cross section of this effect was derived by Breit and Wheeler [62]. The computation of this effect as well as of its inverse, the two-quanta annihilation of a free negaton positon pair, amounts to a second order perturbation calculus in which the term $H^{(1)}$ of the Heisenberg Hamiltonian (cf. p. 19) occurs as the perturbation operator. In the discussion of these effects the terms $H^{(2)}$, $H^{(3)}$ and $H^{(4)}$ of Heisenberg do not come into play, since they do not depend on the occupation numbers of the electron states. A different situation is met, however, when we consider the scattering of light by light [63]. There again the interaction term $H^{(1)}$ gives a contribution in a

fourth order perturbation calculation, due to the virtual formation and annihilation of pairs. If this term only were taken into account we would find the paradoxical result of a non-vanishing cross section for zero frequency. However, it was pointed out by Euler and Kockel [64] and Euler [65] that, whereas the terms $H^{(2)}$ and $H^{(3)}$ do not contribute, the inclusion of $H^{(4)}$ reduces the cross section to a value which goes to zero for small frequencies. This result shows that the higher order terms of the Heisenberg Hamiltonian play a vital role in the present form of hole theory, as already remarked on p. 20. In the region of large wave lengths one obtains [64, 65][26] for the scattering cross section

$$\sigma \sim \left(\frac{e^2}{\hbar c}\right)^4 \left(\frac{\hbar}{mc}\right)^8 \frac{1}{\lambda^6} \qquad \text{(low frequencies)}$$

For high energies of the photons the cross section has been computed by Achieser [65a] who again uses the combined effect of $H^{(1)}$ and $H^{(4)}$ and obtains

$$\sigma \sim \left(\frac{e^2}{\hbar c}\right)^4 \cdot \lambda^2 \qquad \text{(high frequencies)}$$

Evidently the cross section will have a maximal value, $\sigma \sim (e^2/\hbar c)^4 \cdot (\hbar/mc)^2$, for $\lambda \sim \hbar/mc$. Moreover, the cross section for pair production by photon pairs is also small compared to $(\hbar/mc)^2$, viz. [62]

$$\sim \left(\frac{e^2}{\hbar c}\right)^2 \cdot \left(\frac{\hbar}{mc}\right)^2 \cdot \left(\frac{mc^2}{\hbar\omega}\right)^2 \cdot \lg \frac{\hbar\omega}{mc^2}, \qquad \omega = \frac{c}{\lambda}$$

The smallness of both cross sections in comparison with λ^2 means that the deviations from the superposition principle are very small indeed over the entire domain of wave lengths. This result would seem quite appealing in justifying the very approximation procedure by which it has been obtained. However it should not be forgotten that it follows from a theory which, as we have seen, is not consistent. Thus we meet again (cf. p. 17) with an example of an acceptable inference from a formalism which as it stands is not consistent.

In computing such consequences of the non-linearity of hole theory one can start from the linear field equations and push the customary perturbation theory up to adequately high order. One can also follow an alternative method, the idea of which can be understood by the following analogy [65]: the van der Waals forces between two atoms arise from virtual transitions of the atomic electrons to higher energy states. From a computation on these lines one can derive an effective potential energy which represents the action of these forces. Similarly one can proceed in the case of the scattering

[26] The expressions for σ refer to a system in which the photon momenta are equal and opposite.

of light by light and related problems, i.e. one can derive the results from an effective energy of the fourth order in the field quantities. However, the treatment of the van der Waals forces by means of a potential is only adequate as long as the kinetic energy is insufficient to provoke real electronic transitions. Likewise the elimination of the pair field via the introduction of effective energy terms is possible only as long as no actual pairs are created, i.e. as long as the field strengths are smaller than E_{crit}.

To find these fourth (or higher) order effective energies one has arguments of relativistic and gauge invariance as a guide. Indeed to the effective terms in the energy for which we are looking, there must correspond similar terms in the Lagrangian $\mathfrak{L}$. But we know that $\mathfrak{L}$ must be a gauge invariant scalar density. Thus one may start by adding to the customary expression for $\mathfrak{L}$ terms of the fourth, sixth, . . . order in the field strengths with as yet arbitrary coefficients (which depend on such constants as e, $\hbar$ and the mass of the electron), but such that these terms satisfy the invariance requirements. From this pseudo-Lagrangian we then find the energy density in the standard way. In the absence of negatons with positive energy and of positons, the terms additional to the ordinary Maxwellian energy density are then due to the vacuum negatons. The contribution of the latter to the energy density is given by (cf. loc. cit. [54] equ. (6))

$$\sum_i \psi_i\dagger(W_i - eV)\psi_i \tag{13}$$

where the summation goes over all states i of negative energy, characterized by their eigenfunction ψ_i and energy W_i. V is the potential of the external electrostatic field, if any. Now the whole problem consists of course in fixing the unknown coefficients mentioned above. Invariance arguments require that the coefficients be independent of the particular form of the external fields. Consequently one chooses for the fields simple cases in which the computation of (13) is not too complicated. Of course (13) is again infinite; subtractions are performed, and the coefficients are fixed by identification of the remaining finite part of (13) with the additional terms in the energy density written in terms of external field strengths.[27] This procedure was followed by Euler and Kockel [64] and Euler [65] to find the fourth order term describing the scattering of light by light. In the treatment by these authors the additional effective energy terms are taken to be a power series in e. In a later paper by Heisenberg and Euler [66] no such series development is used, and the problem of the additional effective terms in the region limited by (11) is solved exactly.

A more simplified and systematic treatment of the effective terms has been given by Weisskopf [54], who also discusses the same problem for the

[27] It should be noted that the linear modifications occurring in (12b) involve derivatives of the field strengths. These therefore lie beyond the procedure sketched above where it is assumed that the fields are so slowly varying that in the correction terms on the energy density all dependence on derivatives of field strengths may be neglected.

case of spinless charged particles. It should in fact not be forgotten that the various effects which are discussed in this section all have their counterpart in the theory of Bose-Einstein particles. Thus there is an infinite polarization of the vacuum, as was pointed out by Pauli and Weisskopf [31]. Furthermore two interacting beams of light of sufficiently high frequency can create pairs of Bose-Einstein particles. Also there will be a contribution to the scattering of light by light due to virtual pair formation of such particles, etc. (cf. also Heisenberg [80]).

Another characteristic effect of the non-linear type is the coherent scattering of light by electrostatic fields [68, 69]. The actual calculation of this effect is very intricate and has as yet been performed only partially [70, 71].

Prescription 2 (see p. 19) implies that when we ask for the expectation value $\bar{\epsilon}_\Omega$ of the operator $\epsilon_\Omega = e \int_\Omega \psi^\dagger \psi dv$ for the total charge in a field free box with volume Ω, the answer is $\bar{\epsilon}_\Omega = (N_- - N_+)e$, where N_- and N_+ are the number of actually occupied positive energy negaton states and positon states respectively. Thus the charge per unit volume is on the average $\bar{\epsilon}_\Omega/\Omega$. The discussion of the fluctuations around this mean value requires special care. In fact, if we consider a "sharp" box, i.e. a box whose boundary represents an infinitely high potential wall for the enclosed electrons, and then ask for the fluctuations around the mean value of the charge density, the result is infinite. This is of course nothing but a special case of the Klein paradox: pairs will be created due to the potential jump at the boundary. We have therefore to do with a surface effect. In a "smooth" box, i.e. a region in space whose boundary is a smooth potential for the enclosed electrons, the fluctuations can be made arbitrarily small by adequately smoothing out these potential curves [72].

