PARTICLES, FIELDS, AND RIGID BODIES IN THE FORMULATION OF RELATIVITY THEORIES

Thesis submitted for the Degree of

DOCTOR OF PHILOSOPHY

in the

UNIVERSITY OF LONDON

by

J.E. HOGARTH, B.A.Sc., M.A. (Toronto)

Department of Mathematics,
Royal Holloway College,
Englefield Green,
Surrey.

October, 1953.

Acknowledgements

I wish to record my thanks to my supervisor, Prof. W.H. McCrea for his untiring assistance and advice, to the Defence Research Board of Canada, whose generosity made it possible for me to carry out this work, and to the staff and students of the Royal Military College of Science, for their unreserved encouragement. I am also grateful to Miss J. Wilce, who typed the manuscript.

Particles, Fields, and Rigid Bodies in the Formulation of Relativity Theories

(Abstract of Thesis by J. E. Hogarth)

The particle theories of Wheeler and Feynman (electrodynamics) and Whitehead (gravitation) are studied, and the relationships to their field counterparts are examined. An invariant distance and suitable covariant potentials are defined in Riemannian space-time, and by this means the theories are generalized, in particular to the de Sitter space-time of constant curvature. It is shown that the generalization has an interesting significance with respect to the steady-state theory of cosmology.

The electrodynamic generalization consists of finding a covariant vector potential in de Sitter space from which Maxwell's equations can be derived. It is shown that in the flat-space theory of Wheeler and Feynman radiation damping is indeterminate, but that in de Sitter space and in conjunction with steady-state cosmology the irreversibility of radiation is closely related to the phenomenon of the creation of matter.

In de Sitter space Whitehead's theory is shown to yield the Schwarzschild solution of the general theory of relativity with a cosmological constant. A close relationship between Whitehead's theory and the general theory of relativity is suggested. Whitehead's theory is shown to be compatible with steady state cosmology.

Also included in the thesis is a reprint of "The Relativistic Rigid Rod", a paper written jointly with Prof. W.H. McCrea. This paper is presented in the thesis as supporting material.

C O N T E N T S

CHAPTER I

GENERAL INTRODUCTION

Physical theories can generally be divided into two classes, one based primarily on the concept of a field, and the other on that of a particle. Few theories are completely free of either concept, but usually one or other predominates. Pure field theorists will describe a particle as a singularity in the field or as a quantization of a field variable. On the other hand those preferring to think of particles as the only fundamental physical entity will describe a field as being a secondary quantity derived from the behaviour of particles and with no degrees of freedom of its own.

A classical field theory consists of a set of fundamental partial differential equations to be satisfied by the field variables, while the basic laws of a particle theory simply relate the motions of particles. Although the latter approach necessitates the concept of action at a distance, it is usual to define adjunct fields in terms of the particles and their motions, and to determine from the basic laws the partial differential equations satisfied by these fields. The pure field and the adjunct-field equations may be similar or even identical in structure, but the differences that exist between the two kinds of theory need not simply be superfluous ones of fancy. Possible real differences that seem to favour particle or adjunct field theories are the following:

(a) Pure classical field theories invariable encounter difficulties with respect to the structure and self action of particles (or particle like entities). Such difficulties are generally excluded from particle theories at the outset.

(b) Adjunct-field equations are usually derived from the laws of a particle theory by a process of partial differentiation. This means that the number of arbitrary boundary conditions which must be imposed in the solution of a particular problem is fewer than for the pure field counterpart.

(c) Although field theories must eventually admit the concept of a particle in some form or other, solutions to the field equations may exist which bear no relationship to these particles. Such solutions are excluded from an adjunct field theory.

The first successful field theory was that of Maxwell, and a half a century later the particle mechanics of Newton was superceded, in principle though not in practice, by Einstein's field theory of General Relativity. It was in this setting of field physics that quantum theory was born.

Less well known than the works of Maxwell and Einstein are the electrodynamics of Schwarzschild (1903), Tetrode (1922) and Fokker (1929, 1932) and the gravitational theory of A.N. Whitehead (1922), both of which are based on the fundamental concept of a particle and use a Lorentz invariant variational principle to relate the motions of particles. Both claim at least equal status with their field counterparts in the description of nature.

More recently the electrodynamics* of Schwarzschild, Tetrode and Fokker has been examined and further developed by Wheeler and

---

* Hereafter this theory will be referred to as Wheeler and Feynman electrodynamics.

Feynman,[1][2], while an almost forgotten interest in Whitehead's Theory of Relativity[3] has been revived by the work of J.L. Synge.[4][5]

Although almost universally successful as a classical field theory, Maxwellian electrodynamics has failed to give either a completely satisfactory model for an electron or a suitable description of the mechanism of radiation. In the theory of Lorentz, an electron is of finite size and radiates when subjected to a non-uniform acceleration because of the mutual interaction of its parts. This theory is unable to explain the stability of such an explosive self-acting structure.

One modern view with respect to particles is that they have no place in a classical theory. Quantum electrodynamics, however, has not yet proved itself capable of providing a successful electron theory.

Wheeler and Feynman electrodynamics denies self action and attempts to explain radiation in an entirely different way. Here Maxwell's equations are derivable adjunct-field equations. The unique feature of the theory is contained in the fact that the Maxwellian field of a particle is no longer the usual retarded field but a combination of advanced and retarded fields. (However contrary to our notions of cause and effect, the introduction of advanced effects presents no difficulties in a deterministic theory.) Maxwell's equations do not deny the existence of advanced fields; they are usually eliminated in the classical theory by means of appropriate boundary conditions. No such boundary conditions are necessary in a particle theory, the nature of the field being completely prescribed by the theory. In Feynman and Wheeler electrodynamics radiation and radiation damping result from advanced fields.

The difficulties associated with General Relativity are of a more general nature. A particle is described as a singularity in the metric field and (at least) for a single particle at rest the stress-energy-momentum tensor field vanishes at all points not on the world line of the particle. Although no such ideal particle has been found to exist, the Schwarzschild solution satisfies the required conditions except on a narrow tube surrounding the world line of the particle. The initial tremendous success of the theory was due in large part to the existence of this solution of the field equations. After nearly forty years since its conception, General Relativity still fails to throw much light on as simple a configuration as that of two particles. Certainly no solution has been found for the two body problem for which the field equations are satisfied except along two world lines.

The difficulties are twofold. Firstly General Relativity may be described as a closed theory. The metrical structure of the continuum depends on the distribution of mass, but the distribution of mass is a function of the metric and its derivatives. In general, there is no means of entry to this closed loop to gain useful information about the universe. In the case of the Schwarzschild solution this was accomplished by assuming certain symmetry and boundary conditions. In more complicated problems.

however, symmetry conditions are not available in advance.

The existence of the second difficulty is more problematical, but it should not be overlooked. By way of illustration suppose an exact solution of the two-particle problem had been found for which the energy-momentum tensor did not vanish in the surrounding space. What would be the significance of this tensor field? One answer would be that it represented the energy of gravitational radiation. But would this solution necessarily be unique? Surely other solutions would exist with different distributions of continuous matter. Which solution, if any, would be fundamental to the two-body problem. Since the theory is non-linear one cannot simply subtract unrelated fields as in the case of Maxwellian electro-dynamics.

There are strong arguments in favour of a more primitive representation of matter than that of the energy-momentum tensor. Einstein himself has stated that "the energy tensor can be regarded only as a provisional means of representing matter. In reality matter consists of .... particles."(6)

In a relativistic particle theory of gravitation matter is represented as a set of world lines, but it would not be unnatural to encounter the field equations of General Relativity as adjunct-field equations. For if the theory is based on a variational principle and a symmetric second order tensor potential then the final structure is that of a Riemannian manifold with singularities along the world lines of particles. But the theory of General

Relativity is essentially the theory of Riemannian geometry.

A particle theory, in its basic form, is generally devoid of conservation laws. This is an inconvenience and is the main reason for introducing adjunct fields. Now a particle theory that constructs an adjunct Riemannian manifold automatically constructs a symmetric second rank tensor field whose divergence is everywhere zero. The theory must then interpret this conservation law and attach a meaning to the tensor. The problem of interpretation is considerably simplified in the case of a theory for which the conserved tensor vanishes (at least) in the presence of a single particle at rest.* Such a theory is that of A.N. Whitehead, although neither Whitehead himself nor those who have studied his work seem to have attached any importance to or even to have given recognition to this ready made conservation law.

The present work is a study of and an extension to the electrodynamics of Wheeler and Feynman and the relativistic gravitation of Whitehead. The aim of the extension is to incorporate into these theories some of the features of Steady State Cosmology.

Most of the mathematical theorems required are derived in Chapter 2. Chapter 3 and 4, each of which begins with further introductory remarks, are devoted respectively to electrodynamics and gravitation, and the general cosmological aspects of the theories are discussed in Chapter 5.

---

* Except of course on the world line or tube of the particle.

CHAPTER II

MATHEMATICAL PRELIMINARIES

2.1. THEORY OF DISTRIBUTIONS AND ASSOCIATED THEOREMS

Introduction

The use in mathematical physics of the so-called Dirac δ-function, defined by the equations

$$\delta(x) \quad = \quad 0, \qquad\qquad (x \neq 0) \tag{2.101}$$

$$\int_{-\infty}^{\infty} \delta(x)\, dx \quad = \quad 1. \tag{2.102}$$

has long been accepted by physicists as a useful means of describing singularities, although until recently its mathematical meaning remained obscure and even apparently nonsensical.

Certainly the quantity

$$\delta(x)$$

is no function at all, for if it were the well defined properties of functions would demand that if (2.101) were satisfied then the right hand side of (2.102) would be zero. However by assuming certain arbitrary secondary properties of a well behaved function, f, as for example

$$\frac{d}{dx}\int_a^b f(x, y)dy \quad = \quad \int_a^b \frac{\partial f}{\partial x}\, dy.$$

$$\int_a^b f'(x)g(x)dx \; = \; f(x)g(x)\Big]_a^b \; - \int_a^b f(x)g'(x)dx$$

$$\int_a^b f\left[\phi(x)\right] dx \; = \; \int_c^d f(\phi)\,\frac{dx}{d\phi}\, d\phi \;; \quad (\phi(b) = d,\ \phi(a) = c)$$

to apply for the δ-function, a formalism has been worked out which has proved very fruitful in physics, especially when coupled with some degree of geometrical intuition as to the "shape" of the δ-function.

Mathematically such a system is quite unsatisfactory, and even for one primarily interested in physical results, there is always the doubt as to its degree of self-consistency, and hence uncertainty with respect to its validity in new applications.

It was not until recent years that the δ-function technique was put on a firm axiomatic foundation by L. Schwartz in what is called the Theory of Distributions. As this theory is relatively new, it is considered worthwhile to include here a brief outline of its foundations in so far as they are required for the theorems which follow.

Elementary Theory of Distributions*

The great advance made by Schwartz was to define a linear continuous operation on a particular class of functions which assigns a number to each such function. Such an operation he called a Distribution. The most restricted functions to which a distribution is applicable are called testing functions (designated here by ϕ), the most important properties of which are that they are infinitely differentiable and must vanish identically outside some bounded region.

A distribution F can always be defined by an integrable function, $f(x)$, by the equation

$$F(\phi) = \int_{-\infty}^{\infty} f(x)\, \phi(x)\, dx \qquad (2.103)$$

the essential properties of linearity

$$T(\alpha\, \phi_1 + \beta\, \phi_2) = \alpha\, T(\phi_1) + \beta\, T(\phi_2) \qquad (2.104)$$

and of continuity

$$\lim_{n \to \infty} T(\phi_i) = 0 \text{ if } \lim_{n \to \infty} \phi_i = 0; \quad (\phi_i \text{ being a sequence of testing functions})$$

of a distribution T being evident in (2.103). Distributions are in fact designed as an extension to the theory of functions, all integrable functions being included in the sense of (2.103). This inclusion,

---

* The material in this section is in essence taken from Friedman's notes(7), although with considerable shift of emphasis to suit our present purposes.

beside being convenient, is a check on the self-consistency of the theory.

The converse of (2.103) is not true. It is not generally possible to find a function $t(x)$ such that

$$T(\phi) = \int_{-\infty}^{\infty} t(x)\phi(x)\,dx \quad * \qquad (2.105)$$

However it is convenient to use equation (2.105). The quantity $t(x)$ is then called a symbolic function and has no meaning except as contained in (2.105). A calculus will be defined for T, and this calculus and (2.105) will impose a "calculus" on t. By this means the symbolic function t will be used to describe the distribution T.

The rules of manipulation for t will be found to be identical to many of those applicable to the best behaved of functions, a fact which is not surprising since such functions are included in the theory by (2.103). Thus from the manipulative point of view we are led back to the scheme already used in the δ-function technique, but on the firmer foundation of being in correspondence with a self-consistent linear operator theory.

The emphasis in what follows will be directed towards tracing out a few of the threads of this correspondence.

CHANGE OF VARIABLE

We may consider an expression such as

$$\int t(x)\ \phi(y)dy\ ; \quad (y = y(x)) \qquad (2.106)$$

provided we attach a definite meaning to it. For an integrable function $f(x)$, we would have

$$\int f(x)\ \phi(y)dy = \int f(x)\ \phi(y\ (x))\ \frac{dy}{dx}\,dx$$

the right hand side of which defines a distribution on ψ, where

---

* Limits of $-\infty$ and $+\infty$ will hereafter be assumed unless otherwise stated.

$$\psi(x) = \phi \frac{dy}{dx} .$$

We shall therefore consider (2.106) to be a symbolic representation of the quantity

$$\int t(x) \left[\phi(y(x)) \frac{dy}{dx}\right] dx$$

which, by (2.105) is an expression for the distribution

$$T\left(\phi \frac{dy}{dx}\right)$$

Thus, by convention, we may make any well-behaved single-valued substitution we wish in the integrand of the right hand side of (2.105), provided the reverse substitution and (2.105) are used to attach a meaning to the new expression. Or we may start with any symbolic expression we wish provided by standard substitutions we arrive at (2.105) and provided the meaning we attach to the original expression is that contained in (2.105) after substitutions have been made. In the latter case we must of course make certain that the ϕ in (2.105) acquired by this means is a suitable testing function in order that the original expressions have meaning at all.

DIFFERENTIATION

The derivative of a distribution is defined in a manner similar to that for a function, (but with a noticeable change of sign):

$$T'(\phi) = \lim_{h \to 0} \frac{T(\phi(x-h)) - T(\phi(x))}{h} \tag{2.107}$$

Now since T defines a linear continuous operation, we have

$$\begin{aligned} T'(\phi) &= T\left(\lim \frac{\phi(x-h) - \phi(x)}{h}\right) \\ &= T(-\phi'(x)) \end{aligned}$$

or $$T'(\phi) = -T(\phi') \qquad (2.108)$$

again using the property of linearity.

Also, if $f(x)$ is a differentiable function, then using integration by parts

$$\int \frac{df}{dx}\,\phi\,dx = -\int f\,\frac{d\phi}{dx}\,dx$$

since ϕ vanishes at infinity, or, using (2.103) and (2.108)

$$F'(\phi) = \int \frac{df}{dx}\,\phi\,dx$$

Hence the ordinary derivative of a differentiable function is the same whether it is considered as a function or as a distribution. By a converse argument, integration by parts is always valid in (2.105) if we let t' be the symbolic function for T'.

We see by (2.108) that a distribution may be differentiated as often as desired, since ϕ can be differentiated infinitely often. In particular any non-differentiable but integrable function admits an infinite number of derivatives as a distribution.

EXTENSION TO MORE THAN ONE VARIABLE

The above theory is easily extended to any number of independent variables.

Symbolically we write

$$T(\phi(x_1, x_2, \ldots, x_n)) = \int\int\ldots\int t(x_1, x_2, \ldots, x_n)\phi(x_1, x_2, \ldots, x_n)dx_1 \ldots dx_n \qquad (2.109)$$

The partial derivative

$$\frac{\partial T(\phi)}{\partial x_i}$$

is defined by

$$\frac{\partial T(\phi)}{\partial x_i} = \lim_{h \to 0} \frac{T(\phi(x_1 \ldots x_i - h, \ldots x_n)) - T(\phi(x_1 \ldots x_n))}{h} \qquad (2.110)$$

from which it is easily deduced that

$$\frac{\partial T(\phi)}{\partial x_i} = - T\left(\frac{\partial \phi}{\partial x_i}\right) \tag{2.111}$$

The distribution (2.109) assigns a number to each function ϕ. We may however have a partial distribution

$$T_{x_1 \ldots x_m}\ (\phi(x_1 \ldots x_m \ldots x_n)) \qquad (m < n)$$

$$= \int \cdots \int t(x_1 \ldots x_m)\phi(x_1 \ldots x_n)dx_1 \ldots dx_m \tag{2.112}$$

which assigns for each function ϕ a new function θ of n-m variables

$$\theta\ (x_{m+1} \ldots x_n) \equiv T_{x_1 \ldots x_m}\ (\phi(x_1 \ldots x_n)) \tag{2.113}$$

The partial derivative (2.111) holds for (2.113) if

$$i \leqslant m$$

For

$$i > m$$

the ordinary definition

$$\frac{\partial \theta}{\partial x_i} = \lim_{h \to o} \frac{\theta(x_{m+1}, \ldots x_{i+h}, \ldots x_n) - \theta(x_{m+1}, \ldots x_n)}{h}$$

of course applies, and it is easy to see that

$$\frac{\partial T_{x_1 \ldots x_m}(\phi(x_1 \ldots x_n))}{\partial x_i} = T_{x_1 \ldots x_m}\left(\frac{\partial \phi}{\partial x_i}\right); \quad i > m \tag{2.114}$$

Theorem I - Differentiation under an integral sign.

The formula (2.114) clearly amounts to permitting differentiation under the integral sign of (2.112) by the ordinary rule for integrals of functions. That such a rule applies for more general symbolic expressions derived from (2.112) by changes of variables is not at all obvious.*

---

* Although no proof is given by Friedman.

To illustrate and prove this theorem, we will use only two independent variables, x and y. The expression

$$\frac{\partial}{\partial x}\int t(y(x, z))\,\psi(x, z)dz \tag{2.115}$$

has a definite meaning according to our convention with respect to substitution. Expressing the functional dependence of y on x and z by

$$f(x, y, z) = 0$$

and noting that*

$$\frac{\partial x}{\partial y}\cdot\frac{\partial y}{\partial x} = \frac{\partial y}{\partial z}\cdot\frac{\partial z}{\partial y} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial z} = -\frac{\partial x}{\partial y}\cdot\frac{\partial y}{\partial z}\cdot\frac{\partial z}{\partial x} = 1 \tag{2.116}$$

we see that (2.115) is equivalent to

$$\frac{\partial}{\partial x}\int t(y)\,\psi(x, z(x, y))\,\frac{\partial z}{\partial y}\,dy.$$

or

$$\frac{\partial}{\partial x}\int t(y)\,\phi(x, y)\,dy$$

where $\phi(x, y) = \psi(x, z(x, y))\,\frac{\partial z}{\partial y}$

Therefore, using (2.114)

$$\frac{\partial}{\partial x}\int t(y(x, z))\,\psi(x, z)dz$$

$$\equiv \frac{\partial T_y(\phi(x, y))}{\partial x}$$

$$= \int t(y)\left(\frac{\partial \phi}{\partial x}\right)_y dy$$

$$= \int t\left[\left(\frac{\partial\psi}{\partial x}\right)_z + \left(\frac{\partial\psi}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)\right]\left(\frac{\partial z}{\partial y}\right)dy + \int t\,\psi\,\frac{\partial^2 z}{\partial x\,\partial y}\,dy \tag{2.117}$$

---

* I.S. Sokolnikoff, "Advanced Calculus", § 33, Problem 3.

It will now be shown that direct differentiation under the integral sign in (2.115) coupled with an integration by parts is a legitimate procedure in that it yields the equality (2.117).

Thus

$$\frac{\partial}{\partial x} \int t(y)\, \psi(x, z)\, dz$$

$$= \int t'(y)\, \psi(x, z) \frac{\partial y}{\partial x} \frac{\partial z}{\partial y}\, dy + \int t(y) \left(\frac{\partial \psi}{\partial x}\right)_z \frac{\partial z}{\partial y}\, dy$$

$$= - \int t'(y)\, \psi(x, z) \frac{\partial z}{\partial x}\, dy + \int t(y) \left(\frac{\partial \psi}{\partial x}\right)_z \frac{\partial z}{\partial y}\, dy, \text{ using (2.116)}$$

$$= \int t(y) \left(\frac{\partial \psi}{\partial z}\right)_x \frac{\partial z}{\partial y} \frac{\partial z}{\partial x}\, dy + \int t(y)\, \psi \frac{\partial^2 z}{\partial y\, \partial x}\, dy$$

$$+ \int t(y) \left(\frac{\partial \psi}{\partial x}\right)_z \frac{\partial z}{\partial y}\, dy$$

(by a partial integration), which is the same as (2.117).

In effect this theorem indicates a second valid use for partial integration.

δ-distribution

A distribution δ may be defined by

$$\delta(\phi) = \phi(0) \tag{2.118}$$

Now this distribution belongs to a class for which an extension to the class of functions which can be operated on is permissible. This extension is made in the following way:

If $$T(\phi) = 0$$

for every ϕ which vanishes within a certain bounded region A, then T defines a distribution on any function which is differentiable within and on the boundary of A, but which need not vanish anywhere. This extension is evidently not inconsistent with the preceding discussion.

Denoting by σ any function to which δ may be applied, it follows that since the derivatives of σ need only exist in the neighbourhood of the origin, σ may be multivalued. To cover this case we define

$$\delta(\sigma) = \Sigma\, \sigma(0)$$

where Σ indicates summation over the various zeroes of σ.

Symbolically

$$\delta(\sigma) = \int \delta(x)\, \sigma(x)\,dx \qquad (2.119)$$

where the limits implied in (2.119) are any which include the origin, the upper and lower limits being positive and negative respectively. (It is understood that writing limits down in the reverse order is equivalent to a change of sign for the whole expression.)

DIFFERENTIATION

Heaviside's unit step-function is an integrable function and defines a distribution by

$$H(\phi) = \int_{-\infty}^{\infty} h(x)\, \phi(x)\,dx$$

$$\equiv \int_{0}^{\infty} \phi(x)\,dx$$

It follows that

$$H'(\phi) = -\, H(\phi')$$

$$= -\int_{0}^{\infty} \phi'(x)\,dx$$

$$= \phi(0), \text{ by partial integration, since } \phi(\infty) = 0$$

Thus $$H' = \delta \qquad (2.121)$$

Another important property of the δ-distribution is required.

Noting that

$$\int \delta'(x)\; x\,\phi(x)dx = -\int \delta(x)\,\frac{d}{dx}\left[x\,\phi(x)\right]dx$$

$$= -\int \delta(x)\left[\phi(x) + x\,\phi(x)\right]dx$$

$$= -\int \delta(x)\;\phi(x)dx$$

we may write symbolically

$$x\,\delta'(x) = -\delta(x).$$

More generally, by the same method, it can be shown that

$$x\,\delta^{n}(x) = -n\,\delta^{n-1}(x). \qquad (2.122)$$

MULTI-DIMENSIONAL δ-DISTRIBUTIONS

In more than one dimension

$$\delta(\phi(x_1, x_2, \ldots, x_n)) \equiv \phi(0, 0, \ldots, 0)$$

$$= \int\ldots\int \delta(x, \ldots x_n)\phi(x_1 \ldots x_n)dx_1 \ldots dx_n$$

is the product of n commuting partial distributions, for

$$\delta_{x_n}\,(\delta_{x_{n-1}}\,(\ldots\,\delta_{x_1}(\phi(x, \ldots, x_n))\,\ldots\,))$$

$$= \phi(0, 0, \ldots, 0)$$

$$= \delta(\phi(x_1, x_2, \ldots, x_n))$$

Symbolically

$$\delta(x_1, \ldots, x_n) = \delta(x_1)\;\delta(x_2)\;\ldots\;\delta(x_n).$$

Consider the distribution

$$\delta(\phi(x,y,z)) = \phi(0,0,0) = \iiint \delta(x,y,z)\phi(x,y,z)dx\;dy\;dz. \qquad (2.123)$$

Transferring to generalized coordinates, u, v, w, in the usual manner, we obtain

$$\delta(\phi) = \iiint \delta\left\{x(u,v,w),y(u,v,w),z(u,v,w)\right\}\;\phi\;J\;du\;dv\;dw$$

where J is the Jacobian of the transformation.

Now if the origin in the original system has coordinates u_o, v_o, w_o in the new, we may write

$$\delta(\phi) = \phi(x = 0,\ y = 0,\ z = 0)$$

$$= \phi(u = u_o,\ v = v_o,\ w = w_o)$$

$$= \iiint \delta(u-u_o,\ v-v_o,\ w-w_o)\phi(u,v,w)\,du\,dv\,dw.$$

Thus, symbolically

$$\frac{\delta(u-u_o,\ v-v_o,\ w-w_o)}{J} = \delta(x,y,z). \qquad (2.1231)$$

This equality simply means that the usual coordinate transformations are justified in symbolic equations provided that the above prescription is used to transform the symbolic function.

More generally it follows that if x^u are Cartesian coordinates (or in Riemannian space, local cartesians), and if $x_o{}^u$ is a given point (which in Riemannian space must be the origin of the local Cartesians), then if $y_o{}^u$ is the same point in any other coordinate system, we may write

$$\delta(x^1 - x_o{}^1)\delta(x^2 - x_o{}^2)\ \ldots\ \delta(x^n - x_o{}^n)$$

$$= \frac{\delta(y^1 - y_o{}^1)\delta(y^2 - y_o{}^2)\ \ldots\ \delta(y^n - y_o{}^n)}{\sqrt{g}} \qquad (2.124)$$

where g is the determinant of the second order covariant metric tensor (g_{mn}) which is the unit tensor (at $x_o{}^u$) in the x coordinate system.

Clearly $\delta(y^1 - y_o{}^1)\delta(y^2 - y_o{}^2)\ \ldots\ \delta(y^n - y_o{}^n)$ may be regarded as a scalar density.

The formula (2.124) breaks down if g_o is singular. However returning to equation (2.123) we may write in spherical polar coordinates (r, θ, ϕ) centred at the origin of the x,y,z coordinate system,

$$\delta(\Phi) = \Phi(x=0,\ y=0,\ z=0)$$

$$= \Phi(r=0)$$

$$= \int \delta(r)\ \Phi(r,\theta,\phi)dr$$

$$= \iiint \frac{\delta(r)\ \Phi(r,\theta,\phi)}{4\pi r^2} \sin\theta r^2\ dr\ d\theta\ d\phi$$

But $\delta(\Phi) = \iiint \delta(x,y,z)\ \Phi(x,y,z)\ dx\ dy\ dz$

$$= \iiint \delta\left\{x(r,\theta,\phi),y(r,\theta,\phi),z(r,\theta,\phi)\right\} \Phi r^2 \sin\theta\ dr\ d\theta\ d\phi$$

obtaining for this special case

$$\frac{\delta(r)}{4\pi r^2} = \delta(x)\ \delta(y)\ \delta(z) \qquad (2.125)$$

Another useful identity follows from the same transformation of coordinates. Consider the distribution defined by

$$T(\psi) = \iiiint \delta(t^2 - (x^2+y^2+z^2))\psi(x,y,z,t)dt\ dx\ dy\ dz \qquad (2.1251)$$

Transferring to spherical polar coordinates,

$$T(\psi) = \int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi}\int_{-\infty}^{\infty} \delta(t^2-r^2)h(r)\psi(r,\theta,\phi,t)r^2\sin\theta\ dt d\phi d\theta dr$$

Here the Heaviside step function $h(r)$ is inserted to restrict the range of values of r to positive numbers. To evaluate this distribution, the substitution is made

$$t^2 - r^2 = u$$

yielding

$$T(\psi) = \int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi}\int_{-r^2}^{\infty} \delta(u)\, \frac{h(r)\psi(r,\theta,\phi,(u+r^2)^{1/2})}{2(u+r^2)^{1/2}} r^2 \sin\theta\, du\, d\phi\, d\theta\, dr$$

$$+\int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi}\int_{\infty}^{-r^2} \delta(u)h(r)\, \frac{\psi(r,\theta,\phi,-(u+r^2)^{1/2})}{-2(u+r^2)^{1/2}} r^2 \sin\theta\, du\, d\phi\, d\theta\, dr \qquad (2.1252)$$

$$= \int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi} \frac{h(r)}{2} \left[\psi(r,\theta,\phi,r) + \psi(r,\theta,\phi,-r)\right] r \sin\theta\, d\phi\, d\theta\, dr \qquad (2.1253)$$

Similarly it may be shown that

$$\iiiint \left[\delta(t+r) + \delta(t-r)\right] \frac{h(r)}{2r}\, \psi(x,y,z,t)\, dt\, dx\, dy\, dz \qquad (2.1254)$$

$$\equiv \int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi}\int_{-\infty}^{\infty} \left[\delta(t+r)+\delta(t-r)\right] \frac{h(r)}{2}\, \psi(r,\theta,\phi,t) r \sin\theta\, dt\, d\phi\, d\theta\, dr$$

$$= \int_{-\infty}^{\infty}\int_{0}^{\pi}\int_{0}^{2\pi} \frac{h(r)}{2} \left[\psi(r,\theta,\phi,r)+\psi(r,\theta,\phi,-r)\right] r \sin\theta\, d\phi\, d\theta\, dr$$

which is the same as (2.1253). By thus identifying (2.1251) and (2.1254) there obtains

$$\delta(t^2 - (x^2+y^2+z^2)) = \frac{h(r)}{2r} \left[\delta(t+r) + \delta(t-r)\right] \qquad (2.1255)$$

It is worth noting that in (2.1252) the quantity

$$\frac{\psi}{(u+r^2)^{1/2}}$$

is not by itself a suitable testing function, but that the combined quantity

$$\frac{\psi\, r^2}{(u+r^2)^{1/2}}$$

is such a function. Noting that the factor r^2 came from the Jacobian of the coordinate transformation, it is easy to see that the identity (2.1255) does not have a two dimensional analogue. That is, a similar relationship in the two variables t and x,

$$\delta(t^2-x^2) = \frac{h((x^2)^{1/2})}{2x}\left[\delta(t+x)+\delta(t-x)\right]$$

is devoid of meaning in the theory of distributions, although without due caution to note its lack of meaning it could have been derived in a manner similar to that for (2.1255). This simple example is an illustration of the fact that although the techniques developed from the axioms of the Theory of Distributions correspond almost identically with those originally used with respect to the Dirac δ-function, in the former case, pitfalls are clearly marked, while in the latter case, where symbolic functions are treated as though they were ordinary functions, such is not always the case.

