Statistical Geometry and Fundamental Particles

NATHAN ROSEN

University of North Carolina, Chapel Hill, North Carolina

(Received May 10, 1947)

To take account of the fact that the existence of a fundamental length a, the classical electron radius, sets a limit to the accuracy with which the position of a point can be measured, it is proposed to introduce two spaces, an "abstract" space consisting of points, and an "observable" space in which one deals with elementary volumes correlated to the points of the former by means of a statistical distribution function in the form of a three-dimensional Gaussian error function. Such a function is not Lorentz invariant, but one can obtain Lorentz covariance in the observable space by carrying out the usual Lorentz transformation in the abstract space. If one assumes that the usual equations for wave fields, in which the fundamental particles are regarded as points, are valid in the abstract space, then one can obtain corresponding equations in the observable space, with the particles behaving as if they had finite volumes. The difficulties associated with infinite self-energies and singularities in the interactions between particles, as calculated by the usual perturbation method, disappear, but the difficulty associated with the divergence of the series of successive orders of perturbations remains.

1. GENERAL CONCEPTS

THE constant a, the "classical electron radius" of the order of 10^{-13} cm, is considered to play an important role in limiting the range of validity of the present form of the quantum theory.[1] It has been suggested by various authors[2] that this constant represents the smallest measurable distance, and that it is therefore necessary to alter the ordinary concepts of geometry when one is dealing with regions having dimensions of the order of a. Accordingly, March[2] introduced a "granular" geometry in which the smallest distinguishable element of space is a sphere of radius a. Recently there appeared a very interesting paper by Snyder,[3] in which the coordinates and time are treated as non-commuting operators, the coordinates having eigenvalues which are integral multiples of a.

In the present paper the point of view adopted is somewhat related to that of March. It is assumed that the constant a determines the lower limit to the error in the measurement of the position of a point. If one measures the x-coordinate of a point under the most favorable conditions, because there does not exist an infinitesimally small measuring rod in nature, one will not obtain an exact value, in general. On the basis of the theory of random errors of measurement it is reasonable to expect that repeated measurements will give values dis-

tributed about the mean value in a normal or Gaussian distribution. It will be assumed, then, that the probability of getting a value lying between ξ and $\xi+d\xi$ is given by

$$(2\pi)^{-\frac{1}{2}}a^{-1}\exp[-(\xi-x)^2/2a^2]d\xi,$$

where x is the mean value. Extending this idea to the other coordinates, we will assume that, if one measures the position of a point, for example, that of an electron at a certain instant of time, the probability of obtaining values in the ranges $d\xi$, $d\eta$, $d\zeta$ near ξ, η, ζ is given by

$$dW = \chi(\xi-x, \eta-y, \zeta-z)d\xi d\eta d\zeta, \qquad (1)$$

where

$$\chi(u, v, w) = (2\pi a^2)^{-\frac{3}{2}} \times \exp[-(u^2+v^2+w^2)/2a^2], \qquad (2)$$

so that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi(u, v, w)dudvdw = 1. \qquad (3)$$

In this way we arrive at the idea of a geometry dealing with small regions, "elementary volumes," instead of points, provided we think of the elementary volume associated with mean values of the coordinates x, y, z, as made up of those points (values of ξ, η, ζ) for which the function $\chi(\xi-x, \eta-y, \zeta-z)$ has an appreciable value. Such an elementary volume differs from the kind considered by March in that it does not have a definite boundary.

The basic physical principle associated with such a geometry is that only a physical quantity

[1] W. Heisenberg, Zeits. f. Physik **120**, 513 (1943).
[2] A. March, Naturwiss. **26**, 649 (1938), where references are given to papers by himself and other authors.
[3] H. S. Snyder, Phys. Rev. **71**, 38 (1947).

associated with an elementary volume can be observed, while the value at a point is not directly observable. If one wishes to have physical laws in terms of observable quantities, it is therefore necessary to express them in terms of physical quantities which refer to elementary volumes.