These related topics are dealt with in detail in a paper by Bohr and Rosenfeld which is in course of preparation. The main aim of this paper is to establish the domain of correspondence of relativistic quantum-theory with classical theory, similar to the correspondence argument for the non-relativistic quantum case (cf. section IIIb). A preliminary abstract of this work is included at the end of this paper.

(d) *The self-energy problem in relativistic quantum theory* [78]

The fluctuation phenomenon can also be formulated by saying that $\overline{\epsilon_\Omega^2} \neq \bar{\epsilon}_\Omega^2$. This effect is of particular interest for the discussion of the self-energy of the electron, for the electrostatic self-energy is the expectation value of an operator closely connected with ϵ_Ω^2. Evidently the exchange effects of hole theory introduce new aspects of the self-energy problem. In order to see clearly the contrast with non-relativistic quantum theory, it is advantageous first to consider briefly the academic case of the one particle theory.

In this case the Hamiltonian of the system consisting of one electron and its electromagnetic field is given by[28]

$$H = \int \psi^\dagger (c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m_o c^2)\psi dv + V + \frac{1}{8\pi}\int (E^2 + H^2)dv + \int \phi(\operatorname{div} E - 4\pi\rho)dv$$

$$V = \int (\rho\phi - \mathbf{s}\cdot\mathbf{A})dv \tag{14}$$

with $\rho = e\psi^\dagger\psi$, $\mathbf{s} = c\psi^\dagger\boldsymbol{\alpha}\psi$; ϕ and $\mathbf{A}$ are the scalar and vector potential respectively. The computation of that part W of the self-energy which is proportional to e^2 is performed by considering V as a perturbation and using the well known formula for the second order contribution to the energy:

$$W = \sum_{ij}{}' \varepsilon_j \frac{(0,\,0|V|i,\,\mathbf{k}_j)(i,\,\mathbf{k}_j|V|0,\,0)}{E_o - E_i - \hbar\nu_j} \tag{15}$$

Here $(0,\,0|V|i,\,\mathbf{k}_j)$ is the matrix element for a transition from the initial state in which an electron with momentum $\mathbf{p}_o$, energy E_o and no quanta are present to an intermediate state in which the electron has a momentum $\mathbf{p}_i$ and energy E_i and one quantum with momentum $\hbar\mathbf{k}$ and state of polarization j is present.[29] On account of momentum conservation $\mathbf{p}_o = \mathbf{p}_i + \hbar\mathbf{k}_j$, whence the intermediate states may be characterized by $\mathbf{p}_i$, spin and sign of energy of the electron, and state of polarization of the quantum. The computation of W has been discussed by Oppenheimer [75] (contribution of the term $-\int \mathbf{s}\mathbf{A}_\perp dv$ for electrons bound in an atom), Waller [76], (same for free electron), and Weisskopf [77, 78]. For $\mathbf{p}_o = 0$, $E_o = m_o c^2$ the result is, cutting off provisionally in momentum space at a large $P = \frac{\hbar}{a}$,

$$\frac{W}{c^2} = \mu_{rel} + \frac{e^2\hbar}{\pi m_o c a^2} \tag{16}$$

which is identical with the non-relativistic result (8). Of course, the use of a non zero a in relativistic theory has no actual meaning any more, as rela-

[28] The last integral on the right hand side has the value zero. The occurrence of this term is connected with the special role of the relation div $\mathbf{E} = 4\pi\rho$ in the Hamiltonian scheme.

[29] It should be noted that also as regards the electrostatic self-energy—which is the expectation value of $\int(\rho\phi - \mathbf{s}\mathbf{A}_{11})dv$ where $\mathbf{A}_{11}$ is the longitudinal (curl free) part of the vector potential—we use (14) for convenience instead of following the more customary procedure of separating first the Coulomb energy operator by means of a canonical transformation. This means that j takes the values 0, 1, 2, 3, where $j = 2, 3$ denotes the usual transverse photons, $j = 0, 1$ the "ghost quanta" corresponding to the quantization of ϕ and of A_{11} respectively. In (15) $\epsilon_j = +1$ for $j = 1, 2, 3$; $\epsilon_j = -1$ for $j = 0$. For details of this mode of quantizing the electromagnetic field see ref. [73].

tivistic quantum theory in its current form is based on the concept of the point electron [79]. Equation (16) therefore serves only to indicate that W involves a quadratic and a linear infinity (as $\mu_{rel} \sim a^{-1}$).

Turning now from one particle theory to hole theory, one defines the self-energy proportional to e^2 of one negaton (of positive energy) as the difference in energy of the system "vacuum + one negaton": W_{vac+1}, and the energy of the vacuum: W_{vac}. Describing the self-energy, as in (15), by means of virtual emission and absorption of quanta, one sees that the difference of W_{vac+1} and W_{vac} is due to the fact that in computing W_{vac+1} the electrons in the sea may (as a consequence of the exclusion principle) not jump into the one actually occupied positive energy state. On the other hand in W_{vac+1} all transitions of our one negaton to other positive energy states appear which transitions are not present in W_{vac}. Thus it follows immediately that, in the notation of (15), we now have[30]

$$W_{hole\,th.} = W_{vac+1} - W_{vac} = \sum_{i,j} \varepsilon_j \frac{(0,\,0|V|i_+,\,\mathbf{k}_j)(i_+,\,\mathbf{k}_j|V|0,\,0)}{E_o - E_{i_+} - \hbar\nu_j} - \sum_{i,j} \varepsilon_j \frac{(i_-,\,0|V|0,\,\mathbf{k}_j')(0,\,\mathbf{k}_j'|V|i_-,\,0)}{E_{i_-} - E_o - \hbar\nu_j}. \quad (17)$$

Here $i_\pm$ denotes an electron state with momentum $\mathbf{p}_i$, specified spin and energy $E_{i_+} = \pm c\sqrt{p_i^2 + m_o^2c^2}$; momentum conservation yields:

$$\mathbf{p}_o = \mathbf{p}_{i_+} + \hbar\mathbf{k}_j = \mathbf{p}_{i_-} - \hbar\mathbf{k}_j'$$

$W_{hole\,th.}$ was first computed by Weisskopf [78]. The result is,[31] again for $\mathbf{p}_o = 0$

$$W_{hole\,th.} = \mu c^2 = \frac{3e^2}{2\pi\hbar c} m_o c^2 \ln \frac{\hbar}{m_o c a}\left(1 + \sqrt{1 + \frac{m_o^2c^2a^2}{\hbar^2}}\right) - \frac{1}{4\pi}\cdot\frac{e^2}{\hbar c}\cdot m_o c^2 \quad (18)$$

In view of the fact that, as pointed out previously, relativistic quantum theory gives to a certain extent reasonable results if one puts m_o equal (or practically equal) to the experimental mass of the electron, it is suggestive

[30] (15) and (17), with appropriate V, are general formulae for the self-energy of any Dirac particle coupled with Bose-Einstein fields [74].

