EXTENDED PARTICLES AND THE INTERPRETATION OF QUANTUM MECHANICS

C K Raju

RESTRICTED COLLECTION

Thesis submitted to the

Indian Statistical Institute

in partial fulfilment of the requirements

for the award of the degree of

Doctor of Philosophy

under the supervision of

Professor Pratul Bandyopadhyay

Calcutta

August 1980

ACKNOWLEDGMENTS

First of all, I must thank Professor Pratul Bandyopadhyay for agreeing to supervise this work, for his extremely tolerant and helpful attitude, and for the benefit of several discussions.

I am indebted to Professor B. Ramachandran for taking the initiative in circulating my work, at the research-program stage, for expert comments. For taking the trouble of going through the research program for this thesis, and offering their comments, I am grateful to Professor P.A.M. Dirac, Professor C.L. Mehta, Professor A. N. Mitra, Professor J.V. Narlikar, Professor A.K. Raychaudhuri and Dr. P.C.W. Davies. Thanks are due to many of the delegates to the Einstein Centenary Symposium, Ahmedabad, for encouragement, criticism, and lively discussions on some of the ideas put forward in this thesis. Particularly, I would like to thank Professor A. Papapetrou for a discussion and detailed comments pertaining to the problem of junction conditions.

I must thank Professor J.W. Moffat for an interesting discussion on the relationship between relativity and quantum mechanics, and Professor G. Bandyopadhyay, Professor J.K. Beem, Professor J.R. Rao, and Professor A.K. Raychaudhuri for further discussions on the problem of junction conditions.

Thanks are due to many friends and colleagues - particularly Dr. N. Kothari for listening patiently on numerous occasions and helping with some calculations ; Dr. R.K. Raychaudhuri for discussions ; and S. Roy for his extensive criticism and for suggesting some useful references. For the work on delta-functions, I must thank S. Ramakrishna for helping with the nonstandard analysis, and B. Bagchi for pointing out an error in an earlier manuscript.

This thesis would not have been possible without the patience and fortitude displayed by my wife, Dr. Jaya Mehta, despite her own research, and despite adverse circumstances.

Finally, I must thank Shri D.K. Bardhan for his quick typing of the manuscript, and Shri Dilip Chatterjee for his efficient duplicating.

PREFACE

This thesis is based on the following papers, details of which are given in the list of references.

1. Relativistic and Statistical Foundations of Quantum Mechanics
2. Products and Compositions with the Dirac Delta-Function
3. Junction Conditions in General Relativity
4. A Relativistic Model of the Electron
5. Interpretation of the Indeterminacy Relations
6. Interpretation of Quantum Mechanics as a Theory of Extended Particles
7. Classical Time-Symmetric Electrodynamics

Except for certain minor alterations, four of the papers (3 - 6) have been reproduced almost verbatim. (2) has been attached as an appendix, not because it does not belong to the main body of the work, but because it could not be included anywhere without breaking the continuity of the presentation. (7) has been sketched in another appendix, insofar as it has a direct bearing on the logical development of the thesis. (1) contained the research program for this work, along with some preliminary results.

Since the basic problems considered are somewhat difficult, and an altogether new line of thought is offered, it is natural that some unsolved problems remain. Nevertheless, significant headway has been made in obtaining an internally consistent theory. Since this theory makes no new assumptions, and leads to empirical predictions, its implications are of importance, either way.

CONTENTS

INTRODUCTION

This thesis deals with the following two problems

(a) The structure, stability and dynamics of spatially-extended, elementary particles

(b) The interpretation of quantum mechanics

Since these problems have traditionally been thought to be unrelated, we sketch the logical nexusus involved.

0. The word 'particle' will denote a spherically-symmetric, elementary particle of nonzero rest mass, such as the electron.

Chapters 1 and 2

1. Real particles are not mass points

2. Given the imbalance between gravitational and electromagnetic forces, an extended particle cannot be stable, within the existing theory of relativity, unless the usual junction conditions are modified and the particle corresponds to a shell-like material distribution.

3. To modify the junction conditions, it is necessary to define products and compositions with the Dirac delta-function in such a manner that our continued belief in the Einstein equations is justified.

4. Products and compositions with the Dirac delta-distribution are defined, using nonstandard analysis. It is ensured that our continued belief in the Einstein equations would not be phenomenological.

5. With the modified junction conditions, the results obtained by the present methods, for (uncharged) spherically-symmetric surface layers, are compared with the results obtained by earlier authors. These agree although the equations appear to be overdetermined, because only restricted coordinate transformations of flat spaces have been used to obtain the internal metric.

6. By admitting spaces obtained from flat space by more general coordinate transformations, it is shown that the equations of motion for the surface layer are manifestly underdetermined. Thus, it is possible for extended particles to exist in relativity.

7. This procedure is generalised to include charge by taking the external metric to be the Reissner-Nordstrom metric.

Chapters 3 and 4

1. It is observed that an interpretation of the precise form of the indeterminacy relations necessarily leads to the conclusion that the particles described by quantum mechanics must have some finite spatial extension.

2. It is proposed to interpret quantum mechanics as the semiclassical description of the dynamics of extended particles of the type constructed in Chapter 2.

3. The frequency of oscillation, of the particle, is chosen to satisfy $E = h\nu$ in the rest frame.

4. Statistical considerations are introduced by observing that the particle is statistically coupled to the rest of the universe. The (now) random extension of the particle is related to the wavefunction.

5. It is shown that this wavefunction satisfies the Schrodinger equation in statistical equilibrium.

6. It is shown that this wavefunction admits the usual probability interpretation as an excellent approximation and that a representation of observables by positive-operator valued measures is naturally admitted.

7. The empirical consequences of this theory are pointed out along with some unsolved problems. Therefore, this theory is empirically verifiable.

The scheme of presentation does not adhere to the strict logical outline given above. Rather, each problem, encountered on the way, has been studied for its own sake. It is hoped that the effects of the resulting fallout would be interesting enough to offset the consequent breaks in continuity. The need for extended particles has been elaborated below, although some of the stronger points are apparent only a posteriori.

2. <u>The need for extended particles</u>.

It is a historical paradox that, whereas the assumption of point masses goes unquestioned, it is necessary to justify the hypothesis that real particles have some small, but finite, spatial extension. Since it may never be possible to demonstrate, as an element of physical reality, that real particles are akin to geometric points, the principal virtue that can be attributed to this assumption is its simplifying value. But, does this assumption really have great simplifying value ? In Newtonian mechanics, one could definitely answer the preceeding question in the affirmative. But, already, in classical field theory, this assumption leads to difficulties. In quantum theory, these singularities become unabashed confessions of ignorance in the form of infinite renormalization constants. These difficulties cannot simply be wished away by the pious hope that they would be resolved in the not-too-distant future.

If the extensive, and, so far, unsuccessful, efforts in the past are any guide, they call for a bleakly pessimistic view.

Apart from the (essentially) aesthetic criterion of simplicity, in the absence of any definite empirical evidence, any extra hypothesis should be anchored on theoretical necessity. But, there is no theoretical necessity for point masses. On the contrary, as the arguments in Chapters 1 and 3 show, a very definite case can be made out for the opposite point of view. Thus, it is asserted that the quantum mechanical electron cannot, in principle, be localised, and, hence, must correspond to an extended distribution of charged matter. Similarly, according to relativity, real particles can never be point masses, because point masses are black holes. Although this last argument is not applicable to charged particles, it is difficult to believe that charged and uncharged particles are essentially different in structure. This would, further, make it difficult to account for decays of the type $n \longrightarrow p + e + \nu_e$.

In short, the assumption of point masses could wreak havoc on some of the fundamental principles of physics.

While there is no guarantee that these principles will continue to survive at the attoscopic level, one cannot just brush them aside for the sake of a single unsubstantiated assumption.

Now, the hypothesis of extended particles brings in its wake its own brand of difficulties. These difficulties pertain both to the structure and the dynamics of extended particles. And, it is the purpose of this thesis to show that these difficulties are not insurmountable, as has been thought in the past.

One final point that needs mention. The structure assigned to the particle is not quite arbitrary. With the assumption of spherical symmetry in the rest frame, essentially two types of structures are possible. The restriction to an infinitesimally thin shell seems to be required by relativity, quantum mechanics, and time-symmetric electrodynamics.

ABSTRACTS

I JUNCTION CONDITIONS IN GENERAL RELATIVITY

A new formulation of the problem of junction conditions is given. It is pointed out that, if the existing theory of relativity is to be consistent with the existence of matter in the form of particles, then the $g_{\mu\nu}$ cannot be continuously differentiable everywhere. The mathematical part of the problem of junction conditions is solved by using nonstandard analysis to define products and compositions with distributions. The definitions are such that continued belief in the equations of relativity is justified. As an application, the equations of motion for the spherically-symmetric surface layer, at the Schwarzschild-Minkowski junction, are derived. These agree with the equations derived by earlier authors in being under-determined. Applications to singularity theory are pointed out.

II A RELATIVISTIC MODEL OF THE ELECTRON

The problem of the motion of surface layers in relativity is considered in its most general form using the techniques developed in the earlier chapter. By using general coordinate transformations it is shown that the equations of motion are manifestly underdetermined. This indeterminacy can be overcome by prescribing an equation of state, in the macroscopic case. It is concluded that surface layers, in the microscopic case, can evolve in an essentially arbitrary manner, and that it is possible to construct shell-like models of elementary charged particles satisfying the restrictions in Chapter IV.

III INTERPRETATION OF THE INDETERMINACY RELATIONS

It is pointed out that, within the axiomatic formulation of quantum mechanics, the precise form of the indeterminacy relations introduces some qualitatively new features. As a result, the notion of simultaneous measurement, which is an integral part of the usual interpretation of the indeterminacy relations, becomes redundant and misleading. It is shown that the precise form, of the indeterminacy relations, necessarily leads to the conclusion that the particles described by quantum mechanics have some finite (as opposed to infinitesimal) spatial extension.

IV INTERPRETATION OF QUANTUM MECHANICS AS A THEORY OF EXTENDED PARTICLES

Developing on the conclusions reached in the previous chapter, it is proposed to interpret quantum mechanics as a theory of extended particles. Certain restrictions are placed on the underlying model for extended particles. Wave-particle duality is interpreted in the context of the pulsations of the particle. The wavefunction is related to the (random) extension of the particle. It is shown that this wavefunction satisfies the Schrodinger equation. In this theory, the peculiarities of quantum probabilities are related to the assumption that the particle is shell-like. It is shown that a representation of dynamical variables by positive-operator valued measures is possible. The empirical predictions of this theory are pointed out, along with some unsolved problems. It is concluded that it is, at least partially, possible to interpret quantum mechanics as a semiclassical description of the dynamics of extended particles. If this interpretation is correct, quantum mechanics would fail at very high energies, and, possibly, at very low energies.

Appendix A - PRODUCTS AND COMPOSITIONS WITH THE DIRAC DELTA-FUNCTION

The need for defining pointwise products and compositions with distributions is pointed out, in the context of the problems of junction conditions and curved shock waves. Earlier definitions are briefly reviewed, and new definitions are proposed using nonstandard analysis. Basic properties are established, and some products and compositions with the delta distribution are explicitly evaluated. With these definitions the domain of validity of certain equations of physics can be extended to include discontinuous fields, without introducing new phenomenology. As an example, the Rankine-Hugoniot equations are derived from the Euler equations.

NOTATION

C^0	... ...	space of continuous functions
C^k	... ...	space of k-times continuously differentiable functions
D	... ...	space of test functions
D'	... ...	space of distributions
$\mathbb{R}$	... ...	Real line
L^2	... ...	space of square integrable functions (w.r.t. some probability measure)
$,\upsilon$	... ...	differentiation with respect to υ
$;\upsilon$	... ...	covariant differentiation with respect to υ
δ	... ...	Dirac delta-function
χ^+	... ...	Heaviside function
$[\ ]$	... ...	defined on p 12
$\vert$	... ...	defined on p 12
${}^*a$	... ...	element corresponding to a in nonstandard superstructure
${}^*\mathbb{R}$	... ...	nonstandard reals
$\overset{*}{=}$	... ...	nonstandard equality

$\otimes$	... ...	convolution
$(\Omega, \mathcal{B}, P)$	...	Standard Borel probability space (in the sense of measure theory)
E	... ...	Expectation
δ_p	... ...	Standard deviation
Var	... ...	Variance
$\Vert \ \Vert$	... ...	Hilbert norm
$\Vert \ \Vert_o$	... ...	supremum norm
$\Vert \ \Vert_1$	... ...	C^1 seminorm
$\Vert \ \Vert_2$	... ...	L^2 norm
$\langle f, g \rangle$	... ...	inner product of f and g or the value of the functional f at g

CHAPTER I

JUNCTION CONDITIONS IN GENERAL RELATIVITY

1.1 Introduction

The problem of junction conditions in general relativity is usually stated thus : at a boundary between matter and spacetime (called a junction) how smooth should the ten distinct functions, of the coordinates, $g_{\mu\nu}$ be ? A number of papers have been written on this problem since Lanczos (1924), and we shall not go into the merits or demerits of these papers, except to point out that a general consensus was reached with the publication of Lichnerowicz's (1955) treatise. According to this consensus, the junction conditions may be written in the form used by Synge (1966) : if Σ is the hypersurface of discontinuity, with equation $f(x_i) = 0$, and $T_{\mu\nu}$ is the material energy tensor, then

$$T^{\mu}_{\nu} \; f_{,\mu} = \; (C) \; , \qquad (1.1.1)$$

(C) indicating any quantity that is continuous across Σ. The above conditions state, in effect, that only the <u>second</u> derivative of the $g_{\mu\nu}$ may have a discontinuity across a junction.

More recently, this consensus has been modified by the work of Dautcourt (1964), Israel (1966), Papapetrou and Hamoui (1968, 1979), and Evans (1977). In effect, the junction conditions have been altered to permit the occurrence of discontinuities in the first derivative of the $g_{\mu\nu}$, i.e., $g_{\mu\nu} \in C^0$. This alteration, however, has been proposed as a matter of convenience, for the practical purpose of solving the Einstein equations - for theoretical purposes, the junction conditions (1.1.1) continue to be valid.

In the present analysis of the problem, we propose to probe a little deeper. To begin with, the above problem is really equivalent to the following two problems :

(a) To determine the highest order of continuity, for the $g_{\mu\nu}$, that can be taken to represent physical reality.

(b) To determine the lowest order of continuity, for the $g_{\mu\nu}$, that is mathematically feasible and consistent with our belief in the equations of relativity.

We will attempt to answer these questions within Einstein's theory of relativity, for definiteness, although the conclusions are equally applicable to other classical theories.

1.2 Relativistic characterisation of matter

In attempting to solve problem (a) we observe that the existing constructions of the material energy tensor, corresponding to a real material distribution, are macroscopic and outmoded. According to the modern understanding of matter, matter consists of various combinations of elementary particles (such as electrons, quarks, light quarks, mesons and neutrinos), although the only particle that can be called elementary, with certainity, is the electron.

Since the theory of relativity is formulated without regard to scale, it is quite essential that it should be consistent with the empirical fact that matter exists in discrete clusters at the microscopic level. In particular, it should be possible for particles to exist in the theory. Thus, although some attempts have been made (for instance, Synge, 1966) to incorporate the particle nature of matter, in the representation of matter by the material energy tensor, these are not satisfactory, because it is not clear, a priori, that particles can exist in the theory.

In trying to characterise the distribution of matter, corresponding to an elementary particle, with non-zero rest mass, the simplest possibility which arises is that of a point mass. This possibility has been rejected by Dirac (1962b) on the following grounds ; a point mass would behave like a black hole, real particles do not behave like black holes, hence real particles are not point masses. However, given the empirically determined masses and charges of real elementary particles, the Reissner-Nordstrom solution assures us that the above argument is not applicable to real charged particles. Given, further, that real particles are often charged, the argument loses much of its significance. Nevertheless, even in the Reissner-Nordstrom metric, there is a singularity at the origin, so if we accept the point of view that real particles are pointlike, the $g_{\mu\nu}$ cannot be continuous everywhere.

The other possibility, for which there is some empirical support, is that of an extended mass distribution. The stability of such a mass-distribution is a difficult problem, especially in the case of charged particles, because of the inordinate imbalance between gravitational and electromagnetic forces. Given that electrons, for instance, are reasonably stable, if we subscribe to the view that real particles are extended, we seem to be faced with the alternatives of introducing new phenomenology or abandoning the existing theory of relativity at the microscopic level. However, in view of the earlier results on stationary surface layers (for instance, Papapetrou and Hamoui, 1968) there is a third possibility, viz., the $g_{\mu\nu}$ may not be of class C^1 as is usually supposed. If we accept the theory of relativity, and Occam's razor, we are forced to accept the third possibility.

The conclusions reached above may be summarised by saying that the existing theory of relativity (without any additional hypotheses) is consistent with the existence of particles only if the $g_{\mu\nu}$ are at most of class C^0.

In considering discontinuities in the $g_{\mu\nu}$, or their first derivatives, as a part of physical reality, and not as an approximation, the conventional methods of correlating a material distribution with a material energy tensor (via the analogy with Newtonian hydrodynamics) break down, if only because of the possible appearance of 'functions' of the form δ^2 (square of the Dirac δ-function) in the material energy tensor. There is, of course, nothing 'unphysical' about the corresponding matter densities, since we have just demonstrated that the conventional prejudice concerning the existence of smooth matter-densities has no empirical basis. However, to avoid any problems regarding the physical interpretation of such matter densities, here we will adopt the alternative of considering the $g_{\mu\nu}$ as logically primitive. In other words, the notion of spacetime would be accepted as basic, and the notion of matter would appear as a derived concept. If a shell-like model of elementary particles is contemplated, as in Dirac (1962a, 1962b) or Raju (1979, 1980a), the occurrence of all matter would be related to the occurrence of discontinuities in the $g_{\mu\nu}$, or their derivatives.

The process of accepting the $g_{\mu\nu}$ as basic does have a fluid-mechanical analogue in conventional shock-wave theory. Here, the Rankine-Hugoniot equations characterise a normal shock (of infinite extent) in terms of the flow-field behind, and in front of the shock. However, the present methods are more general, and instead of using fluid mechanics as an analogy for deriving results in relativity one can apply the present methods to derive the general equations for a curved shock. Finally, taking the $g_{\mu\nu}$ as basic might appear to reduce the Einstein equations to simple algebraic equations. But it will be seen in the sequel that this does not happen.

1.3 Products and compositions with distibutions

We now go on to consider problem (b) and decide whether the occurrence of discontinuities in the $g_{\mu\nu}$, or their first derivatives, is mathematically feasible and consistent with our belief in the Einstein equations. The effect of such discontinuities on the smoothness of the Christoffel symbols, Ricci tensor etc., can be easily seen from the usual equations (1.4.16). It is, thus, clear that the mathematical aspect of the problem consists of the following :

(1) Defining products of distributions (i.e., defining entities such as δ^2, $\delta\,\delta'$ etc.) in a logically consistent manner, and interpreting equations containing such products. Investigating the validity of the usual algebraic laws for the product.

(2) Defining compositions of distributions with ordinary functions (i.e., defining entities such as $\delta(g(x))$ and investigating the possible validity of a chain rule.

The problem, however, is not a purely mathematical one, because, as will be seen below, many definitions have been given, all of which suffer from the drawback that our belief in the Einstein equations is reduced to a purely phenomenological one. Therefore, the physical aspect of the problem introduces the additional constraint of defining the above entities in such a manner that our belief in the Einstein equations would continue to be justified. A new solution to this problem is given in Appendix A , following Raju (1980b).

The simple results obtained there, effectively solve the mathematical problem associated with the problem of junction conditions - the $g_{\mu\nu}$ may be chosen to be discontinuous. We observe that no new phenomenology has been introduced, and the results derived using nonstandard techniques could very well have been derived without using them (Robinson, 1966). However, the solution is not quite complete because (i) the Ricci tensor need not remain symmetric ; and (ii) in case the $g_{\mu\nu}$ are chosen as discontinuous, there would be four expressions for the Ricci tensor because of the failure of the commutative law, and the correct expression remains to be determined.

One final point that needs mention : according to the above techniques, one cannot arbitrarily prescribe the $g_{\mu\nu}$ on the two sides of a hypersurface of discontinuity. This is because the hypersurface would evolve according to the equations $T^{\mu\nu}{}_{;\nu} = 0$, which, by virtue of (1.3.3) and (1.3.4), would be an overdetermined system of differential equations, involving the defining function f of the hypersurface. Thus, only those hypersurfaces and functions $g_{\mu\nu}$ can be chosen that satisfy the consistency requirements. Because the above equations are nonlinear, the consistency conditions, in the general case, cannot be written down, and they have to be written down seperately for specific problems. However, as an illustration (from fluid mechanics) we observe that the normal shock equations are valid, provided $\mu(P^{-} - P^{+})/\rho^{+} U = 0$, i.e., strong shocks would curve due to the effect of viscosity (μ = viscosity, P = pressure, ρ = density, U = shock velocity, and the superscript + refers to the undisturbed fluid ahead of the shock).

1.4 Applications

As an illustration, and as one of the applications of the above methods, we now propose to check, for consistency, the results obtained by Papapetrou and Hamoui (1968, 1979). The problem is to determine the motion of the spherically - symmetric layer of matter, at the junction between the Schwarzschild and Minkowski metrics. It is assumed that $g_{\mu\nu}$ are continuous across the hypersurface, and that some of their first derivatives have essential discontinuities.

Since different coordinate systems can sometimes give rise to physical differences, we start with the radiative coordinates used by the above authors : $x^{0} = r$, $x^{1} = u$ $x^{2} = \theta$, $x^{3} = \emptyset$. We assume that the hypersurface Σ is spherically symmetric and spacelike, so that its equation can always be written (locally) in the form

$$r = f(u) \tag{1.4.1}$$

The external and internal metrics are respectively given by

$$ds_+^2 = (1 - \frac{2m}{r})\, du^2 + 2dudr - r^2(d\theta^2 + \sin^2\theta d\phi^2),\ r > f(u)$$

$$dS_-^2 = d\,U^2 + 2d\,U\,dR - R^2(d\theta^2 + \sin^2\theta d\phi^2),\ R < F(U), \tag{1.4.2}$$

where $R = F(U)$ is the equation of the hypersurface in the interior coordinates (R, U, θ, ϕ). Carrying out the changes of variables $r \longrightarrow z = r - f(u)$, $R \longrightarrow Z = R - F(U)$, the internal and external metrics may be written

$$dS_-^2 = (1+2F')dU^2 + 2dUdZ - (Z+F)^2(d\theta^2 + \sin^2\theta d\phi^2),\ Z < 0,$$

$$ds_+^2 = (1 - \frac{2m}{z+f} + 2f')du^2 + 2dudz - (z+f)^2(d\theta^2 + \sin^2\theta d\phi^2),\ z > 0. \tag{1.4.3}$$

We suppose that the coordinates (U, Z, θ, ϕ) can be obtained from the coordinates (u, z, θ, ϕ) by the change of variables

$$U = \alpha(u),$$

$$Z = \gamma(u)z,\ \gamma > 0, \tag{1.4.4}$$

and that the junction conditions[1]

$$[g_{\mu\nu}] = 0, \qquad (1.4.5)$$

are satisfied. (1.4.4) and (1.4.5) together yield

$$\begin{aligned} F(U) &= f(u), \\ \alpha' \gamma &= 1, \qquad (1.4.6) \\ \alpha'(\alpha' + 2f') &= 1 - \frac{2m}{f} + 2f'. \end{aligned}$$

These equations were obtained by Papapetrou and Hamoui (1968)

In the present calculation, we require the same coordinates throughout, and in the coordinates (u, z, θ, ϕ) the external and internal metrics are given, (using the junction conditions) by

$$ds_+^2 = \left(1 - \frac{2m}{(z+f)} + 2f'\right)du^2 + 2dudz - (z+f)^2(d\theta^2 + \sin^2\theta d\phi^2), \quad z > 0,$$

$$ds_-^2 = \left(1 - \frac{2m}{f} + 2f' + 2\alpha'\gamma' z\right)du^2 + 2dudz - (\gamma z + f)^2(d\theta^2 + \sin^2\theta d\phi^2), \quad z < 0.$$

$$(1.4.7)$$

[1]See further, equations (1.4.8), (1.4.10) and (1.4.15).