When one attempts to formulate laws in this way, taking account of the basic indeterminacy of position, it is convenient to work with two spaces, an "abstract" or point space and an "observable" space, made up of elementary volumes, as described above. To every point of the abstract space one can correlate an elementary volume of the observable space by taking the mean values of the coordinates of the latter equal to the corresponding coordinates of the former. In the observable space, no meaning will be assigned to the value of a physical quantity, such as a field variable, at a point, but only to the mean value over an elementary volume. The purpose of introducing the abstract space is to help one to obtain such mean values.

In attempting to formulate physical laws in the framework of this statistical geometry it is natural at the outset to try to keep them as nearly as possible in the same form as at present. Thus, in the case where one is dealing with physical variables satisfying linear equations, one might try to retain these equations, but to take them as holding in the *abstract space*. The corresponding equations in the observable space would then be obtained by an averaging process from the latter, so that, for example, a function $f(x, y, z, t)$ would go over into[4]

$$\bar{f}(x, y, z, t) \equiv \langle f(x, y, z, t) \rangle$$

according to the relation

$$\bar{f}(x, y, z, t) = \int \int \int f(\xi, \eta, \zeta, t)$$

$$\times \chi(\xi - x, \eta - y, \zeta - z) d\xi d\eta d\zeta. \quad (4)$$

Such a transformation has some interesting properties. In the case of a constant, C, it follows obviously that

$$\bar{C} = C. \quad (5)$$

One also sees readily that for a coordinate x and

---

[4] Owing to typographical difficulties in setting bars over multiple symbols or boldface symbols, the notation with angular brackets will be used where required. *Ed.*

the time t

$$\bar{x} = x, \quad \bar{t} = t. \quad (6)$$

In the case of any well-behaved function $f(x, y, z, t)$ it follows that

$$\langle Cf \rangle = C\bar{f}, \quad \langle f_1 + f_2 \rangle = \bar{f}_1 + \bar{f}_2 \quad (7)$$

$$\langle \partial f / \partial x \rangle = \partial \bar{f} / \partial x, \quad \langle \partial f / \partial t \rangle = \partial \bar{f} / \partial t, \quad (8)$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{f} dx dy dz = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f dx dy dz. \quad (9)$$

On the other hand one finds that

$$\langle xf \rangle = x\bar{f} + a^2 (\partial \bar{f} / \partial x). \quad (10)$$

It follows that for a function $f(x, y, z, t)$ which satisfies a linear homogeneous partial differential equation with constant coefficients, the equation for $\bar{f}(x, y, z, t)$ will have the same form as that for f. It should be noted however that, while we have obtained the form of the equation in the observable space by an averaging process, this does not mean necessarily that each solution $\bar{f}$ must be obtained by first determining a suitable solution f and then getting its mean value over an elementary volume. Once the appropriate equation in the observable space has been found, it is possible to get its solutions directly. Each such solution, however, is subject to the restriction that it must be possible to express it as the mean value of some function over the elementary volume.

2. LORENTZ TRANSFORMATION

Before considering applications of the preceding ideas, it is necessary to examine them from the standpoint of the special theory of relativity. If we have two frames of reference moving with a uniform relative velocity, it is to be expected that the same uncertainty in the measurement of position will exist in each of them. In each system one will have "elementary volumes," and an observer in each one will be able to introduce an "abstract" and an "observable" space. Any physical law must be expressible in the same form in each frame of reference.

Now, the special relativity theory was developed without taking account of the limitation on measurement imposed by the existence of the

fundamental constant a. Hence one should be prepared to find that the introduction of this constant into the theory leads to changes in the usual transformation relations between quantities in the two frames of reference.

In order to be sure of obtaining transformation equations that will lead to physical laws of the same form in the two systems, one can make the following assumption: The usual Lorentz transformation will be assumed to hold in the transformation from the abstract space of one reference frame to that of the other frame moving relative to it with a uniform velocity. The transformation of a physical quantity in the observable space can then be determined from the corresponding transformation in the abstract space, if one knows how to go from the abstract to the observable space.