[31] Weisskopf uses an alternative form for V:

$$V = C - \int \mathfrak{s}\mathfrak{A}_\perp dv; \qquad C = \frac{1}{2}\int \frac{\rho(\mathbf{x})\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\,dvdv'$$

In *loc. cit.* [78] the expectation value of C is called electrostatic self-energy, that of $-\int \mathfrak{s}\mathfrak{A}_\perp dv$ the electrodynamic self-energy which is subdivided in the "spin" and "fluctuation" energy; it is of course only the sum of all these parts which enters in the final results. Taking for m_o the nuclear mass, as we would have to do in computing the electromagnetic self-energy of the proton, it is remarkable that the "finite term" in (18) is of the order of magnitude of the proton-neutron mass difference.

that, whereas $\mu_{qu}{}^{NR} \gg m_o$ for $a \sim e^2/mc^2$, $\mu \ll m_o$ for this same a. Indeed one possibility of understanding the still unjustified identification (or approximate identification) of m_o with the experimental mass would be that a consistent theory would give a field contribution to the mass which is $\ll m_o$.

A comparison of (18) and (16) clearly demonstrates the profound difference between hole theory and non-relativistic (or one particle) theory: just as there is no correspondence between non-relativistic quantum theory and classical theory as far as problems like that of the self-energy are concerned, there is, with respect to such problems, again no correspondence between relativistic and non-relativistic quantum theory.

It is noteworthy that the existence of the logarithmic divergence in (18) can be understood solely from the fact that the self-energy for a negaton and a positon with same momentum should be equal; using such general arguments it can be shown that the higher order contributions to the self-energy also diverge at most logarithmically [78, 74]. Such arguments can moreover be extended to the coupling between Dirac particles and other fields of the Bose-Einstein type [74],[32] where again the symmetry of the theory between particle and anti-particle is a guiding principle. It may be remarked here incidentally that symmetry considerations are quite generally of great help in hole theory in simplifying calculations and in order of magnitude estimates. In particular a theorem due to Furry [87] may be mentioned; this states that in the calculation of electromagnetic interaction effects, in which one follows the usual perturbation procedure by starting from the free particle picture as a zeroth approximation and in which negatons and positons appear only virtually, the odd order contributions vanish identically.

A result like (18) recalls the great caution which is needed in judging any approach to the problems of elementary particle physics in which the many-body character of the problem is not taken into account from the outset. But it is clear that considerations based on only the first approximation hardly constitute a trustworthy basis for a complete assessment of what precisely the divergence difficulties are. In fact Racah has suggested [82] that taking into account the higher order approximations of the self-energy might ameliorate rather than worsen the situation and that, on hole theory, the self-energy even might become finite if a certain condition is imposed on $e^2/\hbar c$.[33] It should not be forgotten, however, that the legitimacy of the

[32] Non-linear terms in the Hamiltonian due to such couplings are discussed by Heisenberg [80].

[33] Racah has partly computed the second order self-energy ($\sim e^4$). In the present notation his calculations amount to computing the fourth order contribution due to $\int(\rho\phi - sA_\parallel)dv$. There is of course also a cross term of that operator and $-\int sA_\perp dv$ and a fourth order term of the latter operator. Qualitative discussions of the higher order approximations have also been given by Mercier [83] which, inasfar as his conclusions on the types of divergences occurring in higher order are concerned, are not correct however.

procedure of computing the self-energy by a power series development in $e^2/\hbar c$ still awaits proof or disproof (cf. p. 12).

It is interesting to compare the self-energy problem in hole-theory with the same problem for charged Bose-Einstein particles. Calling their mechanical mass M_o, the electromagnetic self-energy is in first approximation $\sim e^2\hbar/M_o c a^2$, irrespective of whether these particles are of the (pseudo) scalar or (pseudo) vector type [78, 84].[34] As in the case of the electron, the $1/a$ infinity, occurring in classical and non-relativistic quantum theory, again disappears from the scene. As pointed out by Bohr [85], it is noteworthy that the electromagnetic self-energy is again $\ll M_o c^2$ if: (a) the mechanical mass M_o is roughly of the order of 100 times the electron mass, a value which is suggested from evidence on the meson, and (b) one introduces as an orientation a cut-off distance $a \sim e^2/mc^2$ as for the electron.

Also in the Bose-Einstein case the quantum description of the interaction of charged particles and the electromagnetic field presents some basic difficulties. Indeed it is characteristic for charge-bearing Bose-Einstein fields that upon quantization there occurs in the commutation relations for these fields the time component of the electromagnetic four vector potential. In the case of the self-energy this will be the electrostatic proper field potential. If the approximation procedure is followed one can look apart from this electrostatic potential in the commutation relations when one goes to the order e^2 in the self-energy. But this can no longer be done in higher approximations. Now the proper field potential is highly singular, and it follows from the work of Schiff, Snyder, and Weinberg [86], who investigated some simple cases of external electrostatic potentials, that under such circumstances a consistent quantization of the formalism meets with basic difficulties. Thus it is questionable whether in the present stage we are able to formulate the self-energy problem in a consistent way.

But let us get back to the electron. Thus far we have only considered the self-energy for an electron with momentum zero, and the question now arises what the quantum mechanical analogs are to the classical transformation relations (4) for an electron in uniform motion. This question must be put somewhat differently in quantum as compared to classical theory: in the latter case we ask for the energy and momentum of an electron with prescribed (constant) velocity v; in quantum theory on the other hand (at least in the customary representation which we have used here) the momentum is prescribed in zeroth approximation. Thus we have to ask for the energy, the velocity, and the correction to the momentum corresponding to a given zero order momentum p. On the lines of the discussion followed hitherto we will view this question again in the light of the approximation procedure, and thus we have to ask in the first place what the self-energy in the e^2 approximation shall be for an electron with momentum p. Now by

[34] For contributions to the self-energy of such particles due to couplings with other than the electromagnetic fields, see [84].

applying an adiabatic transformation it can be seen[35] that W is given by

$$\text{—————(2)—————}$$

$$\tfrac{1}{2}\int(\rho\phi - \mathfrak{s}\mathfrak{A})dv \tag{19}$$

where the symbol —(2)— means that that part of the expectation value of the operator should be taken which is $\sim e^2$. The integrand in (12) is a scalar[36] which (as we are considering expectation values) does not explicitly depend on time. Thus the transformation properties of W are determined by those of the volume element for which $dv = dv_o\sqrt{1-\beta^2}$. According to the foregoing reasoning this has to be translated into momentum language, while to the approximation considered we may put for v the "unperturbed" velocity $cp/\sqrt{p^2 + m_o{}^2c^2}$. Hence the energy E of the electron with momentum p is

$$E = c\sqrt{p^2 + m_o{}^2c^2} + \frac{\mu m_o c^3}{\sqrt{p^2 + m_o{}^2c^2}} \tag{4a'}$$

It can be shown that an equation of the type (4a') also holds for the coupling of Dirac particles with other than the electromagnetic fields [74] and for the self-energy of Bose-Einstein particles [84].

Equation (4a') looks entirely different from (4a) but yet is intimately connected with it. In fact all corresponding quantities occurring in both formulae have the same structural definition from a covariance point of view. Thus the self-energy is, classically as well as quantum mechanically, defined as the spatial integral over the 44-component of a tensor density, whatever the basic differences of quantum mechanical origin between μ and μ_{rel} are. The only further difference between these two quantities is that the former is—whereas the latter is not—computed by using an approximation procedure. This implies that the perturbation energy should be small compared with the unperturbed energy, that is—an over-optimistic assumption—in the present stage

$$\frac{\mu}{m_o} \ll 1 \tag{20}$$

With this in mind we can from (5) immediately derive (4a'): express v in p by using (5b) and neglect terms of higher than the first order in μ/m_o. Substitute the expression thus obtained for v in (5a) and go again to the first order in μ/m_o. This gives (4a').

[35] Cf. *loc. cit.* [77] p. 30, *loc. cit.* [78] p. 80.

