RELATIVITY:

THE SPECIAL THEORY

BY

J. L. SYNGE

School of Theoretical Physics
Dublin Institute for Advanced Studies

Von Stund an sollen Raum für sich und Zeit für sich völlig zu Schatten herabsinken und nur noch eine Art Union der beiden soll Selbstständigkeit bewahren.

H. Minkowski, Raum und Zeit
[Physikalische Zeitschrift 10 (1909) 104]



1956
NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM

RELATIVITY: THE SPECIAL THEORY

SERIES IN PHYSICS

General Editors:

J. DE BOER, H. BRINKMAN and H. B. G. CASIMIR

L. ROSENFELD, Nuclear Forces
E. A. GUGGENHEIM, Thermodynamics
BØRGE BAK, Elementary Introduction to Molecular Spectra
L. ROSENFELD, Theory of Electrons
W. ELENBAAS, The High Pressure Mercury Vapour Discharge
S. R. DE GROOT, Thermodynamics of Irreversible Processes
E. A. GUGGENHEIM and J. E. PRUE, Physicochemical Calculations
E. A. GUGGENHEIM, Boltzmann's Distribution Law

J. BOUMAN (editor), Selected Topics in X-Ray Crystallography
J. G. WILSON (editor), Progress in Cosmic Ray Physics Volume I
J. G. WILSON (editor), Progress in Cosmic Ray Physics Volume II
C. J. GORTER (editor), Progress in Low Temperature Physics I
KAI SIEGBAHN (editor), Beta- and Gamma-Ray Spectroscopy

SOLE DISTRIBUTORS FOR U.S.A.
INTERSCIENCE PUBLISHERS INC., NEW YORK

PRINTED IN THE NETHERLANDS

TO ELIZABETH, BESSIE
AND EILIS

PREFACE

This book originated in the notes of lectures given over a number of years in graduate courses at the University of Toronto and elsewhere. The basic idea is to present the essentials of relativity from the Minkowskian point of view, that is, in terms of the geometry of space-time. This geometrical approach is used to some extent in all expositions of relativity, but I have emphasised it more than is customary, because it is to me (and I think to many others) the key which unlocks many mysteries. My ambition has been to make space-time a real workshop for physicists, and not a museum visited occasionally with a feeling of awe.

As originally planned, the book was to cover both the special and general theories of relativity. But as it was being written, the charm of the special theory so worked on me that I found it impossible to confine it to the required limits and, in the end, the general theory had to be omitted. This is not wholly regrettable, because the special theory is by far the more firmly embedded in modern physics and should not be overshadowed by the general theory, as tends to be the case. However, I have left in Chapter I a foundation strong enough to support both the special and general theories.

To understand a subject, one must tear it apart and reconstruct it in a form intellectually satisfying to oneself, and that (in view of the differences between individual minds) is likely to be different from the original form. This new synthesis is of course not an individual effort; it is the result of much reading and of countless informal discussions, but for it one must in the end take individual responsibility. Therefore I apologise, if apology is necessary, for departing from certain traditional approaches which seemed to me unclear, and for insisting that the time has come in relativity to abandon an historical order and to present the subject as a completed whole, completed, that is, in its essentials. In this age of specialisation, history is best left to historians.

I have tried to include most of the famous relativistic results and to develope those relativistic formulae which may be regarded as basic, but, having done so, I have felt at liberty to go a little further along

unfamiliar paths where the conclusions may, or may not, be physically significant. In every case I hope that I have made clear the assumptions from which the conclusions spring. In this category I may mention the reconstruction in relativistic form of the discrete system of Newtonian mechanics with action and reaction between its particles (Chapter VII), leading to a statistical concept of the energy tensor (Chapter VIII); the elastic collisions of particles (point-particles, that is) with conservation of angular momentum, leading to a surprising multiple determinacy and some rather intricate algebra (Chapter VII); and the models of uncharged "particles" of finite energy constructed out of singularity-free solutions of Maxwell's equations in vacuo, and model photons constructed according to the same plan (Chapter IX).

I would like to express my debt to the late Dr. L. Silberstein whose lectures at the University of Toronto over thirty years ago started my interest in relativity, to colleagues past and present, in particular Professors L. Infeld, C. Lanczos and E. Schrödinger, for many friendly discussions, to Professor N. L. Balazs who read a considerable portion of the manuscript and was fruitful in suggestions, and to Dr. G. H. F. Gardner who supplied an essential and rather subtle mathematical argument (see p. 239). I am most grateful to Dr. F. A. E. Pirani for his painstaking collaboration in proof-reading and checking formulae, and for suggestions which have unkinked the exposition at many points.

Dublin, June 1955 J. L. S.

CONTENTS

Preface vii

CHAPTER I. THE SPACE-TIME CONTINUUM AND THE SEPARATION BETWEEN EVENTS

1. Concepts 1
2. Events and particles 5
3. Space-time 6
4. The assignment of space-time coordinates 7
5. Notation 8
6. World lines and space-time diagrams 9
7. The motion of a material particle 10
8. Past, present and future 11
9. Standard clocks 14
10. The separation between events. 15
11. The fundamental quadratic form 16
12. Finsler space-time and Hamiltonian methods 19
13. Space-time as a Riemannian space 22
14. Measurement of spacelike separation 24
15. The physical meaning of orthogonality 26
16. Distance between particles 29
17. Rigid rods 30
18. The world lines of free particles 32
19. The special and general theories of relativity 34
20. Rigid motions 36

CHAPTER II. INTRODUCTION TO THE SPECIAL THEORY

1. Basis of the special theory of relativity 38
2. Finite separations 39
3. How to draw a straight line in space-time 42
4. Pairs of straight lines in space-time, parallel and skew . 44
5. The physical meaning of the special coordinates 47
6. Splitting space-time into space and time 49

7. Galileian frames of reference 52
8. Proper time and the speed of light 54
9. Minkowskian coordinates 56

CHAPTER III. SPACE-TIME DIAGRAMS

1. Some elements of the geometry of flat space-time . . . 59
2. Orthogonal projections 61
3. Space-time diagrams 63
4. Space-time diagram of the null cone 64
5. M-geometry and E-geometry 65
6. Pseudospheres 67

CHAPTER IV. THE LORENTZ TRANSFORMATION

1. The general Lorentz transformation 69
2. Restrictions on Lorentz transformations 73
3. The two ways of interpreting transformations 75
4. Geometrical meaning of the Lorentz transformation . 76
5. Eulerian angles and pseudoangles 79
6. Lorentz transformations regarded as rigid body displacements 84
7. The Lorentz 4-screw 86
8. Reduction of any Lorentz transformation to a 4-screw 90
9. Correspondence between triads of null rays and unit orthogonal tetrads 94
10. Lorentz transformations represented by arbitrary transformations of triads of null rays 98
11. Spinors 103
12. The two spin transformations corresponding to a given Lorentz transformation 107
13. The simple Lorentz transformation between two frames of reference 110
14. Lorentz transformations with Hermitian (or symmetric) matrix 114

CHAPTER V. APPLICATIONS OF THE LORENTZ TRANSFORMATION

1. Apparent contraction of a moving body and apparent retardation of a moving clock 118
2. Snapshots 120

3. Space-time diagrams of contraction and retardation . . 122
4. Composition of velocities 126
5. The velocity 4-vector and the acceleration 4-vector . 130
6. Transformation of a wave motion 133
7. Reflection at moving mirrors 138
8. Fresnel's convection coefficient 142
9. Aberration 146
10. The expanding universe in special relativity 150
11. The red-shift 152
12. Luminosity and distance 153
13. The dependence of red-shift on apparent distance and the age of the universe 156
14. The Michelson-Morley experiment 158

CHAPTER VI. MECHANICS OF A PARTICLE AND COLLISION PROBLEMS

1. Force. Action and reaction. A philosophical digression 163
2. Particles and mass 165
3. Equations of motion 166
4. Is proper mass constant? 167
5. Interpretation of the equations of motion 168
6. Motion under a constant relative force and in a constant magnetic field 171
7. Momentum 4-vector for a photon 172
8. Collision and disintegration problems 173
9. Space-time diagrams of collisions 176
10. The triangle inequality in space-time 177
11. Mass-centre reference system. Release of energy in disintegration 180
12. Some numerical values 182
13. Inelastic collision of two particles 183
14. Disintegration of one particle into two 185
15. Emission of a photon from an atom 187
16. The sameness of photons 189
17. The emission and absorption of a photon 191
18. The Compton effect 193
19. The annihilation and creation of matter 199
20. Elastic collisions 205

CHAPTER VII. MECHANICS OF A DISCRETE SYSTEM

1. Discrete and continuous systems 208
2. Impulses and continuous forces 209
3. Internal impulses 210
4. The conservation of 4-momentum for a system. . . . 213
5. Angular momentum and its conservation 216
6. The mass-centre of a system 218
7. Intrinsic angular momentum of a particle 220
8. The geometrical representation of a skew-symmetric tensor 223
9. Elastic collisions with unchanged intrinsic angular momentum invariants. The case of identical material particles 227
10. Example of an elastic collision with intrinsic angular momentum invariants unchanged 235
11. General treatment of elastic collision with intrinsic angular momentum 237
12. Summary of procedure for solving a collision problem . 246
13. Particular cases of collisions 248
14. External impulses and impulsive torques acting on a system 251
15. The two-body problem 254

CHAPTER VIII. MECHANICS OF A CONTINUUM

1. Density 261
2. Fundamental laws of relative momentum and relative energy for a system 263
3. Impact of a stream of particles on a target 265
4. Pressure in a relativistic gas 267
5. Pressure due to the impact of photons 269
6. World tubes and their cross-sections 272
7. Green's theorem and the expansion of world tubes . . 276
8. The energy tensor of a continuous medium 281
9. The physical meaning of the energy tensor 285
10. The energy tensor for an incoherent stream of material particles 288
11. Eigen values of the energy tensor 290
12. Mean density, mean velocity and stress 296
13. Equations of motion of a continuous medium 300

14. The perfect fluid in relativity 302
15. Incompressible fluids 306
16. Isolated systems and the energy tensor 309

CHAPTER IX. THE ELECTROMAGNETIC FIELD IN VACUO

1. The electromagnetic tensor F_{rs}. 317
2. Lorentz transformations of the electric and magnetic 3-vectors 320
3. The energy tensor 322
4. Eigen values and principal directions for the electromagnetic energy tensor 325
5. The canonical forms for an electromagnetic field at an event 331
6. Eigen properties of the tensors F_{rs} and F^*_{rs} 336
7. The tensors F_{rs} and F^*_{rs} expressed in terms of invariants and principal null vectors 339
8. The 4-potential 345
9. Plane electromagnetic waves 350
10. Some special systems of plane waves 354
11. Reduction of a pair of sinusoidal plane wave systems. Interference 356
12. Some scalar wave functions 359
13. Generation of a Maxwellian field from a scalar wave function 363
14. An electromagnetic model of a material particle . . . 366
15. Superposition of elementary wave functions 372
16. A nearly static electromagnetic particle (β large) . . . 374
17. Model of a photon with $\beta = 0$ 375
18. Model of a photon with β small 379
19. Null 3-spaces and Green's theorem 383
20. Electromagnetic shock waves 385

CHAPTER X. FIELDS AND CHARGES

1. The discrete and continuous methods 387
2. The Coulomb field of an electric charge 388
3. The field of an accelerated charge 391
4. The ponderomotive force 394
5. The electromagnetic Kepler problem 396

6. Radiation of energy and third-order equations of motion 399
7. Maxwell's equations with current 403
8. Explicit formula for the 4-potential 405
9. The energy tensor of a field with current 410
10. Maxwell's equations derived from a variational principle 413
11. Maxwell's equations in moving matter 415

APPENDIX

A. 3-waves and 2-waves 419
B. Radiation of energy from an accelerated charge . . . 422
C. Scattering and capture by a fixed nucleus 426
D. The absolute 2-content of a 3-cell on a null cone . . . 430
E. Calculations for retarded potential 432
References 435
Index 439

NOTATION

Latin suffixes take the values 1, 2, 3, 4 and Greek suffixes the range 1, 2, 3, with summation over the appropriate range of values in the case of a repeated suffix. Any exceptions to this rule are explicitly noted.

Real coordinates are written with superscripts, x^r. Coordinates with an imaginary time are written with subscripts, x_r, with $x_4 = ict$.

Partial derivatives are often indicated by a comma ($f_{,r} = \partial f/\partial x^r$). Covariant derivatives are indicated by a vertical stroke, $f_{r|s}$.

The signs of the components of the fundamental tensor g_{mn} are chosen so that (for real coordinates) the diagonal form is (1, 1, 1, − 1) and *not* (− 1, − 1, − 1, 1).

CHAPTER I

THE SPACE-TIME CONTINUUM AND THE SEPARATION BETWEEN EVENTS

§ 1. CONCEPTS

A scientific theory may be divided into three parts: (*a*) foundations, (*b*) accepted dogma, (*c*) excursions. The foundations are axioms, principles or laws (e.g. Newton's laws of motion or the first and second laws of thermodynamics). The accepted dogma consists of deductions from the foundations confirmed by observation and experiment, linking reason with nature in a satisfying way (e.g. Newtonian mechanics as it stood before relativity was thought of). The excursions wander out of the domain of accepted dogma, sometimes arousing in cautious minds a feeling that they have more imagination than solid fact in them (e.g. Maxwell's electromagnetic theory of light at the time when he put it forward — all excursions are not so successful!). A scientific theory is a living thing which grows and changes; fruitful excursions extend the body of accepted dogma and critical scrutiny of the foundations clarifies and sometimes modifies them.

If someone, otherwise well equipped in knowledge of mathematics and physics, wants to understand the theory of relativity, by what door is he to enter in? As the result of what process is he to find himself comfortably at home in the accepted dogma of relativity, capable of appreciating the foundations critically and able to discuss the excursions of others and to make his own?

Such questions are of course not pertinent to relativity only — they apply to any branch of theoretical physics, or indeed to any branch of science. But they are particularly difficult to answer satisfactorily in the cases of relativity and quantum mechanics, not merely because these subjects are comparatively young, but because they both uproot concepts usually accepted without question.

The ancient Greeks had their answer. The door by which they entered a subject was a set of axioms; granted these, all one had to do

was to follow the processes of logical thought. To the works of Euclid and Archimedes, written in this axiomatic spirit, modern science owes its being. But it is not as simple as it looks. Axioms, we now realise, are not the self-evident truths they were long supposed to be, but rather the rules of a game, the pieces of which are elements or concepts which remain and must ever remain undefined because there is nothing in terms of which to define them. This was brought to light by the researches of Hilbert into the foundations of geometry at the end of the nineteenth century, and the knowledge that any theory with a claim to logical structure must start with undefined elements and unproved propositions is slowly permeating through science.

In fact, although it seems most natural to start with the foundations and build on them, we now realise that the foundations of a theory are actually the most elusive and confusing part of it. Anyone who tries to put a physical theory on an impeccable axiomatic basis soon realises that he has undertaken a major task, absorbing all his energy and leaving none for the body of the theory in which his real interest lies. The axiomatisation of physics is of great interest, but it is a job for the specialist in axiomatics, and the fruits of his labour are likely to be enjoyed rather by fellow specialists than by theoretical physicists at large. In brief, axiomatics do not provide the door we are seeking.

In modern works on theoretical physics axiomatisation has been largely abandoned. Instead, the entry to a new subject is by what may be called the "cuckoo-process". The eggs are laid, not on the bare ground to be hatched in the clear light of Greek logic, but in the nest of another bird, where they are warmed by the body of a foster mother, which, in the case of relativity, is the Newtonian physics of the nineteenth century. The student is first thoroughly indoctrinated with Newtonian physics, and he accepts its concepts as true to physical reality. Then, step by step, the concepts are modified, until finally he bites off the head of his foster mother and flies from the nest a full-fledged relativitist.

This cuckoo-process follows the true order of historical development in science and it has the advantage that at every stage of the transformation the learner has the comforting support of familiar surroundings. As each support falls away, it is replaced by another, constructed to the new pattern. But it is confusing. The concepts of Newtonian physics interlock with one another (e.g. force, acceleration, inertial mass and gravitational mass), and until one has finally reviewed all Newtonian

concepts, there is always present a suspicion that the same word is being used with meanings which differ with the context.

The plan adopted in this book is a compromise. Formal axiomatics are avoided, but a serious effort is made to state the assumptions with enough clarity to show that relativity is essentially a logical structure for which any interested logician might seek a system of axioms if he wanted to. The appeal to undefined elements, essential in an axiomatic treatment but repellent to most physicists as an unnecessary mystification, is avoided by taking over concepts from Newtonian physics and setting them in a new background. There is, of course, a danger in this, for even in Newtonian physics different people have slightly different ways of looking at concepts, and so the theory of relativity, created in the mind of the individual, may have a slightly subjective character depending on who the individual is. This is unavoidable; only by the give-and-take of scientific conversation can anyone be sure that his concepts are the same as the concepts of others, or, if they are not the same, find out how they differ.

It might seem that this is the cuckoo-process over again, but it is not. We shall from the first turn a cold and sceptical eye on Newtonian physics, never admitting a bunch of interlocking Newtonian concepts but only concepts one at a time, alone and disinfected.

Let us now get to work. We start with a tabula rasa, a clean sheet, a mind in a state of intellectual nudity.

Into this void we admit at once the whole body of *pure mathematics*, or at least those portions of pure mathematics which we may have occasion to use later. *Applied* mathematics on the other hand is excluded, for almost all applied mathematics deals with Newtonian physics, and the words used in it evoke Newtonian concepts and these we are prepared to admit only singly and under scrutiny.

This embargo on applied mathematics is serious, for it excludes the dynamics of particles and rigid bodies, celestial mechanics, Lagrangian and Hamiltonian methods, hydrodynamics, elasticity and electrodynamics. The trouble with these subjects is that they all involve the Newtonian concept of *time*, and that, as we shall see in due course, is one of those Newtonian concepts which we shall *not* take over into relativity. If the subjects listed above are to appear in the theory of relativity, they must appear in a revised form. However the student of relativity has not wasted his time in the study of Newtonian subjects, for in relativity we frequently seek contact with Newtonian physics

in certain limiting cases, and such contacts are understandable only to one familiar with Newtonian theories.

The subject of *geometry* deserves special mention. Is it pure mathematics or applied mathematics? There is considerable confusion on this question, because the word is used to cover two entirely different things. In so far as it is axiomatic (logical deduction from axioms dealing with undefined elements), it is pure mathematics and as such admissible into our relativistic scheme, provided we do not subconsciously define the undefined elements physically and accept the axioms concerning them. In so far as geometry deals with the form of actual things and their measurement, it is definitely applied mathematics, and as such excluded from our relativistic scheme, at least until such time as we are ready to consider the possibility of admitting it.

There is another way of looking at this question of geometry. The positive integers, which correspond to the primitive physical operation of *counting*, form the basis of a considerable portion of mathematics. From them we derive negative numbers, irrational numbers and complex numbers, and hence the body of mathematical analysis. Relativity is not so revolutionary as to question the validity of counting, and in fact it accepts all mathematics based on counting. Thus the relationship of geometry to relativity is most satisfactorily established when geometry is regarded *analytically*, a "point" being nothing but a set of numbers (its coordinates) and a "line" a set of points. This is a most fruitful way to look at geometry, and we shall make extensive use of it. It is only when geometry purports to deal with "physical space" that we must view it with extreme caution. This does not mean that the geometry of physical space will not be discussed, but only that we reserve the right to discuss it at our own time and in our own terms.

For historical accounts of the theory of relativity, the reader may consult DUGAS [1950] and WHITTAKER [1953], or for more mathematical detail, PAULI [1920].[1] Out of the work of Lorentz and Poincaré the special theory of relativity emerged, EINSTEIN [1905] clearing up philosophical difficulties by destroying the concepts of ether and absolute simultaneity and MINKOWSKI [1909] giving the theory a clear mathematical form in terms of the geometry of space-time. After some tentative approaches, the general theory of relativity (the new theory

[1] For these and other references, see p. 435.

of gravitation) took final shape in 1915–1916 (EINSTEIN [1915, 1916])[1].

In the present chapter historical order is abandoned and a basis laid down wide enough to support both the special and the general theories. This avoids a cuckoo-process by which the special theory is first developed and later eaten up by the general theory; if this delays a little the reader's contact with the details of relativity, it will, it is hoped, save him from headaches later on.

We shall now start to take over Newtonian concepts into relativity, beginning with the concepts of *event* and *particle.*

§ 2. EVENTS AND PARTICLES

The word *event* does not occur frequently in Newtonian physics, but this is accidental, because the concept is quite clear. Anything that happens is an *event*, but (just as in geometry we sharpen the concept of a point) we sharpen the concept of an event to mean an occurrence which takes up no room and has no duration. To emphasise this sharpening we may call it a *point-event*, but it is unnecessary to do so because we shall always use the word *event* in this sharpened sense.

To stimulate the imagination we may think of an event as an explosion or collision, but this dramatic or catastrophic association is not essential. *Any* occurrence, sharpened as aforesaid, is an event, and we can of course imagine possible events as well as those which we think of as actually occurring. We shall presently consider the totality of all possible events (space-time).

The concept of a *material particle* also is taken over from Newtonian physics. We are familiar with the way in which it appears in physical theory — a moving point with a number (mass) associated with it, and perhaps a second number (electric charge). This concept we accept in relativity, with one slight reservation; perhaps the mass is not constant. We exercise the same discrimination as in Newtonian physics regarding the circumstances under which a real piece of matter may be treated as a particle; it may on occasion be an electron, an atom, a billiard ball, a planet, a star, or even a nebula.

It is convenient to introduce also the *particle of light* or *photon*, the properties of which will be discussed later. For the moment we think of it as a moving point. There is nothing un-Newtonian in the idea of a small parcel of light.

[1] English translations of a number of fundamental papers are contained in LORENTZ [1923].

We have now taken over from Newtonian physics the concepts of *event* and *particle* (material particle and photon). They are linked together by the fact that *the history of a particle is a continuous sequence of events.*

§ 3. SPACE-TIME

In Newtonian physics an event may be identified by four numbers (x, y, z, t), where (x, y, z) are the rectangular Cartesian coordinates of the place where it occurs and t the time at which it occurs. But we do not have to use these numbers. If we define (X, Y, Z, T) as four functions of (x, y, z, t), then the values of (X, Y, Z, T) serve to identify an event. The essential thing is that an event needs *four* numbers to identify it, and for that reason we say that in Newtonian physics *the totality of all possible events form a four-dimensional continuum.*

This italicised statement we take over into relativity, without necessarily taking over the chain of thought leading up to it. We are not yet ready to discuss whether the concepts of rectangular Cartesian coordinates (x, y, z) and time t are acceptable in relativity. Actually, we shall accept them later with important reservations. For the present let us not think about them, but accept as a fundamental hypothesis of relativity the statement that *the totality of all possible events form a four-dimensional continuum.* This continuum we call *space-time*; we are not at all in a position to remove the hyphen and speak of *space* and *time* separately.

In regard to space-time, the emphasis is on the word *four*. It is four-dimensional, not three-dimensional or five-dimensional. In saying that it is four-dimensional, we mean that an event is identified by four numbers, say (x^1, x^2, x^3, x^4), which numbers we call the *coordinates* of the event, or *space-time coordinates* if we want to be emphatic.

It is well at this point to interject a cautionary remark which does not affect the general line of thought. In Newtonian physics the space-time of all possible events is covered once over by the coordinates (x, y, z, t), each of these four coordinates ranging from $-\infty$ to $+\infty$. There is in fact a one-to-one correspondence between events and tetrads of numbers (x, y, z, t). It is not asserted that the space-time of relativity is covered once over by the coordinates (x^1, x^2, x^3, x^4) ranging from $-\infty$ to $+\infty$. It is enough to think that a portion of space-time is covered once over by these coordinates with suitable ranges. To cover the whole of space-time it may be necessary to use

several coordinate systems, one system for each of several portions of space-time, with an overlap where the portions meet.

§ 4. THE ASSIGNMENT OF SPACE-TIME COORDINATES

To keep physical theory in contact with physical reality it is desirable, when introducing into the theory something represented by a number, to say how that number is to be measured. In this way physical quantities are essentially defined by the operations used to measure them, and consequently this method has been called the *operational method* by BRIDGMAN [1927] [1952] (see also EINSTEIN [1949, p. 335]). The use of this method was advocated by EINSTEIN [1905] [1920] in the early days of relativity; it enabled him to challenge the validity of Newton's absolute time on the ground that no operational procedure had been, or could be, laid down for its measurement.

If we are to follow the operational method, we must not be satisfied to say that it is possible to assign space-time coordinates (x^1, x^2, x^3, x^4) to events; we must say *how* they may be assigned. The operational method does not demand a detailed description of an experiment; it is better to be simple, and it does not matter if the experiment sounds rather fantastic from a practical standpoint, provided it is correct in principle. Such experiments are called *ideal experiments.*

In this spirit, let us see how the space-time coordinates (x^1, x^2, x^3, x^4) may be assigned to an event. Suppose that the event is the explosion of a rocket in mid-air. Let there be four observers, flying about in aeroplanes, not on any particular courses, but turning and diving and climbing in an arbitrary way. Let each observer carry a clock, not necessarily an accurate clock, but perhaps an old battered clock — the one essential is that it keeps going.

Each observer notes the reading of his clock when he hears the explosion of the rocket. Let these four readings be denoted by (x^1, x^2, x^3, x^4); these four numbers may be taken as the coordinates of the event. It is clear that if there are several events (several explosions), they will in general give distinct tetrads of numbers, and in fact there is a one-to-one correspondence between possible events and possible tetrads of coordinates, with perhaps the exception of certain critical events.

It is of course essential that all four clocks keep going. If any one of them stopped, the corresponding x would have a fixed value, and the totality of number-tetrads would have only three variables, not four.

They would then be inadequate to cover the four-dimensional continuum of events.

The above operational procedure appears less fantastic and more practical if we replace the rocket by an earthquake and the four observers in aeroplanes by four seismographs. The essential point is that it is possible to give operational procedures for the assignment of coordinates (x^1, x^2, x^3, x^4) to events in space-time, and that there is an infinite variety in the ways in which this can be done. If two different procedures are used, the first giving coordinates (x^1, x^2, x^3, x^4) and the second giving coordinates (y^1, y^2, y^3, y^4) to the same event, then there will exist formulae of transformation of coordinates

$$\begin{aligned} x^1 &= x^1(y^1, y^2, y^3, y^4), \\ x^2 &= x^2(y^1, y^2, y^3, y^4), \\ x^3 &= x^3(y^1, y^2, y^3, y^4), \\ x^4 &= x^4(y^1, y^2, y^3, y^4). \end{aligned} \tag{1}$$

These may also be written the other way round, expressing the y's in terms of the x's.

§ 5. NOTATION

The following notation will be used throughout the book, unless there is special mention to the contrary:

Latin suffixes take the values 1, 2, 3, 4, with summation over this range understood for a repeated suffix.

Greek suffixes take the values 1, 2, 3, with summation over this range understood for a repeated suffix.

Thus, for example, x^r stands for the four numbers x^1, x^2, x^3, x^4; y_σ stands for y_1, y_2, y_3;

$$x^r x^r = (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2,$$

$$x_\sigma y_\sigma = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

The reader is expected to be familiar with the techniques of tensor calculus. The notation used in this book is that of SYNGE and SCHILD [1952].

When there is no risk of confusion we may use x to stand for x^r, suppressing the suffix and regarding it as understood. Thus, in this notation, the coordinates of an event are x or y, and the transformation (1) may be written briefly as

$$x^r = x^r(y) \text{ or simply } x = x(y). \tag{2}$$

§ 6. WORLD LINES AND SPACE-TIME DIAGRAMS

The history of a particle consists of a continuous sequence of events, and so this history may be expressed by four equations

$$x^r = x^r(w), \tag{3}$$

where w is a parameter. (Note that here w is a *single* quantity, whereas in (2) y stands for the four quantities y^r.) In geometrical language, the history of a particle is a curve in space-time; we shall call it a *world line.*

If we eliminate w from (3), we get three equations of the form

$$f_\sigma(x) = 0; \tag{4}$$

if we solve these equations for x^σ in terms of x^4, we get equations of the form

$$x^\sigma = g^\sigma(x^4). \tag{5}$$

In (3), (4) and (5) we have three alternative ways of expressing the equations of a world line.

A *collision* between two particles corresponds to an intersection of their world lines.

We shall deal in some detail with space-time diagrams in Ch. III, but it is useful to introduce them briefly at this point.

To show a world line graphically, we may use (5) to plot x^1, x^2 and x^3 against x^4, so obtaining the three graphs shown in Fig. 1. The full lines represent completely the history of one particle, and the broken lines the history of a second particle. A collision is shown by a heavy dot, appearing in all three diagrams; naturally x^4 has the same value at all three dots.

The purpose of such diagrams is not to replace formal calculations, but rather to concentrate the attention and give suggestions. It is seldom necessary to go to the trouble of drawing three diagrams as in Fig. 1; one diagram may suffice, and that diagram a rough sketch. We must of course remember that in a single diagram an apparent

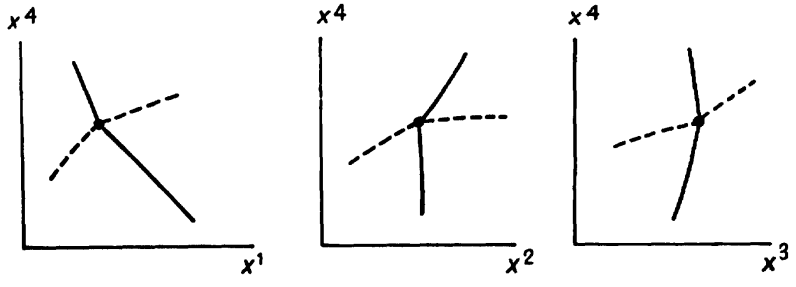


Fig. 1 – Space-time diagram for two colliding particles

intersection does not necessarily imply a collision, and so it is well to mark collisions with some sign like a heavy dot.

Everyone knows how important diagrams are in Newtonian physics — not necessarily accurate diagrams, but quite rough sketches. It is almost impossible to conceive of anyone doing physics without them. All the reader is invited to do now is to carry over into relativity the practice of making rough diagrams, with the difference that in relativity the diagrams are usually diagrams of space-time, not of space; indeed we have not yet recognised anything we can call *space*, the splitting of space-time into space and time being a matter to be considered later with some care. Thus, although (5) may look like equations expressing "position" in terms of "time", there is no reason to regard x^4 as "time"; it is just one of the four space-time coordinates which we happened to put in a privileged position.

§ 7. THE MOTION OF A MATERIAL PARTICLE

In Newtonian mechanics the equations of motion of a material particle are

$$m\frac{d^2x}{dt^2} = X, \quad m\frac{d^2y}{dt^2} = Y, \quad m\frac{d^2z}{dt^2} = Z, \tag{6}$$

where X, Y, Z are the components of the force acting on the particle. If we are dealing with a field of force, fixed or changing, the components (X, Y, Z) are given functions of (x, y, z, t), and if frictional resistance is present then (X, Y, Z) may depend also on the velocity $(dx/dt, dy/dt, dz/dt)$.

The essential point to emphasise about (6) is that they form a set of *three ordinary differential equations of the second order.* They determine a motion (a world line) if we are given initial values of $x, y, z, t, dx/dt, dy/dt, dz/dt$, i.e. if we are given an initial event (x, y, z, t) and an initial space-time direction there defined by the ratios $dx : dy : dz : dt$.

We carry over into relativity the general features of this Newtonian scheme, but not its details. We think of space-time, with coordinates x^r assigned in some way, and of a material particle with a world line in space-time. We suppose that there is some field of force, electromagnetic or gravitational or both; we shall make our ideas about such fields more precise later. What we accept is the statement that *the world line of a material particle is determined by an event x^r on it and by a space-time direction at that event, i.e. by the ratios $dx^1 : dx^2 : dx^3 : dx^4$.*

This demand is met if we assume that the world line of a particle satisfies differential equations of the form

$$\frac{d^2x^\sigma}{(dx^4)^2} = X^\sigma, \tag{7}$$

where X^σ are functions of x^1, x^2, x^3, x^4, dx^1/dx^4, dx^2/dx^4, dx^3/dx^4, or in briefer notation, functions of x^r, dx^π/dx^4. It is easy to see that if we change from coordinates x to coordinates y by a transformation (2), the above equations retain their general form, i.e. they may be written

$$\frac{d^2y^\sigma}{(dy^4)^2} = Y^\sigma, \tag{8}$$

where Y^σ are functions of y^r, dy^π/dx^4. These are not of course the same functions of y^r, dy^π/dx^4 as X^σ were of x^r, dx^π/dx^4.

If we were told the form of the functions X^σ for some particular coordinate system x, we could proceed to solve (7), obtaining theoretical world lines which could be compared with observed world lines. *Our problem is to formulate rules by which we may be able to write down explicitly equations of motion in the form* (7) *or in some equivalent form when the field of force is specified.* Only after this is done can we make contact between our mathematical scheme and actual physical phenomena. So far all we have done is to extract from Newtonian physics certain concepts for use in relativity, and we have hardly committed ourselves to any definite physical prediction which could be put to experimental test. Before adding the necessary definiteness to relativity, we must discuss the meaning of *time*.

§ 8. PAST, PRESENT AND FUTURE

In comparison with some of the "particles" with which physics deals (e.g. a planet in the solar system), a human being is very small indeed, whereas in comparison with others (e.g. an electron) he is very large. Nevertheless we shall take the liberty of reducing a human being to a moving point, and we shall call him an *observer* when so idealised. He has a world line, with equations as in (3), (4) or (5).

We take it as a matter of human experience that an observer can *order* the events which occur on his world line. This means that, given two events on his world line, he can say that one event occurs *before* the other. As far as events on his world line are concerned, he has a past, a present and a future, the present consisting of a single event

which is *after* the past and *before* the future. He knows yesterday, today and tomorrow, and never doubts that yesterday occurred before today and that today occurs before tomorrow.

That is the view we take in relativity, but note that we assert the existence of an intrinsic time-order *only for events occurring on the world line of an observer.* We do *not* assert that there is an intrinsic time-order for any arbitrary pair of events. The existence of such a time-order is asserted in Newtonian physics, for it is implicit in the association of an absolute time t with any event. Thus, *at this precise point the way of relativity departs from the way of Newton, by making a weaker assertion about the ordering of events.*

It is a serious thing to abandon Newtonian time, for the main structure of Newtonian physics must be abandoned with it, and one may well ask why we are taking this step. It may be answered that we are laying the foundation of the theory of relativity, and that the justification lies in the correct prediction of physical phenomena. But, not to lose contact completely with Newtonian ideas, let us see what the rival assumptions amount to.

Consider two events, B and A. If it is possible for an observer to move so that both these events occur on his world line, then, in view of what has been said above, their time-order is known by that observer. Newtonian theory and relativity agree with regard to that time-order saying "B occurs before A" or "B occurs after A", or "B and A are simultaneous", in which last case B and A are really the same event.

But perhaps it is not possible for an observer to include both the given events on his world line. To explore this question, we take an event O (Fig. 2) and suppose that an observer arranges to catch O on

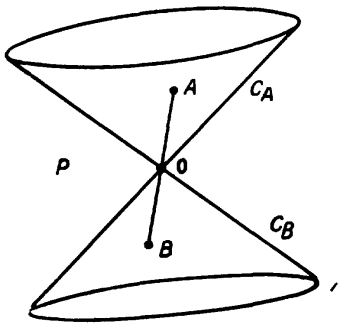


Fig. 2 – The past B, the present P and the future A in relativity

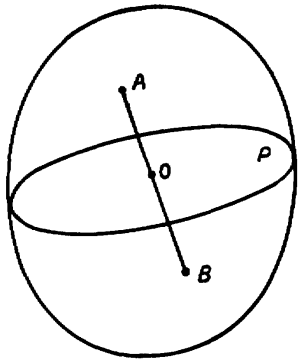


Fig. 3 – The past B, the 3-dimensional present P and the future A in Newtonian physics

his world line. This world line is then split into two parts at O, into a past B (B for Before) and a future A (A for After).

Consider now all possible observers with world lines passing through the event O. These world lines fill at least a portion of space-time adjacent to O, and each event E in this portion may be labelled B or A according as it is before or after O as judged by the observer with world line OE. This means that the whole of space-time adjacent to O may be divided into three parts consisting of

(i) events labelled B (before O) — say, the region B;
(ii) events labelled A (after O) — say, the region A;
(iii) events which cannot be labelled B or A because they cannot be reached from O by any observer's world line — say, the region P.

In addition to these 4-dimensional regions, there is a 3-dimensional "surface" dividing B from P and A from P; this is shown as the cone with sheets C_A and C_B in Fig. 2.

Thus we are led to consider a division of space-time relative to O into

the past B,

the present P.

the future A.

The cone with sheets C_B and C_A we call a *null cone.*

If we now expand the regions B and A of Fig. 2 towards one another, the region P will ultimately be squashed flat, changing from a 4-dimensional region to a 3-dimensional one. The whole neighbourhood of O will be filled by the past B and the future A, divided from one another by a 3-dimensional region, the relic of P, or, equivalently, the 3-space obtained by opening out the null cone until it becomes flat (Fig. 3).

This last is the picture given by the Newtonian concept of time. The present P is only 3-dimensional, with equation $t = \text{const.}$, where t is Newtonian time. So viewed, the Newtonian idea of past, present and future is a particular limiting case of the more general situation shown in Fig. 2, and we may therefore feel with justification that the relativistic view is more powerful than the Newtonian, since the latter is only a particular case of the former. The fact that Newtonian physics gives such a good approximation to reality in most eases is due to the fact that, in a rough manner of speaking, the present P in

Fig. 2 is actually rather thin from top to bottom, the sheets C_B and C_A being close to one another, so that the relativistic present P is not far from being a 3-dimensional Newtonian present P, as in Fig. 3.

But while relativity seems to show itself here as a more general theory than Newton's, viewed in another way it is more restrictive. In Newtonian physics the world line of an observer (equivalently, the world line of a material particle) can pass from O in any space-time direction except the directions lying in the 3-dimensional present P. This restriction amounts to the accepted fact that a man cannot be "in two places at the same time", a restriction which has been accepted without question because it does not seem to be a very serious one. But in relativity the whole of the 4-dimensional present P is forbidden; this restriction is of serious physical import, for, as we shall see later, it amounts to the assertion that no material particle can travel as fast as light.

§ 9. STANDARD CLOCKS

We accept then, as stated above, that an observer is able to order events which occur on his world line. He has, in fact, an intuitive sense of time-order which enables him to say "A is before B is before C is before D", provided the four events occur on his world line. But this intuitive sense is not quantitative; it does not enable the observer to assert that $AB = CD$, meaning thereby that two time-intervals are equal, although in the above case he might say $BC < AD$, since the former interval is contained inside the latter.

To measure time one must use a *clock*, a mechanism of some sort in which a certain process is repeated over and over again under the same conditions, as far as possible. The mechanism may be a pendulum, a balance wheel with a spring, an electric circuit, or some other oscillating system, and out of these one passes by idealisation to the concept of a *standard clock*, bearing to a real clock much the same relationship as a geometrical right-angled triangle bears to the set-square of a draughtsman.

Let us however make the concept of a standard clock more definite by thinking of it as an atom of some specified element emitting a certain specified spectral line, the "ticks" of the clock corresponding to the emission of the crests of successive waves of radiation from the atom. By associating numbers 1, 2, 3,... with these ticks, we have a scale for assigning times to events occurring in the history of the atom,

or (since that would give us a very small unit of time) we may more conveniently associate with the ticks the numbers $a, 2a, 3a, \ldots$, where a is some chosen small number, to be used universally for all clocks.

To base the measurement of time, as above, on a standard atomic frequency is very like the plan of basing the measurement of length on a standard wave length. But in relativity the concept of length is not an easy one, and it seems best to start with an atomic clock as the basic concept and introduce the idea of wave length at a later stage.

§ 10. THE SEPARATION BETWEEN EVENTS

Consider an observer carrying a standard clock along his world line from an event B to an event A, B being before A. Let the clock read s_B at B and s_A at A. Then we call $s = s_A - s_B$ the *separation* between the events B and A, measured along the chosen world line.

If the event A is taken close to B, the separation becomes infinitesimal and we denote it by ds. Then ds is a function of the coordinates of B and A, say x and $x + dx$, and we may write

$$ds = f(x, dx). \tag{9}$$

Now if the infinitesimals dx are all multiplied by the same positive factor k, ds will be multiplied by that factor, and so

$$f(x, kdx) = kf(x, dx) \text{ for } k > 0. \tag{10}$$

Hence we say that $f(x, dx)$ is *positive homogenous of degree unity* in the differentials.

If we write the equations of the world line in the form

$$x^r = x^r(w), \tag{11}$$

where w is a parameter increasing from past to future, then

$$dx^r = x^{r\prime}\, dw \tag{12}$$

where $x^{r\prime} = dx^r/dw$, and so

$$ds = f(x, dx) = f\left(x, \frac{dx}{dw}\, dw\right) = f(x, x')dw. \tag{13}$$

Thus the separation between finitely separated events, say B with $w = w_1$ and A with $w = w_2$, is given by

$$s = \int_B^A f(x, dx) = \int_{w_1}^{w_2} f(x, x')dw, \tag{14}$$

where $f(x, x')$ is positive homogenous of degree unity in the derivatives.

Now s is a physically measurable quantity, and so its value is independent of the particular system of coordinates x employed. Hence the function $f(x, dx)$ must be an *invariant* in the sense of tensor calculus when we transform from coordinates x to coordinates y. This does *not* mean that the function appropriate to y is $f(y, dy)$, the same form of function as before. It means that if $\bar{f}(y, dy)$ is the function appropriate to y, then

$$\bar{f}(y, dy) = f(x, dx) \tag{15}$$

for arbitrary values of x and dx, the values of y and dy being given by

$$y^r = y^r(x), \quad dy^r = \frac{\partial y^r}{\partial x^s} dx^s. \tag{16}$$

§ 11. THE FUNDAMENTAL QUADRATIC FORM

To build a physical theory which can be tested by experiment, we must specify the function $f(x, dx)$ for some coordinate system x. Here various paths open out, one leading back to Newtonian physics, one leading on to Einstein's theory of relativity, and others leading to more general theories.

Consider first the assumption

$$ds = f(x, dx) = \frac{\partial t}{\partial x^r} dx^r, \tag{17}$$

where t is some invariant function of the coordinates. Then (14) integrates at once to give

$$s = t_A - t_B; \tag{18}$$

this means that the separation between two events depends only on the events and not on the world line joining them. This choice of $f(x, dx)$ leads us right back to Newtonian physics, t being the absolute time of Newton[1].

Consider next the choice

$$ds = f(x, dx) = (-g_{mn} dx^m dx^n)^{\frac{1}{2}}, \tag{19}$$

the square root of a quadratic form. (The minus sign is of no particular importance — a mere matter of notation.) Here g_{mn} are functions of the

[1] The dependence of the integral (14) on the path of integration in space-time is sometimes called the "clock paradox" (cf. MØLLER [1952, pp. 49, 258]). But it presents no paradox here. Indeed, the fact that ds is *not* an exact differential constitutes the essential difference between relativity and Newtonian physics.

coordinates; there are ten of them, since without loss of generality we can assume $g_{nm} = g_{mn}$; they are the components of a covariant tensor, since ds is invariant. This is the choice of $f(x, dx)$ appropriate to Einstein's relativity, provided the quadratic form satisfies a certain algebraic requirement, which we shall now discuss.

We suppose that the form is non-singular, i.e. we suppose

$$\det (g_{mn}) \neq 0.$$

It is known[1] that we can always choose coordinates so that at some specified event a non-singular quadratic form reduces to

$$g_{mn}dx^m dx^n = \varepsilon_1(dx^1)^2 + \varepsilon_2(dx^2)^2 + \varepsilon_3(dx^3)^2 + \varepsilon_4(dx^4)^2, \tag{20}$$

where the ε's stand for $+1$ or -1. The numbers of $+$ and $-$ signs occurring here are characteristic of the form and cannot be changed by changing the coordinates. (It is understood that only *real* coordinates are considered here; we can of course change the signs by using imaginary coordinates, and we shall make extensive use of these later, but only as notational device).

In view of what has been said earlier about past, present and future, it is possible to decide on the signs which should occur in (20). To see this, we note that, since coordinates may be interchanged, there are only five choices of signs to be considered:

$$\begin{array}{ll} a) & + + + + \\ b) & + + + - \\ c) & + + - - \\ d) & + - - - \\ e) & - - - - \end{array} \tag{21}$$

Since ds, as in (19), is a real measured quantity, those space-time directions in which a clock can move must satisfy

$$g_{mn}dx^m dx^n < 0. \tag{22}$$

We have seen earlier (cf. Fig. 2) that the region of space-time adjacent to an event can be broken up into a past B, a present P, and a future A. The past and future contain the allowed directions for world lines, and since these allowed direction are to satisfy (22), we may write

$$\begin{array}{ll} \text{past and future:} & g_{mn}dx^m dx^n < 0, \\ \text{present:} & g_{mn}dx^m dx^n > 0, \\ \text{null cone:} & g_{mn}dx^m dx^n = 0. \end{array} \tag{23}$$

[1] Cf. SYNGE and SCHILD [1952 pp. 58, 313].

Past, present and future must all exist, even though the present should consist of only a 3-space as in Fig. 3 (Newtonian case). The past must be divided from the future by the present. A denial of these facts would be a denial of our most primitive intuitions about time-order.

We must therefore rule out (21*a*) (it gives no past and no future) and also (21*e*) (it gives no present). As for (21*c*), it gives

$$\begin{aligned} &\text{past and future: } (dx^1)^2 + (dx^2)^2 - (dx^3)^2 - (dx^4)^2 < 0, \\ &\text{present: } \qquad (dx^1)^2 + (dx^2)^2 - (dx^3)^2 - (dx^4)^2 > 0. \end{aligned} \tag{24}$$

Here the present is 4-dimensional; but it is not hard to see that it fails to divide the region "past and future" into a past and a future, for it is possible to pass continuously from any event in "past and future" to any other without encountering the present. Thus (21*c*) must be ruled out, and so must (21*d*) for similar reasons. Thus we are left with (21*b*) as the only possible choice:

$$g_{mn}dx^m dx^n = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2. \tag{25}$$

There are three + and one − sign. We have

$$\begin{aligned} &\text{past and future: } (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 < 0, \\ &\text{present: } \qquad (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 > 0, \\ &\text{null cone: } \qquad (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 = 0. \end{aligned} \tag{26}$$

These regions are as shown in Fig. 2, the past being distinguished from the future by the condition that dx^4 is positive for one and negative for the other, but which is which depends on the way in which x^4 is chosen. We shall choose it so that x^4 increases ($dx^4 > 0$) as we pass into the future from O.

At the chosen event and for the special coordinates used, we have then

$$ds = f(x, dx) = [-(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2]^{\frac{1}{2}}. \tag{27}$$

In general this reduction of (19) can be made for only one chosen event, although that event may be chosen arbitrarily. If ds has the simple form (27) at some event E for some coordinate system x, then at other events this simplicity is lost and can be obtained at a second event E' only by changing the coordinate system and sacrificing the simplicity at E. That, at least, is the general situation; but, as we shall see later, the simple form (27) may be obtained for one coordinate

system throughout all space-time provided gravitational fields are absent.

Vectors drawn into the past or future are called *timelike*; they are possible tangents to the world lines of standard clocks. Vectors drawn into the present are called *spacelike*; and those on the null cone are called *null vectors*. Thus vectors are classified as follows:

$$\begin{aligned} &\text{timelike:} && g_{mn}V^mV^n < 0, \\ &\text{spacelike:} && g_{mn}V^mV^n > 0, \\ &\text{null:} && g_{mn}V^mV^n = 0. \end{aligned} \tag{28}$$

The quadratic form $g_{mn}dx^m dx^n$ is basic in relativity. It is, so to speak, the core of the subject. No matter what chain of thought is used to introduce it (and different people prefer different chains of thought), the quadratic form is a goal to which all such chains must lead.

§ 12. FINSLER SPACE-TIME AND HAMILTONIAN METHODS

The idea of measuring time by standard clocks led us to think of an invariant function $f(x, dx)$, positive homogeneous of degree unity in the differentials. In the language of geometry this makes space-time a *Finsler space* with line element $ds = f(x, dx)$. If we choose f as in (19), we get a Riemannian space, which is a particular case of a Finsler space, and that is the choice we have to make if we are to develope Einstein's theory of relativity. But if the sole demand is that space-time should be Finslerian, we might choose, for example,

$$ds = (g_{mnrs}dx^m dx^n dx^r dx^s)^{\frac{1}{4}}, \tag{29}$$

where g_{mnrs} is a tensor as indicated, or we might define (cf. STEPHENSON and KILMISTER [1953]) f as a root of the quadratic equation

$$f^2 - 2fA_r\,dx^r + g_{mn}dx^m dx^n = 0, \tag{30}$$

where A_r is a covariant vector and g_{mn} a covariant tensor, both functions of x; if we put $A_r = 0$ we recover (19), and, if A_r is small, we have approximately

$$f(x, dx) = A_r\,dx^r \pm (-g_{mn}dx^m dx^n)^{\frac{1}{2}}. \tag{31}$$

However the interest of Finsler space-time lies not so much in making special choices of f as in a general connection which exists between Finsler space and Hamiltonian methods, so important in modern

physics. This connection exists in particular for Riemannian space-time, but the theory is as easy to follow in the general case, and although we shall not use these ideas later in the book, we shall make a brief excursion here to examine them.

In Finsler space-time with $ds = f(x, dx)$ the curves satisfying the variational principle $\delta \int ds = 0$ or

$$\delta \int f(x, dx) = 0 \tag{32}$$

are of particular interest; they are the *geodesics*. Introducing a parameter u and putting $dx^r/du = x^{r'}$, we may write (32) as

$$\delta \int f(x, x')du = 0 \tag{33}$$

on account of the homogeneity of f, and we are at once led to the *Lagrangian equations* for the geodesics

$$\frac{d}{du}\frac{\partial f}{\partial x^{r'}} - \frac{\partial f}{\partial x^r} = 0. \tag{34}$$

Let us now define σ_r by

$$\sigma_r = -\frac{\partial f}{\partial x^{r'}}; \tag{35}$$

it is easily seen to be a covariant vector. These derivatives are homogeneous of degree zero in x', and so the three ratios $x^{1'} : x^{2'} : x^{3'} : x^{4'}$ can be eliminated, giving an equation of the form

$$\Omega(\sigma, x) = 0, \tag{36}$$

Ω being an invariant. By (35) and the homogeneity of f we have, on any curve $x^r = x^r(u)$,

$$\sigma_r dx^r = -\frac{\partial f}{\partial x^{r'}} x^{r'}\, du = -f(x, x')du = -f(x, dx), \tag{37}$$

and so the variational equation (33) is equivalent to the variational equation plus side condition

$$\delta \int \sigma_r dx^r = 0, \quad \Omega(\sigma, x) = 0. \tag{38}$$

Hence we have

$$0 = \int (\delta\sigma_r dx^r + \sigma_r \delta dx^r) = \int (\delta\sigma_r dx^r - \delta x^r d\sigma_r) \tag{39}$$

for all variations satisfying

$$\frac{\partial \Omega}{\partial \sigma_r}\delta\sigma_r + \frac{\partial \Omega}{\partial x^r}\delta x^r = 0, \tag{40}$$

and so we are led to the equations

$$\frac{dx^r}{du} = \lambda \frac{\partial \Omega}{\partial \sigma_r}, \qquad \frac{d\sigma_r}{du} = -\lambda \frac{\partial \Omega}{\partial x^r}, \tag{41}$$

where λ is a Lagrange multiplier. Defining a new parameter w by $dw = \lambda\, du$, we get

$$\frac{dx^r}{dw} = \frac{\partial \Omega}{\partial \sigma_r}, \qquad \frac{d\sigma_r}{dw} = -\frac{\partial \Omega}{\partial x^r}. \tag{42}$$

These *Hamiltonian equations* are equivalent to the Lagrangian equations (34).

Since the equations (42) can be written out explicitly if $\Omega(\sigma, x)$ is given, it is clear that, just as a function $f(x, dx)$ establishes a Finsler geometry, so does the equation (36) establish a Hamiltonian geometry, and these are merely different ways of looking at the same geometry if f and Ω are related as above. We pass from Finsler to Hamilton as described, and the reverse passage is effected by writing down the five equations

$$\frac{x^{r'}}{f} = -\frac{\dfrac{\partial \Omega}{\partial \sigma_r}}{\sigma_n \dfrac{\partial \Omega}{\partial \sigma_n}}, \qquad \Omega(\sigma, x) = 0, \tag{43}$$

and eliminating the four σ's from them, thus obtaining f as a function of x and x' (for a proof, see SYNGE [1954*a*, p. 139]).

To illustrate the passage from f to Ω, let us take the Riemannian case

$$f(x, x') = (-g_{mn}x^{m'}x^{n'})^{\frac{1}{2}}. \tag{44}$$

Then by (35)

$$\sigma_r = g_{rp}x^{p'}(-g_{mn}x^{m'}x^{n'})^{-\frac{1}{2}}, \tag{45}$$

and so, if g^{rs} is defined by $g^{rm}g_{rn} = \delta^m_n$,

$$g^{rs}\sigma_r\sigma_s = -1. \tag{46}$$

Thus we obtain an equation as in (36)

$$\Omega(\sigma, x) = \tfrac{1}{2}(g^{rs}\sigma_r\sigma_s + 1) = 0, \tag{47}$$

the factor $\frac{1}{2}$ being of course of no importance.

To illustrate the converse procedure, let us start from a slightly more general form of Ω:

$$\Omega(\sigma, x) = \tfrac{1}{2}[g^{rs}(\sigma_r + A_r)(\sigma_s + A_s) + 1] = 0, \tag{48}$$

g^{rs} and A_r being functions of x. Then (43) gives

$$\frac{x^{r'}}{f} = -\frac{g^{rs}(\sigma_s + A_s)}{\sigma_n g^{mn}(\sigma_m + A_m)}. \tag{49}$$

Define V^r by

$$V^r = g^{rs}(\sigma_s + A_s), \tag{50}$$

so that by (48)

$$g_{mn}V^mV^n = -1. \tag{51}$$

Then (49) gives

$$fV^r = x^{r'}(A_nV^n + 1), \tag{52}$$

and hence

$$A_nV^n(f - A_rx^{r'}) = A_r x^{r'}, \ (A_nV^n + 1)\,(f - A_r x^{r'}) = f,$$

so that (52) gives, on multiplication by $(f - A_m x^{m'})$ and division by f,

$$V^r(f - A_m x^{m'}) = x^{r'}. \tag{53}$$

Then by (51)

$$(f - A_m x^{m'})^2 = -g_{rs}x^{r'}x^{s'}, \tag{54}$$

and thus we obtain

$$f(x, x') = A_m x^{m'} \pm (-g_{rs}x^{r'}x^{s'})^{\frac{1}{2}}, \tag{55}$$

which is actually the same as (31).

§ 13. SPACE-TIME AS A RIEMANNIAN SPACE

From this excursion into Finsler space-time, we return to the Einstein expression for separation (19):

$$ds = (-g_{mn}dx^m dx^n)^{\frac{1}{2}}. \tag{56}$$

Space-time possesses, then, an invariant quadratic form

$$\Phi = g_{mn}dx^m dx^n, \tag{57}$$

and so we may say that space-time is a *Riemannian 4-space* V_4 in which Φ defines the *metric*.

In saying this, we at once make available in relativity the great body of mathematical knowledge contained in the subject of Riemannian geometry. But here a warning is necessary, particularly in re-

lation to the words *distance* and *length*. To the pure mathematician working in Riemannian geometry, the "distance" between two points in V_4 and the "length" of a curve in V_4 are merely integrals ($\int ds$). He is not troubled by the fact that these words have already got well-established physical meanings. But, to the physicist, distance and length are quantities to be measured by means of rigid measuring rods or by chains consisting of rigid links. We shall presently discuss the possibility of such measurements in relativity, but we have not, merely by calling space-time a Riemannian space, set up a scheme by which distance or length can be measured. To avoid confusion on this score, we shall continue to refer to the ds of (56) as *separation*, and avoid the words *distance* and *length* until they are properly introduced by means of ideal experiments.

We can assign a real separation ds to any pair of adjacent events, x and $x + dx$, whether the vector dx is timelike, spacelike or null, by writing

$$ds^2 = \varepsilon\, g_{mn} dx^m dx^n; \tag{58}$$

here ε (called the *indicator* of dx) is chosen equal to $+1$ or -1 so as to make ds^2 positive (unless it is zero). Thus we have the scheme:

$$\begin{aligned} &\text{timelike:} && \varepsilon = -1;\ ds^2 = -g_{mn} dx^m dx^n; \\ &\text{spacelike:} && \varepsilon = +1;\ ds^2 = g_{mn} dx^m dx^n; \\ &\text{null:} && ds = 0,\quad g_{mn} dx^m dx^n = 0. \end{aligned} \tag{59}$$

We recall that a timelike separation is measured by a standard clock. By carrying out measurements of ds and dx for ten different world lines through an event, we can calculate the ten g_{mn}. Then we can substitute these values in (58), and so, taking $\varepsilon = +1$, find ds for any spacelike dx. But we shall describe in the next section a more direct ideal experiment for the measurement of a spacelike separation.

Returning to Fig. 2, we recall that space-time directions which correspond to the motions of material particles lie in the conical regions B and A (past and future). The present P is a forbidden region as far as the world lines of material particles are concerned, and the question arises as to the physical meaning of those directions which lie on the null cone dividing the present from the past and future. We make the hypothesis that *the directions of the world lines of photons lie on the null cone*, so that an infinitesimal vector lying along the history of a photon is a null vector, satisfying

$$ds = 0,\ g_{mn} dx^m dx^n = 0. \tag{60}$$

We note that this hypothesis makes a photon in a sense the limiting case of a material particle.

§ 14. MEASUREMENT OF SPACELIKE SEPARATION

The substitution of geometrical language for the language of physics (i.e. language related to experiments, real or ideal) is called the *geometrisation* of physics. From the time Minkowski[1] proposed that space-time be viewed as a four-dimensional space, geometrisation has played an essential part in the theory of relativity. But there is also the converse process, which we may call the *physicisation* of geometry. Having set up a geometrical picture of four-dimensional space-time, geometry suggests to us various relationships and constructions, and these must be translated into the language of physics in order that we, as physicists, may see what they mean.

The active interplay of geometrical and physical ideas plays a fundamental role in relativity, and one soon developes an automatic facility in passing from one point of view to the other. Here, partly as an exercise and partly for its own sake, we shall physicise the geometrical element ds of separation between two adjacent events, P and Q.

First we find out by trial whether it is possible to make a material particle contain these two events in its history. If it is possible, then ds is measured by the difference in the readings at P and Q of a standard clock carried by the particle which includes them in its history.

If it is impossible for a material particle to include both the events in its history, we know that PQ either lies on the null cone or is spacelike. We test whether it lies on the null cone by emitting photons from P in all the space-time directions permitted to photons. If one of these photons includes Q in its history, then PQ is null and we have $ds = 0$.

Suppose now that neither material particle nor photon can include both P and Q in its history. Then the space-time displacement PQ is spacelike; we proceed to describe a simple ideal experiment for the measurement of ds in this case.

Fig. 4 shows the events $P(x)$ and $Q(x + dx)$; PQ is spacelike and so

$$ds = (g_{mn}dx^m dx^n)^{\frac{1}{2}}. \tag{61}$$

This is the quantity we have to measure. Consider a material particle which includes P in its history, i.e. the world line L of the particle passes through P. If we draw the null cone with vertex Q, it will cut L

[1] MINKOWSKI [1909], translated in LORENTZ [1923, p. 75].

in two events, B (before P) and A (after P). The construction of B and A is realised physically by letting the particle carry a lamp which emits photons in all possible ways. One of these photons encounters the event Q, and B is fixed as the event on L corresponding to its emission. If properly reflected at Q, this photon will return to the particle, and A is the event corresponding to its arrival at the particle. (This simple operation takes place when you look at your own eye in a mirror, L being the world line of your eye and Q an event in the history of the mirror.)

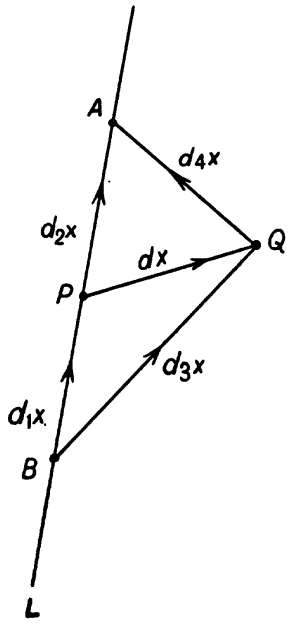


Fig. 4 – Measurement of a spacelike separation

We now have the diagram shown in Fig. 4, consisting of two infinitesimal triangles in space-time. We shall use the following notation for infinitesimal vectors:

$$PQ: dx^r, \qquad BP: d_1x^r, \qquad PA: d_2x^r, \qquad BQ: d_3x^r, \qquad QA: d_4x^r. \tag{62}$$

We have then

$$d_1x^r + dx^r = d_3x^r, \; d_2x^r = dx^r + d_4x^r. \tag{63}$$

But d_3 and d_4 are null, and so

$$\begin{aligned} 0 &= g_{mn}d_3x^m d_3x^n = g_{mn}d_1x^m d_1x^n + 2g_{mn}dx^m d_1x^n + g_{mn}dx^m dx^n, \\ 0 &= g_{mn}d_4x^m d_4x^n = g_{mn}d_2x^m d_2x^n - 2g_{mn}dx^m d_2x^n + g_{mn}dx^m dx^n. \end{aligned} \tag{64}$$

Let us write d_1s and d_2s for the separations BP and PA respectively. These are measured by a standard clock carried by the particle. Since they are timelike separations, we have

$$d_1s^2 = -g_{mn}d_1x^m d_1x^n, \; d_2s^2 = -g_{mn}d_2x^m d_2x^n, \tag{65}$$

and so (64) gives

$$\begin{aligned} -d_1s^2 + 2g_{mn}dx^m d_1x^n + ds^2 &= 0, \\ -d_2s^2 - 2g_{mn}dx^m d_2x^n + ds^2 &= 0. \end{aligned} \tag{66}$$

The vectors d_1x^r, d_2x^r point in the same direction in space-time, both being infinitesimal vectors tangent to L at P. Hence a positive scalar k exists such that

$$d_2x^r = kd_1x^r, \tag{67}$$

圖書館

and so

$$d_2s^2 = k^2 d_1s^2. \tag{68}$$

If we now multiply the first of (66) by k and add it to the second, the middle terms disappear and we get

$$-(k + k^2)d_1s^2 + (k + 1)ds^2 = 0. \tag{69}$$

Since $k + 1$ is certainly not zero, this gives

$$ds^2 = kd_1s^2 = d_1sd_2s. \tag{70}$$

This simple formula expresses the spacelike separation ds in terms of the timelike separations d_1s and d_2s. The latter may be measured by a standard clock, and so we have a way of measuring a spacelike separation, using as apparatus only a standard clock and photons.

§ 15. THE PHYSICAL MEANING OF ORTHOGONALITY

Two vectors, U^r and V^r, are *orthogonal* if their *scalar product* $g_{mn}U^mV^n$ vanishes:

$$g_{mn}U^mV^n = 0. \tag{71}$$

Thus it is meaningful to say that two infinitesimal displacements in space-time, d_1x^r and d_2x^r, are (or are not) orthogonal, the condition of orthogonality being

$$g_{mn}d_1x^m d_2x^n = 0. \tag{72}$$

The concept of orthogonality is so basic geometrically that we naturally ask by what physical tests we can determine whether two infinitesimal displacements are orthogonal. Of course, we could calculate the g_{mn} and make direct trial by substituting in (72), but that is a clumsy way to go about it.

First, it is unnecessary to test a pair of timelike vectors, for *a pair of timelike vectors can never be orthogonal.* We shall now prove this by showing that

$$g_{mn}U^mU^n < 0, \; g_{mn}V^mV^n < 0 \tag{73}$$

imply

$$g_{mn}U^mV^n \neq 0. \tag{74}$$

Since we are dealing with vectors at a certain event, we can use those special coordinates for which at the event we have, as in (25),

$$g_{mn}dx^m dx^n = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2. \tag{75}$$

Then the matrix g_{mn} has only diagonal terms and they are (1, 1, 1, −1). Hence, with Greek suffixes ranging 1, 2, 3, (73) reads

$$U^\varrho U^\varrho - (U^4)^2 < 0, \quad V^\sigma V^\sigma - (V^4)^2 < 0, \tag{76}$$

and so

$$| U^4 V^4 | > (U^\varrho U^\varrho V^\sigma V^\sigma)^{\frac{1}{2}}. \tag{77}$$

Now, for any real X,

$$(U^\varrho + XV^\varrho)\,(U^\varrho + XV^\varrho) \geqslant 0, \tag{78}$$

since this expression is the sum of three squares, and so the quadratic equation

$$X^2 V^\varrho V^\varrho + 2XU^\varrho V^\varrho + U^\varrho U^\varrho = 0 \tag{79}$$

has not got distinct real roots. Hence we have the Schwarz inequality

$$| U^\varrho V^\varrho | \leqslant (U^\varrho U^\varrho V^\sigma V^\sigma)^{\frac{1}{2}}. \tag{80}$$

But

$$g_{mn} U^m V^n = U^\varrho V^\varrho - U^4 V^4 \tag{81}$$

and it is clear from (77) and (80) that this cannot vanish. Hence the result is proved: two timelike vectors cannot be orthogonal.

We can go further if both the timelike vectors point into the future ($U^4 > 0$, $V^4 > 0$) or both point into the past ($U^4 < 0$, $V^4 < 0$). For then

$$U^\varrho V^\varrho \leqslant (U^\varrho U^\varrho V^\sigma V^\sigma)^{\frac{1}{2}} < U^4 V^4,$$

and therefore

$$g_{mn} U^m V^n < 0. \tag{82}$$

We shall return to this in II-(43) and VI-(51), obtaining a stronger inequality.

We have now to consider the orthogonality of (*a*) a pair of spacelike vectors, and (*b*) a spacelike vector and a timelike vector.

Consider two infinitesimal displacements AP $(d_1 x^r)$ and $AQ(d_2 x^r)$, both spacelike (Fig. 5). Then the displacement QP is

$$d_3 x^r = d_1 x^r - d_2 x^r, \tag{83}$$

and so

$$g_{mn} d_3 x^m d_3 x^n = g_{mn} d_1 x^m d_1 x^n - 2 g_{mn} d_1 x^m d_2 x^n + g_{mn} d_2 x^m d_2 x^n. \tag{84}$$

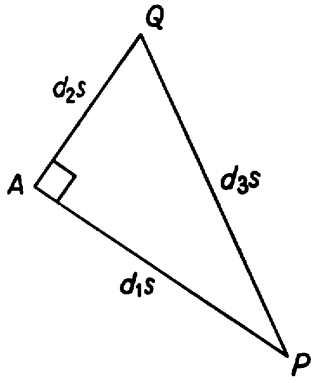


Fig. 5 – Orthogonal infinitesimal displacements, both spacelike $(d_3 s^2 = d_1 s^2 + d_2 s^2)$

If we denote the separations AP, AQ, PQ by d_1s, d_2s, d_3s, respectively, we have then

$$d_1s^2 + d_2s^2 - \varepsilon_3 d_3s^2 = 2g_{mn}d_1x^m d_2x^n, \tag{85}$$

where ε_3 is the indicator of PQ. *The condition of orthogonality of the spacelike displacements* d_1x^r, d_2x^r *is therefore that* PQ *should be spacelike* ($\varepsilon_3 = 1$) *and*

$$d_3s^2 = d_1s^2 + d_2s^2, \tag{86}$$

which is in fact the theorem of Pythagoras. Since, as we have seen above, spacelike separations can be measured physically, we have here a physical test for the orthogonality of spacelike displacements in space-time.

We pass now to the case where AP is spacelike and AQ timelike. We do not test these vectors for orthogonality directly. We note that the orthogonality condition (71) is homogeneous in the components of each of the two vectors, and thus, if U^r and V^r are orthogonal, so also are kU^r, V^r, where k is any scalar. We start then by adjusting one of the vectors (without changing its direction) so as to make the separation AP (spacelike) equal to the separation AQ (timelike); let the common value be ds (Fig. 6). This change is physically meaningful, because we know how to measure separations. When this has been done, we emit photons from the event P. Suppose that the event Q occurs in the history of one of these photons. Then, if AP is d_1x^r and AQ is d_2x^r, the vector $d_1x^r - d_2x^r$ is null, and so

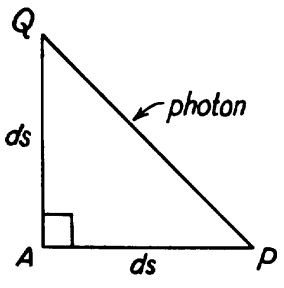


Fig. 6 – Orthogonal infinitesimal displacements, one spacelike and the other timelike

$$\begin{aligned} 0 &= g_{mn}d_1x^m d_1x^n - 2g_{mn}d_1x^m d_2x^n + g_{mn}d_2x^m d_2x^n \\ &= ds^2 - 2g_{mn}d_1x^m d_2x^n - ds^2, \end{aligned} \tag{87}$$

since AP is spacelike and AQ timelike. Hence

$$g_{mn}d_1x^m d_2x^n = 0, \tag{88}$$

and so AP and AQ are orthogonal. *Thus we know that these displacements are orthogonal if a photon from* P *passes through* Q. This is a physical test for the orthogonality of a spacelike displacement and a timelike displacement.

The orthogonality of null vectors with other vectors can easily be

discussed, and it can be shown that (*a*) a null vector is orthogonal to itself (this is obvious) and (*b*) that a null vector cannot be orthogonal to a timelike vector.

In connection with these results, it is a great help to think of a cone in ordinary space to represent the null cone and conjugacy of spatial directions with respect to this cone to represent orthogonality. In most cases this intuitive approach leads to correct results, but of course conclusions reached in this way should be checked by calculation.

Note that although we have discussed orthogonality, and hence *right angles* in space-time, we have not attempted to define angles in general. It is not worth while to do so, because we run into complications, and it is best to think of acute and obtuse angles only between spacelike vectors which are all orthogonal to a single timelike vector. Then "angle" has its ordinary meaning. In other cases we encounter "pseudoangles" [cf. IV–(39)].

§ 16. DISTANCE BETWEEN PARTICLES

In Newtonian physics the distance between two particles at any time t is a definite thing. We cannot take over this concept of distance into relativity because it requires the acceptance of an absolute time t at which to observe simultaneous positions of the particles. In relativity "two particles" means "two world lines", and if we are to speak of the "distance" between two particles, we must define the word in terms of the geometry of the world lines.

Consider, then, two timelike world lines, L and M, which are close together and so form, as it were, the edges of a thin ribbon in space-time (Fig. 7). Let the equations of L be

$$x^r = x^r(s), \tag{89}$$

where ds is the element of separation on L. The unit tangent vector to L is $\lambda^r = dx^r/ds$, and it satisfies

$$g_{mn}\lambda^m\lambda^n = -1. \tag{90}$$

We generate M from L by assigning to each event P on L an infinitesimal vector η^r, a function of s, the event P' on M having the coordinates $x^r + \eta^r$.

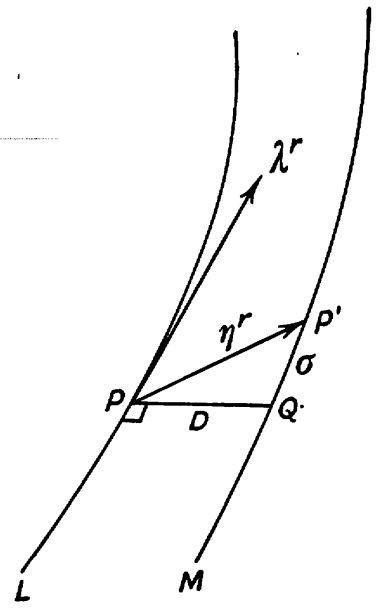


Fig. 7 – Distance D between adjacent particles, or world lines

We wish to draw a vector across from P to M, making this vector orthogonal to L at P. To do this, we travel back from P' along M through a separation σ, thus arriving at an event Q with coordinates $x^r + \eta^r - \sigma\lambda^r$, since to the first order of infinitesimals λ^r is the unit tangent vector to M at P'. Then PQ is orthogonal to L if

$$g_{mn}(\eta^m - \sigma\lambda^m)\lambda^n = 0, \tag{91}$$

and this is satisfied if we choose

$$\sigma = -g_{mn}\eta^m\lambda^n. \tag{92}$$

Thus we get a unique event Q on M such that PQ is orthogonal to L at P. We recall that such orthogonality can be tested physically (cf. Fig. 6).

Writing D for the spacelike separation PQ, we call D *the distance of M from L, measured at the event P*. We have

$$\begin{aligned} D^2 &= g_{mn}(\eta^m - \sigma\lambda^m)\,(\eta^n - \sigma\lambda^n) \\ &= g_{mn}\eta^m\eta^n - 2\sigma g_{mn}\lambda^m\eta^n - \sigma^2 \\ &= g_{mn}\eta^m\eta^n + (g_{mn}\lambda^m\eta^n)^2. \end{aligned} \tag{93}$$

We note that D is not in general a constant; it changes as we move P along L.

To the order of infinitesimals considered, PQ is orthogonal to M at Q, and so PQ is also the distance of L from M, measured at Q. Thus the concept of *the distance between adjacent particles, or world lines,* has been given an operational meaning; for infinitesimal distances it is not necessary to state from which particle the measurement is made (but see II–§ 4 for finite distances).

The word *length* may be used instead of *distance,* if, for example, we think of the two particles as the ends of a rod.

§ 17. RIGID RODS

Consider two adjacent world lines, L and M in Fig. 7. If the distance D is constant, i.e. independent of the choice of P on L, we say that the particles are *rigidly connected,* or form the ends of a *rigid rod* or rod of constant length.

In developing the theory of relativity, we seek to fit into the relativistic framework all those familiar and elementary operations in physics about which we are accustomed to think in the manner approved by Newtonian physics. It is therefore instructive to examine, in terms of

a space-time diagram and the above definition of rigid connection, what we do when we compare by measurement the lengths of two lines drawn on a sheet of paper — say a line drawn from a point A to a point B and a second line drawn from a point C to a point D. These four points are actually pencil marks, and we think of them as particles with world lines which we may, without confusion, call A, B, C, D (Fig. 8). We think of A as being close to B (so that we may use our differential expressions) and of C as being close to D. It is not necessary that the pair A, B should be close to the pair C, D.

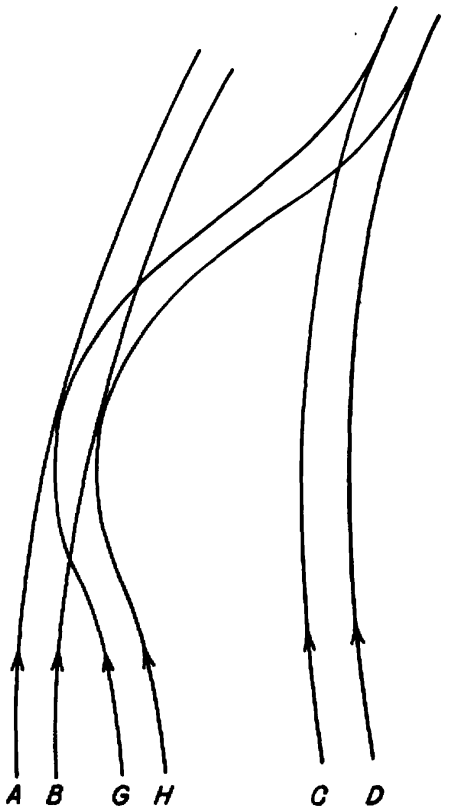


Fig. 8 – Comparison of the lengths of two lines by means of a rigid measuring rod. The arrows indicate the passage from past to future for each of the six particles involved

Unless A and B are rigidly connected, there is no question of measuring *the* length AB, for that length is changing all the time. Let us suppose then that A is rigidly connected to B, and C rigidly connected to D. Let there be a rigid measuring rod situated somewhere or other, the world lines of its ends being the lower parts of the world lines G and H in Fig. 8.

To compare the lengths AB and CD, we move the measuring rod, bending its world lines so that G runs along A. We try to make H run along B, but in general that is impossible ($AB \neq GH$). But if we think of the measuring rod, not as just two particles but as a continuum of particles, each with a world line, then one of these world lines may be made to run along B while G runs along A. If the measuring rod is graduated, we get in this way a measurement of the length of AB.

Let us suppose for simplicity that actually $AB = GH$, so that we can make G run along A and H along B. Then we take up the measuring rod and try to bring the world lines G and H into coincidence with C and D. If we can, then $CD = GH = AB$. If we cannot, we get in this way a measurement of CD from the graduations on the measuring rod.

The problem of measuring lengths is thus a problem of bringing pairs

of world lines into coincidence, so that the space-time diagram looks like a drawing of connecting railway tracks, the tracks being each of constant gauge (condition of rigid connection) in the technical sense described above; we must not expect the gauge to look constant in a space-time diagram.

In such measurements the infinitesimal rigid rod plays a fundamental part. Do such rods exist in nature? In one sense, we can say at once: Certainly not! None of the sharpened idealisations of theoretical physics is to be thought of as actually existing — they are like the point of Euclidean geometry. But in the world of these sharpened concepts, the idea of a rigid rod is useful and admissible.

It is *useful* because we expect that under suitable circumstances (absence of stress and equality of temperature) a bar of steel will act as if it were rigid to a high degree of approximation, recognising however that it needs a lot of intuitive physical sense to know when to expect actual things to behave like the idealised models we make of them. The concept of an infinitesimal rigid rod is *admissible* because it involves no contradiction in the relativistic scheme.