<u>Theorem II - A δ-function Identity</u>

We now establish the theorem* that if x, y, and z are cartesian coordinates, and t an independent variable, then

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right)\delta(t^2 - (x^2 + y^2 + z^2))$$

$$= 4\pi\ \delta(x)\ \delta(y)\ \delta(z)\ \delta(t) \qquad (2.126)$$

Transferring to spherical polar coordinates, (r, θ, φ) in the usual way, and using (2.1255) the left hand side of (2.126) becomes

---

* The physical significance of this theorem in particle electrodynamics and the necessity of a mathematical proof of it are discussed in Section 2.1.

$$\left(\frac{\partial^2}{\partial t^2} - \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r}\right) \frac{h(r)}{2r} \quad (\delta(t-r) + \delta(t+r))$$

$$= \frac{h(r)}{2r} \quad (\delta''(t-r) + \delta''(t+r))$$

$$- \frac{1}{2r^2}\frac{\partial}{\partial r}\Big\{(\delta(t-r) + \delta(t+r)) \; (r\delta(r) - h(r))$$

$$+ r\, h(r) \; (-\delta'(t-r) + \delta'(t+r))\Big\}, \text{ using } (2.121)$$

$$= \frac{h(r)}{2r} \; (\delta''(t-r) + \delta''(t+r))$$

$$- \frac{1}{2r}\Big\{(\delta(t-r) + \delta(t+r)) \quad \left(\frac{\delta(r)}{r} + \delta'(r) - \frac{\delta(r)}{r}\right)$$

$$+ \left(\delta(r) - \frac{h(r)}{r}\right) (-\delta'(t-r) + \delta'(t+r))$$

$$+ \left(\frac{h(r)}{r} + \delta(r)\right) \quad (-\delta'(t-r) + \delta'(t+r))$$

$$+ h(r) \; (\delta''(t-r) + \delta''(t+r))\Big\}$$

$$= \frac{-\delta(r)}{r}\Big[-\delta'(t-r) + \delta'(t+r)\Big]$$

$$- \frac{1}{2r}\Big[\delta(t-r) + \delta(t+r)\Big] \quad \delta'(r)$$

$$= \frac{-\delta(r)}{r} (-\delta'(t) + \delta'(t)) \quad + \frac{\delta(r)}{2r^2} (\delta(t-r) + \delta(t+r)) \text{ using } (2.122)$$

$$= \frac{1}{r^2} \delta(t) \; \delta(r)$$

$= 4\pi \; \delta(x) \; \delta(y) \; \delta(z) \; \delta(t)$ by (2.125), and the theorem is proved.

(Use has been made here of the fact that

$$\delta(t \pm r)\delta(r) = \delta(t)\,\delta(r)$$

which follows directly from (2.1231).

It is worth noting here that while the users* of this theorem did not produce a proof of it, Friedman does in effect prove it by solving the wave equation

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right) w = \delta(x)\,\delta(y)\,\delta(z)\,\delta(t).$$

The procedure required for solving this equation is much more lengthy and elaborate and requires more concepts than for the simple proof given above.

---

* Feynman and Wheeler.

2.2. A DEFINITION OF SPATIAL DISTANCE IN RIEMANNIAN (1-3) SPACE-TIME

Introduction

Various attempts have been made suitably to define spatial distance in Riemannian space-time, in general each definition being constructed to serve some particular purpose. One important example[8] is the distance (Δ) of Kermack, McCrea and Whittaker.

A definition is given here of a distance (D) from a general point to a time-like (world) line along either of the two null geodesics that intersect the point and line. D has the special property that for a given world line it is an invariant function of the coordinates of the general point, and is therefore a suitable quantity from which to construct a potential theory. Although D and Δ are physically distinct, they will be shown to be mathematically equivalent.

In his Theory of Relativity,[3] Whitehead defined a Cartesian scalar to provide for the law of diminishing effects in his gravitational potential. This same scalar appears in the electromagnetic potentials[1] of Lienard-Wiechert. It will be shown that this scalar, defined in flat space-time, is a particular case of D.

Definition of Spatial Distance in Riemannian Space-time.

Referring to Figure (2.21) let X be a general point (with coordinates x^μ) and Γ the required null geodesic joining this point to a given time-like line L, and meeting L at X_0 (x_0^μ); and let v_0^μ be the unit velocity 4-vector of L at x_0. At any point (and in particular at X) on Γ other than X_0, the vector formed by Levi-Civita parallel transport of v_0^μ along Γ will be designated by v^μ.

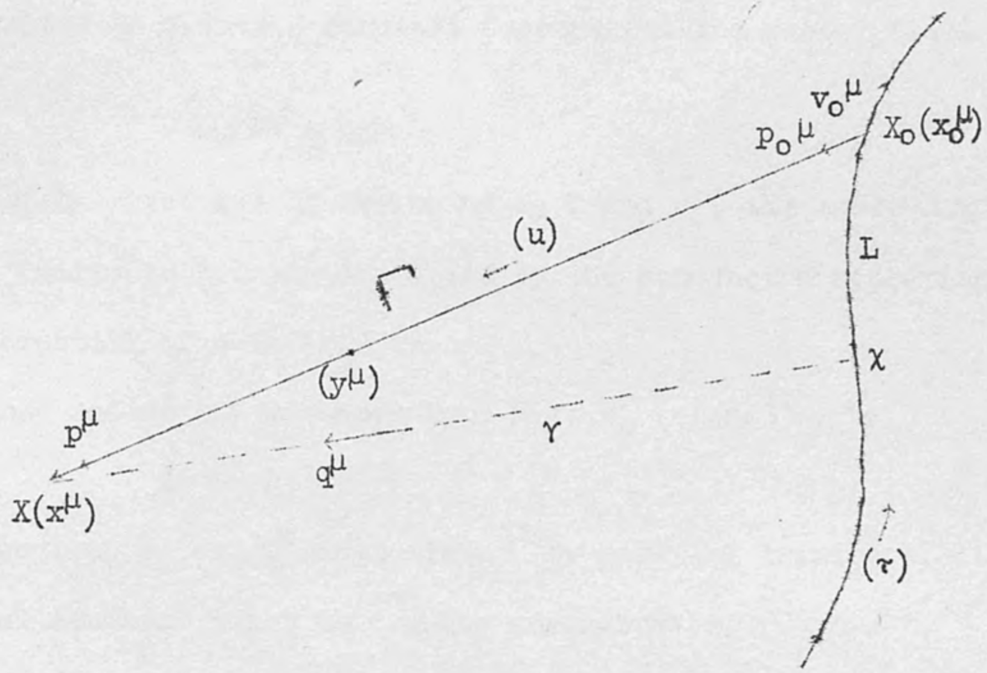


Fig. 2.21

Now consider a parameter u, defined along Γ by the differential equations of the geodesic

$$\frac{dp^\mu}{du} + \left\{ {\mu \atop \sigma\,\rho} \right\} p^\sigma p^\rho = 0 \qquad (2.201)$$

where p^μ is the tangent vector

$$p^\mu \equiv \frac{dy^\mu}{du} \qquad (2.202)$$

along Γ, y^μ being coordinates on Γ .

Because of the homogeneity of (2.201), the parameter u is thereby defined to within a linear transformation. If we specify the value of u at X by u(X) and at X_o by $u(X_o)$, then

$$U \equiv u(X) - u(X_o) \qquad (2.2021)$$

is determined to within a constant factor, and the vector field along Γ

$$P^\mu \equiv U p^\mu \qquad (2.203)$$

is completely specified by choice of X, L and Γ, the undetermined constant factor in U being cancelled by the same factor appearing in the differential of u in (2.202).

We now define the distance from X to X_o (along Γ) by

$$D = P_\mu v^\mu \qquad (2.204)$$

Since both P_μ and v^μ move along Γ by parallel transport, it is immaterial at which point on Γ the invariant

$$P_\mu v^\mu$$

is formed. (Hereafter the vectors P_μ and v^μ will refer to the point X, while a subscript $_o$ will relate any quantity to X_o).

Now for a given L, D is a double-valued invariant function of x^μ. If we let Γ_R and Γ_S respectively specify the appropriate geodesics on the past and future halves of the null surface through X, and name the magnitudes of the distances associated with these geodesics D_R and D_S, then D_R and D_S are each single valued functions of x^μ.

If we adopt at X_o local Cartesian coordinates oriented so that the time-like axis is tangent to L, then $v_o^4 = 1$, $v_o^i = 0$ $(i = 1,2,3)$, and

$$D = P_o^4 \equiv U\, p_o^4 \equiv U \frac{dy_o^4}{du}$$

for this choice of coordinates. It follows that with reference to Γ_R, (since dy_o^4 is then positive), the scalar

$$P_\mu v^\mu$$

is always positive, while it is negative for Γ_S. We will therefore define

$$\frac{1}{D_R} = \frac{h(D)}{D}$$
$$\frac{1}{D_S} = \frac{-h(-D)}{D} \qquad (2.205)$$

instead of (2.204).

COMPARISON WITH Δ.

If, again adopting oriented local cartesian at x_0, we choose the scale of u at X_0 so that

$$d\,u_0 = d\,y_0^4$$

then for this particular choice

$$D = U.$$

Also for a point X infinitesimally close to X_0, so that

$$x^\mu - x_0^\mu = d\,y_0^\mu$$

we have

$$D = d\,u_0 = d\,y_0^4.$$

These results are exactly the requirements set down by Kermack, McCrea, and Whittaker for the definition of Δ. Thus Δ and D are mathematically equivalent. However in applying D to physical problems L is to be the world line and x_0^μ the coordinates of a "source" or "emitter", while x^μ are either the coordinates of a "receptor" or simply coordinates in terms of which a potential field due to a "source" is defined. The roles of X and X_0 are exactly the reverse of this in the physical definition of Δ.

D IN TERMS OF CARTESIAN TENSORS

We now consider the special form which equation (2.204) takes in flat space-time and using Cartesian coordinates. The metric tensor is then denoted by

$$\eta_{mn} \equiv \begin{pmatrix} -1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & 1 \end{pmatrix} \qquad (2.206)$$

In Cartesian tensors a possible choice of U is

$$U = x^4 - x_o^4 \equiv xx_o^4$$

using the notation

$$ab^\mu \equiv a^\mu - b^\mu$$

Also, for Cartesian tensors

$$v^\mu = v_o^\mu.$$

Thus (2.204) becomes

$$D = \left(xx_o^4 . \frac{dy_\mu}{dy^4}\right)\left(v_o^\mu\right) \qquad (2.207)$$

But here $\frac{dy_\mu}{dy^4}$ is constant along Γ

and

$$\frac{dy_\mu}{dy^4} = \frac{xx_{o\mu}}{xx_o^4} \qquad (2.208)$$

Therefore (2.207) becomes

$$D = xx_{o_\mu} v_o^\mu \qquad (2.209)$$

which is the form(1) used for the Lienard - Wiechert potentials.

Again, returning to (2.207), we note that if X and X_o lie on Γ_R, then

$$xx_o^4 = r \qquad (2.210)$$

where

$$r = \left[(xx_o^1)^2 + (xx_o^2)^2 + (xx_o^3)^2\right]^{1/2} .$$

Also, if τ is arc length along L, we have

$$v_o^{\mu} = \frac{dx_o^{\mu}}{d\tau} = \beta \frac{dx_o^{\mu}}{dx_o^4} \tag{2.211}$$

where

$$\beta = \left(1 - \left[\left(\frac{dx_o^1}{d\tau}\right)^2 + \left(\frac{dx_o^2}{d\tau}\right)^2 + \left(\frac{dx_o^3}{d\tau}\right)^2\right]\right)^{-1/2}$$

Using (2.210) and (2.211) in (2.209) we have

$$D_R = \beta\,(r - \xi) \tag{2.212}$$

where here

$$\xi = xx_o^1 \frac{dx_o^1}{dx_o^4} + xx_o^2 \frac{dx_o^2}{dx_o^4} + xx_o^3 \frac{dx_o^3}{dx_o^4}\,.$$

The retarded distance (2.212) is the same as that used by Whitehead* to "express the gravitational law of fading intensity".

<u>Equivalent Definition of Distance in Riemannian Space-Time Using δ-functions.</u>

In order to employ δ-functions to define D, our notation is extended in the following manner:

Again referring to Fig. (2.21), a general point on L is denoted by χ, and its coordinates by $\chi^{\mu}(\tau)$. Also an arbitrary zero is assumed for τ between the intersections of L with Γ_R and Γ_S, so that at Γ_S τ has some positive value and at Γ_R some negative value. The geodesic joining X and χ is designated by γ (in general not a null geodesic), coordinates along γ by y^{μ}, and geodesic arc length from χ to X

---

* See Ref. (3), Ch. IV. equations (7), (8), (13) and (15).

$$\int_{\chi \atop \text{along } \gamma}^{X} ds \equiv \fint_{\chi}^{X} (g_{mn}\, dy^m\, dy^n)^{1/2} \qquad (2.213)$$

by S. (Here g_{mn} is the metric tensor of the space and $\fint$ indicates geodesic line integration.) We purposely avoid the use of an indicator here, so that S^2 may be either positive or negative, being negative for points χ between Γ_R and Γ_S and positive for χ outside these boundaries.

Now in Cartesian coordinates in flat space-time,

$$S^2 = x\chi^\mu\, x\chi_\mu$$

and

$$\frac{\partial}{\partial \tau}(S^2) = 2\chi x^\mu \frac{d\chi_\mu}{d\tau} \qquad (2.214)$$

Therefore in this case we have the following identity:

$$\int_{-\infty}^{\infty} \delta(S^2)\, d\tau$$

$$= \int_{-\infty}^{\infty} \delta(x\chi^\mu\, x\chi_\mu)\, d\tau$$

$$= \int_{-\infty}^{0} \delta(x\chi^\mu\, x\chi_\mu)\, d\tau + \int_{0}^{\infty} \delta(x\chi^\mu\, x\chi_\mu)\, d\tau$$

$$\overset{*}{=} \int_{\infty}^{-\infty} \delta(x\chi^\mu\, x\chi_\mu) \frac{d(x\chi^\mu\, x\chi_\mu)}{2\chi x^\mu \frac{d\chi_\mu}{d\tau}} \qquad (\tau < 0)$$

$$+ \int_{-\infty}^{\infty} \delta(x\chi^\mu\, x\chi_\mu) \frac{d(x\chi^\mu\, x\chi_\mu)}{2\, \chi x^\mu \frac{d\chi_\mu}{d\tau}} \qquad (\tau > 0)$$

Noting that $\int_{\infty}^{-\infty} \delta(x)\, dx = - \int_{-\infty}^{\infty} \delta(x)\, dx$

---

* Bearing in mind that the limits need only include the zeroes of the δ-function argument, the limits $\pm\infty$ are equivalent to although not identical with those derived from the substitution.

and that $x\chi^{\mu} = -\chi x^{\mu}$,

and also that $S^2 = 0$ when $\chi^{\mu} = x_o^{\mu}$

we then have

$$\int_{-\infty}^{\infty} \delta(s^2)d\tau = \left(\frac{1/2}{xx_o{}^{\mu}v_{o_{\mu}}}\right)_{\Gamma_R} + \left(\frac{1/2}{x_o x^{\mu}v_{o_{\mu}}}\right)_{\Gamma_S} \qquad (2.2141)$$

where the subscripts Γ_R and Γ_S indicate on which geodesics the bracketed quantities are evaluated. Using (2.209) this last equation becomes

$$\int_{-\infty}^{\infty} \delta(S^2)d\tau = \frac{1}{2}\left(\frac{1}{D_R} + \frac{1}{D_S}\right) \qquad (2.215)$$

Recalling that by convention the point $\tau = 0$ lies on a space-like surface through X, it is also clear that

$$\int_{-\infty}^{o} \delta(S^2)d\tau = \frac{1}{2D_R} \qquad (2.216)$$

$$\int_{o}^{\infty} \delta(S^2)d\tau = \frac{1}{2D_S}$$

It will now be shown that these last three equations which were written down for flat space-time, equally well apply in general Riemannian space-time.

A THEOREM ON GEODESICS

We adopt the convention here that coordinates along a general geodesic γ are designated by y^{μ}, and u is a parameter along γ satisfying (2.201). To avoid confusion, however, we use q^{μ} and Q^{μ} instead of p^{μ} and P^{μ} to indicate the corresponding tangent vectors when γ is not necessarily null. That is

$$q^\mu = \frac{dy^\mu}{du} \quad \text{along } \gamma$$

$$Q^\mu = Uq^\mu \equiv \left[u(X) - u(\chi)\right] q^\mu$$

It will now be established that

$$\frac{\partial}{\partial\tau}(S^2) = -2\, Q_\mu \frac{d\chi^\mu}{d\tau} \tag{2.218}$$

First we notice that there are several equivalent expressions for S^2, other than that given by (2.213). For

$$S = \int_\chi^X (g_{mn}\, dy^m\, dy^n)^{1/2}$$

$$= \int_\chi^X (q^m\, q_m)^{1/2}\, du$$

But for a given γ, $q^m\, q_m$ is a constant scalar, and therefore

$$S^2 = q^m\, q_m \left(\int_\chi^X du \right)^2$$

$$= Q^m\, Q_m \tag{2.219}$$

$$= U \int_\chi^X g_{mn} \frac{dy^n}{du} \frac{dy^n}{du}\, du \tag{2.220}$$

Now consider two neighbouring points χ and χ' on L, separated by the vector $\Delta\, \chi^\mu$ and arc length $\Delta\tau$. Then indicating quantities on $X\,\chi'$ by dashes

$$\frac{\partial}{\partial\tau}(S^2) = \lim_{\Delta\tau\to 0} \frac{S'^2 - S^2}{\Delta\tau} \tag{2.221}$$

Using the formula (2.220) for S^2, we note that since u is arbitrary to the extent of a constant factor and an additive constant, we may adjust these factors so that u has the same value at X for both γ and γ' and also so that

$$U' = U.$$

Thus

$$S'^2 = U \int_{\chi'}^{X} g_{mn'} \left(\frac{dy^m}{du}\right)'\left(\frac{dy^n}{du}\right)' du \tag{2.222}$$

Now if Δy^μ is the vector

$$y'^\mu - y^\mu$$

where y'^μ is the point on γ' with the same value of u as for y^μ on γ, we have to the first power in Δy^r

$$g'_{mn} \frac{dy'^m}{du} \frac{dy'^n}{du} - g_{mn} \frac{dy^m}{du} \frac{dy^n}{du}$$

$$= \frac{\partial}{\partial y^r} \left(g_{mn} \frac{dy^m}{du} \frac{dy^n}{du} \right) \Delta y^r$$

$$= \frac{\partial g_{mn}}{\partial y^r} \frac{dy^m}{du} \frac{dy^n}{du} \Delta y^r + 2 g_{mn} \frac{d\Delta y^m}{du} \frac{dy^n}{du} \tag{2.223}$$

Now equation (2.201) (with p^μ replaced by q^μ) is equivalent to

$$\frac{\partial g_{mn}}{\partial y^r} \frac{dy^m}{du} \frac{dy^n}{du} = \frac{d}{du} \left(g_{rn} \frac{dy^n}{du} \right)$$

Hence (2.223) may be re-written

$$g'_{mn} \frac{dy^{m'}}{du} \frac{dy'^n}{du} - g_{mn} \frac{dy^m}{du} \frac{dy^n}{du}$$

$$= 2 \frac{d}{du} \left(g_{rn} \frac{dy^n}{du} \right) \Delta y^r + 2 g_{rn} \frac{dy^n}{du} \frac{d}{du} (\Delta y^r)$$

$$= 2 \frac{d}{du} \left(g_{rn} \frac{dy^n}{du} \Delta y^r \right) \tag{2.224}$$

Thus by (2.220), (2.222) and (2.224)

$$S'^2 - S^2 = 2U \int_{\chi}^{X} \frac{d}{du} \left(g_{rn} \frac{dy^n}{du} \Delta y^r\right) du$$

$$= 2U\, g_{rn} \frac{dy^n}{du} \Delta y^r \Big|_{\chi}^{X}$$

$$= -\, 2\, Q_r\, \Delta\, \chi^r$$

Dividing through by $\Delta\tau$ and taking the limit in (2.221) we obtain (2.218).

Using (2.218) we find, following the same argument as for flat space, ((2.218) being a generalization of (2.214)):

$$\int_{-\infty}^{0} \delta(S^2) d\tau = \left(\frac{1/2}{P_\mu\, v^\mu}\right)_{\Gamma_R} = \frac{1}{2\, D_R} \qquad (2.225)$$

$$\int_{0}^{\infty} \delta(S^2) d\tau = -\left(\frac{1/2}{P_\mu\, v^\mu}\right)_{\Gamma_S} = \frac{1}{2\, D_S} \qquad (2.226)$$

An important corollary follows from (2.218). If we hold χ fixed and let X vary, we see that

$$\frac{\partial}{\partial x^\mu} (S^2) = 2\, Q_\mu \qquad (2.227)$$

or, using (2.219)

$$\frac{\partial}{\partial x^\mu} (Q_m\, Q^m) = 2\, Q_\mu$$

or again

$$Q_{m|\mu}\, Q^m = Q_\mu \qquad (2.228)$$

where $|\mu$ represents absolute differentiation with respect to the coordinates x^μ.

If we differentiate both sides of (2.228) once more we obtain

$$Q_{m|\mu n} Q^m + Q_{m|\mu} Q^m{}_{|n} = Q_{\mu|n} \tag{2.229}$$

Now $Q_{m|\mu n} Q^m = Q_{m|n\mu} Q^m + R_{sm\mu n} Q^s Q^m$ (where $R_{sm\mu n}$ is the curvature tensor)

$$= Q_{m|n\mu} Q^m$$

because of the skew-symmetry of the first two indices of the curvature tensor. Hence the first term on the left hand side of (2.229) is symmetric in μ and n. Clearly the second term has the same symmetry and thus we have

$$Q_{\mu|n} = Q_{n|\mu} \tag{2.230}$$

Differentiation of Adjunct Field Quantities

The purpose of this section has been to define a suitable spatial distance in Riemannian space-time in terms of which later to define a potential theory. Once such a theory has been set up in principle, it is essential to determine the differential equations which are satisfied by the potentials involed. The fields derived from these potentials are called adjunct fields.

Two equivalent defintions of D have been given, one of which is described in terms of the δ-functions. The other quantities available for use in constructing a vector or tensor potential are the vectors P^μ and v^μ, which are defined only in terms of parallel transfer.

To illustrate the procedure to be followed later, we shall consider a potential V defined in a given space and in terms of a given (single) world line L by

$$V = \frac{1}{D_R}$$

METHOD I.

$$V = \frac{1}{D_R} = \left(\frac{1}{P_\mu \ v^\mu}\right)_{\Gamma_R}$$

Clearly for given L, V is a function of the coordinates x^μ of the point X at which the potential is being considered. This functional dependence on x^μ however will in practice not be explicit; the quantity $P_\mu \ v^\mu$ for a given coordinate frame and given L is a function of the end points X and X_0 with the dependence of X_0 on X being contained in the equations of the null surface through X and those of the line L. We therefore write

$$\frac{\partial V}{\partial x^\mu} = \left(\frac{\partial V}{\partial x^\mu}\right)_{(x_0^\mu)} + \left(\frac{\partial V}{\partial x_0^n}\right)_{(x^\mu)} \frac{dx_0^n}{d\tau} \frac{\partial \tau}{\partial x^\mu} \qquad (2.231)$$

(Here $(\ )_{(x)}$ indicates that the quantity x is held constant).

A general formula for

$$\frac{\partial \tau}{\partial x^\mu}$$

is derived in the following manner:

Consider two neighbouring null geodesics Γ and Γ' with end points X, X_0, and X', X_0' respectively. Let XX' be separated by the vector Δx^μ and X_0X_0' by Δx_0^μ, as shown in Figure 2.22.

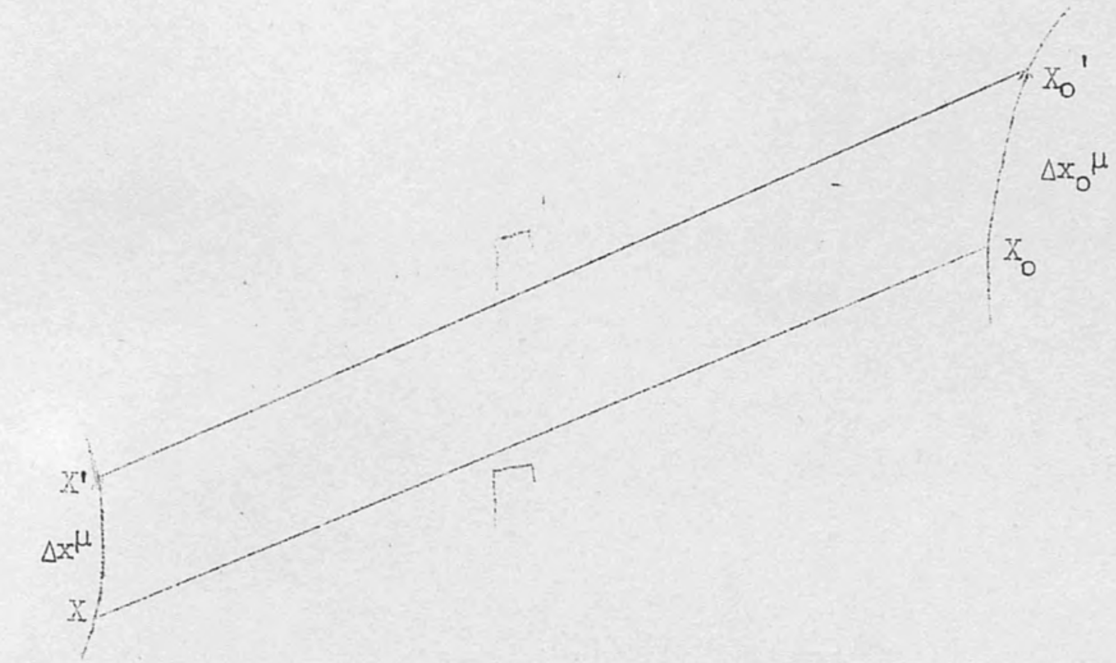


Fig. 2.22.

Then the theorem on null Geodesics in (8) establishes that

$$p_{o_\mu} \Delta x_o^\mu = p_n \Delta x^n$$

or

$$p_{o_\mu} v_o^\mu \Delta \tau_o = p_n \Delta x^n .$$

Now if X' is chosen such that only one component (say Δx^m) is non-zero, we have dividing through by Δx^m

$$p_{o_\mu} v_o^\mu \frac{\Delta \tau_o}{\Delta x^m} = p_m \qquad (2.2311)$$

or taking limits;

$$\frac{\partial \tau_o}{\partial x^m} = \frac{p_m}{p_{o_\mu} v_o^\mu} = \frac{P_m}{P_\mu v^\mu} . \qquad (2.232)$$

Instead of (2.231) we have

$$\frac{\partial V}{\partial x^\mu} = \left(\frac{\partial V}{\partial x^\mu}\right)_{(x_o^\mu)} + \left(\frac{\partial V}{\partial x_o^{\ n}}\right)_{(x^\mu)} \frac{v_o^{\ n} P_\mu}{P_m v^m} \tag{2.232}$$

This formula holds for any function V defined at X in terms of Γ and L.

METHOD II

$$V = \frac{1}{D_R} = 2 \int_{-\infty}^{0} \delta(S^2)\, d\tau \tag{2.233}$$

Since $$\frac{\partial (S^2)}{\partial \tau} = -2\, Q_\mu v^\mu \tag{2.234}$$

it follows that

$$V = 2 \int_{-\infty}^{\infty} \frac{\delta(S^2)}{2 Q_\mu v^\mu}\, d(S^2) \qquad (\tau < 0)$$

and, using (2.114)

$$\frac{\partial V}{\partial x^\mu} = \int_{-\infty}^{\infty} \delta(S^2) \frac{\partial}{\partial x^\mu}\left(\frac{1}{Q_n v^n}\right) d(S^2)$$

which is equivalent to (2.231).