1.4.2 Notation and the general formalism

For simplifying the calculations, we introduce the following notation

$$\chi^{+} = \chi_{(0,\infty)}(z) = H(z), \text{ the Heaviside function,}$$
$$\chi^{-} = \chi_{(-\infty,0)}(z) = 1 - H(z)\,. \tag{1.4.8}$$

If $h(u, z)$ is any function such that

$$h(u, z) = h^{+}(u, z)\,\chi^{+} + h^{-}(u, z)\,\chi^{-}, \tag{1.4.9}$$

then we define

$$[h] = \lim_{z \to 0^{+}} h^{+}(u,z) - \lim_{z \to 0^{-}} h^{-}(u,z), \tag{1.4.10}$$

$$h| = \frac{1}{2}\lim_{z \to 0^{+}} h^{+}(u,z) + \frac{1}{2}\lim_{z \to 0^{-}} h^{-}(u,z). \tag{1.4.11}$$

We observe the following properties

$$\chi^{+}_{,1} = \delta(z), \quad \chi^{-}_{,1} = -\delta(z),$$
$$\chi^{+}.\,\chi^{+} = \chi^{+}, \quad \chi^{-}.\,\chi^{-} = \chi^{-}, \quad \chi^{+}.\,\chi^{-} = 0 \tag{1.4.12}$$

$$\chi^{+}.\,\delta = \chi^{-}.\,\delta = \frac{1}{2}\,\delta,$$
$$h(u,z)\delta = h|\,\delta\,. \tag{1.4.13}$$

Further,

$$
\begin{aligned}
[h_1 + h_2] &= [h_1] + [h_2], \\
[h_1 h_2] &= h_1| \, [h_2], \text{ if } [h_1] \neq 0, \\
(h_1 + h_2)| &= h_1| + h_2|, \\
(h_1 h_2)| &= h_1| \, h_2|, \text{ if } [h_1] = 0 \text{ or } [h_2] = 0 \\
[h], \overset{\circ}{} &= [h, \overset{\circ}{}]
\end{aligned}
\qquad (1.4.14)
$$

With the above notation, we may put

$$g_{\mu\nu} = \overset{+}{g}_{\mu\nu} \, \chi^{+} + \overset{-}{g}_{\mu\nu} \, \chi^{-} . \qquad (1.4.15)$$

With the usual formulae

$$
\begin{aligned}
\Gamma_{\mu\nu\sigma} &= \frac{1}{2} (g_{\mu\nu,\sigma} + g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu}), \\
R_{\mu\nu} &= \Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\nu} \Gamma^{\beta}_{\alpha\beta} + \Gamma^{\alpha}_{\mu\beta} \Gamma^{\beta}_{\nu\alpha}, \\
T^{\mu\nu} &= R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R,
\end{aligned}
\qquad (1.4.16)
$$

we obtain, in view of (1.4.5) and (1.4.12)

$$\Gamma^{\mu}_{\nu\sigma} = \Gamma^{\mu +}_{\nu\sigma} \, \chi^{+} + \Gamma^{\mu \pm}_{\nu\sigma} \, \chi^{-} . \qquad (1.4.17)$$

It follows that the Ricci tensor is given by

$$R_{\mu\nu} = \overset{+}{R}_{\mu\nu} \, \chi^{+} + \overset{-}{R}_{\mu\nu} \, \chi^{-} + S_{\mu\nu} , \qquad (1.4.18)$$

where $S_{\mu\nu} = [\Gamma^{\alpha}_{\mu\alpha}] \, \chi^{+}_{,\nu} - [\Gamma^{1}_{\mu\nu}] \delta ,$

The material energy tensor is given by

$$T^{\mu\nu} = T_+^{\mu\nu} \chi^+ + T_-^{\mu\nu} \chi^- + t^{\mu\nu} ,$$
$$t^{\mu\nu} = S^{\mu\nu} - \frac{1}{2} g^{\mu\nu} S , \qquad (1.4.19)$$
$$S = g^{\mu\nu} S_{\mu\nu} .$$

Since the interior and exterior fields have been chosen to be free, the only equations to be satisfied are

$$t^{\mu\nu}{}_{:\nu} = 0 . \qquad (1.4.20)$$

We observe that while some of the Christoffel symbols are discontinuous, and, therefore, not defined at $z = 0$, (1.4.20) still makes sense. This is because only products (of δ-functions) with the Christoffel symbols enter into (1.4.20), and these are well defined. Hence, it is not necessary to rewrite (1.4.20) in any other, more complicated form, as has been done by earlier authors. To further simplify the derivation of the equations of motion, we put

$$a^+ = a^+(u,z) = 1 - \frac{2m}{(z+f)} + 2 f' ,$$
$$a^- = a^-(u,z) = 1 - \frac{2m}{f} + 2 f' + 2\alpha'\gamma' z ,$$
$$b^+ = b^+(u,z) = -(z+f)^2 , \qquad (1.4.21)$$
$$b^- = b^-(u,z) = -(\gamma z + f)^2 ,$$
$$a = a^+ \chi^+ + a^- \chi^- ,$$
$$b = b^+ \chi^+ + b^- \chi^- .$$

It follows that, denoting $a_{,1}$ by a_1 etc.,

$$[a] = 0, \qquad [b] = 0,$$
$$[a_0] = 0, \qquad [b_0] = 0, \qquad (1.4.22)$$
$$[a_1] = 2\left(\frac{m}{f^2} - \alpha' \gamma'\right), \qquad [b_1] = 2f(\gamma - 1),$$

$$a| = 1 - \frac{2m}{f} + 2f', \qquad b| = -f^2,$$
$$a_0| = \frac{2m}{f^2} f' + 2f'', \qquad b_0| = -2ff', \qquad (1.4.23)$$
$$a_1| = \frac{m}{f^2} + \alpha' \gamma', \qquad b_1| = -f(\gamma+1).$$

The junction conditions (1.4.5) have been used in writing down the above equations.

1.4.3 The equations of motion

Because of the new notation and techniques, the various steps in the derivation of the equations of motion are given below, in detail. In the new notation, the non-vanishing components of the metric tensor are given by

$$\overset{\pm}{g}_{00} = \overset{\pm}{a}, \qquad \overset{\pm}{g}_{22} = \overset{\pm}{b},$$
$$\overset{\pm}{g}_{01} = 1, \qquad \overset{\pm}{g}_{33} = \overset{\pm}{b} \operatorname{Sin}^2 \theta, \qquad (1.4.24)$$

$$\underset{\pm}{g}^{01} = 1, \qquad \underset{\pm}{g}^{22} = b_{\pm}^{-1},$$
$$\underset{\pm}{g}^{11} = -\overset{\pm}{a}, \qquad \underset{\pm}{g}^{33} = b_{\pm}^{-1} \operatorname{Sin}^{-2} \theta. \qquad (1.4.25)$$

Similarly, the nonvanishing Christoffel symbols are given by

$$\Gamma^{0\,\pm}_{00} = -\frac{1}{2} a^{\pm}_{1}, \qquad \Gamma^{2\,\pm}_{02} = \frac{1}{2} b^{-1}_{\pm} b^{\pm}_{0},$$

$$\Gamma^{0\,\pm}_{22} = -\frac{1}{2} b^{\pm}_{1}, \qquad \Gamma^{2\,\pm}_{12} = \frac{1}{2} b^{-1}_{\pm} b^{\pm}_{1},$$

$$\Gamma^{0\,\pm}_{33} = \Gamma^{0\,\pm}_{22} \operatorname{Sin}^2 \theta, \qquad \Gamma^{2\,\pm}_{33} = -\operatorname{Sin}\theta \operatorname{Cos}\theta,$$

$$\Gamma^{1\,\pm}_{00} = \frac{1}{2} (a^{\pm} a^{\pm}_{1} + a^{\pm}_{0}), \qquad \Gamma^{3\,\pm}_{03} = \frac{1}{2} b^{-1}_{\pm} b^{\pm}_{0}, \tag{1.4.26}$$

$$\Gamma^{1\,\pm}_{01} = \frac{1}{2} a^{\pm}_{1}, \qquad \Gamma^{3\,\pm}_{13} = \frac{1}{2} b^{-1}_{\pm} b^{\pm}_{1},$$

$$\Gamma^{1\,\pm}_{22} = \frac{1}{2} (a^{\pm} b^{\pm}_{1} - b^{\pm}_{0}), \qquad \Gamma^{3\,\pm}_{23} = \operatorname{Cot}\theta .$$

$$\Gamma^{1\,\pm}_{33} = \Gamma^{1\,\pm}_{22} \operatorname{Sin}^2 \theta,$$

The distribution part of the Ricci tensor is given, according to (1.4.19), by

$$S_{00} = -[\Gamma^{1}_{00}]\delta = -\frac{1}{2} a| \, [a_1]\delta,$$

$$S_{01} = [\Gamma^{0}_{00} + 2\,\Gamma^{2}_{02}]\delta = -\frac{1}{2} [a_1]\delta = S_{10},$$

$$S_{11} = 2\,[\Gamma^{2}_{12}]\delta = b^{-1}| \, [b_1]\delta, \tag{1.4.27}$$

$$S_{22} = -[\Gamma^{1}_{22}]\delta = -\frac{1}{2} a| \, [b_1]\delta,$$

$$S_{33} = -[\Gamma^{1}_{33}]\delta = S_{22} \operatorname{Sin}^2 \theta .$$

The same tensor, with raised suffices is given by

$$S^{00} = b^{-1}|\,[b_1]\delta,$$
$$S^{01} = -\tfrac{1}{2}[a_1 + 2ab^{-1}\,b_1]\delta = S^{10},$$
$$S^{11} = \tfrac{1}{2}[aa_1 + 2a^2b^{-1}\,b_1]\delta, \qquad (1.4.28)$$
$$S^{22} = -\tfrac{1}{2}ab^{-2}|\,[b_1]\delta,$$
$$S^{33} = S^{22} / \mathrm{Sin}^2\,\theta,$$

and its trace by

$$S = -2ab^{-1}|\,[b_1]\delta - 2\,[a_1]\delta. \qquad (1.4.29)$$

The components of the material energy tensor are given by

$$t^{00} = b^{-1}|\,[b_1]\delta,$$
$$t^{01} = 0 = t^{11},$$
$$t^{22} = \tfrac{1}{2}ab^{-2}|\,[b_1]\delta + \tfrac{1}{2}b^{-1}|\,[a_1]\delta, \qquad (1.4.30)$$
$$t^{33} = t^{22} / \mathrm{Sin}^2\,\theta.$$

The last two of the equations (1.4.20) are identically satisfied, and the other two are

$$t^{00}{}_{:\upsilon} = t^{00}{}_{,0} + (\Gamma^0_{00} + \Gamma^{\upsilon}_{0\upsilon})\,t^{00} + 2\,\Gamma^0_{22}\,t^{22} = 0,$$
$$t^{1\upsilon}{}_{:\upsilon} = \Gamma^1_{00}\,t^{00} + 2\,\Gamma^1_{22}\,t^{22} = 0, \qquad (1.4.31)$$

Explicitly, we have to solve

$$(b^{-1}|\,[\,b_1\,])_{,0} - \frac{1}{2}\,a_1 b^{-1}|\,[\,b_1\,] - \frac{1}{2}\,ab^{-2}b_1|\,[\,b_1\,]$$

$$+\, b^{-2}b_0|\,[\,b_1\,] - \frac{1}{2}\,b^{-1}b_1|\,[\,a_1\,] = 0, \qquad (1.4.32)$$

$$(aa_1 + a_0)\,|\,[\,b_1\,] + (ab_1 - b_0)\,|\,[\,ab^{-1}b_1 + a_1\,] = 0,$$

along with $\alpha'\,\gamma = 1$,

$$\alpha'(\alpha' + 2f') = a|\,.$$

The validity of constant solutions, which can be easily checked, implies that these equations are consistent. In fact, following a suggestion made by Papapetrou (1980; private communication) we show below that the first two of (1.4.32) reduce to identities by virtue of the last two.

Thus, the first of (1.4.32) can be written in the form

$$a|\,(\gamma^2 - 1) - 2\,(\gamma - 1)\,f' = \frac{2m}{f}\,, \qquad (1.4.33)$$

where we have used the values of the bars and brackets, given in (1.4.22) and (1.4.23). From the last two of (1.4.32)

$$a|\,\gamma^2 - 2\,f'\gamma - 1 = 0\,. \qquad (1.4.34)$$

Substituting in (1.4.33), we have an identity by virtue of the value of $a|$ given in (1.4.23). Similarly, the second of

(1.4.32) can be shown to be an identity by repeated use of (1.4.34).

It follows that the hypersurface of discontinuity corresponds to a material distribution. The defining function f can evolve arbitrarily, and some additional conditions, like an equation of state, are necessary to make the motion of the hypersurface determinate. This point is discussed in greater detail in section 2.4.

These results, therefore, agree with the results obtained by Papapetrou and Hamoui (1968 ; 1979) as an extra equation is required for specifying the motion of the hypersurface. However, the techniques developed here are more general and are easier to handle. Thus, earlier authors could not write the equations of motion in the simple form (1.4.20), because the $\Gamma^{\mu}_{\nu\sigma}$ are undefined at $z = 0$ and (1.4.20) makes sense only after the product of a discontinuous function with a delta function has been defined. Again, with the theories developed earlier, it is computationally very difficult to handle the axisymmetric case, and these theories are not designed to deal with discontinuities in the $g_{\mu\nu}$. With the present theory, a solution to both these cases is feasible, but would be given elsewhere.

1.4.4 Other applications

One of the reasons for the apparently overdetermined nature of the equations (1.4.32) is the restriction that the functions α and γ, defined by (1.4.4), be functions of the single variable u. In the next chapter, it is proposed to drop this restriction, and use the Reissner-Nordstrom metric to construct a shell-like model of extended charged particles.

Other applications, to singularity theory, are obvious. Much of the work in singularity theory uses hypotheses which imply that the $g_{\mu\nu}$ are continuously differentiable. Therefore, according to Section 2, these results are not consistent with the existence of matter in the form of particles, and, hence, are not applicable to the real universe.

1.5 Conclusions

In the existing theory of relativity, the $g_{\mu\nu}$ cannot be everywhere continuously differentiable. Discontinuities in the $g_{\mu\nu}$ are permissible. The spherically-symmetric surface layer at the Schwarzschild-Minkowski junction can be stationary.

CHAPTER II

A RELATIVISTIC MODEL OF THE ELECTRON

2.1 Introduction

In this chapter we consider the following two problems :

(i) Can a charged, spherically symmetric, surface layer exist in a more-or-less stable manner ?

(ii) Can such a surface layer oscillate ?

More specifically, these problems are considered in the context of the problem of constructing extended, shell-like models of elementary charged particles. In this context, the significance of these problems rests on the fact that relativity theory deals with matter, and matter consists of particles. Therefore, it would seem essential, for external consistency, that particles should be allowed to exist in the theory of relativity. Various reasons favouring an extended model of the electron, for example, are well known (for instance, Dirac, 1962a, 1962b), however Raju (1980a, 1980d, 1980e) has put forward some new arguments. Although these new arguments favour a shell-like structure for elementary particles, such structures are also required on purely relativistic grounds if extended particles are to exist in relativity theory.

At this stage there arises a very fundamental question. Since the nexus between relativity and nuclear physics is, at present, so hazy as to be almost imperceptible, can it be seriously asserted that 'particles' according to relativity have any resemblance to the particles encountered in nuclear physics ? In fact, is there any serious justification for extrapolating relativity to the domain of microphysics ? Many physicists believe that there is no such justification, and that relativity can be applied at the level of microphysics only after it has been integrated with quantum mechanics. However, it must be emphasised that these beliefs are not born out of any theoretical necessity or empirical compulsion. Moreover, the theory of relativity, as it exists today, is valid without reference to scale and is supported by firmer evidence than mere dimensional considerations. Finally, if we were to accept the interpretation of quantum mechanics, proposed in the next two chapters, as a theory of extended particles - since it is claimed that this interpretation is forced by the existing formalism - then there is every reason to construct models of elementary particles based purely on relativistic considerations.

In the last chapter a necessary condition for the existence of particles in relativity, without the introduction of new phenomenology, was pointed out. According to this necessary condition, the components of the metric tensor cannot be differentiable everywhere. Here it is pointed out that the above condition is also sufficient for the existence of a realistic model of elementary particles.

2.2 Specification of the hypersurface

We will assume that the shell, or surface layer, is quasi-isolated and spherically symmetric. Both uncharged and charged surface layers will be considered together. The external metric will be taken to be of the Schwarzschild type, in the case of uncharged surface layers, or of the Reissner-Nordstrom type, in the case of charged surface layers. It will be understood that the curvature of the shell is larger than the critical curvature (Schwarzschild radius) so that the external metric has no singularity. The interior metric will be assumed to be any metric that can be transformed to the Minkowski type by means of some coordinate transformations. The components of the metric tensor will be taken to be continuous across the shell, with essential discontinuities in some of their first derivatives.

We will use curvature coordinates (t, r, θ, ϕ), as opposed to the radiative coordinates (u, r, θ, ϕ) used earlier, in Chapter I. In the case of charged surface layers, no particular simplification is obtained by the use of radiative coordinates. We will suppose that the equation of the hyper-surface, Σ, corresponding to the shell-like distribution of (charged or uncharged) matter, is given by

$$r = f(t) . \qquad (2.2.1)$$

Transforming to the coordinates (t, z, θ, ϕ) by means of the transformation $r \longrightarrow z = r - f(t)$, the equation of Σ is simply $z = 0$.

The external metric is given by

$$ds_+^2 = g^{-1}(g^2 - f'^2)dt^2 - 2f' g^{-1} dz dt - g^{-1} dz^2 - (z+f)^2 d\Omega^2 \qquad (2.2.2)$$

where $g^{-1} = \frac{1}{g}$, $g = g(z,t) = \left(1 - \frac{2m}{(z+f)} + \frac{k}{(z+f)^2}\right)$ (2.2.3)

k being zero in the uncharged case. The internal metric is given by

$$dS_-^2 = (1 - F'^2)\, dT^2 - 2F'\, dZ\, dT - dZ^2 - (Z+F)^2 d\Omega^2 , \qquad (2.2.4)$$

where $(T, Z, \theta, \emptyset)$ are obtained from $(t, z, \theta, \emptyset)$ by means of the transformations*

$$T = \alpha\,(t, z)$$
$$Z = \beta\,(t, z) \qquad (2.2.5)$$

Here, $R = F(T)$ is the equation of the hypersurface in the interior curvature coordinates $(T, R, \theta, \emptyset)$, in which the metric is Minkowskian. As before, $Z = R - F(T)$.

We suppose that the components of the metric tensor are continuous across the hypersurface. That is

$$[g_{\mu\nu}] = 0 \qquad (2.2.6)$$

In writing down (2.2.6) we are using the earlier notation given in section 1.4.2.

---

*In Papapetrou and Hamoui (1968) and Chapter I the transformation functions α, β in (2.2.5) were restricted to be functions of the single null coordinate.

2.3 The equations of motion

Writing $g_{\mu\nu}$ as $\overset{+}{g}_{\mu\nu} \chi^{+} + g_{\mu\nu} \chi^{-}$, and using the properties derived earlier / we have, from the usual expressions for the Christoffel symbols and the Ricci tensor,

$$\Gamma^{\mu}_{\nu\sigma} = \Gamma^{\mu}_{\nu\sigma} \chi^{+} + \Gamma^{\mu}_{\nu\sigma} \chi^{-} ,$$

$$R_{\mu\nu} = \overset{+}{R}_{\mu\nu} \chi^{+} + \overset{-}{R}_{\mu\nu} \chi^{-} + S_{\mu\nu} , \qquad (2.3.1)$$

where $S_{\mu\nu} = [\Gamma^{\alpha}_{\mu\alpha}] \chi_{,\nu} - [\Gamma^{1}_{\mu\nu}] \delta(z)$ (2.3.2)

from (1.2.13). Since the interior and exterior fields have been chosen to be free, the equations of motion are, as before,

$$t^{\mu\nu}{}_{:\nu} = 0 \qquad (2.3.3)$$

where $t^{\mu\nu} = S^{\mu\nu} - \frac{1}{2} g^{\mu\nu} S$ (2.3.4)

$S^{\mu\nu}$ and S being defined as usual.

To simplify the derivation of the equations of motion, the internal and external metric tensors will be written in the form

$$ds^2_{\pm} = a^{\pm} dt^2 + e^{\pm} dtdz + c^{\pm} dz^2 + b^{\pm} d\Omega^2 \qquad (2.3.5)$$

where

$$a^+ = g^{-1}(g^2 - f'^2) \qquad a^- = (1-F'^2)\alpha_t^2 - \beta_t^2 - 2F'\alpha_t\beta_t$$

$$b^+ = -(z+f)^2 \qquad b^- = -(\beta+F)^2$$

$$c^+ = -g^{-1} \qquad c^- = (1-F'^2)\alpha_z^2 - \beta_z^2 - 2F'\alpha_z\beta_z$$

$$e^+ = -2f'g^{-1} \qquad e^- = 2\left[(1-F'^2)\alpha_t\alpha_z - \beta_t\beta_z - F'(\alpha_t\beta_z + \alpha_z\beta_t)\right]$$

(2.3.6)

α_t, α_z, β_t, etc. denoting the partial derivatives of the functions α and β.

With this notation, the Christoffel symbols are given by

$$\Gamma_{00}^{0^\pm} = g_{1\pm}^{-1}(e^\pm e_0^\pm - e^\pm a_1^\pm - 2c^\pm a_0^\pm) \qquad \Gamma_{00}^{1^\pm} = g_{1\pm}^{-1}(2a^\pm a_1^\pm + e^\pm a_0^\pm - 2a^\pm e_0^\pm)$$

$$\Gamma_{01}^{0^\pm} = g_{1\pm}^{-1}(e^\pm c_0^\pm - 2c^\pm a_1^\pm) \qquad \Gamma_{01}^{1^\pm} = g_{1\pm}^{-1}(a_1^\pm e^\pm - 2a^\pm c_0^\pm)$$

$$\Gamma_{11}^{0^\pm} = g_{1\pm}^{-1}(2c^\pm c_0^\pm + e^\pm c_1^\pm - 2c^\pm e_1^\pm) \qquad \Gamma_{11}^{1^\pm} = g_{1\pm}^{-1}(e^\pm e_1^\pm - e^\pm c_0^\pm - 2a^\pm c_1^\pm)$$

$$\Gamma_{22}^{0^\pm} = g_{1\pm}^{-1}(2c^\pm b_0^\pm - e^\pm b_1^\pm) \qquad \Gamma_{22}^{1^\pm} = g_{1\pm}^{-1}(2a^\pm b_1^\pm - c^\pm b_0^\pm)$$

$$\Gamma_{33}^{0^\pm} = \Gamma_{22}^{0^\pm} \text{Sin}^2\theta \qquad \Gamma_{33}^{1^\pm} = \Gamma_{22}^{1^\pm} \text{Sin}^2\theta$$

$$\Gamma_{02}^{2^\pm} = \frac{1}{2} b_\pm^{-1} b_0^\pm \qquad \Gamma_{03}^{3^\pm} = \frac{1}{2} b_\pm^{-1} b_0^\pm$$

$$\Gamma^{2\pm}_{12} = \frac{1}{2}\, b_{\pm}^{-1} b_1^{\pm} \qquad \Gamma^{3\pm}_{13} = \frac{1}{2}\, b_{\pm}^{-1} b_1^{\pm}$$

$$\Gamma^{2\pm}_{33} = -\,\mathrm{Sin}\ \theta\ \mathrm{Cos}\ \theta \qquad \Gamma^{3\pm}_{23} = \mathrm{Cot}\ \theta \tag{2.3.7}$$

where

$$g_1^{\pm} = e_{\pm}^2 - 4\, a^{\pm}\, c^{\pm} \tag{2.3.8}$$

and $a_{,1}$ is denoted by a_1, etc.

The components of the distribution part of the Ricci tensor are given by

$$S_{00} = -\,[\,\Gamma^1_{00}\,]\delta = -\frac{1}{2}\, a\,|\,[\,a_1\,]\delta$$

$$S_{01} = [\,2\Gamma^2_{02} + \Gamma^0_{00}\,]\delta = -\frac{1}{4}\, e\,|\,[\,a_1\,]\delta$$

$$S_{10} = -\,[\,\Gamma^1_{10}\,]\delta = -\frac{1}{4}\, e\,|\,[\,a_1\,]\delta$$

$$S_{11} = [\,2\Gamma^2_{12} + \Gamma^0_{10}\,]\delta = \frac{1}{2}\,[\,2b^{-1}b_1 - c a_1\,]\delta$$

$$S_{22} = -\,[\,\Gamma^1_{22}\,]\delta = -\frac{1}{2}\, a\,|\,[\,b_1\,]\delta$$

$$S_{33} = -\,[\,\Gamma^1_{33}\,]\delta = S_{22}\ \mathrm{Sin}^2\ \theta \tag{2.3.9}$$

In arriving at (2.3.9) we have already used the junction conditions (2.2.6) to conclude that $[a_o] = [b_o] = [c_o] = [e_o] = 0$. Thus, $[a]$ is a function of t that is identically zero, hence, $[a_o] = [a]_{,o} = 0$.