[36] The approximation procedure rests on the development with respect to the relativistically invariant parameter e^2/hc. Hence in discussing covariance properties it is allowed to consider separately the contributions to the self-energy in a given approximation. Moreover the subtractions performed on hole theory do on account of their very invariance not affect the present argument so that one can deal with W and $W_{hole\ th.}$ at the same time.

It should be pointed out that the computation of W as an integral over momentum space involves an ambiguity, as there are different ways in which one can take together the contributions from the various parts of this space. (This statement of course has a meaning only when one confines oneself to a finite part of momentum space, as is for example useful in comparing two different self-energies.) A natural way of procedure in the case of zero electron momentum is to take first the contribution of a spherical shell with center at the origin and then take the sum over all shells. If one then performs the explicit calculation for a moving electron one should, in order to obtain (7a'), take the sum of the contributions of the Lorentz-transforms of these shells. In computations of the self-energy for $p \neq 0$ one often takes a sum over spherical shells [70, 77]. The latter procedure is related to the former in a manner similar to the relation of the classical Abraham to the classical Lorentz electron. Moreover, as has been pointed out in [74], p. 45, caution is also needed in computing the self-energy for a moving electron from a formalism to which one previously has applied a contact transformation to separate off the electrostatic self-energy. As this transformation itself is not invariant the connection between domains in momentum space referring to the electron at rest on the one hand, and in motion on the other, is no longer given by a Lorentz transformation.

The p occurring in (4a') is still the zero order momentum. On account of the coupling of the electron with the electromagnetic field there will however be an electromagnetic contribution not only to the energy, but also to the momentum. The electromagnetic momentum has been discussed by Kemmer [87].[37] His results can not directly be correlated with the expression (4a') for the self-energy. The reason for this is that Kemmer uses the method of integrating over spheres also in the moving system so that the covariance properties do not come into evidence. Finally, there is also an "electromagnetic velocity," defined as the contribution $\sim e^2$ to the expectation value of the velocity operator α. Between these three quantities, electromagnetic energy, momentum, and velocity there exist relations which are determined by the tensor transformation laws (5). We hope to come back elsewhere to this problem, and will merely state here that in relativistic quantum theory the problem of the self-stress will again occur along with the self-energy question. One will have finite self-energy and zero self-stress as requirements for a consistent relativistic quantum theory, quite similar to the situation in classical relativity theory.

Thus far we have mainly discussed the self-energy of free electrons, the effect of binding, say in a Coulomb field, has as yet not been touched upon. This latter problem is of considerable interest especially in view of recent experiments which show that one of the most important consequences of the Dirac theory needs refinement: it was found by Lamb and Retherford

[37] The electromagnetic momentum is the quantity $P^{(1)}$ of Kemmer; cf. *loc. cit.* p. 695 for the one electron theory and p. 703 for the positon theory.

[88] that the $2S\frac{1}{2}$ level of the hydrogen atom which should, according to Dirac theory, be degenerate with the $2P\frac{1}{2}$-level, actually lies higher by an amount of about 1000 mc/sec. ($= 0.033$ cm.$^{-1}$). Moreover it turns out that there is a similar shift for the deuterium atom [89] which within the limits of accuracy of the experiment is equal to that of the hydrogen case. The interpretation of these phenomena would therefore seem to involve features which are intimately linked up with, or are uniquely accompanying, electromagnetic phenomena. While the deviation from Coulomb's law due to the polarization of the vacuum would not seem to give an account of these experimental results.[38] this effect may, apart from other possible non-Coulombic interactions between electron and proton, be linked up with a difference in self-energy of the electron in the $2S\frac{1}{2}$ and $2P\frac{1}{2}$ state.[39] Recent calculations by Bethe [90] and Kramers [91] seem to disclose the possibility of interpreting this shift along these lines. The question is at present not quite settled, since the above mentioned calculations are made with the help of a one particle picture. The detailed investigation of a relativistic quantum treatment is under investigation by various physicists. The underlying idea of the relativistic treatment may briefly be indicated below.[40]

When one computes the self-energy of an electron bound in a hydrogen atom, the result is a logarithmically divergent integral. The singular part of this integral is identical with that occurring in the expression for the self-energy of a packet of free electron waves with the same momentum distribution as in the corresponding bound state. Now one decomposes the self-energy of the bound electron into two parts; the first is the self-energy of the free wave packet. This part is discarded as it can be considered to renormalize the mass of free electrons from its value m_o, occurring in (9), from which we started in a zeroth approximation (mechanical mass) to its true value (experimental mass). It is only this part which is singular. The remaining part is finite and is identified with an observable level shift.

The procedure thus followed can be described as an attempt to formulate a structure independent theory[41] on relativistic quantum lines. It is quite important for this procedure that the self-energy transforms in the way indicated in (4a'). For (4a') can in the approximation e^2 be written $c\sqrt{p^2 + (m + \mu)^2c^2}$, and this would at first sight seem to make possible an

[38] The exchange polarization has as yet not been discussed completely, cf. p. 23.

[39] Discussion remarks by Oppenheimer, Weisskopf, and Schwinger during the Conference on the Foundations of Quantum Mechanics at Shelter Island, sponsored by the National Academy of Sciences.

[40] The results mentioned in the following paragraph refer to independent investigations by Professors Lamb, Schwinger and Weisskopf. Without kind communications of their preliminary results the writing of this paragraph would not have been possible. I am also greatly indebted to Professor J. R. Oppenheimer for clarifying discussions on the subject.

[41] See p. 3 for the notion of structure independence.

unambiguous identification of the self-energy μ with a mass. It should not be forgotten, however, that the concept of electromagnetic mass is ambiguous in a relativistic theory. This we have seen in classical relativity. (Note that (4b) "looks covariant," but that (4a) spoils the covariance.) The corresponding situation in quantum theory can only be understood by studying the behavior not only of the electromagnetic self-energy, but also of the electromagnetic velocity and momentum under Lorentz transformations. In particular the investigation of the self-stress will be of importance for an analysis of the ambiguities in the subtractions discussed in the previous paragraph (cf. p. 6).

However this may be, it is clear that great progress can be expected from the developments outlined above. On the one hand we will be better able to judge whether a more refined definition of observable quantities may help us out of many difficulties. On the other hand it will be a severe test to any proposed modification of the theory of the electron whether it will be in agreement with such phenomena as the line shift in the hydrogen atom. Also new evidence may be obtained for the degree of reliability of the power series developments in $e^2/\hbar c$.

A more refined investigation of the singularity occurring in the radiative correction for scattering of an electron in an electrostatic field of force (cf. the footnote on p. 12 and also p. 35) has shown that this singularity too is identical with that of the electromagnetic mass. Hence the same "renormalization of the mass" mentioned above will remove the infinite part of the radiative correction [109]. It is also clear that the infinities in the line shift calculation and in that of the radiative correction are intimately related to each other. The only difference is that in the one case we have to do with the self-energy correction to a bound state, in the other with a corresponding correction to a continuum state.

Recent measurements of the hyperfine structure of the hydrogen and deuterium atomic ground state [92] seem to reveal that, there too, deviations from the theoretically expected values exist. It seems possible to account for this effect on similar lines, as in the case of the fine structure mentioned previously [92a], that is, one can evaluate the self-energy of a bound electron in an external magnetic field and compare it with the self-energy for the same state in the absence of that field. The difference is then interpreted as an observable level shift.