The right hand side of (2.233) may, however be differentiated directly in accordance with Theorem I of Section 2.1. Using (2.227) and (2.234) there obtains

$$\frac{\partial V}{\partial x^\mu} = 2 \int_{-\infty}^{0} \delta'(S^2)\; 2\, Q_\mu\, d\tau$$

$$= -2 \int_{-\infty}^{0} \delta'(S^2) \frac{\partial (S^2)}{\partial \tau} \frac{Q_\mu}{Q_n v^n}\, d\tau$$

$$= 2 \int_{-\infty}^{0} \delta(S^2) \frac{\partial}{\partial \tau}\left(\frac{Q_\mu}{Q_n v^n}\right) d\tau$$

by partial integration,

$$= \int_{-\infty}^{\infty} \frac{\delta(S^2)\, \frac{\partial}{\partial \tau}\left(\frac{Q_\mu}{Q_n v^n}\right) d(S^2)}{Q_m v^m} \qquad (\tau < 0)$$

$$= \frac{1}{D_R} \frac{\partial}{\partial \tau}\left(\frac{Q_\mu}{Q_n v^n}\right)_{(S^2=0)}$$

$$= \frac{1}{(D_R)^2}\left(\frac{\partial Q_\mu}{\partial \tau}\right)_{S^2=0} \quad \frac{P_\mu}{D_R^3}\left(\frac{\partial (Q_n v^n)}{\partial \tau}\right)_{S^2=0} \tag{2.235}$$

A comparison of (2.235) with (2.232) yields

$$\frac{\partial Q_\mu}{\partial \tau} = - \frac{\partial (Q_n v^n)}{\partial x^\mu} \tag{2.236}$$

On the surface (2.236) seems to be a meaningless equation, for the left hand side does not appear to be a covariant vector, being the ordinary derivative of a vector. However the parameter τ does not involve the coordinate x^μ, the vector Q^μ being a function of the five independent variables x^μ and τ. The differentiation with respect to τ does not involve the infitesimal displacements which require the use of Christoffel symbols in order to maintain the tensor character of derivatives.

The derivation resulting in equation (2.227), (2.230) and (2.236) illustrates the power of the techniques resulting from the theory of distributions, but as yet it has not been found possible to use these results or to obtain others by means of which to derive adjunct-field equations in general Riemannian space. Instead, in the next section use is made of Method I to establish field equations in a particular type of space-time, utilizing as the method implies, a particular coordinate system.

2.3. MATHEMATICAL EXERCISES IN DE SITTER SPACE-TIME

Introduction

The notation for this section is carried over from the last. We again consider a world line L and a general point $X(x^\mu)$. Γ is a null line through X intersecting L at $X_0(x_0{}^\mu)$ and $v_0{}^\mu$ is the velocity vector of L at X_0. v^μ is the vector formed at X by parallel transfer of the vector $v_0{}^\mu$ from X_0 along Γ to X.

The quantities available in terms of which to define a tensor potential at X* in terms of L and Γ are the vectors P^μ and v^μ and the scalar D. The purpose of the present section is to evaluate the covariant derivatives of these quantities in de Sitter space-time, to make available adjunct-field equations for any potential defined in terms of them. In particular we wish to devise a vector potential A^μ such that if a skew-symmetric tensor A_{mn} is defined by

$$A_{mn} = \frac{\partial A_m}{\partial x^n} - \frac{\partial A_n}{\partial x^m} \qquad (2.301)$$

then the divergence of the corresponding contravariant tensor

$$A^{mn}{}_{|n} \qquad (2.302)$$

vanishes at any point X not on L. (The purpose of this scheme is of course to incorporate Maxwell's equations into a de Sitter space particle electrodynamics).

---

* Until near the end of the section points on L are excluded, and the results are to apply using either Γ_R or Γ_S.

The mathematical operations in this section will be carried out with explicit reference to the metric

$$dS^2 = dt^2 - ((dx^1)^2 + (dx^2)^2 + (dx^3)^2) \exp 2kt \qquad (2.303)$$

but the final results will be expressed in tensor form. (Here k is the constant parameter of this space of constant curvature).

The Christoffel symbols of the second kind for the metric (2.303) are given by

$$\left\{ {4 \atop i\ i} \right\} = k \exp 2kt \qquad (i = 1, 2, 3) \qquad (2.304)$$

$$\left\{ {i \atop 4\ i} \right\} = k \qquad (2.305)$$

all others being zero.

The calculations of the covariant derivatives of D, P^μ, and v^μ is a very lengthy process, and only sufficient detail is included here to illustrate the method. A list of the quantities used and the results obtain is to be found in (2.356) to (2.359) near the end of the section.

<u>Null Surface</u>

Letting y^μ again be coordinates along Γ, the differential equations satisfied by y^μ are

$$T \equiv \frac{1}{2} g_{mn} \frac{dy^m}{du} \frac{dy^n}{du} = 0 \qquad (2.306)$$

and $$\frac{dp^\mu}{du} + \left\{ {\mu \atop m\ n} \right\} p^m p^n = 0$$

where $p^{\mu} = \frac{dy^{\mu}}{du}$, or equivalently

$$\frac{d}{du}\left(\frac{\partial T}{\partial p^n}\right) - \frac{\partial T}{\partial y^n} = 0 \qquad (2.307)$$

With respect to the metric (2.303), (2.306) becomes

$$\left(\frac{dy^4}{du}\right)^2 = \delta_{ij} \frac{dy^i}{du} \frac{dy^j}{du} \exp 2ky^4 \quad (i,j, = 1,2,3) \qquad (2.308)$$

where δ_{ij} is the ordinary Kronecker delta, and the first integral of (2.307), for $n = i$, is

$$\frac{dy^i}{du} \exp 2ky^4 = a^i, \text{ where } a^i \text{ is a constant.} \qquad (2.309)$$

Putting $\delta_{ij}\, a^i a^j = A^2$, (2.308) reduces to

$$\frac{dy^4}{du} = A \exp(-ky^4) \qquad (2.310)$$

with the solution

$$\exp ky^4 - \exp kt_o = kAu$$

(where $y^4 = t_o$ when $u = 0$)

or $\exp kt - \exp kt_o = kAU$ by (2.2021). (2.311)

Using (2.310) in (2.309) we obtain

$$\frac{dy^i}{dy^4} = \frac{a^i}{A} \exp(-ky^4) \qquad (2.312)$$

or upon integrating

$$y^i - x_o{}^i = \frac{a^i}{kA} \left(\exp(-kt_o) - \exp(-ky^4)\right) \qquad (2.313)$$

Thus if X and X_o lie on a common null geodesics then

$$x^i - x_o^i = \frac{a^i}{kA}\left[\exp(-kt_o) - \exp(-kt)\right] \qquad (2.314)$$

must be satisfied.

Squaring both sides of (2.314) and summing through i, we have

$$r^2 = \frac{1}{k^2}\left[\exp(-kt_o) - \exp(-kt)\right]^2 \qquad (2.315)$$

where $r^2 = \sum_i (x^i - x_o^i)^2 \equiv \delta_{ij}(xx_o^i)(xx_o^j)$. (2.316)

Equation (2.315) may be written

$$\pm r = \frac{1}{k}\left[\exp(-kt_o) - \exp(-kt)\right] \qquad (2.317)$$

where now we must distinguish between past and future null lines through X, the positive sign in (2.317) referring to Γ_R and the negative sign to Γ_S. By avoiding the use of (2.317) in what follows, we avoid the immediate need of distinguishing between Γ_R and Γ_S.

Evaluation of the vectors P^μ and P_o^μ.

By (2.311)

$$A = \frac{\exp kt - \exp kt_o}{kU}.$$

Substituting this value of A in (2.310) at X and X_o, there obtains

$$Up^4 = \left[\frac{\exp kt - \exp kt_o}{k}\right]\exp(-kt)$$

$$Up_o^4 = \left[\frac{\exp kt - \exp kt_o}{k}\right]\exp(-kt_o)$$

Also substituting in (2.312) and using (2.313), we have

$$Up^i = (xx_o^{\ i}) \exp k(t_o - t)$$

$$Up_o^{\ i} = (xx_o^{\ i}) \exp k(t - t_o).$$

Also noting that

$$Up^\mu \equiv P^\mu$$

we have in summary

$$\left.\begin{aligned} P^4 = P_4 &= \frac{1 - \exp k(t_o - t)}{k} \\ P_o^{\ 4} = P_{o_4} &= \frac{\exp k(t - t_o) - 1}{k} \\ P^i &= (xx_o^{\ i}) \exp k(t_o - t) \\ P_o^{\ i} &= (xx_o^{\ i}) \exp k(t - t_o) \\ P_{o_i} = P_i &= -(xx_o i) \exp k(t + t_o) \end{aligned}\right\} \qquad (2.318)$$

Parallel transfer

The differential equations of the vector

$$v^\mu$$

which vector is obtained by parallel transfer* of $v_o^{\ \mu}$ at X_o to X along Γ, are

$$\frac{dv^\mu}{du} + \left\{ {\mu \atop m\ n} \right\} v^m p^n = 0 .$$

Using (2.304) and (2.305), these reduce to

$$\left.\begin{aligned} \frac{dv^4}{du} + k\, \delta_{ij}\, v^i p^j \exp 2ky^4 &= 0 \\ \frac{dv^i}{du} + k(v^i p^4 + v^4 p^i) &= 0 \end{aligned}\right\} \qquad (2.319)$$

Now $\delta_{ij}\, v^i p^j \exp 2ky^4 = -v^i p_i = -(v^\mu p_\mu - v^4 p_4)$.

It is convenient to use this last equation in the first of (2.319), since $v^\mu p_\mu$ is constant along Γ. Thus

$$\frac{dv^4}{du} + kv^4 p_4 = kv^\mu p_\mu$$

or, using (2.310)

$$\frac{dv^4}{du} + k\,Av^4 \exp(-ky^4) = k\,v^\mu p_\mu$$

the solution of which is

$$v^4 = B\exp(-ky^4) + \frac{v^\mu p_\mu}{2A}\exp ky^4$$

where B is a constant of integration. Noting that

$$v^4 = v_o{}^4 \text{ when } y^4 = t_o$$

we are able to eliminate the constant B. Also, eliminating A by (2.311), we have for the point X,

$$v^4 = \frac{k\,v^\mu p_\mu}{2} + \left(v_o{}^4 + \frac{kv^\mu p_\mu}{2}\right)\exp k(t_o - t). \tag{2.320}$$

Now we are able to use (2.320) with (2.310) to (2.313) in the last of (2.319) to obtain the following differential equation for v^i:

$$\frac{dv^i}{dy^4} + kv^i + \frac{p^\mu v_\mu}{2A^2} + \left[\frac{a^i}{A}v_o{}^4 \exp kt_o - \frac{p^\mu v_\mu}{2A^2}\exp 2kt_o\right]\exp(-2ky^4) = 0$$

the solution of which is

$$v^i = C^i\exp(-ky^4) - \frac{p_\mu v^\mu}{2kA^2}$$

$$+ \left[\frac{a^i}{A}v_o{}^4 \exp kt_o - \frac{p_\mu v^\mu}{2A^2}\exp 2kt_o\right]\frac{\exp(-2ky^4)}{k}$$

where C^i is a constant of integration. Eliminating C^i by noting that

$$v^i = v_o^i \text{ when } y^4 = t_o$$

and also eliminating a^i and A by (2.313) and (2.311) respectively, we find that at X

$$v^i = \left[v_o^i - k(xx_o^i)(v_o^4 + \frac{kv^\mu P_\mu}{2})\right]\exp k(t_o-t). \qquad (2.321)$$

The formulae (2.320) and (2.321) hold, of course, for any vector formed at X by parallel transfer along Γ from X_o. In particular if we form the acceleration vector

$$f_o^\mu \equiv \frac{dv_o^\mu}{d\tau} + \left\{ \begin{matrix} \mu \\ m\ n \end{matrix} \right\} v_o^m v_o^n \qquad (2.322)$$

or, using (2.304) and (2.303)

$$f_o^4 = \frac{dv_o^4}{d\tau} - kg_{o\,ij}\, v_o^i v_o^j \quad \begin{matrix}(i = 1,2,3)\\(j = 1,2,3)\end{matrix}$$

$$= \frac{dv_o^4}{d\tau} + k\left[(v_o^4)^2 - 1\right] \qquad (2.323)$$

(since v^μ is a unit vector)

and using (2.305)

$$f_o^i = \frac{dv_o^i}{d\tau} + 2k\, v_o^i v_o^4 \qquad (2.324)$$

then upon parallel transfer to X we have

$$f^4 = \frac{kf^\mu P_\mu}{2} + \left(f_o^4 + \frac{kf^\mu P_\mu}{2}\right)\exp k(t_o-t) \qquad (2.325)$$

$$f^i = \left[f_o^i - k(xx_o^i)(f_o^4 + \frac{kf^\mu P_\mu}{2})\right]\exp k(t_o-t) \qquad (2.326)$$

Formula for distance

We have defined

$$D \equiv P_\mu v^\mu$$

$$= P_{o_\mu} v_o^{\ \mu}.$$

Using (2.318) we see that

$$D = \frac{\exp k(t-t_o) - 1}{k} v_o^{\ 4} - \delta_{ij} v_o^{\ i}(xx_o^{\ j}) \exp k(t+t_o). \qquad (2.327)$$

Also if we define the scalar λ by

$$\lambda \equiv P_\mu f^\mu \qquad (2.3271)$$

$$= P_{o_\mu} f_o^{\ \mu}$$

then

$$\lambda = \frac{\exp k(t-t_o) - 1}{k} f_o^{\ 4} - \delta_{ij} f_o^{\ i}(xx_o^{\ j}) \exp k(t+t_o) \qquad (2.328)$$

Derivatives

We now proceed to determine the absolute derivatives of some of the quantities evaluated in the preceding section. Ordinary derivatives of a quantity (say z) are found according to the formula (2.232).

$$\frac{\partial z}{\partial x^\mu} = \left(\frac{\partial z}{\partial x^\mu}\right)_{(x_o^{\ \mu})} + \left(\frac{\partial z}{\partial \tau}\right)_{(x^\mu)} \frac{P_\mu}{D}. \qquad (2.329)$$

If z is a component of a tensor, then of course the appropriate quantities will be added to this ordinary derivative to obtain the absolute derivative.

DIFFERENTIATION OF DISTANCE

We seek here a covariant expression for

$$\frac{\partial D}{\partial x^{\mu}}$$

in terms of the vectors P_μ, v_μ, f_μ, and the scalars D and λ. Noting the explicit formula (2.327) for D, and using (2.329), we have:

$$\frac{\partial D}{\partial t} = v_o{}^4 \exp k(t-t_o) - (v_o{}^4)^2 \frac{P_4}{D} \exp k(t-t_o)$$

$$+ \frac{\exp k(t-t_o) - 1}{k} \frac{dv_o{}^4}{d\tau} \frac{P_4}{D} - k\delta_{ij}\, v_o{}^i(xx_o{}^j) \exp k(t-t_o)$$

$$- k\,\delta_{ij}\, v_o{}^i(xx_o{}^j)v_o{}^4 \frac{P_4}{D} \exp k(t+t_o) - \delta_{ij} \frac{dv_o{}^i}{d\tau} (xx_o{}^j)\frac{P_4}{D} \exp k(t+t_o)$$

$$+ \delta_{ij}\, v_o{}^i\, (v_o{}^4) \frac{P_4}{D} \exp k(t+t_o) \qquad (2.3291)$$

$$= \frac{P_4\lambda}{D} + \frac{P_4}{D}\left[- 1 + (v_o{}^4)^2 + \delta_{ij}\, 2k\, v_o{}^i\, v_o{}^4(xx_o{}^j) \exp k(t+t_o)\right]$$

$$+ \frac{P_4}{D}\left[- (v_o{}^4)^2 \exp k(t-t_o) + \delta_{ij}\, 2kv_o{}^i\, v_o{}^4\, (xx_o{}^j) \exp k(t+t_o)\right]$$

$$- k\,\delta_{ij}\, v_o{}^i(xx_o{}^j) \exp k(t+t_o) - v_o{}^4 \exp k(t-t_o) \qquad (2.330)$$

The first two terms of (2.330) are derived from the third and the last two of (2.3291) using (2.323) and (2.328). Now, using (2.327), (2.330) reduces to

$$\frac{\partial D}{\partial t} = \frac{P_4\lambda}{D} - \frac{P_4}{D} - k\, v_o{}^4\, P_4$$

$$+ v_o{}^4 \exp k(t-t_o) - k\,\delta_{ij}\, v_o{}^i(xx_o{}^j) \exp k(t+t_o)$$

With the help of (2.318), (2.320) and (2.327) this finally becomes

$$\frac{\partial D}{\partial t} = \frac{P_4}{D}(\lambda - 1) + v_4 + \frac{k^2}{2} D P_4 \qquad (2.331)$$

By a similar procedure it can be shown that

$$\frac{\partial D}{\partial x^i} = \frac{P_i}{D}(\lambda - 1) + v_i + \frac{k^2}{2} D P_i \qquad (2.332)$$

or, combining (2.331) and (2.332)

$$\frac{\partial D}{\partial x^\mu} = \frac{P_\mu}{D}(\lambda - 1) + v_\mu + \frac{k^2}{2} D P_\mu . \qquad (2.333)$$

DIFFERENTIATION OF P^μ

We require here the absolute derivative

$$P^m_{\ |n} \equiv \frac{\partial P^m}{\partial x^n} + \left\{ {m \atop n\ \sigma} \right\} P^\sigma .$$

Noting that $\left\{ {4 \atop 4\ \sigma} \right\} = 0$, and using (2.318) and (2.329)

$$P^4_{\ |4} = (1 - v_o^{\ 4} \frac{P_4}{D}) \exp k(t_o - t)$$

$$= \exp k(t_o - t) - \frac{v^4 P_4}{D} + \frac{k}{2} P_4 \, (1 + \exp k(t_o - t))$$

using (2.320)

$$= 1 - \frac{v^4 P_4}{D} - \frac{k^2}{2} P^4 P_4 \qquad (2.334)$$

using (2.318).

Also, noting (2.304)

$$P^4{}_{|i} = -v_o{}^4 \frac{P_i}{D} \exp k(t_o-t) + k\,P^i \exp 2kt$$

$$= -\frac{v^4P_i}{D} + \frac{kP_i}{2}\left(1 + \exp k(t_o-t)\right) - kP_i$$

$$= -\frac{v^4P_i}{D} - \frac{k^2}{2} P^4P_i \qquad (2.335)$$

again using (2.320) and (2.318).

Furthermore noting (2.305)

$$P^i{}_{|4} = -v_o{}^i \frac{P_4}{D} \exp k(t_o-t) - (xx_o{}^i)k \exp k(t_o-t)$$

$$+ (xx_o{}^i)k^2 v_o{}^4 \frac{P_4}{D} P^i \exp k(t_o-t)$$

$$= -v^iP_4 - \frac{k^2}{2}(xx_o{}^i)P_4 \exp k(t_o-t) - k(xx_o{}^i)\exp k(t_o-t) + kP^i$$

$$= -\frac{v^iP_4}{D} - \frac{k^2}{2} P^iP_4 \qquad (2.336)$$

Similarly $P^i{}_{|j} = \delta^i_j - \frac{v^iP_j}{D} - \frac{k^2}{2} P^iP_j.$ (2.337)

Equations (2.334) to (2.337) are equivalent to

$$P^m{}_{|n} = \delta^m_n - \frac{v^mP_n}{D} - \frac{k^2}{2} P^mP_n \qquad (2.338)$$

DIFFERENTIATION OF v^{μ}

It is convenient here to express v^{μ} as the sum of two vectors. We have in (2.320)

$$v^4 = \frac{kD}{2}\left[\exp k(t_o-t) - 1\right] + kD + v_o^{\,4}\exp k(t_o-t)$$

$$= \frac{-k^2D}{2}P^4 + kD + v_o^{\,4}\exp k(t_o-t)$$

using (2.318), and we also have in (2.321),

$$v^i = -\frac{k^2D}{2}(xx_o^{\,i})\exp k(t_o-t) + \left[v_o^{\,i} - k(xx_o^{\,i})v_o^{\,4}\right]\exp k(t_o-t)$$

$$= -\frac{k^2D}{2}P^i + \left[v_o^{\,i} - k(xx_o^{\,i})v_o^{\,4}\right]\exp k(t_o-t)$$

also using (2.318).

Hence we may write

$$v^{\mu} = -\frac{k^2D}{2}P^{\mu} + m^{\mu} \tag{2.339}$$

where $m^4 = kD + v_o^{\,4}\exp k(t_o-t)$ (2.340)

and $m^i = \left[v_o^{\,i} - k(xx_o^{\,i})v_o^{\,4}\right]\exp k(t_o-t).$ (2.341)

We now proceed to find the covariant derivatives

$$m^{\mu}_{\ |n} \equiv \frac{\partial m^{\mu}}{\partial x^n} + \left\{ {\mu \atop n\ \sigma} \right\} m^{\sigma}$$

by using (2.329), (2.304) and (2.305). Noting (2.333) we have

$$m^4_{\ 44} = k\left[\frac{P_4}{D}(\lambda-1) + v_4 + \frac{k^2}{2} D P_4\right]$$

$$+ k\,(v_o{}^4)^2\,\frac{P_4}{D}\exp k(t_o-t) - k\,v_o{}^4 \exp k(t_o-t)$$

$$+ \frac{dv_o{}^4}{d\tau}\,\frac{P_4}{D}\exp k(t_o-t)$$

$$= \frac{f^4P_4}{D} - \frac{k\lambda}{2}\left[1 + \exp k(t_o-t)\right]\frac{P_4}{D} + \frac{kP_4}{D}\exp k(t_o-t)$$

$$+ k\left[\frac{P_4}{D}\lambda - \frac{P_4}{D} + v_4 + \frac{k^2}{2} D P_4\right] - k\,v_o{}^4 \exp k(t_o-t)$$

using (2.323), (2.325) and (2.3271)

$$= \frac{f^4P_4}{D} + \frac{k^2\lambda}{2D} P^4P_4 + k^2D - \frac{k^2}{D} P^4P_4$$

using (2.320) and (2.318).

Also $m^i_{\ 4j} = km^i\,\frac{P_j}{D}\,v_o{}^4 + \left[\frac{dv_o{}^i}{d\tau}\,\frac{P_j}{D} - k(xx_o{}^i)\,\frac{dv_o{}^4}{d\tau}\,\frac{P_j}{D}\right.$

$$\left. - \delta^i_j\,kv_o{}^4 + k\,v_o{}^i v_o{}^4\,\frac{P_j}{D}\right]\exp k(t_o-t) + \delta^i_j\,km^4$$

$$= \frac{km^iP_j}{D}\,v_o{}^4 + \frac{f^iP_j}{D} + \delta^i_j\,km^4$$

$$+ \left[\frac{k^2}{2}(xx_o{}^i)\lambda\,\frac{P_j}{D} - 2kv_o{}^i v_o{}^4\,\frac{P_j}{D} + k^2(xx_o{}^i)\left\{(v_o{}^4)^2 - 1\right\}\frac{P_j}{D}\right.$$

$$\left. - \delta^i_j\,kv_o{}^4 + kv_o{}^i v_o{}^4\,\frac{P_j}{D}\right]\exp k(t_o-t)$$

using (2.323), (2.324), (2.326) and (2.3271).

$$= \frac{f^iP_j}{D} + \frac{k^2\lambda}{2D} P^iP_j + \delta^i_j\,k^2D - \frac{k^2}{D} P^iP_j$$

using (2.340), (2.341) and (2.318).

The other derivatives can be likewise determined, and we find that

$$m^{\mu}{}_{|n} = \frac{f^{\mu}P_n}{D} + \frac{k^2\lambda}{2D} P^{\mu}P_n + \delta^{\mu}_n k^2 D - \frac{k^2}{D} P^{\mu}P_n . \tag{2.342}$$

It follows from (2.339), (2.338), (2.333) and (2.342) that

$$v^{\mu}{}_{|n} = \frac{f^{\mu}P_n}{D} + \delta^{\mu}_n \frac{k^2}{2} D - \frac{k^2}{2D} P^{\mu}P_n + \frac{k^2}{2}(v^{\mu}P_n - P^{\mu}v_n). \tag{2.343}$$

DIFFERENTIATION OF λ

We require the derivatives of the scalar λ, as given by the expression (2.328). But first we define the super-acceleration vector s^{μ} by

$$s_o{}^{\mu} \equiv \frac{df_o{}^{\mu}}{d\tau} + \left\{ {\mu \atop m\ n} \right\} f_o{}^m v_o{}^n \tag{2.344}$$

or, by (2.304) and (2.305)

$$s_o{}^4 = \frac{df_o{}^4}{d\tau} + k\ \delta_{ij}\ v_o{}^i f_o{}^j \exp 2kt_o \tag{2.345}$$

$$\text{and } s_o{}^i = \frac{df_o{}^i}{d\tau} + k(v_o{}^i f_o{}^4 + f_o{}^i v_o{}^4). \tag{2.346}$$

Now since $v_o{}^{\mu} v_{o_{\mu}} = 1$

and since $f_o{}^{\mu}$ is the absolute derivative of $v_o{}^{\mu}$ with respect to the parameter τ, we have

$$v_o{}^{\mu} f_{o_{\mu}} \equiv v^{\mu} f_{\mu} = 0. \tag{2.347}$$

Hence, using (2.303), (2.345) becomes

$$s_o{}^4 = \frac{df_o{}^4}{d\tau} + kv_o{}^4 f_o{}^4 \tag{2.348}$$

We also define a scalar σ by

$$\sigma = s^{\mu}P_{\mu}$$
$$\equiv s_{o}^{\mu}P_{o_{\mu}}. \qquad (2.349)$$

Now, using (2.328) and (2.329),

$$\frac{\partial \lambda}{\partial x^4} = f_o^{\ 4} \exp k(t-t_o) - v_o^{\ 4} \frac{P_4}{D} f_o^{\ 4} \exp k(t-t_o)$$

$$+ P_{o_4} \frac{df_o^{\ 4}}{d\tau} \frac{P_4}{D} - k\ \delta_{ij}\ f_o^{\ i}(xx_o^{\ j}) \exp k(t+t_o)$$

$$- k\ \delta_{ij}\ f_o^{\ i}(xx_o^{\ j}) v_o^{\ 4} \frac{P_4}{D} \exp k(t+t_o)$$

$$- \delta_{ij} \frac{df_o^{\ i}}{d\tau} \frac{P_4}{D} (xx_o^{\ j}) \exp k(t+t_o)$$

$$+ \delta_{ij}\ f_o^{\ i}\ v_o^{\ j} \frac{P_4}{D} \exp k(t+t_o)$$

$$= \frac{P_4}{D}\sigma + f_o^{\ 4} \exp k(t-t_o) - v_o^{\ 4} \frac{P_4}{D} f_o^{\ 4} \exp k(t-t_o)$$

$$- k\ \delta_{ij}\ f_o^{\ i}(xx_o^{\ j}) \exp k(t+t_o) + \frac{P_4}{D} v_o^{\ 4} f_o^{\ 4}$$

$$+ k \frac{P_4}{D} \delta_{ij}(xx_o^{\ i}) v_o^{\ j}\ f_o^{\ 4} \exp k(t+t_o)$$

using (2.318), (2.346), (2.348) and (2.349)

$$= \frac{P_4}{D}\sigma + f_4 + \frac{k^2}{2}\lambda P_4$$

using (2.325), (2.3271) and (2.318).

Similarlythe other derivatives can be evaluated and we find that

$$\frac{\partial \lambda}{\partial x^{\mu}} = \frac{P_{\mu}\sigma}{D} + f_{\mu} + \frac{k^2}{2}\lambda P_{\mu} \qquad (2.350)$$

DIVERGENCE OF f^{μ}

Finally we seek to evaluate the divergence

$$f^{\mu}_{\ |\mu} .$$

We note from (2.325), (2.326) and (2.318) that the vector of f^{μ} may be expressed as

$$f^{\mu} = - \frac{k^2}{2} \lambda P^{\mu} + n^{\mu} \qquad (2.351)$$

where

$$n^4 = k\lambda + f_o^{\ 4} \exp k(t_o - t) \qquad (2.352)$$

$$n^i = \left[f_o^{\ i} - k(xx_o^{\ i}) \ f_o^{\ 4} \right] \exp k(t_o - t) \qquad (2.353)$$

Now

$$n^{\mu}_{\ |\mu} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{\mu}} (\sqrt{g} \ n^{\mu})$$

$$= 3k \ n^4 + \frac{\partial n^{\mu}}{\partial x^{\mu}}$$

from (2.303)

$$= 3kn^4 + k \frac{P_4 \sigma}{D} + k \ f_4 + \frac{k^3}{2} \lambda P_4$$

$$- k \ f_o^{\ 4} \exp k(t_o - t) + k \ f_o^{\ 4} \frac{P_4}{D} v_o^{\ 4} \exp k(t_o - t)$$

$$+ \frac{df_o^{\ 4}}{d\tau} \frac{P_4}{D} \exp k(t_o - t) + k \ n^i \frac{P_i}{D} v_o^{\ 4} \exp k(t_o - t)$$

$$+ \frac{df_o^{\ i}}{d\tau} \frac{P_i}{D} \exp k(t_o - t) - 3 \ k \ f_o^{\ 4} \exp k(t_o - t)$$

$$+ k \ v_o^{\ i} f_o^{\ 4} \frac{P_i}{D} \exp k(t_o - t) - (xx_o^{\ i}) \frac{df_o^{\ 4}}{d\tau} \frac{kP_i}{D}$$

$$\times \ \exp k(t_o - t) .$$

Here we have once again used (2.329). Now using (2.346), (2.348), (2.349), (2.3271), (2.318), (2.352) and (2.353), and noting that

$$P^{\mu}P_{\mu} = 0 \tag{2.3531}$$

the above reduces to

$$n^{\mu}_{\ |\mu} = 4k^2\lambda + \frac{\sigma}{D} \tag{2.354}$$

We can now derive the divergence of f^{μ} from (2.351), (2.350), (2.338) and (2.354), giving

$$f^{\mu}_{\ |\mu} = 2k^2\lambda + \frac{\sigma}{D}\ . \tag{2.355}$$

Summary of Formulae

For convenience a list of formulae for this section are now written down.