The Ricci tensor with raised suffices is given by

$$S^{00} = \frac{1}{4} [2ca_1 + e^2 b^{-1} b_1] \delta$$

$$S^{01} = -\frac{1}{4} [ea_1 + 2aeb^{-1} b_1] \delta$$

$$S^{11} = \frac{1}{4} [2aa_1 + 4a^2 b^{-1} b_1] \delta$$

$$S^{22} = -\frac{1}{2} ab^{-2} | [b_1] \delta$$

$$S^{33} = S^{22} / \mathrm{Sin}^2 \theta \qquad (2.3.10)$$

and its trace by

$$S = -\frac{1}{2} [a_1 + 2ab^{-1} b_1] \delta \qquad (2.3.11)$$

The material energy tensor is given by

$$t^{00} = b^{-1} | [b_1] \delta$$

$$t^{01} = t^{10} = t^{11} = 0$$

$$t^{22} = \frac{1}{2} b^{-1} | [a_1 + ab_1] \delta$$

$$t^{33} = t^{22} / \mathrm{Sin}^2 \theta \qquad (2.3.12)$$

Two of the equations of motion are identically satisfied and the remaining two are given by

$$t^{00}_{,0} + (\Gamma^{0}_{00} + \Gamma^{0}_{00}) \mid t^{00} + 2\Gamma^{0}_{22} \mid t^{22} = 0$$

$$\Gamma^{1}_{00} \mid t^{00} + 2\Gamma^{1}_{22} \mid t^{22} = 0 \qquad (2.3.13)$$

These equations have to be satisfied along with the junction conditions

$$[a] = [b] = [c] = [e] = 0 \qquad (2.3.14)$$

Using α_0, α_1 etc. to denote the boundary values of $\alpha_t(t, z)$, $\alpha_z\,(t, z)$ at $(t, 0)$ and letting

$$A = 1 - \frac{2m}{f} + \frac{k}{f^2}$$

$$B = \frac{1}{\alpha_o^2}\,(\alpha_o^2 - f^2) \qquad (2.3.15)$$

$$D = \frac{1}{\alpha_o^2}\,(f^{''}\alpha_o - f'\alpha_o')$$

the values of the various bars and brackets, involved in (2.3.13), are given below. We observe that <u>the junction condition $[b] = 0$ has been used to eliminate the extra variable F</u>.

$$[a] = A^{-1}(A^2 - f'^2) - \alpha_o^2\,B$$

$$[b] = f^2 - F^2 = 0 \quad \text{(assumed throughout)}$$

$$[c] = -A^{-1} - B\alpha_1^2 - \beta_1\alpha_o^{-1}(\beta_1\alpha_o + 2f'\alpha_1)$$

$$[e] = -2f'A^{-1} - 2B\alpha_o\alpha_1 + 2f'\beta_1 \qquad (2.3.16)$$

$$[a_o] = [b_o] = [c_o] = [e_o] = 0 \qquad (2.3.17)$$

$$[a_1] = A'A^{-2}(A^2 + f'^2) + 2f'\alpha_1 D - 2\alpha_o\alpha_1' B + 2f'\beta_1'$$

$$[b_1] = 2f(\beta_1 + f'\alpha_1\alpha_o^{-1} - 1)$$

$$[c_1] = A'A^{-2} + 2\alpha_1^3\alpha_o^{-2} f'D - 2\alpha_1\alpha_{11}B + 2\beta_1\beta_{11}$$
$$+ 2\alpha_1^2\alpha_o^{-1}\beta_1 D + 2f'\alpha_o^{-1}(\alpha_{11}\beta_1 + \alpha_1\beta_{11})$$

$$[e_1] = 2f'A'A^{-2} + 4f'\alpha_1^2\alpha_o^{-1}D - 2D(\alpha_1'\alpha_1$$
$$+ \alpha_1\beta_{11}) + 2\beta_1'\beta_1 + 2\alpha_1\beta_1 D$$
$$+ 2f'\alpha_o^{-1}(\alpha_1'\beta_1 + \alpha_o\beta_{11} + \alpha_1\beta_1') \qquad (2.3.18)$$

$$a| = \alpha_o^2 B$$

$$b| = -f^2$$

$$c| = -A^{-1}$$

$$e| = -2f'A^{-1} \qquad (2.3.19)$$

$$a_o| = \frac{1}{2} f' A' A^{-2}(A^2+f'^2) - f' f'' A^{-1} - f' D\alpha_o + B\alpha_o \alpha'_o$$

$$b_o| = - 2ff'$$

$$c_o| = \frac{1}{2} f' A' A^{-2} - f' \alpha_1^2 \alpha_o^{-1} D + \alpha_1 \alpha'_1 B - \beta_1 \beta'_1 - \alpha_1 \beta_1 D$$

$$- f' \alpha_o^{-1} (\alpha'_1 \beta_1 + \alpha_1 \beta'_1)$$

$$e_o| = f'' A^{-1} + f'^2 A' A^{-2} - f' \alpha_1 D + B(\alpha'_o \alpha_1 + \alpha_o \alpha'_1)$$

$$- \alpha_o \beta_1 D - f' \alpha_o^{-1} (\alpha'_o \beta_1 + \alpha_o \beta'_1) \qquad (2.3.20)$$

$$a_1| = \frac{1}{2} A' A^{-2}(A^2 + f'^2) - f' \alpha_1 D + \alpha_o \alpha'_1 B - f' \beta'_1$$

$$b_1| = - f(1 + \beta_1 + f' \alpha_1 \alpha_o^{-1})$$

$$c_1| = \frac{1}{2} A' A^{-2} - f' \alpha_o^{-2} \alpha_1^3 D + \alpha_1 \alpha_{11} \alpha_o^{-2} B$$

$$- \beta_1 \beta_{11} - \alpha_1^2 \beta_1 \alpha_o^{-1} D$$

$$- f' \alpha_o^{-1} (\alpha_{11} \beta_1 + \alpha_1 \beta_{11}) \qquad (2.3.21)$$

It follows that we have 5 equations involving 6 variables $f, \alpha_o, \alpha_1, \beta_1, \alpha_{11}, \beta_{11}$, since the equation $[b] = 0$ has been used to eliminate the extra variable F.

2.4 Discussion

It is apparent that the above equations are underdetermined, and that the presence or absence of charge does not, in any way, affect the underdetermined nature of the equations. The physical interpretation of this indeterminacy is simple. In the macroscopic case, the matter in the shell would consist of particles, and the motion of the shell becomes determinate only after the interaction between these particles has been specified. That is, the equations are underdetermined because of the absence of an equation of state. Once an equation of state has been specified the motion of the shell becomes completely determinate since only one more equation is required.

In the microscopic case the situation becomes a little more difficult. This happens because the concept of an equation of state is essentially a macroscopic concept deriving its justification from the kinetic theory of fluids. This theory, however, is based on the assumption that matter consists of particles interacting amongst each other in a certain manner. Therefore, the very mention of an equation of state would be based on the assumption that the matter in the shell (elementary particle) can be further subdivided into other 'particles' which interact amongst themselves like gas molecules. Neither do we have any justification for such an assumption nor does this process of subdividing elementary particles have a logical conclusion. It follows that, in the microscopic case, we must look for some alternative procedure to make the motion of shell determinate.

Before pointing out this alternative procedure, we re-examine the old argument concerning the effect of the relative strengths of electromagnetic and gravitational forces on the stability of electrons, for example. The point is that this argument is based on the assumption of a particular equation of state for the material in the shell. As seen above, this implicitly assumes a subdivision of elementary particles into smaller particles interacting amongst themselves by means of gravitational and electromagnetic forces. Discarding the concept of an equation of state, rather than introducing new phenomenology, seems, therefore, the best method of countering this argument.

The correct procedure in the microscopic case is, simply, to prescribe an equation of motion. In the macroscopic case, this is equivalent to prescribing an equation of state. But, in the microscopic case this equation of motion must be independently verified without reference to an equation of state.

Thus, according to this theory, matter in its ultimate form is closely related to the occurrence of essential discontinuities in the first derivatives of the field functions. In general, the occurrence of essential discontinuities in the $g_{\mu\nu}$ is only a necessary condition, but for the shell discussed above this condition is also sufficient. We reiterate that by prescribing an equation of motion we have not introduced any new forces. Instead we have rejected a particular extrapolation from the macroscopic to the microscopic level.

Next, it is necessary to consider the problem of ascertaining the validity of a particular equation of motion that may be prescribed. The procedure we propose is necessarily a little indirect. Newtonian dynamics can be expected to give an accurate description of the motion of the centre of mass of the particle insofar as this motion remains unaffected by the motion of the shell in the centre-of-mass frame. But, in general, the motion of the shell, in the centre-of-mass frame, is bound to modify the motion of the centre of mass. The form of this modification may not be immediately apparent. This will become clear in Chapter IV where we show that a description of the dynamics of such particles leads to many of the main features of quantum mechanics.

One last point that needs mention. We have assumed that relativity theory is valid in the interior of the electron. Is such an assumption justified ? Thus, Dirac (1938, 1962b) has asserted that the interior of the electron does not lie in the domain of physics and one can expect a breakdown of physical laws there. Similarly, Yukawa (1964, and the references cited therein) has suggested a breakdown of Lorentz invariance in the 'interior' (region of extension) of an extended particle. Our results indicate that this issue is not significant, as long as we assume that the particle interacts at its boundary with the external world. Whether or not we assume the continued validity of physical laws in the interior of the electron, we cannot say anything about the motion of its boundary. A definite statement concerning the motion of the boundary can only be made in the wake of some new empirical or theoretical development.

Actually, the breakdown of Lorentz invariance in the interior of the electron can affect the external behaviour of the electron only in the sense that an external stimulus applied at one 'point' on the boundary instantaneously affects the motion of the entire boundary. But, this can also be expected to happen because of the infinite 'mass density' at the boundary. Naturally, this discussion of microscopic phenomena using macroscopic notions is bound to be somewhat misleading. Nevertheless, breakdown of Lorentz invariance in the above sense would imply the existence of advanced radiation that can be detected empirically (Appendix B and enclosure).

2.5 Conclusions

It is possible to construct shell-like models of extended charged particles, without significantly modifying relativity at the micro-level, and without introducing new phenomenology. Shell-like structures with a δ-function mass density can evolve in an essentially arbitrary manner.

CHAPTER III

INTERPRETATION OF THE INDETERMINACY RELATIONS

3.1 Introduction :

Heisenberg (1927) initially formulated the indeterminacy relations in the form

$$\delta_q \, \delta_p \geq \hbar, \tag{3.1.1}$$

where δ_q and δ_p denote the 'uncertainities', or standard deviations, in the (i th) position coordinate q and the canonically conjugate momentum p (of an electron, say) and $\hbar$ is the Planck constant divided by 2π.

A number of interpretations of the indeterminacy relations (3.1.1) have been proposed (see, for instance, Jammer, 1974). Motivated by Heisenberg's thought experiment and the above way of stating the indeterminacy relations, the conventional interpretation asserts that it is impossible, in principle, to specify the simultaneous values of canonically conjugate variables to an arbitrary degree of accuracy. Alternatively, the so-called statistical interpretation asserts that the product of the standard deviation of two canonically conjugate variables has a lower bound. In this paper, it is pointed out that these assertions are misleading, if they are considered as interpretations of the precise form of the indeterminacy relations. The correct interpretation necessarily leads to the conclusion that the particles described by quantum mechanics have some finite extension in any state. A somewhat similar, but less general, approach was developed by Yukawa (1950a, 1950b) based on Born's reciprocity relations.

3.2 The formal statement

Within the classical axiomatic framework for quantum mechanics (von Neuman, 1955), the indeterminacy relations take on the more precise form

$$\mathrm{Var}\ (\hat{q} \mid \Psi)\ \ \mathrm{Var}\ (\hat{p} \mid \Psi) \geq \hbar^2 / 4,$$

for any state Ψ, (3.2.1)

where $\hat{q}$ and $\hat{p}$ are hypermaximal operators, corresponding to position and momentum, on a Hilbert space H. The state, Ψ, is an element of H common to the domains of $\hat{q}$ and $\hat{p}$, and Var $(. \mid \Psi)$ denotes the variance of an observable in the state Ψ. The above inequality was first established by Robertson (1929), and it was pointed out by Ditchburn (1930) that the apparent discrepancy in the right hand side is due to the assumption, by Heisenberg, of a gaussian distribution for position.

3.3 Qualitative differences

The left hand side of (3.2.1), however, introduces some qualitatively new features, which, though seemingly inconsequential, lead to a real discrepancy. The first new feature lies in the concept of position : the form (3.1.1) suggests that the position is a classical random variable, whereas the 'position' appearing in (3.2.1) is an operator. Following Wigner's (1932) observation, it is known that the quantum mechanical position cannot be a classical random variable if linearity is to be retained.

The second new feature is the explicit appearance of the state in (3.2.1). Although Heisenberg (1927) used the analogous concept of the wavefunction, (3.1.1) does not explicitly include the state. On the other hand, in the axiomatic formulation, it is meaningless to speak of expectations and variances, without first specifying the state.

The third, and most important, new feature is that the variances in (3.2.1) depend only on the observable and the state; because, by definition

$$\mathrm{Var}\,(\hat{q} \mid \Psi) = \langle \hat{q}^2 \Psi, \Psi \rangle - \langle \hat{q} \Psi, \Psi \rangle^2, \qquad (3.3.1)$$

$\langle \cdot, \cdot \rangle$ denoting inner product in H. Thus, once an operator $\hat{q}$, or a unitary equivalent, has been assigned to the position observable, the variance in position, in a fixed state, is a fixed quantity governed by equation (3.3.1). In particular, the variance in position, in a given state, remains fixed, regardless of any intention to measure the momentum, simultaneously, or otherwise. Finally, in contrast to (3.1.1), (3.2.1) holds even if the position and momentum are measured at different times, provided the particle is in the same state on both occasions, i.e., if $\hat{q}$ is measured at time t and $\hat{p}$ is measured at time t', the (3.2.1) holds, provided $\Psi(t) = \Psi(t')$. The last equation is trivially true if $t = t'$, but $\Psi(t) = \Psi(t')$ does not, in general, imply that $t = t'$. Hence, simultaneity, which was explicit in (3.1.1), is not logically necessary for (3.2.1).

3.4 Falsity of the conventional interpretation

Due to the presence of these new features, a part of the usual interpretations of the indeterminacy relations stands falsified. To see this, with regard to the conventional interpretation, we observe that there are two possibilities :

(a) The term 'simultaneous' is inessential, and it is the case that it is impossible, in principle, to specify the values of an observable, to an arbitrary degree of accuracy.

(b) The term 'simultaneous' is essential. Hence, if either of the canonically conjugate variables is measured separately, it can be measured, in principle, to an arbitrary degree of accuracy. This point of view is based on Heisenberg's thought experiment, which suggests that δ_q and δ_p are dependent on the design of the measurement process, hence, ultimately, on the intentions of the observer - if the observer chooses to measure the position of the electron more accurately, he can do so, without qualitatively altering the electron, provided he gives up all hope of measuring its momentum.

The conventional viewpoint has generally been thought to be represented by viewpoint (b). However, as equation (3.3.1) shows, in axiomatic quantum mechanics, δ_q depends only on the state (or the instant at which the measurement is carried out) and is independent of the intentions of the observer. It follows that δ_q cannot be altered, leave alone made arbitrarily small. Hence, the implication of viewpoint (b), that δ_q can, in principle, be made arbitrarily small, is false. Consequently, viewpoint (b) itself must be false. Similarly, the 'statistical interpretation' which seems to be no more than a verbal restatement of (3.2.1) needs to be supplemented with the statement (3.3.1). Otherwise, there is the danger of committing the customary fallacy that only the product, and not the variances individually, must be bounded below.

The above reasoning has demonstrated that the usual interpretations of the indeterminacy relations are fallacious within the framework of axiomatic quantum theory. It is possible that this incompatibility arises from an unfortunate choice of the axiomatic framework. However, with a historical perspective, it is clear that the axiomatic formalism has contributed a great deal to the success of quantum mechanics in practice, whereas the indeterminacy relations, in the form (3.1.1), have usually been reserved for philosophical disputes.

An additional reason for believing in the falsity of the usual interpretation is that it makes certain unconscious assumptions. Thus, when it is suggested that δ_q can be made arbitrarily small (by letting δ_p become large), it is assumed, a priori, that the particle is a point mass, and that it is meaningful to speak of its exact coordinates. Such an assumption is usually justified in classical physics because of its great simplifying value, and because, in classical physics, it can be demonstrated that this simplifying assumption is frequently of no consequence. However, in quantum mechanics the assumption can at best be justified, a posteriori, as a consequence of the formalism. Such justification is unlikely to be forthcoming within the existing formalism, in view of the discussion above.

3.5 Extended particles

While keeping open the possibility of a drastic revision of the axiom scheme for quantum mechanics, here we will consider only the possibility of providing an alternative interpretation, more suited to the existing formalism. To begin with, it might be thought desirable, from the operational point of view, to relate equation (3.3.1) to the outcome of some measurement process. In that case, it is clear that equation (3.3.1) implicitly defines, and refers to the outcome of, an ideal or optimal measurement process. The last assertion is valid because, by deliberately introducing errors, the observed spread in position, in a given state, can be made larger than that prescribed by (3.3.1). In any case, as a consequence of (3.2.1) and (3.3.1),

$$\mathrm{Var}\,(\hat{q}\,|\,\Psi) \geq K\,(\Psi) = K > 0, \qquad (3.5.1)$$

where K is independent of the process used for measuring the variance.

For a single particle, of non-zero rest mass, the inequality (3.5.1) necessarily implies that the particle has some finite (as opposed to infinitesimal) spatial extension in the state Ψ. For, suppose this is not the case, i.e., suppose the mass of the particle is concentrated at a point. Then, by (3.5.1), no process of measurement will ever serve to reveal the exact coordinates of this point. Operationally this is absurd, because the essence of operationalism lies in the assertion that there is no dichotomy between physical reality and the results of measurement processes. Even otherwise, it would not be possible to assign any specific coordinates to this point (at which the mass has been supposed to be concentrated), because the validity of (3.5.1) would, then, have to be accounted for in terms of secondary (hidden) causes, or hidden variables. Without going into the current status of hidden variable theories (reviewed in Clauser and Shimony, 1978), it should be pointed out that this assertion would lead to a denial of (3.5.1) under some conditions, and hence to an extension or rejection of the existing formalism. As such, the assertion, of point masses, would be a statement of a belief about empirical facts, and would not constitute an interpretation of the existing formalism, which it contradicts.

To sum up, if the existing formalism is accepted, then it is operationally meaningless, and theoretically inconsistent, to assert that the mass of the particle is concentrated at a point. Since the mass of the particle is assumed to exist, and since it cannot be said to exist at a point, it must necessarily be distributed in some manner, i.e., the particle has some spatial extension depending on the state Ψ.

3.6 Indefinite localisation

In considering the manner in which the extension depends on the state, we observe that it is still possible to have an analogue to the process of letting δ_q tend to zero. The analogous procedure consists of measuring the position of the system (particle) in states Ψ_n with

$$\text{Var}\,(\hat{q} \mid \Psi_n) \leq 1/n,\; n = 1,2,3,\ldots \qquad (3.6.1)$$

However, in admitting the possibility of states with arbitrarily small dispersion, we also have to admit

$$\text{Var}\,(\hat{p} \mid \Psi_n) \geq n\hbar^2 / 4,\; n = 1,2,3,\ldots \qquad (3.6.2)$$

Now, because $\hat{p}$ is self adjoint,

$$\text{Var}\,(\hat{p} \mid \Psi_n) = \int_{\sigma(\hat{p})} x^2 \langle E_{\hat{p}}(dx)\Psi_n, \Psi_n\rangle - \left[\int_{\sigma(\hat{p})} x \langle E_{\hat{p}}(dx)\Psi_n, \Psi_n\rangle\right]^2$$

(3.6.3)

where $E_{\hat{p}}(dx)$ is the spectral measure induced by $\hat{p}$. Since, for any state Ψ, $(E_{\hat{p}}(dx)\,\Psi, \Psi)$ is a probability measure, $\text{Var}\,(\hat{p} \mid \Psi_n)$ can satisfy inequalities of the form (3.6.2) only if $\sigma(\hat{p})$ is unbounded. The implication of this fact is that inequalities of the form (3.6.1) are possible only if the momentum is allowed to assume unbounded values. In the usual interpretation, it is only the 'measurement error' in momentum, and not the momentum, per se, which has to assume unbounded values in the process of letting δ_q tend to zero.

This observation is not unique to the axiomatic formalism - it is also true if p is regarded as a classical random variable. Even in the framework of de Broglie waves with wavelength $\lambda = h/p$, $1/2\,\lambda$, regarded as a measure of the non-localisability of the particle, can be made arbitrarily small only by letting the momentum become arbitrarily large.

Now, the possibility, of the momentum becoming arbitrarily large, is customarily admitted in quantum mechanics, although it is possible to choose one of the operators, corresponding to position and momentum, as bounded. The customary choice has the unfortunate consequence that the kinetic energy, or the free-particle hamiltonian, becomes unbounded. One of the arguments against such a possibility is that the energy of a 'free' particle can increase only at the expense of the energy of that part of the universe which is causally connected with the particle. If the energy of this part is assumed to be conserved and finite, the energy of the particle can increase indefinitely only if the energy of the rest of the universe can be allowed to decrease indefinitely, i.e., the energy of the rest of the universe must be allowed to be unbounded below. Alternatively, the potential energy of the particle must be allowed to be unbounded below, this potential being created by other particles in the universe. Hence, a beam of free electrons should behave like a beam of electrons in an infinite potential well!

Thus, apart from the obvious risks, faced by a theory which claims to be fully operational, there are certain theoretical difficulties in admitting the possibility of indefinite localisation, in principle. In short, the conclusion, that particles (having non-zero rest mass) described by quantum mechanics must necessarily have some finite spatial extension, is independent of the state unless the possibility of producing an indefinite amount of energy is explicitly admitted.

Although this conclusion is interesting in itself, it is made even more interesting by the fact that similar conclusions can be drawn using the theory of relativity. Thus, according to relativity, a mass point, of mass m, will be surrounded by the Schwarzschild singularity of radius $r = 2Gm/c^2$. Hence, a mass point, in its interactions with the external world, should behave like a black hole. Since real particles cannot be said to exhibit such behaviour, according to relativity, real particles are not mass points.

Apparently, the only known way of preventing an extended mass distribution from collapsing, without introducing altogether new phenomenology, is to attribute a shell-like structure to the distribution, as was done in the earlier chapter. The components of the metric tensor, or, at least, their first derivatives, must have a discontinuity at the surface of the shell.

Once such a shell model for extended particles has been constructed, it is natural to think of interpreting quantum mechanics in the context of such extended particles. This apparently naive expectation is partially borne out in the next chapter which proposes to interpret quantum mechanics as just the approximate classical statistical dynamics of such extended particles.

3.7 Conclusions

The usual interpretation of the indeterminacy relations is fallacious within axiomatic quantum mechanics. Within the axiomatic formalism, the indeterminacy relations necessarily lead to the conclusion that the particles (having some non-zero rest mass) described by quantum mechanics have some spatial extension, in any state. In no state can such particles be indefinitely localised, without explicitly admitting the possibility of producing an infinite amount of energy.

CHAPTER IV

INTERPRETATION OF QUANTUM MECHANICS AS A THEORY OF EXTENDED PARTICLES

4.1 Introduction

The problem of interpreting quantum mechanics is well known, and a review of the better known interpretations can be found in Jammer (1974). More recently, a large class of (local) hidden variable theories have apparently been falsified (Clauser and Shimony, 1978), although the relevance of Bell's inequalities has been questioned - for instance, Lochak (1977).

Here, we propose to adopt an altogether new approach. The logical basis of this approach is the assertion (Raju, 1980d) that an interpretation of the precise form of the indeterminacy relation necessarily leads to the following conclusions :

(i) The usual interpretation of the indeterminacy relations is fallacious within the axiomatic framework for quantum mechanics.

(ii) The particles (of non-zero rest mass) described by quantum mechanics cannot be localised, and, hence, must correspond to extended mass distributions.