For a discussion of rigid motions, see § 20.

§ 18. THE WORLD LINES OF FREE PARTICLES

We described earlier a fantastic way in which coordinates might be assigned to events by four observers flying about in aeroplanes. The fantastic character of this serves to emphasise that we have in general no simple standard way of assigning coordinates. The situation is very different from what we have in Newtonian physics, for there at the very beginning we may use rectangular Cartesian coordinates (x, y, z) and time t. In relativity we formulate physical laws in a way that is meaningful no matter what space-time coordinates are used, and indeed we took a step in that direction when we insisted that in (14) the function $f(x, dx)$ must be an invariant and in (19) that g_{mn} must be a covariant tensor.

In laying down laws for the motion of free material particles and photons we must observe the same invariance, and this we do in making the following *geodesic hypotheses*:

(i) *The world line of a free material particle is a timelike geodesic in space-time, regarded as a Riemannian 4-space with fundamental form Φ as in* (57); *thus the world line satisfies the variational equation*

$$\delta \int ds = 0. \tag{94}$$

(ii) *The world line of a photon is a null geodesic with respect to* Φ.

This assumption about photons is consistent with the one already made in § 13.

Although we shall accept these geodesic hypotheses as working rules to determine the behaviour of material particles and photons, nevertheless we may keep in mind the possibility of *proving* these hypotheses on the basis of others.

From (94) we obtain the Lagrangian *equations of motion* for a free material particle (equations of its *trajectory*)

$$\frac{d}{dw}\frac{\partial f}{\partial x^{r'}} - \frac{\partial f}{\partial x^r} = 0, \quad f = (-g_{mn}x^{m'}x^{n'})^{\frac{1}{2}}, \tag{95}$$

w being any parameter which increases steadily along the world line. If we choose $w = s$, the integrated separation along the world line, then $f = 1$ and (95) may be written

$$\frac{d}{ds}\frac{\partial}{\partial x^{r'}}(f^2) - \frac{\partial}{\partial x^r}(f^2) = 0, \quad f(x, x') = 1, \tag{96}$$

or equivalently

$$\frac{d}{ds}\frac{\partial L}{\partial x^{r'}} - \frac{\partial L}{\partial x^r} = 0, \quad L(x, x') = -\tfrac{1}{2}, \tag{97}$$

where

$$L(x, x') = \tfrac{1}{2}g_{mn}x^{m'}x^{n'}; \tag{98}$$

the equations written on the right in (96) and (97) are particular first integrals of the equations of motion.

The form (97) is usually the most convenient for calculation, but we can also use the equivalent equations, more familiar in Riemannian geometry,

$$\frac{d^2x^r}{ds^2} + \left\{ {r \atop mn} \right\}\frac{dx^m}{ds}\frac{dx^n}{ds} = 0, \quad g_{mn}\frac{dx^m}{ds}\frac{dx^n}{ds} = -1, \tag{99}$$

the Christoffel symbols being

$$\left\{ {r \atop mn} \right\} = g^{rp}[mn, p], \quad 2[mn, p] = \frac{\partial g_{mp}}{\partial x^n} + \frac{\partial g_{np}}{\partial x^m} - \frac{\partial g_{mn}}{\partial x^p}. \tag{100}$$

If we change the parameter in (97), (98) from s to $w = s/k$, where k is any constant, then the equations read

$$\frac{d}{dw}\frac{\partial L}{\partial x^{r'}} - \frac{\partial L}{\partial x^r} = 0, \quad L(x, x') = -\tfrac{1}{2}k^2. \tag{101}$$

Now it is clear from (99) that a trajectory is determined by an initial event and an initial direction in space-time. If we let the initial direction tend to a null direction and at the same time let k tend to zero, we get as limiting form of (101)

$$\frac{d}{dw}\frac{\partial L}{\partial x^{r'}} - \frac{\partial L}{\partial x^r} = 0, \qquad L(x, x') = 0, \tag{102}$$

where L is as in (98), the derivatives of course being with respect to w. These are the equations of null geodesics, i.e. *the world lines of photons*, the equation on the right of (102) being a particular first integral. These equations retain their form if w is transformed to $\overline{w}$ according to

$$\overline{w} = aw + b, \tag{103}$$

where a and b are any constants. An equivalent form of (102) is

$$\frac{d^2x^r}{dw^2} + \left\{ \begin{matrix} r \\ mn \end{matrix} \right\} \frac{dx^m}{dw}\frac{dx^n}{dw} = 0, \qquad g_{mn}\frac{dx^m}{dw}\frac{dx^n}{dw} = 0. \tag{104}$$

We cannot of course use s as parameter on a null geodesic, since $ds = 0$ on it.

§ 19. THE SPECIAL AND GENERAL THEORIES OF RELATIVITY

Relativistic physics, it might be said, reduces itself to the geometry of 4-dimensional Riemannian space-time. Putting the simplest case first, we should start with *flat space-time*, for which the curvature tensor vanishes,

$$R_{mnrs} = 0; \tag{105}$$

then there exist coordinates x^r such that the fundamental form (57) reads

$$\Phi = g_{mn}dx^m dx^n = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2, \tag{106}$$

not at one event only but throughout space-time. *The special theory of relativity is the theory of flat space-time, interpreted physically.*

What is the range of physical applicability of this theory?

It must be said at once that quantum theory is not considered in this book, except incidentally; we deal with deterministic physics, and the physical validity of statements must be qualified by the uncertainties of quantum theory. As a practical limitation this gains more importance the smaller the masses of the particles under consideration are.

But, laying this aside and accepting a deterministic model as physi-

cally interesting, there is an important omission: *we leave out gravitational phenomena.*

According to current dogma, gravitation manifests itself by producing a curvature of space-time, so that (105) is not true in a gravitational field and in consequence no coordinates exist such that (106) holds throughout space-time. *The theory of gravitation based on the curvature of space-time is Einstein's general theory of relativity.* In this theory (105) is replaced by field equations connecting the curvature of space-time with the distribution of matter. A free material particle (free, that is, except for the universal effect of gravity) moves according to the geodesic hypothesis (94) and a photon also obeys the geodesic hypothesis. Except for the foundations already laid and some remarks in the next section, the general theory of relativity lies outside the scope of this book.

In familiar elementary problems of Newtonian mechanics gravity plays an important part. A simple pendulum moves under the action of two forces, the tension in the string and the force of gravity, and Newtonian mechanics treats these two forces on the same footing mathematically. We cannot admit this Newtonian treatment in relativity because it uses an absolute time.

What, then, is the present status of such simple problems? Unfortunately they have none. If, as stated above, gravity is excluded from the special theory, it cannot embrace such problems, and the machinery of the general theory is too ponderous to cope with them. They are forgotten problems, uninteresting to the practical physicist because he has good reason to suppose that the prediction of any relativistic treatment would differ from the Newtonian prediction by amounts too small to measure. Still, as a matter of principle, one would like to see the gap filled.

Attempts have been made to fit gravity into special relativity, the most successful being that of A. N. Whitehead; for his theory and some of its developments the reader is referred to WHITEHEAD [1922], SYNGE [1951, 1952*c*] and RAYNER [1954].[1]

Whitehead's theory is much simpler mathematically than the general theory of relativity and might be used in such problems as that of the simple pendulum, but the translation of Newtonian rigid dynamics

[1] Whitehead's theory may also be set in a space-time of constant Riemannian curvature, which comes next after flat space-time in order of mathematical simplicity; cf. TEMPLE [1923], RAYNER [1953].

into special relativity presents difficulties apparently insurmountable, as described in the next section.

§ 20. RIGID MOTIONS

In § 17 we discussed the concept of an infinitesimal rigid rod, but that is only the first step in a discussion of the concept of rigidity, which plays such an important part in Newtonian mechanics. Any continuum, and in particular a rigid solid, must appear in space-time as a congruence of timelike world lines (the histories of its particles), and any definition of rigidity must be such as to impose conditions on this congruence.

Before the invention of the general theory of relativity, so that only flat space-time was under consideration, BORN [1909] gave a very natural definition of rigidity: *in a rigid motion every pair of adjacent particles are rigidly connected in the sense of* § 17 (*D constant in Fig. 7*). But further investigation by HERGLOTZ [1910] and NOETHER [1910] showed that this definition is not satisfactory, since it gives only three degrees of freedom to a relativistic rigid body instead of the usual six. To supply the missing three degrees of freedom, GARDNER [1952] (cf. SYNGE [1952*a*, *b*]) suggested another definition involving a privileged particle in the body, but this modification is contrary to experiment (cf. DITCHBURN and HEAVENS [1952]). It seems impossible to carry the concept of rigidity into the special theory of relativity except by taking the body to be a discrete set of particles (GARDNER [1953], MATRAI [1953]), or by confining the rigidity to the surface of the body (POUNDER [1954], SYNGE [1954b]). For rigidity in connection with the Michelson-Morley experiment, see V-§ 14.

The fact is that in flat space-time the Born conditions of rigidity restrict the congruence of world lines too much. There is more freedom in curved space-time (general relativity), and SALZMAN and TAUB [1954] have shown that the Born conditions there yield conditions of rigidity of the form

$$\lambda_{m|n} + \lambda_{n|m} + \lambda_m \lambda_{n|p} \lambda^p + \lambda_n \lambda_{m|p} \lambda^p = 0. \tag{107}$$

Here λ^r is the unit tangent vector to the world line, regarded as a function of the space-time coordinates x^r, and the stroke indicates covariant differentiation. There are ten equations in (107), but when we take into consideration the equation

$$g_{mn} \lambda^m \lambda^n = -1, \tag{108}$$

there are only seven independent equations in (107) and (108). For a rigid motion in general relativity we have 21 equations in all (7 as above, 10 field equations, 4 coordinate conditions) and 21 unknowns (4 components of λ^r, 10 components of g_{mn}, 1 density function, 6 components of stress), and so the problem of rigid motion appears to be soluble and definite.

But in the special theory, instead of field equations we have the much more restrictive conditions (105). These tell us that coordinates exist for which g_{mn} have the special values (1, 1, 1, — 1) as in (106), and then we see in (107) and (108) *seven* equations for only *four* unknowns. That is why the Born definition fails in flat space-time.

From this point on, we confine our attention to the special theory of relativity and the concept of rigidity will be encountered only in two simple forms. First, we shall consider bodies in which the world lines are straight and parallel in flat space-time. This corresponds to motion without acceleration, and the Born conditions are satisfied. Secondly in IV-§ 6 we shall treat Lorentz transformations as rigid body displacements in space-time. But there the word *rigid* has quite a different meaning, involving the constancy of the separation of two *events* and not the constancy of the normal separation of two *world lines*, which is what we have been considering in § 17 and above.

CHAPTER II

INTRODUCTION TO THE SPECIAL THEORY

§ 1. BASIS OF THE SPECIAL THEORY OF RELATIVITY

We shall now set down certain basic formulae of the special theory of relativity (physics without gravity), to be regarded either as deductions from the theory of Chapter I or as a new and sufficient set of axioms for the use of anyone who wishes to study the special theory without reference to the general theory — quite a natural wish since the special theory is easier to understand and much richer in contacts with physical reality.

The space-time of special relativity is a four-dimensional continuum of events to which coordinates may be assigned in a great variety of ways, but there exist coordinates x^r such that the separation ds between any pair of adjacent events is given by

$$ds = [-(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2]^{\frac{1}{2}}, \tag{1}$$

this separation being the time measured by a standard clock having a world line which includes the two events.

Relative to any event, the neighbouring part of space-time is divided into

$$\begin{aligned} &\text{past:} && (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 < 0, \ dx^4 < 0, \\ &\text{present:} && (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 > 0, \\ &\text{future:} && (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 < 0, \ dx^4 > 0. \end{aligned} \tag{2}$$

We have here chosen the sense of increasing x^4 to be the time-sense from past to future. The future is divided from the present by one sheet of the null cone, with equation

$$(dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 = 0, \ dx^4 > 0, \tag{3}$$

and the past is divided from the present by the other sheet,

$$(dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 = 0, \ dx^4 < 0, \tag{4}$$

The equations of motion of a free material particle are given by the variational equation I-(94)

$$\delta \int ds = 0, \tag{5}$$

leading by I-(99), in which the Christoffel symbols are now zero, to the differential equations and particular first integral

$$\frac{d^2x^r}{ds^2} = 0, \quad \left(\frac{dx^1}{ds}\right)^2 + \left(\frac{dx^2}{ds}\right)^2 + \left(\frac{dx^3}{ds}\right)^2 - \left(\frac{dx^4}{ds}\right)^2 = -1. \tag{6}$$

These equations integrate at once to

$$x^r = a^r s + b^r, \quad (a^1)^2 + (a^2)^2 + (a^3)^2 - (a^4)^2 = -1, \tag{7}$$

where the a's and b's are eight constants of integration, the a's being bound by the relationship shown.

If we regard space-time as a flat 4-space with fundamental form

$$\Phi = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 \tag{8}$$

— in other words, if we use the language of geometry — then the world lines of free particles, as shown in (7), are to be described as *straight lines* in space-time. Spacelike geodesics are also straight lines, with equations

$$x^r = a^r s + b^r, \quad (a^1)^2 + (a^2)^2 + (a^3)^2 - (a^4)^2 = 1. \tag{9}$$

The world line of a photon is a null geodesic and so by I-(104) has the differential equations and particular first integral

$$\frac{d^2x^r}{dw^2} = 0, \quad \left(\frac{dx^1}{dw}\right)^2 + \left(\frac{dx^2}{dw}\right)^2 + \left(\frac{dx^3}{dw}\right)^2 - \left(\frac{dx^4}{dw}\right)^2 = 0, \tag{10}$$

and hence, by integration,

$$x^r = a^r w + b^r, \quad (a^1)^2 + (a^2)^2 + (a^3)^2 - (a^4)^2 = 0. \tag{11}$$

Thus the world lines of photons are also straight lines in space-time, distinguished from the world lines of free material particles by being null lines, as indicated by the equation on the right in (11).

§ 2. FINITE SEPARATIONS

Consider two events, P and Q, with coordinates y^r and z^r respectively. Through P and Q we can draw a unique straight line in space-time, and it will be timelike, spacelike or null according to the choice of P and Q. Suppose it to be timelike. Then this straight line has the

equations (7), and since it passes through P and Q we have

$$y^r = a^r s_P + b^r, \quad z^r = a^r s_Q + b^r, \tag{12}$$

where s_P and s_Q are the values of s at P and Q respectively. Hence, writing $s_P - s_Q = \Delta s$, we have

$$y^r - z^r = a^r \Delta s, \tag{13}$$

and so by (7)

$$\Delta s^2 = -(y^1 - z^1)^2 - (y^2 - z^2)^2 - (y^3 - z^3)^2 + (y^4 - z^4)^2. \tag{14}$$

Similar reasoning applies for spacelike and null straight lines, and so we may say that *the finite separation between two events, y^r and z^r, is*

$$\Delta s = \int_P^Q ds = [\varepsilon\{(y^1 - z^1)^2 + (y^2 - z^2)^2 + (y^3 - z^3)^2 - (y^4 - z^4)^2\}]^{\frac{1}{2}}, \tag{15}$$

where ε is the indicator of the displacement PQ, being -1 for a timelike displacement and 1 for a spacelike displacement. In the case of a null displacement, $\Delta s = 0$.

The great simplicity of the special theory, as contrasted with the general theory, comes from the fact that the equations of the geodesics in space-time can be integrated immediately and simply. Thus we get a finite algebraic geometry in space-time instead of a differential geometry. We recognise that (15) gives, in the language of geometry, a finite "distance" between two "points", but, in view of possible confusion arising from other uses of the word "distance" it will be wiser to stick to the word *separation.*

If, from any event b^r, we draw all the null geodesics, they form, by (11), a *finite null cone* with equation

$$(x^1 - b^1)^2 + (x^2 - b^2)^2 + (x^3 - b^3)^2 - (x^4 - b^4)^2 = 0. \tag{16}$$

It has two sheets: a past sheet with $x^4 - b^4 < 0$, and a future sheet with $x^4 - b^4 > 0$. Thus relative to any chosen event b^r the whole of space-time is divided in a simple way into past, present and future as follows (Fig. 1):

past:

$$(x^1 - b^1)^2 + (x^2 - b^2)^2 + (x^3 - b^3)^2 - (x^4 - b^4)^4 < 0, \quad x^4 - b^4 < 0,$$

present:

$$(x^1 - b^1)^2 + (x^2 - b^2)^2 + (x^3 - b^3)^2 - (x^4 - b^4)^2 > 0, \tag{17}$$

future:

$$(x^1 - b^1)^2 + (x^2 - b^2)^2 + (x^3 - b^3)^2 - (x^4 - b^4)^2 < 0, \quad x^4 - b^4 > 0.$$

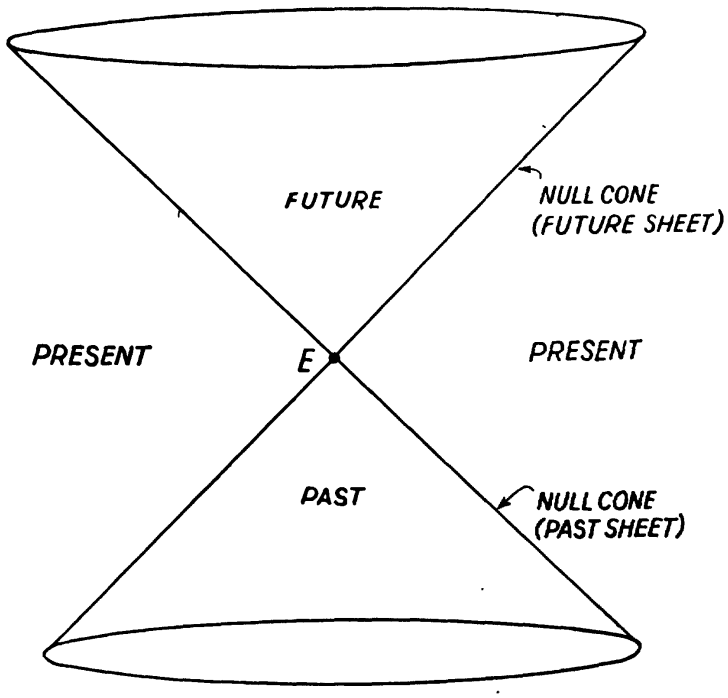


Fig. 1 – The finite past, present and future relative a chosen event E

What was said in Ch. I about the measurement of infinitesimal spacelike separations (cf. Fig. 1–4, p. 25) now applies to *finite* separations, the infinitesimals being replaced by finite differences of coordinates. The discussion of the orthogonality of infinitesimal vectors now applies to finite space-time vectors, considered as segments joining events; the vector drawn from the event a^r to the event b^r has components

$$b^1 - a^1,\ b^2 - a^2,\ b^3 - a^3,\ b^4 - a^4. \tag{18}$$

This means that timelike and spacelike separations are physically measurable by means of standard clocks and photons as in Ch. I, and orthogonality can be tested physically with the same apparatus.

If a vector V^r carries us from an event P to an event Q, then the separation PQ is

$$\begin{aligned} PQ &= [-(V^1)^2 - (V^2)^2 - (V^3)^2 + (V^4)^2]^{\frac{1}{2}} \text{ if } V^r \text{ timelike,} \\ PQ &= [(V^1)^2 + (V^2)^2 + (V^3)^2 - (V^4)^2]^{\frac{1}{2}} \text{ if } V^r \text{ is spacelike.} \end{aligned} \tag{19}$$

The condition of orthogonality of U^r and V^r is

$$U^1V^1 + U^2V^2 + U^3V^3 - U^4V^4 = 0. \tag{20}$$

There are several notations which may be used in the special theory of relativity. The preceding formulae have been written out explicitly, because they are most easily understood in that way. But it is

economical of writing to bring back the tensor g_{mn} of I–(106) with the simple constant components

$$g_{mn} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{21}$$

We use this g_{mn} to lower superscripts in the usual manner of tensor calculus, and we raise subscripts by means of the conjugate tensor g^{mn}, the components of which are precisely as in (21). All the formulae so far written in this Chapter may be put more compactly in this notation. For example, the fundamental form Φ of (8) is now

$$\Phi = g_{mn}dx^m dx^n, \tag{22}$$

and the straight lines (7) and (9) are

$$x^r = a^r s + b^r, \quad a_m a^m = \varepsilon \tag{23}$$

($\varepsilon = -1$ for timelike straight line,
$\varepsilon = +1$ for spacelike straight line).

In what follows we shall use this compact notation. Later, we shall gain even more simplicity of form by introducing an imaginary time-coordinate, but for the present all quantities are real.

§ 3. HOW TO DRAW A STRAIGHT LINE IN SPACE-TIME

A boxer striking at his opponent's jaw, a three-quarter back racing to intercept an opponent on the football field, a captain steering his ship at night in crowded waters — all these have space-time problems to solve. There is a world line (of fist, body, or ship) which is either to intersect another world line or deliberately avoid it, and the solution of such problems in action is instinctive, owing little if anything to Newtonian ideas of space and time and velocity. These ideas we too avoid for the present because our instinctive knowledge of the properties of space-time is better gained by establishing direct connections between space-time and simple ideal physical operations, as we have already tried to do in Ch. I.

The desired instinctive approach to space-time is comparable to that which we all have with regard to Euclidean 3-space, many of the properties of which are obvious to us, although we are ready to prove them in case of doubt. There are two outstanding differences — space-

time is four-dimensional and its metric is not positive-definite. Only by experience can we learn to what extent these things are important, and in this section and the next we shall pursue some ideas partly for their own sake and partly to develope the instinctive approach to space-time.

To draw a straight line on a piece of paper to join two points P and Q, one lays down a straight edge so that P and Q lie on it and then runs a pencil along the edge. Let us see how to draw a straight line in space-time to join two events P and Q.

Suppose first that the vector PQ is timelike. Then the two events can be included in the history of a material particle, and the straight line is drawn by projecting a particle from the event P in such a way that it includes Q in its history. The straight line consists of the events in the history of the particle; it may be produced into the future by following the further history of the particle, and into the past by following its history before it reached P, supposing it to have passed through P without disturbance. We have assumed that PQ points into the future; if it points into the past, P and Q must be interchanged in the description.

If the vector PQ is not timelike, it must be null or spacelike. If it is null, a photon can be projected from P so as to include Q in its history, and this history gives us the straight line PQ.

The case where PQ is spacelike is not so easy, for now neither material particle nor photon can include both P and Q in its history. There are several ways of drawing the straight line PQ in this case; we shall consider one, leaving it to the reader to devise others.

In I-§ 14 we saw that infinitesimal spacelike separations can be measured by photons and clocks, and this method is available for finite separations in the special theory of relativity. We shall use this to give a physical meaning to the spacelike straight line PQ. Let the equations of this straight line be, as in (23),

$$x^r = a^r s + b^r, \quad a_m a^m = 1, \tag{24}$$

and let $s = s_1$ at P, $s = s_2$ at Q, so that

$$(x^r)_P = a^r s_1 + b^r, \quad (x^r)_Q = a^r s_2 + b^r. \tag{25}$$

Then for any third event R, with $s = s_3$, on the straight line PQ we have

$$\begin{aligned} (x^r)_P - (x^r)_Q &= a^r(s_1 - s_2), \\ (x^r)_P - (x^r)_R &= a^r(s_1 - s_3), \\ (x^r)_Q - (x^r)_R &= a^r(s_2 - s_3), \end{aligned} \tag{26}$$

and so the squares of the separations are

$$PQ^2 = (s_1 - s_2)^2, \quad PR^2 = (s_1 - s_3)^2, \quad QR^2 = (s_2 - s_3)^2. \tag{27}$$

It follows that one of the following equations is satisfied:

$$PQ = PR + RQ, \quad PR = PQ + QR, \quad QR = QP + PR, \tag{28}$$

these being (positive) spacelike separations, physically measurable. These equations give us a physical test by which we can determine whether an event R lies, or does not lie, on the straight line PQ, the first equations giving us events R between P and Q and the other equations events outside the segment PQ (Fig. 2).

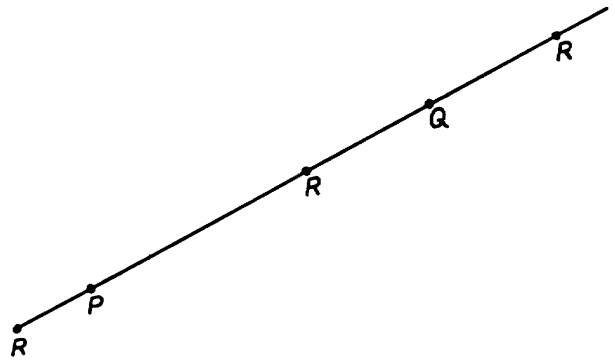


Fig. 2 – Construction of spacelike straight line through P and Q

Although theoretically correct, the equations (28) give a poor practical definition of the straight line from the standpoint of surveyor's technique, because a small displacement of R off the line will give only a second-order error in these equations; this is due to the geodesic character of a straight line.

§ 4. PAIRS OF STRAIGHT LINES IN SPACE-TIME, PARALLEL AND SKEW

Two straight lines in space-time with equations

$$x^r = a^r s + b^r, \quad x^r = \bar{a}^r \bar{s} + \bar{b}^r \tag{29}$$

are *parallel* if $\bar{a}^r = \pm a^r$. Two parallel lines cannot intersect unless they happen to coincide entirely. If $\bar{a}^r \neq \pm a^r$, the two lines either intersect or are *skew*.

Parallel straight lines are of interest because it is by means of them that we break down space-time into space and time, as we shall see later. A pair of skew straight lines represents an interesting physical situation — the histories of two free particles coming in from infinity, passing close to one another, and then going off to infinity again.

Let us take two straight lines L and $\bar{L}$ as in (29), both timelike so that

$$a_m a^m = -1, \quad \bar{a}_m \bar{a}^m = -1, \tag{30}$$

and with the vectors a^r, $\bar{a}^r$ pointing into the future so that a^4 and $\bar{a}^4$ are positive. Let $\bar{x}^r$ be any event on $\bar{L}$ and x^r the foot of the perpendicular dropped from $\bar{x}^r$ on L (Fig. 3), the vector $\eta^r = \bar{x}^r - x^r$ being the *deviation* of $\bar{L}$ from L at x^r. We have

$$\eta^r = \bar{a}^r \bar{s} + \bar{b}^r - a^r s - b^r \tag{31}$$

and, since this is orthogonal to L,

$$0 = a_r \eta^r = a_r \bar{a}^r \bar{s} + s + a_r(\bar{b}^r - b^r). \tag{32}$$

(Remember that we are lowering superscripts by means of the tensor (21).) Hence

$$\bar{s} = -\frac{s}{a_n \bar{a}^n} - \frac{a_r(\bar{b}^r - b^r)}{a_n \bar{a}^n}, \tag{33}$$

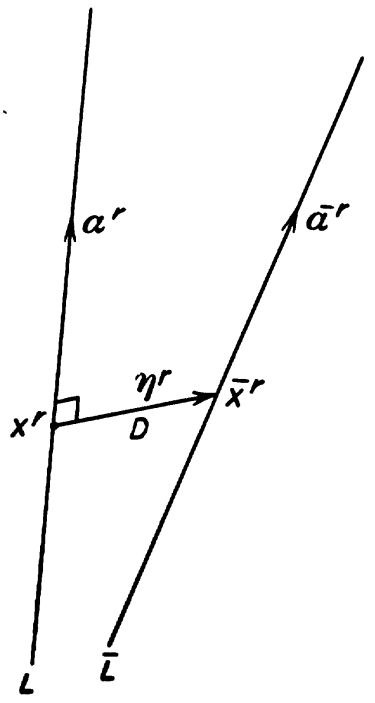


Fig. 3 – Deviation of two straight lines in space-time

and substitution of this in (31) gives η^r as a linear vector function of s:

$$\eta^r = -s\left(a^r + \frac{\bar{a}^r}{a_n \bar{a}^n}\right) + \beta^r \tag{34}$$

where β^r is a constant vector.

Now η^r is spacelike, being orthogonal to the timelike L, and so the distance D of $\bar{L}$ from L at x^r is given by

$$D^2 = \eta_r \eta^r = s^2\left[1 - \frac{1}{(a_n \bar{a}^n)^2}\right] - 2s\beta_r\left(a^r + \frac{\bar{a}^r}{a_n \bar{a}^n}\right) + \beta_r \beta^r. \tag{35}$$

This formula is like what we have in Euclidean 3-space, a quadratic function.

If $\bar{L}$ and L are parallel, we have $\bar{a}^r = a^r$ and therefore

$$D^2 = \beta_r \beta^r, \tag{36}$$

a constant. Thus *the distance between two parallel timelike straight lines is constant.* The particles whose histories these lines represent are therefore to be regarded as rigidly connected in the sense of I–§ 17, although the distance here is finite, not infinitesimal.

We shall now show that the converse is true, namely that *two timelike straight lines at a constant distance from one another (as judged by*

one of them) are necessarily parallel. This is not obvious, for it means that we have to establish a vector condition (parallelism) from a scalar condition (constancy of distance). The constancy of distance has a real physical meaning, since this distance is essentially what is measured by optical interference methods.

Various proofs may be given [cf. VI–(51)]; the following proof gives us an opportunity to introduce polar coordinates in space-time.

We are given that L and $\bar{L}$ are at a constant distance from one another as judged by L, and therefore D^2 in (35) is independent of s. Hence

$$(a_n \bar{a}^n)^2 = 1 . \tag{37}$$

Here a^r and $\bar{a}^r$ are two timelike unit vectors, both pointing into the future; by I–(82) their scalar product is negative and so the condition (37) is equivalent to

$$a_n \bar{a}^n = -1 . \tag{38}$$

Now for any two timelike unit vectors pointing into the future there exist polar coordinates $\chi, \theta, \phi, \bar{\chi}, \bar{\theta}, \bar{\phi}$ such that

$$\begin{aligned} a^1 &= \sinh\chi \sin\theta \cos\phi, & \bar{a}^1 &= \sinh\bar{\chi} \sin\bar{\theta} \cos\bar{\phi}, \\ a^2 &= \sinh\chi \sin\theta \sin\phi, & \bar{a}^2 &= \sinh\bar{\chi} \sin\bar{\theta} \sin\bar{\phi}, \\ a^3 &= \sinh\chi \cos\theta, & \bar{a}^3 &= \sinh\bar{\chi} \cos\bar{\theta}, \\ a^4 &= \cosh\chi, & \bar{a}^4 &= \cosh\bar{\chi}; \end{aligned} \tag{39}$$

we verify that

$$\begin{aligned} a_n a^n &= (a^1)^2 + (a^2)^2 + (a^3)^2 - (a^4)^2 = -1, \\ \bar{a}_n \bar{a}^n &= -1 . \end{aligned}$$

Then

$$a_n \bar{a}^n = \sinh\chi \sinh\bar{\chi} \cos\psi - \cosh\chi \cosh\bar{\chi}, \tag{40}$$

where ψ is a real angle ($0 \leqslant \psi \leqslant \pi$) such that

$$\cos\psi = \cos\theta \cos\bar{\theta} + \sin\theta \sin\bar{\theta} \cos(\phi - \bar{\phi}) . \tag{41}$$

We may write (40) in the form

$$\begin{aligned} -a_n \bar{a}^n &= \cosh\chi \cosh\bar{\chi}(\cos^2 \tfrac{1}{2}\psi + \sin^2 \tfrac{1}{2}\psi) \\ &\quad - \sinh\chi \sinh\bar{\chi}(\cos^2 \tfrac{1}{2}\psi - \sin^2 \tfrac{1}{2}\psi) \\ &= \cosh(\chi - \bar{\chi}) \cos^2 \tfrac{1}{2}\psi + \cosh(\chi + \bar{\chi}) \sin^2 \tfrac{1}{2}\psi . \end{aligned} \tag{42}$$

This expression is bounded by $\cosh(\chi + \bar{\chi})$ and $\cosh(\chi - \bar{\chi})$, each of

which is at least equal to unity. It follows that for two timelike unit vectors pointing into the future

$$a_n \bar{a}^n \leqslant -1, \tag{43}$$

the sign of equality holding only under one of the following circumstances:

$$\chi = \bar{\chi} = 0;\ \chi = \bar{\chi},\ \psi = 0;\ \chi = -\bar{\chi},\ \psi = \pi. \tag{44}$$

It is easy to see from (39) and (41) that any one of these conditions implies

$$a^r = \bar{a}^r. \tag{45}$$

We conclude than that (38) implies (45), and so the constancy of distance implies parallelism, which is what we had to prove.

§ 5. THE PHYSICAL MEANING OF THE SPECIAL COORDINATES

Existence theorems, as such, are of little interest in physics. The physicist, informed that there exist space-time coordinates which give the fundamental form the simple expression

$$\Phi = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2, \tag{46}$$

wants to know how to assign such coordinates to events. We must remember that there is no difficulty at all in thinking of ways of assigning coordinates to events; the problem is to pick out, from the vast variety of all possible coordinates, those which produce the simplicity of (46).

We proceed as follows. Choose any event O. Choose any other event D, such that OD is timelike. Adjust D on the straight line OD so that the separation OD is unity. Actually, this is very simple. We project a free material particle from O and take as the event D that event in the history of the particle at which a standard clock carried by it reads unity, the clock having been set to zero reading at O.

We now add three more events, A, B, C, making OA, OB, OC orthogonal to OD and to one another. These are physical operations (I-§ 15). We further make the separations OA, OB, OC each unity. The result is an orthogonal tetrad of unit vectors in space-time

$$OA,\ OB,\ OC,\ OD,$$

the first three being spacelike and OD being timelike.

There exists (unknown to us) a special coordinate system x^r for which (46) holds. In this coordinate system, the four vectors just

described have components, say

$$a^r,\ b^r,\ c^r,\ d^r,$$

these sixteen numbers being unknown to us. But from the unit character of the vectors and their orthogonality, we know that the following ten relations must be satisfied:

$$\begin{aligned} a_r a^r = b_r b^r = c_r c^r = -\, d_r d^r = 1\,, \\ a_r b^r = a_r c^r = a_r d^r = b_r c^r = b_r d^r = c_r d^r = 0\,. \end{aligned} \tag{47}$$

Take any sixth event E (Fig. 4). We can pass from O to E by drawing a broken line in space-time, starting along OA and then going parallel to OB, OC, OD in turn. Although we are using geometrical words, these are actually physical operations in view of the physicisation in § 3 above. The separations (space-time "lengths") of these four steps are physically measurable; let their values be

$$y^1, y^2, y^3, y^4.$$

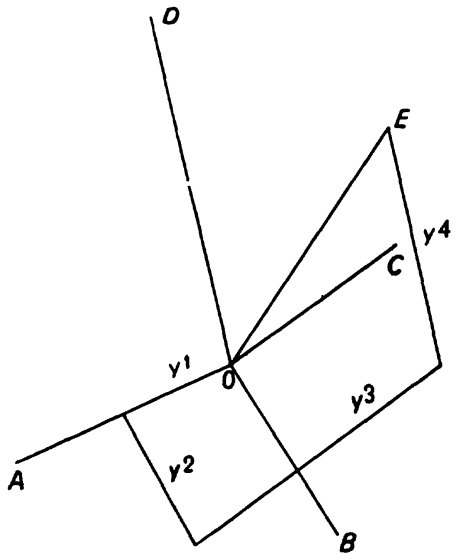


Fig. 4 – Construction for special coordinates

Now in the (unknown) special coordinate system, the vector OE will have some components e^r, and by virtue of the construction just carried out, the measured y's are connected with the unmeasured and unknown components by the four equations

$$e^r = y^1 a^r + y^2 b^r + y^3 c^r + y^4 d^r. \tag{48}$$

If, instead of $E(e^r)$, we take another event $\bar{E}(\bar{e}^r)$, then there will be new values of the y's, say $\bar{y}^r$, and from (48) and the similar equation we get by subtraction

$$e^r - \bar{e}^r = (y^1 - \bar{y}^1)a^r + (y^2 - \bar{y}^2)b^r + (y^3 - \bar{y}^3)c^r + (y^4 - \bar{y}^4)d^r. \tag{49}$$

Hence we obtain, making use of (47) and denoting by ε the indicator of $E\bar{E}$,

$$\begin{aligned} \varepsilon\, E\bar{E}^2 &= (e_r - \bar{e}_r)\,(e^r - \bar{e}^r) \\ &= (y^1 - \bar{y}^1)^2 + (y^2 - \bar{y}^2)^2 + (y^3 - \bar{y}^3)^2 - (y^4 - \bar{y}^4)^2, \end{aligned} \tag{50}$$

or, for an infinitesimal $E\bar{E}$,

$$\varepsilon\, ds^2 = (dy^1)^2 + (dy^2)^2 + (dy^3)^2 - (dy^4)^2. \tag{51}$$

Thus the numbers y^r, which are physically measurable, form a special coordinate system for which Φ is as in (46).

This coordinate system is of course not unique. We observe, and this is most important, that in its choice we have used *ten* degrees of freedom: four degrees of freedom in the choice of the event O and $16 - 10 = 6$ degrees of freedom in the choice of the orthogonal unit tetrad $OABCD$.

§ 6. SPLITTING SPACE-TIME INTO SPACE AND TIME

We are all accustomed to thinking in the Newtonian manner of space and time. When an event occurs in the laboratory, we record its position by measuring its distances x, y, z from two walls and the floor and its time t by reading a clock. We must now see, from the relativistic standpoint, the meaning of these operations.

Let us think of a laboratory in which there is no gravity, and let us put in that laboratory an observer equipped with suitable apparatus — material particles carrying standard clocks, and devices for projecting photons from these particles and reflecting the photons from the particles. The observer takes two material particles and projects them in any arbitrary way, screening them from all forces so that they are free particles. He can measure the distance D between these particles by the method of I–(93), extended from infinitesimal to finite separations as in (35). He will of course in general not find D to be a constant. But let him try again and again until he succeeds in projecting two free particles so that the distance between them is constant. The particles are then rigidly connected in the sense of I–§ 17.

The observer then projects more and more free particles, making sure as he projects each of them that it is rigidly connected in the above sense to those already projected. Having projected a vast number of particles in this way, he may think of their number as infinite and regard their world lines as filling space-time, or at least the portion of space-time available to him. Since these are the world lines of free material particles with constant distances between them, it follows from § 4 that they are parallel. Thus by trial and error the observer has succeeded in constructing a set of parallel straight lines in space-time.

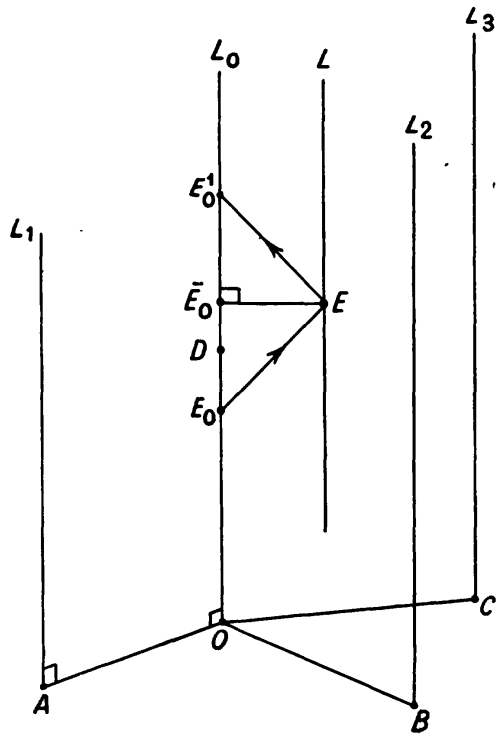


Fig. 5 – A Galileian frame of reference in space-time

Now let the observer jump on to one of these free particles and ride with it, so that his own world line is one of the set of parallel straight lines. He is now in a position to separate space-time into space and time, as we shall see below.

Let L_0 (Fig. 5) be the world line of the observer. He selects some event O on L_0 and calls it *the origin of space and time.* He sets his standard clock so that it reads zero at O. He is then able to assign a time x^4 to any event occurring on L_0, simply by reading his clock when the event occurs. He notes in particular the event D on L_0 for which $x^4 = 1$. (OD is in fact a timelike unit vector, as in the orthogonal tetrad of Fig. 4.)

To attach a time x^4 to any event E occurring on some other world line L of the parallel set, he emits a photon from some suitably chosen event E_0 on L_0 so that the photon catches E in its history. The photon is then reflected back from E, arriving at the observer at the event E_0'. He then assigns to E the time

$$x^4(E) = \tfrac{1}{2}[x^4(E_0) + x^4(E_0')]. \tag{52}$$

It is easy to see that this is precisely the measured time of $\bar{E}_0$, the event obtained by dropping a straight perpendicular from E on L_0, so that

$$x^4(E) = x^4(\bar{E}_0). \tag{53}$$

In accordance with § 4 the distance of L from L_0 is the separation $\bar{E}_0E$, and we have

$$\bar{E}_0E = E_0\bar{E}_0 = \bar{E}_0E_0'. \tag{54}$$

Thus, by observing the times at E_0 and E_0', the observer in fact measures the distance between L and L_0; it is

$$L_0L = \tfrac{1}{2}[x^4(E_0') - x^4(E_0)]. \tag{55}$$

By using light signals (photons) between any two of the particles and standard clocks carried by them, it is possible in this way to

measure the distance between any two of the world lines. These distances are all constant, because the world lines of the particles are parallel straight lines.

The observer needs, in addition to his own world line L_0, three more world lines as a basis for a coordinate system. So he selects a world line L_1 at unit distance from L_0 and notes that event A on it which occurs at zero time. This means, in view of what has been said above, that OA is a common perpendicular to L_0 and L_1. He then selects another world line L_2, also at unit distance from L_0 and notes the event B on it occurring at zero time. Then he has $OA = OB = 1$, where OA and OB are distances between world lines, or equivalently spacelike separations. AB is also a spacelike separation. Now we saw in Fig. 1–5 that the theorem of Pythagoras holds for an infinitesimal spacelike triangle; in flat space-time it holds for a finite triangle also. Thus **OA** and **OB** will be orthogonal if $AB = \sqrt{2}$. (Where there is danger of confusion between a space-time vector and its magnitude, heavy type will be used for the vector.) In general AB will differ from $\sqrt{2}$; but by changing his choice of L_2 the observer can make $AB = \sqrt{2}$, and so make **OA** and **OB** orthogonal to one another.

Then, choosing a third world line L_3 with the event C on it occurring at zero time, he arranges in the same way to make **OC** of unit magnitude and orthogonal to both **OA** and **OB**. He then has an orthogonal tetrad of unit vectors:

$$\mathbf{OA},\ \mathbf{OB},\ \mathbf{OC},\ \mathbf{OD}.$$

Now the vector $\bar{\mathbf{E}}_0\mathbf{E}$, being orthogonal to L_0, can be resolved uniquely into components according to the vector equation

$$\bar{\mathbf{E}}_0\mathbf{E} = x^1\mathbf{OA} + x^2\mathbf{OB} + x^3\mathbf{OC}, \tag{56}$$

the numbers x^1, x^2, x^3 being determined by this resolution. The reader should have no difficulty in supplying an operational technique for this resolution.

The result is that the observer assigns coordinates x^r to any event E, and since this method of assignment is based on a unit orthogonal tetrad as in § 5, these coordinates are in fact the special coordinates for which Φ has the simple form (46). All we have done is to make the process of assigning coordinates a little more physical. Let us now consider in what sense space-time has been split into space and time.

When we speak of *space* in Newtonian physics, we think of a 3-dimensional manifold of points. When we speak of space-time in

relativity, we think of a 4-dimensional manifold of events. The link between the two lies in the fact that the set of parallel world lines we have been considering form a 3-dimensional manifold of world lines, and it is these *world lines* that are to be identified with the *points* of familiar space. The observer is justified in speaking of "space" as a Newtonian physicist might speak of it, provided that he admits two things. First, a "point" of his "space" is really a world line in space-time, and, secondly, that it is *his* "space" and not a universal "space", since the directions in space-time of the parallel world lines may be chosen arbitrarily.

We note that any one of the parallel world lines (fixed points of the observer's space) has the equations

$$x^1 = \text{const.}, \quad x^2 = \text{const.}, \quad x^3 = \text{const.}, \tag{57}$$

while the particular world lines L_0, L_1, L_2, L_3 have the following equations:

$$\begin{aligned} L_0&: x^1 = 0, \quad x^2 = 0, \quad x^3 = 0, \\ L_1&: x^1 = 1, \quad x^2 = 0, \quad x^3 = 0, \\ L_2&: x^1 = 0, \quad x^2 = 1, \quad x^3 = 0, \\ L_3&: x^1 = 0, \quad x^2 = 0, \quad x^3 = 1. \end{aligned} \tag{58}$$

§ 7. GALILEIAN FRAMES OF REFERENCE

A *Galileian frame of reference* means a set of parallel timelike straight lines in space-time, and hence a system of coordinates x^r assigned as above. The name is given in honour of Galileo, the discoverer of what is commonly known as Newton's First Law of Motion, namely, that a free particle travels in a straight line with constant speed.

Note that the frame of reference consists essentially of a congruence[1] of the parallel straight lines and that, after this congruence has been chosen, there is still considerable liberty in the choice of coordinates. First, O may be chosen arbitrarily (a fourfold arbitrariness), and, secondly, there is a threefold arbitrariness in the choice of **OA**, **OB**, **OC**, in view of the conditions of orthogonality and unit magnitude. In fact, for a given Galileian frame of reference, the possible coordinates are subject to the group of transformations

$$\begin{aligned} x^{\varrho'} &= a^\varrho + A^\varrho_\mu x^\mu, \\ x^{4'} &= a^4 + x^4, \end{aligned} \tag{59}$$

[1] A *congruence* in N-space is a set of curves, one curve passing through each point of the space.

where a^ϱ and a^4 are arbitrary constants and A^ϱ_μ any constants satisfying the conditions[1]

$$A^\varrho_\mu A^\sigma_\mu = \delta^{\varrho\sigma}, \tag{60}$$

these being the conditions for an orthogonal transformation in Euclidean 3-space. There are seven degrees of freedom in (59). (Remember that Greek suffixes range 1, 2, 3.)

We shall call any such system of coordinates x^r a *Galileian coordinate system.*

By (57) the spatial coordinates of any one of the particles making up a Galileian frame of reference are constant, and so we speak of them as being *at rest.* Thus a Galileian frame of reference may be thought of as a cloud of free material particles at rest, the distances between them remaining constant. They form a rigid system in the sense of I–§ 17, and this rigidity may be tested by carrying round an infinitesimal rigid measuring rod, making the world lines of its ends coincide with the world lines of adjacent particles. In the geometry of these particles at rest, the element of distance is

$$ds_0^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2. \tag{61}$$

All the experimental verifications of the special theory of relativity are made in terrestrial laboratories, which are regarded as Galileian frames of reference. The world lines L_0, L_1, L_2, L_3 of Fig. 5 are the histories of a corner of the floor of the room and of three points situated on the edges of the room passing through that corner. In that corner a master clock is situated, and the time of any event is the time recorded by that clock, allowance being made for the time taken by a signal to travel from the event to the clock. Apart from unimportant changes of detail, that agrees with standard laboratory technique.

Now an essential criterion for the Galileian character of this frame of reference is that a particle, placed anywhere in the volume of the room should remain permanently at rest relative to the walls. But we all know that a body falls, and this brings us back to the neglect of gravity, discussed in I–§ 19. The practical physicist shrugs his shoulders, points to the very short interval of time involved in the experiment he is performing, and says "A body hardly falls at all in that time!" Thus he accepts the terrestrial laboratory as Galileian for his immediate purposes, although he knows well that for other purposes it is not Galileian.

[1] $\delta^{\varrho\sigma}$ is the Kronecker delta; it is unity if $\varrho = \sigma$, and zero if $\varrho \neq \sigma$.

But it is not gravity alone that violates the Galileian character of the room. Even if we could eliminate gravity, the rotation of the earth would prevent free particles from floating permanently in the volume of the room; they would drift against the wall nearest to the equator. However, the rotation of the earth is so slow that it, like gravity, may be neglected in the consideration of those experiments likely to be of interest in connection with the theory of relativity. On the other hand, it would be foolish to regard a terrestrial laboratory as Galileian if we were studying the motion of a pendulum relativistically (it is gravity that makes the pendulum vibrate) or if we were studying the motion of a gyrocompass relativistically (its oscillations are due to the earth's rotation). Actually one does not study such motions relativistically; a Newtonian treatment is satisfactory in the sense that it predicts the results of experiments within the margin of experimental error, and indeed the theory of relativity has not been developed far enough to enable us to formulate these problems in a relativistic way (cf. I–§ 19).

§ 8. PROPER TIME AND THE SPEED OF LIGHT

Changes of physical units are always confusing, and we shall avoid them as much as possible. But there are some changes of units which demand attention.

In ordinary space we might, if we so wished, measure distances in the directions of the three coordinate axes in three different units, and we might use a fourth unit for distances oblique to the axes. Then the line element ds_0 of space would be such that

$$k^2 ds_0^2 = k_1^2 dx^2 + k_2^2 dy^2 + k_3^2 dz^2, \tag{62}$$

the k's being constants depending on the units chosen. Such a practice is commonly employed in the case of maps of the earth, heights being recorded in feet or metres and horizontal distances in miles or kilometres.

In space-time we shall refrain from using different units for measurements in the directions of the three spatial axes (i.e. measurements of x^1, x^2, x^3), but we shall sometimes change the unit of time, the old measure x^4 and the new measure t being related by

$$x^4 = ct, \tag{63}$$

so that one unit of t is equivalent to c units of x^4. If we use x, y, z

instead of x^1, x^2, x^3, to avoid the writing of indices, then the fundamental form Φ of (46) reads

$$\Phi = dx^2 + dy^2 + dz^2 - c^2dt^2. \tag{64}$$

It is still open to us to change the unit for the separation ds, and indeed we can use different units for timelike separations and spacelike separations. A convenient plan is to write (with the same constant c as in (63))

$$ds = cd\tau \tag{65}$$

for timelike separations. Then $d\tau$ is called the *element of proper time*, and we have

$$\begin{aligned} &\text{Timelike: } d\tau = ds/c = [dt^2 - (dx^2 + dy^2 + dz^2)/c^2]^{\frac{1}{2}}, \\ &\text{Spacelike: } ds = (dx^2 + dy^2 + dz^2 - c^2dt^2)^{\frac{1}{2}}. \end{aligned} \tag{66}$$

The history of a photon satisfies $ds = 0$ or equivalently

$$dx^2 + dy^2 + dz^2 = c^2dt^2, \tag{67}$$

and so, ds_0 being the element of space as in (61), we have

$$c = ds_0/dt. \tag{68}$$

Thus the constant c, introduced in (63), is to be identified physically as the *speed of light*.

Although everything to be said later is true no matter what units are used for space (x, y, z) and time (t), it is well to keep in touch with standard practice by measuring space in cm. and time in sec. Then in (66) we regard $d\tau$ and dt as expressed in sec. and ds, dx, dy, dz as expressed in cm. If x^4 is used, it will be in cm. The constant c is expressed in cm./sec. The experimentally determined value is approximately

$$c = 3.00 \times 10^{10} \text{ cm./sec.} \tag{69}$$

Are we to be impressed by the largeness of this number? Certainly not in any absolute sense, for c was introduced as a mere change of scale in (63), and there is much to be said for keeping $c = 1$. The largeness of c in (69) is impressive only when we contrast it with the velocities to which we are accustomed. All velocities in the laboratory (except for electrons and other ultimate particles, and of course for light itself) are small compared with c, and so also are astronomical velocities. Had they been comparable with c, the theory of relativity would have been forced on physicists long ago.

§ 9. MINKOWSKIAN COORDINATES

A complex number $a + ib$ (where $i = \sqrt{-1}$) is regarded mathematically as an ordered pair of real numbers (a, b), but, strangely enough, although this mathematical rationalisation of what had previously been mysterious is now well over a century old, some faint odour of mystery seems still to haunt the square root of minus one in physics. It should not.

We have no occasion here to introduce into relativity complex numbers in the full sense, i.e. numbers with both real and imaginary parts. As a matter of fact, if we were to think of a space-time with each event specified by four complex numbers, we would be dealing with an *eight*-dimensional manifold, which would be something entirely different from the *four*-dimensional manifold of our experience. All we shall do is to introduce, following Minkowski, a *purely imaginary* time, to use the singularly unfortunate terminology of mathematics. Regarded as a number-pair, a real number a is $(a, 0)$ and a pure imaginary ib is $(0, b)$; each is defined by *one* real number, which is dropped into the front pocket in the case of a and into the back pocket in the case of ib. The purpose of this imaginary time is to secure a great economy in writing, and also to suggest certain analogies between space-time and ordinary space.

The imaginary time coordinate x_4 is defined by

$$x_4 = ix^4 = ict. \tag{70}$$

It is convenient to lower the superscripts on the spatial coordinates, writing

$$x_1 = x^1 = x, \; x_2 = x^2 = y, \; x_3 = x^3 = z. \tag{71}$$

Then the fundamental form Φ of (46) becomes

$$\begin{aligned} \Phi &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2 \\ &= (dx_1)^2 + (dx_2)^2 + (dx_3^2) + (dx_4)^2 \\ &= dx_r dx_r, \end{aligned} \tag{72}$$

or

$$\Phi = g_{mn} dx_m dx_n, \; g_{mn} = \delta_{mn}, \tag{73}$$

δ_{mn} being the Kronecker delta. Thus, for these coordinates x_r, which we call *Minkowskian coordinates*, the fundamental tensor becomes the

unit matrix

$$g_{mn} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{74}$$

It is clear then that the operation of raising or lowering a suffix produces no change at all, and so, for those transformations leaving Φ invariant (i.e. a sum of four squares), covariant and contravariant tensors are transformed in the same way and there is no necessity to distinguish between them. They are like the "Cartesian" tensors of Euclidean 3-space. They may all be written with subscripts (X_r, Y_{rs}).

We can always pass back to real coordinates by using the formulae (70), (71) and the general rules of tensor transformation. It is easy to see that in Minkowskian coordinates every vector has its first three components real and the last imaginary, and for a tensor of the second rank the real and imaginary components are distributed as follows, r standing for a real component and i for an imaginary one:

$$T_{mn} = \begin{pmatrix} r & r & r & i \\ r & r & r & i \\ r & r & r & i \\ i & i & i & r \end{pmatrix} \tag{75}$$

The only trouble arises when we use Minkowskian coordinates in relativistic quantum mechanics, for then $\sqrt{-1}$ occurs in a different way. It is advisable to use two different symbols, say i as in (70) and j for the imaginary element of quantum mechanics. But this will not concern us in this book, since it does not deal with quantum mechanics.

Minkowskian coordinates will henceforth be used throughout, and so it will be well to restate here in terms of them some basic formulae previously written for real coordinates:

$$\text{Element of separation: } ds^2 = \varepsilon\, dx_r dx_r \quad (\varepsilon = \pm 1) \tag{76}$$

$$\left.\begin{aligned} &\text{Past:} && dx_r dx_r < 0, && dx_4/i < 0, \\ &\text{Present:} && dx_r dx_r > 0, && \\ &\text{Future:} && dx_r dx_r < 0, && dx_4/i > 0 \\ &\text{Null cone:} && dx_r dx_r = 0. && \end{aligned}\right\} \tag{77}$$

World line of free material particle:

$$\frac{d^2x_r}{ds^2} = 0, \quad x_r = a_r s + b_r, \quad a_r a_r = -1. \tag{78}$$

World line of photon:

$$\frac{d^2x_r}{dw^2} = 0, \quad x_r = a_r w + b_r, \quad a_r a_r = 0. \tag{79}$$

Finite separation between y_r and z_r:

$$\Delta s = [\varepsilon(y_r - z_r)\,(y_r - z_r)]^{\frac{1}{2}}. \tag{80}$$

Unit vector:

$$V_r V_r = \varepsilon = \pm 1. \tag{81}$$

Orthogonality of vectors:

$$U_r V_r = 0, \quad dx_r \delta x_r = 0. \tag{82}$$

The simplicity of these formulae should show the efficiency of Minkowskian coordinates. It is true that the notation introduced at (21) is not much more complicated, but Minkowskian coordinates relieve us entirely from having to juggle with subscripts and superscripts.

CHAPTER III

SPACE-TIME DIAGRAMS

§ 1. SOME ELEMENTS OF THE GEOMETRY OF FLAT SPACE-TIME

In I–§ 6 we mentioned space-time diagrams and we have made some use of them. Since such diagrams play an important role in the quick understanding of relativity, and since the reader is likely to feel a sense of unreality in using them if they are treated too casually, we shall devote the present chapter to them. The difficulties are not unlike those which face the student when he begins to study solid geometry for the first time and is puzzled by the presence of the third dimension. Those difficulties can be overcome by the construction of models, a procedure which we cannot adopt in space-time, but we can learn much from the analogies which exist between space-time and our familiar Euclidean 3-space.

In Euclidean 3-space we have a line element

$$ds^2 = dx^2 + dy^2 + dz^2,$$

and in flat space-time we have a fundamental form Φ and a separation ds given by

$$\varepsilon\, ds^2 = \Phi = dx_r dx_r, \qquad (1)$$

if we use Minkowskian coordinates, with $x_4 = ict$. The Euclidean form is the sum of *three* squares and the form (1) is the sum of *four* squares, and this makes space-time more difficult to study than Euclidean 3-space. But there is another difference. The Euclidean form is *positive-definite*, whereas (1) is *indefinite*, since the fourth square is the square of a pure imaginary, and hence negative. In passing intuitively from Euclidean 3-space to space-time, these are the two differences which we have to allow for.

As a basis for the construction of space-time diagrams, we shall first discuss a few useful definitions and ideas.

A *3-flat* in space-time has the equation

$$A_r x_r + B = 0, \qquad (2)$$

where A_r and B are constants. A 3-flat is itself a 3-dimensional domain, since this single equation leaves x_r with three degrees of freedom.

A *2-flat* is the intersection of two 3-flats, and so has the two equations

$$A_r x_r + B = 0, \quad A'_r x_r + B' = 0. \tag{3}$$

If we solve these two equations for x_1 and x_2, they appear as linear functions of x_3 and x_4. If then for symmetry we introduce two parameters u, v, which are linear functions of x_3 and x_4, we may write the equations of a 2-flat in the parametric form

$$x_r = a_r u + b_r v + c_r, \tag{4}$$

where a_r, b_r, c_r are constants. It is best to take u and v to be real, the imaginary character of x_4 being taken care of by the imaginary character of the fourth components of a_r, b_r, c_r.

The word *hyperplane* is sometimes used for a 3-flat, and the word *plane* for a 2-flat. However in some ways the 3-flat has the closer analogy to the Euclidean plane, and we shall use the word *plane* indifferently for a 2-flat or a 3-flat when it is clear from the context which is meant.

A *1-flat* is the intersection of three 3-flats. It is most convenient to express it parametrically, as we expressed the 2-flat in (4). The parametric equation is

$$x_r = a_r u + b_r, \tag{5}$$

where u is the parameter and a_r, b_r constants. A 1-flat is in fact the same thing as a *straight line*, and so represents the history of a free material particle if $a_r a_r < 0$, and the history of a photon if $a_r a_r = 0$.

In Euclidean 3-space a plane passing through the origin has a unique straight line through the origin orthogonal to it. Similarly in space-time a 3-flat through the origin with equation $A_r x_r = 0$ has a unique orthogonal 1-flat through the origin, with equations $x_r = A_r u$, where u is a parameter.

A 2-flat through the origin with equations

$$A_r x_r = 0, \quad B_r x_r = 0, \tag{6}$$

has an infinity of 1-flats through the origin orthogonal to it, and these 1-flats (or straight lines) form the 2-flat with parametric equations

$$x_r = A_r u + B_r v. \tag{7}$$

Thus (6) and (7) represent *a pair of 2-flats orthogonal to one another.*

The existence of these orthogonal pairs, one member of the pair completely determined by the other, is a very important feature of four-dimensional geometry. We do not find such orthogonal pairs of like things in three dimensions, and to get the analogue we have to go back to the geometry of the plane, where we find straight lines occurring in such orthogonal pairs.

The condition that (7) should contain a null line passing through the origin is that the equation

$$A_r A_r u^2 + 2A_r B_r uv + B_r B_r v^2 = 0$$

should have real roots for u/v. We can thus classify 2-flats with equation (7) as follows:

$$\left.\begin{matrix}\text{timelike}\\ \text{null}\\ \text{spacelike}\end{matrix}\right\} \text{ if } (A_r B_r)^2 \left\{\begin{matrix}>\\ =\\ <\end{matrix}\right\} A_m A_m B_n B_n. \tag{8}$$

A timelike 2-flat cuts the null cone in two distinct straight lines, a null 2-flat touches it, and a spacelike 2-flat does not meet it at all except at the origin.

If we leave null 2-flats out of consideration, then it is easy to see that in any given pair of orthogonal 2-flats, one is timelike and the other spacelike.

It is worth noting that, in general, 2-flats through the origin have the origin as their only common event. For the four coordinates x_r of an event common to the two 2-flats have to satisfy four linear homogeneous equations, and the only solution is $x_r = 0$ unless the determinant of the four equations vanishes.

§ 2. ORTHOGONAL PROJECTIONS

An architect or engineer makes a plan of a 3-dimensional structure by projecting the structure on a horizontal plane by means of vertical lines. To get an elevation, he projects on to a vertical plane by lines at right angles to that plane. We shall adopt a similar procedure for space-time. Let us start by considering orthogonal projection on a straight line L with equations (5).

Let y_r be any event (Fig. 1). The orthogonal projection of y_r on L is that event z_r on L which is such that the vector $z_r - y_r$ is orthogonal to L. Now the direction of L is given by $dx_r/du = a_r$, and so z_r has to satisfy

$$z_r = a_r u + b_r, \quad (z_r - y_r)a_r = 0. \tag{9}$$

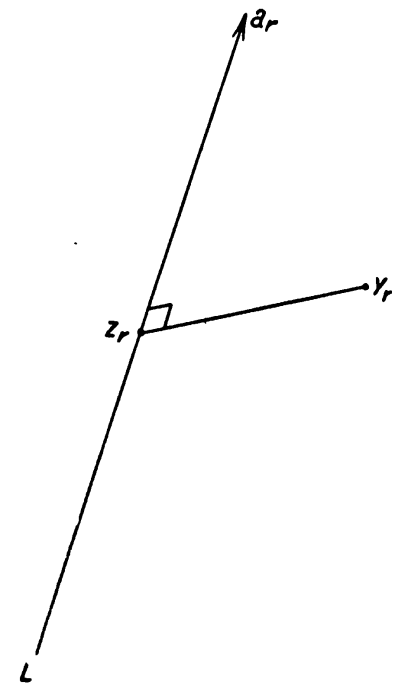


Fig. 1 – Orthogonal projection on a straight line

Substitution from the first in the second gives

$$ua_r a_r + (b_r - y_r)a_r = 0,$$

and if, as we shall suppose, L is not a null line (i.e. $a_r a_r \neq 0$), then this determines u uniquely, and z_r is then given uniquely by (9).

This means that all events in space-time can be *mapped* on any given straight line (not a null line) by the method of orthogonal projection. We can then take a sheet of paper, rule a straight line on it, mark off u according to some scale, and proceed to map any given set of events on this straight line which we have drawn on the paper. In some ways it is quite a good map, but it suffers from the fact that it is a 1-dimensional map of a 4-dimensional domain, and so a triple infinity of events are mapped on a single point of the line.

We can do better by projecting on a 2-flat instead of on a straight line. Let us take a 2-flat Π with the parametric equations (4) and an event y_r anywhere in space-time. We seek the orthogonal projection z_r of y_r on Π. This means that the vector $z_r - y_r$ is to be orthogonal to every direction in Π, and this is achieved by making it orthogonal to both $\partial x_r/\partial u$ and $\partial x_r/\partial v$, i.e. to a_r and to b_r. Thus we require that z_r shall satisfy the equations

$$z_r = a_r u + b_r v + c_r, \quad (z_r - y_r)a_r = 0, \quad (z_r - y_r)b_r = 0.$$

Substitution from the first in the others gives the following two equations for u and v:

$$\begin{aligned} a_r a_r u + a_r b_r v + a_r(c_r - y_r) &= 0, \\ b_r a_r u + b_r b_r v + b_r(c_r - y_r) &= 0. \end{aligned} \tag{10}$$

If, as we shall suppose, the determinant of the coefficients of u and v does not vanish, we get a unique pair of values for u, v, and hence a unique event z_r.

We can now take a sheet of paper, rule on it a pair of axes at right angles, label them u and v, and so map on the paper any event y_r

by plotting the point on the paper with coordinates u, v having the values given by (10). This *plane map* is much better than the linear map discussed above, because each point of the map has now to take care of only a double infinity of events, instead of a triple infinity.

§ 3. SPACE-TIME DIAGRAMS

We shall adopt plane maps for our *space-time diagrams*. It is true that we could project orthogonally on a 3-flat, and so obtain a 3-dimensional map of space-time, in which each point would have to take care of only a single infinity of events. Sometimes it is advisable to do this (or imagine it done). But on the whole it is a luxury we can dispense with, even as the architect or engineer dispenses with models, plans being so much easier to use. The plans and elevations of architects and engineers suffer from the fact that each point has to represent a whole line of points in space, but this does not bother them unduly. Neither need we worry too much that each point in one of our space-time diagrams represents actually a 2-flat in space-time.

We are not embarking on a programme of "graphical relativity". Our space-time diagrams are to be used as a mathematician or physicist uses rough sketches, rather than as an architect or engineer uses blueprints. The diagrams are to serve as guides for the mind. Anyone who studies relativity without understanding how to use simple space-time diagrams is as much inhibited as a student of functions of a complex variable who does not understand the Argand diagram.

The orthogonal projection of a set of events on to a 2-flat gives us a map of those events. That map will be changed if we use a different 2-flat, and it will also be changed if we use different parameters u, v in any 2-flat. It is good to have a standard practice, and we shall in general use the 2-flat which contains the axes of x_1 and x_4, i.e. the 2-flat with equation

$$x_2 = x_3 = 0.$$

In this 2-flat we shall take the parameters

$$u = x_1 = x, \quad v = x_4/i = ct, \tag{11}$$

it being most convenient to use real parameters. The projection of an event (x_1, x_2, x_3, x_4) is then the event $(x_1, 0, 0, x_4)$ and we plot in our diagram the real coordinates $(x = x_1, ct = x_4/i)$. Such a space-time diagram is shown in Fig. 2. Actually there is little loss of generality in

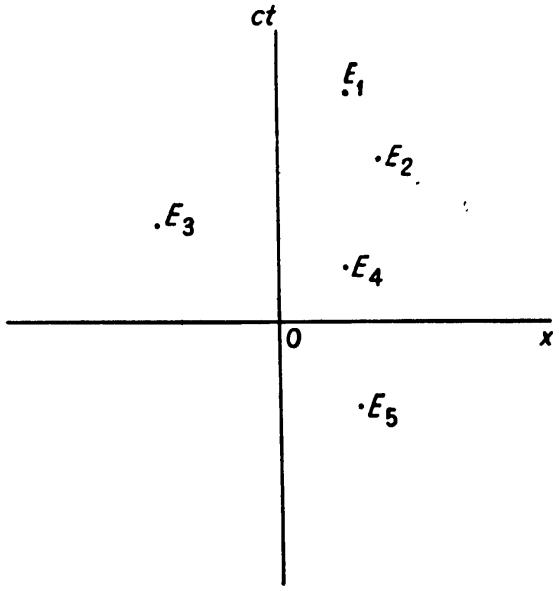


Fig. 2 – Space-time diagram of events

adopting this practice, since the axis of x_4 (i.e. the straight line along which $x_1 = x_2 = x_3 = 0$) may be any timelike straight line, and the axis of x_1 may be taken in any (spacelike) direction orthogonal to it. Thus if for some choice of axes we had an unfortunate coincidence, two important events falling on top of one another in the diagram, we could resolve this coincidence by changing to new axes.

A *curve* in space-time is mapped as a curve. A *straight line* is mapped as a straight line, unless we should be so unfortunate as to catch it end on, in which case it would be mapped as a point. The map of a 2-flat in general covers the whole space-time diagram, but it may be foreshortened into a straight line or into a single point.

§ 4. SPACE-TIME DIAGRAM OF THE NULL CONE

The null cone with vertex at the event b_r has the equation

$$(x_r - b_r)\,(x_r - b_r) = 0, \tag{12}$$

and if we take the vertex at the origin this becomes simply

$$x_r x_r = 0, \tag{13}$$

or equivalently in real explicit notation

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{14}$$

Let us project orthogonally on the (x, ct) plane. Then the event (x, y, z, ct) is mapped at (x, ct), no matter what y and z may be. To find the projections of those events which satisfy (14), we write it in the form

$$c^2t^2 - x^2 = y^2 + z^2 \geqslant 0, \tag{15}$$

and recognise that the map of the null cone covers that part of the plane for which $c^2t^2 - x^2$ is positive, but does not go outside that region. Now the equation

$$c^2t^2 - x^2 = 0 \tag{16}$$

gives us a pair of straight lines inclined at 45° to the axes (null lines), and so the null cone (14) maps on to those two infinite wedges bounded by these null lines which contain the ct-axis. This is the shaded region in Fig. 3, in which two conventional curves have been inserted to suggest the "cone".

A single infinity of events on the null cone map into a single point in the diagram; that is because, when x and ct are given, (14) leaves one degree of freedom in the choice of y and z.

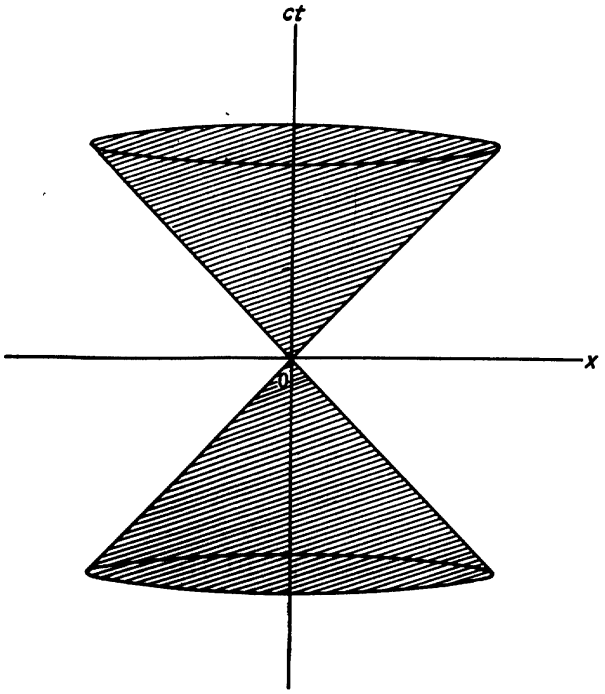


Fig. 3 – Space-time diagram of the nul cone with vertex at the origin

§ 5. M-GEOMETRY AND E-GEOMETRY

The geometry of space-time is Minkowskian geometry, or briefly *M-geometry*. On the other hand, when we have constructed a space-time diagram on a sheet of paper, we begin very naturally to think in terms of the ordinary geometry of the sheet of paper, that is, Euclidean geometry, or *E-geometry*. We must not confuse the two geometries, which are quite different, the M-geometry being based on the indefinite form (1) and the E-geometry on the usual positive-definite metric. Whenever there is any danger of confusion, we should be careful to prefix M or E to indicate which geometry we mean, speaking of M-distance or M-orthogonality when we are using (1) and E-distance or E-orthogonality when we are thinking in terms of the E-geometry of the diagram. Thus, when in § 4 we spoke of the null lines making an angle of 45° with the axes, that was an E-angle.

Vectors which are M-orthogonal are not in general E-orthogonal. Consider two vectors with components X_r, X'_r in Minkowskian coordinates, satisfying the condition of M-orthogonality

$$X_r X'_r = 0. \tag{17}$$

Let us represent these vectors by lines drawn in space-time from the origin to the points x_r, x'_r, and map them in a space-time diagram as the

points (x, ct), (x', ct'), where

$$x = x_1, \quad ct = x_4/i, \quad x' = x_1', \quad ct' = x_4'/i.$$

The condition (17) for M-orthogonality may be written

$$xx' + yy' + zz' - ct.ct' = 0. \tag{18}$$

Contrast this with the condition for E-orthogonality in the map:

$$xx' + ct.ct' = 0. \tag{19}$$

Suppose now for simplicity that the two vectors lie in the (x_1, x_4) 2-flat, so that $y = z = y' = z' = 0$. Then (18) becomes

$$xx' - ct.ct' = 0, \tag{20}$$

and this we recognise in the E-geometry of the space-time diagram as the condition that the vectors (x, ct), (x', ct') should be *conjugate* with respect to the conics

$$x^2 - c^2t^2 = \text{const.}, \tag{21}$$

or in particular with respect to the pair of lines (16). To represent such vectors faithfully in our diagram, we must show them making equal E-angles with the null lines, as in Fig. 4, in which the two M-orthogonal vectors appear as OA and OB.

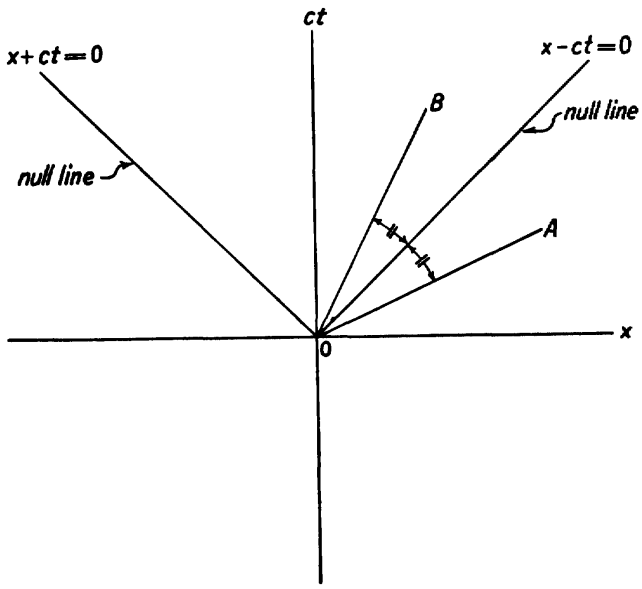


Fig. 4 – The vectors OA and OB are M-orthogonal, but not E-orthogonal

The null lines are M-orthogonal to themselves. That is, $x - ct = 0$ is M-orthogonal to itself, and so is $x + ct = 0$; but $x - ct = 0$ is *not* M-orthogonal to $x + ct = 0$, although they are obviously E-orthogonal.

The use of an E-metric in a space-time diagram suggests the use of an E-metric

$$ds_E^2 = dx^2 + dy^2 + dz^2 + c^2dt^2 \tag{22}$$

in 4-dimensional space-time. If we do this, then we can say that *any two vectors which are M-orthogonal are at the same time E-conjugate with respect to the null cone.* This follows from (14) and (18).

The use of (22) may help us to grapple with the difficulties presented to our intuition by the indefinite character of the metric in space-time. But of course it does not help us when we are confused by its 4-dimensionality. Then we have recourse to a space-time diagram, or to a set of such diagrams, one for the 2-flat (x, ct), another for (y, ct) and a third for (z, ct). It is also wise on occasion to project orthogonally on the 3-flat $t = 0$.

§ 6. PSEUDOSPHERES

The space-time analogue of the sphere in Euclidean space is the *pseudosphere*. If we assign a centre at the origin and a radius R, there are two pseudospheres with this centre and radius, with equations

$$x_r x_r = \varepsilon R^2, \tag{23}$$

with $\varepsilon = 1$ for one pseudosphere and $\varepsilon = -1$ for the other. If we choose $R = 1$, we have in explicit notation for the two unit pseudospheres

$$\begin{aligned} \Sigma&: x^2 + y^2 + z^2 - c^2t^2 = 1, \\ \Sigma'&: x^2 + y^2 + z^2 - c^2t^2 = -1, \end{aligned} \tag{24}$$

or equivalently

$$\begin{aligned} \Sigma&: x^2 - c^2t^2 = 1 - y^2 - z^2, \\ \Sigma'&: x^2 - c^2t^2 = -1 - y^2 - z^2. \end{aligned} \tag{25}$$

When mapped on the (x, ct) plane, the maps of Σ and Σ' are related to the E-hyperbolae

$$\begin{aligned} H&: x^2 - c^2t^2 = 1, \\ H'&: x^2 - c^2t^2 = -1. \end{aligned} \tag{26}$$

In fact, since by (25) we have

$$\begin{aligned} \Sigma&: x^2 - c^2t^2 < 1, \\ \Sigma'&: x^2 - c^2t^2 < -1, \end{aligned} \tag{27}$$

we see that Σ maps on to the region of the plane bounded by H and containing the origin, while Σ' maps on to the two portions of the plane bounded by H' and not containing the origin. These are indicated by the shading in Fig. 5.

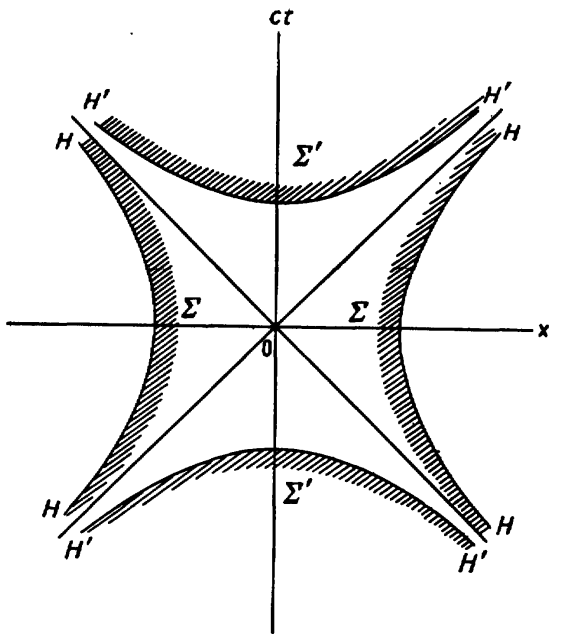


Fig. 5 – Space-time diagram of the unit pseudospheres Σ, Σ' mapped on regions bounded by the E-hyperbolae H, H'

It is easy to see that Σ is a connected 3-space but that Σ' consists of two sheets, disconnected from one another, one in the past and the other in the future relative to the origin of space-time. For Σ is the locus of the extremities of all spacelike unit vectors drawn from the origin, and Σ' the locus of the extremities of all timelike unit vectors drawn from the origin, the upper sheet corresponding to those drawn into the future and the lower sheet to those drawn into the past.

