Brandeis
Summer Institute
in
Theoretical Physics

1964
VOLUME ONE

lectures on GENERAL RELATIVITY

A. TRAUTMANN / F. A. E. PIRANI / H. BONDI

Trautman

lectures on GENERAL RELATIVITY

Brandeis
Summer Institute
in
Theoretical Physics

1964
VOLUME ONE

lectures on GENERAL RELATIVITY

A. TRAUTMAN / F. A. E. PIRANI / H. BONDI

G. Dixon
P. Florides
G. Lemmer

J. Marek

J. Hyde

Prentice-Hall, Inc. Englewood Cliffs, New Jersey

Typing by Judith Gray and Evangeline Goodwin

Printed in The United States of America
Library of Congress Catalog Card Number 65—16531

C—52716

PRENTICE-HALL, INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY., LTD., *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA (PRIVATE) LTD., *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*

FOREWORD

The notes contained in this volume are based on the 1964 Brandeis summer lectures of three specialists in relativity theory. To a greater extent than is common in summer school proceedings, the different contributions fit together to make a coherent volume. Various cross-references will be found in the notes.

Professor Trautman and Professor Bondi have not had an opportunity to proofread the final version of their lecture notes. We are grateful to all three lecturers for their assistance with the initial preparation of the notes. Especially we wish to acknowledge the valuable help of Professor Pirani who contributed to all phases of the editorial work during his stay at Brandeis in the fall of 1964. Editorial assistance was also provided by Miss Christine Denny.

The final copy for this volume was prepared at Brandeis with non-professional assistance. Prentice-Hall is not to be held accountable for the quality of the equations and figures. On the contrary, their editor James Walsh is to be commended for agreeing to publish this volume with speed and at moderate cost.

The 1964 Brandeis Summer Institute in Theoretical Physics was made possible by the generous support of the National Science Foundation and of NATO.

This is the first of two volumes of notes based on the 1964 Institute. The contents of Volume 2 will be found on the next page.

Stanley Deser
Kenneth W. Ford
Editors

Contents of Volume 2

K. Johnson
Quantum Electrodynamics

D. Lichtenberg
Spectroscopy of Strongly Interacting Particles

J. Schwinger
A Field Theory of Particles

S. Weinberg
Composite Particles and Perturbation Theory

S. Weinberg
Quantum Theory of Massless Particles

CONTENTS

VOLUME 1

FOREWORD v

A. TRAUTMAN

FOUNDATIONS AND CURRENT PROBLEMS OF GENERAL RELATIVITY 1

Notes by Graham Dixon, Petros Florides, and Gerald Lemmer

F.A.E. PIRANI

INTRODUCTION TO GRAVITATIONAL RADIATION THEORY 249

Notes by J.J.J. Marek and the Lecturer

H. BONDI

SOME SPECIAL SOLUTIONS OF THE EINSTEIN EQUATIONS 375

Notes by John Hyde

FOUNDATIONS AND CURRENT PROBLEMS OF GENERAL RELATIVITY

ANDRZEJ TRAUTMAN

Polish Academy of Sciences, Warsaw

Notes by Graham Dixon, Petros Florides, and Gerald Lemmer

A. Trautman

CONTENTS

1. GENERAL DISCUSSION 7

1.1 Introduction 7

1.2 Gravitational and Inertial Masses 7

1.3 First Order Corrections to Particle Motion 10

1.4 Gravitational Radiation 13

1.5 Quantum Effects Connected with Gravitation 19

2. VECTOR AND TENSOR ALGEBRA 25

2.1 Introduction 25

2.2 Algebraic Structures 25

2.3 Equivalence Relations and Morphisms 29

2.4 Vector Spaces 30

2.5 The Dual Space 33

2.6 Linear Subspaces 36

2.7 Tensor Product of Vector Spaces 36

2.8 Multiforms and Multivectors 40

2.9 Orientation of Vector Spaces and Subspaces 43

2.10 Densities 46

2.11 Complexification of a Real Vector Space 47

2.12 Euclidean Vector Spaces 48

3. MINKOWSKI SPACE AND LORENTZ TRANSFORMATIONS 54

3.1 Minkowski Vector Spaces 54

3.2 Lorentz Transformations 56

3.3 Null Tetrad and the Special Linear Group SL(2,C). 57

3.4 Lorentz Transformations and Bilinear Transformations of the Complex Plane 61

3.5 Spinors 63

4. DIFFERENTIAL GEOMETRY 67

4.1 Topological Spaces 67

4.2 Products of Topological Spaces 68

4.3 Continuous Functions and Homeomorphisms 69

4.4 Differentiable Manifolds 69

4.5 Coordinate Transformations 71

4.6 Differentiable Mappings 73

4.7 Products of Differentiable Manifolds 73

4.8 Tangent Vectors 74

4.9 The Dragging Along of a Vector by a Differentiable Mapping 78

4.10 Vector Fields 79

4.11 Differential Forms 80

4.12 Tangent Tensors and Tensor Fields 83

4.13 Geometric Objects 84

4.14 One-Parameter Groups of Diffeomorphisms 87

4.15 Lie Derivatives 89

4.16 Affine Connections on a Differentiable Manifold 89

4.17 The Torsion Tensor 94

4.18 The Curvature Tensor 95

4.19 Metric Tensor Fields. Riemannian Geometry 96

4.20 Integration in a Differentiable Manifold 98

5. THEORIES OF SPACE, TIME, AND GRAVITATION 101

5.1 Space-Time as a Differentiable Manifold 101

5.2 The Affine Connection in Physics 103

5.3 Newtonian Mechanics in the Absence of Gravitation 104

5.4 Newtonian Mechanics in the Presence of Gravitation 105

5.5 Axiomatic Formulation of the Geometry of Newtonian Gravitation Theory 106

5.6 The Interpretation of g^{ab}. Privileged Classes of Coordinate Systems 115

5.7 Relativistic Theories of Space, Time, and Gravitation 118

6. FUNDAMENTAL PROPERTIES OF GENERAL RELATIVITY 122

6.1 The Principle of General Covariance 122

6.2 The Principle of Equivalence 123

6.3 Correspondence with Newtonian Theory 127

6.4 Stationary and Static Gravitational Fields 130

6.5 Propagation of Light in a Gravitational Field 134

6.6 Local Reference Frames and Fermi Transport 137

6.7 The Physical Distinction between Static and Stationary Space-Times 143

6.8 Evaluation of the Curvature Tensor at the Earth's Surface 145

6.9 The Gravitational Field Equations 148

6.10 The Perihelion Precession 151

7. INVARIANCE PROPERTIES AND CONSERVATION LAWS 158

7.1 Lie Groups 158

7.2 Lie Groups of Transformations 162

7.3 Infinitesimal Coordinate Transformations 164

7.4 The Variational Principle 165

7.5 Conservation Laws in Riemannian Space-Time 176

7.6 Conservation Laws in General Relativity 182

7.7 Relation between Field Equations and Equations of Motion

7.8 The Motion of Particles in a Given Field 192

7.9 The Canonical Formalism for Relativistic Particles 196

8. RIGID MOTION IN RELATIVITY THEORY 201

8.1 Introduction 201

8.2 Mathematical Formulation of the Criteria 201

8.3 Integration of the Isometry Condition in M_4 207

8.4 The Isometry Condition in V_4 213

8.5 The Born Condition in V_4 215

8.6 The Herglotz-Noether Theorem 222

8.7 Dynamics 224

8.8 Integrability Conditions 226

9. COSMOLOGY 228

9.1 Introduction 228

9.2 Newtonian Cosmology 230

9.3 Relativistic Cosmology 238

9.4 Propagation of Light in Relativistic Cosmology 240

9.5 Distance in Cosmology

9.6 The Event Horizon of the de Sitter Universe 244

9.7 Epilogue 248

1. GENERAL DISCUSSION

1.1. INTRODUCTION

One of the many unsolved problems connected with the general theory of relativity is whether the theory belongs to physics or rather to mathematics. One of my colleagues at this Summer School said that those who work in the theory of relativity do so because of its mathematical beauty rather than because they want to make predictions which could be checked against experiment. I think there is some truth in this statement, and probably I am no exception to it.

Before I indulge in my favorite formalism I should like to give you a simple-minded analysis of the orders of magnitude of possible relativistic and quantum-mechanical effects of gravitation.

I shall pretend that we don't know anything about general relativity and shall try to discover, by approximate analysis and using rough arguments, what can be said about the magnitude of the corrections to Newtonian gravitation theory necessitated by relativity and quantum theory.

We shall assume the following:

(1) Newton's law of gravitational attraction.

(2) That in any given gravitational field, when non-gravitational forces may be neglected, the motion of a body depends only on its initial position and velocity and is independent of the constitution of the body.

(3) In the absence of gravitation the theory of special relativity is valid.

1.2. GRAVITATIONAL AND INERTIAL MASSES

Assumption 1 above implies that the gravitational force acting on a body is expressible in terms of a scalar potential ϕ given by

$$\phi(\vec{x}) = - \sum_{\text{bodies}} \frac{\mu_i^{\ A}}{|\vec{x}-\vec{x}_i|}, \tag{1.1}$$

where i is an index labeling the bodies, $\vec{x}_i$ is the position vector of the ith body, and $\mu_i^{\ A}$ is a constant associated with the ith body and called its active gravitational mass. The gravitational force $\vec{F}$ acting on any body is then given in terms of a constant m^P associated with that body and called its passive gravitational mass by

$$\vec{F} = -m^P \operatorname{grad} \phi. \tag{1.2}$$

To obtain the motion of the body, this value of $\vec{F}$ is to be substituted into the equation of motion

$$m^I \frac{d^2\vec{r}}{dt^2} = \vec{F}, \tag{1.3}$$

where m^I is another constant associated with the body and called its inertial mass, and $\vec{r}$ is the position vector of the body.

At this point one may wish for operational definitions of the constants μ^A, m^P, m^I that we have introduced. We may give them as follows:

(1) The inertial masses of two bodies, 1 and 2, are compared by joining them by a standard spring, pulling them apart so that the spring is taut, and releasing them simultaneously. The accelerations $\vec{f}_1$ and $\vec{f}_2$ of both bodies are measured, and, since by Newton's Third Law of Motion the spring exerts equal and opposite forces on both bodies, we have from (1.3)

$$m_1^{\ I}\, \vec{f}_1 = -m_2^{\ I}\, \vec{f}_2.$$

In this way any body can be compared with a standard body defined as having unit inertial mass. The consistency of this definition (i.e. if we compare bodies 1 and 2, and then 2 and 3, we get the same mass ratio as if we compared 1 and 3 directly) is an empirical fact that substantiates Newton's Third Law.

(2) If, by means of suitable shielding, all non-gravitational forces are removed, then the motion of a body in the gravitational field of a given mass distribution is given, from Assumption 2, by

$$m^I \frac{d^2\vec{r}}{dt^2} = m^P\vec{g}, \qquad (1.4)$$

where $\vec{g}(\vec{r},t)$ is the gravitational vector field. The constants m^P and the field $\vec{g}$ are determined by the experiment only up to transformations

$$m^P \to \frac{m^P}{\alpha}, \quad \vec{g} \to \alpha\vec{g}, \quad \alpha = \text{const.}$$

Moreover, from Assumption 2 above it follows that the ratio m^I/m^P is the same for all bodies. This fact is said to have been noticed and verified by Galileo. Eötvös[1] and more recently Dicke[2] have checked it with an accuracy of at least one part in 10^{10}. By an appropriate choice of α, this ratio can always be made unity, so that we can take

$$m_I = m_P = m, \text{ say,}$$

for all bodies.

(3) Newton's law of attraction tells us that the gravitational field $\vec{g}$ derives from a potential ϕ given by (1.1). By choosing $m^I = m^P$ we fix $\vec{g}$ and, therefore, the constants μ appearing in ϕ.

Now consider two bodies moving in each other's gravitational field. Then the forces acting on the two bodies are:

$$\vec{F}_1 = -\frac{m_1\mu_2^{\ A}}{r^3}\,\vec{r}, \qquad \vec{F}_2 = +\frac{m_2\mu_1^{\ A}}{r^3}\,\vec{r}, \qquad (1.5)$$

---

1. R. V. Eötvös, D. Pekar, and E. Fekete, Ann. der Physik 68, 11 (1922).
2. R. H. Dicke, P.G. Roll, and R. Krotkov, Ann. Phys.(N.Y.) 26, 442 (1964).

where $\vec{F}_i$ is the force on the i-th body and $\vec{r}$ is the position vector of body 1 relative to body 2. By Newton's Third Law of Motion these forces are equal in magnitude and opposite in direction, $\vec{F}_1 = -\vec{F}_2$, so that (1.5) gives

$$m_1 \mu_2{}^A = m_2 \mu_1{}^A, \quad \text{i.e.} \quad \frac{\mu_1{}^A}{m_1} = \frac{\mu_2{}^A}{m_2} .$$

The ratio μ^A/m is thus independent of the body and is a universal constant, k say, called the Newtonian constant of gravitation. We thus have

$$\mu^A = km$$

and so have reduced the three masses originally defined to only one independent one, simply called the mass m of the body.

1.3. FIRST ORDER CORRECTIONS TO PARTICLE MOTION

Let us now consider possible first order corrections to the Newtonian equations of particle motion in the gravitational field of a single massive body that might be given by a relativistic theory of gravitation. We shall try to find a Lagrangian that will include such effects. This does not necessarily mean that the exact relativistic theory must be expressible in Lagrangian form, however, since it is often possible to describe small corrections to a theory approximately by means of an equivalent Lagrangian even when the exact corrections cannot be so expressed.

We shall assume that the body producing the field can be described with sufficient accuracy for our purposes by giving its mass M, velocity $\vec{V}$, angular momentum $\vec{J}$ and charge Q. Out of these, the universal constants k (gravitational constant) and c (velocity of light), and the position vector $\vec{r}$ of the particle, relative to the body producing the field, one can form the dimensionless quantities

$$\frac{km}{c^2 r}, \quad \frac{k\vec{J}}{c^3 r^2}, \quad \frac{kQ^2}{c^4 r^2}.$$

Now using the fact that a Lagrangian must have the dimensions of energy, we can hypothesize that the particle can be described to first order relativistic corrections by a Lagrangian L of form

$$\frac{L}{m} = \frac{1}{2}v^2 + \frac{kM}{r} + \frac{1}{8}\frac{v^4}{c^2} + \alpha\left(\frac{kM}{cr}\right)^2 + \beta\frac{kM}{r}\frac{v^2}{c^2} + \tag{1.6}$$
$$+ \gamma\frac{kM\vec{V}}{c^2 r}\cdot\frac{d\vec{r}}{dt} + \delta\frac{k\vec{J}x\vec{r}}{c^2 r^3}\cdot\frac{d\vec{r}}{dt} + \varepsilon\frac{kQ^2}{c^2 r} + \ldots$$

where m is the mass of the particle, $\vec{v} = d\vec{r}/dt$, and α, β, γ, δ, ε, are dimensionless constants. Of these terms, the first is the Newtonian kinetic energy of the particle, the second gives the Newtonian gravitational field, and the third is the first order correction for the increase of mass with velocity, obtained by expanding the Lagrangian

$$L = -mc^2\sqrt{1-\frac{v^2}{c^2}}$$

for a free particle in special relativity. The term containing the angular momentum has to have the form of a triple scalar product, since angular momentum is an axial vector and has to be combined with other quantities in such a way as to give a scalar term rather than a pseudoscalar one. A term of the form $\zeta\frac{\sqrt{k}Q}{r}$ could have been added, as it would have had the correct dimensions, but unless the coefficient ζ were exceedingly small, it would violate the equality of inertial and active gravitational mass. For an electron

$$\frac{\frac{\sqrt{k}Q}{r}}{\frac{kM}{r}} \sim 10^{21}$$

and so for any reasonable coefficient ζ the gravitational field due to its charge $\sqrt{k}Q/r$ would be far larger than that due to its mass, kM/r. On the grounds that such extremely small dimensionless coefficients do not occur in reasonable physical theories, we shall not include such a term.

Of the correction terms we have hypothesized, the only ones for which there are any observational evidence are those occurring in the first row of (1.6). Observations of the precession of the perihelion of planetary orbits, after taking into account all the perturbations due to the other planets, indicate a small residual precession which can be explained by taking $\alpha \neq 0$. All the other correction terms give effects so slight, for values of the coefficients of the order of unity, that they would have no observable effects on planetary orbits. The term $(kM/cr)^2$ will become appreciable compared with the Newtonian term kM/r if

$$\frac{kM}{c^2 r} \sim 1.$$

On an atomic scale, we have, for example

$$\frac{km}{c^2 r} = \frac{km^2}{e^2} \sim 10^{-42}$$

on the surface of an electron, taking for r the classical radius of the electron, $r = e^2/mc^2$. So clearly it is not on the atomic scale we must look. We must take M as large as possible and r as small as possible for that M. Expressing M in terms of the density ρ, for a sphere of radius r, we have $M = (4/3)\pi\rho r^3$, and then

$$\frac{kM}{c^2 r} \sim \frac{k\rho r^2}{c^2}.$$

So alternatively we can say that we want the largest possible size for a given density. This suggests that we look to astrophysics. Such a term would play an appreciable part in the hypothesized neutron stars. Alternatively we can look on an even larger scale, to cosmology. A typical length scale for cosmology can be taken to be

$$r = cT$$

where $1/T$ is Hubble's constant, and $T \sim 10^{10}$ years. Then we have

$$\frac{k\rho r^2}{c^2} \sim k\rho T^2 \sim 1$$

on substituting in numerical estimates for ρ, which here can be taken as the mean density of matter in the universe. So we expect relativistic gravitational effects to play an important part in cosmology.

If we evaluate the effect of the other corrections in (1.6) on the motion of the planets, we find that some of those additional terms give rise to a precession of the perihelion , and the term with coefficient δ causes a precession of the plane of motion with an angular velocity of precession proportional to $\vec{J}/r^3$. However, as mentioned above, none of these terms, apart from the one with coefficient α, would cause an observable effect if they occurred with coefficients of the order of unity.

1.4. GRAVITATIONAL RADIATION

In a similar way, using rough arguments, we shall now discuss the magnitude of effects connected with gravitational radiation.

In Newtonian theory the gravitational potential ϕ satisfies Poisson's equation

$$\nabla^2\phi = 4\pi k\rho, \tag{1.7}$$

where ρ is the mass density. We do not want to use any particular relativistic theory of gravitation but we can expect that if there are field equations for the gravitational field in a relativistic theory, some of the components of the field will satisfy either the natural relativistic generalization of (1.7), namely

$$\Box\phi = 4\pi k\rho, \tag{1.8}$$

where $\Box \overset{\text{def}}{=} \nabla^2 - \partial^2/\partial t^2$ is the D'Alembertian operator, or else some similar equation. A solution of (1.8) is given by the retarded potential

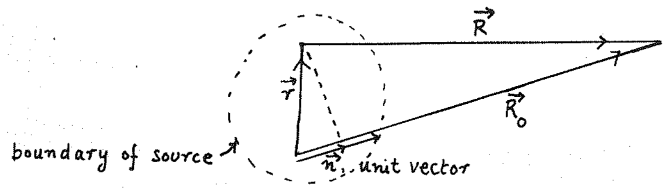
$$\phi(\vec{R}_o,t) = -k\int \frac{\rho(\vec{r},\ t-\frac{R}{c})}{R}\,dV \tag{1.9}$$

where $\vec{R} = \vec{R}_o-\vec{r}$, $R = |\vec{R}|$, $\vec{r}$ is the variable of integration, and the integral is over the 3-space t = const. The general solution of (1.8) consists of a mixture of advanced and retarded potentials plus any solution of the free field equations

$$\Box\phi = 0,$$

but on physical grounds we reject the advanced potential solution and we are not interested in the free field part of the solution. So we shall only consider the solutions given by (1.9).

Consider the field in the wave zone, i.e. at distances from the source, (assumed to be of finite extent), which are large compared with the dimensions of the system and also compared with the wavelength of the radiation. Let $\vec{n}$ be a



unit vector in the direction of $\vec{R}_o$, and choose an origin of coordinates inside the source. Then we see from the diagram that at large R,

$$R \simeq R_o - \vec{n}\cdot\vec{r}. \qquad (1.10)$$

We thus have

$$\rho(\vec{r},\ t-\frac{R}{c}) \simeq \rho(\vec{r},\ t-\frac{R_o}{c} + \frac{\vec{n}\cdot\vec{r}}{c}).$$

and we expand this in a Taylor series in powers of $\vec{n}\cdot\vec{r}/c$ thus:-

$$\rho(\vec{r},\ t-\frac{R}{c}) \simeq \rho(\vec{r},\ t-\frac{R_o}{c}) + \frac{\vec{n}\cdot\vec{r}}{c}\ \dot{\rho}(\vec{r},\ t-\frac{R_o}{c})$$
$$+ \frac{1}{2}\left(\frac{\vec{n}\cdot\vec{r}}{c}\right)^2 \ddot{\rho}(\vec{r},\ t-\frac{R_o}{c}) + \dots \qquad (1.11)$$

where $\dot{}$ denotes differentiation with respect to t. Substitute this expansion into (1.9), remembering that it is valid in the wave-zone for all $\vec{r}$ inside the source, and that the integrand of (1.9) and all terms in (1.11) vanish for $\vec{r}$ outside the source. So we get

$$\phi(\vec{R}_o,t) \simeq -\frac{k}{R_o}\int \rho(\vec{r},\ t-\frac{R_o}{c})\ dV - \frac{k}{R_o}\int \dot{\rho}(\vec{r},\ t-\frac{R_o}{c})\ \frac{\vec{n}\cdot\vec{r}}{c}\ dV$$

$$-\frac{k}{R_o}\int \ddot{\rho}(\vec{r},\ t-\frac{R_o}{c})\ (\frac{\vec{n}\cdot\vec{r}}{c})^2\ dV - \ldots \tag{1.12}$$

where we have, with little error, replaced R by R_o in the denominator of the integrand of (1.9). This is the multipole expansion of the potential in the wave zone that is familiar from electrodynamics. The terms are successively called the monopole, dipole, quadrupole, ... terms.

We now invoke the laws of conservation of mass and momentum to show that, in the absence of an external non-gravitational field causing the momentum of the matter to change, there can be no monopole or dipole radiation. For if M is the total mass of the source, $\vec{P}$ its total momentum, and $\vec{v}$ the velocity field of the matter, then

$$\int \rho(\vec{r},\ t-\frac{R_o}{c})\ dV = M,$$

and from the conservation equation

$$\dot{\rho} + \mathrm{div}\ (\rho\vec{v}) = 0$$

we have

$$\int \dot{\rho}\ \vec{r}\ dV = -\int \mathrm{div}\ (\rho\vec{V})\ \vec{r}\ dV$$

$$= \int \rho\ \vec{V}\ dV + \left\{\begin{array}{l}\text{surface integral which}\\ \text{vanishes as } \rho = 0 \text{ on the}\\ \text{boundary of volume of in-}\\ \text{tegration.}\end{array}\right.$$

$$= \vec{P}.$$

The contributions to ϕ from the monopole and dipole terms are thus

$$\phi(\vec{R}_o,t) = -\frac{kM}{R_o} - \frac{k\,\vec{n}\cdot\vec{P}}{cR_o} + \text{quadrupole and higher terms.} \tag{1.13}$$

But by analogy with electromagnetic theory we expect the radiated power P to be given by an expression of the form of a surface integral of a quantity quadratic in the first derivatives of the field variables, e.g.

$$P = \frac{c}{k} \oint_S (\nabla\phi)^2 \, dS, \tag{1.14}$$

where S is a surface at large distance from the source, and c/k is a factor necessary to give P the correct dimensions. Consequently, the only terms in ϕ that will contribute are those for which $\nabla\phi = \mathcal{O}(1/R)$, and from (1.13), by the constancy of M and $\vec{P}$, for the monopole and dipole terms $\nabla\phi = O(1/R^2)$. Thus we expect gravitational radiation to be predominantly quadrupole. We notice here the analogy with electromagnetism, where monopole radiation does not occur because of charge conservation, and dipole radiation does not occur for a system composed of charges all with the same ratio of e/m.

$\frac{k}{c^5}\dddot{D}^2$

We shall now use (1.12) and the expression (1.14) for the radiated power to estimate the intensity of gravitational radiation from a system of two equal masses moving about one another in circular orbits. Let r be the radius of the orbits, m the mass of the bodies and ω the angular velocity of the bodies in the orbit. Then Newtonian mechanics gives us

$$\frac{km^2}{4r^2} = m\omega^2 r. \tag{1.15}$$

The quadrupole contribution to the potential in the wave zone will then be

$$\phi \sim \frac{k}{R_o} \frac{m\omega^2 r^2}{c^2}$$

so that

$$\nabla\phi \sim \frac{km\omega^3 r^2}{R_o c^3}.$$

Putting this into (1.14) gives

$$P \sim \frac{c}{k}\left(\frac{km\omega^3 r^2}{c^3}\right)^2 \tag{1.16}$$

or, using (1.15),

$$P \sim \frac{mc^2}{r/c}\left(\frac{km}{c^2 r}\right)^4. \tag{1.17}$$

Let us compare this with the analogous electromagnetic case, of two bodies of equal mass m and opposite charges e and -e moving in circular orbits of radius r about one another with a velocity much less than c, so that we can use a non-relativistic approximation. Then the Newtonian equation of motion (1.15) is now replaced by

$$\frac{e^2}{4r^2} = m\omega^2 r, \tag{1.18}$$

as gravitational forces are assumed negligible compared with the electrostatic Coulomb force. The power radiated by electromagnetic (dipole) radiation is

$$P_{e.m.} \sim \frac{mc^2}{r/c}\left(\frac{e^2}{mc^2 r}\right)^3 \tag{1.19}$$

and by substituting (1.18) into (1.16) we get the power radiated by the same system as gravitational (quadrupole) radiation to be

$$P_{grav} \sim \frac{mc^2}{r/c}\left(\frac{e^2}{mc^2 r}\right)^3\left(\frac{km}{c^2 r}\right). \tag{1.20}$$

For motion of particles on an atomic scale, the additional factor in (1.20) compared with (1.19) is

$$\frac{km}{c^2 r} \sim 10^{-47},$$

and so we see that gravitational radiation from atoms must be very small indeed. On an astronomical scale, too, the effect of gravitational radiation is negligible. For example, the gravitational radiation for the Jupiter-Sun system, calculated by a formula similar to (1.17) but for two unequal masses, is about 450 watts, a negligible quantity compared with the total energy of the system.

It may be possible to obtain larger effects from gravitational radiation if external forces are present which destroy the conservation of mechanical momentum: for example, a charged particle moving in a magnetic field. This is not certain, however, since we know that if we take into account the momentum of the external field, total momentum is conserved. This momentum is not localized; it is distributed throughout space, but so is the energy of the field, which acts as a source of the gravitational field. In this case the multipole expansion mode above is not valid, and the situation becomes much more complicated. A discussion of this case has been given by Postovoit and Gercenstein.[3] If we neglect these complicating factors we can estimate the expected amount of gravitational dipole radiation, if any occurs at all, from a particle of mass m and charge e moving in a circular orbit of radius r in a uniform magnetic field of strength H. Using our multipole expansion in exactly the same way as before, we have the equation of motion

$$\frac{Hev}{c} = m\omega^2 r,$$

and we obtain for the power of the gravitational and electromagnetic dipole radiation respectively

$$P_{grav} \sim \frac{mc^2}{\frac{r}{c}} \left(\frac{v}{c}\right)^4 \left(\frac{km}{c^2 r}\right)$$

$$P_{e.m.} \sim \frac{mc^2}{\frac{r}{c}} \left(\frac{v}{c}\right)^4 \left(\frac{e^2}{mc^2 r}\right)$$

in the non-relativistic limit $v \ll c$, where v is the velocity of the particle. Again we see that on an atomic scale the gravitational radiation is negligible when compared with the electromagnetic radiation, e.g. for an electron,

3. V.I. Postovoit and M.E. Gercenstein, J.E.T.P. 42, 163 (1962). (In Russian.)

$$\frac{P_{grav}}{P_{e.m.}} \sim \frac{km^2}{e^2} \sim 10^{-42}.$$

1.5. QUANTUM EFFECTS CONNECTED WITH GRAVITATION

We shall next discuss some of the quantum effects connected with gravitation. But before doing so we give a table of typical lengths which will be useful in discussing orders of magnitude (see p.20). Here k is the gravitational constant, m the mass of the electron, e the electronic charge, $\hbar = h/2\pi$ where h is Planck's constant, and c is the velocity of light. The coupling constant f occurs in the interaction Lagrangian of weak interactions through its square $(f^2\bar{\psi}\psi\bar{\psi}\psi)$.

It is interesting to note that the classical radius of the electron, which is a characteristic length of atomic dimensions, is roughly the geometric mean of the gravitational radius of the electron, which characterizes the distances at which gravitational effects become important in elementary particle physics, and the radius of the Universe, which characterizes cosmology and is also connected with gravitational phenomena.

The ratio

$$\frac{r_g}{r_e} = \frac{\left(\frac{km}{c^2}\right)}{\left(\frac{e^2}{mc^2}\right)} \sim 0.25 \times 10^{-42}$$

characterizes the magnitude of gravitational forces compared with electromagnetic forces on an atomic scale, and we saw it occur in this context in the previous section. The ratio

$$\frac{\ell}{\lambda} = \frac{\sqrt{\frac{k\hbar}{c^3}}}{\left(\frac{\hbar}{mc}\right)} \sim 1.7 \times 10^{-23}$$

characterizes for an electron the quantum mechanical effects due to gravitational forces. Of course, instead of considering an electron we can replace m and e by the mass and charge of any other particle to get the corresponding values for the particle, but for all elementary particles these numbers will

<u>Table of a few typical lengths occurring in physics (given in cm).</u>

Gravitational radius of electron, r_g	ℓ	Weak interactions coupling constant f	Classical radius of electron, r_c	Compton wavelength of electron, λ	'Radius' of Universe
$r_g = \frac{km}{c^2}$	$\ell = \sqrt{\frac{kh}{c^3}}$		$r_c = \frac{e^2}{mc^2}$	$\lambda = \frac{\hbar}{mc}$	$\frac{\text{vel. of light}}{\text{Hubble's constant}}$
0.7×10^{-55}	1.6×10^{-33}	0.7×10^{-16}	2.8×10^{-13}	2.4×10^{-10}	$\sim 10^{28}$

$\longleftarrow$ 42 orders of magnitude $\longrightarrow$ $\longleftarrow$ 41 orders of magnitude $\longrightarrow$

have the same order of magnitude.

Let us now compare the hydrogen atom with a gravitationally bound atom consisting of two neutrons in a bound orbit, and calculate for both the ground state energy E_{ground} and the radius of the orbit in the ground state, a_{ground}, using the old Bohr quantization principles. Then we get

For the hydrogen atom.

$$E_{\text{ground}} = -\frac{1}{2} mc^2\alpha^2 \sim 13.6 \text{ eV}$$

where $\alpha = e^2/\hbar c \approx 1/137$ is the fine structure constant and m is the electron mass.

$$a_{\text{ground}} = \frac{e^2}{mc^2}\frac{1}{\alpha^2} \sim 0.5 \times 10^{-8} \text{cm}$$

For the gravitational atom.

$$E_{\text{ground}} = -\frac{1}{4} Mc^2 \left(\frac{\ell}{\Lambda}\right)^4 \sim 10^{-77} \text{eV}$$

where M is the neutron mass and $\Lambda = \hbar/Mc$.

$$a_{\text{ground}} = \frac{kM}{c^2}\left(\frac{\Lambda}{\ell}\right)^4 \sim 10^{28} \text{ cm.}$$

We notice the odd coincidence that the ground state of the gravitational atom has a radius of the order of the radius of the universe! radius of Universe $\sim \Lambda^3/\ell^2$

Nonrelativistic quantum effects of gravitation may be expected to play a significant rôle for a system for which

$$a_{\text{ground}} = \Lambda\left(\frac{\Lambda}{\ell}\right)^2 \sim \Lambda$$

and where classical dimensions are not much larger than this. But $\Lambda = \ell$ corresponds to a system of mass

$$M \sim 1.5 \times 10^{23} \text{ electron masses.}$$

This could only be realized macroscopically, in which case, however, quantum mechanics would play no rôle. Consequently one can doubt whether the gravitational potential can be meaningfully introduced into the nonrelativistic Schrödinger equation.

The question arises as to whether there are significant relativistic quantum effects of gravitation and whether the gravitational field may be "quantized" by a procedure similar to that applied in electrodynamics. It should not be taken for granted that this is so. For example, we know that in a certain sense statistical mechanics can be derived from the laws governing elementary systems, and that this can be done on both the classical and the quantum level. For simple mechanical systems the quantum description may be obtained from the classical one by a well-defined quantization procedure. The relations between the classical and quantum theories of microscopic and macroscopic systems can be summarized in the diagram:

$$\begin{array}{ccc} \text{Classical mechanics} & \longrightarrow & \text{Classical statistical mechanics} \\ \downarrow & & \\ \text{Quantum mechanics} & \longrightarrow & \text{Quantum statistical mechanics} \end{array}$$

But it would be rather foolish to try to quantize classical statistical mechanics to obtain quantum statistical mechanics; that is, to take temperature, pressure and volume and represent these quantities by operators. It may be argued that when we try to quantize the gravitational field, we are doing something analogous to "quantizing" classical statistical mechanics. Some people advance this as an argument against quantizing general relativity. I prefer not to take sides in this dispute. Instead, I shall suppose that the gravitational field may somehow be quantized and will try to estimate when the quantum effects are noticeable. Again, the arguments will be rough and based on analogies with electrodynamics.

Let ψ be the function describing the matter that is the source of the gravitational field, and let ϕ be a function describing the gravitational field. ψ and ϕ need not necessarily be scalar functions. Then one can expect that the combined matter and gravitation can be described by an actual integral of the symbolic form

$$\frac{1}{\hbar kc}\int(\nabla\phi)^2\,dx + \frac{4\pi}{\hbar c}\int\phi\rho\,dx + \frac{1}{\hbar c}\int(\nabla\psi)^2\,dx,$$

where dx denotes an element of volume in a 4-dimensional space-time manifold. The coefficients involving $\hbar$, k, and c have been introduced so as to make the action dimensionless. The first and last terms describe respectively the free gravitational and matter fields, while the second term describes the interaction, ρ being some bilinear function of ψ and/or $\nabla\psi$ that in some sense can be interpreted as the density of matter.

Write

$$\Phi \overset{\text{def}}{=} \left(\frac{\phi}{\ell c^2}\right)$$

where $\ell = \sqrt{\hbar k/c^3}$ is the characteristic length defined above, and then choose units in which numerically $\hbar = 1$ and $c = 1$. The action integral then takes the form

$$\int \{(\nabla\Phi)^2 + 4\pi\ell\rho\Phi + (\nabla\psi)^2\}\, dx.$$

This has the typical form of an action integral for interacting fields in quantum field theory, with $4\pi\ell$ playing the role of a coupling constant with the dimensions of length.

The dimension of the coupling constant is an important quantity characteristic of any theory of interacting quantized fields, as the renormalizability properties of the theory depend on it in an essential way.

Suppose one wishes to apply the standard perturbation method of calculating transition amplitudes. The n th order term of the S-matrix is, symbolically,

$$S_{(n)} \sim (4\pi i\ell)^n \int \cdots\cdots \int T(\rho_1\Phi_1 \cdot \rho_2\Phi_2 \cdots \rho_n\Phi_n)\, dx_1 \ldots dx_n \,.$$

This is proportional to ℓ^n, and so one can expect that the corresponding contribution to a transition amplitude will be

$$\langle f|S_{(n)}|\, i\rangle \sim \left(\frac{\ell}{\lambda}\right)^n, \qquad (1.21)$$

where $\langle f|$ and $|i\rangle$ are state vectors of the final and initial states respectively, and λ is a wavelength characteristic of the process under consideration. The amplitude (1.21) is significant only for very small λ, $\lambda \sim \ell$. This means that quantum gravitational effects may be expected to be significant only at very high energies.

As an example, consider the transition probabilities for the annihilation of an electron-positron pair into

(i) two photons
(ii) two gravitons.

The cross section σ for reaction (i) is given by

$$\sigma \sim \left(\frac{e^2}{mc^2}\right)^2 \frac{\log\left(\frac{E}{mc^2}\right)}{\left(\frac{E}{mc^2}\right)} \quad \text{for } E>>m,$$

and according to the calculations of Vladimirov,[4] based on a quantized version of the linearized Einstein theory, that for reaction (ii) is given by

$$\sigma \sim \left(\frac{km}{c}\right)^2 \left(\frac{E}{mc^2}\right)^2 \quad \text{for} \quad E \gg m.$$

So we see that the latter process might become significant only at extremely high energies.

As a last comment I should like to say that one should be rather carful about orders of magnitude and should not discuss everything on the assumption that if a dimensionless characteristic number of some process is small then the process is negligible. For example, if m and rare respectively the mass and radius of the Earth, then

$$\frac{km}{c^2 r} \sim 10^{-9}.$$

But nevertheless, gravitation plays an important role in geophysics and in our life! Gravitation plays an important part in astronomy and cosmology because to a very high accuracy large bodies are electrically neutral.

4. J.S. Vladimirov, Proceedings on the Theory of Gravitation, Gauthier-Villars, Paris, and PWN, Warsaw (1964).

2. VECTOR AND TENSOR ALGEBRA

2.1. INTRODUCTION

In this and the following two chapters we shall develop in outline the portions of vector and tensor algebra and differential geometry that will be useful in our study of general relativity later in the course. Most of the theory that we develop here will be familiar to students of general relativity, but not in the form in which we shall obtain it. We shall be using the notation that is nowadays used by pure mathematicians in the fields of group theory and differential geometry, which is sufficiently different from that familiar to physicists that most physicists would have great difficulty in reading the current mathematical literature on differential geometry. Our reason for this approach is threefold:

(a) To enable you to become sufficiently familiar with the language of contemporary differential geometry to be able to read modern books and articles on the subject and to understand the connection between the work done there and that done by physicists working in general relativity;

(b) To clarify the precise meanings of terms such as tensor, manifold, Riemannian space, which are used so freely in physics;

(c) Because it considerably clarifies the role of coordinates in physics.

In this chapter we shall mainly be concerned with vector spaces, but we shall begin by considering more general algebraic structures.

2.2. ALGEBRAIC STRUCTURES

An algebraic structure is an entity consisting of:

(i) A set E, called the basic set, which is non-empty.

(ii) A set Ω, which may be empty, of operators which act on E.

(iii) A number of functions of two variables, called <u>operations</u> which may be of one or two types:
  (a) Internal operations, which map $E \times E \to E$ or $\Omega \times \Omega \to \Omega$.
  (b) External operations, which map $\Omega \times E \to E$

(iv) A set of axioms.

In the above, $\Omega \times E$ etc denotes the <u>Cartesian product</u> of the sets Ω and E, i.e. the set consisting of all ordered pairs, the first member of which lies in Ω and the second in E. Elements of E will be denoted by small Latin letters and elements of Ω by small Greek letters. The operations will be denoted by a symbol placed between the elements on which they act, e.g. an internal operation on E may be denoted by $\top$, and we write $x = y \top z$. We shall denote an algebraic structure by $(E\top, \Omega\overline{\top},\perp)$ etc., where the first number in the bracket denotes the basic set, followed by the internal operations on it; then the set of operators, followed by the internal operations on it; and finally the external operations. If the set of operators is empty, we shall omit it.

Three simple algebraic structures are:

<u>Groupoid</u>. This consists of a set E on which an operation $\cdot$ is defined. There are no axioms, and we denote it by $(E\cdot)$.

<u>Semi-group</u>. This is a groupoid $(E\cdot)$ in which the operation$\cdot$ is associative, i.e.

$$x\cdot(y\cdot z) = (x\cdot y)\cdot z.$$

<u>Group</u>. This is a semi-group $(E\cdot)$ containing a neutral element e (called the unit element unless the group operation is denoted by +, addition, when it is called the zero element), such that

$$e\cdot x = x\cdot e = x \quad \text{for all } x \in E,$$

and such that to every $x \in E$ there corresponds an inverse element $x^{-1} \in E$ such that

$$x^{-1}\cdot x = x\cdot x^{-1} = e.$$

In <u>this section</u> we shall use the following notation:

$\cdot$ will always denote an operation obeying the groupoid axioms, (i.e. no axioms).

$\top$ will always denote an operation obeying the group axioms.

$+$ will always denote an operation obeying the group axioms which is also commutative, i.e. $x + y = y + x$.
$\perp$ will always denote an external operation.

Furthermore, whenever they are used together in an algebraic structure in such a way that the following axioms are meaningful, then they will be assumed to obey those axioms. These axioms are:

$$(1) \quad \begin{cases} (x + y)\cdot z = x\cdot z + y\cdot z \\ x\cdot(y + z) = x\cdot y + x\cdot z \end{cases}$$

$$(2) \quad \begin{cases} (\alpha + \beta)\cdot\gamma = \alpha\cdot\gamma + \beta\cdot\gamma \\ \alpha\cdot(\beta + \gamma) = \alpha\cdot\beta + \alpha\cdot\gamma \end{cases}$$

$$(3) \quad \begin{cases} (\alpha + \beta)\perp x = (\alpha\perp x) + (\beta\perp x) \\ \alpha\perp(x + y) = (\alpha\perp x) + (\alpha\perp y) \end{cases}$$

$$(4) \quad \alpha\perp(x\cdot y) = (\alpha\perp x)\cdot y = x\cdot(\alpha\perp y)$$

$$(5) \quad \alpha\perp(\beta\perp x) = (\alpha\cdot\beta)\perp x \quad \text{or} \quad (\alpha\top\beta)\perp x$$

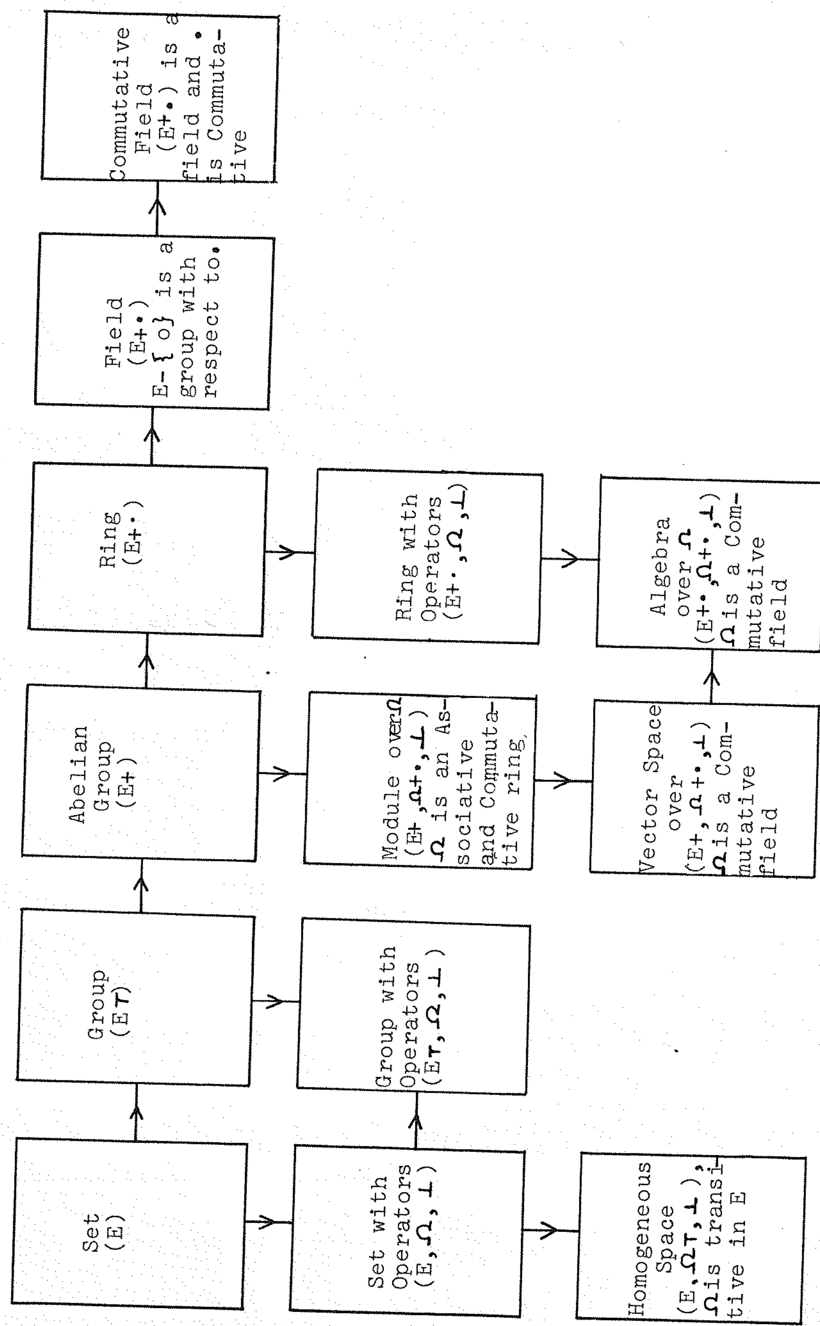
$$(6) \quad \alpha\perp(x\top y) = (\alpha\perp x)\top(\alpha\perp y)$$

(7) If $\varepsilon \in \Omega$ is the neutral element of Ω with respect to $\cdot$ or $\top$, then for every $x \in E$, $\varepsilon\perp x = x$.

The set of operators Ω is said to be <u>transitive in E</u> if, for every $x, y \in E$, there exists an $\alpha \in \Omega$ such that $\alpha\perp x = y$. A group Ω acting in E is said to be effective if $\alpha\perp x = x$ for all $x \in E$ implies that α is the neutral element.

We are now in a position to define easily most of the algebraic structures that are commonly encountered in physics. These are given in the accompanying table, in which the arrows denote the direction of increasing complexity. The axioms of any of the structures listed can be read off immediately from the above list, e.g. a homogeneous space $(E, \Omega\top, \perp)$ satisfies (5) and (7), $(\Omega\top)$ is a group acting transitively in E. An example of a homogeneous space is the affine space $(F, E+, \perp)$ associated with a vector space $(E+, \Omega+\cdot, \perp)$ and defined simply by requiring E to act transitively and effectively on the set F. Well-known examples of commutative fields as the real and complex numbers. An example of a non-commutative field of interest to physicists is the quaternion field. This may be considered as consisting of all 2 by 2 matrices of the form

$$\alpha + i \sum_{k=1}^{3} a_k\sigma_k,$$


Set
(E)
Group
(E T)
Abelian
Group
(E+)
Ring
(E+·)
Field
(E+·)
E-{o} is a
group with
respect to·
Commutative
Field
(E+·) is a
field and ·
is Commuta-
tive
Set with
Operators
(E, Ω, ⊥)
Group with
Operators
(E T, Ω, ⊥)
Module over Ω
(E+, Ω+·, ⊥)
Ω is an As-
sociative
and Commuta-
tive ring
Ring with
Operators
(E+·, Ω, ⊥)
Homogeneous
Space
(E, Ω T, ⊥),
Ω is transi-
tive in E
Vector Space
over
(E+, Ω+·, ⊥)
Ω is a Com-
mutative
field
Algebra
over Ω
(E+·, Ω+·, ⊥)
Ω is a Com-
mutative
field

where α and a_k are real numbers and the σ_k are the Pauli spin matrices.

2.3. EQUIVALENCE RELATIONS AND MORPHISMS

Let E be a set and $E_R \subset E\times E$ be a subset of the Cartesian product ExE. Then E_R is said to define an <u>equivalence relation</u> R on E if it is:

(i) Reflexive, i.e. if $x \in E$ then $(x,x) \in E_R$
(ii) Symmetric, i.e. if $x,y \in E$ and $(x,y) \in E_R$ then $(y,x) \in E_R$
(iii) Transitive, i.e. if $x,y,z \in E$, $(x,y) \in E_R$ and $(y,z) \in E_R$, then $(x,z) \in E_R$.

We then write:

$x \equiv y$ (mod R), or simply $x \equiv y$, (read as x congruent to y) to mean $(x,y) \in E_R$.

Given an equivalence relation R on E it can be used to divide E into a set of disjoint subsets called equivalence classes whose union is the whole of E. $F \subset E$ is called an <u>equivalence class</u> of R if

(i) $x,y \in F$ implies $x \equiv y$ (mod R)
(ii) $x \in F$ and $x \equiv y$ (mod R) implies $y \in F$.

It immediately follows that if F_1 and F_2 are two equivalence classes of R, then either $F_1 = F_2$ or F_1 and F_2 are disjoint, i.e. have no common element. The set of all equivalence classes of R is called the <u>quotient of E by R</u>, written E/R, and the mapping that associates to each $x \in E$ the equivalence class containing it is called the <u>canonical mapping of R</u>.

Now consider two algebraic structures which have the same operations and the same set of operators, but different basic sets, for example $(E\top, \Omega, \perp)$ and $(E'\top, \Omega, \perp)$. Then a mapping f of E into E' which preserves the structure of these algebraic structures is called a <u>homomorphism</u>. In our example f is a homomorphism if

(i) $f(x\top y) = f(x)\top f(y)$ for all $x,y \in E$ and
(ii) $f(\alpha\perp x) = \alpha\perp f(x)$ for all $\alpha \in \Omega$, $x \in E$.

Note that the operations $\top,\perp$ on the left-hand sides of these equations are acting in $(E\top, \Omega, \perp)$ while on the right-hand sides they act in $(E'\top, \Omega, \perp)$.

A homomorphism of E into itself is called an endomorphism of E. A homomorphism f of E into E' satisfying

(i) $x, y \in E$ and $f(x) = f(y)$ implies $x = y$ (i.e. f is one-to-one), and
(ii) if $x' \in E'$ then there exists an $x \in E$ such that $x' = f(x)$ (i.e. f is onto)

is called an isomorphism. An isomorphism of E onto itself is called an automorphism.

Now consider an algebraic structure with base space E and set of operators Ω which has an equivalence relation R defined on E. Then if, for every internal operation $\top$ on E and for every external operation $\perp$,

$$x \equiv x', \qquad y \equiv y', \qquad \text{and} \quad \alpha \in \Omega$$

imply

$$x \top y \equiv x' \top y' \quad \text{and} \quad \alpha \perp x \equiv \alpha \perp x',$$

R is said to be compatible with the structure. In this case one can form the quotient structure which has base space E/R, set of operators Ω, and the same operations, the action of the operations being defined by

$$\text{(i)} \quad \phi(x) \top \phi(y) = \phi(x \top y) \quad \text{for} \quad x, y \in E$$
$$\text{(ii)} \quad \alpha \perp \phi(x) = \phi(\alpha \perp x) \quad \text{for} \quad x \in E, \ \alpha \in \Omega$$

where ϕ is the canonical mapping of R. The consistency of these definitions follows immediately from the compatibility of R with the structure. In general the quotient structure will not satisfy the same axioms as the original structure, but it does in some cases, as for example if the original structure is a vector space (c.f. §2.6). We see that the mapping $x \to \phi(x)$ is a homomorphism of E onto E/R.

2.4. VECTOR SPACES

As defined in Section 2.2, a vector space is an algebraic structure $(\top +, K + \cdot, \perp)$ such that $(K + \cdot)$ is a commutative field, usually the real or complex number field, $(\top +)$ is an abelian group, and the axioms (3), (5) and (7) of the list given there are obeyed. Elements of K will now be denoted by small Latin letters, a, b, ...; these are known as scalars. The neutral elements of K with respect to + and $\cdot$ are denoted respectively by 0 and 1. Elements of $\top$ are called vectors; they will be denoted by small Latin letters with an arrow $\rightarrow$

above them, e.g. $\vec{u}$, $\vec{v}$, ... The neutral element of $\mathcal{T}$, with respect to + will be denoted by $\vec{0}$. For simplicity the operations $\cdot$ and $\perp$ will be denoted by juxtaposition, i.e. we shall write ab instead of a$\cdot$b, and $a\vec{u}$ instead of $a\perp\vec{u}$, where $a,b \in K$, $\vec{u} \in \mathcal{T}$. With this notation the axioms (3), (5) and (7) become

$$(a + b)\vec{u} = a\vec{u} + b\vec{u}$$

$$a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$$

$$a(b\vec{u}) = (ab)\vec{u}$$

$$1\vec{u} = \vec{u}.$$

The n vectors $\vec{u}_1, \ldots, \vec{u}_n$ are said to be <u>linearly independent</u> if

$$\sum_{\alpha=1}^{n} a_\alpha \vec{u}_\alpha = \vec{0},$$

where $a_\alpha \in K$, implies $a_\alpha = 0$, all α. If this is not so, the vectors are called <u>linearly dependent</u>. The least upper bound of n taken over all possible sets of n linearly independent vectors is called the <u>dimension</u> of the vector space. If a vector space $\mathcal{T}$ has dimension n, written

$$\dim \mathcal{T} = n,$$

an ordered set of n linearly independent vectors of $\mathcal{T}$ is called a <u>basis</u> of $\mathcal{T}$. Let $\{\vec{e}_\alpha\}$, $\alpha = 1,2,\ldots n$ be such a basis, and take any $\vec{u} \in \mathcal{T}$. Then by definition of n, the set $\{\vec{u}, \vec{e}_\alpha\}$ is linearly dependent and hence there exist scalars u^α such that

$$\vec{u} - \sum_{\alpha=1}^{n} u^\alpha \vec{e}_\alpha = \vec{0}$$

i.e.

$$\vec{u} = \sum_{\alpha=1}^{n} u^\alpha \vec{e}_\alpha.$$

Thus every vector of $\mathcal{T}$ can be written as a linear combination of basis vectors $\{\vec{e}_\alpha\}$, and the scalars u^α are called the <u>components</u> of $\vec{u}$ with respect to the basis $\{\vec{e}_\alpha\}$.