BASIC FORMULAE

$$\left.\begin{array}{l} \dfrac{dp^{\mu}}{du} + \left\{ {\mu \atop m\ n} \right\} p^m p^n = 0 \quad \text{defines scalar } u. \\[2ex] U \equiv u(X) - u(X_0) \\ P^{\mu} \equiv U\, p^{\mu} \\ P^{\mu}P_{\mu} = 0 \\ d\tau^2 \equiv dx_0^{\ \mu} dx_{0_\mu} \\[1ex] v_0^{\ \mu} \equiv \dfrac{dx_0^{\ \mu}}{d\tau} \quad \text{is velocity vector on L.} \\[2ex] f_0^{\ \mu} \equiv \dfrac{dv_0^{\ \mu}}{d\tau} + \left\{ {\mu \atop m\ n} \right\} v_0^{\ m} v_0^{\ n} \\[2ex] s_0^{\ \mu} \equiv \dfrac{df_0^{\ \mu}}{d\tau} + \left\{ {\mu \atop m\ n} \right\} v_0^{\ m} f_0^{\ n} \end{array}\right| \tag{2.356}$$

$$
\left.\begin{aligned}
v^m_{|n} P^n &= 0 \\
f^m_{|n} P^n &= 0 \\
s^m_{/n} P^n &= 0
\end{aligned}\right\} \quad \text{equations of parallel transfer}
$$

$$
\begin{aligned}
&P^m_{/n} P^n = P^m && \text{follows from the first three equations} \\
&D \equiv P^\mu v_\mu \\
&\lambda \equiv P^\mu f_\mu \\
&\sigma \equiv P^\mu s_\mu \\
&v^\mu v_\mu = v_o{}^\mu v_{o_\mu} = 1 \\
&f^\mu v_\mu = f_o{}^\mu v_{o_\mu} = 0 && \text{follows from the previous equation} \\
&m^\mu \equiv v^\mu + \frac{k^2 D}{2} P^\mu \\
&n^\mu \equiv f^\mu + \frac{k^2 \lambda}{2} P^\mu \\
&M^\mu \equiv m^\mu / D \\
&V^\mu \equiv v^\mu / D \\
&M_{\mu n} \equiv \frac{\partial M_\mu}{\partial x^n} - \frac{\partial M_n}{\partial x^\mu} = M_{\mu|n} - M_{n|\mu} \\
&V_{\mu n} = \frac{\partial V_\mu}{\partial x^n} - \frac{\partial V_n}{\partial x^\mu}
\end{aligned}
\tag{2.356}
$$

COVARIANT DERIVATIVES

The following are the list of covariant derivatives evaluated in this section

$$
\begin{aligned}
\frac{\partial D}{\partial x^\mu} &= \frac{P_\mu}{D} (\lambda - 1) + v_\mu + \frac{k^2}{2} D P_\mu \\
\frac{\partial \lambda}{\partial x^\mu} &= \frac{P_\mu}{D} \sigma + f_\mu + \frac{k^2}{2} \lambda P_\mu
\end{aligned}
\tag{2.357}
$$

$$\left.\begin{aligned}
P^{m}_{|n} &= \delta^{m}_{n} - \frac{v^{m}P_{n}}{D} - \frac{k^{2}}{2} P^{m}P_{n} \\
m^{\mu}_{|n} &= \frac{f^{\mu}P_{n}}{D} + \frac{k^{2}\lambda}{2D} P^{\mu}P_{n} + \delta^{\mu}_{n}\, k^{2}D - \frac{k^{2}}{D} P^{\mu}P_{n} \\
v^{m}_{|n} &= \frac{f^{m}P_{n}}{D} + \delta^{m}_{n}\frac{k^{2}}{2} D - \frac{k^{2}}{2D} P^{m}P_{n} + \frac{k^{2}}{2}\left(v^{m}P_{n} - P^{m}v_{n}\right) \\
f^{n}_{|n} &= 2k^{2}\lambda + \frac{\sigma}{D}
\end{aligned}\right\} \qquad (2.357)$$

COROLLARIES

The following equalities can be derived from (2.356) and (2.357) by a straightforward though sometimes lengthy process of substitution.

$$\left.\begin{aligned}
g^{mn}\left(\frac{1}{D}\right)_{|mn} &= \frac{2k^{2}}{D} \\
V^{\mu}_{|\mu} &= \frac{3}{2} k^{2} \\
P^{\mu}_{/\mu} &= 3 \\
M^{\mu}_{|\mu} &= 3k^{2} \\
V_{\mu\sigma} &= \delta^{nm}_{\mu\sigma}\left[\frac{P_{m}f_{n}}{D^{2}} - \frac{P_{m}v_{n}(\lambda-1)}{D^{3}} - v_{m}P_{n}\frac{k^{2}}{2D}\right] \\
V^{\mu n}_{|n} &= -\frac{k^{4}}{2} P^{\mu} + \frac{k^{2}}{D} v^{\mu}
\end{aligned}\right\} \qquad (2.358)$$

$$\left.\begin{aligned}
M_{\mu\sigma} &= \delta^{nm}_{\mu\sigma}\left[\frac{P_{m}f_{n}}{D^{2}} - \frac{P_{m}v_{n}(\lambda-1)}{D^{3}}\right] \\
M^{\mu n}_{|n} &= 0
\end{aligned}\right\} \qquad (2.359)$$

To obtain the corresponding formulae for flat space, $(g_{mn} = \eta_{mn})$, we simply have to put

$$k = 0.$$

It is interesting to note that in de Sitter space, the tensor M_{mn} and its divergence are given by the same formulae as are V_{mn} (or M_{mn}) in flat space-time.

In view of (2.359) the vector

$$M^{\mu}$$

is seen to be a suitable quantity for use in defining an electromagnetic potential in de Sitter space.

Formulae for M_{μ} containing k implicitly

It is well known (and easy to show, using the metric (2.303)) that de Sitter space is of constant curvature, satisfying[9]

$$R_{rsmn} = k^2(g_{rm}g_{sn} - g_{rn}g_{sm}) \tag{2.360}$$

where R_{rsmn} is the covariant curvature tensor. It follows immediately that we may write

$$\begin{aligned} \frac{1}{2} R_{rsmn} v^r P^s P^m &= \frac{k^2}{2} V_r P^r P_n \\ &= \frac{k^2}{2} D P_n \end{aligned} \tag{2.361}$$

Hence from (2.356)

$$M_{\mu} = \frac{v^r}{D} (g_{\mu r} + \frac{1}{2} R_{rsm\mu} P^s P^m) \tag{2.362}$$

This is a vector which makes no explicit reference to the type of space, although of course the theorems (2.359) apply only for de Sitter space and for flat space.

Other alternatives to (2.362) are contained in the following de Sitter 4-space identities.[9]

$$\frac{1}{2} R_{rsmn} v^r P^s P^m = -\frac{D}{3!} P^m R_{mn}$$

$$= -\frac{RDP_n}{4!}$$

$$= \frac{D}{3!} P^m G_{mn}$$

where R_{mn}, R, and G_{mn} are respectively the Ricci tensor, the curvature invariant and the Einstein tensor.

Singularities

The foregoing has applied equally well for either Γ_R or Γ_S, and for all points excluding those on L. We now distinguish between the retarded and the advanced, and at the same time include points on L in the work, by writing

$$M^\mu_{(R)} = \frac{m^\mu}{D_R} \equiv \frac{m^\mu}{D} h(D) = M^\mu h(D) \qquad (2.363)$$

$$M^\mu_{(S)} = \frac{m^\mu}{D_S} \equiv -\frac{m^\mu}{D} h(-D) = -M^\mu h(-D) \qquad (2.364)$$

$$A^\mu = \frac{1}{2} (M^\mu_{(R)} + M^\mu_{(S)}) \qquad (2.365)$$

The quantities $\frac{h(D)}{D}$ and $\frac{h(-D)}{D}$ do not represent ordinary functions for D equal to zero, but they do represent symbolic functions of four-dimensional distributions.

Referring only to (2.363), we have then

$$M^{\mu\sigma}{}_{|\sigma(R)} \equiv g^{\mu n} g^{\sigma m} \left[(M_n h(D))_{|m} - (M_m h(D))_{|n} \right]_{|\sigma}$$

$$= M^{\mu\sigma}{}_{|\sigma} h(D) - M^{\mu\sigma} \frac{\partial h(D)}{\partial x^{\sigma}}$$

$$- g^{\mu m} g^{\sigma n} \left(M_{n|\sigma} \frac{\partial h(D)}{\partial x^m} - M_{m|\sigma} \frac{\partial h(D)}{\partial x^n} \right)$$

$$- g^{\mu m} g^{\sigma n} M_n \left[h(D) \right]_{|m\sigma}$$

$$- g^{\mu m} g^{\sigma n} M_m \left[h(D) \right]_{|n\sigma} \qquad (2.366)$$

Now the first term on the right hand side of (2.366) vanishes by virtue of (2.359), and using (2.121)

$$M^{\mu\sigma}{}_{|\sigma(R)} = - M^{\mu\sigma} \delta(D)\, D_{|\sigma}$$

$$- g^{\mu m} g^{\sigma n} \left(M_{n|\sigma}\, \delta(D)\, D_{|m} - M_{m|\sigma}\, \delta(D) D_{|n} \right)$$

$$- g^{\mu m} g^{\sigma n} M_n \left(\delta'(D) D_{|m}\, D_{|\sigma} + \delta(D) D_{|m\sigma} \right)$$

$$- g^{\mu m} g^{\sigma n} M_m \left(\delta'(D) D_{|n}\, D_{|\sigma} + \delta(D) D_{|n\sigma} \right) \qquad (2.367)$$

The derivatives of M^{μ} and D can be written down from (2.537), (2.538) and (2.539), and the right hand side of (2.367) evaluated. However the work can be considerably reduced by anticipation of the answer.

A common factor

$$\frac{\delta(D)}{D^2}$$

is to be taken from the entire right hand side of (2.367), the terms containing $\delta'(D)$ having been suitably adjusted by means of (2.122). We have then

$$M^{\mu n}{}_{|n\,(R)} = \frac{\delta(D)}{D^2} Z^{\mu} \tag{2.368}$$

where Z^{μ} is a vector to be determined.

Now since the right hand side of (2.368) vanishes for any X not on L, we employ local Cartesian coordinates with origin at the point X_0 on L, and oriented in such a way that

$$v_0{}^i = 0$$

$$v_0{}^4 = 1 .$$

For this coordinate system, in the neighbourhood of X_0,

$$g_{\mu n} = \eta_{\mu n}$$

$$\begin{aligned} D &\equiv P_{\mu} v^{\mu} = P_4 = (-P_i\, P^i)^{1/2} \\ &= U\left[\left(\frac{dy^1}{du}\right)^2 + \left(\frac{dy^2}{du}\right)^2 + \left(\frac{dy^3}{du}\right)^2\right]^{1/2} \\ &= \left((x^1)^2 + (x^2)^2 + (x^3)^2\right)^{1/2} \end{aligned}$$

using the third and fourth equation of (2.356) and the fact that here

$$x^i = U\,\frac{dy^i}{du}$$

Thus in the neighbourhood of X_o, D plays the role of r in (2.125), and we have

$$\frac{\delta(D)}{4\pi D^2} = \delta(x^1)\delta(x^2)\delta(x^3) \qquad (2.369)$$

If we use (2.369) in (2.368), we see that we need only evaluate Z^μ at X_o (i.e. for D = 0). This is a considerable simplification, and we find that

$$Z^\mu(X_o) = v_o{}^\mu$$

and (2.368) becomes

$$M^{\mu n}{}_{|n(R)} = 4\pi\delta(x^1)\delta(x^2)\delta(x^3)v^\mu. \qquad (2.370)$$

Implied in (2.370) is the condition that XX_o lies on a null geodesic. Hence when

$$x^i = 0$$

we also have

$$x^4 = 0.$$

This condition can be explicitly contained in (2.370) in the following way. Letting χ^μ be coordinates along L, then (2.370) is equivalent to

$$M^{\mu n}{}_{|n(R)} = 4\pi\delta(x^1-\chi^1)\delta(x^2-\chi^2)\delta(x^3-\chi^3)\int_{-\infty}^{\infty} v^\mu \ \delta(x^4-\chi^4)d\ \chi^4.$$

Now since $v_o{}^4 = 1$ at X_o, therefore

$$M^{\mu n}{}_{|n(R)} = 4\pi\int_{-\infty}^{\infty} \delta(x^1-\chi^1)\delta(x^2-\chi^2)\delta(x^3-\chi^3)\delta(x^4-\chi^4)v^\mu d\tau$$

or, in a general coordinate frame

$$= 4\pi \int_{-\infty}^{\infty} \frac{\delta(x^1-\chi^1)\,\delta(x^2-\chi^2)\,\delta(x^3-\chi^3)\,\delta(x^4-\chi^4)}{\sqrt{g}}\, v^{\mu} d\tau \tag{2.371}$$

by (2.124).

This same result can be shown to hold for the advanced field. That is

$$M^{\mu n}{}_{|n(S)} = M^{\mu n}{}_{|n(R)}. \tag{2.3711}$$

Thus, using (2.365), and defining

$$A_{\mu n} = A_{\mu|n} - A_{n|\mu} \tag{2.372}$$

we finally have the following general result for de Sitter space

$$A^{\mu n}{}_{|n} = 4\pi \int_{-\infty}^{\infty} \frac{\delta(x\chi^1)\,\delta(x\chi^2)\,\delta(x\chi^3)\,\delta(x\chi^4)}{\sqrt{g}}\, v^{\mu} d\tau \quad . \tag{2.373}$$

C H A P T E R III

E L E C T R O D Y N A M I C S

3.1. INTRODUCTION

The work in this chapter evolves from two joint papers (1) (2) by J.A. Wheeler and R.P. Feynman in which a theory of electrodynamics (originally ascribed to Schwarzschild, Tetrode, and Fokker) is reviewed and developed. The basic principles of this theory are as follows:

(a) It is set in the Minkowski space-time of special relativity and is Lorentz invariant.

(b) It is a theory of direct interparticle action at a distance, action taking place between charged particles along the null lines joining them.

(c) It is governed by a single variational principle.

(d) There is no action of any particle on itself.

(e) Past and future are symmetrical in every respect, action being defined identically with respect to either branch of a null cone.

The chief claims of the authors for this theory are:

(f) Equations of motion derivable from the basic variational principle are identical in form with those of Lorentz.

(g) A skew-symmetric second rank tensor field (called an adjunct field) obtained in the derivation of the equations of motion (corresponding to the Maxwellian field of the Lorentz equations) can be shown to satisfy Maxwell's equations.

(h) This tensor field is equal to one half the sum of the retarded and advanced fields obtained by Lienard-Wiechert for Maxwell's theory.

(i) Provided one further assumes the universe to contain a sufficient number of particles to be a "perfect absorber",* then this field at the same time is equal to the full retarded Lienard-Wiechert field of experience plus an additional field claimed by Dirac correctly to account for the phenomenon of radiative damping.

Accordingly then the particle theory of Wheeler and Feynman claims equal status with the field theory of Maxwell in the classical description of nature, and in addition contains the following advantages:

(j) (Adjunct) field equations and equations of motion are both derived from a single variational principle.

(k) All fields are ascribed to and derivable from the motions of charged particles. In Maxwell's field theory free space solutions exist which are not in fact a part of our observable experience.

(l) By being able to account for radiative damping without self action, the structural difficulties encountered by Lorentz in requiring an electron of finite size are completely avoided.

The theory of Wheeler and Feynman, however, is subject to a number of criticisms:

(m) The authors do not correctly prove (g), as claimed. Their proof depends on an equation named by them Dirac's identity, which was in fact written down by Dirac as a corollary to Maxwell's equations.** Thus Maxwell's equations are used to prove Maxwell's equations!

---

* A perfect absorber is defined by the authors to be such that a test (charged) particle placed "outside" the absorber suffers no electrodynamical impetus.

** Which equations are, of course, a physical assumption rather than a mathematical identity.

(n) The concept of a "perfect absorber" is artificial and unrealistic.

(o) The basic theory is too time-symmetrical, and the asymmetry in contention (i) is not adequately accounted for.

In this section a brief outline is given of the particle electrodynamics of Wheeler and Feynman, containing a detailed account of (m) and a correct proof of (g). The equations of planetary motion are derived, and shown to have an interesting significance in the Special Theory of Relativity. Following this is an extension of the theory to the de Sitter Space of Constant Curvature. It will be shown that in de Sitter space, with a Steady-State Cosmology distribution of mass, the theory is no longer symmetrical in time and a new approach to the problem of radiative reaction and advanced effects is obtainable.

3.2. WHEELER AND FEYNMAN ELECTRODYNAMICS

Notation

The various particles are labelled by the letters a, b, ... and their masses and charges in c.g.s. e.s. units by m_a, e_a, etc. respectively. A test particle of vanishing mass and charge is represented by the letter ξ.

The background space is Minkowski, with metric tensor

$$\eta_{mn}{}^{*} = \begin{pmatrix} -1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & 1 \end{pmatrix}$$

so that particle "a" has at a given instant the Cartesian coordinates

$$a^i = -a_i \qquad (i = 1, 2, 3)$$

$$a^4 = a_4 = ct \qquad \text{(velocity of light X time)}$$

and the labelling letter "a" is also used to designate arc length along the particle's world line.

Again as in Section 2.2, if x^μ and y^μ are vectors, the vector $x^\mu - y^\mu$ will be represented for simplicity by xy^μ.

Fundamentals

The single postulate of this electrodynamics is that the world line of each particle (say a) conforms with the following principle of least action:

$$\text{Var}\left[\int(\eta_{mn}\, da^m da^n)^{1/2} + \frac{e_a}{m_a c^2}\int F_r(a)\, da^r\right] = 0 \qquad (3.201)$$

where the integrals in (3.201) are formed along the world line of a.

---

* The components of this metric tensor are of opposite sign to that used by Wheeler and Feynman, hence there will be a discrepancy in sign between some of the formula given by them and as written down here.

The electromagnetic potential is given by

$$F_r(a) = \sum_{b \neq a} e_b \int \delta(ab_\mu ab^\mu) db_r \qquad (3.202)$$

We may also define a potential field (due to all particles) by

$$F_r(x) = \sum_{\text{all } b} e_b \int \delta(ab_\mu ab^\mu) db_r \qquad (3.203)$$

By calculating the variation on the left hand side of (3.201) and equating it to zero, the following equations of motion are established:

$$\frac{d^2 a_m}{da^2} = \frac{e_a}{m_a c^2} F_{nm} (a) \frac{da^n}{da} \qquad (3.204)$$

with

$$F_{mn} = \frac{\partial F_m}{\partial x^n} - \frac{\partial F_n}{\partial x^m} \qquad (3.205)$$

Equations (3.204) are identical in form with the Lorentz equations of motion for an electron. In the theory of Lorentz the F_{mn} are field quantities which satisfy Maxwell's equations. Wheeler and Feynman seek to show that the adjunct field

$$F_{mn}(x)$$

as given by (3.205) and (3.203) also satisfy Maxwell's equations, with the current four-vector at the point x^μ defined by

$$j_m(x) = \sum_{\text{all } b} e_b \int \delta(x\, b_1)\delta(xb_2)\delta(xb_3)\delta(xb_4) db_m \, . \qquad (3.206)$$

The one set of Maxwell's equations

$$\frac{\partial F_{mn}(x)}{\partial x^r} + \frac{\partial F_{nr}(x)}{\partial x^m} + \frac{\partial F_{rm}(x)}{\partial x^n} = 0 \qquad (3.207)$$

follows immediately from the definition given in (3.205).

Before discussing the other set, namely

$$\frac{\partial F^{mn}}{\partial x^n}(x) = 4\pi\ j^m(x) \tag{3.208}$$

we note that if we write

$$F_m{}^{(b)}(x) \equiv e_b \int \delta(xb_\mu xb^\mu)\, db_m \tag{3.209}$$

and denote the corresponding second rank skew symmetric tensor by

$$F_{mn}{}^{(b)}(x)$$

then since $F_{mn}(x) \left(= \sum_b F_{mn}{}^{(b)}(x) \right)$ is linear in the particles,

it follows that if $F_{mn}{}^{(b)}(x)$ satisfies Maxwell's equations (which are themselves linear), so also does $F_{mn}(x)$. It is required then to prove that

$$\frac{\partial F^{mn(b)}(x)}{\partial x^n} = 4\pi\ j^{m(b)}(x) \tag{3.210}$$

or equivalently, by (3.205) and (3.206)

$$\frac{\partial^2 F^{m(b)}(x)}{\partial x_n\ \ x^n} - \frac{\partial^2 F^{n(b)}(x)}{\partial x_m \partial x^n} = 4\pi e_b \int \delta(xb_1)\delta(xb_2)\delta(xb_3)\delta(xb_4) X db^m \tag{3.212}$$

METHOD OF PROOF GIVEN AND CORRECTED PROOF

Using (3.209) in (3.212), it follows that sufficient conditions for the proof of (3.212) are:

$$\frac{\partial}{\partial x^n} \int \delta(xb_\mu xb^\mu)\, db^n = 0 \tag{3.213}$$

and

$$\frac{\partial}{\partial x_n \partial x^n} \int \delta(xb_\mu xb^\mu)\, db^m \doteq 4\pi \int \delta(xb_1)\delta(xb_2)\delta(xb_3)\delta(xb_4)\, db^m \tag{3.214}$$

It is next assumed* that

$$\frac{\partial}{\partial x^n} \int \delta(xb_\mu xb^\mu)\, db^n = \int \frac{\partial}{\partial x^n} \left[\delta(xb_\mu xb^\mu)\right] db^n \qquad (3.215)$$

and

$$\frac{\partial}{\partial x_n \partial x^n} \int \delta(xb_\mu xb^\mu)\, db^m = \int \frac{\partial}{\partial x_n \partial x^n} \left[\delta(xb_\mu xb^\mu)\right] db^m \qquad (3.216)$$

The legitimacy of (3.215) and (3.216) is not an obvious corollory of generally prescribed properties of δ-functions. A proof of the validity of (3.215) and (3.216) is contained in Theorem I of Section 2.1 of this present work.

Again if we use (3.215) and (3.216) in (3.213) and (3.214) respectively, (3.212) can be shown to be true by proving that

$$\int \frac{\partial}{\partial x^n} \left[\delta(xb_\mu xb^\mu)\right] db^n = 0 \qquad (3.217)$$

$$\frac{\partial}{\partial x_n \partial x^n} \left[\delta(xb_\mu xb^\mu)\right] = 4\pi\, \delta(xb_1)\delta(xb_2)\delta(xb_3)\delta(xb_4) \qquad (3.218)$$

Equation (3.218) is named by Wheeler and Feynman "Dirac's Identity", and a reference is given. Such an identity does not appear explicitly in the reference, but is implied by Dirac in the following manner:

Dirac accepts the right hand side of equation (3.209) as the correct expression for the electromagnetic potential of an electron, and the right hand side of (3.212) as a suitable expression for the

---

* By Wheeler and Feynman.

current due to an electron. Then on the physical assumption of Maxwell's equations, the validity of (3.212) and also of (3.213) and (3.214) is implied.

Such an argument does not constitute a mathematical proof, although it was evidently accepted by Wheeler and Feynman as such for the establishment of (3.218).

A mathematical proof of (3.218) is contained in Theorem II of Section 2.1 of this work.

It is interesting to note that while the justification given by Wheeler and Feynman for (3.218) equally well applies to (3.217), they do offer a separate proof of (3.217). It is worthwhile examining this proof, for it illustrates a dangerous liberty in the use of the δ-function technique. Their proof is as follows:

$$\begin{aligned}
&\int_{-\infty}^{\infty} \frac{\partial}{\partial x^n}\left[\delta(xb_\mu xb^\mu)\right] db^n \\
= &\int_{-\infty}^{\infty} \delta'(xb_\mu xb^\mu)\; 2xb_n \frac{db^n}{db}\, db \\
= - &\int_{-\infty}^{\infty} \frac{\partial}{\partial b}\left[\delta(xb_\mu xb^\mu)\right] db \\
= \;&\delta(xb_\mu xb^\mu)\Big|_{-\infty}^{\infty} \quad \text{by partial integration} \\
= \;&0
\end{aligned} \qquad (3.219)$$

This proof seemingly uses only well established techniques justified by the Theory of Distribution. Before discussing it further, we apply similar techniques to the left hand side of (3.218) to obtain

$$\left.\begin{aligned} \frac{\partial}{\partial x_n \partial x^n}\left[\delta(xb_\mu xb^\mu)\right] &= \frac{\partial}{\partial x_n}\left[\delta'(xb_\mu xb^\mu)2xb_n\right] \\ &= \delta''(xb_\mu xb^\mu)4xb^n xb_n + 8\delta'(xb_\mu xb^\mu) \\ &= 0 \text{ by } (2.122) \end{aligned}\right] \quad (3.220)$$

which is obviously false, in view of (3.218).

The arguments in (3.219) and (3.220) only apply for points not on the world line of b. For the quantity

$$\int \delta(xb_\mu xb^\mu)\,db$$

is equivalent to

$$\int \frac{\delta(xb_\mu xb^\mu)\,d(xb_\mu xb^\mu)}{\dfrac{\partial(xb_\mu xb^\mu)}{\partial b}} \equiv \int \frac{\delta(xb_\mu xb^\mu)\,d(xb_\mu xb^\mu)}{2\,xb_\mu \dfrac{db^\mu}{db}}$$

This is a δ-distribution therefore operating on the function

$$\frac{1}{2xb_\mu \dfrac{db^\mu}{db}} \quad (3.221)$$

of the variable

$$(xb_\mu xb^\mu).$$

In terms of this single variable, however, the quantity (3.221) is only a suitable testing function if it is bounded and possesses derivatives when

$$xb_\mu xb^\mu = 0$$

---

* See Section 2.1.

This condition is clearly not satisfied for

$$xb_\mu = 0$$

or, in other words, when the point X lies on the world line of b.

The difficulty is removed by using* a 4-dimensional distribution in the variables

$$(xb_\mu xb^\mu)$$

and $\quad x^i \qquad (i = 1, 2, 3)$

or in terms of

$$xb_\mu xb^\mu$$

and $\quad r \qquad (r = (-xb_i xb^i)^{1/2})$

and using (2.125).

The requirement that

$$r > 0$$

for the expression

$$\int \delta(xb_\mu xb^\mu)db \tag{3.222}$$

to have meaning can be stated previsely by inclusion of the Heaviside stepfunction. We therefore write

$$\int \delta(xb_\mu xb^\mu)h(r)db$$

instead of (3.222). (It should be noted that h(r) is an invariant, in spite of the fact that r is not).

It is now an easy matter to prove (3.217). Firstly we notice that if we choose coordinates so that

$$\frac{db^4}{db} = 1$$

$$\frac{db^i}{db} = 0$$

---

* As is done in Theorem II of Section 2 and in the proof of 2.373

then

$$\int \delta(xb^{\mu}xb^{\mu})h(r)db^{m}$$

$$= \int \delta(xb^{\mu}xb^{\mu}) \frac{h(r)}{2r} d(xb^{\mu}xb_{\mu}).$$

Now since $\frac{h(r)}{2r}$ is an integrable function of the variables x^i, the quantity

$$\delta(xb^{\mu}xb^{\mu})\frac{h(r)}{2r}$$

is the symbolic function of a distribution for all values of the variables $xb^{\mu}xb_{\mu}$ and x^i.

We have the, in the same coordinate frame,

$$\frac{\partial}{\partial x^n} \int \delta(xb^{\mu}xb_{\mu})h(r)db^n$$

$$= \int \delta'(xb^{\mu}xb_{\mu})2\, xb_n\, h(r)db^n$$

$$+ \int \delta(xb^{\mu}xb_{\mu})\; \delta(r)\; \frac{\partial r}{\partial x^4}\, db^4$$

The first term on the right hand side of this last equation is equal to zero by virtue of (3.219). The second term is also zero, since r is a function of x^i only, and thus (3.217) is proved.

Electromagnetic Potentials

It was shown in Section 2.2 (equations 2.2141 and 2.215) that the potential (3.209) may be expressed in the following equivalent forms:

$$F_n^{(b)}(x) = e_b \int \delta(xb_{\mu}xb^{\mu})db_m \qquad (3.209)$$

$$= \frac{e_b}{2}\left[\left(\frac{db_m/db}{xb^{\mu}\,\frac{db_{\mu}}{db}}\right)_{\Gamma_R} + \left(\frac{db_m/db}{bx^{\mu}\,\frac{db_{\mu}}{db}}\right)_{\Gamma_S}\right] \tag{3.223}$$

$$= \frac{e_b}{2}\left[\left(\frac{v_m}{D}\right)_R + \left(\frac{v_m}{D}\right)_S\right] \tag{3.224}$$

$$= \frac{1}{2} F_m{}^{(b)}(x)_R + \frac{1}{2} F_m{}^{(b)}(x)_S \text{ say} \tag{3.2241}$$

where D is the distance from the point x to the world line of b as given by (2.205), and v_m is the unit velocity 4-vector of the particle b.