(e) *Approaches to a consistent relativistic quantum theory*

Finally we will briefly survey how the various attempts to attain a relativistic theory of the electron (discussed in section IIc) have developed in relativistic quantum theory. The foregoing discussion has shown that a consistent classical relativistic theory does not give any guarantees for its quantum counter part. Moreover, in so far as the following considera-

tions deal with a one particle picture, no conclusions can be derived therefrom concerning a quantum theoretical description which is truly relativistic.

Theory of Poincaré. Since a model with non-zero radius is essential here, and no description of the cohesive forces in terms of field equations is given, it is difficult to see how Poincaré's idea can be incorporated in a quantum theoretical description. No attempts in this direction have been made.

λ-*Limiting process; theory of Dirac.* (a) *One particle theory.* As already pointed out by Wentzel [9], the λ-process enables us also in quantum theory to eliminate the $1/a$ electrostatic infinity in (8). The theory was brought into Hamiltonian form, suitable for quantization, by Dirac [11, 12]. In order to eliminate the $1/a^2$ (or $\int_0^\infty kdk$) infinity in (8), Dirac introduces [93] what amounts to photons of negative energy, as a result of which integrals of the type $\int_0^\infty kdk$ are replaced by $\int_{-\infty}^{+\infty} kdk = 0$. Higher order quantum contributions have been considered by Eliezer [94]. While this recipe enables one to eliminate certain divergences at the high frequency end, it brings new trouble in the domain of low frequencies. A naïve treatment of the scattering probability of electrons in an electrostatic field under the emission of quanta of large wave length gives a divergent result, due to a singularity at the infrared end. A more refined procedure shows that this singularity is due only to an inadequate mathematical handling of the problem on the one hand, but brings in a new singularity at the ultraviolet end on the other, thus reducing whatever inconsistencies there are in this problem to the ultraviolet catastrophe, ([26], see also p. 12 and p. 34). This step means a definite advance because an infrared catastrophe would imply a failure of the correspondence principle in a domain where its validity is unquestionable. The negative photon device now eliminates also in this problem the ultraviolet, but brings in again the infrared catastrophe [95].

A similar situation is met in the *theory of Heitler* [96] where, in aiming at a description of radiation damping in quantum theory, an algorithm is developed in which all those virtual transitions ("round-about transmissions") leading to infinite self-energies and to infinite contributions to transition probabilities are discarded. Again, no ultraviolet catastrophe, but instead an infrared catastrophe [97].

(b) *Hole theory.* Here a basic reformulation is needed [93] if one introduces negative energy photons for if these were coupled to negative energy electrons, the electron sea would evaporate. However, as observed by Pauli [98], to whose beautiful review article we refer for a detailed exposition of Dirac's ideas, the combined application of negative energy photons and λ-process does not enable us to eliminate divergences of the type (18); cf. *loc. cit.*, p. 204. The same is true in the equivalent procedure of combining negative energy photons with Riesz's method [99].

In a later paper Dirac has suggested a program [100] in which also the λ-process and negative energy photons are employed, but in which power series developments in e^2/hc are no longer necessary. This interpretation is in a too early stage to judge to what extent it may meet the difficulties outlined above.

Action at a distance. While a detailed investigation of the self-energy problem on quantum theory has not been carried out so far, an interesting result has been obtained by Plass [101], who treated the black body radiation from this point of view. It turns out that one can find a transcription of the degrees of freedom of the electromagnetic field, as treated by Rayleigh and Jeans, in the "corporate" degrees of freedom. These degrees of freedom are due to the existence of couplings, requiring a finite time for their propagation, between the particles in the wall of the black body cavity. They appear superposed on the three intrinsic degrees of freedom attributed by classical mechanics to each particle. The corporate degrees of freedom (of which there are infinitely many) occur only if the distance between the particles is less than a critical amount which increases indefinitely with the strength and number of interactions between the particles; for two electrons the critical distance is the classical electron radius.

Two-field theories. These have been investigated in [74]. The general idea is first to quantize the theory and then to determine the singular parts in the electromagnetic self-energy on the one hand and in the self-energy due to the short range field ("f-field") on the other. One then examines whether compensation is possible such that the total self-energy becomes finite. As f-fields the various types proposed by Kemmer [102] were investigated. On a one particle theory no compensation is feasible. On hole theory, relativistic compensation can, at least in first approximation, be achieved by assuming the f-field to be of the scalar or pseudovector type, but not if it is of the vector or pseudoscalar type. (In the latter cases one can eliminate the singularities by taking subtractive fields (cf. p. 10) but the non-positive definiteness is, especially on hole theory, a serious difficulty.) In view of (20) it is interesting to note that the self-energy is $\ll m_0c^2$ if the range of the field is $\sim 10^{-13}$ cm. A general feature of theories of this type is that the quanta of the f-field have a very short life ($\sim 10^{-21}$ sec.) for decay into a negaton-positon pair. The theory has no direct correspondence to the classical two-field theories (cf. p. 10) as a result of the reversal of the order in which the elimination of the divergences on the one hand and the quantisation on the other hand is performed.

Also investigated were the consequences of assuming the inseparability of the electromagnetic and the f-field, so that for example the proton also is the source of such a field. On these lines a reasonable theoretical value for the proton-neutron mass difference could be obtained. The value of the f-field range which is needed for this purpose (a few times e^2/mc^2 would, due to a corresponding deviation from Coulomb's law, yield a contribution to the

line shift in the hydrogen spectrum of the order observed by Lamb and Retherford [88, 89]. However, such a value for the range would seem to be incompatible with our present notions of symmetry properties of atomic nuclei [103]. But if one restricts the application of the f-field to the electron only the formalism is consistent.

Non-linear theories. In so far as it is required that the theory be unitary in the restricted sense (see p. 10), the non-linear theories have failed. Considering the electron to be a point particle (in a suitable representation) the coordinates of which are taken as quantum variables, Pryce [104] has shown that the commutation relations between position and momentum variables of the electron are incompatible with those of the field variables if one defines the electron momentum as the spatial integral of the field momentum density. Thus on the present scheme momentum and energy cannot be entirely of electromagnetic origin.[42] Moreover, it is not possible as was sometimes believed to deduce the spin of the electron from its electromagnetic properties, as one might be tempted to try on a unitary program. Finally, it is impossible to derive electromagnetically the existence of states of matter having negative energy [104], so that the theory would not involve a mechanism for the production of pairs similar to that of the Dirac theory. Indeed the negative energy states are of course intimately related to the occurrence of a term βm in the wave equation for the electron, and the situation is, so to say, that on unitary theory one can construct the factor m, but not the factor β. Thus in equation (9) neither α nor β can be reduced to electromagnetic entities, which emphasizes once more the deep dualism of the present theory of the electron.

Another possibility is to abandon the idea of unitarity but stick to that of non-linearity, thus retaining the βm_0 term in the Dirac equation but making the electromagnetic field term in the Lagrangian non-linear [105].

V. Concluding Remarks

Along with the development and progress of our description of physical phenomena, the fundamental problems concerning the electron as an elementary particle thus appear to require new points of view for their solution. It may be hoped that further experimental data, especially on high energy phenomena, will further deepen our insight into the questions pertaining to the electron as well as other elementary particles and especially to the question what one really has to mean by the concept of elementarity itself. The lack of our understanding of the relative properties of the growing number of fundamental particles is indeed in a sense as unsatisfactory as chemistry would be without a periodic table. Moreover, the recognition of fundamental connections between the particles might well clarify apparent inconsistencies of our present picture. As observed by Bohr [85], a sug-

[42] Quite apart from the divergence difficulties, similar inconsistencies would also arise in attempting to attain a unitary theory within the framework of a linear formalism.

gestive example of such a situation is found in the neutralization of the electron sea by the proton sea[43] and in the possibility of compensating the negative energy densities in space due to seas of Fermi-Dirac particles by the zero point field energies of Bose-Einstein assemblies.