It is, therefore, natural to think that the peculiar pattern of similiarities and differences between classical and quantum mechanics arises because of the extended nature of real particles. In fact, by virtue of (ii), there is a definite theoretical necessity for interpreting quantum mechanics as a theory of extended particles. In this paper, it is pointed out that many significant concepts of quantum mechanics have a natural counterpart in the context of a semiclassical description of the dynamics of extended particles. Looking at it in another way, the ideas presented below also have a direct bearing on the problem of describing the dynamics of extended particles in a manner that is Lorentz covariant and compatible with quantum mechanics.

4.2 The model for extended particles

The further development of this theory requires a model for extended particles. Apart from spherical symmetry in the rest frame, the main restrictions to be imposed on such a model would be the following :

(A) The particle interacts at its boundary with the external world.

(B) The particle pulsates uniformly in the rest frame.

Stated more mathematically, restriction (A) asserts that the mass of the particle is distributed (in the rest frame) over a spherically-symmetric hypersurface, i.e., the particle is shell-like. From the point of view of relativity, such shell-like models seem to be necessary to overcome the classical phenomenological argument concerning the imbalance between gravitational and electromagnetic forces. A similar assumption has also been used in time-symmetric electrodynamics by Dirac (1938) and Raju (1980a), although the empirical consequences cannot be said to have been conclusively verified.

Such shell-like models have been constructed by Dirac (1962b), for instance, by introducing new phenomenology to offset the inordinate imbalance between gravitational and electromagnetic forces. On the other hand, as seen earlier such shell-like models can be obtained by suitably altering the usual junction conditions in relativity. In fact, the results indicate that charged surface layers can exist and that oscillating solutions are possible. Although oscillating solutions also occur in Dirac's (1962) model, the advantage of this approach is that no new phenomenology is introduced.

Thus, we have a picture of an extended particle as an oscillating surface layer, obtained with or without additional phenomenology. With this picture we proceed with the interpretation of quantum mechanics.

4.2.1 Planck's constant

The first step is to introduce an analogue of Planck's constant into the theory, and connect it with the particular model, of extended particles, under consideration. This is done by defining a constant h_o by

$$E_o = h_o \upsilon_o , \qquad (4.2.1)$$

where E_o is the energy, and υ_o is the frequency of oscillation, measured in the rest frame. However, if (4.2.1) is to agree with the usual quantum mechanical relationship, some more restrictions are necessary on the model of extended particles :

(C) The oscillations of the particle are linear.

(D) The frequency of oscillation, υ_o, is proportional to the proper mass.

(E) The constant of proportionality is the Planck's constant (with $c = 1$).

As the results in Chapter II show, these restrictions are not unreasonable for a realistic model of extended particles. With Dirac's (1962) model, in the linear approximation, restrictions (A) - (D) are satisfied though (E) is not. The consequences of any possible nonlinearity are evaluated in Section 7.

Thus, in the context of extended particles, the wave-particle duality, implicit in (4.2.1), is interpreted as arising from the pulsations of the particle. Incidentally, we observe that the frequency of oscillation also leads to a Lorentz covariant description of the energy of the centre-of-mass of the particle.

4.3 The wavefunction

The next step is to introduce statistical considerations into the theory. This does not require any further restrictions because a particle in the real universe is never isolated. The external field, therefore, is at best statistically determined. For instance, the brownian motion of the stars (Chandrashekhar, 1943) would induce fluctuations in the external graviatational field. Similiarly, the random motion of nearby charged particles may be expected to produce small fluctuations in the external metric.

As a result of these fluctuations, the extension of the particle (i.e., the curvature of the spherical shell) and the phase of its oscillations, at any instant[2], are random variables. We can combine the two to obtain a single, complex valued random variable Ψ. $E\, 4\pi\, |\Psi|^2$ is, then, just the mean surface area of the particle ($|\Psi|$ being the extension), and this must, presumably, be finite. Therefore, $\Psi \varepsilon L^2 = L^2(\underline{\Omega}, \underline{B}, P)$, $(\underline{\Omega}, \underline{B}, P)$ being a standard borel probability space. Ψ would be taken to correspond to the quantum mechanical wavefunction.

---

[2] It is possible to speak of 'the curvature of the shell at any instant', under the usual restriction that the normal to the shell be time-like.

This assignment of a random variable to the state can be viewed classically in terms of incomplete information. But, because the entire cosmos is responsible for keeping this information incomplete, it is concievable that it is impossible, even in principle, to have complete information about the state.

4.4 The Schrodinger equation

Suppose the particle is in equilibrium with its surroundings, then the random process $\Psi(t)$ is stationary in the narrow sense, i.e., if $p(t_1, t_2, \ldots, t_k)$ represents the joint distribution of $\Psi(t_1)$, $\Psi(t_2), \ldots, \Psi(t_k)$ then

$$p(t_1 + s, t_2 + s, \ldots, t_k + s) = p(t_1, t_2, \ldots, t_k). \tag{4.4.1}$$

(4.4.1) merely states that in statistical equilibrium a change of the time-origin has no physical significance.

Because $\Psi \varepsilon L^2$, stationarity in the narrow sense implies stationarity in the wide sense. That is, $E\,\Psi(t)$ is independent of t, and the covariance $E\,\Psi(t)\,\overline{\Psi(s)}$ is a function only of the difference $t-s$. It follows (Rozanov, 1967) that the map $U(t)$, defined by

$$U(t)\ \Psi(s) = \Psi(t+s) \tag{4.4.2}$$

extends to a group of unitary operators on the subspace

$$H_\infty = \text{span}\left\{\Psi(s),\ s \in \mathbb{R}\right\}. \qquad (4.4.3)$$

$U(t)$ extends trivially to L^2, and, by Stone's theorem, there exists a densely defined self-adjoint operator H_1 on L^2, such that

$$U(t) = \exp(-\,i\,H_1 t) \qquad (4.4.4)$$

The operator H_1 satisfies the differential equation

$$\frac{\partial}{\partial t} U(t) = -\,i\,H_1\,U(t)\,, \qquad (4.4.5)$$

or, since $\Psi(t) = U(t)\,\Psi(0)$,

$$i\,\frac{\partial \Psi}{\partial t} = H_1 \Psi. \qquad (4.4.6)$$

Let $H = h_o\,H_1$, $\qquad (4.4.7)$

where h_o is defined by (4.2.1), then the spectrum of H just consists of the energy values of the particle, hence H is the Hamiltonian operator of quantum mechanics. H is trivially bounded below, and if h_o is indeed the Planck's constant then (4.4.7) is just the Schrodinger equation.

Thus, in the present theory, the Schrodinger equation appears as a consequence of (4.2.1) and some very general statistical laws.

4.5 Quantum probabilities and the operator representation

Wigner, in 1932, observed, and later proved (Wigner, 1972), that a joint distribution for position and momentum, consistent with linearity, does not exist in quantum mechanics. This observation has led to numerous attempts at formalising the notion of quantum probabilities, in the belief that they are significantly different from classical probabilities (in the sense of measure theory). On the other hand, one can very well adopt the point of view that Wigner's theorem asserts a failure of the operator representation, and that it is possible to formulate and work with quantum mechanics without introducing specifically 'quantum' probabilities. Further, at the very heart of the problem lies the insistence that the position and momentum be random variables. In the extended particle case, this insistence may simply not be justified.

Before considering the extended particle case, we first consider some peculiarities of the probabilities appearing in quantum mechanics. For definiteness, we consider the probabilities regarding the ith position coordinate $\hat{q}$ of a single particle.

(i) The state Ψ itself generates a probability, since $\| \Psi \|^2$ corresponds to the probability that $\hat{q}$ takes on some value.

(ii) Corresponding to each state Ψ we construct a random measure (i.e., a hilbert-space-valued measure) E ,

$$E_{\Psi} (A) = E(A)\Psi \qquad (4.5.1)$$

E being the spectral measure induced by the self-adjoint operator corresponding to $\hat{q}$.

(iii) Representing the state space as an L^2 space of random variables with zero means, we see that the probability $Q_{\Psi}(A)$, that $\hat{q}$ takes on some value in the region $A \subset \mathbb{R}$

$$Q_{\Psi} (A) = \frac{\| E(A) \Psi \|^2}{\| \Psi \|} \qquad (4.5.2)$$

is essentially a ratio of the variances of two complex valued random variables.

As Wiener (1958) has observed, a substantial part of the mystery of quantum mechanics lies in the fact that these peculiarities have never been satisfactorily explained[3]. This mystery can be resolved, at least partially, in the extended particle context. Firstly, it is quite meaningless to speak of the 'position' of the particle, since the particle simultaneously has several 'positions'. It might be a little more meaningful to speak of a portion of the particle lying in some region A. But, if the particle happens to be very small, say 10^{-30} cms across, a tremendous amount of energy would be required to resolve a portion of the particle. So, for all practical puposes, and for the range of energies for which quantum mechanics has been tested, it is quite meaningless to speak of a specific portion of the particle.

The next best thing one can do is to speak of the probability that the particle, on observation, would be found in some region A. Since observation involves an interaction with the external world, and since the particle interacts at its

---

[3]Mathematically, a single probability measure, P, can always be represented as the variance measure of the random measure induced by a stationary stochastic process with covariance function $\hat{P}$, where $\hat{P}$ is the Fourier transform of P.

surface, this probability would be proportional to the total surface area of the particle in A. But, for a given region A, the surface area of the particle actually in A is a random variable dependent on the state, and its mean value would be taken to represent the above probability.

To summarise, one can speak meaningfully only of the probability of finding the particle in a certain region, and the probability of finding the ith position coordinate in region A is given by

$$P_\Psi (A) = \frac{E\, X_\Psi (A, w)}{E\, X_\Psi (\mathbb{R}, w)} , \qquad (4.5.5)$$

where $X_\Psi (A, w)$ is the surface area of the sphere with centre 0 (say) and radius $|\Psi|$ that lies in the cylinder set over A in three dimensions. Thus, we see that $\|\Psi\|$ is indeed proportional to the probability of finding the particle somewhere, and, to evaluate the probability of finding the ith position coordinate in A, we do have to construct a random measure. Moreover, as is customary in quantum mechanics, only the probability measure, and not a specific random variable distributed according to it, can be assigned to the dynamical variable.

Peculiarity (iii) requires a closer study. We first evaluate the probability $P_\Psi(A)$. Since $E\,X_\Psi(A)$ is additive, it is sufficient to evaluate this probability for regions A of the form $(-\infty, s]$. This is done in Appendix C, and leads to the expression

$$\begin{aligned} E\,X_\Psi(s, w) &= 2\pi\sigma^2 e^{-s^2/\sigma^2} + \pi\sqrt{\pi}\,\sigma\, s\, \mathrm{Erfc}(-s/\sigma), \quad s < 0, \\ &= 4\pi\sigma^2 - 2\pi\sigma^2 e^{-s^2/\sigma^2} + \pi\sqrt{\pi}\,\sigma\, s\, \mathrm{Erfc}(s/\sigma), \quad s > 0, \end{aligned} \tag{4.5.4}$$

where $\mathrm{Erfc}(z)$ is the complementary error function. Using the asymptotic expansion

$$\mathrm{Erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-y^2} dy = \frac{e^{-z^2}}{\sqrt{\pi}\, z} \left\{ 1 - \frac{1}{2z^2} + \ldots \right\}, \tag{4.5.5}$$

(4.5.4) may be written in the form

$$\begin{aligned} E\,X_\Psi(s, w) &= \pi\sigma^2 e^{-s^2/\sigma^2} \left\{ 1 + \frac{\sigma^2}{2s^2} - \ldots \right\}, \quad s < 0, \\ &= 4\pi\sigma^2 - \pi\sigma^2 e^{-s^2/\sigma^2} \left\{ 1 + \frac{\sigma^2}{2s^2} - \ldots \right\}, \quad s > 0. \end{aligned} \tag{4.5.6}$$

In arriving at (4.5.4), it has been assumed that Ψ has a complex gaussian distribution with mean zero and variance σ^2. There is some virtue to the gaussian distribution; however, the choice of a different distribution would have only a slight effect on the empirical consequences of the theory. Also, the parameter σ entering into (4.5.4) is not arbitrary, and is restricted by physical considerations. σ would be, approximately, at most half the order of magnitude of the mean extension of the particle. With the choice of a complex gaussian distribution for Ψ, for example, σ would be of the same order of magnitude as the mean extension of the particle. In general, for a realistic model of, say, the electron, σ would be quite small.

We now claim that these probabilities do vary in an approximately quadratic manner with the state. More precisely, a map, $\Psi \longrightarrow \mu_\Psi$, from the set of states to the set of finite positive measures, on the real line, would be said to vary quadratically if

1. $\mu_{\alpha\Psi} = |\alpha|^2 \mu_\Psi$,

2. Parallelogram Law :

$$\mu_{\Psi+\xi} + \mu_{\Psi-\xi} = 2\mu_\Psi + 2\mu_\xi \ , \qquad (4.5.7)$$

3. $\mu_\Psi(\mathbb{R}) = \|\Psi\|^2$.

With $\mu_{\Psi}(A) = \frac{1}{4\pi} E\, X_{\Psi}(A, w)$, we see that property 1, for *instance, is satisfied if* $|\alpha|$ is large compared to σ/s. Since σ/s is small, we can, for an approximate theory, assume that 1 is true for all values of α.

We now observe that, for each borel set A, there exists a sesquilinear map

$$(\Psi, \xi) \longrightarrow \mu_{\Psi, \xi}(A), \tag{4.5.8}$$

where

$$\mu_{\Psi, \xi} = \frac{1}{4}\left\{\mu_{\Psi+\xi} - \mu_{\Psi-\xi} + i\mu_{\Psi+i\xi} - i\mu_{\Psi-i\xi}\right\} \tag{4.5.9}$$

is a complex borel measure. Hence, there exist operators E(A) such that

$$\mu_{\Psi, \xi}(A) = \langle E(A)\Psi, \xi \rangle. \tag{4.5.10}$$

$E(A) \geq 0$, since

$$\langle E(A)\Psi, \Psi\rangle = \mu_{\Psi, \Psi}(A) = \mu_{\Psi}(A) \geq 0. \tag{4.5.11}$$

If A_1 and A_2 are disjoint then

$$\begin{aligned}\langle E(A_1 \cup A_2)\Psi, \Psi\rangle &= \mu_\Psi(A_1 \cup A_2) \\ &= \mu_\Psi(A_1) + \mu_\Psi(A_2) \\ &= \langle E(A_1)\Psi, \Psi\rangle + \langle E(A_2)\Psi, \Psi\rangle \\ &= \langle (E(A_1) + E(A_2))\Psi, \Psi\rangle . \qquad (4.5.12)\end{aligned}$$

Since (4.5.12) holds for all Ψ

$$E(A_1 \cup A_2) = E(A_1) + E(A_2) , \qquad (4.5.13)$$

for disjoint A_1 and A_2. Further

$$E(\emptyset) = 0 , \qquad (4.5.14)$$

trivially, and

$$\langle E(\mathbb{R})\Psi, \Psi\rangle = \mu_\Psi(\mathbb{R}) = \|\Psi\|^2 = \langle \Psi, \Psi\rangle = \langle I\Psi, \Psi\rangle , \qquad (4.5.15)$$

implies

$$E(\mathbb{R}) = I , \qquad (4.5.16)$$

where I is the identity operator.

It follows that $E(.)$ is a positive-operator valued measure. $E(.)$ is a projection valued measure if

4. $$\mu_{E(A_1)\Psi, \xi}(A_2) = \mu_{\Psi, \xi}(A_1 A_2) . \qquad (4.5.17)$$

An alternative formulation of condition 4, which is closer to the spirit of statistical mechanics, is obtained by noting that E(.) is a projection valued measure iff $\sqrt{E}$ is additive, i.e., iff the random process

$$Y(s, w) = \sqrt{E(-\infty, s]}\Psi \qquad (4.5.18)$$

is a martingale. These conditions are however not easy to interpret physically. On the other hand, since projections are weakly dense in the convex set of positive contractive operators, a reformulation of quantum mechanics in terms of positive operator valued measures (Davies, 1974) has certain advantages, including the existence of joint distributions.

4.6 Further problems

The correspondence proposed in the preceding sections indicates that an interpretation of quantum mechanics as a semiclassical description of the dynamics of extended particles is, at least partially, feasible. However, a final decision on the possibility of a complete interpretation of quantum mechanics, along these lines, must be deferred till the following issues are resolved :

(i) to determine whether the probabilities for momentum also vary quadratically, and to obtain, explicitly, the relationship between position, momentum and the hamiltonian ;

(ii) to determine whether such extended particles can have angular momentum with properties analogous to spin angular momentum ;

(iii) to determine whether the behaviour of such extended particles can be described in a Lorentz covariant manner.

Although no definite solutions to these problems are available, at present, some possibilities are suggested below.

As far as (i) is concerned, we observe that the various (semiclassical) dynamical variables connected with the extended particle are, essentially, a little fuzzy around the corresponding classical values for the centre of mass of the particle. Since, one can obtain the Schrodinger equation from the Newtonian equations (for instance, Nelson, 1966), it is likely that the present methods would lead, approximately, to the usual equations of quantum mechanics, for a fairly large class of potentials. Slightly different methods would be required, however, since the present theory ascribes only a probability distribution, and not a specific random variable, to a dynamical variable.

Regarding (ii) we observe that there are the widespread misconceptions that intrinsic angular momentum is intrinsically quantum mechanical (for instance, Landau and Lifshitz, 1957, p 186) and that a semiclassical explanation of spin must necessarily involve a rotating extended particle (for instance, McGregor, 1978). Intrinsic angular momentum can be defined for a classical 'point' particle (Synge, 1965), and for an axisymmetric extended particle the net angular momentum need not be zero. Further, a semiclassical explanation for the Davisson-Germer experiment is impossible only if the structure of the dipole is assumed to be independent of the external field. Since the last assumption is quite false in the present theory, an explanation of spin may be difficult, but can not be regarded as impossible, a priori.

(iii) does not appear to be an excessively difficult problem, since, by using a multicomponent wavefunction, an ellipsoid can be described in much the same way as a sphere,

4.7 Empirical tests

If this approach to the interpretation of quantum mechanics is correct, quantum theory would fail in certain situations. Quantitative predictions in concrete experimental situations may take some time to develop. Qualitatively, however, some situations, in which some of the axioms of quantum theory would fail, are immediately discernible.

(a) According to the present theory, the probabilities given by quantum mechanics are approximately correct for regions that are large compared to the mean extension of the particle. So, one can expect failures when $| s |$, in equation (4.5.4), is small. In practice, such situations would occur only when two particles interact at very high energies. It may not be feasible to test the other possibility, viz., that of σ being large.

(b) With extended particles, at high energies, departures from spherical symmetry are bound to occur. One way of testing this would be to look for quadrupole moments in the case of charged particles with spin.

(c) <u>In case the oscillations of the particle are non-linear</u>, failures at low energies are also possible. In such a case, the pulsations of the particle may be considered as a superposition of oscillations at different frequencies. This would imply that the de Broglie relationship, $\lambda \upsilon = c^2/v$, is only approximately correct, and that other 'wavelengths' can be associated with the particle, for the same value of the energy. These wavelengths would be observable as $\lambda \rightarrow \infty$, i.e., at low energies.

4.8 Conclusions

It is, at least partially, possible to interpret quantum mechanics as a semiclassical description of the dynamics of extended particles. If this interpretation is correct, then quantum mechanics would fail at very high, and, possibly, at very low energies.

APPENDIX A

PRODUCTS AND COMPOSITIONS WITH THE DIRAC DELTA-FUNCTION

A 1 Introduction

The notion of the Dirac delta-function has been rigorously formulated in Schwartz's (1951) theory of distributions and Mikusinski's (1959) theory of operators. In both these theories, pointwise products (of two distributions, or two operators) and compositions (of a distribution, or operator, with an ordinary function) are irregular operations in the sense of Mikusinski (1961). However, in many concrete situations in physics, such irregular operations arise, and are dealt with, without due regard to rigour. Although this problem has been known for nearly three decades, it remains incompletely solved. The earlier attempts at defining, or using, pointwise products of distributions (Konig 1953, Guttinger 1955, Gonzalez-Dominguez and Scarfiello 1956, Mikusinski 1961, 1966, Fisher 1971, Thurber and Katz 1974) have largely been inspired by possible applications (Guttinger 1955, Thurber and Katz 1974, Takahashi 1954) to the renormalisation problem in quantum field theory. In particular, most of the effort seems to have gone into proving the formula

$$\delta . x^{-1} = -\frac{1}{2}\delta' , \qquad (A\ 1.1)$$

which may be of some use in quantum mechanics. (A 1.1) was first established by Gonzalez-Dominguez and Scarfiello (1956).

Similarly, compositions have been defined in only a few simple cases (Lojasiewicz 1957, Fisher 1974, Tewari 1977) and many interesting expression lie outside the scope of those definitions. Naturally, these operations, if suitably defined, have a much wider range of applicability. The two examples given below illustrate the general situation that might arise.

(a) Junction conditions : In general relativity, the exact degree of smoothness that can be assigned to the components of the metric tensor, $g_{\mu\lambda}$, is not known. In certain situations it may be physically permissible to choose the $g_{\mu\lambda} \in C^0$ (for instance, Lanczos 1924, Papapetrou and Hamoui 1968, 1979, Evans 1977). Mathematically, however, this leads to difficulties in view of the usual formulae

$$\Gamma_{\mu\lambda\sigma} = \frac{1}{2}(g_{\mu\lambda,\sigma} + g_{\mu\sigma,\lambda} - g_{\lambda\sigma,\mu}) ,$$

$$R_{\mu\lambda} = \Gamma^{\alpha}_{\mu\alpha,\lambda} - \Gamma^{\alpha}_{\mu\lambda,\alpha} - \Gamma^{\alpha}_{\mu\lambda}\Gamma^{\beta}_{\alpha\beta} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\lambda\alpha} , \qquad \text{(A 1.2)}$$

$$R^{\mu\lambda} - \frac{1}{2} g^{\mu\lambda} R = - K T^{\mu\lambda} ,$$

$$T^{\mu\lambda}_{\ \ ;\lambda} = 0$$

In particular, if the $g_{\mu\lambda}$ were chosen to be discontinuous, as suggested by Raju (1979), the components of the Ricci tensor would involve functions of the form δ^2. Thus, either one has to solve the problem mentioned in the first paragraph, or abandon the formulae (A 1.2) and develop altogether new techniques, that may or may not be reliable.

(b) Curved shocks : Shock waves arising in practice are usually curved, and the equations of continuity and momentum

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \bar{v}) = 0 ,$$

$$\rho \frac{\partial \bar{v}}{\partial t} + \rho(\bar{v} . \nabla) \bar{v} = - \nabla P - \rho \nabla F + \operatorname{div} S_v \qquad \text{(A 1.3)}$$

(where S_v is the viscous stress tensor), immediately lead to the above problem if the density ρ, pressure P and velocity $\bar{v}$ are chosen to be discontinuous at an arbitrary hypersurface.

Mathematically, the above problems can be solved by suitably defining entities such as δ^2 ; but, physically, there is a subtler problem. An arbitrary definition would reduce our belief in the equations (A 1.2) and (A 1.3) to a phenomenological one, and, thus, would not be of much value to the physicist. Further, one has to bear in mind the fact that the 'discontinuity' that occurs across a shock front, for instance, may not be a discontinuity in the mathematical sense of the word, i.e., sometimes the representation by discontinuous functions is chosen just to make the problem tractable, and because such a representation may be approximately valid.

In view of this, certain elementary concepts from non-standard analysis are useful, and lead to a very simple solution of the problem of defining irregular operations.

A 2.1 Products

We let D, D' denote the space of test functions and distributions respectively. It is well known that the product of $T \in D'$ and $f \in D$ can be defined by

$$\langle Tf, g \rangle = \langle T, fg \rangle , \qquad \text{(A 2.1)}$$

for any $g \in D$, $\langle T, h \rangle$ denoting the value of the functional T at h. This product is well defined, and the Leibniz formula holds (Rudin 1974). Konig (1953) has constructed product spaces of distributions, and mapped these spaces back into D' in a manner that preserves the formula (A 2.1) in the form

$$\langle TS, g \rangle = \langle T, Sg \rangle . \qquad \text{(A 2.2)}$$

It is asserted that (A 2.2) makes sense on the null space of S, with $TS = 0$ there. Thus, we have

$$\begin{aligned} \delta(x-a)\ \delta(x-b) &= c_1\ \delta(x-b) , \\ \delta\delta' &= c_2\ \delta' + c_3\ \delta , \end{aligned} \qquad \text{(A 2.3)}$$

where c_1, c_2, c_3 are arbitrary constants. The product, in general, is neither commutative nor associative. In fact, there is only one possibility concerning the association of factors in a product $T_1, T_2, \ldots, T_n$ either $T_1(T_2(\ldots T_n))$ or $(\ldots(T_1 T_2)\ldots) T_n$. The usual example for the failure of the associative law is

$$x^{-1}\ (x\delta) = 0 \neq \delta = (x^{-1} x)\ \delta , \qquad \text{(A 2.4)}$$

The main problem with the product, so defined, lies in arbitrariness in the choice of the constants c_1, c_2, c_3. In practice, the choice of the constants is tailored to meet the needs of a particular application. Needless to say, the tailoring does not always fit the physical requirements of the problem, rendering our continued belief in equations of the type (A 1.2) and (A 1.3) invalid.