It may turn out in some cases that the transformation in the observable space will differ from the Lorentz transformation. However, as far as the transformation of coordinates and time is concerned, we see from Eq. (6) that the Lorentz transformation will remain valid. From a consideration of a plane monochromatic wave one can readily see that the frequency and wavelength of a light wave transform in the usual way.

At this point it is appropriate to consider briefly another question suggested by relativistic considerations. In the preceding discussion the time coordinate has been treated quite differently from the space coordinates. In particular, we have considered the uncertainty in the measurement of a space coordinate at a particular moment of time. It might be asked whether one should not introduce an uncertainty in the time (in this case of the order of a/c) in the same way as has been done for the coordinates. This would mean having on the right-hand side of Eq. (1) an additional factor

$$(2\pi)^{-\frac{1}{2}} c a^{-1} \exp[-c^2(\tau-t)^2/2a^2] d\tau.$$

However, the objection to such a procedure appears to be that it would be difficult to interpret the formalism from the operational standpoint. In the three-dimensional treatment considered previously the function χ of Eq. (1) represents the probability distribution of measured values of the coordinates of a point, such as the position of an electron, at a given moment

of time, so that the time serves as a parameter to identify what is to be measured. In the four-dimensional treatment, where the time would also be measured and would have an uncertainty, no such parameter would in general be available, so that it would not always be clear what one was measuring or how the measurement could be repeated. The procedure adopted here, of using the time as a parameter is, after all, in agreement with what is generally done in the quantum theory.

3. ELECTRON AND RADIATION FIELD

We next consider the problem of the electron and its interaction with the electromagnetic field. In order to avoid lengthy derivations, reference will be made to the book of Heitler[5] for a discussion of the classical and quantum theories of the electromagnetic field.

The electromagnetic field for a given distribution of charges with a density ρ and convective velocity $\mathbf{v}$ can be described classically in terms of the potentials ϕ, $\mathbf{A}$ by means of the following equations:[6]

$$\Box^2\phi = -4\pi\rho, \tag{11}$$

$$\Box^2\mathbf{A} = -(4\pi/c)\rho\mathbf{v}, \tag{12}$$

$$\nabla\cdot\mathbf{A}+\dot{\phi}/c=0, \tag{13}$$

with a dot denoting differentiation with respect to the time. The electric and magnetic field intensities $\mathbf{E}$ and $\mathbf{H}$ are then given by

$$\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}}/c, \quad \mathbf{H} = \nabla\times\mathbf{A}. \tag{14}$$

If we assume that these equations are to remain valid in the abstract space and then go over to the observable space by transformations of the type of Eq. (4), we get in place of (11)

$$\Box^2\bar{\phi} = -4\pi\bar{\rho}, \tag{15}$$

while the other equations are changed in the same way. We see that in the observable space the equations for the electromagnetic field have the same form as in the abstract space, except for the fact that the charge and current densities refer to mean values over an elementary volume. Hereafter, we shall drop the bars from the field

[5] W. Heitler, *The Quantum Theory of Radiation*, second edition (Oxford Press, England, 1944).
[6] The d'Alembertian operator $\Box^2 \equiv \nabla^2 - \partial^2/c^2\partial t^2$.

variables, but it will be understood that they refer to the observable space.

It has been suggested[7] that an electron, because of its fundamental character, should be represented by the simplest of all structures, a point. Let us adopt this standpoint, but let us represent the electron by a point in the *abstract space*. The corresponding representation in the observable space, however, will no longer be a point but an elementary volume.