CHAPTER IV

THE LORENTZ TRANSFORMATION

§ 1. THE GENERAL LORENTZ TRANSFORMATION

If we use any arbitrary system of coordinates in space-time (each coordinate being either real or pure imaginary), the element of separation and the fundamental form are given by

$$\varepsilon\, ds^2 = \Phi = g_{mn} dx^m dx^n, \tag{1}$$

where g_{mn} are functions of the coordinates. It is only for special systems of coordinates that we have

$$\Phi = dx_r dx_r. \tag{2}$$

Such coordinates are the Minkowskian coordinates which we have been using, the first three real and the last a pure imaginary. In real coordinates, the equivalent simple form is

$$\Phi = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 - (dx^4)^2. \tag{3}$$

The transformation from one set of Minkowskian coordinates to another, or, equivalently, from one set of real coordinates as in (3) to another, is called a *Lorentz transformation*. These transformations are basic in the special theory of relativity, which might almost be called "the theory of the Lorentz transformation".

We shall discuss the Lorentz transformation for Minkowskian coordinates. From this we can at once pass, when necessary, to the Lorentz transformation for real coordinates as in (3).

We shall first show that the Lorentz transformation is *linear*. Let x and x' be two systems of Minkowskian coordinates. Then

$$dx'_r = \frac{\partial x'_r}{\partial x_s} dx_s, \tag{4}$$

and so

$$\Phi = dx'_r dx'_r = \frac{\partial x'_r}{\partial x_s} \frac{\partial x'_r}{\partial x_t} dx_s dx_t,$$

$$\Phi = dx_r dx_r = \delta_{st} dx_s dx_t. \tag{5}$$

Since these two expressions are to be equal for arbitrary values of the differentials, we have

$$\frac{\partial x'_r}{\partial x_s}\frac{\partial x'_r}{\partial x_t} = \delta_{st}. \tag{6}$$

Let us write for brevity

$$X^{r'}_s = \frac{\partial x'_r}{\partial x_s}, X^{r'}_{ps} = \frac{\partial^2 x'_r}{\partial x_p \partial x_s} = X^{r'}_{sp}. \tag{7}$$

Then differentiation of (6) with respect to x_p gives

$$X^{r'}_{ps}X^{r'}_t + X^{r'}_{pt}X^{r'}_s = 0. \tag{8}$$

A change of suffixes $p \to s$, $s \to t$, $t \to p$ makes this read

$$X^{r'}_{st}X^{r'}_p + X^{r'}_{sp}X^{r'}_t = 0;$$

subtracting (8) from this, we get

$$X^{r'}_{st}X^{r'}_p - X^{r'}_{tp}X^{r'}_s = 0.$$

Now interchange p and t, obtaining

$$X^{r'}_{sp}X^{r'}_t - X^{r'}_{pt}X^{r'}_s = 0.$$

If we add (8) to this and divide by 2, we have

$$X^{r'}_{ps}X^{r'}_t = 0. \tag{9}$$

Now by (6) we have

$$\det X^{r'}_s = \pm 1, \tag{10}$$

and so (9) implies

$$X^{r'}_{ps} = \frac{\partial^2 x'_r}{\partial x_p \partial x_s} = 0, \tag{11}$$

and thus it is proved that *the Lorentz transformation is linear.*

Let us then write the transformation as

$$x'_r = A_r + A_{rs}x_s, \tag{12}$$

where A_r and A_{rs} are constants. In view of the imaginary character of the fourth coordinates, x'_4 and x_4, A_r must be of the form (r, r, r, i) and A_{rs} of the form

$$\begin{pmatrix} r & r & r & i \\ r & r & r & i \\ r & r & r & i \\ i & i & i & r \end{pmatrix} \tag{13}$$

where r stands for "real" and i for "imaginary". There is of course no reason to suppose A_{rs} symmetric; in general $A_{rs} \neq A_{sr}$.

We have then

$$dx'_r = A_{rs}dx_s, \tag{14}$$

and *the conditions that the linear transformation* (12) *should be a Lorentz transformation are, as in* (6),

$$\boxed{A_{rs}A_{rt} = \delta_{st}.} \tag{15}$$

These conditions are both necessary and sufficient. They are formally the same as the conditions for an orthogonal transformation in Euclidean 4-space, but the presence of an imaginary coordinate makes Lorentz transformations very different from orthogonal transformations of real variables. This "orthogonal character" disappears if we use real coordinates as in (3).

Let A^*_{rs} be the cofactor of A_{rs} in det (A_{mn}), divided by that determinant. Then

$$A_{rs}A^*_{rt} = \delta_{st}. \tag{16}$$

These equations may be regarded as *defining* A^*_{rt}, and so, comparing them with (15), we see that

$$A^*_{rt} = A_{rt}. \tag{17}$$

If then we wish to solve (12) for x_r, and proceed in the usual way, multiplying by A^*_{rt}, we obtain, on making some changes in the suffixes,

$$x_r = A'_r + A_{sr}x'_s, \tag{18}$$

where

$$A'_r = -A_sA_{sr}. \tag{19}$$

We note that (18) is much the same as (12), A_r being replaced by A'_r and the suffixes of A_{rs} being interchanged.

The transformation (12) may be carried out in two steps:

$$(x \to x''): \; x''_r = A_{rs}x_s, \tag{20}$$

$$(x'' \to x'): \; x'_r = A_r + x''_r, \tag{21}$$

each of which is obviously a Lorentz transformation. Following an obvious analogy with change of axes in Euclidean 3-space, we call (20) a *rotation*[1] and (21) a *translation*. Thus *a Lorentz transformation may be broken down into a rotation followed by a translation*. But this

[1] The word "rotation", as used here, may include a reflection; cf. § 2.

order can be reversed. For the transformation (12) may also be broken down into the two steps:

$$(x \to x''): \; x''_r = -A'_r + x_r, \tag{22}$$

$$(x'' \to x'): \; x'_r = A_{rs} x''_s, \tag{23}$$

where A'_r is as in (19). (The proof of this requires (28) below.) Now the translation precedes the rotation.

We note that a Lorentz transformation has 4 degrees of freedom in respect of the translation and 6 degrees of freedom in respect of the rotation, the 16 constants A_{rs} being connected by the 10 relations (15).

Although we must not forget that a translation is included in the general Lorentz transformation, it plays a secondary role, and we shall concentrate on the rotational part, i.e. on Lorentz transformations of the form

$$x'_r = A_{rs} x_s, \quad A_{rs} A_{rt} = \delta_{st}. \tag{24}$$

It is a matter of taste whether one works with indicial notation or with matrices. In matrix notation we write

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \tag{25}$$

and denote the transpose of $\mathbf{A}$ by $\tilde{\mathbf{A}}$ ($\tilde{A}_{rs} = A_{sr}$).Then the conditions in (24) may be written

$$\tilde{\mathbf{A}}\mathbf{A} = \mathbf{1}, \tag{26}$$

where $\mathbf{1}$ is the unit matrix. The reciprocal matrix of $\mathbf{A}$ is therefore $\mathbf{A}^{-1} = \tilde{\mathbf{A}}$. Hence

$$\mathbf{A}\tilde{\mathbf{A}} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{1}. \tag{27}$$

Thus the conditions for a Lorentz transformation may also be expressed as

$$A_{sr} A_{tr} = \delta_{st}, \tag{28}$$

this differing from (24) in that the summation is now taken for the second suffixes. This result is also easily established in the indicial notation, for, since a determinant may be expanded either by rows or by columns, we have

$$A_{sr} A^*_{tr} = \delta_{st}, \tag{29}$$

as well as (16), and then (17) leads at once to (28). A matrix satisfying (26) is called an *orthogonal* matrix (cf. JEFFREYS and JEFFREYS [1946], p. 110).

§ 2. RESTRICTIONS ON LORENTZ TRANSFORMATIONS

From (26) we have $(\det \mathbf{A})^2 = 1$, and so $\det \mathbf{A} = \pm 1$, as in (10). We shall here consider only *proper* Lorentz transformations, for which

$$\det \mathbf{A} = 1. \tag{30}$$

This is not a serious restriction. If we have a Lorentz transformation $(x \to x')$ with $\det \mathbf{A} = -1$, then the transformation $(x \to x'')$ is a proper Lorentz transformation if x''_r are related to x'_r by

$$x''_1 = -x'_1,\ x''_2 = x'_2,\ x''_3 = x'_3,\ x''_4 = x'_4; \tag{31}$$

the change from x' to x'' amounts only to a reversal of the sense of one of the spatial axes. Just as a rotation in Euclidean 3-space preserves the orientation of the axes of coordinates, so a proper Lorentz transformation preserves the orientation of the axes, whereas this orientation is reversed if $\det \mathbf{A} = -1$ (see also § 9).

An advantage of using only proper Lorentz transformations is that, with respect to them, the permutation symbol ε_{mnrs} transforms as a tensor. This permutation symbol is defined by

$$\varepsilon_{mnrs} \begin{cases} = 0 \text{ if two suffixes are equal,} \\ = 1 \text{ if the sequence } mnrs \text{ is 1234, or an} \\ \quad \text{even permutation of 1234,} \\ = -1 \text{ if the sequence } mnrs \text{ is an odd} \\ \quad \text{permutation of 1234.} \end{cases} \tag{32}$$

There is a second restriction we shall impose. This will rule out a transformation such as

$$x'_1 = -x_1,\ x'_2 = x_2,\ x'_3 = x_3,\ x'_4 = -x_4. \tag{33}$$

which is a proper Lorentz transformation, but reverses the time-sense. For if the positive sense of the x_4-axis (defined by $dx_4/i > 0$) points into the future, then (33) tells us that the positive sense of the x'_4-axis points into the past. This we shall not allow, because it is desirable to have the positive sense of the axis of the fourth coordinate always pointing into the future. The restriction ensuring this is very simple. On the x'_4-axis we have $dx'_1 = dx'_2 = dx'_3 = 0$, and so by (18)

$$dx_4/i = A_{44}\, dx'_4/i. \tag{34}$$

To make dx_4/i and dx'_4/i both positive at the same time, we must have A_{44} positive.

To sum up, the restricted Lorentz transformations we shall study are such that

$$\begin{aligned} &x'_r = A_{rs}x_s, && x_r = A_{sr}x'_s, \\ &A_{rs}A_{rt} = \delta_{st}, && A_{sr}A_{tr} = \delta_{st}, \\ &\det \mathbf{A} = 1, && A_{44} > 0. \end{aligned} \tag{35}$$

We note here two simple Lorentz transformations, which are important because any Lorentz transformation (35) can be broken down into a succession of transformations of these types. With the second we shall be much concerned later. They are

$$\begin{aligned} x'_1 &= x_1 \cos\theta + x_2 \sin\theta, \\ x'_2 &= -x_1 \sin\theta + x_2 \cos\theta, \\ x'_3 &= x_3, \\ x'_4 &= x_4, \end{aligned} \tag{36}$$

and

$$\begin{aligned} x'_1 &= x_1, \\ x'_2 &= x_2, \\ x'_3 &= x_3 \cos i\chi + x_4 \sin i\chi, \\ x'_4 &= -x_3 \sin i\chi + x_4 \cos i\chi, \end{aligned} \tag{37}$$

θ and χ being any real constants. We recognise (36) as an ordinary plane rotation through an angle θ, and so, by analogy, we may regard (37) as a rotation through an imaginary angle $i\chi$. But this analogy must be handled with caution, and it is better to use the fact that

$$\begin{aligned} \cos i\chi &= \cosh\chi, \\ \sin i\chi &= i \sinh\chi, \end{aligned} \tag{38}$$

and rewrite (37) in the form

$$\begin{aligned} x'_1 &= x_1, \\ x'_2 &= x_2, \\ x'_3 &= x_3 \cosh\chi + ix_4 \sinh\chi, \\ x'_4 &= -ix_3 \sinh\chi + x_4 \cosh\chi. \end{aligned} \tag{39}$$

The real quantity χ is called a "pseudoangle".

One verifies immediately that (36) and (39) are Lorentz transformations in the general sense ($dx'_r dx'_r = dx_r dx_r$), and further that they both satisfy the restrictions in (35).

§ 3. THE TWO WAYS OF INTERPRETING TRANSFORMATIONS

If, in a Euclidean plane, we write

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta, \\ y' &= -x\sin\theta + y\cos\theta, \end{aligned} \tag{40}$$

we have a transformation which we can interpret in two different ways. First, we may think of fixed axes, and of a point which at first has coordinates (x, y) and later has coordinates (x', y'), all coordinates being measured with respect to the fixed axes. In this way of looking at the transformation (40), the *axes* are fixed and the *plane* rotates through an angle θ, in the sense from Oy towards Ox. Secondly, we may think of a fixed point, referred first to axes Oxy and secondly to axes $Ox'y'$, the axes $Ox'y'$ being obtained from Oxy by rotating them through an angle θ in the sense from Ox towards Oy. Now the *plane* is fixed and the *axes* rotate.

In such a simple case, where all the operations are easily followed intuitively, it matters little which view we adopt, and there is no considerable confusion in changing from the one to the other. But in space-time our intuition is not so active, and it is very easy to get confused if we mix the two analogous ways of looking at a Lorentz transformation. The first approach (axes fixed and events in space-time changed under the transformation) has been used extensively by E. Cartan and his school. The second approach (events fixed and coordinates changed) is that of the Princeton school of tensor calculus (Eisenhart and Veblen). The Cartan method is extremely powerful, but the other appeals more to anyone accustomed to thinking of the points of ordinary geometry as fixed things, the positions of which may be described by many different systems of coordinates. The traditional order in elementary mathematics (Euclid first and Descartes later) makes it easier for us to think of a transformation of coordinates than of a transformation of points.

In the general theory of relativity the second approach is commonly used; events are thought of as "fixed" things, described by many different systems of coordinates. And that is the best way to start in the special theory also, the Lorentz transformation being thought of as a transformation of coordinates rather than as a transformation of events. But as soon as we begin to think of the *axes* of Minkowskian coordinates, we recognise these axes as being themselves composed of

events, and so a transformation of coordinates (being essentially a transformation of axes) inevitably presents itself to us as a transformation of events. Thus in the end we use both viewpoints, just as we use both viewpoints in relation to the simple transformation (40). Sometimes one is simpler, sometimes the other; the great essential is to try to be quite clear which view we are taking in any particular argument, because otherwise great confusion may result.

§ 4. GEOMETRICAL MEANING OF THE LORENTZ TRANSFORMATION

Let x_r be a system of Minkowskian coordinates. The *origin* of space-time is the event $x_r = 0$, and the *coordinate axes* are those four straight lines through the origin along each of which three of the coordinates are zero. The positive senses are those in which x_1, x_2, x_3 and x_4/i respectively increase, the fourth axis pointing into the future. If we lay off unit vectors along the axes in the positive senses, then these four vectors have the following components:

$$\begin{aligned} &x_1\text{-axis: } (1,\ 0,\ 0,\ 0),\\ &x_2\text{-axis: } (0,\ 1,\ 0,\ 0),\\ &x_3\text{-axis: } (0,\ 0,\ 1,\ 0),\\ &x_4\text{-axis: } (0,\ 0,\ 0,\ i). \end{aligned} \tag{41}$$

These coordinate vectors are orthogonal to one another; if U_r and V_r are any two of them, then obviously

$$U_r V_r = 0. \tag{42}$$

Now let y_r be any other system of Minkowskian coordinates with the same origin. Let us write for the four vectors (41), when referred to the y-coordinates,

$$\begin{aligned} &x_1\text{-axis: } X_r^{(1)},\\ &x_2\text{-axis: } X_r^{(2)},\\ &x_3\text{-axis: } X_r^{(3)},\\ &x_4\text{-axis: } X_r^{(4)}. \end{aligned} \tag{43}$$

Then from the unit character of these vectors (the last being timelike) and from their orthogonality, we have

$$X_r^{(\varrho)} X_r^{(\sigma)} = \delta_{\varrho\sigma}, \quad X_r^{(\varrho)} X_r^{(4)} = 0, \quad X_r^{(4)} X_r^{(4)} = -1, \tag{44}$$

Greek suffixes having the values 1, 2, 3 as always.

The unit orthogonal tetrad (43) defines the x-coordinates in the

sense that if we are given y_r for an event we can determine x_r, provided this tetrad is given. To see this, we note that the position-vector y_r of the event can be resolved along the tetrad; let the resolution be

$$y_r = a_1 X_r^{(1)} + a_2 X_r^{(2)} + a_3 X_r^{(3)} + a_4 X_r^{(4)}. \tag{45}$$

This is a vector equation, true for all Minkowskian coordinate systems (y), and we recognise that the a's are invariants. Let us make the y-coordinate system coincide with the x-coordinate system. Then the components of the vectors (43) take the simple values (41), and, since now $y_r = x_r$, (45) gives

$$x_1 = a_1, \quad x_2 = a_2, \quad x_3 = a_3, \quad x_4 = ia_4. \tag{46}$$

But these quantities, as we have seen, are invariants with respect to transformations of the y's, and so, going back to a general Minkowskian coordinate system y_r as in (45), we have

$$\begin{aligned} y_r &= x_1 X_r^{(1)} + x_2 X_r^{(2)} + x_3 X_r^{(3)} - ix_4 X_r^{(4)} \\ &= x_\varrho X_r^{(\varrho)} - ix_4 X_r^{(4)}. \end{aligned} \tag{47}$$

Multiplying first by $X_r^{(\sigma)}$ and secondly by $X_r^{(4)}$ and using (44), we get

$$x_\sigma = y_r X_r^{(\sigma)}, \quad x_4 = -iy_r X_r^{(4)}, \tag{48}$$

thus expressing the x-coordinates of an event y_r by means of the orthogonal tetrad (43).

Now let x'_r be another system of Minkowskian coordinates, the Lorentz transformation $(x \to x')$ being, as in (35),

$$x'_r = A_{rs} x_s. \tag{49}$$

Just as the x-coordinates have a tetrad of coordinate vectors $X_r^{(s)}$ as in (43), so x'-coordinates have a tetrad $X_r^{(s)'}$, both sets being referred to a general Minkowskian coordinate system y_r, which gives us, so to speak, an impartial view of the two other coordinate systems. As in (47) we have for any event

$$y_r = x'_\sigma X_r^{(\sigma)'} - ix'_4 X_r^{(4)'}. \tag{50}$$

The transformation (49) may be regarded as a transformation of one coordinate tetrad into the other (this view we shall expand later). But it is certain that, since the coordinate systems are determined by their respective tetrads, the coefficients A_{rs} can depend only on the eight vectors $X_r^{(s)}$ and $X_r^{(s)'}$. Let us express A_{rs} in terms of these vectors.

Equating the two expressions (47) and (50) for y_r, we have

$$x'_\sigma X_r^{(\sigma)'} - ix'_4 X_r^{(4)'} = x_\sigma X_r^{(\sigma)} - ix_4 X_r^{(4)}. \tag{51}$$

Now multiply first by $X_r^{(\rho)'}$ and secondly $X_r^{(4)'}$, and use the relationships of the form (44) for the tetrad $X_r^{(s)'}$. We get

$$\begin{aligned} x'_\varrho &= x_\sigma X_r^{(\sigma)} X_r^{(\varrho)'} - ix_4 X_r^{(4)} X_r^{(\varrho)'}, \\ x'_4 &= -ix_\sigma X_r^{(\sigma)} X_r^{(4)'} - x_4 X_r^{(4)} X_r^{(4)'}. \end{aligned} \tag{52}$$

Comparison of this with (49) gives

$$\begin{aligned} A_{\varrho\sigma} &= X_r^{(\varrho)'} X_r^{(\sigma)}, & A_{\varrho 4} &= -iX_r^{(\varrho)'} X_r^{(4)}, \\ A_{4\sigma} &= -i X_r^{(4)'} X_r^{(\sigma)}, & A_{44} &= -X_r^{(4)'} X_r^{(4)}. \end{aligned} \tag{53}$$

We sum up as follows: *The Lorentz transformation (49) corresponds to a rotation of one orthogonal tetrad of space-time axes into another, and the coefficients in this transformation are given in terms of scalar products of the unit coordinate vectors by the formulae (53).*

The scalar product of two vectors is always real, and so we verify that the matrix A has the real and imaginary elements shown in (13).

Let us simplify the notation, suppressing suffixes and writing $\mathbf{V}$ for any vector V_r. The scalar product will be indicated by a dot:

$$\begin{aligned} U_r V_r &= \mathbf{U}.\mathbf{V}, \\ V_r V_r &= \mathbf{V}.\mathbf{V} = \mathbf{V}^2. \end{aligned} \tag{54}$$

Then, for a unit vector $\mathbf{V}$,

$$\begin{aligned} \mathbf{V}^2 &= 1 \text{ if } \mathbf{V} \text{ is spacelike}, \\ \mathbf{V}^2 &= -1 \text{ if } \mathbf{V} \text{ is timelike}. \end{aligned} \tag{55}$$

For the two tetrads of unit coordinate vectors let us write, instead of $X_r^{(s)}$ and $X_r^{(s)'}$,

$$\begin{aligned} &x\text{-coordinate vectors: } \mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}, \\ &x'\text{-coordinate vectors: } \mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}'. \end{aligned} \tag{56}$$

They satisfy, as in (44),

$$\begin{aligned} \mathbf{I}^2 = \mathbf{J}^2 = \mathbf{K}^2 = 1, \quad \mathbf{L}^2 = -1, \\ \mathbf{I}.\mathbf{J} = \mathbf{I}.\mathbf{K} = \mathbf{I}.\mathbf{L} = \mathbf{J}.\mathbf{K} = \mathbf{J}.\mathbf{L} = \mathbf{K}.\mathbf{L} = 0, \end{aligned} \tag{57}$$

with the same conditions for $\mathbf{I}'$, $\mathbf{J}'$, $\mathbf{K}'$, $\mathbf{L}'$.

Then (53) may be written in matrix form:

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} = \begin{pmatrix} \mathbf{I}'.\mathbf{I} & \mathbf{I}'.\mathbf{J} & \mathbf{I}'.\mathbf{K} & -i\mathbf{I}'.\mathbf{L} \\ \mathbf{J}'.\mathbf{I} & \mathbf{J}'.\mathbf{J} & \mathbf{J}'.\mathbf{K} & -i\mathbf{J}'.\mathbf{L} \\ \mathbf{K}'.\mathbf{I} & \mathbf{K}'.\mathbf{J} & \mathbf{K}'.\mathbf{K} & -i\mathbf{K}'.\mathbf{L} \\ -i\mathbf{L}'.\mathbf{I} & -i\mathbf{L}'.\mathbf{J} & -i\mathbf{L}'.\mathbf{K} & -\mathbf{L}'.\mathbf{L} \end{pmatrix} \tag{58}$$

§ 5. EULERIAN ANGLES AND PSEUDOANGLES

In the kinematics of a rigid body with a fixed point, the most useful scheme for describing the position of the body is that based on the Eulerian angles (cf. WHITTAKER [1937a], p. 9, SYNGE and GRIFFITH [1949], p. 288). In this scheme we transform one orthogonal triad into another in three steps, each step being a rotation about an axis.

Rotations in space-time differ from rotations in Euclidean 3-space, partly on account of its 4-dimensionality and partly on account of its indefinite metric. Nevertheless we can carry out a scheme closely analogous to the Eulerian scheme, using six steps instead of three, to transform one orthogonal tetrad into another.

In the notation of § 4, let **I**, **J**, **K**, **L** be any unit orthogonal tetrad, satisfying (57). If θ is any real number, the equations

$$\begin{aligned} \mathbf{I}' &= \mathbf{I}\cos\theta + \mathbf{J}\sin\theta, \\ \mathbf{J}' &= -\mathbf{I}\sin\theta + \mathbf{J}\cos\theta, \\ \mathbf{K}' &= \mathbf{K}, \\ \mathbf{L}' &= \mathbf{L}, \end{aligned} \tag{59}$$

define a new tetrad **I**′, **J**′, **K**′, **L**′, which is easily seen to be unit-orthogonal (i.e. it satisfies (57)). The passage

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$$

may be called a *spacelike rotation in the 2-flat* (**I**, **J**).

The position-vector of any arbitrary event may be written

$$\mathbf{E} = \alpha\mathbf{I} + \beta\mathbf{J} + \gamma\mathbf{K} + \delta\mathbf{L}, \tag{60}$$

where α, β, γ, δ are real numbers. If we hold these numbers fixed and change **I**, **J**, **K**, **L** to **I**′, **J**′, **K**′, **L**′ according to (59), we get a new vector **E**′. Thus (59) transforms all events in space-time. If in particular

$$\mathbf{E} = \gamma\mathbf{K} + \delta\mathbf{L}, \tag{61}$$

then **E** is unchanged by (59), and so we may call the 2-flat (**K**, **L**) the *axis* of the rotation (59). This "axis" is analogous to the fixed axis present in any rotation in Euclidean 3-space.

Similarly, if χ is any real number, the equations

$$\begin{aligned} \mathbf{I}' &= \mathbf{I}, \\ \mathbf{J}' &= \mathbf{J}, \\ \mathbf{K}' &= \mathbf{K}\cosh\chi + \mathbf{L}\sinh\chi, \\ \mathbf{L}' &= \mathbf{K}\sinh\chi + \mathbf{L}\cosh\chi, \end{aligned} \tag{62}$$

define a new unit orthogonal tetrad. This is a *timelike rotation in the 2-flat* (**K**, **L**), having (**I**, **J**) for *axis*.

The transformations (59), (62) should be compared with the Lorentz transformations (36), (39); the relationship between (39) and (62) will be recognised if (39) is regarded as connecting (x_3', ix_4') with (x_3, ix_4).

We are now going to carry out a transformation

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}') \tag{63}$$

from any unit orthogonal tetrad to any other, subject to the restrictions that (*a*) they have the same orientation (i.e. the corresponding Lorentz transformation is proper), and (*b*) **L** and **L**′ are both timelike and point into the future. The transformation is carried out in six steps, involving three timelike rotations of the type (62) through "pseudoangles" χ_1, χ_2, χ_3, and three spacelike rotations of the type (59) through "angles" θ_1, θ_2, θ_3. To shorten the writing, we shall put

$$\begin{aligned} C_\varrho &= \cosh \chi_\varrho, \quad S_\varrho = \sinh \chi_\varrho \\ c_\varrho &= \cos \theta_\varrho, \quad s_\varrho = \sin \theta_\varrho \end{aligned} \qquad (\varrho = 1, 2, 3) \tag{64}$$

The first three steps are the successive timelike rotations:

$$\begin{aligned} &(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}_1, \mathbf{J}_1, \mathbf{K}_1, \mathbf{L}_1)\text{: axis } (\mathbf{J}, \mathbf{K})\text{: pseudoangle } \chi_1; \\ &(\mathbf{I}_1, \mathbf{J}_1, \mathbf{K}_1, \mathbf{L}_1) \to (\mathbf{I}_2, \mathbf{J}_2, \mathbf{K}_2, \mathbf{L}_2)\text{: axis } (\mathbf{K}_1, \mathbf{I}_1)\text{: pseudoangle } \chi_2; \\ &(\mathbf{I}_2, \mathbf{J}_2, \mathbf{K}_2, \mathbf{L}_2) \to (\mathbf{I}_3, \mathbf{J}_3, \mathbf{K}_3, \mathbf{L}_3)\text{: axis } (\mathbf{I}_2, \mathbf{J}_2)\text{: pseudoangle } \chi_3. \end{aligned} \tag{65}$$

Note that $\mathbf{I}_1$ is *not* the first component of the vector **I**; it is itself a vector. Explicitly, these rotations are, as in (62),

$$\begin{aligned} &\mathbf{I}_1 = \mathbf{I}C_1 + \mathbf{L}S_1, \quad \mathbf{I}_2 = \mathbf{I}_1, && \mathbf{I}_3 = \mathbf{I}_2, \\ &\mathbf{J}_1 = \mathbf{J}, && \mathbf{J}_2 = \mathbf{J}_1C_2 + \mathbf{L}_1S_2, && \mathbf{J}_3 = \mathbf{J}_2, \\ &\mathbf{K}_1 = \mathbf{K}, && \mathbf{K}_2 = \mathbf{K}_1, && \mathbf{K}_3 = \mathbf{K}_2C_3 + \mathbf{L}_2S_3, \\ &\mathbf{L}_1 = \mathbf{I}S_1 + \mathbf{L}C_1, && \mathbf{L}_2 = \mathbf{J}_1S_2 + \mathbf{L}_1C_2, && \mathbf{L}_3 = \mathbf{K}_2S_3 + \mathbf{L}_2C_3. \end{aligned} \tag{66}$$

The result of these three rotations is

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}_3, \mathbf{J}_3, \mathbf{K}_3, \mathbf{L}_3), \tag{67}$$

where

$$\begin{aligned} \mathbf{I}_3 &= \mathbf{I}C_1 + \mathbf{L}S_1, \\ \mathbf{J}_3 &= \mathbf{I}S_1S_2 + \mathbf{J}C_2 + \mathbf{L}C_1S_2, \\ \mathbf{K}_3 &= \mathbf{I}S_1C_2S_3 + \mathbf{J}S_2S_3 + \mathbf{K}C_3 + \mathbf{L}C_1C_2S_3, \\ \mathbf{L}_3 &= \mathbf{I}S_1C_2C_3 + \mathbf{J}S_2C_3 + \mathbf{K}S_3 + \mathbf{L}C_1C_2C_3. \end{aligned} \tag{68}$$

We easily check that this is a unit orthogonal tetrad, cf. (57).

Let us now choose χ_1, χ_2, χ_3 so that $\mathbf{L}_3 = \mathbf{L}'$, the fourth vector of the second tetrad in (63). This means that, on account of the last of (68), we are to satisfy

$$\begin{aligned} S_1C_2C_3 &= \mathbf{L}'.\mathbf{I}, \\ S_2C_3 &= \mathbf{L}'.\mathbf{J}, \\ S_3 &= \mathbf{L}'.\mathbf{K}, \\ C_1C_2C_3 &= -\mathbf{L}'.\mathbf{L}. \end{aligned} \tag{69}$$

We can find χ_1, χ_2, χ_3, to satisfy the first three, and then the last is satisfied automatically by virtue of the identity (easily proved)

$$(\mathbf{L}'.\mathbf{L})^2 = 1 + (\mathbf{L}'.\mathbf{I})^2 + (\mathbf{L}'.\mathbf{J})^2 + (\mathbf{L}'.\mathbf{K})^2. \tag{70}$$

We have thus, in three rotations, brought $\mathbf{L}$ into coincidence with $\mathbf{L}'$. It remains to bring $\mathbf{I}_3$, $\mathbf{J}_3$, $\mathbf{K}_3$ into coincidence with $\mathbf{I}'$, $\mathbf{J}'$, $\mathbf{K}'$. Since both triads lie in the 3-flat orthogonal to $\mathbf{L}'$, this is merely a matter of a Euclidean rotation in 3-space, and that can be handled by means of the Eulerian angles. But we may as well carry out the whole job from the present point of view, and so we proceed to give spacelike rotations as follows:

$$\begin{aligned} &(\mathbf{I}_3, \mathbf{J}_3, \mathbf{K}_3, \mathbf{L}') \to (\mathbf{I}_4, \mathbf{J}_4, \mathbf{K}_4, \mathbf{L}')\text{: axis } (\mathbf{K}_3, \mathbf{L}')\text{: angle } \theta_1; \\ &(\mathbf{I}_4, \mathbf{J}_4, \mathbf{K}_4, \mathbf{L}') \to (\mathbf{I}_5, \mathbf{J}_5, \mathbf{K}_5, \mathbf{L}')\text{: axis } (\mathbf{J}_4, \mathbf{L}')\text{: angle } \theta_2. \end{aligned} \tag{71}$$

Taking the second rotation in the sense $\mathbf{K}_4 \to \mathbf{I}_4$ for positive θ_2, to agree with the usual convention for Eulerian angles, we have then ($\mathbf{L}'$ being unchanged)

$$\begin{aligned} \mathbf{I}_4 &= \mathbf{I}_3c_1 + \mathbf{J}_3s_1, & \mathbf{I}_5 &= \mathbf{I}_4c_2 - \mathbf{K}_4s_2, \\ \mathbf{J}_4 &= -\mathbf{I}_3s_1 + \mathbf{J}_3c_1, & \mathbf{J}_5 &= \mathbf{J}_4, \\ \mathbf{K}_4 &= \mathbf{K}_3, & \mathbf{K}_5 &= \mathbf{I}_4s_2 + \mathbf{K}_4c_2, \end{aligned} \tag{72}$$

and so

$$\begin{aligned} \mathbf{I}_5 &= \mathbf{I}_3c_1c_2 + \mathbf{J}_3s_1c_2 - \mathbf{K}_3s_2, \\ \mathbf{J}_5 &= -\mathbf{I}_3s_1 + \mathbf{J}_3c_1, \\ \mathbf{K}_5 &= \mathbf{I}_3c_1s_2 + \mathbf{J}_3s_1s_2 + \mathbf{K}_3c_2. \end{aligned} \tag{73}$$

We choose θ_1, θ_2 to make $\mathbf{K}_5 = \mathbf{K}'$, i.e. we satisfy

$$c_1s_2 = \mathbf{K}'.\mathbf{I}_3, \quad s_1s_2 = \mathbf{K}'.\mathbf{J}_3, \quad c_2 = \mathbf{K}'.\mathbf{K}_3, \tag{74}$$

which equations are easily seen to be consistent.

We have now made the transformation

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}_5, \mathbf{J}_5, \mathbf{K}', \mathbf{L}'), \tag{75}$$

and we complete the transformation (63) by a third spacelike rotation

$$(\mathbf{I}_5, \mathbf{J}_5, \mathbf{K}', \mathbf{L}') \to (\mathbf{I}_6, \mathbf{J}_6, \mathbf{K}', \mathbf{L}'): \text{axis } (\mathbf{K}', \mathbf{L}'): \text{angle } \theta_3, \tag{76}$$

which gives

$$\begin{aligned} \mathbf{I}_6 &= \mathbf{I}_5 c_3 + \mathbf{J}_5 s_3, \\ \mathbf{J}_6 &= -\mathbf{I}_5 s_3 + \mathbf{J}_5 c_3. \end{aligned} \tag{77}$$

We choose θ_3 to make $\mathbf{I}_6 = \mathbf{I}'$, $\mathbf{J}_6 = \mathbf{J}'$, i.e. we satisfy

$$\begin{aligned} c_3 &= \mathbf{I}'.\mathbf{I}_5, & s_3 &= \mathbf{I}'.\mathbf{J}_5, \\ s_3 &= -\mathbf{J}'.\mathbf{I}_5, & c_3 &= \mathbf{J}'.\mathbf{J}_5, \end{aligned} \tag{78}$$

these equations being consistent because $(\mathbf{I}', \mathbf{J}')$ and $(\mathbf{I}_5, \mathbf{J}_5)$ both lie in the 2-flat orthogonal to the pair $(\mathbf{K}', \mathbf{L}')$.

We have now completed the transformation (63). To find the coefficients A_{rs} of the corresponding Lorentz transformation, we have to form the matrix (58) of scalar products, and this is easy. By (77) and (73) we have

$$\begin{aligned} \mathbf{I}' = \mathbf{I}_6 &= \mathbf{I}_3(c_1c_2c_3 - s_1s_3) + \mathbf{J}_3(s_1c_2c_3 + c_1s_3) - \mathbf{K}_3 s_2 c_3, \\ \mathbf{J}' = \mathbf{J}_6 &= \mathbf{I}_3(-c_1c_2s_3 - s_1c_3) + \mathbf{J}_3(-s_1c_2s_3 + c_1c_3) + \mathbf{K}_3 s_2 s_3, \\ \mathbf{K}' = \mathbf{K}_5 &= \mathbf{I}_3 c_1 s_2 + \mathbf{J}_3 s_1 s_2 + \mathbf{K}_3 c_2, \end{aligned} \tag{79}$$

and so, for example, by (68),

$$\begin{aligned} \mathbf{I}'.\mathbf{I} &= \mathbf{I}_3.\mathbf{I}(c_1c_2c_3 - s_1s_3) + \mathbf{J}_3.\mathbf{I}(s_1c_2c_3 + c_1s_3) - \mathbf{K}_3.\mathbf{I} s_2 c_3 \\ &= C_1(c_1c_2c_3 - s_1s_3) + S_1S_2(s_1c_2c_3 + c_1s_3) - S_1C_2S_3 s_2 c_3. \end{aligned} \tag{80}$$

The complete table of the coefficients A_{rs} is then as follows, by (58).

TABLE I

Coefficients A_{rs} of Lorentz transformation $x'_r = A_{rs}x_s$ in terms of the Eulerian pseudoangles χ_1, χ_2, χ_3, and angles θ_1, θ_2, θ_3, in the notation of (64).

$A_{11} = \mathbf{I}'.\mathbf{I} =$ $C_1(c_1c_2c_3 - s_1s_3)$ $+ S_1S_2(s_1c_2c_3 + c_1s_3)$ $- S_1C_2S_3s_2c_3$	$A_{12} = \mathbf{I}'.\mathbf{J} =$ $C_2(s_1c_2c_3 + c_1s_3)$ $- S_2S_3s_2c_3$	$A_{13} = \mathbf{I}'.\mathbf{K} =$ $- C_3s_2c_3$	$A_{14} = -i\,\mathbf{I}'.\mathbf{L} =$ $i[S_1(c_1c_2c_3 - s_1s_3)$ $+ C_1S_2(s_1c_2c_3 + c_1s_3)$ $- C_1C_2S_3s_2c_3]$
$A_{21} = \mathbf{J}'.\mathbf{I} =$ $- C_1(c_1c_2s_3 + s_1c_3)$ $- S_1S_2(s_1c_2s_3 - c_1c_3)$ $+ S_1C_2S_3s_2s_3$	$A_{22} = \mathbf{J}'.\mathbf{J} =$ $- C_2(s_1c_2s_3 - c_1c_3)$ $+ S_2S_3s_2s_3$	$A_{23} = \mathbf{J}'.\mathbf{K} =$ $C_3s_2s_3$	$A_{24} = -i\,\mathbf{J}'.\mathbf{L} =$ $i[-S_1(c_1c_2s_3 + s_1c_3)$ $- C_1S_2(s_1c_2s_3 - c_1c_3)$ $+ C_1C_2S_3s_2s_3]$
$A_{31} = \mathbf{K}'.\mathbf{I} =$ $C_1c_1s_2 + S_1S_2s_1s_2$ $+ S_1C_2S_3c_2$	$A_{32} = \mathbf{K}'.\mathbf{J} =$ $C_2s_1s_2 + S_2S_3c_2$	$A_{33} = \mathbf{K}'.\mathbf{K} =$ C_3c_2	$A_{34} = -i\,\mathbf{K}'.\mathbf{L} =$ $i[S_1c_1s_2 + C_1S_2s_1s_2$ $+ C_1C_2S_3c_2]$
$A_{41} = -i\,\mathbf{L}'.\mathbf{I} =$ $-i\,S_1C_2C_3$	$A_{42} = -i\,\mathbf{L}'.\mathbf{J} =$ $-i\,S_2C_3$	$A_{43} = -i\,\mathbf{L}'.\mathbf{K} =$ $-iS_3$	$A_{44} = -\mathbf{L}'.\mathbf{L} =$ $C_1C_2C_3$

To pass from (**I**, **J**, **K**, **L**) to all orthogonal tetrads (**I**′, **J**′, **K**′, **L**′), with the same orientation and with **L** and **L**′ both timelike and pointing into the future, we give to the six parameters the ranges

$$\begin{aligned} &-\infty < \chi_1 < \infty, \quad -\infty < \chi_2 < \infty, \quad -\infty < \chi_3 < \infty, \\ &0 \leqslant \theta_1 < 2\pi, \qquad 0 \leqslant \theta_2 \leqslant \pi, \qquad 0 \leqslant \theta_3 < 2\pi. \end{aligned} \tag{81}$$

Fig. 1 shows the six steps by which the tetrad (**I**, **J**, **K**, **L**), indicated by points enclosed in single circles, is transformed into the tetrad (**I**′, **J**′, **K**′, **L**′), indicated by points enclosed in double circles. This is a symbolic diagram; if we wished, we could construct one to scale by projecting the extremities of the vectors on a 2-flat.

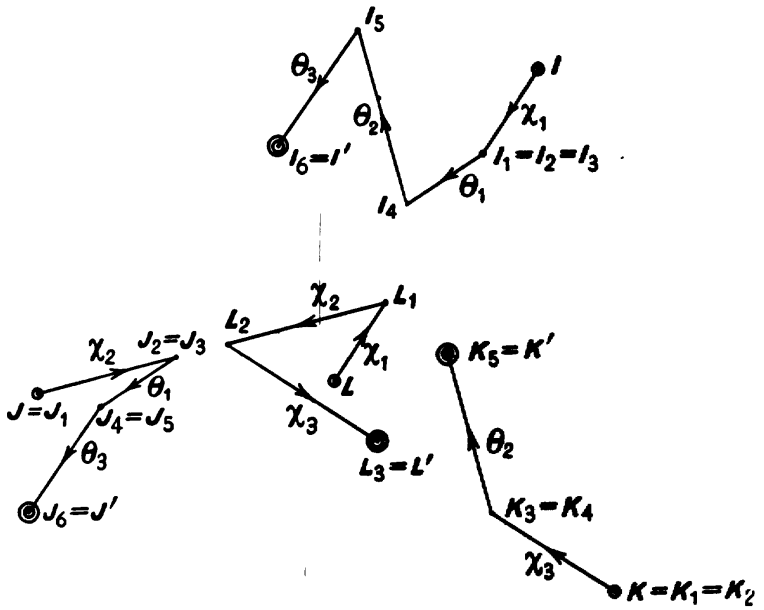


Fig. 1 – The six steps in the transformation $(I, J, K, L) \to (I', J', K', L')$

We have just seen how to transform a unit orthogonal tetrad (**I**, **J**, **K**, **L**) into another unit orthogonal tetrad (**I**′, **J**′, **K**′, **L**′). This transformation may be used to carry any event **E** into a new event **E**′; this is done by means of (60), in which we change (**I**, **J**, **K**, **L**) into (**I**′, **J**′, **K**′, **L**′) but retain *the same values* of $\alpha, \beta, \gamma, \delta$. Thus the whole of space-time is transformed into itself, and, when we look at it in this way, we think of a transformation of events rather than a transformation of coordinates. But if we give to A_{rs} the values shown in Table I, then

$$x_r' = A_{rs} x_s \tag{82}$$

is a Lorentz transformation of coordinates, (**I**, **J**, **K**, **L**) being the

coordinate vectors for the x-coordinates and $(\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$ for the x'-coordinates.

Thus the two viewpoints discussed in § 3 are present at the same time; we have either a transformation of events or a transformation of coordinates, according to the way we choose to look at it.

§ 6. LORENTZ TRANSFORMATIONS REGARDED AS RIGID BODY DISPLACEMENTS

Let ζ_r be some assigned system of Minkowskian coordinates in space-time. We shall regard the Lorentz transformation (82) as four equations which change an event $\zeta_r = x_r$ into an event $\zeta_r = x'_r$. This means that we are keeping the coordinate system fixed, and transforming events.

Let y_r change into y'_r and z_r into z'_r. Then

$$\begin{aligned} y'_r &= A_{rs}y_s, \quad z'_r = A_{rs}z_s, \\ y'_r - z'_r &= A_{rs}(y_s - z_s), \\ (y'_r - z'_r)\,(y'_r - z'_r) &= A_{rs}(y_s - z_s)A_{rt}(y_t - z_t) \\ &= \delta_{st}(y_s - z_s)\,(y_t - z_t) = (y_s - z_s)\,(y_s - z_s). \end{aligned} \tag{83}$$

Thus the Minkowskian separation between two events is unchanged by the Lorentz transformation. This we knew already, but it is well to emphasise it here, because we want to recognise a Lorentz transformation as a "rigid body transformation" or "distance-preserving transformation", "distance" being of course measured in this sense by Minkowskian separation. We do this in order to create analogies with the displacements of rigid bodies in Euclidean 3-space.

In Euclidean 3-space any rotation of a rigid body about a fixed point has the remarkable property that there exists an axis of rotation or line of fixed points (Euler's theorem). But in a general Lorentz transformation (82) it is not hard to see that there is no fixed event except the origin, and so the analogy fails. This is because Euclidean 3-space is of *odd* dimensionality and space-time of *even* dimensionality. The analogy would be better if we went back to rotations in the Euclidean plane, but then the situation is too simple to be of much help in suggesting analogues in space-time.

But if we think of a *free* rigid body in Euclidean 3-space, we get a better analogy with the Lorentz transformation (82). Each has six degrees of freedom. This suggests that there might exist for a Lorentz

transformation a reduction analogous to the reduction of a rigid body displacement to a *screw* (Chasles' theorem). In a screw there is one line (the axis of the screw) which transforms into itself as a whole. We shall see that any Lorentz transformation (82) possesses *two* straight lines in space-time (and only two) which have this property of transforming into themselves.

Before proceeding to discuss these lines, and the reduction of a Lorentz transformation to what we shall call a *4-screw*, let us develope some simple features of Lorentz transformations.

Lorentz transformations form a *group*. A succession of two Lorentz transformations

$$\begin{aligned} &(x \to x') && x'_r = A_{rs}x_s, \\ &(x' \to x'') && x''_r = A'_{rs}x'_s, \end{aligned} \tag{84}$$

gives the transformation

$$(x \to x'') \qquad x''_r = A'_{rs}A_{st}x_t = A''_{rt}x_t, \tag{85}$$

where

$$A''_{rt} = A'_{rs}A_{st}. \tag{86}$$

It is easy to see that $(x \to x'')$ is a Lorentz transformation. For

$$\begin{aligned} A''_{rt}A''_{rv} &= A'_{rs}A_{st}A'_{rq}A_{qv} \\ &= A'_{rs}A'_{rq}A_{st}A_{qv} \\ &= \delta_{sq}A_{st}A_{qv} \\ &= A_{st}A_{sv} = \delta_{tv}. \end{aligned} \tag{87}$$

But this is the condition (15) that $(x \to x'')$ is a Lorentz transformation, and so *a succession of two Lorentz transformations is itself a Lorentz transformation*. In the reduction (87) we made use of the conditions (15) for the two Lorentz transformations $(x \to x')$ and $(x' \to x'')$.

To complete the proof that Lorentz transformations form a group, we have to show that they contain the identity $x'_r = x_r$ and that each transformation $(x \to x')$ possesses a unique inverse $(x' \to x)$. If we take $A_{rs} = \delta_{rs}$ we get the identity and it is a Lorentz transformation because $\delta_{rs}\delta_{rt} = \delta_{st}$. As for the unique inverse, it is written in (35).

Just as infinitesimal rotations in Euclidean 3-space are much easier to think of than finite rotations, so infinitesimal Lorentz transformations are comparatively simple. An infinitesimal Lorentz transformation differs only infinitesimally from the identity, and so we write for its coefficients

$$A_{rs} = \delta_{rs} + \alpha_{rs}, \tag{88}$$

where α_{rs} are infinitesimals, the products of which are to be discarded. Substituting in the general conditions for a Lorentz transformation, viz.

$$A_{rs}A_{rt} = \delta_{st}, \tag{89}$$

we get

$$\delta_{rs}\delta_{rt} + \delta_{rs}\alpha_{rt} + \alpha_{rs}\delta_{rt} = \delta_{st},$$

or

$$\alpha_{st} + \alpha_{ts} = 0. \tag{90}$$

Thus *in an infinitesimal Lorentz transformation* (88) *the matrix* α_{rs} *is skew-symmetric.*

§ 7. THE LORENTZ 4-SCREW

We define a *Lorentz 4-screw* as follows: It is a Lorentz transformation consisting of a rotation in a timelike 2-flat Π, followed by (or preceded by — we shall see that it does not matter) a rotation in a spacelike 2-flat Π^*, *the 2-flats* Π *and* Π^* *being orthogonal to one another* [cf. III–(6) and (7) for the orthogonality of 2-flats].

To describe the 4-screw in symbols, let $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$ be a unit orthogonal tetrad ($\mathbf{L}$ timelike and pointing into the future), with $(\mathbf{K}, \mathbf{L})$ chosen in Π and $(\mathbf{I}, \mathbf{J})$ in Π^*. We have then the conditions as in (57):

$$\begin{gathered}\mathbf{I}^2 = \mathbf{J}^2 = \mathbf{K}^2 = 1, \ \mathbf{L}^2 = -1, \\ \mathbf{I}.\mathbf{J} = \mathbf{I}.\mathbf{K} = \mathbf{I}.\mathbf{L} = \mathbf{J}.\mathbf{K} = \mathbf{J}.\mathbf{L} = \mathbf{K}.\mathbf{L} = 0.\end{gathered} \tag{91}$$

Let us call the timelike rotation R and the spacelike rotation R^*, so that the operation of the 4-screw is R^*R, read from right to left. The rotation R gives the transformation

$$R\colon\ (\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}'', \mathbf{J}'', \mathbf{K}'', \mathbf{L}''), \tag{92}$$

where, as in (62),

$$\begin{aligned}\mathbf{I}'' &= \mathbf{I}, \\ \mathbf{J}'' &= \mathbf{J}, \\ \mathbf{K}'' &= \mathbf{K}\cosh\chi + \mathbf{L}\sinh\chi, \\ \mathbf{L}'' &= \mathbf{K}\sinh\chi + \mathbf{L}\cosh\chi,\end{aligned} \tag{93}$$

and R^* then gives the transformation

$$R^*\colon\ (\mathbf{I}'', \mathbf{J}'', \mathbf{K}'', \mathbf{L}'') \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}'), \tag{94}$$

where, as in (59),

$$\begin{aligned} \mathbf{I}' &= \mathbf{I}'' \cos\theta + \mathbf{J}'' \sin\theta, \\ \mathbf{J}' &= -\mathbf{I}'' \sin\theta + \mathbf{J}'' \cos\theta, \\ \mathbf{K}' &= \mathbf{K}'', \\ \mathbf{L}' &= \mathbf{L}''. \end{aligned} \tag{95}$$

Here χ is the pseudoangle of R and θ the angle of R^*.

The result is the transformation

$$R^*R\colon (\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}'), \tag{96}$$

where

$$\begin{aligned} \mathbf{I}' &= \mathbf{I}\cos\theta + \mathbf{J}\sin\theta, \\ \mathbf{J}' &= -\mathbf{I}\sin\theta + \mathbf{J}\cos\theta, \\ \mathbf{K}' &= \mathbf{K}\cosh\chi + \mathbf{L}\sinh\chi, \\ \mathbf{L}' &= \mathbf{K}\sinh\chi + \mathbf{L}\cosh\chi. \end{aligned} \tag{97}$$

It is clear that we get the same final result if we take the rotations in the opposite order. Thus

$$R^*R = RR^*; \tag{98}$$

in fact, the operations R and R^* *commute.*

In Euclidean 3-space a translation and a rotation do not in general commute; they do commute if they are the translation and rotation of a screw. In space-time a timelike rotation and a spacelike rotation do not in general commute; they do commute if their 2-flats are orthogonal, i.e. if they form a 4-screw. Thus we see a good analogy between the ordinary screw and the 4-screw. We shall now come to what, for the 4-screw, is the analogue of the axis of the Euclidean screw.

If we write, as in (60),

$$\begin{aligned} \mathbf{E} &= \alpha\mathbf{I} + \beta\mathbf{J} + \gamma\mathbf{K} + \delta\mathbf{L}, \\ \mathbf{E}' &= \alpha\mathbf{I}' + \beta\mathbf{J}' + \gamma\mathbf{K}' + \delta\mathbf{L}', \end{aligned} \tag{99}$$

the coefficients being the same in both equations, then the 4-screw (97) gives us a transformation $\mathbf{E} \to \mathbf{E}'$ of the whole of space-time into itself. The holding fixed of the coefficients in (99) means that the vector $\mathbf{E}$ is "rigidly attached" to the tetrad $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$, as it must be under a Lorentz transformation.

It is clear that the 4-screw transforms the 2-flat Π into itself as a whole, and it also transforms Π^* into itself as a whole. Let us now take

two vectors in Π, namely,

$$\mathbf{M} = p(\mathbf{K} + \mathbf{L}), \quad \mathbf{N} = q(-\mathbf{K} + \mathbf{L}), \tag{100}$$

where p and q are positive real numbers. Then, by virtue of (91),

$$\mathbf{M}^2 = 0, \quad \mathbf{N}^2 = 0, \quad \mathbf{M}.\mathbf{N} = -2pq. \tag{101}$$

Hence **M** and **N** are null vectors, and it is easy to see that they both point into the future, since p and q are positive.

The 4-screw (97) transforms **M** and **N** into

$$\begin{aligned} \mathbf{M} \to \mathbf{M}' &= p(\mathbf{K}' + \mathbf{L}') = p(\mathbf{K} + \mathbf{L})e^{\chi} = \mathbf{M}e^{\chi}, \\ \mathbf{N} \to \mathbf{N}' &= q(-\mathbf{K}' + \mathbf{L}') = q(-\mathbf{K} + \mathbf{L})e^{-\chi} = \mathbf{N}e^{-\chi}. \end{aligned} \tag{102}$$

Thus the null vectors **M** *and* **N** *have the remarkable property that they retain their directions under application of the 4-screw, the components of* **M** *being multiplied by* e^{χ} *and the components of* **N** *by* $e^{-\chi}$.

By giving to p in (100) all positive values, the extremity of **M** traces out a *null ray*, i.e. that half of a complete null line which runs from the origin into the future (we shall always understand *null ray* in that sense). Similarly, by giving all positive values to q, the extremity of **N** traces out a second null ray. It follows from (102) that *each of these null rays is transformed into itself as a whole by the 4-screw, the events on one of these null rays being pulled out and the events on the other being pushed in.* We shall call these two null rays the *axial null rays* of the 4-screw (Fig. 2).

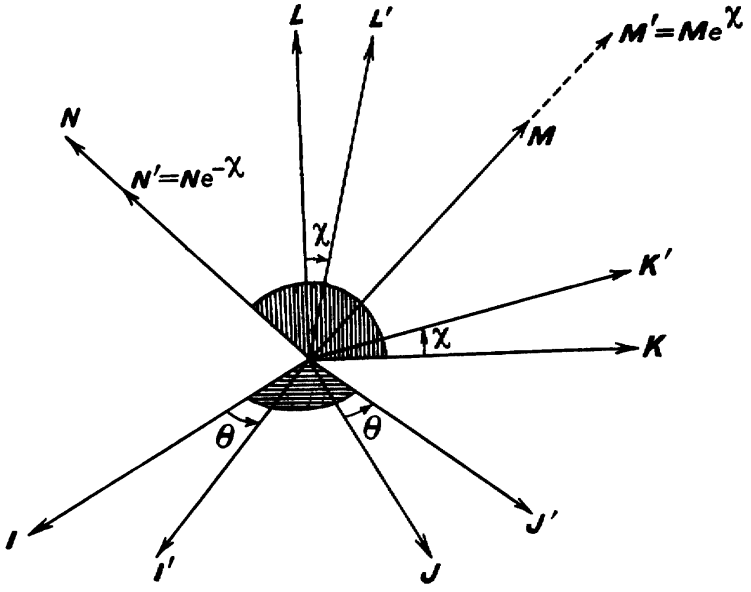


Fig. 2 – A 4-screw and its axial null rays. The two orthogonal 2-flats are indicated by the shading

A 4-screw is determined by two orthogonal 2-flats, Π and Π^*, and the two scalars, χ and θ, for the transformation (97) is the same no matter what orthogonal pair (**K**, **L**) we select in Π and no matter what orthogonal pair (**I**, **J**) we select in Π^* (provided of course that **L** is always timelike, pointing into the future, and the orientation of the tetrad is not reversed). Now Π is determined by the two null rays in which it cuts the null cone, and Π^* is determined when Π is known. Hence we know that, if we are given its axial null rays and the scalars χ and θ, a 4-screw is completely determined.

We shall now prove that *if, as the result of a Lorentz transformation T, each of a pair of null rays is transformed into itself as a whole, then T is a 4-screw having those null rays for axial null rays.*

To show this, take null vectors **M** and **N** on the two null rays. Put

$$\mathbf{K} = p(\mathbf{M} - \mathbf{N}), \quad \mathbf{L} = q(\mathbf{M} + \mathbf{N}). \tag{103}$$

Then

$$\mathbf{K}^2 = -2p^2\mathbf{M}.\mathbf{N}, \quad \mathbf{L}^2 = 2q^2\mathbf{M}.\mathbf{N}, \quad \mathbf{K}.\mathbf{L} = 0. \tag{104}$$

Now, for any positive values of p and q, it is clear that $p\mathbf{M} + q\mathbf{N}$ is timelike (**M** and **N** both pointing into the future), and so

$$0 > (p\mathbf{M} + q\mathbf{N}).(p\mathbf{M} + q\mathbf{N}) = 2pq\mathbf{M}.\mathbf{N}. \tag{105}$$

Therefore

$$\mathbf{M}.\mathbf{N} < 0; \tag{106}$$

in fact, *the scalar product of two distinct null vectors, both pointing into the future, is negative.* Let us then choose

$$p = q = (-2\mathbf{M}.\mathbf{N})^{-\frac{1}{2}}, \tag{107}$$

so that, by (104),

$$\mathbf{K}^2 = 1, \quad \mathbf{L}^2 = -1, \quad \mathbf{K}.\mathbf{L} = 0; \tag{108}$$

then (**K**, **L**) is a pair of orthogonal unit vectors, with **L** timelike; it is easy to see that **L** points into the future.

Now since, by hypothesis, the two null rays transform into themselves, the Lorentz transformation T must give

$$\mathbf{M} \to \mathbf{M}' = k\mathbf{M}, \quad \mathbf{N} \to \mathbf{N}' = k'\mathbf{N}, \tag{109}$$

where k and k' are positive. But scalar products are conserved under

a Lorentz transformation, and so

$$\mathbf{M}.\mathbf{N} = \mathbf{M}'.\mathbf{N}' = kk'\mathbf{M}.\mathbf{N}, \quad kk' = 1. \tag{110}$$

We can therefore write (109) as

$$\mathbf{M}' = k\mathbf{M}, \quad \mathbf{N}' = k^{-1}\mathbf{N}. \tag{111}$$

Then

$$\begin{aligned}
&\mathbf{K} = p(\mathbf{M} - \mathbf{N}), \quad \mathbf{L} = p(\mathbf{M} + \mathbf{N}),\\
&\mathbf{M} = \frac{1}{2p}(\mathbf{K} + \mathbf{L}), \quad \mathbf{N} = \frac{1}{2p}(-\mathbf{K} + \mathbf{L}),\\
&\mathbf{K}' = p(\mathbf{M}' - \mathbf{N}') = p(k\mathbf{M} - k^{-1}\mathbf{N})\\
&\qquad = \tfrac{1}{2}(k + k^{-1})\mathbf{K} + \tfrac{1}{2}(k - k^{-1})\mathbf{L},\\
&\mathbf{L}' = p(\mathbf{M}' + \mathbf{N}') = p(k\mathbf{M} + k^{-1}\mathbf{N})\\
&\qquad = \tfrac{1}{2}(k - k^{-1})\mathbf{K} + \tfrac{1}{2}(k + k^{-1})\mathbf{L},
\end{aligned} \tag{112}$$

and so

$$\begin{aligned}
\mathbf{K}' &= \mathbf{K}\cosh\chi + \mathbf{L}\sinh\chi,\\
\mathbf{L}' &= \mathbf{K}\sinh\chi + \mathbf{L}\cosh\chi,
\end{aligned} \tag{113}$$

where $\chi = \log k$.

If $(\mathbf{I}, \mathbf{J})$ make up an orthogonal tetrad with $(\mathbf{K}, \mathbf{L})$, then the transformed pair $(\mathbf{I}', \mathbf{J}')$ must lie in the same 2-flat as $(\mathbf{I}, \mathbf{J})$, since, by (113) the 2-flats $(\mathbf{K}, \mathbf{L})$ and $(\mathbf{K}', \mathbf{L}')$ are the same. It follows that the transformation T which carries $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$ into $(\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$ must be of the form (97). Thus T is a 4-screw, and it follows from the second line of (112) that its axial null rays lie along $\mathbf{M}$ and $\mathbf{N}$.

The italicised statement preceding (103) is therefore proved. Its significance is that, if we know that a certain Lorentz transformation transforms two null rays into themselves, then we know that it is a 4-screw. This result is a step towards the major theorem of the next section.

§ 8. REDUCTION OF ANY LORENTZ TRANSFORMATION TO A 4-SCREW

Let us apply a 4-screw R^*R and *follow* it with an *infinitesimal* Lorentz transformation T, so that the complete operation is TR^*R, read from the right. This is of course a Lorentz transformation. Our purpose is to show that it can be reduced to a 4-screw, i.e. to prove the existence of another 4-screw $\bar{R}^*\bar{R}$, differing infinitesimally from R^*R, such that

$$\bar{R}^*\bar{R} = TR^*R. \tag{114}$$

The operation R^*R effects the transformation

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}'), \tag{115}$$

as in (97); and then T gives the transformation

$$(\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}') \to (\mathbf{I}'', \mathbf{J}'', \mathbf{K}'', \mathbf{L}''), \tag{116}$$

where

$$\begin{aligned}
\mathbf{I}'' &= \mathbf{I}' + b_{11}\mathbf{I}' + b_{12}\mathbf{J}' + b_{13}\mathbf{K}' + b_{14}\mathbf{L}', \\
\mathbf{J}'' &= \mathbf{J}' + b_{21}\mathbf{I}' + b_{22}\mathbf{J}' + b_{23}\mathbf{K}' + b_{24}\mathbf{L}', \\
\mathbf{K}'' &= \mathbf{K}' + b_{31}\mathbf{I}' + b_{32}\mathbf{J}' + b_{33}\mathbf{K}' + b_{34}\mathbf{L}', \\
\mathbf{L}'' &= \mathbf{L}' + b_{41}\mathbf{I}' + b_{42}\mathbf{J}' + b_{43}\mathbf{K}' + b_{44}\mathbf{L}',
\end{aligned} \tag{117}$$

the b's being infinitesimal. Since this is a Lorentz transformation, $(\mathbf{I}'', \mathbf{J}'', \mathbf{K}'', \mathbf{L}'')$ is a unit orthogonal tetrad, satisfying conditions as in (91), and so (since the products of b's are to be discarded), we are led immediately to the conditions

$$b_{\rho\sigma} = -b_{\sigma\rho}, \quad b_{\rho 4} = b_{4\rho}, \quad b_{44} = 0. \tag{118}$$

Note the formal difference between these conditions and (90); they are different ways of saying the same thing.

Combining (115) and (116), we get the transformation

$$TR^*R\colon\ (\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}'', \mathbf{J}'', \mathbf{K}'', \mathbf{L}''). \tag{119}$$

We now ask whether this transformation leaves some pair of null rays unchanged. Since the 4-screw R^*R certainly does do this, we naturally investigate null rays adjacent to the axial null rays of R^*R. Let $\mathbf{M}$ and $\mathbf{N}$ be vectors along these axial null rays, given as in (100), with $p = q = 1$ for convenience, by the formulae

$$\mathbf{M} = \mathbf{K} + \mathbf{L}, \quad \mathbf{N} = -\mathbf{K} + \mathbf{L}, \tag{120}$$

so that

$$\mathbf{M}.\mathbf{N} = -2. \tag{121}$$

Let us see what TR^*R does to a pair of null vectors of the form

$$\overline{\mathbf{M}} = \mathbf{M} + \mathbf{m}, \quad \overline{\mathbf{N}} = \mathbf{N} + \mathbf{n}, \tag{122}$$

$\mathbf{m}$ and $\mathbf{n}$ being infinitesimal vectors.

Since $\mathbf{M}$ and $\mathbf{N}$ are orthogonal to $\mathbf{I}$ and $\mathbf{J}$, we ensure the null character of $\overline{\mathbf{M}}$ and $\overline{\mathbf{N}}$ (quantities of the second order being discarded)

by taking $\mathbf{m}$ and $\mathbf{n}$ in the 2-flat $(\mathbf{I}, \mathbf{J})$, so that we have

$$\begin{aligned} \overline{\mathbf{M}} &= \alpha\mathbf{I} + \beta\mathbf{J} + \mathbf{M}, \\ \overline{\mathbf{N}} &= \gamma\mathbf{I} + \delta\mathbf{J} + \mathbf{N}, \end{aligned} \tag{123}$$

$\alpha, \beta, \gamma, \delta$ being any infinitesimal scalars. We shall try to choose these scalars so that $\overline{\mathbf{M}}$ and $\overline{\mathbf{N}}$ retain their directions under the transformation TR^*R.

Under application of TR^*R we get

$$\begin{aligned} \overline{\mathbf{M}} \to \overline{\mathbf{M}}'' &= \alpha\mathbf{I}'' + \beta\mathbf{J}'' + \mathbf{M}'', \\ \overline{\mathbf{N}} \to \overline{\mathbf{N}}'' &= \gamma\mathbf{I}'' + \delta\mathbf{J}'' + \mathbf{N}'', \end{aligned} \tag{124}$$

where, by (117) and (118),

$$\begin{aligned} \mathbf{M}'' &= \mathbf{K}'' + \mathbf{L}'' \\ &= \mathbf{K}' + \mathbf{L}' + (b_{41} + b_{31})\mathbf{I}' + (b_{42} + b_{32})\mathbf{J}' + b_{34}(\mathbf{K}' + \mathbf{L}') \\ &= (b_{41} + b_{31})\mathbf{I}' + (b_{42} + b_{32})\mathbf{J}' + (1 + b_{34})\mathbf{M}', \\ \mathbf{N}'' &= -\mathbf{K}'' + \mathbf{L}'' = (b_{41} - b_{31})\mathbf{I}' + (b_{42} - b_{32})\mathbf{J}' + (1 - b_{34})\mathbf{N}'. \end{aligned} \tag{125}$$

Since $\mathbf{I}''$, $\mathbf{J}''$ differ only infinitesimally from $\mathbf{I}'$, $\mathbf{J}'$, (124) gives

$$\begin{aligned} \overline{\mathbf{M}}'' &= \alpha'\mathbf{I}' + \beta'\mathbf{J}' + (1 + b_{34})\mathbf{M}', \\ \overline{\mathbf{N}}'' &= \gamma'\mathbf{I}' + \delta'\mathbf{J}' + (1 - b_{34})\mathbf{N}', \end{aligned} \tag{126}$$

where we have put for brevity

$$\begin{aligned} \alpha' &= \alpha + b_{41} + b_{31}, \quad \beta' = \beta + b_{42} + b_{32}, \\ \gamma' &= \gamma + b_{41} - b_{31}, \quad \delta' = \delta + b_{42} - b_{32}. \end{aligned} \tag{127}$$

Let us further put

$$\cos\theta = c, \quad \sin\theta = s. \tag{128}$$

Then, since R^*R is as in (97) and (102) holds, we have from (126)

$$\begin{aligned} \overline{\mathbf{M}}'' &= (\alpha'c - \beta's)\mathbf{I} + (\alpha's + \beta'c)\mathbf{J} + (1 + b_{34})e^{\chi}\mathbf{M}, \\ \overline{\mathbf{N}}'' &= (\gamma'c - \delta's)\mathbf{I} + (\gamma's + \delta'c)\mathbf{J} + (1 - b_{34})e^{-\chi}\mathbf{N}. \end{aligned} \tag{129}$$

These two vectors will have their original directions (those of $\overline{\mathbf{M}}$ and $\overline{\mathbf{N}}$) if a positive constant k exists such that

$$\overline{\mathbf{M}}'' = k\overline{\mathbf{M}}, \quad \overline{\mathbf{N}}'' = k^{-1}\overline{\mathbf{N}}. \tag{130}$$

If T were absent, so that we had merely the 4-screw, then $k = e^{\chi}$,

and so we naturally try to satisfy (130) by taking

$$k = (1 + \lambda)e^{\chi}, \quad k^{-1} = (1 - \lambda)e^{-\chi}, \tag{131}$$

λ being infinitesimal. Then by (123), to the first order,

$$\begin{aligned} k\overline{\mathbf{M}} &= e^{\chi}[\alpha\mathbf{I} + \beta\mathbf{J} + (1 + \lambda)\mathbf{M}], \\ k^{-1}\overline{\mathbf{N}} &= e^{-\chi}[\gamma\mathbf{I} + \delta\mathbf{J} + (1 - \lambda)\mathbf{N}]. \end{aligned} \tag{132}$$

Thus we have to find $\alpha, \beta, \gamma, \delta, \lambda$ to satisfy

$$\begin{aligned} &e^{\chi}[\alpha\mathbf{I} + \beta\mathbf{J} + (1 + \lambda)\mathbf{M}] \\ &\qquad = (\alpha' c - \beta' s)\mathbf{I} + (\alpha' s + \beta' c)\mathbf{J} + (1 + b_{34})e^{\chi}\,\mathbf{M}, \\ &e^{-\chi}[\gamma\mathbf{I} + \delta\mathbf{J} + (1 - \lambda)\mathbf{N}] \\ &\qquad = (\gamma' c - \delta' s)\mathbf{I} + (\gamma' s + \delta' c)\mathbf{J} + (1 - b_{34})e^{-\chi}\,\mathbf{N}. \end{aligned} \tag{133}$$

Taking scalar products by $\mathbf{I}$, $\mathbf{J}$, $\mathbf{M}$, $\mathbf{N}$, we get the following equivalent equations to be satisfied:

$$\alpha e^{\chi} = \alpha' c - \beta' s, \quad \beta e^{\chi} = \alpha' s + \beta' c, \tag{134}$$

$$\gamma e^{-\chi} = \gamma' c - \delta' s, \quad \delta e^{-\chi} = \gamma' s + \delta' c, \tag{135}$$

$$\lambda = b_{34}. \tag{136}$$

Now (134) is a pair of linear equations for α and β with determinant

$$(e^{\chi} - c)^2 + s^2,$$

which vanishes only in the trivial case where $\theta = \chi = 0$; we are not concerned with this, for then the 4-screw disappears. The equations (135) are similar. Hence we get unique values for $\alpha, \beta, \gamma, \delta, \lambda$ satisfying (133).

Thus the transformation TR^*R leaves a pair of null rays unchanged, and so by § 7 *we know that TR^*R is a 4-screw.* In fact, we have established the existence of $\bar{R}^*\bar{R}$ in (114), T being any *infinitesimal* Lorentz transformation.

Consider now any *finite* Lorentz transformation T, proper and future-preserving. As in § 5, it can be broken down into a succession of six rotations,

$$T = R_6 R_5 R_4 R_3 R_2 R_1. \tag{137}$$

Each of these finite rotations may be broken down into an infinite sequence of infinitesimal rotations, and so the whole transformation

T may be regarded as the successive applications of infinitesimal Lorentz transformations. Suppose R_1 has been carried out; this is a particular case of a 4-screw. By the result established above, the property of being equivalent to a 4-screw is preserved under the application of each infinitesimal Lorentz transformation, and so this property is preserved right through the infinite sequence of infinitesimal operations which complete T. *Therefore any finite Lorentz transformation (of the restricted class defined in (35)) is equivalent to a 4-screw.*

We have thus carried over into space-time the equivalent of Chasles' theorem in Euclidean 3-space, and indeed we could make the analogy closer by regarding a Lorentz transformation as a translation and rotation of a "rigid body" defined by the extremities of three orthogonal unit vectors, **I J**, **K**, moving in a *curved* 3-space, namely, the pseudosphere.

We note that by virtue of (58) the matrix **A** of the 4-screw (97), regarded as a transformation $x'_r = A_{rs}x_s$ from coordinate vectors (**I**, **J**, **K**, **L**) to coordinate vectors (**I'**, **J'**, **K'**, **L'**), is

$$\mathbf{A} = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cosh\chi & i\sinh\chi \\ 0 & 0 & -i\sinh\chi & \cosh\chi \end{pmatrix} \tag{138}$$

§ 9. CORRESPONDENCE BETWEEN TRIADS OF NULL RAYS AND UNIT ORTHOGONAL TETRADS

We have been thinking of a Lorentz transformation as a rigid displacement of space-time, with fixed origin. We describe the displacement by saying that a certain unit orthogonal tetrad (**I**, **J**, **K**, **L**) is transformed into a certain other unit orthogonal tetrad (**I'**, **J'**, **K'**, **L'**), just as we describe a rotation in Euclidean 3-space by saying that a certain orthogonal triad is transformed into a certain other orthogonal triad. The law for the transformation of any event is

$$\begin{gathered} \mathbf{E} \to \mathbf{E}', \\ \mathbf{E} = \alpha\mathbf{I} + \beta\mathbf{J} + \gamma\mathbf{K} + \delta\mathbf{L}, \\ \mathbf{E}' = \alpha\mathbf{I}' + \beta\mathbf{J}' + \gamma\mathbf{K}' + \delta\mathbf{L}'. \end{gathered} \tag{139}$$

To describe any given displacement of space-time, we are not obliged to select a particular initial tetrad (**I**, **J**, **K**, **L**). We can choose

it as we like, and the matrix **A** of the transformation depends on the initial tetrad we choose. If we choose the tetrad with (**I**, **J**) in the spacelike 2-flat of the equivalent 4-screw, and (**K**, **L**) in the orthogonal 2-flat, then the matrix **A** has the simple form (138).

But an orthogonal tetrad is a rather clumsy thing to think about, for it involves 4 vectors, having 16 components in all, with 10 relations (91) connecting them (so that there are 6 degrees of freedom left). It would be simpler, but unsymmetric, to use only the triad (**I**, **J**, **K**), and that would suffice, since **L** is completely determined (except for sense) when this triad is known. Then we would have $3 \times 4 = 12$ components to deal with, connected by the 6 relations

$$\mathbf{I}^2 = \mathbf{J}^2 = \mathbf{K}^2 = 1, \quad \mathbf{J}.\mathbf{K} = \mathbf{K}.\mathbf{I} = \mathbf{I}.\mathbf{J} = 0. \tag{140}$$

There is another way of describing unit orthogonal tetrads (and hence Lorentz transformations) which has certain advantages. This is based on *triads of null rays*. A null vector **M** defines a null ray, which we shall denote by M. But M determines **M** only to within a positive factor, for the null vectors $k\mathbf{M}$ ($k > 0$) all belong to the same null ray M. All rays and null vectors considered here will be chosen pointing into the future (so that $M_4/i > 0$).

Any ray (not necessarily null) is determined by the three ratios $x_1 : x_2 : x_3 : x_4$ of the coordinates of any event on it. On a null ray these three ratios are connected by virtue of the relation $x_r x_r = 0$, and so it follows that *a null ray is determined by two numbers*. Hence a triad of null rays (M, N, P) are determined by (and determine) *six* numbers, and so *such a triad has the same number of degrees of freedom as a unit orthogonal tetrad.*

We shall now show how a triad of null rays determines a unit orthogonal tetrad, and conversely. Suppose we are given the three null rays (M, N, P). Let (**M**, **N**, **P**) be any three null vectors along them, so that

$$\mathbf{M}^2 = \mathbf{N}^2 = \mathbf{P}^2 = 0. \tag{141}$$

We shall now *normalise* (**M**, **N**, **P**) by the demand that

$$\mathbf{M}.\mathbf{N} = \mathbf{N}.\mathbf{P} = \mathbf{P}.\mathbf{M} = -1. \tag{142}$$

This we do by replacing (**M**, **N**, **P**) by ($m\mathbf{M}$, $n\mathbf{N}$, $p\mathbf{P}$), where m, n, p are any positive numbers, and then demanding that the relations (142) be satisfied by the three new null vectors, which of course lie on the

given null rays (M, N, P). The conditions are

$$mn\mathbf{M}.\mathbf{N} = np\mathbf{N}.\mathbf{P} = pm\mathbf{P}.\mathbf{M} = -1, \tag{143}$$

and these lead to

$$m^2 = \frac{-\mathbf{N}.\mathbf{P}}{(\mathbf{P}.\mathbf{M})\,(\mathbf{M}.\mathbf{N})}, \quad n^2 = \frac{-\mathbf{P}.\mathbf{M}}{(\mathbf{M}.\mathbf{N})\,(\mathbf{N}.\mathbf{P})}, \quad p^2 = \frac{-\mathbf{M}.\mathbf{N}}{(\mathbf{N}.\mathbf{P})\,(\mathbf{P}.\mathbf{M})}. \tag{144}$$

Since, by (106), all these scalar products are negative, these equations determine unique positive values of m, n, p. Thus the normalisation is possible, and it is unique.

Let us then write $(\mathbf{M}, \mathbf{N}, \mathbf{P})$ for the normalised triad of null vectors lying on the given triad of null rays (M, N, P), so that (141) and (142) are satisfied. We seek a plan for associating a unit orthogonal tetrad $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$ with (M, N, P). A symmetric plan would be best, but none exists. Of the many unsymmetric plans, we shall describe one which is perhaps the simplest.