In discussing the synthesis of relativity and quantum ideas we have confined ourselves to special relativity. General relativity has thus far found its main application in phenomena concerning matter in bulk. But we will sooner or later be confronted with the problem of the role of gravitational effects in elementary particle phenomena. Considering the limitations inherent in the present special relativistic quantum theory, it would seem, from well-known order of magnitude considerations, that we have not yet been able to probe deep enough into the phenomena in which small distances are involved to be able to judge this question at its full value.

In conclusion, we have to see in the very occurrence of the infinities a starting point for new developments. The future will tell whether further progress will reveal the necessity of a further renunciation of the unambiguous use of some of our physical concepts in the region of high energies and correspondingly small distances [108].

December 1947.

References

1. H. A. Lorentz, *The Theory of Electrons*, Teubner, 1916. Contributions by Lorentz of later date are all found in *The Collected Papers of Lorentz*, M. Nyhoff, the Hague, 1937.
2. J. J. Thomson, *Phil. Mag.*, (5), *11*, 229, 1881.
3. H. A. Kramers, Quantentheorie des Elektrons und der Strahlung, §§ 89, 90, *Hand und Jahrbuch der Chemischen Physik*, Leipzig, 1938.
   H. A. Kramers, *Il. Nuov. Cim.*, 1938, p. 1.
   W. Opechowski, *Physica*, *8*, 226, 1941.
   J. Serpe, *Physica*, *7*, 133, 1940; *8*, 161, 1941.
4. (a) P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *114*, 243, 1927.
   (b) P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *114*, 710, 1927.
5. Cf. M. Abraham, *Theorie der Elektrizität II*, §19, Teubner, 1908.
6. H. Poincaré, *C. R. Paris*, *140*, 1504, 1905; *Rend. Palermo*, *21*, 1906; *La Mécanique Nouvelle*, Gauthier-Villars, 1924.
7. M. Laue, *Ann. der Phys.*, (4) *35*, 524, 1911.
8. H. A. Lorentz, *Versl. Kon. Ac. Amsterdam*, *26*, 981, 1917; *Coll. Works*, Vol. V, p. 314.
9. G. Wentzel, *Z. f. Phys.*, *86*, 479, 635, 1933; *87*, 726, 1934 (the latter especially concerning classical theory). Cf. also G. Wentzel, *Quantentheorie der Wellenfelder*, Edwards Bros., 1946.
10. S. T. Ma, *Phys. Rev.*, *71*, 787, 1947, giving also references to the papers dealing with the Riesz method.
11. P. A. M. Dirac, *Proc. Roy. Soc.*, *167*, 148, 1938; *Ann. de l'Inst. H. Poincaré*, *9*, 13, 1939.
12. P. A. M. Dirac, *Quantum Electrodynamics*, Comm. Dublin Inst. for Adv. Studies, *A1*, 1943; mimeographed lectures, Princeton 1946.

[43] This of course implies the description of the proton by a Dirac equation. Actually little is known about the detailed relativistic properties of the proton and the only purpose which the proton sea has been used for so far is to attain very short life times of certain particles of the Bose-Einstein type [106]. For the possible existence of the anti-proton in cosmic radiation, see [107].

13. Cf. also H. J. Bhabha, *Phys. Rev.*, *70*, 759, 1946.
14. G. J. Eliezer, *Proc. Cambr. Phil. Soc.*, *39*, 173, 1943.
15. J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.*, *17*, 157, 1945.
16. M. Schoenberg, *Phys. Rev.*, *69*, 211, 1945; *Summa Brazil Math.*, *1*, 41, 77, 1946.
17. E. C. G. Stueckelberg, *Nature*, *144*, 118, 1939; *Helv. Phys. Acta*, *14*, 51, 1941.
18. Proposed independently by: F. Bopp, *Ann. der Phys.*, *38*, 345, 1940; *42*, 575, 1943; and A. Landé and L. Thomas, *Phys. Rev.*, *60*, 121, 514, 1940; *65*, 175, 1944. Bopp actually starts from an electromagnetic Lagrangian involving derivatives of the field strengths; cf. also B. Podolsky, *Phys. Rev.*, *62*, 68, 1942. For the problems on hand this is equivalent to introducing a new field.
19. A thorough survey is given by M. Born, *Ann. de l'Inst. H. Poincaré*, *7*, 155, 1937.
20. a. M. Born, *Proc. Roy. Soc.*, *A*, *143*, 410, 1934.
    b. M. Born and L. Infeld, *Proc. Roy. Soc.*, *A*, *144*, 425, 1934.
    c. E. Schroedinger, *Proc. Roy. Soc.*, *A*, *150*, 465, 1935.
    d. L. Infeld, *Proc. Cambr. Phil. Soc.*, *32*, 127, 1936; *33*, 70, 1937.
21. M. H. L. Pryce, *Proc. Cambr. Phil. Soc.*, *31*, 625, 1935. For the two-dimensional case the electrostatic problem was completely solved by Pryce, *Proc. Cambr. Phil. Soc.*, *31*, 50, 1935.
22. M. Born, *Mem. Indian Ac. of Sc.*, *3*, 85, 1935.
23. For a critical survey of Mie's ideas see W. Pauli, Relativitätstheorie, *Enz. der Math. Wiss.*, *5*, 539, 1921, §64.
24. L. Rosenfeld and J. Solomon, *Jour. de Phys.*, *7*, 139, 1931.
25. W. Pauli, *Hdb. der Physik*, XXIV/1, p. 255.
26. W. Pauli and M. Fierz, *Il. Nuov. Cim.*, 1937, p. 167.
27. H. A. Kramers, private communication.
28. N. Bohr and L. Rosenfeld, *Kgl. Dansk Vid. Selskab*, *12*, No. 8, 1932.
29. Cf. L. Landau and R. Peierls, *Z. f. Phys.*, *69*, 56, 1931; cf. also *loc. cit.* [25], p. 246.
30. W. Pauli, *loc. cit.* [25], p. 216.
31. W. Pauli and V. Weisskopf, *Helv. Phys. Acta*, *7*, 719, 1934; W. Pauli, *Ann. de l'Inst. H. Poincaré*, *6*, 137, 1936.
32. P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *117*, 610; *118*, 351, 1928.
33. For a survey of the properties of the Dirac matrices cf. W. Pauli, *loc. cit.* [25], pp. 219ff.; *Ann. Inst. H. Poincaré*, *6*, 109, 1936.
34. For the theoretical discussion of the Klein-Nishina formula as well as deceleration radiation, pair formation and annihilation, etc., cf. W. Heitler, *The Quantum Theory of Radiation*, 2nd ed., Oxford University Press, 1944, where the references to the original publications are also given.
35. P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *126*, 360, 1929.
36. Cf. *loc. cit.* [34], p. 191.
37. J. R. Oppenheimer, *Phys. Rev.*, *35*, 562, 1930.
38. I. Tamm, *Z. f. Phys.*, *62*, 545, 1930.
39. H. Weyl, *Gruppentheorie und Quantenmechanik*, 2nd ed., p. 234, Hirzel Verlag, 1931.
40. P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *133*, 60, 1931.
41. C. D. Anderson, *Science*, *76*, 238, 1932; *Phys. Rev.*, *43*, 491, 1933.
42. P. M. S. Blackett and G. P. S. Occhialini, *Proc. Roy. Soc.*, *A*, *139*, 699, 1933.
43. V. Fock, *C. R. Leningrad*, *1*, 267, 1933.
44. E. Majorana, *Il. Nuov. Cim.*, *14*, 171, 1937.
45. G. Racah, *Il. Nuov. Cim.*, *14*, 322, 1937.
46. H. A. Kramers, *Proc. Ak. Wetensch. Amsterdam*, *40*, 813, 1937.
47. W. Pauli, *Rev. Mod. Phys.*, *13*, 203, 1941.
48. W. Heisenberg, *Ann. der Phys.*, (5), *10*, 888, 1931.
49. R. Peierls, *Proc. Roy. Soc.*, *146*, 420, 1934.
50. W. Furry and J. R. Oppenheimer, *Phys. Rev.*, *45*, 245, 343, 1934.
51. P. A. M. Dirac, Rapports du 7-ième Conseil Solvay 1933, *Proc. Cambr. Phil. Soc.*, *30*, 150, 1934.
52. W. Heisenberg, *Z. f. Phys.*, *90*, 209, 1934; *92*, 692, 1934; some corrections are given in *loc. cit.* [66].
53. P. Jordan and E. Wigner, *Z. f. Phys.*, *47*, 631, 1928.
54. V. Weisskopf, *Kgl. Danske Vidensk. Selskab*, *14*, Nr. 6, 1936.