It is clear that for a particle b at rest $\frac{db^i}{db} = 0$,

$$D_R = D_S = r \tag{3.225}$$

where $r = (xb_i bx^i)^{1/2}$.

Wheeler and Feynman further show the advanced and retarded potentials due to a fixed linear conductor (at rest) to be equal. This is not a surprising result, for at each fixed point in the conductor there is a constant current vector. The advanced and retarded potentials of each individual moving charged particle are not equal, since the particles move, but the configuration as a whole is stationary, and the sum of the advanced potentials of all the particles of the conductor is equal to the sum of their retarded potentials. At each instant at the point of observation the advanced field of each particular particle is equal to the retarded field of some other particle.

A discussion of the significance of advanced potentials in more complex cases is reserved for Section 3.4.

General Tensor Formulation of Theory.

Although Wheeler and Feynman expressed Fokker's electrodynamics in terms of Cartesian tensors, it is easy (in view of Section 2.2) to write down the corresponding equations in terms of general tensors in flat space. The fundamental principle is contained in

$$\mathrm{Var}\left[\int (g_{mn}\, da^m da^n)^{1/2} + \frac{e_a}{m_a c^2}\int F_r(a)\, da^r\right] = 0 \qquad (3.226)$$

where here a^μ are general coordinates, g_{mn} is the metric tensor, and using (2.2141) and (2.215)

$$\begin{aligned} F_r(a) &= \sum_{b\neq a}\int \delta(S^2)\, db_r \\ &= \sum_{b\neq a}\frac{e_b}{2}\left[\left(\frac{v_r}{D}\right)_R + \left(\frac{v_r}{D}\right)_S\right] \end{aligned} \qquad (3.227)$$

Here v^r is the velocity 4-vector of particle "b" and D the distance from "a" to the world line of "b".

The equations of motion resulting from (3.226) will be

$$\frac{\delta^2 a_m}{\delta a^2} = \frac{e_a}{m_a c^2} F_{nm}(a)\,\frac{da^n}{da} \qquad (3.228)$$

where here the symbol δ represents absolute differentiation, and

$$F_{mn} = \frac{\partial F_m}{\partial x^n} - \frac{\partial F_n}{\partial x^m} \qquad (3.229$$

as before.

We will also have (in addition to (3.207))

$$F^{mn}{}_{|n}(x) = 4\pi \sum_{\text{all } b} e_b \int \frac{\delta(xb_1)\,\delta(xb_2)\,\delta(xb_3)\,\delta(xb_4)\,db^m}{\sqrt{g}} \tag{3.230}$$

instead of (3.208) and (3.206). (Use is made here of (2.124)).

The tensor equations (3.228) and (3.230) are valid deductions from (3.226) in a particular frame (Cartesian coordinates), and hence hold in any coordinate frame.

3.3 PLANETARY MOTION

In this section the equations of motion (using Wheeler and Feynman electrodynamics) are found for a charged test particle moving in the field of a single stationary charged particle.

As has been shown, (equations (3.224) and (3.225)), if a single particle (b) is at rest in a Galilean frame of reference, its advanced and retarded potentials at any point x^μ are equal, and the total potentials are given by

$$\left.\begin{aligned} F_4^{(b)}(x) &= \frac{e_b}{r} \\ F_i^{(b)}(x) &= 0, \qquad (i = 1, 2, 3) \end{aligned}\right\} \tag{3.301}$$

where $r = (-xb_i xb^i)^{1/2}$.

The equation of motionfor a test particle ξ as given by (3.228) are

$$\frac{d^2\xi_\rho}{d\xi^2} = \frac{e_\xi}{m_\xi c^2}\left(\frac{\partial F_\mu}{\partial x^\rho} - \frac{\partial F_\rho}{\partial x^\mu}\right)\frac{d\xi^\mu}{d\xi} \tag{3.302}$$

It will be assumed that the motion takes place in the plane $\xi^3 = 0$.

Now replacing the left hand side of this last equation by $\frac{\delta}{\delta\xi}\left(\frac{d\xi_\rho}{d\xi}\right)$, and transforming to polar coordinates in the plane of the motion, centred at b^μ

$$r = (\xi^{1^2} + \xi^{2^2})^{1/2}, \quad \theta = \tan^{-1}\frac{\xi^2}{\xi^1}, \quad t = \frac{\xi^4}{c},$$

we have from (3.301) and (3.302)

$$\frac{d^2r}{d\xi^2} - r\left(\frac{d\theta}{d\xi}\right)^2 = \frac{e_\xi e_b}{m_\xi c r^2}\frac{dt}{d\xi} \tag{3.303}$$

$$\frac{1}{r}\frac{d}{d\xi}\left(r^2\frac{d\theta}{d\xi}\right) = 0 \tag{3.304}$$

$$c^2\frac{d^2t}{d\xi^2} = \frac{e_\xi e_b}{m_\xi c r^2}\frac{dr}{d\xi}. \tag{3.305}$$

We may write

$$\frac{dt}{d\xi} = \frac{\gamma}{c} \tag{3.306}$$

where $$\gamma = \left[1 - \left\{\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2\right\}/c^2\right]^{1/2} \tag{3.307}$$

since $d\xi^2 = c^2dt^2 - dr^2 - r^2d\theta^2$.

The parameter ξ may be eliminated from equations (3.303) to (3.305) with the help of (3.306), yielding respectively

$$\frac{d}{dt}\left(\gamma\frac{dr}{dt}\right) - \gamma r\left(\frac{d\theta}{dt}\right)^2 + \frac{U'(r)}{m_\xi} = 0 \tag{3.308}$$

$$\frac{d}{dt}\left(\gamma r^2\frac{d\theta}{dt}\right) = 0 \tag{3.309}$$

$$m_\xi c^2(\gamma - 1) + U(r) = B. \tag{3.310}$$

Here

$$U(r) \doteq \frac{e_b e_\xi}{r}$$

and equation (3.310) is a first integral of (3.305), the quantity B being a suitable chosen constant of integration.

We need go no further. These are identically (see McCrea,[17] Relativity Physics, equations (61) and (62)) the differential equations in Special Relativity for the general central orbit, where the force is derivable from the potential U(r).

This is an interesting result, because from the point of view of Special Relativity, equations (3.308) to (3.310) were derived from a Lorentz invariant energy-momentum tensor and a potential energy specified only in a single frame. From the present point of view the same equations have been developed as a particular result from the more general Lorentz invariant electrodynamics contained in equation (3.226).

3.4. ABSORBER THEORY OF RADIATION

Introduction

It is a well known physical fact that when a charged particle is accelerated it suffers a seemingly intrinsic damping force given at non-relativisitc velocities by the following expression

$$\text{Damping force} = \frac{2(\text{charge})^2(\text{time rate of charge of acceleration})}{3(\text{velocity of light})^3} \qquad (3.401)$$

A relativistically invariant generalization of this formula has been proposed by Dirac.[10] Here force is represented by a skew-symmetric tensor (multiplied by the charge of the particle). The tensor proposed by Dirac to describe the damping force acting a charge particle (say a) is given by

$$F_{mn}(a)_{\text{damping}} = \frac{1}{2}\left(F_{mn}{}^{(a)}_{R}(a) - F_{mn}{}^{(a)}_{S}(a)\right) \qquad (3.402)$$

that is, by one half the difference of the retarded and advanced Maxwellian fields of the particle itself, as evaluated at its own position. This difference is evidently finite, although the individual terms are not.

In Wheeler and Feynman electrodynamics, the total force tensor acting on a particle is given by

$$F_{mn}(a) \equiv \sum_{b \neq a} \frac{1}{2}\left(F_{mn}{}^{(b)}_{R}(a) + F_{mn}{}^{(b)}_{S}(a)\right) \qquad (3.403)$$

(See equations (3.2241), (3.228), and (3.229)).

To correspond to the retarded fields of experience and to include radiation damping as given by (3.402), we would require instead of (3.403) the following formula:

$$F_{mn}(a) = \sum_{b \neq a} F_{mn\,R}^{(b)}(a) + \frac{1}{2}\left(F_{mn\,R}^{(a)}(a) - F_{mn\,S}^{(a)}(a)\right) \qquad (3.404)$$

Wheeler and Feynman, in their "Absorber Theory of Radiation", attempt to show that in a universe with a sufficient number of charged particles equation (3.404) is a natural consequence of the time symmetrical identity (3.403). Thus the second term on the right hand side of (3.404) does not represent any self action or intrinsic force, but is essentially the result of summing the advanced fields of all the other particles of the universe which have been accelerated by the retarded field of particle a.

A major difficulty in the Wheeler and Feynman presentation is that equation (3.403) is in every way symmetrical with respect to time, while the damping forces (3.401) and (3.402) have a unidirectional aspect which correctly lead to the irreversibility of radiation. In their derivations of radiative reaction the authors either arbitrarily omit certain effects or describe the irreversibility of radiation as a statistical effect due to unnamed initial conditions which are unsymmetrical in time. No indication is given as to how these initial conditions permanently affect the time symmetry of radiation damping; in fact such an effect conceptually seems to be most unlikely, particularly in a universe where other independent (e.g. gravitational) forces exist. It is interesting to note that throughout their arguments the authors frequently make use of forces* which are independent of the electromagnetic field.

---

* In spite of the claim that "All of mechanics... is contained in (their) single variational principle."

Four derivations (with increasing generality) are given of radiative reaction, two of which (the first and last) will be critically examined here.

Derivation I

This is the simplest derivation, and considers a single accelerated charged particle "a" (called the source) and an infinite number of other charged particles, essentially at rest with respect to "a", separated by large distances and uniformly distributed. The particles other than "a" comprise the absorber. The technique here is to calculate the retarded field due to "a" at each of these particles, to determine the resultant acceleration of each particle, and then to derive the resultant advance field at "a" due to these accelerated charges. This advanced field arrives at "a" simultaneously with its acceleration, and constitutes the radiative reaction.

If the ordinary classical 3-vector acceleration of "a" is denoted by $\underline{f}$, then the radiative reaction force in the direction of $\underline{f}$ due to the presence of a single particle "b" at distance (3-space) r from a is found to be given by

$$\left(\underline{f} \quad e_a^2/2c^4 \right) \left(e_b^2/m_b r_b \right) \sin^2\left(\underline{f}, \underline{r}_b \cdot\right)$$

If there are N particles per unit volume, then the number in a spherical shell of thickness dr will be

$$4\pi N \, r^2 dr$$

For the particles in this shell, the factor

$$\sin^2 (\underline{f}, \underline{r})$$

has the average value of 2/3. Thus the total force of reaction would be the integral (over r) of the quantity

$$\left(\underset{\sim}{2f} \; {e_a}^2/_3\, c^3\right)\left(2\pi N {e_b}^2/m_b c\right) dr$$

Wheeler and Feynman point out that this derivation is incorrect in that it yields a reactive force which is in phase with and proportional to the acceleration, and increases without limit with the thickness of the absorber. The reason for this is that interactions between the particles of the absorber themselves have not been taken into account. To account for this, the acceleration is analysed into its Fourier components, and each frequency is considered separately. This permits a usage of the wave theory of the propogation of electromagnetic disturbances (since the fields satisfy Maxwell's equations) through an ionized medium. It is found then that provided some small factor of absorption is included the integral representing the radiative reaction converges to a limit which is independent of N, and that the limit involves a phase shift and a factor proportional to the frequency of the acceleration in such a way that the radiative reaction is given by (3.401).

The most important criticism claimed in the present work for this simple derivation is the arbitrariness of considering only the retarded field of the source and the advanced field of the absorber. If the same calculation were carried out using the advanced field of the source and the simultaneous retarded field of the absorber, then we should arrive at the same value for radiative reaction, but with a change in sign. The sum of these two effects would of course provide for zero radiative reaction (at least on the average), which is what is to be expected from a completely time-symmetrical theory. In effect the authors have ignored the existence of an absorber along the pastnull cone of the source.

Derivation IV

This is a general derivation, and depends on the conecption of a perfect absorber defined by the fact that beyond it the electromagnetic field due to the totality of particles in the universe vanishes. That is, outside the absorber we have

$$\sum_{\text{all } b} \frac{1}{2} F_{mn\,R}^{(b)}(x) + \sum_{\text{all } b} \frac{1}{2} F_{mn\,S}^{(b)}(x) = 0 \qquad (3.405)$$

Since the two terms on the left hand side of (3.405) represent respectively outgoing and incoming waves, they cannot cancel each other by destructive interference, but must individually vanish. It follows that

$$\sum_{\text{all } b} \frac{1}{2} \left(F_{mn\,R}^{(b)}(x) - F_{mn\,S}^{(b)}(x)\right) = 0 \text{ outside the absorber} \qquad (3.406)$$

Now, since the divergence of the left hand side of (3.406) vanishes everywhere, it follows from (3.406) that

$$\sum_{\text{all } b} \frac{1}{2} \left(F_{mn\,R}^{(b)}(x) - F_{mn\,S}^{(b)}(x)\right) = 0 \quad \text{everywhere} \qquad (3.407)$$

For particle "a" we have then

$$F_{mn}^{(a)} = \sum_{b \neq a} \frac{1}{2}\left(F_{mn}{}^{(b)}_{R}(a) + F_{mn}{}^{(b)}_{S}(a)\right)$$

$$\equiv \sum_{b \neq a} F_{mn}{}^{(b)}_{R}(a) + \frac{1}{2}\left(F_{mn}{}^{(a)}_{R}(a) - F_{mn}{}^{(a)}_{S}(a)\right)$$

$$- \sum_{\text{all } b} \frac{1}{2}\left(F_{mn}{}^{(b)}_{R}(a) - F_{mn}{}^{(b)}_{S}(a)\right)$$

$$= \sum_{b \neq a} F_{mn}{}^{(b)}_{R}(a) + \frac{1}{2}\left(F_{mn}{}^{(a)}_{R}(a) - F_{mn}{}^{(a)}_{S}(a)\right) \qquad (3.408)$$

by virtue of (3.407).

It is thus that Wheeler and Feynman claim the equivalence of (3.403) and (3.404) for a completely absorbing universe.

This derivation of the Absorber Theory of Radiation possesses the attractiveness of generality and apparent mathematical simplicity. However it contains certain inherent difficulties.

Firstly we note that the definition contained in (3.405) of a perfect absorber or electromagnetic shield does not aptly describe the universe as a whole. The absorber must consist of ordinary particles and as such must be surrounded by a further absorber, and so on ad infinitum, leaving nowhere for (3.405) to be satisfied.

Equation (3.405) can only be used to describe the particular case where it is assumed that all particles are at rest except as affected by a localized disturbance. Equation (3.405) then implies that this disturbance remains localized in that it vanishes for all time outside some region containing sufficient particles to constitute an absorber. It is thus claimed here that equation (3.407) is more general than (3.405) and as such should be regarded as the absorber postulate. But it is agreed that the use of (3.405) constitutes a powerful argument in favour of looking upon (3.407) as a natural consequence of a sufficiency of charged particles in the universe.

A more controversial aspect of this derivation (IV) is the physical significance of the two terms on the right hand side of equation (3.408). It was noted by the authors that equation (3.407) equally well leads to a result different in form but equivalent to (3.408), namely

$$F_{mn}(a) = \sum_{b \neq a} F_{mn\,S}^{(b)}(a) - \frac{1}{2}\left(F_{mn\,R}^{(a)}(a) - F_{mn\,S}^{(a)}(a)\right) \qquad (3.409)$$

The physical significance of the two equations is said to be made clear by considering the particle "a" to be disturbed by a non-electromagnetic force. Then the first term on the right hand side of (3.408) is taken to be zero, while the second term represents radiative reaction. Under similar conditions the first term on the right hand side of (3.409) is not zero, but is just twice the second term (and of opposite sign). This is because the particles of the absorber were excited by the retarded field of "a", and their advanced fields converge on "a" to give the required result.

Thus while the two equations are mathematically equivalent, (3.408) is to have the physical significance of naturally separating out radiative reaction from other electromagnetic forces. It is stated that unsymmetrical initial conditions are responsible for being able (on a statistical basis) to neglect the first term on the right hand side of (3.408) in estimating radiative reaction.

This is not a strong argument, and in particular it is difficult to see how initial electromagnetic conditions affect the radiation reaction suffered by a particle which has been stimulated by an independent force.

In this, as in the earlier derivations, the authors in effect omit to take into account the existence of an absorber (A_P) lying along the past null cone of the source. For in the absence of such an absorber the first term on the right hand side of (3.408) automatically vanishes in the calculation of radiative reaction. We will now put forward arguments which indicate that the presence of A_P may be neglected provided it is **not a "perfect absorber"!**

Perfect and Imperfect Absorbers and the Irreversibility of Radiation

We denote respectively by A_F and A_P the totality of particles lying on the future (σ_F) and past (σ_P) branches of the null cone centred on particle "a" (the source). As before, we consider all particles to be at rest except as disturbed by the source. We define A_F to constitute a perfect absorber if (3.407) holds in the absence of A_P.* Under such circumstances (3.408) becomes

$$F_{mn}(a) = \frac{1}{2}\left(\underset{R}{\overset{(a)}{F_{mn}}}(a) - \underset{S}{\overset{(a)}{F_{mn}}}(a) \right) \qquad (3.410F)$$

Because of (3.407) we would also have

$$\sum_{A_F} \underset{S}{\overset{(b)}{F_{mn}}}(x) = - \underset{S}{\overset{(a)}{F_{mn}}}(x) \qquad (3.411F)$$

---

* In view of the first three derivations of Wheeler and Feynman it is evident that such a definition is self-consistent and requires only a sufficient number of particles with a large scale uniformity of distribution.

holding in σ_P, and

$$\sum_{A_F} \overset{(b)}{F_{mn}}{}_{S}(x) = \overset{(a)}{F_{mn}}{}_{R}(x) \qquad (3.412F)$$

in any region in σ_F between a and A_F.

Similarily A_P is defined to be a perfect absorber if (3.407) holds in the absence of A_F. Then instead of (3.410F), (3.411F) and (3.412F) we would have

$$F_{mn}(a) = -\frac{1}{2}\left(\overset{(a)}{F_{mn}}{}_{R}(a) - \overset{(a)}{F_{mn}}{}_{S}(a)\right) \qquad (3.410P)$$

$$\sum_{A_P} \overset{(b)}{F_{mn}}{}_{R}(x) = -\overset{(a)}{F_{mn}}{}_{R}(x) \text{ in } \sigma_F \qquad (3.411P)$$

$$\sum_{A_P} \overset{(b)}{F_{mn}}{}_{R}(x) = \overset{(a)}{F_{mn}}{}_{S}(x) \text{ between a and } A_P.$$

The left hand sides of equations (3.410) to (3.412) will be called the response to the stimulus of the source.

We now define an imperfect absorber of the first kind to be one whose response is some factor (less than unity) times the response of a perfect absorber for the same stimulus. We name these factors f and p respectively for A_F and A_P. With only a single imperfect absorber of the first kind equation (3.407) would not be satisfied and the right hand side of equations (3.410) to (3.412) would be multiplied by the appropriate factor for p.

It is clear from equations (3.410) to (3.412) that the type of absorber under consideration qualitatively produces no new fields, but only changes the magnitude and/or sign of the fields due to the source. It is reasonable to expect such also to be the case if both A_F and A_P exist simultaneously. In this case if we represent the stimulus of the source by unity the stimulus applied to A_F and A_P will respectively be

$$1 - \alpha_P$$

and

$$1 - \alpha_F$$

where α_P is the response of A_P and α_F is that of A_F. We have then

$$\begin{aligned} \alpha_P &= p(1 - \alpha_F) \\ \alpha_F &= f(1 - \alpha_P) \end{aligned} \qquad (3.413)$$

In terms of this notation, equation (3.407) evidently requires that

$$\alpha_P + \alpha_F = 1 \qquad (3.414)$$

We consider three possibilities:

(a) $p = f = 1$ (Both A_F and A_P are perfect absorbers). Here equations (3.413) are indeterminate, and lead only to

$$\alpha_P + \alpha_F = 1$$

which is the same as (3.414). This is the case of the Wheeler and Feynman derivations, where it was assumed on the basis of initial conditions that

$$\begin{aligned} \alpha_P &= 0 \\ \alpha_F &= 1 \end{aligned} \qquad (3.4141)$$

It will be seen that (3.4141) is obtained unambiguously in (b).

(b) $p<1$, $f = 1$. (A_F only is a perfect absorber). Equations (3.413) now have the unique solution

$$\alpha_P = 0$$

$$\alpha_F = 1$$

which is compatible with (3.414). We thus have the remarkable result that the response of the perfect absorber A_F is the same as if the imperfect absorber A_P were non-existent, while the response of the latter is zero.

(c) $p<1$, $f<1$. Here (3.413) have a unique solution which is not compatible with (3.414).

We have

$$\alpha_F = \frac{f(1 - p)}{1 - fp}$$

$$\alpha_P = \frac{p(1 - f)}{1 - fp}$$

This result is useful in studying the case where A_F only approximates to a perfect absorber. We write

$$f = 1 - \Delta$$

$$p = 1 - \xi$$

where $\Delta << \xi < 1$,

obtaining

$$\frac{\alpha_F}{\alpha_P} = \frac{(1-\Delta)\xi}{(1-\xi)\Delta} > \frac{\xi}{\Delta}$$

$$\alpha_P + \alpha_F = 1 - \frac{\Delta\xi}{\Delta+\xi-\Delta\xi} \approx 1 - \Delta$$

We see then that the response of A_F is very nearly that of a perfect absorber in the absence of A_P, while the response of A_P is very small, and that (3.414) is approximately satisfied. These results still hold if A_P approximates to a perfect absorber, provided only that the approximation of A_F to a perfect absorber is of a higher degree than that of A_P. That is

$$\alpha_F \approx 1$$

$$\alpha_P \approx 0$$

provided $\Delta << \xi$

even though $\xi << 1$

A weakness in the foregoing arguments is contained in the fact that it is unreasonable to expect the response of an imperfect absorber to be of the same form as its stimulus. It may however be assumed from symmetry conditions that the response to a point source is equivalent to a disturbance emanating from the same point. We therefore define an imperfect absorber of the second kind to be one whose response to the non-zero stimulus

$$F_{mn}{}^{(a)}(x)$$

is equal to $\quad B_{mn}{}^{(a)}(x)$

where F_{mn} and B_{mn} are different. If A_P is such an absorber, then B_{mn} is to replace F_{mn} on the right hand side (only) of equations (3.410P), (3.411P) and (3.412P).

Consider again the case of a perfect absorber A_F, an imperfect absorber of the second kind A_P and a source "a". Denoting the response of A_P by $\beta_{mn}{}^{(a)}$, then in view of equation (3.411P) and the last paragraph the stimulus applied to A_F is given by

$$F_{mn_R}^{(a)}(x) - \beta_{mn_R}^{(a)}(x)$$

and since the response of A_F is the same as its stimulus, the stimulus applied to A_P is (using 3.411F)

$$F_{mn_S}^{(a)}(x) - \left(F_{mn_S}^{(a)}(x) - \beta_{mn_S}^{(a)}(x)\right)$$

which is the same as its response, and therefore must vanish. We conclude then that the response (and stimulus) of A_P is zero also in this general case, and that the equation

$$F_{mn}(a) = \sum_{b\neq a} F_{mn_R}^{(b)}(a) + \frac{1}{2}\left(F_{mn_R}^{(a)}(a) - F_{mn_S}^{(a)}(a)\right) \qquad (3.408)$$

derived by Wheeler and Feynman from (3.407) and (3.403) has the physical significance of having radiative reaction correctly represented by the second term on the right hand side, provided A_F only is a perfect absorber. We thus claim that the irreversibility of radiation is a result of a change (from imperfection to perfection) in the properties of the absorber as it progresses from past to future.

It will be shown in the next section, after generalizing the electrodynamics of Wheeler and Feynman to the de Sitter space of constant curvature, that with a Steady-State Cosmology distribution of mass and charge A_F and A_P appear to have the required difference in their properties.

3.5. EXTENSION OF WHEELER AND FEYNMAN PARTICLE ELECTRODYNAMICS TO THE DE SITTER SPACE-TIME OF CONSTANT CURVATURE

Introduction

A natural extension of the electrodynamics contained in equation (3.201) to general Riemannian space is obtained by writing as the fundamental physical principle

$$\mathrm{Var}\left[\int\left(g_{mn}\,da^m da^n\right)^{1/2} + \frac{e_a}{m_a c^2}\int F_r(a)\,da^r\right] = 0 \qquad (3.501)$$

where now the g_{mn} comprise the metric tensor of the Riemannian space, a^μ are general coordinates, and Fr is some (as yet unspecified) co-variant vector potential field.

It is easily shown that the equations of motion[11] resulting from (3.501) are

$$\frac{d^2a^r}{da^2} + \left\{ {r \atop m\;n} \right\} \frac{da^m}{da}\frac{da^n}{da} = g^{rn}\,\frac{e_a}{m_a c^2}\,F_{mn}(a)\,\frac{da^m}{da} \qquad (3.501)$$

where $$F_{mn}(x) = \frac{\partial F_m(x)}{\partial x^n} - \frac{\partial F_n(x)}{\partial x^m} \qquad (3.502)$$

and $da^2 = g_{mn}\,da^m da^n$. (3.503)

To complete this generalized theory we require only a definition of the vector

$$F_r(x)$$

The properties that will be demanded of this vector field are as follows:

(a) In the special case of flat space-time it must reduce to that given by the time-symmetrical defintion (3.227).

(b) It must be determined by the world lines of the various charged particles.

(c) In order to maintain the Principle of Superposition we will require that the potential due to a group of particles is equal to the sum of the potentials of the individual particles. We write

$$F_r(a) = \sum_{b \neq a} F_r^{(b)}(a) \tag{3.504}$$

$$F_r(x) = \sum_{\text{all } b} F_r^{(b)}(x) \tag{3.505}$$

Equation (3.504) ensues the property of Omission of Self Action in (3.501).

We add to the foregoing list the tentative* requirement that in addition to the one set of Maxwell's equations which are derivable from (3.502):

$$F_{mn,r} + F_{nr,m} + F_{rm,n} = 0 \tag{3.506}$$

(d) The skew symmetric tensor defined by (3.502) must satisfy the second set of Maxwell's generalized equations

$$F^{mn}{}_{/n} = 4\pi j^m \tag{3.507}$$

where the current four-vector j^m is given by

$$j^m(x) = \sum_b e_b \int \frac{\delta(xb_1)\,\delta(xb_2)\,\delta(xb_3)\,\delta(xb_4)\,db^m}{\sqrt{g}} \tag{3.508}$$

Bearing in mind equation (2 . 124) it is seen that equations (3.507) and (3.508) are identical in form to the corresponding equations (3.230) of the flat space theory.

---

* This requirement is discussed in Chapter IV.

In view of the invariant spatial distance D defined in Section 2.2. it is a simple matter to find a vector potential satisfying (a), (b) and (c) above. The simplest of these is given by (3.227) itself. (No explicit reference to the metric is contained in (3.227)). However it can be shown that with this choice of potential requirement (d) is not in general satisfied in non-flat Riemannian space, it is not in fact satisfied in de Sitter space.

Although it has not been found possible in this work to find a general potential satisfying requirements (a) to (d), a suitable potential has been found for the special case of de Sitter space. For the remainder of this chapter only this special case will be considered.

Spatial Distance in de Sitter Space.

It is of interest to note that the invariant spatial distance D which is to be used in the definition of an electromagnetic potential, has a special physical significance in the theory of Steady-state Cosmology. Using the metric

$$ds^2 = dt^2 - (dx^2 + dy^2 + dz^2) \exp 2kt$$

then in Steady State theory large scale movements of particles are along the geodesics

$$x = \text{const}$$

$$y = \text{const}$$

$$z = \text{const}$$

The spatial distance from a point X with coordinates

$$(0, 0, 0, t)$$

to the world line (L_0) of one of these particles (X_0) is given by (2.327). We have then, assuming X_0 to lie on the past branch of the null cone through X.

$$D = \frac{\exp k(t-t_0) - 1}{k} \qquad (3.509)$$

$$= r \exp kt$$

by (2.317), where r is given by (2.316). Here D has the physical significance of being the luminosity distance of relativistic and steady state cosmology(12) (from the source X_0 to the observer X).

Electromagnetic Potential in de Sitter Space

We can now define, for de Sitter space;

$$F_r^{(b)}(a) = \frac{1}{2} e_b \left(\frac{m_r}{D_R} + \frac{m_r}{D_S} \right) \qquad (3.510)$$

where m_r is defined in (2.356) and D_R^{-1} and D_S^{-1} in (2.205). (The point a^μ and the world-line of b are respectively denoted in Chapter II by X and L).

Noting equations (2.363) to (2.365) and (2.372), it is clear in view of (2.373) that (3.507) is satisfied by the choice of potential (3.510). We also see that if we put

$$k = 0 \quad \text{(degeneration to flat space)}$$

in the definition of m_r, then (3.510) becomes equivalent to (3.227).

Radiative Reaction by a Single Charged Particle

We consider now the case of two charged particles "a" and "b" initially unaccelerated and linked by a single family of null geodesics. Particle "a" is then stimulated into acceleration by a non-electromagnetic force, producing a radiation field at "b". Particle "b" is accelerated by this field and in turn produces a radiation field of the opposite sense (with respect to advancement or retardation) which arrives at "a" simultaneously with its original acceleration. We call this latter field (calculuated at "a") the advanced or retarded radiative reaction of "b" on "a", the reaction being termed advanced if "b" lies on the future branch of a's null cone, and retarded if "b" lies on the past branch of a's null cone.