Now take another basis $\{\vec{e}_{\alpha'}\}$. Then there exist scalars $A^\beta_{\alpha'}$ such that

$$\vec{e}_{\alpha'} = A^\beta_{\alpha'}\,\vec{e}_\beta, \tag{2.1}$$

where we are using the summation convention, that a repeated index is to be summed over its range of values. (This convention will always be used below, unless the contrary is explicitly stated.) However, as $\{\vec{e}_{\alpha'}\}$ is also a basis, there exist scalars $A^{\alpha'}_\beta$ such that

$$\vec{e}_\alpha = A^{\beta'}_\alpha\,\vec{e}_{\beta'}. \tag{2.2}$$

It follows from (2.1) and (2.2) that

$$A^\alpha_{\beta'}\,A^{\beta'}_\gamma = \delta^\alpha_\gamma \tag{2.3}$$

where

$$\delta^\alpha{}_\beta \overset{\text{def}}{=} \begin{cases} 1 \text{ if } \alpha = \beta \\ 0 \text{ if } \alpha \neq \beta \end{cases}$$

is the Kronecker δ symbol. From (2.3) we see that the matrix $(A^\alpha_{\beta'})$ is non-singular and its inverse is $A^{\beta'}_\gamma$. Conversely, given any basis $\{\vec{e}_\alpha\}$ and a non-singular matrix $A^\alpha_{\beta'}$, the vectors $\vec{e}_{\alpha'}$ defined by

$$\vec{e}_{\alpha'} \overset{\text{def}}{=} A^\beta_{\alpha'}\,\vec{e}_\beta$$

form a basis of $\mathcal{T}$.

Let u^α, $u^{\alpha'}$ be the components of $\vec{u} \in \mathcal{T}$ with respect to the bases $\{\vec{e}_\alpha\}$, $\{\vec{e}_{\alpha'}\}$ respectively. Then

$$\begin{aligned} \vec{u} &= u^{\alpha'}\,\vec{e}_{\alpha'} \\ &= u^{\alpha'}\,A^{\beta}_{\alpha'}\,\vec{e}_{\beta} \qquad \text{by (2.1)} \end{aligned} \tag{2.4}$$

But also

$$\vec{u} = u^{\beta}\,\vec{e}_{\beta}. \tag{2.5}$$

Comparing (2.4) and (2.5) and remembering the linear independence of the $\vec{e}_{\alpha}$, we get

$$u^{\alpha} = A^{\alpha}_{\beta'}\,u^{\beta'}$$

and similarly (2.6)

$$u^{\beta'} = A^{\beta'}_{\alpha}\,u^{\alpha}$$

which relates the components of a vector with respect to two bases related by (2.1).

2.5. THE DUAL SPACE

Consider the vector space $(T+, K+\cdot, \perp)$. A mapping ω of T into K is called a <u>linear form on T</u> if for all $\vec{u}, \vec{v} \in T$ and all $a, b \in K$,

$$\omega(a\vec{u} + b\vec{v}) = a\omega(\vec{u}) + b\omega(\vec{v}). \tag{2.7}$$

If ω and π are linear forms on T, we can define $\omega + \pi$ and $a\omega$ by

$$(\omega + \pi)(\vec{u}) \overset{\text{def}}{=} \omega(\vec{u}) + \pi(\vec{u}) \tag{2.8}$$

$$(a\omega)(\vec{u}) \overset{\text{def}}{=} a(\omega(\vec{u})) \tag{2.9}$$

and these are easily seen to be linear forms on T also. Then with these definitions we see that linear forms on T form a vector space over K, called the <u>dual space</u> of T, denoted by T*.

It follows from (2.7) that to determine ω completely it is sufficient to give the scalars ω_α defined by

$$\omega_\alpha \overset{\text{def}}{=} \omega(\vec{e}_\alpha), \quad \alpha = 1,2,\ldots n. \tag{2.10}$$

where $\{\vec{e}_\alpha\}$ is any basis of T, and conversely, any set of n scalars ω_α determine a unique linear form ω through (2.10). We can therefore define n linear forms e^α, $\alpha = 1,2,\ldots n$, by

$$e^\alpha (\vec{e}_\beta) = \delta_\beta{}^\alpha, \tag{2.11}$$

and then for any form ω and any vector $\vec{u}$ with components u^α with respect to the basis $(\vec{e}_\alpha)$ we have

$$e^\alpha(\vec{u}) = e^\alpha(u^\beta\vec{e}_\beta) = u^\beta e^\alpha(\vec{e}_\beta) = u^\alpha \text{ by (2.7) and (2.11)}$$

and (2.12)

$$\omega(\vec{u}) = \omega(u^\beta\vec{e}_\beta) = u^\beta\omega(\vec{e}_\beta) = \omega_\beta u^\beta \text{ by (2.7) and (2.10).}$$

(2.13)

Thus $\omega(\vec{u}) = \omega_\alpha e^\alpha(\vec{u})$ for any vector $\vec{u}$, and so

$$\omega = \omega_\alpha e^\alpha. \tag{2.14}$$

We then see that the n forms e^α are linearly independent and span T*, i.e. every member of T* can be expressed as a linear combination of the e^α. They therefore form a basis of T*, called the <u>basis dual to $(\vec{e}_\alpha)$</u>, and we see that

$$\dim T = \dim T^*.$$

Let $(\vec{e}_{\alpha'})$ be another basis of T related to $(\vec{e}_\alpha)$ by

$$\vec{e}_{\alpha'} = A^\beta_{\alpha'}\vec{e}_\beta, \quad \vec{e}_\beta = A^{\alpha'}_\beta \vec{e}_{\alpha'}$$

and let $(e^{\alpha'})$ be the basis of T* dual to $(\vec{e}_{\alpha'})$. Then

$$\begin{aligned} e^{\alpha'}(\vec{e}_\beta) &= e^{\alpha'}(A^{\gamma'}_\beta \vec{e}_{\gamma'}) \\ &= A^{\gamma'}_\beta e^{\alpha'}(\vec{e}_{\gamma'}) \qquad (2.15) \\ &= A^{\gamma'}_\beta \delta^{\alpha'}_{\gamma'} = A^{\alpha'}_\beta \end{aligned}$$

But also

$$A^{\alpha'}_\gamma e^\gamma(\vec{e}_\beta) = A^{\alpha'}_\beta, \qquad (2.16)$$

and so from (2.15) and (2.16) we get

$$e^{\alpha'} = A^{\alpha'}_\beta e^\beta \qquad (2.17)$$

which shows that the dual basis transforms with the inverse transformation to the basis of T.

Now since T* is a vector space over K, we can form its dual, T**. Let elements of T** be denoted by $\underset{\sim}{u}$, etc. and and let $(\underset{\sim}{e}_\alpha)$ be the basis of T** dual to (e^α) of T*. Then if $\underset{\sim}{u} \in$ T** and

$$\underset{\sim}{u} = u^\alpha \underset{\sim}{e}_\alpha$$

we can associate with it the vector

$$\vec{u} \overset{\text{def}}{=} u^\alpha \vec{e}_\alpha \in T.$$

We easily see that this correspondence between $\underset{\sim}{u} \in$ T** and $\vec{u} \in$ T is independent of the particular basis chosen, and it is a natural isomorphism of T** onto T such that if $\omega \in$ T*,

$$\underset{\sim}{u}(\omega) = \omega(\vec{u}).$$

Because of this we do not regard T** as a new vector space at all, and we can say that the dual space of T* is T.

2.6. LINEAR SUBSPACES

Let T be a vector space and S be a subset of T. Then S is said to be a linear subspace of T if, for all $\vec{u}$, $\vec{v} \in S$ and all scalars a, b, $a\vec{u} + b\vec{v} \in S$. Given any set D of vectors, not necessarily containing a finite number of vectors, the set G(D) of all vectors of T that can be expressed as a finite linear combination of vectors of D is a linear subspace of T called the subspace spanned by D.

Now let S be a linear subspace of T, when T need not necessarily be finite-dimensional. Then we can define an equivalence relation on T by

$$\vec{u} \equiv \vec{v} \quad \text{if and only if} \quad \vec{u} - \vec{v} \in S$$

and then we write

$$\vec{u} \equiv \vec{v} \ (\text{mod } S).$$

We easily verify that this is compatible with the structure of T as a vector space, and then in the manner explained in §2.3 (Equivalence Relations and Morphisms) we can construct the quotient structure (T/S+, K+·, ⊥). Further, we easily verify that the structure satisfies the axioms of a vector space. It is called the quotient space of T by S and is denoted shortly by T/S. If ϕ is the canonical mapping of the equivalence relations, then we see that the zero element of T/S is $\phi(\vec{u})$, any $\vec{u} \in S$.

2.7. TENSOR PRODUCT OF VECTOR SPACES

Let T and U be two vector spaces over the same field K. Form the Cartesian product TxU, i.e., the set of all ordered pairs $(\vec{t}, \vec{u})$, $\vec{t} \in T$, and $\vec{u} \in U$, and from this construct the set G(TxU) of all finite formal linear combinations of elements of TxU, so that the elements of G(TxU) have the form

$$\sum_{(\vec{t},\vec{u}) \in T \times U} a_{\vec{t},\vec{u}} (\vec{t},\vec{u}) \qquad (2.18)$$

where the coefficients are elements of K and the sum contains only a finite number of terms. G(T x U) is thus an infinite-dimensional vector space.

Now let D be the set of all elements of G(T xU) of the form

$$(a\vec{u} + b\vec{v},\ c\vec{w} + d\vec{x}) - ac\,(\vec{u},\vec{w}) - ad(\vec{u},\vec{x})$$
$$- bc(\vec{v},\vec{w}) - bd(\vec{v},\vec{x})$$

and construct the subspace G(D) of G(T x U) spanned by D. We may now form the quotient space G(T x U)/G(D). This is called the tensor product of T and U, and is denoted by $T \otimes U$. Let ϕ be the canonical mapping of G(T x U) onto $T \otimes U$. Then we write

$$\phi(\vec{t},\ \vec{u}) \quad = \quad \vec{t} \otimes \vec{u} \tag{2.19}$$

and we easily see that for any finite sums $\sum a_t \vec{t}$, $\sum b_u \vec{u}$ of vectors of T and U respectively, we have

$$(\sum a_t \vec{t}) \otimes (\sum b_u \vec{u}) \quad = \quad \sum a_t b_t \cdot \vec{t} \otimes \vec{u}. \tag{2.20}$$

Now let dim T = n, dim U = m, and let $(\vec{e}_\alpha)$, $\alpha = 1,2,..n$, and $(\vec{f}_a)$, $a = 1,2,...m$ be respectively bases of T and U. The general member of $T \otimes U$ has the form

$$\phi(\sum a_{\vec{t},\vec{u}}\ (\vec{t},\vec{u})), \tag{2.21}$$

i.e. the image of (2.18) under ϕ. $\vec{t}$ and $\vec{u}$ can be written as

$$\vec{t} \;=\; t^\alpha\, \vec{e}_\alpha, \qquad \vec{u} \;=\; u^a\, \vec{f}_a,$$

and so

$$\sum a_{\vec{t},\vec{u}}\ (\vec{t},\vec{u}) \quad = \quad \sum a_{\vec{t},\vec{u}}\ (t^\alpha\, \vec{e}_\alpha,\ u^a\, \vec{f}_a).$$

But by construction,

$$\sum a_{\vec{t},\vec{u}}(t^{\alpha}\,\vec{e}_{\alpha},u^{a}\vec{f}_{a}) - \sum a_{\vec{t},\vec{u}}\,t^{\alpha}u^{a}(\vec{e}_{\alpha},\vec{f}_{a}) \in G(D).$$

Hence

$$\phi(\sum a_{\vec{t},\vec{u}}(\vec{t},\vec{u})) = \phi(\sum a_{\vec{t},\vec{u}}\,t^{\alpha}u^{a}(\vec{e}_{\alpha},\vec{f}_{a}))$$

$$= \sum a_{\vec{t},\vec{u}}\;t^{\alpha}u^{a}\;\phi(\vec{e}_{\alpha},\vec{f}_{a})$$

by definition of the quotient structure (Section 2.3). Using (2.19) we see that this can be written as

$$\sum a_{\vec{t},\vec{u}}\,t^{\alpha}u^{a}\;\vec{e}_{\alpha}\otimes\vec{f}_{a}$$

and so $\vec{e}_{\alpha}\otimes\vec{f}_{a}$ span $T\otimes U$. Clearly they are linearly independent, and so they form a basis of $T\otimes U$, which thus has dimension mn. Now take any $S \in T\otimes U$. Then there exists a unique set of mn scalars $S^{\alpha a}$ such that

$$S = S^{\alpha a}\,\vec{e}_{\alpha}\otimes\vec{f}_{a}. \tag{2.22}$$

Make a transformation to new bases $(\vec{e}_{\alpha'})$, $(\vec{f}_{a'})$ of T and U respectively, and such that

$$\vec{e}_{\alpha'} = A^{\beta}_{\alpha'}\,\vec{e}_{\beta},\quad \vec{e}_{\beta} = A^{\alpha'}_{\beta}\,\vec{e}_{\alpha'}$$

and

$$\vec{f}_{a'} = B^{b}_{a'}\,\vec{f}_{b},\quad \vec{f}_{b} = B^{a'}_{b}\,\vec{f}_{a'}.$$

If S is now expanded in terms of the new basis $\vec{e}_{\alpha'}\otimes\vec{f}_{a'}$ of $T\otimes U$ as

$$S = S^{\alpha' a'}\vec{e}_{\alpha'}\otimes\vec{f}_{a'},$$

then

$$S^{\alpha' a'} = A^{\alpha'}_{\beta} B^{a'}_{b} S^{\beta b}. \tag{2.23}$$

This is the usual definition of a tensor in terms of the transformation law of its components.

We can form the tensor product of more than two vector spaces by repeating this process. Thus if U, V, W are vector spaces over the same field K, we can form $(U \otimes V) \otimes W$ and $U \otimes (V \otimes W)$. However, just as for the dual of a space T we saw that T^{**} and T were related by a natural isomorphism and need not be considered as distinct, there is also a natural isomorphism between $(U \otimes V) \otimes W$ and $U \otimes (V \otimes W)$, and we do not consider these as distinct either. We can then simply write $U \otimes V \otimes W$, and this is naturally extended to the product of any number of vector spaces.

Now let T be a vector space and let T^* be its dual space. We can then form the repeated tensor product

$$(\otimes^k T) \otimes (\otimes^\ell T^*) \tag{2.24}$$

in which the factor T appears k times and the factor T^* appears ℓ times. Elements of this vector space are called tensors of <u>valence (k,ℓ)</u>, and are said to be k times contravariant and ℓ times covariant. If $(\vec{e}_\alpha)$ is a basis of T and (e^α) is the dual basis of T^*, then the general element S of (2.24) can be written as

$$S = S^{\alpha_1 \ldots \alpha_k}{}_{\beta_1 \ldots \beta_\ell}\, \vec{e}_{\alpha_1} \otimes \ldots \otimes \vec{e}_{\alpha_k} \otimes e^{\beta_1} \otimes \cdots \otimes e^{\beta_\ell}.$$

If we change to a new basis $(\vec{e}_{\alpha'})$ of T and the corresponding dual basis $(e^{\alpha'})$, and if

$$\vec{e}_{\alpha'} = A^{\beta}_{\alpha'} \vec{e}_\beta \quad \text{so that} \quad e^{\alpha'} = A^{\alpha'}_{\beta} e^\beta .$$

then the transformation law for the components of S is

$$S^{\alpha_1' \ldots \alpha_k'}{}_{\beta_1' \ldots \beta_\ell'} = A^{\alpha_1'}_{\gamma_1} \cdots A^{\alpha_k'}_{\gamma_k} A^{\delta_1}_{\beta_1'} \cdots A^{\delta_\ell}_{\beta_\ell'} S^{\gamma_1 \ldots \gamma_k}{}_{\delta_1 \ldots \delta_\ell} \tag{2.25}$$

Any tensor $S \in T \otimes U$ defines a homomorphism of U^* into T by

$$U^* \ni U \to S \cdot U \overset{\text{def}}{=} S^{\alpha a} U_a \vec{e}_\alpha \in T,$$

where $S^{\alpha a}$, U_a are respectively the components of S and U with respect to the bases $(\vec{e}^\alpha)$ of T and $(\vec{f}^a)$ of U. Clearly this homomorphism is independent of the particular bases chosen, and conversely any homomorphism of U^* into T defines a unique element of $T \otimes U$. In particular a tensor of valence (1,1), i.e. an element of $T \otimes T^*$, defines an endomorphism of T, i.e. a linear operator on T. Thus linear operators on T are elements of $T \otimes T^*$.

The tensor product $T \otimes U$ is sometimes defined as the space of homomorphisms of U^* into T.

2.8. MULTIFORMS AND MULTIVECTORS

Let us define the generalized Kronecker δ-symbol by

$$\delta^{\beta_1 \cdots \beta_m}_{\alpha_1 \cdots \alpha_m} = \begin{cases} +1 \text{ if } \beta_1 \cdots \beta_m \text{ is an even permutation of } \alpha_1 \cdots \alpha_m \\ -1 \text{ if } \beta_1 \cdots \beta_m \text{ is an odd permutation of } \alpha_1 \cdots \alpha_m \\ 0 \text{ in all other cases.} \end{cases}$$

Then if $\overset{1}{\omega}, \overset{2}{\omega}, \ldots, \overset{m}{\omega}$ are m linear forms on a vector space T, i.e. are elements of T^*, we define their <u>exterior product</u> (also called the wedge product) by

$$\overset{1}{\omega} \wedge \overset{2}{\omega} \wedge \cdots \wedge \overset{m}{\omega} \overset{\text{def}}{=} \frac{1}{m!} \delta^{12 \cdots m}_{\alpha_1 \alpha_2 \cdots \alpha_m} \overset{\alpha_1}{\omega} \otimes \overset{\alpha_2}{\omega} \otimes \cdots \otimes \overset{\alpha_m}{\omega}. \tag{2.26}$$

It is an element of $\otimes^m T^*$, and the set of all elements of $\otimes^m T^*$ of the form $\overset{1}{\omega} \wedge \cdots \wedge \overset{m}{\omega}$ span a subspace of $\otimes^m T^*$, denoted by $\wedge^m T^*$. One can easily show that if dim T = n then $\wedge^m T^*$ has dimension

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

and that the set $\{e^{\alpha_1}\wedge \ .. \ \wedge e^{\alpha_m}; \quad 1 \leq \alpha_1 < \alpha_2 \ \dots < \alpha_m \leq n\}$ forms a basis of it. Elements of $\wedge^m T^*$ are called m-forms, or generically, multiforms. Multiforms which can be expressed as the exterior product of one-forms (elements of T^*) are called simple.

The tensor

$$\omega \equiv \omega_{\alpha_1 \dots \alpha_m} e^{\alpha_1} \otimes \dots \otimes e^{\alpha_m}$$

belongs to $\wedge^m T^*$ if and only if

$$\omega_{\alpha_1 \dots \alpha_m} = \omega_{[\alpha_1 \dots \alpha_m]}, \tag{2.27}$$

where the square brackets around the indices denote the completely antisymmetric part, i.e.

$$\omega_{[\alpha_1 \dots \alpha_m]} \overset{\text{def}}{=} \frac{1}{m!} \delta^{\beta_1 \dots \beta_m}_{\alpha_1 \dots \alpha_m} \omega_{\beta_1 \dots \beta_m}$$

We shall use this notation extensively later on; and also the completely symmetric part of a quantity will be denoted by round brackets placed round the indices. When (2.27) is satisfied, we easily see that

$$\omega = \omega_{\alpha_1 \dots \alpha_m} e^{\alpha_1} \wedge e^{\alpha_2} \wedge \dots \wedge e^{\alpha_m}.$$

In an analogous manner to the above we can define the exterior product of vectors and the vector space $\wedge^m T$, whose elements are called m-vectors, or generically, multivectors.

Now let S be an m-dimensional linear subspace of the n-dimensional vector space T, and let $\{\vec{f}_a\}$, a = 1,2,...,m be basis of S. Then it can be shown that we can always extend this to a basis $\{\vec{f}_1, \vec{f}_2, \quad , \vec{f}_n\}$ of T by adjoining to it (n-m) more vectors $\vec{f}_{m+1}, \dots, \vec{f}_n$. A basis of T of this form is said to be <u>adapted to the subspace S</u>. Let $\{f^\alpha\}$ be the dual basis to $\{\vec{f}_\alpha\}$, $= 1,2,\ldots n$. We shall use the following ranges for different types of indices:

$\alpha, \beta, \ldots$ run from 1 to n

a,b,... run from 1 to m

$k, \ell, \ldots$ run from (m + 1) to n.

Let $\{\vec{f}_{a'}\}$ be another basis of S, such that

$$\vec{f}_{a'} = B^b_{a'}\vec{f}_b, \quad \vec{f}_b = B^{a'}_b\vec{f}_{a'}, \qquad (2.28)$$

and extend it to a basis $\{\vec{f}_{\alpha'}\}$ of T with corresponding dual basis $\{f^{\alpha'}\}$. Then there exists a non-singular matrix $A^{\alpha'}_\beta$ with inverse $A^\beta_{\alpha'}$, such that

$$\vec{f}_{\alpha'} = A^\beta_{\alpha'} f_\beta, \quad f^{\alpha'} = A^{\alpha'}_\beta f^\beta \qquad (2.29)$$

and from (2.28)

$$A^b_{a'} = B^b_{a'}, \quad A^k_{a'} = 0.$$

Now we have

$$f^{\alpha'}(\vec{f}_{\beta'}) = \delta^{\alpha'}_{\beta'}, \qquad f^\alpha(\vec{f}_\beta) = \delta^\alpha_\beta \qquad (2.30)$$

Hence

$$\begin{aligned} 0 &= f^{k'}(\vec{f}_{a'}) = (A^{k'}_c f^c + A^{k'}_\ell f^\ell)(B^b_{a'}\vec{f}_b) \text{ by (2.29) and (2.28)} \\ &= A^{k'}_c B^b_{a'} f^c(\vec{f}_b) + A^{k'}_\ell B^b_{a'} f^\ell(\vec{f}_b) \\ &= A^{k'}_b B^b_{a'} \quad \text{by (2.30).} \end{aligned}$$

But the matrix $B^b_{a'}$ is non-singular. So we must have

$$A^{k'}_{b} = 0$$

and then we have from (2.29)

$$f^{k'} = A^{k'}_{\ell} f^{\ell}.$$

We thus see that under a change of adapted basis, the m vectors $\vec{f}_a$ and the (n-m) forms f^k transform among themselves according to

$$\vec{f}_{a'} = B^{b}_{a'} \vec{f}_b, \qquad f^{k'} = A^{k'}_{\ell} f^{\ell}. \tag{2.31}$$

Now construct the m-vector $\vec{f}_1 \wedge \vec{f}_2 \wedge \ldots \vec{f}_m$ and the (n-m)-form $f^{m+1} \wedge \ldots \wedge f^n$. From (2.31) and (2.26) we then get that under a change of basis these transform thus:

$$\begin{aligned}
\vec{f}_{1'} \wedge \vec{f}_{2'} \wedge \ldots \wedge \vec{f}_{m'} &= B^{a_1}_{1'} B^{a_2}_{2'} \ldots B^{a_m}_{m'} \vec{f}_{a_1} \wedge \ldots \wedge \vec{f}_{a_m} \\
&= B^{a_1}_{1'} B^{a_2}_{2'} \ldots B^{a_m}_{m'} \delta^{1 \ldots m}_{a_1 \ldots a_m} \vec{f}_1 \wedge \cdots \wedge \vec{f}_m \\
&= (\det B^{a}_{b'}) \vec{f}_1 \wedge \ldots \wedge \vec{f}_m
\end{aligned} \tag{2.32}$$

and similarly $f^{(m+1)'} \wedge \ldots \wedge f^{n'} = (\det A^{k'}_{\ell}) f^{m+1} \wedge \ldots \wedge f^n$.

They are thus multiplied by a non-zero scalar factor under such a change of basis, and so we see that S determines to within a scalar factor an m-vector and an (n-m)-form. Either of these is sufficient completely to determine S, as it is easily seen that two different m-dimensional subspaces determine different m-vectors and (n-m)-forms.

2.9. ORIENTATION OF VECTOR SPACES AND SUBSPACES

Let $(T+, R+\cdot, \perp)$ be an n-dimensional vector space over the field R of real numbers, and let S be an m-dimensional linear subspace of T. Let $\mathcal{F}$ be the set of all bases of S, and let $\{\vec{f}_a\}$, $\{\vec{f}_{a'}\}$, $a = 1,2,\ldots m$ be two such bases, related by

$$\vec{f}_{a'} = B^b_{a'}\,\vec{f}_b .$$

Then we define an equivalence relation R+ on $\mathfrak{F}$ by writing

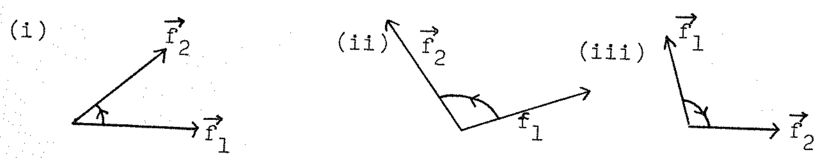
$$\{\vec{f}_{a'}\} \equiv \{\vec{f}_a\} \quad (\text{mod } R+)$$

to mean

$$\det B^b_{a'} > 0.$$

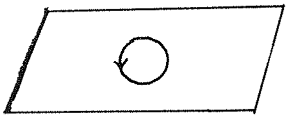
We can then form the quotient set $\mathfrak{F}$/R+, which clearly has just two elements. If one of these elements is singled out and called the positive sense, the other element being called the negative sense, S is said to be provided with an outer orientation if the quotient space T/S has an inner orientation.

To illustrate this we shall consider the case of n = 3, m = 2. S is then a plane in 3-dimensional space, and a basis of S consists of an ordered pair of non-parallel vectors. We easily see that two such bases are contained in the same equivalence class if we have to rotate in the same direction to get from the first vector to the second by a rotation through an angle less than π.

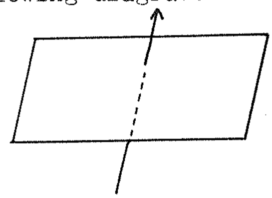


Thus in the diagrams (i) and (ii) have the same orientation while (iii) has the opposite orientation. An inner orientation in a plane is thus provided by giving a direction of rotation.

Now consider an outer orientation in the same example. We must first construct T/S. Two vectors $\vec{a}, \vec{b} \in T$ are congruent (mod S) if $\vec{a}-\vec{b} \in S$. Hence the equivalence classes of S can be represented by a family of planes parallel to S. All we need is one representative of each equivalence class, and this is provided by a line cutting each plane precisely once, and a basis of T/S consists of one element, a directed segment of this line. An inner orientation of T/S is thus seen to consist of a direction along this line. Hence an outer orientation of a plane in 3-dimensional space is provided by a directed vector not lying in the plane. The two concepts are illustrated in the following diagrams.



An inner orientation on a plane.



An outer orientation on a plane in 3 dimensions.

The whole space T cannot be given an outer orientation, but may be given an inner orientation, or simply, an orientation. If an orientation of T is given, then with any inner orientation of S we can associate an outer orientation as follows: Let $\{\vec{f}_1, \ldots, \vec{f}_m\}$ be a basis of S with positive inner orientation. Extend it to a basis $\{\vec{f}_1, \ldots, \vec{f}_n\}$ of T with positive orientation. Let ϕ be the canonical mapping of T to T/S. Then we define the basis $\{\phi(\vec{f}_{m+1}), \ldots, \phi(\vec{f}_n)\}$ of T/S as having positive orientation, and this provides an outer orientation of S.

An example of this is the familiar process of defining the positive normal direction to a directed loop by the 'right-handed corkscrew rule." What this is doing is associating with an inner orientation of a 2-dimensional space the outer orientation induced by the orientation of the 3-dimensional space provided by a right-handed triad.

An orientation of T can be specified by giving a function δ of $\mathfrak{J}$ to R such that $\delta = \pm 1$ and

$$\delta\{\vec{e}_\alpha\} = \delta\{\vec{e}_{\alpha'}\} \quad \text{if and only if} \quad \{\vec{e}_\alpha\} \equiv \{\vec{e}_{\alpha'}\} \ (\text{mod } R+).$$

Then we see that if

$$\vec{e}_{\alpha'} = A^{\beta}_{\alpha'}\, \vec{e}_\beta \quad \text{and} \quad A \overset{\text{def}}{=} \det A^{\beta}_{\alpha'}$$

then

$$\delta\{\vec{e}_{\alpha'}\} = \frac{A}{|A|}\,\delta\{\vec{e}_\alpha\}.$$

$$(\vec{u},\vec{v}) = (\vec{u},\vec{0}) + i(\vec{v},\vec{0}). \tag{2.35}$$

If we now agree that the element $(\vec{u},\vec{0}) \in T \times T$ will simply be written as $\vec{u}$, then from (2.35) we get

$$(\vec{u},\vec{v}) = \vec{u} + i\vec{v}. \tag{2.36}$$

If, as usual, we now omit the operation symbols $\cdot$ and $\perp$, then we can say that the complexification of T consists of of all vectors of the form $\vec{u} + i\vec{v}$, $\vec{u}$, $\vec{v} \in T$, obeying the rules

$$(\vec{u} + i\vec{v}) + (\vec{w} + i\vec{x}) = (\vec{u} + \vec{w}) + i(\vec{v} + \vec{x})$$

$$(a + ib)(\vec{u} + i\vec{v}) = (a\vec{u} - b\vec{v}) + i(b\vec{u} + a\vec{v}),$$

which is the form taken by (i) and (ii) in the new notation. These are just the expressions we would expect by formal expansion of the left-hand sides. But this can only be justified by the above argument.

If $\vec{u}$, $\vec{v} \in T$, the complex conjugate of $(\vec{u},\vec{v}) \in T \times T$ is defined to be $(\vec{u}, -\vec{v})$, i.e.

$$(\vec{u} + i\vec{v})^* = \vec{u} - i\vec{v}.$$

2.12. EUCLIDEAN VECTOR SPACES

Let $(T+, R+\cdot, \perp)$ be a vector space over the real field R. A <u>scalar product</u> on T is a mapping of T x T to R, denoted by

$$(\vec{u},\vec{v}) \longrightarrow \vec{u} \cdot \vec{v},$$

which satisfies

(i) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (symmetry)

(ii) $\vec{u} \cdot (a\vec{v} + b\vec{w}) = a\vec{u} \cdot \vec{v} + b\vec{u} \cdot \vec{w}$ (linearity in the second factor), where $\vec{u}, \vec{v}, \vec{w} \in T$ and $a, b \in R$. We note that (i) and (ii) together imply linearity in the first factor also. If the scalar product also satisfies

(iii) $\vec{u} \cdot \vec{v} = 0$ for all $\vec{v} \in T$ if and only if $\vec{u} = \vec{0}$

it is said to be <u>non-degenerate</u>. If this is not so, it is said to be degenerate. A vector space provided with a scalar product is called a <u>Euclidean vector space</u>.

Let $\{\vec{e}_\alpha\}$ be a basis of T and define the scalars $g_{\alpha\beta}$ by

$$g_{\alpha\beta} \stackrel{\text{def}}{=} \vec{e}_\alpha \cdot \vec{e}_\beta \tag{2.37}$$

If $\{\vec{e}_{\alpha'}\}$ is another basis of T and

$$\vec{e}_{\alpha'} = A^{\beta}_{\alpha'} \vec{e}_\beta \quad , \quad \vec{e}_\alpha = A^{\beta'}_{\alpha} \vec{e}_{\beta'} \tag{2.38}$$

and we also write

$$g_{\alpha'\beta'} \stackrel{\text{def}}{=} \vec{e}_{\alpha'} \cdot \vec{e}_{\beta'} \tag{2.39}$$

then

$$g_{\alpha'\beta'} = A^{\gamma}_{\alpha'} A^{\delta}_{\beta'} g_{\gamma\delta} . \tag{2.40}$$

It follows from this that the scalars $g_{\alpha\beta}$ can be interpreted as the components with respect to the basis $\{\vec{e}_\alpha\}$, of a tensor of valence (0,2). If $\{e^\alpha\}$ is the basis of T* dual to $\{\vec{e}_\alpha\}$, then the tensor itself is given by

$$g = g_{\alpha\beta}\, e^\alpha \otimes e^\beta \quad \in \quad T^* \otimes T^* \tag{2.41}$$

This tensor is called the <u>covariant metric tensor</u> of the Euclidean vector space T. The matrix $g_{\alpha\beta}$ is non-singular if, and only if, the scalar product is non-degenerate.

A non-singular covariant metric tensor defines a natural isomorphism between T and T* in the manner described in Section 2.7. In this isomorphism there corresponds to the vector $\vec{u} = u^\alpha \vec{e}_\alpha \in T$ the linear form $u \in T^*$ defined by

$$u \stackrel{\text{def}}{=} g_{\alpha\beta}\, u^\alpha e^\beta \tag{2.42}$$

If the inverse of the matrix $g_{\alpha\beta}$ is denoted by $g^{\alpha\beta}$, so that

$$g^{\alpha\beta} g_{\beta\gamma} = \delta^{\alpha}_{\gamma},$$

then conversely, to the linear form $v = v_\alpha e^\alpha \in T^*$ corresponds the vector $\vec{v} \in T$ given by

$$\vec{v} = g^{\alpha\beta} v_\alpha \vec{e}_\beta . \tag{2.43}$$

We are here using the convention that vectors and forms which correspond in this natural isomorphism are denoted by the same kernel letter.

More generally we see that g can be used to define a natural isomorphism between $(\otimes^k T) \otimes (\otimes^\ell T^*)$ and $(\otimes^m T) \otimes (\otimes^n T^*)$, provided $k + \ell = m + n$, and again two tensors related by this isomorphism are denoted by the same kernel letter. In terms of components we see that this corresponds to the usual convention on raising and lowering indices with the metric tensor.

Given a non-degenerate scalar product on T we can define a scalar product on T* as follows: if $u, v \in T^*$ and if they correspond in the natural isomorphism to $\vec{u}, \vec{v} \in T$, then we define

$$u \cdot v \stackrel{def}{=} \vec{u} \cdot \vec{v} . \tag{2.44}$$

This clearly satisfies the axiom of a scalar product.

Let $e_\alpha \in T^*$ be the linear form that corresponds to $\vec{e}_\alpha \in T$ in the natural isomorphism. Then since

$$\vec{e}_\alpha = u^{\beta}_{(\alpha)} \vec{e}_\beta \qquad \text{where} \qquad u^{\beta}_{(\alpha)} = \delta^{\beta}_{\alpha},$$

applying (2.42) gives

$$e_\alpha = g_{\beta\gamma} u^{\beta}_{(\alpha)} e^\gamma = g_{\alpha\gamma} e^\gamma ,$$

so that

$$e^\alpha = g^{\alpha\beta} e_\beta . \tag{2.45}$$

But from (2.44),

$$e_\alpha \cdot e_\beta = \vec{e}_\alpha \cdot \vec{e}_\beta = g_{\alpha\beta}$$

and hence using (2.45),

$$e^\alpha \cdot e^\beta = g^{\alpha\gamma} g^{\beta\delta} e_\gamma \cdot e_\delta$$
$$= g^{\alpha\gamma} g^{\beta\delta} g_{\gamma\delta} .$$

Hence

$$e^\alpha \cdot e^\beta = g^{\alpha\beta}$$

which is the formula for T* analogous to (2.37) for T. We see that under the change of basis given by (2.38), $g^{\alpha\beta}$ transforms according to

$$g^{\alpha'\beta'} = A^{\alpha'}_{\gamma} A^{\beta'}_{\delta} g^{\gamma\delta},$$

and so the scalars $g^{\alpha\beta}$ can be considered as the components of a tensor $G \in T \otimes T$, defined by

$$G \stackrel{\text{def}}{=} g^{\alpha\beta} \vec{e}_\alpha \otimes \vec{e}_\beta .$$

and called the <u>contravariant metric tensor</u>.

Given an <u>n-dimensional Euclidean</u> vector space T it can be shown that it is always possible to choose a basis $\{i_\alpha\}$ of it such that

$$i_\alpha \cdot i_\beta = \begin{cases} +1 & \text{if } 1 \leq \alpha = \beta \leq k \leq n, \\ -1 & \text{if } k < \alpha = \beta \leq l \leq n, \\ 0 & \text{in all other cases.} \end{cases}$$

for some integers k, l which are characteristic of T, i.e. are the same for all such bases $\{i_\alpha\}$. If $l < n$ the scalar product is degenerate and the space is said to be <u>singular</u>. If $l = n$, the scalar product is non-degenerate, and such a basis is called <u>orthonormal</u>. We further subdivide this case: If $k = 0$ or $k = n$, the space is called an <u>ordinary Euclidean space</u>. If $0 < k < n$, the space is called a <u>pseudo-Euclidean space</u> and $2k-n$ is called its <u>signature</u>. The particular case of a pseudo-Euclidean space which has $n = 4$, $k = 1$ or 3 is called a <u>Minkowski vector space</u>. This is the space used in the special theory of relativity. We shall use a metric for this space which has $k = 1$ but will change the numbering of the basis vectors so that for an orthonormal basis,

$$g_{11} = g_{22} = g_{33} = -1 \quad , \quad g_{44} = +1 \; , \quad g_{ab} = 0 \quad \text{for } a \neq b$$

and we shall use small Latin letters to run from 1 to 4.

If we have a Euclidean vector space T over the real field R, and we complexify it in the manner explained in Section 2.11, we can introduce in a natural way a scalar product in the complexified space by

$$(\vec{u} + i\vec{v}) \cdot (\vec{w} + i\vec{z}) \overset{\text{def}}{=} (\vec{u}\cdot\vec{w} - \vec{v}\cdot\vec{z}) + i(\vec{v}\cdot\vec{w} + \vec{u}\cdot\vec{z}).$$

This is easily seen to satisfy the requirements of symmetry and bilinearity, and it is non-degenerate if, and only if, the space T is non-singular.

Now let S be a linear subspace of the Euclidean vector space T. Then we can define on S the scalar product <u>induced</u> by T thus:

$$\text{If } \vec{u}, \vec{v} \in S \quad \text{then} \quad (\vec{u}\cdot\vec{v})_S = \vec{u}\cdot\vec{v}.$$

If T is an ordinary Euclidean space, the scalar product induced on S is always non-degenerate, since in such a space, $\vec{u} \neq \vec{0}$ implies $\vec{u}\cdot\vec{u} \neq 0$. But if T is a pseudo-Euclidean space it is possible for the induced metric to be degenerate, i.e. S can contain a vector $\vec{k} \neq \vec{0}$ such that $\vec{k}\cdot\vec{u} = 0$ for all $\vec{u} \in S$. Such an S is called a <u>null subspace</u> of T. We note that in particular the vector $\vec{k}$ satisfies $\vec{k}\cdot\vec{k} = 0$, i.e. $\vec{k}$ is a null vector.

It was shown in Section 2.8 that an m-dimensional subspace S of T can be characterized by an m-vector determined by S up to a scalar factor. Let $f^{\alpha_1 \cdots \alpha_m}$ be the components of such an m-vector characterizing S, so that

$$f^{\alpha_1 \cdots \alpha_m} = f^{[\alpha_1 \cdots \alpha_m]}.$$

Then it can be shown that S is null if and only if there exists a vector $\vec{k} \neq 0$ such that

(i) $$f^{\alpha_1 \cdots \alpha_m} k_{\alpha_m} = 0$$

and

(ii) $$f^{[\alpha_1 \cdots \alpha_m} k^{\alpha]} = 0 \, ;$$

(ii) is the condition that $\vec{k} \in S$ and (i) implies that $\vec{k}$ is orthogonal to every element of S. (Two vectors $\vec{u}, \vec{v} \in T$ are said to be orthogonal if $\vec{u} \cdot \vec{v} = 0$.)

Reference

1. N. Bourbaki, Éléments de Mathématique, Livre II, Algèbre, Ch. 1,2.

3. MINKOWSKI SPACE AND LORENTZ TRANSFORMATIONS

3.1. MINKOWSKI VECTOR SPACES

We have already defined a Minkowski vector space M as a 4-dimensional pseudo-Euclidean vector space over the real field R of signature ± 2, and have said that we shall use the signature -2. A vector $\vec{u} \in M$ is called

time-like if $\vec{u} \cdot \vec{u} > 0$

null if $\vec{u} \cdot \vec{u} = 0$

space-like if $\vec{u} \cdot \vec{u} < 0$.

A linear subspace S of M is said to be

space-like if all its vectors are space-like

time-like if it contains a time-like vector.

We have already defined a null subspace S of M as one containing a null vector $\vec{k}$ such that $\vec{k} \cdot \vec{u} = 0$, all $\vec{u} \in S$. So a null subspace can never be space-like. We shall now show that a time-like vector cannot be orthogonal to a null vector, so that a null subspace cannot be time-like either. So every subspace S of M is precisely one of the following: space-like, null or time-like.

Let $\vec{t}$ be any time-like vector. Then there exists an orthonormal basis of M in which $t^a = 0$ if $a \neq 4$

$$g_{11} = g_{22} = g_{33} = -1, \quad g_{44} = +1$$

Let $t^a = (0,0,0,t)$ and let $\vec{u}$ be any vector orthogonal to $\vec{t}$. Then if $u^a = (u,v,w,x)$, we have

$$0 = \vec{u}\cdot\vec{t} = tx \text{ , so that } x=0.$$

Hence

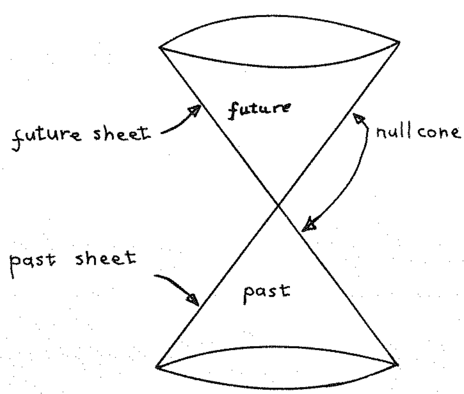
$$\vec{u}\cdot\vec{u} = -u^2 - v^2 - w^2 < 0,$$

and so $\vec{u}$ is space-like. Thus, as stated above, a time-like vector cannot be orthogonal to a null vector or to another time-like vector. We also can easily show that two null vectors are orthogonal if and only if they are proportional (i.e. parallel).

Let $\mathfrak{T}$ be the set of time-like vectors of M. Define an equivalence relation $\uparrow$ on $\mathfrak{T}$ by writing, for $\vec{u}, \vec{v} \in \mathfrak{T}$

$$\vec{u} \equiv \vec{v} \pmod{\uparrow} \quad \text{to mean} \quad \vec{u}\cdot\vec{v} > 0$$

We can now form the quotient $\mathfrak{T}/\uparrow$, and it clearly contains just two elements. If now one of these elements is singled out and called the positive time direction, the Minkowski vector space is said to be <u>oriented in time</u>. The singled-out element of $\mathfrak{T}/\uparrow$ is called the <u>future</u> and the other element is called the <u>past</u>. The set of all null vectors forms a cone, called the <u>null cone</u>, as indicated in the diagram. The future-pointing time-like vectors all lie inside one sheet of the cone, called the future sheet, and the past-pointing time-like vectors all lie inside the other sheet, called the past sheet. We use the notation $\uparrow$ for the equivalence relation because of this connection with the direction of time.



3.2. LORENTZ TRANSFORMATIONS

We showed in section 2.7 that an element of the tensor product $T \otimes T^*$ defines an endomorphism of the vector space T. Now for T take the Minkowski vector space M, and let $L \in M \otimes M^*$.

We shall write the endomorphism of M defined by L thus:

$$M \ni \vec{u} \to L\vec{u} \in M.$$

L is called a <u>Lorentz transformation</u> if, for every $\vec{u} \in M$,

$$L\vec{u} \cdot L\vec{u} = \vec{u} \cdot \vec{u}. \tag{3.1}$$

If we now let $\vec{u}, \vec{v} \in L$, then from (3.1) we must have

$$L(\vec{u}+\vec{v}) \cdot L(\vec{u}+\vec{v}) = (\vec{u}+\vec{v}) \cdot (\vec{u}+\vec{v})$$

i.e.

$$L\vec{u} \cdot L\vec{u} + 2L\vec{u} \cdot L\vec{v} + L\vec{v} \cdot L\vec{v} = \vec{u} \cdot \vec{u} + 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v}$$

and hence by (3.1) this gives

$$L\vec{u} \cdot L\vec{v} = \vec{u} \cdot \vec{v}. \tag{3.2}$$

Thus a Lorentz transformation of M is a linear transformation which preserves scalar products.

Now let L_1 and L_2 be two Lorentz transformations. We define their product $L_1 L_2$ by

$$L_1 L_2 \vec{u} \stackrel{\text{def}}{=} L_1 (L_2 \vec{u}) \tag{3.3}$$

for $\vec{u} \in M$, and we easily verify that this is also a Lorentz transformation. Clearly the identity transformation is a Lorentz transformation, and from (3.1) it follows that L is a non-singular mapping and so has an inverse L^{-1}. So with respect to the product defined by (3.3), Lorentz transformations form a group, called the <u>Lorentz group</u>.

Using the notation of sections 3.1 and 2.9, a Lorentz transformation L is said to be <u>orthochronous</u> if, for every $n \in \mathfrak{J}$,

$$L\vec{u} \equiv \vec{u} \pmod{\uparrow},$$

and is said to be <u>orientation-preserving</u> if for every basis $\{\vec{e}_\alpha\}$ of M,

$$\{L\vec{e}_\alpha\} \equiv \{\vec{e}_\alpha\} \pmod{R_+}.$$

If it is both orthochronous and orientation-preserving, then L is called <u>proper</u>. A proper Lorentz transformation thus preserves the direction in time of time-like vectors and preserves the inner orientation defined by a triad of space-like vectors. We easily see that proper Lorentz transformations form a subgroup of the Lorentz group, called the <u>proper Lorentz group</u>.

3.3. NULL TETRAD AND THE SPECIAL LINEAR GROUP SL(2,C)

Let $\{\vec{x},\vec{y},\vec{z},\vec{t}\}$ be an orthonormal basis of Minkowski space M such that

$$\vec{x}^2 = \vec{y}^2 = \vec{z}^2 = -\vec{t}^2 = -1, \tag{3.4}$$

where we write $\vec{u}^2 \overset{\text{def}}{=} \vec{u}\cdot\vec{u}$, and let $\overline{M}$ be the complexification of M as defined in Section 2.11, provided with the scalar product induced by M as defined in Section 2.12. Define vectors† $\vec{m}, \vec{l}, \vec{n} \in \overline{M}$ by

$$\begin{aligned} \sqrt{2}\,\vec{m} &= \vec{x} + i\vec{y} \\ \sqrt{2}\,\vec{l} &= \vec{t} + \vec{z} \\ \sqrt{2}\,\vec{n} &= \vec{t} - \vec{z} \end{aligned} \tag{3.5}$$

so that $\sqrt{2}\,m^* = \vec{x} - i\vec{y}$

---

† These are the same as the vectors introduced by Professor Pirani in his course of lectures, but corresponding to our vectors $\vec{m}, \vec{l}, \vec{n}$ he used the notation $\vec{t}, \vec{k}, \vec{m}$.

where * denotes complex conjugate. Then $\vec{l}$, $\vec{n}$ are real, $\vec{m}$, $\vec{m}^*$ are complex, and together they form a basis of M. We easily see that they are all null vectors, and that the only non-zero scalar products between them are

$$\vec{l}\cdot\vec{n} = 1 \qquad\qquad \vec{m}\cdot\vec{m}^* = -1 \tag{3.6}$$

A set of four null vectors with these properties is called a <u>null tetrad</u>.

Let $\vec{a} \in$ M. As explained in Section 2.11 we can also consider it as being in $\overline{M}$, and hence it can be expressed in terms of the basis $\vec{l}$, $\vec{n}$, $\vec{m}$, $\vec{m}^*$ of $\overline{M}$ thus:

$$\vec{a} = p\vec{n} + q\vec{l} + r\vec{m} + s\vec{m}^*, \quad p, q, r, s \in \mathbb{C}$$

and by scalarly multiplying this separately by $\vec{n}$, $\vec{l}$, $\vec{m}$, $\vec{m}^*$ and using (3.6) we obtain p, q, r, s so that

$$\vec{a} = (\vec{a}\cdot\vec{l})\vec{n} + (\vec{a}\cdot\vec{n})\vec{l} - (\vec{a}\cdot\vec{m})\vec{m}^* - (\vec{a}\cdot\vec{m}^*)\vec{m} \tag{3.7}$$

Squaring both sides of this equation gives

$$\vec{a}^2 = 2(\vec{a}\cdot\vec{l})(\vec{a}\cdot\vec{n}) - 2(\vec{a}\cdot\vec{m})(\vec{a}\cdot\vec{m}^*) \tag{3.8}$$

Now let us define a function $\vec{a} \to A(\vec{a})$ on M mapping M into the group of (2 x 2) complex matrices by

$$A(\vec{a}) \stackrel{def}{=} \sqrt{2}\begin{pmatrix} \vec{a}\cdot\vec{l} & \vec{a}\cdot\vec{m}^* \\ \vec{a}\cdot\vec{m} & \vec{a}\cdot\vec{n} \end{pmatrix} \tag{3.9}$$

If we define symbolically the 2 x 2 matrix $\vec{\sigma}$ whose elements are vectors of $\overline{M}$ by

$$\vec{\sigma} \stackrel{def}{=} \sqrt{2}\begin{pmatrix} \vec{l} & \vec{m}^* \\ \vec{m} & \vec{n} \end{pmatrix}$$

then (3.9) can be symbolically written as

$$A(\vec{a}) = \vec{a}\cdot\vec{\sigma} \tag{3.10}$$

Now $\vec{a} \in M$, and hence $\vec{a}\cdot\vec{\ell}$, $\vec{a}\cdot\vec{n}$ are real numbers and $\vec{a}\cdot\vec{m}$, $\vec{a}\cdot\vec{m}^*$ are complex conjugate numbers. So from (3.9) the matrix $A(\vec{a})$ is hermitian. Also, from (3.7) and (3.9) we see that any 2 x 2 hermitian matrix H determines a vector $\vec{a} \in M$ by $A(\vec{a}) = H$. Equation (3.9) thus defines a one-to-one correspondence between vectors of M and hermitian (2 x 2) matrices.

From (3.8) and (3.9) we get

$$\vec{a}^2 = \det A(\vec{a}). \tag{3.11}$$

Let SL(2,C) be the special linear group of order 2 over the complex field C, i.e. the group of 2 x 2 complex matrices whose determinants are unity (i.e. they are <u>unimodular</u>). Let $U \in SL(2,C)$ and $\vec{a} \in M$. Then $UA(\vec{a})U^{\dagger}$ is a hermitian matrix, where † denotes hermitian conjugate, and hence there exists a unique $\vec{a}' \in M$ such that

$$A(\vec{a}') = U A(\vec{a}) U^{\dagger} \tag{3.12}$$

Now det U = 1, and hence

$$\begin{aligned} \det A(\vec{a}') &= \det U \det A(\vec{a}) \det U^{\dagger} \\ &= \det A(\vec{a}), \end{aligned}$$

and so by (3.11)

$$(\vec{a}')^2 = \vec{a}^2 . \tag{3.13}$$

But clearly the correspondence $\vec{a} \to \vec{a}'$ defined by (3.12) is linear. If we write

$$\vec{a}' = L(U)\vec{a} \tag{3.14}$$

then (3.13) gives

$$L(U)\vec{a}\cdot L(U)\vec{a} = \vec{a}\cdot\vec{a}\ , \quad \text{all } \vec{a} \in M.$$

Thus from (3.1), $L(U)$ is a Lorentz transformation. We see from (3.12) that $L(U) = L(-U)$, and hence both $\pm U$ give rise to the same Lorentz transformation. It can be shown that $L(U)$ is a proper Lorentz transformation, that every proper Lorentz transformation L can be obtained in this way from some $U \in SL(2,C)$, and that L determines this U to within a sign.

Now let $U_1, U_2 \in SL(2,C)$. Then

$$\begin{aligned} A(L(U_1U_2)\vec{a}) &= U_1U_2A(\vec{a})U_2^{+}U_1^{+} \\ &= U_1A(L(U_2)\vec{a})U_1^{+} \\ &= A(L(U_1)L(U_2)\vec{a}) \end{aligned}$$

for all $\vec{a} \in M$, by repeated use of (3.12) and (3.14). Thus

$$L(U_1U_2) = L(U_1)L(U_2).$$

The mapping $U \to L(U)$ is thus a homomorphism of the group SL(2,C) onto the proper Lorentz group which preserves the group structure. If we define an equivalence relation X on SL(2,C) by writing

$$U \equiv V (\mathrm{mod}\, X) \quad \text{to mean} \quad U=V \text{ or } U=-V$$

and let ϕ be the canonical mapping of SL(2,C) to SL(2,C)/X, then we see that the mapping $\phi(u) \to L(u)$ for $u \in SL(2,C)$ is an isomorphism between the proper Lorentz group and the group SL(2,C)/X.

As an example of this correspondence, consider

$$U = \pm\begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix}.$$

Then we easily see that

$$L(\vec{l})=\vec{l}, \quad L(\vec{n})=\vec{n}, \quad L(\vec{m})=e^{-i\phi}\vec{m}, \quad L(\vec{m}^{*})=e^{i\phi}\vec{m}^{*}$$

and so from (3.5)

$$L(\vec{x}) = \vec{x}\cos\phi + \vec{y}\sin\phi$$

$$L(\vec{y}) = -\vec{x}\sin\phi + \vec{y}\cos\phi$$

$$L(\vec{z}) = \vec{z}\ , \quad L(\vec{t}) = \vec{t}$$

This is a rotation through an angle ϕ in the $\vec{x},\vec{y}$ plane.