First, we define $\mathbf{K}$ and $\mathbf{L}$ by

$$\mathbf{K} = \frac{1}{\sqrt{2}}(\mathbf{M} - \mathbf{N}), \quad \mathbf{L} = \frac{1}{\sqrt{2}}(\mathbf{M} + \mathbf{N}). \tag{145}$$

By virtue of (141) and (142) we have then

$$\mathbf{K}^2 = 1, \ \mathbf{L}^2 = -1, \ \mathbf{K}.\mathbf{L} = 0, \tag{146}$$

so that $(\mathbf{K}, \mathbf{L})$ is a unit orthogonal pair, with $\mathbf{L}$ timelike and pointing into the future.

Then we put

$$\mathbf{J} = a\mathbf{M} + b\mathbf{N} + c\mathbf{P}, \tag{147}$$

and choose the coefficients so that

$$\mathbf{J}^2 = 1, \ \mathbf{J}.\mathbf{K} = \mathbf{J}.\mathbf{L} = 0. \tag{148}$$

The result is easily found to be

$$\mathbf{J} = \frac{1}{\sqrt{2}}(\mathbf{M} + \mathbf{N} - \mathbf{P}). \tag{149}$$

We now have a unit orthogonal triad $(\mathbf{J}, \mathbf{K}, \mathbf{L})$, connected with

(**M**, **N**, **P**) by the equations

$$\mathbf{J} = \frac{1}{\sqrt{2}}(\mathbf{M} + \mathbf{N} - \mathbf{P}), \quad \mathbf{K} = \frac{1}{\sqrt{2}}(\mathbf{M} - \mathbf{N}),$$

$$\mathbf{L} = \frac{1}{\sqrt{2}}(\mathbf{M} + \mathbf{N}); \tag{150}$$

$$\mathbf{M} = \frac{1}{\sqrt{2}}(\mathbf{L} + \mathbf{K}), \quad \mathbf{N} = \frac{1}{\sqrt{2}}(\mathbf{L} - \mathbf{K}), \quad \mathbf{P} = \sqrt{2}(\mathbf{L} - \mathbf{J}).$$

It is clear that any orthogonal tetrad of unit vectors (**I**, **J**, **K**, **L**), with **L** timelike and pointing into the future, determines a unique triad of null rays in this way. Conversely, any triad of null rays (M, N, P), pointing into the future, determines an orthogonal tetrad of unit vectors almost completely; for it determines the triad (**J**, **K**, **L**) and then **I** is known except for sense. We could give **I** either of the two values

$$I_r = \pm i\, \varepsilon_{rsmn} J_s K_m L_n, \tag{151}$$

where ε_{rsmn} is the permutation symbol of (32) and I_r, J_s, ... are the components of the vectors in some Minkowskian coordinate system x_r; the factor i is needed to make the first three components of I_r real. The ambiguous sign in (151) brings us face to face with the question of the *orientation* of tetrads in space-time, and this we shall now discuss briefly.

A Galileian observer always takes his time-axis pointing into the future, so that $x_4/i > 0$ for events which, to him, occur after the event which is his origin. He may choose his space-axes with either of the two orientations which are familiar to us, right-handed and left-handed. Now if (**I**, **J**, **K**, **L**) is an orthogonal tetrad, with **L** timelike and pointing into the future, there exists a Galileian observer for whom **L** is the time-axis, and for whom (**I**, **J**, **K**) are spatial vectors; we shall say that the tetrad (**I**, **J**, **K**, **L**) is *right-handed* or *left-handed* according as the triad (**I**, **J**, **K**) is right-handed or left-handed to this observer.

A Lorentz transformation T carries an orthogonal tetrad (**I**, **J**, **K**, **L**) into another orthogonal tetrad (**I**′, **J**′, **K**′, **L**′), the vectors **L** and **L**′ being both timelike and pointing into the future. The transformation T is *proper* (cf. § 2) if the two tetrads have the same orientation (both right-handed or both left-handed), and *improper* if the orientations are different.

Let us now make two decisions to avoid getting lost in a plethora of possibilities:

a) We shall always use right-handed Minkowskian coordinates; thus $x_4/i > 0$ indicates the future and the space-axes are right-handed.

b) We shall choose the ambiguous sign in (151) so that

$$I_r = -i\varepsilon_{rsmn} J_s K_m L_n. \tag{152}$$

This makes the tetrad (**I**, **J**, **K**, **L**) right-handed. For, if we take **L** for time-axis and **J**, **K** for axes Ox_2, Ox_3 respectively, we have

$$J_r = (0, 1, 0, 0),\ K_r = (0, 0, 1, 0),\ L_r = (0, 0, 0, i),$$

and (152) gives $I_r = (1, 0, 0, 0)$, so that **I** coincides with the axis Ox_1.

Thus, supplementing (150) with (152), we have now established a one-to-one correspondence

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \leftrightarrow (M, N, P) \tag{153}$$

between right-handed orthogonal tetrads of unit vectors and triads of null rays. Note that if we use this to obtain an orthogonal tetrad from a triad of null rays, the triad of null rays is quite arbitrary, except that they are taken pointing into the future and no two may coincide. There is no question of making them linearly independent, because it is impossible for three distinct null rays (all pointing into the future) to be linearly dependent, i.e. to lie in a 2-flat.

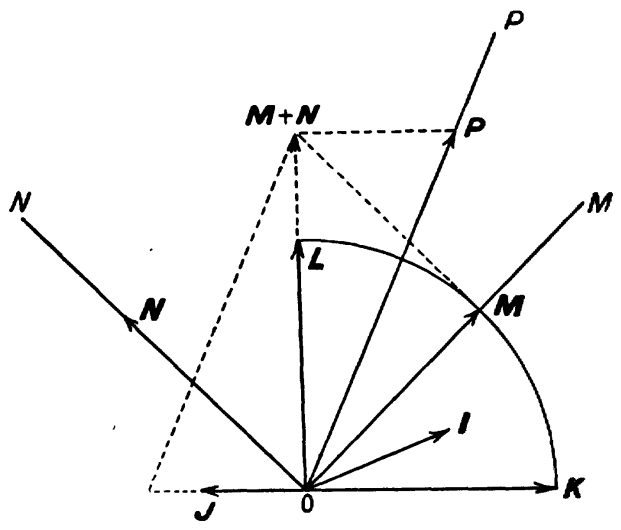


Fig. 3 – The triad of null rays (M, N, P), with normalised null vectors (**M**, **N**, **P**) along them, and the associated unit orthogonal tetrad (**I**, **J**, **K**, **L**)

The arrangement is as shown in the space-time diagram Fig. 3. The four vectors (**K**, **L**, **M**, **N**) lie in a 2-flat, and the six vectors (**J**, **K**, **L**, **M**, **N**, **P**) lie in a 3-flat, to which **I** is orthogonal.

§ 10. LORENTZ TRANSFORMATIONS REPRESENTED BY ARBITRARY TRANSFORMATIONS OF TRIADS OF NULL RAYS

Choose *any two* triads of null rays, (M, N, P) and (M', N', P'), and associate with each a unit orthogonal right-handed tetrad, as in

(153). We may indicate this association by

$$(M, N, P) \ldots (\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$$
$$(M', N', P') \ldots (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}').$$

Then, if we make the mental passage,

$$(M, N, P) \to (M', N', P'),$$

in other words, if we *transform* the triad of null rays, then we make the transformation

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \to (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}'),$$

which is a Lorentz transformation on account of the unit orthogonal character of the tetrads. Further, it is proper transformation, since both tetrads are right-handed.

The connections may be indicated by the following chart:

$$\begin{array}{ccc} (\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) & \leftrightarrow & (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}') \\ \vdots & & \vdots \\ (M, N, P) & \leftrightarrow & (M', N', P') \end{array} \tag{154}$$

Here the arrows indicate transformations (the double head signifying their reversibility) and the vertical dotted lines indicate one-to-one correspondences. Should we wish to obtain improper Lorentz transformations, we would leave one of the correspondences (say that on the left) as described above, and change the other correspondence by writing

$$I'_r = i\,\varepsilon_{rsmn} J'_s K'_m L'_n$$

instead of using the minus sign as in (152); then (**I**, **J**, **K**, **L**) and (**I**′, **J**′, **K**′, **L**′) would have opposite orientations, the former right handed, the latter left handed.

We can make this statement: *Any proper Lorentz transformation (with det* $\mathbf{A} = +1$*) which preserves the future may be generated by a transformation of a triad of null rays. And to every* ARBITRARY *transformation of a triad of null rays there corresponds a unique proper Lorentz transformation, preserving the future.*

We can make this representation of a Lorentz transformation more vivid to the senses by cutting space-time across by the hyperplane $x_4/i = 1$. The section of the null cone $x_r x_r = 0$ by this hyperplane is the sphere

$$x_1^2 + x_2^2 + x_3^2 = 1. \tag{155}$$

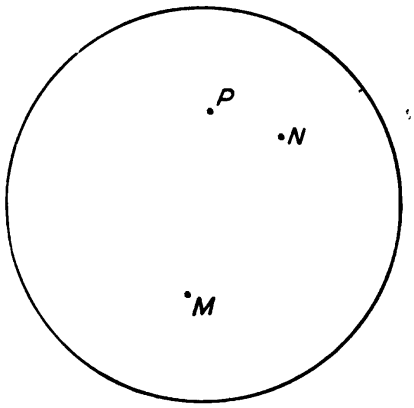


Fig. 4 – A triad of points on the unit sphere representing a triad of null rays

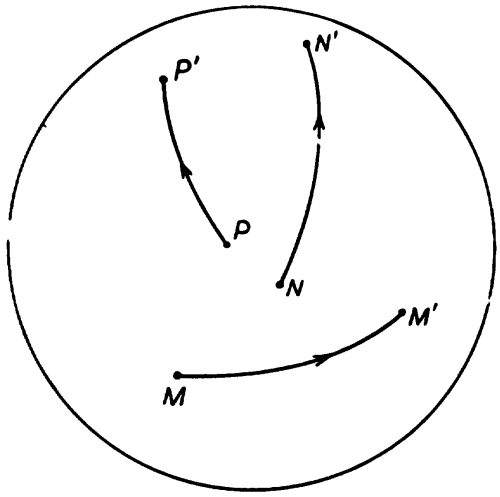


Fig. 5 – Representation of a Lorentz transformation by arbitrary displacements $M \to M'$, $N \to N'$, $P \to P'$ on the unit sphere

To each null ray there corresponds a unique point on the sphere, and so the triad (M, N, P) of null rays gives us a triad of points on the sphere, which we label (M, N, P) (Fig. 4). Conversely, a triad of points on the sphere defines a triad of null rays. The points may be chosen arbitrarily, but no two may coincide.

We can give any three points (M, N, P) on the sphere *completely arbitrary* displacements to new positions (M', N', P') on the sphere (Fig. 5), and so generate a Lorentz transformation. It seems rather remarkable that a transformation as algebraically involved as the Lorentz transformation is should be representable by something so simple as an *arbitrary* transformation of three points on a sphere.

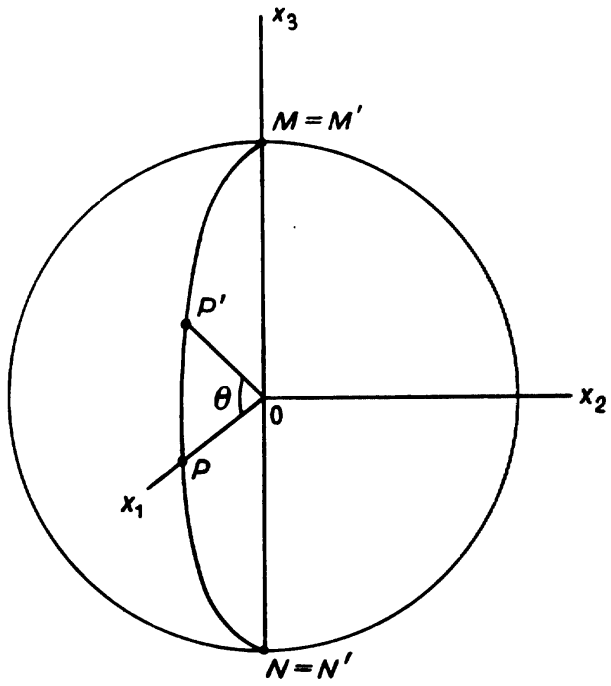


Fig. 6 – Example of the spherical representation of a Lorentz transformation

As an illustration, let us take (see Fig. 6) the following three points on the unit sphere:

$$M(0, 0, 1),\ N(0, 0, -1),\ P(1, 0, 0).$$

Let us fix M and N and

move P round the meridian great circle through an angle θ, so that we have as our new points

$$M'(0, 0, 1), \; N'(0, 0, -1), \; P'(\cos\theta, 0, \sin\theta).$$

We shall take θ in the range

$$-\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi,$$

to avoid a coincidence of points.

These points on the unit sphere are the intersections of certain null rays with the 3-flat $x_4 = i$; unnormalised null vectors along these null rays are

$$\mathbf{M} = (0, 0, 1, i), \; \mathbf{N} = (0, 0, -1, i), \; \mathbf{P} = (1, 0, 0, i),$$
$$\mathbf{M}' = (0, 0, 1, i), \; \mathbf{N}' = (0, 0, -1, i), \; \mathbf{P}' = (\cos\theta, 0, \sin\theta, i).$$

These we must now normalise by (144), so as to have $\mathbf{M}.\mathbf{N} = \mathbf{N}.\mathbf{P} = \mathbf{P}.\mathbf{M} = -1$; the normalised null vectors are

$$\mathbf{M} = \frac{1}{\sqrt{2}}(0, 0, 1, i), \; \mathbf{N} = \frac{1}{\sqrt{2}}(0, 0, -1, i),$$

$$\mathbf{P} = \sqrt{2}\,(1, 0, 0, i),$$

$$\mathbf{M}' = \frac{\lambda}{\sqrt{2}}(0, 0, 1, i), \; \mathbf{N}' = \frac{1}{\lambda\sqrt{2}}(0, 0, -1, i),$$

$$\mathbf{P}' = \frac{\sqrt{2}}{\cos\theta}(\cos\theta, 0, \sin\theta, i),$$

where

$$\lambda = \frac{\cos\tfrac{1}{2}\theta + \sin\tfrac{1}{2}\theta}{\cos\tfrac{1}{2}\theta - \sin\tfrac{1}{2}\theta},$$

the factors outside the parentheses multiplying all the components inside. Then by (150) and (152) the tetrad $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$ corresponding to (M, N, P) is

$$\mathbf{I} = (0, 1, 0, 0), \; \mathbf{J} = (-1, 0, 0, 0), \; \mathbf{K} = (0, 0, 1, 0), \; \mathbf{L} = (0, 0, 0, i),$$

and the tetrad $(\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$ corresponding to (M', N', P') is

$$\mathbf{I}' = (0, 1, 0, 0), \; \mathbf{J}' = (-1, 0, 0, 0),$$
$$\mathbf{K}' = (0, 0, \sec\theta, i\tan\theta), \; \mathbf{L}' = (0, 0, \tan\theta, i\sec\theta)$$

Thus, by moving the point P on the sphere, we have generated the Lorentz transformation

$$\begin{gathered}\mathbf{I}' = \mathbf{I},\ \mathbf{J}' = \mathbf{J},\\ \mathbf{K}' = \sec\theta\,\mathbf{K} + \tan\theta\,\mathbf{L},\\ \mathbf{L}' = \tan\theta\,\mathbf{K} + \sec\theta\,\mathbf{L},\\ \tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi.\end{gathered} \tag{156}$$

This we recognise as the transformation (62) with $\cosh\chi = \sec\theta$, $\sinh\chi = \tan\theta$. By moving P all along the great circle between M and N we generate this whole set of transformations with

$$-\infty < \chi < \infty.$$

To enlarge on the meaning of the equivalence of the transformations $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \leftrightarrow (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$ and $(M, N, P) \leftrightarrow (M', N', P')$ of (154), let us look at things in a different way. Any Lorentz transformation may be written (as usual we consider only those which leave the origin unchanged)

$$\Lambda\colon\ X_r' = A_{rs}X_s,\ A_{rs}A_{rt} = \delta_{st}; \tag{157}$$

here we think of the coordinate axes as fixed and Λ as a process which moves the event X_r to the event X_r'. Now Λ ,having only six parameters in it, is determined when we know what it does to a few rays, a ray (not necessarily a null ray) being a set of events kX_r for $0 < k < \infty$. We need consider only rays with $X_4/i > 0$. Then the three numbers $x_\varrho = iX_\varrho/X_4$ determine a ray before transformation and the three numbers $x_\varrho' = iX_\varrho'/X_4'$ determine it after transformation. By (157) we have the transformation

$$T\colon\ x_\varrho' = \frac{A_{\varrho\sigma}x_\sigma + iA_{\varrho 4}}{-iA_{4\sigma}x_\sigma + A_{44}}, \tag{158}$$

a six-parameter group of projective transformation of Euclidean 3-space into itself, equivalent to the Lorentz transformation Λ of space-time.

In particular, T transforms the unit sphere (155) into itself. Moreover, T (having six parameters) is determined when we are told what it does to three points on the unit sphere, and that is where the points M, N, P of Fig. 5 come in.

We note that, since Λ leaves two null rays unchanged [cf. § 8], T leaves two points on the unit sphere unchanged.

§ 11. SPINORS

Spinors play a fundamental part in relativistic quantum mechanics, appearing in Dirac's equations for an electron in an electromagnetic field (DIRAC [1930], Ch. XIII). Quantum theory lies outside the scope of this book and no extensive treatment of spinors will be presented here, but as they are intimately connected with the Lorentz transformation, it is well to give some account of them, explaining what they are and how a spin transformation generates a Lorentz transformation.[1]

Take a pair of complex numbers, ξ and η. Write

$$\begin{aligned} x_1 &= \bar{\xi}\eta + \xi\bar{\eta}, \\ ix_2 &= \bar{\xi}\eta - \xi\bar{\eta}, \\ x_3 &= \bar{\xi}\xi - \bar{\eta}\eta, \\ x_4 &= i(\bar{\xi}\xi + \bar{\eta}\eta), \end{aligned} \tag{159}$$

where the bars denote complex conjugates. Then, to any choice of ξ and η there corresponds a unique tetrad of values of x_r, with the first three real and x_4 a pure imaginary with $x_4/i > 0$. A simple calculation gives

$$x_r x_r = 0. \tag{160}$$

Thus, taking x_r to be the Minkowskian coordinates of an event in space-time, we may say that *any complex pair* (ξ, η) *defines a unique event on the future sheet of the null cone at the origin.*

Now introduce into space-time the coordinates

$$\begin{aligned} p &= \tfrac{1}{2}(x_1 + ix_2), \\ q &= \tfrac{1}{2}(x_1 - ix_2), \\ r &= \tfrac{1}{2}(x_3 - ix_4) = \tfrac{1}{2}(x_3 + ct), \\ s &= \tfrac{1}{2}(-x_3 - ix_4) = \tfrac{1}{2}(-x_3 + ct). \end{aligned} \tag{161}$$

Note that p and q are complex (with $q = \bar{p}$) and r and s are real. The one complex coordinate p, together with the real r and s, would suffice to determine an event, but it is more convenient to use all the four (p, q, r, s).

We note that the conditions

$$pq - rs = 0, \quad r + s > 0 \tag{162}$$

[1] For the theory of spinors, see WEYL [1931], INFELD [1932], MURNAGHAN [1938], BADE and JEHLE [1953]; this last paper contains a bibliography. For the connection between spinors and skew-symmetric tensors, see WHITTAKER [1937b] and RUSE [1937].

imply that the event (p, q, r, s) lies on the future sheet of the null cone. We further note that the fundamental form of space-time is

$$dx_r\, dx_r = 4(dp\, dq - dr\, ds). \tag{163}$$

The equations (159) are equivalent to

$$p = \bar{\xi}\eta,\ q = \xi\bar{\eta},\ r = \bar{\xi}\xi,\ s = \bar{\eta}\eta. \tag{164}$$

The existence of (ξ, η) satisfying these equations implies the truth of (162).

Suppose now that we are given an event on the future sheet of the null cone. This means that we are given p, q, r, s (with $q = \bar{p}$), satisfying (162). Let us see whether there exists a pair (ξ, η) satisfying (164). We try then to solve (164) by writing

$$\xi = ae^{i\theta},\ \eta = be^{i\phi}, \tag{165}$$

where a, b, θ, ϕ are real, a and b being positive. The *four* equations (164) are equivalent to the *three* equations

$$abe^{i(\phi-\theta)} = p,\ a^2 = r,\ b^2 = s. \tag{166}$$

Thus a and b are determined uniquely, and so is the angle $\theta - \phi$, but there remains a degree of arbitrariness in (ξ, η); if (ξ, η) satisfy (164), then so do $(\xi \exp i\omega, \eta \exp i\omega)$, where ω is any real number. This indeterminacy is of course evident when we examine (164).

The ordered pair (ξ, η) is called a *spinor*. If we resolve ξ and η into their real and imaginary parts, writing

$$\xi = \mu + i\nu,\ \eta = \varrho + i\sigma,$$

then the spinor is the tetrad of real numbers $(\mu, \nu, \varrho, \sigma)$.

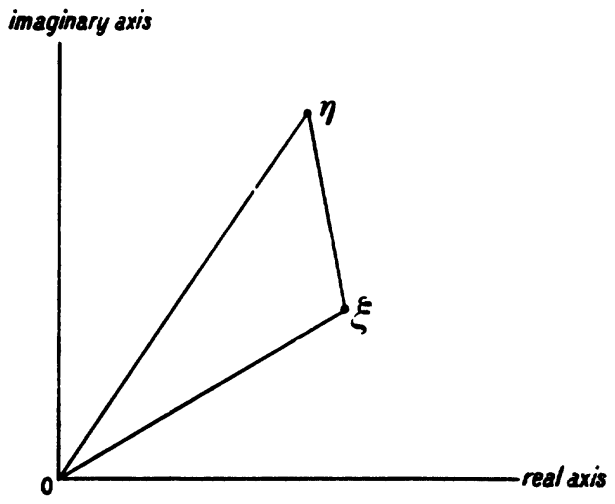


Fig. 7 – Triangle in the spinor plane

We may represent a spinor by *two* points in a plane, the complex numbers ξ and η being plotted as in the usual Argand diagram (Fig. 7). We call this the *spinor plane* and the triangle $O\xi\eta$ a *spin triangle*, the vertices being ordered as indicated.

A spin triangle determines a unique future-pointing null vector by (159); *but a future-pointing null vector does not*

determine a unique spin triangle — it determines a set of them obtained from one by rotation about O through all angles.

Let us now subject spinors to a linear transformation

$$(\xi, \eta) \to (\xi', \eta') \begin{cases} \xi' = \alpha\xi + \beta\eta, \\ \eta' = \gamma\xi + \delta\eta, \end{cases} \tag{167}$$

where α, β, γ, δ are any four complex constants, subject to the *unimodular* condition

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \alpha\delta - \beta\gamma = 1. \tag{168}$$

We note that the *spin transformation* (167) resembles a Lorentz transformation in that it has *six* degrees of freedom; for there are eight real numbers in α, β, γ, δ, and (168) provides two relations between these real numbers.

If we write (p', q', r', s') for the result of making the substitution (167) in (164), i.e. if we write

$$p' = \bar{\xi}'\eta', \quad q' = \xi'\bar{\eta}', \quad r' = \bar{\xi}'\xi', \quad s' = \bar{\eta}'\eta', \tag{169}$$

then

$$\begin{aligned} p' &= (\bar{\alpha}\bar{\xi} + \bar{\beta}\bar{\eta})(\gamma\xi + \delta\eta) \\ &= \bar{\alpha}\delta\bar{\xi}\eta + \bar{\beta}\gamma\xi\bar{\eta} + \bar{\alpha}\gamma\bar{\xi}\xi + \bar{\beta}\delta\bar{\eta}\eta. \end{aligned} \tag{170}$$

If we now substitute from (164) in this and the three similar equations, we get

$$\begin{aligned} p' &= \bar{\alpha}\delta p + \bar{\beta}\gamma q + \bar{\alpha}\gamma r + \bar{\beta}\delta s, \\ q' &= \beta\bar{\gamma}p + \alpha\bar{\delta}q + \alpha\bar{\gamma}r + \beta\bar{\delta}s, \\ r' &= \bar{\alpha}\beta p + \alpha\bar{\beta}q + \bar{\alpha}\alpha r + \bar{\beta}\beta s, \\ s' &= \bar{\gamma}\delta p + \gamma\bar{\delta}q + \bar{\gamma}\gamma r + \bar{\delta}\delta s, \\ &\alpha\delta - \beta\gamma = 1. \end{aligned} \tag{171}$$

(We check that $\bar{p}' = q'$.)

The spin transformation (167) *generates* the linear transformation (171) in the sense that if the coefficients α, β, γ, δ of (167) are given, then the coefficients in (171) are determined. Note that these coefficients are the 16 products obtained by taking one factor from the set $(\alpha, \beta, \gamma, \delta)$ and one from the set $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\delta})$. The transformation may also be written in matrix form:

$$\begin{pmatrix} r' & q' \\ p' & s' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} r & q \\ p & s \end{pmatrix} \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix}. \tag{172}$$

In passing from (167) to (171) we used (164), which ties the event (p, q, r, s) to the null cone, and we might therefore think that (171) must be regarded only as a transformation of the null cone into itself. But this is not so. We may think of (171) as a transformation in its own right:

$$(p, q, r, s) \to (p', q', r', s'),$$

i.e. by (161) a transformation $(x \to x')$ of the whole of space-time. By direct calculation from (171), or by taking determinants in (172), we find

$$dp'dq' - dr'ds' = dp\,dq - dr\,ds, \tag{173}$$

by virtue of the unimodular condition (168), and hence by (163)

$$dx_r'dx_r' = dx_r dx_r. \tag{174}$$

This tells us that the transformation $x_r' = A_{rs}x_s$, equivalent to (171) by (161), is a Lorentz transformation, and it is a proper one for direct calculation gives

$$\det A_{rs} = (\alpha\delta - \beta\gamma)^2\,(\bar{\alpha}\bar{\delta} - \bar{\beta}\bar{\gamma})^2 = 1 \tag{175}$$

by (168). Further, it preserves the future. To see this, it suffices to examine timelike directions with $x_1 = x_2 = 0$, pointing into the future; for them we have $p = q = 0$, $r > 0$, $s > 0$, and then by (161)

$$ct' = x_4'/i = r' + s' = (\bar{\alpha}\alpha + \bar{\gamma}\gamma)r + (\bar{\beta}\beta + \bar{\delta}\delta)s > 0.$$

We conclude then that, *for any values of the complex numbers* α, β, γ, δ, *subject only to the condition* $\alpha\delta - \beta\gamma = 1$, (171) *gives a proper future-preserving Lorentz transformation of whole of space-time and not merely of the null cone.*

We might regard (171) as a factory for the mass-production of Lorentz transformations. A Lorentz transformation, as in (35), contains 16 coefficients A_{rs} connected by the 10 relations $A_{rs}A_{rt} = \delta_{st}$. To make sure that a given set of A_{rs} are actually the coefficients of a Lorentz transformation, a considerable amount of algebra may be involved. On the other hand, in (171) all we have to do is to choose any four complex numbers, making sure that they satisfy $\alpha\delta - \beta\gamma = 1$; this is very easy, for we can choose three of them and then make the fourth one fit.

Here are two simple examples of the generation of Lorentz transformations from spin transformations. In each case we give the values of

$\alpha, \beta, \gamma, \delta$, the spin transformation (167), and the Lorentz transformation, first in the form (171), and then in terms of Minkoswkian coordinates.

$$\alpha = e^{i\theta/2},\ \beta = 0,\ \gamma = 0,\ \delta = e^{-i\theta/2},\ \alpha\delta - \beta\gamma = 1,$$
$$\xi' = \xi e^{\frac{1}{2}i\theta},\ \eta' = \eta e^{-\frac{1}{2}i\theta}, \tag{176}$$
$$p' = e^{-i\theta}p,\ q' = e^{i\theta}q,\ r' = r,\ s' = s,$$
$$x_1' = x_1 \cos\theta + x_2 \sin\theta,$$
$$x_2' = -x_1 \sin\theta + x_2 \cos\theta,$$
$$x_3' = x_3,$$
$$x_4' = x_4.$$

$$\alpha = e^{-\chi/2},\ \beta = 0,\ \gamma = 0,\ \delta = e^{\chi/2},\ \alpha\delta - \beta\gamma = 1,$$
$$\xi' = \xi e^{-\frac{1}{2}\chi},\ \eta' = \eta e^{\frac{1}{2}\chi}, \tag{177}$$
$$p' = p,\ q' = q,\ r' = e^{-\chi}r,\ s' = e^{\chi}s,$$
$$x_1' = x_1,$$
$$x_2' = x_2,$$
$$x_3' = x_3 \cosh\chi + ix_4 \sinh\chi,$$
$$x_4' = -ix_3 \sinh\chi + x_4 \cosh\chi.$$

We recognise (36) and (39).

§ 12. THE TWO SPIN TRANSFORMATIONS CORRESPONDING TO A GIVEN LORENTZ TRANSFORMATION

We have seen that a given spin transformation (167) generates a unique Lorentz transformation (171). Does a given Lorentz transformation generate a unique spin transformation? We shall answer that question by proving the following theorem: *Any proper future-preserving Lorentz transformation can be expressed in the form* (171), *with values of* $\alpha, \beta, \gamma, \delta$ (*subject to* $\alpha\delta - \beta\gamma = 1$) *uniquely determined except for a change of sign in them all.* In other words, a given Lorentz transformation generates *two* (*and only two*) spin transformations, the coefficients $(\alpha, \beta, \gamma, \delta)$ of one being related to the coefficients $(\alpha', \beta', \gamma', \delta')$ of the other by

$$\alpha = -\alpha',\ \beta = -\beta',\ \gamma = -\gamma',\ \delta = -\delta'. \tag{178}$$

That such a duality should exist is of course evident from (171), for the coefficients are not changed if we reverse the signs of all the four numbers $(\alpha, \beta, \gamma, \delta)$.

To prove the theorem, let us use the result of § 9: any future-preserving Lorentz transformation is equivalent to an arbitrary transformation of a triad of null rays.

Now a null *vector* corresponds to a spin triangle (Fig. 7), with the understanding that all triangles formed from one by rotation about the origin are equivalent. Hence a null *ray* corresponds to what we might call a *spin shape*, using this term to mean the totality of all triangles $O\xi\eta$ in the spinor plane which can be formed from one by rotation round the origin and uniform expansion. This is the same as multiplying (ξ, η) by some undetermined complex constant, such a multiplication having the effect of moving the corresponding event along the null ray. Hence a triad of null rays corresponds to three spinors of the form

$$(m_1\xi_1, m_1\eta_1), \ (m_2\xi_2, m_2\eta_2), \ (m_3\xi_3, m_3\eta_3), \tag{179}$$

where (ξ_1, η_1), (ξ_2, η_2), (ξ_3, η_3) are three definite spinors and m_1, m_2, m_3 three arbitrary complex numbers.

Suppose that a certain Lorentz transformation is given; let it change the triad of null rays (179) into the triad

$$(m_1'\xi_1', m_1'\eta_1'), \ (m_2'\xi_2', m_2'\eta_2'), \ (m_3'\xi_3', m_3'\eta_3'). \tag{180}$$

The question before us is this: Can the transformation from (179) to (180) be brought about by a spin transformation of the form (167), and, if so, to what extent is the spin transformation determined? Equivalently, we inquire into the existence of α, β, γ, δ satisfying equations of the form

$$\begin{aligned} k_1\xi_1' &= \alpha\xi_1 + \beta\eta_1, \quad k_1\eta_1' = \gamma\xi_1 + \delta\eta_1, \\ k_2\xi_2' &= \alpha\xi_2 + \beta\eta_2, \quad k_2\eta_2' = \gamma\xi_2 + \delta\eta_2, \\ k_3\xi_3' &= \alpha\xi_3 + \beta\eta_3, \quad k_3\eta_3' = \gamma\xi_3 + \delta\eta_3, \end{aligned} \tag{181}$$

$$\alpha\delta - \beta\gamma = 1, \tag{182}$$

where for brevity we have written $k_1 = m_1'/m_1$, $k_2 = m_2'/m_2$, $k_3 = m_3'/m_3$. In these equations the twelve quantities

$$\xi_1, \eta_1, \xi_2, \eta_2, \xi_3, \eta_3, \xi_1', \eta_1', \xi_2', \eta_2', \xi_3', \eta_3'$$

are regarded as known, and the seven quantities

$$\alpha, \beta, \gamma, \delta, k_1, k_2, k_3$$

as unknown, so that we have seven equations for seven unknowns.

If we eliminate α, β from the equations on the left in (181) and γ, δ

from the equations on the right, we get the following two equations for k_1, k_2, k_3:

$$\begin{vmatrix} k_1\xi_1' & \xi_1 & \eta_1 \\ k_2\xi_2' & \xi_2 & \eta_2 \\ k_3\xi_3' & \xi_3 & \eta_3 \end{vmatrix} = 0, \qquad \begin{vmatrix} k_1\eta_1' & \xi_1 & \eta_1 \\ k_2\eta_2' & \xi_2 & \eta_2 \\ k_3\eta_3' & \xi_3 & \eta_3 \end{vmatrix} = 0. \tag{183}$$

The ratios of the k's are therefore determined uniquely, and so we may write $k_2 = \theta k_1$, $k_3 = \phi k_1$, where θ and ϕ are known. Then (181) gives $\alpha, \beta, \gamma, \delta$ in the form

$$\alpha = \alpha_0 k_1, \ \beta = \beta_0 k_1, \ \gamma = \gamma_0 k_1, \ \delta = \delta_0 k_1, \tag{184}$$

where $\alpha_0, \beta_0, \gamma_0, \delta_0$ are known, and finally (182) gives k_1^2 uniquely. Thus *$\alpha, \beta, \gamma, \delta$ are determined uniquely except for a common change of sign when the Lorentz transformation is given*, and so the theorem is proved.

This means that every proper future-preserving Lorentz transformation may be written in the form (171); in fact, (171) gives us *all* such Lorentz transformations.

By changing from Minkowskian coordinates x_r to (p, q, r, s) by (161), we can throw any proper future-preserving Lorentz transformation into the form

$$\begin{aligned} p' &= Ap + Bq + Cr + Ds, \\ q' &= \bar{B}p + \bar{A}q + \bar{C}r + \bar{D}s, \\ r' &= Ep + \bar{E}q + Rr + Ss, \\ s' &= Fp + \bar{F}q + Tr + Us, \end{aligned} \tag{185}$$

where A, B, C, D, E, F are complex constants (the bars indicating conjugates) and R, S, T, U are real positive constants. Now we know that (185) must be actually of the form (171), and so $\alpha, \beta, \gamma, \delta$ exist so that

$$\begin{gathered} \bar{\alpha}\delta = A, \ \bar{\beta}\gamma = B, \ \bar{\alpha}\gamma = C, \ \bar{\beta}\delta = D, \\ \bar{\alpha}\beta = E, \ \bar{\alpha}\alpha = R, \ \bar{\beta}\beta = S, \\ \bar{\gamma}\delta = F, \ \bar{\gamma}\gamma = T, \ \bar{\delta}\delta = U. \end{gathered} \tag{186}$$

A product formed from two members of $(\alpha, \beta, \gamma, \delta)$ and two members of $(\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\delta})$ can be factorised in two ways into two factors, with each factor in the set of products occurring in (186). For example,

$$\bar{\alpha}\bar{\beta}\alpha\gamma = \bar{\alpha}\alpha . \bar{\beta}\gamma = \bar{\alpha}\gamma . \bar{\beta}\alpha. \tag{187}$$

From this we deduce that $RB = C\bar{E}$. In this way we can obtain 21 simple connections between the coefficients of a Lorentz transformation

when written in the form (185). These connections are as follows:

$$\begin{aligned} RU &= \bar{A}A, \quad ST = \bar{B}B, \\ RS &= \bar{E}E, \quad RT = \bar{C}C, \\ SU &= \bar{D}D, \quad TU = \bar{F}F, \end{aligned} \tag{188}$$

$$\begin{aligned} RB &= C\bar{E}, \quad RD = A\bar{E}, \quad RF = A\bar{C}, \\ SA &= DE, \quad SC = BE, \quad SF = \bar{B}D, \\ TA &= CF, \quad TD = BF, \quad TE = \bar{B}C, \\ UB &= D\bar{F}, \quad UC = A\bar{F}, \quad UE = A\bar{D}, \\ AB &= CD, \quad A\bar{B} = EF, \quad C\bar{D} = E\bar{F}. \end{aligned}$$

These 21 connections are, of course, not independent. They embrace the orthogonality conditions $A_{rs}A_{rt} = \delta_{rt}$ with the future-preserving condition $A_{44} > 0$.

§ 13. THE SIMPLE LORENTZ TRANSFORMATION BETWEEN TWO FRAMES OF REFERENCE

In studying the mathematics of Lorentz transformations we have wandered away from physics. Let us now recall the splitting of space-time into space and time (II–§ 6), the meaning of a Galileian frame of reference (II–§ 7), and the connection between Galileian coordinates and Minkowskian coordinates (II–§ 9). We get back to physics when we remember that a Galileian frame of reference is a set of parallel world lines (the histories of a cloud of free particles), that the real coordinate x^4 is measured by a standard clock carried on one of these particles, and that the passage to Minkowskian coordinates involves only a notational dodge — the introduction of the imaginary time

$$x_4 = ix^4 = ict.$$

The vector **L** of the unit orthogonal tetrad of coordinate vectors (**I**, **J**, **K**, **L**) introduced in (56) lies along the world line of one of the free particles defining the frame of reference. The Galileian observer who uses this frame has therefore no option about the choice of **L**; the frame determines it. But it is otherwise with (**I**, **J**, **K**). They must be unit vectors, orthogonal to **L** and to one another, but that leaves them with as much freedom as an orthogonal triad of unit vectors in Newtonian physics. Subject to the restrictions mentioned, the Galileian observer can choose them as he thinks fit.

A Lorentz transformation

$$(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}) \rightarrow (\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}') \tag{189}$$

carries us from one system of Minkowskian coordinates to another. But it is only the part $\mathbf{L} \rightarrow \mathbf{L}'$ that can be truly regarded as describing a change in frame of reference. If, then, two Galileian observers are willing to cooperate in order to give to the transformation (189) the simplest form possible, they can do so by adjusting their space vectors $(\mathbf{I}, \mathbf{J}, \mathbf{K})$ $(\mathbf{I}', \mathbf{J}', \mathbf{K}')$ to that end. They cannot change $\mathbf{L}$ and $\mathbf{L}'$ without changing their frames of reference.

To see how (189) is to be given its simplest form by this cooperation, consider two Galileian observers, S and S', whose timelike coordinate vectors are $\mathbf{L}$ and $\mathbf{L}'$ respectively. The vectors $\mathbf{L}$, $\mathbf{L}'$ are contained in a 2-flat Π with equations of the form

$$a_r y_r = 0, \quad b_r y_r = 0, \tag{190}$$

y_r being any system of Minkowskian coordinates. There is a 3-flat orthogonal to $\mathbf{L}$ with an equation of the form

$$c_r y_r = 0. \tag{191}$$

The three equations (190), (191) define the ratios $y_1 : y_2 : y_3 : y_4$, and so determine a straight line in space-time which lies both in Π and in the 3-flat, and consequently is orthogonal to $\mathbf{L}$. *Let S choose his space vector $\mathbf{I}$ on this line*; he can do so, because he can choose $\mathbf{I}$ in any direction orthogonal to $\mathbf{L}$. Then $\mathbf{I}$ is in Π. Let S' cooperate by choosing $\mathbf{I}'$ in Π (Fig. 8).

We now have $\mathbf{I}$, $\mathbf{L}$, $\mathbf{I}'$, $\mathbf{L}'$ all in one 2-flat Π. Then $\mathbf{J}$, $\mathbf{K}$, $\mathbf{J}'$, $\mathbf{K}'$ must lie in the 2-flat Π^* orthogonal to Π. However S may choose $\mathbf{J}$ and $\mathbf{K}$, S' can choose $\mathbf{J}' = \mathbf{J}$, $\mathbf{K}' = \mathbf{K}$. This completes the specification of the two tetrads, $(\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L})$ and $(\mathbf{I}', \mathbf{J}', \mathbf{K}', \mathbf{L}')$.

A reversal in the sense of $\mathbf{I}$ or $\mathbf{I}'$ would change the orientation of one of the tetrads, and so would change the Lorentz transformation from proper to improper, or vice versa. Apart from this ques-

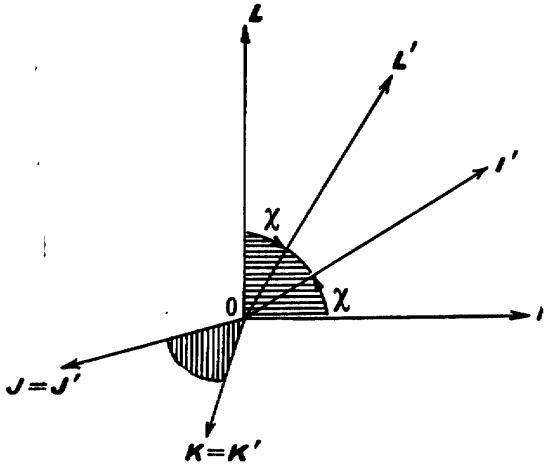


Fig. 8 – Coordinate tetrads for the simple Lorentz transformation between two frames of reference, obtained by the cooperation of the two observers

tion of reversals of sense, we may say that *two given frames of reference (i.e.* **L** *and* **L**′ *given) determine the two tetrads* (**I**, **J**, **K**, **L**) *and* (**I**′, **J**′, **K**′, **L**′) *to within one degree of freedom, that degree of freedom corresponding to rotation of* (**J**, **K**), *and* (**J**′, **K**′) *with them, in the 2-flat* Π^*.

With the tetrads specified as above, the transformation (189) is a timelike rotation through some pseudoangle χ [cf. (62)], provided we choose the senses of **I** and **I**′ in such a way that (**I**, **L**) can be continuously transformed into (**I**′, **L**′) without crossing the null cone. Thus, *merely by cooperation between the two Galileian observers in the matter of choice of space axes, a general Lorentz transformation can be expressed as a timelike rotation*

$$\begin{aligned} \mathbf{I}' &= \mathbf{I}\cosh\chi + \mathbf{L}\sinh\chi, \\ \mathbf{L}' &= \mathbf{I}\sinh\chi + \mathbf{L}\cosh\chi, \\ \mathbf{J}' &= \mathbf{J}, \qquad \mathbf{K}' = \mathbf{K}. \end{aligned} \tag{192}$$

By (58) the components of the matrix A_{rs} of the transformation $x_r' = A_{rs}x_s$ are as follows, those not shown being zero:

$$\begin{aligned} A_{11} &= \cosh\chi, \qquad A_{14} = i\sinh\chi, \\ A_{41} &= -i\sinh\chi, \quad A_{44} = \cosh\chi, \\ A_{22} &= A_{33} = 1. \end{aligned} \tag{193}$$

We note that $\det \mathbf{A} = 1$, so that this is a proper Lorentz transformation. The transformation between the Minkowskian coordinates of the two observers then reads

$$\boxed{\begin{aligned} x_1' &= x_1\cosh\chi + ix_4\sinh\chi, \\ x_4' &= -ix_1\sinh\chi + x_4\cosh\chi, \\ x_2' &= x_2, \quad x_3' = x_3, \end{aligned}} \tag{194}$$

or in terms of real coordinates without suffixes,

$$\begin{aligned} x' &= x\cosh\chi - ct\sinh\chi, \\ ct' &= -x\sinh\chi + ct\cosh\chi, \\ y' &= y, \quad z' = z. \end{aligned} \tag{195}$$

It remains to give a kinematical meaning to the constant χ. Let S observe the cloud of free particles which form the frame of S'. Since

each of these is fixed in the frame of S', we have for each of them

$$dx' = dy' = dz' = 0, \tag{196}$$

and so by (195)

$$\frac{dx}{dt} = c \tanh \chi, \quad \frac{dy}{dt} = 0, \quad \frac{dz}{dt} = 0, \tag{197}$$

Thus S sees every particle of the frame of S' moving parallel to his x-axis with velocity

$$v = c \tanh \chi. \tag{198}$$

We have then

$$\tanh \chi = v/c, \ \cosh \chi = \gamma, \ \sinh \chi = \gamma v/c, \tag{199}$$

where

$$\gamma = \frac{1}{\sqrt{(1 - v^2/c^2)}}, \tag{200}$$

and the matrix A_{rs} of (193) becomes

$$\begin{matrix} A_{11} = \gamma, \ A_{14} = i\gamma v/c, \\ A_{41} = -i\gamma v/c, \ A_{44} = \gamma. \end{matrix} \quad A_{22} = A_{33} = 1, \tag{201}$$

The Lorentz transformation (195) for real coordinates may be written

$$\boxed{\begin{aligned} x' &= \gamma(x - vt), \\ t' &= \gamma(t - vx/c^2), \\ y' &= y, \ z' = z. \end{aligned}} \tag{202}$$

If we solve these equations for x, y, z, t, we get

$$\begin{aligned} x &= \gamma(x' + vt'), \\ t &= \gamma(t' + vx'/c^2), \\ y &= y', \ z = z'. \end{aligned} \tag{203}$$

If we now follow a particle fixed in S, we have $dx = dy = dz = 0$ and so $dx'/dt' = -v$, $dy'/dt' = dz'/dt' = 0$. Therefore we may say that *the velocity of S' relative to S has components* $(v, 0, 0)$, *and the velocity of S relative to S' has components* $(-v, 0, 0)$.

The formulae (202) for the simple Lorentz transformation between two Galileian frames of reference are most important formulae in the special theory of relativity. We shall use them extensively in Chapter V. In considering them, one can afford to forget the details of their

derivation and satisfy oneself by a simple direct calculation that they make

$$dx'^2 + dy'^2 + dz'^2 - c^2dt'^2 = dx^2 + dy^2 + dz^2 - c^2dt^2, \quad (204)$$

and consequently form a Lorentz transformation. The purpose of the argument leading to them was to show that, although they are very simple in form, they are quite general, in the sense that any Lorentz transformation (proper and future-preserving) can be given this form by a mere cooperation between the two observers involved.

If c is large, or, more correctly, if the dimensionless combinations v/c and $x/(ct)$ are small, then we may as an approximation discard terms from (202) on account of their smallness. If we discard terms of the second order, then $\gamma = 1$ and the second term on the right of the formula for t' disappears, with the result that (202) reads in approximate form

$$x' = x - vt, \; y' = y, \; z' = z, \; t' = t. \quad (205)$$

These are the familiar Newtonian formulae connecting two sets of axes in uniform relative motion, including the identical transformation for the time, always understood in Newtonian physics. The approximate collapse of (202) into (205) is most important physically, since it indicates in a general way that relativistic effects are likely to be below the limit of observation unless velocities are involved which are comparable with the velocity of light, in the sense that $(v/c)^2$ or $vx/(c^2t)$ is large enough to be considered different from zero, taking into consideration the accuracy of the experimental technique available.

§ 14. LORENTZ TRANSFORMATIONS WITH HERMITIAN (OR SYMMETRIC) MATRIX

It is not always convenient to demand from two Galileian observers the full measure of cooperation necessary to secure the simple form (202) for the Lorentz transformation between their frames.

We shall now suppose that S retains his axes unchanged, the cooperation being left to S' in his choice of space axes.

Let v_ϱ be the velocity of S' (i.e. of the particles forming his frame) relative to S. Let v'_ϱ be the velocity of S relative to S'. Since S is not going to change his space axes, v_ϱ are given numbers. But S' will change his space axes and so change v'_ϱ, this change being the same as what we get in Newtonian mechanics when we change the axes along

which a certain velocity is measured. It is clear that S' can in this way give to v'_ϱ any *ratios* he pleases; let him choose his space axes so that

$$v'_\varrho = kv_\varrho, \tag{206}$$

k being some unknown constant.

Let us write

$$v = (v_\varrho v_\varrho)^{\frac{1}{2}}, \quad v' = (v'_\varrho v'_\varrho)^{\frac{1}{2}}. \tag{207}$$

Then a rotation of space axes by S leaves v unaltered and a rotation of space axes by S' leaves v' unaltered. But, with the full cooperation of § 13, we have

$$v'_1 = -v_1, \quad v'_2 = -v_2\,(=0), \quad v'_3 = -v_3\,(=0), \tag{208}$$

and so we have $v' = v$ for the vectors in (206), since this relation holds for (208). Therefore $k = \pm 1$. But (208) is a special case of (206), and in (208) we have $k = -1$. By continuity we must have $k = -1$ in (206), and so that equation becomes

$$v'_\varrho = -v_\varrho. \tag{209}$$

Let us see what this relation tells us about the Lorentz transformation $x'_r = A_{rs}x_s$, or, as we may write it,

$$\begin{aligned} x'_\varrho &= A_{\varrho\sigma}x_\sigma + iA_{\varrho 4}\,ct, \\ ct' &= -iA_{4\sigma}x_\sigma + A_{44}\,ct. \end{aligned} \tag{210}$$

Following a particle of S, we have

$$dx_\varrho = 0, \quad \frac{dx'_\varrho}{dt'} = v'_\varrho, \tag{211}$$

and so

$$v'_\varrho = icA_{\varrho 4}/A_{44}. \tag{212}$$

If we use the inverse transformation $x_r = A_{sr}x'_s$ and follow a particle of S', we get similarly

$$v_\varrho = icA_{4\varrho}/A_{44}. \tag{213}$$

Hence, by (209),

$$A_{\varrho 4} = -A_{4\varrho}. \tag{214}$$

The orthogonality conditions $A_{rs}A_{rt} = \delta_{st}$ may be written

$$A_{\varrho\sigma}A_{\varrho\tau} + A_{4\sigma}A_{4\tau} = \delta_{\sigma\tau}, \tag{215}$$

$$A_{\varrho\sigma}A_{\varrho 4} + A_{4\sigma}A_{44} = 0, \tag{216}$$

$$A_{\varrho 4}A_{\varrho 4} + A_{44}^2 = 1. \tag{217}$$

From (212) and (217) we find

$$A_{44} = \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}, \tag{218}$$

and then by (215) and (216) we have for the nine quantities $A_{\varrho\sigma}$ the nine equations

$$\begin{aligned} A_{\varrho\sigma} A_{\varrho\tau} &= \delta_{\sigma\tau} + \gamma^2 v_\sigma v_\tau/c^2, \\ A_{\varrho\sigma} v_\varrho &= \gamma v_\sigma. \end{aligned} \tag{219}$$

The solution is not unique, because S' can rotate his axes about the vector v_ϱ. Let us seek a *symmetric* solution of the form

$$A_{\varrho\sigma} = a\delta_{\varrho\sigma} + b\, v_\varrho v_\sigma/v^2; \tag{220}$$

this makes

$$\begin{aligned} A_{\varrho\sigma} A_{\varrho\tau} &= a^2 \delta_{\sigma\tau} + (2ab + b^2) v_\sigma v_\tau/v^2, \\ A_{\varrho\sigma} v_\sigma &= (a + b) v_\sigma, \end{aligned} \tag{221}$$

and (219) are satisfied provided a and b satisfy

$$a^2 = 1, \quad 2ab + b^2 = \gamma^2 v^2/c^2, \quad a + b = \gamma. \tag{222}$$

These three equations are consistent, since, from the definition of γ,

$$1 + \gamma^2 v^2/c^2 = \gamma^2. \tag{223}$$

Hence (219) have the solutions

$$A_{\varrho\sigma} = a\delta_{\varrho\sigma} + (\gamma - a) v_\varrho v_\sigma/v^2, \tag{224}$$

where $a = \pm 1$. To get the identical transformation when $v_\varrho = 0$, we choose $a = 1$, and so we have for the matrix of the Lorentz transformation

$$\begin{aligned} &A_{\varrho\sigma} = \delta_{\varrho\sigma} + (\gamma - 1) v_\varrho v_\sigma/v^2, \quad A_{\varrho 4} = i\gamma v_\varrho/c, \\ &A_{4\sigma} = -i\gamma v_\sigma/c, \quad A_{44} = \gamma. \end{aligned} \tag{225}$$

A Hermitian matrix H_{rs} satisfies $\bar{H}_{rs} = H_{sr}$, the bar denoting the complex conjugate. Since $A_{\varrho\sigma}$ and A_{44} are real and $A_{\varrho 4}$ and $A_{4\varrho}$ pure imaginaries, the conditions that A_{rs} be Hermitian read

$$A_{\varrho\sigma} = A_{\sigma\varrho}, \quad A_{\varrho 4} = -A_{4\varrho}. \tag{226}$$

We note that the matrix in (225) is Hermitian.

If A_{rs} is Hermitian, then the corresponding matrix for the transformation of the real coordinates x^r is symmetric. To see this, we have merely to look at (210).

Thus, by cooperation on the part of S' alone, we can make the matrix A_{rs} Hermitian and the corresponding real matrix symmetric.

The reader should have no difficulty in dealing with the following

problem: Let S and S' be two Galileian frames of reference such that the Lorentz transformation $x'_r = A_{rs}x_s$ connecting their Minkowskian coordinates has the matrix (225). Show that the space axes of S' appear to S, when viewed instantaneously ($x_4 = \text{const.}$), to be inclined to one another at angles α_{12}, α_{23}, α_{31}, the angle between Ox'_1 and Ox'_2 satisfying

$$\cos \alpha_{12} = -\frac{v_1 v_2/c^2}{(1 - v_1^2/c^2)^{\frac{1}{2}}\,(1 - v_2^2/c^2)^{\frac{1}{2}}}, \tag{227}$$

with similar equations for the other two angles. How do the space axes of S appear to S'?

CHAPTER V

APPLICATIONS OF THE LORENTZ TRANSFORMATION

§ 1. APPARENT CONTRACTION OF A MOVING BODY AND APPARENT RETARDATION OF A MOVING CLOCK

We are now about to make some applications of the simple formulae IV–(202) for a Lorentz transformation between two Galileian frames of reference:

$$\begin{aligned} x' &= \gamma(x - vt), \\ t' &= \gamma(t - vx/c^2), \\ y' &= y, \ z' = z, \\ \gamma &= (1 - v^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{1}$$

It will be recalled that v is the velocity of the frame S' relative to the frame S, positive when S' moves in the positive sense along the x-axis.

The results we are going to establish seem very queer to most people at first sight. That is because it is so hard for us, reared on the absolute time of Newton, to break away from that concept and recognise (as we must in relativity) that *there is no absolute simultaneity*. If the subscripts 1 and 2 refer to two events, then

$$t_2 - t_1 = 0 \tag{2}$$

does *not* imply

$$t_2' - t_1' = 0. \tag{3}$$

Suppose that S has under observation a body fixed in the frame of S', and wishes to measure its length without interfering with its motion. He adopts the natural procedure of noting the *simultaneous* positions of two particles of the body, and then, at his leisure, measures the distance between these two positions, the body of course having passed on. Note (and this is essential) that S will use S-simultaneity (2) and not S'-simultaneity (3).

For simplicity we shall take the body to be a rod lying along the

x'-axis, of length l' when measured by S'. The histories of its two ends may then be written

$$\begin{aligned} x_1' &= a, \qquad y_1' = 0, \; z_1' = 0, \\ x_1' &= a + l', \; y_2' = 0, \; z_2' = 0, \end{aligned} \tag{4}$$

where a is some constant. Let the two observations be made at time t. Then, by (1),

$$\begin{aligned} x_1' &= \gamma(x_1 - vt), \; y_1' = y_1, \; z_1' = z_1, \\ x_2' &= \gamma(x_2 - vt), \; y_2' = y_2, \; z_2' = z_2, \end{aligned} \tag{5}$$

if (x_1, y_1, z_1) and (x_2, y_2, z_2) are the two positions noted by S. Hence

$$x_2 - x_1 = (x_2' - x_1')/\gamma, \; y_1 = y_2 = z_1 = z_2 = 0, \tag{6}$$

and so the two positions lie on the x-axis at a distance l from one another, where

$$l = l'/\gamma = l'(1 - v^2/c^2)^{\frac{1}{2}}. \tag{7}$$

Therefore $l < l'$; *the length of the rod is apparently reduced in the ratio* $(1 - v^2/c^2)^{\frac{1}{2}}$ *on account of its motion.*

This apparent shortening is called the *FitzGerald-Lorentz contraction,* having been first put forward as an empirical assumption independently by G. F. FitzGerald and H. A. Lorentz, before the creation of the special theory of relativity.

Since (1) gives

$$y_2 - y_1 = y_2' - y_1', \; z_2 - z_1 = z_2' - z_1', \tag{8}$$

it follows that no such apparent change in length occurs for a rod lying at right angles to the x'-axis. Hence we see that a volume V' in the frame of S' undergoes an apparent contraction given by

$$V = V'/\gamma = V'(1 - v^2/c^2)^{\frac{1}{2}}, \tag{9}$$

when viewed by S, using simultaneous observations in the above manner.

Consider now a standard clock carried by S', its history being

$$x' = \text{const.}, \; y' = \text{const.}, \; z' = \text{const.} \tag{10}$$

We consider two events, E_1 and E_2, in the history of the clock. Let t_1 and t_2 be times the attributed to these events by S, and t_1' and t_2' the times attributed to them by S'. These last are actual readings taken from the face of the clock.

As judged by S' these two events have the coordinates

$$E_1\colon\ (x', y', z', t_1'), \quad E_2\colon\ (x', y', z', t_2'). \tag{11}$$

But by the inverse Lorentz transformation IV–(203) we have

$$t = \gamma(t' + vx'/c^2), \tag{12}$$

and so the times attributed by S to these events are related to the quantities in (11) by

$$\begin{aligned} t_1 &= \gamma(t_1' + vx'/c^2), \\ t_2 &= \gamma(t_2' + vx'/c^2). \end{aligned} \tag{13}$$

Therefore

$$t_2 - t_1 = \gamma(t_2' - t_1') = (t_2' - t_1')\,(1 - v^2/c^2)^{-\frac{1}{2}}. \tag{14}$$

Thus $t_2 - t_1 > t_2' - t_1'$. If $t_2' - t_1'$ is 60 sec., then S judges that more than 60 sec. have elapsed between the two events, basing his judgment on readings of his own standard clock. Thus the moving clock appears to be *slowed down* or *retarded*, the value of the retardation being given by (14).

In mentioning seconds here we have in mind a unit of time which is defined as a multiple of the period of some fundamental atomic vibration. A direct experimental verification of this retardation, using "atomic clocks", was made by IVES and STILWELL [1938, 1941].

§ 2. SNAPSHOTS

To get an intuitive grasp of the phenomena just discussed, we may construct space-time diagrams, and that we shall do in § 3. But there is a more primitive and more natural approach, and that is by forming mental moving pictures of what is going on. These moving pictures we break down into a sequence of "stills" or "snapshots", just as a cinematograph film is broken down into a sequence of separate photographs.

There is no need to tell the reader that physicists form such pictures in their minds, for he certainly forms such pictures himself, as others have done since the dawn of science and before that dawn. What is needed is a warning that in relativity the formation of such pictures must be undertaken a little more circumspectly, lest in their very formation some Newtonian non-relativistic concept should creep in.

In taking an ordinary photograph, there is some finite exposure. That we eliminate in our ideal *snapshot*, making the exposure zero.

In an ordinary photograph there is (theoretically at least) a distortion due to the finiteness of the velocity of light. The photographic plate records the object as it was when the light left it, and not as it is when the light reaches the plate. This is a complication we wish to avoid, and so our snapshots will be supposed to be taken *instantaneously*, as if light travelled with infinite velocity. This may seem a queer thing to do in relativity, where the finiteness of the velocity of light is so important, but the snapshot is merely a device to help our understanding. It must not be thought of as being an optical photograph, but as something which could be constructed only by a laborious synthesis on the part of a Galileian observer, who notes the (to him) simultaneous positions of particles and plots them in his frame of reference. The picture he obtains in this way (or the model, for he can give it three dimensions) is the same as he would obtain if he had an optically perfect camera and could use light (or some other signal) travelling with infinite velocity. It is not for a moment suggested that anything like that exists in nature.

Another point, and a very important one, is that there is no *absolute* observer who can, so to speak, take a snapshot of things "as they really are". The only possible observers are Galileian observers, and one is as good as another. The appearance of phenomena depends on the observer, and so, in speaking of any snapshot, we must be careful to say who "photographed" it. A snapshot is, in fact, a section $t =$ const. of space-time; MILNE [1935, p. 107] calls it a *world-map*.

Fig. 1 shows a snapshot made by S of a rod AB of unit length fixed in his frame and of a rod $A'B'$, also of unit length, carried by S' with velocity v relative to S. It is contracted according to (7). In Fig. 1 we have $A'B' = \frac{1}{2}AB$, and this corresponds to $\gamma = 2$, $v/c = \sqrt{3}/2 = 0.866$. If S waits and takes further snapshots, they will show AB the same in all of them, and $A'B'$ always of the same length but translated in the direction of its motion. If S' takes snapshots, they will show $A'B'$ fixed and AB moving and contracted, *because of course the FitzGerald-Lorentz contraction occurs (with the same contraction ratio) whether the two rods are viewed by S or by S'.* We cannot show both these contractions in a single snapshot, because

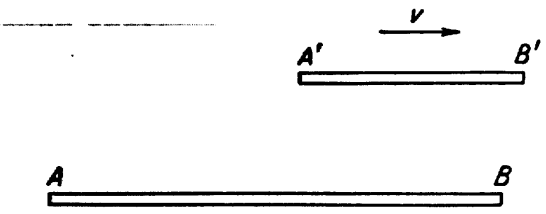


Fig. 1 – Snapshot by S of the FitzGerald-Lorentz contraction, drawn for $v/c = 0.866$

the snapshot must be taken by S or by S'. To enlist a third Galileian observer would not be helpful.

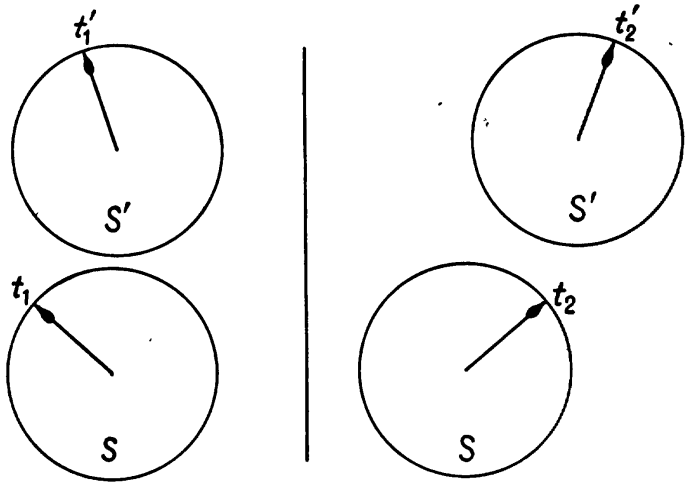


Fig. 2 – Two snapshots by S of an S-clock and an S'-clock. The hand of the S-clock has moved further on its dial than the hand of the S'-clock

Fig. 2 shows two snapshots made by S. On the left we have two clocks, one belonging to S and the other to S'. They are situated side by side, the S-clock reading t_1 and the S'-clock reading t_1'. On the right we have a second snapshot. The S'-clock has moved away. The S-clock reads t_2 and the S'-clock reads t_2'. In accordance with (14), the hand of the S-clock has moved further round the dial than the hand of the S'-clock. If S' had taken two snapshots of a clock fixed in the S-frame, together with a clock of his own, the hand of the S'-clock would have moved further, *for clock retardation, like the FitzGerald-Lorentz contraction, works both ways.* This reciprocity is due to the fact that the transformations IV–(202) and IV–(203) differ only in the sign attached to v, and the contraction and retardation depend on v^2.

§ 3. SPACE-TIME DIAGRAMS OF CONTRACTION AND RETARDATION

Let us now view the apparent contraction and retardation by means of space-time diagrams. The great advantage of a space-time diagram is this: a snapshot must be taken by someone, and what it shows depends on who that is, whereas a space-time diagram is absolute and impartial. For that reason Minkowski proposed that we should think in terms of the *absolute world* of space-time, rather than in the kinematical terms appropriate to the "snapshots" of § 2.

In dealing with the simple Lorentz transformation (1), in which the transformation of y and z is merely the identity, it is natural to make space-time diagrams by orthogonal projection on the 2-flat which contains the axes of x, ct, x', ct'. We take, then, a sheet of paper, and draw on it a pair of perpendicular axes, which we label OX, OT (Fig. 3). Any event (x, t) is plotted at P, where $AP = x$, $OA = ct$, AP being

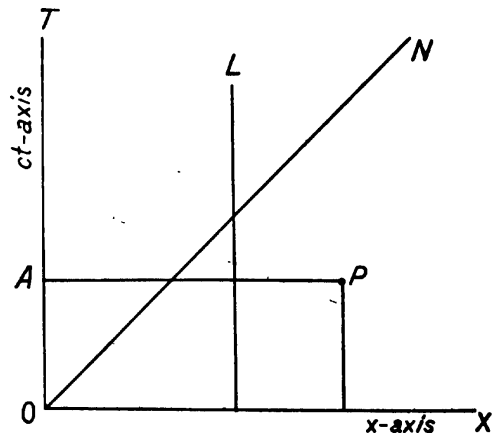


Fig. 3 – Space-time diagram showing OX and OT as axes of x and ct respectively

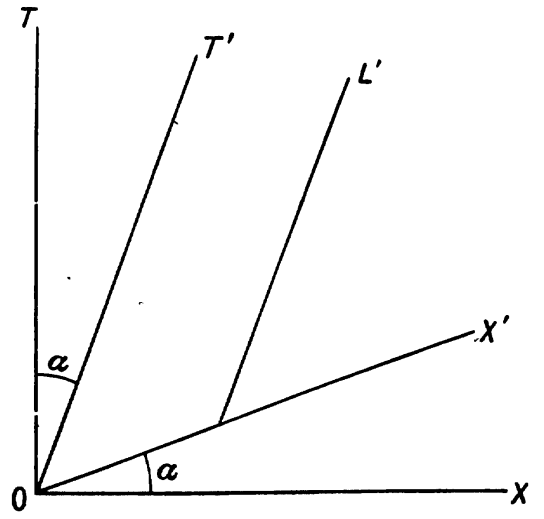


Fig. 4 – The axes of x, ct, x', ct' as they appear in the space-time diagram

drawn parallel to OX. Thus the x-axis in space-time is represented as OX and the ct-axis as OT.

The null line $x = ct$ appears as ON, inclined at 45° to OX and OT.

The histories of all particles fixed in the frame S appear as straight lines parallel to OT, like L in Fig. 3. Two events which are simultaneous, as judged by S, lie on a line parallel to OX (A and P, for example).

We have now to consider the x'-axis and the ct'-axis. On the x'-axis we have $t' = 0$, and so by (1) the equation of this axis is

$$ct/x = v/c; \tag{15}$$

thus the x'-axis appears in the space-time diagram as the line OX' (Fig. 4) making with OX the angle α such that[1]

$$\tan \alpha = v/c. \tag{16}$$

The ct'-axis, shown as OT', has the equation $x - vt = 0$, and makes with OT the same angle α. It is easily seen that for v positive (and we shall take it so), OX' and OT' both lie *between* OX and OT, as shown in the diagram.

The histories of particles fixed in the frame of S' appear as straight lines parallel to OT', as L' in Fig. 4; the criterion for S'-simultaneity is that the line joining the two events should be parallel to OX'.

How are we to plot an event for which (x', t') are assigned? The answer is given by (1), but we have to interpret it geometrically. In

[1] Remember the distinction made in III–§ 5 between M-geometry and E-geometry; the angle α is of course an E-angle.

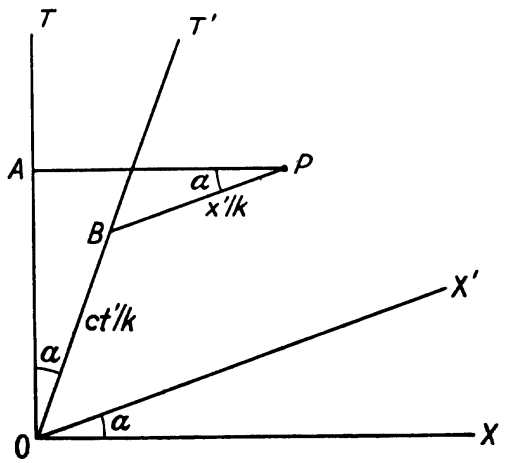


Fig. 5 – The coordinates x', ct' as modified oblique Cartesian coordinates

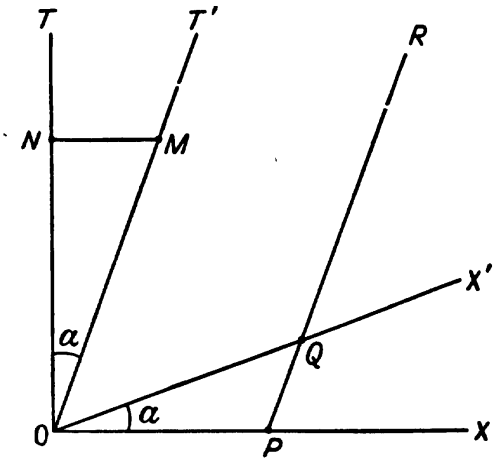


Fig. 6 – Space-time diagram for the FitzGerald-Lorentz contraction and clock-retardation

the following discussion all distances are E-distances, i.e. distances as measured by a ruler on the paper of the diagram.

Let P be the event (x', t') (Fig. 5). Then

$$x = AP, \quad ct = OA, \tag{17}$$

AP being drawn parallel to OX. Draw BP parallel to OX', with B on OT'. Then

$$\begin{aligned} x &= AP = OB \sin\alpha + BP \cos\alpha, \\ ct &= OA = OB \cos\alpha + BP \sin\alpha, \end{aligned} \tag{18}$$

and so, by (16),

$$\begin{aligned} BP &= \sec 2\alpha\,(x \cos\alpha - ct \sin\alpha) \\ &= \sec 2\alpha \cos\alpha\,(x - vt), \\ OB &= \sec 2\alpha\,(ct \cos\alpha - x \sin\alpha) \\ &= \sec 2\alpha \cos\alpha\,(ct - vx/c). \end{aligned} \tag{19}$$

It follows from (1) that

$$\begin{aligned} x' &= \gamma(x - vt) = kBP, \\ ct' &= \gamma(ct - vx/c) = kOB, \end{aligned} \tag{20}$$

where

$$k = \gamma \cos 2\alpha \sec\alpha = \sqrt{\frac{1 - v^2/c^2}{1 + v^2/c^2}}. \tag{21}$$

The result (20) is interesting. *It means that (x', ct') are the oblique Cartesian coordinates of P for the axes $OX'T'$, modified by the factor k.*

Let us now consider the FitzGerald-Lorentz contraction in terms of the space-time diagram. In Fig. 6, OT' and PQR are the histories of the

ends of a rod fixed in the frame S'. By (20) its length, as judged by S', is

$$l' = k\,OQ. \tag{22}$$

The other observer S takes simultaneous observations of its ends at the events O and P, and the length, as judged by him, is

$$l = OP. \tag{23}$$

But by the Euclidean trigonometry of the triangle OPQ, we have

$$OP/OQ = \sin(90° - 2\alpha)/\sin(90° - \alpha) = \cos 2\alpha \sec \alpha = k/\gamma, \tag{24}$$

using (21), and so, by (22) and (23),

$$l/l' = OP/(kOQ) = \gamma^{-1} = (1 - v^2/c^2)^{\frac{1}{2}}, \tag{25}$$

which is the FitzGerald-Lorentz contraction (7).

Consider now a clock carried by S' with OT' for world line (Fig. 6). By (20) the time interval from O to M, as measured by S', is

$$t' = k\,OM/c. \tag{26}$$

The events O and M are also observed by S, and he records the time interval between them as

$$t = ON/c, \tag{27}$$

where MN is drawn parallel to OX to meet OT. Now

$$ON/OM = \cos \alpha, \tag{28}$$

and so, by (26) and (27) with the help of (21),

$$\begin{aligned} t/t' &= k^{-1} \cos \alpha \\ &= \gamma^{-1} \sec 2\alpha \cos^2 \alpha \\ &= \gamma^{-1}(1 - \tan^2 \alpha)^{-1} \\ &= \gamma^{-1}(1 - v^2/c^2)^{-1} \\ &= \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{29}$$

This agrees with the retardation (14).

It may seem curious that, in terms of the Euclidean metric of the paper, we have $OM > ON$ in Fig. 6, since this would appear to indicate an *acceleration* of the moving clock instead of a *retardation*. But to make such an interpretation would be to assign to the Euclidean metric a significance it does not possess. What we have to compare (and what we have actually compared above) are the Minkowskian separations OM, ON, and we have

$$\begin{aligned} OM_m^2 &= ON_m^2 - NM_m^2 \\ &= ON_e^2 - NM_e^2 \\ &< ON_e^2 = ON_m^2, \end{aligned} \tag{30}$$

where the subscripts m and e indicate Minkowskian and Euclidean measures respectively.

The above derivation of the contraction and retardation shows that, on account of the intrusion of the Euclidean metric of the paper, it would be easy to get confused in a space-time diagram when the diagram is used for the calculation of quantitative results. It can be used in this way, but on the whole its chief function is as a source of suggestion and a rough guide to precise formal calculations.

§ 4. COMPOSITION OF VELOCITIES

Suppose that a particle moves along the x'-axis with a velocity u' as measured by S'. What is its velocity as measured by S?

In Newtonian mechanics the answer would be

$$u = u' + v, \tag{31}$$

the sum of the velocity u' of the particle relative to S' and the velocity v of S' relative to S. But that formula is not correct in relativity. To find the correct formula, we put

$$u = dx/dt, \quad u' = dx'/dt', \tag{32}$$

for the velocities of the particle relative to S and S' respectively. We then appeal to the Lorentz transformation (1), or rather to the inverse transformation IV–(203); this gives

$$u = \frac{dx}{dt} = \frac{dx' + vdt'}{dt' + vdx'/c^2} = \frac{u' + v}{1 + u'v/c^2}. \tag{33}$$

This is the *relativistic rule for the composition of velocities*, in the case where the particle under observation moves in the direction of the relative velocity of the two observers (the x-axis). We shall later discuss more general cases, but for the present let us explore the simple and interesting formula (33).

This formula may be expressed in another way by introducing parameters χ, ϕ, ϕ' defined by

$$\tanh \chi = v/c, \quad \tanh \phi = u/c, \quad \tanh \phi' = u'/c. \tag{34}$$

This is the same χ as we met in IV–(192), the pseudoangle of rotation of the Lorentz transformation (1). We can of course give similar meanings to ϕ and ϕ', if we introduce a third Galileian observer riding on the particle.

Substitution from (34) in (33) gives

$$\tanh \phi = \tanh (\phi' + \chi), \quad \phi = \phi' + \chi, \tag{35}$$

and so we may say that *velocities are compounded by adding the corresponding pseudoangles.*

The formula (33), regarded as a transformation $u' \to u$, has an interesting property, implicit in (35). For we know that $|\tanh \theta| \leqslant 1$ for all real values of θ, the equality sign holding only in the limits $\theta \to \pm \infty$. This means that no matter what values we give to u' and v, subject only to

$$|u'| \leqslant c, \quad |v| \leqslant c,$$

then the value of $|u|$ given by (33) cannot exceed c; it becomes equal to c if $|u'| = c$ or $|v| = c$. For example, if $u' = c$, then by (33)

$$u = \frac{c + v}{1 + cv/c^2} = c. \tag{36}$$

Thus *a particle which travels with the velocity of light relative to S' also travels with the velocity of light relative to S,* a result known to us already, of course, by the invariance of $dx^2 - c^2dt^2$.

Indeed we may note that

$$c^2dt'^2 - dx'^2 = c^2dt^2 - dx^2 \tag{37}$$

may be written

$$c^2dt'^2(1 - u'^2/c^2) = c^2dt^2(1 - u^2/c^2), \tag{38}$$

and this shows us directly that $u'^2 < c^2$ implies $u^2 < c^2$, and also that $u'^2 = c^2$ implies $u^2 = c^2$.

As an illustration of the composition of velocities, consider the situation shown in Fig. 7; it is an S-snapshot of two particles moving towards one another, each with the speed u. What is their *relative velocity*?

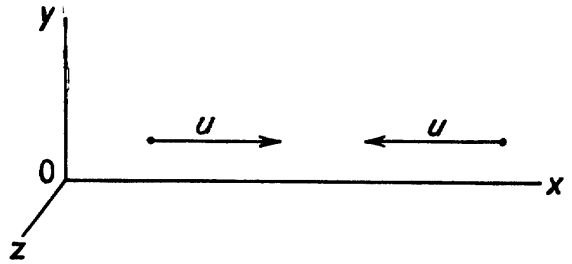


Fig. 7 – S-snapshot of two particles approaching one another. What is their relative velocity?

It is true that, under such circumstances, the observer S might find it convenient to *define* the relative velocity as the difference of the velocities he observes, i.e. $u - (-u) = 2u$. But that is not what we have in

mind. Rather let us imagine that a second Galileian observer S' rides on one of the particles, and in that situation observes the velocity of the other.