The velocity 4-vector* of particle "a" for simplicity will be denoted by

$$\dot{a}^m$$

and the acceleration vector

$$\frac{d^2a^m}{da^2} + \left\{ {m \atop r\ s} \right\} \frac{da^r}{da} \frac{da^s}{da}$$

by

$$\ddot{a}^m .$$

A similar notation will be used for particle b's velocity and acceleration. Otherwise the notation will be that used for equation (2.356), in which cases a prefix $_o$ will be used to distinguish those quantities for which b is to be considered the source. We have for example, from (2.204),

$$D = P_\mu \dot{a}^\mu \qquad (3.511)$$

$$_oD = {_oP_\mu} \dot{b}^\mu \qquad (3.512)$$

---

* We are reminded that all vectors under consideration are defined by parallel transfer along the null geodesic joining "a" and "b".

It is easily shown from (2.318) that

$$P_\mu = - {}_oP_\mu \, . \qquad (3.513)$$

We now calculate the field

$$F_{mn}{}^{(a)}(b)$$

which will be one half of the full retarded or advanced field of "a". Using (3.510), (2.356) and (2.359) we have

$$F_{\mu r}{}^{(a)}(b) = \frac{1}{2}\, e_a\, \delta^{nm}_{\mu r} \left[\frac{P_m \ddot{a}_n}{D^2} - \frac{P_m \dot{a}_n P_r \ddot{a}^r}{D^3} + \frac{P_m \dot{a}_n}{D^3} \right] \qquad (3.513)$$

The first two terms in the bracket on the right hand side of (3.513) constitute the radiation field, each being essentially proportional to

$$\frac{1}{D} \, .$$

We now assume that "a" and "b" are separated by a large spatial distance, and therefore drop the last bracketed term, which is essentially proportional to

$$\frac{1}{D^2}$$

We have then, using (3.5011)

$$\ddot{b}_r = \frac{1}{2}\, \frac{e_a e_b}{m_b c^2}\, \delta^{nm}_{\mu r} \left[\frac{P_m \ddot{a}_n}{D^2} - \frac{P_m \dot{a}_n P_r \ddot{a}^r}{D^3} \right] \qquad (3.514)$$

The radiation field of "b" at "a" is given by

$$F_{pq}{}^{(b)}(a) = \frac{1}{2}\, e_b\, \delta^{r\sigma}_{pq} \left[\frac{{}_oP_\sigma \ddot{b}_r}{{}_oD^2} - \frac{{}_oP_\sigma \dot{b}_r P_s \ddot{b}^s}{{}_oD^3} \right] \qquad (3.515)$$

On substituting (3.514) in (3.515) and using (2.3531), and (3.511) to (3.513) there obtains

$$F_{pq}^{(b)}(a) = \frac{1}{4}\frac{e_a e_b{}^2}{m_b c^2}\left[\frac{P_p \ddot{a}_q - P_q \ddot{a}_p}{{}_oDD^2} - \left(\frac{P_p \dot{a}_q - P_q \dot{a}_p}{{}_oDD^3}\right) P_r \dot{a}^r\right] \tag{3.516}$$

Particle "a" is subject to two forces, the non-electromagnetic disturbing force and that contained in (3.516). The portion of "a" 's acceleration arising from the tensor (3.516) will be denoted by

$${}_o\ddot{a}_n$$

so that we have from (3.501)

$${}_o\ddot{a}_n = \frac{e_a}{m_a c^2} F_{mn}^{(b)}(a)\, \dot{a}^m \tag{3.517}$$

We now employ the metric (2.303) with "a" at the origin of the space coordinates, and consider the special case where

$$\left.\begin{aligned} \dot{a}^4 &= 1 \\ \dot{a}^i &= 0 \qquad (i = 1, 2, 3) \\ \ddot{a}^4 &= 0 \end{aligned}\right\} \tag{3.518}$$

(The last of (3.518) follows from the first two).

Substituting (3.516) in (3.517) and applying (3.518) we find

$${}_o\ddot{a}_4 = 0 \tag{3.519}$$

$${}_o\ddot{a}_i = -\frac{1}{4}\frac{e_a{}^2 e_b{}^2}{m_a m_b c^4}\left(\frac{P_4 \ddot{a}_i}{{}_oDD^2} + \frac{P_i P^j \ddot{a}_j}{{}_oDD^3}\right) (j = 1,2,3) \tag{3.520}$$

Using the values given in (2.318) for P_m and employing (2.317) and (3.511) with (3.518) there obtains

$$ {}_o\ddot{a}_i = \left(\frac{-\ddot{a}_i}{{}_oDD} + \frac{\delta_{ik} b^k b^j \ddot{a}_j}{{}_oDDr^2} \right) \frac{1}{4} \frac{e_a{}^2 e_b{}^2}{m_a m_b c^4} $$

where $r^2 = \delta_{ij}\ b^i b^j$.

If we now choose one of the coordinates axis (say the n^{th}) along the direction of the acceleration vector a^i, so that

$$ \ddot{a}_i = 0 \quad \text{for} \quad i \neq n $$

then it is easy to show that

$$ {}_o\ddot{a}_n m_a = -\frac{1}{4} \frac{e_a{}^2 e_b{}^2}{m_b c^4} \frac{\ddot{a}_n \sin^2\theta}{{}_oDD} \qquad (3.521) $$

where θ is given by

$$ \cos\theta = \frac{b^n}{r} $$

The vector component

$$ {}_o\ddot{a}_n \quad m_a $$

is clearly the component of the force of radiative reaction in the direction of the acceleration.

Steady State Cosmology(12)

Before applying the foregoing to a large number of particles "b", some assumption must be made with regard to the distribution of charged particles in the universe. We adopt here the large scale distributions and motions demanded by the Steady State theory of Cosmology.

This theory is set in de Sitter 4-space, and if we emply the metric

$$ ds^2 = dt^2 - (dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2) \exp 2kt \qquad (3.522) $$

then the large scale motions of particles are along the geodesics defined by

$$dr = d\theta = d\phi = 0 \tag{3.523}$$

The large scale proper density of particles is taken to be a constant, so that the number of particles lying on the null cone through the origin of spatial coordinates

$$r = 0$$

and having radial coordinate between r and r+dr is

$$4\pi r^2 n dr \exp 3kt \tag{3.524}$$

where n is a constant.

The Steady State theory demands a steady creation of matter to maintain a constant proper density. The creation of charged particles will not produce an appreciable electromagnetic field provided the average rate of creation of charge is zero, and provided the initial velocities of newly created particles are not random but are the same as the large scale velocities of the existing matter as given by (3.523)

Radiative Reaction with Steady State Distribution of charged particles

The case is now considered (using the metric (3.522)) of a source "a" whose world line is along

$$r = 0$$

and a distribution of charged particles "b" given by (3.524) with notion (3.523). We assume each charged particles "b" to have equal mass (m) and equal magnitude of charge (e).

Because of symmetry conditions the force of radiative reaction acting on "a" will be along the direction of its total acceleration. The factor $\sin^2\theta$ in (3.521) will have the average value 2/3, and the total force of reaction due to those particles with radial coordinate between r and r+dr will be

$$-\frac{2}{3}\frac{e_a{}^2 e^2}{mc^4}\pi n \frac{r^2}{{}_0DD} dr \exp 3kt\, \ddot{a}^\mu \qquad (3.525)$$

The values giving for ${}_0D$ and D in (2.327) are in the present circumstances

$${}_0D = \frac{\exp k(t-a^4)-1}{k}$$

$$D = \frac{\exp k(a^4-t)-1}{k}$$

Using (2.317) and denoting a^4 by t_0 we have

$${}_0DD = -r^2 \exp k(t+t_0)$$

Thus (3.525) becomes

$$(2\ddot{a}^\mu e_a/3c^2)(\pi ne^2/mc)dr \exp(-kt_0)\exp 2kt \qquad (3.526)$$

Excepting the presence of the exponential terms, this is essentially* the formula of Wheeler and Feynman for the corresponding calculation (derivation I) in flat space-time.

By (2.317)

$$dr = \pm \exp(-kt)dt$$

---

* The formula of Wheeler and Feynman contains an additional factor of 2, which is present because of their assumption of the full retarded field of the source being experienced by the absorber.

therefore (3.526) can be expressed as

$$\pm\ (2\ddot{a}^{\mu}e_a/3c^2)(\pi ne^2/mc)\ \exp k(t-t_o)dt \qquad (3.527)$$

(The positive sign in (3.527) corresponds to $t > t_o$, the negative sign to $t < t_o$).

The simplified radiative reaction of the future absorber (A_F) is found by integrating (3.527) from t_o to ∞, and of the past (A_P) by similarly integrating from $-\infty$ to t_o. The former integral is infinite, but the latter has the finite value

$$-\frac{1}{k}\ (2\ddot{a}^{\mu}e_a/3c^2)(\pi ne^2/mc) \qquad (3.528)$$

These values bear little relationship to a correct derivation of radiative reaction, because we have not taken into account the interactions of the particles of the absorber. However the following physical significance may be attached to the above results: the simplified radiative reaction of A_F is equal to that of an infinite number of particles at a finite distance, while that of A_F corresponds to the reaction of a finite number of particles at a finite distance. It seems reasonable to assume that A_P cannot provide complete absorption.

In their Derivation I, Feynman and Wheeler showed, using a theory based on Maxwell's equations, that when interactions between particles were taken into account the integral representing radiative reaction converged to give the desired quantity. In view of the fact that the exponential in (3.526) may be regarded as a change of particle density with distance, and the fact that Maxwell's equations are satisfied in the present theory, one might reasonably expect the same result to hold here for the case of A_F.

Radiative Reaction - General Derivation

A general derivation of radiative reaction in de Sitter space and with Steady-state distribution of charged particles will now be given.

We have

$$F_{mn}(a) = \sum_{b \neq a} \frac{1}{2} \left(F_{mn}{}^{(b)}_{R}(a) + F_{mn}{}^{(b)}_{S}(a) \right) \qquad (3.529)$$

Now it follows from (2.3711) and the definition of electromagnetic potential (3.510) that

$$\sum_{\text{all } b} \left(F^{mn}{}^{(b)}_{R}(x) - F^{mn}{}^{(b)}_{S}(x) \right)_{|n} = 0 \qquad (3.530)$$

It is now assumed that the solution of (3.530) appropriate to a universe containing a sufficient number of charged particles is

$$\sum_{\text{all } b} \left(F^{mn}{}^{(b)}_{R}(x) - F^{mn}{}^{(b)}_{S}(x) \right) = 0 \qquad (3.531)$$

If (3.531) is applied to (3.529) there obtains

$$F_{mn}(a) = \sum_{b \neq a} F_{mn}{}^{(b)}_{R}(a) + \frac{1}{2} \left(F_{mn}{}^{(a)}_{R}(a) - F_{mn}{}^{(a)}_{S}(a) \right) \qquad (3.532)$$

In view of implications of (3.528) A_P may not be regarded as a perfect absorber, and by the arguments in Section 3.4. under the heading "Perfect and Imperfect Absorbers", (which arguments equally well apply here), the second term on the right hand side of (3.532) constitutes radiative reactions, leaving as the remaining field the full retarded field of all the other particles.

It is noted that Maxwell's equations are a necessary but not sufficient condition for (3.530) to be satisfied. For Maxwell's equations are a necessary condition for (3.530), which in turn is necessary for (3.531). The assumption (3.531) would appear to be a weakness in the foregoing derivation. However it should be pointed out that since the fields in this theory satisfy Maxwell's equations the arguments used by Wheeler and Feynman and contained in equation (3.405) equally well apply here.

The essential difference between the present theory and that of Wheeler and Feynman is the means whereby the irreversibility of radiation is derived. No unsymmetrical element exists in their theory, hence they are obliged to introduce unsymmetrical initial conditions, but without explaining either the meaning of these conditions or the means by which the desired result is achieved. In the present theory the irreversibility of radiation follows naturally from an unsymmetrical metric and a constant proper density of charged matter.

It is interesting to note that the unsymmetrical factor in (3.527) is the exponential in the variable t. If the argument of this exponential were a negative multiple of t, then radiation would take place in the opposite sense of t. Now in (3.525) (from which (3.527) was derived) the cube of this exponential appears, being required by the assumption of constant proper density. Other factors in the metric reduce this cube to the exponential in (3.527). However these other factors are not powerful enough to change the sign of the exponential. Now since the proper density is kept constant by a creation of matter, it seems reasonable to associate the time direction of radiation with the time direction of the creation of matter, and to claim that the irreversibility of both of these processes are closely related.

CHAPTER IV

GRAVITATION

4.1. INTRODUCTION

While the two major developments in gravitation - Newton's particle theory of instantaneous action at a distance according to an inverse square law, and Einstein's revolutionary idea of relating matter to the curvature tensor of a four dimensional space-time continuum - differ considerably in concept, the observational differences are extremely minute, being in fact just sufficient in magnitude to bring favour on the latter. Indeed the amazingly accurate predictions of astronomy are made almost without exception on the basis of Newton's theory.

However, after the general acceptance of the principle of the special theory of relativity there were reasons probably stronger than the second order inaccuracies of Newtonian gravitation in its disfavour. If we accept from the special theory that casual effects are propagated with a speed not exceeding that of light,* we should expect in the case of a test particle (or a planet) revolving about a gravitating mass to find an aberration effect proportional to the velocity perpendicular to the radius vector. Such an effect, not found in nature, would be considerably larger than the second order inaccuracies already mentioned. Thus from the observational point of view, Newton's theory of gravitation is more compatible with fact and with the results of the general

---

* That such an interpretation of the special theory of relativity is not universal is exemplified by (13).

theory of relativity if we do not add to it a property which may seem inescapable from the point of view of special relativity.

After approximately forty years of contemplation by scientists, the special theory of relativity stands in considerably higher universal favour than its generalization in the general theory. It is therefore of some importance to have a formula which will more directly include gravitation in the special theory, which will be invariant under its transformations, and which will give results at least as acceptable as the pre-relativity theory.

Whitehead[3] proposed such a theory as early as 1922. His work yields the classical results of gravitation, and in addition second order effects which correctly predict the three famous so-called confirmations of the general theory of relativity. He pointed out that while Einstein's gravitational law is extremely limited in its application to actual problems, a theory such as his can be applied to any configuration.

In the present chapter the structure of Whitehead's theory is examined, particularly with respect to its ability to predict the advance of perihelion and its position with respect to the gravitational mass of classical energy. The last section of the Chapter is devoted to an extension (based on the distance D of Section 2.2.) of Whitehead's theory to de Sitter space-time of constant curvature.* The more general considerations of the theory with respect to General Relativity, Cosmology, and the problem of inertia

---

* The generalization to be proposed here is quite distinct from that of Temple[9] (1923).

are reserved for the following chapter.

In contrast to Wheeler and Feynman electrodynamics, Whitehead's theory of gravitation is based on the idea of antecedence of cause over effect. This is certainly not demanded by the principle of special relativity around which the theory is framed, for every mathematical equation describing that principle is symmetrical with respect to positive and negative time. It is of some interest then to devise and examine a time-symmetrical Whitehead theory.* This is done in Section 4.5.

An interesting example of a Lorentz invariant gravitational theory with dependence upon advanced effects is that proposed by Synge[14] in 1935. This theory includes non-antecedence effects because it is based on the emission of momentum particles from each gravitating body in such directions as to be received by the others. A fore-knowledge by each body of the future positions of the others is implied, which is indeed an advanced effect, since each body at present is influencing the past histories of the rest.

---

* Although such a theory will later be abandoned on cosmological grounds.

4.2. WHITEHEAD'S THEORY OF GRAVITATION

Noting that Whitehead's theory of gravitation is framed in terms of Cartesian tensors in the Minkowski space-time of Special Relativity, and using the notation of Section 3.2, the fundamental principle of the theory is as follows:

Along the world line of each particle (say a) the following principle of least action is satisfied;

$$\text{Var} \int (dJ(a)^2)^{1/2} = 0 \qquad (4.201)$$

$$\text{where } dJ(a)^2 = da^2 - \frac{2}{c^2} \sum_{b \neq a} \frac{\gamma\, m_b}{\Omega(r-\zeta)}\, dG_b^{\,2} \qquad (4.202)$$

Here

$$\zeta = (ab^1)\frac{db^1}{db^4} + (ab^2)\frac{db^2}{db^4} + (ab^3)\frac{db^3}{db^4}$$

$$\Omega = \left(1 - \frac{(db^1)^2 + (db^2)^2 + (db^3)^2}{(db^4)^2}\right)^{-1/2}$$

$$r = \left[(ab^1)^2 + (ab^2)^2 + (ab^3)^2\right]^{1/2}$$

γ is the gravitational constant

dG_b is that element of the world line of particle "b" which is antecedent to the element da of the world line of particle "a"; that is, two null cones extending from the extreme ends of the element of da into its past cut off dG_b of the world line of b.

The quantity $\Omega(r-\zeta)$ is an invariant cartesian scalar which is evaluated on the light cone extending from "a" into its past.

For a test particle (with coordinates x^{μ}) moving in the field of a particle of mass m at rest at the origin,

$$dJ^2 = \left(1 - \frac{2\gamma m}{c^2 r}\right)(dx^4)^2 + \frac{4\gamma m}{c^2 r^2}\,\delta_{ij}\,x^i dx^j dx^4 - \left(\delta_{ij} + \frac{2\gamma m}{c^2 r^3}\delta_{ik}\delta_{j\ell}\,x^k x^\ell\right) dx^i dx^j \qquad (4.2021)$$

(Here δ_{ij} is the Kronecker delta, and i, j, k, ℓ = 1,2,3.)

Eddington[15] pointed out that the metric of the Schwarzschild spherically symmetric static solution of the general relativity field equations can be put in this form by a simple transformation. Although Whitehead had not noted this point, he had derived the same advance of perihelion as for general relativity.

The mathematical complexities associated with Whitehead's derivations of specific results from (4.202) have undoubtedly been a major factor deterring a more universal investigation into his work . What is required in order to avoid these complexities is an alternative formulation of (4.202) which permits the direct usage of well known mathematical theorems.

General Tensor Formulation of Whitehead's Gravitational Principle

Noting (2.212) and (2.209), we may write instead of (4.202)

$$dJ^2(a) = da^2 - \frac{2}{c^2} \sum_{b \neq a} \frac{\gamma m_b}{\frac{db^\mu}{db} a b_\mu} \left(\frac{dG_b}{da}\right)^2 da^2 \qquad (4.203)$$

where retarded quantities only are assumed.

The quantity $\frac{dG_b}{da}$ is easily evaluated. Since we are concerned with points on the world line of "b" which are linked to "a" by null lines, we have

$$(a^\mu - b^\mu)(a_\mu - b_\mu) = 0$$

and by the definition of dG_b,

$$\left[\left(a^\mu + \frac{da^\mu}{da}da\right) - \left(b^\mu + \frac{db^\mu}{db}dG_b\right)\right]\left[\left(a_\mu + \frac{da_\mu}{da}da\right) - \left(b_\mu + \frac{db_\mu}{db}dG_b\right)\right] = 0$$

The difference between these two zero quantities yields, upon neglecting second order differentials

$$\frac{dG_b}{da} = \frac{\frac{da^\mu}{da}\, ab_\mu}{\frac{db^m}{db}\, ab_m} \qquad (4.204)$$

This result is a particular case of the theorem on null geodesics in Riemannian space by Kermack, McCrea and Whittaker.[8] Using (4.204) in (4.203) there obtains

$$dJ^2(a) = \eta_{mn}\, da^m da^n - \frac{2}{c^2} \sum_{b \neq a} \gamma m_b \frac{(ab_m da^m)^2}{\left(\frac{db^\mu}{db} ab_\mu\right)^3}$$

$$= g_{mn}(a) da^m da^n \qquad (4.205)$$

$$\text{where } g_{mn}(a) = \eta_{mn} - \frac{2}{c^2} \sum_{b \neq a} \gamma m_b \frac{ab_m ab_n}{\left(\frac{db^\mu}{db}\, ab_\mu\right)^3} \qquad (4.2051)$$

This is the form of Whitehead's principle used by Synge in a recent paper[4], although he did not include his derivation of it.

Hitherto Whitehead's formula has been expressed only in terms of Cartesian tensors. A general tensor formulation is made readily available by the existence of the distance D of Section 2.2. Using (2.212) we have

$$dJ^2(a) = da^2 - \frac{2}{c^2} \sum_{b\neq a} \frac{\gamma m_b}{D_R} dG_b{}^2 \tag{4.206}$$

$$= \alpha_{mn} da^m da^n - \frac{2}{c^2} \sum_{b\neq a} \frac{\gamma m_b}{D_R} \left(\frac{dG_b}{da}\right)^2 da^2 \tag{4.207}$$

where a^μ are now general coordinates and α_{mn} is the corresponding flat-space metric tensor. The theorem referred to in (8) yields

$$\frac{dG_b}{da} = \frac{\frac{da^\mu}{da} P_\mu}{\frac{db^m}{db} P_m} = \frac{\frac{da^\mu}{da} P_\mu}{\frac{db^m}{db} P_m} = \frac{\frac{da^\mu}{da} P_\mu}{D} \tag{4.208}$$

using the definitions (2.203) and (2.204). Applying this result to (4.207) there obtains

$$dJ^2(a) = g_{mn}(a)\, da^m da^n \tag{4.209}$$

where now

$$g_{mn}(a) = \alpha_{mn} - \frac{2}{c^2} \sum_{b\neq a} \gamma m_b \left(\frac{P_m P_n}{D^3}\right)_R \tag{4.210}$$

Alternatively by (2.216), we may use in (4.210)

$$\left(\frac{P_m P_n}{D^3}\right)_R = 2 \int_0^\infty \delta(S^2) \frac{P_m P_n}{D^2}\, db \tag{4.211}$$

We will call the tensor field

$$g_{mn}(x)_a \equiv a_{mn} - \frac{2}{c^2} \sum_{b \neq a} \gamma m_b \left(\frac{P_m P_n}{D^3} \right)_R \qquad (4.212)$$

the pseudo-metric tensor of particle "a", and the similar tensor field which includes summation over all the particles, namely

$$g_{mn}(x) \equiv a_{mn} - \frac{2}{c^2} \sum_{\text{all } b} \gamma m_b \left(\frac{P_m P_n}{D^3} \right)_R \qquad (4.213)$$

the associated metric tensor. The continuum will be described as the background space when referred to the metric a_{mn}, and as the associated Riemannian space when referred to the associated metric tensor.

It is easy to see from (4.213) that the world lines of particles are singularities in the associated metric tensor, while (4.201) with (4.209) precisely defines the world line of each particle to be a geodesic with respect to its pseudo-metric tensor field. The world line of a test particle which makes no contribution to the associated metric tensor will be a geodesic in the associated Riemannian space.

4.3. ADVANCE OF PERIHELION

We have noted that Whitehead's theory gives the same advance of perihelion as does the general theory of relativity. It will now be shown from what part of the structure of Whitehead's theory this advance of perihelion arises. This is accomplished by an investigation of a theory which differs from that of Whitehead by the omission of the factor

$$\left(\frac{dG_b}{da}\right)^2$$

from the second term on the right hand side of (4.203).

We then have

$$\mathrm{Var} \int dL(a) = 0 \tag{4.301}$$

where

$$dL^2(a) = da^2 - \frac{2}{c^2} \sum_{b \neq a} \frac{\gamma m_b}{\left(\frac{db^\mu}{db} ab_\mu\right)_R} da^2 \tag{4.302}$$

For a test particle (with cartesian coordinates x^μ) moving in the field of a single particle of mass m at rest at the origin, (4.302) simplifies to

$$d\,L^2 = \psi dS^2 \tag{4.303}$$

where
$$\psi = 1 - \frac{2\gamma_m}{rc^2} \tag{4.304}$$

and
$$dS^2 = \eta_{mn} dx^m dx^n$$

Assuming the motion to take place in the plane

$$x^3 = 0$$

and transferring to polar coordinates in the plane,

$$x^1 = r\cos\theta$$

$$x^2 = r\sin\theta$$

$$x^4 = ct$$

(4.303) becomes

$$dL = \psi^{1/2}(c^2dt^2 - dr^2 - r^2d\theta^2)^{1/2}.$$

The Lagrangian equations of motion resulting from (4.301) are

$$\frac{d}{dS}\left(\psi^{1/2}\frac{dr}{dS}\right) = -\frac{1}{2}\psi^{-1/2}\frac{\partial\psi}{\partial r} + \psi^{1/2}r\left(\frac{d\theta}{dS}\right)^2$$

$$\frac{d}{dS}\left(\psi^{1/2}r^2\frac{d\theta}{dS}\right) = 0$$

and

$$\frac{d}{dS}\left(\psi^{1/2}\frac{dt}{dS}\right) = 0$$

Substituting the value of ψ in (4.304) into these equations, there obtains

$$\frac{d^2r}{dS^2} - r\left(\frac{d\theta}{dS}\right)^2 = -\frac{\gamma m}{r^2c^2}\left(1+\left(\frac{dr}{dS}\right)^2\right)\Bigg/1-\frac{2\gamma m}{rc^2} \tag{4.305}$$

$$\frac{d}{dS}\left(r^2\frac{d\theta}{dS}\right) + \frac{d\theta}{dS}\frac{dr}{dS}\frac{\gamma m}{c^2}\Bigg/1-\frac{2\gamma m}{rc^2} = 0 \tag{4.306}$$

$$\frac{d^2t}{dS^2} + \frac{dt}{dS}\frac{dr}{dS}\frac{\gamma m}{r^2c^2}\Bigg/1-\frac{2\gamma m}{rc^2} = 0 \tag{4.307}$$

Equation (4.306) has the first integral

$$r^2\frac{d\theta}{dS} = \frac{h}{c}\left(1 - \frac{2\gamma m}{rc^2}\right)^{-1/2} \qquad (4.308)$$

where $\frac{h}{c}$ is a constant of integration. Using (4.308) to eliminate the parameter S from (4.305), and substituting $u = \frac{1}{r}$, we obtain without the use of any approximation

$$\frac{d^2u}{d\theta^2} + u = \frac{\gamma m}{h^2} \qquad (4.309)$$

which gives identically the classical path; the advance of the perihelion is exactly zero.

We can go father than this to show that the motion is the same as for classical gravitation. Equation (4.307) has the first integral

$$\frac{dt}{dS} = \frac{A}{c}\left(1 - \frac{2\gamma m}{rc^2}\right)^{-1/2} \qquad (4.310)$$

$\frac{A}{c}$ being a constant of integration. If (4.310) is now used to eliminate the parameter S from (4.305) and (4.306) there obtains

$$\left.\begin{aligned} \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 &= -\frac{\gamma m}{r^2A^2} \\ \frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) &= 0 \end{aligned}\right\} \qquad (4.311)$$

These equations of motion, derived from a Lorentz invariant theory for an observer at rest at the origin without recourse to any approximation, are identical to the corresponding classical equations of motion, except that the "effective gravitational constant", $\frac{-\gamma}{A^2}$ depends in a second order way on the configuration.

We conclude then that the advance of the perihelion in Whitehead's theory is due to the factor

$$\left(\frac{dG_b}{da}\right)^2$$

appearing in equation (4.203).

4.4. GRAVITATIONAL MASS AND CLASSICAL ENERGY

In his analysis of planetary motion, Whitehead assumed the central gravitating body (e.g. the sun) to be a single particle at rest at the origin of cartesian coordinates. This is, of course, the simplest representation of such a body. In a recent paper Synge(4) made an improvement on this approximate representation by considering the field of a finite spherically symmetric body consisting of a group of particles at rest but spatially separated from each other. This of course is not a precise use of Whiteheads theory of gravitation, for interactions between the particles of the body have been neglected. A precise use of Whitehead's theory would require the formidable task of first determining the motions of the particles of the group as demanded by that theory before one could discuss the associated metric field of the body as a whole. Synge's representation therefore uses Whitehead's theory in estimating the gravitational field of a body, but neglects the theory with respect to the behaviour of the particles of the body itself.

One might of course assume that the particles of a body are held in place by other (e.g. electromagnetic) forces. Such an assumption is undoubtedly implied (though not stated) in the case of Synge's work. Nevertheless the effect of possible velocities of particles about their mean positions is not taken into account. That such velocities enter into the calculation of gravitational potential in Whitehead's theory is made clear by a cursory examination of equation (4.202).

We shall now consider the gravitational field of a particle (of mass m) which makes arbitrarily small displacements about a fixed point (the origin of Cartesian coordinates), taking into account the velocity of the particle, which velocity need not be small. We need not be concerned with the forces which keep the particle near the origin - there might for example be another similar particle in close interaction, the pair of which would form a doublet.