Mikusinski (1961) on the other hand, has proposed a general theory of irregular operations for distributions. If R is an operation defined for test functions, R can be extended to distributions by defining

$$R(f,g,\ldots) = \lim_{n \to \infty} R(\xi_n, \eta_n, \ldots) \qquad (A\ 2.5)$$

(where ξ_n, η_n are fundamental sequences converging to $f, g, \ldots$), provided the sequence $R(\xi_n, \eta_n, ,,,)$ is fundamental. The sequence ξ_n can be obtained as $f \circledast \delta_n$, where δ_n is a sequence converging to δ, and $\circledast$ denotes convolution. It is asserted that the extension of the operation R, so defined, exists and is unique. Mikusinski (1966) uses this definition to obtain the formula

$$\delta^2 - \frac{1}{\pi^2} \left(\frac{1}{x}\right)^2 = - \frac{1}{\pi^2} \frac{1}{x^2} , \qquad \text{(A 2.6)}$$

and also the equality (A 1.1). The l.h.s. of (A 2.6) is considered as the distributional limit of $\delta_n^2 - \frac{1}{\pi^2} (\frac{1}{x} \circledast \delta_n)^2$, no meaning being assigned to the individual terms.

Fisher (1971) also obtains the formula (A 1.1), using essentially Mikusinski's definition restricted to binary operations. Thus, the product of three distributions (f.g.).h, even if it exists, is not necessarily equal to the limit of the sequence $f_n . g_n . h_n$, but, is given as the limit of the sequence $p_n . h_n$, where p is the distribution f.g. Applications are to be found in Fisher (1972, 1973). The product again fails to be associative although it is commutative. The last two theories do not ascribe any general meaning to the symbol δ^2, and, hence, are not applicable to the sort of problems proposed in the introduction.

Thurber and Katz (1974) do not really define products using the nonstandard extension, $*D'$, of D'. Instead they seem to consider

$$\triangle^{\rho}(x-a) = \left(\frac{n}{\pi}\right)^{\rho/2} e^{-n\rho(x-a)^2} , \qquad \text{(A 2.7)}$$

where n is a positive infinite constant, as a fractional power of the delta function. Naturally, there are various types of delta functions in this theory, i.e., the theory of Thurber and Katz deals with nonstandard extensions of sequences converging to the delta-distribution, and not with the delta distribution, per se.

A 2.2 Definition of f.g

Consider the nonstandard spaces $^*D$ and $^*D'$ (Stroyan and Luxemburg, 1976). Define, for $f, g \varepsilon D'$,

$$f_n = f(\underline{x})\ \delta_n , \qquad \delta_n(x) = n\ \sigma(nx) , \tag{A 2.8}$$

$$f.g \overset{*}{=} \lim_{n = w} {}^*(f_n . g) \tag{A 2.9}$$

σ being a symmetric, infinitely differentiable function with $\int_{-\infty}^{+\infty} \sigma(x)\ dx = 1$, with $\sigma(0) \neq 0$, and with support contained in the interval $[-1,1]$. The $*$ in (A 2.9) denotes the nonstandard extension of the sequence of distributions $f_n . g$, and the notation $\lim_{n = w}$ refers to an evaluation of the w-th term of this sequence for a fixed positive infinite integer w.

The nonstandard representation of a given distribution is nearly unique, in that any two representations would differ by an infinitesimal distribution. If two (nonstandard) distributions, h_1, h_2, differ by an infinitesimal distribution, we write $h_1 \overset{*}{=} h_2$.

The product of two distributions, defined by (A 2.9), always exists in $^*D'$. In case f is a function, the product defined by (A 2.9) differs from the one defined by (A 2.1) by an infinitesimal distribution. (A 2.9) extends (A 2.1), and, in particular, we have

$$\delta^2 \overset{*}{=} \delta_w(0)\,\delta . \qquad \text{(A 2.10)}$$

δ^2 turns out to be an infinite distribution, i.e., $\langle \delta^2 , g \rangle$ is infinite for $g \in \text{fin}\ {}^*D'$ (Stroyan and Luxemburg, 1976).

Naturally, the choice of the infinite subscript w and the sequence δ_n is non-unique, and different choices will lead to different distributions. However, this does not affect the final result, which should be a standard one, and which should not involve the nonstandard elements. For the problems proposed in the introduction, this comes about in the following manner : for standard real numbers a, b, c,

$$a\,\delta^2 + b\,\delta + c \overset{*}{=} 0, \qquad \text{(A 2.11)}$$

$$\text{iff} \qquad a = b = c = 0. \qquad \text{(A 2.12)}$$

Proof : $c = 0$ trivially, and (A 2.11) implies that $a \overset{*}{=} \frac{-b}{\delta_w(0)}$ is an infinitesimal. Since a has been assumed to be a standard real number, $a = 0$, leading to $b = 0$.

In the present theory, also, it is possible to define fractional powers of the delta function by

$$\delta^\rho \overset{*}{=} \delta_w^{\rho-1}(0)\,\delta, \qquad \text{(A 2.13)}$$

leading, in particular, to the infinitesimal distribution $\sqrt{\delta} \overset{*}{=} \delta_w^{-1/2}(0)\,\delta$.

A 2.3 Properties

The commutative law fails, since

$$\delta\delta' \overset{*}{=} \lim_{n=w} \delta_n \cdot \delta' \overset{*}{=} -\delta_w'(0)\delta + \delta_w(0)\,\delta', \qquad \text{(A 2.14)}$$

whereas

$$\delta'\delta \overset{*}{=} \delta_w'(0)\,\delta. \qquad \text{(A 2.15)}$$

The associative law also fails, in general, since

$$(f\delta)\delta' \overset{*}{=} f(0)\ (\delta\delta') \overset{*}{=} -f(0)\ \delta_w'(0)\delta + f(0)\ \delta_w(0)\delta' , \tag{A 2.16}$$

whereas

$$f(\delta\delta') \overset{*}{=} -(f(0)\delta_w'(0) + f'(0))\delta + f(0)\ \delta_w(0)\delta' . \tag{A 2.17}$$

As is obvious from (A 1.2) and (A 1.3), a situation frequently encountered in applications is the multiplication of a delta function by a discontinuous function. To cover this situation, we have the following theorem.

<u>Theorem 1</u> : If f is a function with a simple discontinuity at 0 then

$$f.\delta \overset{*}{=} \frac{1}{2}\left[f(0^+) + f(0^-)\right]\delta , \tag{A 2.18}$$

where $f.\delta$ is defined by (A 2.9).

<u>Proof</u> : It is sufficient to prove that

$$(f \circledast \delta_w)(0) \overset{*}{=} \frac{1}{2}\left[f(0^+) + f(0^-)\right]. \tag{A 2.19}$$

Now, from (A 2.8)

$$(f \circledast \delta_n)(0) = \int_{-\infty}^{\infty} f(-y)\ n\,\sigma(ny)\,dy . \tag{A 2.20}$$

which gives, by a simple change of variables $(ny = x)$

$$(f \circledast \delta_n)(0) = \int_0^1 \left[f\left(\frac{-x}{n}\right) + f\left(\frac{x}{n}\right) \right] \sigma(x)\, dx\,, \qquad \text{(A 2.21)}$$

since σ is symmetric, with support contained in $[-1,1]$.
Since f is continuous in a neighbourhood of zero

$$\begin{aligned} f\left(\frac{-x}{n}\right) &= f(0^-) + \varepsilon_1(n,x)\,, \\ f\left(\frac{x}{n}\right) &= f(0^+) + \varepsilon_2(n,x) \end{aligned} \qquad 0 \leq x \leq 1 \qquad \text{(A 2.22)}$$

where $|\varepsilon_1(n,x)| \leq \bar{\varepsilon}_1(n)$ and $|\varepsilon_2(n,x)| \leq \bar{\varepsilon}_2(n)$, for $0 \leq x \leq 1$

and $\lim_{n \longrightarrow \infty} \bar{\varepsilon}_1(n) = \lim_{n \longrightarrow \infty} \bar{\varepsilon}_2(n) = 0$. Hence

$$\left|(f \circledast \delta_n)(0) - \frac{1}{2}\left[f(0^+) + f(0^-)\right]\right| \leq \bar{\varepsilon}_1(n) + \bar{\varepsilon}_2(n)\,. \qquad \text{(A 2.23)}$$

Since $\bar{\varepsilon}_1(w)$ and $\bar{\varepsilon}_2(w)$ are infinitesimals, the theorem holds.

<u>Corollary</u> 1 : If H is the Heaviside function, $H(x) = 1$, for $x > 0$ and $H(x) = 0$ otherwise,

$$H\,.\,\delta \overset{*}{=} \frac{1}{2}\,\delta\,. \qquad \text{(A 2.24)}$$

(A 2.24) is also valid with Fisher's (1971) definition.

<u>Theorem 2 (Leibniz rule)</u> : For $f, g \in D'$,

$$(f . g)^{(k)} \overset{*}{=} \sum_{i=0}^{k} \binom{k}{i} f^{(i)} g^{(k-i)} . \qquad (A\ 2.25)$$

<u>Proof</u> : By definition, $(f.g)^{(k)} \overset{*}{=} (f_w . g)^{(k)}$. Since, f_w is a function, the validity of the Leibniz rule for the product (2.1) implies that

$$(f_w . g)^{(k)} \overset{*}{=} \sum_{i=0}^{k} \binom{k}{i} f_w^{(i)} . g^{(k-i)} . \qquad (A\ 2.26)$$

Since, by definition, $f^{(i)} . g^{(k-i)} \overset{*}{=} f_w^{(i)} . g^{(k-i)}$, (A 2.25) holds.

<u>Corollary 2</u> : If H is the Heaviside function,

$$H . \delta' \overset{*}{=} \frac{1}{2} \delta' - \delta^2 . \qquad (A\ 2.27)$$

<u>Proof</u> : $(H . \delta)' \overset{*}{=} (\frac{1}{2} \delta)' \overset{*}{=} H' . \delta + H . \delta' \overset{*}{=} \delta^2 + H . \delta'$.

<u>Theorem 3</u> : If a_{ij}, b_k, and c are standard real numbers

$$\Sigma\, a_{ij}\, \delta^{(i)}(x-x_o)\, \delta^{(j)}(x-x_o) + \Sigma\, b_k\, \delta^{(k)}(x-x_o) + c \overset{*}{=} 0,$$

iff $a_{ij} = b_k = c = 0$.

Proof : The proof is similar to that of (A 2.11) - (A 2.12) and will be omitted.

A 2.4 Extensions

Products of the form $\delta^{(k)} . x^{-n}$ are apparently useful in quantum renormalization theory, and we will now explicitly evaluate such products, starting from the product $\delta . x^{-1}$. The definition (A 2.9) cannot immediately be used, since the function x^{-1}, not being integrable in a neighbourhood of zero, does not induce a distribution. However, on the subspace $D(0) = \{g \in D,\ g(0) = 0\}$, x^{-1} induces a linear functional

$$\langle x^{-1}, g \rangle = \int_{-\infty}^{\infty} g(x) . x^{-1}\, dx < \infty, \ \forall\, g \in D(0) . \qquad \text{(A 2.29)}$$

This linear functional is continuous since, for any compact K, $K \subseteq [-n, n]$ and for any $g \in D(0)$, Supp $g \subseteq K$,

$$\int_{-\infty}^{\infty} g(x) . x^{-1}\, dx = \left(\int_{-\infty}^{-1} + \int_{-1}^{1} + \int_{1}^{\infty} \right)(g(x) . x^{-1}\, dx) \qquad \text{(A 2.30)}$$

so

$$\left| \int_{-\infty}^{\infty} g(x) . x^{-1}\, dx \right| \leq \int_{-1}^{1} |g(x) . x^{-1}|\, dx + 2 \log n \ \|g\|_0$$

(A 2.31)

Now, for $0 \leq x \leq 1$

$$|g(x).x^{-1}| = |g'(\theta x)| , \text{ for some } \theta, \ 0 \leq \theta \leq 1$$

$$\leq \ ||g||_1 . \qquad \text{(A 2.32)}$$

Hence,

$$\left| \int_{-\infty}^{\infty} g(x).x^{-1} \, dx \right| \leq 2 \, ||g||_1 + 2 \ \log \ n \, ||g||_0 . \qquad \text{(A 2.33)}$$

It follows from the Hahn-Banach continuous extension theorem, the topology of D being locally convex, that x^{-1} admits a continuous extension to D'. If $\lambda_{1/x}$ and $\lambda'_{1/x}$ are two extensions of x^{-1}, $\lambda_{1/x} - \lambda'_{1/x}$ vanishes on $D(0)$. Hence, the most general extension would be of the form $\lambda_{1/x} + c\,\delta$, where c is an arbitrary constant and $\lambda_{1/x}$ is any particular extension of x^{-1}. A particular extension, $\lambda_{1/x}$, is given by

$$< \lambda_{1/x}, g > = \int_{-\infty}^{\infty} [\, g(x) - g(0) \,] \, x^{-1} \, dx . \qquad \text{(A 2.34)}$$

Hence, $< x^{-1}, g > = c \, g(0) + \int_{-\infty}^{\infty} [\, g(x) - g(0) \,] \, x^{-1} \, dx .$

(A 2.35)

Similarly,

$$< x^{-n}, g > = \sum_{i=0}^{n-1} (-1)^{i} c_i^n g^{(i)}(0) + \int_{-\infty}^{\infty} \left[g(x) - \sum_{i=1}^{n-1} \frac{g^{(i)}(x)}{i!} \right] x^{-n} dx. \tag{A 2.36}$$

Now, $< \delta^{(k)}.x^{-n}, g >$ is finite provided $\lim_{x \to 0} [g(x).x^{-n}]^{(k)}$ exists, and, in that case,

$$< \delta^{(k)}.x^{-n}, g > \overset{*}{=} \lim_{x \to 0} (-1)^k [g(x).x^{-n}]^{(k)} \tag{A 2.37}$$

$$\overset{*}{=} \lim_{x \to 0} (-1)^k \sum_{i=0}^{k} \binom{k}{i} (-1)^i \frac{(n+i-1)!}{(n-1)!} \frac{g^{(n-1)}}{x^{n+i}}$$

The limit of any of the terms in the above summation exists, iff $g \, \varepsilon^* D(0,1,\ldots,k+n-1) \overset{*}{=} \left\{ \phi \, \varepsilon \, D, \ \phi(0) = \ldots = \phi^{(k+n-1)}(0) = 0 \right\}$, and, in that case

$$\lim_{x \to 0} \frac{g^{(k-i)}(x)}{x^{n+i}} \overset{*}{=} \frac{g^{(k+n)}(0)}{(n+i)!} . \tag{A 2.38}$$

That is

$$< \delta^{(k)}.x^{-n}, g > \overset{*}{=} \frac{(-1)^k g^{(k+n)}(0)}{(n-1)!} \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{n+i}, \tag{A 2.39}$$

$$g \, \varepsilon \, D(0,1,\ldots,k+n-1) .$$

It follows that

$$\delta^{(k)} . x^{-n} \stackrel{*}{=} \sum_{i=0}^{k+n-1} c^n_{ki} \, \delta^{(i)} + \frac{(-1)^n}{(n-1)!} \, \delta^{(k+n)} . \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{(n+i)} , \tag{A 2.40}$$

with the added proviso that all the constants $c^n_{ki} \in {}^*R$ be infinite, if the constants c^n_1 in equation (A 2.36) are chosen to be finite.

The Leibniz rule, in the form

$$[\delta^{(k)} . x^{-n}]' \stackrel{*}{=} \delta^{(k)} . -n \, x^{-(n+1)} + \delta^{(k+1)} . x^{-n} , \tag{A 2.41}$$

is satisfied on the subspace $^*D(0,1,\ldots,k+n)$, because

$$\sum_{i=0}^{k+1} \binom{k}{i} \frac{(-1)^i}{n+i} = \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{(n+1+i)} + \sum_{i=0}^{k+1} \binom{k+1}{i} \frac{(-1)^i}{n+i} . \tag{A 2.42}$$

We have, thus, achieved a legitimate method of subtracting infinities. The method is not fully satisfactory because the constants c^n_i can only be fixed phenomenologically. Further, the formula (A 1.1) is not valid, since

$$\delta . x^{-1} \stackrel{*}{=} -\delta' , \tag{A 2.43}$$

on $^*D(0)$.

A 3. Compositions

If g is a C^1 function, $g(x_1) = 0$, $g'(x_1) \neq 0$, $\delta(g(x))$ is usually defined (Gelfand and Shilov 1964) by carrying out a formal change of variables

$$\langle \delta(g(x)), h \rangle = \frac{h(x_1)}{|g'(x_1)|} , \quad \forall h \in D ,$$

i.e., $$\delta(g(x)) = \frac{1}{|g'(x_1)|} \delta(x-x_1). \qquad (A\ 3.1)$$

Here, we will define, for any $f \in D'$, $g \in C^\infty$, $g' \neq 0$

$$f(g(x)) \overset{*}{=} \lim_{n=w} {}^*(f_n(g(x)) ,$$

$$f_n = f\ (\underline{x})\ \delta_n . \qquad (A\ 3.2)$$

The distribution defined by (A 3.2) always exists and is given, using the change of variables formula for ordinary functions, by

$$\langle f(g(x)), h \rangle \overset{*}{=} \langle f_w , \frac{h(g^{-1}(x))}{|g'(g^{-1}(x))|} \rangle$$

$$\overset{*}{=} \langle f , \frac{h(g^{-1}(x))}{|g'(g^{-1}(x))|} \rangle . \qquad (A\ 3.3)$$

Moreover, the distribution defined by (A 3.2) is nearly unique in that a different sequence δ_n or a different infinite constant w' would lead to a distribution $f_{w'}(g(x))$ which would differ from $f_w(g(x))$ by an infinitesimal distribution.

The chain rule is valid.

Theorem 4 (chain rule) : If $f \in D'$, $g \in C^\infty$, $g' \neq 0$,

$$[f(g(x))]' \overset{*}{=} f'(g(x))\, g'(x). \qquad \text{(A 3.4)}$$

Proof : The nonstandard proof follows immediately from the usual chain rule. One can also see this directly, since, for any $h \in D$,

$$\langle [f(g(x))]', h \rangle \overset{*}{=} - \langle f(g(x)), h' \rangle$$

$$\overset{*}{=} - \langle f, \frac{h'(g^{-1}(x))}{|g'(g^{-1}(x))|} \rangle ; \qquad \text{(A 3.5)}$$

and

$$\langle f'(g(x)).g'(x), h \rangle \overset{*}{=} \langle f'(g(x)), g'(x).h \rangle$$

$$= \langle f', \frac{g'(g^{-1}(x))\, h(g^{-1}(x))}{|g'(g^{-1}(x))|} \rangle$$

$$= \langle -f, \pm [h(g^{-1}(x))]' \rangle$$

$$= - \langle f, \frac{h'(g^{-1}(x))}{|g'(g^{-1}(x))|} \rangle , \qquad \text{(A 3.6)}$$

since $(g^{-1})' = 1 / g'(g^{-1})$.

In case g has several real roots $x_1, x_2, \ldots, x_n$, $g'(x_i) \neq 0$, the definition (A 3.2) reduces to the usual definition for a distribution of finite order, provided $g'(g^{-1}(x)) \neq 0$, for $x \in \operatorname{supp} f$, since

$$\langle \delta_W^{(k)}(g(x)), h \rangle \stackrel{*}{=} \sum_{i=1}^{n} (-1)^k \left[\frac{h(g^{-1}(x))}{|g'(g^{-1}(x))|} \right]^{(k)} \Bigg|_{x=x_i}$$

(A 3.7)

The chain rule continues to be valid.

A 3.2 Multiple roots

In case $g \in C^\infty$ has multiple roots at x_1, then (A 3.3) can no longer be used, since $g'(x_1) = 0$. In this situation, Fisher (1974) and Tewari (1977) have adopted a limiting procedure to define $\delta^{(r)}(x^{2m+1})$, yielding

$$\delta^{(r)}(x^{2m+1}) = \frac{r!\ \delta^{2mr+2m+r}(x)}{(2m+1)(2mr+2m+r)!} + \sum_{i=0}^{2mr+2m+r-1} C_i^{rm}\, \delta^{(i)} ,$$

(A 3.8)

where C_i^{rm} are arbitrary constants, the functional having been extended in the usual manner. The distributions corresponding to $\delta^{(r)}(x^{2m})$ are not defined by Fisher (1974), because x^{2m} is not invertible in a neighbourhood of zero.

If the definition (A 3.2) is used, $\delta^{(r)}(g(x))$ is defined for $g \in C^{\infty}$, regardless of the nature of the roots of $g(x)$. Thus,

$$< \delta^{(r)}(g(x)), \phi(x) > \overset{*}{=} \lim_{n=w} \int_{-\infty}^{\infty} \delta_n^{(r)}(g(x)) \; \phi(x)dx , \qquad (A\ 3.9)$$

and

$$\text{if} \quad I_n^r = \int_{-\infty}^{\infty} \delta_n^{(r)}(g(x)) \; \phi(x)\,dx , \qquad (A\ 3.10)$$

then I_w^r is * finite, since I_n^r is finite, the integrand being continuous with compact support. If g is invertible in a neighbourhood of zero, I_w^r is finite provided $\lim_{x \to 0} \phi(g^{-1}(x))/|g'(g^{-1}(x))|$ exists. In particular, if $g(x) = x^{2m+1}$, the definition (A 3.2) agrees with (A 3.8) on the subspace $D(0,1,\ldots,2mr+2m+r-1)$. However, the constants C_i^{rm}, in this theory, are not arbitrary. To evaluate these constants, we select functions $h_i \in D$ which behave like $x^i/i!$ in a neighbourhood of zero. Thus,

$$\begin{aligned} h_i^{(k)}(0) &= 0 , \quad i \neq k \\ h_i^{(i)}(0) &= 1 \end{aligned} \qquad (A\ 3.11)$$

Further,

$$< \delta^{(r)}(x^{2m+1}), h_i(x) > \overset{*}{=} (-1)^i C_i^{rm} \overset{*}{=} \lim_{n=w} \int_{-\infty}^{\infty} \delta_n^{(r)} (x^{2m+1}) h_i(x)\,dx$$

$$\overset{*}{=} \lim_{n=w} n^{r+1} \int_{-\infty}^{\infty} \sigma^{(r)} (nx^{2m+1}) h_i(x)\,dx$$

$$\overset{*}{=} \lim_{n=w} n^{r+1} \int_{-n^{-1/2m+1}}^{n^{-1/2m+1}} \sigma^{(r)} (nx^{2m+1}) h_i(x)\,dx \,, \qquad \text{(A 3.12)}$$

since

$$\sigma^{(r)} (nx^{2m+1}) = 0$$

$$\text{if} \quad x > n^{-1/2m+1} \qquad \text{(A 3.13)}$$

$$\text{or} \quad x < -n^{-1/2m+1}$$

(A 3.13) holds since σ, hence $\sigma^{(r)}$, is identically zero outside $[-1,1]$. For sufficiently large n, $h_i(x)$ may be replaced by $x^i/i!$, and carrying out the change of variables $y = n^{1/2m+1} x$, in I_n

$$I_n = \int_{-n^{-1/2m+1}}^{n^{-1/2m+1}} n^{r+1} \sigma^{(r)} (nx^{2m+1}) \frac{x^i}{i!}\,dx$$

$$= \int_{-1}^{1} \frac{n^{(2m+r-i)}}{i!} \sigma^{(r)}(y)\, y^i\,dy\,. \qquad \text{(A 3.14)}$$

So,

$$(-1)^i \, I_w \overset{*}{=} C_i^{mr} \overset{*}{=} (-1)^i \frac{w^{(2m+r-i)}}{i!} \int_{-1}^{1} \sigma^{(r)}(y) y^i \, dy \qquad (A\ 3.15)$$

A 3.3 Even functions

Compositions with even functions occur in many situations in physics. Apart from situations arising out of the examples mentioned in the introduction, we may mention the Schwarzschild-Tetrode-Fokker action (Hoyle and Narlikar 1964)

$$J = -\sum_i m_i \int (z^{\mu}_{(i)} \; z_{(i)\mu})^{1/2} \, d\tau_i$$

$$- \frac{1}{2} \sum_{i \neq j} e_i e_j \int z^{\mu}_{(i)} \delta [z^{\mu}_{(i)} - z^{\mu}_{(j)}) (z_{(i)\mu} - z_{(j)\mu})] \, z_{(j)\mu} d\tau_i d\tau_j \, ,$$

(A 3.16)

where e_i = charge, m_i = mass and $z_i = z_i(\tau_i)$ = world line of ith particle.