Let S' ride on the particle on the left. Then, as in (1), we have the transformation

$$\begin{aligned} x' &= \gamma(x - ut), \\ t' &= \gamma(t - ux/c^2), \\ \gamma &= (1 - u^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{39}$$

Consider now the motion of the other particle. For it we have

$$dx/dt = -u, \tag{40}$$

and so, by (39),

$$\frac{dx'}{dt'} = \frac{dx - udt}{dt - udx/c^2} = -\frac{2u}{1 + u^2/c^2}. \tag{41}$$

This is the relative velocity of the two particles in the relativistic sense. We verify immediately that its magnitude cannot exceed c by noting the algebraic inequality

$$1 + u^2/c^2 \geqslant 2u/c. \tag{42}$$

Although our Newtonian intuition may continue to insist that if we let u approach the value c in Fig. 7, then *surely* the relative velocity must approach $2c$, in relativity it just is not so. For example, if $u = 9c/10$, then

$$\frac{dx'}{dt} = -\frac{180}{181}c. \tag{43}$$

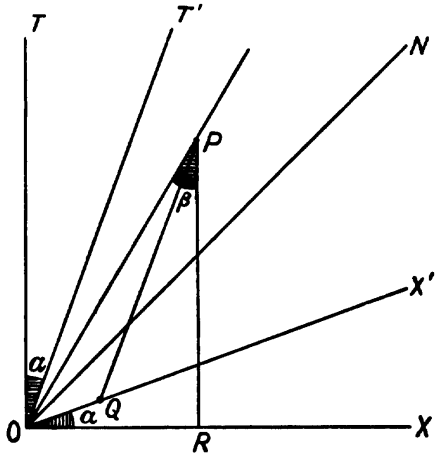


Fig. 8 – Space-time diagram for the composition of velocities

Let us now consider the composition of velocities with the aid of a space-time diagram (Fig. 8). This diagram is essentially Fig. 4, with the addition of OP to represent the world line of the moving particle and the null line ON to remind us that, since $u < c$, OP must make with OX an angle greater than 45°.

Draw PQ parallel to OT', with Q on OX', and draw PR parallel to OT, with R on OX. Denote the angle OPR by β. Then it is easy to see, if we recall

(20), that

$$u'/c = OQ/QP, \quad u/c = \tan\beta. \tag{44}$$

(The measures are Euclidean.) We have also $\tan\alpha = v/c$, and so by elementary trigonometry

$$\frac{u'}{c} = \frac{OQ}{QP} = \frac{\sin(\beta-\alpha)}{\cos(\beta+\alpha)} = \frac{\tan\beta - \tan\alpha}{1 - \tan\beta\tan\alpha} = \frac{u/c - v/c}{1 - uv/c^2}, \tag{45}$$

which is the same as (33), written in the equivalent form

$$u' = \frac{u - v}{1 - uv/c^2}. \tag{46}$$

So far we have considered only the case where the two Galileian observers are connected by the simple Lorentz transformation (1), and the particle under observation moves along the common x-axis. Let us now remove these restrictions, taking a general Lorentz transformation and allowing the particle to have, relative to S', a velocity with components

$$u'_\varrho = dx'_\varrho/dt'. \tag{47}$$

We take the Lorentz transformation $x_r = A_{sr}x'_s$, as in IV–(35), or, more explicitly,

$$\begin{aligned} x_\varrho &= A_{\sigma\varrho}x'_\sigma + A_{4\varrho}x'_4, \\ x_4 &= A_{\sigma 4}x'_\sigma + A_{44}x'_4. \end{aligned} \tag{48}$$

Then the velocity observed by S is

$$\begin{aligned} u_\varrho = dx_\varrho/dt &= ic\,dx_\varrho/dx_4 \\ &= ic\,\frac{A_{\sigma\varrho}dx'_\sigma + A_{4\varrho}ic\,dt'}{A_{\mu 4}dx'_\mu + A_{44}ic\,dt'} \\ &= ic\,\frac{A_{\sigma\varrho}u'_\sigma + icA_{4\varrho}}{A_{\mu 4}u'_\mu + icA_{44}}. \end{aligned} \tag{49}$$

This is the most general formula for the composition of velocities.

In the particular case of a Hermitian Lorentz transformation as in IV–(225), (49) becomes

$$\begin{aligned} u_\varrho &= \frac{u'_\varrho + (\gamma-1)v_\varrho v_\sigma u'_\sigma/v^2 + \gamma v_\varrho}{\gamma(1 + v_\mu u'_\mu/c^2)} \\ \gamma &= (1 - v^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{50}$$

Here v_ϱ is the relative velocity of the two frames, S and S'. If it is small, then $\gamma = 1$ approximately, and if we discard small quantities of the second order, (50) gives approximately

$$u_\varrho = \frac{u'_\varrho + v_\varrho}{1 + v_\mu u'_\mu / c^2}. \tag{51}$$

Note that it is not assumed that u'_ϱ is small. It is interesting to compare (51) with (33).

We recall that the Hermitian Lorentz transformation IV–(225) results from a special choice of space axes by one of the observers.

§ 5. THE VELOCITY 4-VECTOR AND THE ACCELERATION 4-VECTOR

The formula (49) for the transformation of velocity u_ϱ is unpleasantly complicated. It is a projective transformation, and as such much harder to handle than the simple affine transformation for a 4-vector,

$$V'_r = A_{rs} V_s. \tag{52}$$

To gain simplicity in relativity we must cease to regard the ordinary velocity u_ϱ (the 3-*velocity*, we may call it) as fundamental, and substitute a 4-vector for it. Accordingly we define the *velocity 4-vector* or *4-velocity* as

$$\lambda_r = dx_r/ds, \tag{53}$$

ds being the element of separation along the world line of the particle under consideration. It is obvious that this is indeed a 4-vector, transforming like V_r in (52) when we change the frame of reference. *The 4-velocity is in fact simply the unit tangent vector to the world line of the particle.*

Since the world line of a particle must be timelike (we think only of material particles and not of photons, for which $ds = 0$), we have

$$ds^2 = -\, dx_r dx_r, \tag{54}$$

and so the unit character of λ_r is expressed by

$$\lambda_r \lambda_r = -\, 1. \tag{55}$$

Thus, although λ_r has four components, the fourth component (always taken pointing into the future) is expressed in terms of the other three by

$$\lambda_4 = i(1 + \lambda_\varrho \lambda_\varrho)^{\frac{1}{2}}. \tag{56}$$

When we think in terms of space-time, we think in terms of the 4-

velocity λ_r; when we think in terms of space and time, we think in terms of the 3-velocity u_ϱ. The connection is very simple. We have

$$ds = (c^2dt^2 - dx_\varrho dx_\varrho)^{\frac{1}{2}} = cdt(1 - u^2/c^2)^{\frac{1}{2}}, \tag{57}$$

or

$$c\,dt = \gamma\,ds,\ \gamma = (1 - u^2/c^2)^{-\frac{1}{2}}. \tag{58}$$

Hence

$$\boxed{\begin{aligned} \lambda_\varrho &= \frac{dx_\varrho}{ds} = \frac{\gamma}{c}\frac{dx_\varrho}{dt} = \frac{\gamma u_\varrho}{c}, \\ \lambda_4 &= \frac{dx_4}{ds} = i\gamma. \end{aligned}} \tag{59}$$

These formulae are important, because they form the link between the space-time and the kinematical viewpoints.

We may resolve the vector λ_r into two vectors as follows:

$$(\lambda_\varrho, \lambda_4) = (\gamma u_\varrho/c, 0) + (0, i\gamma). \tag{60}$$

This resolution enables us to show the velocity 4-vector as in Fig. 9. We see the unit vector tangent to the world line C at A resolved into a vector $\gamma u_\varrho/c$ in the observer's space and a vector of magnitude γ parallel to the time-axis.

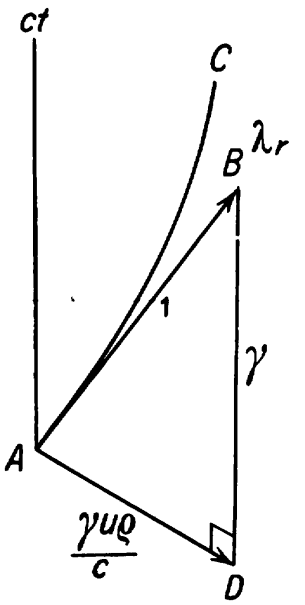


Fig. 9 – Space-time diagram of the 4-velocity λ_r

This resolution suggests that, if we insist on thinking kinematically, we should regard as fundamental, not the 3-velocity u_ϱ, but the 3-vector

$$\lambda_\varrho = \gamma u_\varrho/c, \tag{61}$$

the components of which are the projections of the 4-velocity on the space axes of the observer. If we denote the magnitude of this 3-vector by λ, so that

$$\lambda = (\lambda_\varrho\lambda_\varrho)^{\frac{1}{2}}, \tag{62}$$

then

$$\begin{aligned} \lambda^2 = \gamma^2u^2/c^2 = \gamma^2(1 - \gamma^{-2}) &= \gamma^2 - 1, \\ \gamma &= (1 + \lambda^2)^{\frac{1}{2}}, \end{aligned} \tag{63}$$

and so we express u_ϱ in terms of λ_ϱ by

$$u_\varrho = \frac{c\lambda_\varrho}{(1 + \lambda^2)^{\frac{1}{2}}}. \tag{64}$$

We can use this idea to discuss the relative velocity of two Galileian observers, S and S', whose time-axes are given as the two unit vectors T_r, T'_r, respectively, referred to a third impartial Galileian observer. We project T'_r orthogonally on the 3-flat orthogonal to T_r, i.e. on the 3-space of S. If that projection has components λ_ϱ along the space axes of S, then the components along these axes of the velocity of S' relative to S are given by (64).

As a matter of fact, we would gain a clearer idea of the concept of *the relative velocity of two observers* by abandoning the kinematical viewpoint altogether. If we draw the 4-vectors T_r, T'_r to represent the 4-velocities of the particles fixed in the frames of S, S' respectively, then we may regard the 4-vector (Fig. 10)

$$V_r = T'_r - T_r \tag{65}$$

as the proper space-time representation of the velocity of S' relative to S, and $-V_r$ as the proper representation of the velocity of S relative to S'. But we shall not pursue this idea further.

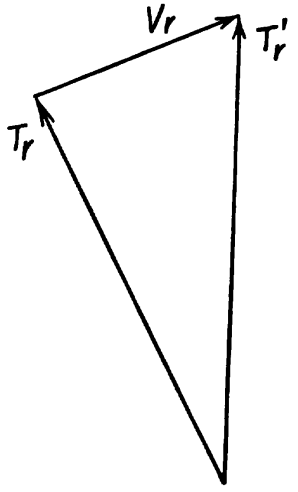


Fig. 10 – Space-time representation of the 4-velocity of S' relative to S

The *acceleration 4-vector* or *4-acceleration* of a particle is defined as

$$\frac{d\lambda_r}{ds} = \frac{d^2x_r}{ds^2}. \tag{66}$$

Hence, by (59), its components are expressed in terms of the 3-velocity and its rate of change as follows:

$$\begin{aligned} \frac{d\lambda_\varrho}{ds} &= \frac{d}{ds}\left(\frac{\gamma u_\varrho}{c}\right) = c^{-2}\gamma \frac{d}{dt}(\gamma u_\varrho), \\ \frac{d\lambda_4}{ds} &= \frac{d}{ds}(i\gamma) = ic^{-1}\gamma \frac{d\gamma}{dt}. \end{aligned} \tag{67}$$

Since λ_r is a unit vector, satisfying (55), we have

$$\lambda_r \frac{d\lambda_r}{ds} = 0. \tag{68}$$

Thus the 4-velocity and the 4-acceleration are not independent. When we substitute in (68) from (59) and (67), we get

$$u_\varrho \frac{d}{dt}(\gamma u_\varrho) - c^2 \frac{d\gamma}{dt} = 0. \tag{69}$$

This is of course merely an identity, which we can easily verify by using $\gamma = (1 - u^2/c^2)^{-\frac{1}{2}}$. But it is an important identity for mechanics, in connection with the concept of energy, as we shall see in Ch. VI.

It will be noticed that the 4-velocity is dimensionless and that the 4-acceleration has the dimensions $[L^{-1}]$. If we thought it desirable to have the usual dimensions for velocity and acceleration, we could achieve this by modifying the definitions, incorporating the factor c in the definition of the 4-velocity and the factor c^2 in that of the 4-acceleration, or, equivalently, by differentiating the coordinates with respect to proper time τ [cf. II–(66)] instead of with respect to separation s.

§ 6. TRANSFORMATION OF A WAVE MOTION

The formula

$$\phi = \phi^{(0)} \cos 2\pi\nu \left(\frac{l_\varrho x_\varrho}{u} - t \right) \tag{70}$$

represents the propagation of a *scalar* disturbance in a Galileian frame S. Here $\phi^{(0)}$ is the constant amplitude, ν the frequency, u the phase velocity (or velocity of propagation), and l_ϱ the direction cosines of the normals to the plane waves, drawn in the direction of propagation; we have

$$l_\varrho l_\varrho = 1 . \tag{71}$$

For a *vector* disturbance the corresponding formula would be

$$\phi_m = \phi_m^{(0)} \cos 2\pi\nu \left(\frac{l_\varrho x_\varrho}{u} - t \right), \tag{72}$$

where $\phi_m^{(0)}$ is a constant amplitude-vector, and for a *tensor* disturbance we would have

$$\phi_{mn} = \phi_{mn}^{(0)} \cos 2\pi\nu \left(\frac{l_\varrho x_\varrho}{u} - t \right), \tag{73}$$

where $\phi_{mn}^{(0)}$ is a constant amplitude-tensor.

For the purposes of the present discussion, it does not matter which of the three types of disturbance we consider, because it is the phase factor that interests us, not the amplitude. Let us take the scalar disturbance (70), thinking of the waves as sound waves. In relativity light waves, as we shall see later, are of the type (73). What we have to say about the phase factor in (70) is applicable to the phase factor in

(73), and so we shall be able to apply our conclusions to light and other forms of electromagnetic radiation as well as to sound.

We want to find out how the disturbance (70) appears to a second Galileian observer S', whose Minkowskian coordinates are connected with those of S by

$$x_r = A_{sr}x'_s. \tag{74}$$

We start by writing (70) in terms of Minkowskian coordinates:

$$\phi = \phi^{(0)} \cos \frac{2\pi}{c} f_r x_r, \tag{75}$$

where

$$f_\varrho = l_\varrho \nu c/u, \quad f_4 = i\nu. \tag{76}$$

Then by (74), $\phi^{(0)}$ being invariant by assumption, we have

$$\phi = \phi^{(0)} \cos \frac{2\pi}{c} f_r A_{sr} x'_s = \phi^{(0)} \cos \frac{2\pi}{c} f'_r x'_r, \tag{77}$$

where

$$f'_s = A_{sr} f_r. \tag{78}$$

But this is the formula for the transformation of a vector in space-time, and so we recognise f_r as a 4-vector: we call it the *frequency vector*, since its fourth component is $i\nu$ by (76).

It is clear that f_r is orthogonal (in the space-time sense) to the 3-flats

$$f_s x_s = \text{const.}, \tag{79}$$

on each of which the disturbance ϕ is constant; these are in fact the *phase-planes* or 3-*waves* of the disturbance, the crests of the waves being given by

$$f_s x_s = nc, \tag{80}$$

where n is an integer.

If we change (77) back into the form (70) in terms of the space and time coordinates of S', we get the frequency ν', the phase velocity u', and the direction cosines l'_ϱ of the normals to the waves, all as observed by S'. As in (76) we have

$$f'_\varrho = l'_\varrho \nu' c/u', \quad f'_4 = i\nu'. \tag{81}$$

The simplest and most methodical way of making the transformations

$$\begin{aligned} &(\nu, u, l_\varrho) \to (\nu', u', l'_\varrho) \\ &(\nu', u', l'_\varrho) \to (\nu, u, l_\varrho) \end{aligned} \tag{82}$$

is by means of (78), which tells us in fact that f_r transforms like x_r.

Let us examine the transformation of a wave motion in the simplest possible case, S and S' being connected by the simple Lorentz transformation (1) and the direction of propagation being along the x'-axis for S'. It will be best to keep to Minkowskian coordinates, as used in the above discussion, and so we shall use, instead of the real Lorentz transformation (1), the matrix A_{rs} of IV-(201). We have then

$$\begin{aligned} f_1 &= A_{11}f'_1 + A_{41}f'_4 = \gamma(f'_1 - ivf'_4/c), \\ f_4 &= A_{14}f'_1 + A_{44}f'_4 = \gamma(ivf'_1/c + f'_4), \\ f_2 &= f'_2, \quad f_3 = f'_3, \end{aligned} \tag{83}$$

where v is the velocity of S' relative to S and $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$.

For propagation along the x'_1-axis, we have by (81)

$$f'_1 = \nu' c/u', \quad f'_2 = f'_3 = 0, \quad f'_4 = i\nu', \tag{84}$$

and so (83) gives

$$\begin{aligned} f_1 &= \gamma\nu'\left(\frac{c}{u'} + \frac{v}{c}\right), \\ f_4 &= i\gamma\nu'\left(1 + \frac{v}{u'}\right), \\ f_2 &= f_3 = 0. \end{aligned} \tag{85}$$

Thus the propagation is along the x_1-axis, as we would expect by symmetry, and since by (76) we have

$$f_1 = \nu c/u, \quad f_4 = i\nu, \tag{86}$$

it follows from (85) that *the frequency ν and phase velocity u for S are given in terms of the frequency ν' and phase velocity u' for S' by the formulae*

$$\nu = \nu'\gamma\left(1 + \frac{v}{u'}\right), \tag{87}$$

$$u = \frac{u' + v}{1 + u'v/c^2}. \tag{88}$$

There is nothing new in (88); it is the formula (33) for the composition of velocities. But we have not met (87) before; it is the formula for the *relativistic Doppler effect.* If we let c tend to infinity, (87) gives the usual

non-relativistic Doppler formula

$$\nu = \nu' \left(1 + \frac{v}{u'}\right). \tag{89}$$

The above results are valid no matter what the phase velocity u' may be. There is no reason to take it less than c, and indeed in the case of de Broglie waves, to which the above reasoning is applicable, u' exceeds c.

But the most interesting case is that of *light waves* in vacuo, and we shall devote the rest of this section to them. For light waves we have $u' = c$; then by (88) $u = c$ and (87) gives

$$\nu = \nu'\gamma\left(1 + \frac{v}{c}\right) = \nu' \sqrt{\frac{1 + v/c}{1 - v/c}}. \tag{90}$$

We note that though the Doppler effect for light is a first-order effect, depending on v/c, the relativistic correction, due to the presence of γ in (90), is a second-order effect, depending on $(v/c)^2$. The difficulty in testing relativity experimentally arises from the fact that so many relativistic corrections to Newtonian physics are of the second order, and consequently very small, since it is hard to get two Galileian frames with a relative velocity at all comparable with the velocity of light.

The frequency vector f_r for light is very important, for it is closely connected with the dynamics of radiation, as we shall see later in Chapter VI. Putting $u = c$ in (76), we have for its components

$$\boxed{f_\varrho = l_\varrho \nu, \; f_4 = i\nu.} \tag{91}$$

If space-time is full of plane waves of light, we are to think of these waves as defining a system of parallel vectors in space-time, with components as in (91).

For light, the vector f_r is a null vector, for we have

$$f_r f_r = \nu^2(l_\varrho l_\varrho - 1) = 0. \tag{92}$$

We have already seen that in general (i.e. not for light only) f_r is orthogonal to the 3-waves (79). But the vector f_r for light has also the property of *lying in the 3-waves*. For the equation of the 3-wave through the origin of space-time is

$$f_r x_r = 0, \tag{93}$$

and this is satisfied by $x_r = f_r$ on account of (92).

We recognise (93) as the equation of the tangent 3-flat to the null cone $x_r x_r = 0$ at the event f_r. Thus *the 3-wave through the origin touches the null cone along the generator defined by* f_r (Fig. 11).

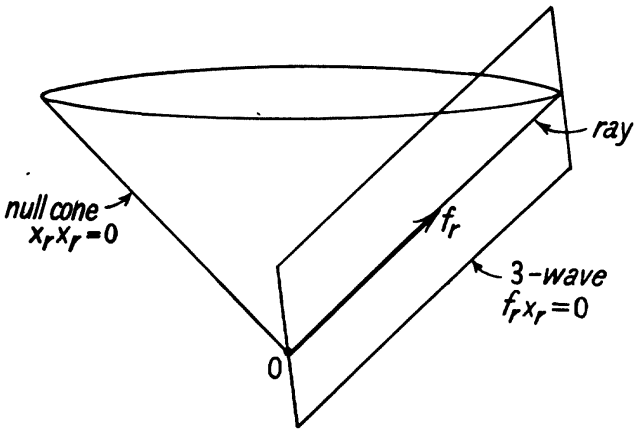


Fig. 11 – The null cone and the 3-wave tangent to it along the ray

An infinite straight line in space-time drawn in the direction of f_r is a null line. We may call it a *ray*. It represents the history of a point, carried forward with velocity c on a phase wave of the system

$$\phi = \phi^{(0)} \cos 2\pi\nu \left(\frac{l_\varrho x_\varrho}{c} - t \right) = \phi^{(0)} \cos \frac{2\pi}{c} f_r x_r . \qquad (94)$$

We have been thinking about light, but actually our subject has been what we may call the kinematics of waves in space-time, first for plane waves propagated with arbitrary velocity and then for plane waves propagated with the fundamental velocity c. These last we called light waves but we might as well have called them radio waves, for all electromagnetic waves in vacuo are propagated with velocity c (see Ch. IX). Thus what has been said applies to radio waves (long and short), infra-red waves, visible light, ultra-violet light, X-rays and γ-rays; these are to be regarded as differing from one another only in respect of amplitude and frequency. If we leave amplitude out of account, it is only by their differing frequencies that we distinguish these various radiations from one another. But is it a real distinction? It hardly seems so, since (90) tells us that the frequency of any given radiation changes with the frame of reference, taking all values from zero to infinity. In that sense, *all the types of radiation mentioned above are essentially the same from a relativistic standpoint, for any one can be changed into any other by a suitable change of frame of reference.* One man's X-ray is another man's visible light, so to speak.

An interesting case arises when the frequency vectors $f_r^{(1)}$, $f_r^{(2)}$ of two radiations have the same null direction in space-time (Fig. 12), this being an absolute (Lorentz-invariant) relationship. By (91) it is easy to see that every Galileian observer reports that the two radiations are

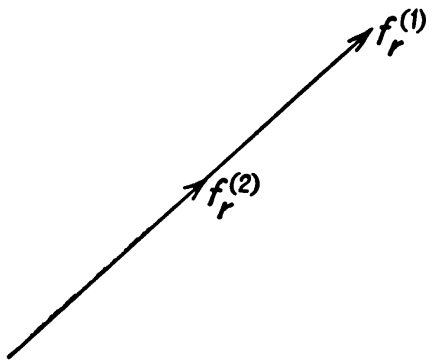


Fig. 12 – Two frequency 4-vectors with a common direction in space-time

propagated in the same spatial direction ($l_\varrho^{(1)} = l_\varrho^{(2)}$), but that they have different frequencies ($\nu^{(1)} \neq \nu^{(2)}$). This distinction in frequency is absolute, the ratio $\nu^{(1)}/\nu^{(2)}$ being preserved under Lorentz transformation. The conclusion is that, although a single radiation cannot be labelled for all Galileian observers (as visible light or radio waves, for example), the *distinction* between two radiations is absolute in the sense that all Galileian observers would find the same ratio for the two frequencies.

The relationship of two null vectors to one another is further discussed in VI–§ 16, and plane sinusoidal electromagnetic waves, satisfying Maxwell's equations, are treated in IX–§ 10.

§ 7. REFLECTION AT MOVING MIRRORS

When light is reflected by a fixed plane mirror, there is no change in frequency, and the reflection takes place according to the simple law

$$\text{angle of reflection} = \text{angle of incidence.}$$

Accepting this as valid for a mirror fixed in a Galileian frame of reference, we can find how light behaves when reflected from a mirror which has a uniform motion of translation. The device (which can be used in other problems also) is to use a frame of reference S' in which the mirror is fixed, write down the formulae for the phenomenon as viewed by S', and then pass by a Lorentz transformation to a description of the phenomenon as viewed by a second Galileian observer S, relative to whom the mirror is in motion.

Let us carry this out in a special case as an exercise. S observes light of frequency ν travelling in the direction of the x_1-axis (Fig. 13). For these waves the frequency vector is, by (91),

$$f_1 = \nu, \quad f_2 = f_3 = 0, \quad f_4 = i\nu. \qquad (95)$$

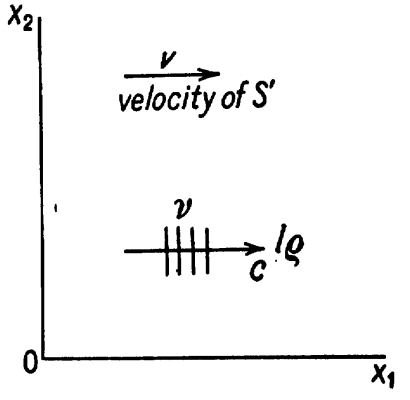


Fig. 13 – S observes incident light with wave-normal l_ϱ and frequency ν

Let S' move relative to S along the x_1-axis with velocity v. We have then the formulae of transformation as in (83):

$$\begin{aligned} f_1 &= \gamma(f'_1 - ivf'_4/c), \\ f_4 &= \gamma(ivf'_1/c + f'_4), \\ f_2 &= f'_2,\ f_3 = f'_3, \end{aligned} \tag{96}$$

and also

$$\begin{aligned} f'_1 &= \gamma(f_1 + ivf_4/c), \\ f'_4 &= \gamma(-ivf_1/c + f_4). \end{aligned} \tag{97}$$

(To pass quickly from (96) to (97), interchange f and f', and replace v by $-v$). For the light (95), as viewed by S', we have therefore

$$\begin{aligned} f'_1 &= \gamma\nu(1 - v/c), \\ f'_4 &= i\gamma\nu(1 - v/c), \\ f'_2 &= f'_3 = 0. \end{aligned} \tag{98}$$

In (95) and (98) we have two alternative descriptions of the incident light.

Now suppose that there is a mirror fixed in the S'-frame. We shall for simplicity take it parallel to the x'_3-axis and inclined at 45° to the x'_1-axis (Fig. 14). This mirror appears to S to be moving in the x_1-direction. (It does *not* appear to S to be inclined at 45° to the x_1-axis.)

Let f'_r be the frequency vector for S' of the light after reflection at the mirror. By the simple law of reflection for a fixed mirror, we have

$$f'_1 = 0,\quad f'_2 = f'_1,\quad f'_3 = 0,\quad f'_4 = f'_4, \tag{99}$$

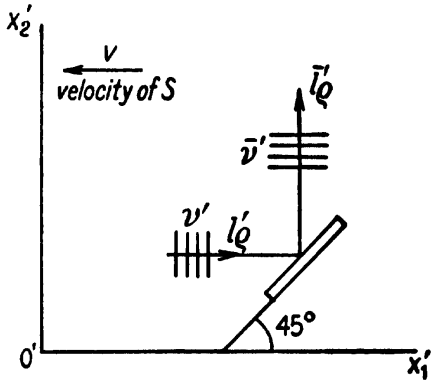


Fig. 14 – S' observes incident light with wave-normal l'_ϱ and frequency ν', and reflected light with wave-normal $\bar{l}'_\varrho$ and frequency $\bar{\nu}'$

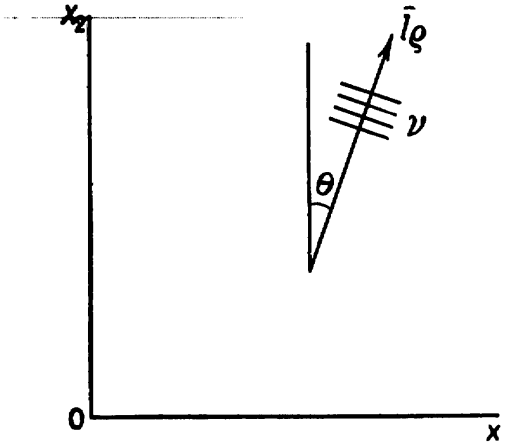


Fig. 15 – S observes reflected light with wave-normal $\bar{l}_\varrho$ and frequency $\bar{\nu}$

or by (98)

$$\bar{f}_1' = 0, \quad \bar{f}_2' = \gamma\nu(1 - v/c), \quad \bar{f}_3' = 0, \quad \bar{f}_4' = i\gamma\nu(1 - v/c). \tag{100}$$

We now transform back to S, to see how the reflected light appears to him. Let its frequency vector be $\bar{f}_r$. Then by (96)

$$\begin{aligned} \bar{f}_1 &= \gamma(\bar{f}_1' - iv\bar{f}_4'/c) = \gamma^2\nu\frac{v}{c}\left(1 - \frac{v}{c}\right), \\ \bar{f}_2 &= \bar{f}_2' = \gamma\nu\left(1 - \frac{v}{c}\right), \\ \bar{f}_3 &= 0, \\ \bar{f}_4 &= \gamma(iv\bar{f}_1'/c + \bar{f}_4') = i\gamma^2\nu\left(1 - \frac{v}{c}\right). \end{aligned} \tag{101}$$

If $\bar{l}_\varrho$ is the unit normal to the reflected wave and $\bar{\nu}$ its frequency (as judged by S), we have by (91)

$$\bar{\nu} = \nu\gamma^2\left(1 - \frac{v}{c}\right) = \frac{\nu}{1 + v/c}, \tag{102}$$

$$\bar{l}_1 = i\bar{f}_1/\bar{f}_4 = \frac{v}{c}, \quad \bar{l}_2 = i\bar{f}_2/\bar{f}_4 = \frac{1}{\gamma}, \quad \bar{l}_3 = 0. \tag{103}$$

The frequency of the reflected light is given by (102); as for its direction of propagation, it follows from (103) that it makes with the x_2-axis an angle θ (Fig. 15) such that

$$\tan\theta = \frac{\bar{l}_1}{\bar{l}_2} = \frac{\gamma v}{c} = \frac{v/c}{\sqrt{(1 - v^2/c^2)}}. \tag{104}$$

The change in frequency and the angle θ are both of order v/c when this ratio is small.

How do the above results compare with the corresponding results in Newtonian physics? Fifty years ago that would have been an important question, because at that time most physicists thought in the Newtonian way. Now it is only of historical interest and need not detain us. It suffices to say that Newtonian physics gives no simple formulae like (102) and (103), involving only the observer S and the mirror; Newtonian physics requires an *ether*, relative to which the velocity of light is c, and so offers an infinity of solutions to the problem of reflection at a moving mirror, depending on the choice of the ether. Relativity eliminates this "absolute" frame of reference, using instead an "absolute" space-time.

As a second problem, suppose that S' sets up a pair of parallel plane mirrors, normal to the x_1'-axis and facing one another, and arranges to have light oscillating to and fro between the two mirrors. If the disturbances are (we take a scalar[1] disturbance for simplicity)

$$\begin{aligned} &\phi^{(0)} \cos \frac{2\pi}{c} f_1^{(1)\prime} x_r', \quad \phi^{(0)} \cos \frac{2\pi}{c} f_1^{(2)\prime} x_r', \\ &f_1^{(1)\prime} = \nu', \quad f_1^{(2)\prime} = -\nu', \\ &f_4^{(1)\prime} = i\nu', \quad f_4^{(2)\prime} = i\nu', \end{aligned} \tag{105}$$

for the waves passing in the two directions, then the total disturbance is

$$\begin{aligned} \phi &= \phi^{(0)} \cos \frac{2\pi\nu'}{c} (x_1' - ct') + \phi^{(0)} \cos \frac{2\pi\nu'}{c} (x_1' + ct') \\ &= 2\phi^{(0)} \cos 2\pi\nu' t' . \cos 2\pi\nu' x_1'/c, \end{aligned} \tag{106}$$

which represents standing waves. Let us see how these waves appear to a second Galileian observer S, relative to whom S' has a velocity v along the x_1-axis. By (96) we have for the frequency vectors $f_r^{(1)}$, $f_r^{(2)}$ relative to S

$$\begin{aligned} &f_1^{(1)} = \gamma\nu'(1 + v/c), \quad f_1^{(2)} = \gamma\nu'(-1 + v/c), \\ &f_4^{(1)} = i\gamma\nu'(1 + v/c), \quad f_4^{(2)} = i\gamma\nu'(1 - v/c), \end{aligned} \tag{107}$$

and the total disturbance is

$$\begin{aligned} \phi &= \phi^{(0)} \cos \frac{2\pi}{c} f_r^{(1)} x_r + \phi^{(0)} \cos \frac{2\pi}{c} f_r^{(2)} x_r \\ &= \phi^{(0)} \cos \frac{2\pi\nu'\gamma}{c} \left(1 + \frac{v}{c}\right)(x_1 - ct) \\ &\qquad + \phi^{(0)} \cos \frac{2\pi\nu'\gamma}{c} \left(1 - \frac{v}{c}\right)(x_1 + ct) \\ &= 2\phi^{(0)} \cos \frac{2\pi\nu'\gamma}{c} (x_1 - vt) . \cos \frac{2\pi\nu'\gamma}{c} \left(\frac{vx_1}{c} - ct\right). \end{aligned} \tag{108}$$

If we write this in the form

$$\phi = a \cos \frac{2\pi\nu'\gamma}{c} \left(\frac{vx_1}{c} - ct\right), \tag{109}$$

[1] See IX-§ 11 for Maxwellian theory.

it appears as a wave motion with velocity c^2/v, frequency $\nu'\gamma$, and varying amplitude

$$a = 2\phi^{(0)} \cos \frac{2\pi\nu'\gamma}{c} (x_1 - vt). \tag{110}$$

On the other hand, if we write it in the form

$$\phi = b \cos \frac{2\pi\nu'\gamma}{c} (x_1 - vt), \tag{111}$$

it appears as a wave motion with velocity v, frequency $\nu'\gamma v/c$, and varying amplitude

$$b = 2\phi^{(0)} \cos \frac{2\pi\nu'\gamma}{c} \left(\frac{vx_1}{c} - ct\right). \tag{112}$$

If v/c is small, then (110) becomes approximately

$$a = 2\phi^{(0)} \cos 2\pi\nu' x_1/c, \tag{113}$$

a varying amplitude depending only on position, and (112) becomes approximately

$$b = 2\phi^{(0)} \cos 2\pi\nu' t, \tag{114}$$

a varying amplitude depending only on time.

§ 8. FRESNEL'S CONVECTION COEFFICIENT

Consider a transparent medium, such as glass or water, at rest in a Galileian frame S'. If light of wave length λ' in vacuo enters this medium, it travels with velocity

$$u' = \frac{c}{n(\lambda')}, \tag{115}$$

where $n(\lambda')$ is the refractive index for vacuum wave length λ'. This is simply the *definition* of refractive index, the dependence of n on λ' indicating that the medium is dispersive.

Note that in (115) u' is the velocity *inside* the medium, but λ' is the wave length *outside* it. We shall put up with this source of possible confusion in order to conform to the practice of physicists who are accustomed to treat *wave length* as basic; if we used *frequency*, this possible confusion would not arise, since the frequency is the same inside and outside the medium.

Suppose now that the medium has a velocity v relative to a second Galileian observer S (Fig. 16). Our problem is to find the velocity u of

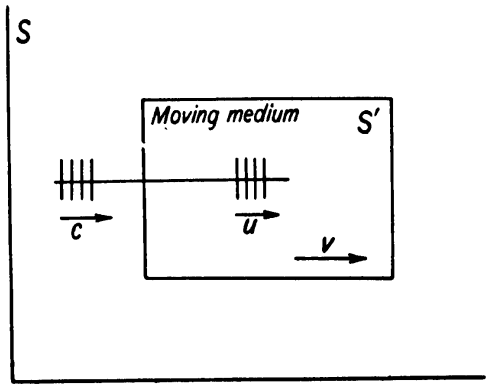


Fig. 16 – Snapshot of light passing through a moving medium

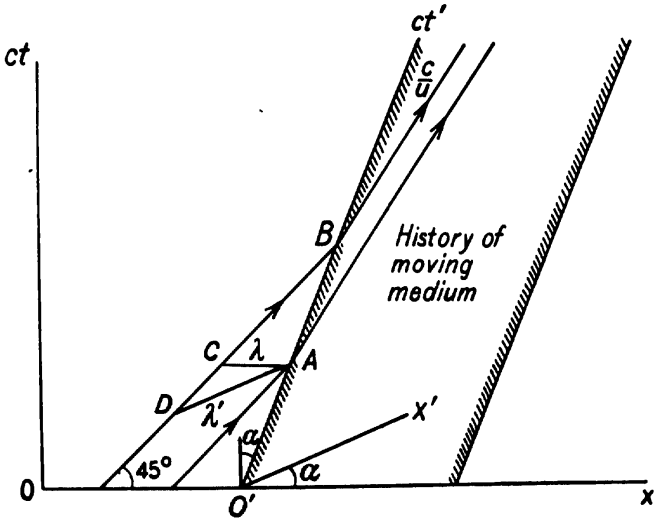


Fig. 17 – Space-time diagram of light passing through a moving medium (tan $\alpha = v/c$)

light in this moving medium, as measured by S. To specify the colour of the light, it is natural for S to use λ, the wave length in vacuo as measured by him.

Let us use a space-time diagram (Fig. 17). The axes of x and ct are shown. The shaded lines indicate the histories of the front and back of the medium; since its velocity is v, the slopes of these lines are $d(ct)/dx = c/v$. The diagram shows also two adjacent waves (i.e. histories of points of equal phase) entering the medium at the events A and B; before entering, the slope of each is unity, and, after entering, it is c/u, all these slopes being measured in the E-metric of the diagram.

The lines $O'x'$ and $O'ct'$ are drawn in the directions of the space-axis and the time-axis of S', the latter coinciding with the world line representing the history of the front of the medium. The slopes (in E-metric) are v/c and c/v (see Fig. 4, p. 123).

By the law of composition of velocities (33), we have

$$u = \frac{u' + v}{1 + u'v/c^2}, \tag{116}$$

where

v = velocity of medium relative to S,

u = velocity of light in medium relative to S,

u' = velocity of light in medium relative to S'.

Hence by (115)

$$u = \frac{\dfrac{c}{n(\lambda')} + v}{1 + \dfrac{v}{cn(\lambda')}}, \tag{117}$$

where $n(\lambda')$ is the refractive index of the medium for the *vacuum* wave length λ' (LAUE [1907]).

The formula just given is not quite what we want: we need a formula involving λ instead of λ'. These two lengths are shown in Fig. 17: $\lambda = (CA)_M$, $\lambda' = (DA)_M$, these being Minkowskian separations measured parallel to Ox, $O'x'$ respectively. We could obtain λ' in terms of λ by using the trigonometry of this Euclidean diagram, but it is simpler to refer to (90), which gives us, as relationship between the frequencies,

$$\nu = \nu' \sqrt{\frac{1 + v/c}{1 - v/c}}, \tag{118}$$

since we are at present dealing with light in vacuo.

Now $c = \lambda\nu = \lambda'\nu'$, and so

$$\frac{\lambda'}{\lambda} = \frac{\nu}{\nu'} = \sqrt{\frac{1 + v/c}{1 - v/c}}. \tag{119}$$

Thus *the exact expression for the velocity u of light of (vacuum) wave length λ in a medium of refractive index $n(\lambda)$, moving with velocity v, is given by*

$$\frac{u}{c} = \frac{\dfrac{1}{n(\lambda\phi)} + \dfrac{v}{c}}{1 + \dfrac{v}{cn(\lambda\phi)}}, \quad \phi = \sqrt{\frac{1 + v/c}{1 - v/c}}. \tag{120}$$

In actual experiments, the ratio v/c is small, and we may use an approximation. We have, approximately,

$$\lambda\phi = \lambda + \lambda v/c, \tag{121}$$

and so

$$\frac{1}{n(\lambda\phi)} = \frac{1}{n(\lambda)} + \lambda \frac{v}{c} \frac{d}{d\lambda} \frac{1}{n(\lambda)}. \tag{122}$$

Hence, neglecting v^2/c^2, we get from (120)

$$\frac{u}{c} = \frac{\dfrac{1}{n(\lambda)} + \lambda \dfrac{v}{c}\dfrac{d}{d\lambda}\dfrac{1}{n(\lambda)} + \dfrac{v}{c}}{1 + \dfrac{v}{c}\dfrac{1}{n(\lambda)}}$$

$$= \frac{1}{n(\lambda)} + \frac{v}{c}\left\{1 - \frac{1}{[n(\lambda)]^2} + \lambda \frac{d}{d\lambda}\frac{1}{n(\lambda)}\right\}. \tag{123}$$

This formula for the velocity of light in a moving medium may be written

$$u = \frac{c}{n(\lambda)} + kv, \tag{124}$$

where

$$\begin{aligned} k &= 1 - \frac{1}{[n(\lambda)]^2} + \lambda \frac{d}{d\lambda}\frac{1}{n(\lambda)} \\ &= 1 - \frac{1}{[n(\lambda)]^2} - \frac{\lambda}{[n(\lambda)]^2}\frac{dn(\lambda)}{d\lambda}. \end{aligned} \tag{125}$$

This quantity k is called Fresnel's "drag coefficient" or "convection coefficient" for the following reason. In (124), the first term represents the value u would have if the medium were at rest ($v = 0$). To this is added a certain fraction (k) of the velocity of the medium. Fresnel gave the formula (124) with

$$k = 1 - \frac{1}{[n(\lambda)]^2}, \tag{126}$$

and explained it in terms of a dragging of the ether by the moving medium, which imparted to the ether a fraction (k) of its velocity (v). This result is closely connected with the observed fact that aberration is independent of refractive index (see § 9).

The relativistic derivation of (125), as given above, is really very simple, although allowance for dispersion adds a complication. If we are prepared to treat n as a constant in (120), and neglect v^2/c^2, we get

$$u = \frac{c}{n} + kv, \quad k = 1 - \frac{1}{n^2}. \tag{127}$$

Ordinary media are not highly dispersive; for water at 20°C, n varies only from 1.3308 to 1.3428 as the wave length varies from 6708 A° to 4047 A°, roughly the range of the visible spectrum. Consequently the simple formula (127) gives fairly good agreement with observation.

The formula (125), which is due to LAUB [1908], is valid for a moving block, as in Fig. 16, and must be modified if the block is replaced by water flowing through a tube fixed in S, the S-velocity of the water being zero at the ends of the tube, where it enters and leaves laterally. For the block, the S'-frequencies are the same inside and outside; for the water, it is the S-frequencies that are the same. To modify (125) accordingly, go back to (117) and write $n(\nu')$ instead of $n(\lambda')$, ν' being the S'-frequency in the water. Then, by (87), the S-frequency in the water is $\nu'' = \nu'\gamma(1 + v/u')$; here $\gamma = 1$ to the first order in v/c. But $\nu'' = \nu$, the S-frequency of the incident light, and so $\nu' = \nu(1 - v/u')$. Substitute in (117), expand in powers of v/c, and finally use $\nu\lambda = c$ to replace ν by λ. The resulting formula for the convection coefficient in flowing water is[1]

$$k = 1 - \frac{1}{[n(\lambda)]^2} - \frac{\lambda}{n(\lambda)} \frac{d\, n(\lambda)}{d\, \lambda}. \tag{128}$$

The experimental verification of (128) by P. ZEEMAN [1914, 1916] is striking and here are his results:

Wave length of light in angstroms (1 A° = 10^{-8} cm)	Convection coefficient k calculated from (128)	Convection coefficient k observed
4500	0.464	0.465
4580	0.463	0.463
5461	0.454	0.451
6870	0.447	0.445

Earlier experiments had been made by Fizeau and by Michelson and Morley. For an historical account, see E. BUCHWALD [1951].

§ 9. ABERRATION

Another important application of the formulae for the transformation of the frequency vector is in connection with the phenomenon of astronomical aberration. Let there be two Galileian systems, S, S', with space axes so chosen that the transformation of coordinates is the special Lorentz transformation, the velocity of S' relative to S being v along the x_1-axis. S' has under observation a system of plane light

[1] This is the formula given by LORENTZ [1895, p. 101], LAUB [1907], PAULI [1920, p. 564], CUNNINGHAM [1921, p. 44], SILBERSTEIN [1924, p. 71], MØLLER [1952, p. 64]. I am indebted to Prof. H. P. Robertson for clarifying the relationship between (125) and (128).

waves in vacuo with wave-normal l'_ϱ and frequency ν'; the direction-cosines l'_ϱ determine the direction in which the observer in S' must point a telescope in order that the image of the distant source may be found on the cross-wires. We wish to find the direction in which S is to point his telescope, say l_ϱ.

By (83) and (91) we have

$$\begin{aligned} l_1\nu &= \gamma l'_1\nu' + \frac{\gamma v}{c}\nu', \\ l_2\nu &= l'_2\nu', \\ l_3\nu &= l'_3\nu', \\ \nu &= \frac{\gamma v}{c} l'_1\nu' + \gamma\nu'. \end{aligned} \tag{129}$$

We have then for the direction-cosines of the wave normal

$$\begin{aligned} l_1 &= \frac{\gamma\nu'}{\nu}\left(l'_1 + \frac{v}{c}\right) = \frac{l'_1 + v/c}{1 + l'_1 v/c}, \\ l_2 &= l'_2\frac{\nu'}{\nu} = \frac{l'_2}{\gamma(1 + l'_1 v/c)}, \\ l_3 &= l'_3\frac{\nu'}{\nu} = \frac{l'_3}{\gamma(1 + l'_1 v/c)}. \end{aligned} \tag{130}$$

We observe that if $l'_1 = 1$, so that $l'_2 = l'_3 = 0$, then the telescopes have the same direction in the two systems, namely along the x_1-axis. Otherwise the apparent directions will be different.

Let us suppose that the apparent S-direction of the source (i.e. the direction of the telescope from eye-piece to object glass) makes an angle θ with the x_1-axis, and that the apparent S'-direction makes an angle θ' with the x'_1-axis (Fig. 18). Then

$$\begin{aligned} l_1 &= -\cos\theta, \quad \sqrt{l_2^2 + l_3^2} = \sin\theta, \\ l'_1 &= -\cos\theta', \quad \sqrt{l'^2_2 + l'^2_3} = \sin\theta', \end{aligned} \tag{131}$$

and (130) gives

$$\cos\theta = \frac{\cos\theta' - v/c}{1 - (v/c)\cos\theta'}. \tag{132}$$

The phenomenon of aberration consists in the fact that $\theta \neq \theta'$, and (132) is the *relativistic formula for aberration*, obtained without approximation. The difference $\theta - \theta'$ is the *angle of aberration*.

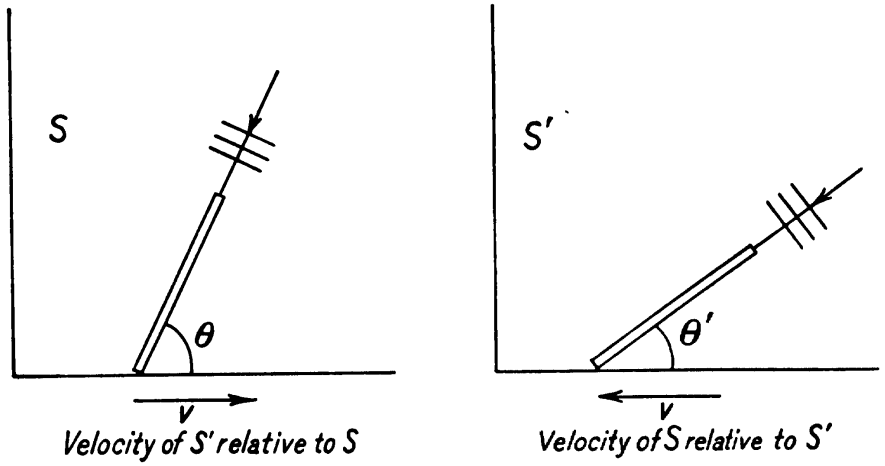


Fig. 18 – Snapshots of telescope and light waves made by S and S'

The formula (132) may also be written

$$\cot\theta = \frac{\gamma(\cos\theta' - v/c)}{\sin\theta'}, \quad \gamma = \frac{1}{\sqrt{(1 - v^2/c^2)}}. \tag{133}$$

It is easy to see from Fig. 19, which shows the vector composition of velocities, that on a Newtonian basis we have

$$\begin{aligned} \cot\theta' &= \frac{\cos\theta + v/c}{\sin\theta} \text{ if } S \text{ is fixed in the ether,} \\ \cot\theta &= \frac{\cos\theta' - v/c}{\sin\theta'} \text{ if } S' \text{ is fixed in the ether.} \end{aligned} \tag{134}$$

For observational purposes, the frames of reference S and S' consist of the earth itself at two different positions in its orbit round the sun. Since aberration is a small effect, depending on v/c, we may neglect the square of this ratio and replace γ in (133) by unity. Then (133) and

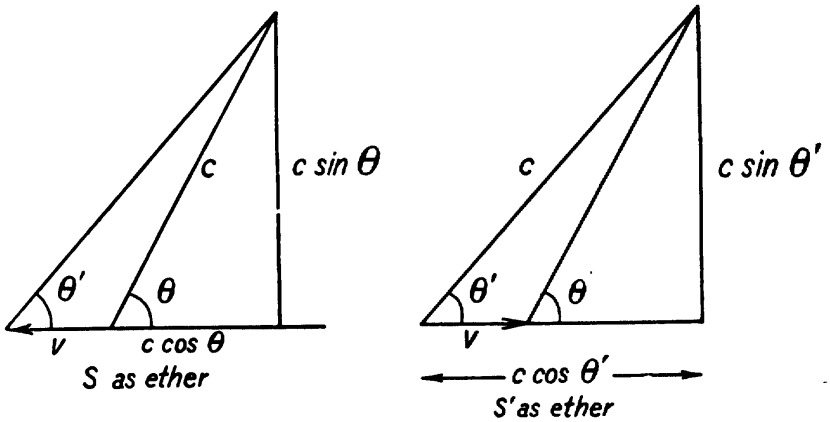


Fig. 19 – Newtonian velocity diagrams for aberration

both of (134) must be regarded as equally valid, because they all reduce to

$$\theta - \theta' = \frac{v}{c} \sin \theta. \tag{135}$$

The approximation is amply justified, since the maximum angle of aberration for the earth is only 41″ (cf. RUSSELL, DUGAN and STEWART [1926, p. 140]).

We can express (133) in symmetrical form. We have

$$\cos \theta \sin \theta' = \gamma \left(\sin \theta \cos \theta' - \frac{v}{c} \sin \theta \right), \tag{136}$$

and a similar equation, obtained by interchanging θ, θ' and changing the sign of v:

$$\cos \theta' \sin \theta = \gamma \left(\sin \theta' \cos \theta + \frac{v}{c} \sin \theta' \right). \tag{137}$$

Multiplying these equations by $\sin \theta' \cos \theta$ and $\sin \theta \cos \theta'$ respectively, and subtracting, we have

$$\sin^2 \theta' \cos^2 \theta - \sin^2 \theta \cos^2 \theta' = - \gamma \frac{v}{c} \sin \theta \sin \theta' (\cos \theta + \cos \theta'), \tag{138}$$

or

$$\sin(\theta' + \theta) \sin(\theta' - \theta)$$

$$= - 2 \frac{\gamma v}{c} \sin \theta \sin \theta' \cos \tfrac{1}{2}(\theta' + \theta) \cos \tfrac{1}{2}(\theta' - \theta), \tag{139}$$

so that

$$2 \sin \tfrac{1}{2}(\theta - \theta') = \gamma \frac{v}{c} \frac{\sin \theta \sin \theta'}{\sin \tfrac{1}{2}(\theta + \theta')}, \tag{140}$$

the required symmetrical expression.

The case of a telescope filled with water is easily explained on relativistic grounds, but not on Newtonian. It is known that the angle of aberration in unaltered if water is put in the tube.[1] This independence of refractive index is obvious from relativity, because the sole condition requisite for the production of the image of the star on the cross-wires is that the wave normal just outside the telescope should be parallel to its axis, the frame of reference being that in which the telescope is at rest.

[1] This was observed by Airy and Hoek (cf. PRESTON [1912, p. 538]).

§ 10. THE EXPANDING UNIVERSE IN SPECIAL RELATIVITY

The spectra of distant nebulae show a consistent shift of spectral lines towards the red end of the spectrum, i.e. an apparent increase in wave length (or decrease in frequency) when comparison is made with spectra from terrestrial sources. The apparent increase in wave length is proportional to the wave length and to the distance of the nebula.

It is usual to explain this phenomenon in terms of the general theory of relativity, but it is possible to construct a very simple model within the special theory which will give the observed effect.[1]

We suppose all stars and nebulae (we shall call them *particles*) to have been generated in a primitive explosion which occurred at some event O (Fig. 20), and shot out in all directions with all speeds up to that of light. We leave gravity out of account, and suppose the particles to have straight world lines, so that what we see in our space-time diagram is the null cone ON drawn into the future from O, the interior of this null cone being filled with a vast number of straight world lines, all passing through O.

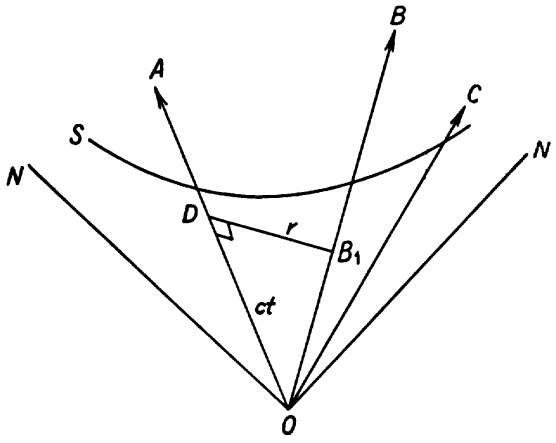


Fig. 20 – Space-time diagram of the expanding universe in special relativity

No world line is privileged above another. The three world lines OA, OB, OC in Fig. 20 are on a parity. Any one of them may be used as time-axis (i.e. the particle is reduced to rest), and in any one of the frames of reference so obtained the kinematical picture is essentially the same.

Suppose we use the frame of OA. Let B_1 be any event on OB and B_1D the perpendicular (in the Minkowskian sense) dropped on OA. Then

$$r = DB_1 \tag{141}$$

is the distance of the particle OB from the particle OA at the instant t, t being the time of D in the frame of OA. The particle OB appears to

[1] KERMACK and MCCREA [1933], MILNE [1935, p. 127].

the particle OA to be receding with a speed

$$v = \frac{dr}{dt} = c\frac{DB_1}{OD} = \frac{r}{t}. \tag{142}$$

Thus OA observes around him particles receding from him according to this law, the furthest particles having a speed approaching c. This picture is the same no matter what particle is used as a base for observations.

If we wished to talk about the distribution of particles statistically, we would draw a pseudosphere S with centre O (Fig. 20), and assign a distribution of particles over it, each such particle having a world line passing through O. The simplest distribution is a uniform one, and then we would have a completely isotropic universe, in the sense that it would look the same to all observers carried on the different particles of the system. But we shall not be concerned here with the question of distributions.

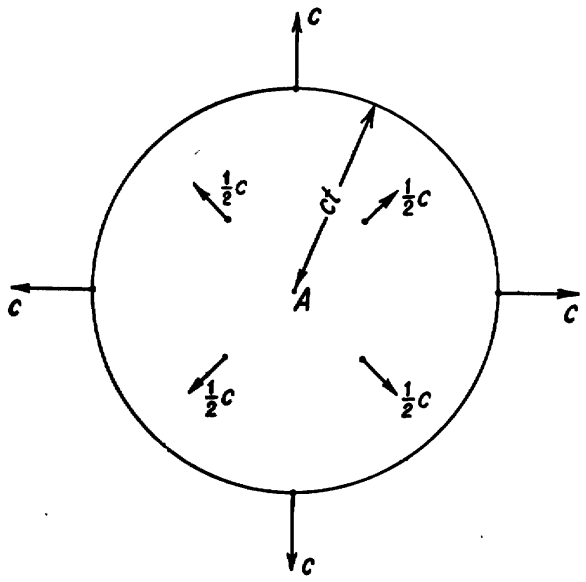


Fig. 21 – Snapshot of the expanding universe

To get a kinematical picture corresponding to the space-time diagram in Fig. 20, we record a snapshot made by an observer on the particle OA at his time t. This snapshot shows a large sphere of radius ct, the radius increasing at the rate c (Fig. 21). The particle OA is the centre of this sphere, and the other particles are moving directly away from the centre with speeds proportional to their distances from it, according to (142). Thus the universe is finite, as judged by OA, with radius ct.

But if we look at it another way, it is infinite. Instead of picking out some particle to set an observer on, let us take those events in the histories of the particles when they are all of the same age, measuring the age of a particle by its proper time, starting from O. The events so obtained give us a pseudosphere S as shown in Fig. 20. It is of infinite extent, and so in this view the universe is infinite. It is, after all, only a question of words, but there is something to be said for this way of

looking at the matter. The distances involved are so great that the primitive idea of simultaneity (defined by $t = \text{const.}$) loses its intuitive significance and is little more than a hangover from Newtonian physics which we have not yet been able to shake off.

§ 11. THE RED-SHIFT

We assume[1] that any atom emits radiation which has certain definite frequencies in the rest-frame of the atom, whether the atom is situated on the earth or on a distant star. If ν_0 is such a *proper frequency* and ν_a the corresponding *apparent frequency* (when the light from the star is observed on the earth), then it is evident that, in the model just described, we shall have $\nu_a < \nu_0$ (and hence a shift of spectral lines towards the red) on account of a Doppler effect due to the outward radial motion of all stars and nebulae relative to the sun. We have now to calculate the amount of this red-shift.

If a source recedes from an observer at speed v, then we have

$$\nu_a = \nu_0 \gamma \left(1 - \frac{v}{c}\right) = \nu_0 \left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}} \tag{143}$$

by (87), in which we have put $u' = -c$ for light proceeding from the source to the observer. If r_0 and t_0 are respectively the distance of the source and the time at the instant of emission, as measured in the frame of the observer, we have $v = r_0/t_0$ by (142) and hence

$$\frac{\nu_a}{\nu_0} = \left(\frac{ct_0 - r_0}{ct_0 + r_0}\right)^{\frac{1}{2}}. \tag{144}$$

This formula is exact. If r_0/ct_0 is small, we have approximately

$$\frac{\nu_a}{\nu_0} = 1 - \frac{r_0}{ct_0}, \tag{145}$$

or, if we write $\delta\nu = \nu_0 - \nu_a$ for the decrease in frequency,

$$\frac{\delta\nu}{\nu_0} = \frac{r_0}{ct_0}. \tag{146}$$

Let us derive (144) in a different way, using a space-time diagram. In Fig. 22, L is the world line of the observer (we may take it to be the world line of the sun) and L^* the world line of the source — a distant

[1] This fundamental assumption was made on p. 14, in setting up the concept of a standard clock.

star or nebula.[1] Let E^* and F^* be two events on L^*, the interval of proper time between then being τ^*, so that $E^*F^* = c\tau^*$. Draw null cones into the future from these two events and let these cones cut L at E and F respectively. Write $EF = c\tau$.

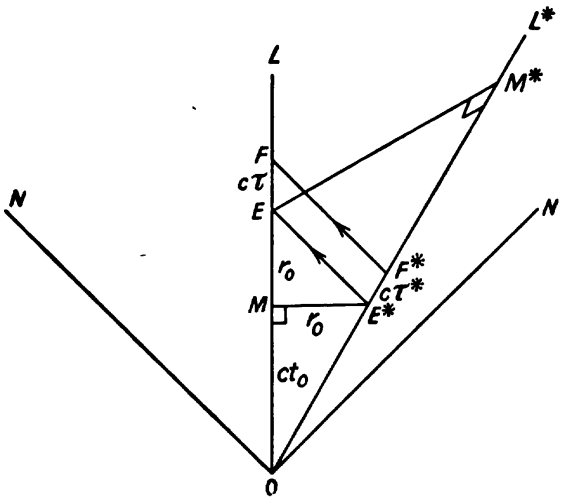


Fig. 22 – Space-time diagram for red-shift in special relativity

If E^* and F^* represent the emission of two successive crests of monochromatic light waves, then E and F represent the reception of these crests; τ^* is the proper period of the light and τ its apparent period. In terms of proper frequency ν_0 and apparent frequency ν_a we have

$$\frac{\nu_a}{\nu_0} = \frac{\tau^*}{\tau} = \frac{E^*F^*}{EF}. \tag{147}$$

From similar triangles[2] in Fig. 22 we have

$$\frac{E^*F^*}{EF} = \frac{OE^*}{OE}, \tag{148}$$

and so

$$\frac{\nu_a}{\nu_0} = \frac{OE^*}{OE} = \frac{(c^2t_0^2 - r_0^2)^{\frac{1}{2}}}{ct_0 + r_0} = \left(\frac{ct_0 - r_0}{ct_0 + r_0}\right)^{\frac{1}{2}}, \tag{149}$$

as in (144), the line ME^* being drawn orthogonal to L, so that

$$ME^* = ME = r_0, \quad OM = ct_0, \quad OE^* = (c^2t_0^2 - r_0^2)^{\frac{1}{2}}. \tag{150}$$

Note that the ratio (149) is independent of ν_0.

§ 12. LUMINOSITY AND DISTANCE

The ratio ν_a/ν_0 can be measured spectroscopically, but r_0 and t_0 are not directly measurable quantities. However astronomers can estimate *apparent distance*, based on luminosity in a way which we shall now describe.

[1] The nebulae in question are extra-galactic nebulae, which are other galaxies like our own; they are called *nebulae* (clouds) because they can be resolved into stars only by extremely powerful telescopes.

[2] In this work we use M-geometry [cf. III–§ 5].

The *intrinsic luminosity* of a source is defined as the total energy of photons[1] emitted by it per unit proper time of the source; the *apparent luminosity* of a source is the amount of energy received by an observer on a target of unit area (normal to the direction of the source) per unit proper time of the observer. It is usual in astronomy to choose the units so that the intrinsic luminosity of the sun is unity and the apparent luminosity of the sun is unity when viewed from the earth; but for our purposes it is simpler to use c.g.s. units, and we shall do so.

To carry through the argument we here anticipate VI–(34) by stating that the relative energy of a photon of frequency ν is $h\nu$, where h is Planck's constant.

Let a source emit $n(\nu_0)\tau^*$ photons of proper frequency ν_0 in an interval τ^* of proper time; then its intrinsic luminosity is

$$L = h\sum n(\nu_0)\nu_0, \tag{151}$$

summed over all frequencies.

To calculate the apparent luminosity of the source, we use Fig. 22, with the slight change that E^* and F^* now correspond to the beginning and end of a burst of photons, so that $E^*F^* = c\tau^*$ where τ^* is the (proper) duration of the burst; $EF = c\tau$ where τ is the interval of proper time of the observer during which the burst is received. We add an extra line to the diagram, EM^* drawn orthogonal to L^*. Then, in Minkowskian measure, the right angled triangles OME^* and OM^*E are similar, and we have

$$\frac{EM^*}{E^*M} = \frac{OE}{OE^*} = \frac{\nu_0}{\nu_a}, \tag{152}$$

by (149).

We suppose the photons to be thrown out isotropically when viewed in the rest-frame of L^*. When some of them reach E, they appear to L^* to form a thin spherical shell of radius EM^*, and so the number of photons of proper frequency ν_0 caught on a target of area $d\sigma$, normal to the line of approach, is

$$\frac{d\sigma}{4\pi EM^{*2}}\, n(\nu_0)\tau^*. \tag{153}$$

The measure of area $d\sigma$ is the same in the rest-frames of L and L^*,

[1] Usually one thinks of light, but in a relativistic argument it is best to include all frequencies.

because instantaneous lines drawn in it are orthogonal to both L and L^* and so undergo no FitzGerald-Lorentz contraction.

Each photon of proper frequency ν_0 has an energy $h\nu_a$ relative to L, and thus the total energy received on the target from the burst of photons is

$$\frac{d\sigma}{4\pi EM^{*2}} h\tau^* \sum n(\nu_0)\nu_a, \tag{154}$$

summed for all frequencies, ν_a corresponding to ν_0. The burst lasts for proper time τ as judged by L, and so we get the apparent luminosity l by dividing (154) by $\tau d\sigma$:

$$l = \frac{1}{4\pi EM^{*2}} h \frac{\tau^*}{\tau} \sum n(\nu_0)\nu_a. \tag{155}$$

But $\tau^*/\tau = \nu_a/\nu_0$ (the same for all frequencies) and therefore

$$l = \frac{1}{4\pi EM^{*2}} \left(\frac{\nu_a}{\nu_0}\right)^2 h \sum n(\nu_0)\nu_0 = \frac{L}{4\pi EM^{*2}} \left(\frac{\nu_a}{\nu_0}\right)^2 \tag{156}$$

by (151).

We now *define* the *apparent distance* D of the source by the formula (suggested by the inverse square law for small distances)

$$D = \sqrt{\frac{L}{4\pi l}} \tag{157}$$

and (156) gives

$$D = EM^* \frac{\nu_0}{\nu_a}, \tag{158}$$

or, by (150) and (152),

$$D = r_0 \left(\frac{\nu_0}{\nu_a}\right)^2. \tag{159}$$

As in (149) we have

$$\left(\frac{\nu_a}{\nu_0}\right)^2 = \frac{ct_0 - r_0}{ct_0 + r_0}. \tag{160}$$

In (159) and (160) we have two equations connecting D, r_0, t_0 and ν_a/ν_0. We can express ν_a/ν_0 in terms of any two of D, r_0 and t_0, but it is better to bring in t, the time of reception (cf. § 13).

Note that D involves nothing but luminosities, and might therefore properly be called *luminosity distance*. However, BONDI [1952, p. 108] has defined a luminosity distance q_1 in a different way in general relativity, incorporating a Doppler shift factor so that $q_1 = D\nu_a/\nu_o$;

in flat space-time $q_1 = EM^*$ (Fig. 22). All this is a mere matter of definition and of secondary importance, but, if we admit the factor v_a/v_o into the definition at all, we might as well define the distance (qualified by a suitable adjective) as $D(v_a/v_o)^2 = r_o$ [cf. (159)].

§ 13. THE DEPENDENCE OF RED-SHIFT ON APPARENT DISTANCE AND THE AGE OF THE UNIVERSE

Let t be the time of reception in the frame of L, so that t is fact the age of the sun, in its own proper time, from the initial explosion O. Then (Fig. 22) we have

$$ct = OE = ct_0 + r_0, \tag{161}$$

and elimination of r_0 and t_0 from (159), (160) and (161) gives

$$\boxed{\frac{\nu_a}{\nu_0} = \left(1 + \frac{2D}{ct}\right)^{-\frac{1}{2}}.} \tag{162}$$

This formula expresses accurately the ratio of apparent frequency ν_a to proper frequency ν_0 in terms of apparent distance D, defined by (157), and the age of the universe (the proper age of the sun at the present time).

With $\delta\nu = \nu_0 - \nu_a$, we have

$$\frac{\delta\nu}{\nu_0} = 1 - \left(1 + \frac{2D}{ct}\right)^{-\frac{1}{2}} = \frac{D}{ct} - \tfrac{3}{2}\left(\frac{D}{ct}\right)^2 + \ldots. \tag{163}$$

In terms of wave length ($c = \lambda\nu$), the apparent wave length λ_a is related to the proper wave length λ_0 by

$$\frac{\lambda_a}{\lambda_0} = \left(1 + \frac{2D}{ct}\right)^{\frac{1}{2}}, \tag{164}$$

or, with $\delta\lambda = \lambda_a - \lambda_0$,

$$\frac{\delta\lambda}{\lambda_0} = -1 + \left(1 + \frac{2D}{ct}\right)^{\frac{1}{2}} = \frac{D}{ct} - \tfrac{1}{2}\left(\frac{D}{ct}\right)^2 + \ldots. \tag{165}$$

Except for very distant nebulae, the linear terms dominate, and the important astronomical quantity is *Hubble's constant*:

$$\sigma = \lim_{D\to 0}\left(\frac{\delta\nu}{\nu_0}\,\frac{1}{D}\right) = \lim_{D\to 0}\left(\frac{\delta\lambda}{\lambda_0}\,\frac{1}{D}\right) = \frac{1}{ct}, \tag{166}$$

which is the reciprocal of a length. The evaluation of this constant is of great interest. The ratio $\delta\lambda/\lambda_0$ can be observed directly and so can the apparent luminosity l, but to get D from l we have to know the intrinsic luminosity L in order to use (157). We cannot enter here into the complex of facts and assumptions by which astronomers estimate D except to say that it would be easy if there existed a set of stars all of the same intrinsic luminosity L, one of these stars being so close that its distance could be measured by other means, and at least one such star being observable in each nebula whose distance is sought.

For many years the value

$$\sigma = 6.0 \times 10^{-28}\,\mathrm{cm}^{-1} \tag{167}$$

was accepted (cf. TOLMAN [1934, p. 359]), but recent work by W. Baade and others indicates that the values of D for extragalactic nebulae was underestimated by a factor of 2, so that the present acceptable value of Hubble's constant is[1]

$$\sigma = 3 \times 10^{-28}\,\mathrm{cm}^{-1}. \tag{168}$$

Inserting the value (168) in (166), we get

$$\begin{aligned} ct = \sigma^{-1} &= 3.3 \times 10^{27}\,\mathrm{cm}, \\ t &= 1.1 \times 10^{17}\,\mathrm{sec}, \end{aligned} \tag{169}$$

or, since 1 sidereal year $= 3.1558 \times 10^7$ sec, *we find the age of the universe to be*

$$t = 3.5 \times 10^9 \text{ years}. \tag{170}$$

This value is in general agreement with estimates based on the age of the earth, on the composition of meteorites, and on the evolution of stars and of our galaxy.

It is interesting to note that, if we move the world line L^* over towards the null cone in Fig. 22, so that $(r_0 - ct_0)$ tends to zero, then by (160) v_a tends to zero and by (159) D tends to infinity. This means that although the model universe we have been considering is in a sense finite, being of radius $r = ct$ at time t, yet *observationally it is infinite*; for as we view nebulae further and further away on the null

[1] See HUBBLE [1951], [1953], SHARPLESS [1953], STRUVE [1953] and THACKERAY and WESSELINK [1953], for which references I am indebted to Professor H. Brück and Dr. G. J. Whitrow.

Still more recent work indicates that the correction factor should be 3, rather than 2, making $\sigma = 2 \times 10^{-28}\mathrm{cm}^{-1}$.

cone drawn into the past from our present instant E, the apparent frequency tends to zero (light changes through radio waves to still lower frequencies) and the apparent luminosity tends to zero, so that the radiation could not be detected. No telescope, however large, could reach the boundary of this model universe.

§ 14. THE MICHELSON-MORLEY EXPERIMENT

Let L_0 be the world line of a source of light (Fig. 23) and L, L^* the world lines of two mirrors. In general these are curves in space-time, not straight lines. Light starts from an event a_r on L_0 and is at once split into two parts. One part travels to L from which it is reflected at the event b_r and returns to L_0 at the event d_r; the other part is reflected from L^* at b_r^* and returns to L_0 at d_r^*. There is no reason to suppose that d_r coincides with d_r^*.

The Michelson interferometer was designed to measure the interval of proper time between d_r^* and d_r in such a situation by observing the pattern of interference fringes formed by the light on its reunion after traversing the two different paths. With such an instrument Michelson and Morley performed their famous experiment [MICHELSON and MORLEY, 1887], L and L^* being the world lines of mirrors fixed to a slab of stone which could be rotated slowly in a horizontal plane, and L_0 being the world line of the centre of the slab, a fixed point on the earth's surface. A complete account of the Michelson-Morley experiment, and other experiments of the same type performed by him, was given by D. C. MILLER [1933]. Similar experiments were performed by KENNEDY [1926], ILLINGWORTH [1927], MICHELSON, PEASE and PEARSON [1929], and Joos [1930].

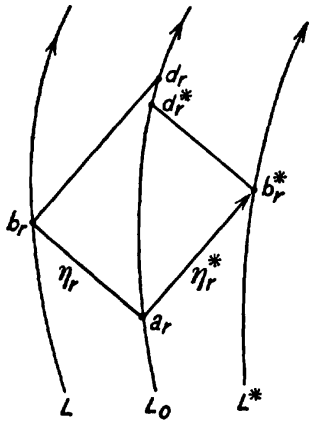


Fig. 23 – Space-time diagram for Michelson-Morley experiment

The apparatus employed in a Michelson-Morley experiment (we may use that name, to cover all experiments of this type) involves many mirrors but the simplification shown in Fig. 23 (a source and two mirrors) is usually regarded as adequate for the discussion of the theory of the instrument (SILBERSTEIN [1924, pp. 71, 88]). We shall not go into that theory here, but use the simple formula

$$\frac{\delta}{w} = \frac{c(T - T^*)}{\lambda} = \nu(T - T^*), \tag{171}$$

where

δ = fringe shift,

w = fringe width,

λ, ν = wave length and frequency of monochromatic light used,

T, T^* = intervals of proper time for L_0 from a_r to d_r and from a_r to d_r^* respectively.

We may call T and T^* the *trip-times* for light pursuing the two paths indicated. We note that $T - T^*$ is the interval of proper time from d_r^* to d_r. This interval is connected by (171) with the measurable quantities c, w, λ (or ν) and δ.

Let us now calculate the trip-time T, taking for L_0 the equations $x_r = x_r(s)$ and supposing L generated by drawing given null vectors from the events on L_0; this method of describing L is convenient and involves no restriction on it. We have then

$$b_r = a_r + \eta_r, \tag{172}$$

where η_r is the null vector drawn from a_r. Since $(d_r - b_r)$ is a null vector, we have

$$(d_r - b_r)\,(d_r - b_r) = 0. \tag{173}$$

Hence by (172)

$$(d_r - a_r - \eta_r)\,(d_r - a_r - \eta_r) = 0, \tag{174}$$

and so, since η_r is null,

$$(d_r - a_r)\,(d_r - a_r) - 2\eta_r(d_r - a_r) = 0. \tag{175}$$

Denoting by s the separation between a_r and d_r, measured on L_0, we write

$$d_r = a_r + \sum_{n=1}^{\infty} \frac{s^n}{n!}\left(\frac{d^n x_r}{ds^n}\right)_{a_r}, \tag{176}$$

and, substituting in (175), obtain for s an equation of infinite degree.

Supposing that L_0 is nearly straight and that consequently the terms in (176) decrease rapidly, we cut off the series, and, retaining only the more important terms, obtain from (175) the equation

$$(sx_r' + \tfrac{1}{2}s^2x_r'')\,(sx_r' + \tfrac{1}{2}s^2x_r'') - 2\eta_r(sx_r' + \tfrac{1}{2}s^2x_r'') = 0, \tag{177}$$

the primes denoting derivatives with respect to s at a_r. We have the identities

$$x_r'x_r' = -1, \quad x_r'x_r'' = 0, \tag{178}$$

and so, rejecting terms in s^4, we get

$$s^2(1 + \eta_r x''_r) = -2s\eta_r x'_r \tag{179}$$

and hence

$$s = -\frac{2\eta_r x'_r}{1 + \eta_n x''_n}. \tag{180}$$

Thus the two trip-times are

$$T = -\frac{2\eta_r x'_r}{c(1 + \eta_n x''_n)}, \quad T^* = -\frac{2\eta^*_r x'_r}{c(1 + \eta^*_n x''_n)}, \tag{181}$$

where $\eta^*_r = b^*_r - a_r$, and what the interferometer measures is

$$\lambda\frac{\delta}{w} = c(T - T^*) = 2\left(\frac{\eta^*_r x'_r}{1 + \eta^*_n x''_n} - \frac{\eta_r x'_r}{1 + \eta_n x''_n}\right). \tag{182}$$

This is an invariant expression. Let us evaluate it in that Galileian frame S in which the source is at rest at a_r (i.e. we take the time-axis tangent to L_0 at a_r). We have then

$$x'_\varrho = 0, \quad x'_4 = i, \quad x''_4 = 0, \tag{183}$$

to be substituted in (182). If u_ϱ and f_ϱ denote respectively the 3-velocity and 3-acceleration of the source, we have in general

$$x'_\varrho = \frac{\gamma u_\varrho}{c}, \quad x''_\varrho = \frac{\gamma}{c}\frac{d}{dt}\left(\frac{\gamma u_\varrho}{c}\right), \quad \gamma = (1 - u^2/c^2)^{-\frac{1}{2}}, \tag{184}$$

and thus at a_r we have

$$u_\varrho = 0, \quad f_\varrho = \frac{du_\varrho}{dt} = c^2 x''_\varrho. \tag{185}$$

In the frame S, η_ϱ are the space coordinates of the mirror L at the event b_r. Let us put $\eta = (\eta_\varrho\eta_\varrho)^{\frac{1}{2}}$, so that $\eta_4 = i\eta$ since η_r is a null vector. With a similar notation for η^*, (182) reads

$$\lambda\frac{\delta}{w} = 2\left(\frac{\eta}{1 + \eta_\varrho f_\varrho/c^2} - \frac{\eta^*}{1 + \eta^*_\varrho f_\varrho/c^2}\right). \tag{186}$$

Note that η and η^* are distances from the source to the mirrors, measured in the frame S, but the positions of the mirrors are not the positions simultaneous with the emission of the light from the source, but a little later, to allow for the passage of light from source to mirror.

What does relativity predict that an observer will see when he looks into the interferometer, the source L_0 being fixed on the earth's surface and the mirrors L, L^* being attached to a slab of stone, rotating slowly round the source as centre? Will the fringes move? Equivalently, will δ change? (λ and w are fixed constants, depending only on the nature of the light used and the adjustment of the interferometer.)

Relativity can give no answer to this question until an answer is given to another question: How does η change when the slab of stone is rotated? We can put the responsibility more squarely on η and η^* by noting that, although we may have difficulty in saying *precisely* what f_ϱ is, nevertheless we naturally attribute it to the orbital and diurnal rotations of the earth and conclude that $\eta_\varrho f_\varrho/c^2$ and $\eta_\varrho^* f_\varrho/c^2$ are extremely small and may be dropped, so that we have

$$\lambda \frac{\delta}{w} = 2(\eta - \eta^*). \tag{187}$$

When the Michelson-Morley experiment was performed, no doubts were entertained about the validity of the concept of a rigid body, and a slab of stone was regarded as rigid in the Newtonian sense, at least when made as free as it could be of stresses and temperature variations. In fact, the experiment was made to enlarge our knowledge of the behaviour of light, the behaviour of bodies called "rigid" not being questioned. But actually the experiment involved both things — the behaviour of light and the behaviour of bodies.