55. R. Serber, *Phys. Rev.*, *49*, 545, 1936.
56. O. Klein, *Z. f. Phys.*, *53*, 157, 1929. Cf. also W. Pauli, *loc. cit.* [25] p. 242; F. Sauter, *Z. f. Phys.*, *69*, 742, 1931.
57. R. Serber, *Phys. Rev.*, *48*, 49, 1935.
58. W. Pauli and M. E. Rose, *Phys. Rev.*, *49*, 462, 1936.
59. E. Uehling, *Phys. Rev.*, *48*, 55, 1935.
60. H. J. Bhabha, *Proc. Roy. Soc.*, *A*, *154*, 195, 1936.
61. Ho Zah Wei, *C. R. Paris*, *222*, 1168, 1946; *Phys. Rev.*, *70*, 224, 1946.
62. G. Breit and J. A. Wheeler, *Phys. Rev.*, *46*, 1087, 1934.
63. O. Halpern, *Phys. Rev.*, *44*, 855, 1934.
64. H. Euler and B. Kockel, *Naturw.*, *12*, 246, 1935.
65. H. Euler, *Ann. der Phys.*, *26*, 398, 1936.
65a. A. Achieser, *Phys. Zs. der Soviet Union*, *11*, 263, 1937.
66. W. Heisenberg and H. Euler, *Z. f. Phys.*, *98*, 714, 1936.
67. See *loc. cit.* [54], section II.
68. M. Delbrück, *Z. f. Phys.*, *84*, 144, 1933.
69. N. Kemmer and V. Weisskopf, *Nature*, *137*, 659, 1936.
70. N. Kemmer, *Helv. Phys. Acta*, *10*, 112, 1937.
71. N. Kemmer, and G. Ludwig, *Helv. Phys. Acta*, *10*, 182, 1937.
72. W. Heisenberg, *Ber. Sächs. Akad. Wiss.*, *86*, 317, 1934; J. R. Oppenheimer, *Phys. Rev.*, *45*, 903, 1934.
73. E. Fermi, *Rend. Lincei*, *9*, 881, 1939; L. Rosenfeld, *Z. f. Phys.*, *76*, 729, 1932; L. Broer and A. Pais, *Proc. Ac. Amsterdam*, *48*, 190, 1945.
74. A. Pais, On the Theory of Elementary Particles, *Verh. Kon. Ac. Amsterdam*, Vol. 19, 1947.
75. J. R. Oppenheimer, *Phys. Rev.*, *35*, 461, 1930.
76. I. Waller, *Z. f. Phys.*, *62*, 673, 1930.
77. V. Weisskopf, *Z. f. Phys.*, *89*, 27, 1934; *90*, 817, 1934.
78. V. Weisskopf, *Phys. Rev.*, *56*, 72, 1939.
79. Cf. W. Pauli, *loc. cit.* [25], p. 271.
80. W. Heisenberg, *Z. f. Phys.*, *113*, 61, 1939; see also: *Cosmic Radiation*, §12, ed. by W. Heisenberg, Dover, New York 1946.
81. W. Furry, *Phys. Rev.*, *51*, 125, 1937.
82. G. Racah, *Phys. Rev.*, *70*, 406, 1946.
83. A. Mercier, *Helv. Phys. Acta*, *12*, 55, 1938; cf. also *loc. cit.* [78], p. 83.
84. A. Pais, *Physica*, *12*, 81, 1946.
85. N. Bohr, *Proc. Phys. Soc. London*, in press.
86. H. Snyder and J. Weinberg, *Phys. Rev.*, *57*, 307, 1940; L. I. Schiff, H. Snyder, and J. Weinberg, *Phys. Rev.*, *57*, 315, 1940.
87. N. Kemmer, *Ann. der Phys.*, *22*, 674, 1935.
88. W. Lamb and R. Retherford, *Phys. Rev.*, *72*, 241, 1947.
89. W. E. Lamb, private communication.
90. H. A. Bethe, *Phys. Rev.*, *72*, 339, 1947.
91. H. A. Kramers, private communication.
92. J. E. Nafe, E. B. Nelson, and I. I. Rabi, *Phys. Rev.*, *71*, 917, 1947.
92a. J. Schwinger, unpublished results.
93. P. A. M. Dirac, *Proc. Roy. Soc.*, *A*, *180*, 1, 1942.
94. J. Eliezer, *Proc. Roy. Soc.*, *A*, *187*, 197, 1946.
95. W. Pauli, *Helv. Phys. Acta*, *19*, 234, 1946.
96. Cf. *loc. cit.* [34], pp. 240–252.
97. H. A. Bethe and J. R. Oppenheimer, *Phys. Rev.*, *70*, 451, 1946.
98. W. Pauli, *Rev. Mod. Phys.*, *15*, 175, 1943.
99. T. Gustafson, *Kgl. Fys. Sallsk. Lund Forhandl.*, *16*, No. 2, 1946.
100. P. A. M. Dirac, *Comm. Dublin Inst. for Adv. Studies*, *A*, 3, 1946.
101. G. N. Plass, Black body radiation in the theory of action at a distance, Thesis, Princeton, 1946.
102. N. Kemmer, *Proc. Roy. Soc.*, *A*, *166*, 127, 1938.
103. A. S. Wightman, *Phys. Rev.*, *71*, 447, 1947.
P. R. Zilsel, J. S. Blair, and J. L. Powell, *Phys. Rev.*, *72*, 225, 1947.