As pointed out earlier, compositions with even functions have not been defined previously, because a change of variables cannot immediately be carried out. However, expressions such as $\delta_w(x^{2m})$ still make sense, and these still induce distributions/in *D' because $\delta_n(x^{2m})$ induces a distribution for each n.

Further,

$$\langle \delta_w (x^{2m}), g \rangle \stackrel{*}{=} \lim_{n=w} \int_{-\infty}^{\infty} \delta_n(x^{2m}) \; g(x)dx$$

$$\stackrel{*}{=} \lim_{n=w} \int_{-\infty}^{\infty} n \, \sigma(nx^{2m}) \; g(x)dx$$

$$= \lim_{n=w} \int_{-n^{-1/2m}}^{n^{-1/2m}} n \, \sigma(n \, x^{2m}) \; g(x)dx \qquad (A \; 3.17)$$

since supp $\sigma \subseteq [-1,1]$. Now

$$I_n = \int_{-n^{-1/2m}}^{n^{-1/2m}} n \, \sigma(n \, x^{2m}) \; g(x)dx = \int_{-n^{-1/2m}}^{0} + \int_{0}^{n^{-1/2m}} \qquad (A \; 3.18)$$

In the first integral we carry out the change of variables

$y = x$, yielding

$$I_n = \int_{0}^{n^{-1/2m}} n \, \sigma(n \, x^{2m}) \; [g(x) + g(-x)]dx . \qquad (A \; 3.19)$$

Substituting $x = (z/n)^{1/2m}$, $0 \leq x \leq n^{-1/2m}$

$$I_n = \frac{1}{2m} \int_0^1 \sigma(z) \left\{ g[(z/n)^{1/2m}] + g[-(z/n)^{1/2m}] \right\} (z/n)^{(1-2m)/2m} dz$$

$$(A \; 3.20)$$

The integral in (A 3.20) is finite provided

$\lim_{y \to 0^+} [g(y^{1/2m}) + g(-y^{1/2m})] . y^{\frac{(1-2m)}{2m}}$ exists. The limit in question exists and is zero provided $g \in D(0,1,2,\ldots,2m-2)$. Since g has compact support, $\lim_{n \to \infty} I_n = 0$, i.e., $\delta_W(x^{2m})$ is infinitesimal for $g \in \text{fin}\ {}^*D(0,1,2,\ldots,2m-2)$. It follows that

$$\delta_W(x^{2m}) \stackrel{*}{=} \sum_{i=0}^{2m-2} a_i^{om}\, \delta^{(i)} . \qquad (A\ 3.21)$$

To evaluate the constants a_i^{om} we select, as before, functions h_i so that

$$(\delta_W(x^{2m}), h_i > \stackrel{*}{=} (-1)^i a_i^{om}$$

$$\stackrel{*}{=} \lim_{n = w} \int_0^{n^{-1/2m}} n\, \sigma(n\, x^{2m}) [\frac{x^i}{i!} + \frac{(-x)^i}{i!}] dx .$$

(A 3.22)

Substituting $z = n^{1/2m}\, x$

$$I_n = \frac{n^{2m-i-1}}{i!} \int_0^1 \sigma(z^{2m}) [z^i + (-z)^i] dz . \qquad (A\ 3.23)$$

So,

$$a_i^{om} \overset{*}{=} (-1)^i \, I_w = \frac{(-1)^i \, w^{2m-i-1}}{i!} \int_0^1 \sigma(z^{2m}) \left[z^i + (-z)^i \right] dz \tag{A 3.24}$$

More generally, we obtain by the above procedure

$$\delta^{(r)} \, (x^{2m}) \overset{*}{=} \sum_{i=0}^{2mr+2m-2} a_i^{rm} \, \delta^{(i)} \tag{A 3.25}$$

$$a_i^{rm} \overset{*}{=} (-1)^i \, \frac{w^{2mr+2m-i-1}}{i!} \int_0^1 \sigma^{(r)} (z^{2m}) \left[z^i + (-z)^i \right] dz \tag{A 3.26}$$

Compositions with oscillating functions also occur, for instance, in the study of oscillating surface layers in relativity. For this purpose, we record

$$\delta(1-\mathrm{Sin}\, x) \overset{*}{=} k \sum_n \delta\left(x - \frac{(4n+1)\pi}{2}\right) \tag{A 3.27}$$

$$k \overset{*}{=} \sqrt{\frac{w}{2}} \int_0^{\pi/2} \sigma(1-\mathrm{Sin}\, y) \sqrt{1+\mathrm{Sin}\, y} \; dy \, . \tag{A 3.28}$$

Before concluding this section, we observe that the hypothesis $g \in C^\infty$ is not essential for defining $f(g(x))$, for $f \in D'$. $f(g(x))$ exists in $*D'$ provided $f_n(g(x)) \in D'$ for each n, i.e., provided $f_n(g(x))$ is locally integrable.

A 4. Derivation of the Rankine-Hugoniot equations

It is usually asserted that the Euler equations (A 1.3) are not valid at discontinuities, where one must use, for instance, the Rankine-Hugoniot equations. The Rankine-Hugoniot equations are usually stated in the form

$$\rho_o(U - u_o) = \rho(U - u) ,$$
$$\rho_o(U - u_o)\ (u - u_o) = P - P_o \qquad \text{(A 4.1)}$$

where U = shock velocity, and the subscript o refers to the undisturbed fluid ahead of the shock. We propose to derive these equations from the usual equations of continuity and momentum (A 1.3) thereby demonstrating that the equations (A 1.3) are indeed valid at discontinuities. To this end, we observe that the equations (A 4.1) are valid for normal shocks of infinite extent, and a two point flow-field in two dimensions, with a simple discontinuity at the surface of the shock.

To derive these equations, using the present methods, we suppose that the hypersurface of discontinuity is given by

$$y = \bar{y}(t) = Ut , \quad U = \frac{\partial \bar{y}}{\partial t} = \text{constant}. \qquad \text{(A 4.2)}$$

Further, let

$$\rho = \rho^- \chi_- + \rho^+ \chi_+ ,$$
$$u_i = u_i^- \chi_- + u_i^+ \chi_+ , \qquad i = 1,2 \qquad (A\ 4.3)$$
$$P = P^- \chi_- + P^+ \chi_+$$

where

$$\chi_+ = \chi_{(\bar{y}(t),\infty)}(x) = H(x - \bar{y}(t)) ,$$
$$\chi_- = 1 - \chi_+ , \qquad (A\ 4.4)$$

H being the Heaviside function.

We observe that

$$\frac{\partial \chi_+}{\partial x} = \delta(x - \bar{y}(t)) , \qquad \frac{\partial \chi_-}{\partial x} = - \delta(x - \bar{y}(t)) ,$$
$$\frac{\partial \chi_+}{\partial t} = -U\ \delta(x - \bar{y}(t)), \qquad \frac{\partial \chi_-}{\partial t} = + U\ \delta(x - \bar{y}(t)) . \qquad (A\ 4.5)$$

The equation of continuity, in two dimensions, is

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho\ u_1)}{\partial x} + \frac{\partial (\rho\ u_2)}{\partial y} = 0 \qquad (A\ 4.6)$$

Substituting (A 4.3) in (A 4.6), we have

$$\rho^- \frac{\partial \chi_-}{\partial t} + \rho^+ \frac{\partial \chi_+}{\partial t} + \rho^-\ u_1^- \frac{\partial \chi_-}{\partial x} + \rho^+\ u_1^+ \frac{\partial \chi_-}{\partial x} = 0 . \qquad (A\ 4.7)$$

Substituting (A 4.5) in (A 4.7) leads to

$$-(\rho^- - \rho^+)\, U\, \delta(x-\bar{y}(t)) = (\rho^+ u_1^+ - \rho^- u_1^-)\, \delta(x-\bar{y}(t)) \qquad \text{(A 4.8)}$$

Hence, by Proposition 3,

$$\rho^+ (U - u_1^+) = \rho^- (U - u_1^-) \qquad \text{(A 4.9)}$$

which is the same as the first of (A 4.1) with $\rho_o = \rho^+$, $\rho = p^-$, $u_o = u_1^+$, $u = u_1^-$.

The momentum equation is

$$\rho \frac{\partial u_1}{\partial t} + \rho\, u_1 \frac{\partial u_1}{\partial x} = \frac{\partial P}{\partial x} = -\frac{\partial P}{\partial y} \qquad \text{(A 4.10)}$$

Substituting (A 4.3) and (A 4.5) in (A 4.10), we have

$$-(Uu_1^+ - Uu_1^-)\, \rho\, \delta(x-\bar{y}(t)) + (\rho^+ u_1^+ X_+ + \rho^- u_1^- X_-)(u_1^+ - u_1^-)\, \delta(x-\bar{y}(t))$$

$$= (P^- - P^+)\, \delta(x-\bar{y}(t)) \,. \qquad \text{(A 4.11)}$$

Hence, by Propositions 1 and 3,

$$\tfrac{1}{2}(u_1^+ - u_1^-)\, [\rho^+ u_1^+ + \rho^- u_1^- - \rho^+ U - \rho^- U] = P^- - P^+ . \qquad \text{(A 4.12)}$$

Using, (A 4.9) leads to the second Rankine-Hugoniot equation

$$\rho^+ (U - u_1^+)\, (u_1^- - u_1^+) = P^- - P^+ . \qquad \text{(A 4.13)}$$

Sometimes a third Rankine-Hugoniot equation is used, and this can be derived similarly from the energy equation.

A 4.2 The effect of viscosity

If turbulence is present behind, and in front of, the shock, the effect of (eddy) viscosity may be taken into account by including a term of the form $\mu \nabla^2 u$ in place of the viscous stress tensor in (A 1.3). Now, (A 4.5) yields

$$\frac{\partial u_1}{\partial x} = (u_1^+ - u_1^-)\ \delta(x-\bar{y}(t)) \qquad \text{(A 4.14)}$$

$$\frac{\partial^2 u_1}{\partial x^2} = (u_1^+ - u_1^-)\ \delta'(x-\bar{y}(t)) \qquad \text{(A 4.14)}$$

Hence, by Theorem 3, we would have an additional equation of the form

$$\mu(u_1^+ - u_1^-) = 0 \qquad \text{(A 4.15)}$$

which must be regarded as a consistency condition. Using equation (A 4.13) this can be put in the form

$$N = \frac{\mu(P^- - P^+)}{\rho^+(u_1^+ - U)} = 0\,. \qquad \text{(A 4.16)}$$

If we only have $N \sim 0$, the Rankine-Hugoniot equations are only approximately applicable.

The physical interpretation of this consistency condition is that the Rankine-Hugoniot equations are not valid for strong shocks - strong shocks would curve due to the effect of viscosity. It is proposed to develop a general theory of curved shocks, using the above methods.

APPENDIX B

CLASSICAL TIME-SYMMETRIC ELECTRODYNAMICS

B 1. Detailed abstract

Maxwell's equations are time symmetric, but the advanced solutions are rejected on semi-empirical grounds. However, it was pointed out by Dirac (1938) that a Lorentz-covariant derivation of radiative damping necessitates the use of time-symmetric solutions. The problem of the absorber theory of radiation is to recover the usual retarded solutions starting from time-symmetric solutions. Wheeler and Feynman (1945) proposed that this could be done by incorporating the response of the 'absorber' (rest of the universe) with the elementary time-symmetric field of a particle. Hogarth (1962) and Hoyle and Narlikar (1964) modified this theory to include the expansion of the universe. The above two theories suffer from serious drawbacks.

In the present theory, these drawbacks are removed by attributing a 'rigid' shell-like structure to elementary particles. According to this theory, retarded radiation will be approximately consistent in the expansion-phase of a universe evolving from an initial singularity. The theory predicts the existence of small amounts of advanced radiation and compares this prediction with the results of Partridge's (1973) experiment.

APPENDIX C

C 1. Evaluation of the distribution function

Let $X(s, w) = X((-\infty, s], w)$ (C 1.1)

and $F_s(t) = P\left\{ w,\ X(s, w) \le t \right\}$. (C 1.2)

Clearly $F_s(t) = 0$ for $t < 0$, (C 1.3)

and for $t \geq 0$.

$$\left\{ w,\ X(s, w) \le t \right\} = \left\{ w,\ s \le -|\Psi| \right\} \cup \left\{ w, -|\Psi| \le s \le |\Psi| \right.$$
$$\left. \text{and } 2\pi |\Psi| (s + |\Psi|) \le t \right\} \cup \left\{ w, |\Psi| \le s \text{ and } 4\pi |\Psi|^2 < t \right\} \qquad \text{(C 1.4)}$$

since,

$$X(s, w) = 0 \text{ if } s \le -|\Psi| ,$$
$$= 2\pi |\Psi| (s + |\Psi|), \text{ if } -|\Psi| \le s \le |\Psi| ; \qquad \text{(C 1.5)}$$
$$= 4\pi |\Psi|^2, \text{ if } |\Psi| \le s .$$

The probability of each of the sets in (C 1.4) is evaluated below.

(1) $$P\left\{ s \le -|\Psi| \right\} = P\left\{ |\Psi| \le -s \right\}$$
$$= F_{|\Psi|}(-s) \qquad \text{(C 1.6)}$$

where $F_{|\Psi|}$ is the distribution function of $|\Psi|$

(2) $P\{|\Psi| \leq s \text{ and } 4\pi|\Psi|^2 \leq t\} = 0$, if $s < 0$ (C 1.7)

$$= F_{|\Psi|}(\sqrt{t/4\pi}), \text{ if } t \leq 4\pi s^2$$
$$s \geq 0 \qquad \text{(C 1.8)}$$
$$= F_{|\Psi|}(s), \quad \text{if } t \geq 4\pi s^2$$

(3) $P\{-|\Psi| \leq s \leq |\Psi| \text{ and } 2\pi|\Psi|^2 + 2\pi s|\Psi| - t \leq 0\}$.

The quadratic equation $2\pi x^2 + 2\pi s x - t = 0$ always has two real roots since the discriminant $\Delta = 4\pi^2 s^2 + 8\pi t > 0$, since $t > 0$.

Let

$$s_{\pm}(t) = \frac{-2\pi s \pm \sqrt{4\pi^2 s^2 + 8\pi t}}{4\pi}$$

$$= -s/2 \pm \frac{1}{2}\sqrt{s^2 + 2t/\pi} \qquad \text{(C 1.9)}$$

then $2\pi|\Psi|^2 + 2\pi s|\Psi| - t \leq 0$ iff

$$s_- \leq |\Psi| \leq s_+ \qquad \text{(C 1.10)}$$

If $s \geq 0$ then

$$s_- = \frac{s}{2} - \frac{1}{2}\sqrt{s^2 + 2t/\pi} \leq -\frac{s}{2} \leq s, \qquad \text{(C 1.11)}$$

and $s_+ \geq s$ if

$$- \frac{s}{2} + \frac{1}{2} \sqrt{s^2 + 2t/\pi} \geq s ,$$

i.e., iff $t \geq 4\pi s^2$. (C 1.12)

Thus the above probability is zero if $t \leq 4\pi s^2$ and is otherwise

$$F_{|\Psi|}(s_+) - F_{|\Psi|}(s). \qquad \text{(C 1.13)}$$

If $s \leq 0$ then

$$s_- = - \frac{s}{2} - \frac{1}{2} \sqrt{s^2 + 2t/\pi} \leq - \frac{s}{2} \leq - s \qquad \text{(C 1.14)}$$

and, $s_+ \geq -s$, since $t > 0$.

It follows that the above probability is $F_{|\Psi|}(s_+) - F_{|\Psi|}(-s)$ for all $t \geq 0$.

Thus

$$\begin{aligned} F_s(t) &= F_{|\Psi|}(\sqrt{t/4\pi}) \quad \text{if} \quad t \leq 4\pi s^2 \\ &= F_{|\Psi|}(s_+(t)) \quad \text{if} \quad t > 4\pi s^2 \end{aligned} \qquad s \geq 0 \qquad \text{(C 1.16)}$$

and $F_s(t) = F_{|\Psi|}(s_+(t))$ if $s \leq 0$. (C 1.17)

$$E\, X_\Psi(s, w) = \int_0^\infty t \, dF_s(t) \qquad \text{(C.1.18)}$$

has been evaluated below assuming that Ψ has a complex gaussian distribution with mean zero and variance σ^2.

C 2. <u>Evaluation of the integral</u>

$$I = \int_0^\infty t \, d\, F_s(t) \qquad \text{(C 2.1)}$$

<u>Case I</u> If $s < 0$, then

$$F_s(t) = F_{|\Psi|}(s_+(t)) \qquad \text{(C 2.2)}$$

where

$$d\, F_{|\Psi|}(x) = \frac{1}{\sigma^2} x \, e^{-x^2/2\sigma^2} dx\,, \quad x \geq 0, \qquad \text{(C 2.3)}$$

$$= 0 \qquad\qquad x \leq 0$$

and $s_+(t)$ is defined by (C 1.9) and for simplicity we take $\|\Psi\|_2^2 = 2\sigma^2$.

Integrating by parts, making the transformation

$$t \longrightarrow v = +\sqrt{s^2 + \frac{2t}{\pi}} \qquad \text{(C 2.4)}$$

and observing that when $t = 0$, $v = -s$ since $s < 0$

$$I = \int_{-s}^{\infty} \pi v \, e^{-(v-s)^2/8\sigma^2} dv \qquad \text{(C 2.5)}$$

with

$$v \longrightarrow z = v - s \qquad \text{(C 2.6)}$$

$$I = 4\pi\sigma^2 \, e^{-s^2/2\sigma^2} + \pi\sqrt{\pi}\sqrt{2}\,\sigma\, s \ \mathrm{Erfc}\,(-s/\sqrt{2}\sigma) \qquad (C\ 2.7)$$

<u>Case II</u> If $s \geq 0$, then

$$\begin{aligned} F(t) &= F_{|\Psi|}\,(\sqrt{t/4\pi}) \quad \text{if} \quad t \leq 4\pi s^2 \\ &= F_{|\Psi|}\,(s_+(t)) \quad \text{if} \quad t \geq 4\pi s^2 \end{aligned} \qquad (C\ 2.8)$$

with $F_{|\Psi|}$ and $s_+(t)$ as before and

$$I = I_1 + I_2 \qquad (C\ 2.9)$$

$$\begin{aligned} I_1 &= \int_0^{4\pi s^2} t \, d\, F_{|\Psi|}\,(\sqrt{t/4\pi}) \\ &= -4\pi s^2 \, e^{-s^2/2\sigma^2} + 8\pi\sigma^2 \, [1 - e^{-s^2/2\sigma^2}] \end{aligned} \qquad (C\ 2.10)$$

and

$$I_2 = \int_{4\pi s^2}^{\infty} t \, d\, F_{|\Psi|}\,(s_+(t)) \qquad (C\ 2.11)$$

reduces as in case I above to give

$$I_2 = 4\pi s^2 \, e^{-s^2/2\sigma^2} + 4\pi\sigma^2 \, e^{-s^2/2\sigma^2} + \pi\sqrt{\pi}\sqrt{2}\sigma \ \mathrm{Erfc}\ (s/\sqrt{2}\sigma) \qquad (C\ 2.12)$$

BIBLIOGRAPHY

Akutowicz E and Wiener N 1957 Rend. Circ. Mat. Palermo 6, 1

Babbitt D G 1963 J. Math. Phys. 4, 36.

Bandyopadhyay P and Roy S 1976 Int. J. Theo. Phys. 15, 323.

Belinfante F J 1973 A Survey of Hidden Variable Theories (New York : Pergamon)

Bell J S 1964 Physics 1, 195.

_______ 1966 Rev. Mod. Phys. 38, 444.

Bohm D 1951 Quantum Theory (New York : Prentice Hall)

_______ 1952a Phys. Rev. 85, 180.

_______ 1952b Phys. Rev. 85, 186.

_______ 1953a Phys. Rev. 89, 458.

_______ 1953b Prog. Theor. Phys. 9, 273.

_______ and Akaranov Y 1957 Phys. Rev. 108, 1070.

_______ 1960 Nuovo Cim. 17, 964.

Bohm D and Bub J 1966a Rev. Mod. Phys. 38, 453.

_______ 1966b Rev. Mod. Phys. 38, 470.

Bohr N 1935 Phys. Rev. 48, 696.

Brittin W E and Chappel W R 1962 Rev. Mod. Phys. 34, 620.

de Broglie L 1951 C.r. Acad. Sci. Paris 233, 64.

_______ 1952 C.r. Acad. Sci. Paris 234, 265 ; 235, 557 ; 235, 1345.

CameronR H 1960 J. Math. and Phys. 39, 126.

________ and Martin W T 1945a Trans. Amer. Math. Soc. 58, 184

________ 1945b Bull. Amer. Math. Soc. 51, 73.

Capper D M and Duff M J 1974 Nucl. Phys. B62, 147.

Chandrashekhar S 1943 Rev. Mod. Phys. 15, 1

Clauser J F 1976 Phys. Rev. Lett. 36, 1223.

________ and Shimony A 1978 Rep. Prog. Phys. 41, 1881

Clauser J F, Horne M A, Shimony A and Holt R A 1969 Phys. Rev. Lett. 23, 880.

Colodny R G (ed) 1972 Paradigms and Paradoxes ; The Philosophical Challenge of the Quantum Domain

Dautcourt G 1964 Math. Nachr. 74, 518.

Davies E B 1974 The Quantum Mechanics of Open Systems (New York : Academic Press).

Deser S and van Nieuwenhuizen 1974 Phys. Rev. D 10, 401 ; 10, 411.

De Witt B S 1964 Relativity Groups and Topology (London : Gordon and Breach)

________ 1967 Phys. Rev. 162, 1195.

________ and Graham R N 1972 The Many Worlds Interpretation of Quantum Mechanics

Dinculeanu N 1967 Vector Measures (Oxford : Pergamon Press)

Dirac P A M 1930 The Principles of Quantum Mechanics (Oxford : Clarendon)

________ 1938 Proc. Roy Soc. A 167, 148.

________ 1945 Rev. Mod. Phys. 17, 195.

________ 1962a Proc. Roy. Soc. A 268, 57.

________ 1962b Proc. Roy. Soc. A 270, 354.

________ 1975 General Theory of Relativity (New York : John Wiley)

Ditchburn R W 1930 Proc. Roy. Irish Acad. 39, 73.

Einstein A, Podolsky B and Rosen N 1935 Phys. Rev. 47, 777.

Emch G 1972 Algebraic Methods in Statistical Mechanics and Quantum Field Theory (New York : Wiley Interscience).

d' Espagnat B 1971 Conceptual Foundations of Quantum Mechanics (New York : Benjamin).

Evans A B 1977 Gen. Rel. Grav. 8, 155.

Everett H 1957 Rev. Mod. Phys. 29, 454.

Faraci G, Guttkowski D, Notarugo S and Pennisi A R 1974 Lett. Nuovo Cim. 9, 607

Feynman R P 1948 Rev. Mod. Phys. 20, 367.

________ 1949 Phys. Rev. 76, 749.

Feynman R P and Hibbs A R 1965 Quantum Mechanics and Path Integrals (New York : Mcgraw Hill).

Fisher B 1971 Quart. J. Math. 22, 291.

________ 1972 Proc. Camb. Phil. Soc. 71, 123

________ 1973 Math. Ann. 203, 103

________ 1974 Math. Student 42, 28.

Freedman S J and Clauser J F 1972 Phys. Rev. Lett. 28, 938.

Frenkel A 1977 in Leite-Lopes and Paty (ed) p

Fry E S and Thompson R C 1976 Phys. Rev. Lett. 37, 465.

Gelfand and Shilov G 1964 Generalised Functions Vol. I.

Gelfand I M and Yaglom A M 1960 J. Math. Phys. 1, 48.