The concept of a rigid body involves great difficulties in relativity (cf. p. 36) and it appears impossible to make any satisfactory prediction as to the result of an ideal Michelson-Morley experiment performed with mirrors attached to a rigid three-dimensional base (the slab of stone), for the reason that such a base has not been satisfactorily defined in the terms of mathematical physics.

Except for a small effect reported by MILLER [1933], the results of all Michelson-Morley experiments have been *null* (i.e. δ does not change when the base supporting the mirrors is rotated). This result has been of tremendous historical importance because it helped to loosen the hold of the ether-theory, but that does not concern us here since the concept of the ether is dead and gone. We must now look on the null result as information about the behaviour of the bodies used to support the mirrors. If the small effect observed by Miller is discarded

as due to some systematic error[1], we conclude that, under the circumstances of the experiments and for the materials used in the bases supporting the mirrors, η and η^* change by amounts too small to produce measurable optical effects; and if the small effect were validated in a repetition of the experiment, we would not interpret it as evidence for or against the theory of relativity, but rather as information about the way in which η and η^* changed under the circumstances of the experiment.

[1] In a recent paper, R. S. SHANKLAND, S. W. MCCUSKEY, F. C. LEONE and G. KUERTI [1955] attribute it to random fluctuations in Miller's readings of the fringe positions and to local temperature anomalies.

CHAPTER VI

MECHANICS OF A PARTICLE AND COLLISION PROBLEMS

§ 1. FORCE. ACTION AND REACTION. A PHILOSOPHICAL DIGRESSION

The concept of *force* is basic in Newtonian mechanics. It is introduced as a statical idea, and one first studies the theory of systems in equilibrium. As long as we consider only equilibrium relative to a Galileian frame, there is no reason why we should not take over into relativity the whole body of the classical theory of statics. Thus *statical* force presents no problem as far as relativity is concerned.

But *dynamical* force is a very different matter, particularly in connection with that law of Newton which asserts the equality of action and reaction. Consider two particles connected by a light extended spring. In Newtonian mechanics we would say that the forces exerted by the spring on the particles *at any instant* are equal and opposite, depending on the length of the spring *at that instant*. But the length of the spring at any instant means the distance between *simultaneous* events in the histories of its two ends, and so we see that the idea of absolute simultaneity is included in Newton's law of action and reaction. In the theory of relativity events which are simultaneous for one observer are not in general simultaneous for another, and so we must abandon the law of the equality of action and reaction as a general law of nature. We can keep it when the action and reaction arise from contact, because in that case the simultaneity of events occurring at different places is not involved.

If we consider the fundamental roles played in Newtonian mechanics by the concept of force and the equality of action and reaction, we realise how revolutionary and destructive the theory of relativity is. It destroys the physical validity of Newtonian mechanics as developed throughout the eighteenth and nineteenth centuries, since the independent variable t which appears in the equations of Euler, Lagrange and Hamilton can no longer be given a physical meaning.

There are two ways of salvaging the wreck. The first way is practical but intellectually unsatisfying. We know that Newtonian mechanics gives good predictions in physics and astronomy, and IV–(205) tells us why this is so; the relative speeds are small compared with the speed of light and the Lorentz transformation reduces in the limit to the Newtonian transformation. We may then decide to use Newtonian mechanics on all ordinary occasions, and bring in relativity as a special correction when high speeds are involved.

This plan conflicts with an unspoken hypothesis which underlies all the physical theories so far created, namely that behind physical phenomena there lies a unique mathematical structure which it is the purpose of theory to reveal. According to this hypothesis, the mathematical formulae of physics are *discovered* not *invented*, the Lorentz transformation, for example, being as much a part of physical reality as a table or a chair.

This hypothesis, if we accept it, makes the salvage of Newtonian mechanics impossible. The unique mathematical structure of nature cannot be both Newtonian and relativistic — there is either an absolute time t or there is not, — and we must declare ourselves for Newton or for Einstein; on the basis of experimental evidence we have no choice but to accept Einstein's theory and reject Newton's.

But this hypothesis of a unique mathematical structure for nature is actually very naive. It is a product of the eighteenth century, a period when mathematics was understood much less than it is today, and its survival is an anachronism, unacceptable to any physicist who has thought about mathematics or any mathematician who has thought about physics. Modern mathematics is based on the acceptance of certain simple fundamental operations, such as the putting of two sets of things into correspondence, and on certain laws of reasoning, and out of these it constructs such concepts as *irrational number, limit, continuity, derivative, integral.* When understood properly (i.e. as mathematicians understand them) these concepts exist in the human mind and not in nature; it is a meaningless waste of time to debate whether the ratio of two measured lengths is rational or irrational, or whether matter is continuous or discontinuous, because the concepts of irrationality and continuity belong to a world of the intellect, a world of mathematics, and not to the real world in which phenomena occur and are measured by pieces of apparatus.

The second way of salvaging Newtonian mechanics is to reject the

old hypothesis of the underlying mathematical structure of nature, and to accept Newtonian mechanics (or for that matter relativistic mechanics) as a *mathematical model* which we have constructed to assist us in our exploration of nature, and nothing more. This means that a theory in mathematical physics is to be judged primarily on the basis of its logical consistency and only secondarily on the basis of the truth of its physical predictions. Newtonian mechanics and relativistic mechanics can stand as permanent logical structures, proof against anything which may be found by experiment, and we can add to them in the future other mathematical models when they are needed for the interpretation of experiments.

This is no new thing. Geometry, a fundamental branch of physics, has been through this mill, and two things have emerged where one went in — on the one hand the mathematical model (Euclidean, non-Euclidean or Riemannian, a purely logical structure) and on the other hand the physical science of measurement — replacing the old mixture of axiomatics and physical reality.

Let us then proceed with relativistic mechanics, free from the thought that we are destroying Newtonian mechanics and also from the thought that what we are now developing can ever be destroyed.

§ 2. PARTICLES AND MASS

In Newtonian mechanics a particle is an idealisation — a concentration of finite mass at a point. In the physical application of the idea, a piece of matter is treated as a particle if its size is very small compared with the sizes of other quantities involved in the particular problem under discussion. Thus, in comparison with the centimetre, electrons, protons, atoms or molecules may for most purposes be treated as particles; in comparison with a range of twenty miles, an artillery projectile may be regarded as a particle; in comparison with the radius of its orbit, the earth may be regarded as a particle; and so on.

In Newtonian mechanics a particle possesses a mass which does not change. In relativistic mechanics the question of mass is more complicated, and we have to distinguish between *relative mass* and *proper mass*. These ideas will become clearer after we have developed the mechanical theory into which they fit, but for the moment we may remark that in relativity *mass* and *energy* are identified with one another, except for a constant factor, and thus a change of energy

means a change of mass. As we shall see, the relative mass of any particle depends on its velocity relative to the observer, and so is variable. The question of the constancy or variability of proper mass is more delicate. Ultimate particles of matter (protons and electrons) are assigned unchanging proper masses, but the proper mass of a structure (atom, molecule, or body of ordinary size treated as a particle) depends on its energy. The proper mass of a hydrogen atom cannot be the sum of the proper masses of the proton and electron which compose it, for the proper mass of the atom changes with its energy when it radiates.

We regard proper mass as a primary concept; for our purposes it is enough to say that it is a positive number associated with a material particle.

§ 3. EQUATIONS OF MOTION

Consider a particle of proper mass m, which may for the present be regarded as constant or variable. Let the particle move on some world line with equations $x_r = x_r(s)$, s being the Minkowskian arc length or separation. Then, as in v–(53), the 4-vector dx_r/ds is the 4-velocity. We write

$$M_r = m\frac{dx_r}{ds}, \tag{1}$$

and call this (for reasons which will appear presently) the *momentum-energy 4-vector*, or more briefly the *momentum 4-vector* or *4-momentum*.

If u_ϱ is the 3-velocity of the particle (its velocity as measured by the Galileian observer using the Minkowskian coordinates x_r), then the components of the 4-velocity are as in v-(59), and so the components of the 4-momentum are

$$M_\varrho = \frac{1}{c}m\gamma u_\varrho, \quad M_4 = im\gamma. \tag{2}$$

Note that for very small velocity, γ is nearly unity, and so for that case we have approximately

$$M_\varrho = \frac{1}{c}mu_\varrho, \quad M_4 = im. \tag{3}$$

Except for the factor $1/c$, the three components M_ϱ thus have the Newtonian form for momentum; the fourth component is proportional to the proper mass in this approximation.

Now M_r is a 4-vector, i.e. it transforms like dx_r under Lorentz

transformation. So also is dM_r/ds. Thus if we introduce a 4-*force* X_r (a 4-vector) and write the equations

$$\frac{dM_r}{ds} = X_r, \tag{4}$$

we have *equations of motion* which are invariant under Lorentz transformation, and have at least some formal resemblance to Newton's law of motion. These equations of motion may also be written

$$\frac{d}{ds}\left(m\frac{dx_r}{ds}\right) = X_r. \tag{5}$$

§ 4. IS PROPER MASS CONSTANT?

Since the particle's world line is timelike (or, equivalently, its speed is less than that of light), we have

$$\frac{dx_r}{ds}\frac{dx_r}{ds} = -1 \tag{6}$$

from the definition of ds. Thus, in (5) and (6), we have five equations to determine the five quantities m, x_r as functions of s, if the 4-force X_r is given. This makes the problem of motion mathematically definite, but it also implies that we are not at liberty to assume a priori that the proper mass is a constant, for the proper mass is to be found from the equations of motion. Indeed we can find its rate of change. Write (5) in the form

$$\frac{dm}{ds}\frac{dx_r}{ds} + m\frac{d^2x_r}{ds^2} = X_r. \tag{7}$$

It follows from (6) that

$$\frac{dx_r}{ds}\frac{d^2x_r}{ds^2} = 0, \tag{8}$$

and so, on multiplying (7) by dx_r/ds, we get

$$\frac{dm}{ds} = -X_r\frac{dx_r}{ds}. \tag{9}$$

Thus *proper mass remains constant if, and only if, the 4-force* X_r *is orthogonal to the world line of the particle,* dx_r/ds *being the unit tangent vector to that world line.*

This condition of orthogonality is actually fulfilled (as we shall see later) in the most important application of the equations of motion,

i.e. to the case of a charged particle moving in an electromagnetic field. So, having satisfied ourselves that we could, if necessary, deal with the case of variable proper mass, we shall in what follows below assume that

$$X_r \frac{dx_r}{ds} = 0, \tag{10}$$

so that m is a constant by virtue of (9). When we come to problems of collision and disintegration in § 8, we shall re-open the question of changes in m.

§ 5. INTERPRETATION OF THE EQUATIONS OF MOTION

If we substitute from (2) in (5), the equations of motion take the form

$$\begin{aligned} \frac{d}{dt}(m\gamma u_\varrho) &= \frac{c^2 X_\varrho}{\gamma}, \\ \frac{d}{dt}(m\gamma c^2) &= \frac{c^3 X_4}{i\gamma}. \end{aligned} \tag{11}$$

We have used the fact that $ds = cdt/\gamma$ [cf. v–(58)].

Let us define the *relative 3-force* P_ϱ by

$$P_\varrho = \frac{c^2 X_\varrho}{\gamma}; \tag{12}$$

then the first line of (11) gives

$$\frac{d}{dt}(m\gamma u_\varrho) = P_\varrho. \tag{13}$$

This we shall interpret presently.

Although the 4-force has four components, they are connected by (10), since we have agreed to discuss the case of constant proper mass. This relationship may also be written

$$X_\varrho \frac{u_\varrho}{c} + iX_4 = 0, \tag{14}$$

and so

$$\frac{X_4}{i} = \frac{X_\varrho u_\varrho}{c} = \frac{\gamma P_\varrho u_\varrho}{c^3}, \quad \frac{c^3 X_4}{i\gamma} = P_\varrho u_\varrho. \tag{15}$$

Thus the last of (11) may be written

$$\frac{d}{dt}(m\gamma c^2) = P_\varrho u_\varrho. \tag{16}$$

We have, in (13) and (16), reduced the equations of motion (4) to a form which involves only the 3-vectors u_ϱ (relative velocity) and P_ϱ (relative force). By defining certain terms we can give verbal statements of these equations which reveal them as generalisations of the Newtonian equations of motion.

We make the following definitions:

Relative mass $= m^* = m\gamma = m(1 - u^2/c^2)^{-\frac{1}{2}}$.

Relative momentum $= m^*u_\varrho = m\gamma u_\varrho = mu_\varrho(1 - u^2/c^2)^{-\frac{1}{2}}$.

Relative energy $= E = m^*c^2 = m\gamma c^2 = mc^2(1 - u^2/c^2)^{-\frac{1}{2}}$.

Relative kinetic energy $= T = (m^* - m)c^2 = (\gamma - 1)mc^2$
$= mc^2[(1 - u^2/c^2)^{-\frac{1}{2}} - 1]$.

Activity of relative force $= P_\varrho u_\varrho$.

We may then write (13) and (16) in the forms

$$\frac{d}{dt}(m^*u_\varrho) = P_\varrho, \tag{17}$$

$$\frac{dE}{dt} = P_\varrho u_\varrho; \tag{18}$$

since m is constant, we may also write (18) as

$$\frac{dT}{dt} = P_\varrho u_\varrho. \tag{19}$$

In words, we may say:

Rate of change of relative momentum = relative force.

Rate of change of relative energy (or relative kinetic energy) = activity of relative force.

Except for the insertion of the word *relative*, these agree with the verbal statements of Newtonian mechanics.

The agreement is even better than might appear at first sight. For if we make a binomial expansion, we get

$$\begin{aligned} T &= mc^2\left(1 + \tfrac{1}{2}\frac{u^2}{c^2} + \tfrac{3}{8}\frac{u^4}{c^4} + \ldots - 1\right) \\ &= \tfrac{1}{2}mu^2\left(1 + \tfrac{3}{4}\frac{u^2}{c^2} + \ldots\right); \end{aligned} \tag{20}$$

if u/c is small, we have approximately $T = \frac{1}{2}mu^2$, which is the familiar expression for kinetic energy in Newtonian mechanics.

If u/c tends to zero, the distinction between relative mass and proper mass disappears, and we recognise the full agreement between the relativistic and the Newtonian equations of motion in this limit.

Since m^* depends on u through γ, it is customary to say that "mass depends on velocity." But it must be remembered that this m^* is the *relative* mass; the *proper* mass does not depend on velocity, and the "dependence on velocity" is introduced merely for the sake of obtaining formal agreement between the equations of relativity and those of Newtonian mechanics.

There is more to be said about energy. The quantity which occurs naturally in (18) is E, not T. This may seem unimportant, since $E - T = mc^2$, which is a constant, disappearing when we differentiate. But the difference between E and T is not trivial. If the particle is at rest ($u = 0$), then $T = 0$, but $E = mc^2$. Thus a particle, even when at rest, is to be thought of as carrying energy, the amount being directly proportional to the proper mass. This identification of mass and energy (to within the constant factor c^2) is one of the most important contributions of the theory of relativity to physics. We write

$$E_0 = mc^2, \tag{21}$$

and call E_0 the *proper energy* of the particle. It is Lorentz invariant. It remains constant as long as the proper mass remains constant. But if m changes even by a very small amount, E_0 changes and its change may be considerable, in view of the factor c^2 which is roughly 10^{21} in c.g.s. units. Equation (21) is a famous equation of modern physics, embodying as it were the source of atomic energy, since that may be regarded as arising from changes in proper mass in nuclear disintegrations (for the history of this equation, see WHITTAKER [1953, p. 51]).

We called M_r the momentum-energy 4-vector; we can now see the justification for including the word *energy* if we write (2) in the form

$$\begin{gathered} M_\varrho = \frac{m\gamma u_\varrho}{c}, \quad M_4 = im\gamma = \frac{iE}{c^2}, \\ M_r M_r = -m^2 = -\frac{E_0^2}{c^4}. \end{gathered} \tag{22}$$

§ 6. MOTION UNDER A CONSTANT RELATIVE FORCE AND IN A CONSTANT MAGNETIC FIELD

To illustrate the application of the relativistic equations of motion, consider the motion of a particle under the action of a constant relative force P acting along the x-axis. There is only one equation to consider:

$$\frac{d}{dt}\frac{mu}{\sqrt{1-u^2/c^2}} = P. \tag{23}$$

This gives at once

$$\frac{mu}{\sqrt{1-u^2/c^2}} = Pt, \quad \frac{u}{c} = \frac{Pt}{\sqrt{m^2c^2+P^2t^2}}, \tag{24}$$

the constant of integration having been chosen so that the particle starts from rest at $t = 0$. We see that as t tends to infinity, u tends to c, but never attains this value. This simple example illustrates the fact that no particle can ever acquire a velocity as great as that of light.

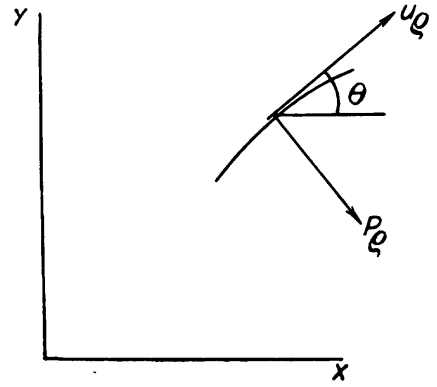


Fig. 1 – Particle moving in a constant magnetic field

As a second illustration, consider a particle bearing an electric charge e and moving in the xy-plane under the influence of a uniform magnetic field which acts in the direction of the z-axis. We assume [cf. x-(30)] that the particle is acted on by a (relative) force in a direction perpendicular to its velocity, the magnitude of the force being euH, where u is the magnitude of the velocity and H the intensity of the field in electromagnetic units. If θ denotes the inclination of the velocity to the x-axis (Fig. 1), the equations of motion (13) give

$$\begin{cases} m\dfrac{d}{dt}\left(\dfrac{u\cos\theta}{\sqrt{1-u^2/c^2}}\right) = euH\sin\theta, \\ m\dfrac{d}{dt}\left(\dfrac{u\sin\theta}{\sqrt{1-u^2/c^2}}\right) = -euH\cos\theta. \end{cases} \tag{25}$$

Multiplying these equations by

$$\frac{2u\cos\theta}{\sqrt{1-u^2/c^2}}, \qquad \frac{2u\sin\theta}{\sqrt{1-u^2/c^2}},$$

respectively and adding, we get

$$m \frac{d}{dt} \frac{u^2}{\sqrt{1 - u^2/c^2}} = 0, \tag{26}$$

and so $u = \text{const.}$, as may also be seen directly from (16). Then the first of (25) gives

$$- \frac{mu \sin \theta}{\sqrt{1 - u^2/c^2}} \frac{d\theta}{dt} = euH \sin \theta, \tag{27}$$

so that

$$\frac{d\theta}{dt} = - \frac{e}{m} H \sqrt{1 - \frac{u^2}{c^2}}. \tag{28}$$

Thus the direction of the velocity changes at a constant rate, and so, since its magnitude is constant, the particle describes a circle. If the radius of the circle is r and the periodic time is τ, then

$$u = \frac{2\pi r}{\tau}, \qquad 2\pi = \frac{e}{m} \tau H \sqrt{1 - \frac{u^2}{c^2}}, \tag{29}$$

and so

$$\frac{e}{m} = \frac{1}{rH} \frac{u}{\sqrt{1 - u^2/c^2}}. \tag{30}$$

This is a very important practical formula, because it enables us to compute the ratio (charge)/(proper mass) for a particle if r, H and u are known.

§ 7. MOMENTUM 4-VECTOR FOR A PHOTON

Light, as an electromagnetic wave phenomenon, possesses momentum and energy according to the theory of Maxwell. In the modern view light is propagated in quanta or *photons*, a photon being regarded (somewhat vaguely) as a localised set of plane waves. To incorporate the photon into the scheme of relativistic mechanics, we think of it as something akin to a material particle, and, like the material particle, endowed with a momentum 4-vector. However, we remember that in some sense the photon is associated with a system of plane waves moving with velocity c, and so we ask whether such a system of waves provides us with a 4-vector which will serve as something from which to construct the momentum 4-vector of a photon.

We met in v-(91) precisely the 4-vector we want, viz. the frequency 4-vector f_r with components $(l_\varrho \nu, i\nu)$, where l_ϱ are the direction cosines of the normal to the waves and ν is the frequency. We now formally define the momentum 4-vector of a photon as

$$M_r = \frac{h}{c^2} f_r, \tag{31}$$

where h is Planck's constant. The components of M_r are then

$$M_\varrho = \frac{h}{c^2} l_\varrho \nu, \qquad M_4 = \frac{ih\nu}{c^2}. \tag{32}$$

Now for a material particle we have, as in (22),

$$M_\varrho = \frac{m\gamma u_\varrho}{c}, \qquad M_4 = \frac{iE}{c^2}, \tag{33}$$

and so we are led to the following definitions by analogy with the definitions preceding (17):

$$\begin{aligned} &\text{Relative momentum of photon} = \frac{h\nu}{c} l_\varrho. \\ &\text{Relative energy of photon} = h\nu. \end{aligned} \tag{34}$$

In the kinematical picture, we think of a photon as a point moving with velocity c in the direction l_ϱ and carrying along with it a certain number $h\nu$ (its energy).

Since f_r is a null vector, so also is M_r. Since for a particle the proper mass is the magnitude of M_r, we say by analogy that *a photon has zero proper mass.*

§ 8. COLLISION AND DISINTEGRATION PROBLEMS

We have already remarked on the difficulty of extending into the theory of relativity the Newtonian theory of a system of interacting particles, on account of the absence of absolute simultaneity. This difficulty does not arise in the case of a collision between particles, and we can handle problems of collision (and disintegration) very satisfactorily in the theory of relativity.

Let us first review the Newtonian theory of collisions. The basic Newtonian law is the law of conservation of momentum; this may be written

$$\Sigma m u'_\varrho = \Sigma m u_\varrho, \tag{35}$$

where m indicates mass, u_ϱ velocity before collision, u'_ϱ velocity after collision, and the summations Σ extend over all the particles taking part in the collision. If n particles are involved, then $3n$ final velocity components occur in (35), and so in general these three equations give us only partial information about the result of the collision.

If, however, only two particles collide, and they coalesce into a single particle as the result of the collision (this we call an *inelastic* collision), then (35) tells us everything. Let $m_{(1)}$, $m_{(2)}$ be the masses of the colliding particles and $u_{(1)\varrho}$, $u_{(2)\varrho}$ their velocities before collision. If they coalesce to form a single particle of mass $m_{(3)}$ and velocity $u_{(3)\varrho}$, then (35) gives

$$m_{(3)}\, u_{(3)\varrho} = m_{(1)}\, u_{(1)\varrho} + m_{(2)}\, u_{(2)\varrho}, \tag{36}$$

as the expression of the conservation of momentum. Further, in Newtonian mechanics we postulate the conservation of mass, and so

$$m_{(3)} = m_{(1)} + m_{(2)}. \tag{37}$$

Thus the *four* equations (36) and (37) give us the *four* unknowns $m_{(3)}$, $u_{(3)\varrho}$, if we assume that $m_{(1)}$, $m_{(2)}$, $u_{(1)\varrho}$, $u_{(2)\varrho}$ are given.

In elastic or semi-elastic collisions (in which the particles rebound and do not coalesce), the result of the collision is not determined by the conservation of momentum and mass.

Before leaving the Newtonian theory of collisions, let us recall that energy is lost in a collision, unless it is perfectly elastic. Or rather it is converted into heat, and so disappears from the mechanical picture. This leakage of mechanical energy into heat (in which form it cannot be further discussed by the methods of mechanics) is an unsatisfactory feature of Newtonian theory. It is avoided in a simple and natural way in relativistic mechanics, as we shall see presently, by allowing a hot particle to have a greater proper mass than a cold one.

In relativity we lay down, as the basic *law of conservation of* 4-*momentum*, the equation

$$\Sigma M'_r = \Sigma M_r. \tag{38}$$

Here M_r denotes the 4-momentum of a particle (this word covering both material particle and photon) before collision, and M'_r the 4-momentum of a particle after collision, the summation Σ in each case extending over all the particles involved. It is not implied that the number of particles is the same before and after the collision. Although we shall continue to use the word *collision* for convenience, the same

law covers an explosion or disintegration, and indeed on any scale (large or small), provided we are prepared to accept the validity of treating the bodies involved as particles.

Let us write again, in a slightly different form, the fundamental expressions (33) and (32) for 4-momentum:

For a particle of proper mass m:

$$M_\varrho = \frac{m\gamma u_\varrho}{c}, \quad M_4 = im\gamma, \quad \gamma = \frac{1}{\sqrt{1 - u^2/c^2}}. \tag{39}$$

For a photon of frequency ν:

$$M_\varrho = \frac{h}{c^2} l_\varrho \nu, \quad M_4 = \frac{ih\nu}{c^2}. \tag{40}$$

Neither proper mass nor frequency is necessarily conserved in a collision, and so we shall write m, ν for these quantities before collision, and m', ν' after collision. Then (38) gives the four equations

$$\Sigma m'\gamma' u'_\varrho + \frac{h}{c}\Sigma l'_\varrho \nu' = \Sigma m\gamma u_\varrho + \frac{h}{c}\Sigma l_\varrho \nu, \tag{41}$$

$$\Sigma m'\gamma' + \frac{h}{c^2}\Sigma\nu' = \Sigma m\gamma + \frac{h}{c^2}\Sigma\nu. \tag{42}$$

Equation (41) expresses the relativistic law of the conservation of relative momentum (3-momentum), and (42) expresses the relativistic law of the conservation of energy (relative energy, *not* proper energy), if we multiply by c^2 and write it

$$\Sigma m'\gamma' c^2 + \Sigma h\nu' = \Sigma m\gamma c^2 + \Sigma h\nu. \tag{43}$$

Thus the four Newtonian equations ((35) plus conservation of mass) are replaced by the four equations contained in (38), which four equations may also be written as (41) together with (42) or (43). All we shall say in the present chapter about collision or disintegration problems is based on these equations.

We may also write (38) in the form

$$\Sigma m'\lambda'_r + \frac{h}{c^2}\Sigma f'_r = \Sigma m\lambda_r + \frac{h}{c^2}\Sigma f_r, \tag{44}$$

where λ_r, λ'_r are 4-velocities and f_r, f'_r frequency 4-vectors.

Through all the complexities of quantum mechanics, the law of conservation (38) remains a clear and simple guiding principle in the

discussion of atomic phenomena. In particular this applies to problems of nuclear disintegration and the formation of mesons in the phenomena of cosmic radiation.

We shall discuss some simple basic problems, using only the law of conservation. In each case we have a choice between a description in terms of kinematics and a description in terms of the geometry of space-time, or indeed we may mix the two. Before proceeding to particular problems, let us see how collisions look in terms of the geometry of space-time.

§ 9. SPACE-TIME DIAGRAMS OF COLLISIONS

Figs. 2 and 3 are space-pictures [snapshots as in v-§ 2, p. 120] of a collision, as viewed by some Galileian observer, the former just before and the latter just after the collision occurs. In the case shown, the collision of two particles has given rise to three.

Consider now a space-time diagram of this collision. Fig. 4 shows the whole affair, both before (below) and after (above) the collision-event. The arrows represent 4-momenta and also the world lines of the particles. To show graphically the conservation law (38), we make the



Fig. 2 – Snapshot of collision about to occur between two particles. Arrows indicate 3-velocities or 3-momenta

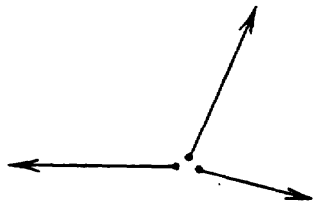


Fig. 3 – Snapshot just after collision, showing three particles

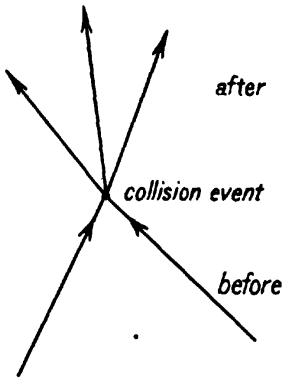


Fig. 4 – Space-time diagram of the collision shown in Figs. 2 and 3

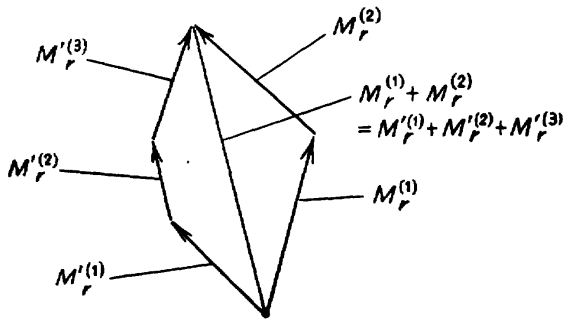


Fig. 5 – Space-time representation of the conservation law (38)

same sort of construction as we would make if we were compounding forces in ordinary statics. We draw the 4-vector $M_r^{(1)}$ from some base point in space-time, and then draw $M_r^{(2)}$ from its extremity (these are the 4-momenta of the two particles before collision). In this way we construct the 4-vector $M_r^{(1)} + M_r^{(2)}$ (Fig. 5).

If we construct similarly from the same base point the vector sum $M'^{(1)}_r + M'^{(2)}_r + M'^{(3)}_r$ (these being the three 4-momenta after collision), the law (38) tells us that we arrive at the same final point. The 4-vector

$$M_r^{(1)} + M_r^{(2)} = M'^{(1)}_r + M'^{(2)}_r + M'^{(3)}_r \tag{45}$$

is thus constructed in two different ways. The Minkowskian lengths of the constituent 4-vectors are the proper masses of the particles (zero in the case of photons).

Everyone knows how useful geometrical intuition can be in Newtonian statics where forces are compounded as we have compounded 4-momenta in Fig. 5. With some caution (the metric is indefinite!) we can use the same intuition in dealing with these space-time diagrams. If we remember that "timelike" means "lying inside the null cone", the following statement is suggested intuitively:

The sum of any number of 4-vectors, all timelike or null and all pointing into the future, is timelike, except when all the given vectors are null and have a common direction in space-time.

The formal proof of this is easy (cf. SYNGE [1934]) and is left to the reader. This theorem implies that we can never get a single photon and nothing more as the product of a collision, unless all the colliding particles are photons travelling in the same direction.

§ 10. THE TRIANGLE INEQUALITY IN SPACE-TIME

Euclid proved that any side of a triangle is less than the sum of the other two. This is the *Euclidean triangle inequality*

$$OB \leqslant OA + AB,$$

the sign of equality occurring only if the triangle collapses into a straight line with A between O and B. Euclid's proof is not suggestive for our purposes. Instead we may prove this triangle inequality by a vectorial method, using the fact that the scalar product of two unit vectors (the cosine of an angle) cannot exceed unity. The

proof reads

$$\begin{aligned} \mathbf{OB} &= \mathbf{OA} + \mathbf{AB}, \\ OB^2 = \mathbf{OB}.\mathbf{OB} &= OA^2 + AB^2 + 2\,\mathbf{OA}.\mathbf{AB} \\ &= OA^2 + AB^2 + 2\,OA\,AB\cos OAB \\ &\leqslant OA^2 + AB^2 + 2\,OA\,AB \\ &= (OA + AB)^2 \quad \text{q.e.d.} \end{aligned}$$

This gives us the key to the triangle inequality in space-time. We examine for boundedness the scalar product $\lambda_r\mu_r$ of two unit 4-vectors, λ_r and μ_r, both timelike and pointing into the future[1]. This means that

$$\begin{aligned} \lambda_r\lambda_r &= -1, \quad \lambda_4/i > 0, \\ \mu_r\mu_r &= -1, \quad \mu_4/i > 0. \end{aligned} \tag{46}$$

Let us choose λ_r for time axis: then

$$\lambda_\varrho = 0, \quad \lambda_4 = i, \tag{47}$$

and so

$$\lambda_r\mu_r = \lambda_4\mu_4. \tag{48}$$

But μ_r is a timelike unit vector, and so

$$\mu_4^2 = -1 - \mu_\varrho\mu_\varrho, \quad \mu_4 = i\sqrt{1 + \mu_\varrho\mu_\varrho}, \tag{49}$$

where we take $+i$, not $-i$, since μ_r points into the future. Then

$$\lambda_r\mu_r = i.i\sqrt{1 + \mu_\varrho\mu_\varrho} = -\sqrt{1 + \mu_\varrho\mu_\varrho}. \tag{50}$$

Thus the scalar product of two unit vectors, both timelike and pointing into the future, is negative and less than -1:

$$\lambda_r\mu_r \leqslant -1. \tag{51}$$

The sign of equality holds only when the two vectors coincide.

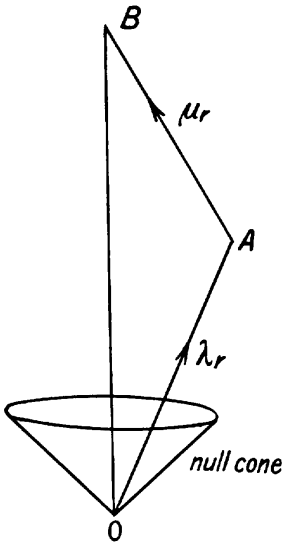


Fig. 6 – The triangle inequality in space-time: $OB \geqslant OA + AB$

Consider now a triangle OAB as in Fig. 6. Here OA_r, AB_r are timelike vectors pointing into the future; it follows that OB_r is also a timelike vector pointing into the future. Write

$$OA_r = OA\lambda_r, \quad AB_r = AB\mu_r, \tag{52}$$

where λ_r and μ_r are timelike unit vectors pointing into the future, and OA, AB are positive Minkowskian lengths. Then (note how the

[1] See also p. 46 for a different treatment.

Euclidean argument is modified)

$$\begin{aligned} OB_r &= OA_r + AB_r, \\ OB^2 = -\,OB_r OB_r &= -\,(OA_r + AB_r)\,(OA_r + AB_r) \\ &= OA^2 + AB^2 - 2\,OA_r AB_r \\ &= OA^2 + AB^2 - 2\,OA\;AB\;\lambda_r \mu_r \\ &\geqslant OA^2 + AB^2 + 2\,OA\;AB \end{aligned} \tag{53}$$

by (51). Thus we have *the triangle inequality in space-time*

$$OB \geqslant OA + AB. \tag{54}$$

The equality sign holds only if A lies on OB. We note that this Minkowskian triangle inequality reverses the sense of the Euclidean triangle inequality, and for this we have to thank the indefinite metric of space-time. The case where OA_r or AB_r is a null vector may be approached as a limit; (54) remains true with OA or AB replaced by zero.

It is easy to extend the triangle inequality (54) to give the *polygonal inequality*

$$OB \geqslant OA_1 + A_1A_2 + \ldots + A_nB, \tag{55}$$

where $OA_1, A_1A_2, \ldots A_nB$ are 4-vectors, all pointing into the future and all timelike or null (Fig. 7). The proof is left to the reader.

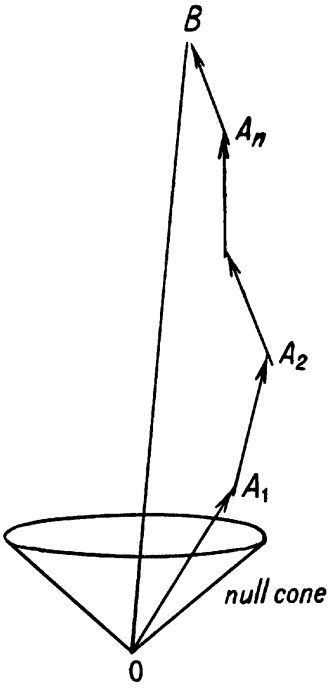


Fig. 7 – The polygonal inequality in space-time: $OB \geqslant OA_1 + A_1A_2 + \ldots + A_nB$

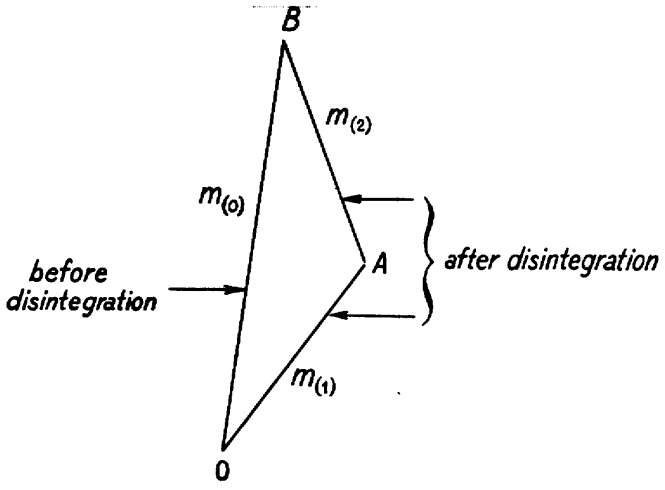


Fig. 8 – Decrease in total proper energy as the result of disintegration: $m_{(0)}c^2 > m_{(1)}c^2 + m_{(2)}c^2$.

We shall draw an immediate physical conclusion from the triangle inequality (54). Let a particle of proper mass $m_{(0)}$ disintegrate into two particles of proper masses $m_{(1)}$ and $m_{(2)}$. Fig. 8 is the appropriate space-time diagram. The proper energy before the disintegration is $E_0 = m_{(0)}c^2$; the sum of the proper energies after disintegration is $E_0' = m_{(1)}c^2 + m_{(2)}c^2$. The triangle inequality (54) tells us that

$$m_{(0)} > m_{(1)} + m_{(2)}, \quad E_0 > E_0'. \tag{56}$$

There is a *decrease* of total proper energy as a result of the disintegration. We shall examine this question more fully below.

§ 11. MASS-CENTRE REFERENCE SYSTEM. RELEASE OF ENERGY IN DISINTEGRATION

In any collision there is a vector of *total 4-momentum* (ΣM_r) and this is unchanged by the collision ($\Sigma M_r' = \Sigma M_r$). Since this vector is timelike, we may take it for time axis. If we do this, we say that we are using a *mass-centre reference system*; its use simplifies the algebra of collisions since $\Sigma M_\varrho' = 0$ in this frame. To investigate the result of a collision, we have then from (38)

$$\Sigma M_\varrho' = 0, \quad \Sigma M_4' = \Sigma M_4, \tag{57}$$

or, as in (41) and (42),

$$\begin{aligned} \Sigma m'\gamma' u_\varrho' + \frac{h}{c}\Sigma l_\varrho' \nu' &= 0, \\ \Sigma m'\gamma' + \frac{h}{c^2}\Sigma\nu' &= \Sigma m\gamma + \frac{h}{c^2}\Sigma\nu. \end{aligned} \tag{58}$$

We have considered briefly above the case where a particle disintegrates into two particles. Let us now be more general and suppose that a single particle of proper mass m disintegrates into any number of particles (material particles and photons). Let us use the mass-centre reference system S, viz. that in which the original particle was at rest. Then the last of (58) reads, on multiplication by c^2,

$$\Sigma m'\gamma' c^2 + \Sigma h\nu' = mc^2. \tag{59}$$

Each side of this equation represents the relative energy, the left hand side after, and the right hand side before, disintegration; the right hand side also represents the proper energy of the particle which disintegrates.

It is apparently a fact of nature that there exist particles of constant proper masses (the ultimate particles — electron, proton and mesons). Hence it is sensible physically to think of catching such particles shot off in the disintegration and bringing them to rest in the frame S *without changing their proper masses.* We let the photons go; they cannot be brought to rest in any frame. Collecting the pieces thus, we have a set of particles of total proper mass $\Sigma m'$ and total proper energy $\Sigma m'c^2$; this is also their total relative energy, since they are now at rest in S. Relative energy has been lost, and the amount of the loss is

$$Q = \Sigma m'\gamma'c^2 + \Sigma h\nu' - \Sigma m'c^2. \tag{60}$$

In the definitions of p. 169 the relative kinetic energy of a particle was given as $T = mc^2(\gamma - 1)$, and so we can write (60) in the form

$$Q = \Sigma T' + \Sigma h\nu'. \tag{61}$$

Or again we may write it

$$Q = mc^2 - \Sigma m'c^2. \tag{62}$$

When a particle at rest disintegrates, it releases an amount of relative energy Q given by any one of the above formulae.

Atomic nuclei are regarded as composed of protons and neutrons. If m is the proper mass of the nucleus and $\Sigma m'$ the sum of the proper masses of the protons and neutrons which compose it, then the expression (62) is called the *binding energy* of the nucleus. It is clear from (61) that if a nucleus is capable of exploding spontaneously into its constituent protons and neutrons, together with photons, the binding energy must be positive; such a nucleus would be unstable. If the binding energy is negative, a spontaneous explosion cannot occur, because it would violate the polygonal inequality (55).

For example, the deuteron nucleus consists of one proton and one neutron. The proper masses are as follows, expressed in units for which the oxygen atom is 16:[1]

$$\text{deuteron nucleus: } 2.01417 = m,$$

$$\left.\begin{array}{ll}\text{proton:} & 1.00757\\ \text{neutron:} & 1.00893\end{array}\right\} \Sigma m' = 2.01650 > m.$$

[1] Cf. BETHE [1947, pp. 5, 23–25]; in the numbers given above, the mass of an orbital electron (0.00055) is not included. I follow Bethe in the definition of binding energy; some writers prefer to define it as $-Q$, so that it is positive for a stable nucleus.

The binding energy is negative, and so it is impossible for the deuteron nucleus to explode spontaneously into a proton and a neutron. Further, if a proton and a neutron collide to form a deuteron nucleus, this must be accompanied by the emission of a material particle or photon to carry away the excess energy.

§ 12. SOME NUMERICAL VALUES

Although it is not the purpose of this book to go into practical details, the laws governing collisions and disintegrations are so simple and fundamental that some numerical values will be quoted here in order that the reader may have them to hand if he wants to see

TABLE 1

Planck's constant: $h = 6.62 \times 10^{-27}$ gm cm^2 sec^{-1}
Velocity of light: $c = 3.00 \times 10^{10}$ cm sec^{-1}

Material particle	Proper mass m (gm)	Proper energy mc^2 (ergs)	Relative kinetic energy T in ergs (cf. p. 169) when moving with velocity 0.1 c	0.995 c
Electron	9.11×10^{-28}	8.20×10^{-7}	4.13×10^{-9}	7.39×10^{-6}
Proton	1.67×10^{-24}	1.51×10^{-3}	7.61×10^{-6}	1.36×10^{-2}
α-particle (He nucleus)	6.64×10^{-24}	5.98×10^{-3}	3.01×10^{-5}	5.39×10^{-2}
Atom of U 235	3.90×10^{-22}	3.51×10^{-1}	1.77×10^{-3}	3.16
One gram of matter	1	9.00×10^{20}	4.53×10^{18}	8.11×10^{21}

Photon	Wave length λ (cm)	Frequency ν (sec^{-1})	Relative energy $h\nu$ (ergs)
X-rays (radium L-series)	1.01×10^{-8}	2.97×10^{18}	1.97×10^{-8}
Visible light (sodium D line)	5.89×10^{-5}	5.09×10^{14}	3.37×10^{-12}
(first line of Balmer series for hydrogen)	6.56×10^{-5}	4.57×10^{14}	3.03×10^{-12}
Short wave radar	5	6.00×10^{9}	3.97×10^{-17}
Medium wave radio	5×10^{4}	6.00×10^{5}	3.97×10^{-21}

the magnitudes which may be involved. All quantities are expressed in c.g.s. units.[1]

§ 13. INELASTIC COLLISION OF TWO PARTICLES

Consider two particles which collide and coalesce. Their proper masses and velocities before collision are supposed to be known; our problem is to find the proper mass and velocity of the single particle formed by them.

Let

$m_{(1)}$ = proper mass of first particle before collision,
$m_{(2)}$ = proper mass of second particle before collision,
$m_{(3)}$ = proper mass of joint particle after collision,
$\lambda_{(1)r}$ = 4-velocity of first particle before collision,
$\lambda_{(2)r}$ = 4-velocity of second particle before collision,
$\lambda_{(3)r}$ = 4-velocity of joint particle after collision.

The conservation law (44) gives

$$m_{(3)}\lambda_{(3)r} = m_{(1)}\lambda_{(1)r} + m_{(2)}\lambda_{(2)r}. \tag{63}$$

Now

$$\lambda_{(3)r}\lambda_{(3)r} = \lambda_{(1)r}\lambda_{(1)r} = \lambda_{(2)r}\lambda_{(2)r} = -1, \tag{64}$$

and so (63) gives

$$m_{(3)}^2 = m_{(1)}^2 + m_{(2)}^2 - 2m_{(1)}m_{(2)}\lambda_{(1)r}\lambda_{(2)r}. \tag{65}$$

Here we have the proper mass of the joint particle in terms of known quantities; then (63) gives $\lambda_{(3)r}$ and the problem is solved.

Let us express $m_{(3)}$ in terms of the 3-velocities $u_{(1)\varrho}$, $u_{(2)\varrho}$ before collision.

We have

$$\begin{aligned} \lambda_{(1)\varrho} &= \frac{\gamma_{(1)}u_{(1)\varrho}}{c}, \quad \lambda_{(1)4} = i\gamma_{(1)}, \\ \lambda_{(2)\varrho} &= -\frac{\gamma_{(2)}u_{(2)\varrho}}{c}, \quad \lambda_{(2)4} = i\gamma_{(2)}, \end{aligned} \tag{66}$$

[1] The values shown are rounded off to three significant figures. For more accurate values, see Smithsonian Physical Tables (Ninth Edition, Washington, 1954); for example,

$$c = 2.99776 \pm .00004 \times 10^{10} \text{ cm sec}^{-1}$$

where

$$\gamma_{(1)} = \frac{1}{\sqrt{1 - u_{(1)}^2/c^2}}, \quad \gamma_{(2)} = \frac{1}{\sqrt{1 - u_{(2)}^2/c^2}}. \tag{67}$$

Thus we have for the proper mass of the joint particle

$$m_{(3)}^2 = m_{(1)}^2 + m_{(2)}^2 + 2m_{(1)}m_{(2)}\gamma_{(1)}\gamma_{(2)}\left(1 - \frac{u_{(1)\varrho}u_{(2)\varrho}}{c^2}\right). \tag{68}$$

If the ratios $u_{(1)}/c$, $u_{(2)}/c$ are small, this gives approximately

$$\begin{aligned} m_{(3)}^2 &= m_{(1)}^2 + m_{(2)}^2 + 2m_{(1)}m_{(2)} \\ &\qquad + m_{(1)}m_{(2)}\frac{u_{(1)}^2 + u_{(2)}^2 - 2u_{(1)\varrho}u_{(2)\varrho}}{c^2} \\ m_{(3)} &= m_{(1)} + m_{(2)} + \tfrac{1}{2}\frac{m_{(1)}m_{(2)}}{m_{(1)} + m_{(2)}}\frac{u_{(1)}^2 + u_{(2)}^2 - 2u_{(1)\varrho}u_{(2)\varrho}}{c^2} \\ &= m_{(1)} + m_{(2)} + \tfrac{1}{2}\frac{m_{(1)}m_{(2)}}{m_{(1)} + m_{(2)}}\frac{v^2}{c^2}, \end{aligned} \tag{69}$$

where v is the magnitude of the relative velocity of the two particles before collision; that is, $v^2 = v_\varrho v_\varrho$, $v_\varrho = u_{(1)\varrho} - u_{(2)\varrho}$. Hence the gain in proper energy as a result of the collision is approximately

$$(m_{(3)} - m_{(1)} - m_{(2)})c^2 = \tfrac{1}{2}\frac{m_{(1)}m_{(2)}}{m_{(1)} + m_{(2)}}v^2. \tag{70}$$

Let us compare this gain in proper energy with the energy dissipated into heat according to Newtonian theory. According to Newtonian theory (the m's being Newtonian masses), we have

$$(m_{(1)} + m_{(2)})u_{(3)\varrho} = m_{(1)}u_{(1)\varrho} + m_{(2)}u_{(2)\varrho}, \tag{71}$$

and hence

$$(m_{(1)} + m_{(2)})^2 u_{(3)}^2 = m_{(1)}^2 u_{(1)}^2 + m_{(2)}^2 u_{(2)}^2 + 2m_{(1)}m_{(2)}u_{(1)\varrho}u_{(2)\varrho}. \tag{72}$$

Thus the loss of kinetic energy (turned into heat) is

$$\begin{aligned} &\tfrac{1}{2}m_{(1)}u_{(1)}^2 + \tfrac{1}{2}m_{(2)}u_{(2)}^2 - \tfrac{1}{2}(m_{(1)} + m_{(2)})u_{(3)}^2 \\ &= \tfrac{1}{2}m_{(1)}u_{(1)}^2 + \tfrac{1}{2}m_{(2)}u_{(2)}^2 \\ &\quad - \frac{1}{2(m_{(1)} + m_{(2)})}(m_{(1)}^2 u_{(1)}^2 + m_{(2)}^2 u_{(2)}^2 + 2m_{(1)}m_{(2)}u_{(1)\varrho}u_{(2)\varrho}) \\ &= \tfrac{1}{2}\frac{m_{(1)}m_{(2)}}{m_{(1)} + m_{(2)}}(u_{(1)}^2 + u_{(2)}^2 - 2u_{(1)\varrho}u_{(2)\varrho}), \end{aligned} \tag{73}$$

agreeing with the approximate relativistic formula (70) for the increase in proper energy.

§ 14. DISINTEGRATION OF ONE PARTICLE INTO TWO

Consider the spontaneous disintegration of an atom, in which a single particle (electron, proton, neutron, or α-particle) is thrown off. Let $m_{(3)}$ be the proper mass of the parent particle, $m_{(1)}$ the proper mass of the particle shot off, and $m_{(2)}$ the proper mass of the residual particle.

This problem of disintegration is similar to the problem of inelastic collision discussed above, but now the problem is approached from the other end, the knowns and unknowns being interchanged. We shall use a different method.

Take a Galileian frame in which the parent particle is at rest, and choose the x_1-axis in the direction in which $m_{(1)}$ is shot off. We have then

$$u_{(3)} = 0, \quad \gamma_{(3)} = 1, \quad u_{(1)1} = u_{(1)}, \quad u_{(1)2} = u_{(1)3} = 0. \tag{74}$$

The conservation equation (41) gives $u_{(2)2} = u_{(2)3} = 0$, $u_{(2)1} = -u_{(2)}$, and hence with (42)

$$\begin{aligned} m_{(1)}\gamma_{(1)}u_{(1)} &= m_{(2)}\gamma_{(2)}u_{(2)}, \\ m_{(1)}\gamma_{(1)} + m_{(2)}\gamma_{(2)} &= m_{(3)}. \end{aligned} \tag{75}$$

If the three proper masses are known, then $u_{(1)}$ and $u_{(2)}$ can be calculated; this can be done if we know the atomic weights of the parent particle and the residual particle, and also the mass of the particle shot off. But we can look at (75) more generally as two equations connecting five quantities, and we can compute them all if three are given.

The algebra is most conveniently handled by introducing pseudo-angles ϕ_1, ϕ_2 defined by

$$\tanh\phi_1 = u_{(1)}/c, \quad \tanh\phi_2 = u_{(2)}/c, \tag{76}$$

so that

$$\begin{aligned} \cosh\phi_1 &= \gamma_{(1)}, \quad \cosh\phi_2 = \gamma_{(2)}, \\ \sinh\phi_1 &= \gamma_{(1)}u_{(1)}/c, \quad \sinh\phi_2 = \gamma_{(2)}u_{(2)}/c. \end{aligned} \tag{77}$$

Then (75) read

$$\begin{aligned} m_{(1)}\sinh\phi_1 &= m_{(2)}\sinh\phi_2, \\ m_{(1)}\cosh\phi_1 + m_{(2)}\cosh\phi_2 &= m_{(3)}. \end{aligned} \tag{78}$$

Elimination of ϕ_2 gives

$$\cosh \phi_1 = \frac{m_{(3)}^2 + m_{(1)}^2 - m_{(2)}^2}{2m_{(3)} m_{(1)}}, \tag{79}$$

a formula analogous to the formula of elementary trigonometry for the cosine of an angle of a triangle in terms of the sides — indeed what we are doing here may be called the pseudotrigonometry of a Minkowskian triangle. From (79) we get

$$\sinh \phi_1 = \frac{A}{2m_{(3)} m_{(1)}}, \tag{80}$$

$$u_{(1)}/c = \tanh \phi_1 = \frac{A}{m_{(3)}^2 + m_{(1)}^2 - m_{(2)}^2}, \tag{81}$$

where

$$A^2 = (m_{(3)} + m_{(1)} + m_{(2)})\,(m_{(3)} - m_{(1)} + m_{(2)}) \times (m_{(3)} + m_{(1)} - m_{(2)})\,(m_{(3)} - m_{(1)} - m_{(2)}).$$

There is a similar formula for $u_{(2)}/c$, differing from (81) only by interchange of $m_{(1)}$ and $m_{(2)}$ in the denominator. Thus we have explicit formulae for the speeds of the two particles after the disintegration in terms of the three proper masses; but the directions of the velocities are undetermined, except that we know that they are opposite to one another.

If $m_{(1)}$, $m_{(2)}$ and $u_{(1)}$ are given, the determination of $m_{(3)}$ from (79) involves only the solution of a quadratic equation. There is a simple approximate solution if we suppose that $(m_{(3)} - m_{(1)} - m_{(2)})$ is small and write

$$m_{(3)} - m_{(1)} - m_{(2)} = (m_{(1)} + m_{(2)})\delta, \tag{82}$$

where δ is small. Then (81) gives approximately

$$\frac{u_{(1)}}{c} = \left(\frac{2m_{(2)}\,\delta}{m_{(1)}}\right)^{\frac{1}{2}}, \quad \delta = \frac{m_{(1)}}{2m_{(2)}}\left(\frac{u_{(1)}}{c}\right)^2, \tag{83}$$

and so, by (82),

$$m_{(3)} = m_{(1)} + m_{(2)} + \tfrac{1}{2}\frac{m_{(1)}}{m_{(2)}}(m_{(1)} + m_{(2)})\left(\frac{u_{(1)}}{c}\right)^2. \tag{84}$$

This is a simple approximate formula for the proper mass of the parent particle in terms of the proper masses of the two resulting particles and the speed of emission of one of them (small in comparison

with c), the frame of reference being that in which the parent particle was at rest.

§ 15. EMISSION OF A PHOTON FROM AN ATOM

If we think of an atom as a particle in the sense in which we have been using the word, and suppose the atom capable of having certain discrete energies determined by quantum conditions, then we must consider those energy levels to be levels of *proper* energy, since no privileged frame of reference is to hand except the frame of the atom itself. Thus when an atom radiates, sending out a photon, it changes its proper mass according to the equation

$$mc^2 - m'c^2 = W, \tag{85}$$

where

m = proper mass before emission,

m' = proper mass after emission,

W = a quantity of energy determined by the rules of quantum mechanics.

The conservation law gives

$$M_r = M'_r + P_r, \tag{86}$$

where

M_r = 4-momentum of atom before emission,

M'_r = 4-momentum of atom after emission,

P_r = 4-momentum of emitted photon.

By (22) and (31) we have

$$M_r M_r = -m^2, \quad M'_r M'_r = -m'^2, \quad P_r P_r = 0, \tag{87}$$

and so, by (86),

$$0 = P_r P_r = M_r M_r + M'_r M'_r - 2M_r M'_r. \tag{88}$$

If we regard m and W as known, then (85) gives m',

$$m' = m - W/c^2, \tag{89}$$

and by (88) we have for the scalar product of the two momentum 4-vectors, before and after emission,

$$M_r M'_r = -\tfrac{1}{2}(m^2 + m'^2). \tag{90}$$

We know only two things about the final 4-momentum M_r', viz. that its magnitude is m' as in (89) and that it satisfies (90). These facts do not determine it; it has still two degrees of freedom. To explore the situation geometrically, we draw (with some centre O) the pseudospheres Σ, Σ' with radii m, m', respectively (Fig. 9). Drawing vectors from O, we note that the extremity of M_r lies on Σ and the extremity of M_r' on Σ', and (86) tells us that, since P_r is a null vector, the extremity of M_r' also lies on the null cone drawn into the past from the extremity of M_r. Thus all we know is that the extremity of M_r' lies in a certain 2-space, namely, the intersection of the null cone just described and the pseudosphere Σ'.

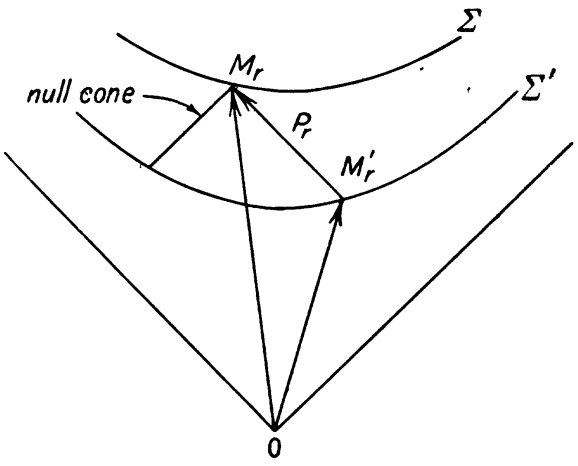


Fig. 9 – Space-time diagram of the emission of a photon from an atom

To investigate the emitted photon, we recall that by (32)

$$P_\varrho = hvl_\varrho/c^2, \quad P_4 = ihv/c^2, \tag{91}$$

where v is its frequency and l_ϱ are the direction cosines of its line of propagation, referred to any Galileian frame of reference we choose to use. By (86) and (90) we have

$$\begin{aligned} P_r M_r &= M_r M_r - M_r M_r' \\ &= -m^2 + \tfrac{1}{2}(m^2 + m'^2) \\ &= -\tfrac{1}{2}(m^2 - m'^2), \\ P_r M_r' &= M_r' M_r - M_r' M_r' \\ &= -\tfrac{1}{2}(m^2 + m'^2) + m'^2 \\ &= -\tfrac{1}{2}(m^2 - m'^2). \end{aligned} \tag{92}$$

Thus *these two scalar products are equal*, an invariant result, independent of the frame of reference.

If we use the frame of reference in which the atom is at rest *before* emission, then we have $M_\varrho = 0$, $M_4 = im$, and so, by (91) and the

first of (92),

$$h\nu = \frac{c^2}{2m}(m^2 - m'^2). \tag{93}$$

On the other hand, if we use the frame of reference in which the atom is at rest *after* emission, then similarly

$$h\nu = \frac{c^2}{2m'}(m^2 - m'^2). \tag{94}$$

The difference between (93) and (94) is very small and unobservable spectroscopically. If we use neither of these frames of reference, then we shall get other frequencies on account of the Doppler effect (p. 136). It was by the observation of such Doppler effects that the clock retardation described on p. 120 was verified experimentally (IVES and STILWELL [1938, 1941], IVES [1952]).

§ 16. THE SAMENESS OF PHOTONS

The frequency ν of a photon emitted in a transition from one energy level to another, in accordance with (85), may have any value at all from zero to infinity, depending on the frame of reference in which it is observed. It is only when we consider emissions from atoms having fairly small velocities relative to the frame of the observer that we can speak of a small range of frequencies corresponding to a definite transition.

This brings out an interesting difference between material particles and photons. A material particle has a proper mass which is Lorentz-invariant, and consequently independent of the observer. If we say that a particle A has a greater proper mass than a particle B, this statement is absolute, i.e. independent of the observer. It is a much more significant statement than a statement to the effect that a particle A is moving faster than another particle B. For there exists a frame in which A is at rest, and another in which B is at rest, and in fact there are not, in any absolute sense, fast particles and slow particles.

We cannot make an absolute comparison between two photons on the basis of proper mass, because the proper mass of every photon is zero. In fact, if we regard a photon as something completely described by its frequency vector f_r, there are no absolute numbers which we can attach to two photons for purposes of comparison. If we meet a photon,

we cannot say that it was generated by any particular atomic emission, because all photons are essentially the same in the sense that any assigned frequency vector can be transformed into any other assigned frequency vector by a Lorentz transformation.[1]

Suppose now that we have two frequency vectors with different directions in space-time (Fig. 10). We draw these vectors, say f_r and f'_r, from some base event O, and then form the vectors

$$U_r = f_r - f'_r, \quad V_r = f_r + f'_r. \tag{95}$$

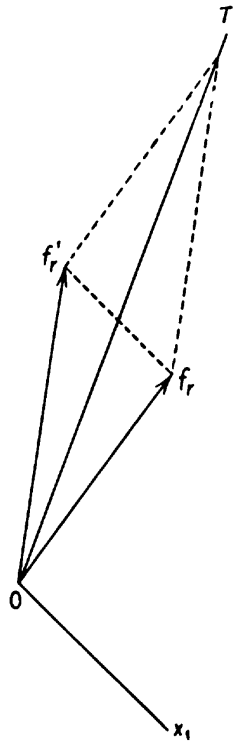


Fig. 10 – Mass-centre reference system for two photons

These vectors are orthogonal to one another; they lie in the 2-flat of the two frequency vectors, with U_r spacelike and V_r timelike and pointing into the future. We may take our time-axis in the direction of V_r and our x_1-axis in the direction of U_r. Then for these special coordinates we have

$$\begin{aligned} &U_2 = U_3 = U_4 = 0, \\ &V_1 = V_2 = V_3 = 0, \quad V_4/i > 0, \end{aligned} \tag{96}$$

and so, by (95),

$$\begin{aligned} &f_1 + f'_1 = 0, \quad f_4 - f'_4 = 0, \\ &f_2 = f_3 = f'_2 = f'_3 = 0. \end{aligned} \tag{97}$$

Thus, by v-(91), the frequencies are the same and the two photons move parallel to the x_1-axis, one in one direction and the other in the opposite direction. This reduction is always possible unless f_r and f'_r have the same direction in space-time, a case already considered on p. 138 (Fig. v–12).

To sum up:

(*a*) If two photons appear to some Galileian observer to travel in the same direction with different frequencies, then this holds for all Galileian observers, the observed values of the frequencies ranging from zero to infinity according to the observer, but with a ratio of frequencies the same for all observers.

(*b*) If two photons appear to some Galileian observer to travel in

[1] This simple approach, in which f_r is regarded as a complete description of a photon, leaves out of account the *polarisation* of the electromagnetic field which accompanies or directs the photon; this polarisation should in some way appear in its description.

different directions, then there exists another Galileian observer to whom the photons appear to have a common frequency and to travel in opposite directions.

What we have actually done above is to introduce the mass-centre reference system of § 11. In case (*b*) the use of this special frame produces standing waves for the two wave-systems with frequency vectors f_r, f'_r, provided the two amplitudes are equal (cf. RISCO [1952]). For a discussion in terms of Maxwellian theory, see IX–§ 11.

§ 17. THE EMISSION AND ABSORPTION OF A PHOTON

Consider two like atoms, say of hydrogen. Let atom A be in an excited state, with proper mass m, and let atom B be in the ground state, with proper mass m'. We understand that m is greater than m', and that the two are connected by the quantum condition (85), in which W is known.

Let A fall to the ground state (proper mass m'), emitting a photon. Under what circumstances can this photon be absorbed by B, raising it to the excited state with proper mass m? We shall now discuss this question on the basis of the law of conservation of 4-momentum and the quantum law (85), both being assumed absolutely accurate.[1]

Let

M_r = 4-momentum of A before emission,
M'_r = 4-momentum of A after emission,
$\bar{M}'_r$ = 4-momentum of B before absorption,
$\bar{M}_r$ = 4-momentum of B after absorption,
P_r = 4-momentum of photon emitted by A and absorbed by B.

We have the following equations:

$$M_r M_r = -m^2, \quad \bar{M}'_r \bar{M}'_r = -m'^2, \tag{98}$$

$$\left.\begin{aligned} M'_r M'_r &= -m'^2, \quad \bar{M}_r \bar{M}_r = -m^2, \\ M_r &= M'_r + P_r, \quad \bar{M}_r = \bar{M}'_r + P_r, \\ P_r P_r &= 0. \end{aligned}\right\} \tag{99}$$

Regarding m and m' as known, we ask: What are the restrictions on the five vectors M_r, M'_r, $\bar{M}_r$, $\bar{M}'_r$, P_r satisfying these thirteen equations?

We shall proceed by regarding M_r and $\bar{M}'_r$ as known; i.e. we suppose

[1] Line breadth is neglected in this precise mathematical argument.

that we know all about the atom A before emission and the atom B before absorption. This means that (98) are satisfied, and we need pay no more attention to those equations. Then the condition (necessary and sufficient) for the existence of M'_r and $\bar{M}_r$ satisfying (99) is that M_r, $\bar{M}'_r$, P_r should satisfy

$$\begin{aligned}(M_r - P_r)\,(M_r - P_r) &= -\,m'^2,\\ (\bar{M}'_r + P_r)\,(\bar{M}'_r + P_r) &= -\,m^2,\\ P_r P_r &= 0,\end{aligned} \tag{100}$$

or equivalently

$$\begin{aligned}P_r M_r = -\tfrac{1}{2}(m^2 - m'^2), \quad P_r \bar{M}'_r = -\tfrac{1}{2}(m^2 - m'^2),\\ P_r P_r = 0.\end{aligned} \tag{101}$$

These are three conditions on the four components of P_r, if we regard M_r and $\bar{M}'_r$ as known.

So far the frame of reference is unspecified. Let us now choose it so that atom A is at rest before emission. Then

$$M_\varrho = 0, \quad M_4 = im, \tag{102}$$

and so the first of (101) gives

$$P_4 = \frac{i}{2m}(m^2 - m'^2), \tag{103}$$

and the remaining equations read

$$\begin{aligned}P_\varrho \bar{M}'_\varrho &= -\frac{1}{2m}(m^2 - m'^2)\,(m + i\bar{M}'_4),\\ P_\varrho P_\varrho &= \frac{1}{4m^2}(m^2 - m'^2)^2.\end{aligned} \tag{104}$$

The existence of real numbers P_ϱ satisfying these equations implies a restriction on $\bar{M}'_r$. For the first of (104) puts the extremity of the 3-vector P_ϱ on a plane and the second equation puts it on a sphere. For a solution to exist, the plane must cut the sphere, and the condition for this is

$$\frac{1}{2m}(m^2 - m'^2)\,|m + i\bar{M}'_4|\,(\bar{M}'_\varrho \bar{M}'_\varrho)^{-\frac{1}{2}} < \frac{1}{2m}(m^2 - m'^2), \tag{105}$$

or

$$(m + i\bar{M}'_4)^2 < \bar{M}'_\varrho \bar{M}'_\varrho, \tag{106}$$

or

$$m^2 + 2im\bar{M}'_4 < \bar{M}'_r \bar{M}'_r = -\,m'^2. \tag{107}$$

Hence the condition (necessary and sufficient) for the possibility of emission and absorption as described is

$$-2i\bar{M}_4' m > m^2 + m'^2. \tag{108}$$

If v_ϱ is the velocity of B (before absorption) relative to A (before emission) and $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$, $v^2 = v_\varrho v_\varrho$, then $\bar{M}_4' = im'\gamma$, and the condition (108) is

$$\gamma > \frac{m^2 + m'^2}{2mm'}. \tag{109}$$

Thus the phenomenon certainly cannot take place if B is at rest ($\gamma = 1$). B must have, relative to A, a velocity satisfying

$$1 - \frac{v^2}{c^2} < \left(\frac{2mm'}{m^2 + m'^2}\right)^2, \tag{110}$$

or equivalently

$$\frac{v}{c} > \frac{(m - m')(m + m')}{m^2 + m'^2}. \tag{111}$$

Since $mc^2 - m'c^2 = W$ and W is small compared with mc^2, this condition is approximately

$$\frac{v}{c} > \frac{W}{mc^2} = \frac{h\nu}{mc^2}, \tag{112}$$

where ν is the frequency of the photon. This gives a lower bound for v if B is to absorb the photon emitted by A.

For the first line of the Balmer series for hydrogen, we have (p. 182) $h\nu = 3.03 \times 10^{-12}$, and for the proton $mc^2 = 1.51 \times 10^{-3}$. Hence, under the sharp assumptions made, a second hydrogen atom can absorb this photon only if it is moving relative to the emitting atom with a velocity satisfying

$$\frac{v}{c} > \frac{3.03 \times 10^{-12}}{1.51 \times 10^{-3}} = 2.01 \times 10^{-9},$$

$$v > 2.01 \times 10^{-9} \times 3.00 \times 10^{10} = 60.3 \text{ cm sec}^{-1}. \tag{113}$$

§ 18. THE COMPTON EFFECT

If a photon collides with an electron and rebounds from it, we have what is called the *Compton effect*. If

M_r = 4-momentum of electron before collision,
M_r' = 4-momentum of electron after collision,
P_r = 4-momentum of photon before collision,
P_r' = 4-momentum of photon after collision,

we have the following eight equations

$$M_r M_r = -m^2, \quad P_r P_r = 0, \tag{114}$$

$$\left.\begin{aligned} M'_r M'_r = -m^2, \quad P'_r P'_r = 0, \\ M'_r + P'_r = M_r + P_r. \end{aligned}\right\} \tag{115}$$

Here m is the proper mass of the electron, treated as an absolute constant.

If we regard M_r, P_r as given and seek to determine M'_r, P'_r, we have in (115) only six equations for eight unknowns. Hence the solution has two degrees of indeterminacy.[1]

There are a number of different ways of looking at the Compton effect. The first is the space-time way, in which we do not tie ourselves down to any particular frame of reference; the other ways involve special choices of the frame of reference.

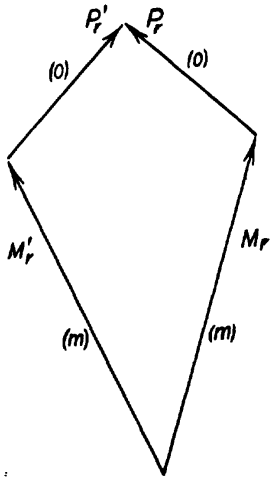


Fig. 11 – Space-time diagram of the Compton effect

Let us first consider the space-time way. Fig. 11 is a space-time diagram of the Compton effect. We see the two timelike vectors M_r, M'_r, each of Minkowskian length m, and the two null vectors P_r, P'_r drawn from their respective extremities. The figure is closed by virtue of the conservation law. If M_r, P_r are given, the construction of the rest of the figure has two degrees of freedom, as indicated above.

Let us now consider snapshots of the Compton effect, using the following frames of reference:

a) A frame in which the electron is at rest before the collision.

b) The mass-centre reference system.

c) Frame with time-axis along $(M_r + M'_r)$.

d) Frame with time-axis along $(P_r + P'_r)$.

In discussing these we shall use the following notation:

v_e = velocity of electron before collision,

v'_e = velocity of electron after collision,

[1] In VII–§ 13 we shall take intrinsic angular momentum into account, and then this indeterminacy disappears, leaving us with a finite number of solutions, at least two and at most four, if we include among the solutions the trivial case where the particle and the photon do not affect one another at all.

ν = frequency of photon before collision,
ν' = frequency of photon after collision,
l_ϱ = direction cosines of velocity of photon before collision,
l'_ϱ = direction cosines of velocity of photon after collision.

a) Electron at rest before collision.

Instead of (115) we shall use the conservation equations in the form (41), (42) with appropriate change in notation. By our choice of frame, we have $v_\varrho = 0$, and so the conservation equations read

$$\begin{aligned} m\gamma' v'_\varrho + h\nu' l'_\varrho/c &= h\nu l_\varrho/c, \\ m\gamma' + h\nu'/c^2 &= m + h\nu/c^2. \end{aligned} \tag{116}$$

Here the unknowns are v'_ϱ, l'_ϱ, ν', but these equations do not suffice to determine them. We can impose two more conditions, and this we shall do by assigning the direction of the photon's motion after collision, i.e. by assigning l'_ϱ, subject of course to $l'_\varrho l'_\varrho = 1$.

We now proceed to solve (116) for v'_ϱ and ν'. To simplify the algebra, we write

$$\xi = \nu'/\nu, \quad k = h\nu/(mc^2), \qquad \cos\theta = l_\varrho l'_\varrho, \tag{117}$$

θ being the angle through which the photon is deflected (Fig. 12). Then (116) may be written

$$\gamma' v'_\varrho/c = kl_\varrho - \xi k l'_\varrho, \tag{118}$$

$$\gamma' - 1 = k(1 - \xi). \tag{119}$$

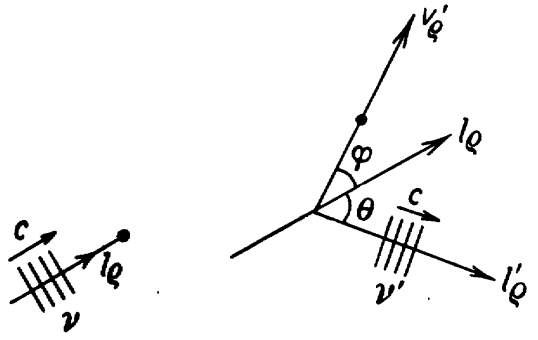


Fig. 12 – Snapshots of the Compton effect, with the electron at rest before collision[1]

Now $v'^2/c^2 = 1 - \gamma'^{-2}$, and so squaring (118) gives

$$\gamma'^2 - 1 = \gamma'^2 v'^2/c^2 = k^2(1 - 2\xi\cos\theta + \xi^2). \tag{120}$$

Then, by (119),

$$\gamma' + 1 = \frac{\gamma'^2 - 1}{\gamma' - 1} = \frac{k(1 - 2\xi\cos\theta + \xi^2)}{1 - \xi}. \tag{121}$$

Subtraction of (119) from (121) gives us an equation for ξ, which reduces to

$$2k\,\xi\sin^2\tfrac{1}{2}\theta = 1 - \xi. \tag{122}$$

[1] In this pair of snapshots, and likewise in those which follow, the one on the left is taken *before* and the one on the right *after*.

Hence we have the frequency after collision:

$$\nu' = \xi\nu = \frac{\nu}{1 + 2k \sin^2 \tfrac{1}{2}\theta}; \quad k = \frac{h\nu}{mc^2}. \tag{123}$$

From (122) we get also a neat formula for the change in wave length λ, since $\lambda\nu = c$:

$$\lambda' - \lambda = \frac{2h}{mc} \sin^2 \tfrac{1}{2}\theta. \tag{124}$$

To find the speed of the electron's recoil, we have by (122)

$$\xi = \frac{1}{1 + 2k \sin^2 \tfrac{1}{2}\theta}, \tag{125}$$

and so by (119)

$$\gamma' = \frac{1 + 2k(1 + k) \sin^2 \tfrac{1}{2}\theta}{1 + 2k \sin^2 \tfrac{1}{2}\theta}. \tag{126}$$

This gives

$$\frac{v'}{c} = \left(1 - \frac{1}{\gamma'^2}\right)^{\frac{1}{2}} = \frac{2k \sin \tfrac{1}{2}\theta \,[1 + k(2 + k) \sin^2 \tfrac{1}{2}\theta]^{\frac{1}{2}}}{1 + 2k(1 + k) \sin^2 \tfrac{1}{2}\theta}. \tag{127}$$

As for the direction of recoil, let ϕ be the angle between the velocity of the electron and that of the photon before collision, with sense as shown in Fig. 12. (It is clear from (116) that v'_ϱ, l_ϱ, l'_ϱ are coplanar.) Then by (118) we have

$$\begin{aligned} (\gamma' v'/c) \sin \phi &= \xi k \sin \theta, \\ (\gamma' v'/c) \cos \phi &= k - \xi k \cos \theta, \end{aligned} \tag{128}$$

and so, by (125),

$$\tan \phi = \frac{\cot \tfrac{1}{2}\theta}{1 + k}. \tag{129}$$

In view of the indeterminacy in the problem, it is interesting to examine the bounds on the quantities when the direction of the photon after collision is regarded as arbitrary. Thus, the frequency after collision passes from a maximum $\nu' = \nu$ for $\theta = 0$ to a minimum $\nu' = \nu(1 + 2k)^{-1}$ for $\theta = \pi$, in which last case the photon is thrown straight back. In that case (124) gives

$$\lambda' - \lambda = \frac{2h}{mc}; \tag{130}$$

the quantity $h/(mc)$ is called the *Compton wave length* of the electron ($h/mc = 2.4 \times 10^{-10}$ cm).[1]

By (129) $\tan\phi$ is positive; the electron is always thrown forward. As for the bounds on the speed of recoil, these are best investigated by means of (126). It is easy to see that γ' is a monotonically increasing function of θ for $0 \leqslant \theta \leqslant \pi$, and consequently so is v'. Therefore v' ranges from zero for $\theta = 0$ (in which case the photon does nothing to the electron) to a maximum for $\theta = \pi$, given by

$$\frac{v'}{c} = \frac{2k(1+k)}{1+2k+2k^2}. \tag{131}$$

b) Mass-centre reference system.