3.4. LORENTZ TRANSFORMATIONS AND BILINEAR TRANSFORMATIONS OF THE COMPLEX PLANE

Let $\vec{k}$ be any real null vector, so that $\vec{k}^2 = 0$. Then if $\vec{k}$ is not parallel to $\vec{\ell}$, $\vec{k}\cdot\vec{\ell} \neq 0$. Define a mapping $\vec{k} \to \zeta(\vec{k})$ of real null vectors not parallel to $\vec{\ell}$ into C by

$$\zeta(\vec{k}) \overset{\text{def}}{=} -\frac{\vec{k}\cdot\vec{m}^*}{\vec{k}\cdot\vec{\ell}}\ . \tag{3.15}$$

Then if $\zeta \neq 0$, $\vec{k}\cdot\vec{m} \neq 0$ and from (3.8) for $\vec{a} = \vec{k}$ we get

$$\zeta(\vec{k}) = -\frac{\vec{k}\cdot\vec{m}^*}{\vec{k}\cdot\vec{\ell}} = -\frac{\vec{k}\cdot\vec{n}}{\vec{k}\cdot\vec{m}} \tag{3.16}$$

If $\vec{k}\cdot\vec{\ell} = 0$ then $\vec{k}\cdot\vec{n} \neq 0$ and from (3.8) $\vec{k}\cdot\vec{m} = 0$. The second term in (3.16) then becomes the indeterminate 0/0 while the last term becomes $\vec{k}\cdot\vec{n}/0 = \infty$. So if we let $\hat{C}$ denote the extended complex plane, i.e. C together with ∞, we can extend our map $\vec{k} \to \zeta(\vec{k})$ to include the case $\vec{k}\cdot\vec{\ell} = 0$ by then defining $\zeta(\vec{k}) = \infty$. It is then a mapping of all real null vectors into the extended complex plane $\hat{C}$.

Now from (3.7) and (3.16), if $\zeta \neq \infty$

$$\vec{k} = (\vec{k}\cdot\vec{\ell})(\vec{n} + \zeta\zeta^*\vec{\ell} + \zeta\vec{m} + \zeta^*\vec{m}^*) \tag{3.17}$$

and so ζ determines the direction of $\vec{k}$, which we write as $G(\vec{k})$. $G(\vec{k})$ denotes the one-dimensional vector subspace of $\overline{M}$ given by

$$G(\vec{k}) \overset{\text{def}}{=} \{a\vec{k};\ a \in C\}.$$

If $\zeta(\vec{k}) = \infty$, then by definition $\vec{k} \in G(\vec{\ell})$. Hence in all cases ζ determines the direction of $\vec{k}$. We therefore have defined a one-to-one correspondence between directions of real null vectors and the extended complex plane.

Now let $L(u)$ be the proper Lorentz transformation corresponding to the unimodular matrix

$$U = \begin{pmatrix} a^* & -b^* \\ -c^* & d^* \end{pmatrix}, \quad ad - bc = 1$$

Let $\vec{k}$ be a real null vector, and put

$$\zeta = \zeta(\vec{k}), \quad \zeta' = \zeta(L(u)\vec{k}).$$

Then using (3.12) and

$$A(\vec{k}) = \sqrt{2}(\vec{k}\cdot\vec{\ell})\begin{pmatrix} 1 & -\zeta \\ -\zeta^* & \zeta\zeta^* \end{pmatrix}$$

which comes from (3.9) and (3.17), we get

$$\zeta' = \frac{c + d\zeta}{a + b\zeta} \tag{3.18}$$

which is a bilinear (fractional linear) transformation of the extended complex plane.

We have thus obtained a mapping of proper Lorentz transformations into the group of bilinear transformations of the extended complex plane. But also, given any bilinear transformation (3.18) it determines the ratios of a:b:c:d, and together with the condition ad-bc = 1, we see that a, b,c,d are determined to within a common sign, and so U is determined to within a sign. But this then determines a unique proper Lorentz transformation. We have thus shown that:

There exists a one-to-one correspondence between proper Lorentz transformations and bilinear transformations of the extended complex plane.

Now a bilinear transformation which is not the identity has just two fixed points, which may be coincident, and such a fixed point corresponds to a real null direction which is left invariant by the corresponding proper Lorentz transformation. We thus see that a proper Lorentz transformation leaves unaltered two real null directions, which may be coincident.

We leave the following as an exercise:

Exercise

Prove that for every proper Lorentz transformation L which is not the identity transformation, there exists a null tetrad $\vec{n}$, $\vec{\ell}$, $\vec{m}$, $\vec{m}^*$ such that:

(i) if L leaves invariant two distinct real null directions,

$$L\vec{m} = e^{i\phi}\vec{m}$$
$$L\vec{\ell} = e^{\psi}\vec{\ell}$$
$$L\vec{n} = e^{-\psi}\vec{n}$$

where ϕ, ψ are real numbers

or (ii) if L leaves invariant only one real null direction,

$$L\vec{m} = \vec{m} + z\vec{n}$$
$$L\vec{\ell} = \vec{\ell} + z\vec{m}^* + z^*\vec{m} + zz^*\vec{n}$$
$$L\vec{n} = \vec{n}$$

where z is a complex number.

3.5. SPINORS

In Section 3.3 we defined a matrix $\vec{\sigma}$ of vectors of $\overline{M}$ by

$$\vec{\sigma} = \sqrt{2}\begin{pmatrix} \vec{\ell} & \vec{m}^* \\ \vec{m} & \vec{n} \end{pmatrix} .$$

Using (3.5) this can be written as

$$\vec{\sigma} = \begin{pmatrix} \vec{t}+\vec{z} & \vec{x}-i\vec{y} \\ \vec{x}+i\vec{y} & \vec{t}-\vec{z} \end{pmatrix}$$

or

$$\vec{\sigma} = I\vec{t} + \sigma_x\vec{x} + \sigma_y\vec{y} + \sigma_z\vec{z} \tag{3.19}$$

where I is the 2 x 2 unit matrix and σ_x, σ_y, σ_z are the Pauli spin matrices defined by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.20}$$

Now from (3.10), (3.12) and (3.14) we have

$$(L(u)\vec{a})\cdot\vec{\sigma} = U(\vec{a}\cdot\vec{\sigma})U^{+}$$

But from (3.2),

$$L(u)\vec{a}\cdot\vec{\sigma} = \vec{a}\cdot L^{-1}(u)\vec{\sigma}$$

where in the term $L^{-1}(u)\vec{\sigma}$, $\vec{\sigma}$ is considered as a vector whose components are 2 x 2 matrices. So we now have

$$\vec{a}\cdot L^{-1}(u)\vec{\sigma} = \vec{a}\cdot U\vec{\sigma}U^{+}$$

and this holds for all $\vec{a} \in M$. Therefore

$$L^{-1}(u)\vec{\sigma} = U\vec{\sigma}U^{+}$$

or

$$L(u)\vec{\sigma} = U^{-1}\vec{\sigma}(U^{-1})^{+} \tag{3.21}$$

This gives the transformation law of the matrix $\vec{\sigma}$ when we make a Lorentz transformation of the vectors $\vec{l}$, $\vec{m}$, $\vec{n}$ used in defining it.

Now in Section 3.3 we show that

$$L(U_1U_2) = L(U_1)L(U_2). \tag{3.22}$$

If we insert the dependence and write U(L) to denote the 2 x 2 unimodular matrix determined by L to within a sign, it follows from (3.22) that

$$U(L_1)U(L_2) = \pm U(L_1L_2) . \tag{3.23}$$

The matrices U(L) are said to form a two-valued representation of the proper Lorentz group. So do the complex conjugate matrices $U^*(L)$, as we see from (3.23) that we also have

$$U^*(L_1)U^*(L_2) = \pm U^*(L_1L_2) . \tag{3.24}$$

Now define two complex 2-dimensional vector spaces V and $\overline{V}$, such that when we transform M with the Lorentz transformation L, then V and $\overline{V}$ are transformed respectively with U(L) and $U^*(L)$. Elements of V are called spinors of valence (1,0) and elements of $\overline{V}$ are called conjugate spinors of valence (1,0). Then, just as for tensors, we can define spinors of higher valence as elements of repeated tensor products of V, $\overline{V}$ and their duals V^*, $\overline{V}^*$. From (3.21) we have

$$\vec{\sigma} = L(\omega) U \vec{\sigma} U^{+}$$

and so we see that if we regard $\vec{\sigma}$ as an element of $M \otimes V \otimes \overline{V}$, it is invariant under a Lorentz transformation. We can then use $\vec{\sigma}$ to define an isomorphism between $V^* \otimes \overline{V}^*$ and M in the manner explained in Section 2.7.

Closely related to the matrix $\vec{\sigma}$ is the matrix $\vec{\tau}$ defined by

$$\vec{\tau} = \sqrt{2} \begin{pmatrix} \vec{n} & -\vec{m}^* \\ -\vec{m} & \vec{l} \end{pmatrix}$$

This is related to $\vec{\sigma}$ by

$$\vec{\tau} = \epsilon \, \vec{\sigma}^* \epsilon^{-1}$$

where

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

In the same way that $\vec{\sigma}$ can be considered as an invariant element of $M \otimes V \otimes \overline{V}$, it can be shown that ϵ can be considered as an invariant element of $V^* \otimes V^*$ or $\overline{V}^* \otimes \overline{V}^*$, and $\vec{\tau}$ as an invariant element of $M \otimes \overline{V}^* \otimes V^*$.

We leave the following results as exercises.

Excercise 1.

Prove that

$$\vec{\sigma} \otimes \vec{\tau} + (\vec{\tau} \otimes \vec{\sigma})^{\dagger} = 2G \cdot I$$

where $\otimes$ denotes both the tensor and matrix product, G is the contravariant metric tensor of Minkowski space, and I is the 2 x 2 unit matrix.

Exercise 2.

If $u \in \bar{V}^*$ and $u^{\dagger}$ is its hermitian conjugate, u being considered as a column vector so that $u^{\dagger}$ is a row vector, then

$$\vec{k} \stackrel{\text{def}}{=} u^{\dagger} \vec{\sigma} u$$

is a real null vector and

$$u u^{\dagger} = \tfrac{1}{2} \vec{k} \cdot \vec{\tau}.$$

This is a particular example of the isomorphism between M and $V^* \otimes V^*$ determined by $\vec{\sigma}$.

Exercise 3.

Prove that $\vec{\tau}$ satisfies the equation

$$L(U)\vec{\tau} = U^{\dagger} \vec{\tau} U.$$

This proves that $\vec{\tau}$ is invariant under Lorentz transformations when considered as an element of $M \otimes V^* \otimes V^*$.

Reference

1. R. Penrose and W. Rindler, Spinor Calculus, in preparation.

3

4. DIFFERENTIAL GEOMETRY

4.1. TOPOLOGICAL SPACES

We begin this Chapter by outlining that portion of point set topology that we shall need in our study of differentiable manifolds later in the Chapter.

Let X be any set and θ be a collection of subsets of X such that

(i) The empty set $\phi \in \theta$ and $X \in \theta$

(ii) If $V_\alpha \in \theta$ for all $\alpha \in I$, where I is an arbitrary set, then $\bigcup_{\alpha \in I} V_\alpha \in \theta$

(iii) If $V_1, V_2, \ldots, V_n \in \theta$ then $\bigcap_{i=1}^{n} V_i \in \theta$

where $\bigcup$ and $\bigcap$ respectively denote the union and intersection of sets. Then θ is said to define a topology for X, and X is then called a topological space. The subsets of X contained in θ are called open sets of X.

We give now some examples of topologies for an arbitrary set X. Each of the following definitions of θ gives a topology for X:

(i) θ contains all subsets of X;

(ii) θ contains the empty set ϕ and all subsets of X that contain all but a finite number of elements of X;

(iii) $\theta = \{\phi, X\}$.

If a set X is provided with two topologies θ and θ', θ' is said to be a stronger topology than θ if every subset of X contained in θ is also contained in θ'. In this case we also say that θ is a weaker topology than θ'. In the above example (i) is stronger than (ii) which in turn is stronger than (iii). (i) is the strongest possible topology for X and (iii) is the weakest possible topology.

A neighborhood of $p \in X$ is any open set containing p. The sequence of points $p_i \in X$, i = 1,2,... is said to converge to the point $p \in X$ if every neighborhood of p contains all but a finite number of the points p_i, and then p is called limit of the sequence. It is possible for a sequence to have several different limits. In the examples above, with topology (iii) every sequence converges to every point of X. With topology (i), no sequence containing an infinite number of different points can converge to any limit at all. For this reason, the topology (i) for X is called the discrete topology.

A topological space $\{X, \theta\}$ that satisfies the Hausdorff axiom, namely: If p,q are distinct points of X then there exist neighborhoods U_p, U_q, of p,q respectively such that $U_p \cap U_q = \phi$ is called a Hausdorff space. It is easily seen that in a Hausdorff space a sequence cannot have two distinct limits.

A collection $B \subset \theta$ of open sets of X is called a base of the topological space if every open set can be expressed as the union of (possibly an infinite number) elements of B. A base is said to be countable if it contains a denumerable number of elements.

If R is the real field, and $a, b \in R$, with $a < b$, then the subset of R given by $a < x < b$ is called an open interval of R. We define the natural topology θ of R by taking as a base the set of all open intervals of R.

A topological space $\{X, \theta\}$ is called connected if X is not the union of two disjoint non-empty open sets.

4.2. PRODUCTS OF TOPOLOGICAL SPACES

Let $\{X, \theta\}$, $\{X', \theta'\}$ be two topological spaces. Define a set B of subsets of the Cartesian product $X \times X'$ by

$$B = \{ u \times u' : u \in \theta , u' \in \theta' \}.$$

Then we can define a topology on $X \times X'$ by saying that B forms a base of this topology, so that the open sets are the union of elements of B. Clearly this satisfies all the axioms for a topology, and $X \times X'$ endowed with this topology is called the topological product space of the topological spaces $\{X, \theta\}$ and $\{X', \theta'\}$. We can similarly form the product space of any number of topological spaces.

In this way we can put a topology on n-dimensional number space $R^n \overset{\text{def}}{=} R \times R \times \ldots \times R$ to n factors by defining it to have the product topology when R has its natural topology. This is called the natural topology of R^n.

4.3. CONTINUOUS FUNCTIONS AND HOMEOMORPHISMS

Let $\{X,\theta\}$ and $\{X',\theta'\}$ be two topological spaces, and let f be a function on X with values in X', written

$$X \ni p \rightarrow f(p) \in X'.$$

f is said to be continous at $p \in X$ if, corresponding to every neighborhood $V' \in \theta'$ of f(p), there is a neighborhood $V \in \theta$ of p such that

$$q \in V \quad \text{implies} \quad f(q) \in V'.$$

f is said to be continuous on a subset $D \subset X$ if it is continuous at every point of D.

If f is a one-to-one mapping of X onto X' such that both f and its inverse f^{-1} are continuous, then f is called a homeomorphism of X onto X'. If V is a subset of X, we define f(V) as a subset of X' by

$$q \in f(V) \text{ if there exists a } p \in V \text{ such that } q = f(p).$$

Then f is a homeomorphism of X onto X' if and only if : f(V) is open in X' if and only if V is open in X.

4.4. DIFFERENTIABLE MANIFOLDS

A function f mapping R^n to R is said to be of class C^ℓ in some region of R^n if it is ℓ times differentiable and its ℓth derivatives are continuous in that region. It is said to be of class C^∞ in a region of R^n if it is differentiable an infinite number of times in that region.

Let $\{X,\theta\}$ be a connected Hausdorff space with a countbasis. Let n be a positve integer and ℓ be a positve integer or ∞, and let $\mathfrak{D}$ be a family of real-valued functions defined on X satisfying the following axioms:

(i) if f is a real-valued function defined on X such that to every point $a \in X$ there corresponds a neighborhood U of a and a function $g \in \mathfrak{D}$ such that $f(b) = g(b)$ for all $b \in U$, then $f \in \mathfrak{D}$.

(ii) if k is a positive integer and $g^1, g^2, \ldots, g^k \in \mathcal{D}$, and if f is a real-valued function defined on R^k of class C^ℓ, then $f(g^1, g^2, \ldots, g^k) \in \mathcal{D}$.

(iii) to every point $a \in X$ there corresponds a neighborhood U of a and n functions $x^1, x^2, \ldots, x^n \in \mathcal{D}$ such that

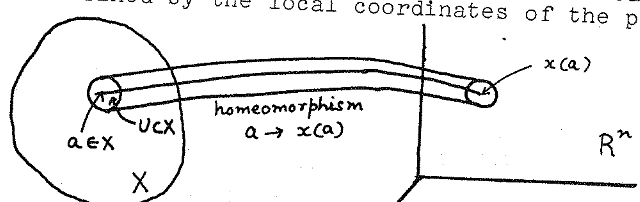
(a) the mapping $U \ni b \to (x^1(b), x^2(b), \ldots, x^n(b)) \in R^n$ is a homeomorphism of U onto a subset of R^n

(b) to every $f \in \mathcal{D}$ corresponds a function F defined on R^n, of class C^ℓ such that if $b \in U$ then

$$f(b) = F(x^1(b), x^2(b), \ldots x^n(b))$$

Then the triple $\{X, \theta, \mathcal{D}\}$ is said to be an <u>n-dimensional differentiable manifold</u>, of class C^ℓ and the functions contained in $\mathcal{D}$ are called <u>differentiable functions</u> on X to R. The set $\{x^1, \ldots, x^n\}$ of differentiable functions defined by axiom (iii) are called <u>local coordinates</u> in U (or around a). If they further satisfy axiom (iii) for U = X, they are called <u>global coordinates</u>. Global coordinates need not necessarily exist.

When dealing with differentiable manifolds, one must always keep clear in one's mind the distinction between the space X and the space R^n to which it is locally homeomorphic. The points of the manifold are points of X, but the coordinates of the points of X are points of R^n. If for $a \in X$ we write x(a) as an abbreviation for the set $\{x^1(a), \ldots, x^n(a)\} \in R^n$, the diagram below illustrates the relationship between X and R^n. The space X is a topological space whose elements are of an unspecified nature, while the space R^n is just an ordinary (Euclidean) space whose elements are n- tuples of real numbers. Between these two spaces there is a local homeomorphism defined by the local coordinates of the points of X.



X, the differentiable manifold

R^n, n-dimensional number space

As a simple example, the real line R can be defined as a C^ℓ differentiable manifold by taking for $\mathfrak{D}$ the set of all C^ℓ differentiable functions on R (in the usual sense of real analysis), and defining global coordinates by $x(a) = a$, $a \in R$, $x \in R$.

Analytic manifolds can also be defined, but this requires more care. We cannot just replace C^ℓ by C^ω in the above definition, C^ω being the class of analytic functions.

4.5. COORDINATE TRANSFORMATIONS

Let $\{X, \theta, \mathfrak{D}\}$ be an n-dimensional differentiable manifold. From now on we shall for simplicity only consider differentiable manifolds of class C^∞, but the work can easily be modified to apply also for class C^ℓ, ℓ finite. Let $p \in X$, and let $\{x^a\}$, $\{x^{a'}\}$ be two sets of local coordinates around p, defined respectively on neighborhoods U, U' of p. Then since x^a, $x^{a'} \in \mathfrak{D}$, by axiom (iiib) there exist C^∞-functions F^a, $F^{a'}$ on R^n to R, for a, a' = 1,2,...,n, such that

$$x^{a'}(q) = F^{a'}(x^1(q), x^2(q), \cdots, x^n(q)) \quad \text{for } q \in U \qquad (4.1)$$

and

$$x^a(q) = F^a(x^{1'}(q), x^{2'}(q), \cdots, x^{n'}(q)) \quad \text{for } q \in U'. \qquad (4.2)$$

Now the intersection $U \cap U'$ of U and U' is not empty, as $p \in U \cap U'$, and it is open as U and U' are open. It is therefore a neighborhood of p. Let Ξ, Ξ' be respectively the image of $U \cap U'$ under the mappings $q \to x^a(q)$, $q \to x^{a'}(q)$. Since these are homeomorphisms, Ξ and Ξ' are open subsets of R^n. We see that both (4.1) and (4.2) are valid in $U \cap U'$, and hence on substituting (4.1) into (4.2) we get

$$x^a = F^a(F^{1'}(x^b), F^{2'}(x^b), \cdots F^{n'}(x^b)) \quad \text{for all } \{x^b\} \in \Xi$$

Differentiating this with respect to x^b gives

$$\delta^a_b = \left(\frac{\partial F^a}{\partial x^{c'}}\right)_{x^{c'} = F^{c'}(x^d)} \left(\frac{\partial F^{c'}}{\partial x^b}\right)_{x^d} \qquad \text{for all } \{x^d\} \in \Xi$$

and hence we see that the Jacobian matrices

$$\frac{\partial F^a}{\partial x^{b'}}, \qquad \frac{\partial F^{a'}}{\partial x^b}$$

are non-singular for $\{x^{a'}\} \in \Xi'$, $\{x^a\} \in \Xi$ respectively. These Jacobian matrices are usually written as

$$\frac{\partial x^a}{\partial x^{b'}} \quad \text{and} \quad \frac{\partial x^{a'}}{\partial x^b}$$

respectively, and the transformation in R^n mapping Ξ to Ξ' defined by

$$x^a \to x^{a'} = F^{a'}(x^b)$$

is called a <u>coordinate transformation</u> in the neighborhood $U \cap U'$ of p.

Conversely, suppose one is given a set x^a of local coordinates in the neighborhood U of $p \in X$ whose image under $q \to \{x^a(q)\} \in R^n$ is Ξ. Let $F^{a'}$, $a' = 1', 2', \ldots, n'$ be n C^∞ functions on R^n to R such that the Jacobian matrix $\partial F^{a'}/\partial x^b$ is non-singular on Ξ. Then the n functions $q \to x^{a'}(q)$ on X to R defined by

$$x^{a'}(q) = F^{a'}(x^1(q), \cdots, x^n(q))$$

are, by axiom (ii), contained in $\mathfrak{D}$, and we see that they satisfy the conditions for local coordinates on U. The transformation in R^n given by

$$x^a \to x^{a'} = F^{a'}(x^b)$$

is thus a coordinate transformation in U.

It should be noted that a coordinate transformation is a transformation in the auxiliary space R^n, not in the manifold X.

4.6. DIFFERENTIABLE MAPPINGS

Let $(X,\theta,\mathfrak{D})$ and $(Y,\mathcal{P},\mathfrak{E})$ be two differentiable manifolds, and let h be a mapping of X into Y. If f is any function defined on Y, with values in some set Z, we define the function f∘h of X into Z by

$$f\circ h(p) = f(h(p)) \quad \text{for } p \in X.$$

Then the mapping h is said to be <u>differentiable</u> if

$$g \in \mathfrak{E} \text{ implies } g\circ h \in \mathfrak{D}.$$

If h is a homeomorphism of X into Y such that both h and h^{-1} are differentiable, h is called a <u>diffeomorphism</u>. In this case both manifolds must have the <u>same dimension</u>.

4.7. PRODUCTS OF DIFFERENTIABLE MANIFOLDS

Let $(X,\theta,\mathfrak{D})$ and $(Y,\mathcal{P},\mathfrak{E})$ be two differentiable manifolds of dimensions m, n respectively. Define the projection operators π, χ on the Cartesian set product XxY by

$$X\times Y \ni (p,q) \to \pi(p,q) = p \in X$$

$$X\times Y \ni (p,q) \to \chi(p,q) = q \in Y.$$

Form the topological product space $(X\times Y, \mathcal{Q})$ of the topological spaces (X,θ) and $(Y,\mathcal{P})$ and define a set $\mathfrak{F}$ of real-valued functions on XxY as follows: The function f,

$$X\times Y \ni (p,q) \to f(p,q) \in R$$

is to be in $\mathfrak{F}$ if and only if, for any $(p,q) \in X\times Y$ and any local coordinate systems $\{x^\alpha\}$ on a neighborhood U of p and $\{y^a\}$ on a neighborhood V of q, there exists a function K on R^{m+n} of class C^∞ such that

$$f = K(x^1\circ\pi, \ldots, x^m\circ\pi\ ,\ y^1\circ\chi, \ldots\ y^n\circ\chi)$$

for all $(p,q)\in U\times V$. Then it can be shown that $(X\times Y, \mathcal{Q}, \mathcal{F})$ satisfies all the axioms of an (m+n)-dimensional differentiable manifold. It is called the product of the differentiable manifolds $(X, \theta, \mathfrak{D})$ and $(Y, \mathcal{P}, \mathcal{E})$. It can easily be generalized to the product of any finite number of differentiable manifolds.

4.8. TANGENT VECTORS

Let $(X, \theta, \mathfrak{D})$ be an n-dimensional differentiable manifold and let $p\in X$. Let $\vec{u}$ be a linear mapping of $\mathfrak{D}$ into R such that for any $f,g \in\mathfrak{D}$,

$$\vec{u}(fg) = f(p)\vec{u}(g) + g(p)\,\vec{u}(f) \tag{4.3}$$

where fg denotes the ordinary product of the functions f and g. Then $\vec{u}$ is called a tangent vector to X at p. If $\vec{u},\vec{v}$ are tangent vectors to X at p, and $a\in R$, we define the mappings $\vec{u}+\vec{v}$ and $a\vec{u}$ of $\mathfrak{D}$ to R by

$$\begin{aligned}(\vec{u}+\vec{v})(f) &= \vec{u}(f) + \vec{v}(f)\\ (a\vec{u})(f) &= a(\vec{u}(f))\end{aligned} \tag{4.4}$$

for all $f\in\mathfrak{D}$. Clearly these are linear mappings and they satisfy (4.3). Then with these definitions we see that the set of all tangent vectors at $p\in X$ forms a vector space over R, denoted by T_p.

Now let $\{x^a\}$, a = 1,2,...,n, be a system of local coordinates around p. Then $x^a\in\mathfrak{D}$. Define

$$u^a \overset{\text{def}}{=} \vec{u}(x^a) \in R \quad , \qquad x_0^a \overset{\text{def}}{=} x^a(p) \in R \tag{4.5}$$

Let $f \in \mathfrak{D}$. Then there exists a function F on R^n to R of class C^∞ such that

$$f(q) = F(x^1(q), \cdots, x^n(q)) \tag{4.6}$$

for all q in some neighborhood of p. But as F is of class C^∞, there exist n^2 functions G_{ab}, a,b = 1,2,...,n of R^n to R and of class C^∞ such that

$$F(x) = F(x_0) + (x^a - x_0^a)\left.\frac{\partial F}{\partial x^a}\right|_{x_0} + (x^a - x_0^a)(x^b - x_0^b)G_{ab}(x) \tag{4.7}$$

where we have abbreviated $x^1, \ldots, x^n$ to x. Now any constant $a \in R$ can also be regarded as a function in $\mathfrak{D}$, $p \to a \in R$ for all $p \in X$. Putting $g = a \in R$ in (4.3) gives

$$\vec{u}(af) = f(p)\vec{u}(a) + a\vec{u}(f) \quad , \quad \text{all } f \in \mathfrak{D}$$

But by the linearity of $\vec{u}$,

$$\vec{u}(af) = a\vec{u}(f)$$

and hence

$$\vec{u}(a) = 0 , \qquad a \in R . \tag{4.8}$$

Now in (4.7) remember that $x^a \in \mathfrak{D}$ so that F(x), $G_{ab}(x) \in \mathfrak{D}$, and then let $\vec{u}$ act on both sides of the equation, using (4.3), (4.5) and (4.8). We get

$$\vec{u}(F(x)) = u^a \left.\frac{\partial F}{\partial x^a}\right|_{x_0}$$

and since from (4.6), f = F(x), we have

$$\vec{u}(f) = u^a \left.\frac{\partial F}{\partial x^a}\right|_{x=x(p)} \tag{4.9}$$

Now consider the n linear maps of $\mathfrak{D}$ to R defined by

$$\mathfrak{D} \ni f \to \vec{e}_a(f) \stackrel{\text{def}}{=} \left.\frac{\partial F}{\partial x^a}\right|_{x=x(p)}, \quad a=1,2,\cdots n \tag{4.10}$$

This satisfies (4.3) and hence the e_a are tangent vectors at p. Equation (4.9) can then be written as

$$\vec{u}(f) = u^a \vec{e}_a(f) \quad \text{for all } f \in \mathfrak{D}$$

and hence

$$\vec{u} = u^a \vec{e}_a . \tag{4.11}$$

The $\vec{e}_a$ are clearly independent and so they form a basis of the tangent vector space at p, T_p. The basis $\{\vec{e}_a\}$ of T_p is called the natural basis associated with the local coordinates $\{x^a\}$. One can write

$$\vec{e}_a = \frac{\partial}{\partial x^a}$$

so that

$$\vec{e}_a(f) = \frac{\partial f}{\partial x^a} \tag{4.12}$$

but this must be interpreted in the sense of (4.10). From (4.11) we see that the real numbers u^a defined by (4.5) are the components of u with respect to the basis $\{\vec{e}_a\}$.

Now let $x^{a'}$ be new local coordinates around p. Then, as in section 4.5, there exist n C^∞-functions $F^{a'}$ on R^n to R such that

$$x^{a'}(q) = F^{a'}(x^1(q), x^2(q), \cdots x^n(q))$$

for q in some neighborhood of p. Then if $\{\vec{e}_{a'}\}$ is the natural basis of T_p associated with the local coordinates $\{x^{a'}\}$, and if $u^{a'}$ are the components of $\vec{u}$ with respect to this basis, (4.5) and (4.9) give

$$u^{a'} = \vec{u}(x^{a'}) = u^b \frac{\partial F^{a'}}{\partial x^b} . \tag{4.13}$$

As explained in section 4.5, one often writes

$$\frac{\partial F^{a'}}{\partial x^b} \quad \text{as} \quad \frac{\partial x^{a'}}{\partial x^b} .$$

Equation (4.13) can then be written as

$$u^{a'} = \frac{\partial x^{a'}}{\partial x^b} u^b \tag{4.14}$$

which is the usual definition of a vector on a differentiable manifold in terms of the transformation law of its components. From (4.10) one can easily show that

$$\vec{e}_{a'} = \frac{\partial x^b}{\partial x^{a'}} \vec{e}_b \tag{4.15}$$

which gives the transformation law of a natural basis under a change of local coordinates. From (4.14) and (4.15) we then verify the consistency of (4.11), as we see that

$$\vec{u} = u^a \vec{e}_a = u^{a'} \vec{e}_{a'}$$

A differentiable mapping

$$C: \qquad R \ni t \to p(t) \in X$$

of R to X is called a <u>differentiable curve</u> in X. Let $t_o \in R$ and let $p_o = p(t_o)$. If $f \in \mathfrak{D}$, then $f \circ p$ is a differentiable function on R. We can therefore define a mapping u of $\mathfrak{D}$ to R by

$$\vec{u}(f) \overset{\text{def}}{=} \frac{d}{dt} f \circ p(t) \Big|_{t=t_o} . \tag{4.16}$$

We see that u is linear and satisfies (4.3) for $p = p_o$. It is therefore a tangent vector to X at p_o, called the <u>tangent vector to the curve C at p_o</u>. Putting $f = x^a$ in (4.16) and using (4.5) gives

$$u^a = \frac{d}{dt} x^a (p(t)) \Big|_{t=t_0} \tag{4.17}$$

which is the usual definition of a tangent vector to a curve in terms of its components.

4.9. THE DRAGGING ALONG OF A VECTOR BY A DIFFERENTIABLE MAPPING

Let $(X, \theta, \mathfrak{D})$ and $(Y, \rho, \mathfrak{E})$ be two differentiable manifolds and let h be a differentiable mapping of X into Y. Then corresponding to a tangent vector $\vec{u}$ to X at $p \in X$ we can define a tangent vector $\overrightarrow{hu}$ to Y at $h(p) \in Y$ by

$$\overrightarrow{hu}(g) = \vec{u}(g \circ h) \tag{4.18}$$

for any $g \in \mathfrak{E}$. This is easily shown to satisfy the axioms for a tangent vector to $\mathfrak{E}$ at h(p), and the mapping $\vec{u} \to \overrightarrow{hu}$ is a linear mapping of the tangent vector space to X at p into the tangent vector space to Y at h(p). We say that $\overrightarrow{hu}$ is obtained from $\vec{u}$ by <u>dragging along</u> by the differentiable mapping h.

If X and Y have the same dimension and h is a diffeomorphism, then corresponding to any $f \in \mathfrak{D}$ we can define a function $hf \in \mathfrak{E}$ by

$$hf = f \circ h^{-1} \tag{4.19}$$

We say that hf is obtained from f by dragging along by h. In this case (4.18) can be written as

$$\overrightarrow{hu}(hf) = \vec{u}(f) \quad \text{for } f \in \mathfrak{D}.$$

4.10. VECTOR FIELDS

Let $T_p(X)$ denote the tangent vector space to the differentiable manifold $(X, \theta, \mathfrak{D})$ at $p \in X$. A vector field $\vec{u}$ on X is a correspondence that associates to every $p \in X$ an element $\vec{u}_p$ of $T_p(X)$. Then if $f \in \mathfrak{D}$, $\vec{u}(f)$ is a real-valued function on X,

$$X \ni p \to \vec{u}_p(f) \quad \in R$$

In general this function will not be differentiable, i.e. will not be in $\mathfrak{D}$. But if, for every $f \in \mathfrak{D}$, $\vec{u}(f) \in \mathfrak{D}$, then the vector field $\vec{u}$ is said to be differentiable. The set of all differentiable vector fields on X will be denoted by $\mathfrak{X}$. If $\vec{u}, \vec{v} \in \mathfrak{X}$ and $a \in R$, we define $\vec{u}+\vec{v}$ and $a\vec{u}$ as differentiable vector fields on X by

$$(\vec{u} + \vec{v})_q = \vec{u}_q + \vec{v}_q \tag{4.20}$$

$$(a\vec{u})_q = a\vec{u}_q \tag{4.21}$$

for all $q \in X$. With addition, and multiplication by a real number, so defined, $\mathfrak{X}$ is an infinite-dimensional vector space over the real field R. We also define the Lie bracket or commutator of $\vec{u}, \vec{v} \in \mathfrak{X}$ by

$$[\vec{u},\vec{v}](f) = \vec{u}(\vec{v}(f)) - \vec{v}(\vec{u}(f)) \tag{4.22}$$

for all $f \in \mathfrak{D}$. We easily show that this satisfies the axioms for a vector field (whereas $\vec{u}(\vec{v}(f))$ and $\vec{v}(\vec{u}(f))$ separately do not), and furthermore, it is a differentiable vector field, so that $\mathfrak{X}$ is closed under the Lie bracket operation. With this additional structure, $\mathfrak{X}$ forms a Lie algebra over the real field R. A Lie algebra is an algebra (defined in section 2.1) in which the internal operation usually denoted by [,] satisfies

(i) $[\vec{u},\vec{v}] = -[\vec{v},\vec{u}]$.

(ii) $[\vec{u},[\vec{v},\vec{w}]] + [\vec{v},[\vec{w},\vec{u}]] + [\vec{w},[\vec{u},\vec{v}]] = 0$.

These are easily seen to be satisfied by the definition (4.22). (ii) above is called the Jacobi identity.

The set $\mathfrak{D}$ of differentiable functions forms a ring (defined in section 2.1) with respect to the operations of addition and ordinary multiplication of functions. If $f \in \mathfrak{D}$ and $\vec{u} \in \mathfrak{X}$, we define $f\vec{u} \in \mathfrak{X}$ by

$$(f\vec{u})_p = f(p)\vec{u}_p \quad , \qquad p \in X . \tag{4.23}$$

Then $\mathfrak{X}$ forms a module (defined in section 2.1) over the ring $\mathfrak{D}$, with respect to the operations of addition defined by (4.20) and multiplication by $f \in \mathfrak{D}$, defined by (4.23).

Now let h be a diffeomorphism of X onto itself, and let $\vec{u} \in \mathfrak{X}$. Then we can drag along $\vec{u}$ by h to form a vector field $\overrightarrow{hu} \in \mathfrak{X}$ defined by

$$\overrightarrow{hu}_{h(p)}(hf) = \vec{u}_p(f) \tag{4.24}$$

for all $p \in X$ and all $f \in \mathfrak{D}$. Replacing hf by f and h(p) by p, so that f and p become respectively $h^{-1}f$ and $h^{-1}(p)$, we get

$$\overrightarrow{hu}_p(f) = \vec{u}_{h^{-1}(p)}(h^{-1}f)$$

and by (4.19) this can be written as

$$\overrightarrow{hu}_p(f) = \vec{u}_{h^{-1}(p)}(f \circ h)$$

If $\overrightarrow{hu} = \vec{u}$, the vector field $\vec{u}$ is said to be <u>invariant</u> with respect to h.

4.11. DIFFERENTIAL FORMS

Let $T_p(X)$ be the tangent vector space at p to the differentiable manifold $X \ni p$, and let $T_p{}^*(X)$ be the dual space to $T_p(X)$. The elements of $T_p{}^*(X)$ are called <u>covectors</u> or <u>differential forms</u> at p. A field ω of covectors or differential forms on X is a correspondence

$$X \ni p \to \omega_p \in T_p^*(X)$$

which associates with every element p of X a differential form ω_p at p. If ω is a field of differential forms and $\vec{u}_p \in T_p(X)$, then $\omega_p(\vec{u}_p) \in R$. Hence if $\vec{u}$ is a vector field on X, $\omega(\vec{u})$ is a real-valued function on X defined by

$$\omega(\vec{u})(p) = \omega_p(\vec{u}_p) \quad \text{for } p \in X.$$

If $\omega(\vec{u}) \in \mathfrak{D}$ for all $\vec{u} \in \mathfrak{X}$, ω is said to be <u>differentiable</u>, and is called a <u>differentiable form field</u> on X. If ω, π are differentiable form fields on X and $a \in R$, we define $\omega + \pi$ and $a\omega$ as differentiable form fields on X by

$$(\omega + \pi)_p = \omega_p + \pi_p, \tag{4.25}$$

$$(a\omega)_p = a\omega_p. \tag{4.26}$$

With these definitions of addition and of multiplication by a real number, the set of all differentiable form fields on X forms an infinite-dimensional vector space over R.

Now take any $f \in \mathfrak{D}$ and define a mapping of $T_p(X)$ to R by

$$T_p(X) \ni \vec{u}_p \to \vec{u}_p(f) \in R.$$

Then by (4.4) this is a linear mapping, and so it determines a differential form at p, denoted by df. We then have

$$df(\vec{u}_p) = \vec{u}_p(f). \tag{4.27}$$

This mapping is defined for all p, and so df is in fact a field of differential forms on X. Further, if $\vec{u} \in \mathfrak{X}$, then $df(\vec{u}) = \vec{u}(f) \in \mathfrak{D}$, and hence df is a differentiable form field on X. So with every differentiable function f on X there is associated a differentiable form field df defined by (4.27).

If $f, g \in \mathfrak{D}$ and $\vec{u} \in \mathfrak{X}$, then using (4.27), (4.3) can be written in the form

$$d(fg)(\vec{u}) = g\,df(\vec{u}) + f\,dg(\vec{u})$$

and this holds for all vector fields $\vec{u}$, so that

$$d(fg) = f.\,dg + g.\,df \qquad (4.28)$$

Now let $p \in X$ and let $\{x^a\}$ be a system of local coordinates around p. Let us use the notation of section 4.8. Then since the $x^a \in \mathcal{D}$, if $\vec{u} \in T_p(x)$ then (4.27) gives

$$\begin{aligned} dx^a(\vec{u}) &= \vec{u}(x^a) \\ &= u^a \quad \text{by (4.5).} \end{aligned} \qquad (4.29)$$

But by (4.11),

$$\vec{u} = u^a \vec{e}_a$$

and so

$$dx^a(\vec{u}) = u^b\, dx^a(\vec{e}_b).$$

Comparing with (4.29) this gives

$$dx^a(\vec{e}_b) = \delta^a_b$$

and so $\{dx^a\}$ is the dual basis of the natural basis $\{\vec{e}_a\}$ associated with the local coordinates $\{x^a\}$. If we write, as is often done, as explained in section 4.8, $\partial/\partial x^a$ for $\vec{e}_a$, then we have that:

$\{dx^a\}$ is the basis dual to $\{\partial/\partial x^a\}$.

So any differential form ω at p can be written .

$$\omega = \omega_a\, dx^a \qquad \text{with} \qquad \omega_a = \omega(\vec{e}_a)$$

In particular take $\omega = df$, $f \in \mathcal{D}$. Then

$$\omega_a = df(\vec{e}_a) = \vec{e}_a(f) = \frac{\partial f}{\partial x^a} \qquad \text{by (4.12).}$$

Hence

$$df = \frac{\partial f}{\partial x^a}\, dx^a . \tag{4.30}$$

Let h be a diffeomorphism of X and let ω be a differentiable form field on X. Then we define the differentiable form field $h\omega$ obtained from ω by dragging along by h by

$$(h\omega)_{h(p)}(\overrightarrow{hu}) = \omega_p(\vec{u}) \tag{4.31}$$

for all $p \in X$ and all $\vec{u} \in \mathfrak{X}$. This is easily seen to satisfy all the axioms of a differentiable form field. For the particular case of $\omega = df$, $f \in \mathfrak{D}$, it gives

$$\begin{aligned} hdf(\overrightarrow{hu}) &= df(\vec{u}) = \vec{u}(f) \\ &= \overrightarrow{hu}\,(hf) \qquad \text{from section 4.9} \\ &= dhf\,(\overrightarrow{hu}). \end{aligned}$$

But $\overrightarrow{hu}$ is arbitrary, and hence

$$hdf = dhf .$$

4.12. TANGENT TENSORS AND TENSOR FIELDS

Let $T_p(X)$ be the tangent vector space at $p \in X$ to the n-dimensional differentiable manifold X, and let $T_p{}^*(X)$ be its dual space. Then if we drop the X in $T_p(X)$ for simplicity, a <u>tangent tensor</u> to X at p of valence (k,ℓ) is an element of the tensor product

$$(\otimes^k T_p) \otimes (\otimes^\ell T_p^*).$$

A <u>tensor field</u> on X of valence (k,ℓ) is a correspondence that assigns to every point $p \in X$ a tangent tensor to X at p of valence (k,ℓ).

Let $\{x^a\}$ be a system of local coordinates around p.

Then if $\{\vec{e}_a\}$ is the natural basis at p of T_p associated with these local coordinates, and $\{dx^a\}$ is the dual basis, then any tangent tensor S of valence (k, ℓ) at p can be written uniquely in the form

$$S = S^{a_1 \cdots a_k}{}_{b_1 \cdots b_\ell}\, \vec{e}_{a_1} \otimes \cdots \otimes \vec{e}_{a_k} \otimes dx^{b_1} \otimes \cdots \otimes dx^{b_\ell}$$

the scalars $S^{a_1 \cdot\cdot a_k}{}_{b_1 \cdot\cdot b_\ell}$ are called the natural components of S with respect to the coordinates $\{x^a\}$.

If S is a tensor field, the components $S^{a_1 \cdot\cdot a_k}{}_{b_1 \cdot\cdot b_\ell}$ are real-valued functions on that neighborhood u of p for which the $\{x^a\}$ are local coordinates. If there exist $n^{(k+\ell)}$ functions $f^{a_1 \cdots a_k}{}_{b_1 \cdots b_\ell} \in \mathcal{D}$ such that

$$S^{a_1 \cdots a_k}{}_{b_1 \cdots b_\ell}(q) = f^{a_1 \cdots a_k}{}_{b_1 \cdots b_\ell}(q)$$

for all $q \in U$, S is said to be differentiable on U. If every point $p \in X$ is contained in a neighborhood on which S is differentiable, then S is said to be a differentiable tensor field. It is easily seen that this definition agrees with those given above of differentiable vector fields ($k = 1$, $\ell = 0$) and differentiable form fields ($k = 0$, $\ell = 1$).

If h is a diffeomorphism of X, the differentiable tensor field hS obtained from S by dragging along by h is defined by

$$hS = S^{a_1 \cdots a_k}{}_{b_1 \cdots b_\ell} \circ h^{-1} \cdot h\vec{e}_{a_1} \otimes \cdots \otimes h\vec{e}_{a_k} \otimes h\,dx^{b_1} \otimes \cdots \otimes h\,dx^{b_\ell}$$

This is also easily seen to agree with the above definitions of dragging along of differentiable vector and form fields.

4.13. GEOMETRIC OBJECTS

Let X be an n-dimensional differentiable manifold. A vector field $\vec{v}$ on X may be considered as a correspondence which associates to every point $p \in X$ and every system of local coordinates $\{x^a\}$ around p, n real numbers $(u^1, u^2, \ldots, u^n)$

written

$$\vec{u}: \quad (p, \{x^a\}) \rightarrow (u^1, u^2, \cdots u^n) \in R^n$$

such that if also $\{x^{a'}\}$ is another system of local coordinates around p, and

$$\vec{u}: \quad (p, \{x^{a'}\}) \rightarrow (u^{1'}, u^{2'} \cdots u^{n'}) \in R^n$$

then

$$u^{a'} = \left.\frac{\partial x^{a'}}{\partial x^b}\right|_p u^b. \tag{4.32}$$

A definition of this type can clearly be given for tensors of any valence. But since tensors are not sufficient for all purposes in geometry and physics, for example scalar densities are not tensors, to avoid having to expand definitions and theorems whenever we need a new type of entity, it is convenient to define a more general entity, the geometric object, which includes nearly all the entities needed in geometry and physics, so that definitions and theorems can be given in terms of geometric objects so as to hold for all the more specialized cases that we may require.

Let $p \in X$ be an arbitrary point of X and let $\{x^a\}$, $\{x^{a'}\}$ be two systems of local coordinates around p. A geometric object field y is a correspondence

$$y: \quad (p, \{x^a\}) \longrightarrow (y_1, y_2, \cdots y_N) \in R^N$$

which associates with every point $p \in X$, and every system of local coordinates $\{x^a\}$ around p, a set of N real numbers, together with a rule which determines $(y_{1'}, \ldots, y_{N'})$, given by

$$y: \quad (p, \{x^{a'}\}) \longrightarrow (y_{1'}, \cdots y_{N'})$$

in terms of the $(y_1, y_2, \ldots, y_N)$ and the values of p of the functions and their partial derivatives which relate the coordinate systems $\{x^a\}$ and $\{x^{a'}\}$ as in (4.1). The N numbers $(y_1, \ldots, y_N)$ are called the components of y at p with respect to the coordinates $\{x^a\}$. When y is a one-component geometric object, we write y and y' to denote its values corresponding to $\{x^a\}$ and $\{x^{a'}\}$, respectively.

The particular type of geometric object is determined by this rule. For example, for a vector this rule has the form (4.32).

At first it may appear that this definition is so general that there can be nothing of geometric significance which is not a geometric object. However, this is not so. Both parts of the definition may break down, e.g.:

(i) spinors are not geometric objects as components of a spinor can only be associated with an orthonormal basis and not with the natural basis of a general coordinate system, so that in this case the correspondence does not exist.

(ii) Let u be a differentiable vector field with components u^a with respect to the natural basis associated with the local coordinate system x^a. Then with p and x^a we may associate the n^2 numbers $(\partial u^a/\partial x^b)p$. Under a change of coordinate system these transform thus:

$$\frac{\partial u^{a'}}{\partial x^{b'}} = \frac{\partial x^{a'}}{\partial x^c}\frac{\partial x^d}{\partial x^{b'}}\frac{\partial u^c}{\partial x^d} + \frac{\partial x^d}{\partial x^{b'}}\frac{\partial^2 x^{a'}}{\partial x^c \partial x^d}u^c$$

This transformation law involves the components u^c of the original vector, and hence the n^2 numbers $\partial u^a/\partial x^b$ do not form the components of a geometric object as the transformation law has an unallowed form. However, the $n^2 + n$ numbers

$$\left(u^a , \frac{\partial u^a}{\partial x^b} \right)$$

do form the components of a geometric object.

If y is a geometric object, and h is a diffeomorphism of X, then we define the geometric object hy obtained from y by dragging along by h by

if $\quad y: (p, \{x^a\}) \to (y_1, y_2, \cdots y_N)$

then $\quad hy: (h(p), \{hx^a\}) \to (y_1, y_2, \cdots y_N)$

so that hy associates with h(p) and $\{hx^a\}$ the <u>same numbers</u> as y associates with p and $\{x^a\}$. We write

$$hy: \quad (p, \{x^a\}) \rightarrow (hy_1, \dots hy_N)$$

so defining hy_A, $A = 1,2,\dots,N$.

This definition is easily seen to be consistent with the particular case of a tensor field defined above.

4.14. ONE-PARAMETER GROUPS OF DIFFEOMORPHISMS

Let $(X, \theta, \mathfrak{D})$ be a differentiable manifold. A family h_t, $t \in R$, of mappings of X onto X is said to be a <u>one-parameter group G_1 of diffeomorphisms of X</u> if

(i) For every $t \in R$ and $p \in X$, $p \rightarrow h_t(p)$ is a diffeomorphism of X onto X

(ii) The mapping

$$R \times X \ni (t, p) \rightarrow h_t(p) \in X$$

of the product manifold RxX onto X is differentiable

(iii) $h_{t+s} = h_t \circ h_s$ for all $t, s \in R$.

From (iii) we see that the h_t form an abelian group under the functional combination operation $\circ$, whose identity is h_o and such that the inverse of h_t is h_{-t}.

Through any point $p \in X$ we can define a differentiable curve $p(t)$ by

$$p(t) = h_t(p).$$

This curve is called the <u>orbit</u> of p generated by the group, and the set of orbits of <u>all</u> the points of X forms the <u>trajectories</u> of the G_1. We see that from (iii) there is exactly one trajectory through each point of X.

The tangent vector field $\vec{u}$ to the trajectories of the group is defined by

$$\vec{u}_p(f) = \frac{d}{dt} f(h_t(p))\Big|_{t=0} \qquad (4.33)$$

It is a differentiable vector field and is called the vector field <u>induced</u> by the G_1.

One may pose the inverse problem, namely given a differentiable vector field $\vec{u}$ on X, is there a one-parameter group G_1 of diffeomorphisms which induces $\vec{u}$? The answer is that in general such a G_1 will not exist, but instead there will exist a <u>local</u> one-parameter group of diffeomorphisms which induces $\vec{u}$. A local G_1 means that h_t will not everywhere be defined for all t; instead, for any $p \in X$ there is a neighborhood U of p, a number $\varepsilon > 0$ and a family $\{h_t, |t| < \varepsilon\}$ of mappings of X into X such that

(i) h_t is a diffeomorphism of U onto $h_t(u)$ for any t for which $|t| < \varepsilon$

(ii) the mapping $(t,p) \to h_t(p)$ of the product manifold IxU into X is differentiable, where I is the interval $(-\varepsilon, +\varepsilon)$ of the real line,

(iii) if $|t|$, $|s|$ and $|t+s| < \varepsilon$ and $q \in U$, then

$$h_t \circ h_s = h_{t+s} .$$

The existence of such a local G_1 follows from the theory of ordinary differential equations. Let $\{x^a\}$ be local coordinates around $p \in X$. Then if a local G_1 exists which induces $\vec{u}$, and we define

$$p(t) \stackrel{def}{=} h_t(p)$$

and let $\phi^a(t)$ be the local coordinates of p(t), then by putting x^a for f in (4.33), it can be written in the form

$$\frac{d\phi^a}{dt} = u^a(\phi^b)$$

where $u^a(\phi^b)$ denotes the natural components of $\vec{u}$ at p(t) with respect to the local coordinates $\{x^a\}$. The existence of the local G_1, called the local G_1 <u>generated</u> by $\vec{u}$, can then be shown to follow from the existence theorem on solutions of systems of ordinary first-order differential equations.

4.15. LIE DERIVATIVES

Let $\vec{u}$ be a differentiable vector field on a differentiable manifold X, and let h_t be the local one-parameter group of diffeomorphisms of X generated by $\vec{u}$. If s is a differentiable tensor field on X, we define the Lie derivative of s with respect to $\vec{u}$, written $\pounds_{\vec{u}} s$, by

$$\pounds_{\vec{u}} s \overset{\text{def}}{=} -\frac{d}{dt} h_t s \Big|_{t=0} \tag{4.34}$$

and it is a tensor of the same valence as s.

If y is a geometric object field on X with components y_A, then the Lie derivative of y is a geometric object field on X with the same number of components as y, defined by

$$(\pounds_{\vec{u}} y)_A = -\frac{d}{dt} h_t y_A \Big|_{t=0}$$

but it is not necessarily of the same type as y. For example, the Lie derivative of an affine connection (defined below) is a tensor.