Gihman II and Skorohod A V 1973 Stochastic Processes Vol. I (Berlin : Springer).

Gleason A M 1957 J. Math. Mech. 6, 885.

Gonzalez-Dominguez A and Scarfiello A 1956 Rev. de la Union Matem. Argen. 1, 53.

Gudder S P 1977 in Price and Chissick (eds) p 247.

Guttinger W 1955 Prog. Theor. Phys. 13, 612.

Haag and Kastler D 1964 J. Math. Phys. 5, 848.

Heisenberg W 1927 Zeit. Phys. 43, 172.

Holt R A 1973 Ph.D thesis, Harvard.

't Hooft G and Veltman M 1974 Ann. Inst. H. Poincare 20, 69.

Hooker C A 1972 in Colodny (ed).

Hoyle F and Narlikar J V 1964 Proc. Roy. Soc. A 277, 1

Hurd A and Loeb P (eds) 1974 Victoria Symposium on Nonstandard Analysis (Berlin : Springer).

Israel W 1966 Nuovo Cim. 44B, 1.

Jammer M 1974 The Philosophy of Quantum Mechanics (New York : John Wiley).

Jauch J 1968 Foundations of Quantum Mechanics (Reading : Addison Wesely).

Jordan P, von Neuman J and Wigner E P 1934 Ann. Math. 35, 29.

Kac M 1951 Proc. of the Second Berkely Symposium on Mathematical Statistics and Probability (Berkeley : UCLA Press).

Kasday L 1971 in Proc. Enrico Fermi Summer School IL (New York : Academic Press).

Kocher C A and Commins E D 1967 Phys. Rev. Lett. 18, 575.

Konig H 1953 Math. Ann. 128, 420.

Kuryshkin V V 1977 in Price and Chissick (eds) p 61.

Lamehi-Rachti M and Mittig W 1977 in Leite-Lopes and Paty (eds).

Lanczos C 1924 Ann. Phys. 74, 518.

Landau L D and Lifshitz E M 1957 Quantum Mechanics; Non-Relativistic Theory (Oxford : Pergamon).

Lanz L 1977 in Price and Chissick (eds) p 187.

Leite-Lopes J and Paty M (eds) 1977 Quantum Mechanics, a Half Century Later (Dordrecht : D Reidel).

Lichnerowicz A 1955 Theories Relativistes de la Gravitation et de l'Electromagnetisme (Paris : Masson).

Lochak G 1977 in Leite-Lopes and Paty (eds) p 245.

MacGregor M H 1978 The Nature of the Elementary Particle (Berlin : Springer).

Mackey G W 1963 The Mathematical Foundations of Quantum Mechanics (New York : Benjamin).

Masani P 1973 in Tucker and Maynard (eds) p 217.

Mikusinski J 1959 Operational Calculus (London : Pergamon).

________ 1961 Studia Math. 20, 163.

________ 1966 Bull. de la Acad. Polon. Sci. Ser. Sci. Math. Astr. Phys. 14, 511.

Miller K S 1974 Complex Stochastic Processes.

Misner C W Wheeler J A and Thorne K S 1973 Gravitation (New York : John Wiley).

Morette C 1951 Phys. Rev. 81, 848.

Moyal J E 1949 Proc. Camb. Phil. Soc. 45, 99.

von Neuman J 1955 Mathematical Foundations of Quantum Mechanics (Princeton UP).

Nouri-Moghadam M and Taylor J G 1975 Proc. Roy. Soc. A344, 87.

Nelson E 1964 J. Math. Phys. 5, 332.

Papaliolos C 1967 Phys. Rev. Lett. 18, 622.

Papapetrou A and Hamoui A 1968 Ann. Inst. H. Poincare 9, 179.

——— 1979 Gen. Rel. Grav. (preprint).

Park J L and Margenau H 1968 Int. J. Theor. Phys. 1, 211.

Paty M 1977 in Leite-Lopes and Paty (eds) p 261.

Pearle P 1967 Amer. J. Phys. 36, 463.

Prasanna A R, Narlikar J V and Vishweshara C V (eds) 1980 Gravitation, Quanta and the Universe (New Delhi : Wiley Eastern).

Raju C K 1979 " Relativistic and Statistical Foundations of Quantum Mechanics " ISI, Delhi, discussion paper No.7802; revised version presented at the Einstein Centenary Symposium, Ahmedabad. Abstract in Prasanna et al (eds).

Raju C K 1980a "Classical Time-Symmetric Electrodynamics" to appear J. Phys. A.

________ 1980b "Products and Compositions with the Dirac Delta Function" ISI Tech. Report No. P and E/E/06-80, submitted J. Phys. A.

________ 1980c "Junction Conditions in General Relativity" ISI Tech. Rep. No. P and E/E/07-80, submitted Gen. Rel. Grav.

________ 1980d "Interpretation of the Indeterminacy Relations " ISI Tech. Rep. No. P and E/E/08-80, submitted Int. J. Theo. Phys.

________ 1980e "Interpretation of Quantum Mechanics as a Theory of Extended Particles" ISI Tech. Rep. No.P and E/E/09-80, submitted Int. J. Theo. Phys.

________ 1980f "A Relativistic Model of the Electron" ISI Tech. Rep. No. P and E/E/10-80, to be submitted J. Phys. A.

Robertson H P 1929 Phys. Rev. 34, 163.

Robinson A 1966 Non-Standard Analysis (Amsterdam : North Holland)

Rozanov Yu A 1967 Stationary Random Processes (San Francisco : Holden Day).

Rudin W 1974 Functional Analysis (New York : McGraw Hill).

Segal I E 1947 Ann. Math. 48, 930.

Schwartz L 1951 Theories des Distributiones (Paris : Hermann)

Skorohod A V 1974 Integration in Hilbert Space (Berlin Springer Verlag).

Stroyan K D and Luxemburg W A J 1976 Introduction to the theory of Infinitesimals (New York : Academic Press).

Synge J L 1966 Relativity : the General Theory (Amsterdam : North Holland).

________ 1965 Relativity : the Special Theory (Amsterdam : North Holland).

Tewari A K 1977 Math. Student 45, 55.

Thurber J K and Katz J 1974 in Hurd and Loeb (eds) p 272.

Tucker D H and Maynard H B (eds) 1973 Vector and Operator Valued Measures (New York : Academic Press).

Varadarajan V S 1968 *Geometry of Quantum Theory* (Princeton : van Nostrand).

Wiener N 1958 Nonlinear Problems in Random Theory (Cambridge : MIT Press).

Wiener N and Siegal A 1953 Phys. Rev. 91, 1551.

——— 1955 Nuovo Cim. Suppl. 2, 982.

Wigner E P 1932 Phys. Rev. 40, 749.

——— 1971 in Perspectives in Quantum Theory, p 25 (Yourgrau A and van der Merwe A eds).

Wilson A R, Lowe J and Butt D K 1976 J. Phys. G 2, 613.

Wu C S and Shaknov I 1950 Phys. Rev. 77, 136.

Yukawa H 1950 Phys. Rev. 77, 219 ; 80, 1047.

——— 1953 Phys. Rev. 91, 415.

——— 1964 Prog. Theor. Phys. 31, 1167.

PUNCH DISK NO. 2 FILE NAME 976 ,7

5871 (M4)

Physical A — Paper 976/1

Classical time-symmetric electrodynamics

Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics
Classical time-symmetric electrodynamics

C K Raju

C K Raju
C K Raju
C K Raju
C K Raju
C K Raju
C K Raju
C K Raju
C K Raju
C K Raju

Classical time-symmetric electrodynamics

C K Raju†

Indian Statistical Institute, 7 SJS Sansanwal Marg, New Delhi 110 029, India

Received 2 July 1979, in final form 10 March 1980

Abstract. A brief review of the classical aspects of the absorber theory of radiation is presented. Difficulties in the arguments used by earlier authors are discussed. The divergences which arise from the use of time-symmetric electrodynamics are pointed out. It is shown that the earlier difficulties can be removed by attributing differing signal velocities to advanced and retarded interactions. This difference in signal velocities is interpreted as arising from the extended, shell-like structure of charged particles. This leads to a new calculation of the absorber response. Absorption due to the time-symmetry normalisation factor is described. It is concluded that retarded radiation is approximately consistent in the Einstein–de Sitter model, whereas in the closed Friedman model, it is likely that retarded radiation is dominant during expansion, and advanced radiation during contraction. The theory predicts that advanced radiation exists in small amounts and can be detected experimentally.

† Present address: Physical and Earth Sciences Division, Indian Statistical Institute, 203 B T Road, Calcutta 700 035, India.

1. Introduction

The scalar wave equation in flat space

$$\Box^2\psi(r,t) = 4\pi\rho(r,t) \tag{1.1}$$

(where $\Box^2 = \nabla^2 - (\partial^2/\partial t^2)$ is the wave operator, $\psi(r,t)$ is the wave amplitude at the space–time point (r,t) and $\rho(r,t)$ is the source density at (r,t)) has two types of solution, (Davies 1974): $\psi_r = \psi(r,t^+)$ and $\psi_a = \psi(r,t^-)$ where

$$\psi(r,t) = \int_{R^3} (\rho(r',t')/R)\, d^3r' \qquad R = |r-r'| \qquad t^{\pm} = t \pm R. \tag{1.2}$$

These are known (with obvious notation) as the retarded and the advanced solutions, and represent waves propagating into the future and the past respectively. In curved space, with metric tensor $g^{\mu\nu}$, the corresponding retarded and advanced solutions for the scalar and vector wave equations

$$g^{\mu\nu}\psi_{;\mu\nu} = 0 \tag{1.3}$$

$$g^{\mu\nu}A^{\sigma}_{;\mu\nu} + R^{\sigma\alpha}A_{\alpha} = 0 \tag{1.4}$$

where $R^{\sigma\alpha}$ is the Ricci tensor, have been studied by De Witt and Brehme (1960).

Mathematically, any normalised combination

$$\psi = \theta\psi_a + (1-\theta)\psi_r \tag{1.5}$$

is also a solution, and the choice of the correct solution depends on the boundary conditions imposed. But, physically, boundary conditions cannot be imposed at will, and the existence of solutions with an advanced component would contradict our usual ideas of causality. The advanced solutions are rejected for this reason. However, it was pointed out by Dirac (1938) that this semi-empirical rejection may not be well founded, because a covariant derivation of the radiative damping force leads to the expression

$$\tfrac{2}{3}e^2(\dddot{z}^{\mu} + \dot{z}^{\mu}\ddot{z}^2) = e\dot{z}\tfrac{1}{2}(F^{\mu}_{r\nu} - F^{\mu}_{a\nu}) \tag{1.6}$$

(where e is the charge, z^{μ} the world line, $F^{\mu}_{r\nu}$, $F^{\mu}_{a\nu}$ the retarded and advanced fields of the particle and dots denote differentiation with respect to the parameter), apparently necessitating the use of advanced solutions.

Thus the problem is to specify the physical nature of the boundary conditions which give rise to the retarded solutions of experience, starting from solutions of the form (1.5). This problem is tackled in the absorber theory of radiation, initiated by Wheeler and Feynman (1945, 1949), and developed by Hogarth (1962), Hoyle and Narlikar (1964, 1969, 1971, 1972), Davies (1970, 1971, 1972a) and others. This direct particle interaction theory uses the Schwarzschild–Tetrode–Fokker action

$$J = -\sum_i m_i \int (\dot{z}^{\mu}_{(i)}\dot{z}_{(i)\mu})^{1/2}\, d\tau_i - \tfrac{1}{2}\sum_{i\neq j} e_i e_j \int \dot{z}^{\mu}_{(i)}\,\delta[(z^{\mu}_{(i)} - z^{\mu}_{(j)})(z_{(i)\mu} - z_{(j)\mu})]\,\dot{z}_{(j)\mu}\, d\tau_i\, d\tau_j \tag{1.7}$$

(where $z_i = z_i(\tau_i)$ is the world line of the ith particle with charge e_i and mass m_i, τ_i is the ith particle proper time and δ is the Dirac delta function).

In analogy with field theory, the last term of (1.7) can be used to define the four-potential

$$A_{\mu}(x) = \sum_j e_j \int \delta[(x^{\mu} - z^{\mu}_{(j)})(x_{\mu} - z_{(j)\mu})]\,\dot{z}_{(j)\mu}\, d\tau_j. \tag{1.8}$$

This four-potential satisfies the Lorentz condition $A^{\mu}{}_{,\mu} = 0$ and the electromagnetic wave equation

$$\Box^2 A_{\mu}(x) = 4\pi \sum_j e_j \int \delta(x - z_{(j)\mu})\,\dot{z}_{(j)\mu}\, d\tau_j \stackrel{\text{def}}{=} 4\pi j_{\mu}(x). \tag{1.9}$$

0643

0644 For consistency with (1.7), the unique solution of (1.9) is

0654

0654 $$A^\mu = \tfrac{1}{2}(A^\mu_r + A^\mu_a) \tag{1.10}$$

0656 where $A^\mu_r = A^\mu_*(t^-)$, $A^\mu_a = A^\mu_*(t^+)$ and

0659

$$A^\mu_*(t') = \int 4 j^\mu(r', t')/R \, \mathrm{d}^3 r'. \tag{1.11}$$

0666 If we re-introduce fields $F^\mu{}_\nu$ defined as usual by

0669

0670 $$F^\mu_{r\nu} = A^\mu_{r,\nu} - A^\nu_{r,\mu}$$

0671 $$F^\mu_{a\nu} = A^\mu_{a,\nu} - A^\nu_{a,\mu} \tag{1.12}$$

0672 then the electromagnetic fields obtained in this theory are given by

0681 $$F^{\mu\nu} = \tfrac{1}{2}F^{\mu\nu}_r + \tfrac{1}{2}F^{\mu\nu}_a. \tag{1.13}$$

0686 Hoyle and Narlikar (1964) have generalised this procedure to curved space, using the
0699 appropriate Green functions in the action (1.7). Equation (1.13) can also be obtained
0712 from the physical argument that Maxwell's equations, being time symmetric, should not
0724 by themselves impose an arrow of time on the solutions.

0734 The usual retarded fields are now obtained by adding to (1.13) the radiative
0747 damping term

0749 $$F_{rad} = \tfrac{1}{2}F_r - \tfrac{1}{2}F_a. \tag{1.14}$$

0752 (Tensor indices will be dropped henceforth as they are not necessary.) For a given
0766 particle, the term (1.14) was interpreted by Wheeler and Feynman (1945; WF) as the
0780 response of an ideal absorber to the elementary time-symmetric field of the charged
0793 particle. Thus the existence of purely retarded radiation gives a condition on the real
0807 universe, namely it should be totally absorbing (example—static, Euclidean universe).
0817 Hoyle and Narlikar (1964; HN) gave the differing condition that the universe should be
0831 opaque along the future null cone and transparent along the past null cone (example—
0844 steady-state model). In the following, shortcomings in the arguments of both WF and
0857 HN are pointed out, and it is shown that a more plausible condition is that the universe
0874 should be opaque along the past null cone, but not totally absorbing.

0886 One other possibility that needs to be mentioned here is the common presumption
0899 that purely retarded radiation exists in the universe may simply not be true. Attempts
0913 to confirm the presence or absence of advanced radiation experimentally have been
0925 made, and are continuing (Partridge 1973, Herron and Pegg 1974, Pegg 1975a, Davies
0938 1975). However, it is hardly possible to interpret the results of these experiments in the
0951 absence of a satisfactory theory. The present attempt to construct a satisfactory theory
0966 arises from an attempt to remove the difficulties faced by the earlier theories of WF and
0982 HN. Hence these difficulties will be considered first.

0990

2. Theory   Theoretical difficulties

0991

0992 *2.1. Self-consistency*

0993 The fundamental problem of time-symmetric electrodynamics (TSE) is to reconcile the
1004 following two facts; the action principle (1.7) permits only time-symmetric fields, while
1016 the usual fields of experience are, at least approximately, retarded. In WF it is proposed
1031 that this problem could be looked upon either from a general point of view or as a
1048 matter of explicit calculation. Only the general point of view is used in HN, and we
1064 consider this first.

1067 The total field acting on a charged particle i is given, according to (1.13), by

1082 $$F_{tot} = \tfrac{1}{2}\sum_{j\neq i} F^j_r + \tfrac{1}{2}\sum_{j\neq i} F^j_a \tag{2.1}$$

1085 where the field of the jth particle is indexed by j, and the summation ranges over all
1102 other particles j in the universe. On the other hand, to account for the observed fully
1118 retarded fields and radiation damping, the total field should be of the form

1131 $$F_{tot} = \sum_{j\neq i} F^j_r + \tfrac{1}{2}(F^i_r - F^i_a). \tag{2.2}$$

1133

1134

The problem of self-consistency is, then, to decide the circumstances under which (2.1) is consistent with (2.2). A necessary and sufficient condition for (2.1) to be consistent with (2.2) is, clearly,

$$\sum_{i} (F^i_r - F^i_a) = 0. \tag{2.3}$$

The general point of view, mentioned above, aims to show that (2.3) is valid under various plausible physical conditions. However, by subtracting (2.3) from (2.1)

$$F_{tot} = \sum_{j \neq i} F^j_a + \tfrac{1}{2}(F^i_a - F^i_r) \tag{2.4}$$

and, conversely, (2.1) and (2.4) together imply (2.3). It follows that (2.2) is consistent with (2.1) if and only if (2.4) is, i.e., retarded fields with radiative damping are consistent if and only if advanced fields with radiative anti-damping are simultaneously consistent.

Two explanations have been offered for this apparently paradoxical situation. According to WF, the particles on the past null cone of i may be assumed to be in a state of random motion, i.e., their motion is uncorrelated with the motion of particle i. Hence, $\Sigma_{j\neq i} F^j_r$ is small compared with the radiative damping term. On the other hand, the fields F^j_a are highly correlated with the motion of i, and it may be imagined that $\Sigma_{j\neq i} F^j_a = -(F^i_a - F^i_r)$. According to HN the retarded field is attentuated; hence, $\Sigma_{j\neq i} F^j_r$ is small compared with $\Sigma_{j\neq i} F^j_a$. Both these arguments appear to be unacceptable—the first because the assumption of random motion, when signals can be propagated along the past null cone, appears to be unrealistic, and the second because it involves an application of the refractive index only to response fields.

Moreover, it is easily seen that these arguments make sense only after the world lines have been prescribed according to the principle of retarded causality. Thus, the question arises: 'Why should the world lines be determined in this manner?' Since the action (1.7) determines both the fields as well as the world lines, it has to be shown that the world lines obtained by using purely retarded fields are identical with the world lines given by the action principle $\delta J = 0$. In particular, given that $\ddot{z}_i(\tau_{i0}) \neq 0$, we have to show that the solution to the constrained problem

$$\delta J = 0 \qquad \ddot{z}_j(\tau^-_j(\tau_{i0})) = 0 \qquad \ddot{z}_i(\tau_{i0}) \neq 0 \tag{2.5}$$

(where $\tau^-_j(\tau_{i0})$ corresponds to the value of τ_j at which the past null cone at $z_i(\tau_{i0})$ meets z_j) is also a solution to the 'unconstrained' problem

$$\delta J = 0 \qquad \ddot{z}_i(\tau_{i0}) \neq 0. \tag{2.6}$$

Neither WF nor HN have demonstrated this, and it is not clear how this can be possible without having all the Lagrangian multipliers equal to zero. Since the Lagrangian multiplier corresponding to a constraint can be interpreted as the sensitivity to that constraint, it would follow that the real assumption is not just that accelerations along the past null cone are zero, but that the solution to the variational problem (2.6) remains unaffected by small accelerations along the past null cone. In this context the explanation due to HN might seem more appealing; however, this explanation does not appear to be convincing to the author and some others (for instance, Davies 1978, private communication), as it allows the refractive index to distinguish between stimulus and response fields.

The other method proposed in WF was that of explicit calculation. Here the problem is considered as follows. Suppose particle i is disturbed (non-electromagnetically) and radiates the time-symmetric field

$$F^i_{particle} = \tfrac{1}{2}(F^i_r + F^i_a). \tag{2.7}$$

This field, in the course of propagation, disturbs other particles in the universe, which, in turn radiate time-symmetric fields. It is required to show, by explicit calculation, that these elementary response fields up to produce the absorber response field

$$F^i_{response} = \tfrac{1}{2}(F^i_r - F^i_a). \tag{2.8}$$

The total field attributed to particle i is then

$$F^i_{tot} = F^i_{particle} + F^i_{response} = F^i_r. \tag{2.9}$$

In the calculations given in WF (derivations I–III) to calculate the absorber response the fully retarded field of the particle i is used to arrive at the expression (2.8). However, to calculate the total field attributed to particle i, the expression (2.7) is used for the field of the particle. Thus the 'cycle of reasoning' used in WF is potentially circular unless

$$F_r^i - \tfrac{1}{2}(F_r^i + F_a^i) = \tfrac{1}{2}(F_r^i - F_a^i) = 0, \tag{2.10}$$

i.e., unless radiative damping vanishes.

In an attempt to show that the argument is not circular, in WF derivation II, the total outgoing disturbance is denoted by (?) F_r^i, and equation (20) of WF reads

total disturbance diverging from source	= proper retarded field of source itself	+ field apparently diverging from source actually composed of parts converging on individual absorber particles.

(2.11)

The method used by WF is such that results for the advanced field can be obtained by replacing 'retarded' everywhere by 'advanced', and 'diverging' by 'converging'. Hence, if it is not assumed, *a priori*, that (?) = 1, we also have

total disturbance converging on source	= proper advanced field of source itself	+ response field of past absorber apparently converging on source	+ response field of future absorber apparently converging on source

(2.12)

leading to

$$(1 - ?) = (\tfrac{1}{2}) + \tfrac{1}{2}(1 - ?) + \tfrac{1}{2}(?), \tag{2.13}$$

which implies (?) = 0. Thus, the argument in WF is circular, unless radiative damping vanishes.

2.2. *The divergences of* TSE

The source of the above inconsistencies can be traced to the fact that explicit calculation, even in the two-particle case, leads to divergences, and these divergences are bound to persist in the n-particle case. Thus, consider two charged particles i and j with charges e_i and e_j and masses m_i and m_j. Suppose a disturbance acts on particle i giving it a non-relativistic acceleration

$$\dot{v}_i(t) = A\,e^{-i\omega t} \tag{2.14}$$

where A is the amplitude and ω is the periodicity of the disturbance.

As a result of this acceleration, particle i radiates the time-symmetric fields $\tfrac{1}{2}(F_r^i + F_a^i)$. The field F_r^i interacts with the particle j at a later time giving it an acceleration

$$\dot{v}_j^r(t) = \frac{1}{2}\frac{e_i e_j}{m_j} A \sin(\dot{v}_j, R_r) \exp[-i\omega(t - R_r)e] \tag{2.15}$$

where R_r is the interparticle separation in the retarded case and e is the unit electric polarisation vector, the direction of which is taken to be negative if it has a positive component along $\dot{v}_j(t)$.

The corresponding advanced field of j interacts with i *simultaneously* with the original disturbance to produce an additional acceleration

$$\Delta_a^r \dot{v}_i(t) = \frac{1}{2}\frac{e_i^2 e_j^2}{m_i m_j} \sin^2(\dot{v}_i, R_r) A\,e^{-i\omega t}. \tag{2.16}$$

Similarly, the advanced/retarded interaction between i and j leads to an additional acceleration

$$\Delta_r^a \dot{v}_i(t) = \frac{1}{2}\,\frac{e_i^2 e_j^2}{m_i m_j}\sin^2(\dot{v}_i, R_a) A\, e^{-i\omega t} \tag{2.17}$$

where R_a is the interparticle separation in the advanced case.

Thus, starting from the assumption that the acceleration of i at time t is $\dot{v}_i(t)$, we have reached the conclusion that the acceleration of i at a time infinitesimally later than t is $\dot{v}_i(t) + \Delta_a^r \dot{v}_i(t) + \Delta_r^a \dot{v}_i(t)$. It is useless to try to sum up all the changes arising as a result of these stimulus and response fields because, if as a result we arrive at the value $\dot{v}_i'(t) \neq 0$ for the acceleration of i, the above reasoning goes through with $\dot{v}_i'(t)$ in place of $\dot{v}_i(t)$. In fact, singolarities are present at all points along the world lines of i and j where the interaction actually takes place. Thus, as a result of the interaction between i and j, the field at any point of space–time due to an arbitrary (non-zero) initial acceleration of i is indeterminate.