Consider a point charge e_k having at the time t coordinates x_k, y_k, z_k, and an instantaneous velocity v_k in the x direction. The charge density ρ at a point x, y, z, at the time t is then given by

$$\rho = e_k(1 - v_k^2/c^2)^{-\frac{1}{2}}\delta([x - x_k][1 - v_k^2/c^2]^{-\frac{1}{2}}) \times \delta(y - y_k)\delta(z - z_k), \quad (16)$$

where δ is the Dirac delta-function. It follows that

$$\bar{\rho} = e_k\chi(x - x_k, y - y_k, z - z_k), \quad (17)$$

so that the "effective charge density" is now given by a Gaussian distribution. It is seen that $\bar{\rho}$ has spherical symmetry about the center. Thus this effective charge distribution does not undergo a Lorentz contraction when it is in motion. This is, of course, attributable to the fact that the charge distribution arises from the geometry of the space and not from any assumed internal structure of the electron. A similar behavior was pointed out by March.[2]

In the same way one obtains in the observable space

$$\langle\rho v\rangle = e_k v_k\chi(x - x_k, y - y_k, z - z_k), \quad (18)$$

where v_k is the velocity of the point charge in the abstract space, or that of the center of the charge distribution in the observable space.

We see then that, on the basis of statistical geometry, one obtains a description of the electron in which many of the previous difficulties have been removed. The electron behaves like a distributed charge without any singularities, as far as its interaction with the electromagnetic field is concerned. At the same time any questions concerning the stability of the electron or the nature of the cohesive forces holding the charge together drop out. It is meaningless to ask about

the force exerted on or by a part of the electron, but only the electron as a whole.

The equations for the electromagnetic field associated with a system of electrons can now be written

$$\Box^2\phi = -4\pi \sum_k e_k\chi(x - x_k, y - y_k, z - z_k), \quad (19)$$

$$\Box^2 A = -(4\pi/c) \sum_k e_k v_k \times \chi(x - x_k, y - y_k, z - z_k), \quad (20)$$

with the remaining equations having the same form as (13) and (14).

Let now expand ϕ and A in series of plane monochromatic waves, following Heitler (reference 5, p. 47) and using his notation. For example, if we write

$$\phi = \sum_\sigma a_\sigma(t)\phi_\sigma(x, y, z), \quad (21)$$

then from (19) we obtain

$$\ddot{a}_\sigma + \omega_\sigma^2 a_\sigma = \sum_k e_k\bar{\phi}_\sigma(k), \quad (22)$$

where ω_σ is the (circular) frequency, and $\bar{\phi}_\sigma(k)$ is given by

$$\bar{\phi}_\sigma(k) = \int \phi_\sigma(\xi, \eta, \zeta) \times \chi(\xi - x_k, \eta - y_k, \zeta - z_k)d\xi d\eta d\zeta, \quad (23)$$

which is the mean value of ϕ_σ over the elementary volume of the k'th particle. Equation (22) is the same as the corresponding equation given by Heitler (reference 5, p. 49) except for the fact that here we have $\bar{\phi}_\sigma$ instead of ϕ_σ on the right-hand side. In the same way, in the expansion of the vector potential A in series of transverse and longitudinal plane waves, one gets equations for the coefficients which differ from those given there by Heitler only with respect to this averaging over the elementary volume.

Now the effect of averaging over the elementary volume becomes important as one goes to high frequencies. Consider a plane monochromatic wave, say, of the form

$$\phi = B\exp[i(kx - \omega t)], \quad (24)$$

with $\omega = kc$. Then one finds

$$\bar{\phi} = B[\exp(-k^2a^2/2)]\exp[i(kx - \omega t)] = [\exp(-k^2a^2/2)]\phi = [\exp(-\omega^2a^2/2c^2)]\phi. \quad (25)$$

[7] J. Frenkel, Zeits. f. Physik **32**, 518 (1925); P. A. M. Dirac, Proc. Roy. Soc. **A167**, 148 (1938).

This means that the coupling of an electron with the high frequency components of the electromagnetic field is now weaker than in the classical theory.