We leave it as an exercise to show that for a scalar field $f \in \mathfrak{D}$ and a vector field $\vec{v} \in \mathfrak{X}$

$$\pounds_{\vec{u}} f = \vec{u}(f)$$

and

$$\pounds_{\vec{u}} \vec{v} = [\vec{u}, \vec{v}] .$$

4.16. AFFINE CONNECTIONS ON A DIFFERENTIABLE MANIFOLD

Let $(X, \theta, \mathfrak{D})$ be an n-dimensional differentiable manifold. An affine connection at a point $p \in X$ is a mapping

$$T_p(X) \times \mathcal{X} \ni (\vec{u}, \vec{v}) \to \vec{u} \cdot \nabla \vec{v} \in T_p(X)$$

of the Cartesian product $T_p(X) \times \mathcal{X}$ into the tangent vector space $T_p(X)$ at p, which satisfies

(i) Linearity in both arguments with respect to addition and multiplication by a real number, i.e. if $\vec{u}_1, \vec{u}_2, \vec{u} \in T_p(X)$, $\vec{v}_1, \vec{v}_2, \vec{v} \in \mathcal{X}$, $a, b \in R$, then

$$(a\vec{u}_1 + b\vec{u}_2) \cdot \nabla \vec{v} = a(\vec{u}_1 \cdot \nabla \vec{v}) + b(\vec{u}_2 \cdot \nabla \vec{v})$$

$$\vec{u} \cdot \nabla (a\vec{v}_1 + b\vec{v}_2) = a(\vec{u} \cdot \nabla \vec{v}_1) + b(\vec{u} \cdot \nabla \vec{v}_2)$$

(ii) If $u \in T_p(X)$, $\vec{v} \in \mathcal{X}$ and $f \in \mathfrak{D}$, then

$$\vec{u} \cdot \nabla f \vec{v} = f(p) \vec{u} \cdot \nabla \vec{v} + \vec{u}(f) \vec{v}_p$$

An affine connection on X is said to be given if an affine connection is given at every point of X. Then if ∇ is an affine connection on X and $\vec{u}, \vec{v} \in \mathcal{X}$, $\vec{u} \cdot \nabla \vec{v}$ is a vector field on X defined by

$$(\vec{u} \cdot \nabla \vec{v})_p = \vec{u}_p \cdot \nabla \vec{v} \quad \text{for } p \in X.$$

This affine connection is said to be <u>differentiable</u> if $\vec{u} \cdot \nabla \vec{v} \in \mathcal{X}$ for all $\vec{u}, \vec{v} \in \mathcal{X}$.

Now consider fixed $p \in X$ and fixed $\vec{v} \in \mathcal{X}$. Then the mapping

$$T_p(X) \ni \vec{u} \to \vec{u} \cdot \nabla \vec{v} \in T_p(X)$$

is linear, and hence it defines, as explained in section 2.7, a tensor of valence (1,1), an element of $T_p{}^*(X) \otimes T_p(X)$, denoted by $(\nabla \vec{v})_p$. This tensor is called the <u>covariant derivative</u> of $\vec{v}$ at $p \in X$. It is defined at all $p \in X$, and hence $\nabla \vec{v}$ is a tensor field on X. If the affine connection is differentiable, we see that $\nabla \vec{v}$ is a differentiable tensor field.

Now let $\vec{e}_a$, a = 1,2,...,n be n differentiable vector fields on X such that $\{\vec{e}_a\}$ forms a basis of $T_q(X)$ for q in some neighborhood U of p. Then the numbers Γ^c_{ab} defined on U by

$$\vec{e}_a \cdot \nabla \vec{e}_b = \Gamma^c_{ab} \vec{e}_c \qquad (4.35)$$

are called the <u>components (or coefficients) of the affine connection</u> in U with respect to the basis $\{\vec{e}_a\}$.

Let $\vec{u}, \vec{v} \in \mathfrak{X}$ and let

$$\vec{u} = u^a \vec{e}_a \quad , \qquad \vec{v} = v^a \vec{e}_a$$

in U. Then from the axioms for the affine connection we have

$$\begin{aligned}\vec{u} \cdot \nabla \vec{v} &= u^a \vec{e}_a \cdot \nabla (v^b \vec{e}_b) \\ &= u^a \vec{e}_a (v^b) \vec{e}_b + u^a v^b \vec{e}_a \cdot \nabla \vec{e}_b \\ &= u^a \vec{e}_b \{ \vec{e}_a (v^b) + \Gamma^b_{ac} v^c \} \quad \text{by (4.35).}\end{aligned}$$

Write

$$\nabla_a v^b \overset{\text{def}}{=} \vec{e}_a (v^b) + \Gamma^b_{ac} v^c \qquad (4.36)$$

Then

$$\vec{u} \cdot \nabla \vec{v} = u^a \vec{e}_b \cdot \nabla_a v^b$$

and hence we see that the numbers $\nabla_a v^b$ are the components of the covariant derivative of $\vec{v}$ with respect to the basis $\{\vec{e}_a\}$.

If the $\{\vec{e}_a\}$ form the natural basis associated with the local coordinate system $\{x^a\}$ in U, then (4.12) enables us to re-write (4.36) as

$$\nabla_a v^b = \frac{\partial v^b}{\partial x^a} + \Gamma^b_{ac} v^c \qquad (4.37)$$

which is the usual definition of the covariant derivative in terms of its components.

If ω is a differentiable form field on X, then we define the covariant derivative $\nabla\omega$ of ω at p as a tensor of valence (0,2) such that if $\vec{u},\vec{v} \in T_p(X)$ then

$$\nabla\omega(\vec{u},\vec{v}) \overset{def}{=} \vec{u}(\omega(\vec{v})) - \omega(\vec{u}\cdot\nabla\vec{v}).$$

The components $\nabla_a\omega_b$ of $\nabla\omega$ with respect to the basis $\{\vec{e}_a\}$ are then given by

$$\nabla_a\omega_b = \vec{e}_a(\omega_b) - \Gamma_{ab}{}^c\,\omega_c$$

where $\omega_a = \omega(\vec{e}_a)$, or if the $\{\vec{e}_a\}$ form the natural basis associated with the local coordinate system $\{x^a\}$, thus

$$\nabla_a\omega_b = \frac{\partial\omega_b}{\partial x^a} - \Gamma_{ab}^{\;\;c}\,\omega_c\,. \tag{4.38}$$

Covariant differentiation can now easily be extended to a general tensor field. Let s be a differentiable tensor field of valence (k,ℓ). Then if $\vec{u} \in T_p(X)$, the affine connection at p applied to s, written $\vec{u}\cdot\nabla s$, is defined to be a tensor at p of valence (k,ℓ) such that

(i) it is linear in $\vec{u}$, i.e. if $a,b \in R$, $\vec{u},\vec{v} \in T_p(X)$, then

$$(a\vec{u}+b\vec{v})\cdot\nabla s = a(\vec{u}\cdot\nabla s) + b(\vec{v}\cdot\nabla s)$$

(ii) it is linear in s, i.e. if s,t are differentiable tensor fields of valence (k,ℓ) and $a,b \in R$, then

$$\vec{u}\cdot\nabla(as+bt) = a\vec{u}\cdot\nabla s + b\vec{u}\cdot\nabla t$$

(iii) it obeys the Leibnitz rule with respect to tensor products, i.e. if s,t are differentiable tensor fields, not necessarily of the same valence, then

$$\vec{u}\cdot\nabla(s\otimes t) = (\vec{u}\cdot\nabla s)\otimes t + s\otimes(\vec{u}\cdot\nabla t)$$

(iv) if $f \in \mathfrak{D}$, we define

$$\vec{u}\cdot\nabla f = \vec{u}(f).$$

Since any tensor can be expressed as the sum of tensor products of vectors and differential forms, we see that these axioms completely determine $\vec{u}\cdot\nabla s$ for any tensor s. Since for fixed s this is linear in $\vec{u}$, it defines a tensor of valence $(k, \ell+1)$, called the <u>covariant derivative</u> of s at p, denoted by ∇s.

If

$$R \ni t \to p(t) \in X$$

is a differentiable curve C in X, and $\vec{u}$ is its tangent vector field defined by (4.16), then if $\vec{v} \in \mathcal{X}$ we define the <u>absolute derivative</u> of $\vec{v}$ along the curve, written $D\vec{v}/dt$, by

$$\frac{D\vec{v}}{dt} \overset{\text{def}}{=} \vec{u}\cdot\nabla\vec{v}$$

If $D\vec{v}/dt = 0$ along C, we say that $\vec{v}$ is <u>parallel</u> along C, and the value of $\vec{v}$ at one point on C determines the value at any other point on C by <u>parallel transport</u>. If C is such that $D\vec{u}/dt = 0$, C is said to be <u>geodesic</u> and t is called an <u>affine parameter</u> on C. If we change to a new parameter λ, $\lambda = \lambda(t)$ being a monotonic differentiable function of t, the equation of the geodesic will take the form

$$\frac{D\vec{u}}{d\lambda} = f(\lambda)\vec{u}$$

for some function $f(\lambda)$. λ will also be an affine parameter, i.e. $f(\lambda) = 0$, if and only if $\lambda = at + b$, a,b numbers, $a \neq 0$. If $\{x^a\}$ is a system of local coordinates, the geodesic equations may be written in the form

$$\frac{d^2x^a}{dt^2} + \Gamma^a_{bc}\frac{dx^b}{dt}\frac{dx^c}{dt} = 0 \qquad (4.39)$$

obtained by using (4.37) and remembering that if u^a are the natural components of $\vec{u}$, $u^a = dx^a/dt$.

4.17. THE TORSION TENSOR

Let X by a differentiable manifold provided with an affine connection ∇. This affine connection is said to be <u>symmetric</u> if, for every $\vec{u},\vec{v} \in \mathcal{X}$,

$$\vec{u}\cdot\nabla\vec{v} - \vec{v}\cdot\nabla\vec{u} = [\vec{u},\vec{v}]. \tag{4.40}$$

If $\{\vec{e}_a\}$ is the natural basis associated with the local coordinate system $\{x^a\}$, then

$$[\vec{e}_a, \vec{e}_b] = 0$$

and by (4.35), (4.40) becomes

$$\Gamma^c_{ab} = \Gamma^c_{ba} \tag{4.41}$$

which is the origin of the name 'symmetric' for this property. Note, however, that (4.41) only holds for a natural basis, not for a general basis.

If ∇ is a non-symmetric affine connection, we may define from it another affine connection ∇' by

$$\vec{u}\cdot\nabla'\vec{v} \overset{\text{def}}{=} \tfrac{1}{2}\left(\vec{u}\cdot\nabla\vec{v} + \vec{v}\cdot\nabla\vec{u} + [\vec{u},\vec{v}]\right).$$

This is easily verified to satisfy the axioms of an affine connection, and it is symmetric. It is called the symmetric part of ∇. If ${}'\Gamma^c_{ab}$ are the natural components of ∇' with respect to the local coordinates $\{x^a\}$, then we see that

$${}'\Gamma^c_{ab} = \tfrac{1}{2}\left(\Gamma^c_{ab} + \Gamma^c_{ba}\right).$$

From (4.39) we then see that the geodesics of both ∇ and ∇' are the same, as the equation of a geodesic only involves the symmetric part of the affine connection.

We can also define from ∇ a tensor field S of valence (1,2) by

$$\mathcal{X}\times\mathcal{X} \ni (\vec{u},\vec{v}) \to S(\vec{u},\vec{v}) = \tfrac{1}{2}\left(\vec{u}\cdot\nabla\vec{v} - \vec{v}\cdot\nabla\vec{u} - [\vec{u},\vec{v}]\right)$$

This is easily seen to satisfy the axioms of a tensor field, i.e. it is linear in both arguments with respect to addition and multiplication by a real-valued function on X. In terms of natural components, if we write

$$S(\vec{u},\vec{v}) = S_{ab}^{\ \ c}\, u^a v^b \vec{e}_c$$

then

$$S_{ab}^{\ \ c} = \tfrac{1}{2}\left(\Gamma_{ab}^{\ \ c} - \Gamma_{ba}^{\ \ c}\right).$$

S is called the <u>torsion tensor</u> of the affine connection, and its vanishing is the condition for the affine connection to be symmetric.

4.18. THE CURVATURE TENSOR

Let X be a differentiable manifold provided with a differentiable affine connection ∇, and let $\vec{u},\vec{v},\vec{w} \in \mathcal{X}$. Define the mapping

$$\mathcal{X}\times\mathcal{X}\times\mathcal{X} \ni (\vec{u},\vec{v},\vec{w}) \to R(\vec{u},\vec{v},\vec{w})$$

$$= \vec{w}\cdot\nabla(\vec{v}\cdot\nabla\vec{u}) - \vec{v}\cdot\nabla(\vec{w}\cdot\nabla\vec{u}) + (\vec{v}\cdot\nabla\vec{w})\cdot\nabla\vec{u} - (\vec{w}\cdot\nabla\vec{v})\cdot\nabla\vec{u} \in \mathcal{X}$$

Then one can easily show that this mapping is linear in each argument vector with respect to addition and multiplication by a differentiable function, i.e. if f, $g \in \mathfrak{D}$ and $\vec{t},\vec{u},\vec{v},\vec{w} \in \mathcal{X}$, then

$$R(f\vec{t} + g\vec{u}, \vec{v}, \vec{w}) = f\,R(\vec{t},\vec{v},\vec{w}) + g\,R(\vec{u},\vec{v},\vec{w})$$

and similarly for the other two arguments. It therefore determines a differentiable tensor field R of valence (1,3), called the <u>Riemann-Christoffel tensor field</u>, or the <u>curvature tensor field</u>, of the affine connection. If $\{\vec{e}_a\}$ is a basis of $T_p(X)$ and u^a, v^a, w^a are the components of $\vec{u}_p$, $\vec{v}_p$, $\vec{w}_p$ respectively with respect to this basis, then we write

$$R_p(\vec{u}, \vec{v}, \vec{w}) = \vec{e}_a R^a{}_{bcd} u^b v^c w^d$$

and the $R^a{}_{bcd}$ are the components of the curvature tensor. If $\Gamma_{ab}^{\ c}$ are the components of the affine connection and $\{\vec{e}_a\}$ is the natural basis associated with the local coordinate system $\{x^a\}$ around p, then

$$R^a{}_{bcd} = \partial_d \Gamma^a_{cb} - \partial_c \Gamma^a_{db} + \Gamma^a_{de}\Gamma^e_{cb} - \Gamma^a_{ce}\Gamma^e_{db} \tag{4.42}$$

where

$$\partial_a \equiv \partial/\partial x^a .$$

It can be shown that in any neighborhood U there exists a local coordinate system for which the natural components of the affine connection vanish, $\Gamma_{ab}^{\ c} = 0$, if and only if $R^a{}_{bcd} = 0$ in U.

4.19. METRIC TENSOR FIELDS. RIEMANNIAN GEOMETRY

The differentiable manifold X is said to be provided with a metric tensor field g if a scalar product

$$T_p(X) \times T_p(X) \ni (\vec{u}, \vec{v}) \to \vec{u}\cdot\vec{v} = g(\vec{u}, \vec{v}) \in R$$

is defined in the tangent space at every point $p \in X$. g is a tensor field of valence (0,2) and it is said to be differentiable if $g(\vec{u}, \vec{v}) \in \mathcal{D}$ whenever $\vec{u}, \vec{v} \in \mathcal{X}$.

An affine connection ∇ on X is called metric if for any differentiable curve $t \to p(t)$ in X and any vector field $\vec{v}$ which is parallel along this curve.

$$\frac{d}{dt} g(\vec{v}, \vec{v}) = 0$$

i.e. if parallel vector fields along a curve have constant length. It can easily be shown that given a non-singular differentiable metric tensor field on a differentiable manifold X, there exists exactly one symmetric, metric, affine connection on X, and that in natural components with respect to a local coordinate system $\{x^a\}$, it is given by

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad}(\partial_b g_{cd} + \partial_c g_{bd} - \partial_d g_{bc}) \qquad (4.43)$$

where the g^{ad} are the components of the contravariant metric tensor,

$$g^{ad} g_{bd} = \delta^a_b$$

as in section 2.12.

A differentiable manifold provided with a non-singular differentiable metric tensor field and a symmetric, metric, affine connection is called a Riemannian space, and the above theorem is known as the Fundamental Theorem of Riemannian Geometry.

An isometry or a motion of a Riemannian space is a diffeomorphism h of X such that for every $p \in X$ and every pair of vectors $\vec{u}, \vec{v} \in T_p(X)$,

$$h\vec{u} \cdot h\vec{v} = \vec{u} \cdot \vec{v} \qquad (4.44)$$

i.e. scalar products of vectors are unchanged when the vectors are dragged along by h. From the definition of dragging along we have

$$\vec{u} \cdot \vec{v} = g(\vec{u}, \vec{v}) = hg(\overrightarrow{hu}, \overrightarrow{hv})$$

and hence with (4.44) this gives

$$g(\overrightarrow{hu}, \overrightarrow{hv}) = hg(\overrightarrow{hu}, \overrightarrow{hv})$$

Since this must hold for all $\vec{u}, \vec{v}$ we have as the condition for an isometry

$$hg = g \qquad (4.45)$$

i.e. the metric must be invariant under dragging along by h. If h_t, $t \in R$, is a one-parameter group of isometries, then (4.45) gives, with (4.34),

$$\underset{\vec{u}}{£}\, g = 0 \tag{4.46}$$

where $\vec{u}$ is the vector field induced by the group. Equation (4.46) is known as Killing's equation, and in terms of components it takes the form

$$\underset{\vec{u}}{£}\, g_{ab} \equiv \nabla_a u_b + \nabla_b u_a = 0$$

where $u_a = g_{ab} u^b$ and u^b are the components of $\vec{u}$.

4.20. INTEGRATION IN A DIFFERENTIABLE MANIFOLD

Let $(X, \theta, \mathfrak{D})$ be an n-dimensional differentiable manifold and let E be a geometric object field defined on X with one component,

$$E: \quad (p, \{x^a\}) \rightarrow E(x^a) \quad \in R.$$

Let $U \in \theta$ be an open set of X on which two local coordinate systems $\{x^a\}$ and $\{x^{a'}\}$ are defined. Then these coordinate systems map U into open sets of R^n, say x(U), x'(U) respectively, and $E(x^a)$, $E'(x^{a'})$ are real-valued functions which may be considered as defined on x(U), x'(U) respectively. We may therefore form the integrals

$$\int_{x(U)} E(x^a)\, d^n x, \quad \int_{x'(U)} E'(x^{a'})\, d^n x'. \tag{4.47}$$

Now if J is the Jacobian of the coordinate transformation from $\{x^a\}$ to $\{x^{a'}\}$,

$$J = \det \frac{\partial x^{a'}}{\partial x^b}$$

then

$$\int_{x'(u)} E'(x^{a'})\, d^n x' = \int_{x(u)} E'(x^{a'})\, |J|\, d^n x .$$

So the two integrals in (4.47) are equal for all possible choices of U if and only if

$$E(x^a) = E'(x^{a'})\,|J| \qquad (4.48)$$

From section 2.10 we see that such a geometric object is a pseudoscalar density. In this case, since $\int_{x(u)} E(x^a) d^n x$ is independent of the local coordinate system chosen in U, we may consistently call it the integral of E over U.

Now let Y be any subset of X. Then we can subdivide it into subsets of X each of which can be covered by a single local coordinate system. Then the integral of E over each subset can be defined as above, and the integral of E over Y is defined to be the sum of the integrals over these subsets of Y.

We have thus shown how to integrate a pseudoscalar density field over an arbitrary region of a differentiable manifold by reducing it to integration over a Euclidean space, which we know how to do.

The region Y of X is said to be <u>oriented</u> if

(i) the tangent space at each point of Y has an orientation, and

(ii) if $\{x^a\}$ is a local coordinate system in $U \subset Y$, then the corresponding natural bases of the tangent spaces at the points of U all have the same orientation. We then may call this orientation the orientation of the local coordinates $\{x^a\}$.

Since the Jacobian connecting two local coordinate systems of the same orientation is positive, if we restrict ourselves to local coordinate systems with positive orientation, (4.48) becomes

$$E(x^a) = E'(x^{a'})\, J$$

which is satisfied by scalar densities also. Now, proceeding as before, we see that one can define the integral of a scalar density over an oriented region of a differentiable manifold. We should note, however, that it is not always possible to provide a manifold <u>globally</u> with an orientation.

Finally, we show that a field of differential n-forms can be integrated over an oriented region of an n-dimensional differentiable manifold. For if A is a field of differential n-forms and $\{x^a\}$ a local coordinate system in $U \subset X$, then with A we can associate a one-component geometric object E by

$$E: \quad (p, \{x^a\}) \to E(x^a)$$

$$A = E(x^a)\, dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$$

Using the results of section 2.8 we easily see that E is a scalar density, and the integral of A over any region is defined to be the integral of E over that region as defined above.

The exposition of elements of differential geometry given in this chapter follows rather closely that of K. Nomizu, Lie Groups and Differential Geometry, Publ. Math. Soc. Japan, No. 2, 1956.

References:

N. Bourbaki, Topologie Générale. Hermann and Cie, Paris.

K. Nomizu, Lie Groups and Differential Geometry. Math. Soc. of Japan (1956).

J. A. Schouten, Ricci-Calculus (2nd ed.). Springer-Verlag, Berlin (1954).

E. Cartan, Leçons sur La Géométrie des espaces de Riemann, Gauthier-Villars, Paris (1928).

E. Cartan, La Théorie des groupes finis et continus et la Géométrie Differentielle. Gauthier-Villars, Paris (1937).

5. THEORIES OF SPACE, TIME AND GRAVITATION

5.1. SPACE-TIME AS A DIFFERENTIABLE MANIFOLD

The most fundamental principle of physics, common to all physical theories so far put forward, is that space and time can be represented by a 4-dimensional differentiable manifold. This is often considered as so obvious that it is hardly worth mentioning or analyzing, but we consider it to be worthy of some attention, as did Einstein who, in his exposition of the theory of relativity, devoted considerable attention to the problem: why do we consider space and time to be a continuum (by which he meant what is nowadays called a differentiable manifold)?

If one ignores quantum phenomena, including the atomic structure of matter, and assumes matter to be infinitely divisible so that there is no inherent lower limit on the extension of bodies, we shall show that it is plausible that a differentiable manifold is an appropriate model of space-time. One can imagine that there is in space a very large number of very small clocks, i.e. small bodies with a mechanism that can be used to indicate time. They need not be 'good' clocks in any sense of the word 'good.' All that is required is that they associate a number with every instant of time, and that two instants of time are never given the same number by the same clock. Further, each clock has three identification numbers engraved on it, such that no two clocks bear the same three numbers. The fact that three numbers are both necessary and sufficient to distinguish the clocks is the meaning of space being 3-dimensional, and is to be taken as an experimental fact. Time is only one-dimensional, as the clocks need indicate only one number completely to specify an instant of time. Together, we see that we need 4 numbers to specify a point in space and time, and so space-time is 4-dimensional.

If the identification numbers and the time readings are such that the neighboring clocks in space bear numbers differing very little from one another and indicate times differing very little, then they can be considered as defining a coordinate system in a differentiable manifold representation of space-time. Such a coordinate system can be used to analyze all physical phenomena in terms of coincidences of events with these clocks. For example, the motion of a particle may be described by saying that the particle coincides with clock (1,3,-5) at time 6 by that clock, with clock (2,3,-6) at time 7 by that clock, with clock (3,3,-6) at time 8 ..., filling in as many events along the path of the clock as is desired. But to carry through this procedure it must be possible to associate clocks with as many points of space as one desires, which requires that we be able to make the clocks arbitrarily small, or we would not be able to find room for them at all; and furthermore it must be possible to read the dial of each clock with arbitrarily high accuracy. Perhaps we should point out that such a system of clocks has nothing to do with the <u>metrical</u> properties of space-time; it is necessary in order for space and time to be considered even as a differentiable manifold.

The situation becomes essentially more complicated when one takes into account the atomic, quantum nature of matter. It does not seem to be possible to have clocks whose linear dimensions are smaller than, say, 10^{-13} cm. Moreover, such small clocks would be subject to quantum laws, so that if one knows with great accuracy where one such clock is at a given instant of time, at a later time its position will be very indeterminate. Even worse, such elementary clocks are, by quantum principles, indistinguishable, so we can neither engrave them with numbers to identify one from another, nor provide them with dials to indicate the time, etc! In fact, they would be completely useless for our purposes. This shows that one really has to use macroscopic clocks, composed of very large numbers of atoms, and then we cannot pack enough such clocks together to form a very useful coordinate system. And if it is not possible to construct, even in principle, a coordinate system in space and time, does it make sense to assume that space and time can be represented by a differentiable manifold? The question seems to be open, but one should not assume it as obvious that it is so.

Many people have expressed doubts along the above lines, and some have tried to construct physical theories from new assumptions about the structure of space and time, but up to now, no such attempt has met with much success. The first idea that suggests itself is to represent space and time by a regular lattice of points like that of atoms in a crystal, but one can easily see that this will not work. The symmetry group of the lattice would have to approximate to the Lorentz group, and one cannot find a lattice which would admit a symmetry transformation corresponding to a Lorentz transformation with very small velocity. A recent discussion containing arguments against the continuity of space and time is given by Bohm.

It has been said that the modern approaches to elementary particle theory based on S-matrix theory and related theories--such as dispersion relation theories--are not based on the assumption of a space-time continuum. But actually this statement is unfounded, as they require invariance under the inhomogeneous Lorentz group, from which we can construct the usual Minkowski space-time of special relativity.

From now on we shall always assume that space-time can be represented by a 4-dimensional differentiable manifold. This is why the differentiable manifold concept was defined with care and discussed in detail in the preceding chapter. Any changes in this assumption would result in a very profound revolution in physics.

5.2. THE AFFINE CONNECTION IN PHYSICS

There is another fundamental principle of physics that is common to all physical theories so far put forward. This is that the differentiable manifold of space and time is endowed with an affine connection whose geodesics form a privileged set of world-lines in space-time. The particular affine connection depends on the theory under consideration and may also depend on the particular solution of the theory we are considering, but the existence of an affine connection is common to all theories. It is necessary in order that the fundamental laws of physics can be expressed in the form of differential equations, which is certainly true of all physical theories.

Since one must be able to determine by physical experiment every mathematical construct introduced in a physical theory, we must look for a method of physically determining the affine connection. The symmetric part of an affine connection is determined when the totality of geodesics in the manifold is known. Not every family of curves in a differentiable manifold can be interpreted as the geodesics of an affine connection, so we must look for a privileged set of world-lines in space-time which can be so interpreted. Of course, this is not the only way of determining physically the symmetric part of the affine connection, but it is the most natural way. If a theory postulates a non-symmetric affine connection, the antisymmetric part, i.e. the torsion tensor, must be separately determined.

So we see that the first question we should ask of a physical theory is: What physically determined privileged class of world-lines is to be interpreted as the totality of geodesics of the affine connection of the space-time manifold? To this question each theory gives its own answer, and to proceed further we must specialize to a particular theory. This we shall now do.

5.3. NEWTONIAN MECHANICS IN THE ABSENCE OF GRAVITATION

The privileged class of world-lines in Newtonian mechanics in the absence of gravitation is provided by Newton's First Law of Motion. This law is sometimes said to be trivial and to follow from the Second Law, but we shall see that it is in fact one of the most important laws of physics. Before considering it further, however, we should first briefly discuss the concept of absolute time which lies at the very foundations of Newtonian physics.

In the language of differentiable manifolds, if $\{X, \Theta, \mathfrak{F}\}$ is the space-time manifold, then an absolute time is a real-valued function $t \in \mathfrak{F}$ with the following mathematical properties

(i) it is determined up to a linear transformation $t \to at+b$, $a,b \in R$, $a > 0$

(ii) the subspaces t = const are 3-dimensional manifolds, homeomorphic to $\underline{R}^3$.

(iii) through each point of $\overline{X}$ passes exactly one such subspace.

It is also assumed that there exists a family of ideal clocks which measure t, i.e. which indicate, say by means of a pointer on a dial, the value of t at the point in space-time at which the clock is, regardless of the motion which the clock has undergone in the time intervening since the clock was correctly set at some time in the past. This is the physical meaning of absolute time. It is up to the theory to specify how these ideal clocks may in principle be constructed.

We are now in a position to state Newton's First Law of Motion. It may conveniently be formulated in two parts:

(i) In the absence of gravitation there exists a privileged class of motions of bodies, called <u>free motions</u>, followed by bodies not acted on by any force.

(ii) There exists a global coordinate system $\{x^a\}$, $a = 1,2,3,4$, with $x^4 = t$, an absolute time, in which free motions are characterized by

$$\frac{d^2x^\alpha}{dt^2} = 0 \qquad \alpha = 1,2,3 \tag{5.1}$$

From this we get

$$\frac{d^2x^a}{dt^2} = 0 \qquad a = 1,2,3,4$$

for free motions, which shows that the trajectories of free motions may be considered as the geodesics of an affine connection whose components vanish in this particular coordinate system. Consequently the curvature tensor of the affine connection vanishes. Such an affine connection is called <u>integrable</u>. The coordinate systems specified in (ii) above are called <u>inertial</u>. They are characterized by the vanishing of the components of the affine connection with respect to them, and are so determined up to an arbitrary linear transformation of the x^a which leaves x^4 an absolute time, i.e. transformations of the form

$$\left.\begin{array}{l} x^\alpha \to A^\alpha{}_\beta x^\beta + c^\alpha t + d^\alpha \\ t \to at + b \end{array}\right\} \qquad (5.2)$$

where $A^\alpha{}_\beta$, a, b, c^α, d^α are real constants, $a \neq 0$ and the matrix $A^\alpha{}_\beta$ non-singular.

5.4. NEWTONIAN MECHANICS IN THE PRESENCE OF GRAVITATION

In the presence of gravitation, Newton's First Law of Motion cannot be postulated in the above form as there can be no completely free motions, i.e. there can be no bodies not acted on by any force whatsoever, since gravitation is a force with infinite range which cannot be shielded by any known means. We are therefore forced to seek a new privileged class of motions. The best we can do is to take for this privileged class the <u>free falls</u>, i.e. the trajectories of bodies acted on only by a gravitational force. The modified formulation of the First Law then takes the form:

(i) There exists a privileged class of motions of bodies, called <u>free falls</u>, followed by bodies acted on only by a gravitational force.

(ii) There exists a global coordinate system $\{x^a\}$, $a = 1,2,3,4$, and a real-valued function $\phi \in \mathfrak{F}$, such that free falls are characterized by

$$\frac{d^2x^\alpha}{dt^2} = -\frac{\partial\phi}{\partial x^\alpha}, \qquad \alpha = 1,2,3. \qquad (5.3)$$

The trajectories of free falls can now be taken as the geodesics of an affine connection on the space-time manifold, but now the affine connection will have a non-vanishing curvature tensor. We shall study this in more detail below.

We note that (5.3) is invariant under a larger group of transformations than (5.2). In fact, it is easy to show that the most general transformation leaving the form of (5.3) invariant consists of the composition of a rotation with a transformation of the form

$$\left.\begin{aligned} x^\alpha &\to x^\alpha + a^\alpha(t) \\ \phi &\to \phi - \sum_{\alpha=1}^{3} \frac{d^2 a^\alpha}{dt^2} x^\alpha \end{aligned}\right\} \qquad (5.4)$$

where the real-valued function $a^\alpha(t)$ are arbitrary. As a consequence of this extension of the invariance group, the motion of an inertial reference frame is not as well defined in the presence of gravitation as it is in its absence. In the case of an isolated system of bodies we can return to the group (5.2) and so completely recover the usual concept of an inertial frame, by imposing the additional condition that ϕ should tend to zero at large distances from the system. However, in practice we cannot have an isolated system; the distant matter in the universe is always present and there is nothing we can do about it. We shall see in Chapter 9 that this weakening of the concept of an inertial frame has importance in the formulation of Newtonian cosmology. It is interesting to note that this weakening of the concept of inertial frames in the presence of gravitation occurs in Newtonian physics, and is not peculiar to general relativity, as is commonly supposed.

One could now determine from (5.3) the affine connection implied by interpreting the trajectories of free falls as geodesics, and from this one could build up the whole geometric structure of Newtonian gravitation theory. We shall not follow this approach, however, and instead we shall propose a set of geometric axioms and shall show that they lead to the familiar form of Newtonian gravitation theory.

5.5. AXIOMATIC FORMULATION OF THE GEOMETRY OF NEWTONIAN GRAVITATION THEORY

We suppose that space-time can be represented by a differentiable manifold $\{X, \Theta, \mathcal{D}\}$, homeomorphic to R^4 and endowed with

(i) A symmetric affine connection with components Γ^a_{bc}.

(ii) A symmetric tensor field g of valence (2,0).

(iii) A real-valued function $t \in \mathfrak{F}$.

These structures are to satisfy a certain set of axioms. Let $\{x^a\}$ be a global coordinate system. Let Γ^a_{bc}, $R^a{}_{bcd}$, g^{ab} be the corresponding natural components of the affine connection, its curvature tensor, and the tensor field g, and let $t_a \overset{\text{def}}{=} \partial t/\partial x^a$. Then the axioms are:

I. $R^c{}_{cab} = 0$. This is equivalent to the existence of a covariant constant scalar density W, $\nabla_a W = 0$, which can be used to define a unit of volume throughout the manifold (c.f. Schouten,[2] p. 155), or equivalently, a covariant constant 4-vector $f^{abcd} = f^{[abcd]}$.

II. $t_{[e}R^a{}_{b]cd} = 0$.

III. $R^a{}_b{}^c{}_d = R^c{}_d{}^a{}_b$, where $R^a{}_b{}^c{}_d \overset{\text{def}}{=} g^{ce}R^a{}_{bed}$.

IV. $\nabla_c g^{ab} = 0$.

V. The rank of the matrix g^{ab} is 3, $g^{ab}t_b = 0$, and the quadratic form $g^{ab}x_a x_b$ is positive semi-definite.

VI. In vacuo the gravitational field equations are $R_{ab} = 0$, where $R_{ab} = R^c{}_{abc}$ is the Ricci tensor of the affine connection.

VII. Free falls follow the geodesics of Γ^a_{bc}.

We shall now derive some consequences of these axioms. From V, $g^{ab}t_b = 0$, and covariantly differentiating this and using IV gives

$$g^{ab}\nabla_c t_b = 0$$

from which V gives the existence of a vector S_a such that

$$\nabla_a t_b = S_a t_b \tag{5.5}$$

But the connection is symmetric, and so

$$\nabla_{[a}t_{b]} = \partial_{[a}\partial_{b]}t = 0$$

which with (5.5) gives

$$S_{[a}t_{b]} = 0 \quad \text{so that} \quad S_a = T t_a \tag{5.6}$$

for some scalar function T. Putting this back in (5.5) gives

$$\nabla_a t_b = T t_a t_b \tag{5.7}$$

Now II implies the existence of a tensor $R^a{}_{cd}$ such that

$$R^a{}_{bcd} = t_b R^a{}_{cd} \tag{5.8}$$

Hence

$$t_d R^d{}_{cab} = t_d t_c R^d{}_{ab} = t_c R^d{}_{dab} = 0 \quad \text{by I} \tag{5.9}$$

But we have, by the Ricci identity,

$$2\nabla_{[a}\nabla_{b]}t_c = t_d R^d{}_{cab}$$

and so from (5.9)

$$\nabla_{[a}\nabla_{b]}t_c = 0$$

Substituting (5.7) into this gives

$$\partial_{[a}T\, \partial_{b]}t = 0$$

so that $\partial_a T$ and $\partial_a t$ are everywhere parallel. The surfaces T = const. and t = const. must therefore coincide, and so T must be a function of t only. So

$$T = T(t) \tag{5.10}$$

Now none of the axioms would be affected if t were replaced by a differentiable function f(t), $df/dt \neq 0$. If we write $\bar{t} = f(t)$, then

$$\nabla_a \bar{t}_b = f'' t_a t_b + f' \nabla_a t_b$$

where $\bar{t}_a = \partial\bar{t}/\partial x^a$ and ' denotes d/dt. Using (5.7) and (5.10) this gives

$$\nabla_a \bar{t}_b = (f'' + T(t) f') t_a t_b$$

Now choose f so that

$$f''(t) + T(t) f'(t) = 0$$

This can always be done. Then we get $\nabla_a \bar{t}_b = 0$. Since $\bar{t}$ satisfies the axioms for t, we see that without loss of generality we can always assume

$$\nabla_a t_b = 0. \tag{5.11}$$

From now on we shall assume t to be so chosen. Then t is determined up to a linear transformation $t \to at+b$, $a, b \in R$, $a \neq 0$.

Now let $x^a = x^a(t)$ be a geodesic in X, with t not necessarily assumed to be an affine parameter. Then from section 4.16

$$\frac{D}{dt}\frac{dx^a}{dt} = \lambda(t)\frac{dx^a}{dt} \tag{5.12}$$

for some scalar function $\lambda(t)$. Now

$$t_a \frac{dx^a}{dt} = \frac{\partial t}{\partial x^a}\frac{dx^a}{dt} = 1$$

So multiplying (5.12) by t_a and using (5.11) gives

$$\lambda(t) = t_a \frac{D}{dt}\frac{dx^a}{dt} = \frac{D}{dt}\left(t_a \frac{dx^a}{dt}\right) = 0$$

Hence we see that in fact t <u>is</u> an affine parameter on all geodesics in X.

Let us now return to (5.8). We know that

$$R^{a}{}_{[bcd]} = 0$$

for the curvature tensor of any symmetric connection, and so

$$t_{[b} R^{a}{}_{cd]} = 0$$

which implies the existence of a tensor $\phi^{a}{}_{c}$ such that

$$R^{a}{}_{cd} = 2\phi^{a}{}_{[c} t_{d]} \tag{5.13}$$

and $\phi^{a}{}_{c}$ is determined up to a transformation

$$\phi^{a}{}_{c} \to \phi^{a}{}_{c} + S^{a} t_{c} \tag{5.14}$$

for any vector field S^{a}. So now we have reduced $R^{a}{}_{bcd}$ to the form

$$R^{a}{}_{bcd} = 2t_{b}\phi^{a}{}_{[c} t_{d]}. \tag{5.15}$$

To proceed further, we shall state without proof that from (5.8) it can be shown that there exist three linearly independent covariant constant vector fields $\underset{K}{\xi}{}^{a}$, where K = 1,2,3 is a label for the three fields, all lying in surfaces t = const., so that

$$t_{a}\underset{K}{\xi}{}^{a} = 0, \qquad \nabla_{a}\underset{K}{\xi}{}^{b} = 0 \tag{5.16}$$

and by applying the Bianchi identity to (5.15) it can be shown that the S^{a} in (5.14) can be chosen to make $\phi^{a}{}_{c}$ take the form

$$\phi^{a}{}_{c} = \nabla_{c}\phi^{a}, \qquad \phi^{a} t_{a} = 0 \tag{5.17}$$

This vector field ϕ^{a} is so determined up to a transformation of the form

$$\phi^a \to \phi^a + \overset{K}{u}(t)\underset{K}{\xi}^a \tag{5.18}$$

where the three scalar functions $\overset{K}{u}(t)$ of t are arbitrary.

Now define a new affine connection in the manifold, with components ${}^\circ\Gamma^a_{bc}$, by

$${}^\circ\Gamma^a_{bc} = \Gamma^a_{bc} - \phi^a t_b t_c . \tag{5.19}$$

Then using (5.15) and (5.17) it is easy to show that ${}^\circ R^a{}_{bcd}$, the curvature tensor of this new connection, vanishes identically, so that the connection ${}^\circ\Gamma^a_{bc}$ is integrable. The geodesic equation

$$\frac{d^2x^a}{dt^2} + \Gamma^a_{bc}\frac{dx^b}{dt}\frac{dx^c}{dt} = 0$$

which is, by VII, the equation of the trajectory of a free fall, can now by (5.19) be written as

$$\frac{d^2x^a}{dt^2} + {}^\circ\Gamma^a_{bc}\frac{dx^b}{dt}\frac{dx^c}{dt} = -\phi^a \tag{5.20}$$

As stated in section 4.18, since ${}^\circ\Gamma^a_{bc}$ is integrable, there exists a coordinate system $\{x^{a'}\}$ which may be called 'inertial,' in which $\Gamma^{a'}_{b'c'} = 0$, and then (5.20) becomes

$$\frac{d^2x^{a'}}{dt^2} = -\phi^{a'}.$$

We interpret $-\phi^{a'}$ as the gravitational force on the body in free fall. The terms involving ${}^\circ\Gamma^a_{bc}$ in (5.20), present when a general coordinate system is used, are then interpreted as the inertial forces, i.e. centrifugal, Coriolis forces, etc., and again $-\phi^a$ is identified as the gravitational force.

It is important to note that, because of the possibility of making a transformation of the form (5.18), neither the gravitational field ϕ^a nor the 'inertial' affine connection ${}^\circ\Gamma^a_{bc}$ are determined uniquely by the geometry of Newtonian space-time. These concepts may be uniquely defined only by a global assumption, e.g. that at large distances from an isolated system of bodies the field ϕ^a tends to zero. As was pointed out in the previous section, this assumption cannot always be made, and so we again arrive at the lack of determinacy of inertial reference frames caused by the presence of gravitation. Note that the combined inertial and gravitational fields, represented by the original affine connection Γ^a_{bc}, is uniquely defined and can be determined by local experiments on freely falling test particles. The splitting of this field into an inertial and a gravitational part is only a convenience; it is not necessary for the meaningful interpretation of the theory, and we have made it here in order to establish the connection with the usual formulation of the Newtonian theory of gravitation.

So far, we have not considered the field equations satisfied by ϕ^a. Now substituting (5.15) into III gives, on using V,

$$g^{ac}\phi^b{}_c = g^{bc}\phi^a{}_c. \tag{5.21}$$

Remembering that $\phi^a{}_c = \nabla_c\phi^a$, (5.21) is the integrability condition for the equation

$$\phi^a = g^{ab}\partial_b\phi \tag{5.22}$$

so that III gives the existence of a scalar gravitational potential ϕ. But on using (5.15), VI gives

$$\begin{aligned} 0 &= 2t_c\phi^a{}_{[a}t_{d]} \\ &= t_c t_d \nabla_a\phi^a \qquad \text{by (5.17) and (5.11)} \\ &= t_c t_d\, g^{ab}\nabla_a\nabla_b\phi \qquad \text{by (5.22) and IV}. \end{aligned}$$

So the field equation satisfied by the potential ϕ is

$$g^{ab}\nabla_a\nabla_b\phi = 0\,. \tag{5.23}$$

Now let $\overset{\circ}{\nabla}_a$ denote covariant differentiation with respect to the 'inertial' affine connection ${}^\circ\Gamma^a_{bc}$. Then by (5.19)

$$\nabla_a\nabla_b\phi = \overset{\circ}{\nabla}_a\overset{\circ}{\nabla}_b\phi - t_a t_b\phi^c\partial_c\phi\,;$$

multiplying this by g^{ab} and using V gives

$$g^{ab}\nabla_a\nabla_b\phi = g^{ab}\overset{\circ}{\nabla}_a\overset{\circ}{\nabla}_b\phi \tag{5.24}$$

so that the field equations (5.23) can be written as

$$g^{ab}\overset{\circ}{\nabla}_a\overset{\circ}{\nabla}_b\phi = 0. \tag{5.25}$$

We have now obtained the complete set of equations of particle motion (5.20) and empty space field equations (5.22) and (5.25). These are suggestively like the usual equations of Newtonian gravitation theory, but we have not yet shown that they are exactly equivalent. This is our next task, and to do it we set up a special coordinate system. We first note that, by (5.19),

$$\begin{aligned}\overset{\circ}{\nabla}_a t_b &= \nabla_a t_b + t_a t_b \phi^c t_c \\ &= 0 \quad \text{by (5.11) and (5.17).}\end{aligned} \tag{5.26}$$

and

$$\begin{aligned}\overset{\circ}{\nabla}_a g^{bc} &= \nabla_a g^{bc} - t_a(\phi^b t_d g^{dc} + \phi^c t_d g^{bd}) \\ &= 0 \quad \text{by IV and V.}\end{aligned} \tag{5.27}$$

But, as we have shown above, there exists a special coordinate system $\{x^{a'}\}$ in which ${}^\circ\Gamma_{b'}{}^{a'}{}_{c'} = 0$, and in this coordinate system the covariant differentiation $\overset{\circ}{\nabla}_a$ reduces to partial differentiation ∂_a. Equations (5.26) and (5.27) then tell us that the components $t_{a'}$, $g^{a'b'}$ are constants, i.e. independent of position. Since $t_{a'} = \partial_{a'}t$, t must be a linear function of the $x^{a'}$, and since the condition ${}^\circ\Gamma_{b'}{}^{a'}{}_{c'} = 0$ is unaffected by a linear coordinate transformation, the coordinates $\{x^{a'}\}$ may be chosen so that

$$t = x^{4'}, \tag{5.28}$$

and then

$$t_{a'} = \delta^4_{\alpha'} . \tag{5.29}$$

V then gives

(i) $g^{a'4'} = g^{4'a'} = 0$

(ii) the 3 x 3 matrix $g^{\alpha'\beta'}$, $\alpha', \beta' = 1,2,3$, is non-singular.

Since by V, the quadratic form $g^{a'b'} x_{a'} x_{b'}$ is positive semi-definite, the quadratic form $g^{\alpha'\beta'} x_{\alpha'} x_{\beta'}$ must be positive definite. It is therefore possible, by a linear transformation of the coordinates $x^{1'}$, $x^{2'}$, $x^{3'}$ (which leaves (5.28) unaffected), to reduce the constant matrix $g^{\alpha'\beta'}$ to the unit matrix.

So we have now shown that coordinates $\{x^a\}$ may be found such that

$$\left.\begin{aligned} {}^{\circ}\Gamma^a_{bc} &= 0, \\ t_a &= \delta^4_a, \\ g^{ab} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \end{aligned}\right\} \tag{5.30}$$

In these coordinates (5.20), (5.22), and (5.25) give

$$\left.\begin{aligned} \frac{d^2x^\alpha}{dt^2} &= -\frac{\partial\phi}{\partial x^\alpha}, \qquad \alpha = 1,2,3 \\ \Delta\phi &= 0 \end{aligned}\right. \tag{5.31}$$

where Δ is the usual Laplacian operator

$$\Delta \overset{\text{def}}{=} \frac{\partial^2}{(\partial x^1)^2} + \frac{\partial^2}{(\partial x^2)^2} + \frac{\partial^2}{(\partial x^3)^2} .$$

These are precisely the usual form of the equations of Newtonian gravitation theory, and our identification of the geometric formulation with the standard formulation is complete. If we wish to include the source terms due to matter in the gravitational field equations, we simply replace VI by

$$R_{ab} = 4\pi k \rho t_a t_b$$

so that instead of (5.23) we get

$$g^{ab} \nabla_a \nabla_b \phi = 4\pi k \rho$$

which leads to Poisson's equation

$$\Delta\phi = 4\pi\rho$$

as required. Here ρ is the mass density and k is the Newtonian constant of gravitation.

5.6. THE INTERPRETATION OF g^{ab}. PRIVILEGED CLASSES OF COORDINATE SYSTEMS

The tensor g^{ab}, which cannot be interpreted as a metric tensor in the whole space-time manifold because its rank is 3, does determine a non-singular metric tensor in each of the surfaces t=const., which are the instantaneous (in terms of absolute time) 3-spaces of Newtonian mechanics. Let $\vec{u}$ be a vector tangent to a surface t = const., with components u^a, so that $u^a t_a = 0$. Then it determines a differential form u with components u_a such that

$$u^a = g^{ab} u_b$$

by virtue of V, to within a transformation

$$u_a \to u_a + \lambda t_a, \qquad \lambda \text{ scalar.}$$

Now let $\vec{u}$, $\vec{v}$ be two such vectors with u,v corresponding forms. Then, since $g^{ab}t_b=0$, the scalar product defined by

$$\vec{u}\cdot\vec{v} = g^{ab} u_a v_b$$

is unique. We thus see that the physical interpretation of g^{ab} is as a scalar product on the instantaneous 3-spaces of Newtonian physics.

We note that in the Newtonian theory the affine connection is not determined completely by the 'metric' g^{ab}. The affine connection and the 'metric' g^{ab} are related by IV, but this is not sufficient to determine Γ_{bc}^{a} completely because the matrix g^{ab} is singular. (Note that Γ_{bc}^{a} and ${}^{\circ}\Gamma_{bc}^{a}$ are two different solutions of $\nabla_a g^{bc} = 0$, considered as an equation for the components of the affine connection.) A preferred coordinate system in a theory is one which puts some geometrical structure in the theory in a particularly simple form. The multiplicity of geometrical structures present in Newtonian theory thus enables us to have many different classes of preferred coordinate systems, some more useful than others.

The most restrictive set of conditions we can impose on the coordinates is (5.30). It is easily seen that the general coordinate transformation preserving all these conditions is

$$x^{\alpha} \to R^{\alpha}{}_{\beta}\, x^{\beta} + v^{\alpha} x^{4} + a^{\alpha}, \qquad \alpha, \beta = 1,2,3$$

$$x^{4} \to x^{4} + b$$

where $R^{\alpha}{}_{\beta}$, v^{α}, a^{α}, b are all constants and independent of time, and the matrix $R^{\alpha}{}_{\beta}$ is orthogonal. These, of course, are the Galilean transformations, as expected.

The next most useful set of coordinates is that obtained by dropping the condition ${}^{\circ}\Gamma_{bc}^{a} = 0$, but retaining the other two conditions. The allowed transformations are then

$$x^{\alpha} \to R^{\alpha}{}_{\beta}(t)\, x^{\beta} + a^{\alpha}(t)$$

$$x^{4} \to x^{4} + b$$

where the coefficients $R^{\alpha}{}_{\beta}(t)$ and $a^{\alpha}(t)$ are now arbitrary functions of t, subject to $R^{\alpha}{}_{\beta}(t)$ being an orthogonal matrix for all t. These are the coordinates used when we have axes fixed in an arbitrarily moving rigid body, and they require the introduction of inertial forces in the equations of motion, represented by the now non-zero ${}^{o}\Gamma^{a}_{bc}$.

Other classes of coordinates, e.g. (5.2), obtained by dropping the condition on g^{ab} and allowing $t_a = a\delta_a{}^4$, $a \neq 0$, are of less importance.

Remember that in addition to coordinate transformations, we also have transformations of ϕ^a given by (5.18) and equivalent to the transformations of ϕ given by (5.4); and linear transformations of t as stated above, (5.11). The significance of the transformations of ϕ has been discussed above, and that of the transformations of t is simply a change in the origin and the unit of absolute time.

To complete our study of Newtonian gravitation theory we shall just say a few words about the relationship between it and relativistic theory. In special relativity the coordinates may be chosen so that the metric tensor takes the Minkowski form

$$\eta_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}$$

so that

$$\eta^{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1/c^2 \end{pmatrix} \tag{5.32}$$

where c is the velocity of light. If we take the formal limit $c \to \infty$, by which we know that special relativity reduces to Newtonian mechanics, we see that η^{ab} reduces to the g^{ab} of (5.30), while η_{ab} becomes meaningless as it contains an infinite element. This shows the origin of the 'metric' g^{ab} of the Newtonian theory. The change of rank of g^{ab} in this limiting process is the cause of the loss of the completely deterministic relation between the metric tensor and the affine connection that is present both in special relativity and in general relativity. The multiplicity of independent geometric structures in Newtonian theory, which results in the many possible classes of privileged coordinate systems present in Newtonian theory, is also missing in relativity, where the metric tensor is the only independent geometric structure, and the Lorentz group of special relativity is the group that preserves the special form (5.32) of the metric tensor.

We shall now leave Newtonian gravitation theory to go on to consider possible relativistic theories of gravitation.

5.7. RELATIVISTIC THEORIES OF SPACE, TIME, AND GRAVITATION

After Einstein had formulated the special theory of relativity, it became clear that Newton's theory of gravitation could not be reconciled with the postulates of that theory, as Newton's law of attraction is based on the principle of instantaneous action at a distance, which is irreconcilable with special relativity. It therefore became necessary to look for a new theory of gravitation which was consistent with the postulates of special relativity.

The first idea that comes to mind is to represent gravitation by some geometric object field ϕ, not necessarily scalar, in the Minkowski flat space-time of special relativity, just as the electromagnetic field can be represented by a bivector field. Since the field ϕ must in some sense approximate to the Newtonian gravitational potential in the non-relativistic limit, as discussed in section 1.4, we expect ϕ to satisfy an equation of the form

$$\Box\phi + \text{possible nonlinear terms} = 4\pi k\rho \qquad (5.33)$$

where $\Box$ is the D'Alembertian operator $\Box = \nabla^2 - \partial^2/\partial t^2$. However, further consideration shows that this approach probably isn't very sound, for, if we accept the equality of inertial and passive gravitational masses as a basic law of nature, it should be impossible uniquely to separate gravitational and inertial effects by local experiments on test-particles, and the inertial reference frames implicitly assumed in writing an equation of the form (5.33) cannot be well-defined. As we saw above, these phenomena occur in Newtonian gravitation theory, although they are not usually recognized as so doing. It may be argued that when other phenomena are considered besides the motion of test-particles, it may then be possible to determine an inertial reference frame by local experiments. But, following Einstein, it seems aesthetically better to generalize from the fact that local dynamical experiments on test-particles cannot distinguish inertial effects from gravitational effects to the principle that <u>no</u> local experiments should be able to distinguish inertial effects from gravitational effects. This principle is known as the <u>Principle of Equivalence</u>. More mathematically, we may say that in the presence of gravitation it is not possible by local experiments to determine a unique integrable affine connection in the space-time manifold.

At this point we could give the line of reasoning which led Einstein to his General Theory of Relativity. But this can be found in many books (see the references for this Chapter at the end of the chapter), and so instead we shall present a more usual approach. The relativistic theory of gravitation that we are looking for has to agree, in two different limits, with both special relativity (in the absence of gravitation) and Newtonian gravitation theory (in the weak-field and 'non-relativistic' limits), as is indicated in the following diagram:

$$\begin{tikzcd}
\boxed{\text{Relativistic theory of gravitation}} \arrow[r] \arrow[d] & \boxed{\text{Newtonian theory of gravitation}} \arrow[d] \\
\boxed{\text{Special theory of relativity}} \arrow[r] & \boxed{\text{Newtonian theory in absence of gravitation}}
\end{tikzcd}$$

where arrows indicate increasing specialization. So we shall tabulate the geometric form of the postulates of both special relativity and Newtonian gravitation, and see how far it is possible to guess the axioms of a theory of which both these theories are special cases. In doing so one should remember that a physical theory can never be proved; it can only be made plausible or disproved. We hope to show how the general theory of relativity can be made plausible.