The contribution of the particle to the associated metric tensor (4.213) is (in contravariant form)

$$\frac{2\gamma m}{c^2}\left(\frac{P^m P^n}{D^3}\right) \equiv V^{mn} \text{ say} \qquad (4.401)$$

Denoting by y^m the cartesian coordinates of the particle (y^i is assumed to be small) and using identities equivalent to those used in Section 4.2., namely

$$\left.\begin{aligned}
P^m &= xy^m \\
D &= \Omega(r-\zeta) \\
\Omega &= \left(1-\frac{v^2}{c^2}\right)^{-1/2} \\
v &= c\left[\frac{(dy^1)^2+(dy^2)^2+(dy^3)^2}{(dy^4)^2}\right]^{1/2} \\
r &= \left[(xy^1)^2+(xy^2)^2+(xy^3)^2\right]^{1/2} \\
\zeta &= \frac{1}{c}\,(\underline{xy}.\underline{v}) \\
(\underline{xy}.\underline{v}) &= c\left[(xy^1)\frac{dy^1}{dy^4}+(xy^2)\frac{dy^2}{dy^4}+(xy^3)\frac{dy^3}{dy^4}\right] \\
v^i &= \frac{dy^i}{dy^4} \quad (i = 1, 2, 3)
\end{aligned}\right\} \qquad (4.402)$$

the potential (4.401) can be expressed as

$$V^{mn} = \frac{2\gamma m}{c^2} \frac{(xy^m)(xy^n)}{\Omega^3(r-\zeta)^3}$$

$$= \frac{2Vm}{c^2} \frac{xy^m/r \;\; xy^n/r}{\Omega^3 r\left[1 - \frac{1}{c}(\underline{v}.\underline{xy}/r)\right]^3} \qquad (4.403)$$

It is seen that the effect of the coordinate y^i ($i = 1,2, 3$) can be made arbitrarily small by considering only regions x^i for which $\frac{y^i}{r}$ is arbitrarily small. We shall therefore approximate by neglecting y^i in (4.403), and write

$$V^{mn} \approx \frac{2\gamma m}{c^2} \frac{P^m P^n}{\Omega^3 r^3\left[1 - \frac{1}{c}\, v \cos\theta\right]^3} \qquad (4.404)$$

where now P^m refers not to the coordinates of particle but to the origin,

$$r = \left[(x^1)^2 + (x^2)^2 + (x^3)^2\right]^{1/2}$$

and θ is the angle between the 3-space vectors $\underline{v}$ and $\underline{x}$.

The only factor in (4.404) which distinguishes the particle from one fixed (without motion) at the origin is the classical velocity $\underline{v}$.

Expanding the quantity

$$\frac{1}{\Omega^3(1 - \frac{v}{c}\cos\theta)^3} \qquad (4.405)$$

there obtains

$$\left\{1 - \frac{3}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \frac{1}{8}\frac{v^6}{c^6} + \frac{3}{128}\frac{v^8}{c^8} + \dots\right\}\left\{1 + 3\frac{v}{c}\cos\theta + 6\frac{v^2}{c^2}\cos^2\theta + 10\frac{v^3}{c^3}\cos^3\theta \dots\right\} \quad (4.406)$$

Because of the presence of the variable quantity $\cos\theta$ in (4.406) the potential under consideration depends not only on the velocity of the particle but on the direction of this velocity with respect to the point of observation. We will now use an averaging process, where we consider all directions of the vector v as being equally probable, the average value of $\cos\theta$ being the average of that for all possible directions.

On this basis the probability of θ having a value between ϕ and $\phi+d\phi$ will be proportional to $\sin\phi$, in fact it will be equal to $\frac{\sin\phi}{2}$, since the range of ϕ is from 0 to π. Thus the expected or mean value of $\cos^n\theta$ is given by

$$\overline{\cos^n\theta} = \frac{1}{2}\int_0^\pi \cos^n\phi\sin\phi d\phi$$

$$= \frac{1}{n+1} \text{ if n is even}$$

$$= 0 \quad \text{if n is odd.}$$

Using the average values of $\cos^n\theta$ in (4.406), we find that the average or expected value of (4.405) is given by

$$\left\{1 - \frac{3}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \frac{1}{8}\frac{v^6}{c^6} + \frac{3}{128}\frac{v^8}{c^8} + \dots\right\}\left\{1 + 2\frac{v^2}{c^2} + 3\frac{v^4}{c^4} + 4\frac{v^6}{c^6} + \dots\right\} \quad (4.407)$$

A term by term multiplication of this product reveals that each resultant term is positive, and thus the product is greater than unity for any v less than c. Using this result in (4.404) it is seen that on the average at least, the gravitational potential of the particle under consideration is greater than that of a particle of equal mass at rest at the origin.

If the velocity v is small enough that terms in $\frac{v^2}{c^2}$ higher than the first order may be neglected, then (4.407) becomes simply

$$1 + \frac{1}{2}\frac{v^2}{c^2}$$

and applying this result to (4.404)

$$V^{mn} = 2\gamma m \left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) \frac{P^m P^n}{r^3} \qquad (4.408)$$

Here, on the average the effective gravitational mass is equal to the sum of the constant mass m plus the mass equivalent of the classical kinetic energy.

Gravitational Mass and Potential Energy

The identification of inertial mass with energy is an achievement of the special theory of relativity. Gravitation, however, is not an essential part of that theory, and the precise meaning of gravitational mass will depend on the particular theory of gravitation. Except in the simplest cases the general theory of relativity is necessarily rather vague on the meaning of gravitational mass and its relationship to the energy momentum tensor.

Experiment indicates that at least in the case of nuclear physics a configuration of particles does not have a gravitational (or inertial) mass equal to the sum of those of the individual particles, but that the total mass is altered according to the energy of the configuration. We might expect a theory of gravitation to account for this effect even though we do not know the nature of the binding forces, or alternatively to predict it qualitatively in terms of known forces.

The type of averaging process (or a probability basis) used in the derivation of (4.408) is not inconsistent with the spirit of quantum theory. It was found that Whitehead's theory includes kinetic but not potential energy in gravitational mass. An alteration of Whitehead's theory will now be proposed which includes classical potential energy in gravitational mass. Before stating this change, we note that if we add Wheeler and Feynman electrodynamics to Whitehead's theory, we would require the equations of motion to be derived from

$$\mathrm{Var} \int d\tau(a) = 0 \qquad (4.409)$$

where $$d\tau(a) = dJ(a) + \frac{e_a}{m_a c^2} F_m(a) da^m \qquad (4.410)$$

$F_m(a)$ being the electromagnetic vector potential. (4.410) corresponds to equation (3.501).

The proposed change for Whitehead's theory is to replace the quantity

$$dG_b^{\,2}$$

in equation (4.202) by

$$dG_b \, d\tau(b).$$

Noting that dG_b may be identified with db, this alteration is equivalent to replacing the constant mass m_b in (4.202) by the scalar μ_b, where

$$\mu_b = m_b \frac{d\tau(b)}{db} \tag{4.411}$$

We now calculate this new gravitational mass for the case where the particle "b" is in the gravitational and electromagnetic fields of particles which are at rest. Then using (4.205), (3.301) and (4.410),

$$\mu_b = m_b \left[1 - \sum_{a \neq b} \frac{2\gamma\mu_a}{r^3 c^2} \left(b a_m \frac{db^m}{db} \right)^2 \right]^{1/2} + \frac{e_a}{c^2} \sum_{a \neq b} \frac{e_b}{r} \frac{db^4}{db}$$

Here r is the spatial distance from "b" to "a". If the second term in the square bracket is small compared to unity, we will have

$$\mu_b = m_b \left[1 - \sum_{a \neq b} \frac{\gamma\mu_a}{r^3 c^2} \left(b a_m \frac{db^m}{db} \right)^2 \right] + \frac{e_b}{c^2} \sum_{a \neq b} \frac{e_a}{r} \frac{db^4}{db}$$

If we neglect higher powers of $\frac{v^2}{c^2}$ than the first, where v is the classical velocity of "b", then this expression reduces to

$$\mu_b = m_b - \frac{\gamma m_b}{c^2} \sum_{a \neq b} \frac{\mu_a}{r} + \frac{e_b}{c^2} \sum_{a \neq b} \frac{e_a}{r} \tag{4.412}$$

which is the Whitehead mass plus the equivalent mass of the classical potential energy. Although we assumed the particles "a" to be at rest their motions (if not too large) would not affect this approximation. It is interesting to note that no new kinetic energy terms arise here to spoil the previous result (4.408).

4.5. TIME SYMMETRICAL THEORY

It has been emphasised that Whitehead's gravitational potentials are entirely of a retarded nature. A corresponding time-symmetrical theory can be constructed simply by replacing (4.213) by

$$\overset{*}{g}_{mn}(x) = a_{mn} - \frac{1}{c^2} \sum_b \gamma m_b \left[\left(\frac{P_m P_n}{D^3} \right)_R + \left(\frac{P_m P_n}{D^3} \right)_S \right] \qquad (4.501)$$

$$= a_{mn} - \frac{2}{c^2} \sum_b \gamma m_b \int_{-\infty}^{\infty} \delta(S^2) \frac{P_m P_n}{D^2} \, db \qquad (4.502)$$

using (2.215) in (4.501) to obtain (4.502). In terms of cartesian tensors (4.501) or (4.502) may be equivalently expressed as

$$g_{mn}(x) = \eta_{mn} - \frac{1}{c^2} \sum_b \gamma m_b \left[\left(\frac{xb_m xb_n}{\left\{ \frac{db^\mu}{db} xb_\mu \right\}^3} \right)_R + \left(\frac{xb_m xb_n}{\left\{ \frac{db^\mu}{db} bx_\mu \right\}^3} \right)_S \right] \qquad (4.503)$$

Use is made here of (2.2141) and the fact that using cartesian coordinates

$$P^\mu = xb^\mu$$

This follows directly from (2.318) (by taking the limit as k tends to zero).

For a single particle (of mass m) at rest at the origin, (4.503) reduces to

$$\begin{aligned} g_{44} &= 1 - \frac{2\gamma m}{rc^2} \\ g_{4i} &= 0 \\ g_{ij} &= -\delta_{ij} - \frac{2\gamma m}{r^3 c^2} x_i x_j \qquad (i, j = 1,2,3) \end{aligned} \qquad (4.504)$$

where $r = (\delta_{ij}\, x^i x^j)^{1/2}$.

---

* Use of this time-symmetrical metric is confined to the present section.

This is the same as for Whitehead's theory (see (4.2021)) except for the absence here of the components g_{4i}. (4.504) is also a very good approximation to the Schwarzschild solution of the general theory of relativity.

The differential equations satisfied by the path of a test particle moving in the field (4.504) can be determined using a method identical to that used by Bergmann[16] (Chapter XIV) for the Schwarzschild field. If the orbit of the test particle is in the plane $x^3 = 0$, it can be shown that

$$\frac{d^2u}{d\phi^2} + u = \frac{\gamma m}{h^2}\left(1 + \frac{3h^2u^2}{c^2}\right) + \frac{4\gamma^2 m^2 u}{c^4}\left(u\frac{d^2u}{d\phi^2} + \left(\frac{du}{d\phi}\right)^2\right) \qquad (4.505)$$

where $u = \frac{1}{r}$,

$$\tan\phi = x^2/x^1$$

and h/c is the constant angular momentum.

The second term on the right hand side of (4.505) is small enough to be neglected in any practical problem, and the remaining equation is the same as that obtained for the Schwarzschild field. Thus advance of perihelion (which arises through the term $\frac{3h^2u^2}{c^2}$ in the first bracket) is essentially the same for Whitehead's theory (which gives the same advance as does the general theory of relativity) and for this time-symmetrical modification of Whitehead's theory.

This is not a surprising result, since the phenomenon of advance of perihelion is in every way symmetrical with respect to positive and negative time.

4.6. EXTENSION OF WHITEHEAD'S THEORY OF GRAVITATION TO DE SITTER SPACE-TIME

Whitehead's fundamental gravitational equation, (4.202) may be written in the following form:

$$dJ(a)^2 = da^2 - \frac{2}{c^2} \sum_{b \neq a} \gamma m_b \; V dG_b{}^2 \qquad (4.601)$$

where $2\gamma m_b V/c^2$ is the gravitational potentisl of particle "b". In order to generalize the theory to a general Riemannian background space, it is merely necessary to require the potential V to be an invariant function of the coordinates which in the special case of Cartesian coordinates in flat space makes (4.601) equivalent to (4.203). In this way Temple(9) generalized Whitehead's theory to de Sitter space, using a potential defined by

$$S^{mn} \; V_{(T)}{}_{|mn} = 0 \qquad (4.602)$$

where S_{mn} is the de Sitter metric tensor. A complete specification of $V_{(T)}$ requires appropriate boundary conditions in addition to (4.602).

We now propose a more direct generalization of Whitehead's theory by writing

$$V = \frac{1}{D_R} \qquad (4.603)$$

where D_R is defined by (2.205) and (2.204). That $V_{(T)}$ and V are different in de Sitter space is made clear by the first of (2.358).

(4.603) may be used as a generalization to any Riemannian background space, but we shall confine our attention in what follows to de Sitter space.

The scalar D_R was used in Section 4.2. to put Whitehead's theory into general tensor form in flat space, and equations (4.209) to (4.213) are also valid in our de Sitter space generalization (with a_{mn} being replaced by S_{mn}). In particular, we re-write (4.213)

$$g_{mn}(x) = S_{mn} - \frac{2}{c^2} \sum_b \gamma m_b \left(\frac{P_m P_n}{D} \right)_R \qquad (4.604)$$

Field of a Particle at Rest

Using the de Sitter background metric (2.303) we now calculate the field (4.604) of a single particle of mass m at rest at the origin of spatial coordinates and with time coordinate t_o. We have in (2.318), (2.327) and (2.317) (with $x_o{}^i = v_o{}^i = 0$, $v_o{}^4 = 1$)

$$\left.\begin{aligned} P_4 &= (1 - \exp k(t_o - t))/k \\ P_i &= x_i \exp k(t+t_o) \\ D &= (\exp k(t-t_o) - 1)/k \\ r &= (\exp(-kt_o) - \exp(-kt))/k \\ \text{where} \quad r &= (\delta_{ij}\, x^i x^j)^{1/2}, \quad (i, j = 1, 2, 3). \end{aligned}\right] \qquad (4.605)$$

Using (4.605) in (4.604) there obtains

$$\left.\begin{aligned} g_{44} &= 1 - \frac{2\gamma m \exp(-kt)}{rc^2(1 + kr \exp kt)^2} \\ g_{4i} &= - \frac{2\gamma m\, x_i}{r^2c^2(1 + kr \exp kt)^2} \\ g_{ij} &= - \delta_{ij} \exp 2kt - \frac{2\gamma m\, x_i x_j \exp kt}{r^3c^2(1 + kr \exp kt)^2} \end{aligned}\right] \qquad (4.506)$$

Transferring to spherical polar coordinates, given by

$$\left.\begin{aligned} x^1 &= r\sin\theta\cos\phi \\ x^2 &= r\sin\theta\sin\phi \\ x^3 &= r\cos\theta \end{aligned}\right\} \quad (4.6061)$$

and writing $d\tau^2 = g_{mn}\, dx^m dx^n$

the associated metric is given by

$$\begin{aligned} d\tau^2 = {} & \left[1 - \frac{2\gamma m \exp(-kt)}{rc^2(1 + kr\exp kt)^2}\right] dt^2 \\ & + \left[\frac{4\gamma m}{rc^2(1 + kr\exp kt)^2}\right] dt\,dr \\ & - (dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2)\exp 2kt \\ & - \left[\frac{2\gamma m \exp kt}{rc^2(1 + kr\exp kt)^2}\right] dr^2 \end{aligned} \quad (4.607)$$

The exponentials can be eliminated by substituting

$$\rho = r\exp kt \quad (4.608)$$

yielding

$$\begin{aligned} d\tau^2 = {} & (1 - k^2\rho^2 - 2\gamma m/\rho c^2)dt^2 \\ & + (2k\rho + 4\gamma m/\rho c^2(1+k\rho))dt\,d\rho \\ & - \rho^2(d\theta^2 + \sin^2\theta\, d\phi^2) \\ & - (1 + 2\gamma m/\rho c^2(1 + k\rho)^2)d\rho^2 \end{aligned} \quad (4.609)$$

The cross term containing the product of differentials of the first and fourth coordinates can be eliminated by the further substitution

$$t = t' + f(\rho)$$

where f is a function of ρ satisfying

$$\frac{df}{d\rho} = -\frac{k\rho + 2\gamma m/\rho c^2(1 + k\rho)}{1 - k^2\rho^2 - 2\gamma m/\rho c^2}$$

(4.609) then simplies to

$$\begin{aligned} d\tau^2 &= (1 - k^2\rho^2 - 2\gamma m/\rho c^2)dt'^2 \\ &\quad - \rho^2(d\theta^2 + \sin^2\theta d\phi^2) \\ &\quad - (1 - k^2\rho^2 - 2\gamma m/\rho c^2)^{-1} d\rho^2 \end{aligned} \tag{4.610}$$

This is precisely the spherically symmetric exterior solution to Einstein's field equations with a non-zero cosmological constant.[18] Denoting the cosmological constant by Δ, then $\Delta/3$ appears in that solution instead of the k^2 in (4.610). The field equations are

$$G_{mn} = \Delta g_{mn}$$

where G_{mn} is the Einstein tensor. Hence the metric (4.610) satisfies

$$G_{mn} = 3k^2 g_{mn} \tag{4.611}$$

But (4.611) is a necessary (though not sufficient) condition for constant curvature[9], and for the problem under consideration the background metric (2.303) and the associated Riemannian metric both satisfy the same tensor equation (4.611).

It is easy to show, using Bergmann's method[16] that the following differential equation is satisfied by a test particle moving in the plane $\theta = \pi/2$ in the gravitational field (4.607)

$$\frac{d^2u}{d\phi^2} + u = \frac{\gamma m}{h^2}(1 + 3h^2u^2) - \frac{k^2}{h^2u^3} \tag{4.612}$$

where $u = \frac{1}{\rho}$

$\frac{h}{c}$ is the constant angular momentum.

The last term on the right hand side of (4.612) appears here because of the curvature of the background space, and if the curvature is not too large it may be neglected, leaving the same result as for the Schwarzschild solution with no cosmological constant.

The following significance is attached to equation (4.612). Consider a test particle moving essentially in a circular orbit. (The purpose of this requirement is merely to simplify the arguments which follow). Classically, for a circular orbit, the angular momentum is proportional to the square root of distance. Applying this as an approximation in the present case (where u corresponds roughly to inverse distance), we conclude at least that k^2/h^2u^3 is a decreasing function of u. It thus follows that this term is smaller for small orbits than for large orbits. (Other varying terms in (4.612) decrease with increased size of the orbit). Consider an orbit small enough for k^2/h^2u^3 to be neglected. Then neglecting also the relativistic term $3u^2\gamma m$ (which leads to the advance of perihelion), we have left the classicalpath with u and hence ρ being periodic functions of ϕ. Noting (4.508) it follows that r is on the average a decreasing function of the time coordinate t.

Now in steady-state cosmology the average large scale motions of particles are along geodesics of constant r. (These geodesics describe the motion of "ideal" particles. Thus with respect to the continuum of ideal particles, the test particle under consideration converges on the central gravitating body. Equation (4.612) is therefore of some importance in the explanation of the condensation of nebulae in steady-state theory.

Gravitational Potential of a Steady-State Distribution of Mass

Employing the background metric (2.303) we now determine the associated metric (4.604) for a constant proper density of mass (with respect to the background metric), the motion of each particle being given by (3.523).

Denoting the coordinates of a typical particle by x_o^i, t_o, $(i = 1, 2, 3)$, and its mass by m, we first evaluate the potential

$$V_{mn} \equiv \frac{2}{c^2} \sum_b \gamma m_b \left(\frac{P_m P_n}{D^3}\right)_R \tag{4.6121}$$

at the origin of spatial coordinates and at time t. Writing

$$\left.\begin{aligned} x_o^1 &= r \sin\theta \cos\phi \\ x_o^2 &= r \sin\theta \cos\phi \\ x_o^3 &= r \cos\theta \\ r^2 &= \delta_{ij}\, x_o^i x_o^j \end{aligned}\right] \tag{4.613}$$

and using (2.317), (2.318) and (2.327), we obtain

$$\left.\begin{aligned} D &= r \exp kt \\ P_4 &= r \exp kt_o \\ P_i &= x_o^i \exp k(t+t_o) \end{aligned}\right] \tag{4.614}$$

If ρ is the constant proper density then

$$V_{mn_{(x^i=0)}} = \frac{2\gamma\rho}{c^2} \int_0^\infty \int_0^\pi \int_0^{2\pi} \left(\frac{P_m P_n}{D^3}\right)_R \exp 3kt_0 \, r^2 \sin\theta \, d\phi \, d\theta \, dr \tag{4.615}$$

The limits of the variable r are made evident by (2.317) and the factor $\exp 3kt_0$ is present because we are integrating over proper volume.

Using (4.613) and (4.614) in (4.615), and noting from (2.317) that

$$r = (\exp(-kt_0) - \exp(-kt))/k$$

$$dr = -dt_0 \exp(-kt_0)$$

there obtains

$$V_{44_{(x^i=0)}} = \frac{8\pi\gamma\rho}{kc^2} \int_{-\infty}^{t} \left[\exp 3k(t_0-t) - \exp 4k(t_0-t)\right] dt_0$$

$$= \frac{2\pi\gamma\rho}{3k^2c^2}$$

$$V_{4i_{(x^i=0)}} = 0$$

$$V_{ij_{(x^i=0)}} = \delta_{ij} \frac{8\pi\gamma\rho}{3kc^2} \exp 2kt \int_{-\infty}^{\infty} \left[\exp 3k(t_0-t) - \exp 4k(t_0-t)\right] dt_0$$

$$= \delta_{ij} \frac{2\pi\gamma\rho}{9k^2c^2} \exp 2kt$$

This procedure would not work in a time-symmetrical theory, for the integrals just evaluated would not converge for advanced potentials. Using the above results in (4.6121) and thus in (4.604) we find also using (2.303) that at $x^i = 0$ the associated metric tensor is given by

$$\left.\begin{aligned} g_{44} &= 1 - \frac{2\pi\gamma\rho}{3k^2c^2} \\ g_{i4} &= 0 \\ g_{ij} &= -\delta_{ij}\left(1 + \frac{2\pi\gamma\rho}{9k^2c^2}\right)\exp 2kt \end{aligned}\right] \quad (4.616)$$

This result is independent of the choice of origin of spatial coordinates and it is therefore easy to show that (4.616) is the required associated metric tensor at any point. The associated metric tensor for a steady-state distribution of mass is given in terms of spherical polar coordinates by

$$d\tau^2 = (1 - 2\pi\gamma\rho/3k^2c^2)dt^2$$
$$- (1 + 2\pi\gamma\rho/9k^2c^2)(dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2)\exp 2kt \quad (4.617)$$

while the background metric is

$$dS^2 = dt^2 - (dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2)\exp 2kt \quad (4.618)$$

We thus have the remarkable result that a steady-state distribution of mass in a de Sitter background space (B) produces a de Sitter associated Riemannian space (A), subject to the condition that

$$2\pi\gamma\rho/3k^2c^2 < 1. \quad (4.619)$$

Using the value[12]

$$1/k = 1.5 \times 10^{27} \text{ secs.}$$

and the accurately known value of γ

$$\gamma = 6.7 \times 10^{-8} \text{ cm}^3/\text{gm.sec.}^3$$

We find that (4.619) requires

$$\rho < 3 \times 10^{-27} \text{ gms./cm}^3 \tag{4.6191}$$

The best observational value of the mean density of matter in space seems to be within a factor of 10 from 10^{-27} gm/cm^3,[12] so that (4.6191) is neither confirmed nor contradicted by experiment. Such a comparison is not very accurate, however, because if light is more closely associated with A than with B (see Chapter V) then the observed density of matter (and the observed radius of curvature of space) would be those values which pertain to A.

Substituting

$$\begin{aligned} t' &= (1 - 2\gamma\rho/3k^2c^2)^{1/2}\, t \\ r' &= (1 + 2\pi\gamma\rho/9k^2c^2)^{1/2}\, r \\ k' &= (1 - 2\gamma\rho/3k^2c^2)^{-1/2} k \end{aligned} \tag{4.620}$$

in (4.617), there obtains

$$d\tau^2 = dt'^2 - (dr'^2 + r'^2 d\theta^2 + r'^2 \sin^2\theta d\phi^2) \exp 2k't' \tag{4.621}$$

Proper 3-space volumes in B and in A may be compared by the following equation,

$$r'^2 \sin\theta dr' d\theta d\phi \exp k't' = (1 + 2\pi\gamma\rho/9k^2c^2)^{3/2} r^2 \sin\theta dr d\theta d\phi \exp kt \tag{4.622}$$

which is obtained from (4.620), or

$$dv' = (1 + 2\pi\gamma\rho/9k^2c^2)^{3/2}\, dv$$

where dv is an element of proper volume. The density will be inversely proportional to the proper volume of a given element of 3-space, so we may write

$$\rho' = (1 + 2\pi\gamma\rho/9k^2c^2)^{-3/2}\, \rho \tag{4.623}$$

where ρ' is the density of matter in A.

It is interesting to note the form taken in A by the null surfaces of B. We have for these null surfaces (see (2.317)),

$$kr = \exp(-kt_0) - \exp(-kt)$$

or, using (4.620)

$$k'r' = (1 + 2\pi\gamma\rho/9k^2c^2)^{1/2}(1 - 2\gamma\rho/3k^2c^2)^{-1/2} \times \left[\exp(-k't'_0) - \exp(-k't')\right]$$

or

$$k'r' > \exp(-k't'_0) - \exp(-k't')$$

But the null surfaces of A are given by

$$k'r' = \exp(-k't'_0) - \exp - 9k't')$$

It follows immediately that the null surfaces of B are space-like surfaces in A. But gravitational effects are defined along null lines in B, and if electromagnetic effects are defined along null lines in A, then we would conclude that gravitational effects travel faster than the speed of light.

CHAPTER V

GENERAL COSMOLOGICAL CONSIDERATIONS

In Chapter III it was found that a particle theory of electrodynamics could be formulated in a steady-state universe (and with the de Sitter metric demanded by steady state theory) for which no particle acts on itself and for which radiation damping is not an intrinsic property of a particle but is a large-scale property of the universe as a whole. Such a theory is most attractive from a cosmological point of view, and explicitly satisfies the electrodynamical equivalent of Mach's principle. It is worth pointing out that the means of acquiring irreversibility of radiation as suggested by Wheeler and Feynman in their flat-space formulation of the theory does not depend upon the universe as a whole. In that theory, initial conditions, which applied for some unknown reason at some time in the past, are to account for the phenomenon of radiation damping throughout all of space-time. The initial conditions constitute an additional postulate or principle in their theory - a principle which is not compatible with the perfect cosmological principle. The main contribution of the present work to the theory of Wheeler and Feynman has been to remove the initial condition postulate.

Electrodynamics and Metric

Metrical space-time plays a fundamental role in the theory of particle electrodynamics. Electromagnetic potential is defined along null geodesics, and in that sense at least, we may say that light travels along the null geodesics of the metrical space-time. Proper time for electrodynamic phenomenon is also defined by the metric tensor.

In the gravitation of Chapter IV two metrics are encountered, namely, the metric of the background space B and that of the associated space A. Both of these metrics are defined in the same continuum; the former is used to define gravitational potential, and the latter, derived from the former and the gravitational potential, essentially defines the inertial frames of particles. Which then is the metric to be associated with electrodynamics.

Experiment gives a clear answer. The phenomenon of the red-shift of the light from heavy bodies and the bending of the path of light in the neighbourhood of heavy bodies clearly indicate that our electrodynamics should be framed in A. And since our measurements of time and distance are dependent upon phenomena which are essentially of an electrodynamic nature, our experience is directly related not to B, but to A.

Whitehead attached great importance to the spatial and temporal meaning of the cartesian coordinates of the Minkowski background space of his theory. But what significance are we to attach to a time we do not experience and a space we cannot measure. The only real significance of B, from the present point of view, is that it provides a (metric) tensor in terms of which Whitehead's principle describes gravitational effects.

The use of A as electrodynamic metric poses a new problem. It was found in Section 4.6. that both B and A could be spaces of constant curvature. However local irregularities were not taken into account; in nature A has local deviations from a de Sitter metric. Although we were able to define an electrodynamics for general Riemannian space-time, we were only able to devise one for which Maxwell's equations were satisfied in de Sitter space and in flat space. This is not a serious difficulty because the experimental evidence of Maxwell's equations applies in our theory only to small scale (local) phenomenon, and to the cosmological phenomenon of radiation. Locally flat-space is applicable in the former case, and de Sitter space in the second. For orders of magnitude between these two extremes Maxwell's equations will at least be approximately satisfied. Nevertheless it would be most desirable to find an electromagnetic potential which satisfies Maxwell's equations generally.

Inertia

The origin of inertia is a subject of great concern to a large number of scientists. A clear and concise statement of the problem has been given by Bondi[12]. One view - that of Newton and Whitehead - regards inertia as a fundamental property of the background space, while the other, due to Mach, regards inertia as a relative property determined by the distribution of mass in the universe.

An explicit theory of relativity of inertia faces great inherent difficulties. Because of the seemingly constant nature of the gravitational "constant", inertia must depend largely on distant matter with very little contribution from local irregularities.

Now in its final form a gravitational theory must provide a metric tensor in terms of which to define inertial frames, and this metric must be determined by the distribution of mass. But until a metric is defined the terms "distribution" of mass and "distant" matter have little or no meaning.

Probably the most pertinent physical fact to be explained by a theory of inertia is why dynamical experiments and observations on the motions of stars yield very accurately the same rate of rotation of the earth, or in other words, why the stars do not appear to rotate with respect to our inertial frame. Neglecting the unsatisfactory explanation of pure coincidence, the only answer Whitehead's theory can give to this problem is that large scale rotations with respect to the background space (which determines local inertia) are not compatible with the theory. This will most certainly be true in the case of pure roation, but Bondi(12) has pointed out that an apparent rotation could be obtained by a uniform shear, each particle moving say with a uniform velocity in the direction of the x axis, the velocity being proportional to the y coordinate. Bondi's claim is that for a theory which is Lorentz invariant, if a configuration where each particle is at rest is possible, so also is this configuration of shear, where each particle has uniform motion.