If again one follows Heitler (reference 5, p. 50) in writing down a Hamiltonian function for the system and then calculating the energy present in the longitudinal field components, one finds that this can be expressed as a sum of terms of two kinds: (1) interaction energies of pairs of particles, and (2) self-energies of single particles. For the energy of interaction of two particles one obtains, on integrating over the various frequencies and directions of the plane waves,

$$V_{ik} = (e_i e_k/r)\Phi(r/2a),\qquad(26)$$

where e_i, e_k are the charges of the particles, r is the distance between them, and

$$\Phi(x) = 2\pi^{-\frac{1}{2}}\int_0^x [\exp(-u^2)]du.\qquad(27)$$

For $r\gg a$, this goes over into the coulomb interaction, but as r approaches zero, it remains finite, instead of going to infinity as does the coulomb interaction.

The self-energy of a particle W_k one obtains either by going through the same kind of calculation as that leading to Eq. (26) or by getting $\frac{1}{2}V_{kk}$ from (26), letting r approach zero. One finds for the self-energy

$$W_k = \frac{1}{2}V_{kk} = e_k^2/2\pi^{\frac{1}{2}}a.\qquad(28)$$

Incidentally, the same results for V_{ik} and W_k can be obtained by calculating the electrostatic interaction and electrostatic energy of charge distributions of the form given by Eq. (17).

If one equates the expression (28) for an electron to its rest energy m_0c^2 one finds

$$a = 0.79\times10^{-13}\text{ cm}$$

based on the way a was defined in Eq. (2). However, it is not clear at present that such a procedure of equating the electrostatic self-energy to the total rest energy is justified.

Let us now consider the question of quantization. In view of what was found for the equation of the electromagnetic field in the observable space it is plausible to take for the quantum-theoretical equations for the system

consisting of electrons and radiation field the same equations used up to now (reference 5, Chapters II and III) with the sole modification that in the interaction of an electron with the field, the mean value of the latter over an elementary volume must now be used.

The Hamiltonian for the system can be written

$$H = \sum_k H_k + \sum_\lambda H_\lambda + \sum_{i>k} V_{ik} + \sum_k W_k + H'.\qquad(29)$$

Here the Hamiltonian for a particle is given on the basis of the Dirac theory by

$$H_k = c\boldsymbol{\alpha}_k\cdot(\mathbf{p}_k - e_k\langle\mathbf{A}^e\rangle(k)) + m_kc^2\beta_k + e_k\bar{\phi}^e(k),\qquad(30)$$

$\bar{\phi}^e(k)$, $\langle\mathbf{A}^e\rangle(k)$ being the mean values of the potentials of the external field at the k'th particle. The Hamiltonian for a transverse radiation wave is given by

$$H_\lambda = 2\omega_\lambda^2 q_\lambda^* q_\lambda\qquad(31)$$

in terms of the quantized amplitudes q_λ (a discussion of which will be found in reference 5, p. 59). The term H' represents the interaction between the electrons and the transverse, or radiation, field. It is given in the present case by

$$H' = -\sum_k e_k\boldsymbol{\alpha}_k\cdot\langle\mathbf{A}\rangle(k)$$
$$= -\sum e_k\boldsymbol{\alpha}_k\cdot(q_\lambda\langle\mathbf{A}_\lambda\rangle(k) + q_\lambda^*\langle\mathbf{A}_\lambda^*\rangle(k)),\qquad(32)$$

where

$$\langle\mathbf{A}_\lambda\rangle(k) = \int\mathbf{A}_\lambda(\xi,\eta,\zeta)$$
$$\times\chi(\xi-x_k,\eta-y_k,\zeta-z_k)d\xi d\eta d\zeta,\qquad(33)$$

and $\mathbf{A}_\lambda$ is the vector potential for a plane monochromatic transverse wave, suitably normalized (reference 5, p. 59). It follows from Eq. (25) that

$$\langle\mathbf{A}_\lambda\rangle(k) = \exp[-\omega_\lambda^2 a^2/2c^2]\mathbf{A}_\lambda(x,y,z).\qquad(34)$$

It might be pointed out that in (31) we are taking the Hamiltonian of the radiation field in the usual form. An alternative procedure is based on the supposition that the quantization of the radiation field is carried out in the abstract space instead of the observable space. This, however, necessitates giving up the proportionality between energy and frequency and leads to the existence of a maximum value for the energy of a photon. Such a situation does not appear satisfactory.