Table of Axioms

Newtonian Theory of Gravitation

N1. Space and time can be represented by a 4-dimensional differentiable manifold endowed with
(a) a symmetric affine connection Γ^a_{bc}
(b) a metric tensor g^{ab} of rank 3 such that $\nabla_a g^{bc} = 0$, and of signature (1,1,1,0)

Special Relativity

S1. Space and time can be represented by a 4-dimensional differentiable manifold endowed with
(a) a symmetric affine connection
(b) a metric tensor g^{ab} of rank 4 such that $\nabla_a g^{bc} = 0$, and of signature (1,1,1,-1).

(c) a differentiable function t, the absolute time.	Along any time-like world-line a proper-time τ is defined by $d\tau^2 = -g_{ab}dx^a dx^b$
N2. There exist ideal clocks which measure absolute time t.	S2. There exist ideal clocks which measure proper time τ along their world-lines.
N3. Certain restrictions are imposed on the curvature tensor, given in section 5.5.	S3. Restrictions on the curvature tensor follow from $\nabla_a g^{bc} = 0$, and from rank of g^{ab} being 4.
N4. In absence of gravitation $R^a{}_{bcd} = 0$.	S4. In absence of gravitation $R^a{}_{bcd} = 0$.
N5. In presence of gravitation, $R_{ab} = 0$ in vacuo, $R_{ab} = 4\pi k\rho t_a t_b$ in matter.	S5. -
N6. The world-lines of freely falling test bodies are the geodesics of Γ^a_{bc}.	S6. The world-lines of free test bodies are the geodesics of Γ^a_{bc}.

Now we know that Newtonian theory in the absence of gravitation, i.e. with N4 satisfied rather than N5, follows in the formal limit $c \to \infty$ from special relativity. So we want to modify the postulates of special relativity as laid out above as little as possible consistent with N5 appearing in the limit in place of N4. Clearly we must put in S5 that in vacuo in the presence of gravitation $R_{ab} = 0$. S1, S2, S5 and S6 then give the general relativity postulates for the gravitational field in the absence of matter. We shall return to the problem of the field equations in the presence of matter later in the course, as the second part of N5 has no obvious consistent generalization to the relativistic theory.

We hope that we have shown above that by presenting special relativity and Newtonian gravitation theory in geometric terms, general relativity theory is the most natural generalization of special relativity theory to include gravitation in a manner which will reduce to Newtonian gravitation theory in the non-relativistic limit.

At this point we should perhaps mention that most of the other postulates S1 to S6 in the above table could be modified also in the passage to a relativistic gravitation theory, and that many theories have been proposed in which such generalizations are made. Examples of such generalizations are:

(a) use of a non-symmetric affine connection
(b) use of a non-symmetric metric tensor
(c) dropping the condition $\nabla_a g^{bc} = 0$ which says that the affine connection is metric.

Having mentioned their existence, we shall not pursue them any further.

References:

On the geometric formulation of the Newtonian theory of gravitation:

E. Cartan, Ann. Éc. Norm. Sup. 40, 325 (1923).

E. Cartan, Ann. Éc. Norm. Sup. 41, 1 (1924).

K. Friedrichs, Math. Ann. 98, 566 (1962).

P. Havas, (preprint).

A. Trautman, C.R. Acad. Sc. Paris 257, 617 (1963).

6. FUNDAMENTAL PROPERTIES OF GENERAL RELATIVITY

6.1. THE PRINCIPLE OF GENERAL COVARIANCE

This is a principle which is often considered to be at the foundations of general relativity, but we have managed to obtain general relativity by a (we hope) fairly convincing chain of reasoning without ever mentioning such a principle. So perhaps we should now discuss the status of this principle.

It is a principle that was first put forward by Einstein, and it is often stated in the following forms, which are not exactly equivalent.

A. All coordinate systems are equally good for stating the laws of physics, and they should be treated on the same footing.
B. The equations of physics should have tensorial form.
C. The equations of physics should have the same form in all coordinate systems.

We shall consider the meaning that is to be attached to each form.

When one uses form A, one wishes to emphasize that, since there is no preferred integrable affine connection in the theory, it is not possible to distinguish a set of inertial coordinate systems. This is in some way a negative statement, and should not be overemphasized. In particular cases when we have a space-time manifold that refers to a definite physical situation, there are certain coordinates which can be preferred to others. For example, when we consider the field of a spherically symmetric body, we expect the field itself also to have spherical symmetry, and certainly coordinates that are adapted to spherical symmetry are to be preferred over other coordinate systems. But this should not be confused with the contents of statement A, which really says that there are no inertial coordinate systems in general relativity.

When the form B is used it means that if, in physics, we have a solution of an equation in one coordinate system and then transform this solution to any other coordinate system, then the transformed solution must satisfy the transformed equation. Tensorial equations have precisely this property.

We now come to the form C of the principle of general covariance. Suppose that in a certain coordinate system we have an equation

$$A_1 = 0 \tag{6.1}$$

where A_a are the components of a differentiable form field. Introduce a vector field $\vec{u}$ such that $\vec{u} = \vec{e}_1$, where $\{\vec{e}_a\}$ is the natural basis associated with the coordinates. Then in any coordinate system the equation (6.1) can be written in the form

$$u^a A_a = 0 .$$

This procedure, however, involves the introduction of an auxiliary vector field $\vec{u}$. I think that one should not introduce such additional structures in addition to those already present in the axioms of the theory (e.g. the metric tensor, affine connection) and to those that are necessary to describe the physical system. I think this statement is rather important, but it is denied by some people. Fock, for example, considers that Einstein's theory as formulated above is incomplete. He claims that, in order to make the theory meaningful and complete, one has to introduce in addition to the metric g_{ab} and the variables describing the physical system, privileged harmonic coordinate systems satisfying the condition $\partial_b(\sqrt{-g}\, g^{ab}) = 0$. Again, Møller claims that unless one introduces a privileged set of tetrads (orthonormal basis vectors) at every point, one cannot meaningfully talk about energy. In my opinion these additional structures are not really necessary, and a physical theory with these structures is not correct. The reason for my claim is that I don't think one can give a physical interpretation to these additional structures. If such a physical interpretation could be given, then I agree that one could introduce such structures. But the possibility of giving a physical interpretation to them includes them in those allowed by the principle I have already stated.

6.2. THE PRINCIPLE OF EQUIVALENCE

We shall now discuss further the meaning and use of the principle of equivalence, which was briefly mentioned in section 5.7. The principle of equivalence states that the local effects of a gravitational field are indistinguishable from those of inertial forces, and the principle acts as a guide for obtaining the equations of motion of systems in general relativity when the corresponding equations in special relativity are known. We shall illustrate the procedure by considering Maxwell's equations in vacuo. If f^{ab} is the electromagnetic field tensor, then Maxwell's equations in special relativity may be written in Minkowski coordinates as

$$\partial_b f^{ab} = 0, \qquad \partial_{[a} f_{bc]} = 0 \tag{6.2}$$

Still considering special relativity, we may transform (6.2) to an arbitrary coordinate system by making a coordinate transformation, and the effect on the equations is to replace partial derivatives ∂_a by covariant derivatives ∇_a, and the Minkowski metric tensor η_{ab}, occurring implicitly in the raising and lowering of indices, by the metric tensor g_{ab}:

$$\nabla_b f^{ab} = 0, \qquad \nabla_{[a} f_{bc]} = 0 \tag{6.3}$$

Now inertial forces occur in these equations in the form of the affine connection used in forming the covariant derivatives. The principle of equivalence tells us that the local effects of a gravitational field must occur in the same way. We may thus adopt (6.3) as the equations of the electromagnetic field in curved space, i.e. when a gravitational field is present; the difference from (6.3) being that the affine connection in the covariant derivative will no longer be integrable. The principle of equivalence, however, does not force us to adopt (6.3). We could add terms explicitly containing the curvature tensor $R^a{}_{bcd}$, and write, for example,

$$\nabla_b f^{ab} + R^a{}_{bcd} \nabla^b f^{cd} = 0 \tag{6.4}$$

since the second term vanishes in the transition to special relativity, where $R^a{}_{bcd} = 0$.

To avoid this ambiguity, it has been suggested that in the transition to general relativity we should adopt a <u>Principle of Minimal Gravitational Coupling</u>, i.e. we should not add any terms explicitly containing the curvature tensor. But this must be treated with care. For suppose we had started, not from (6.2), but from the equations in terms of the 4-potential A^α in the form

$$\partial_b\partial^b A_a = 0 \qquad \partial_a A^a = 0 \tag{6.5}$$

Then by the above procedure we are led to take

$$\nabla_b\nabla^b A_a = 0 \qquad \nabla_a A^a = 0 \tag{6.6}$$

for the electromagnetic field equations in curved space-time. Now using

$$f_{ab} = \nabla_a A_b - \nabla_b A_a \tag{6.7}$$

and the Ricci identity, we get

$$\begin{aligned} \nabla_b f^{ab} &= \nabla_b(\nabla^a A^b - \nabla^b A^a) \\ &= -\nabla_b\nabla^b A^a + \nabla^a(\nabla_b A^b) + R^a{}_b A^b \\ &= R^a{}_b A^b \qquad \text{by (6.6)} \end{aligned} \tag{6.8}$$

This agrees with (6.3) in empty space, $R_{ab} = 0$, but not in the presence of matter. There are, however, several reasons lending us to adopt (6.3) rather than (6.6) as the correct equations of electromagnetism in general relativity:

(i) The equations (6.6) are not gauge-invariant under the restricted gauge transformation $A_a \to A_a + \partial_a\Lambda$, $\nabla_a\nabla^a\Lambda = 0$, under which (6.5) are invariant, and consequently we do not know in which gauge (6.6) should be considered as holding.

(ii) If we consider a system containing charges, described by a current vector j^a, (6.5) becomes

$$\partial_b\partial^b A_a = j_a \qquad \partial_a A^a = 0 \tag{6.9}$$

and so (6.6) becomes

$$\nabla_b\nabla^b A_a = j_a \qquad \nabla_a A^a = 0 \tag{6.10}$$

from which (6.8) gives

$$\nabla_b f^{ab} = R^a{}_b A^b - j^a \tag{6.11}$$

Acting on this with ∇_a and using the identity $\nabla_a \nabla_b f^{ab} = 0$ gives

$$\nabla_a j^a = \nabla_a (R^a{}_b A^b).$$

which in general does not vanish, and hence charge is not conserved by these equations. Not only is this aesthetically distasteful, but also the law of conservation of charge is one of the most accurately tested laws of physics, and we could use the experimental accuracy to set an upper bound on the curvature of space.

On these grounds we adopt (6.3) as Maxwell's equations in general relativity.

In the case of Maxwell's equations there is one formulation, namely (6.2), on which the principle of minimal gravitational coupling may be used. We now consider another example which shows that even this principle may not necessarily be applicable. The equations of motion of a spinning particle in general relativity have been derived in many ways, e.g. one may first consider an extended body described by an energy-momentum tensor T^{ab} and obtain equations of motion from the conservation equation $\nabla_b T^{ab} = 0$, and then consider the limiting case as the body is shrunk to a point while maintaining a finite internal angular momentum. The equations obtained by this and every other method show that the spin, described by a skew tensor S^{ab}, interacts with the curvature of space to produce a force $\frac{1}{2}R_{abcd}\, u^b S^{cd}$ which denotes the particle from a geodesic, where u^a is the velocity vector of the particle. But if one started from the equations of motion of a spinning particle in special relativity, the principle of minimal gravitational coupling would prohibit the inclusion of such a term in the general relativistic equations.

A further difficulty arises in the transition from special to general relativity if the special relativity equations contain second partial derivatives, since the corresponding covariant derivatives do not commute. So should we, for example, replace a term $\partial_a \partial_b u_c$ with $\nabla_c \nabla_b u_c$, $\nabla_b \nabla_a u_c$, $\nabla_{(a}\nabla_{b)} u_c$ or some other term? What term? There is no simple answer to this question; all these forms being equal in special relativity, where $R^a{}_{bcd} = 0$.

Having now seen what problems can arise in the translation of equations of motion from special to general relativity, one may ask what method can one use to be sure of obtaining a correct result. The only certain method is to analyze a complicated system into simpler ones whose equations of motion in general relativity are already known. This may be done for the spinning particle, for example, as follows: First consider an extended body and write down equations of motion for an element of it, taking into account the forces exerted on it by the rest of the body. Then integrate over the whole body so that the internal forces integrate to zero, and finally shrink it to a point. The particle itself cannot be considered as a single element since, in the limiting process of shrinking to a point while maintaining a non-zero internal angular momentum, the internal velocities in the body must tend to infinity. In cases where such an analysis is impossible, we use the principles of equivalence and of minimal gravitational coupling as guides, but they must be used with caution. Besides Maxwell's equations, another problem requiring great care is the Dirac equation, but we shall not consider this here.

6.3. CORRESPONDENCE WITH NEWTONIAN THEORY*

A requirement of any relativistic theory of gravitation is that it should reduce to Newtonian theory in the case of slowly moving bodies in a weak field. We shall now show that, together with the assumption that space-time may be represented by a 4-dimensional Riemannian manifold of signature (---+) in which free falls are represented by geodesics, this requirement leads to a condition on the weak field form of the metric tensor.

Consider a space-time containing a weak gravitational field, so that there exists a coordinate system $\{x^a\}$ in which the components of the metric tensor, g_{ab}, differ only slightly from the Minkowski metric

$$\eta_{ab} = \mathrm{diag}(-1,-1,-1,+1). \tag{6.13}$$

---

* Lecturer's note: This section has been considerably expanded by one of the notetakers (W.G.D.) from the presentation given in the lectures.

Further, suppose that this field is produced by bodies whose velocities in this coordinate systems are small compared with that of light, c. Let λ be a small dimensionless parameter of order v/c, where v is a typical particle velocity, and suppose that g_{ab} can be expanded in powers of λ thus:

$$g_{ab} = \eta_{ab} + \lambda h_{ab} + O(\lambda^2) \tag{6.14}$$

where $x^4 = ct$. Then since in 'time' δx^4 the bodies producing the field move only distances of order $\lambda\delta x^4$, derivatives of the field variables h_{ab} with respect to x^4 are of order λ times the spatial derivatives of these variables. To make this distinction appear explicitly in the mathematics, we may introduce an auxiliary time-variable

$$x^0 = \lambda x^4 \tag{6.15}$$

and write all time derivatives in terms of this, so that then, say, $\partial_1 h_{ab}$ and $\partial_0 h_{ab}$ are of the same order in λ.

Now consider the motion in this field of a free test-particle for which also v/c is of order λ. Let $x^a = x^a(s)$ be the world-line of the particle, s being the proper time measured along it. Then

$$c\frac{dt}{ds} = 1 + O(\lambda), \qquad \frac{1}{c}\frac{dx^\alpha}{dt} = O(\lambda), \qquad \alpha = 1,2,3. \tag{6.16}$$

$$\frac{dx^a}{ds} = \frac{1}{c}\frac{dx^a}{dt}\left(c\frac{dt}{ds}\right) \tag{6.17}$$

$$\frac{d^2x^a}{ds^2} = \frac{1}{c^2}\frac{d^2x^\alpha}{dt^2}\left(c\frac{dt}{ds}\right)^2 + \frac{1}{c}\frac{dx^a}{dt}\frac{d^2(ct)}{ds^2} = \frac{1}{c^2}\frac{d^2x^a}{dt^2}\left(\frac{cdt}{ds}\right)^2(1+O(\lambda)) \tag{6.18}$$

Using these, the geodesic equation for the particle path,

$$\frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0$$

can be put in the form

$$\frac{1}{c^2}\frac{d^2x^a}{dt^2} + \frac{1}{c^2}\Gamma^a_{bc}\frac{dx^b}{dt}\frac{dx^c}{dt}(1 + O(\lambda)) = 0. \tag{6.19}$$

Now let α = 1,2,3, put a = α in this, and use (6.14) to give

$$\frac{1}{c^2}\frac{d^2x^\alpha}{dt^2} - \frac{\lambda}{2c^2}(\partial_b h_{c\alpha} + \partial_c h_{b\alpha} - \partial_\alpha h_{bc})\frac{dx^b}{dt}\frac{dx^c}{dt}(1 + O(\lambda)) = 0. \tag{6.20}$$

Using (6.15) and (6.16) we see that apart from the term

$$-\frac{\partial_\alpha h_{44}}{c^2}\frac{dx^4}{dt}\frac{dx^4}{dt}$$

from the parentheses in (6.20), all the terms get a λ-factor either from a $dx^\alpha/c\,dt$ or from a $\partial_4 h_{ab} = \lambda\,\partial_0 h_{ab}$.
So (6.20) simplifies to

$$\frac{1}{c^2}\frac{d^2x^\alpha}{dt^2} + \frac{1}{2}\lambda\,\partial_\alpha h_{44}(1 + O(\lambda)) = 0,$$

which can be written

$$\frac{d^2x^\alpha}{dt^2} = -\frac{1}{2}c^2\frac{\partial g_{44}}{\partial x^\alpha}(1 + O(\lambda)). \tag{6.21}$$

But the Newtonian equation of motion for a test-particle in a gravitational field of potential ϕ is

$$\frac{d^2x^\alpha}{dt^2} = -\frac{\partial\phi}{\partial x^\alpha}. \tag{6.22}$$

Comparing this with (6.21) and noting that at large distances from the sources of the field $\phi \to 0$ and $g_{44} \to 1$, we get

$$g_{44} = 1 + \frac{2\phi}{c^2} + O(V/c) \tag{6.23}$$

which is the promised condition.

Now let us consider the effect of a small coordinate transformation

$$x^a \to x^{a'} = x^a + \lambda \xi^a(x) \tag{6.24}$$

which preserves (6.14) and $\partial_4 h_{ab} = O(\lambda) \times \partial_\alpha h_{ab}$. We easily show that

$$g_{a'b'} = g_{ab} - \lambda(\partial_a \xi_b + \partial_b \xi_a) + O(\lambda^2) \tag{6.25}$$

where $\xi_a = \eta_{ab}\xi^b$. To preserve $\partial_4 h_{ab} = O(\lambda) \times \partial_\alpha h_{ab}$, we see that we must have $\partial_4 \xi_a = O(\lambda) \times \partial_\alpha \xi_a$. From (6.25), we thus see that the only component of the metric tensor which is unaltered to first order in λ by the transformation (6.24) is g_{44}, since this is the only one in which ξ_a occurs in (6.25) only in the form of x^4 derivatives.

We have therefore shown that the only component of the metric tensor which is well-defined to first order for a slowly varying weak gravitational field, is determined to this order by the requirement that the theory should agree with Newtonian theory to this order. It is given by

$$g_{44} = 1 + \frac{2\phi}{c^2} + O(v/c)$$

ϕ being the Newtonian gravitational potential.

6.4. STATIONARY AND STATIC GRAVITATIONAL FIELDS

A naive definition of a stationary gravitational field would be : one for which a coordinate system $\{x^a\}$ exists in which $\partial g_{ab}/\partial x^4 = 0$, x^4 being a time-like coordinate. We shall try to formulate this in a more mathematical way. In this coordinate system define a vector field ξ^a by $\xi^a = \delta_4{}^a$. Then

$$\underset{\xi}{\pounds}\, g_{ab} = \xi^c \partial_c g_{ab} + g_{ac}\partial_b \xi^c + g_{cb}\partial_a \xi^c$$

$$\Rightarrow \quad \underset{\xi}{\pounds}\, g_{ab} = \partial_4 g_{ab} \text{ on putting } \xi^a = \delta^a_4$$

$$= 0 \text{ by the stationary condition.}$$

so that the metric is invariant under dragging along by ξ^a, or in other words, ξ^a is a Killing vector field. Conversely, if a space possesses a time-like Killing vector field ξ^a, there always exists a coordiante system for which $\xi^a = (0,0,0,1)$, and then $\partial g_{ab}/\partial x^4 = 0$. We therefore say that a space-time is <u>stationary</u> if it admits a time-like Killing vector field.

A special case of a stationary space-time is one for which the trajectories of ξ^a are orthogonal to a family of hypersurfaces. Such a space-time is said to be <u>static</u>, and it implies the existence of functions χ and σ such that

$$\xi_a = \chi \frac{\partial \sigma}{\partial x^a} \tag{6.26}$$

from which we easily obtain

$$\xi_{[a}\, \partial_b \xi_{c]} = 0 \tag{6.27}$$

So for a static space-time we have a ξ^a for which

$$\nabla_a \xi_b + \nabla_b \xi_a = 0 \qquad \text{(Killing's equations)} \tag{6.28}$$

$$\xi_{[a} \nabla_b \xi_{c]} = 0 \qquad \text{(Equivalent to (6.27))} \tag{6.29}$$

From (6.28) and (6.29),

$$\xi_a \nabla_b \xi_c + \xi_b \nabla_c \xi_a + \xi_c \nabla_a \xi_b = 0$$

Multiply this by ξ^c, put $\xi^2 = \xi^c \xi_c$ and use (6.28) to give

$$\xi_a \nabla_b \xi^2 - \xi_b \nabla_a \xi^2 + \xi^2(\nabla_a \xi_b - \nabla_b \xi_a) = 0$$

and hence

$$\frac{\partial}{\partial x^a}\left(\frac{\xi_b}{\xi^2}\right) = \frac{\partial}{\partial x^b}\left(\frac{\xi_a}{\xi^2}\right)$$

Therefore there exists a function σ such that

$$\frac{\xi_a}{\xi^2} = \frac{\partial \sigma}{\partial x^a} \quad , \qquad \text{i.e.} \qquad \xi_a = \xi^2 \partial_a \sigma \tag{6.30}$$

which has the form (6.26) with $\chi = \xi^2$.
 Now choose a coordinate system in which $\xi^a = (0001) = \delta^a_4$. Then by (6.30),

$$g_{a4} = g_{ab}\, \delta^b_4 = g_{ab}\, \xi^b = \xi_a = \xi^2 \partial_a \sigma$$

and

$$\xi^2 = g_{ab}\, \delta^a_4 \delta^b_4 = g_{44}$$

Hence

$$g_{a4} = g_{44}\, \partial_a \sigma. \tag{6.31}$$

Putting a = 4, this gives

$$\partial_4 \sigma = 1 \qquad \text{and therefore} \qquad \sigma = x^4 + f(x^1, x^2, x^3)$$

for some function f. Now make a coordinate transformation

$$x^\alpha \to x^{\alpha'} = x^\alpha, \qquad \alpha = 1,2,3 \quad ; \quad x^4 \to x^{4'} = x^4 + f(x^1, x^2, x^3).$$

Then we still have

$$\xi^{a'} = \frac{\partial x^{a'}}{\partial x^b}\xi^b = \frac{\partial x^{a'}}{\partial x^4} = \delta_4^{a'}$$

and so we may suppose this to be the coordinate system used above, so that (6.31) gives

$$g_{\alpha 4} = 0 . \qquad \alpha = 1,2,3$$

We have already seen that $\xi^a = \delta_4^a$ implies $\partial_4 g_{ab} = 0$, and thus we have proved the following theorem:

In a static space-time there exists a coordinate system, which we shall say is adapted to the Killing vector field ξ^a, in which the metric is time-independent and $g_{14} = g_{24} = g_{34} = 0$.

Now let us see what freedom we have left in the coordinates. Suppose that there is only one time-like Killing vector field ξ^a, and consider two coordinate systems $\{x^a\}$ and $\{x^{a'}\}$ adapted to ξ^a. Then

$$\delta_4^a = \xi^{a'} = \frac{\partial x^{a'}}{\partial x^b}\xi^b = \frac{\partial x^{a'}}{\partial x^b}\delta_4^b = \frac{\partial x^{a'}}{\partial x^4} \qquad (6.32)$$

and $$0 = g_{\alpha 4} = \frac{\partial x^{a'}}{\partial x^\alpha}\frac{\partial x^{b'}}{\partial x^4}g_{a'b'} = \frac{\partial x^{a'}}{\partial x^\alpha}g_{a'4'} \quad \text{by (6.32).}$$

$$= \frac{\partial x^{4'}}{\partial x^\alpha} \quad \text{as} \quad g_{\alpha' 4'} = 0 . \qquad (6.33)$$

From (6.32) and (6.33) we see that the 'allowed' coordinate transformations are

$$x^\alpha \to x^{\alpha'} = f^{\alpha'}(x^\beta), \quad \alpha,\beta = 1,2,3 \ ; \ x^4 \to x^{4'} = ax^4 + b, \quad a, b, \text{ constants.}$$

If further g_{44} tends to a constant value at large distances, we may impose the further condition $g_{44} \to 1$, which fixes $a = \pm 1$, and if we do not allow reversal of the direction of time, then $a > 0$ and we must take $a = 1$. The time is then defined to within an additive constant, and is called a world-time.

If the space-time admits more than one time-like Killing vector field, the allowed coordinate transformations may be less restricted.

6.5. PROPAGATION OF LIGHT IN A GRAVITATIONAL FIELD

The propagation of light is governed by Maxwell's equations

$$\nabla_b f^{ab} = 0, \qquad \nabla_{[a} f_{bc]} = 0 \tag{6.34}$$

Put

$$f_{ab} = \mathrm{Re}(F_{ab} e^{i\psi}) \tag{6.35}$$

where F_{ab} is a complex slowly varying function of position and ψ is a rapidly varying phase factor. This represent a wave whose amplitude and polarization, given by F_{ab}, changes little over a large number of wavelengths.

Substitute (6.35) into (6.34) and neglect derivatives of F_{ab} in comparison with those of ψ. We get

$$F^{ab}\partial_b \psi = 0 \tag{6.36}$$

$$F_{[ab}\,\partial_{c]}\psi = 0 \tag{6.37}$$

which are exact in the infinite-frequency (or optical) limit. From (6.37), there exists a vector field w_a such that

$$F_{ab} = w_{[a}\,\partial_{b]}\psi$$

and on putting this into (6.36) we get

$$g^{ab}\partial_a\psi\,\partial_b\psi = 0 \tag{6.38}$$

which is called the Eikonal Equation.

Put $k_a = \partial_a \psi$,

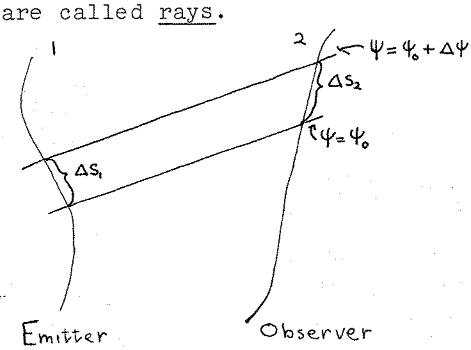
so that

$$\nabla_a k_b = \nabla_b k_a$$

Then (6.38) shows that k^a is null, and so

$$k^b \nabla_b k_a = k^b \nabla_a k_b = \tfrac{1}{2} \nabla_a (k^b k_b) = 0 \qquad (6.39)$$

Hence k^a is tangent to a null geodesic congruence; the trajectories of k^a are called rays.



Now consider the world-lines of an emitter and observer of light, as illustrated, and consider two rays of light joining them. Since $k^a \partial_a \psi = 0$, they will lie in surfaces of constant ψ, say $\psi = \psi_0$ and $\psi = \psi_0 + \Delta\psi$, with $\Delta\psi$ small. Let the interval between the points of intersection on line 1 be Δs_1, and on line 2 be Δs_2, and let $\vec{u}_1$ and $\vec{u}_2$ be the unit tangent vectors to the two lines. Then we have

$$(\partial_a \psi)_1 \Delta s_1 . u_1^a = \Delta\psi = (\partial_a \psi)_2 \Delta s_2 . u_2^a \qquad (6.40)$$

where $(\partial_a \psi)_1$, $(\partial_a \psi)_2$ are evaluated at the points of intersection on 1 and 2 respectively. But if ω_1 and ω_2 are the frequencies assigned to the light by 1 and 2 respectively, then since the surfaces ψ = const are surfaces of constant phase, (6.40) gives

$$\frac{\omega_1}{\omega_2} = \frac{\Delta s_2}{\Delta s_1} = \frac{(k_a u^a)_1}{(k_a u^a)_2} \tag{6.41}$$

This formula is valid in both special and general relativity and gives the combined Doppler and gravitational shift of spectrum lines.

For a static gravitational field we may take both world-lines to be trajectories of the Killing vector field This situation corresponds to the emitter and observer being at rest relative to the field, and the distance between them, measured along a geodesic normal to both lines, is time independent. In this case

$$k^b \nabla_b (k^a \xi_a) = (k^b \nabla_b k^a)\xi_a + k^a k^b (\nabla_b \xi_a) = 0$$

since the first term vanishes by (6.39), and the second term vanishes as $\nabla_{(a}\xi_{b)} = 0$ by Killing's equations. Hence $k^a \xi_a$ is constant along the rays. But also by construction, ξ^a is parallel to u^a, and so $\xi^a = \xi u^a$, where $\xi^2 = \xi^a \xi_a$. Equation (6.41) then gives

$$\frac{\omega_1}{\omega_2} = \frac{(k_a u^a)_1}{(k_a u^a)_2} = \frac{(\xi)_2}{(\xi)_1} .$$

Further, in a coordinate system adapted to $\vec{\xi}$, $\xi^2 = g_{44}$. So now

$$\frac{\omega_1}{\omega_2} = \left\{ \frac{(g_{44})_2}{(g_{44})_1} \right\}^{1/2} . \tag{6.42}$$

This is an exact formula for a static space-time. If we now suppose the gravitational field to be weak, and use the Newtonian approximation given by (6.23), (6.42) gives for the gravitational frequency shift

$$\frac{\Delta\omega}{\omega} = -\frac{\Delta\phi}{c^2} . \tag{6.43}$$

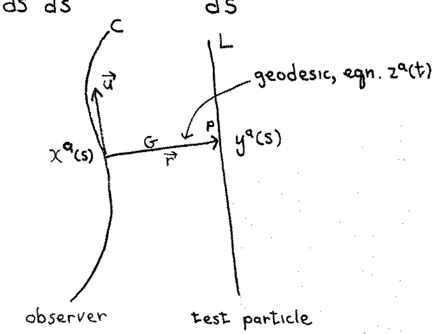
We note that this result has been obtained without using any field equations. It has recently been tested experimentally, to quite high accuracy, in the Earth's gravitational field by using the Mössbauer effect to obtain very narrow spectral lines.

We have been able to separate the gravitational shift from a Doppler shift in a static space-time, because then a meaning can be attached to two observers spatially separated being at rest. In a general space-time this is not so, and the gravitational and Doppler shifts are not so separable.

6.6. LOCAL REFERENCE FRAMES AND FERMI TRANSPORT

We shall now investigate the local reference frames that an observer may use to describe the behavior of matter in his neighborhood. Consider an observer O moving along an arbitrary time-like world-line C, of equation $x^a = x^a(s)$, with s being the proper time measured along C. Suppose he observes a free test particle P in his neighborhood, whose world-line L has equation $x^a = y^a(s)$, where $y^a(s)$ is the point of intersection of L with the instantaneous 3-space of observer through $x^a(s)$. s will not, in general, be an affine parameter on L, and so the geodesic equation for L takes the form

$$\frac{D}{ds}\frac{dy^a}{ds} = \lambda(s)\frac{dy^a}{ds}. \tag{6.44}$$



Now O will describe the position of P by a vector $\vec{r}$ at O whose direction is that of the geodesic G joining O and P, and whose length is the length of that geodesic. Let G have equation $x^a = z^a(t)$, with t the distance measured along G from O, and

$$z^a(0) = x^a(s) \qquad z^a(t_1) = y^a(s) \tag{6.45}$$

Then

$$\frac{d^2z^a}{dt^2} + \Gamma^a_{bc}\frac{dz^b}{dt}\frac{dz^c}{dt} = 0 . \tag{6.46}$$

Expand $z^a(t_1)$ in a Taylor series about $t = 0$:

$$z^a(t_1) = z^a(0) + \left(\frac{dz^a}{dt}\right)_0 t_1 + \frac{1}{2}\left(\frac{d^2z^a}{dt^2}\right)_0 t_1^2 + O(t_1^3) \tag{6.47}$$

Then using (6.45) and (6.46), and noting that

$$r^a = \left(\frac{dz^a}{dt}\right)_0 t_1 , \tag{6.48}$$

(6.47) gives

$$y^a(s) = x^a(s) + r^a - \frac{1}{2}\Gamma^a_{bc}(x^a)r^b r^c + O(r^3) . \tag{6.49}$$

Differentiating twice with respect to s gives

$$\frac{dy^a}{ds} = \frac{dx^a}{ds} + \frac{dr^a}{ds} + O(r) \tag{6.50}$$

$$\frac{d^2y^a}{ds^2} = \frac{d^2x^a}{ds^2} + \frac{d^2r^a}{ds^2} - \Gamma^a_{bc}(x^a)\frac{dr^b}{ds}\frac{dr^c}{ds} + O(r) \tag{6.51}$$

Now remember that dy^a/ds is a vector defined on L, and dx^a/ds and r^a are vectors defined on C. Then we have

$$\frac{D}{ds}\frac{dy^a}{ds} = \frac{d^2y^a}{ds^2} + \Gamma^a_{bc}\frac{dy^b}{ds}\frac{dy^c}{ds}$$

$$\frac{D}{ds}\frac{dx^a}{ds} = \frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds}$$

$$\frac{Dr^a}{ds} = \frac{dr^a}{ds} + \Gamma^a_{bc}\frac{dx^b}{ds}r^c$$

$$\frac{D^2r^a}{ds^2} = \frac{d^2r^a}{ds^2} + 2\Gamma^a_{bc}\frac{dx^b}{ds}\frac{dr^c}{ds} + O(r)$$

Using these and (6.50) we easily see that (6.51) can be put in the form

$$\frac{D}{ds}\frac{dy^a}{ds} = \frac{D}{ds}\frac{dx^a}{ds} + \frac{D^2 r^a}{ds^2} + O(r). \tag{6.52}$$

Now put

$$u^a \stackrel{\text{def}}{=} \frac{dx^a}{ds} \tag{6.53}$$

and use (6.44) and (6.50) to write (6.52) in the form

$$\ddot{\vec{r}} = \lambda(s)(\vec{u} + \dot{\vec{r}}) - \dot{\vec{u}} + O(r) \tag{6.54}$$

where $^\bullet$ denotes D/ds. To determine $\lambda(s)$ we use the condition that $y^a(s)$ lies in the instantaneous 3-space of O, by which we mean

$$\vec{u}\cdot\vec{r} = 0 \tag{6.55}$$

By differentiation this gives

$$\dot{\vec{u}}\cdot\vec{r} + \vec{u}\cdot\dot{\vec{r}} = 0, \qquad \ddot{\vec{u}}\cdot\vec{r} + 2\dot{\vec{u}}\cdot\dot{\vec{r}} + \vec{u}\cdot\ddot{\vec{r}} = 0 \tag{6.56}$$

and

$$\vec{u}\cdot\vec{u} = 1 \qquad \text{gives} \qquad \vec{u}\cdot\dot{\vec{u}} = 0. \tag{6.57}$$

Scalarly multiply (2.54) by $\vec{u}$ and use (6.56) and (6.57) to give

$$\lambda(s) = -2\dot{\vec{u}}\cdot\dot{\vec{r}} + O(r)$$

which, put back into (6.54), gives

$$\ddot{\vec{r}} = -\dot{\vec{u}} - 2\dot{\vec{u}}\cdot\dot{\vec{r}}(\vec{u} + \dot{\vec{r}}) + O(\vec{r}). \tag{6.58}$$

This is the equation of motion of P as seen by O. We shall use it below only in the case $\vec{r} = 0$, i.e. P coincides instantaneously with O, and in this case it is exact. A more detailed investigation of the terms of O(r) shows that they are negligible if $|\vec{r}|$ is small compared to the radius of curvature of space, and $|\dot{\vec{r}}| \ll 1$, i.e. the velocity of the particle is small compared with that of light.

Now suppose O refers $\vec{r}$ to an orthonormal basis of vectors $\vec{e}_\mu$, $\mu = 1,2,3$, spanning his instantaneous 3-space. Write $\vec{e}_4 = \vec{u}$, and let p,q = 1,2,3,4. Then

$$\vec{e}_p \cdot \vec{e}_q = \eta_{pq}, \qquad \text{where } \eta_{pq} = \mathrm{diag}(-1,-1,-1,+1) = \eta^{pq}. \tag{6.59}$$

Write

$$\vec{e}^{\,p} \overset{\text{def}}{=} \eta^{pq}\vec{e}_q, \quad \omega_{\mu\nu} \overset{\text{def}}{=} \dot{\vec{e}}_\mu \cdot \vec{e}_\nu, \quad r^\mu \overset{\text{def}}{=} \vec{r}\cdot\vec{e}^{\,\mu}, \quad \kappa \overset{\text{def}}{=} \dot{\vec{u}}\cdot\vec{e}^{\,\mu} \tag{6.60}$$

Then $\omega_{\mu\nu} = -\omega_{\nu\mu}$ from differentiating (6.59). Also, since r^μ etc are scalars, we shall write $\dot{r}^\mu$ to mean dr^μ/ds and not $\dot{\vec{r}}\cdot\vec{e}^{\,\mu}$. Now

$$\begin{aligned} \dot{\vec{e}}^{\,\mu} &= (\dot{\vec{e}}^{\,\mu}\cdot\vec{e}^{\,\nu})\vec{e}_\nu + (\dot{\vec{e}}^{\,\mu}\cdot\vec{u})\vec{u} \\ &= \omega^{\mu\nu}\vec{e}_\nu - (\dot{\vec{u}}\cdot\vec{e}^{\,\mu})\vec{u} \\ &= \omega^\mu{}_\nu \vec{e}^{\,\nu} - \kappa^\mu \vec{u}. \end{aligned} \tag{6.61}$$

Also

$$\begin{aligned} \dot{r}^\mu &= \dot{\vec{r}}\cdot\vec{e}^{\,\mu} + \vec{r}\cdot\dot{\vec{e}}^{\,\mu} \\ &= \dot{\vec{r}}\cdot\vec{e}^{\,\mu} + \omega^\mu{}_\nu \vec{e}^{\,\nu} \end{aligned} \tag{6.62}$$

and differentiating again gives

$$\ddot{r}^\mu = \ddot{\vec{r}}\cdot\vec{e}^{\,\mu} - \omega^\mu{}_\nu\omega^\nu{}_\rho r^\rho + \dot{\omega}^\mu{}_\nu r^\nu + 2\omega^\mu{}_\nu \dot{r}^\nu \tag{6.63}$$

But from (6.58) and (6.60) we easily get

$$\ddot{\vec{r}}\cdot\vec{e}^{\mu} = -K^{\mu} - 2(\quad{}_{\nu}\dot{r}^{\nu})\dot{r}^{\mu} + O(r) \tag{6.64}$$

and hence from (6.63) and (6.64), at r = 0 we have

$$\ddot{r}^{\mu} = -K^{\mu} - 2\dot{r}^{\mu}(\quad{}_{\nu}\dot{r}^{\nu}) + 2\omega^{\mu}{}_{\nu}\dot{r}^{\nu} \tag{6.65}$$

which relates the velocity and acceleration of the particle, referred to the axes $\vec{e}_{\mu}$, as it passes through O. We note that it is composed of three parts:

(i) $-K^{\mu}$, independent of $\dot{r}^{\mu}$, which can be determined using a particle at rest.

(ii) $-2\dot{r}^{\mu}(K_{\nu}\dot{r}^{\nu})$ parallel to the velocity and proportional to (velocity)^2.

(iii) $2\omega^{\mu}{}_{\nu}\dot{r}^{\nu}$, which depends on the motion of the axes and has the structure of a Coriolis force, proportional to and perpendicular to the velocity.

We thus see that $\omega^{\mu\nu}$ behaves as an angular velocity of the axes of the reference frame, and that the axes are the nearest that can be obtained to Newtonian non-rotating axes when $\omega^{\mu\nu} = 0$. This is further confirmed if we consider the extra terms that arise in (6.65) when we take the case $r \neq 0$. For then from (6.63) we see that we get on the right-hand side of (6.65) the additional terms

$$-\omega^{\mu}{}_{\nu}\,\omega^{\nu}{}_{\rho}r^{\rho} + \dot{\omega}^{\mu}{}_{\nu}r^{\nu}$$

among others. The first of these is the centrifugal force associated with axes of angular velocity $\omega^{\mu\nu}$ and confirms our interpretation above of $\omega^{\mu\nu}$. The second is also familiar as the inertial force associated with non-uniformly rotating axes. All the other terms vanish in the limit of flat space-time and velocities small compared with that of light.

We now show that the axes can always be chosen so that $\omega^{\mu\nu} = 0$. For (6.61) and (6.60) give $\omega^{\mu\nu} = 0$ as equivalent to

$$\dot{\vec{e}}^{\mu} = -(\dot{\vec{u}}\cdot\vec{e}^{\mu})\vec{u} \tag{6.66}$$

which is a set of ordinary differential equations, which always have a solution for given initial $\vec{e}^{\mu}$ at any point of C. Moreover, these equations preserve orthonormality of the tetrad $(\vec{u}, \vec{e}_{\mu})$ defined by (6.66).

A vector $\vec{w}$ defined on C whose components referred to the tetrad $\vec{e}_p$ are independent of s is said to be Fermi-Walker propagated along C. Let $w^p = \vec{w}\cdot\vec{e}^p$. Then we have

$$\frac{D\vec{w}}{ds} = w^a \frac{D\vec{e}_a}{ds} = (\vec{w}\cdot\vec{u})\dot{\vec{u}} + w^{\mu}\frac{D\vec{e}_{\mu}}{ds} \quad \text{as } \vec{e}_4 = \vec{u}$$
$$= (\vec{w}\cdot\vec{u})\dot{\vec{u}} - w^{\mu}(\dot{\vec{u}}\cdot\vec{e}_{\mu})\vec{u} \qquad \text{by (6.66)}$$

and since $\quad w^{\mu}\vec{e}_{\mu} = \vec{w} - (\vec{w}\cdot\vec{u})\vec{u},\quad$ this gives

$$\frac{D\vec{w}}{ds} = (\vec{w}\cdot\vec{u})\dot{\vec{u}} - (\vec{w}\cdot\dot{\vec{u}})\vec{u} \tag{6.67}$$

as the condition for $\vec{w}$ to be Fermi-Walker propagated along C. From (6.67) one may easily verify that $\vec{u}$ itself is Fermi-Walker propagated, and that Fermi-Walker propagation is metric, i.e. it preserves scalar products. Consequently if $\vec{w}\cdot\vec{u} = 0$ at one point on C, it remains so under Fermi-Walker propagation. Then (6.67) reduces to

$$\frac{D\vec{w}}{ds} = -(\vec{w}\cdot\dot{\vec{u}})\vec{u} \tag{6.68}$$

which is known as Fermi propagation. Finally, we note that Fermi-Walker propagation along a geodesic ($\dot{\vec{u}} = 0$) reduces to parallel propagation.

We have thus shown that Fermi propagated axes are the nearest that can be obtained in curved space-time to a Newtonian non-rotating reference frame, and that such axes can be determined by local dynamical experiments.

The vector $\dot{\vec{u}}$ is normal to the curve C, and is called the curvature vector of C. If we write $\dot{\vec{u}} = K\vec{n}$, $|\vec{n}| = 1$, $K > 0$, $\vec{n}$ is called the normal (or first normal) to C, and K the first curvature.

Now consider (6.65) with Fermi propagated axes for an observer on the Earth. Then $\ddot{r}^{\mu} = g^{\mu}/c^2$, where g^{μ} is the gravitational acceleration, the c^2 arising from s = ct. So (6.65) gives $K^\mu = -g^\mu/c^2$, and the term $-2\dot{r}^\mu(K_\nu \dot{r}^\nu)$ is easily seen to be negligible for velocities $\ll$ c. Since $k^\mu = \vec{u}\cdot\vec{e}^{\,\mu}$, we see that the first curvature of our world-lines on the Earth is

$$K = \frac{g}{c^2} \simeq 1.09 \times 10^{-18}\ \mathrm{cm}^{-1}.$$

6.7. THE PHYSICAL DISTINCTION BETWEEN STATIC AND STATIONARY SPACE-TIMES

As we have seen in section 6.4, a stationary space-time is one containing a time-like Killing vector field $\vec{\xi}$, and correspondingly it has a privileged class of observers, which we shall call Copernican observers, whose world-lines are the trajectories of $\vec{\xi}$. Now we have seen in the preceding section that an observer can by local dynamical experiments determine a Fermi propagated local reference frame which is the nearest he can obtain to a Newtonian non-rotating frame. But in a stationary space-time a Copernican observer has an alternative criterion for non-rotation, namely he may consider a frame as non-rotating when neighboring Copernican observers appear to be at rest in that frame. We shall investigate under what conditions these two criteria agree.

At one point let a Copernican observer C set up an orthonormal reference triad $\vec{e}_\mu$ of vectors normal to $\vec{\xi}$, and let them be propagated along his world-line by dragging along by $\vec{\xi}$, i.e. so that

$$\underset{\vec{\xi}}{\pounds}\, \vec{e}_\mu = 0. \tag{6.69}$$

Clearly this is the condition that neighbouring Copernican observers appear to be at rest in the reference frame. Also, since

$$\underset{\vec{\xi}}{\pounds}\, g_{ab} = 0 \ (\text{Killing's equations}) \quad \text{and} \quad \underset{\vec{\xi}}{\pounds}\, \vec{\xi} = 0, \tag{6.70}$$

we see that (6.69) preserves the orthonormality of the axes and their orthogonality to $\vec{\xi}$. Now using the notation of section 6.6 we have $\vec{\xi}$ parallel to $\vec{u}$ for a Copernican observer,

$$\vec{\xi} = \xi \vec{u}, \qquad \xi \overset{\text{def}}{=} |\vec{\xi}|. \tag{6.71}$$

But (6.69) gives

$$\xi^b \nabla_b e^a_\mu - e^b_\mu \nabla_b \xi^a = 0, \tag{6.72}$$

and using (6.71) this becomes

$$\xi \dot{e}^a_\mu - e^b_\mu \nabla_b \xi^a = 0. \tag{6.73}$$

Hence

$$\omega_{\mu\nu} = \dot{\vec{e}}_\mu \cdot \vec{e}_\nu = \xi^{-1} e^a_\nu e^b_\mu \nabla_b \xi_a \tag{6.74}$$

where the skew-symmetry of $\omega_{\mu\nu}$ is maintained by Killing's equations. Also,

$$3\xi^c e^a_\nu e^b_\mu \xi_{[a} \nabla_b \xi_{c]} = \xi^2 e^a_\nu e^b_\mu \nabla_{[a} \xi_{b]}$$

and comparing with (6.74) this gives, on using (6.71),

$$\omega_{\mu\nu} = 3\xi^{-2} e^a_\mu e^b_\nu u^c \xi_{[a} \nabla_b \xi_{c]}, \tag{6.75}$$

from which $\omega_{\mu\nu} = 0$ if and only if

$$\xi_{[a} \nabla_b \xi_{c]} = 0 .$$

From (6.29), this is the condition for a stationary space-time to be static. But $\omega_{\mu\nu} = 0$ is the condition in which the frame is Fermi transported, and hence neighboring Copernican observers in a stationary space-time appear at rest in a Fermi transported local reference frame if and only if the space-time is static.

On a cosmological scale this implies that dynamical and astronomical criteria of non-rotation agree in a stationary universe if and only if it is also static.

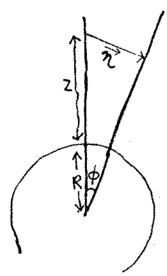
6.8. EVALUATION OF THE CURVATURE TENSOR AT THE EARTH'S SURFACE

We have already seen that agreement with Newtonian theory for a weak slowly varying gravitational field requires $g_{44} = 1 + 2\phi/c^2$. We shall now see that certain components of the curvature tensor may be similarly determined. For this we require the equation of geodesic deviation, which was proved in Prof. Pirani's lectures, and which states that if we have two neighboring, almost parallel geodesics and η^a is the orthogonal connecting vector joining them, then

$$\frac{D^2\eta^a}{ds^2} - R^a{}_{bcd}\, u^b \eta^c u^d = 0 , \qquad (6.76)$$

where u^a is the unit tangent vector to one geodesic and s is the proper time (distance) measured along it. We shall consider two cases for freely falling bodies near the Earth's surface:

Case 1. Slight horizontal separation.



Choose locally Minkowski axes with origin on the Earth's surface, z-axis vertically upwards and x,y axes horizontal. Consider two bodies falling freely toward the center of the Earth, one down the z-axis and one down a line in the (x,z) plane making a small angle ϕ with the z-axis, both being released from rest at the same time t = 0 from the same height z_o. Let $x^o = ct$. Then since the velocities of the bodies are $\ll$ c, we have

$$u^a \simeq (0,0,0,1)$$

$$\eta^a \simeq (\eta,0,0,0)$$

with

$$\eta \cong (R+z)\phi$$

and

$$z = z_0 - \tfrac{1}{2}gt^2$$

where R is the Earth's radius and g is the acceleration of gravity at the surface of the Earth. Then

$$\frac{D^2\eta^a}{ds^2} \doteq \left(\frac{1}{c^2}\frac{d^2\eta}{dt^2}, 0, 0, 0\right) \doteq \left(-\frac{g\phi}{c^2}, 0, 0, 0\right)$$

and so substituting into (6.76) we get approximately

$$(-g\phi/c^2, 0, 0, 0) - (R^1{}_{4'4}, R^2{}_{424}, R^3{}_{424}, 0)(R+z)\phi = 0 \tag{6.77}$$

But if M is the mass of the Earth and k the constant of gravitation, then

$$g = kM/R^2$$

and so putting z = 0 in (6.77) gives

$$R_{1414} \doteq \frac{kM}{c^2R^3} \qquad R_{2414} \doteq 0 \qquad R_{3414} \doteq 0$$

where we have lowered the first index with $g_{ab} \doteq \eta_{ab}$. By symmetry we clearly also have

$$R_{2424} = R_{1414} \qquad\qquad R_{2434} \doteq 0$$

Case 2. Slight vertical separation.

We now take both bodies falling freely down the z-axis, released from heights z_o and $z_o + \delta z$ respectively at t = 0, with δz small. Then

$$u^a \doteq (0,0,0,1)$$

$$\eta^a \doteq (0,0,\eta,0)$$

Now the gravitational acceleration at height z is

$$g(z) = \frac{kM}{(R+z)^2} = \frac{kM}{R^2} - \frac{2kMz}{R^3} + O(z^2), \qquad z \ll R$$

and hence

$$\eta(t) = \left\{ z_0 + \delta z - \frac{1}{2}\left(\frac{kM}{R^2} - \frac{2kM(z_0+\delta z)}{R^3}\right)t^2 + O(t^3)\right\}$$

$$- \left\{ z_0 - \frac{1}{2}\left(\frac{kM}{R^2} - \frac{2kMz_0}{R^3}\right)t^2 + O(t^3)\right\}$$

$$= \delta z + \frac{kM\delta z}{R^3} t^2 + O(t^3).$$

Therefore

$$\frac{D^2\eta^a}{ds^2} \doteq \left(0, 0, \frac{2kM\delta z}{R^3c^2}, 0\right)$$

and so (6.76) takes the form

$$\left(0, 0, \frac{2kM\delta z}{R^3}, 0\right) - (R'_{434}, R^2_{434}, R^3_{434}, 0)\delta z = 0$$

Hence

$$R_{1434} \doteq 0, \qquad R_{2434} \doteq 0 \qquad R_{3434} \doteq -\frac{2kM}{c^2R^3}.$$

So finally we have

$$R_{\alpha 4 \beta 4} \doteq 0 \ \text{ if } \alpha \neq \beta, \quad R_{1414} = R_{2424} \doteq \frac{kM}{c^2R^3}, \quad R_{3434} \doteq -\frac{2kM}{c^2R^3}$$

at the Earth's surface, to first order in k.

We note from this that to first order in k, the Ricci tensor, $R_{ab} \overset{\text{def}}{=} R^c{}_{abc}$, satisfies

$$R_{44} = R_{1414} + R_{2424} + R_{3434} = 0 \tag{6.78}$$

6.9. THE GRAVIATIONAL FIELD EQUATIONS

We have so far not decided what the gravitational field equations should be in the presence of matter, but have suggested that in the absence of matter

$$R_{ab} = 0 \tag{6.79}$$

This is supported by our approximate calculation above, which shows that agreement with Newtonian theory near the Earth's surface requires $R_{44} = 0$ to first order, but it does not necessitate (6.79) even to first order in general, as our result may be due to the special symmetry of the example.

Since Newtonian theory is described by Poisson's equation, a second order differential equation for the potential ϕ in terms of the mass density ρ, and we have seen that $g_{44} \doteq 1 + 2\phi/c^2$ for weak fields, agreement with Newtonian theory for weak fields strongly suggests that our relativistic equations should be second order differential equations in the metric g_{ab} with mass occurring as the source of the field. But because special relativity has shown us that mass and energy are equivalent forms of the same entity, this suggests that not only substantial matter but all forms of energy should be included in the source terms. We also know from special relativity that the way covariantly to describe an energy density is by means of an energy-momentum tensor T^{ab}, and so we now expect T^{ab} to be the source of the gravitational field.

To continue, we must resort to less strong arguments. Poisson's equation involves both ρ and the second derivatives of ϕ linearly. We may therefore expect the relativistic equations to involve T^{ab} (remembering that $T^{44} \sim \rho$) and the second derivatives of g_{ab} linearly. But we also require that the theory should be formulated in terms of tensors, and the only tensors involving T^{ab} linearly are T^{ab} and $T \overset{\text{def}}{=} T^a{}_a$, unless we are prepared to go to higher than second rank tensors (e.g. $g_{ab}\, T_{cd}$). It can also be shown that the only second rank tensors that can be constructed out of g_{ab} and its first and second derivatives, and which contain the second derivatives linearly, are R_{ab} and Rg_{ab}, where R_{ab} is the Ricci tensor and $R \overset{\text{def}}{=} R^a{}_a$ is the curvature scalar. We see that two possibilities are open to us, either a scalar or a second rank tensor theory. A vector theory may be ruled out also on the ground that it would, like Maxwell's theory, lead to a repulsion between two masses of the same sign.