The unpleasant asymmetric algebra involved above was of our own making, for we chose a frame of reference badly suited to the space-time diagram Fig. 11. We shall now take a better frame, viz. that of the mass-centre (p. 180). Instead of (116), we have now

$$m\gamma v_Q + hvl_Q/c = 0, \quad m\gamma' v'_Q + hv'l'_Q/c = 0, \tag{132}$$

$$m\gamma' + hv'/c^2 = m\gamma + hv/c^2. \tag{133}$$

From (132) we get

$$m^2(\gamma^2 - 1) = (hv/c^2)^2, \quad m^2(\gamma'^2 - 1) = (hv'/c^2)^2, \tag{134}$$

and so, by (133),

$$m^2(\gamma'^2 - \gamma^2) = (h/c^2)^2(v'^2 - v^2) = -(h/c^2)(v' + v)(\gamma' - \gamma), \tag{135}$$

or

$$(\gamma' - \gamma)[m^2(\gamma' + \gamma) + (h/c^2)(v' + v)] = 0. \tag{136}$$

The second factor is positive: therefore $\gamma' = \gamma$, and $v' = v$ by (133). In fact we have

$$v' = v, \quad \gamma' = \gamma, \quad v' = v. \tag{137}$$

Thus *the frequency of the photon and the speed of the electron are unchanged by the collision when we use the mass-centre reference system.* The frequency and speed are connected by (134), which gives

$$\frac{mvc}{\sqrt{(1 - v^2/c^2)}} = hv. \tag{138}$$

The snapshots of the collision are as in Fig. 13. The indeterminacy in the problem corresponds to the fact that the direction of motion of

[1] $h/(mc^2)$ is the period of de Broglie waves for an electron at rest; cf. SYNGE [1954*a*, p. 107].

the electron after the collision is arbitrary. Both before and after collision, the velocity of the photon is opposed to that of the electron [cf. (132)].

c) Time-axis chosen along $(M_r + M'_r)$.

If we take the time-axis along the vector $(M_r + M'_r)$, we have

$$M_\varrho + M'_\varrho = 0, \quad (139)$$

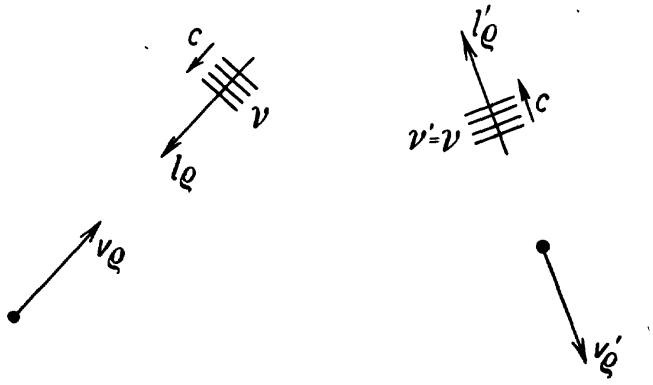


Fig. 13 – Snapshots of the Compton effect for the mass-centre reference system

and hence

$$v'_\varrho = -v_\varrho, \quad \gamma' = \gamma. \quad (140)$$

The conservation equations (41), (42) give

$$2m\gamma v_\varrho = (h\nu/c)\,(l'_\varrho - l_\varrho), \quad \nu' = \nu. \quad (141)$$

Thus the velocity of the electron is reversed, the frequency of the photon is unchanged, and by (141) we have

$$m(\gamma^2 - 1)^{\frac{1}{2}} = (h\nu/c^2) \sin \tfrac{1}{2}\theta, \quad (142)$$

where θ is the angle through which the photon is deflected (Fig. 14). It also follows from (141) that

$$v'_\varrho l'_\varrho = -v_\varrho l'_\varrho = v_\varrho l_\varrho, \quad (143)$$

which tells us that the angle between the directions of motion of the electron and the photon is unchanged by the collision.

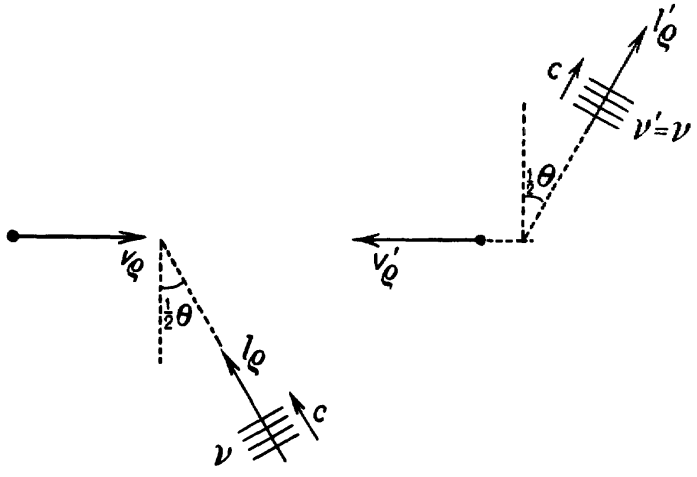


Fig. 14 – Snapshots of the Compton effect, with the time-axis along $(M_r + M'_r)$

d) Time-axis chosen along $(P_r + P'_r)$

If we take the time-axis along $(P_r + P'_r)$ we have

$$P_\varrho + P'_\varrho = 0. \quad (144)$$

From this it follows that the frequency of the photon is unchanged by the collision, and that its di-

rection of motion is reversed. We have $l'_\varrho = -l_\varrho$, and the conservation law (41), (42) gives

$$m\gamma' v'_\varrho = m\gamma v_\varrho + 2(h\nu/c)l_\varrho, \quad \gamma' = \gamma, \quad v' = v. \tag{145}$$

Thus the speed of the electron is unchanged by the collision. As for the change in the direction of the electron's motion, it follows from (145) that

$$v'_\varrho n_\varrho = v_\varrho n_\varrho, \tag{146}$$

where n_ϱ is any 3-vector perpendicular to l_ϱ so that the electron is, as it were, reflected in the plane of the photon's waves. Further (145) gives

$$m\gamma v \sin \tfrac{1}{2}\phi = h\nu/c, \tag{147}$$

where ϕ is the angle through which the electron is deflected (Fig. 15).

The several results worked out above show how differently the same phenomenon appears when viewed from different frames of reference. The great advantage of the space-time approach (Fig. 11) lies in its absolute character. Further, it suggests the various frames of reference for which the kinematical descriptions are particularly simple.

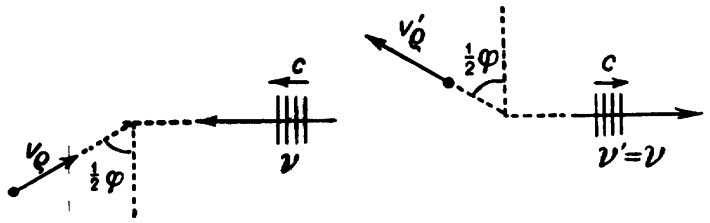


Fig. 15 – Snapshots of the Compton effect, with time-axis along $(P_r + P'_r)$

§ 19. THE ANNIHILATION AND CREATION OF MATTER

For a material particle, and for a photon also, we have assumed that the momentum 4-vector M_r points into the future. This means that

$$M_4/i > 0. \tag{148}$$

Therefore for any assembly of material particles and photons we have

$$\Sigma M_4/i > 0. \tag{149}$$

We may also express this by saying that the relative energy of each particle is positive (148), and that therefore the total relative energy in any assembly is also positive (149).

Thus, on the basis of the conservation law,

$$\Sigma M'_r = \Sigma M_r, \tag{150}$$

it is clearly impossible for any assembly consisting of material particles, or photons, or both, to disappear completely as a result of collisions. Nor is it possible for such an assembly to be created out of nothing.

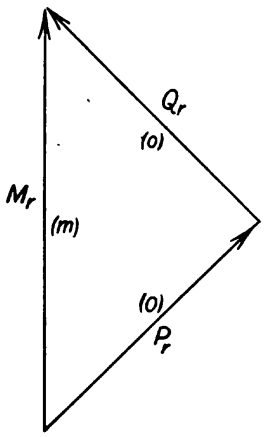


Fig. 16 – Space-time diagram of the annihilation of a single material particle, or its creation out of two photons

But while an assembly cannot disappear completely, it is possible for a set of material particles to turn into a set of photons, or vice versa. In other words, matter can be changed into radiation, and vice versa. This problem of the annihilation or creation of matter we shall discuss here on the basis of the conservation of 4-momentum, paying only slight attention to the conservation of electric charge and none at all to the conservation of spin.

Let us first treat the annihilation of a single material particle, i.e. its replacement by two photons. The space-time diagram is as in Fig. 16, which shows the 4-momentum M_r of the particle and the 4-momenta P_r, Q_r of two photons. This diagram applies also to the creation of a material particle out of two photons. The following equations are to be satisfied:

$$P_r + Q_r = M_r, \quad P_r P_r = 0, \quad Q_r Q_r = 0. \tag{151}$$

Let us use the frame of reference in which the particle is at rest. The equations become

$$P_\varrho + Q_\varrho = 0, \quad P_4 + Q_4 = im, \quad P_r P_r = 0, \quad Q_r Q_r = 0, \tag{152}$$

where m is the proper mass of the particle. If ν, ν' are the frequencies of the photons and l_ϱ, l'_ϱ the unit 3-vectors along their directions of motion, (152) gives

$$l_\varrho \nu + l'_\varrho \nu' = 0, \quad \nu + \nu' = mc^2/h, \tag{153}$$

and it follows at once that

$$\nu' = \nu, \quad l'_\varrho = -l_\varrho. \tag{154}$$

Thus in this frame of reference the photons have the same frequency

$$\nu = \tfrac{1}{2}mc^2/h, \tag{155}$$

and opposite directions of motion, one of the two directions being chosen arbitrarily. The snapshots of the annihilation are as in Fig. 17.

Any two photons (unless their frequency vectors have the same

direction in space-time) can combine to form a single material particle. If P_r, Q_r are their 4-momenta, then that of the material particle replacing them is

$$M_r = P_r + Q_r. \qquad (156)$$

It is easy to see that this vector is timelike and points into the future. For, since P_r and Q_r are future-pointing null vectors, we have by (156) and IV–(106)

$$M_r M_r = 2P_r Q_r < 0, \qquad (157)$$

and

$$M_4/i = P_4/i + Q_4/i > 0. \qquad (158)$$

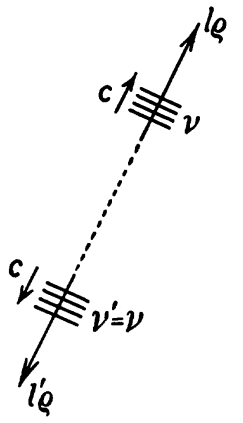


Fig. 17 – Snapshots of the annihilation of a single material particle

From (157) we can obtain the proper mass of the created particle. In terms of the elements ν, l_ϱ, ν', l'_ϱ of the two photons, we have

$$-P_r Q_r = \frac{h^2}{c^4}\nu\nu'(1 - l_\varrho l'_\varrho) = \frac{2h^2}{c^4}\nu\nu' \sin^2 \tfrac{1}{2}\theta, \qquad (159)$$

where θ is the angle between the directions of motion of the two photons. The proper mass m of the particle is given by $m^2 = -M_r M_r$, and we obtain from (157) and (159) the value

$$m = \frac{2h}{c^2}(\nu\nu')^{\frac{1}{2}} \sin \tfrac{1}{2}\theta. \qquad (160)$$

The right hand side here is a Lorentz-invariant of some importance in problems involving two photons; it vanishes if, and only if, the two photons have the same direction of motion.

The simplest frame from which to view the creation of a single particle is the mass-centre reference system for the two photons. This is at the same time the frame in which the created particle is at rest. The snapshots of the creation are then as in Fig. 17, if we interchange "before" and "after", and make the photons on the right approach one another instead of receding from one another.

From the annihilation of a single particle we can pass to the annihilation of any set of particles. Any set can be replaced by two photons, with satisfaction of the law of conservation; the essential equations are as in (151), with M_r replaced by ΣM_r, the total 4-momentum of the set of particles.

No assembly of particles can turn into a single photon, nor can a single photon turn into an assembly of particles. An assembly of photons cannot turn into a single photon, except in the case where all the frequency vectors have a common direction in space-time. In that special case, a set of photons can turn into a single photon, with a frequency equal to the sum of the frequencies of the several photons. This "addition of frequencies" may seem a little strange; but it is of course merely an addition of relative energies, and the addition of energies is something we are familiar with in physics.

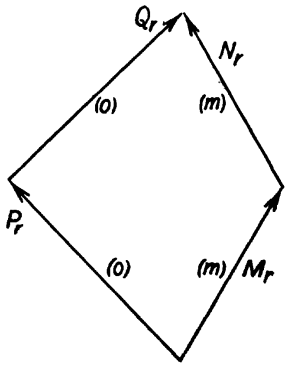


Fig. 18 – Space-time diagram of the annihilation or creation of a positron-electron pair.

The conservation of electric charge prohibits the annihilation of an electron. It also prohibits the annihilation of a positron (positively charged electron). But it does not prohibit the annihilation of an electron and a positron as a result of collision between the two. We shall now discuss this annihilation, and also the corresponding creation of a positron-electron pair out of a pair of photons.

The space-time diagram is as in Fig. 18. Here M_r, N_r are the 4-momenta of the positron and electron, and P_r, Q_r the 4-momenta of the two photons into which they change or out of which they are created. We have

$$M_r M_r = -m^2, \quad N_r N_r = -m^2, \quad P_r P_r = 0, \quad Q_r Q_r = 0, \qquad M_r + N_r = P_r + Q_r, \tag{161}$$

m being the common value of the proper masses of the positron and the electron. It is interesting to compare Fig. 18 with the corresponding diagram (Fig. 11) for the Compton effect. Both diagrams are composed of the same elements — two null vectors and two timelike vectors of M-length m, but the arrangements are different.

Let us discuss annihilation first. The simplest frame of reference is the mass-centre system for the positron and the electron. If we use this, it is easy to see that the two particles approach one another with the same speed v, and that the directions of the two photons are opposed, the common line of motion of the photons being arbitrary (that is the essential indeterminacy in the problem of annihilation).

The frequencies of the photons are equal:

$$\nu = mc^2\gamma/h, \quad \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}. \tag{162}$$

It is interesting, however, to see what happens if we use that frame of reference in which the electron is at rest. Let

$$v_e = \text{velocity of positron},$$

$$\nu, l_e, \nu', l_e' = \text{elements for the two photons.}$$

We regard v_e as known, and the other quantities as unknown. Then, by (41) and (42),

$$m\gamma v_e = \frac{h\nu}{c} l_e + \frac{h\nu'}{c} l_e', \tag{163}$$

$$m\gamma + m = \frac{h\nu}{c^2} + \frac{h\nu'}{c^2}. \tag{164}$$

From (163) we obtain

$$\frac{m^2\gamma^2 v^2}{c^2} = \frac{h^2}{c^4}(\nu^2 + \nu'^2 + 2\nu\nu' \cos\theta), \tag{165}$$

where θ is the angle between the directions of the photons. But $\gamma^2 v^2/c^2 = \gamma^2 - 1$, and so, with (164), we have

$$\begin{aligned} \gamma^2 - 1 &= \frac{h^2}{m^2c^4}(\nu^2 + \nu'^2 + 2\nu\nu' \cos\theta) \\ \gamma + 1 &= \frac{h}{mc^2}(\nu + \nu'). \end{aligned} \tag{166}$$

Here γ is known. We can choose θ arbitrarily [subject to a limitation later in (169)] and solve for ν and ν'. We get

$$(\gamma + 1)^2 - (\gamma^2 - 1) = 2(\gamma + 1) = \frac{h^2}{m^2c^4} 4\nu\nu' \sin^2 \tfrac{1}{2}\theta. \tag{167}$$

This gives us $\nu\nu'$, and so, by the second of (166),

$$\begin{aligned} (\nu - \nu')^2 &= (\nu + \nu')^2 - 4\nu\nu' \\ &= \frac{m^2c^4}{h^2}(\gamma + 1)^2 - \frac{2m^2c^4}{h^2}(\gamma + 1) \operatorname{cosec}^2 \tfrac{1}{2}\theta \\ &= \frac{m^2c^4}{h^2}(\gamma + 1)(\gamma - 1 - 2\cot^2 \tfrac{1}{2}\theta). \end{aligned} \tag{168}$$

We note that we must have

$$2\cot^2\tfrac{1}{2}\theta \leqslant \gamma - 1, \quad \tan\tfrac{1}{2}\theta \geqslant \sqrt{\frac{2}{\gamma - 1}}. \tag{169}$$

This gives a lower bound for the angle between the directions of motion of the two photons. As for their frequencies, we have by (168)

$$\nu - \nu' = \frac{mc^2}{h}(\gamma + 1)^{\frac{1}{2}}(\gamma - 1 - 2\cot^2\tfrac{1}{2}\theta)^{\frac{1}{2}}, \tag{170}$$

and so, by (166), the two frequencies are

$$\nu, \nu' = \frac{mc^2}{2h}[\gamma + 1 \pm (\gamma + 1)^{\frac{1}{2}}(\gamma - 1 - 2\cot^2\tfrac{1}{2}\theta)^{\frac{1}{2}}], \tag{171}$$

with θ chosen arbitrarily except for (169). We can then get (to within a rotation about the direction of v_ϱ) the directions of motion of the photons from (163), or equivalently

$$\begin{aligned} m\gamma v_\varrho l_\varrho &= \frac{h\nu}{c} + \frac{h\nu'}{c}\cos\theta, \\ m\gamma v_\varrho l'_\varrho &= \frac{h\nu}{c}\cos\theta + \frac{h\nu'}{c}. \end{aligned} \tag{172}$$

As a check we note that if $v = 0$ then $\gamma = 1$ and (169) gives $\theta = \pi$, so that the photons recede directly away from one another. Also (171) gives for the frequencies and wave lengths

$$\nu = \nu' = mc^2/h, \quad \lambda = \lambda' = h/mc; \tag{173}$$

this is the Compton wave length (p. 197).

We note that, by (171), the greatest and least frequencies which can occur when a positron, moving with speed v, strikes an electron at rest, correspond to $\theta = \pi$; the values are

$$\begin{aligned} \nu_{max} &= \frac{mc^2}{2h}[\gamma + 1 + \sqrt{(\gamma^2 - 1)}], \\ \nu_{min} &= \frac{mc^2}{2h}[\gamma + 1 - \sqrt{(\gamma^2 - 1)}]. \end{aligned} \tag{174}$$

As for the creation of matter from the collision of two photons, it is easy to see from Fig. 16, combined with the polygonal inequality (55), that two photons with 4-momenta P_r, Q_r can turn into any number of material particles, subject only to the condition that the sum of the

proper masses of the particles shall not exceed the magnitude of the vector $(P_r + Q_r)$. Now the square of this magnitude is, by (159),

$$-(P_r + Q_r)(P_r + Q_r) = -2P_rQ_r = \frac{4h^2}{c^4} \nu\nu' \sin^2 \tfrac{1}{2}\theta, \qquad (175)$$

θ being the angle between the directions of motion of the two photons and ν, ν' their frequencies, and so the restriction on the material particles formed by the collision of two photons is

$$\Sigma m \leqslant \frac{2h}{c^2} (\nu\nu')^{\frac{1}{2}} \sin \tfrac{1}{2}\theta. \qquad (176)$$

This is a statement valid for all Galileian frames of reference.

§ 20. ELASTIC COLLISIONS

Using the word *particle* in the wide sense, including both material particles and photons, we say that a collision between particles is *elastic* (or *perfectly elastic* if we want to be emphatic) if no particles are created or destroyed, and the proper mass of each material particle is conserved. (The proper mass of a photon, being always zero, is automatically conserved.)

We recognise the Compton effect (§ 18) as an elastic collision.

Consider now an elastic collision between two material particles with proper masses m and n. Let M_r, N_r be their 4-momenta before collision, so that

$$M_rM_r = -m^2, \quad N_rN_r = -n^2. \qquad (177)$$

If M'_r, N'_r are the 4-momenta after collision, then by the conservation law

$$M'_r + N'_r = M_r + N_r, \qquad (178)$$

and by the elastic condition

$$M'_rM'_r = -m^2, \quad N'_rN'_r = -n^2. \qquad (179)$$

The space-time diagram is as in Fig. 19.

We have in (178), (179) six equations for the eight unknowns M'_r, N'_r; there are therefore two degrees of freedom in the solution. In Newtonian mechanics (kinetic theory of gases) this indeterminacy may be removed by

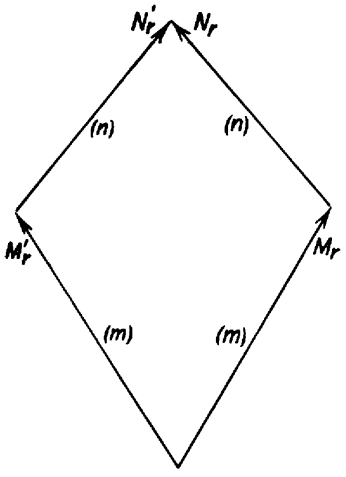


Fig. 19 – Space-time diagram of the elastic collision of two material particles

taking the particles to be small smooth spheres, and assigning the direction of the line of centres at the instant of impact. But any attempt to treat relativistic particles as of finite size is fraught with difficulty, and we are compelled to retain the indeterminacy here. (But see VII–§§ 9–13 for the case of particles with intrinsic angular momentum.)

If we use the mass-centre reference system, so that $M_\varrho + N_\varrho = 0$, we may write, instead of (178), (179),

$$\begin{aligned} m\gamma'_u u'_\varrho + n\gamma'_v v'_\varrho &= 0, \\ m\gamma'_u + n\gamma'_v &= m\gamma_u + n\gamma_v, \end{aligned} \tag{180}$$

where

$$\begin{aligned} u_\varrho,\ v_\varrho &= \text{3-velocities before collision}, \\ u'_\varrho,\ v'_\varrho &= \text{3-velocities after collision}, \\ \gamma_u = (1 - u^2/c^2)^{-\frac{1}{2}}, &\quad \gamma_v = (1 - v^2/c^2)^{-\frac{1}{2}}, \\ \gamma'_u = (1 - u'^2/c^2)^{-\frac{1}{2}}, &\quad \gamma'_v = (1 - v'^2/c^2)^{-\frac{1}{2}}. \end{aligned}$$

It is easy to see that (180) gives unique values to γ'_u and γ'_v. But these equations are satisfied in particular by $u'_\varrho = u_\varrho$, $v'_\varrho = v_\varrho$, and so we conclude that in general

$$\gamma'_u = \gamma_u, \quad \gamma'_v = \gamma_v, \tag{181}$$

and consequently

$$u' = u, \quad v' = v. \tag{182}$$

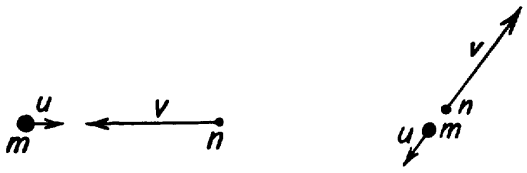


Fig. 20 – Snapshots of an elastic collision, using the mass-centre reference system

Thus the speeds of the particles are unchanged by the elastic collision, if we use the mass-centre reference system. The indeterminacy in the problem leaves the direction of the line joining the particles after collision completely arbitrary (Fig. 20).

By the first of (180), the speeds (before or after collision) are connected by

$$m^2(\gamma_u^2 - 1) = n^2(\gamma_v^2 - 1), \tag{183}$$

or

$$\frac{m^2u^2}{1 - u^2/c^2} = \frac{n^2v^2}{1 - v^2/c^2}, \tag{184}$$

or

$$\frac{u^{-2} - c^{-2}}{v^{-2} - c^{-2}} = \frac{m^2}{n^2}. \tag{185}$$

The more massive particle moves with the smaller speed. If we let m/n tend to infinity, we get either $u \to 0$ or $v \to c$.

In particular, the collision might leave the motions of both particles unchanged. This follows from the basic equations, for (178), (179) admit the solution $M'_r = M_r$, $N'_r = N_r$, on account of (177).

The inclusion of intrinsic angular momentum in VII–§§ 9–13 does not alter the relations obtained above, but it supplements them with others which reduce the indeterminacy from an infinite to a finite one.

CHAPTER VII

MECHANICS OF A DISCRETE SYSTEM

§ 1. DISCRETE AND CONTINUOUS SYSTEMS

In Newtonian mechanics one usually starts by considering the mechanics of a single particle, and then one passes on to a system of particles between which there act forces obeying Newton's law of action and reaction. Finally, one replaces the system of discrete particles by a continuous medium, as in hydrodynamics and elasticity. However, the discrete picture is not entirely abandoned; for example, in the kinetic theory of gases a discrete theory forms the basis of the macroscopic mechanics of a gas, the transition from the discrete to the continuous being carried out by statistical methods.

Is matter really discrete or continuous? As far as we are concerned, that question must be regarded as quite meaningless. For "continuous" is a mathematical word, not a physical word, and has only a very vague bearing on nature; we must not try to attach physical meanings to mathematical concepts which involve infinite processes — for example, we are not to spend time arguing whether the ratio of the length of the edge of a set-square to the length of a standard metre is rational or irrational. Such delicate questions have meaning in the mathematical model we make of nature, but none in nature itself.

We should not say "matter is composed of particles of no size" or "matter is continuous", if, in making such a statement, we are referring to nature. But we can, and shall, create discrete models and continuous models of natural bodies, passing from the former to the latter by a statistical argument. The significance of such models (and models of both types are useful) is that they provide us with theories which can be applied to nature in the way in which all theories have to be applied, namely, with a considerable amount of intuition and common sense.

The above remarks merely underline the philosophical attitude recommended in VI–§ 1. It is a theme that bears repetition.

§ 2. IMPULSES AND CONTINUOUS FORCES

Before dealing with a discrete system of particles, let us go back to the equations of motion of a material particle, as given in VI–(5):

$$\frac{d}{ds}\left(m\frac{dx_r}{ds}\right) = X_r, \tag{1}$$

where m is the proper mass (which need not be constant) and X_r is the 4-force. If we integrate between events P and Q in the history of the particle, we get the following expression for the change in 4-momentum:

$$\left[m\frac{dx_r}{ds}\right]_P^Q = \int_P^Q X_r\,ds. \tag{2}$$

Let us now suppose that $X_r = 0$ except in the interval PQ, and let us make this interval shorter and shorter, at the same time increasing X_r so that the limit

$$Y_r = \lim_{Q\to P}\int_P^Q X_r\,ds \tag{3}$$

exists. Then, in this limit, (1) and (2) give us a world line consisting of two straight parts with a change of direction at P, the instantaneous change in 4-momentum being

$$\Delta\left(m\frac{dx_r}{ds}\right) = Y_r. \tag{4}$$

We call Y_r an *impulse 4-vector* or *4-impulse.*

It is clear intuitively that it does not matter much whether we suppose a particle subject to a continuous 4-force or to a succession of small 4-impulses following one another at short intervals. In the former case, the world line is a smooth curve; in the latter, it consists of a great many short straight portions (Fig. 1).

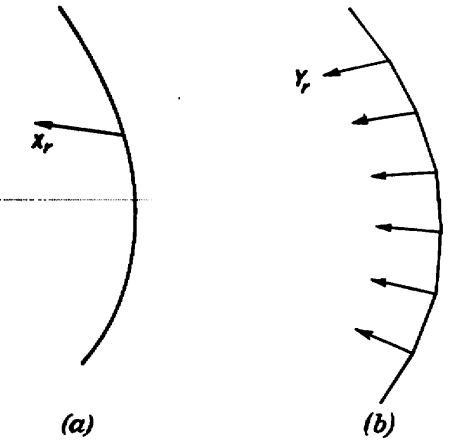


Fig. 1 – World line of a particle under (*a*) a continuous force, (*b*) a succession of small impulses

We saw on p. 167 that the proper mass of a particle remains constant if the 4-force is orthogonal to the world line of the particle. Let us examine what happens to proper mass when an impulse acts, as in (4).

Put $\lambda_r = dx_r/ds$, and let m, λ_r denote quantities before the impulse

and m', λ'_r quantities after the impulse. Then

$$m'\lambda'_r - m\lambda_r = Y_r, \tag{5}$$

and so, since $\lambda_r\lambda_r = \lambda'_r\lambda'_r = -1$,

$$\begin{aligned} m'^2 &= -(m'\lambda'_r)(m'\lambda'_r) = -(m\lambda_r + Y_r)(m\lambda_r + Y_r) \\ &= m^2 - 2m\lambda_r Y_r - Y_r Y_r. \end{aligned} \tag{6}$$

Thus if the proper mass is unchanged, the 4-impulse Y_r must satisfy the condition

$$2m\lambda_r Y_r + Y_r Y_r = 0. \tag{7}$$

If Y_r is small, this is approximately

$$\lambda_r Y_r = 0, \tag{8}$$

agreeing with the orthogonality condition for a continuous force.

§ 3. INTERNAL IMPULSES

Suppose we have two particles which, in some way, exert forces on one another. This is a familiar situation in Newtonian mechanics, but in relativity we encounter a great difficulty on account of the absence of absolute simultaneity, as already noted on p. 163. There is no widely accepted way of meeting this difficulty and consequently there is no widely accepted theory of the relativistic mechanics of a system of particles. However, the idea of a system of particles is such a natural one and so fundamental in Newtonian physics, that one is tempted to try to overcome the aforesaid difficulty and construct an appropriate relativistic mechanics, even at the price of introducing new physical entities for the express purpose of conveying momentum and energy from one particle to another. The reader should understand that the theory which follows is not one of those parts of relativity which are commonly accepted at the present time as adequate representations of nature, but rather a tentative scheme (SYNGE [1934]) from which at a later stage we shall pass by a statistical transition to firmer ground — the energy tensor of a continuous medium.

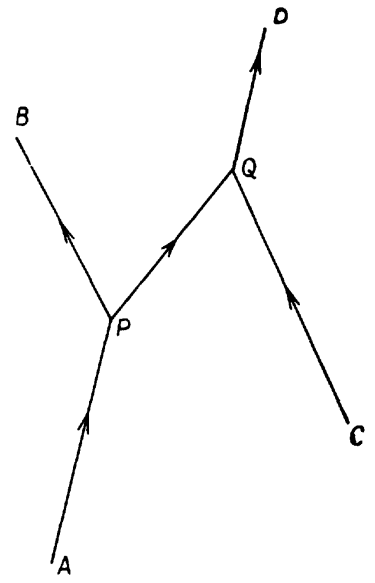


Fig. 2 – Space-time diagram of an internal impulse passing from P to Q

Let us consider two particles, proceeding in free motion with straight world lines AP, CQ (Fig. 2). At the event P some entity possessing momentum and energy (i.e. a momentum-energy 4-vector) leaves the first particle and travels in a straight line with the fundamental velocity c (i.e. its world line is a straight null line) to meet the other particle at the event Q, where it is absorbed. As a result of this occurrence, the world line of the first particle pursues a new direction PB and that of the second particle a new direction QD. The entity which passes from P to Q we call an *internal impulse.*

Let

$M_r^{(1)}$ = 4-momentum of first particle on AP,
$M_r^{(1)'}$ = 4-momentum of first particle on PB,
$M_r^{(2)}$ = 4-momentum of second particle on CQ,
$M_r^{(2)'}$ = 4-momentum of second particle on QD,
M_r = 4-momentum of internal impulse.

We shall assume that the conservation of 4-momentum holds at P and Q. Therefore

$$M_r^{(1)'} + M_r = M_r^{(1)}, \quad M_r^{(2)'} = M_r^{(2)} + M_r. \tag{9}$$

By addition we eliminate the internal impulse, obtaining

$$M_r^{(1)'} + M_r^{(2)'} = M_r^{(1)} + M_r^{(2)}. \tag{10}$$

Thus the internal impulse has fulfilled the purpose of transmitting 4-momentum from one particle to the other, with conservation of total 4-momentum.

We have now to face a difficulty, most easily explained by the following illustration. Let A and B be two boats, floating on a lake, each equipped with a gun. A fires bullets into B, and B fires bullets into A. As a result, the boats drift apart. For the recoil of A's gun forces him away from B, and A is further forced away from B by the momentum of B's bullets. An interchange of bullets produces a *repulsion*, never an *attraction.*

In the case of a material particle or a photon, the momentum 4-vector (by its definition) lies along the world line, pointing into the future; this is equivalent to saying that in each case the energy is positive. If we extend this property to internal impulses, so that M_r points from P to Q in Fig. 2, then it is easy to see that the internal impulses have the same effect as the bullets exchanged between the

two boats — they cause a repulsion, never an attraction. Hence they cannot explain the forces of cohesion in a system of particles.

To overcome this difficulty we now suppose that the internal impulse travelling from P to Q may have a momentum-energy 4-vector pointing in the direction QP (i.e. into the past). Following the pattern of VI–(22), we define relative 3-momentum and relative energy in terms of 4-momentum M_r as follows, these definitions holding for material particles, photons, and internal impulses:

$$\text{relative 3-momentum} = cM_\varrho,$$
$$\text{relative energy} = c^2M_4/i. \tag{11}$$

It follows then that for an *attractive* internal impulse *the 3-momentum points in the opposite direction to the direction of motion, and the relative energy is negative.*[1]

There is nothing objectionable in the idea of a 4-vector which points into the past instead of into the future. Indeed, when we consider electric current as a 4-vector [cf. X–(80)], a current of positive electricity gives a 4-vector pointing into the future, and a current of negative electricity gives a 4-vector pointing into the past. A localisation of negative energy in an internal impulse is something which, perhaps, we would rather not have; but we shall have to put up with it if we are to carry out our scheme.

In building up the relativistic mechanics of a system of particles we shall include internal impulses of both types, repulsive and attractive, and we shall assume that the conservation law of 4-momentum is obeyed at each emission and absorption.

In our space-time diagrams we have now to take care of two possible senses on a world line — the sense from past to future and the direction of 4-momentum. These are the same for material particles and photons, but may differ for internal impulses. To avoid possible confusion, we shall use a double arrow for the sense from past to future, and a single arrow for the direction of the 4-momentum. Fig. 3 shows on the left the exchange of a repulsive internal impulse, and on the right the

[1] An attractive impulse might be called a *negative photon*, since it differs mechanically from an ordinary photon only by reversal of the 4-momentum. Under the name of *gravitinos*, attractive impulses have been introduced into relativistic cosmology by PIRANI [1955]; in his theory there is spontaneous creation of material particles and gravitinos, the total 3-momentum and the total relative energy being zero before and after the creation.

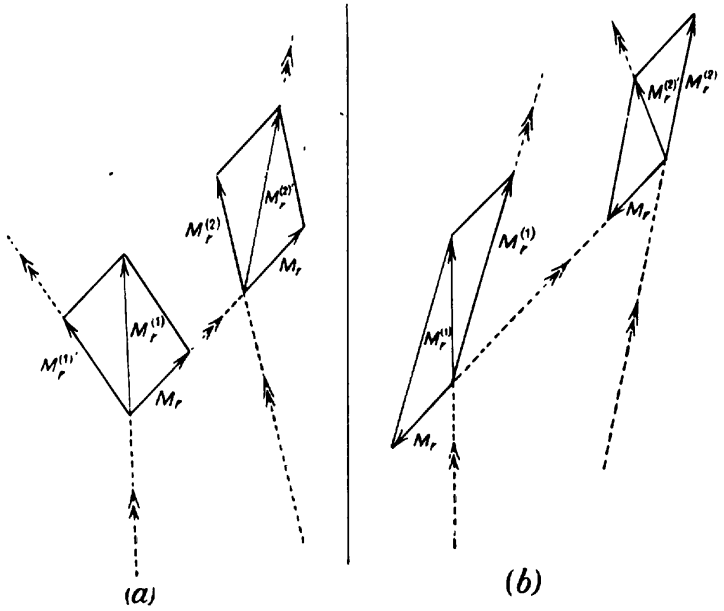


Fig. 3 – Space-time diagram of the exchange of an internal impulse between two particles: (*a*) repulsive impulse, (*b*) attractive impulse

exchange of an attractive internal impulse, complete with the parallelograms for the composition of 4-vectors according to the conservation law.

§ 4. THE CONSERVATION OF 4-MOMENTUM FOR A SYSTEM

Let us now think of a system consisting of any number of particles, with internal impulses passing between them. The corresponding space-time diagram would show a network of world lines, straight segments between their intersections, some of them being the world lines of particles and some the world lines of internal impulses.

If we feel uncomfortable about the internal impulses, introduced ad hoc for the purposes of the present theory, we can withdraw them and think of the system as composed of particles only, interacting with one another by direct elastic collision, as on p. 205. This is in fact a relativistic picture of the molecules of a gas, and all that will be said below is true for such a system.

If, on the other hand, we wish to make the interpretation of the argument as general as possible, we should include internal impulses; we should understand the word "particle" in the wider sense, including photons; we should suppose the particles to explode and to coalesce; we should admit the possibility of the simultaneous emission or absorption of several internal impulses.

The essential thing is that we have a relativistic theory (Lorentz-invariant since constructed in terms of Minkowskian geometry) involving straight world lines between collisions and the law of

conservation of 4-momentum [VI–(38)] at each collision. The conservation laws which we shall discuss are consequences of this, and it is not at all necessary to pin ourselves down with regard to the more physical details of the system under consideration.

Let us slice space-time across by a spacelike 3-space S. In particular we may take S to be the 3-flat $t = \text{const.}$ of any Galileian observer. At each event of intersection of S with a world line of the system, there is a 4-momentum M_r pertaining to the world line in question. We make the following definition:

Total 4-momentum of the system

$$\text{with respect to } S = \Sigma M_r, \tag{12}$$

the summation including all the world lines cut by S.

Now move S in space-time. As long as this motion does not carry S across a collision, ΣM_r remains constant because each individual M_r remains constant. And when it does carry S across a collision, ΣM_r is still unchanged, because each M_r not involved in the collision is unchanged, and the sum of those involved in the collision is unchanged by virtue of the conservation law VI–(38). Thus the right hand side of (12) does not depend on the choice of S, and we may rewrite our definition:

$$\text{Total 4-momentum of the system} = \Sigma M_r. \tag{13}$$

The fact that this is independent of S may be called the *open law of conservation of total 4-momentum.* We use the word "open" to distinguish this law from the "closed" law, discussed below.

There is one case in which (12) and (13) may fail, and that is when the system of particles extends out to infinity. Then (12) may cease to have meaning owing to the divergence of the infinite series ΣM_r. This is quite a real difficulty, for the simplest type of system to discuss is that which extends to infinity in all directions. We shall now obtain a different formulation of the law of conservation which avoids this difficulty of convergence.

Instead of taking S to be a spacelike 3-space, let us now take it to be a *closed* 3-space in space-time. Then S divides space-time into interior and exterior regions, and if a 4-vector is given at any event on S, we may say that it points *into* S or *out of* S, these abbreviations meaning of course *into the interior of* S and *out of the interior of* S, respectively. We need not trouble about the exceptional case where the

vector is tangent to S, since a small change in S will get rid of such an occurrence.

Consider now the intersection of S with the world lines of the system. On each world line there is a double arrow pointing into the future (indicating, we shall say, the *sense of propagation*), and a single arrow representing 4-momentum. All four possible combinations are shown in Fig. 4. For any one of the world lines the sense of propagation points either into S or out of S; equivalently, we may say that a world line *enters* S or *leaves* S.

We make the definition:

Flux of 4-momentum across S =

(sum of 4-momenta for world lines leaving S) — (sum of 4-momenta for world lines entering S). (14)

We shall now establish the following *closed law of conservation of 4-momentum: For any system of particles, interacting by means of internal impulses, with the law* VI–(38) *satisfied at each collision, the flux of 4-momentum across any closed 3-space is zero.*

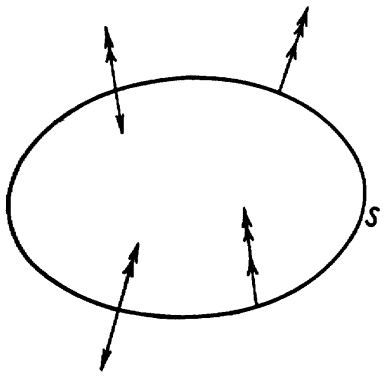


Fig. 4 – World lines cutting a closed 3-space S, with the senses of propagation indicated by double arrows and 4-momenta by single arrows

Let us start by considering the simplest of all systems, viz. a single particle or internal impulse which experiences no collisions at all, and for which consequently the 4-momentum is constant. To prove the above theorem for this simple case, all we have to show is that the world line enters S as often as it leaves it, and this is obviously true.

It is further obvious that the theorem is true for a system consisting of any number of particles and internal impulses, provided that no collisions occur inside S.

To prove the theorem in general, we fix our attention on all the world lines which enter S, and pretend that they experience no collisions. Then the flux of 4-momentum is zero. Now we introduce the collisions one by one. The introduction of a collision does not change the incoming flux, and it does not change the outgoing flux either, since the total 4-momentum resulting from a collision is the same as the total 4-momentum of the colliding entities. The theorem is proved.

This result is basic. It, or something equivalent, lies at the root of all treatments of continuous media, where however the basic fact may be accepted as a hypothesis rather than deduced from something simpler. The merit of the concept of internal impulses lies in the rational foundation which it gives to this law.

Note that S is *any* closed 3-space. It may be made small or large, unlike the previous spacelike S which had to be large enough to catch the world lines of all the particles. The two laws (the *open* law for spacelike S and the *closed* law for closed S) are linked together by the fact that ΣM_r in (12) may be regarded as a flux across the spacelike S, if the future side of S is taken as its exterior and the past side as its interior. When we pass to a continuous medium, the open law gives us an integral law of conservation (unless there should be divergence, as mentioned above), whereas the closed law gives us a set of partial differential equations when we make S shrink to zero. This will be treated later.

§ 5. ANGULAR MOMENTUM AND ITS CONSERVATION

Let x_r be an event and M_r a momentum-energy 4-vector associated with it. We define the corresponding *angular momentum* of this vector-at-an-event (M_r at x_r) to be

$$H_{rs} = x_r M_s - x_s M_r. \tag{15}$$

This is a skew-symmetric tensor, and so has six independent components:

$$\begin{aligned} &H_{23},\ H_{31},\ H_{12} \text{ real},\\ &H_{14},\ H_{24},\ H_{34} \text{ pure imaginary}. \end{aligned} \tag{16}$$

Such a skew-symmetric tensor is sometimes called a *six-vector* (cf. p. 224), but it is a term which may cause confusion; for H_{rs} is a tensor of the second order, not a vector.

The angular momentum tensor has the important property of being *constant along a straight world line between collisions*. This is true for a material particle, a photon, or an internal impulse. For, if x_r and x'_r are two events on the world line and M_r the corresponding 4-momentum, then for these two events we have

$$\begin{aligned} H'_{rs} &= x'_r M_s - x'_s M_r,\\ H_{rs} &= x_r M_s - x_s M_r, \end{aligned} \tag{17}$$

and so

$$H_{rs} - H'_{rs} = (x_r - x'_r)M_s - (x_s - x'_s)M_r. \tag{18}$$

But M_r lies on the world line, in the sense of propagation or opposed to it, and therefore

$$M_r = \theta(x_r - x'_r), \tag{19}$$

where θ is a factor of proportionality. Then (18) gives $H_{rs} = H'_{rs}$, and the result is established.

This constancy of H_{rs} is directly analogous to the constancy of the angular momentum of a free particle about a fixed point in Newtonian mechanics. But whereas Newtonian angular momentum may be regarded either as a skew-symmetric tensor in 3-space or as a vector, in space-time we have no such option; in space-time angular momentum is a tensor, not a vector.

The tensor H_{rs}, as given in (15), should be called the angular momentum of M_r at x_r *with respect to the origin*. The angular momentum of M_r at x_r *with respect to an event* a_r is defined as

$$H_{rs} = (x_r - a_r)M_s - (x_s - a_s)M_r. \tag{20}$$

It is clear that this is equal to the angular momentum of M_r at x_r with respect to the origin, less the angular momentum of M_r at a_r with respect to the origin.

We shall for simplicity take angular momenta with respect to the origin, unless otherwise stated.

Consider now a system of particles and internal impulses. For such a system we can obtain two laws of conservation of angular momentum, an open law and a closed law, as for 4-momentum. The reasoning by which these are established is essentially the same as that used in the case of 4-momentum, and it will suffice to state the definitions involved and the laws.

Let S be any spacelike 3-space. Then we define

Total angular momentum of the system

$$\text{with respect to } S = \Sigma H_{rs}, \tag{21}$$

the summation including the values of H_{rs} at all events where S cuts world lines of the system. With the exception of possible cases of divergence arising for infinite systems, this sum is independent of the choice of S, and so we may write

$$\text{Total angular momentum of the system} = \Sigma H_{rs}. \tag{22}$$

The fact that the total angular momentum is independent of the choice of S is the *open law of conservation of angular momentum.*

Now take any closed 3-space S, and make the definition:

Flux of angular momentum across S

= (sum of angular momenta for world lines leaving S)

− (sum of angular momenta for world lines entering S). (23)

The *closed law of conservation of angular momentum* then reads: *For any system of particles, interacting by means of internal impulses, with the law* VI–(38) *satisfied at each collision, the flux of angular momentum across any closed 3-space is zero.*

To pass from space-time to the kinematical point of view, we note that, according to the definition (15), the components of the angular momentum tensor have the following values for a particle of proper mass m at position x_ϱ at time t, moving with velocity v_ϱ:

$$H_{\varrho\sigma} = m\gamma(x_\varrho v_\sigma - x_\sigma v_\varrho)/c, \quad H_{\varrho 4} = im\gamma(x_\varrho - tv_\varrho). \tag{24}$$

We recognise in the 3×3 matrix $H_{\varrho\sigma}$ the relativistic generalisation of Newtonian angular momentum; except for the trivial factor c, it is the Newtonian angular momentum, with relative mass $m\gamma$ inserted for Newtonian mass. But we must not forget the other components $H_{\varrho 4}$; it is typical of the transition from Newtonian physics to relativity that a 3×3 matrix becomes a 4×4 matrix. Even if the particle is at rest, part of the angular momentum tensor survives; we have then

$$H_{\varrho\sigma} = 0, \quad H_{\varrho 4} = imx_\varrho. \tag{25}$$

For a photon of frequency ν, moving in a direction with direction cosines l_ϱ, we have

$$H_{\varrho\sigma} = h\nu(x_\varrho l_\sigma - x_\sigma l_\varrho)/c^2, \quad H_{\varrho 4} = ih\nu(x_\varrho - ctl_\varrho)/c^2. \tag{26}$$

§ 6. THE MASS-CENTRE OF A SYSTEM

For a system of particles, interacting by means of internal impulses, there exist, as we have seen, a constant 4-vector and a constant skew-symmetric tensor:

$$\begin{aligned} M_r &= \text{total 4-momentum}, \\ H_{rs} &= \text{total angular momentum with respect to the origin}. \end{aligned} \tag{27}$$

Here we have dropped the Σ sign used in § 5; M_r and H_{rs} now refer to the whole system, and not to individual entities in it. To avoid

questions of divergence, in the definitions of these quantities and in the proofs of their constancy, we shall suppose that the system is finite.

The angular momentum with respect to an arbitrary event a_r is, as we see from (20) on summing throughout the system,

$$H'_{rs} = H_{rs} - a_r M_s + a_s M_r. \tag{28}$$

We now ask whether it is possible, by suitable choice of a_r, to make

$$H'_{rs} M_s = 0. \tag{29}$$

This demand looks artificial, but it leads, as we shall see, to an interesting definition of the mass-centre of the system (SYNGE [1935a], MØLLER [1949]).

In view of (28), the equation (29) may be written

$$a_r M_s M_s - a_s M_r M_s = H_{rs} M_s. \tag{30}$$

Here we have four equations for the four quantities a_r. They are vector equations with respect to Lorentz transformations, and so we may use any Galileian frame of reference we please. If the system consisted only of material particles and photons, then M_r would certainly be timelike, and the presence of repulsive internal impulses would not alter this. On the other hand, attractive internal impulses (for which M_r points into the past) tend to destroy this timelike character. But they would have to be strong to do this, and we shall assume that M_r is timelike. Let us then choose the time-axis parallel to M_r, so that we have

$$M_\varrho = 0, \quad M_4 \neq 0. \tag{31}$$

Then the equations (30) become

$$a_\varrho M_4^2 = H_{\varrho 4} M_4, \quad a_4 M_4^2 - a_4 M_4^2 = 0. \tag{32}$$

The last is satisfied identically, and so the solution of (30) is, for this special frame of reference,

$$a_\varrho = H_{\varrho 4}/M_4, \quad a_4 \text{ arbitrary}. \tag{33}$$

Thus the requirement (29) gives us, not a single event a_r, but a straight L in space-time, parallel to M_r. This partial indeterminacy is indeed obvious from (28), for H'_{rs} is unchanged if we add θM_r to a_r, θ being an arbitrary factor.

Let us now define *mass-centre* and sum up what we have done as follows: *The history of the mass-centre of a system of particles and*

internal impulses is a straight line L in space-time such that

$$H'_{rs}M_s = 0, \tag{34}$$

where H'_{rs} is the total angular momentum of the system with respect to any event on L, and M_r is the total 4-momentum of the system.

The history of the mass-centre is uniquely determined by (34), and, in a frame of reference with time-axis parallel to M_r, we may say that the mass-centre is the fixed point with coordinates

$$\bar{x}_\varrho = H_{\varrho 4}/M_4, \tag{35}$$

where $H_{\varrho 4}$ is calculated with respect to the origin of the coordinates used. We may think of the mass-centre of a system as a hypothetical particle, having for world line the straight line L.

It may be observed that the *mass-centre reference system*, introduced in VI–§ 11, was an anticipation of (and consistent with) the more general concept of mass-centre now introduced. The situation in Ch. VI was simpler, because there we were concerned only with a single collision.

To make contact with Newtonian mechanics, let us consider a system consisting of material particles only. In a frame of reference for which (31) is true, we have

$$\Sigma m\gamma v_\varrho = 0, \tag{36}$$

the summation running over all the particles, and so, by (24) and (35), we have

$$\bar{x}_\varrho = \frac{H_{\varrho 4}}{M_4} = \frac{\Sigma m\gamma(x_\varrho - tv_\varrho)}{\Sigma m\gamma} = \frac{\Sigma m\gamma x_\varrho}{\Sigma m\gamma}. \tag{37}$$

This is the Newtonian formula for the mass-centre, changed only by the substitution of relative mass $m\gamma$ for Newtonian mass.

§ 7. INTRINSIC ANGULAR MOMENTUM OF A PARTICLE

In the Newtonian mechanics of a system, two vectors are of prime importance — the linear momentum and the angular momentum, the latter depending on the point with respect to which it is calculated. It is generally most convenient to use the mass-centre, and then attention is focussed on three things:

i) mass-centre,
ii) linear momentum,
iii) angular momentum with respect to the mass-centre.

If we make a Newtonian system shrink to zero size, we turn it into a particle, the position of the particle coinciding with the mass-centre and the linear momentum remaining a finite vector. In Newtonian mechanics we do not as a rule, although we could, associate a finite intrinsic angular momentum with a particle.

In relativistic mechanics, as developed above, a system possesses three important things:

i) a world line L representing the history of the mass-centre,
ii) a 4-momentum M_r, parallel to L,
iii) an angular momentum tensor H_{rs} with respect to any event on L, satisfying

$$H_{rs}M_s = 0. \tag{38}$$

This suggests the concept of a relativistic particle having

i) a world line L,
ii) a 4-momentum M_r tangent to L,
iii) a skew-symmetric *intrinsic* angular momentum tensor H_{rs} satisfying (38).

Our previous concept of a particle is thus generalised by attributing intrinsic angular momentum to it.

A particle at an event x_r, with 4-momentum M_r and intrinsic angular momentum H_{rs}, has, with respect to the origin, an angular momentum[1]

$$H'_{rs} = H_{rs} + x_r M_s - x_s M_r; \tag{39}$$

we add to the intrinsic angular momentum the angular momentum as given by (15).

It is easy to see that the conservation laws for angular momentum (§ 5) hold equally well if we endow the particles with intrinsic angular momentum, provided that the intrinsic angular momentum of each particle is constant between collisions, and the total intrinsic angular momentum is conserved at each collision.

In Newtonian mechanics linear momentum and angular momentum with respect to the mass-centre are dynamically independent, but in relativity 4-momentum and intrinsic angular momentum are connected by (38). If we accept the concept of a particle with intrinsic angular

[1] The part H_{rs} in (39) may be called the *spin* angular momentum, and the part $x_r M_s - x_s M_r$ the *orbital* angular momentum. A particle with spin is to be regarded as an independent concept, rather than as the limit of an extended system (cf. p. 316).

momentum, and think of a 4-force applied to the particle, its world line will be curved and M_r will change; to satisfy (38), H_{rs} must also change. To examine this change, we write

$$\frac{d}{ds}(H_{rs}M_s) = \frac{dH_{rs}}{ds}M_s + H_{rs}\frac{dM_s}{ds} = 0. \tag{40}$$

If X_r is the 4-force, we have the equations of motion VI–(4),

$$\frac{dM_s}{ds} = X_s, \tag{41}$$

and (40) gives

$$\frac{dH_{rs}}{ds}M_s = -H_{rs}X_s. \tag{42}$$

There appear to be four equations here, but actually there are only three, since multiplication by M_r gives the identity $0 = 0$ on account of the skew-symmetry of H_{rs} and (38). Should we wish to make the change in H_{rs} mathematically determinate, we would have to add further equations of motion. But this question we shall not pursue further (cf. FRENKEL [1926, p. 353], WEYSSENHOFF [1947], WEYSSENHOFF and RAABE [1947]).

A particular significance attaches to quantities which are invariant under Lorentz transformations. The 4-momentum M_r yields one such invariant:

$$M_rM_r = -m^2, \tag{43}$$

where m is the proper mass. Any skew-symmetric tensor possesses two Lorentz-invariants; for H_{rs} the first is

$$\Omega^2 = \tfrac{1}{2}H_{rs}H_{rs} = H_{23}^2 + H_{31}^2 + H_{12}^2 + H_{14}^2 + H_{24}^2 + H_{34}^2, \tag{44}$$

and the second is

$$-\tfrac{1}{4}iH_{rs}H_{rs}^* = H_{23}H_{14} + H_{31}H_{24} + H_{12}H_{34}. \tag{45}$$

Here H_{rs}^* is the *dual* of H_{rs}, defined by

$$H_{rs}^* = \tfrac{1}{2}\,i\,\varepsilon_{rsmn}H_{mn}, \tag{46}$$

ε_{rsmn} being the four-dimensional permutation symbol as in IV–(32):

$$\varepsilon_{rsmn}\begin{cases} = 1 \text{ if the set of numbers } (rsmn) \text{ is the set } (1234), \text{ in order, or in an even permutation of that order,} \\ = -1 \text{ if } (rsmn) \text{ is an odd permutation of } (1234), \\ = 0 \text{ otherwise.} \end{cases} \tag{47}$$

Thus

$$\begin{aligned} H^*_{23} &= iH_{14}, \quad H^*_{31} = iH_{24}, \quad H^*_{12} = iH_{34}, \\ H^*_{14} &= iH_{23}, \quad H^*_{24} = iH_{31}, \quad H^*_{34} = iH_{12}. \end{aligned} \tag{48}$$

Now it is easy to see by direct calculation that

$$\det H_{rs} = (H_{23}H_{14} + H_{31}H_{24} + H_{12}H_{34})^2, \tag{49}$$

in view of the skew-symmetry of H_{rs}. But (38) implies $\det H_{rs} = 0$.[1] *Thus the second invariant* (45) *vanishes, leaving* (44) *as the only one to consider.*

To explore this invariant (44), let us take the time-axis parallel to M_r; then $M_\varrho = 0$ and by (38)

$$H_{\varrho 4} = 0. \tag{50}$$

By (44) we have, for this special frame,

$$\Omega^2 = \tfrac{1}{2}H_{\varrho\sigma}H_{\varrho\sigma} = H^2_{23} + H^2_{31} + H^2_{12}. \tag{51}$$

These components are real [cf. p. 57], and so Ω^2 is not negative.

Returning to a general Galileian frame of reference, we see that a real non-negative invariant Ω is defined by

$$\Omega^2 = \tfrac{1}{2}H_{rs}H_{rs}, \quad \Omega \geqslant 0, \tag{52}$$

and we call Ω so defined the *intrinsic angular momentum invariant* of the particle. We might also call H_{rs} the *spin tensor* and Ω the *spin invariant.*

It follows from (51) that $\Omega = 0$ only if all the components of H_{rs} vanish.

§ 8. THE GEOMETRICAL REPRESENTATION OF A SKEW-SYMMETRIC TENSOR

The momentum-energy 4-vector M_r is, as we have seen, extremely easy to represent geometrically. We can think of M_r as a 4-vector in space-time, and sketch it in a space-time diagram, or we can think of a 3-vector M_ϱ in the observer's space and a single number M_4, such that

$$M_\varrho = m\gamma v_\varrho/c, \quad M_4 = im\gamma. \tag{53}$$

But how are we to represent the skew-symmetric angular momentum tensor H_{rs} geometrically?

It is better to ask a more general question: How are we to represent

[1] The rank of a skew-symmetric 4×4 matrix is 0, 2 or 4; H_{rs} is singular and therefore of rank 2, if it does not vanish completely.

any skew-symmetric tensor geometrically? For we recall that H_{rs} (the intrinsic angular momentum of a particle) is a rather special tensor, since it satisfies (38) and hence, as we have seen,

$$H_{rs}H^*_{rs} = 0. \tag{54}$$

The more general treatment is useful in connection with electromagnetic fields, where we have a skew-symmetric tensor not subject to any special condition.

Let C_{rs} be any skew-symmetric tensor ($C_{rs} = -C_{sr}$). It has a dual C^*_{rs} as in (46):

$$C^*_{rs} = \tfrac{1}{2}i\,\varepsilon_{rsmn}C_{mn}. \tag{55}$$

It is easy to see that $C_{\varrho 4}$ and $C^*_{\varrho 4}$ are 3-vectors with respect to orthogonal transformation of the space-axes, the time-axis being held fixed. But the six components of these two 3-vectors include all six components of C_{rs}, and so we get the very useful result: *any skew-symmetric tensor of the second order can be represented by a pair of 3-vectors.*

If we write C_ϱ, C^*_ϱ for the 3-vectors, we have

$$C_\varrho = iC_{\varrho 4}, \quad C^*_\varrho = -\,i\,C^*_{\varrho 4}, \tag{56}$$

the imaginary factors being inserted to make the 3-vectors real. The second of the above equations may be developed as follows:

$$C^*_\varrho = \tfrac{1}{2}\varepsilon_{\varrho 4mn}C_{mn} = \tfrac{1}{2}\varepsilon_{\varrho 4\mu\nu}C_{\mu\nu} = \tfrac{1}{2}\varepsilon_{\varrho\mu\nu}C_{\mu\nu}, \tag{57}$$

where $\varepsilon_{\varrho\mu\nu}$ is the 3-dimensional permutation symbol, defined as in (47) but for the range 1, 2, 3. Explicitly

$$\begin{aligned} &C_1 = iC_{14}, \quad C_2 = iC_{24}, \quad C_3 = iC_{34}, \\ &C^*_1 = C_{23}, \quad C^*_2 = C_{31}, \quad C^*_3 = C_{12}. \end{aligned} \tag{58}$$

We note that

$$C_{\varrho\sigma} = \varepsilon_{\varrho\sigma\tau}C^*_\tau. \tag{59}$$

If we write

$$C^2 = C_\varrho C_\varrho, \quad C^{*2} = C^*_\varrho C^*_\varrho, \tag{60}$$

so that C and C^* are the magnitudes of the two 3-vectors, then for the two invariants of C_{rs} we have

$$\tfrac{1}{2}C_{rs}C_{rs} = C^{*2} - C^2, \quad \tfrac{1}{4}C_{rs}C^*_{rs} = C_\varrho C^*_\varrho. \tag{61}$$

Thus the vanishing of the first invariant implies $C = C^*$, and the vanishing of the second implies that C_ϱ is orthogonal to C^*_ϱ. (Remember that Greek suffixes have the range 1, 2, 3 and that we are

dealing with ordinary 3-vectors, so that orthogonality has its Euclidean meaning.)

There is an interesting combination of these invariants which we may notice in passing:

$$I = \tfrac{1}{2}[(C_{rs}C_{rs})^2 + (C_{rs}C^*_{rs})^2]^{\frac{1}{2}}. \tag{62}$$

By (61) this may be written

$$I = (C^4 + C^{*4} + 2C^2C^{*2}\cos 2\theta)^{\frac{1}{2}}, \tag{63}$$

where θ is the angle between C_ϱ and C^*_ϱ. It is clear that I is always real, and vanishes if, and only if, $C = C^*$ and C_ϱ is orthogonal to C^*_ϱ.

Let us now apply these results to the intrinsic angular momentum H_{rs}, satisfying as in (38),

$$H_{rs}M_s = 0. \tag{64}$$

As in (58) we have two *3-vectors of angular momentum*

$$\begin{gathered} H_1 = iH_{14}, \quad H_2 = iH_{24}, \quad H_3 = iH_{34}, \quad (H_\varrho = iH_{\varrho 4}), \\ H^*_1 = H_{23}, \quad H^*_2 = H_{31}, \quad H^*_3 = H_{12}, \\ (H^*_\varrho = \tfrac{1}{2}\varepsilon_{\varrho\sigma\tau}H_{\sigma\tau}, \ H_{\varrho\sigma} = \varepsilon_{\varrho\sigma\tau}H^*_\tau). \end{gathered} \tag{65}$$

The angular momentum invariant Ω is given by

$$\Omega^2 = H^{*2} - H^2, \quad \Omega \geqslant 0, \tag{66}$$

where H^* and H are the magnitudes of the 3-vectors H^*_ϱ and H_ϱ, respectively. Since, as we have seen above, Ω^2 is non-negative, we have

$$H^* \geqslant H, \tag{67}$$

and since we know that the invariant (45) vanishes, we have [cf. (61)]

$$H_\varrho H^*_\varrho = 0. \tag{68}$$

By (53) and (64) we have

$$H_{\varrho\sigma}v_\sigma/c + iH_{\varrho 4} = 0, \tag{69}$$

or

$$H_\varrho = -\varepsilon_{\varrho\sigma\tau}v_\sigma H^*_\tau/c = c^{-1}(H^* \times v)_\varrho, \tag{70}$$

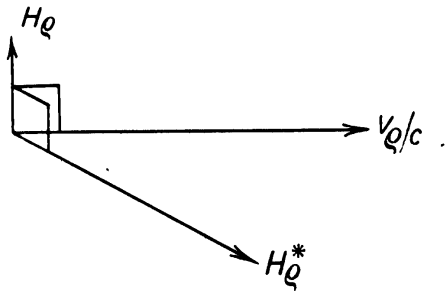


Fig. 5 – Snapshot illustrating relationship $H_\varrho = c^{-1}(H^* \times v)_\varrho$ between the velocity and the two intrinsic angular momentum 3-vectors

the sign $\times$ indicating the usual vector product. *The vector H_ϱ is, to within the factor c^{-1}, the vector product of H^*_ϱ and the velocity of the particle,* as illustrated in Fig. 5.

Here the frame of reference is general. If we take that frame which reduces the particle to rest (time-axis parallel to M_r), then $v_\varrho = 0$ and $H_\varrho = 0$ by (50). The picture as in Fig. 5 then shows nothing but the vector H^*_ϱ, which may have any direction in the observer's space.

Fig. 5 gives a spatial representation of intrinsic angular momentum. To get a space-time representation we may use null vectors (as we shall in IX–§ 7 for the electromagnetic tensor) or we may more simply use M_r and the 4-vector $-H^*_{rs}M_s$, which is orthogonal to M_r since

$$M_r H^*_{rs} M_s = 0 \tag{71}$$

on account of the skew-symmetry of H^*_{rs}. It is easy to see that the pair of 4-vectors $(M_r, -H^*_{rs}M_s)$ determine H_{rs}. For in the frame which makes $M_\varrho = 0$, the components of the second vector are, by (56),

$$-H^*_{\varrho s}M_s = -H^*_{\varrho 4}M_4 = mH^*_\varrho, \quad -H^*_{4s}M_s = 0, \tag{72}$$

and so H^*_ϱ is determined and hence H_{rs} because $H_\varrho = 0$. The pair of 4-vectors is shown in Fig. 6.

Another way of looking at the same thing is to introduce two unit vectors, λ_r along M_r (so that λ_r is the 4-velocity) and μ_r along $-H^*_{rs}M_s$. Then the intrinsic angular momentum is

$$H_{rs} = i\Omega\varepsilon_{rsmn}\lambda_m\mu_n. \tag{73}$$

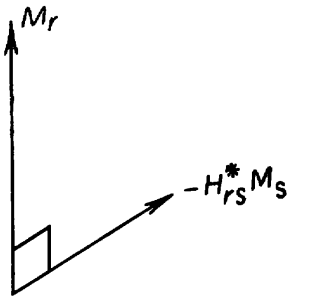


Fig. 6 – Space-time diagram of the 4-momentum M_r and the 4-vector $-H^*_{rs}M_s$ orthogonal to it

To prove this, we note that it is a tensorial relation, and therefore it is true in all Galileian frames if true in one, provided we do not change the orientation of the axes. Make $M_\varrho = 0$ and hence $\lambda_\varrho = 0$, $\lambda_4 = i$. Then (73) gives

$$\begin{aligned} H_{\varrho\sigma} &= -\Omega\varepsilon_{\varrho\sigma 4n}\mu_n = \Omega\varepsilon_{\varrho\sigma\nu}\mu_\nu, \\ H_{\varrho 4} &= \Omega\varepsilon_{\varrho 44n}\mu_n = 0. \end{aligned} \tag{74}$$

But by (66) we have, since $H_\varrho = 0$ when $M_\varrho = 0$,

$$\mu_\nu = H_\nu^*/\Omega, \tag{75}$$

and so (74) gives

$$H_{\varrho\sigma} = \varepsilon_{\varrho\sigma\nu} H_\nu^*, \tag{76}$$

the correct value, by (65). From this, and the second of (74), the equation (73) is established.

§ 9. ELASTIC COLLISIONS WITH UNCHANGED INTRINSIC ANGULAR MOMENTUM INVARIANTS. THE CASE OF IDENTICAL MATERIAL PARTICLES

On p. 205 we studied the elastic collision of two particles, the condition of elasticity being that the proper masses should be unchanged. Let us now consider the elastic collision of two particles possessing intrinsic angular momenta, with the additional elastic condition that the individual angular momentum invariants shall be unchanged by the collision.

In Newtonian mechanics the result of a collision between two smooth elastic spheres is mathematically determinate when we are given the velocities of the centres of the spheres and the direction of the line of centres at the instant of collision. It makes no difference whether the spheres are spinning or not; the angular velocities are unchanged by the collision. If we reduce the spheres to points with finite masses, the solution has a two-fold indeterminacy because the direction of the line of centres at the instant of collision ceases to have a meaning. As we shall see, it is otherwise in relativity. An elastic collision, with conservation of proper masses and intrinsic angular momentum invariants, has at most a finite number of results if we treat the particles as mathematical points. This is due to the connection (64) between 4-momentum and intrinsic angular momentum.

The particles occurring in nature are not, of course, mathematical points, and it is open to question whether a collision between two protons, for example, will have results similar to those obtained from the present theory. But that will not deter us from investigating what the results of this theory are. The essential thing is to note the hypotheses: the particles have no size, a collision means an intersection of world lines, and hence the conservation of the total angular momentum tensor for the two particles implies the conservation of the sum of their intrinsic angular momentum tensors, since we can take

moments about the event of collision and so reduce to zero the part of the total angular momentum tensor due to the 4-momenta.

The algebra of the problem is difficult, but it simplifies somewhat when the two particles are identical, i.e. when they have the same proper mass m and the same intrinsic angular momentum invariant Ω. We shall discuss that case here and return to the general case on p. 237.

Let

$$\left.\begin{matrix} M_r^{(0)}, & M_r'^{(0)} \\ M_r, & M_r' \end{matrix}\right\} = \text{4-momenta of particles} \left\{\begin{matrix} \text{before} \\ \text{after} \end{matrix}\right\} \text{collision},$$

$$\left.\begin{matrix} H_{rs}^{(0)}, & H_{rs}'^{(0)} \\ H_{rs}, & H_{rs}' \end{matrix}\right\} = \text{intrinsic angular momenta} \left\{\begin{matrix} \text{before} \\ \text{after} \end{matrix}\right\} \text{collision}.$$

By the definitions of proper mass and intrinsic angular momentum invariant, by (64), and by the condition of elastic collision, we have then

$$\begin{gathered} M_r^{(0)} M_r^{(0)} = M_r'^{(0)} M_r'^{(0)} = -m^2, \quad H_{rs}^{(0)} H_{rs}^{(0)} = H_{rs}'^{(0)} H_{rs}'^{(0)} = 2\Omega^2, \\ H_{rs}^{(0)} M_s^{(0)} = 0, \quad H_{rs}'^{(0)} M_s'^{(0)} = 0; \end{gathered} \tag{77}$$

$$\begin{gathered} M_r M_r = M_r' M_r' = -m^2, \quad H_{rs} H_{rs} = H_{rs}' H_{rs}' = 2\Omega^2, \\ H_{rs} M_s = 0, \quad H_{rs}' M_s' = 0. \end{gathered} \tag{78}$$

By the conservation of 4-momentum and angular momentum we have

$$M_r + M_r' = M_r^{(0)} + M_r'^{(0)}, \quad H_{rs} + H_{rs}' = H_{rs}^{(0)} + H_{rs}'^{(0)}. \tag{79}$$

To find the result of a collision, we are to assume values of $(M_r^{(0)}, M_r'^{(0)}, H_{rs}^{(0)}, H_{rs}'^{(0)})$ satisfying (77), and then solve the equations (78), (79) for $(M_r, M_r', H_{rs}, H_{rs}')$. On account of the identities $M_r H_{rs} M_s = 0$, $M_r' H_{rs}' M_s' = 0$, there are only six independent equations in the second line of (78), and so we count in all 20 equations for 20 unknowns. This suggests that the result of the collision is determined by the initial data; we have to examine how far this is true.

The equations (78), (79) possess an obvious solution:

$$M_r = M_r^{(0)}, \quad M_r' = M_r'^{(0)}, \quad H_{rs} = H_{rs}^{(0)}, \quad H_{rs}' = H_{rs}'^{(0)}. \tag{80}$$

In this solution the states of the particles are completely unchanged by the collision. There is a second solution

$$M_r = M_r'^{(0)}, \quad M_r' = M_r^{(0)}, \quad H_{rs} = H_{rs}'^{(0)}, \quad H_{rs}' = H_{rs}^{(0)}. \tag{81}$$

We may describe this by saying that the particles are interchanged.

An observer who could not distinguish the particles individually would say that nothing was changed. We shall call (80) and (81) *trivial solutions*. Thus we may say that *the set of equations* (78), (79) *always possess at least two solutions, but they may be trivial.*

It is easy to see that if we had taken two particles with different (m, Ω), the trivial solution (80) would still exist, but not (81).

Looking at the whole system (77), (78), (79), we note their symmetry with respect to initial and final quantities. It is clear that initial and final states are interchangeable — the elastic collision is a reversible process. This suggests a slightly different and more general way of looking at our problem.

Let P_r be any timelike 4-vector and Q_{rs} any skew-symmetric tensor. Write down the equations

$$\left.\begin{aligned} M_r M_r = M'_r M'_r = -m^2, \quad H_{rs}H_{rs} = H'_{rs}H'_{rs} = 2\Omega^2, \\ H_{rs}M_s = 0, \quad H'_{rs}M'_s = 0, \\ M_r + M'_r = P_r, \quad H_{rs} + H'_{rs} = Q_{rs}. \end{aligned}\right\} \tag{82}$$

If we regard P_r and Q_{rs} as assigned, we have here 20 equations for the 20 quantities $(M_r, M'_r, H_{rs}, H'_{rs})$. If there exist two solutions of (82), then the transition from one solution to the other may be regarded as the result of the collision of two particles characterised by the numbers (m, Ω), P_r and Q_{rs} being respectively the total 4-momentum and intrinsic angular momentum, the same before and after the collision.

Now if P_r and Q_{rs} are chosen arbitrarily, there is no reason to suppose that (82) possess any solution at all. But if they are the total 4-momentum and intrinsic angular momentum of a pair of particles characterised by (m, Ω), then, as in (80) and (81), we know that at least two solutions of (82) exist. We shall recall this later; for the present we proceed with the solution of (82), P_r and Q_{rs} being given, arbitrary except for the condition that P_r is timelike.

Choose the time-axis parallel to P_r, so that $P_\varrho = 0$. Then from the equations

$$M_r M_r = M'_r M'_r = -m^2, \quad M_r + M'_r = P_r, \tag{83}$$

we obtain

$$M'_\varrho = -M_\varrho, \quad M_4 = M'_4 = i(m^2 + M^2)^{\frac{1}{2}} = \tfrac{1}{2}P_4, \tag{84}$$

where

$$M^2 = M_\varrho M_\varrho. \tag{85}$$

Let us define a real positive quantity W by

$$W = -\tfrac{1}{2}iP_4; \tag{86}$$

its value is known since P_4 is known. Then we have

$$M'_\varrho = -M_\varrho, \quad M_4 = M'_4 = iW, \quad m^2 + M^2 = W^2. \tag{87}$$

Thus the eight unknowns M_r, M'_r have been reduced to three, viz. M_ϱ, and these last are connected by the last of (87).

We now have the following 15 equations for the 15 unknowns $(M_\varrho, H_{rs}, H'_{rs})$:

$$\left.\begin{gathered} H_{rs}H_{rs} = H'_{rs}H'_{rs} = 2\Omega^2, \\ H_{\varrho\sigma}M_\sigma + iH_{\varrho 4}W = -H'_{\varrho\sigma}M_\sigma + iH'_{\varrho 4}W = 0, \\ H_{rs} + H'_{rs} = Q_{rs}, \\ m^2 + M^2 = W^2. \end{gathered}\right\} \tag{88}$$

Introduce the 3-vectors H_ϱ, H^*_ϱ, H'_ϱ, H'^*_ϱ, Q_ϱ, Q^*_ϱ, defined in terms of the tensors as in (56). Using (65) and (66), write (88) in the form

$$\left.\begin{gathered} H^{*2} - H^2 = \Omega^2, \quad H'^{*2} - H'^2 = \Omega^2, \\ (M \times H^*)_\varrho + WH_\varrho = 0, \quad -(M \times H'^*)_\varrho + WH'_\varrho = 0, \\ H_\varrho + H'_\varrho = Q_\varrho, \quad H^*_\varrho + H'^*_\varrho = Q^*_\varrho, \\ m^2 + M^2 = W^2. \end{gathered}\right\} \tag{89}$$

In these vector products M stands for the vector M_ϱ.