If one treats the coupling term H' as a perturbation and carries out the usual perturbation-theory calculations, one finds that the use of mean values of the radiation field over an elementary volume, as given by Eq. (34), gets rid of the previous "high frequency" divergence difficulties.

We have already seen how one obtains a finite value for the static self-energy, Eq. (28). By means of a second-order perturbation calculation one can calculate the dynamic (or transverse) self-energy (reference 5, p. 181). One finds, on taking account of the negative energy states of the electron, that this is given by

$$W_k' = (e_k{}^2\hbar c/2\pi a^2 E) - (e_k{}^2/2a\pi^{\frac{1}{2}})$$
$$\times [1 - (m_0{}^2 c^3/2pE) \ln\{(E+cp)/(E-cp)\}], \quad (35)$$

where E is the unperturbed energy and p the momentum of the electron. This result, while finite, is unsatisfactory, however. The first term, which is the important one for small velocities, turns out to have a value which is much larger than $m_0 c^2$, if one takes $a \sim 10^{-13}$ cm.

It appears that this is related to another difficulty: if one calculates the contribution to the self-energy of higher order perturbation terms, one finds that, while the value obtained for each order is finite, the terms of even order (the only non-vanishing terms) keep getting larger as one goes to higher orders, for $a \sim 10^{-13}$ cm. A rough estimate indicates that the ratio of an even-order term to the preceding one is of the magnitude of $e^2\hbar c/a^2(mc^2)^2$, or of 137 for a equal to the classical electron radius. This means, of course, that the perturbation calculation is divergent and hence should not be used in determining the energy of interaction of the electron and the field. It appears that the coupling between the electron and the field is too strong for the perturbation theory to be applicable. It might be pointed out that this difficulty is not peculiar to the present approach. In a sense a similar difficulty also exists in the usual "point-electron" theory, but is masked by the presence of infinities in the individual terms of the perturbation calculation.

As has already been remarked, in the observable space the interaction between a particle and the field involves the mean value of the field over the elementary volume of the particle. Since our knowledge of the field can be obtained only from the observation of its effect on particles, it follows that only such a mean value (a mean of the mean, from the standpoint of the abstract space) can be determined. Hence it is to be expected that in such matters as commutation relations among field variables, etc., one should deal with such mean values over elementary volumes, and not the values at a point.

4. OTHER PARTICLES

In the case of other elementary particles one can use the same method as for electrons: the particle is regarded as a point in the abstract space and, therefore, as an elementary volume in the observable space. This leads to the expectation that all fundamental particles should have the same "size."

The expression for the interaction of a particle with a field in the observable space must take account of the fact that the particle is now an extended source. This can be done as above, by writing down the field equations in the abstract space and then going over to the observable space. From the form of the equations the corresponding Hamiltonian can then be deduced. In general it differs from the Hamiltonian for "point" particles in that the interaction terms involve the average values of the field variables over the elementary volumes of the particles.

In the case of nucleons interacting with a meson field, it is evident that the static interaction between two heavy particles (calculated in first approximation) will not have any singularity. Hence, if such an interaction is used as the potential energy in the Schrödinger equation, no "cutting off" of this interaction energy at small distances between the particles is necessary.

Note added in proof:—It has been kindly pointed out to me by Mr. M. F. M. Osborne that in the book by A. S. Eddington, *Fundamental Theory* (Cambridge University Press, 1946), use was made of an "abstract" space and a Gaussian-error transformation function, although in somewhat different applications.