For a scalar theory we see that the only possibility is to take

$$R = \lambda T$$

where λ is a constant determined by requiring Newtonian theory to hold in the weak field limit. This, being only one equation, is not sufficiently restrictive to determine the metric uniquely even with suitable boundary conditions, and it must be supplemented by further conditions. The simplest possibility is to require that the space be conformally flat, i.e. that there exist a scalar field $\phi(x)$ such that the metric can be put in the form

$$g_{ab} = e^{2\phi} \eta_{ab}$$

where η_{ab} is the Minkowski metric tensor η_{ab} = diag.(-1,-1,-1,+1). This condition is expressed by the vanishing of the Weyl tensor

$$C^{a}{}_{bcd} = 0$$

and was proposed by Nordström[1] before the advent of general relativity. However, it may be ruled out on experimental grounds if we accept the experimental tests of general relativity, as it gives a wrong value for the precession of the perihelion of Mercury, and it gives no deflection of light by a gravitational field.

For a tensor theory we can take

$$R^{ab} + a g^{ab} R + b g^{ab} = c T^{ab} \tag{6.80}$$

where for given a and b, c would be determined by correspondence with Newtonian theory. To determine a, we remember that in special reltivity T^{ab} satisfies a conservation law

$$\partial_b T^{ab} = 0.$$

We may generalize this to general relativity by requiring that

1. G. Nordström, Ann. der Physik 42, 533 (1913).

$$\nabla_b(R^{ab} - \tfrac{1}{2}Rg^{ab}) = 0 \tag{6.82}$$

and so taking the divergence of (6.80) and using (6.81) and (6.82) gives

$$(a+\tfrac{1}{2})\nabla_a R = 0 \tag{6.83}$$

But if $\nabla_a R = 0$, contraction and differentiation of (6.80) gives $\nabla_a T = 0$, which we know to be false as T is not an absolute constant. So (6.83) gives $a = -\frac{1}{2}$, and (6.80) becomes

$$R^{ab} - \tfrac{1}{2}Rg^{ab} + bg^{ab} = cT^{ab} \tag{6.84}$$

The term bg^{ab} is the so-called <u>cosmological term</u> originally introduced by Einstein in an abortive attempt to prevent the field equations having an empty space solution, and now usually omitted for simplicity. c is given by correspondence with Newtonian theory to be $c = -8\pi k/c^4$, k being the Newtonian gravitational constant. Equation (6.84) then becomes the standard form of the field equations in general relativity, namely

$$R^{ab} - \tfrac{1}{2}Rg^{ab} = -\frac{8\pi k}{c^4}T^{ab}. \tag{6.85}$$

Einstein also considered another possibility, the case $a = -\frac{1}{4}$, $b = 0$. This arose in his attempts to describe all matter by means of fields and singularities in the fields. The only quantity he was willing to put on the right-hand side of the field equations was the Maxwell stress-energy tensor of the electromagnetic field, which has zero trace. All other matter was to be described by singularities in these two fields. His field equations were then

$$R^{ab} - \tfrac{1}{2}Rg^{ab} = \text{const} \times T^{ab}_{\text{maxwell}}$$

in which both sides identically have zero trace. The attempt did not lead anywhere, and we shall not consider it further.

Many more possibilities present themselves if we are willing to drop either the linearity of the field equations in the second derivatives of g_{ab} and in T^{ab}, or if we are prepared to use equations of higher than second order, neither of which possibilities can be ruled out a priori. Such field equations consistent with the conservation equation (6.81) can be conveniently obtained from an action principle. Consider an action

$$I = \int L(g_{ab}, \partial_a g_{bc}, \partial_a \partial_b g_{cd}, \ldots .) \, d^4x$$

where the only requirement on L is that it be a scalar density depending on its arguments in the same way in all coordinate systems. Then if we define the variational derivative

$$G^{ab} \overset{\text{def}}{=} \frac{\delta I}{\delta g_{ab}}$$

we shall see in the next chapter that it is a tensor which identically satisfies

$$\nabla_b G^{ab} = 0$$

and therefore field equations of the form

$$G^{ab} = \lambda T^{ab}$$

are consistent with (6.81). Our field equations (6.85) can be so obtained by taking $L = R\sqrt{-g}$. Eddington considered taking $L = R^{abcd} R_{abcd} \sqrt{-g}$, and many other possibilities have been suggested, but none has proved superior to conventional general relativity.

6.10. THE PERIHELION PRECESSION

The three discrepancies between Newtonian theory and general relativity that have been subject to experimental test are

(i) The gravitational red shift of spectrum lines.

(ii) Secular motion of the perihelion of planetary orbits.

(iii) The deflection of light rays by a gravitational field.

The first of these has been shown above to follow from correspondence with Newtonian theory; the other two involve more details of the structure of the theory.

To treat the secular motion of planetary orbits, we assume that the planets may be treated as test-particles in the sun's gravitational field, and so move on geodesics. The equations of these geodesics are conveniently found by using the theory of the Hamilton-Jacobi equation which, for the geodesics of a metric with components $g_{ab}(x)$, is

$$g^{ab}\frac{\partial S}{\partial x^a}\frac{\partial S}{\partial x^b} = m^2, \quad m \text{ constant.} \tag{6.86}$$

We now prove two theorems which we shall require:

A. If $S(x^a)$ is a solution of (6.86), then the curves $x^a = x^a(s)$ given by

$$m\frac{dx^a}{ds} = g^{ab}\frac{\partial S}{\partial x^b} \tag{6.87}$$

are geodesic lines, s being the proper time (distance) measured along them.

We first prove that s is the proper time on the curves. For from (6.87) we have

$$m^2 g_{ab}\frac{dx^a}{ds}\frac{dx^b}{ds} = g^{ab}\frac{\partial S}{\partial x^a}\frac{\partial S}{\partial x^b} = m^2 \text{ by (6.86)}$$

and hence

$$ds^2 = g_{ab}\,dx^a dx^b$$

as required. Then, to prove that the curves are geodesics, we must show that

$$\frac{D}{ds}\left(\frac{dx^a}{ds}\right) = 0$$

Now from (6.87) we have

$$m\frac{D}{ds}\left(\frac{dx^a}{ds}\right) = \frac{D}{ds}\left(g^{ab}\frac{\partial S}{\partial x^b}\right)$$

$$= g^{ab}\frac{dx^c}{ds}\nabla_c\nabla_b S, \text{ using } \frac{D}{ds} \equiv \frac{dx^c}{ds}\nabla_c$$

$$= m^{-1}g^{ab}g^{cd}(\nabla_d S)(\nabla_c\nabla_b S) \quad \text{by (6.87)}$$

$$= \tfrac{1}{2}m^{-1}g^{ab}\nabla_b(g^{cd}\partial_c S.\partial_d S)$$

$$= 0 \quad \text{by (6.86)}$$

which completes the proof.

B. If $S(a,x)$, for some range of a continuous parameter $\underline{a}$, is a family of solutions of (6.86), then

$$\frac{\partial S}{\partial a} = b \qquad b \text{ constant},$$

is a first integral of the geodesic equation.

We have to prove that along a geodesic, $\frac{d}{ds}\left(\frac{\partial S}{\partial a}\right) = 0$. Now

$$\frac{d}{ds}\left(\frac{\partial S}{\partial a}\right) = \frac{\partial^2 S}{\partial x^a \partial a}\frac{dx^a}{ds}$$

$$= \frac{\partial^2 S}{\partial a\partial x^a}\cdot\frac{1}{m}g^{ab}\frac{\partial S}{\partial x^b} \quad \text{by (6.87)}$$

$$= \frac{1}{2m}\frac{\partial}{\partial a}\left(\frac{\partial S}{\partial x^a}g^{ab}\frac{\partial S}{\partial x^b}\right)$$

$$= 0 \quad \text{by (6.86)}$$

which completes the proof.

Another result, which we shall not require here but which is of the same type as these, is that if ξ^a is a Killing vector field, then $\xi_a dx^a/ds = a$, const., is a first integral of the geodesic equation. For, along a geodesic,

$$\frac{d}{ds}\left(\xi_a \frac{dx^a}{ds}\right) = \frac{dx^b}{ds} \nabla_b \left(\xi_a \frac{dx^a}{ds}\right)$$

$$= (\nabla_b \xi_a)\frac{dx^a}{ds}\frac{dx^b}{ds} + \xi_a \frac{D}{ds}\frac{dx^a}{ds}$$

$$= 0$$

since the first term vanishes because of Killing's equations and the second term by the geodesic equation.

We now apply this to the theory of planetary orbits. Assuming the sum to be spherically symmetric, its field is given by the Schwarzschild metric, which in units with c = 1 is

$$ds^2 = \left(1 - \frac{2kM}{r}\right)dt^2 - \frac{dr^2}{1 - \frac{2kM}{r}} - r(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{6.88}$$

where M is the mass of the sun. This follows directly from the assumptions of empty space ($R_{ab} = 0$) and spherical symmetry. The Hamilton-Jacobi equation (6.86) for this metric is

$$\left(1 - \frac{2kM}{r}\right)^{-1}\left(\frac{\partial S}{\partial t}\right)^2 - \left(1 - \frac{2kM}{r}\right)\left(\frac{\partial S}{\partial r}\right)^2 - \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 - \frac{1}{r^2\sin^2\theta}\left(\frac{\partial S}{\partial \phi}\right)^2 = m^2 \tag{6.89}$$

The coordinates ϕ and t are cyclic, i.e. do not occur in the coefficients in (6.89), and hence they occur linearly with constant coefficients in the solution for S. Further, by the spherical symmetry of the problem the orbits will be plane. We consider those orbits lying in the plane $\theta = \pi/2$. S now has the form

$$S = -Et + J\phi + W(r, J, E) \tag{6.90}$$

where E and J are parameters, and substitution into (6.89) gives for W the equation

$$\left(1 - \frac{2kM}{r}\right)\frac{\partial W}{\partial r} = +\left\{E^2 - \left(m^2 + \frac{J^2}{r^2}\right)\left(1 - \frac{2kM}{r}\right)\right\}^{1/2} \tag{6.91}$$

where in taking the square root we have arbitrarily chosen the positive sign for convenience. Now by theorem B above, since J is a continuous parameter in (6.90), $\partial S/\partial J$ = const is a first integral of the geodesic equation, and by (6.90) this becomes

$$\phi = -\frac{\partial W}{\partial J} + \text{const.} \tag{6.92}$$

From (6.91),

$$\frac{\partial^2 W}{\partial J \partial r} = -\frac{J}{r^2}\left\{E^2 - \left(m^2 + \frac{J^2}{r^2}\right)\left(1 - \frac{2kM}{r}\right)\right\}^{-1/2}. \tag{6.93}$$

If we integrate this and put it into (6.92) we have

$$\phi = \int \frac{J\,dr}{r^2\left\{E^2 - \left(m^2 + \frac{J^2}{r^2}\right)\left(1 - \frac{2kM}{r}\right)\right\}^{1/2}} \tag{6.94}$$

Now consider a Newtonian mechanics problem of a particle of mass m in a spherically symmetric potential V(r). Use spherical polar coordinates (r, θ, ϕ) and consider the orbit in plane $\theta = \pi/2$. The energy and angular momentum integrals are then

$$\tfrac{1}{2} m \left\{\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2\right\} + V(r) = \mathcal{E}, \text{ constant}$$

and

$$r^2 \frac{d\phi}{dt} = J, \quad \text{constant}$$

which integrate to give

$$\phi = \int \frac{J\,dr}{r^2(2m\mathcal{E} - 2mV - J^2/r^2)^{1/2}}. \tag{6.95}$$

Comparing this with (6.94) we see that (6.94) and (6.95) are identical if

$$2m\mathcal{E} = E^2 - m^2, \qquad V(r) = -\frac{kMm}{r} - \frac{kMJ^2}{mr^3} \tag{6.96}$$

We have thus reduced the problem to one of Newtonian mechanics with a potential given by (6.96). The term -kMm/r in V(r) is just the Newtonian gravitational potential, for which the planetary orbits are ellipses. The other term in V(r) is very small, and has the effect of producing a secular precession of the perihelion of magnitude

$$\Delta\phi = \frac{6\pi k M}{c^2 a(1-e^2)}$$

radians per revolution, where we have reinserted the velocity of light c, and

M is the mass of the sun,
k the Newtonian gravitational constant
a the semi-major axis of the orbit,
e the eccentricity of the orbit.

The deflection of light by the sun may formally be obtained by putting m = 0 in (6.94). For a light ray which passes the sun at a distance ρ from its center, we find a deflection of magnitude

$$\Delta\phi = \frac{4kM}{c^2\rho}.$$

For an account of the observational data on the experimental tests, see the references at the end of this chapter. For the perihelion precession of Mercury it is found that after allowing for the perturbations of the orbit caused by the other planets, there remains a precession of about 43" per century. This is in very good agreement with the precession predicted by general relativity, but there remains the possibility that some effect may have been overlooked that could account for the precession on other grounds. The evidence for the deflection of light by the sun is much less conclusive, as observations can only be made at a total solar eclipse and are very difficult to make even them. Undoubtedly some effect is present, of the order of magnitude predicted by general relativity, but it is difficult to say much more than this.

References:

W. Pauli, Theory of Relativity. Pergamon Press, London (1958).
L. D. Landau and E. M. Lifshitz, The Classical Theory of Fields. (2nd ed.). Addison-Wesley (1961).

For experimental information on the observational tests of general relativity, see:

B. Bertotti, D. Brill and R. Krotkov. Article in Gravitation: An Introduction to Current Research, ed. L. Witten.
M. G. Adam, Proc. Roy. Soc. A 270, 297 (1962).

A very interesting discussion of the equivalence principle, red-shift and bending of light in various theories of gravitation is to be found in:

A. Schild, Proc. of Int. School of Physics 'Enrico Fermi,' Course 20, p. 69. Academic Press, N. Y. and London. (Date unknown.)

7. INVARIANCE PROPERTIES AND CONSERVATION LAWS

In this chapter we shall investigate the consequences of two assumptions:

(1) that the field equations (equations of motion) of a dynamical system can be derived from a variational principle,

(2) that they are invariant under some continuous group of transformations.

But before doing so, we shall summarize some of the theory of Lie groups of transformations.

7.1. LIE GROUPS*

A Lie Group is a set G of elements which constitute

(i) a group

(ii) a differentiable manifold

and is such that the mapping

$$G \times G \ni (a,b) \longrightarrow ab^{-1} \in G \quad ,$$

defined using the group law of composition, is differentiable with respect to the differentiable manifold structure. This last condition relates the otherwise independent structures we have imposed on G. In the following we shall denote the set of differentiable vector fields by χ.

With any $a \in G$ we can associate two mappings of G onto G:

$G \ni x \longrightarrow r_a(x) = xa \in G$, called <u>right translation by a</u>,

and

$G \ni x \longrightarrow l_a(x) = ax \in G$, called <u>left translation by a</u>.

* This section and the next are based on S. Kobayashi and K. Nomizu, <u>Foundations of Differential Geometry</u> (Interscience 1963).

From the axioms of G given above we see that these mappings are diffeomorphisms, and we note that

$$r_a \cdot r_b = r_{ba} \quad , \quad l_a \cdot l_b = l_{ab} . \tag{7.1}$$

Given any vector field $\vec{A} \in X$ we may drag it along by r_a and l_a to form $r_a\vec{A}$, $l_a\vec{A} \in X$. If $r_a\vec{A} = \vec{A}$ (or $l_a\vec{A} = \vec{A}$) for all $a \in G$, we may say that $\vec{A}$ is <u>right (left) invariant</u> respectively. We now prove:

<u>Theorem 1</u>

There is a (1,1) correspondence between right invariant vector fields on G and tangent vectors at the unit element e of G. Further, these right invariant vector fieldson G form a Lie Algebra $\mathfrak{g}$ with respect to the commutator $[\vec{A},\vec{B}]$.

<u>Proof</u>

Let $T_a(G)$ be the tangent space to G at $a \in G$, and let $\vec{A}_e \in T_e(G)$. Define

$$T_a(G) \ni \vec{A}_a = r_a \vec{A}_e . \tag{7.2}$$

Then if $f \in \mathcal{D}$, where $\mathcal{D}$ denotes the set of differentiable functions of G,

$$\begin{aligned} r_b \vec{A}_a(f) &= \vec{A}_a(f \cdot r_b) \\ &= r_a \vec{A}_e(f \cdot r_b) \quad \text{by (7.2)} \\ &= \vec{A}_e(f \cdot r_b \cdot r_a) \\ &= \vec{A}_e(f \cdot r_{ab}) \quad \text{by (7.1)} \\ &= r_{ab} \vec{A}_e(f) \\ &= \vec{A}_{ab}(f) \quad \text{by (7.2)} \end{aligned}$$

and so $\vec{A}_a$ is right invariant. But if two right invariant vector fields $\vec{A}_a$ and $\vec{B}_a$ are such that $\vec{A}_e = \vec{B}_e$, they are equal everywhere, for

$$\vec{A}_a = r_a \vec{A}_e \quad \text{by definition of right invariance}$$

$$= r_a \vec{B}_e = \vec{B}_a .$$

Hence every right invariant vector field can be obtained from its value at e in this manner. This proves the first part of the theorem.

To prove the second part we merely have to show that if $\vec{A}$, $\vec{B}$ are right invariant, so is $[\vec{A},\vec{B}]$. But

$$r_a[\vec{A},\vec{B}](f) = r_a \vec{A}(\vec{B}(f)) - r_a \vec{B}(\vec{A}(f))$$

$$= \vec{A}(\vec{B}(f)) - \vec{B}(\vec{A}(f)) \quad \text{as } r_a\vec{A} = \vec{A},\ r_a\vec{B} = \vec{B}$$

$$= [\vec{A},\vec{B}].$$

This completes the proof.

Because of this theorem the elements of $\mathfrak{g}$ may be considered either as right invariant vector fields or as elements of $T_e(G)$, and we see that the dimension of the vector space $\mathfrak{g}$ is the same as the dimension of the manifold G, say n.

Let $\{\vec{E}_\mu\}$, $\mu = 1,2,\cdots,n$ be a basis of $\mathfrak{g}$. Then as $[\vec{E}_\mu,\vec{E}_\nu] \in \mathfrak{g}$, there exist constants $c^\rho_{\mu\nu}$, called the <u>structure constants</u> of G with respect to the basis $\{\vec{E}_\mu\}$, such that

$$[\vec{E}_\mu,\vec{E}_\nu] = c^\rho_{\mu\nu}\vec{E}_\rho , \qquad (7.3)$$

using the summation convention. From the antisymmetry of the Lie bracket and the Jacobi identity, they satisfy the relations

$$c^\rho_{\mu\nu} = -c^\rho_{\nu\mu} , \quad c^\sigma_{[\lambda\mu}c^\rho_{\nu]\sigma} = 0 \qquad (7.4)$$

and they can be shown to determine the group to within topological differences.

Let us now consider the one-parameter local group G_1 of diffeomorphisms ϕ_t generated by any $\vec{A} \in X$, so that if $f \in \mathcal{D}$, $p \in G$,

$$\vec{A}_p(f) = \frac{d}{dt} f(\phi_t(p))\Big|_{t=0} \tag{7.5}$$

We prove:

Lemma 1

If h is a diffeomorphism of $G \rightarrow G$ and $\vec{A} \in \mathfrak{X}$ generates ϕ_t, then $h\vec{A}$ generates $h \circ \phi_t \circ h^{-1}$.

Proof

We have to show that for any $f \in \mathcal{F}$,

$$\frac{d}{dt} f(h \circ \phi_t \circ h^{-1}(p))\Big|_{t=0} = (h\vec{A})_p(f)$$

But
$$\begin{aligned}
L.H.S. &= \frac{d}{dt} f \circ h \circ \phi_t \circ h^{-1}(p)\Big|_{t=0} \\
&= \frac{d}{dt} h^{-1} f \circ \phi_t (h^{-1}(p)) \\
&= \vec{A}_{h^{-1}(p)}(h^{-1}f) \qquad \text{by (7.5)} \\
&= h\vec{A}_p(f)
\end{aligned}$$

which completes the proof.

Now let $\vec{A} \in \mathfrak{g}$ and consider the curve in G through e defined by

$$a_t = \phi_t(e). \tag{7.6}$$

Take $h = r_{a_s}$ in the above lemma. Then it tells us that $r_{a_s}\vec{A}$ generates $r_{a_s} \circ \phi_t \circ r_{a_s}^{-1}$. But as $\vec{A} \in \mathfrak{g}$, $r_{a_s}\vec{A} = \vec{A}$, and we already know that $\vec{A}$ generates a unique G_1 of diffeomorphisms, namely ϕ_t. So

$$r_{a_s} \circ \phi_t \circ r_{a_s}^{-1} = \phi_t$$

which gives
$$r_{a_s} \circ \phi_t = \phi_t \circ r_{a_s}. \tag{7.7}$$

Apply this at e. We have

$$r_{a_s}(e) = a_s = \phi_s(e)$$

and so

$$r_{a_s} \circ \phi_t(e) = r_{a_s}(a_t) = a_t a_s$$

and

$$\phi_t \circ r_{a_s}(e) = \phi_t \circ \phi_s(e)$$

$$= \phi_{t+s}(e) \quad \text{by the axioms of a } G_1$$

$$= a_{t+s} \quad \text{by (7.6).}$$

(7.7) then gives

$$a_t a_s = a_{t+s} \tag{7.8}$$

so that the curve a_s is a one-parameter subgroup of G. Note that similarity of

$$a_t a_s = a_{t+s} \quad \text{and} \quad \phi_t \circ \phi_s = \phi_{t+s}$$

is deceptive, as the former uses the group law of composition while the latter uses composition of functions.

We shall now leave the general theory of Lie Groups and specialize to Lie Groups of transformations.

7.2. LIE GROUPS OF TRANSFORMATIONS

Let X be a differentiable manifold and G be a Lie Group, with a composition $\perp$ defined between X and G such that

$$G \times X \ni (a,p) \rightarrow a \perp p \in X$$

is a differentiable mapping of GxX into X and satisfies the following axioms:

$$a \perp (b \perp p) = ab \perp p \tag{7.9}$$

$$e \perp p = p \tag{7.10}$$

where a,b $\in$ G, p $\in X$ and e is the unit element of G. Then G is called a Lie Group of Transformations of X.

We shall assume that G acts effectively on X, i.e. a $\perp$ p = p for all p $\in$ X if and only if a = e.

Now take any $\vec{A} \in \mathfrak{g}$ and construct the G_1 of diffeomorphisms ϕ_t generated by $\vec{A}$. Put $a_t = \phi_t(e)$, as in the previous section, and define a one-parameter set h_t of diffeomorphisms of X by

$$X \ni p \rightarrow h_t(p) = a_t \perp p \in X.$$

Then:

(1) $\quad h_0(p) = e \perp p = p \qquad$ by (7.10)

(2) $\quad h_t \cdot h_s(p) = a_t \perp (a_s \perp p)$

$\qquad = a_t a_s \perp p \qquad$ by (7.9)

$\qquad = a_{t+s} \perp p \qquad$ by (7.8).

$\qquad = h_{t+s}(p).$

Hence h_t is a one-parameter group of diffeomorphisms of X, and it induces a differentiable vector field ξ on X, such that if f is a differentiable function on X,

$$\vec{\xi}_p(f) = \frac{d}{dt} f(h_t(p))\Big|_{t=0} . \qquad (7.11)$$

Let $\mathfrak{X}$ now be the set of differentiable vector fields on X. Then the following theorem can be proved:

Theorem 2

The mapping $\vec{A} \to \vec{\xi}$ defined above is a one-to-one homomorphism of $\mathfrak{g}$ into $\mathfrak{X}$, these two sets being considered as Lie algebras.

For proof, see, for example, Nomizu and Kobayashi.

Since $\mathfrak{X}$ is infinite-dimensional and $\mathfrak{g}$ finite-dimensional, we see that the mapping must be into rather than onto. The $\vec{\xi}_\mu$ constructed in this way (by taking for $\vec{A}$ the r elements of a basis $\{\vec{E}_\mu\}$ of $\mathfrak{g}$, where r is the dimension of G), span an r-dimensional Lie subalgebra of $\mathfrak{X}$, and

$$[\vec{\xi}_\mu, \vec{\xi}_\nu] = c^{\rho}_{\mu\nu}\vec{\xi}_\rho ,$$

where the $c^{\rho}_{\mu\nu}$ are the structure constants of G with respect to the basis $\{\vec{E}_\mu\}$. To every set of scalars c^μ, $\mu = 1,2,\ldots,r$ there corresponds a local G_1 of diffeomorphisms of X, namely that generated by $c^\mu\vec{\xi}_\mu$, which is a subgroup of the full Lie group of transformations.

7.3. INFINITESIMAL COORDINATE TRANSFORMATIONS

Let h be a diffeomorphism of X and let $\{x^\alpha\}$ be a coordinate system on X, either locally or globally defined. Remembering that $p \to \{x^\alpha(p)\}$ is a mapping of X into R^n, the n-dimensional Euclidean space, we may define a new coordinate system $\{x^{\alpha'}\}$ on X by

$$x^{\alpha'}(p) = h x^\alpha(p) = (x^\alpha \circ h^{-1})(p)$$

hx^α denoting the function obtained from x^α by dragging along by h. In this way we may associate a coordinate transformation on X with any diffeomorphism of X.

Now take a G_1 of diffeomorphisms h_t of X, and let $\vec{\xi}$ be the vector field on X induced by the G_1. Take the coordinate transformation

$$x^{\alpha}(p) \to x^{\alpha\prime}(p) = h_t\, x^{\alpha}(p) \qquad (7.12)$$

induced by h_t and expand it to first order for small t, say $t = \epsilon$, by Taylor's Theorem:

$$\begin{aligned} x^{\alpha\prime} &= x^{\alpha} + \epsilon \frac{d}{dt}\, h_t\, x^{\alpha}(p)\Big|_{t=0} \\ &= x^{\alpha} + \epsilon \frac{d}{dt}\, x^{\alpha} \circ h_t^{-1}(p)\Big|_{t=0} \end{aligned} \qquad (7.13)$$

But $h_t^{-1}(p) = h_{-t}(p)$ by the axioms of a G_1. Equation (7.13) can then be written as

$$\begin{aligned} x^{\alpha\prime}(p) &= x^{\alpha}(p) - \epsilon \frac{d}{dt}\, x^{\alpha} \cdot h_t(p)\Big|_{t=0} \\ &= x^{\alpha}(p) - \epsilon\, \vec{\xi}_p(x^{\alpha}) \text{ by } (7.11) \\ &= x^{\alpha}(p) - \epsilon\, \xi^{\alpha}(p) \end{aligned}$$

where ξ^{α}, the components of $\vec{\xi}$, are given as usual by $\xi^{\alpha} = \vec{\xi}(x^{\alpha})$. So

$$x^{\alpha\prime} = x^{\alpha} - \epsilon\, \xi^{\alpha}.$$

This is called an infinitesimal coordinate transformation. We shall put $\delta^* x^{\alpha} = -\epsilon\, \xi^{\alpha}$ and write it as

$$x^{\alpha} \to x^{\alpha\prime} = x^{\alpha} + \delta^* x^{\alpha}.$$

7.4. THE VARIATIONAL PRINCIPLE

We assume that the description of a physical system is given completely by a system of functions $Y_A(x)$, where x stands for the local coordinates defined on an n-dimensional differentiable manifold X, called the base space.

Examples

(i) Classical Mechanics: A system with N degrees of freedom is described completely by N generalized coordinates $q_1 \ldots q_N$, each q being a function of time. In this case the base space is simply the time line.

(ii) Quantum Mechanics: The description of the same physical system is conveyed by a state vector $\Psi = \Psi(q_1, \ldots q_N, t)$ and X is an (N + 1) dimensional differentiable manifold.

(iii) Classical Field Theory: In this case the Y_A are components of a field or fields and X is the 4-dimensional space-time. It is sometimes convenient to include the metric tensor components g_{ab} with the field variables Y_A even if the g_{ab} are not dynamical variables.

We assume that the field equations can be derived from the principle of stationary action

$$\delta W = 0$$

where

$$W = \int_\Omega L\, dx$$

and

$$L = L(x, y_A, y_{A,a}, y_{A,ab})$$

The region of integration Ω is that region of R^n which is the map of the domain of X over which the Y_A's are defined. For the moment L is not specified as a scalar or a scalar density etc., but is perfectly general. For simplicity we write

$$L = L(x; y(x))$$

and we assume that variations of Y_A and of $Y_{A,a}$ vanish on the boundary $F\Omega$ of Ω i.e.

$$\delta y_A \big|_{F\Omega} = \delta y_{A,a} \big|_{F\Omega} = 0. \tag{7.14}$$

By definition of δW, we obtain

$$\begin{aligned} \delta W &= \int_\Omega L(x; y+\delta y)\, dx - \int_\Omega L(x; y)\, dx \\ &= \int_\Omega \left(\frac{\delta W}{\delta y_A} \delta y_A + \partial_a T^a \right) dx \end{aligned} \tag{7.15}$$

where
$$\frac{\delta W}{\delta y_A} \overset{\text{def.}}{=} L^A \overset{\text{def.}}{=} \frac{\partial L}{\partial y_A} - \partial_a \frac{\partial L}{\partial y_{A,a}} + \partial_a \partial_b \frac{\partial L}{\partial y_{A,ab}}.$$

Also from the definition of δW, the functions T^a depend linearly on δY_A and $\delta Y_{A,a}$ and hence, on converting the divergence in (7.15) to an integral over the surface $F\Omega$, it is clear that it makes no contribution to δW by virtue of (7.14). Hence

$$\delta W = \int_\Omega L^A \delta y_A \, dx = 0$$

implies

$$L^A = 0$$

provided all the Y_A's are dynamical variables.

We are now free to perform coordinate transformations in X, or transformations of the Y_A's themselves. Such transformations do not alter the physical system but rather provide different descriptions of the same physical system and are generically termed <u>gauge transformations</u>. Gauge transformations include as special cases

(1) Coordinate transformations

$$x^{\alpha\prime} = X^\alpha(x).$$

(2) Transformations of the type

$$y_{A}'(x') = Y_{A}(x\,;y(x))$$

e.g. $$A_{a'}(x') = A_{a}(x)\frac{\partial x^{a}}{\partial x^{a'}}\,.$$

(3) The familiar gauge transformation of electromagnetic theory viz.

$$x^{a'} = x^{a} \quad \text{and} \quad A_{a}'(x') = A_{a}(x) + \partial_{a}\chi\,.$$

In a new gauge the Lagrangian is written as L'(x';y'(x')) and we assume that new field equations are obtained upon varying the new action

$$W' = \int_{\Omega'} L'(x';y'(x'))dx'\,.$$

The prime on L' indicates that, in general, its form is different from that of L. We can derive sufficient conditions for the compatibility of the two descriptions, i.e. sufficient conditions which ensure that the solutions of the field equations derived from L when transformed to the new gauge are the same as those solutions of the field equations derived from L'.

We have the action

$$W' = \int_{\Omega'} L'(x';y'(x'))dx'$$

and the sufficient conditions for compatibility are that there exist functions Q^{a} such that

$$W' = \int_{\Omega}\left[L(x;y(x)) - \partial_{a}Q^{a}(x;y(x))\right]dx\,.$$

Proof. $$\delta W' = \delta W - \int_{F\Omega}\delta Q^{a}n_{a}\,dS\,,$$

and if $$Q^{a} = Q^{a}(x, y_{A}, y_{A,a}) \text{ only, then}$$

$$\int \delta Q^a n_a dS = 0$$

by (7.14). Hence $$\delta W = \delta W'$$

But for arbitrary δY_A, $\delta W = 0$ implies and is implied by $L^A = 0$, and also $\delta W' = 0$ implies and is implied by $L'^A = 0$, and these imply that $L^A = 0$ implies and is implied by $L'^A = 0$. This proof does not hold as a necessary condition.

In general, gauge transformations change the form of the equations of motion. Those gauge transformations which leave the form of the equations of motion unaltered are called symmetry transformations. This simply means that

$$L^A(x; y(x)) \equiv L'^A(x; y(x)).$$

This condition is implied by the sufficient condition that the form of the Lagrangian be unchanged i.e.

$$L'(x; y(x)) = L(x; y(x)).$$

We shall assume that the symmetry transformations form a continuous group and henceforth we shall restrict our considerations to infinitesimal symmetry transformations.

We now write an infinitesimal coordinate transformation in the form

$$x^{a'} = x^a + \delta^* x^a$$

where $$\delta^* x^a = -\epsilon \xi^a$$

and we define $\delta^* Y_A$ and $\bar{\delta} Y_A$ by the following equations:

$$y'_A(x') = y_A(x) + \delta^* y_A$$

and

$$\bar{\delta} y_A = y'_A(x) - y_A(x)$$
$$= y'_A(x') - y_A(x) + y'_A(x) - y'_A(x')$$
$$= \delta^* y_A - y_{A,a}\,\delta^* x^a$$

where we have neglected terms of order ε^2. We also define $\bar{\delta}L$ by

$$\bar{\delta}L = L(x; y'(x)) - L(x; y(x))$$

and we note that $\bar{\delta}$ and partial differentiation commute, i.e.

$$\bar{\delta} y_{A,a} = \partial_a \bar{\delta} y_A .$$

Clearly, for a pure dragging along

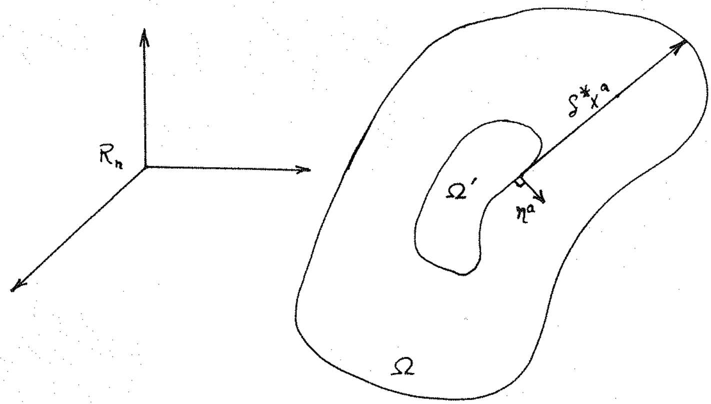
$$\bar{\delta} y_A = \underset{\varepsilon\xi}{\pounds} y_A .$$

The sufficient conditions for compatibility between systems differing by a symmetry transformation can be stated thus. There exist functions $\bar{\delta}Q^a$ of order ε such that

$$\int_{\Omega'} L(x'; y'(x'))dx' = \int_{\Omega}[L(x; y(x)) - \partial_a \bar{\delta} Q^a(x; y(x))]dx$$

where we can regard $\bar{\delta}Q^a$ as the definition of the function Q^a for the special case of symmetry transformations. Alternatively, we may write these conditions as

$$\int_{\Omega'} L(x'; y'(x'))dx' - \int_{\Omega}[L(x; y(x)) - \partial_a \bar{\delta} Q^a]dx = 0 .$$



By a relabeling of x' to x, we obtain

$$\int_\Omega [L(x;y'(x)) - L(x;y(x)) + \partial_a \bar{\delta} Q^a] dx + \int_{F\Omega} L\, \delta^* x^a n_a dS = 0$$

where $\delta^* x^a n_a dS$ is the volume between Ω' and Ω as indicated in the diagram. Hence we obtain

$$\int_\Omega [\bar{\delta} L + \partial_a(\bar{\delta} Q^a + L\, \delta^* x^a)] dx \equiv 0 \qquad (7.16)$$

and Ω being arbitrary, we arrive at

$$\bar{\delta} L + \partial_a(\bar{\delta} Q^a + L\, \delta^* x^a) \equiv 0 \qquad (7.17)$$

Now

$$\begin{aligned} \bar{\delta} L &= L(x; y + \bar{\delta} y) - L(x;y) \\ &= \frac{\partial L}{\partial y_A} \bar{\delta} y_A + \frac{\partial L}{\partial y_{A,a}} \bar{\delta} y_{A,a} + \frac{\partial L}{\partial y_{A,ab}} \bar{\delta} y_{A,ab} \\ &= L^A \bar{\delta} y_A + \partial_a \left[\left(\frac{\partial L}{\partial y_{A,a}} - \partial_b \frac{\partial L}{\partial y_{A,ab}} \right) \bar{\delta} y_A + \frac{\partial L}{\partial y_{A,ab}} \bar{\delta} y_{A,b} \right] \end{aligned}$$

From (7.17) we obtain the <u>basic identity</u>

$$L^A \bar{\delta} y_A + \partial_a \bar{\delta} t^a \equiv 0 \qquad (7.18)$$

where $$\bar{\delta} t^{a} = \left[\left(\frac{\partial L}{\partial y_{A,a}} - \partial_\beta \frac{\partial L}{\partial y_{A,ab}} \bar{\delta} y_{A,b}\right] + \bar{\delta} Q^{a} + L \delta^{*} x^{a}.$$

If all the Y_A's are dynamical variables then

$$L^{A} = 0$$

and we get a conservation law. Also, if $\bar{\delta} y_A = 0$,

Noether Theorem

Case I. Suppose the symmetry group is an r-dimensional Lie group of transformations. Then we can write

$$\delta^{*} x^{a} = -\varepsilon^{i} \xi^{a}_{i} \quad , \quad i = 1, 2, \cdots, r$$

and

$$\bar{\delta} y_A = \varepsilon^{i} \eta_{Ai}$$

$$\bar{\delta} t^{a} = \varepsilon^{i} t^{a}_{i} \qquad \text{etc.}$$

where these equations can be regarded as definitions of η_{Ai} and $t_i{}^a$ etc. In these circumstances the basic identity becomes

$$\varepsilon^{i}\left(L^{A} \eta_{Ai} + \partial_a t^{a}_{i}\right) \equiv 0$$

for any ε^{i}, and hence

$$L^{A} \eta_{Ai} + \partial_a t^{a}_{i} \equiv 0 .$$

If $L^A = 0$, we get r conservation laws

$$\partial_a t^{a}_{i} = 0 \quad , \quad i = 1, 2, \ldots, r.$$

Case II. Suppose the symmetry transformations constitute a group whose "parameters" are s arbitrary functions of x^a.

Such a group is usually called an infinite group.

We assume that the infinitesimal symmetry transformation of y_A is

$$\bar{\delta} y_A = \varepsilon^j(x)\, \gamma_{Aj} - \varepsilon^j{}_{,a}\, \gamma^a{}_{Aj}\,, \quad j = 1, 2, \cdots, s$$

(e.g. in electromagnetic theory $\bar{\delta} A_a = -\varepsilon_{,a}$) and we obtain similar equations for $\bar{\delta} t^a$ etc. On substituting in the basic identity we get an equation involving ε; $\varepsilon_{,a}$; $\varepsilon_{,ab}$ etc. The coefficients of these terms must vanish separately owing to their independent arbitrariness. If we choose ε and its derivatives such that they vanish on $F\Omega$, then by (7.16)

$$0 \equiv \int_\Omega L^A \bar{\delta} y_A \, dx = \int_\Omega L^A (\varepsilon^j \gamma_{Aj} - \varepsilon^j{}_{,\alpha} \gamma^\alpha{}_{Aj})\, dx$$

$$= \int_\Omega \varepsilon^j \left[L^A \gamma_{Aj} + (L^A \gamma^a{}_{Aj})_{,a} \right] dx$$

or, since ε^j is arbitrary

$$L^A \gamma_{Aj} + (L^A \gamma^a{}_{Aj})_{,a} \equiv 0 \qquad j = 1, 2, \cdots, s. \quad (7.19)$$

These s differential identities connecting the field equations are called the Generalized Bianchi Identities. They imply that the field equations are not independent and this gives rise to difficulties when attempting to formulate the Cauchy problem.

Example: Assume

$$L = L(x, y_A, y_{A,a}),$$

then

$$L^A = \frac{\partial L}{\partial y_A} - \partial_a \left(\frac{\partial L}{\partial y_{A,a}} \right) = - \frac{\partial^2 L}{\partial y_{A,a} \partial y_{B,b}} \, y_{B,a,b} + \left(\begin{matrix} 1^{\text{st}} \text{ deri-} \\ \text{vative terms} \end{matrix} \right)$$

Suppose we wish to impose Cauchy data on the surface

$$x^0 = \text{constant}.$$

We can decompose L^A so as to exhibit $y_{B,00}$, thus

$$L^A = -\frac{\partial^2 L}{\partial y_{A,o}\partial y_{B,o}} y_{B,oo} + (\text{terms not involving } y_{B,oo}). \tag{7.20}$$

If the matrix $\partial^2 L/\partial y_{A,o}\partial y_{B,o}$ is singular, then, given y_A and its time derivative on the surface we cannot solve for $y_{B,oo}$.

Now, from the Generalized Bianchi identities, the coefficient of 3rd derivatives of y_A must vanish, i.e.

$$\frac{\partial^2 L}{\partial y_{A,a}\partial y_{B,b}} \gamma^c{}_{Aj}\, y_{B,abc} = 0$$

for any y_A. Since $y_{B,abc}$ is symmetric in a, b and c we need only consider that part of this equation which is symmetric in a, b and c. On putting

$$a = b = c = o$$

we obtain
$$\frac{\partial^2 L}{\partial y_{A,o}\partial y_{B,o}} \gamma^o{}_{Aj} = 0.$$

Whenever there exist non-zero $\gamma^o{}_{Aj}$'s such that this equation is valid, the matrix in question is singular and the Cauchy problem cannot be formulated as usual.

Furthermore the expression $\gamma^o{}_{Aj}L^A$, calculated from (7.20) does not contain time derivatives of y_A and hence the equations

$$\gamma^o{}_{Aj}\, L^A = 0,$$

which are valid in view of the field equations, act as subsidiary or constraint equations on the Cauchy data y_A and $y_{A,o}$.

If the infinite group of gauge transformations possesses a Lie subgroup then we obtain so-called <u>strong conservation laws</u> i.e. conservation laws which are valid whether the field equations are satisfied or not. They may be demonstrated by writing

$$\varepsilon^j(x) = \varepsilon^i \xi^j{}_i(x),$$

where the ε^i are the numerical parameters of the Lie group and consequently the expressions for $\bar{\delta} y_A$ and $\bar{\delta} t^a$ assume the simplified form

$$\bar{\delta} y_A = \varepsilon^i \left(\xi^j{}_i \gamma_{Aj} - \partial_a \xi^j{}_i \gamma^a{}_{Aj} \right)$$

and

$$\bar{\delta} t^a = \varepsilon^i t^a_i \, .$$

On substituting these expressions in the basic identity (7.18) we obtain

$$\xi^i L^A \left(\xi^j{}_i \gamma_{Aj} - \partial_a \xi^j{}_i \gamma^a{}_{Aj} \right) + \varepsilon^i \partial_a t^a{}_i \equiv 0 \qquad (7.21)$$

and the ε^i, being arbitrary, may be ignored. Multiplying the generalized Bianchi identities (7.19) by ξ^j_i yields

$$L^A \xi^j{}_i \gamma_{Aj} + \xi^j{}_i \left(L^A \gamma^a{}_{Aj} \right)_{,a} \equiv 0 \qquad (7.22)$$

and subtracting (7.22) from (7.21) gives us the strong conservation laws

$$\partial_a \left(t^a{}_i - L^A \gamma^a{}_{Aj} \xi^j{}_i \right) = 0. \qquad (7.23)$$

Note that these conservation laws hold independently of the field equations

$$L^A = 0.$$

Weak conservation laws are those which hold only as a consequence of the field equations.

If we write equations (7.22) as

$$\partial_a \Theta^a{}_i = 0 \qquad i = 1, 2, \cdots, r$$

for an r-parameter subgroup, then there exist bivectors $U^{ab}{}_i$ which are skew symmetric in a and b, such that

$$\Theta^{a}{}_{i} = U^{ab}{}_{i,b} \; .$$

The $U^{ab}{}_{i}$'s are sometimes called <u>superpotentials</u>.

Consequently the form of $t^{a}{}_{i}$ (which may be the energy-momentum or current of a physical system) for theories admitting infinite transformations which contain a Lie subgroup can be written as

$$t^{a}{}_{i} = U^{ab}{}_{i,b} + L^{A}\, \gamma^{a}_{Aj}\, \xi^{j}{}_{i} \, .$$

7.5. CONSERVATION LAWS IN RIEMANNIAN SPACE-TIME

For the moment we shall not regard the metric tensor g_{ab} as a dynamical variable but rather as a description of the background space upon which the physical system is imposed. This does not preclude variations of the action with respect to the metric but such variations will not give rise to field equations.

We choose as our variables $_{A}$, the metric tensor g_{ab}, a set of arbitrary field variables $\psi_r (r = 1 \ldots N)$, and a function z^a $(a = 0,1,2,3)$ which describes the world line of a single particle. The action is assumed to be of the form

$$W = \int_{\Omega} \mathcal{L}\, dx = \int_{\Omega} \left(L + \int_{-\infty}^{\infty} \Lambda\, \delta(x - z)\, ds \right) dx$$

where $L = L(g_{ab}, \psi r, \psi r{,a})$ is assumed to be a scalar density, and

$$\Lambda = \Lambda(\psi_r, \dot{z}^a)$$

is assumed to be a scalar. The integral of $\Lambda\delta(z-x)$ is taken over the world line of the particle and since the delta function is a scalar density, $\mathcal{L}$ is also a scalar density and has, therefore, the same form in all coordinate systems.

In this case, therefore, the infinite gauge group is the group of coordinate transformations. There may well be more gauge groups but we restrict our attention to this particular case.

We write the infinitesimal coordinate transformation as

$$x^{a} = x^{a} - \xi^{a}$$

and we deduce immediately that

$$\bar{\delta}\psi_r = \underset{\xi}{\mathcal{L}}\psi_r .$$

It can be readily shown that if ψ is a tensor field, the Lie derivative of ψ can be written as

$$\underset{\xi}{\mathcal{L}}\psi_r = \psi_{r,a}\xi^{a} - F_{ra}{}^{sb}\psi_s\xi^{a}{}_{,b}$$

where the $F_{ra}{}^{sb}$ are constants.

We now perform all the necessary variations of the action and for simplicity, introduce the following notation

$$-2\frac{\delta W}{\delta g_{ab}} = T^{ab} \quad \text{- a tensor density,}$$

$$\frac{\delta W}{\delta \psi_r} = \mathcal{L}^{r}$$

and

$$\frac{\delta W}{\delta z^{a}} = \Lambda_a .$$

The equations

$$\mathcal{L}^{r} = 0$$

are the field equations for ψ_r, and

$$\Lambda_a = 0$$

are the equations of motion of the particle. It should be emphasized again that the variation with respect to g_{ab} does not give rise to field equations since the g_{ab} are not dynamical variables but have been included in L merely to ensure that L has the same form in all coordinate systems.

In these circumstances it can be shown that the basic identity (7.18) assumes the following form

$$-\frac{1}{2} T^{ab} \underset{\xi}{\mathcal{L}} g_{ab} + \mathcal{L}^r \underset{\xi}{\mathcal{L}} \psi_r - \int_{-\infty}^{\infty} \Lambda_a \xi^a \delta(x-z)\, ds + \partial_a t^a - \int_{-\infty}^{\infty} ds\, \delta(x-z) \frac{dp}{ds} \equiv 0 \tag{7.24}$$

where

$$t^a = -L\xi^a + \frac{\partial L}{\partial \psi_{r,a}} \underset{\xi}{\mathcal{L}} \psi_r$$

and

$$p = \left[\left(\Lambda - \frac{\partial \Lambda}{\partial \dot{z}^b} \dot{z}^b\right) \dot{z}^a + \frac{\partial \Lambda}{\partial \dot{z}^a}\right] \xi^a$$

with the dot denoting differentiation with respect to s.

From the basic identity (7.24) it is possible to obtain the generalized Bianchi identities which are

$$\nabla_b \left(T_a{}^b + \mathcal{L}^r F_{ra}{}^{sb} \psi_s\right) + \mathcal{L}^r \nabla_a \psi_r - \int_{-\infty}^{\infty} \Lambda_a \delta(x-z)\, ds \equiv 0. \tag{7.25}$$

If the field equations for ψ_r are satisfied i.e.

$$\mathcal{L}^r = 0$$

and whenever $T_a{}^b$ is covariantly conserved i.e.

$$\nabla_b T_a{}^b = 0$$

then it follows automatically from (7.25) that

$$\Lambda_a = 0.$$

Hence the equations of motion are satisfied if and only if the energy-momentum is covariantly conserved.

In general relativity T^{ab} is always covariantly conserved as a consequence of the Einstein field equations and the contracted Bianchi identities and consequently the equations of motion in general relativity emerge automatically from the field equations.

As a special case we consider only a system of fields

$$W = \int_{\Omega} L\, dx$$

Since L is a scalar density, we have

$$\begin{aligned} 0 &\equiv \underset{\xi}{\mathcal{L}} L - \partial_a (L \xi^a) \\ &= \frac{\partial L}{\partial g_{ab}} \mathcal{L} g_{ab} + \frac{\partial L}{\partial \psi_r} \mathcal{L} \psi_r + \frac{\partial L}{\partial \psi_{r,a}} \mathcal{L} \psi_{r,a} - \partial_a (L \xi^a). \end{aligned}$$

Partial differentiation commutes with Lie differentiation and hence this becomes

$$0 = -\tfrac{1}{2} T^{ab} \mathcal{L} g_{ab} + L^r \mathcal{L} \psi_r + \partial_a \left(\frac{\partial L}{\partial \psi_{r,a}} \mathcal{L} \psi_r - L \xi^a \right).$$

This equation contains $\xi^a{}_{,bc}$, $\xi^a{}_{,b}$ and ξ^a and, because of the arbitrariness of ξ and its derivatives the coefficients of these terms must vanish separately provided we symmetrize over b and c. This gives us three sets of equations corresponding to the coefficients of $\xi^a{}_{,(bc)}$, $\xi^a{}_{,b}$ and ξ^a respectively,

$$S_a{}^{(bc)} \equiv 0 \quad \text{where} \quad S_a{}^{bc} = \frac{\partial L}{\partial \psi_{r,c}} F_{ra}{}^{sb} \psi_s \tag{7.26a}$$

$$T_a{}^b \equiv t_a{}^b + \nabla_c S_a{}^{bc} + L^r F_{ra}{}^{sb} \psi_s \tag{7.26b}$$

where

$$t_a{}^b = \frac{\partial L}{\partial \psi_{r,b}} \nabla_a \psi_r - L \delta^b_a \tag{7.26b'}$$

and $t_a{}^b$ can be recognized as the canonical energy-momentum tensor density. And finally

$$\nabla_b T_a{}^b = -L^r \nabla_a \psi_r + \nabla_b (L^r F_{ra}{}^{sb} \psi_s) \tag{7.26c}$$

which indicates that if the field equations are satisfied ($L^r = 0$), then $T_a{}^b$ is covariantly conserved ($\nabla_b T_a{}^b = 0$).

If we assume that the field equations are satisfied and from (7.26b) calculate the covariant divergence of $t_a{}^b$, then, using the Ricci identity and the fact that $S_a{}^{bc}$ is skew symmetric in b and c, we obtain

$$\nabla_b t_a{}^b = S_a{}^{bc} R^d{}_{abc}. \tag{7.27}$$

In special relativity $T_a{}^b$ and $t_a{}^b$ differ only by a curl, viz $\nabla_c S_a{}^{bc}$ as is clear from (7.26b) and the divergences of both $T_a{}^b$ and $t_a{}^b$ vanish. In curved space however, the divergence of $t_a{}^b$ is not zero but is given by (7.27) and hence $T_a{}^b$ may be regarded as a better description of the energy-momentum of a physical system in curved space.

The equation

$$\nabla_b T_a{}^b = 0$$

can be written

$$\nabla_b T_a{}^b = \partial_b T_a{}^b + T_a{}^c \Gamma^b{}_{cb} - T_c{}^b \Gamma^c{}_{ab} - T_a{}^b \Gamma^c{}_{cb} = 0$$

the last term arising because $T_a{}^b$ is a tensor density. Since the affine connection is symmetric this equation reduces to

$$\partial_b T_a{}^b - T_c{}^b \Gamma^c{}_{ab} = 0$$

which is not a true conservation law in that it is not an ordinary divergence which, when integrated over all space, can be reduced to a surface integral.

If there exists a Killing vector field ξ i.e.

$$\mathop{\pounds}_{\xi} g_{ab} = \nabla_a \xi_b + \nabla_b \xi_a = 0$$

and the field equations are satisfied, then we have

$$\partial_a t^a = \nabla_a t^a = 0 \tag{7.28}$$

where

$$t^a = \frac{\partial L}{\partial \psi_{r,a}} \pounds_\xi \psi_r - L \xi^a$$

is the canonical energy-momentum vector density. Furthermore, if

$$T^a = T^a{}_b \xi^b$$

then

$$\partial_a T^a = \nabla_a T^a \tag{7.29a}$$

since T^a is a vector density, and

$$\nabla_a T^a = (\nabla_a T^a{}_b)\xi^b + T^{ab} \nabla_a \xi_b = 0 \tag{7.29b}$$

because T^{ab} is covariantly conserved and symmetric. Also, from equation (7.26b), if $L^r = 0$ then

$$T^a = t^a + \partial_b(\xi^c S_c{}^{ab})$$

and the conservation laws (7.28) and (7.29) are equivalent in view of the 4-dimensional generalization of Gauss's theorem.

In the case of special relativity i.e. Minkowski space and cartesian coordinates, there exist 10 Killing vectors. If we choose the Killing vectors corresponding to Lorentz transformations then,

$$\xi^a = \omega^a{}_b x^b$$

where

$$\omega_{ab} = -\omega_{ba} = \text{constant}$$

and, on substituting the above expression for ξ^a in t^a, we obtain

$$t^a = \frac{1}{2} j_{bc}{}^a \omega^{cb}$$

where

$$\frac{1}{2} j_{bc}{}^a = x_{[b} t^a{}_{c]} + S_{[bc]}{}^a$$

which is the expression corresponding to total angular momentum. The conservation law (7.28) implies

$$\partial_a j_{bc}{}^a = 0$$

which is the conservation law of angular momentum.