Schulman (1974) encountered a similar problem with respect to the differential equation

$$\ddot{x}(t) + \omega^2 x(t) = \tfrac{1}{2}\alpha x(t-\tau) + \tfrac{1}{2}\beta x(t+\sigma) + \phi(t) \tag{2.18}$$

(where $\alpha, \beta, \tau, \sigma$ are constants, $\tau, \sigma > 0$, and $\phi(t)$ is a given function), and has suggested the use of boundary conditions to make such an equation tractable. In fact, the initial-value problem for such equations remains unsolved. Schulman has also suggested that a difference in the values of τ and σ could, perhaps, account for the suppression of advanced interactions.

Wheeler and Feynman (1949) have attempted to resolve this 'paradox' by saying that impulsive forces do not exist in nature. However, if we were to accept this point of view then the action (1.7) must be abandoned, because it assures the existence of just such impulsive forces. In this paper a somewhat similar point of view is adopted, and this is considered in greater detail below. In particular, time-symmetric fields will be dealt with, without reference to any action principle.

3. Signal velocity for advanced radiation

The roots of the above paradox lie in the assumption that a charged particle is a point charge which responds instantaneously to any incident radiation. Now electromagnetic, gravitational and quantum-mechanical considerations (Dirac 1938, 1962a, b, Raju 1979) indicate that the charges we consider must be distrihuted over a finite region. However, considerations of finiteness alone are not sufficient to remove the ambiguities noted above. The hypothesis of extended charges implies that the velocity with which the actual interaction between two charged particles takes place (signal velocity) is not the same as the wave velocity. Now, in the retarded case the signal velocity is lower than the wave velocity, and, as Kamat (1970) has suggested, this is true in the advanced case as well. As a result, a larger time interval is required for the actual interaction to take place in both cases. Naturally, in the advanced case this time interval is measured in the backward direction. Hence, the interaction with signal velocity takes place *earlier* than the interaction with wave velocity and not later, as suggested by Kamar (1970). By symmetry, this 'advance' in the advanced case just compensates the usual delay in the retarded case, and this brings us back to the situation in § 2.2.

However, let us consider instead a model of an extended charged particle which interacts at its boundary/Shell-like models of this type have been proposed by Dirac (1962a) and Raju (1979). The boundary of the particle in these models may be considered to be rigid in the sense that spherical symmetry is maintained, or in the sense (of Dirac 1962b) that signals can travel instantaneously in the interior of the particle. In this situation the time interval for preliminary interaction in the advanced case is fractionally longer than the corresponding time interval in the retarded case (see figure 1). It follows that, in all cases, the time interval for an actual interaction is longer in the advanced case than in the retarded case. There is, therefore, a systematic bias ensuring that the signal velocity for advanced interaction is smaller than the signal velocity for retarded interaction, i.e., the time interval in which the actual advanced interaction takes place is longer than the time interval in which the actual retarded interaction takes place. The relevant interaction diagram would be as in figure 2 and not as in figure 1(b) of Schulman (1974).

Thus the notion of an extended charged particle as a (rigid) shell leads to a natural physical justification of the hypothesis of lowered signal velocity for advanced radiation. The notion of an extended particle was also used by Dirac (1938) to account for pre-acceleration. However, Dirac's explanation of pre-acceleration is not fully satisfactory because, for the one-dimensional equations of motion,

$$m(\dot{w} - \alpha \ddot{w}) = f(\tau) \tag{3.1}$$

(where m is the mass, $\alpha = \frac{2}{3}e^2/m$, e is the charge, $w(\tau) = \sinh^{-1}(\dot{z}(\tau))$, $z(\tau)$ is the world line of the particle and $f(\tau)$ is the force due to the external field), Dirac (1938) proposed the special value of the initial acceleration, given by

$$\dot{w}(0) = \frac{1}{m\alpha}\int_0^\infty e^{-\tau/\alpha} f(\tau)\, d\tau. \tag{3.2}$$

It is clear that in response to an impulse at $\tau = 0$,

$$f(\tau) = K\delta(\tau), \tag{3.3}$$

the particle acceleration, given by

$$\dot{w}(\tau) = \frac{1}{\alpha} K\, e^{\tau/\alpha} \qquad \tau < 0$$

$$= 0 \qquad \tau > 0, \tag{3.4}$$

is non-zero (though small) for large negative values of τ. Dirac (1938) attempted to explain this by saying that the classical electron has no boundaries and can, therefore, 'feel' the impulse at large times before it is actually applied. This explanation no longer appears reasonable in the context of a definite model for the electron that has a definite size or some size smaller than a definite size.

There is an alternative explanation in the present framework, because the charged particle radiates and interacts with other charged particles if an impulse is applied, at $\tau = 0$, say. Due to the lowered signal velocity of advanced radiation, the effects of this impulse are propagated (in a rapidly decaying manner, if retarded radiation is approximately consistent) along the entire portion of the world line for $\tau \leq 0$. Thus it is not the indefinite size of the electron but the lowered signal velocity for advanced radiation which accounts for pre-acceleration at large times before an impulse is applied.

Finally, we note that there is no serious disadvantage associated with the fact that the hypothesis we are making apparently does not have an immediate generalisation to the quantum-mechanical case (see, however, Raju 1979). This is true, if only because any really successful theory of quantum mechanics will incorporate relativistic considerations, and be able to treat particles with an extended structure.

4. The absorber response

We consider only the case of an isolated charged particle which is disturbed by some non-electromagnetic force and radiates time-symmetric fields. As pointed out by Hogarth (1962), Sciama (1963) and Hoyle and Narlikar (1964), the effective interaction with the absorber takes place across cosmological distances. At these distances, the part of the universe (the past absorber) which interacts with the advanced component of the field of the charged particle must be considered to be physically distinct from the part (the future absorber) which interacts with the retarded component. The situation is further simplified by the fact that the charged particle can receive only retarded radiation from the past absorber, and only advanced radiation from the future absorber.

It will be understood in the following that we are dealing with plane-polarised fields of a fixed frequency ω. Since the theory is linear and since the results are indepdent of ω, the results hold by Fourier superposition for more complicated fields. Following figure 3, we let S_i^P and S_i^F, be the various stimulus fields for the past absorber and the future absorber, while R_i^P and R_i^F denote the corresponding response fields. Because the difference in time intervals for retarded and advanced interactions depends only the (average) size of the charged particles, and not on the interparticle separation, all the absorber particles can be lumped together in the interaction diagram. However, the multiparticle nature of real absorbers might affect another assumption, that of linearity, which, in this context, refers to spherical symmetry (Hogarth 1962). Such possible departures from linearity will not be considered here and we assume that all the response fields are linearly related to the stimulus fields, the response factors being given by p and f for the past and future absorbers respectively.

Denoting the advanced and retarded source fields by F_a and F_r, we have

$$\begin{aligned} S_1^P &= \tfrac{1}{2}F_a = \tfrac{1}{2}(F_a - F_r) \qquad & S_1^F &= \tfrac{1}{2}F_r = \tfrac{1}{2}(F_r - F_a) \\ R_1^P &= \tfrac{1}{2}p(F_r - F_a) & R_1^F &= \tfrac{1}{2}f(F_a - F_r) \end{aligned} \tag{4.1}$$

where the first line of (4.1) is obtained by noting that the field F_r is zero in the region of the past absorber and F_a is zero in the region of the future absorber. In the last line of (4.1) we have temporarily discarded any possible difference between a converging retarded field and an advanced field. However, this does not require any new assumption since we have already assumed the particle to be isolated.

The other stimulus and response fields are given by

$$\begin{aligned} S_2^P &= R_1^F \qquad & S_2^F &= R_1^P \qquad & R_i^P &= pS_i^P \\ S_n^P &= R_{n-1}^F & S_n^F &= R_{n-1}^P & R_i^F &= fS_i^F. \end{aligned} \tag{4.2}$$

We now assume that, from the point of view of macroscopic observation, all the response fields act simultaneously with the initial source fields. As a particular consequence of this assumption, the resulting electromagnetic arrow of time, like the thermodynamic arrow of time, will be applicable only at a macroscopic level.

With the above assumption, the total field

$$F_{tot} = \tfrac{1}{2}F_r + \tfrac{1}{2}F_a + \sum_{i=1}^{\infty} R_i^P + \sum_{i=1}^{\infty} R_i^F. \tag{4.3}$$

Now, ignoring those terms which cancel with $\sum_{i=1}^{\infty} R_i^F$, and assuming $|pf| \neq 1$,

$$\begin{aligned} \sum_{i=1}^{\infty} R_i^P &= [p + pfp + p(fp)^2 + \ldots]\tfrac{1}{2}(F_r - F_a) \\ &= \left(p \sum_{n=0}^{\infty} (pf)^n\right)\tfrac{1}{2}(F_r - F_a) \\ &= \left(\frac{p}{1-pf}\right)\tfrac{1}{2}(F_r - F_a). \end{aligned} \tag{4.4}$$

Similarly,

$$\sum_{i=1}^{\infty} R_i^F = \left(\frac{f}{1-pf}\right)\tfrac{1}{2}(F_a - F_r). \tag{4.5}$$

From (4.4) and (4.5)

$$F_{tot} = \frac{1}{2}\left(1 + \frac{p}{1-pf} - \frac{f}{1-pf}\right)F_r + \frac{1}{2}\left(1 + \frac{f}{1-pf} - \frac{p}{1-pf}\right)F_a. \tag{4.6}$$

Hence the condition for $F_{tot} = F_r$ is

$$(p - f)/(1 - pf) = 1 \tag{4.7}$$

or

$$p(1 + f) = (1 + f). \tag{4.8}$$

Equation (4.8) is valid if either $p = 1$ with f arbitrary, or $f = -1$ regardless of the value of p, given that $|pf| \neq 1$. Similarly, we have for the existence of purely advanced radiation the conditions $p = -1$ or $f = 1$, $|pf| \neq 1$.

If we know *a priori* that $p = f = 1$ then (4.7) is indeterminate. However, in this case $R_n^P + R_n^F = 0$ for each n, i.e., the response fields cancel termwise. Hence, in the expression (4.3) for F_{tot}, we are left only with the original time-symmetric fields of the particle. Thus, in the case $p = f = 1$ the nature of the radiation is time symmetric. On the other hand, if we only have $p \approx 1$, $f \approx 1$ then a mixture of the form $(1 - \delta)F_r + \delta F_a$ can exist, with

$$\delta = \frac{(1-p)-\frac{1}{2}(1-p)(1-f)}{(1-p)+(1-f)-(1-p)(1-f)} \simeq \frac{(1-p)}{(1-f)} \ll 1 \qquad (4.9)$$

provided $(1-p) \ll (1-f)$†.

† This is *not* the same as saying $f \ll p$. For instance, if $p = 1 - 10^{-20}$ and $f = 1 - 10^{-16}$, then $(1-p) \ll (1-f)$ is true, but not $f \ll p$ in the usual sense.

Strictly speaking, since any measurement will involve only a finite portion of the world line of the particle, the observed radiation will always be a mixture, and the empirical testability of this prediction is considered in § 6.

5. The absorber mechanism

It is now necessary to calculate the values of p and f for various cosmological models. The main interest, of course, lies in finding those cases for which $p = 1, f \neq 1$. We first consider qualitatively the mechanism underlying the absorber response. This has never, in any case, been clearly brought out. Thus, for instance, Kamat (1970) questions the assumption that this response is emitted in a 'backwards' direction, and feels that the existence of a 'backwards refractive index' is necessary for the existence of such a response. On the other hand, Pegg (1975b) has maintained that the refractive index arises after consideration of the absorber response, and has quoted the derivations in WF in support of his contention. However, in these derivations the interaction of the advanced component of the elementary response fields with the charged particles in the absorber has not been considered at all, implying that a special choice for the refractive index for advanced radiation has already been made. As such, the origin of the backwards refractive index remains a mystery.

This mystery is, however, easily resolved by considering the phase relations for the elementary response fields. Consider a retarded stimulus field, incident on an absorber which we assume to be a dilute plasma of charged particles. For the purpose of understanding the origin of the backwards response field it is sufficient to consider the case of a plane monochromatic field, which we denote at a plane z in the absorber by $F_z = F_0 e^{i\omega t}$. The phase of the field at the position $z + \delta z_i$ of particles i (see figure 4) is $\omega(t - \delta z_i)$. We assume that the damping is very light and the refractive index is approximately unity, so that the phase of the re-emitted retarded field of particle i at $z + \delta z$ in the direction θ is

$$\omega(t - \delta z_i) - \omega[(\delta z - \delta z_i)/\cos\theta].$$

Thus in the forward direction ($\theta = 0$) the re-emitted retarded fields interfere constructively and are in phase with the original retarded field. In any other direction the interference is destructive because of the random positioning of the particles.

But, for the advanced field the phase at z is $\omega(t - \delta z_i) + \omega\delta z_i \pm \pi$, whereas the phase at $z + \delta z$ is $\omega(t - \delta z_i) - \omega[(\delta z + \delta z_i)/\cos\theta]$. It follows that the re-emitted advanced fields interfere constructively in the backwards direction and destructively elsewhere. Moreover, they are exactly π out of phase with the incident retarded field. The additional phase difference of π appears because the re-emitted retarded fields are in phase with the incident field whereas the re-emitted advanced fields are π out of phase with the re-emitted retarded fields.

This explains the origin of the backward response field, but now there is a new mechanism of absorption at work in the time-symmetry normalisation factor $\frac{1}{2}$. Assuming an extinction theorem of the Ewald–Oseen type (see, for instance, Born and Wolf 1964), only half of the absorbed energy is re-emitted in the forward direction, while the other half is converted into the energy of the absorber response. Thus, it is time-symmetric scattering rather than thermal absorption which gives rise to the absorber response. Thermal absorption, if it occurs, will permanently remove the energy available for the absorber response. It should be observed that secondary and tertiary scattering also contribute to the absorber response, because in each case the response field exactly retraces the path of the incident stimulus field. In actuality, an infinity of interactions might take place within the absorber, but we do not have to consider this, since we can get the magnitude of the response field by straightforward energy considerations. Thus, if there are sufficiently many charged particles in the absorber so that the incident field loses all its energy (by scattering) then all this lost energy must appear in the form of the absorber response, i.e., the absorber is ideal. In this case any significant motion of the charged particles in the absorber is confined to very small time intervals and consequently there is very little thermal absorption.

Applying these considerations to the past and future absorbers we see that $0 \leq p, f \leq 1$ and $p = 1$ ($f = 1$) provided the past (future) absorber has a sufficiently large number of charged particles.

This brings us to the problem of determining those cosmological models that satisfy the conditions for the existence of retarded radiation. Here we will consider the situation only with respect to evolutionary models. The past null cone in this case is opaque ($p \simeq 1$) due to the initial singularity. Along the future null cone we assume that absorption occurs mainly by discrete objects, which absorb their geometric cross section of the incident flux. In that case, models which expand with $R \propto t^{2/5}$ or more are transparent along the future null cone (Davies 1972b). These models, which include the Einstein–de Sitter model, therefore satisfy the requirements for the existence of retarded radiation.

The situation is not so clear with regard to models with a final singularity, because, as seen in § 3, retarded radiation can be consistent even if $p \simeq 1$ and $f \simeq 1$. In particular, the possibility raised by Gold (1967) is not ruled out. For a source at large times from the initial singularity, some of the emitted radiation would undergo large blue shifts and correspondingly greater losses due to (non-linear) thermal absorption and pair production. Hence $(1-p)$ would increase as the source moves away from the initial singularity. In this manner it is possible that p ultimately falls below the value of f. Similarly, due to the peculiarities of absorption in an epoch dominated by advanced radiation, $(1-f)$ could decrease towards the final singularity.

Thus there is a good possibility that, at least in this case, the cosmological arrow of time determines the electromagnetic arrow of time, although a deeper investigation would be required before drawing any firm conclusion.

6. Empirical detection of advanced radiation

In § 4, the conditions for the existence of purely retarded radiation were derived under the assumption that all the stimulus and response fields act simultaneously at any point in space. In actuality, according to the basic hypothesis, the nth response field acts τ seconds earlier than the $(n-1)$th response field, where τ is the characteristic delay associated with the signal velocity of advanced radiation. The precise value of τ would depend on the particular model under consideration. However, for any realistic model of a particle of finite size, τ would be very small. In fact, if the model proposed by Dirac (1962) and Raju (1979) is used, the value of τ would definitely be smaller than 10^{-23} s. A macroscopic observer would carry out an observation only in a finite time interval, T, large compared with τ. A large number, $N = (T/\tau)$, of the response fields ~~compared~~ would act within this time interval, and so the assumption of simultaneity would be justified *provided* the series converges with sufficient rapidity. Since retarded radiation is approximately consistent, i.e., $p \simeq 1$, the series would converge rapidly, unless $f \to 1$. But, by using increasingly efficient local absorbers, in theory it can always be arranged to have $f \to 1$ in the laboratory. Hence, if the theory is correct, advanced radiation can be detected experimentally

Since one experiment of this nature has already been carried out (Partridge 1973) and another proposed (Herron and Pegg 1974), it would be worthwhile discussing these theoretical predictions in this context. Partridge's experiment consisted of measuring the power input to a horn antenna as it radiated alternately into free space and a local absorber. Partridge assumed a relationship of the type

$$P_f = (1-\delta)P_a \qquad (6.1)$$

where P_f and P_a denote the power inputs while radiating into free space and the local absorber respectively. According to the theories in WF and HN, $\delta \geq 0$. According to the present theory, the presence of a local absorber would increase the content of advanced radiation in the mixture, and hence δ should be negative. The mean value of δ obtained by Partridge was $(-1 \cdot 1 \pm 1 \cdot 6) \times 10^{-9}$, and Partridge concluded that this was not significant. However, in obtaining this mean value Partridge took a weighted average over various phase settings of a phase-sensitive detector, and, for the two phase settings $\phi = 0°$ and $\phi = 180°$ for which the detector is most sensitive, the values of δ obtained were significant and negative. Although the possibility of instrumental error can by no means be ruled out, these results are certainly suggestive, and the experiment deserves to be repeated with greater sensitivity.

In the proposed experiment by Herron and Pegg, a 'dynamic' local absorber is to be used, which, it is supposed, would affect only the value of p or that of f. If only the value of p is sought to be increased, then if $p = 1$ there would be no change in the power input, whereas if $p = 1, p \neq 1$, then the power input would increase.

7. Conclusion

Retarded radiation is consistent in those models which satisfy $p = 1$ or $f = -1$, with $|pf| \neq 1$. The Einstein–de Sitter model satisfies these conditions approximately. In the closed Friedman model it is likely that retarded radiation is dominant during expansion and advanced radiation during contraction.

The theory predicts that advanced radiation exists in small amounts, and can be detected experimentally.

Acknowledgments

The author would like to thank Professor P Badyopadhyay, Dr P C W Davies and Professor J V Narlikar for discussions and helpful comments.

References

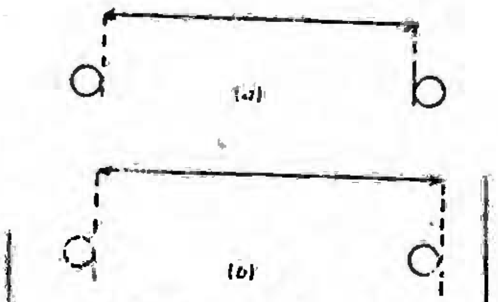
Born M and Wolf E 1964 *Principles of Optics* (Oxford: Pergamon)
Davies P C W 1970 *Proc. Camb. Phil. Soc* 68 751–64
—— 1971 *J. Phys. A: Gen. Phys.* 4 836–45
—— 1972a *J. Phys. A: Gen. Phys.* 5 1025–36
—— 1972b *J. Phys. A: Gen. Phys.* 5 1722–37
—— 1974 *The Physics of Time Asymmetry* (Berkeley: University of California Press)
—— 1975 *J. Phys. A: Math. Gen.* 8 272–80
De Witt B S and Brehme R W 1960 *Ann. Phys., NY* 9 220
Dirac P A M 1938 *Proc. R. Soc.* A 167 148 148–69
—— 1962a *Proc. R. Soc.* A 268 57–67
—— 1962b *Proc. R. Soc.* A
Gold T 1967 *The Nature of Time* (New York: Cornell University Press)
Heron M L and Pegg D T 1974 *J. Phys. A: Math., Nucl. Gen.* 7 1965–9
Hogarth J E 1962 *Proc. R. Soc.* A 267 365–83
Hoyle F and Narlikar J V 1964 *Proc. R. Soc.* A 277 1–23
—— 1969 *Ann. Phys., NY* 54 207–39
—— 1971 *Ann. Phys., NY* 62 44–97
—— 1972 *Nuovo Cim.* 7 242–61
Kamat R V 1970 *J. Phys. A: Gen. Phys.* 3 473–80
Partridge R B 1973 *Nature* 244 263–5
Pegg D T 1975a *J. Phys. A: Math. Gen.* 8 L60
—— 1975b *Rep. Prog. Phys.* 38 1339–83
Raju C K 1979 in *Gravitation, Quanta and the Universe* ed A R Prasanna (New Delhi: Wiley Eastern)
Schulman L S 1974 *J. Math. Phys.* 15 295–8
Sciama W 1963 *Proc. R. Soc.* A 273 484–95
Wheeler J A and Feynman R P 1945 *Rev. Mod. Phys.* 17 157–81
—— 1949 *Rev. Mod. Phys.* 21 425–33

Figure 1. Difference in the time taken for preliminary interaction: (a) retarded interaction; (b) advanced interaction

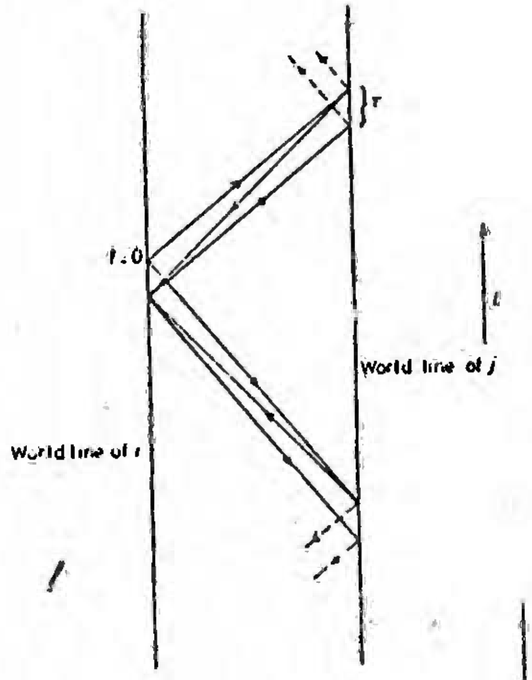
Figure 2. Interaction diagram for two particles, illustrating the effect of lowered signal velocity for [illegible] radiation. The time difference, τ, is grossly exaggerated.

Figure 3. Interaction diagram for an 'isolated' particle.

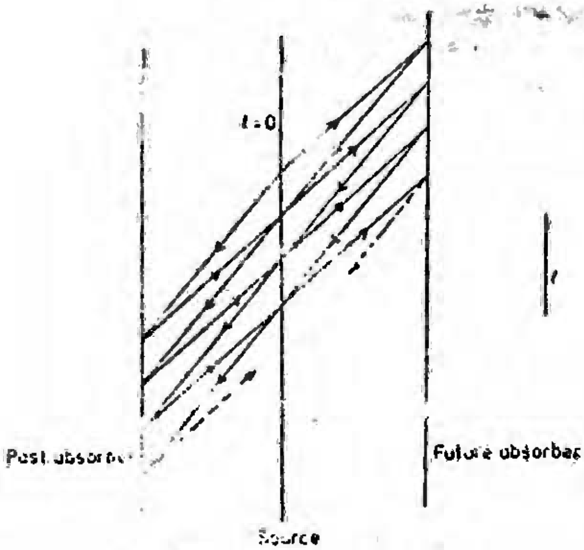
Figure 4. Origin of the backward response field: the re-emitted advanced fields interfere constructively in the backward direction and destructively elsewhere.



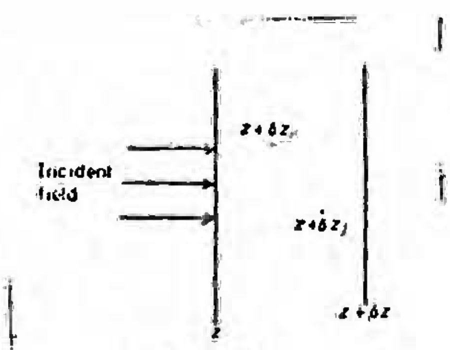
PHYS "A" 946/1 ½



PHYS "A" 946/2 ½



PHYS "A" 946/3 ½



PHYS "A" 946/4 ½