Whitehead's theory is in a somewhat better position with respect to this problem if the background space is not the Minkowski space-time of special relativity, but the space-time of constant curvature of de Sitter. In the latter case large scale Lorentz transformations are no longer possible, and it seems more than likely that large scale apparent rotations due to shear would no longer be a possible configuration in Whitehead's theory. Whether or not other apparent rotations are possible in Whitehead's theory is a matter of considerable importance, and requires further investigation.

Particle Theory of Gravitation and the General Theory of Relativity

The fact that Whitehead's Theory yields exactly the Schwarzschild solution of the general theory of relativity (with or without the cosmological constant) suggests a possible close relationship between these two theories. We now discuss the form that such a relationship might take.

The following features of Whitehead's gravitation are evident:

(a) The theory produces a second-rank symmetric tensor field in terms of which as metric the world-lines of particles are singularities, the world lines of test particles are geodesics, and the "path" of light is along null geodesics.

(b) Since the associated space is Riemannian the identities of Riemannian geometry automatically provide adjunct field conservation equations.

(c) The adjunct field (Einstein tensor) vanishes (if B is flat) or is proportional to the associated metric tensor (if B is de Sitter) for a single particle at rest, except on the world line of the particle.

(a) A steady-state distribution of matter in a de Sitter universe is a possible and self-consistent configuration (provided the background space is also de Sitter and the density of matter is not too large).

These features are all compatible with the general theory of relativity. Provided the Einstein tensor is capable of a suitable physical interpretation as "field" energy, we might say that Whitehead's theory is a special case of general relativity.

It is interesting to note that in General Relativity the interpretation of the stress-energy-momentum tensor is somewhat flexible. It was by a re-interpretation of this tensor that McCrea(19) was able to show (d) to be a feature of the general theory of relativity.

The notion of stress is not explicitly contained in Whitehead's theory. However by the inclusion of potential energy in the gravitational mass of a particle (as is done in Section 4.4.), it may be possible to find an interpretation of the concept of stress.

The suggested relationship between the two theories is illustrated by the following example. A further investigation might provide a general solution to the two body problem in Whitehead's theory. The Einstein tensor could then be calculated. Provided this tensor is capable of a physical interpretation as representing field energy (e.g. gravitational energy), then we would say that the field equations of general relativity have been satisfied.

The concept of field energy is not incompatible with a particle theory. This is illstrated by an electrodynamical analogy. The effect of a source of light is manifested by an excitation of particles of matter at a different place at some later time. From an observational point of view energy disappears and later reappears somewhere else. We do not observe its passage, but it is convenient to think of a field of light moving from one point to another. By this means we are able to effect a conservation of energy.

REFERENCES

(1) J.A. Wheeler and R.P. Feynman, Reviews of Modern Physics, 21, 425 (1949).

(2) J.A. Wheeler and R.P. Feynman, Reviews of Modern Physics, 17, 157 (1945).

(3) A.N. Whitehead, "The Principle of Relativity", Cambridge (1922).

(4) J.L. Synge, Proceedings of the Royal Society, A211, 303, (1952).

(5) J.L. Synge, "The Relativity Theory of A.N. Whitehead", Lecture Series No. 5, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, (1951).

(6) A. Einstein, "The Meaning of Relativity" Princeton University Press, (1945).

(7) B. Friedman, "Techniques in Solving Partial Differential Equations". New York University, Institute for Mathematics and Mechanics (1950).

(8) W.O. Kermack, W.H. McCrea and E.T. Whittaker. Proceedings of the Royal Society of Edinburgh, 53, 31, (1932/33).

(9) G. Temple, Proceedings of the Cambridge Philosophical Society, 36, 176 (1924).

(10) P.A.M. Dirac, Proceedings of the Royal Society, A167, 148 (1938).

(11) G. Stephenson and C.W. Kilmister, Il Nuovo Cimento, 10, 230 (1953).

(12) H. Bondi, "Cosmology", Cambridge University Press, (1952).

(13) C.O. Hines, Nature, 170, 582, (1952).

(14) J.L. Synge, Physical Review, 47, 760, (1935).

(15) A.S. Eddington, Nature, 113, 192 (1924).

(16) P.G. Bergmann, "Introduction to the Theory of Relativity", Prentice-Hall, New York, (1942).

(17) W.H. McCrea, "Relativity Physics", Methuen, (1935).

(18) R.C. Tolman, "Relativity Thermodynamics and Cosmology", Clarendon Press, Oxford, (1934).

(19) W.H. McCrea, Proceedings of the Royal Society, A206, 562, (1951).

Reprinted from the *Proceedings of the Cambridge Philosophical Society*,
Volume 48, Part 4, pp. 616–624, 1952.
PRINTED IN GREAT BRITAIN

THE RELATIVISTICALLY RIGID ROD

BY J. E. HOGARTH* AND W. H. McCREA

Received 15 April 1952

ABSTRACT. A relativistically rigid rod is defined as one along which any disturbance travels with the fundamental speed c without dissipation of energy. A necessary consequence of this definition is that the rod should satisfy a certain 'equation of elasticity'. If this equation is satisfied, it is shown in accordance with relativistic mechanics that a disturbance of any amplitude is propagated along the (unstressed or stressed) rod with speed c and without dissipation. Consequently, the equation provides also a sufficient condition for rigidity in the sense defined.

1. *Introduction*. The term *rod* is here applied to a uniform one-dimensional body, and all motions, whether of translation or of deformation, are restricted to the single spatial dimension in which the body is taken to exist. Consequently the only stresses to be considered are longitudinal.

If desired, the work can be applied to a three-dimensional body if it be supposed capable of sustaining stress in only one direction and if all strains and motions be restricted to this direction. Such application is of no additional physical interest. However, it permits us to employ, when convenient, the mathematical treatment appropriate to a three-dimensional body including, for instance, the ordinary use of a stress-tensor. We hope in due course properly to extend the present work to the treatment of isotropic three-dimensional bodies.

In this paper we are concerned only with *special* relativity theory.

One of us has recently (4) considered the problem of a rigid rod in accordance with the theory of special relativity. He defined *a rigid rod as one along which disturbances are transmitted with speed c*. Here c is the fundamental speed of relativity theory, usually described as the speed of light, though this description happens to be irrelevant in the present paper.

With this definition the rod was shown to be not inextensible. But, since it possesses the maximum degree of rigidity consistent with the requirements of special relativity, the term 'rigid' is appropriate. We are concerned here only with longitudinal rigidity; nothing corresponding to 'flexural rigidity' makes its appearance in the present one-dimensional problems.

As usually developed, the theory of special relativity postulates a rigid rod as the fundamental measuring scale. Therefore it might be argued that there can be no meaning in the theory for a change of length of a rigid rod. As a consequence of the previous work referred to it is seen, however, that it is necessary to take as a measuring scale a rigid rod that is *at rest in an inertial frame and is unstressed*. If a similar rod is in uniform motion relative to the first, it is of course taken as providing an equivalent measuring scale in the inertial frame in which it is at rest. When we say that a rigid rod is extensible we mean that, if it is allowed to attain equilibrium in an inertial frame under an applied tension then its length as measured by an unstressed rod in the same

* Member of the staff of Radio Physics Laboratory, Defence Research Board of Canada.

frame differs from its length before the tension was applied. Also, if the rod is accelerated relative to an inertial frame it is in general thereby stressed, and its length, as measured by an observer in the inertial frame using an unstressed rigid scale at rest in his frame, depends upon the stress-distribution and so upon the manner in which the accelerating forces are applied to the moving rod. Indeed, the rod cannot be accelerated like a classical rigid body but only (roughly speaking) like a classical deformable body. The processes of stopping a moving rod, as discussed in the previous paper, illustrate this feature. It is seen that nothing in all this is inconsistent with the principles of special relativity theory or devoid of observational interpretation in accordance with the theory.

From this discussion it follows that, if a body is accelerated relative to an inertial frame, then an observer attached to the body cannot use a rigid rod unambiguously to measure lengths unless its method of employment is prescribed in considerable detail. For, suppose he holds two such rods side by side, holding them first by the same ends and then shifting his hold upon one of them to its other end. In general, he will observe a relative change of length. In the first case he is pushing both rods or pulling both, while in the second case he is pushing one rod and pulling the other, with the result that there is a difference in their states of strain. This remark is trivial in itself, but it appears to have interesting consequences in regard to the so-called 'principle of equivalence'. The latter will not be pursued here, though they provide one reason for the investigation. The remark is, too, an illustration of the significance of the 'drag-point' defined by Synge and Gardiner (5).

Although the extensibility of a rigid rod is compatible with special relativity theory, our discussion does carry the assumption that, *whenever a particular rigid rod is at rest in an inertial frame, then it has always the same 'length'*. More precisely, if two unstressed rigid rods initially at rest in such a frame have the same length as measured by each other, then they must always have the same length when they are again in this state, whatever different motions they may undergo during the interval. Otherwise a rigid rod could not be used in the usual way to define length in an inertial frame, and the theory of special relativity would have to be re-formulated.

It was shown in effect in (4) that the condition is satisfied when it was remarked that a relativistically rigid rod is also perfectly elastic.

The previous work was, however, incomplete in regard to this feature. The object of the present paper is to state the definition of rigidity more rigorously and to derive several theorems. In a subsequent note, the results will be applied to the problem of a rotating rigid 'hoop'.

2. *Definition of rigid rod.* We shall adopt:

DEFINITION A. *A rigid rod is one* (*a*) *along which any disturbance travels with speed* c *for all possible amplitudes of the accompanying strain and for all possible states of stress existing before the application of the disturbing forces,* (*b*) *in which any change of configuration is reversible and takes place without dissipation of energy.*

The reasons for adopting this definition are contained in § 1 and in (4). Of course, it is not obvious that the definition is self-consistent or consistent with the principles of special relativity. The results about to be proved may be regarded as supplying the needed 'existence theorems'. We proceed by deriving a necessary property and then proving several results which establish its sufficiency for our requirements.

3. *Equation of elasticity.* We may now assert:

THEOREM 1. *The stress in any element of a given rigid rod depends only upon the strain in that element.*

From (4) and from Definition A (a), it follows that the rod is extensible and that the extension does depend upon the stress. The only quantities other than the strain upon which the stress could depend are the time-derivatives of the strain. But, just as in classical mechanics, a dependence upon the latter quantities would make it possible for one to stretch and re-compress the rod at different rates in such a way as to bring it back to its initial configuration with a net gain of work. Such a possibility is disallowed by Definition A (b). This establishes the theorem.

When the rod is unstressed and at rest in an inertial frame, let L_0 be its length and m_0 be the mass per unit length. When it is at rest under tension T, let L be the length and m the mass per unit length. We prove

THEOREM 2.
$$T = \tfrac{1}{2}m_0c^2[1-(L_0/L)^2], \tag{3·1}$$
$$m = \tfrac{1}{2}m_0[1+(L_0/L)^2]. \tag{3·2}$$

If the rod is moving initially with uniform velocity V relative to an inertial frame, and if the leading end is suddenly stopped then ((4), equation (7·1)) the whole rod in due course comes momentarily to rest with length L_1, where

$$L_1 = L_0\left(\frac{1-V/c}{1+V/c}\right)^{\frac{1}{2}}. \tag{3·3}$$

According to Theorem 1, the strain-energy depends only upon the initial and final states, not upon the mode of passage from one to the other. So the energy is equal to the work that would be done in a quasi-static process producing the same change of length in the rod. This is

$$-\int_{L_0}^{L_1} T\,dL, \tag{3·4}$$

where T is the tension when the length is L and, by Theorem 1, T varies only with L.

The kinetic energy lost by the rod is, by a standard formula of special relativity,

$$m_0L_0c^2[(1-V^2/c^2)^{-\frac{1}{2}}-1].$$

Using (3·3), this may be expressed in the form

$$\tfrac{1}{2}m_0c^2(L_0-L_1)^2/L_1. \tag{3·5}$$

According to the definition, there is no dissipation of energy and therefore the energies (3·4), (3·5) are equal. Their equality yields the formula (3·1).

It is easily verified that the same result follows also from the corresponding calculation for the case where the rear end of the rod is stopped, instead of the leading end.

By the well-known relativistic relation between mass and energy, the mass of the rod is increased by the amount of its strain-energy divided by c^2. As we have seen, the strain-energy is given by (3·5). Hence the total mass per unit length in the strained state is m_1, say, where

$$m_1 = \frac{1}{L_1}\left[m_0 L_0 + \tfrac{1}{2}m_0 \frac{(L_0 - L_1)^2}{L_1}\right] = \tfrac{1}{2}m_0\left[1 + \left(\frac{L_0}{L_1}\right)^2\right].$$

This establishes the formula (3·2).

The result (3·1) was stated in (4), equation (9·9), but the derivation was not given. It is interesting that it should be similar in form to a formula for 'incompressibility' given by Clark (2) for a three-dimensional medium. The present derivation is, however, entirely different from Clark's; the interpretation of Clark's formula from the present viewpoint will be considered in subsequent work on the three-dimensional problem.

Equation (3·1) is now seen to be a necessary consequence of Definition A. We proceed to show that it is also a sufficient condition for the definition to be satisfied. That is to say, we shall now adopt

DEFINITION B. *A rigid rod is one for which formula* (3·1) *is valid.*

Then, from the ensuing results we shall infer the equivalence of the two definitions.

4. *Transmission of impulses: infinitesimal amplitudes.*

THEOREM 3. *Longitudinal vibrations and disturbances of infinitesimal amplitude are propagated in the unstressed rod with speed* c.

Writing $L = L_0 + \Delta L$, (3·1) gives

$$T = m_0 c^2 \Delta L/L_0 + O[(\Delta L/L_0)^2].$$

Hence for small ΔL the rod obeys Hooke's law with modulus λ, where

$$\lambda = m_0 c^2. \tag{4·1}$$

For vibrations with infinitesimal amplitude both the velocities of the particles of the rod and the associated stresses are also infinitesimal. In this case there is, therefore, no difference between relativistic and classical mechanics. Hence the wave-velocity has the classical value $(\lambda/m_0)^{\frac{1}{2}}$. With the value of λ in (4·1) this gives velocity c; there being consequently no dispersion, the group-velocity is also c.

THEOREM 4. *If the rod is stretched to any given tension* T_1, *it transmits longitudinal vibrations and disturbances of infinitesimal amplitude with speed* c.

The appropriate relativistic equation of motion is well known (3). For one-dimensional motion under no body-force it takes the form

$$\left(1-\frac{u^2}{c^2}\right)^{-1}\left(m-\frac{T}{c^2}\right)\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)-\frac{\partial T}{\partial x}-\frac{u}{c^2}\frac{\partial T}{\partial t}=0, \tag{4·2}$$

where T is the tension and u is the velocity of the material at position x. If the motion is infinitesimal, this reduces to

$$\left(m_1 - \frac{T_1}{c^2}\right)\frac{\partial u}{\partial t} - \frac{\partial T}{\partial x} = 0, \tag{4·3}$$

where T_1 is the given tension and m_1 the corresponding mass per unit length.

Let L_1 be the length of the whole rod when under tension T_1. Let the origin of x be taken at one end of the rod, and let x_0 be the distance of the particle at x from this end before stretching. Since the motion is infinitesimal, we have therefore

$$\frac{x}{x_0} = \frac{L_1}{L_0}, \quad \text{giving} \quad \frac{dx}{dx_0} = \frac{L_1}{L_0}. \tag{4·4}$$

Let ξ be the (infinitesimal) displacement at time t of the particle at x from its equilibrium position. The element $(x, x+dx)$ is then displaced to

$$(x+\xi, x+dx+\xi+(\partial\xi/\partial x)\,dx)$$

so that it becomes of length $(1+\partial\xi/dx)\,dx$, its unstretched length being dx_0. By (3·1) the tension in this element is accordingly

$$T = \tfrac{1}{2}m_0c^2\left\{1-\left[\frac{dx_0}{(1+\partial\xi/\partial x)\,dx}\right]^2\right\} = T_1+m_0c^2(L_0/L_1)^2\,\partial\xi/\partial x$$

to the first order in ξ, using (4·4). This gives

$$\frac{\partial T}{\partial x} = m_0c^2\left(\frac{L_0}{L_1}\right)^2\frac{\partial^2\xi}{\partial x^2}. \tag{4·5}$$

From (3·1) and (3·2) we obtain

$$m_1-T_1/c^2 = m_0(L_0/L_1)^2. \tag{4·6}$$

Noting that $u = \partial\xi/\partial t$ and using (4·5) and (4·6) the equation of motion (4·3) becomes

$$\partial^2\xi/\partial t^2-c^2\,\partial^2\xi/\partial x^2 = 0.$$

Since this is the equation of wave-propagation with speed c the theorem is proved.

Theorem 3 is, of course, the special case for $T_1 = 0$. It should be noted that the result of Theorem 4 does depend essentially upon the occurrence of the factor $(m-T/c^2)$ in (4·2) in place of the simple factor m that would occur in the classical equation of motion.

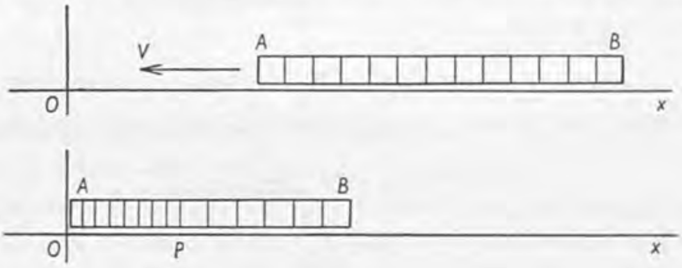
The remarkable property possessed by the formula (3·1) of yielding the speed of propagation c independently of the tension provides a further illustration of the fact that we have the closest possible analogue to the classical rigid rod. For we expect this analogue to yield the fundamental speed c *wherever* we should have infinite speed in the classical case. This is what we are now finding to hold good.

5. *Transmission of impulses: finite amplitudes.* We now consider the propagation along the rod of a disturbance of finite amplitude. Any such disturbance may be regarded as being generated by constraining one or both ends of the rod to move in a particular manner. We consider first the simplest case. The rod being initially at rest in an inertial frame, one end is constrained, from some particular instant, to move with given uniform speed V. This is the case of the rod being struck a 'sledge-hammer' blow at one end. It is physically equivalent to the case of the rod, initially moving with uniform velocity V in an inertial frame S, running into a stationary barrier. This is the form in which it is convenient to discuss the problem.

It may be remarked that there is nothing inconsistent with our assumptions in treating the barrier as unyielding. For the barrier could be composed of material for which the parameter m_0 is arbitrarily large. Alternatively, we could suppose that the rod collides with an exactly similar rod moving with speed V in the opposite direction. By symmetry, the only possibility is that the colliding ends are brought instantaneously to rest.

Let t, x be time and distance measured in frame S. Let the barrier be fixed at the origin O of the x-axis in S and let the rod AB be moving with uniform velocity $-V$ in such a way that the leading end A reaches O and is stopped there at time $t = 0$.

It is taken that the following process will ensue. At time $t = 0$ a disturbance starts travelling from O along the rod with some uniform speed v relative to S. It will reach a point P at distance $x\ (>0)$ at time $t = x/v$. At this instant the material in PB has not yet been 'informed' of the collision and is therefore still moving with its initial velocity $-V$. The material in OP is at rest and is in a state of uniform compression producing (negative) tension T. This material will remain in equilibrium under this tension until the disturbance reaches the end B of the rod. After that, the (negative) tension will push B in the positive x-direction and the process will be reversed. When the return impulse has reached the end A, the rod will be moving as a whole with velocity $+V$; that is to say, it will have rebounded as in a perfectly elastic collision. Here, however, we are concerned only with the first stage of the process.



This is merely a description of the classical process of elastic impact (with no dissipation). But there is nothing in it that does not carry over into special relativity theory. Of course, it assumes that there is no dispersion, i.e. that the impulse does not 'spread' as it traverses the rod. This assumption is justified by the fact that the law of elasticity (3·1) contains no parameter that could produce dispersion. It is verified by Theorems 3 and 4 in cases of infinitesimal disturbances.

We now treat the process quantitatively and prove

THEOREM 5. *The disturbance produced in stopping a moving free rigid rod traverses it with speed c.*

The particle of the rod at P at time $t = x/v$ has travelled a distance Vx/v after time $t = 0$. Hence, at $t = 0$ the portion of the rod OP had length $x(1+V/v)$ as measured in S. Taking account of the FitzGerald contraction corresponding to the fact that it was then moving with speed V relative to S, its proper length was $\beta x(1+V/v)$, where $\beta = (1-V^2/c^2)^{-\frac{1}{2}}$. Therefore, as each element of the rod is brought to rest, it becomes compressed in the ratio

$$1 : \beta(1+V/v). \tag{5·1}$$

Hence, using (3·1), the tension in the portion OP is

$$T = \tfrac{1}{2}m_0c^2[1-\beta^2(1+V/v)^2], \tag{5·2}$$

this being, of course, negative in such circumstances.

The element of the rod coming to rest in time dt is of length $v\,dt$ as measured in its compressed state in S. Using (5·1), the proper mass of this element before coming to rest was dm, where

$$dm = m_0\beta(1+V/v)\,v\,dt. \tag{5·3}$$

By the usual formula of special relativity, the momentum dp (in the negative x-direction) of this element relative to S was therefore

$$dp = \beta V\,dm = m_0\beta^2(v+V)\,V\,dt. \tag{5·4}$$

The fundamental equation of motion, relativistic as well as classical, now yields

$$dp = -T\,dt. \tag{5·5}$$

For this expresses that the momentum destroyed in time dt is equal to the impulse of the force $-T$ acting upon the element during this time.

Using (5·3) and (5·4) in (5·5) we find

$$2\beta^2(v+V)\,V = c^2\beta^2(1+V/v)^2 - c^2,$$

which reduces to

$$(2vV+V^2)(v^2-c^2) = 0.$$

Since V is arbitrary and v is positive this gives uniquely

$$v = c \tag{5·6}$$

and the theorem is proved.

We note also that on putting $v = c$ in (5·1) we recover the value of the compression-ratio in (3·3).

We have now established the converse of the main result of (4) by showing that the law of elasticity (3·1) gives the speed c for the propagation of the disturbance in the process there considered. In that paper it was postulated that the speed is c and it was stated that this required the truth of formula (3·1) as a result of energy considerations. But it was not there shown generally that this formula necessarily yields the required speed of propagation.

We wish next to establish a similar result in the case of a rod that is initially stressed. It will be seen below why this result is needed.

In order to see the meaning of the problem, we may imagine the given rod to be initially stretched (or compressed) between massive end-supports. One of these can then be supposed suddenly set in motion with uniform velocity V relative to the other. We are concerned with the disturbance that then traverses the rod. Or, we may think of the given rod as being pinned at its ends, in a stretched (or compressed) condition, between two other rods identical with each other, and the whole system can then replace the single rod considered in Theorem 5. (In the latter realization of the problem, all three rods must be stressed, and we can regard the discussion as applying to any one of them. Three, rather than two, rods are mentioned merely for the sake of better realism: the reader might say that two would 'bow'!)

As before, we may suppose the system to be moving initially with velocity $-V$ relative to an inertial frame S and suppose that the effect to be studied is achieved by allowing the system to collide with a suitable barrier. So far as it is applicable, we employ the same notation as before.

We shall require

LEMMA. *If the material at some point P of a given rod has proper-mass m_1, per unit proper length and tension T_1, and if its velocity relative to an inertial frame* S *is V, then the momentum relative to* S *per unit length at P measured in* S *is*

$$\beta^2(m_1 - T_1/c^2)\,V. \tag{5·7}$$

This result is, in fact, well known. For the expression (5·7) is the relevant component (usually written T^{14}) of the stress-momentum-energy tensor appropriate to this case (1, 6). The only departure from the standard statement of the result is in our giving it in terms of a line-density instead of a volume-density. This is justified by the remark in §1 that we may, if desired, consider a suitable three-dimensional body instead of the one-dimensional rod; quantities relating to unit volume of the body would then correspond to those relating to unit length of the rod.

Although we take the result (5·7) itself from the standard treatment, it is of interest to give a physical interpretation. In relativity theory, momentum can be treated as energy-flux. Suppose for the moment that the velocities concerned are small. Then, if stressed material is in motion, the energy-flux across unit area of any geometrical surface drawn through the material is composed of the material flux (giving a term of the form mV) together with the work done by the material on one side of the boundary upon that on the other (giving a term of the form $-TV/c^2$, the factor $1/c^2$ being required merely to give the correct mass-units). If the velocity is not small, one factor of the form β will appear in order to take account of the relativistic 'change of mass with velocity', and another factor of the same form to take account of the FitzGerald contraction, since unit length in S corresponds to proper-length β. This interpretation may be taken as accounting physically for the several factors in (5·7), but the actual derivation can be given satisfactorily only by means of the tensor treatment to which reference has been made.

It may also be noticed that the leading term in (4·3) is consistent with (5·7) for the case of a small velocity.

We are now in a position to prove

THEOREM 6. *The disturbance produced in stopping a moving stressed rigid rod traverses it with speed c.*

Let the initial tension be T_1. Then the initial proper-length L_1 is given by (3·1), i.e.

$$T_1 = \tfrac{1}{2}m_0c^2[1-(L_0/L_1)^2]. \tag{5·8}$$

The ratio $L_1 : L_0$ given by (5·8) is that by which every element of the rod is initially stretched (or compressed). Equation (5·1) now gives the further compression produced by the stopping process. Combining these factors, the tension in the portion OP of the rod is

$$T = \tfrac{1}{2}m_0c^2[1-\beta^2(1+V/v)^2(L_0/L_1)^2]. \tag{5·9}$$

Using (5·7) we have here in place of (5·4)

$$dp = \beta^2(m_1 - T_1/c^2)(v+V)\,V\,dt = m_0(L_0/L_1)^2\beta^2(v+V)\,V\,dt, \tag{5·10}$$

on using the value (5·8) for T_1 and the corresponding value from (3·2) for m_1. In place of (5·5) we must now write

$$dp = -(T-T_1)\,dt, \tag{5·11}$$

since $-(T-T_1)$ is here the resultant force acting upon the element of the rod concerned. From (5·8), (5·9), (5·11) we obtain

$$2(L_0/L_1)^2\beta^2(v+V)\,V = (L_0/L_1)^2[c^2\beta^2(1+V/v)^2-c^2].$$

Since the factor $(L_0/L_1)^2$ cancels out, we are then left with exactly the same equation as before to determine v. As we have seen, this gives $v = c$, thus proving the theorem.

In regard to the proofs of Theorems 5 and 6 it should be remarked that v as defined was the speed of advance of the disturbance relative to the disturbed portion of the rod. Had the rod not been 'rigid', we should have found for v some value less than c. This would not then have been also the speed of advance relative to the undisturbed portion of the rod, though the latter could, of course, have then been evaluated using the standard Lorentz-velocity transformation. However, since in our case the velocity of advance has been shown to be the fundamental velocity c, it has this value whether it be taken relative to the disturbed or undisturbed material.

We may now formulate as the general conclusion of the work:

THEOREM 7. *Any disturbance of any amplitude is propagated in the rod with speed* c.

We shall now think of the rod as being initially at rest in an inertial frame.

Any disturbance generated at one end of the rod and traversing it in one direction is determined by the motion produced at that end. In Theorems 5 and 6 we have dealt with the case where that motion is uniform relative to the undisturbed part of the rod, and we have found that the resulting disturbance is propagated with speed c. As we have seen, it produces a uniform stress in the disturbed material. If after a certain time, a different uniform velocity is imparted to the end of the rod, its effect will therefore be propagated through the material already disturbed by the first motion again with speed c. This follows by application of Theorem 6. We can proceed in this manner for any discrete set of changes of velocity of the end of the rod. We can then consider the given motion of the end as the limit of such a set, and we conclude that the whole disturbance traverses the rod with speed c.

The admissibility of this argument here should be contrasted with its failure in the case of a classically compressible rod in which the speed of propagation does depend upon the stress. In the latter case, a single change of velocity of one end is propagated through the rod with some particular speed and produces a stress in the disturbed region. A second ensuing change of velocity must be propagated through this stressed region. But, owing to the stress, the speed is *different* for this second disturbance, and so on for any subsequent ones. Then the propagation of a general disturbance becomes a much more complicated matter than it is in our case, and no simple general conclusion can be reached. This remark applies with even more force to the case of a relativistic non-rigid rod.

REFERENCES

(1) BERGMANN, P. C. *Introduction to the theory of relativity* (New York, 1942), p. 128, equation (8.23).
(2) CLARK, G. L. *Proc. Camb. phil. Soc.* 45 (1949), 405–10.
(3) McCREA, W. H. *Relativity physics*, 3rd ed. (London, 1949), p. 82, equation (206).
(4) McCREA, W. H. *Sci. Proc. R. Dublin Soc.* (N.S.), 26 (1952), 27–36.
(5) SYNGE, J. L. *Sci. Proc. R. Dublin Soc.* (N.S.), 26 (1952), 45–54.
(6) TOLMAN, R. C. *Relativity, thermodynamics and cosmology* (Oxford, 1934), p. 72, equation (37.7).

ROYAL HOLLOWAY COLLEGE
ENGLEFIELD GREEN, SURREY