7.6. CONSERVATION LAWS IN GENERAL RELATIVITY

We choose an action corresponding to the gravitational field to be of the form

$$W_g = \int_\Omega G(g_{ab}, g_{ab,c}, g_{ab,cd})\, dx \qquad (7.30)$$

and we assume that G is a scalar density which is invariant in form and this ensures the tensorial character of the field equations and their compatibility.

Suppose, however, that G is linear in $g_{ab,cd}$ and that $\partial G/\partial g_{ab,cd}$ is independent of $g_{ab,c}$ (this is satisfied if $G \sim \sqrt{-g}\, R$), then replacing G by

$$G - \partial_d \left(\frac{\partial G}{\partial g_{ab,cd}} g_{ab,c} \right)$$

leads to the same field equations since the extra term is a divergence and the whole expression is independent of $g_{ab,cd}$; but if G is a scalar density, the whole expression is not a scalar density.

We shall assume however that G is a scalar density and we introduce the following definition:

$$-\frac{1}{16\pi}G^{ab} = \frac{\delta W_g}{\delta g_{ab}}$$

$$= \frac{\partial G}{\partial g_{ab}} - \partial_c \frac{\partial G}{\partial g_{ab,c}} + \partial_c\partial_d \frac{\partial G}{\partial g_{ab,cd}}. \qquad (7.31)$$

In the absence of matter the field equations are

$$G^{ab} = 0$$

where G^{ab} is a tensor density.

Since G is a scalar density we have the identity

$$0 \equiv \underset{\xi}{\mathcal{L}}\, G - (G\xi^a)_{,a}$$

which can be written in the form

$$0 = -\frac{1}{16\pi}G^{ab}\mathcal{L}\, g_{ab} + \partial_a t^a \qquad (7.32)$$

when we integrate, by parts, the expansion

$$\mathcal{L}\, G = \frac{\partial G}{\partial g_{ab}}\mathcal{L}\, g_{ab} + \frac{\partial G}{\partial g_{ab,c}}\mathcal{L}\, g_{ab,c} + \frac{\partial G}{\partial g_{ab,cd}}\mathcal{L}\, g_{ab,cd}.$$

On integrating (7.32) over a region Ω and choosing $\xi = 0$ on $F\Omega$, we obtain

$$\int_\Omega G^{ab}\nabla_a\xi_b\, dx = 0 \qquad (7.33)$$

where we have used

$$\mathcal{L}\, g_{ab} = \nabla_a\xi_b + \nabla_b\xi_a,$$

and integrating (7.33) by parts yields

$$\int_\Omega (\nabla_b G^{ab})\,\xi_a\, dx \equiv 0$$

which implies the generalized Bianchi identity

$$\nabla_b G^{ab} = 0 \tag{7.34}$$

since ξ^a is arbitrary. In the special case when

$$G = -\frac{\sqrt{-g}\,R}{16\pi}$$

we obtain

$$G^{ab} = \frac{\sqrt{-g}}{16\pi}\left(R^{ab} - \frac{1}{2} g^{ab} R\right)$$

and (7.34) becomes the usual contracted Bianchi identity

$$\nabla_b \left(R^{ab} - \frac{1}{2} g^{ab} R\right) = 0.$$

The identity (7.34) is the same as that obtained in the general case described earlier, viz.

$$L^A \gamma_{Aj} + (L^A \gamma^a{}_{Aj})_{,a} \equiv 0$$

with G^{ab} standing for L^A, and therefore we have difficulty in solving the Cauchy problem. Some of the field equations are independent of $g_{ab,oo}$ and these field equations must be satisfied by the Cauchy data before the Cauchy problem is solved. In this case G^{ao} does not contain $g_{ab,oo}$ and the constraint

$$G^{ao} = 0$$

(being a system of 4 nonlinear partial differential equations which must be solved before attempting the Cauchy problem), gives an indication of the complexity of the Cauchy problem in general relativity. We may note that if

$$\sigma = \text{constant}$$

is the Cauchy surface, then the constraint equations can be expressed covariantly as

$$G^{ab}\sigma_{,b} = 0.$$

We now consider matter in the presence of a gravitational field and we add a corresponding matter term W to the action. On varying with respect to g_{ab} we obtain

$$\frac{\delta}{\delta g_{ab}}(W_g + W) = -\frac{1}{16\pi}G^{ab} - \frac{1}{2}T^{ab} = 0$$

and the field equations thus are

$$G^{ab} = -8\pi T^{ab} \qquad (7.35)$$

Equations (7.34) and (7.35) imply

$$\nabla_b T^{ab} = 0, \qquad (7.36)$$

i.e. T^{ab} is covariantly conserved as a consequence of the gravitational field equations alone, whereas in special relativity (7.36) holds only if the matter field equations are satisfied.

From equations (7.32) and (7.34) we obtain the strong conservation law for gravitation

$$\nabla_a\left(-\frac{1}{8\pi}G^{ab}\xi_b + \tau^a\right) = 0 \qquad (7.37)$$

and since the expression in the bracket is a vector density, the covariant derivative could equally well be replaced by a derivative. If we use the field equations and replace G^{ab} by T^{ab} in (7.37) we obtain a conservation law for matter and gravitation viz.

$$\partial_a(T^a + \tau^a) = 0$$

where

$$T^a = T^{ab}\xi_b$$

and ∂_a and ∇_a are equivalent.

<u>Example</u>: Einstein's Theory of Gravitation

We put

$$G = -\frac{1}{16\pi}\sqrt{-g}\,R \qquad (\kappa = c = 1)$$

and obtain

$$G^{ab} = \sqrt{-g}\,(R^{ab} - \tfrac{1}{2}g^{ab}R).$$

In this case the basic identity becomes

$$-(R^{ab} - \tfrac{1}{2}g^{ab}R)\nabla_a\xi_b + \nabla_a[R^{ab} - \tfrac{1}{2}g^{ab}R]\,\xi_b = 0$$

which is equivalent to equation (7.32). It follows that we can choose the equation

$$\tau^a = \frac{1}{8\pi}\sqrt{-g}\,(R^{ab} - \tfrac{1}{2}g^{ab}R)\xi_b$$

since $\sqrt{-g}$ is covariantly conserved, and on using the field equations (7.35) we obtain

$$T^a + \tau^a = 0.$$

This equation may be interpreted to mean that the gravitational energy annihilates the matter energy so that the total current of energy and momentum is zero.

However, a transformation of τ^a of the form

$$\tau'^a = \tau^a + V^{ab}{}_{,b} \qquad (7.38)$$

where

$$V^{ab} = -V^{ba}$$

preserves

$$\partial_a(T^a + \tau^a) = 0$$

but in this case

$$T^{a} + \tau^{a} \neq 0.$$

The function

$$V^{ab} = V^{ab}(g_{cd}, g_{cd,a}, \xi)$$

need not be a tensor and it is widely believed that a correct choice of V^{ab} could lead to a physically meaningful conservation law. It is believed for example that a good choice of V^{ab} would be one in which $T^{a} + \tau^{a}$ would be independent of $g_{ab,cd}$ and, in the special case of a Schwarzschild field, the expression for total energy E, given by

$$E = \int (T^{0} + \tau^{0})\, dV$$

would reduce to mc^2, with m being the mass of the Schwarzschild particle.

We may observe also that all identities involving the arbitrary field ξ^{a} are valid even when ξ is not a vector but in such a case quantities such as τ^{a} etc. will not have a tensorial character. For example, we may introduce 4 sets of ξ's such that, in all coordinate systems, the ξ's are given by

$$\xi^{a}_{(b)} = \delta^{a}_{b}$$

where b = 0,1,2,3 labels the 4 ξ's.

We consider now a τ^{a} which may be either the initial τ^{a} as defined by (7.32)or a transformed τ^{a} as defined by

$$\tau^{a}{}_{b} = \tau^{a}(\xi_{(b)})$$

The quantity $\tau^{a}{}_{b}$ is independent of the ξ's and is a function of g_{ab} and its derivatives only. We call $\tau^a{}_b$ a pseudo tensor or energy-momentum complex.

Clearly we can obtain arbitrarily many expressions for $\tau^a{}_b$ simply by exploiting the freedom in the choice of the V^{ab} and the ξ^a and this in turn leads to arbitrarily many conservation laws of a non-tensorial character. However, we may restrict the arbitrariness somewhat by requiring that τ^{ab} give a physically reasonable result for an isolated system of particles generating a gravitational field.

A particularly useful choice of V^{ab} was made by Komar which yielded the result

$$T^a + \tau^a = \frac{\sqrt{-g}}{4\pi} \nabla_b \left(\nabla^{[a} \xi^{b]} \right).$$

From this expression, the Møller pseudo energy-momentum tensor can be obtained by an appropriate choice of ξ.

We have seen that the definition of energy in general relativity is extremely difficult owing to the arbitrariness in the choice of V^{ab} and ξ^a. The basic reason for this does not lie in the peculiarities of general relativity but rather in the special and unusual character of the gravitational field.

In general, the notion of energy is closely related to the notion of force i.e. given a force we can calculate work done by the force and thus define changes in energy, but the notion of force cannot be readily utilized when considering gravitational interactions. Even in Newtonian theory we cannot uniquely define the gravitational force acting on a body unless we have an isolated system. When there is a strong gravitational field acting over all space this definition is impossible.

In electrodynamics, for example, the energy-momentum tensor is constructed essentially from forces i.e. from first derivatives of the potentials. In a field theory of gravitation, if we regard the metric g_{ab} as the potentials and try to construct an energy-momentum tensor from the first derivatives of the g_{ab}, we find that it is impossible to do so.

We may also present a more formal argument by a comparison with special relativity where the notions of energy, momentum and angular momentum are closely related to the symmetry properties of Minkowski space i.e. homogeneity gives rise to conservation laws of energy and momentum, and isotropy to conservation of angular momentum. In Riemannian geometry, in general, there are no symmetries and hence we would not expect conservation laws.

However, the situation regarding conservation laws in general relativity is not as serious as it might appear, since the role of conservation laws in any physical theory, whilst often extremely convenient to the solving of certain problems, is not indispensable.

The energy concept in general relativity can really be made meaningful only in special cases e.g. an isolated system of bodies which produce a gravitational field that becomes weak at large distances from the bodies. In order to formulate a meaningful conservation law in the case of a strong gravitational field over all space, Møller has introduced privileged tetrad frames in space time but the physical meaning of such tetrads is far from being clear.

For fields which are asymptotically flat, certain asymptotic symmetries do exist and one can formulate global conservation laws corresponding to these symmetries. Thus far however, the treatment of this problem has been somewhat inelegant in that Einstein's idea of asymptotic cartesian coordinates has been employed. A far better formulation of the symmetries would be in terms of asymptotic Killing vectors and this has not yet been done in full generality.

7.7. RELATION BETWEEN FIELD EQUATIONS AND EQUATIONS OF MOTION

We have seen from general considerations of the variational principle that if all the equations describing a physical system follow from the variational principle, and if the matter field equations are satisfied, i.e.

$$L^r = 0$$

then

$$\nabla_b T^{ab} = 0$$

implies the equations of motion

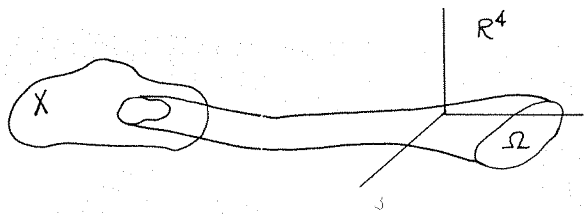
$$\Lambda_a = 0$$

and the gravitational field equations (7.35) imply

$$\nabla_b T^{ab} = 0.$$

Hence we infer that if the matter field equations and the gravitational field equations are satisfied, then the equations of motion are automatically implied. This implication at first sight appears to give us more information ~~than~~ than we should expect from the amount of labor expended. In special relativity, for example, we had to vary the world lines of the particles in order to obtain the equations of motion. In reality, however, we vary with respect to the 10 functions g_{ab}, but the geometry of the manifold is determined by only 6 functions among these 10. The remaining 4 functions represent some description of the inherent freedom of coordinate choice in general relativity, and the variation with respect to these 4 functions gives us the equations of motion.

Consider the following situation:



Suppose the metric field g_{ab} is given in X, and consider the coordinate transformation

$$x^a \rightarrow \bar{x}^a = F^a(x)$$

where F^a is invertible, of class C^∞, and a 1 to 1 map of Ω on Ω. In addition F^a also satisfies

$$F^a \big|_{F\Omega} = x^a$$

$$\frac{\partial F^a}{\partial x^b}\bigg|_{F\Omega} = \delta^a_b \, ,$$

i.e. F is the identity transformation to the first order on the boundary of Ω.

We define a new metric field g_{ab} in the manifold by

$$g_{ab}(x) = \bar{g}_{cd}(\bar{x}) \frac{\partial F^c}{\partial x^a} \frac{\partial F^d}{\partial x^b}$$

and we consider the metric field $\bar{g}_{ab}(x)$ which is a new metric field in the same region and same coordinate system as the original one.

It follows from the properties of F that

$$\bar{g}_{ab}|_{F\Omega} = g_{ab}|_{F\Omega}.$$

Since G is a scalar density, one then has in the action principle that

$$\int_{\Omega} G(g_{ab}(x)\ \dots)\,dx = \int_{\Omega} G(\bar{g}_{ab}(x)\dots)\,dx$$

If $g_{ab}(x)$ is a solution of the field equations i.e. if g_{ab} makes the action an extreme, then so does $\bar{g}_{ab}(x)$, and g_{ab} and $\bar{g}_{ab}$ are virtually indistinguishable. One is simply the dragged along version of the other.

Suppose now we take a fixed curve c:

$$x^a = z^a(\lambda)$$

in Ω and consider a combined action of the form

$$\int_{\Omega} G\,dx - m\int_{c} ds$$

Let $g_{ab}(x)$ be a field rendering $\int_{\Omega} G dx$ stationary. On replacing g_{ab} by $\bar{g}_{ab}(x)$ this integral does not change but $\int_c ds$ does change! Thus the change $g_{ab}(x) \rightarrow \bar{g}_{ab}(x)$ is equivalent to a change of the curve c i.e.

$$z^a(\lambda) \rightarrow \bar{z}^a(\lambda)$$

where

$$F^a(\bar{z}) = z^a.$$

Hence we see that varying with respect to the 10 g_{ab} is equivalent to varying both the geometry and the world lines.

Because of the relation between the field equations and the equations of motion in general relativity it is not possible, even in principle, to solve in a finite number of steps, the problem of motion in general relativity. In electrodynamics, for example, if we have a system of charges then we can write down Maxwell's equations for the field and the Lorentz equation for the motion of the charges. If we assume arbitrary motion, we can solve the field equations and insert the solutions into the Lorentz equation and find the world lines precisely. In general relativity arbitrary motions cannot be assumed since the motions are determined by field equations.

An approximate procedure for determining the world lines was developed by Einstein, Infeld and Hoffmann. In this procedure all functions occurring in the field equations are written as an expansion in c—the velocity of light, e.g.

$$\phi(x^0, x^\alpha, c) = \phi(ct, x^\alpha, c) = \sum_{n=0}^{\infty} \frac{1}{c^n}\, \phi_n(t, x^\alpha).$$

The series expressions are inserted in the field equations and all coefficients of separate powers of c are equated to zero. This yields an infinite system of equations which may be solved step by step; the equations of motion appear as integrability conditions for each successive equation. The process converges rapidly for slow moving bodies but cannot be applied to radiative problems.

7.8. THE MOTION OF PARTICLES IN A GIVEN FIELD

We shall assume here that the particles do not influence the field i.e. they are test particles. The action integral for a single particle in a field ψ_r is given by

$$W = \int_c \Lambda(\psi_r, \dot{x}^a)\, ds \tag{7.39}$$

where Λ is a scalar of invariant form and the integral is evaluated over a portion of the world line of the particle c, parametrized by its length s. As an example, Λ for a charged particle in an electromagnetic field is given by

$$\Lambda = -mc - \frac{e}{c} A_a \dot{x}^a$$

We could easily generalize Λ to

$$\Lambda = \Lambda(\psi_r, \psi_{r,a}, g_{ab}, \dot{x}^a)$$

but shall not do so here.

We may remark here that since $\dot{x}^a$ is a unit vector, $\Lambda(\psi_r, \dot{x}^a)$ is defined only on the unit hyperboloid in the x space and this limited definition prevents us from differentiating with respect to all the components of x. Hence we extend Λ off the unit hyperboloid by considering a more general vector u^a and the corresponding $\Lambda(x, u^a)$. This extension is by no means unique.

If now we wish to vary the world line of the particle then clearly we cannot retain the parameter s since, in general, a neighboring world line with the same endpoints as the original would have different length. Thus we consider a new parameter λ whose values corresponding to the endpoints of the world lines are λ_1 and λ_2. Equation (7.39) thus becomes

$$W = \int_{\lambda_1}^{\lambda_2} \Lambda\left(\psi_r, \frac{dx^a}{d\lambda} \Big/ \frac{ds}{d\lambda}\right) \frac{ds}{d\lambda}\, d\lambda, \tag{7.40}$$

and if we write

$$L(x, x'^a) = \Lambda\left(\psi_r, \frac{dx^a}{d\lambda} \Big/ \frac{ds}{d\lambda}\right) \frac{ds}{d\lambda}$$

where

$$x'^a = \frac{dx^a}{d\lambda}$$

then (7.40) is recognized as a familiar expression in the calculus of variations, and the Euler-Lagrange equations are

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial x'^a}\right) - \frac{\partial L}{\partial x^a} = 0 \tag{7.41}$$

or

$$\frac{d}{d\lambda} \frac{\partial}{\partial x'^a}\left(\Lambda \frac{ds}{d\lambda}\right) - \frac{\partial L}{\partial \psi_r} \partial_a \psi_r = 0. \tag{7.42}$$

We extend Λ as described above, and note that having calculated $\partial\Lambda/\partial u^a$, (on substituting $\dot{x}^a$ for u^a), the symbol $\partial\Lambda/\partial u^a$ really means $\partial\Lambda/\partial u^a$ (ψ_r, $\dot{x}^a$), so that we can write (7.42) as

$$\frac{d}{d\lambda}\left[\left(\Lambda - \frac{\partial\Lambda}{\partial u^b}u^b\right)\dot{x}^a + \frac{\partial\Lambda}{\partial u^a}\right] - \frac{\partial\Lambda}{\partial\psi_r}\partial_a\psi_r\frac{ds}{d\lambda} = 0 \qquad (7.43)$$

Here we have used

$$\frac{ds}{d\lambda} = \sqrt{g_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}}$$

and hence

$$\frac{\partial}{\partial x'^a}\left(\frac{ds}{d\lambda}\right) = \frac{x'^a}{\sqrt{g_{ab}x'^a x'^b}} = \frac{x'^a}{\frac{ds}{d\lambda}} = \dot{x}^a .$$

Moreover, we have chosen a coordinate system in which the affine connection vanishes and g_{ab} is a constant so that the only x dependence occurs in ψ_r.

However, since our final equation of motion must be a vector equation, the general form of (7.43) is

$$\Lambda_a \equiv \frac{D}{ds}\left[\left(\Lambda - \frac{\partial\Lambda}{\partial u^b}\right)\dot{x}^a + \frac{\partial\Lambda}{\partial u^a}\right] - \frac{\partial\Lambda}{\partial\psi_r}\nabla_a\psi_r = 0 \qquad (7.44)$$

The expression within the square bracket is called the generalized momentum p_a, and

$$p_a = \frac{\partial L}{\partial x'^a}. \qquad (7.45)$$

From the assumption that Λ is a scalar we may write

$$\underset{\xi}{\mathcal{L}}\,\Lambda\,(\psi_r, u^a) \equiv \xi^a\partial_a\Lambda$$

and from the assumption that it is of invariant form this becomes

$$\frac{\partial\Lambda}{\partial\psi_r}\mathcal{L}\psi_r + \frac{\partial\Lambda}{\partial u^a}\mathcal{L}u^a \equiv \xi^a\frac{\partial\Lambda}{\partial\psi_r}\nabla_a\psi_r + \xi^a\frac{\partial\Lambda}{\partial u^b}\nabla_a u^b$$

where we have replaced ∂_a by ∇_a since Λ is a scalar. By expressing $\mathcal{L}u^a$ in terms of ξ^a we obtain

$$\frac{\partial \Lambda}{\partial \psi_r} \mathcal{L}_\xi \psi_r \equiv \frac{\partial \Lambda}{\partial \psi_r} \xi^a \nabla_a \psi_r + \frac{\partial \Lambda}{\partial u^a} u^b \nabla_b \xi^a. \tag{7.46}$$

In order to arrive at conservation laws, we first construct the scalar $p_a\xi^a$ and hence

$$\frac{d}{ds}(p_a \xi^a) = \frac{Dp_a}{ds}\xi^a + p_a \frac{D\xi^a}{ds}$$

Using (7.44), (7.45), and (7.46), this equation becomes

$$\frac{d}{ds}(p_a\xi^a) \equiv \xi^a \Lambda_a + \frac{\partial \Lambda}{\partial \psi_r}\mathcal{L}_\xi \psi_r + \tfrac{1}{2}\left(\Lambda - \frac{\partial \Lambda}{\partial u^c}u^c\right)u^a u^b \mathcal{L}_\xi g_{ab} \tag{7.47}$$

where the term $\mathcal{L}_\xi g_{ab}$ comes from $\nabla_b \xi^a$ in (7.46).

We may obtain a conservation law from (7.47) if we can make the right-hand side vanish. We assume therefore:

(i) The equations of motion are satisfied i.e.

$$\Lambda_a = 0$$

(ii) The field ξ is a symmetry of ψ_r i.e. ψ_r is invariant under a dragging along generated by ξ, which implies

$$\mathcal{L}_\xi \psi_r = 0$$

(iii) ξ is a Killing vector i.e.

$$\mathcal{L}_\xi g_{ab} = 0.$$

Hence we deduce that along the world line of the particle

$$p_a \xi^a = \text{constant}.$$

Now consider an r-parameter Lie group of transformations generated by ξ_i^a $(i = 1,\ldots r)$, with

$$[\bar{\xi}_i, \bar{\xi}_j] = c^k_{ij}\, \bar{\xi}_k . \tag{7.48}$$

Assume further that all the $\bar{\xi}_i$'s are Killing vectors and preserve the field ψ_r, i.e.

$$\pounds\, g_{ab} = \pounds\, \psi_r = 0$$

Then, if we define

$$p_i = p_a \xi^a_{\ i}$$

it can be shown that

$$\{p_i, p_j\} = c^k_{ij}\, p_k \tag{7.49}$$

where in general the Poisson bracket $\{u,v\}$ is defined by

$$\{u, v\} = \frac{\partial u}{\partial p_a}\frac{\partial v}{\partial x^a} - \frac{\partial u}{\partial x^a}\frac{\partial v}{\partial p_a} .$$

When we quantize a classical system the Poisson brackets go over into commutators and from equations (7.48) and (7.49) we can see that in these circumstances an isomorphism is established between the Lie algebra of the constants of motion p_i and the Lie algebra of the group of symmetries.

7.9. THE CANONICAL FORMALISM FOR RELATIVISTIC PARTICLES

When attempting to quantize a physical system it is convenient to express the classical equation of motion of the system in Hamiltonian form. In relativity this is accomplished most easily by abandoning a manifestly covariant formalism and singling out the time coordinate from the spatial ones.

We shall show here how to reduce the equations to canonical form without loss of covariance i.e. the time will not play a preferred role. We choose as our Lagrangian

$$L(x,x') = \Lambda \frac{ds}{d\lambda}$$

and on replacing x' by $\mu x'$, we obtain

$$L(x,\mu x') = \mu L(x,x')$$

(i.e. L is homogeneous of degree 1 in x'), and from the Euler identity we obtain

$$\frac{\partial L}{\partial x'^a} x'^a = L. \tag{7.50}$$

We differentiate (7.50) with respect to x'^b, obtaining

$$\frac{\partial L}{\partial x'^b} + \frac{\partial^2 L}{\partial x'^a \partial x'^b} x'^a = \frac{\partial L}{\partial x'^b}$$

or

$$\frac{\partial^2 L}{\partial x'^a \partial x'^b} x'^a = 0 \tag{7.51}$$

which indicates that the matrix $\partial^2 L/\partial x'^a \partial x'^b$ is singular; we assume that its rank is 3, i.e. x'^a is the only null vector.

In the circumstances, the equation

$$p_a = \frac{\partial L}{\partial x'^a}$$

cannot be solved for x'^a as a function of p_a and x^a and, therefore, there exists a function $H(p_a, x^b)$ such that

$$H\left(\frac{\partial L}{\partial x'^a}, x^b\right) = 0. \tag{7.52}$$

We differentiate (7.52) with respect to x'^b and x^b, obtaining

$$\frac{\partial H}{\partial p_a} \frac{\partial^2 L}{\partial x'^a \partial x'^b} \equiv 0 \tag{7.53}$$

and

$$\frac{\partial H}{\partial p_a} \frac{\partial^2 L}{\partial x'^a \partial x^b} + \frac{\partial H}{\partial x^b} \equiv 0. \tag{7.54}$$

Comparing (7.51) and (7.53) and using the assumption that the rank of $\partial^2 L/\partial x'^a \partial x'^b$ is 3, we deduce that there exists a function $\nu(x,x')$ such that

$$x'^a = \nu \frac{\partial H}{\partial p_a} \tag{7.55}$$

and substituting this in (7.54) yields

$$x'^a \frac{\partial^2 L}{\partial x'^a \partial x^b} + \nu \frac{\partial H}{\partial x^b} = 0. \tag{7.56}$$

On differentiating the Euler identity with respect to x^b we get

$$\frac{\partial^2 L}{\partial x'^a \partial x^b} x'^a = \frac{\partial L}{\partial x^b}$$

which, when combined with (7.56) gives

$$\frac{\partial L}{\partial x^b} + \nu \frac{\partial H}{\partial x^b} = 0 .$$

The equations of motion can be written

$$p'^a = \frac{\partial L}{\partial x^a}$$

where

$$p'^a = \frac{d p^a}{d\lambda}$$

and hence we obtain

$$p'^a = -\nu \frac{\partial H}{\partial x^a} \tag{7.57}$$

and

$$x'^a = \nu \frac{\partial H}{\partial p_a} \tag{7.58}$$

together with the constraint $H \equiv 0$.

The function ν, however, involves velocities and hence (7.57) and (7.58) cannot be regarded as true canonical equations. It can be shown, however, that a special choice of λ (corresponding to $\lambda = s$) reduces ν to 1.

Consider an extended $\Lambda = \Lambda(x^a, u^b)$ and define

$$p_a(x,u) = \left(\Lambda - \frac{\partial\Lambda}{\partial u^b}u^b\right)u_a + \frac{\partial\Lambda}{\partial u^a}$$

then on the unit hyperboloid

$$p_a(x,\dot{x}) = \frac{\partial L}{\partial x'^a}.$$

Now, provided that

$$\Lambda - \frac{\partial\Lambda}{\partial u^b}u^b = 0$$

then

$$\det\left|\frac{\partial p_a}{\partial u^b}\right| \neq 0$$

and the system of equations may be inverted to give

$$u^a \equiv u^a\left(\frac{\partial L}{\partial x'}, x\right).$$

We may now propose an explicit construction for H, viz.

$$H(p,x) \equiv \tfrac{1}{2}\left(\Lambda - \frac{\partial\Lambda}{\partial u^c}u^c\right)\left(g_{ab}u^a u^b - 1\right).$$

which vanishes when we substitute $\partial L/\partial x'^a$ for p_a and it may be proved that the equations of motion can be written

$$\dot{p}_a = -\frac{\partial H}{\partial x^a}$$

and

$$\dot{x}^a = \frac{\partial H}{\partial p_a}.$$

References:

E. Noether, Göttinger Nachr. 235 (1918).
E. Bessel-Hagen, Math Ann. 84, 258 (1921).
E. L. Hill, Rev. Mod. Phys. 23, 253 (1951).
P. G. Bergmann, Handbuch der Physik, vol. 4.
P. G. Bergmann, Phys. Rev. 75, 680 (1949).
P. G. Bergmann and R. Thomson, Phys. Rev. 89, 400 (1953).
A. Komar, Phys. Rev. 113, 934 (1959).
A. Komar, Phys. Rev. 127, 1411 (1962).
R. S. Arnowitt, S. Deser, and C. W. Misner; also A. Trautman: articles in Gravitation: An Introduction to Current Research, ed. L. Witten, Wiley, N. Y. and London (1962).
A. Trautman, Lectures given at College de France.
R. S. Arnowitt, C. Møller, and J. Plebanski, Lectures given at Warsaw Conference 1962.

8. RIGID MOTION IN RELATIVITY THEORY

This chapter is based on a lecture given by Prof. F. A. E. Pirani which was included in the lecture course of Prof. A. Trautman. The notation of two-component spinors which will be used somewhat in this chapter is explained in the notes on Prof. Pirani's lectures in this volume.

8.1. INTRODUCTION

In Newtonian mechanics any body at a given time t may be regarded as a set of points R(t) in a 3-dimensional Euclidean space E_3. By following the motion of the individual particles of the body, any motion of the body defines a continuous one-parameter family of mappings ϕ_t of the subset R(0) of E_3, with ϕ_t carrying R(0) onto R(t). Rigid motions are characterized by the mappings ϕ_t preserving distances, i.e. being isometric. The isometric mappings of E_3 constitute the Galilean group, consisting of rotations and translations in E_3, and so the motion of a rigid body is described by a continuous one-parameter family of Galilean transformations. From this point of view the kinematics of rigid bodies in Newtonian mechanics reduces to the theory of the Galilean group.

In relativity theory we cannot consider an extended body at one instant of time, as simultaneity cannot be defined over a finite region of space. We must instead consider the 3-parameter family of world-lines of the points of which the body is composed. These will form a congruence of curves in Minkowski space M_4 in the special theory of relativity, or in a 4-dimensional Riemannian space V_4 in the general theory of relativity. But there is no immeditely obvious condition that can be imposed on this congruence that can naturally be called a condition for the body to be rigid.

Let us first consider the situation in special relativity. In our discussion of the Newtonian concept of rigidity given above, we saw that a prominent part is played by the Galilean group, which is the group of isometries of E_3. We We might therefore expect that the group of isometries of M_4

should play a prominent part in the theory of rigid motion in special relativity. This group is well known to be the inhomogeneous Lorentz group (also known as the Poincare group). Since we are seeking a condition on a congruence of curves, the simplest condition we can impose involving the inhomogeneous Lorentz group is to require that the curves of the congruence form the trajectories of a one-parameter subgroup of this group. This is equivalent to requiring that the space M_4 admit a Killing vector field everywhere tangent to the curves of the congruence. We shall show below that one consequence of this criterion is that if the motion of the body (i.e. the instantaneous position and velocity of all its particles) is known to one observer at one instant of time, then the motion is determined for all observers at all times. This condition is clearly too strict to be named rigidity.

We must now look for a weaker condition. One is provided by that proposed by Born[1] for a rectilinear rigid motion, and then, independently, by Herglotz[2] and by F. Noether[3] for a general rigid motion. This definition states:

> A body is called rigid if the distance between every neighboring pair of particles, measured orthogonal to the world-line of either of them, remains constant along the world-line.

This is equivalent to requiring every element of the body to appear rigid, in the Newtonian sense, to an inertial observer instantaneously at rest relative to the element, which is surely the least we can require of a criterion for rigidity if it is to have any connection with our intuitive ideas of the concept. When, below, we speak of rigid motions, it is to this Born criterion that we shall be referring. However, it was shown by Herglotz and Noether that, according to this criterion,

<u>Every rotating rigid motion in flat space-time is isometric.</u>

A proof of this will be given below. We see that even this definition is unsatisfactory in that it fails to give us enough degrees of freedom for a rigid motion because still the entire motion of a rotating body is known when it is known to one observer at one instant of time.

Now let us turn to the situation in general relativity. The above two attempts at definitions of rigidity in special relativity can be immediately generalized to general relativity. The isometry condition now requires that there exist a Killing vector field everywhere tangent to the curves of the congruence of world-lines of the body. But here we are

---

1. M. Born, Ann. der Physik <u>30</u>, 1 (1909).
2. G. Herglotz, Ann. der Physik <u>31</u> 393 (1910).
3. F. Noether, Ann. der Physik <u>31</u>, 919 (1910).

in an even worse position than before, as we know that in general a Riemannian space-time does not admit a Killing vector field at all, and so in a general Riemannian space-time a rigid motion according to this definition could not exist at all. The Born definition can be taken over into general relativity in exactly the form stated above. We shall see below that in a V_4 the Herglotz-Noether theorem no longer holds, in the sense that there exist, in certain space-times, non-isometric rotating congruences of time-like curves which satisfy the Born criterion, but in a general V_4 again no rigid motions exist. The integrability conditions for the Born criterion in curved space-time have been studied by Pirani and Williams,[4] and a discussion of part of their work is given below. It has been shown by Rayner,[5] and is proved below by the method of Pirani and Williams,[4], that the angular velocity of a rigid heavy body is constant in magnitude along any particle world-line in the body. So again we see that the criterion is unsatisfactory, even in space-times admitting rigid motions.

We shall now give a mathematical treatment of the two criteria, and shall prove the claims made above.

8.2. MATHEMATICAL FORMULATION OF THE CRITERIA

The world-lines of the particles of any body may be written $x^a = x^a(\underset{\sim}{u},\sigma)$, where $\underset{\sim}{u} = (u^1,u^2,u^3)$ are three parameters labelling the world-lines and σ is a parameter varying along each world-line. The coordinate displacement dx^a between the points $x^a(\underset{\sim}{u},\sigma)$ and $x^a(\underset{\sim}{u} + d\underset{\sim}{u}, \sigma + d\sigma)$ is

$$dx^a = \frac{\partial x^a}{\partial u^\alpha} du^\alpha + \frac{\partial x^\alpha}{\partial \sigma} d\sigma \tag{8.1}$$

where small Greek letters run from 1 to 3. The distance between these points is given by

$$ds^2 = g_{ab}\, dx^a\, dx^b. \tag{8.2}$$

4. F. A. E. Pirani and G. Williams, Seminaire Janet 5, 8(1962).
5. C. B. Rayner, C. R. Acad. Sci Paris 248, 929 (1959).

Both criteria of section 8.1 require ds^2 to be constant under certain transformations, and we now consider each case separately.

(a) Isometric motion

The condition that the particle world-lines be the trajectories of a one-parameter group of isometries of V_4 is equivalent to requiring that σ can be chosen on the world-lines in such a way that ds^2 given by (8.2) is independent of σ for all given u^α, du^α, $d\sigma$, i.e.

$$\frac{d}{d\sigma}(g_{ab}dx^a dx^b) = 0 \tag{8.3}$$

Let

$$V^a \stackrel{\text{def.}}{=} \frac{\partial x^a}{\partial \sigma} \tag{8.4}$$

Then for any scalar function $\phi(x)$,

$$\frac{d\phi}{d\sigma} = \underset{V}{\pounds}\,\phi \tag{8.5}$$

so that (8.3) may be written as

$$\pounds(g_{ab}dx^a dx^b) = 0 \tag{8.6}$$

Now we have, remembering (8.4).

$$\begin{aligned} \underset{V}{\pounds}\frac{\partial x^a}{\partial u^\alpha} &= V^b\frac{\partial}{\partial x^b}\left(\frac{\partial x^a}{\partial u^\alpha}\right) - \frac{\partial x^b}{\partial u^\alpha}\frac{\partial V^a}{\partial x^b} \\ &= \frac{\partial}{\partial \sigma}\left(\frac{\partial x^a}{\partial u^\alpha}\right) - \frac{\partial}{\partial u^\alpha}\left(\frac{\partial x^a}{\partial \sigma}\right) \\ &= 0 \end{aligned} \tag{8.7}$$

and also, since du^α and $d\sigma$ are constants,

$$\underset{V}{\pounds}\,du^\alpha = 0\ , \quad \underset{V}{\pounds}\,d\sigma = 0\ , \quad \underset{V}{\pounds}\frac{\partial x^a}{\partial \sigma} = \underset{V}{\pounds}\,V^a = 0 \tag{8.8}$$

Hence from (8.1), (8.7), and (8.8) we get

$$\mathcal{L}_{v}\, dx^a = 0$$

from which (8.6) gives

$$(\mathcal{L}_{v}\, g_{ab})\, dx^a dx^b = 0 \tag{8.9}$$

The isometry condition requires this to hold for all du^α and all $d\sigma$, which by (8.1) means that at any one point dx^a is arbitrary. So (8.9) gives

$$\mathcal{L}_{v}\, g_{ab} = 0 \tag{8.10}$$

as the conditions for an isometric motion, i.e. v^a must be a Killing vector field, as stated in the introduction.

(b) Born rigid motion.

Let σ be the proper time measured along the world-lines, so that v^a given by (8.4) satisfies

$$v^a v_a = 1 \tag{8.11}$$

Then for given du^α, the dx^a given by (8.1) is orthogonal to v^a if

$$d\sigma = -v_b \frac{\partial x^b}{\partial u^\alpha} du^\alpha$$

and then

$$dx^a = h^a{}_b \frac{\partial x^b}{\partial u^\alpha} du^\alpha \tag{8.12}$$

where

$$h^a{}_b \overset{\text{def.}}{=} \delta^a{}_b - v^a v_b \tag{8.13}$$

is the projection operator which projects orthogonally to v^a.

We note that

$$h^a{}_b h^b{}_c = h^a{}_c . \tag{8.14}$$

For the dx^a of (8.12) we have

$$ds^2 = g_{cd} h^c{}_a h^d{}_b \frac{\partial x^a}{\partial u^\alpha} \frac{\partial x^b}{\partial u^\beta} du^\alpha du^\beta$$

which with (8.14) gives

$$ds^2 = h_{ab} \frac{\partial x^a}{\partial u^\alpha} \frac{\partial x^b}{\partial u^\beta} du^\alpha du^\beta . \tag{8.15}$$

The condition for Born rigidity is that, for all given u^α and du^α, this expression be independent of σ. By (8.5) this necessitates

$$\underset{v}{\mathcal{L}} \left(h_{ab} \frac{\partial x^a}{\partial u^\alpha} \frac{\partial x^b}{\partial u^\beta} \right) = 0$$

With (8.7) this gives

$$\underset{v}{\mathcal{L}} (h_{ab}) \frac{\partial x^a}{\partial u^\alpha} \frac{\partial x^b}{\partial u^\beta} = 0 \tag{8.16}$$

But, as $h_{ab}v^b = 0$ identically, we also have

$$\underset{v}{\mathcal{L}} (h_{ab}) v^a = 0 \tag{8.17}$$

So as $\partial x^a/\partial u^\alpha$ and v^a together form a basis of the tangent space at any point, (8.16) and (8.17) together give

$$\underset{v}{\mathcal{L}} h_{ab} = 0$$

as the condition for a rigid motion.

8.3. INTEGRATION OF THE ISOMETRY CONDITION IN M_4

Use Minkowski coordinates in which the metric tensor takes the form

$$g_{ab} = \mathrm{diag}(-1,-1,-1,+1).$$

Then (8.10) gives

$$\underset{v}{\mathcal{L}}\, g_{ab} \equiv \partial_a v_b + \partial_b v_a = 0 \tag{8.18}$$

Differentiating with respect to x^c gives

$$\partial_c \partial_a v_b + \partial_c \partial_b v_a = 0 \tag{8.19}$$

Cyclically permute the indices to obtain

$$-\partial_a \partial_b v_c - \partial_a \partial_c v_b = 0, \tag{8.20}$$

$$\partial_b \partial_c v_a + \partial_b \partial_a v_c = 0. \tag{8.21}$$

and add (8.19), (8.20) and (8.21) to give

$$\partial_b \partial_c v_a = 0$$

Integrating once gives

$$\partial_c v_a = \omega_{ac} \ , \text{ constant} \tag{8.22}$$

and from (8.18),

$$\omega_{ac} = -\omega_{ca} \tag{8.23}$$

Integrating (8.22) again gives

$$v^a = \omega^a{}_b x^b + b^a \tag{8.24}$$

where the b^a are constants. This may be re-written as

$$\frac{dx^a}{d\sigma} = \omega^a{}_b x^b + b^\alpha \tag{8.25}$$

It follows immediately from (8.10) that

$$\frac{d}{d\sigma}\left(g_{ab}\frac{\partial x^a}{\partial \sigma}\frac{\partial x^b}{\partial \sigma}\right) = 0$$

so that σ is a constant multiple of the proper time on every world-line, but this multiple will differ from one world-line to another. So σ may be chosen as the proper time on any one world-line, but it will not then so serve on the other world-lines.

We recognize (8.25) as an infinitesimal inhomogeneous Lorentz transformation, and so the general isometric motion is one in which the particle positions at parameter $\sigma + d\sigma$ are obtained from those at time σ by an infinitesimal inhomogeneous Lorentz transformation which is independent of σ. If we use the term boost for a Lorentz transformation from one frame to another parallel to it but with a uniform velocity relative to it, then we can analyze the motion according to the type of transformation as follows:

When the infinitesimal inhomogeneous Lorentz transformation is:	The motion is:
A pure translation	Rectilinear motion with uniform velocity
A pure rotation	Rotation with uniform angular velocity
A pure boost	Rectilinear motion with uniform acceleration.

The term 'uniform' here indicates a constant value along the world-line of any particle. The value will, however, in general vary from particle to particle since σ is proportional, but not equal to the proper time on their world-lines. The general motion is a combination of all three types given in the table. In integrating the infinitesimal transformations however, we shall see that one singular case occurs, that of a null rotation, which is a certain combination of a rotation and a boost.

We shall now consider the problem of explicitly integrating (8.25) to obtain an explicit expression for the world-lines of the particles in an isometric motion. We write (8.25) in matrix form thus:

$$\frac{d\underset{\sim}{x}}{d\sigma} = \Omega \underset{\sim}{x} + \underset{\sim}{b} \tag{8.26}$$

where

$$\underset{\sim}{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ x^4 \end{pmatrix}, \quad \underset{\sim}{b} = \begin{pmatrix} b^1 \\ b^2 \\ b^3 \\ b^4 \end{pmatrix}, \text{ and } \Omega = \| \omega^a{}_b \|$$

whether or not the matrix Ω is non-singular, $\underset{\sim}{b}$ can always be decomposed thus:

$$\underset{\sim}{b} = \Omega \underset{\sim}{c} + \underset{\sim}{d}, \text{ with } \Omega \underset{\sim}{d} = 0 \tag{8.27}$$

and (8.26) can then be written as

$$\frac{d}{d\sigma}(\underset{\sim}{x} + \underset{\sim}{c}) = \Omega(\underset{\sim}{x} + \underset{\sim}{c}) + \underset{\sim}{d}$$

with solution

$$\underset{\sim}{x}(\sigma) + \underset{\sim}{c} = e^{\Omega\sigma}(\underset{\sim}{x}(0) + \underset{\sim}{c}) + \underset{\sim}{d}\sigma. \tag{8.28}$$

To evaluate $e^{\Omega\sigma}$, we must distinguish between two cases for Ω, most easily done in terms of two-component spinors.
We shall write, e.g.

$$\omega_{ab} \longleftrightarrow \omega_{A\dot{X}B\dot{Y}}$$

to indicate the correspondence

$$\omega_{ab} = \sigma_a{}^{A\dot{X}} \sigma_b{}^{B\dot{Y}} \omega_{A\dot{X}B\dot{Y}}$$

$$\omega_{A\dot{X}B\dot{Y}} = \sigma^a{}_{A\dot{X}} \sigma^b{}_{B\dot{Y}} \omega_{ab}$$

between tensors and two-component spinors, as explained in my lecture course.[6] Then we have, since ω_{ab} is skew,

---

6. F. A. E. Pirani: Gravitational Radiation (this volume).

$$\omega_{ab} \leftrightarrow \omega_{AB}\epsilon_{\dot{X}\dot{Y}} + \epsilon_{AB}\bar{\omega}_{\dot{X}\dot{Y}} \tag{8.29}$$

with $\omega_{AB} = \omega_{BA}$. ω_{AB} can be decomposed into a symmetrized product of eigenspinors thus:

$$\omega_{AB} = \kappa_A\lambda_B + \lambda_A\kappa_B \tag{8.30}$$

The two cases we must distinguish are $\kappa_A\lambda^A \neq 0$ and $\kappa_A\lambda^A = 0$.
(a). The case $\kappa_A\lambda^A \neq 0$.

In this case we may modify the decomposition (8.30) by writing

$$\omega_{AB} = \tfrac{1}{2}(\eta + i\omega)(\kappa_A\mu_B + \mu_A\kappa_B) \tag{8.31}$$

with

$$\kappa_A\mu^A = 1$$

and ω, η real. Define real null vectors k^a, m^a and complex conjugate null vectors t^a, $\bar{t}^a$ by

$$\begin{aligned} k^a &\leftrightarrow \kappa_A\bar{\kappa}_{\dot{X}} \qquad & m^a &\leftrightarrow \mu_A\bar{\mu}_{\dot{X}} \\ t^a &\leftrightarrow \kappa_A\bar{\mu}_{\dot{X}} \qquad & \bar{t}^a &\leftrightarrow \mu_A\bar{\kappa}_{\dot{X}} \end{aligned} \tag{8.32}$$

Then on substituting (8.31) into (8.29) and using (8.32) and

$$\epsilon_{AB} = \kappa_A\mu_B - \mu_A\kappa_B \tag{8.33}$$

we get

$$\omega^{ab} = 2\{\eta\, k^{[a}m^{b]} - i\omega\, t^{[a}\bar{t}^{b]}\} \tag{8.34}$$

Choose a special coordinate system $\{x^{a'}\}$ in which

$$k^{a'} = (1,0,0,0),\ m^{a'} = (0,1,0,0),\ t^{a'} = (0,0,1,0)\ \bar{t}^{a'} = (0,0,0,1) \tag{8.35}$$

Then from (8.32) and (8.33) we find

$$k^a m_a = 1\ ,\ t^a \bar{t}_a = -1 \quad \text{all other scalar products zero,} \tag{8.36}$$

so that (8.34), (8.35), (8.36) together give

$$g_{a'b'} = g^{a'b'} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix},\ \omega^{a'}{}_{b'} = \mathrm{diag}\,(\eta, -\eta, i\omega, -i\omega) \tag{8.37}$$

Then

$$(e^{\Omega\sigma})^{a'}{}_{b'} = \mathrm{diag}\,(e^{\eta\sigma}, e^{-\eta\sigma}, e^{i\omega\sigma}, e^{-i\omega\sigma}) \tag{8.38}$$

Now make a coordinate transformation to coordinates $\{x^a\}$, $x^a = A^a_{b'}\ x^{b'}$, where

$$A^a{}_{b'} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -i & i \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \text{so that } A^{a'}{}_b = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & i & 0 & 0 \\ 1 & -i & 0 & 0 \end{pmatrix} \tag{8.39}$$

Then we find that

$$g_{ab} = A^{c'}{}_a A^{d'}{}_a\, g_{c'd'} = \mathrm{diag}\,(-1, -1, -1, +1)$$

so that the x^a are Minkowski coordinates, and transforming (8.38) gives

$$(e^{\Omega\sigma})^a{}_b = \begin{pmatrix} \cos\omega\sigma & -\sin\omega\sigma & 0 & 0 \\ \sin\omega\sigma & \cos\omega\sigma & 0 & 0 \\ 0 & 0 & \cosh\eta\sigma & \sinh\eta\sigma \\ 0 & 0 & \sinh\eta\sigma & \cosh\eta\sigma \end{pmatrix} \tag{8.40}$$

which we recognize as the matrix of a homogeneous Lorentz transformation consisting of a rotation in the (x^1, x^2)-plane through an angle $\omega\sigma$ and a boost in the x^3-direction to velocity $\tanh\eta\sigma$. The origin of coordinates can always be chosen so that $\underset{\sim}{c} = 0$, and then

$$\underset{\sim}{x}(\sigma) = e^{\Omega\sigma}\underset{\sim}{x}(0) + \underset{\sim}{d}\sigma \tag{8.41}$$

From (8.27) we see that $\underset{\sim}{d}$ is a null eigenvector of Ω, and so we can distinguish the four possible cases:

$\omega \neq 0,\ \eta \neq 0$ when $d^1 = d^2 = d^3 = d^4 = 0$

$\omega \neq 0,\ \eta = 0$ when $d^1 = d^2 = 0$

$\omega = 0,\ \eta \neq 0$ when $d^3 = d^4 = 0$

$\omega = 0,\ \eta = 0$ when d is arbitrary, and the motion is simply translation with uniform velocity.

We note that in the case $\underset{\sim}{d} = 0$, the point (0,0,0) at t = 0 must be excluded from the body, because by (8.41) we see that it remains at the same point <u>in space-time</u>, and so cannot represent a real particle of the body.

We have thus integrated case (a) in a special Lorentz frame. The motion in a general Lorentz frame is obtained by making an arbitrary Lorentz transformation.

We now turn to

(b). The case $\kappa_A\lambda^A = 0$.

In this case κ_A and λ_A are proportional and can, therefore, be chosen so that $\kappa_A = 2\lambda_A$, and so (8.30) becomes

$$\omega_{AB} = \kappa_A \kappa_B \tag{8.42}$$

Now choose any μ_A such that $\kappa_A\mu^A = 1$, and construct the null vectors (8.32) as before. Then in place of (8.34), we now get

$$\omega^{ab} = 2k^{[a}(t^{b]} + \bar{t}^{b]}) \tag{8.43}$$

which in the special coordinate system of (8.35) gives

$$\omega^{a'}{}_{b'} = \begin{pmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{8.44}$$

Then we find that $\Omega^3 = 0$, and

$$\left(e^{\Omega\sigma}\right)^{a'}{}_{b'} = 1 + \Omega\sigma + \tfrac{1}{2}\Omega^2\sigma^2 = \begin{pmatrix} 1 & \sigma^2 & -\sigma & -\sigma \\ 0 & 1 & 0 & 0 \\ 0 & -\sigma & 1 & 0 \\ 0 & -\sigma & 0 & 1 \end{pmatrix}$$

Transforming to the coordinates $\{x^a\}$ of (8.39), we get

$$\left(e^{\Omega\sigma}\right)^a{}_b = \begin{pmatrix} 1 & 0 & \sigma & -\sigma \\ 0 & 1 & 0 & 0 \\ -\sigma & 0 & 1-\frac{1}{2}\sigma^2 & \frac{1}{2}\sigma^2 \\ -\sigma & 0 & -\frac{1}{2}\sigma^2 & 1+\frac{1}{2}\sigma^2 \end{pmatrix} \qquad (8.45)$$

This is the matrix of a finite null rotation, and is the singular case referred to in the introduction. Its nature is more easily seen in the infinitesimal form. We have in these coordinates

$$\omega^a{}_b = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

which are seen to be a combination of the infinitesimal matrices for a rotation about the x^2-axis and a boost in the x^1-direction. The general form of $\underset{\sim}{d}$ satisfying (8.27) is easily seen to be

$$d^1 = 0, \quad d^2 \text{ arbitrary} \quad d^3 = d^4$$

Substitution into (8.41) now completes the integration of case (b) in a special Lorentz frame.

8.4. THE ISOMETRY CONDITION IN V_4

In a general space-time, Killing's equations have no solution, and no isometric motions can exist. If we attempt to proceed as in section 8.3, instead of (8.18) we have

$$\nabla_a v_b + \nabla_b v_a = 0$$

and on using the Ricci identity

$$\nabla_a \nabla_b v_c - \nabla_b \nabla_a v_c = R^d{}_{cab} v_d$$

together with the identity $R^d{}_{[abc]} = 0$ of the curvature tensor, we get instead of

$$\partial_b \partial_c v_a = 0$$

the equation

$$\nabla_b \nabla_c v_a = R^d{}_{bac} v_d$$

We cannot then proceed as before.

For later use we shall now express the condition for an isometric motion in terms of the normalized velocity vector. If we now let v^a be the normalized velocity vector, so that $v^a v_a = 1$, then the condition for an isometric motion is that there should exist a scalar field ξ such that ξv^a is a Killing vector field, i.e.

$$\nabla_a (\xi v_b) + \nabla_b (\xi v_a) = 0$$

Then

$$\xi \nabla_{(a} v_{b)} + v_{(a} \partial_{b)} \xi = 0 \qquad (8.46)$$

Define the <u>acceleration vector</u> α^a by

$$\alpha^a = \frac{Dv^a}{ds} = v^b \nabla_b v^a \qquad (8.47)$$

Then multiply (8.46) by v^b and use $v^b \nabla_a v_b = 0$ to give

$$\xi \alpha_a + \partial_a \xi + v_a \frac{d\xi}{ds} = 0$$

Multiplying this by v^a and using $v^a \alpha_a = 0$ gives

$$\frac{d\xi}{ds} = 0$$

so that

$$\xi \alpha_a = -\partial_a \xi$$

and

$$\alpha_a = -\partial_a (\log \xi) \tag{8.48}$$

Hence

$$\nabla_{[a} \alpha_{b]} = 0 . \tag{8.49}$$

Conversely, (8.49) implies the existence of a scalar ξ satisfying (8.48), and we can then work back to (8.46). Equation (8.49) is thus a necessary and sufficient condition for the motion to be isometric.