104. M. H. L. Pryce, *Proc. Roy. Soc., A, 159*, 355, 1937.
105. Cf. M. Born and H. W. Peng, *Proc. Roy. Soc. Edinburgh, 52*, 40, 92, 127, 1944.
106. S. Sakata and Y. Tanikawa, *Phys. Rev., 57*, 548, 1940.
G. Wentzel, *Rev. Mod. Phys., 19*, 5, 1947.
R. Finkelstein, *Phys. Rev., 72*, 415, 1947.
107. O. Klein, *Arkiv. f. Mat. Ast. och Fys., 31*, A, No. 14, 1944; N. Arley, *Kgl. Danske Vidensk. Selskab, 23*, No. 7, 1945; *Physica, 12*, 177, 1946; *Phys. Rev., 70*, 975, 1946.
108. Cf. W. Heisenberg, *Ann. der Phys., 32*, 20, 1938.
H. S. Snyder, *Phys. Rev., 71*, 38, 1947; *72*, 68, 1947.
109. H. Lewis, *Phys. Rev., 73*, in press; S. T. Epstein, *ibid.*

FIELD AND CHARGE MEASUREMENTS IN QUANTUM THEORY

BY N. BOHR AND L. ROSENFELD

The manuscript of this paper, which will shortly appear in the Mat.-Fys. Med. of the *Kgl. Danske Videnskabernes Selskab*, was written already in 1939, but its publication has been delayed due to the events of later years. We are indebted to Professor Bohr and Professor Rosenfeld for the opportunity of reading the manuscript and making an abstract of it which is here included with their consent.—John A. Wheeler.

THIS paper analyzes the question how far measurements of electromagnetic field quantities and electric charges are limited by typical quantum mechanical fluctuation phenomena. The first half considers the consequences of the field fluctuations of quantum electrodynamics, and the second deals with the charge fluctuations of pair theory. The purpose of the analysis is to clarify the logical aspects of the definition and use of concepts compatible with the mathematical formalism, and thus to bring out some aspects of this formalism which might otherwise have escaped attention, and which are fundamental features of its logical consistency.

The paper begins with a review of the earlier communication of Bohr and Rosenfeld (see p. 13) on the question of measurability of electromagnetic fields. It is emphasized again that suitably constructed test bodies—to the extent that one can look apart in the measurements from their atomic constitution—permit one in idealized experiments to determine field quantities with just the precision claimed by quantum electrodynamic theory. In particular, the average value of a field quantity over a finite space-time interval can be found with arbitrary accuracy. The same is true of determinations of values of two different field quantities over the same space-time interval. However, reciprocal uncertainty relations in general exist between the values of field quantities averaged over two different space-time intervals. About this complementary character of the two quantities in question the analysis of the idealized experiments yields the same conclusions as does the formalism of electrodynamics.

The earlier considerations of simple space-time averages of field quantities are extended. In particular the measurability of the electric charge is considered. For this purpose one considers the case where (1) the test bodies employed envelop a certain closed volume, (2) the thickness b of the test bodies tends to an infinitesimally small value, and (3) the partial test bodies are connected by way of suitable levers with a single external device for the measurement of force or momentum. The connections are made in

such a way that one obtains the value of the normal component of the electric field integrated over the surface in question. It is shown that this integral can in principle be determined with an arbitrary accuracy so far as concerns any possible limitations imposed upon the measurement by the quantum theory of fields. It is therefore concluded that the determination via the theorem of Gauss of the electric charge enclosed within the surface offers no difficulty so long as one disregards the consequences of the phenomenon of pair creation for the measurement.

That the measurement of the charge contained within a region of space cannot be made with arbitrary accuracy is, however, the conclusion reached by Heisenberg (*Sächs. Akad.*, *86*, 317 (1934)) through an analysis in the light of pair theory of the fluctuations to be expected in such a region. From his paper it furthermore follows that the fluctuation of charge depends not only upon the magnitude of the volume itself but also upon the sharpness b with which the boundaries of this volume are defined. In the case where the product of the velocity of light and the interval T during which the mean value of the charge is taken is smaller than both the quantity b and the distance $\hbar/mc$, the mean square deviation $\overline{(\Delta e)^2}$ of the charge from its expectation value is given by the expression:

$$\overline{(\Delta e)^2} = \frac{(\text{electronic charge})^2 \times (\text{surface of enclosure})}{\begin{pmatrix}\text{thickness } b \text{ of test bodies}\\ \text{which define enclosure}\end{pmatrix} \times \begin{pmatrix}\text{velocity of light times}\\ \text{time of measurement}\end{pmatrix}} \qquad (1)$$

The remainder of the paper of Bohr and Rosenfeld is devoted to a discussion of the physical significance of the charge fluctuations implied by this formula and a derivation of the formula directly from an analysis of the measuring process itself. It is thus shown that the possibilities of measurement contained within the framework of the theory of fields and pairs in principle permit a determination of the charge contained within an enclosure with just the accuracy specified by the equation in question, again subject to the condition that one may look apart from the atomic constitution of the measuring devices.

On the question of the origin of the charge fluctuations, attention is drawn to the earlier view expressed by Oppenheimer (*Phys. Rev.*, *47*, 144 (1935)), that these charge fluctuations are not only inseparable from but in fact due to the zero-point fluctuations in the electromagnetic field itself. The contrary conclusion is reached by Bohr and Rosenfeld. They show that the charge fluctuations arise entirely from the pair field. The absence of any contribution to the charge fluctuation from the zero-point fluctuations of the electromagnetic field is indeed already evident from their theorem about the exact measurability—within the framework of pure quantum electrodynamics—of the surface integral of the normal component of the electric field.

The uncertainty in the value of the Gaussian integral—and therefore

in the value of the included charge—is found to be due to the effect of pairs unavoidably created by the measuring device itself. This phenomenon shows up when one considers in detail a device composed as mentioned above of a large number of small charged test bodies arranged to surround completely the volume under study. Specifically, Bohr and Rosenfeld point out:

a. That the typical charged test body receives a normal displacement in the course of the measurement;
b. That the displacement of charge creates a supplementary electric field within the region occupied by the test body itself;
c. That this electric field will in general disturb the infinite distribution of negative energy electrons and bring into existence pairs of positive and negative electrons;
d. That the charge of these pairs will be so distributed as to produce an additional electric polarization;
e. That this electric polarization will produce a supplementary force on the test body which will disturb the result of the measurement;
f. That this disturbance is, however, proportional in a certain approximation to the magnitude of the displacement of the test object;
g. That the perturbation can therefore be compensated by suitable elastic devices;
h. That the perturbation consequently will not in itself impair the possibility of measuring with arbitrary accuracy the charge contained within the given volume element;
i. That the supplementary force created in this way, for example by N pairs of positive and negative electrons, will be subject however to unpredictable statistical fluctuations about its normal average value proportional to $N^{1/2}$.
j. That there will be on this account an unavoidable fluctuation in the value of the surface integral of the normal component of the electric force; and
k. That this fluctuation, when calculated, is found to correspond to an uncertainty in the charge contained within the given volume just equal to the figure of Heisenberg quoted above.

Although there is no direct connection between these charge fluctuations and the field fluctuations discussed in the earlier half of their paper, Bohr and Rosenfeld nevertheless point out the close analogy between the two kinds of fluctuations. In both cases the displacement of the measuring device produces a back reaction on this test object itself, in the one case by way of the radiation field which it creates, in the other case by way of the positive and negative electrons which it generates. In both cases, the average value of the reaction so developed is directly proportional to the magnitude of the displacement and can therefore either be compensated

by a suitable elastic device or otherwise be taken into account in a predictable way. However, there are necessarily quantum fluctuations about this average value because the reaction is transmitted by a finite number of photons in the one case or a finite number of pairs in the other. It is these fluctuations which set a limit to the accuracy which can be obtained in certain field and charge measurements.

From their analysis of the measuring process Bohr and Rosenfeld conclude that existing theory—when used within its proper domain of application and thus in particular employed in a way which does not depend upon the atomic character of instruments of measurement—gives a logically self-consistent account of the possibilities of determining charges and electromagnetic field quantities.