We note also, for later use, that (8.48) and (8.46) give

$$\nabla_{(a} v_{b)} - v_{(a} \alpha_{b)} = 0 . \tag{8.50}$$

8.5. THE BORN CONDITION IN V_4

In section 8.2 we showed that a velocity field v^a satisfies the Born criterion for rigidity in a V_4 if and only if

$$\underset{v}{\mathcal{L}} h_{ab} = 0 \quad \text{where} \quad h_{ab} = g_{ab} - v_a v_b \tag{8.51}$$

and v^a is normalized, $v^a v_a = 1$. We shall now study some consequences of this condition.

For convenience, introduce the projection symbol $\perp$, read as 'perp,' which projects every free index with $h^a{}_b$, e.g.

$$\perp R^a{}_{bcd} \overset{\text{def}}{=} h^a{}_p h^q{}_b h^r{}_c h^s{}_d R^p{}_{qrs}$$

Then we have

$$\nabla_a v_b = (h^c{}_a + v_a v^c)(h^d{}_b + v_b v^d)\nabla_c v_d$$

$$= \perp \nabla_a v_b + v_a v^c \nabla_c v_b \qquad \text{as } v^a \nabla_b v_a = 0$$

$$= \perp \nabla_a v_b + v_a \alpha_b \qquad \text{by (8.47)}$$

Now, following Ehlers,[7] we make a further decomposition of the purely space-like tensor $\perp\nabla_a v_b$ thus:

$$\perp \nabla_a v_b = \omega_{ab} + \sigma_{ab} + \tfrac{1}{3}\theta h_{ab}$$

where

$\omega_{ab} \overset{\text{def.}}{=} \perp \nabla_{[a} v_{b]}$ is the <u>angular velocity tensor</u>,

$\sigma_{ab} \overset{\text{def.}}{=} \perp \nabla_{(a} v_{b)} - \tfrac{1}{3}\nabla_c v^c h_{ab}$ is the <u>shear velocity tensor</u>,

and

$\theta \overset{\text{def.}}{=} \nabla_a v^a$ is the <u>expansion velocity scalar</u>.

Then

$$\nabla_a v_b = \omega_{ab} + \sigma_{ab} + \tfrac{1}{3}\theta h_{ab} + v_a \alpha_b \,. \tag{8.52}$$

We note that ω_{ab} is skew and σ_{ab} is symmetric and trace-free.

Now (8.51) can be written in the form

---

7. P. Jordan, J. Ehlers, W. Kundt and R. K. Sachs, Acad. Wiss. Lit., Mainz, Abh. Math-naturw., Kl., 791 (1961).

$$v^c \nabla_c h_{ab} + h_{ac} \nabla_b v^c + h_{cb} \nabla_a v^c = 0$$

which simplifies on using (8.47) to

$$\nabla_{(a} v_{b)} - \alpha_{(a} v_{b)} = 0 \text{, ie } \perp \nabla_{(a} v_{b)} = 0 \qquad (8.53)$$

With (8.52) this gives

$$\sigma_{ab} + \tfrac{1}{3}\theta h_{ab} = 0$$

and by taking the trace of this and using $\sigma^a{}_a = 0$ we get

$$\underline{\sigma_{ab} = 0} \qquad \underline{\theta = 0} \qquad (8.54)$$

We have thus shown:
A continuous medium is kinematically rigid if and only if its velocity field is shear-free and expansion-free.

From (8.50) and (8.53) we also have:
Every isometric motion is kinematically rigid.

To obtain further consequences of the Born criterion, it is convenient to introduce an auxiliary 3-dimensional space V. In so doing, we are following the treatment given by Pirani and Williams.[4] We first note that, as in section 6.4, it follows from $\cancel{\mathcal{L}}_v h_{ab} = 0$ (from now on for convenience we shall omit the v in $\cancel{\mathcal{L}}_v$) that there exist coordinate systems in which

$$v^a = (0,0,0,1) \; ; \quad \partial h_{ab}/\partial x^4 = 0. \qquad (8.55)$$

Equations which, so far as is shown, hold only in such a coordinate system, will be written with $\overset{*}{=}$ in place of =. Then we easily see that

$$v^a \overset{*}{=} \delta^a{}_4 \, , v_a \overset{*}{=} g_{a4} \, , v_4 \overset{*}{=} g_{44} \overset{*}{=} 1, h_{a4} \overset{*}{=} 0, h^a_4 \overset{*}{=} 0, h^4{}_\alpha \overset{*}{=} -v_\alpha \qquad (8.56)$$

and

$$\not{t}\, T_{ab\ldots}{}^{cd\ldots} \stackrel{*}{=} \frac{\partial}{\partial x^4} T_{ab\ldots}{}^{cd\ldots} \tag{8.57}$$

Now from (8.55) it follows that $h_{\alpha\beta}$, $\alpha, \beta = 1,2,3$, depend only on the three coordinates x^α, and from (8.51) the rank of the matrix $h_{\alpha\beta}$ is three, i.e. it is non-singular. Hence $h_{\alpha\beta}$ may be interpreted as the metric tensor of a three-dimensional Riemannian space $\hat{V}$. Geometric objects in this space defined in terms of this metric (e.g. affine connection, covariant differentiation, curvature tensor) will be indicated by $\wedge$ (e.g. $\hat{\Gamma}^\gamma_{\alpha\beta}$, $\hat{\nabla}_\alpha$, $\hat{R}_{\alpha\beta\gamma\delta}$), and small Greek letters will run from 1 to 3. We shall now relate such objects to those of the original V_4.

We first have

$$\hat{g}_{\alpha\beta} \stackrel{*}{=} h_{\alpha\beta} \quad , \quad \hat{g}^{\alpha\beta} \stackrel{*}{=} g^{\alpha\beta} \tag{8.58}$$

the latter following from

$$h_{\alpha\beta} g^{\beta\gamma} \stackrel{*}{=} h_{\alpha b} g^{b\gamma} = \delta^\gamma_\alpha - v_\alpha v^\gamma \stackrel{*}{=} \delta^\gamma_\alpha .$$

Now we have from (8.51)

$$\nabla_a h_{bc} = - v_b \nabla_a v_c - v_c \nabla_a v_b \tag{8.59}$$

and from (8.54) and (8.52),

$$\nabla_a v_b = \omega_{ab} + v_a \alpha_b \tag{8.60}$$

so that (8.59) can be written

$$\nabla_a h_{bc} = -2\, \omega_{a(b} v_{c)} - 2\, v_a v_{(b} \alpha_{c)} .$$

Equating this to

$$\nabla_a h_{bc} = \partial_a h_{bc} - h_{dc} \Gamma^d_{ab} - h_{bd} \Gamma^d_{ac}$$

and using

$$\hat{\Gamma}^{\gamma}_{\alpha\beta} \overset{*}{=} \tfrac{1}{2} g^{\gamma\delta}\left(\partial_\alpha h_{\beta\delta} + \partial_\beta h_{\alpha\delta} - \partial_\delta h_{\alpha\beta}\right)$$

then gives

$$\hat{\Gamma}^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\alpha\beta} + g^{\gamma\delta}\left(2\omega_{\delta(\alpha} v_{\beta)} - v_\alpha v_\beta \alpha_\delta\right). \qquad (8.61)$$

Now let $T_{ab\ldots c}$ be any covariant tensor which is orthogonal to v^a on all its indices and which satisfies

$$\mathcal{L}_{\xi}\, T_{ab\ldots c} = 0$$

Then by (8.57),

$$\frac{\partial T_{ab\ldots c}}{\partial x^4} \overset{*}{=} 0$$

and further, in these coordinates $T_{ab\ldots c}$ vanishes if any index is 4. Hence $T_{\alpha\beta\ldots\gamma}$ may be considered as a tensor in $\hat{V}$. We now claim that

$$\hat{\nabla}_\alpha T_{\beta\gamma\ldots\delta} \overset{*}{=} \perp \nabla_\alpha T_{\beta\gamma\ldots\delta} \qquad (8.62)$$

and we shall prove it for a vector T_a. We have

$$\nabla_a T_b = \partial_a T_b - \Gamma^{c}_{ab} T_c$$

from which

$$\perp \nabla_\alpha T_\beta = h^{a}_{\alpha} h^{b}_{\beta}\left(\partial_a T_b - \Gamma^{c}_{ab} T_c\right)$$

$$\overset{*}{=} \partial_\alpha T_\beta - \Gamma^{\gamma}_{\alpha\beta} T_\gamma - \left(h^4_\alpha \Gamma^{\gamma}_{4\beta} + h^4_\beta \Gamma^{\gamma}_{4\alpha} + h^4_\alpha h^4_\beta \Gamma^{\gamma}_{44}\right) T_\gamma \qquad (8.63$$

on using (8.56). Now we have

$$\nabla_a v^\beta = \partial_a v^\beta + \Gamma^{\beta}_{ac} v^c \overset{*}{=} \Gamma^{\beta}_{a4}$$

which gives with (8.60)

$$\Gamma^{\beta}_{a4} \stackrel{*}{=} \omega_a{}^{\beta} + v_a \alpha^{\beta} \tag{8.64}$$

and putting this into (8.63), and using (8.56) and (8.61), we get

$$\perp \nabla_\alpha T_\beta \stackrel{*}{=} \partial_\alpha T_\beta - \hat{\Gamma}^{\gamma}_{\alpha\beta} T_\gamma = \hat{\nabla}_\alpha T_\beta \tag{8.65}$$

This completes the proof for a vector T_a. The proof for higher order tensors then follows easily.

We may now apply (8.62) with $T_{ab} = \perp\nabla_a T_b$. This satisfies the necessary conditions by virtue of (8.65) and $\hat{\nabla}_\alpha T$ being, by definition of $\hat{\nabla}$, independent of x^4. Then we get

$$\hat{\nabla}_\alpha \hat{\nabla}_\beta T_\gamma \stackrel{*}{=} \perp \nabla_\alpha \perp \nabla_\beta T_\gamma \tag{8.66}$$

Now
$$\begin{aligned} \nabla_a \perp \nabla_b T_c &= \nabla_a (h^p{}_b h^q{}_c \nabla_p T_q) \\ &= h^p{}_b h^q{}_c \nabla_a \nabla_p T_q + (h^p{}_b \nabla_a h^q{}_c + h^q{}_c \nabla_a h^p{}_b) \nabla_p T_q \end{aligned}$$

Substitute (8.59) into this and use (8.60) to give

$$\perp \nabla_a \perp \nabla_b T_c = \perp \nabla_a \nabla_b T_c - \omega_{ac} h^p{}_b v^q \nabla_p T_q - \omega_{ab} h^q{}_c v^p \nabla_p T_q \tag{8.67}$$

But $v^q T_q = 0$ gives

$$v^q \nabla_p T_q = - T_q \nabla_p v^q$$

and

$$0 = \pounds T_q = v^p \nabla_p T_q + T_q \nabla_q v^p$$

and on using these and (8.60), (8.67) becomes

$$\perp\nabla_a \perp \nabla_b T_c = \perp\nabla_a\nabla_b T_c + \omega_{ac}\omega_b{}^d T_d + \omega_{ab}\omega_c{}^d T_d$$

Substitute this into (8.66), take the part skew in α and β and use the Ricci identity to give

$$\hat{R}_{\alpha\beta\gamma\delta} \overset{*}{=} R_{\alpha\beta\gamma\delta} - (\omega_{\alpha\gamma}\omega_{\beta\delta} - \omega_{\beta\gamma}\omega_{\alpha\delta}) - 2\,\omega_{\alpha\beta}\omega_{\gamma\delta}. \tag{8.68}$$

But we have identically for any 3-dimensional skew tensor $\omega_{\alpha\beta}$,

$$\omega_{\alpha[\beta}\,\omega_{\gamma\delta]} = 0$$

and using this in (8.68) gives

$$\hat{R}_{\alpha\beta\gamma\delta} \overset{*}{=} R_{\alpha\beta\gamma\delta} - 3\,\omega_{\alpha\beta}\omega_{\gamma\delta}. \tag{8.69}$$

We now claim that it follows from (8.62) and (8.69) that

(i) $$\mathcal{L}\,(\perp R_{abcd} - 3\,\omega_{ab}\omega_{cd}) = 0 \tag{8.70}$$

(ii) If $T_{ab\ldots c}$ is a tensor orthogonal to v^a on all its indices and such that

$$\mathcal{L}\ T_{ab\ldots c} = 0$$

then $$\mathcal{L}\ \perp\nabla_a T_{bc\ldots d} = 0. \tag{8.71}$$

For in the adapted coordinate system, on using (8.62) and (8.69), (8.70) and (8.71) simply become

$$\frac{\partial}{\partial x^4}\hat{R}_{\alpha\beta\gamma\delta} = 0 \quad , \quad \frac{\partial}{\partial x^4}\ \nabla_\alpha T_{\beta\gamma\ldots\delta} = 0,$$

equations which are true by definition of $\hat{V}$. But being tensor equations, (8.70) and (8.71) hold in all coordinate systems if they hold in one coordinate system.

8.6. THE HERGLOTZ-NOETHER THEOREM

Let us now consider the situation in flat space-time, so that $R_{abcd} = 0$ and covariant derivatives commute. Then (8.70) gives

$$\mathcal{L}(\omega_{ab}\omega_{cd}) = 0. \tag{8.72}$$

Now assume ω_{ab} to be non-zero, so that the matter is rotating. Then multiplication of (8.72) by ω^{cd} and then by ω^{ab} leads easily to

$$\mathcal{L}(\omega_{ab}) = 0. \tag{8.73}$$

Equation (8.71) then gives

$$\mathcal{L}(\perp \nabla_a \omega_{bc}) = 0$$

from which

$$\mathcal{L} \perp \nabla_{[a}\omega_{b]c} = 0. \tag{8.74}$$

But

$$\omega_{bc} = \nabla_b v_c - v_b \alpha_c$$

from which we easily get

$$\perp \nabla_{[a}\omega_{b]c} = -\omega_{ab}\alpha_c$$

Hence (8.73) and (8.74) give

$$\mathcal{L}\alpha_a = 0. \tag{8.75}$$

Now a simple calculation gives

$$\nabla_a \mathcal{L} v_b - \mathcal{L}\nabla_a v_b = v_c \nabla_a \nabla_b v^c$$

But $\mathcal{L}v_b = \alpha_b$ and $\mathcal{L}\nabla_a v_b = \mathcal{L}(\omega_{ab} + v_a\alpha_b)$. So (8.73), (8.75) and (8.76) give

$$\nabla_a \alpha_b = \alpha_a \alpha_b + v_c \nabla_a \nabla_b v^c$$

Hence

$$\nabla_{[a}\alpha_{b]} = 0$$

and so by (8.49) the motion is isometric.

In the previous section we proved that every isometric motion is rigid. We have now proved a partial converse, the Herglotz-Noether theorem, namely:
Every rotating rigid motion in flat space-time is isometric. We shall now show, however, that there do exist curved space-times which admit non-isometric rotating rigid motions. For example, consider the transformation of the metric given by

$$g_{ab} \rightarrow \bar{g}_{ab} = g_{ab} - x v_a v_b + 2y\, v_{(a} s_{b)} + (1-x)^{-1} y^2 s_a s_b$$

where x,y are arbitrary scalar functions of position with $|x| < 1$, and s_a is an arbitrary vector field orthogonal to v^a, i.e. $v^a s_a = 0$. This necessitates a corresponding transformation

$$v^a \rightarrow \bar{v}^a = (1-x)^{-1/2}\, v^a$$

in the velocity vector to retain its normalization. Then we see that

$$\bar{h}_{ab} \stackrel{\text{def}}{=} \bar{g}_{ab} - \bar{v}_a \bar{v}_b = g_{ab} - v_a v_b = h_{ab}$$

and hence a congruence of particle world-lines which formed a rigid motion in the old metric will do so also in the new metric. But clearly the condition for an isometric motion will not necessarily be preserved, and such a transformation can be made without annihilating the angular velocity. Thus by such a transformation one can construct non-isometric rotating rigid motions. The Ricci tensor of the new space, however, will in general be very complicated and it is not possible to say whether or not it represents a physically possible matter distribution.

Now consider a non-rotating rigid motion in a curved space time. Then

$$\nabla_a v_b = v_a \alpha_b$$

Hence

$$v_{[a} \nabla_b v_{c]} = 0$$

which is the condition for the congruence to be hypersurface-orthogonal. Furthermore, $\perp \nabla_a v_b = 0$, from which we see that these hypersurfaces have zero extrinsic curvature. In flat space-time this means that the particle world-lines of a non-rotating rigid motion are orthogonal to a family of hyper-planes. In any space-time it means that the motion of the whole body is known if the motion of any one particle is known. For we have just to construct the family of hyper-surfaces formed by all the geodesics through and orthogonal to the known particle path, and the other particle paths must be the orthogonal trajectories of these hypersurfaces.

8.7. DYNAMICS

All our work so far has dealt only with the kinematics of the motion; we have not concerned ourselves with the influence of the body on the space-time through which it is moving. We shall now prove one simple result which needs dynamical considerations.

The gravitational field equations are, from section 6.9,

$$G_{ab} \overset{\text{def.}}{=} R_{ab} - \tfrac{1}{2} R g_{ab} = -K T_{ab} \tag{8.76}$$

where $K = 8\pi k/c^4$ and T_{ab} is the energy-momentum tensor of the body. We must now ask what is the connection between T_{ab} and the velocity field v^a? We shall adopt the point of view of Synge[8] concerning the physical interpretation of the

8. J. L. Synge, Relativity: The Special Theory. North-Holland Publ. Co., Amsterdam (1960).

energy-momentum tensor; he identifies the kinematical velocity v^a with the dynamical velocity, i.e. the time-like eigenvector of T_{ab}. Thus we have

$$T_{ab} v^b = \rho v_a \qquad (8.77)$$

defining a scalar field ρ, which may be interpreted as the energy density of the body.

Now for a rigid body, on using (8.60) and the conservation equation $\nabla_b T^{ab} = 0$, we have

$$\begin{aligned} \frac{d\rho}{ds} = v^a \nabla_a \rho &= \nabla_a(\rho v^a) \quad \text{as } \nabla_a v^a = 0 \\ &= \nabla_a(T^{ab} v_b) \quad \text{by } (8.77) \\ &= T^{ab} \nabla_a v_b \qquad (8.78) \\ &= T^{ab} v_a \alpha_b \\ &= \rho v^b \alpha_b = 0 \end{aligned}$$

and so the density is constant along any particle path.

Now we easily see that

$$h^{ad} h^{bc} \perp R_{abcd} = h^{ad} h^{bc} R_{abcd} = -2\, G_{ab} v^a v^b$$

and so, on multiplying (8.70) by $h^{ad} h^{bc}$ we get

$$\mathcal{L}\left(G_{ab} v^a v^b - 3\omega^2\right) = 0 \qquad (8.79)$$

where

$$\omega^2 \overset{\text{def.}}{=} \tfrac{1}{2}\, \omega^{ab} \omega_{ab}$$

is the magnitude of the angular velocity. But by (8.76) and (8.77),

$$G_{ab} v^a v^b = -\kappa\rho$$

and so

$$\mathcal{L}\left(G_{ab} v^a v^b\right) = -\kappa \frac{d\rho}{ds} = 0$$

Equation (8.79) then gives

$$\frac{d\omega}{ds} = 0 \qquad (8.80)$$

and so we see that for a rigid body the angular velocity along any particle path is of constant magnitude. We note that the same result follows if, instead of dealing with a heavy body satisfying (8.77), we consider a test body in vacuo, so that $G_{ab} = 0$. The result for a heavy body was first proved by Rayner.[5]

8.8. INTEGRABILITY CONDITIONS

So far we have studied properties of rigid motions without answering the question: under what conditions do such motions exist? Suppose we are given a particular space-time, so that g_{ab}, R_{abcd}, etc. are known functions of the coordinates. We should like to find what conditions must be satisfied by the geometry and the initial conditions imposed on the body in order that there should exist a rigid motion in the space-time with these initial conditions. This is the problem of constructing integrability conditions for the equations of rigid motion. From (8.53) the equations we are concerned with can be put in the simple form

$$\nabla_{(a} v_{b)} - v^c (\nabla_c v_{(a}) v_{b)} = 0 \quad , \qquad v^a v_a = 1$$

but in this form, where the independent variables are the four v^a, the integrability conditions are very difficult to construct. To make the problem easier, we increase the number of independent variables by treating α^a and ω^{ab}, as well as the v^a, as unknowns. It is then possible to write the equations in a form which expresses the derivatives

$$\nabla_a v_b \ , \ \nabla_a \alpha_b \ , \ \nabla_a \omega_{bc}$$

as algebraic functions of the geometry and the unknowns, and then integrability conditions can be constructed in a well-known manner. The first equation of the new set is very simple; we have

$$\nabla_a v_b = \omega_{ab} + v_a \alpha_b$$

but the others are very complicated. The problem is solved by Pirani and Williams,[4] but we shall not go into details here. The equation (8.70) plays an important part in the work.

9. COSMOLOGY

9.1. INTRODUCTION

Cosmology is concerned with the structure and evolution of the Universe as a whole and attempts to give a description of its geometry and distribution of matter. Local irregularities are neglected despite the fact that these irregularities occur in regions that are rather large. The existence of cosmology as a science rests on the assumption that there are large scale overall regularities in the Universe and this assumption is usually expressed in terms of a Cosmological Principle. The cosmological principle differs in detail from one cosmological theory to another but broadly speaking it expresses the fact that the Universe is more or less the same everywhere.

We shall now present some (more precise) formulations of the cosmological principle.

(i) There exists a "cosmic time" t or a set of privileged hypersurfaces in space and time such that the geometry of these hypersurfaces is the same at all points. In addition to this homogeneity condition we sometimes require the hypersurfaces to be isotropic.

(ii) A second possibility is to start with a privileged family of observers or world lines such that the Universe appears to be the same for all observers at different points of space, at different instances of time and in all directions. In Riemannian space-time, this implies the existence of a cosmic time (the world lines form a normal congruence). This principle implies a "steady state" cosmology.

The simplest model that one can envisage is one whose geometry is Newtonian and whose matter distribution is uniform and at rest in a particular inertial frame. This model however, leads to difficulties which we shall discuss.

(i) Olbers' Paradox (1826)

Assuming the above mentioned model, there is no reason why it should be dark at night! This is seen as follows. If the Universe is populated uniformly with stars emitting radiation at a constant rate then we can calculate the radiation falling on a given area at any point in space. If the intensity of radiation falls off as $1/r^2$ and n is the number of stars per unit volume, then the total amount of radiation per unit area per second is given by

$$\int_0^\infty n \frac{E}{4\pi r^2} \cdot 4\pi r^2 dr,$$ which diverges.

Even when we regard the stars as extended "black bodies" and take into account the absorption of radiation by them, we find that the intensity of light at any point is the same as the average intensity of radiation on the surface of the stars. It does not seem possible to remedy this paradox without abandoning the idea of a static Universe.
(ii) Newton's law of attraction when applied to an infinite distribution of matter of constant density leads to an infinite potential viz.

$$\phi = \int \frac{\kappa\rho}{r} dV \sim \infty$$

This equation, however, follows from Poisson's equation

$$\nabla^2\phi = 4\pi\kappa\rho$$

only under the boundary conditions

$$\phi = O\left(\frac{1}{r}\right) \quad and \quad |\nabla\phi| = O\left(\frac{1}{r^2}\right)$$

which can be made only if ρ vanishes sufficiently rapidly when these boundary conditions are not satisfied; if ρ is a constant, a typical solution of Poisson's equation is

$$\phi = \frac{4\pi\kappa\rho}{6} r^2$$

which unfortunately also contradicts the initial assumptions in that, first, this solution appears to distinguish the origin from other points since the gravitational force vanishes only at the origin, and secondly, in view of the nonvanishing force everywhere but at the origin, the stars cannot be at rest.

Neumann and Seeliger (in 1895) proposed the idea of replacing Poisson's equation by

$$\nabla^2\phi - \lambda\phi = 4\pi\kappa\rho.$$

This corresponds to assuming that the gravitational forces have a finite range with $1/\sqrt{\lambda}$ being the characteristic length for the gravitational interactions. This equation has the constant solution

$$\phi = \frac{4\pi\kappa\rho}{\lambda}$$

for constant ρ, which is in agreement with the cosmological principle and the static nature of the Universe.

With the advent of general relativity, it seemed a natural course of action to apply the Einstein equations

$$R_{ab} - \tfrac{1}{2} g_{ab} R = -8\pi\kappa T_{ab} \tag{9.1}$$

to cosmology. Einstein himself thought that these equations would be unsatisfactory for two reasons:
(i) The equations (9.1), in the Newtonian limit, reduce to Poisson's equation which at that time seemed to be unsatisfactory.
(ii) Under the influence of Mach, Einstein believed that in the absence of matter there should be no geometric structure to space-time, i.e. with $T_{ab} = 0$ there should be no solutions to equations (9.1).

Therefore, Einstein modified the field equations by adding a cosmological term

$$\lambda g_{ab} + R_{ab} - \tfrac{1}{2} g_{ab} R = -8\pi\kappa T_{ab} \tag{9.2}$$

However, these equations are not the analog of the Neumann-Seeliger equation in the Newtonian limit but go over into

$$\nabla^2\phi + \lambda c^2 = 4\pi\kappa\rho.$$

Furthermore, it was found that with $T_{ab} = 0$, equations (9.2) did have solutions. In 1922 Friedmann found a number of non-static solutions of (9.1) which satisfied the cosmological principle. In view of these arguments Einstein later abandoned the cosmological term in the field equations.

9.2. NEWTONIAN COSMOLOGY

We choose as our field equations in Newtonian theory

$$\nabla^2\phi + \lambda c^2 = 4\pi\kappa\rho. \tag{9.3}$$

For constant ρ, a typical solution is

$$\phi \sim r^2$$

and hence

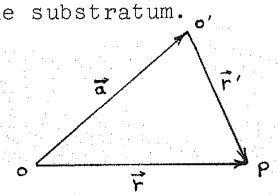
$$\nabla\phi \sim r$$

which appears to violate the cosmological principle. However, we know that in Newtonian theory, the existence of a strong gravitational field extending to infinity forces us to replace the group of Galilean transformations leading from one inertial system to another, by a larger group involving accelerations, and it can be shown that in these circumstances, the cosmological principle is not violated.

From the cosmological principle we infer that

$$\rho = \rho(t)$$

and we introduce a velocity field $\vec{v}(\vec{r},t)$ characterizing the motion of the substratum.



In the figure, O, O' and P are points of the substratum, O and O' are associated with observers, and P is any point.

$$\vec{r}' = \vec{r} - \vec{a}$$

$\vec{v}(\vec{r},t)$ = Velocity of P as observed by O

$\vec{v}'(\vec{r}',t)$ = Velocity of P as observed by O'

and consequently $\vec{v}(\vec{r},t) = \vec{v}(\vec{a},t) + \vec{v}'(\vec{r}',t)$.

The cosmological principle states that if O measures the velocity of a point P to be $\vec{v}$, and if the point P has the same relation to O' as P has to O, then the velocity of P' as measured by O' should also be $\vec{v}$. From this argument we deduce that

$$\vec{v}(\vec{r},t) = \vec{v}'(r,t)$$

or

$$\vec{v}(\vec{r},t) - \vec{v}(\vec{a},t) = \vec{v}(\vec{r}-\vec{a},t).$$

Thus $\vec{v}$ is a linear homogeneous function of $\vec{r}$ and we can write

$$v^{\alpha} = A^{\alpha}{}_{\beta} r^{\beta}$$

where $A^{\alpha}{}_{\beta}$ is a tensor. However, if we assume that the motion of the substratum is isotropic, then since, in general, a tensor distinguishes directions, we must write $A^{\alpha}{}_{\beta}$ as a multiple of the unit tensor i.e.

$$\vec{v} = A(t)\vec{r} \tag{9.4}$$

which are the differential equations of motion for the substratum.

We can always write

$$A = \frac{\dot{R}}{R}$$

and integrate (9.4) to obtain

$$\vec{r}(t) = R(t)\,\vec{r}_0$$

where we have put

$$\vec{r}_0 = \vec{r}(t_0)$$

and

$$R(t_0) = 1.$$

We can now supplement equation (9.3) by writing down the equations of motion for a perfect fluid without pressure, viz.

$$\frac{d\vec{v}}{dt} = -\operatorname{grad}\phi \qquad (9.5)$$

and the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\vec{v}) = 0. \qquad (9.6)$$

Equation (9.6) implies

$$\rho(t) = \frac{\rho(t_0)}{[R(t)]^3}.$$

If we now assume

$$\partial_\alpha \partial_\beta \phi = \delta_{\alpha\beta}\,\Phi(t) \qquad \alpha,\beta = 1,2,3 \qquad (9.7)$$

then this ensures that the curvature tensor is spherically symmetric and constant over the surfaces of constant time, a situation which is necessary to satisfy the requirements of homogeneity and isotropy. It then follows from (9.3), (9.5) and (9.7) that

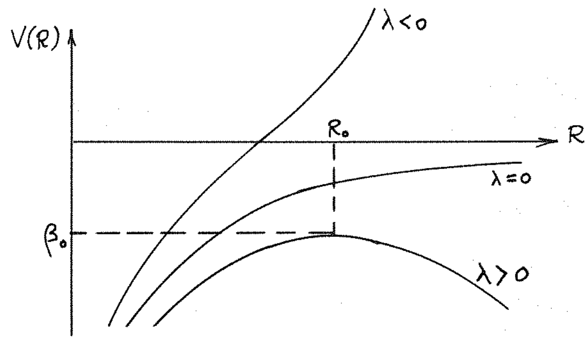
$$\dot{R}^2 - \frac{8\pi\kappa\rho_0}{3}\frac{1}{R} - \frac{\lambda c^2 R^2}{3} = \beta \qquad (9.8)$$

which can be interpreted as the energy integral if we regard R as the coordinate of a particle moving in 1-dimension and if we regard

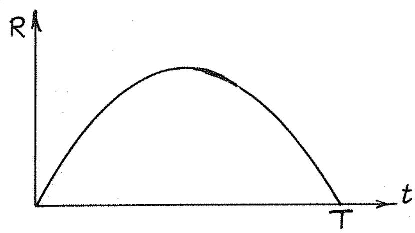
$$-\frac{8\pi\kappa\rho_0}{3}\frac{1}{R} - \frac{\lambda c^2}{3}R^2 = V(R) \qquad (9.9)$$

as the potential energy.

We can represent the solutions of these equations graphically:
First we plot $V(R)$ against R for various λ.



The constant β characterizes the solutions or time development of the system.

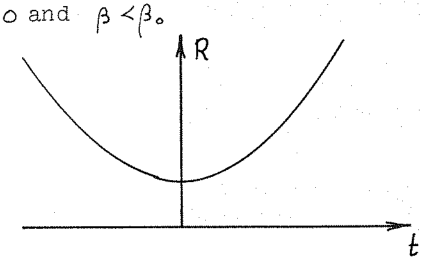


This is a bounded orbit and is valid when

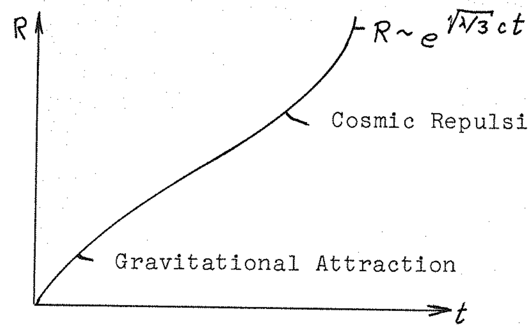
(i) $\lambda < 0$ and all β

(ii) $\lambda = 0$ and $\beta < 0$

(iii) $\lambda > 0$ and $\beta < \beta_0$

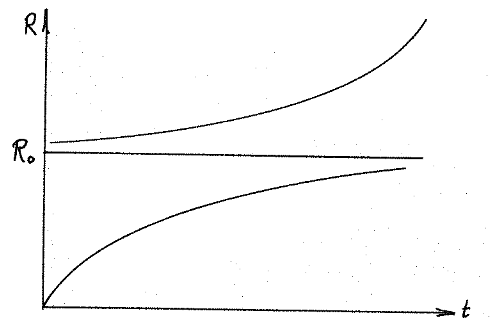


This situation can also prevail when $\lambda > 0$ and $\beta < \beta_0$



This is the situation when $\lambda > 0$, $\beta > \beta_0$.

When 1/R and β are small compared with the cosmic term $R \sim e^{\sqrt{\lambda/3}\,ct}$



These are the solutions when $\lambda > 0$, $\beta = \beta_0$

The solution $R = R_0$ is the Einstein model and is the only static solution. It is not very interesting because of its instability.

The value of β can be shown to be related to the characteristic time of the solutions. In the middle figure on p.234, for $\lambda = 0$, $\beta < 0$, it can be shown that

$$T = \frac{8\pi^3 \kappa \rho_0}{3|\beta|^{3/2}}$$

Suppose we wish to correlate our theoretical models with observations. One of the few things we know about the structure of the Universe is Hubble's constant which tells us the rate of change of the red-shift with distance. Hence we want our theory to give an account of the propagation of light so that we can calculate the red-shift. In order to write down Maxwell's equations in Newtonian theory we must introduce the ether which is a privileged inertial system, at each point of space, absolutely at rest. Alternatively we can define the ether as a rigging of the hypersurface t = const, where t is the absolute time and a surface is said to be rigged if at each point of the surface there is defined a direction not tangent to the surface.

From the hypersurface

$$t = \text{constant}$$

we form

$$t_a = \partial_a t$$

and suppose we are given a rigging characterized by a vector field u^a such that

$$u^a t_a = 1 .$$

Define

$$\gamma^{ab} = u^a u^b - g^{ab}$$

where g^{ab} is the singular metric of Newtonian theory. We can show that γ^{ab} is non-singular and Maxwell's equations can be obtained simply by using γ^{ab} to raise indices. We recall that using g^{ab} to raise indices does not give Maxwell's equations. We have

$$f_{ab} = 2 \ A_{[a,b]}$$

which imply

$$f_{[ab,c]} = 0$$

and these correspond to

$$\operatorname{div} \vec{H} = 0 \qquad \operatorname{curl} \vec{E} + \frac{\partial \vec{H}}{\partial t} = 0$$

but on raising indices with g^{ab} we get

$$f_{a}{}^{b}{}_{,b} = 0$$

which imply

$$\operatorname{div} \vec{E} = 0 = \operatorname{curl} \vec{H}.$$

Owing to the singularity of g^{ab}, it is impossible to obtain the term $\partial E/\partial t$ in the last of Maxwell's equations. If we use γ^{ab} in place of g^{ab} however, we obtain the correct form of Maxwell's equations.

In particular when

$$u^a \stackrel{*}{=} \delta^a_0$$

(where $\stackrel{*}{=}$ here means equal in a particular inertial frame) Maxwell's equations in this frame have the same form as in special relativity.

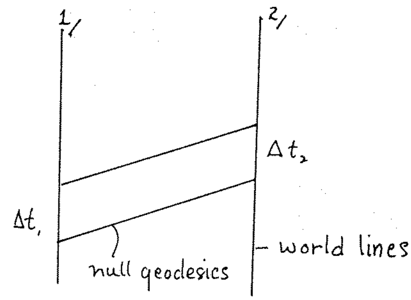
In cosmology a natural rigging is provided by the world lines of the particles of the substratum and in this case (in an inertial frame)

$$u^a = (1, \vec{v})$$

where $\vec{v}$ is the velocity of the substratum and the eikonal equation (derived from Maxwell's equations) describing the propagation of light is

$$(\nabla \psi)^2 - \left(\frac{\partial \psi}{\partial t} + \vec{v} \cdot \nabla \psi \right)^2 = 0. \qquad (c = 1)$$

From this we can deduce the formula for the change in the frequency of light from distant galaxies, viz.



$$\frac{\omega_2}{\omega_1} = \frac{\Delta t_1}{\Delta t_2} = \frac{R(t_1)}{R(t_2)}$$

where ω_1 and ω_2 are the frequencies of the light. This red shift explains Olbers' paradox. Note that there is a red shift only if the Universe expands i.e. if

$$R(t_2) > R(t_1)$$

9.3. RELATIVISTIC COSMOLOGY

We shall give a cosmological principle which leads to the important class of Friedmann solutions.
(i) The world lines of particles of the substratum form a normal congruence (i.e. hypersurface orthogonal) of time-like geodesics. This specifies a "cosmic time" viz. the proper time measured along these geodesics from a given hypersurface.
(ii) The hypersurfaces t = constant are isotropic. It can be shown that this implies that these hypersurfaces are spaces of constant curvature.
(iii) The motion of the substratum is non-shearing. A, B, C are particles of the substratum at a given time and A' B' C' are the same particles at a later time. The condition for no shear is that the triangles ABC and A'B'C' be similar.
We choose coordinates such that

$$t = \text{proper time}$$

along the world lines of the substratum particles and n^α ($\alpha = 1,2,3$) be constant along each world line. It follows then, that

$$ds^2 = dt^2 - [R(t)]^2 h_{\alpha\beta}\, dx^\alpha dx^\beta$$

where $h_{\kappa\beta}\, dx^{\alpha}\, dx^{\beta}$ is the time-independent line element of a space of constant curvature.

Every space of constant curvature can be realized (locally at least) as the surface of an n-dimensional hypersphere in an (n+1) Euclidean or pseudo-Euclidean space. It is easy to calculate the line element of this sphere e.g.

$$dl^2 = \frac{dx^2 + dy^2 + dz^2}{\left[1 + \frac{K}{4}(x^2+y^2+z^2)\right]^2}$$

where

$$\begin{aligned} K &= +1 \quad \text{for curvature} > 0 \\ &= 0 \quad \text{for curvature} = 0 \\ &= -1 \quad \text{for curvature} < 0. \end{aligned}$$

With this expression for $h_{\kappa\beta}$ (or the corresponding expression in polar coordinates) we can calculate the curvature tensor and hence the Einstein tensor.

If we now take the form of the energy-momentum tensor to be that of a perfect fluid without pressure viz.

$$T^{ab} = \rho u^a u^b$$

where u^a is the 4-velocity of the substratum, then the Einstein field equations become

$$\dot{R}^2 - \frac{8\pi K \rho_0}{3}\frac{1}{R} - \frac{\lambda c^2 R^2}{3} = -Kc^2 \qquad (9.10)$$

provided

$$\dot{R} \neq 0$$

and

$$\rho_0 = \rho(t_0), \quad R(t_0) = 1.$$

This is called Friedmann's equation and has exactly the same form as (9.8) in the Newtonian theory, except that in (9.10) the constant on the right-hand side is the value of the curvature of the hypersurface $t = t_o$. If $R = 0$, in addition to (9.10) we must also have

$$\frac{K}{R^2} = \lambda$$

which determines the relation between λ and ρ_o. This case corresponds to the static Einstein Universe and if $\lambda > 0$ then $K > 0$ and the Universe can be thought of as being a product of a 3-sphere and the time line, i.e. the Universe is cylindrical. This Universe however, has been shown to be unstable.

Another special case is the de Sitter Universe in which $\rho = 0$ and $\lambda > 0$. One possibility for writing its line element is to choose $K = 0$, and then

$$ds^2 = dt^2 - e^{2t/T}(dx^2 + dy^2 + dz^2)$$

where

$$T = \sqrt{\frac{3}{\lambda c^2}}.$$

The de Sitter Universe in a space of constant 4-dimensional curvature.

9.4. PROPAGATION OF LIGHT IN RELATIVISTIC COSMOLOGY

We saw above that with the assumptions of section 9.3 the cosmological metric can be written in the form

$$ds^2 = dt^2 - [R(t)]^2 d\sigma^2 \qquad (9.11)$$

where dt is the proper time along the world-lines x^α = const of the particles of the substratum, and $d\sigma^2$ is a time- independent line element of the hypersurfaces t = const. But for studying the propagation of light in this space, as we know that the paths of light rays (i.e. null geodesics) are invariant under a conformal transformation, we may consider instead of (9.12) the metric

$$d\tilde{s}^2 = ds^2/[R(t)]^2 = dt^2/[R(t)]^2 - d\sigma^2$$

Now define a new time-coordinate t' by

$$dt' = dt / R(t). \tag{9.12}$$

Then

$$d\tilde{s}^2 = (dt')^2 - d\sigma^2$$

which is a static metric. We may therefore apply to it the theory of section 6.5. This tells us that if we have two observers 1 and 2 traveling along lines x = const, i.e. with particles of the substratum, then if 1 emits two pulses of light at temporal separation $(d\tilde{s})_1 = (dt')_1$ which are observed by 2 to have temporal separation $(d\tilde{s})_2 = (dt')_2$, then

$$(dt')_1 = (dt')_2. \tag{9.13}$$

But it also tells us that if the frequency of the light emitted by 1 is ω_1 and 2 sees it as having frequency ω_2, then

$$\frac{\omega_1}{\omega_2} = \frac{(ds)_2}{(ds)_1} = \frac{(dt)_2}{(dt)_1}.$$

So using (9.12) and (9.13) we get

$$1 + z \overset{\text{def}}{=} \frac{\omega_1}{\omega_2} = \frac{R(t_2)}{R(t_1)} \simeq 1 + \left.\frac{\dot{R}}{R}\right|_{t_1} (t_2 - t_1) \tag{9.14}$$

where $\dot{R} \overset{\text{def}}{=} dR/dt$ and t_1 is the time of emission by 1, t_2 the time of reception by 2. z is the parameter usually used to describe the red-shift (which it is for $\dot{R} > 0$). If we interpret $c(t_2 - t_1)$ as the distance between the emitter and observer, (9.14) says that the red-shift z increases linearly with distance. This is Hubble's law, and the constant $\dot{R}/R$ can then be identified with Hubble's constant. This interpretation, however, can be made only when $c(t_2 - t_1)$ is (cosmologically speaking) small. For when $(t_2 - t_1)$ is large, not only does the approximation in (9.14) break down, but also the interpretation of $c(t_2 - t_1)$ as distance can no longer be justified.

Exercise: Show that if g^{ab} is the degenerate metric tensor of Newtonian theory and u^a is the velocity 4-vector of the substratum in Newtonian cosmology, then $\gamma^{ab} \stackrel{\text{def}}{=} u^a u^b - g^{ab}$ is a Riemannian metric and that the corresponding time-element can be reduced by a coordinate transformation to the form

$$dt^2 - [R(t)]^2 (dx^2 + dy^2 + dz^2).$$

9.5. DISTANCE IN COSMOLOGY

We are led, from the discussion in the preceding section, to consider how we are to define distances in cosmology. Since we do have, in both Newtonian and relativistic cosmology, a world-time t, clearly the simplest definition of distance would be to say that the distance between two bodies at time t is the length of the geodesic in the hypersurface t = constant joining them. (By this we mean that the hypersurface t = constant is itself to be considered as a manifold, and the geodesic is a geodesic in this manifold.) But such a definition is of no practical use. Instead, we seek a definition of an observational nature, which can actually be used by astronomers to determine the distances of distant nebulae.

If we know the actual size of a distant nebula, then we can define an observational distance by

$$\text{distance} = d/\alpha , \tag{9.15}$$

where d is the actual diameter of the nebula and α is the observed angular diameter. Such a definition would be satisfactory if some means of determining d were known. In practice, however, what is used is a definition based on the apparent luminosity of a nebula or some star in it, and we shall now consider how this may be done.

Let E be the energy radiated per unit time by the distant nebula, and let I be the intensity of radiation received per unit area per unit time by the observer. Then if we neglect the diminution of energy caused by the red-shift, we can reasonably define the distance of the nebula to be $[E/4\pi I]^{1/2}$. But due to the red-shift, the energy of each photon of light is reduced by a factor $\omega_1/\omega_2 = (1 + z)$, and the number of photons per unit time which reach the observer is also reduced by the same effect by another factor of $(1 + z)$. To counterbalance this reduction then, we should replace I in the above formula by $I(1 + z)^2$. We can then define the distance d, the <u>luminosity-distance</u>, by

$$d^2 = \frac{E}{4\pi I (1+z)^2} . \qquad (9.16)$$

This is effectively the definition of distance used by astronomers.

In the relativistic cosmology of section 9.3, we saw that the metric can be put in the form

$$ds^2 = dt^2 - [R(t)]^2 \frac{dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2)}{(1 + \frac{1}{4}Kr^2)^2} \qquad (9.16a)$$

where $K = 0$ or ± 1. It can easily be shown that for this metric, if the observer is at $r = 0$ and is observing a particle of the substratum at $r = r_1$, the light from which is reaching the observer at time $t = t_2$, then

$$d = \frac{r_1 R(t_2)}{1 + \frac{1}{4}Kr_1^2} . \qquad (9.17)$$

For the de Sitter universe, $K = 0$ and $R(t) = e^{t/T}$. Then using (9.17) one can easily show that in this case

$$z = d/T . \qquad (9.18)$$

No approximation is made here, and so for a de Sitter universe we see that Hubble's law is exact and that the value of Hubble's constant is T^{-1}. Interpreting our universe according to (9.18), the best estimates to date of the value of T give $T \sim 1.3 \times 10^{10}$ years.

We note that the definition (9.16) of distance involves E, the absolute luminosity of the source, which is not observationally measurable. This definition thus appears to suffer from the same defects as (9.15), where the diameter of the source was not observationally determinable. But from observations of nearby galaxies, whose distances have been determined by other means and so whose absolute luminosities may be calculated, it appears that all nebulae have roughly the same absolute luminosity. Since the observations of interest in cosmology deal with the number of nebulae which lie in a certain range of distance from us, it is then a good approximation to take all nebulae as having the same absolute luminosity.

There is still a flaw in this argument. We are observing distant nebulae as they were a long time ago, when they were much younger than they are now. So in an evolutionary theory of the universe, the mean age of the distant galaxies as we see them is much less than the mean age of nearby galaxies, and so there is no reason to believe that they will have the same mean luminosity.

There is one cosmological theory based on the assumption that, for privileged observers moving with the substratum, the universe presents the same view, irrespective of where they look around. This is known as the Perfect Cosmological Principle. This assumption, together with the experimentally established expansion of the universe, lead to the de Sitter metric as the background for the steady-state theory. Since the mean density of matter must be constant in time by the perfect cosmological principle, to make up for the decrease that would be caused by the expansion of the universe, matter must be continually created. If ρ_0 is the mean density of matter, the rate of creation may easily be seen to be

$$\frac{3\rho_0}{T} \sim 10^{-46}\,\mathrm{gm\,cm^{-3}\,sec^{-1}} \sim \begin{cases} \text{one proton per liter} \\ \text{every } 5 \times 10^{11} \text{ years.} \end{cases}$$

According to the steady-state theory, the average number of galaxies per unit volume is constant; when galaxies run away new ones get formed in their place from the matter that is being created everywhere.

In contradistinction to this theory, any conservative theory predicts that in the past the density of galaxies was higher, and that more distant galaxies are older than they appear to us to be. This is a point that can be checked by observation. Optical observations are inconclusive in this respect, but observations of radio galaxies tend to indicate thatat large distances there are more of them per unit volume than there are in the neighborhood of our galaxy. This tends to support evolutionary models in preference to the steady-state theory.

6.9. THE EVENT HORIZON OF THE DE SITTER UNIVERSE

We shall now take a closer look at the de Sitter universe to illustrate the occurrence of horizons in cosmological models. Consider the de Sitter metric in the form

$$ds^2 = dt^2 - e^{2t/T}(dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2)). \tag{9.19}$$

The null-cone through the point $r = 0$, $t = t_o$ has an equation of the form $r = r(t)$, and $ds = 0$ for displacements in it. Then (9.19) gives

$$dr = \pm e^{-t/T} dt$$

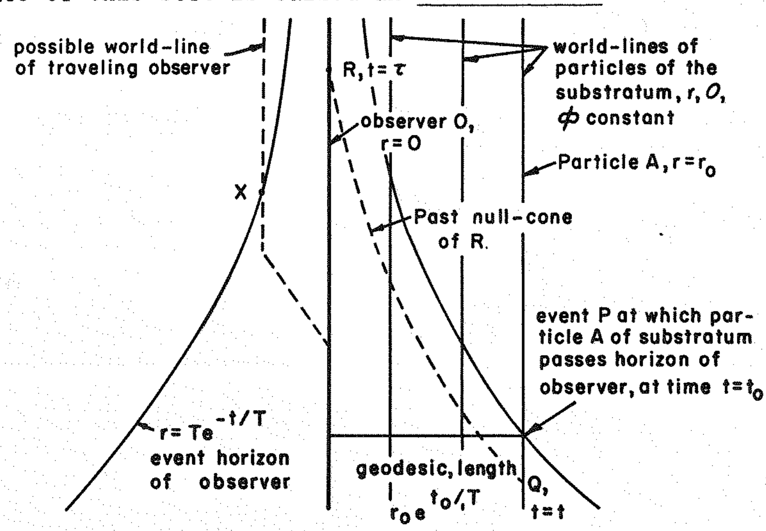
which integrates to give

$$r = \pm T(e^{-t/T} - e^{-t_o/T}). \tag{9.20}$$

The two signs correspond to the past (sign) and future (sign) null-cones respectively. Now at any particular time information can be received only from events inside the past null-cone through the observer. We thus see from (9.20) that an observer whose world-line is $r = 0$ can <u>never</u> receive any information from events occurring outside the surface

$$r = Te^{-t/T}. \tag{9.21}$$

A surface of this sort is called an <u>event-horizon</u>.



The situation is illustrated in the preceding diagram. Vertical world-lines represent particles of the substratum. Horizontal lines lie in surfaces of constant t. Let us consider what the observer O sees while observing particle A, which has $r = r_o$. By (9.21) this particle crosses the observer's event horizon at P in the diagram, at time $t_o = T \log (T/r_o)$. O can see A only at events on A's world-line with $t < t_o$. O sees such an event by light which travels along a null geodesic (dotted line from Q to R in the diagram) and reaches him at time

$$\tau = -T \log \left(e^{-t/T} - e^{-t_o/T} \right)$$

and he then ascribes to A a luminosity-distance

$$d = r_o e^{\tau/T} = \frac{r_o}{e^{-t/T} - e^{-t_o/T}}$$

and a red-shift $z = d/T$. So in fact O sees nothing unusual occur; as $t \to t_o$ the light takes longer and longer to reach O from A, the luminosity-distance of A tends to infinity as $t \to t_o$ and so does the corresponding red-shift. This is quite normal in an expanding universe. The only peculiar phenomenon is that A disappears over O's horizon at a finite proper time to A.

At time t, we see easily that the geodesic distance (in a surface of constant t) of A from O is

$$D = r_o e^{t/T}$$

which is still finite at the point P. The velocity of recession, dD/dt, is

$$\begin{aligned} \frac{dD}{dt} &= \frac{r_o}{T} e^{t/T} \\ &= e^{(t-t_o)/T} \to 1 \quad \text{as } t \to t_o \end{aligned}$$

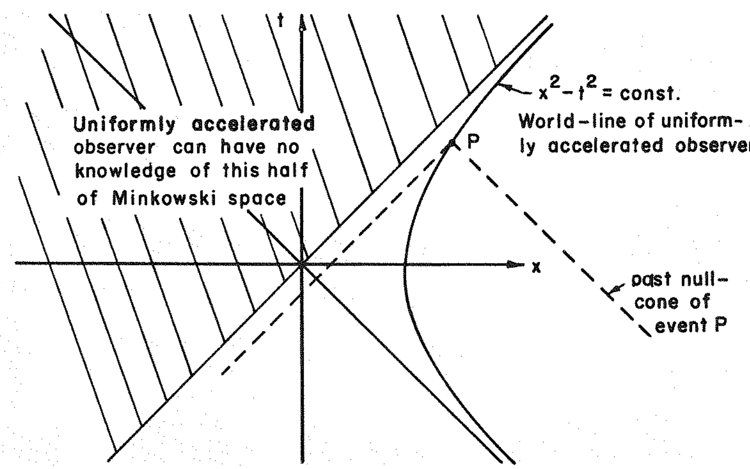
We thus see that this geodesically-measured velocity of recession tends to the velocity of light as the particle approaches the event-horizon.

Now so far we have considered only an observer at $r = 0$. But since the metric can be put in the form (9.19) with $r = 0$ being an arbitrary particle of the substratum (or alternatively, since the de Sitter universe is homogeneous), the above conclusions apply to any observer moving with the substratum.

Finally we shall consider if O can obtain information about what happens outside his horizon by sending a friend out into space, and getting him to send a report on what he sees. The world-line of such a traveling observer is indicated in the diagram. We easily see that although his friend will be able to see past O's horizon, he will not be able to do so until he passes the point X, on this horizon, and after he has passed that point he can neither return home to O nor send any information back to O.

So we see that O can never obtain any information about events outside his horizon, but he cannot neglect their existence, as by himself traveling around, he can change his horizon and find such forbidden knowledge. But once he knows it, he can never return home.

To conclude this section, we note that even in Minkowski space-time some observers have event-horizons.



This figure shows the case of a uniformly accelerated observer traveling in the x-direction. In suitably chosen rectangular coordinates his world-line has equation $x^2 - t^2 =$ const, $y = z = 0$. It is then clear from the diagram that light emitted from events in the shaded region will never reach the observer, who therefore has an event-horizon. But of course no such horizon exists for inertial observers.

For a detailed study of event-horizons, and the associated phenomenon of particle-horizons, in the case of the general metric (9.16a), see the article by Rindler.[1]

1. W. Rindler, M.N.R.A.S. 116, 662 (1956).

9.7. EPILOGUE

I want to finish this discussion of cosmology by saying that in my opinion it is not worthwhile to work in theoretical cosmology at the present time. I think it would be better to sit and wait for the astronomers to get more data on the motion and distribution of distant galaxies. I think it is the opinion of most people who work in that field that one can expect in the next few years to obtain such significant data, probably from radio-astronomy, and there are also hopes of putting a telescope into orbit.

<u>References</u>:

H. Bondi, <u>Cosmology</u>. C.U.P. (1960).
An extended list of further references is to be found there. Also
O. Heckmann and E. Schücking, Handbuch der Physik, vol. 53.
A. R. Sandage, Astrophys. J. <u>133</u>, 355 and <u>134</u>, 916 (1961).