We note that H_ϱ and H'_ϱ are given explicity in terms of the other unknowns by

$$H_\varrho = -W^{-1}(M \times H^*)_\varrho, \quad H'_\varrho = W^{-1}(M \times H'^*)_\varrho. \tag{90}$$

Thus we have reduced the system to a set of 9 equations for the 9 unknowns $(M_\varrho, H^*_\varrho, H'^*_\varrho)$:

$$\left.\begin{gathered} H^{*2} - W^{-2}(M \times H^*)_\varrho (M \times H^*)_\varrho = \Omega^2, \\ H'^{*2} - W^{-2}(M \times H'^*)_\varrho (M \times H'^*)_\varrho = \Omega^2, \\ (M \times H^*)_\varrho - (M \times H'^*)_\varrho = -WQ_\varrho, \\ H^*_\varrho + H'^*_\varrho = Q^*_\varrho, \\ m^2 + M^2 = W^2. \end{gathered}\right\} \tag{91}$$

Now for any pair of vectors X_ϱ, Y_ϱ, we have

$$(X \times Y)_\varrho (X \times Y)_\varrho = X^2Y^2 - (X_\varrho Y_\varrho)^2, \tag{92}$$

and so, making use of the last of (91), we write our set of 9 equations in the form

$$\left.\begin{aligned} m^2H^{*2} + (M_\varrho H^*_\varrho)^2 &= W^2\Omega^2, \\ m^2H'^{*2} + (M_\varrho H'^*_\varrho)^2 &= W^2\Omega^2, \\ (M \times H^*)_\varrho - (M \times H'^*)_\varrho &= -WQ_\varrho, \\ H^*_\varrho + H'^*_\varrho &= Q^*_\varrho, \\ m^2 + M^2 &= W^2. \end{aligned}\right\} \tag{93}$$

We have chosen the time-axis, but the space-axes are at our disposal. Choose Ox_3 in the direction of Q_ϱ and make the plane Ox_1x_3 contain Q^*_ϱ with $Q^*_1 \geqslant 0$ (Fig. 7); then we have, with $Q^2 = Q_\varrho Q_\varrho$,

$$Q_1 = Q_2 = 0, \quad Q_3 = Q, \quad Q^*_2 = 0. \tag{94}$$

The third of (93) then gives

$$\left.\begin{aligned} M_3 = 0, \quad H^*_3 - H'^*_3 &= 0, \\ M_1(H^*_2 - H'^*_2) & \\ - M_2(H^*_1 - H'^*_1) &= -WQ, \end{aligned}\right\} \tag{95}$$

and the fourth of (93) gives

$$\left.\begin{aligned} H^*_1 + H'^*_1 &= Q^*_1, \\ H^*_2 + H'^*_2 &= 0, \\ H^*_3 + H'^*_3 &= Q^*_3. \end{aligned}\right\} \tag{96}$$

We have then

$$H^*_3 = H'^*_3 = \tfrac{1}{2}Q^*_3, \tag{97}$$

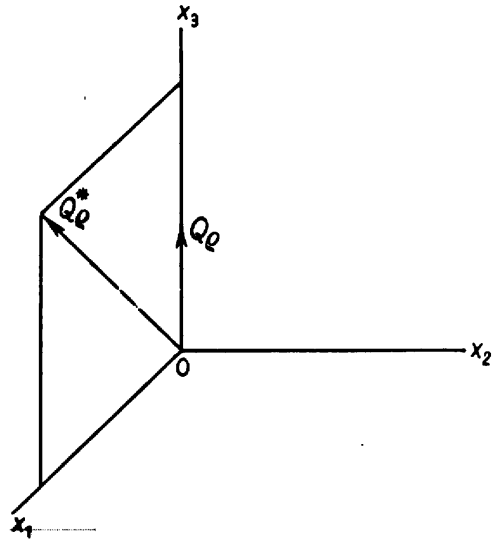


Fig. 7 – Special choice of space-axes

and (93) reduce to the following six equations for the six unknowns $(M_1, M_2, H^*_1, H^*_2, H'^*_1, H'^*_2)$:

$$\left.\begin{aligned} m^2(H^{*2}_1 + H^{*2}_2 + \tfrac{1}{4}Q^{*2}_3) + (M_1H^*_1 + M_2H^*_2)^2 &= W^2\Omega^2, \\ m^2(H'^{*2}_1 + H'^{*2}_2 + \tfrac{1}{4}Q^{*2}_3) + (M_1H'^*_1 + M_2H'^*_2)^2 &= W^2\Omega^2, \\ H^*_1 + H'^*_1 = Q^*_1, \quad H^*_2 + H'^*_2 &= 0, \\ M_1(H^*_2 - H'^*_2) - M_2(H^*_1 - H'^*_1) &= -WQ, \\ m^2 + M^2_1 + M^2_2 &= W^2. \end{aligned}\right\} \tag{98}$$

Put

$$H^*_1 - H'^*_1 = \alpha_1, \quad H^*_2 - H'^*_2 = \alpha_2. \tag{99}$$

Then

$$H_1^* = \tfrac{1}{2}(Q_1^* + \alpha_1), \quad H_2^* = \tfrac{1}{2}\alpha_2,$$
$$H_1'^* = \tfrac{1}{2}(Q_1^* - \alpha_1), \quad H_2'^* = -\tfrac{1}{2}\alpha_2, \tag{100}$$

and our problem reduces to solving for $(M_1, M_2, \alpha_1, \alpha_2)$ the four equations

$$\left.\begin{aligned} m^2[(Q_1^* + \alpha_1)^2 + \alpha_2^2 + Q_3^{*2}] \\ + [M_1 Q_1^* + (M_1\alpha_1 + M_2\alpha_2)]^2 = 4W^2\Omega^2, \\ m^2[(Q_1^* - \alpha_1)^2 + \alpha_2^2 + Q_3^{*2}] \\ + [M_1 Q_1^* - (M_1\alpha_1 + M_2\alpha_2)]^2 = 4W^2\Omega^2, \\ M_1\alpha_2 - M_2\alpha_1 = -WQ, \\ m^2 + M_1^2 + M_2^2 = W^2. \end{aligned}\right\} \tag{101}$$

Addition and subtraction of the first two equations gives, with

$$Q^{*2} = Q_\varrho^* Q_\varrho^* = Q_1^{*2} + Q_3^{*2},$$
$$m^2(Q^{*2} + \alpha_1^2 + \alpha_2^2) + M_1^2 Q_1^{*2} + (M_1\alpha_1 + M_2\alpha_2)^2 = 4W^2\Omega^2, \tag{102}$$
$$m^2 Q_1^*\alpha_1 + M_1 Q_1^*(M_1\alpha_1 + M_2\alpha_2) = 0. \tag{103}$$

Let us assume that $Q_1^* \neq 0$ ($Q_1^* = 0$ represents an exceptional case). Then we have, from (101) and (103), two equations for α_1 and α_2:

$$\left.\begin{aligned} (m^2 + M_1^2)\alpha_1 + M_1 M_2\alpha_2 = 0, \\ M_2\alpha_1 - M_1\alpha_2 = WQ, \end{aligned}\right\} \tag{104}$$

and so

$$\alpha_1 = \frac{M_2 Q}{W}, \qquad \alpha_2 = -\frac{(m^2 + M_1^2)Q}{M_1 W},$$
$$\alpha_1^2 + \alpha_2^2 = \frac{Q^2}{M_1^2 W^2}[(m^2 + M_1^2)^2 + M_1^2 M_2^2], \tag{105}$$
$$M_1\alpha_1 + M_2\alpha_2 = -\frac{m^2 Q M_2}{M_1 W}.$$

Substituting these values in (102) and transcribing the last of (101), we have for M_1 and M_2 the two equations

$$\left.\begin{aligned} m^2\left\{Q^{*2} + \frac{Q^2}{M_1^2 W^2}[(m^2 + M_1^2)^2 + M_1^2 M_2^2]\right\} \\ + M_1^2 Q_1^{*2} + \frac{m^4 Q^2 M_2^2}{M_1^2 W^2} = 4W^2\Omega^2, \\ M_1^2 + M_2^2 = W^2 - m^2. \end{aligned}\right\} \tag{106}$$

Elimination of M_2 gives the following equation of the fourth degree for M_1:

$$M_1^4 Q_1^{*2} + M_1^2[m^2(Q^{*2} + Q^2) - 4W^2\Omega^2] + m^4 Q^2 = 0. \quad (107)$$

When M_1 has been found to satisfy this equation, M_2 has the two values

$$M_2 = \pm (W^2 - m^2 - M_1^2)^{\frac{1}{2}}. \quad (108)$$

When M_1 and M_2 have been thus determined, all the other unknowns are uniquely given in terms of them by equations written above.

For arbitrarily assigned P_r and Q_{rs} we have no assurance that (107) has a real root; or, if it has a real root, that (108) gives a real M_2. But let us now pass to the case where P_r and Q_{rs} are respectively the total 4-momentum and intrinsic angular momentum of a pair of identical particles characterised by the numbers (m, Ω). Then we know that (107) has two real roots, equal and opposite, corresponding to the trivial solutions (80), (81). Therefore (107), regarded as a quadratic equation for M_1^2, has a real positive root, say

$$M_1^2 = p^2; \quad (109)$$

when this is substituted in (108) we get

$$M_2^2 = W^2 - m^2 - p^2, \quad (110)$$

and we know that this is non-negative, on account of the existence of the trivial solutions. Thus we get from (109) the *four* solutions

$$M_1 = \pm p, \quad M_2 = \pm (W^2 - m^2 - p^2)^{\frac{1}{2}}, \quad (111)$$

with any choice of signs. The two trivial solutions are of the form

$$M_1 = \varepsilon p, \quad M_2 = \varepsilon\eta(W^2 - m^2 - p^2)^{\frac{1}{2}}, \quad (112)$$

where η is fixed with one of the values $+1$, -1, and ε takes these two values in turn. Then the non-trivial solutions are

$$M_1 = \varepsilon p, \quad M_2 = -\varepsilon\eta(W^2 - m^2 - p^2)^{\frac{1}{2}}, \quad (113)$$

η being as in (112) and ε taking both signs. Thus there are in general two non-trivial solutions; they can hardly be regarded as distinct, since one follows from the other by interchange of the particles.

It is hard, in such an intricate algebraic problem, to take account of all exceptional cases. For example, we might get $M_2 = 0$ in (110), and then the two non-trivial solutions coincide with the trivial ones.

In addition to the trivial and non-trivial solutions arising from (109), there may be other non-trivial solutions. For since (107) has one real

root for M_1^2, it has a second real root, and it is clear from the coefficients in (107) that it also is positive, say

$$M_1^2 = q^2 = m^4Q^2/(Q_1^{*2}p^2). \tag{114}$$

If (108) then gives a real pair of values to M_2, we have here four non-trivial solutions; but we have no assurance that (108) will give real values. We can at any rate assert that, if we omit exceptional cases, there cannot be more than eight solutions of (78), (79), these eight solutions containing the two trivial ones (80), (81). The interesting fact is that the inclusion of angular momentum introduces *determinacy* into the problem of elastic collisions.

The obstacle to a complete treatment lies in the complexity of the algebra involved. One needs to study particular cases, and one will be worked out in the next section. However, one can simplify the situation by assuming that W/m is large.

This means that the particles are approaching one another very fast in the mass-centre system, since

$$W = m\gamma = m(1 - v^2/c^2)^{-\frac{1}{2}}, \tag{115}$$

where v is the speed of either particle before collision, or after collision (the speed is unchanged by the collision). Then (107) gives, approximately, the following two positive values for M_1^2:

$$M_1^2 = \frac{4W^2\Omega^2}{Q_1^{*2}}, \qquad M_1^2 = \frac{m^4Q^2}{4W^2\Omega^2}, \tag{116}$$

the first being large and the second small. The corresponding values of M_2^2 are approximately, by (108),

$$M_2^2 = W^2\left(1 - \frac{4\Omega^2}{Q_1^{*2}}\right), \qquad M_2^2 = W^2. \tag{117}$$

The first may be negative, making M_2 imaginary; the second is always positive.

Let us summarise briefly what we have found out about the elastic collision of identical particles with proper mass m and angular momentum invariant Ω. Omitting exceptional cases [e.g. $Q_1^* = 0$ in (103)], there can be only a small number of different results of the collision, the initial 4-momenta and intrinsic angular momenta being given. In addition to the trivial results, in which the states of the particles are unaltered or interchanged, there are in general two determinate final states, not essentially distinct from one another;

these are obtained from (107), from which we take that value of M_1^2 which gives the trivial results. But there may also be four more determinate final states, corresponding to the other root M_1^2 of (107). In the case of high energies, the key equations are (116) and (117).

§ 10. EXAMPLE OF AN ELASTIC COLLISION WITH INTRINSIC ANGULAR MOMENTUM INVARIANTS UNCHANGED

Fig. 8 shows two identical particles about to collide. The proper mass of each is m and they are referred to their mass-centre reference system. They are moving on the x_1-axis, and their velocities are

$$\left.\begin{aligned} u_1 &= -v, \quad u_2 = 0, \quad u_3 = 0, \\ v_1 &= v, \quad\quad v_2 = 0, \quad v_3 = 0. \end{aligned}\right\} \tag{118}$$

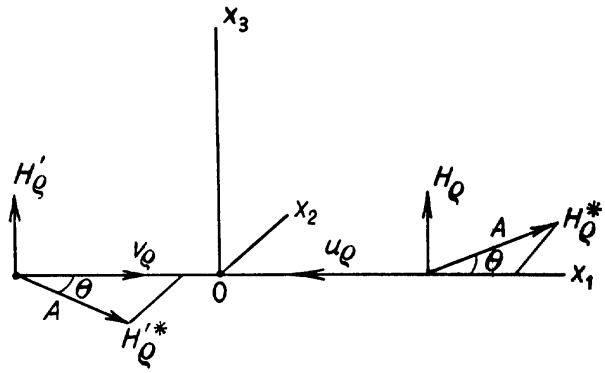


Fig. 8 – Snapshot of two particles with intrinsic angular momenta about to collide

For the total 4-momentum P_r we have then, with W defined as in (86),

$$\left.\begin{aligned} P_\varrho &= 0, \quad P_4 = 2im\gamma, \\ W &= m\gamma = m(1 - v^2/c^2)^{-\frac{1}{2}}. \end{aligned}\right\} \tag{119}$$

In order to get a situation which is not too complicated algebraically, we take the intrinsic angular momentum 3-vectors as shown. The magnitudes of H^*_ϱ and H'^*_ϱ have the common value A, and these vectors lie in the plane Ox_1x_2, making the same angle θ with Ox_1, but lying on opposite sides of this axis. Then, by Fig. 5, the vectors H_ϱ and H'_ϱ are both parallel to Ox_3 and have the same magnitude, $A(v/c)\sin\theta$. Each particle has the same angular momentum invariant Ω, given by [cf. (89)]

$$\Omega^2 = H^{*2} - H^2 = A^2[1 - (v/c)^2 \sin^2\theta]. \tag{120}$$

If we regard Ω as assigned, we are to think of A and θ as chosen to fit this equation.

For this system of two particles we have, then, as vectors of total intrinsic angular momentum [cf. (89)],

$$\left.\begin{aligned} &Q_1 = 0, \quad Q_2 = 0, \quad Q_3 = 2A(v/c)\sin\theta, \quad (Q = 2A(v/c)\sin\theta), \\ &Q_1^* = 2A\cos\theta, \quad Q_2^* = 0, \quad Q_3^* = 0, \quad (Q^* = 2A\,|\cos\theta\,|). \end{aligned}\right\} \tag{121}$$

These agree with (94), and so we are in fact using the space-axes

prescribed as in Fig. 7. In addition to the conditions of (94), we have the equation $Q_3^* = 0$; this is an accidental result, due to the simplicity of the case under consideration.

We are now ready to use the basic equation (107), which reads, after division by $4A^2$,

$$M_1^4 \cos^2\theta + M_1^2 m^2 \left[\cos^2\theta + \frac{v^2}{c^2}\sin^2\theta - \gamma^2\left(1 - \frac{v^2}{c^2}\sin^2\theta\right)\right] + m^4 \frac{v^2}{c^2}\sin^2\theta = 0, \quad (122)$$

or, if we multiply by $\gamma^2 \sec^2\theta$ and remember that $\gamma^2 v^2/c^2 = \gamma^2 - 1$,

$$M_1^4\gamma^2 - M_1^2 m^2(\gamma^4 - \gamma^2 + \tan^2\theta) + m^4(\gamma^2 - 1)\tan^2\theta = 0. \quad (123)$$

This has the roots $M_1 = \pm mp, \pm mq$, where

$$p = (\gamma^2 - 1)^{\frac{1}{2}}, \quad q = \gamma^{-1}\tan\theta. \quad (124)$$

For M_2 we have, as in (108),

$$M_2^2 = W^2 - m^2 - M_1^2, \quad (125)$$

and so

$$M_2 = 0 \text{ for } M_1 = \pm mp, \quad (126)$$

$$M_2 = \pm m(\gamma^2 - 1 - \gamma^{-2}\tan^2\theta)^{\frac{1}{2}} \text{ for } M_1 = \pm mq. \quad (127)$$

If we write v_ϱ' for the velocity of one of the particles after the collision, (126) gives

$$m\gamma' v_1'/c = \pm m(\gamma^2 - 1)^{\frac{1}{2}}, \quad v_2' = 0, \quad v_3' = 0, \quad (128)$$

which leads to

$$v_1' = \pm v_1, \quad v_2' = 0, \quad v_3' = 0. \quad (129)$$

These are the trivial solutions, corresponding to no change at all and to interchange of the particles. Since $M_2 = 0$ in (126), the expected non-trivial solutions are absent here.

But if the values of γ and θ are such as to satisfy

$$\gamma^2(\gamma^2 - 1) > \tan^2\theta, \quad (130)$$

we get a non-trivial result from (127), viz.

$$\left.\begin{aligned} m\gamma' v_1'/c &= \pm m\tan\theta/\gamma, \\ m\gamma' v_2'/c &= \pm m(\gamma^2 - 1 - \gamma^{-2}\tan^2\theta)^{\frac{1}{2}}, \\ v_3' &= 0, \end{aligned}\right\} \quad (131)$$

or, since $\gamma' = \gamma$,

$$\left.\begin{aligned} v_1' &= \pm c \tan\theta/\gamma^2, \\ v_2' &= \pm c(1 - \gamma^{-2} - \gamma^{-4}\tan^2\theta)^{\frac{1}{2}}, \\ v_3' &= 0. \end{aligned}\right\} \tag{132}$$

The arbitrary signs are independent, and so we get four non-trivial results here, provided (130) is satisfied. But these four results are to be regarded as two only, if we allow for interchange of the particles. The directions of rebound lie on the lines AOA and BOB in Fig. 9, these lines lying in the plane Ox_1x_2 and making with the x_1-axis the angle ϕ given by

$$\tan^2\phi = \gamma^2(\gamma^2 - 1)\cot^2\theta - 1. \tag{133}$$

It is interesting to note that the value of Ω plays no part in determining the directions of rebound, since it disappeared, with A, in (122).

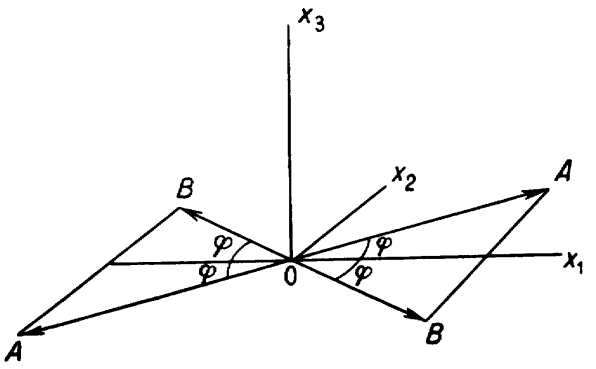


Fig. 9 – Directions of rebound of particles

In the case of a high-energy collision, γ is large, and then (130) is satisfied by any θ not too close to $\frac{1}{2}\pi$. If we fix θ and let γ go to infinity (so that the speed of the colliding particles tends to c), then (133) gives $\phi = \frac{1}{2}\pi$ in the limit, so that the particles rebound at right angles to their line of approach.

This example serves to illustrate the method of § 9; other particular examples might be worked out, and in particular it would be interesting to find the statistical distribution of the directions of rebound, taking as equally probable all directions of the intrinsic angular momentum 3-vectors of the two particles before collision, when referred to those frames of reference in which the particles are then individually at rest. We now proceed to a more general treatment of the problem of collision.

§ 11. GENERAL TREATMENT OF ELASTIC COLLISION WITH INTRINSIC ANGULAR MOMENTUM

We return to the general problem raised at the beginning of § 9, the elastic collision of two particles with conservation of 4-momentum

and angular momentum, with proper masses and intrinsic angular momentum invariants also conserved. The particles may be material particles or photons, it being assumed that for a material particle the proper mass and intrinsic angular momentum invariant are both different from zero, and that for a photon both these vanish. We might also include internal impulses of the repulsive type, regarding them as mechanically the same as photons, but those of the attractive type (with 4-momentum pointing into the past) would cause complications in the argument, and we shall not consider them.

Let a collision take place between two particles with proper masses m, m' and intrinsic angular momentum invariants Ω, Ω', these four quantities being unchanged by the collision. As in § 9, we denote by M_r, M'_r the 4-momenta after collision and by H_{rs}, H'_{rs} the intrinsic angular momenta after collision. Then if P_r is the total 4-momentum before (and after) collision, and Q_{rs} the total intrinsic angular momentum before (and after) collision, we have the following 20 equations to determine the 20 quantities M_r, M'_r, H_{rs}, H'_{rs}:

$$\left.\begin{aligned} &M_rM_r = -m^2, && M'_rM'_r = -m'^2,\\ &H_{rs}H_{rs} = 2\Omega^2, && H'_{rs}H'_{rs} = 2\Omega'^2,\\ &H_{rs}M_s = 0, && H'_{rs}M'_s = 0,\\ &M_r + M'_r = P_r, && H_{rs} + H'_{rs} = Q_{rs}. \end{aligned}\right\} \tag{134}$$

The right hand sides of these equations are given quantities.

In § 9 we used special axes in space-time, but we shall now pursue a different plan in order to get more symmetry in the algebra. Define R_r, S_{rs} by

$$R_r = M_r - M'_r, \quad S_{rs} = H_{rs} - H'_{rs}, \tag{135}$$

so that

$$\begin{aligned} M_r &= \tfrac{1}{2}(P_r + R_r), & H_{rs} &= \tfrac{1}{2}(Q_{rs} + S_{rs}),\\ M'_r &= \tfrac{1}{2}(P_r - R_r), & H'_{rs} &= \tfrac{1}{2}(Q_{rs} - S_{rs}). \end{aligned} \tag{136}$$

Our problem is thus immediately reduced to finding the 10 quantities R_r, S_{rs}, and for them we get 10 equations by substituting (136) in the first three lines of (134). We shall use the notation

$$U_rV_r = (UV), \quad X_{rs}Y_{rs} = (XY) \tag{137}$$

for any 4-vectors U_r, V_r and any skew-symmetric tensors X_{rs}, Y_{rs}.

Then (134) gives

$$\begin{aligned}(PP) + (RR) &= -2(m^2 + m'^2),\\ (PR) &= -(m^2 - m'^2),\\ (QQ) + (SS) &= 4(\Omega^2 + \Omega'^2),\\ (QS) &= 2(\Omega^2 - \Omega'^2),\end{aligned} \tag{138}$$

and also

$$Q_{rs}P_s + S_{rs}R_s = 0, \quad Q_{rs}R_s + S_{rs}P_s = 0. \tag{139}$$

There appear to be eight equations in (139), but actually there are only six independent equations, since two identities follow when we multiply by P_r and R_r. Thus *the solution of the collision problem involves the solution of the* 10 *independent equations in* (138) *and* (139) *for the* 10 *unknowns* (R_r, S_{rs}), *the right hand sides of* (138) *being given and also the quantities* (P_r, Q_{rs}).

We know that this system of equations does possess at least one solution corresponding to the state of affairs before the collision takes place. We shall call that the *trivial solution.*

The essential steps in the argument which follows are due to G. H. F. Gardner.

Let us begin by writing down some formulae for skew-symmetric tensors (RUSE [1936]). We recall the definition (47) of the permutation symbol ε_{rsmn} and note the identity

$$\varepsilon_{rsmn}\varepsilon_{rspq} = 2(\delta_{mp}\delta_{nq} - \delta_{mq}\delta_{np}), \tag{140}$$

which is easy to verify. Any skew-symmetric tensor X_{rs} and its dual X^*_{rs} are related by

$$\begin{aligned}&X^*_{rs} = \tfrac{1}{2}i\varepsilon_{rsmn}X_{mn}, \quad X_{rs} = -\tfrac{1}{2}i\varepsilon_{rsmn}X^*_{mn},\\ &X^*_{23} = iX_{14}, \quad X^*_{31} = iX_{24}, \quad X^*_{12} = iX_{34},\\ &X^*_{14} = iX_{23}, \quad X^*_{24} = iX_{31}, \quad X^*_{34} = iX_{12}.\end{aligned} \tag{141}$$

It is easy to verify directly (by taking first $s = t = 1$ and then $s = 1$, $t = 2$) that, for any pair of skew-symmetric tensors X_{rs}, Y_{rs},

$$\begin{aligned}X_{rs}Y^*_{rt} + X^*_{rt}Y_{rs} &= \tfrac{1}{2}\delta_{st}(XY^*)\\ &= \tfrac{1}{2}\delta_{st}(X^*Y) = X_{rt}Y^*_{rs} + X^*_{rs}Y_{rt}.\end{aligned} \tag{142}$$

In particular, if we put $Y_{rs} = X_{rs}$, we have the very useful formula

$$X_{rs}X^*_{rt} = \tfrac{1}{4}\delta_{st}(XX^*). \tag{143}$$

We note that

$$\det X_{rs} = (X_{23}X_{14} + X_{31}X_{24} + X_{12}X_{34})^2 = -\tfrac{1}{16}(XX^*)^2, \tag{144}$$

so that the condition that a skew-symmetric matrix be *singular* is

$$(XX^*) = 0. \tag{145}$$

Hence, by (143), a singular skew-symmetric tensor satisfies

$$X_{rs}X^*_{rt} = 0. \tag{146}$$

We now turn to (139) and multiply the two equations by S^*_{rt} and Q^*_{rt} respectively; this gives, by (143),

$$\begin{aligned} S^*_{rt}Q_{rs}P_s + \tfrac{1}{4}(SS^*)R_t &= 0, \\ \tfrac{1}{4}(QQ^*)R_t + Q^*_{rt}S_{rs}P_s &= 0. \end{aligned} \tag{147}$$

We now add these equations and use (142), obtaining

$$(SQ^*)P_t + \tfrac{1}{2}[(SS^*) + (QQ^*)]R_t = 0. \tag{148}$$

In order to make a certain deduction from this equation we have to exclude a type of collision which is quite uninteresting, viz. one in which the two particles have the same velocity before impact (equivalently, their 4-momenta M_r, M'_r have the same direction in space-time). It is not hard to see that the result of such a collision is trivial (nothing is changed). Now if P_r and R_r have the same direction after collision, then M_r and M'_r have the same direction after collision, and it follows from this that they had the same direction before collision. Thus, if we exclude (as we shall) the case where the particles have the same velocity before collision, it follows that P_r and R_r have different directions, and so (148) implies

$$(SQ^*) = 0, \quad (SS^*) + (QQ^*) = 0; \tag{149}$$

we shall need the second of these later.

We now introduce the known vector

$$B_r = Q_{rs}P_s, \tag{150}$$

which is obviously orthogonal to P_r, so that

$$(BP) = 0. \tag{151}$$

Writing the first of (139) in the form

$$S_{rs}R_s = -B_r, \tag{152}$$

we see that B_r is orthogonal to R_r:

$$(BR) = 0. \tag{153}$$

We now multiply (152) by S^*_{rt} and use (143) and (149); this gives

$$-\tfrac{1}{4}(QQ^*)R_t = -S^*_{rt}B_r = S^*_{tr}B_r. \tag{154}$$

By (150) and (143)

$$Q^*_{tr}B_r = Q^*_{tr}Q_{rs}P_s = -\tfrac{1}{4}(QQ^*)P_t. \tag{155}$$

With these preliminaries disposed of, we come now to the central point in the argument, which is *a transformation of the unknowns* in (138), (139) from (R_r, S_{rs}) to (R_r, S'_{rs}) where

$$S'_{rs} = S_{rs} - i\lambda\varepsilon_{rsmn}R_mB_n, \tag{156}$$

or equivalently

$$S'^*_{rs} = S^*_{rs} + \lambda(R_rB_s - R_sB_r), \tag{157}$$

λ being a certain invariant, to be assigned in a moment. At the same time we introduce a new (known) tensor

$$Q'_{rs} = Q_{rs} - i\lambda\varepsilon_{rsmn}P_mB_n, \tag{158}$$

or equivalently

$$Q'^*_{rs} = Q^*_{rs} + \lambda(P_rB_s - P_sB_r). \tag{159}$$

We choose λ to make Q'_{rs} singular, i.e. so that

$$(Q'Q'^*) = 0. \tag{160}$$

Now by (158) and (159) we find

$$\begin{aligned}(Q'Q'^*) &= (QQ^*) + 2\lambda Q_{rs}P_rB_s - i\lambda\varepsilon_{rsmn}Q^*_{rs}P_mB_n\\ &= (QQ^*) - 2\lambda(BB) + 2\lambda Q_{mn}P_mB_n\\ &= (QQ^*) - 4\lambda(BB).\end{aligned} \tag{161}$$

Let us then choose

$$\lambda = \tfrac{1}{4}\frac{(QQ^*)}{(BB)}, \tag{162}$$

thus satisfying (160).

We now calculate the invariants $(S'S')$, $(Q'Q')$, $(Q'S')$. The work is a little simplified by noting the fact that for any skew-symmetric tensor $(XX) = -(X^*X^*)$, and then using (157) and (159). The expressions are reduced by means of (154) and (155), and we get

$$\begin{aligned}(S'S') &= (SS) + \tfrac{1}{8}(QQ^*)^2\,(RR)/(BB),\\ (Q'Q') &= (QQ) + \tfrac{1}{8}(QQ^*)^2\,(PP)/(BB),\\ (Q'S') &= (QS) + \tfrac{1}{8}(QQ^*)^2\,(PR)/(BB).\end{aligned} \tag{163}$$

By (156) and (158) we have

$$\left.\begin{aligned} Q'_{rs}P_s = Q_{rs}P_s, \quad S'_{rs}R_s = S_{rs}R_s, \\ Q'_{rs}R_s + S'_{rs}P_s = Q_{rs}R_s + S_{rs}P_s. \end{aligned}\right\} \tag{164}$$

If we now use (163) and (164) in (138) and (139), we have the following set of equations (10 of them independent) to solve for the 10 unknowns (R_r, S'_{rs}):

$$\left.\begin{aligned} (PP) + (RR) &= -2(m^2 + m'^2), \\ (PR) &= -(m^2 - m'^2), \\ (Q'Q') + (S'S') &= 4(\Omega^2 + \Omega'^2) - \tfrac{1}{4}(m^2 + m'^2)\,(QQ^*)^2/(BB), \\ (Q'S') &= 2(\Omega^2 - \Omega'^2) - \tfrac{1}{8}(m^2 - m'^2)\,(QQ^*)^2/(BB), \\ Q'_{rs}P_s + S'_{rs}R_s = 0, \quad & Q'_{rs}R_s + S'_{rs}P_s = 0. \end{aligned}\right\} \tag{165}$$

When we have solved these equations for (R_r, S'_{rs}), we shall have S_{rs} by (156), and then the original unknowns $(M_r, M'_r, H_{rs}, H'_{rs})$ will be given by (136).

It will be observed that the equations (165) are of the same form as (138), (139), and so it might appear that we have made no progress towards the solution; it might seem that we have gone in the wrong direction, since the right hand sides in (165) are more complicated than those in (138). Our achievement consists in having Q'_{rs} singular, whereas Q_{rs} was not singular; this we shall now make use of.

The data of the problem provide us with the following three vectors:

$$P_r, \quad B_r = Q_{rs}P_s = Q'_{rs}P_s, \quad B^*_r = Q'^*_{rs}P_s, \tag{166}$$

the last being now introduced for the first time. It is evident that P_r is orthogonal to B_r and B^*_r; further, B_r is orthogonal to B^*_r, because

$$B_rB^*_r = Q'_{rs}Q'^*_{rt}P_sP_t = 0 \tag{167}$$

by (146). *Thus P_r, B_r, B^*_r form an orthogonal triad; this fact enables us to reduce the formidable algebra of* (165).

In order to simplify the algebra, it is convenient to define p, b, b^* (all positive) by

$$p^2 = -(PP), \quad b^2 = (BB), \quad b^{*2} = (B^*B^*). \tag{168}$$

Then we have

$$(Q'Q') = 2(b^{*2} - b^2)/p^2; \tag{169}$$

this is easily seen by taking P_r for time axis, so that $P_\varrho = 0$, $P_4 = ip$,

and then

$$\begin{aligned} b^2 - b^{*2} &= - Q'_{r4}Q'_{r4}p^2 + Q'^{*}_{r4}Q'^{*}_{r4}p^2 \\ &= - p^2(Q'_{\varrho 4}Q'_{\varrho 4} - Q'^{*}_{\varrho 4}Q'^{*}_{\varrho 4}) \\ &= - \tfrac{1}{2}p^2(Q'Q'). \end{aligned} \tag{170}$$

We define dimensionless (known) invariants as follows:

$$\begin{gathered} C = 1 - 2(m^2 + m'^2)/p^2, \quad D = - (m^2 - m'^2)/p^2, \\ E = \frac{p^2}{b^{*2}}\{2(\Omega^2+\Omega'^2) - \tfrac{1}{8}(m^2+m'^2)(QQ^*)^2/b^2 - (b^{*2}-b^2)/p^2\}, \\ F = \frac{p^2}{b^{*2}}\{\Omega^2 - \Omega'^2 - \tfrac{1}{16}(m^2 - m'^2)\,(QQ^*)^2/b^2\}. \end{gathered} \tag{171}$$

With this notation, the equations (165), which we have to solve, read

$$\left.\begin{aligned} (RR) &= p^2C, & (PR) &= p^2D, \\ (S'S') &= 2b^{*2}E/p^2, & (Q'S') &= 2b^{*2}F/p^2, \end{aligned}\right\} \tag{172}$$

$$Q'_{rs}P_s + S'_{rs}R_s = 0, \tag{173}$$

$$Q'_{rs}R_s + S'_{rs}P_s = 0. \tag{174}$$

We now use a special frame of reference, uniquely determined by the data of the problem, taking the axes $Ox_1x_2x_3x_4$ in the directions of the orthogonal tetrad

$$B_r, \quad B^*_r, \quad G_r, \quad P_r, \tag{175}$$

respectively, where

$$G_r = - i\varepsilon_{smrn}B_sB^*_mP_n; \tag{176}$$

in doing this we use the fact that P_r is timelike, pointing into the future, and in (176) we have ensured that the orientation of the tetrad is correct. We have then

$$\begin{aligned} &B_1 = b, \quad B_2 = 0, \quad B_3 = 0, \quad B_4 = 0, \\ &B^*_1 = 0, \quad B^*_2 = b^*, \quad B^*_3 = 0, \quad B^*_4 = 0, \\ &P_1 = 0, \quad P_2 = 0, \quad P_3 = 0, \quad P_4 = ip. \end{aligned} \tag{177}$$

Then

$$B_r = Q'_{rs}P_s = ipQ'_{r4}, \tag{178}$$

and when we put this into (177) we get

$$ipQ'_{14} = b, \quad Q'_{24} = 0, \quad Q'_{34} = 0. \tag{179}$$

Also

$$B_r^* = Q_{rs}'^* P_s = ipQ_{r4}'^*, \tag{180}$$

and (177) gives

$$Q_{14}'^* = 0, \quad ipQ_{24}'^* = b^*, \quad Q_{34}'^* = 0. \tag{181}$$

We conclude that

$$Q_{23}' = Q_{12}' = Q_{24}' = Q_{34}' = 0, \quad Q_{14}' = -ib/p, \quad Q_{31}' = -b^*/p, \tag{182}$$

so that only two of the six components of Q_{rs}' survive; further

$$(Q'Q') = 2Q_{14}'^2 + 2Q_{31}'^2 = 2(b^{*2} - b^2)/p^2. \tag{183}$$

We recall (153); it tells us that for our special axes

$$R_1 = 0. \tag{184}$$

Then (174) reduce to

$$\begin{aligned} &ipS_{14}' = -Q_{13}'R_3 - Q_{14}'R_4 = (-b^*R_3 + ibR_4)/p, \\ &S_{24}' = 0, \quad S_{34}' = 0, \quad 0 = 0, \end{aligned} \tag{185}$$

and (173) reduce to

$$\begin{aligned} &S_{12}'R_2 + S_{13}'R_3 + S_{14}'R_4 = -ipQ_{14}' = -b, \\ &S_{23}'R_3 = 0, \quad S_{32}'R_2 = 0, \quad 0 = 0. \end{aligned} \tag{186}$$

Now if $S_{23}' \neq 0$, we get $R_2 = R_3 = 0$, and with (184) this gives R_r the same direction as P_r. This is a particular case we already decided to exclude; therefore $S_{23}' = 0$, and so we have only three components of S_{rs}' to find, since

$$S_{23}' = S_{24}' = S_{34}' = 0. \tag{187}$$

By (185) and (186) the surviving components $(S_{12}', S_{13}', S_{14}')$ satisfy

$$\left.\begin{aligned} ip^2S_{14}' &= -b^*R_3 + ibR_4, \\ S_{12}'R_2 + S_{13}'R_3 + S_{14}'R_4 &= -b. \end{aligned}\right\} \tag{188}$$

Of the equations (172), (173), (174), we have now used up (173), (174) — they are equivalent to (187), (188) —, and (172) now read

$$\left.\begin{aligned} R_2^2 + R_3^2 + R_4^2 &= p^2C, \\ ipR_4 &= p^2D, \\ S_{12}'^2 + S_{13}'^2 + S_{14}'^2 &= b^{*2}E/p^2, \\ -ibS_{14}' + b^*S_{13}' &= b^{*2}F/p. \end{aligned}\right\} \tag{189}$$

Our problem is now reduced to (188), (189), these being six equations for the six unknowns $(R_2, R_3, R_4, S'_{12}, S'_{13}, S'_{14})$, the other components of R_r and S'_{rs} being zero by (184), (187).

We have R_4 at once:

$$R_4 = -ip D. \tag{190}$$

Let us now simplify the writing by introducing new unknowns $(Y, Z, \eta, \zeta, \omega)$, putting

$$\begin{gathered} R_2 = pY, \quad R_3 = pZ, \\ S'_{12} = b^*\eta/p, \quad S'_{13} = b^*\zeta/p, \quad S'_{14} = -ib^*\omega/p. \end{gathered} \tag{191}$$

Then for these unknowns we have the five equations (putting $b/b^* = \beta$)

$$\left.\begin{aligned} \omega &= -Z + \beta D, \\ \eta Y + \zeta Z - \omega D &= -\beta, \\ Y^2 + Z^2 &= C + D^2, \\ \eta^2 + \zeta^2 - \omega^2 &= E, \\ -\beta\omega + \zeta &= F. \end{aligned}\right\} \tag{192}$$

From the first and last we express Z and ζ in terms of ω:

$$Z = -\omega + \beta D, \quad \zeta = \beta\omega + F. \tag{193}$$

The other three equations in (192) may then be written

$$\left.\begin{aligned} \eta Y &= (\omega - \beta D)(\beta\omega + F) + \omega D - \beta, \\ Y^2 &= -(\omega - \beta D)^2 + C + D^2, \\ \eta^2 &= \omega^2 - (\beta\omega + F)^2 + E. \end{aligned}\right\} \tag{194}$$

Elimination of Y and η gives the following biquadratic equation for ω:

$$\begin{aligned} &[(\omega - \beta D)(\beta\omega + F) + \omega D - \beta]^2 \\ &\quad + [(\omega - \beta D)^2 - C - D^2][\omega^2 + E - (\beta\omega + F)^2] = 0, \end{aligned} \tag{195}$$

which may also be written

$$\omega^4 + A_2\omega^2 + A_1\omega + A_0 = 0, \tag{196}$$

where

$$\left.\begin{aligned} A_2 &= C(\beta^2 - 1) + 2DF + E - 2\beta^2, \\ A_1 &= 2\beta[D(\beta^2 - 1 - E) + F(C - 1)], \\ A_0 &= \beta^2(1 + 2DF + D^2E) + (F^2 - E)(C + D^2). \end{aligned}\right\} \tag{197}$$

This biquadratic equation is the centre of the algebraic problem; once it has been solved for ω, all the other unknowns follow at once. In the next section we shall summarise the whole procedure.

§ 12. SUMMARY OF PROCEDURE FOR SOLVING A COLLISION PROBLEM

The procedure for solving a collision problem may be summarised in the following steps:

a) We are given the four quantities $(m, \Omega; m', \Omega')$ characterising the two particles.

b) In some frame of reference (F) we are given the initial 4-momenta $(M_r^{(0)}, M_r'^{(0)})$ and the initial intrinsic angular momenta $(H_{rs}^{(0)}, H_{rs}'^{(0)})$, satisfying

$$\left.\begin{aligned} &M_r^{(0)}M_r^{(0)} = -m^2, && M_r'^{(0)}M_r'^{(0)} = -m'^2, \\ &H_{rs}^{(0)}H_{rs}^{(0)} = 2\Omega^2, && H_{rs}'^{(0)}H_{rs}'^{(0)} = 2\Omega'^2, \\ &H_{rs}^{(0)}M_s^{(0)} = 0, && H_{rs}'^{(0)}M_s'^{(0)} = 0. \end{aligned}\right\} \tag{198}$$

c) From these data we calculate the following vectors, tensors and invariants:

$$\left.\begin{aligned} &P_r = M_r^{(0)} + M_r'^{(0)}, \quad Q_{rs} = H_{rs}^{(0)} + H_{rs}'^{(0)}, \\ &B_r = Q_{rs}P_s, \quad b^2 = (BB), \quad b > 0, \\ &(QQ^*) = Q_{rs}Q_{rs}^*, \quad \lambda = \tfrac{1}{4}(QQ^*)/b^2, \\ &Q_{rs}'^* = Q_{rs}^* + \lambda(P_rB_s - P_sB_r), \\ &B_r^* = Q_{rs}'^*P_s, \quad b^{*2} = (B^*B^*), \quad b^* > 0, \\ &\beta = b/b^*, \quad p^2 = -(PP), \quad p > 0. \end{aligned}\right\} \tag{199}$$

d) From these we calculate the invariants C, D, E, F by (171) and then A_2, A_1, A_0 by (197).

e) We solve the biquadratic equation (196) for ω. On account of the existence of the trivial solution in which nothing is changed by the collision, we know that (196) has one real root; therefore it has two real roots; it may have four real roots, but not more. Discussion of exceptional cases is tedious (some will be dealt with in the next section), and we shall here assume that (196) has four distinct roots, none of which is a zero of any of the right hand sides in (194). An imaginary or complex root must be rejected.

f) Having got ω from (196), Z and ζ are given by (193). Then we get η from the last of (194); there are two values (plus and minus) corre-

sponding to any ω; we reject a solution with imaginary η. We get Y from the first of (194). At most there are *eight* solutions for (Y, Z, η, ζ, ω), and hence *at most eight solutions of the collision problem*, including the trivial solution. There are always at least two solutions (one of them the trivial one), but in exceptional cases these may collapse into one only.

g) We calculate an orthogonal tetrad of unit vectors as follows:

$$\begin{gathered} I_r = B_r/b, \quad J_r = B_r^*/b^*, \quad L_r = P_r/p, \\ K_r = -i\varepsilon_{smrn} I_s J_m L_n. \end{gathered} \tag{200}$$

This gives the special frame for which (177) hold; let us call this frame F_0 to distinguish it from the frame F with which we started.

h) Substituting the results (*f*) in (191), we get (R_r, S'_{rs}) in F_0. By (156) and (177) we then have for S_{rs} in F_0

$$\begin{gathered} S_{12} = S'_{12} = b^*\eta/p, \quad S_{13} = S'_{13} = b^*\zeta/p, \\ S_{14} = S'_{14} = -ib^*\omega/p, \\ S_{23} = -\lambda pbD, \quad S_{24} = i\lambda bR_3 = i\lambda bpZ, \\ S_{34} = -i\lambda bR_2 = -i\lambda bpY. \end{gathered} \tag{201}$$

i) The results of the collision, including the trivial result (i.e. the initial values unchanged), are then given in the frame F by calculating[1]

$$\begin{aligned} R_r &= pYJ_r + pZK_r - pDL_r, \\ S_{rs} &= -\lambda pbD(J_rK_s + J_sK_r) - b^*\zeta p^{-1}(K_rI_s - K_sI_r) \\ &\quad + b^*\eta p^{-1}(I_rJ_s - I_sJ_r) - b^*\omega p^{-1}(I_rL_s - I_sL_r) \\ &\quad + \lambda bpZ(J_rL_s - J_sL_r) - \lambda bpY(K_rL_s - K_sL_r), \end{aligned} \tag{202}$$

and substituting in the formulae

$$\begin{aligned} M_r &= \tfrac{1}{2}(P_r + R_r), \quad H_{rs} = \tfrac{1}{2}(Q_{rs} + S_{rs}), \\ M'_r &= \tfrac{1}{2}(P_r - R_r), \quad H'_{rs} = \tfrac{1}{2}(Q_{rs} - S_{rs}). \end{aligned} \tag{203}$$

As remarked earlier, the biquadratic equation (196) is the centre of the algebraic problem, and we must solve it if we are to obtain all the solutions. But Gardner has observed that if we are given one solution (R_r, S_{rs}) — it may be the trivial one — then we can obtain a

[1] We check the correctness of these tensor formulae by taking coordinate axes along the tetrad (I_r, J_r, K_r, L_r), and comparing the resulting values of (R_r, S_{rs}) with the values given in (184), (190) and (201).

second by a process of reflection without solving (196) at all. To see this, we note that if $(Y, Z, \eta, \zeta, \omega)$ form a solution of the equations (192), then so do $(-Y, Z, -\eta, \zeta, \omega)$. Hence we can obtain a second solution from a given solution of the form (202) simply by changing the signs of Y and η; this gives a new solution $(\bar{R}_r, \bar{S}_{rs})$ in the form

$$\begin{aligned} \bar{R}_r &= R_r - 2pYJ_r, \\ \bar{S}_{rs} &= S_{rs} - 2b^*\eta p^{-1}(I_rJ_s - I_sJ_r) + 2\lambda bpY(K_rL_s - K_sL_r). \end{aligned} \tag{204}$$

Now Y and η appearing here are known, since they belong to the known solution; in fact, by (202) we have

$$pY = R_rJ_r, \quad b^*\eta p^{-1} = S_{rs}I_rJ_s, \quad \lambda bpY = S_{rs}K_rL_s, \tag{205}$$

and so we have the second solution

$$\begin{aligned} \bar{R}_r &= R_r - 2J_rR_sJ_s, \\ \bar{S}_{rs} &= S_{rs} - 2(I_rJ_s - I_sJ_r)S_{mn}I_mJ_n + 2(K_rL_s - K_sL_r)S_{mn}K_mL_n. \end{aligned} \tag{206}$$

§ 13. PARTICULAR CASES OF COLLISIONS

Two types of invariants are involved in the algebra of the preceding sections. Those of the first type (m, m', Ω, Ω') depend only on the nature of the colliding particles, and not at all on the way in which they collide; those of the second type (b, b^*, p, C, D, E, F) involve both the nature of the particles and the way in which they collide. In this section we examine the consequences of some particular choices of the invariants of the first type, leaving those of the second type unrestricted except in so far as they depend on the nature of the particles. Three cases will be considered: (*a*) collision of two identical material particles, (*b*) collision of a material particle and a photon, (*c*) collision of two photons. It must be remembered that the conclusions we come to here are consequences of the concepts of 4-momentum and intrinsic angular momentum as developed in this book, the laws of conservation, and the assumption that the world lines of the particles intersect at a collision, i.e. the impact parameter is zero.

(*a*) *Collision of two identical material particles*

This type of collision was already discussed on pp. 227—235 by a different method. To follow the general method, we put

$$m = m' \neq 0, \quad \Omega = \Omega' \neq 0 \tag{207}$$

in (171); this gives

$$D = F = 0, \tag{208}$$

and so (197) reduce to

$$A_2 = C(\beta^2 - 1) + E - 2\beta^2, \quad A_1 = 0, \quad A_0 = \beta^2 - CE. \tag{209}$$

The biquadratic (196) becomes a quadratic in ω^2 [cf. (107)],

$$\omega^4 + \omega^2[C(\beta^2 - 1) + E - 2\beta^2] + \beta^2 - CE = 0, \tag{210}$$

where, by (171),

$$C = 1 - \frac{4m^2}{p^2}, \quad E = \frac{p^2}{b^{*2}}\left\{4\Omega^2 - \tfrac{1}{4}m^2\frac{(QQ^*)^2}{b^2}\right\} + \beta^2 - 1. \tag{211}$$

The equations (193), (194) simplify to

$$\left.\begin{aligned} Z &= -\omega, \quad \zeta = \beta\omega, \\ \eta Y &= \beta(\omega^2 - 1), \\ Y^2 &= C - \omega^2, \\ \eta^2 &= \omega^2(1 - \beta^2) + E. \end{aligned}\right\} \tag{212}$$

(*b*) *Collision of a material particle and a photon*

This type of collision has already been discussed on pp. 193—199 (Compton effect); the theory there was based on 4-momentum only, and there was a two-fold indeterminacy in the result of the collision. By including intrinsic angular momentum, as we now do, that indeterminacy is removed, in the sense that the number of possible results of the collision is now finite.

We put

$$m \neq 0, \quad \Omega \neq 0, \quad m' = 0, \quad \Omega' = 0; \tag{213}$$

then by (171) we have

$$C = 1 + 2D, \quad E = 2F + \beta^2 - 1, \tag{214}$$

and (197) gives

$$\left.\begin{aligned} A_2 &= 2D\beta^2 + 2(D + 1)(F - 1), \\ A_1 &= 0, \\ A_0 &= D^2\beta^4 + 2D\beta^2(D + 1)(F - 1) + (D + 1)^2(F - 1)^2. \end{aligned}\right\} \tag{215}$$

In the case (*a*) (two identical material particles) the biquadratic (196) degenerated to a quadratic in ω^2 as in (210); now in case (*b*) (material particle and photon) we get a greater degeneracy, the biquadratic

becoming a perfect square; for, by (215), (196) becomes

$$[\omega^2 + D\beta^2 + (D + 1)(F - 1)]^2 = 0. \tag{216}$$

This gives two, and only two, values for ω:

$$\omega = \pm [- D\beta^2 - (D + 1)(F - 1)]^{\frac{1}{2}}, \tag{217}$$

necessarily real on account of the existence of the trivial solution. Solving the last of (194) for η, we get two or four solutions, but at least two, the right hand side being certainly positive for the trivial ω (say $\omega = \omega_0$), but not necessarily positive for $\omega = -\omega_0$.

Thus when a material particle collides with a photon, there are certainly two results of the collision and possibly four (but not more), the trivial result being included in these counts.

(*c*) *Collision of two photons*

This is the most degenerate case of all. We put

$$m = \Omega = m' = \Omega' = 0, \tag{218}$$

and (171) gives

$$C = 1, \quad D = 0, \quad E = \beta^2 - 1, \quad F = 0, \tag{219}$$

and by (197)

$$A_2 = -2, \quad A_1 = 0, \quad A_0 = 1. \tag{220}$$

Then the biquadratic equation (196) reads

$$\omega^4 - 2\omega^2 + 1 = 0, \tag{221}$$

and we have

$$\omega = \varepsilon \ (= \pm 1); \tag{222}$$

(193), (194) now read

$$Z = -\varepsilon, \quad \zeta = \beta\varepsilon, \quad \eta Y = 0, \quad Y^2 = 0, \quad \eta^2 = 0. \tag{223}$$

There are therefore just two results of the collision, including the trivial result.

In any frame F these two results read, by (202),

$$\left.\begin{aligned} R_r &= -p\varepsilon K_r, \\ S_{rs} &= -b\varepsilon p^{-1}(K_r I_s - K_s I_r) \\ &\quad - b^*\varepsilon p^{-1}(I_r L_s - I_s L_r) - \varepsilon\lambda b p(J_r L_s - J_s L_r). \end{aligned}\right\} \tag{224}$$

To pass from the one to the other, we merely reverse the signs of all the quantities (R_r, S_{rs}). It is clear then from (136) that the non-

trivial solution is given in terms of the initial values by

$$\begin{aligned} M_r &= M_r'^{(0)}, \quad H_{rs} = H_{rs}'^{(0)}, \\ M_r' &= M_r^{(0)}, \quad H_{rs}' = H_{rs}^{(0)}. \end{aligned} \tag{225}$$

The 4-momenta and intrinsic angular momenta of the two photons are interchanged, and so this second solution, which we have called "non-trivial" in a technical sense, should indeed be regarded as trivial, since nothing happens except a relabelling of the photons. We may say that *two photons cannot collide except trivially.*

To sum up briefly:

In general we have to solve a biquadratic equation; there may be eight results of the collision (not more).

For two identical material particles, the biquadratic becomes a quadratic for the square of the unknown. Again there may be eight solutions (not more).

For a material particle and a photon the biquadratic becomes a perfect square; there may be four results of the collision (not more).

For two photons the biquadratic is again a perfect square; the only solutions are the trivial one (nothing changed) and one in which the photons are interchanged.

Note that, for a photon, we have not only made the usual assumption of vanishing proper mass ($m = 0$); we have also assumed that its intrinsic angular momentum invariant vanishes ($\Omega = 0$). This latter assumption simplifies the algebra but is by no means essential, and the case $\Omega \neq 0$ might be worked out.

In thinking of determinacy in connection with these problems, it must be remembered that the angular momentum tensors (not merely the invariants) are supposed to be known before collision.

§ 14. EXTERNAL IMPULSES AND IMPULSIVE TORQUES ACTING ON A SYSTEM

In the first six sections of this chapter we studied the mechanics of a system of material particles, photons and internal impulses, everything being based on the conservation of 4-momentum at each collision in the system. Then, in § 7, we introduced particles with intrinsic angular momentum and the subsequent sections were devoted to a detailed study of collisions of a pair of such particles.

Let us now return to the mechanics of a system, composed, as

earlier, of material particles, photons and internal impulses. But we shall now enlarge our previous concepts by supposing that each of these entities may possess intrinsic angular momentum, represented by a skew-symmetric tensor which remains constant between collisions. The argument which follows is so general that it can take care of this enlargement without inconvenience.

The only basic laws assumed below are as follows:

i) Between collisions, the 4-momentum and the intrinsic angular momentum of each entity remain constant.

ii) At each collision, the total 4-momentum and the total intrinsic angular momentum of the participants are conserved.

iii) A law concerning the application of external impulses and external impulsive torques, which will be explained below.

Omitting for the moment the third of these, it is easy to see that the open and closed laws of conservation of 4-momentum of pp. 214—215, and also the open and closed laws of conservation of angular momentum of p. 218, hold good even when the entities possess intrinsic angular momentum.

In Newtonian mechanics it is a fundamental practice to isolate (mentally) a system of particles. Thus, to find the reaction across a section of a bar swinging under gravity, we cut it across mentally at the section and take as our "system" the part of the bar below (or above) the section.

This procedure can be carried over into relativity. We think of a system composed of material particles, photons and internal impulses, which not only experience collisions among themselves, but are also subjected to *external impulses* (4-vectors) and *external impulsive torques* (skew-symmetric tensors). These external effects may be thought of as arising from collisions with the entities of another system, but we may also regard them as existing in their own right, just as in Newtonian mechanics we admit the concept of external forces, without necessarily regarding them as due to other particles outside the system we have isolated.

We now fill in the law (iii) above as follows:

iiia) When an external impulse acts, there is an increase of 4-momentum equal to the external impulse.

iiib) When an external impulsive torque acts, there is an increase of angular momentum equal to the external impulsive torque.

Let us slice space-time across by two non-intersecting spacelike 3-spaces, S_1 and S_2, as in Fig. 10, with S_2 after S_1 in time. They might be the 3-flats $t = t_1$ and $t = t_2$ with $t_2 > t_1$. We can then speak of the total 4-momentum of the system with respect to S_1, or with respect to S_2, in the sense of (12). Let us avoid the possibility of divergence by supposing the system finite.

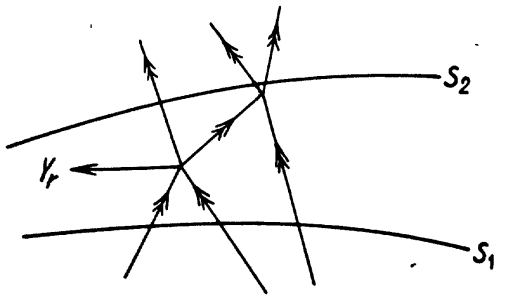


Fig. 10 – The open law of 4-momentum

If external impulses act on the system (Y_r in Fig. 10), it is clear, by the same argument as on p. 214 but with inclusion of (iiia) above, that we have the following *open law of 4-momentum*:

$$\begin{aligned}&\text{(Total 4-momentum with respect to } S_2\text{)}\\ &-\text{ (total 4-momentum with respect to } S_1\text{)}\\ &=\text{ sum of all external impulses applied}\\ &\quad\text{at events between } S_1 \text{ and } S_2.\end{aligned} \tag{226}$$

Similarly, if we take a closed 3-space S in space-time (Fig. 11), and define flux of 4-momentum as in (14), we get the *closed law of 4-momentum*:

$$\begin{aligned}&\text{Flux of 4-momentum across } S\\ &=\text{ sum of all external impulses applied at events inside } S.\end{aligned} \tag{227}$$

In the same way, if we supplement the reasoning of p. 217 with (iiia) and (iiib) above, we get the *open law of angular momentum*:

$$\begin{aligned}&\text{(Total angular momentum with respect to } S_2\text{)}\\ &-\text{ (total angular momentum with respect to } S_1\text{)}\\ &=\text{ (sum of moments of all external impulses}\\ &\quad\text{applied at events between } S_1 \text{ and } S_2\text{)}\\ &+\text{ (sum of all external impulsive torques applied}\\ &\quad\text{at events between } S_1 \text{ and } S_2\text{)}.\end{aligned} \tag{228}$$

Here all angular momentum are taken relative to some specified event, perhaps relative to the origin, in which case the moment of an external impulse Y_r is

$$x_r Y_s - x_s Y_r. \tag{229}$$

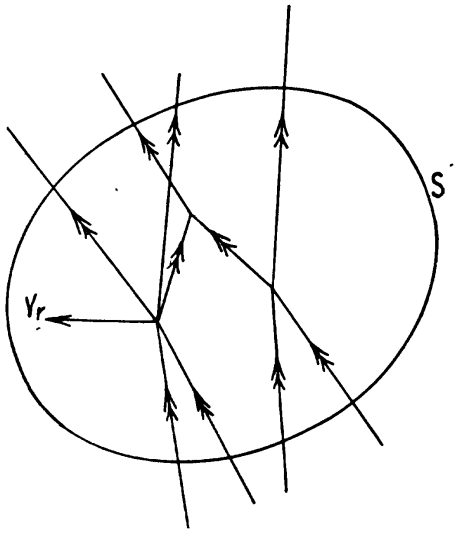


Fig. 11 – The closed law of 4-momentum

Also, for a closed 3-space S, we have the *closed law of angular momentum*:

$$\begin{aligned}&\text{Flux of angular momentum across } S\\ &= (\text{sum of moments of all external impulses applied at events inside } S)\\ &+ (\text{sum of all external impulsive torques applied at events inside } S).\end{aligned} \tag{230}$$

In these laws of angular momentum, intrinsic angular momentum has been lumped in with the ordinary angular momentum of the type defined in (15). But it is possible to separate it out. For if, as we shall suppose, the intrinsic angular momentum of any entity is unchanged between collisions, and if the total intrinsic angular momentum is unchanged by a collision, then it is easy to see that we have laws for intrinsic angular momentum as follows:

Open law of intrinsic angular momentum:

$$\begin{aligned}&(\text{Total intrinsic angular momentum with respect to } S_2)\\ &- (\text{total intrinsic angular momentum with respect to } S_1)\\ &= \text{sum of all external impulsive torques applied between } S_1 \text{ and } S_2.\end{aligned} \tag{231}$$

Closed law of intrinsic angular momentum:

$$\begin{aligned}&\text{Total flux of intrinsic angular momentum across } S\\ &= \text{sum of all external impulsive torques applied at events inside } S.\end{aligned} \tag{232}$$

In these statements an external impulsive torque is thought of as a skew-symmetric tensor, say Y_{rs}, the effect of which is to produce an increase in intrinsic angular momentum equal to Y_{rs}.

§ 15. THE TWO-BODY PROBLEM

Of all the systems we can conceive, the simplest is one consisting of two material particles, and indeed we have considered such systems, but only in respect of collisions between the particles. Such collisions were discussed, first in Ch. VI without consideration of intrinsic

angular momentum and then in the present chapter with this included, the result of this inclusion being to create determinacy where previously there was indeterminacy.

Now if the particles do not collide, they will simply proceed along straight world lines, unless we make some provision for their interaction. Such provision was made by Newton in astronomy in the form of the inverse square law of attraction, and this is carried over into atomic physics as Coulomb's law of attraction between the nucleus and a satellite electron in a hydrogenic atom. In each of these cases a hypothesis (the inverse square law) is added to the Newtonian laws of motion.

In the treatment of a discrete system, as developed in the present chapter, no provision has been made for interaction between particles except direct collision and the internal impulses, which travel with the speed of light and convey momentum and energy. We shall now see what results follow when we attack the two-body problem with these primitive weapons. We shall consider two material particles with a single internal impulse (of attractive type) shuttling to and fro between them; it will be understood that 4-momentum is conserved at each collision of the impulse with a particle, and that these collisions are elastic, in the sense that each particle retains the same proper mass throughout and the proper mass of the impulse remains always zero. Intrinsic angular momentum will not be taken into account in this work.

At first sight it would appear that the problem of finding the motion of this system (two material particles and one internal impulse) must have a high degree of indeterminacy, because the impulse plays the part of Newtonian action and reaction (made over into relativity), and we know that in Newtonian mechanics the three laws of motion by no means give a determinate motion in the two-body problem — we must add the inverse square law or some other. But this does not prove to be the case, and it is rather remarkable that the inverse square law comes out as a *deduction*, at least when we take an approximation in which the relative speeds of the particles are small compared with the speed of light. We do not make any approximation based on the size of the ratio of the masses.

Fig. 12 is a space-time diagram of the two-body problem. The broken line *abcd* is the history of a material particle of proper mass m, and $ABCD$ the history of a second particle of proper mass M. The

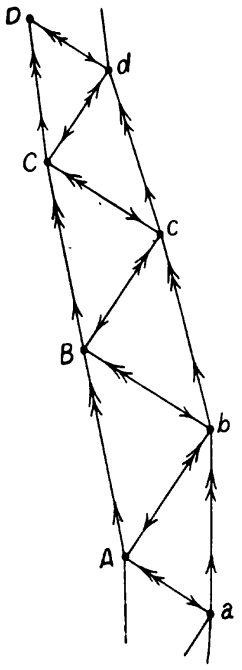


Fig. 12 – Space-time diagram of the two-body problem

broken line $aAbBcCdD$ is the history of the internal impulse oscillating between the particles; it consists of straight null segments. The whole arrangement may be called a *ladder* in space-time. The double arrows indicate sense of propagation (past to future) and the single arrows sense of 4-momentum, the two types of arrow having the same direction for the particles but opposite senses for the impulse, since we suppose it to be of the attractive type.

We shall first show that, given the two proper masses m and M, the ladder is determined by the following elements:

(i) The event A and the direction in space-time of AB.

(ii) The null line Ab and the event b on it.

(iii) The direction in space-time of ab.

(iv) The energy of the internal impulse on Ab (equivalently, the fourth component of its 4-momentum).

We note that these data determine the 4-momenta on AB, ab and Ab, the 4-momentum on Ab following from the fact that it must be a 4-vector on this assigned null line, and so it is determined by its fourth component, given by (iv) above. Further, the event B is known because it is the unique intersection of the given straight line AB and the null cone drawn into the future from the assigned event b. Consider now the collision at b. The 4-momenta of the particle on ab and the impulse on Ab are known; let their sum be P_r. Then if m_r and I_r are respectively the 4-momenta of the particle and impulse after collision, we have the equations

$$m_r + I_r = P_r, \quad m_r m_r = -m^2, \quad I_r = \theta V_r, \tag{233}$$

where V_r is the known null vector bB and θ an undetermined factor. Substituting for I_r from the third of these in the first, and then substituting for m_r in the second, we get

$$(P_r - \theta V_r)\,(P_r - \theta V_r) = -m^2, \tag{234}$$

or, since $V_r V_r = 0$,

$$2\theta V_r P_r = m^2 + P_r P_r. \tag{235}$$

Thus θ is determined, and hence, by (233), I_r and m_r also.

We now know the following:

(i) The event b and the direction in space-time of bc.
(ii) The null line bB and the event B on it.
(iii) The direction in space-time of AB.
(iv) The energy of the internal impulse on bB.

But these are the same data as were assigned originally, except that we have moved on a step and interchanged the roles of the two particles. Clearly, we can go on step by step and so construct the whole ladder from the data (i) to (iv), either in their original form or in the new form.

Thus we see that the motion in this two-body problem is determined by certain initial data, which we might roughly describe as the initial positions and velocities of the particles and the initial energy of the internal impulse. It is by no means necessary to assign beforehand the way in which the energy of the impulse changes in the course of the motion — that is a matter of deduction from the initial data, and in this respect the relativistic theory deviates at once from the Newtonian, in which we would have (instead of the energy of the impulse) a potential energy of the system, and we would be told in advance that it varied inversely as the distance between the particles.

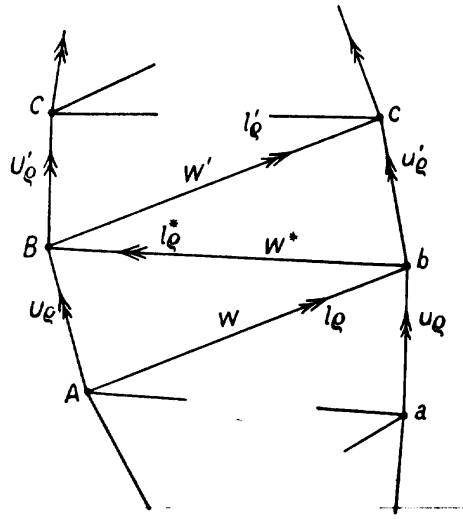


Fig. 13 – Space-diagram of the two-body problem

To discuss the two-body problem in more detail, we replace the space-time diagram by a space-diagram (Fig. 13), drawn in the space of some Galileian observer. Here a, b, c, A, B, C are the positions of the corresponding events in Fig. 12; the straight lines are tracks of the particles and the impulse, the double arrows indicating sense of propagation. We shall use the following notation:

u_ϱ, u'_ϱ = velocities on ab, bc respectively,
U_ϱ, U'_ϱ = velocities on AB, BC respectively,
W, W^*, W' = energies of internal impulse on Ab, bB, Bc respectively,
l_ϱ, l^*_ϱ, l'_ϱ = direction cosines of Ab, bB, Bc respectively (pointing in the sense of propagation).

Thus the 4-momentum of the impulse on Ab, for example, is $(Wl_\varrho/c^2, iW/c^2)$; W is negative, since the impulse is assumed to be attractive.

Application of the open law of conservation (p. 214) to the whole system gives an equation of 3-momentum

$$M\Gamma' U'_\varrho + m\gamma' u'_\varrho + W' l'_\varrho/c = M\Gamma U_\varrho + m\gamma u_\varrho + W l_\varrho/c, \tag{236}$$

where

$$\gamma = (1 - u^2/c^2)^{-\frac{1}{2}}, \quad \Gamma = (1 - U^2/c^2)^{-\frac{1}{2}}, \qquad u^2 = u_\varrho u_\varrho, \quad U^2 = U_\varrho U_\varrho, \tag{237}$$

with similar notation for γ' and Γ', and an equation of energy

$$M\Gamma' c^2 + m\gamma' c^2 + W' = M\Gamma c^2 + m\gamma c^2 + W. \tag{238}$$

There is also an equation of angular momentum, but we shall not write it down.

If we apply the conservation law to the collision at b, we get

$$\left.\begin{aligned} m\gamma' u'_\varrho + W^* l^*_\varrho/c &= m\gamma u_\varrho + W l_\varrho/c, \\ m\gamma' c^2 + W^* &= m\gamma c^2 + W. \end{aligned}\right\} \tag{239}$$

To discuss the motion methodically we would use difference equations and a more refined notation. But we shall not attempt that, confining our attention rather to slow motions (compared with light); then it is possible to replace the difference equations by differential equations, and so compare the motion with Newtonian motion.

As a basis for approximation we shall regard c as large and the following quantities as finite: the speeds of the particles, the distance between them, and the energy of the internal impulse. We shall start by neglecting terms of the order c^{-2}. Then (236) and (238) become

$$\left.\begin{aligned} MU'_\varrho + mu'_\varrho + W' l'_\varrho/c &= MU_\varrho + mu_\varrho + W l_\varrho/c, \\ \tfrac{1}{2}MU'^2 + \tfrac{1}{2}mu'^2 + W' &= \tfrac{1}{2}MU^2 + \tfrac{1}{2}mu^2 + W, \end{aligned}\right\} \tag{240}$$

and (239) becomes

$$\left.\begin{aligned} mu'_\varrho + W^* l^*_\varrho/c &= mu_\varrho + W l_\varrho/c, \\ \tfrac{1}{2}mu'^2 + W^* &= \tfrac{1}{2}mu^2 + W. \end{aligned}\right\} \tag{241}$$

Now we are to remember that, while the particle m goes from b to c, the internal impulse travels the path bBc; the time taken for this passage (say τ) is small of order c^{-1}, and is in fact

$$\tau = 2r/c \tag{242}$$

approximately, where $r = Ab$. Also $l_\varrho + l^*_\varrho$ is small of order c^{-1}.

It follows from the first of (241) that the differences $u_\varrho - u'_\varrho$ are small of order c^{-1}, and hence from the second of (241) that $W^* - W$ is also small of this order; we can then write the first of (241) in the form

$$m(u'_\varrho - u_\varrho) = 2Wl_\varrho/c, \tag{243}$$

with an error of order c^{-2}. If we divide this by τ, we get

$$\frac{m(u'_\varrho - u_\varrho)}{\tau} = \frac{Wl_\varrho}{r}, \tag{244}$$

with an error of order c^{-1}.

Now the path of either of the particles is a broken line, the velocity changing abruptly at the corners. This we replace by a continuous motion, for which, by (244),

$$m\frac{du_\varrho}{dt} = \frac{Wl_\varrho}{r}, \tag{245}$$

and similarly, working with the other particle, we get

$$M\frac{dU_\varrho}{dt} = -\frac{Wl_\varrho}{r}. \tag{246}$$

With a similar replacement of the discontinuous by the continuous, the second of (240) gives

$$\frac{d}{dt}(\tfrac{1}{2}MU^2) + \frac{d}{dt}(\tfrac{1}{2}mu^2) + \frac{dW}{dt} = 0, \tag{247}$$

or

$$MU_\varrho\frac{dU_\varrho}{dt} + mu_\varrho\frac{du_\varrho}{dt} + \frac{dW}{dt} = 0, \tag{248}$$

and when we substitute in this from (245) and (246), we get

$$\frac{W}{r}l_\varrho(u_\varrho - U_\varrho) + \frac{dW}{dt} = 0. \tag{249}$$

But the rate of change of the distance between the particles in this continuous motion is

$$\frac{dr}{dt} = l_\varrho(u_\varrho - U_\varrho), \tag{250}$$

and when we substitute this in (249) and integrate, we obtain

$$W = -\frac{k}{r}, \tag{251}$$

k being some constant, positive since we chose W negative (but of course the same argument applies to repulsion, with W positive and k negative). The approximate equations of motion (245), (246) now read

$$m\frac{du_\varrho}{dt} = -\frac{kl_\varrho}{r^2}, \quad M\frac{dU_\varrho}{dt} = \frac{kl_\varrho}{r^2}, \tag{252}$$

and these have the form of the Newtonian equations of motion under an attraction varying as the inverse square of the distance. It should be noted, however, that the constant k which appears in (251) is merely a constant of integration, and in no sense an absolute constant.

It is interesting that so fundamental a law as that of the inverse square results from what are really very simple assumptions, but one would hesitate to assume that this mechanism might be used to make valid predictions in any branch of physics. Its claim to attention rests on the fact that it embodies the simplest available generalisation to relativity of Newton's law of action and reaction.

The inverse square law is also obtained by a similar argument applied to a model in which the interaction between two material particles consists of a continuous stream of internal impulses, instead of only one (cf. SYNGE [1935a]).

CHAPTER VIII

MECHANICS OF A CONTINUUM

§ 1. DENSITY

In this chapter we shall make the transition from a discrete system to a continuum, using a statistical argument which relies to some extent on intuition. As a preliminary we shall discuss some simple special cases to illustrate the general ideas underlying the transition, but first we must deal with the question of *density*.

Imagine a stream of material particles, all moving with the same 3-velocity v_ϱ relative to some Galileian observer. We suppose the particles to be very numerous and close together.

The observer fixes his attention on a small cloud of these particles, and counts the particles in the cloud; let their number be n, a large number in a statistical sense. The observer measures the volume V which this cloud occupies at time t (this is the *relative volume* of the cloud). The quotient

$$N = n/V \tag{1}$$

is, by definition, the *relative numerical density* of the particles.

Next, the observer abandons his original frame of reference and travels along with the cloud (the same cloud as above). He again measures its volume, now denoted by V_0 (this is the *proper volume* of the cloud). The quotient

$$N_0 = n/V_0 \tag{2}$$

is, by definition, the *proper numerical density* of the particles. The number n in (1) and (2) is, of course, the same number, since it is the same cloud in each case.

We recall from p. 119 the relationship between relative volume and proper volume:

$$V = V_0/\gamma, \quad \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}. \tag{3}$$

We have therefore

$$N = \gamma N_0; \tag{4}$$

the relative numerical density is always greater than the proper numerical density.

Numerical density is, perhaps, the most fundamental type of density, but for purposes of mechanics *mass-density* is more important. Here we have an embarrassing number of possibilities in definition, because we have not only the relative and proper volumes of the cloud of particles, but also the relative and proper masses to consider. We recall from p. 169 that relative mass m^* is related to proper mass m by

$$m^* = m\gamma. \tag{5}$$

Rejecting one of the four possible definitions of mass-density as too artificial, we note the following three definitions:

$$\begin{aligned} &\text{Relative density of relative mass} = \mu^{**} = \Sigma m^*/V, \\ &\text{Relative density of proper mass} = \mu^* = \Sigma m/V, \\ &\text{Proper density of proper mass} = \mu = \Sigma m/V_0. \end{aligned} \tag{6}$$

In each case Σ denotes summation over all the particles in the cloud, which has a relative volume V and a proper volume V_0. It follows from (3) and (5) that

$$\mu^{**} = \gamma\mu^* = \gamma^2\mu. \tag{7}$$

If we multiply the expressions in (6) by c^2, we get at once appropriate definitions for the *relative density of relative energy*, the *relative density of proper energy*, and the *proper density of proper energy* for a cloud of particles.

If we have to deal with particles streaming in all directions, instead of a cloud of particles with a common velocity, there is no frame of reference in which all the particles are at rest. In such cases definitions involving proper volume V_0 are no longer available, because there is no proper volume for the system.

Something similar occurs in the case of photons. Consider a set of photons, very numerous and close together, and all moving in the same direction. There is no Galileian frame of reference in which the photons are at rest, and so there are no "proper" things which we can associate with them. Of the three definitions (6), modified by changing from mass to energy, only the first can be applied to a cloud of photons; it reads

$$\text{Relative density of relative energy} = \Sigma h\nu/V, \tag{8}$$

where ν is the frequency of a photon and the summation is taken over all photons in the relative volume V.

The *relative density of relative momentum* may be similarly defined, for material particles or for photons, by taking the sum of the relative momenta and dividing by V.

From the above discussion the reader may get the impression that, in relativity, density is a rather confusing concept. But that is because we have tried to force a relativistic concept into the Newtonian mould, by splitting space-time into the space and time of some particular Galileian observer. A real insight into the meaning of density is gained only from the space-time point of view, and this will be dealt with later. Nevertheless we cannot afford to omit presentations in space and time, because that is the way in which experimental techniques are commonly discussed, and however beautiful the theory may be, it must be brought into contact with the ideas of the laboratory.

§ 2. FUNDAMENTAL LAWS OF RELATIVE MOMENTUM AND RELATIVE ENERGY FOR A SYSTEM

Let us now interpret in terms of space and time the open law of 4-momentum, as expressed in VII–(226), p. 253.

We take the spacelike 3-spaces S_1 and S_2 to be the instantaneous spaces of some Galileian observer for times t and $t + dt$ respectively. If M_r is the 4-momentum of a material particle or a photon, then, as on pp. 169, 170, 173, we have

$$\begin{aligned} &\text{Relative momentum} = cM_\varrho, \\ &\text{Relative energy} = -ic^2M_4. \end{aligned} \tag{9}$$

The relative momentum and relative energy of an internal impulse may be defined as above in terms of its 4-momentum M_r.

We now begin the transition from the discrete to the continuous by supposing the external 4-impulses to be numerous and small, so that they can be replaced by 4-forces, acting continuously. We use (9) to write VII–(226) in the form

$$\begin{aligned} &\text{Increase in relative momentum of system in time } dt \\ &\qquad = c\,\Sigma\, dY_\varrho, \\ &\text{Increase in relative energy of system in time } dt \\ &\qquad = -ic^2\,\Sigma\, dY_4, \end{aligned} \tag{10}$$

the expressions on the right coming from the external 4-impulses applied in the interval $(t, t + dt)$.

By VII–(3) a small 4-impulse dY_r applied to a particle moving with speed v is equivalent to a continuous 4-force X_r applied during the time dt, the value of X_r being given by

$$dY_r = X_r ds = cX_r dt/\gamma, \quad \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}. \tag{11}$$

But by VI–(12) the three components X_ϱ are related to the relative 3-force P_ϱ by

$$\gamma P_\varrho = c^2 X_\varrho, \tag{12}$$

and so (11) gives

$$cdY_\varrho = P_\varrho dt. \tag{13}$$

As for dY_4, we have by VI–(11)

$$dY_4 = X_4 ds = id(m\gamma), \tag{14}$$

and so

$$-ic^2 dY_4 = d(m\gamma c^2) = dE, \tag{15}$$

where E is the relative energy ($E = m\gamma c^2$) of the particle on which the impulse acts.

We now substitute from (13) and (15) in (10), and so obtain, as deductions from the open law of 4-momentum VII–(226), the following laws:

Law of relative momentum:

$$\begin{aligned}&\text{Rate of increase of relative momentum of system}\\ &\quad = \text{total external relative force } (\Sigma P_\varrho)\\ &\quad\quad \text{acting on system.}\end{aligned} \tag{16}$$

Law of relative energy:

$$\begin{aligned}&\text{Rate of increase of relative energy of system}\\ &= \text{rate at which energy is being supplied to system}\\ &\quad \text{by external agencies.}\end{aligned} \tag{17}$$

These statements appear very natural to the Newtonian physicist, for (apart from the word *relative*) they are precisely what he would write down, (16) being the law of momentum and (17) the first law of thermodynamics. It is an important contribution of relativity that these two laws are linked together as parts of a single law of 4-momentum.

We shall now illustrate these laws by some simple examples.

§ 3. IMPACT OF A STREAM OF PARTICLES ON A TARGET

A Galileian observer sets up a small plane target of area dS, and bombards it with a stream of particles, all moving with the same velocity v_ϱ. The particles, we shall assume, are absorbed by the target.

What relative external force must be applied to the target to keep it at rest? And at what rate must heat-energy be conducted off in order to keep the target at constant temperature?

Let n_ϱ be the unit normal to the target, drawn away from the bombarding particles (Fig. 1). For our system we take the target, those particles already absorbed by it up to time t, and those particles which strike it in the interval $(t, t + dt)$. Let μ^{**} be the relative density of relative mass of the particles as they advance on the target. Those particles which strike in the interval $(t, t + dt)$ occupy a volume $v_\sigma n_\sigma dt\, dS$, and so their total relative momentum is

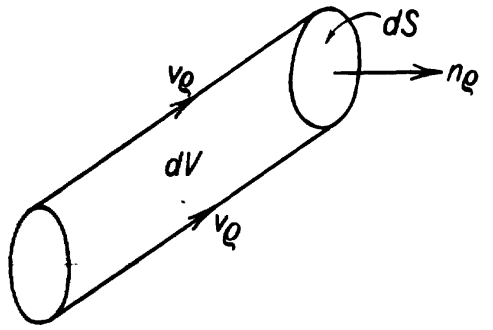


Fig. 1 – Target bombarded by particles which are absorbed by it

$$\mu^{**} v_\varrho v_\sigma n_\sigma dt dS.$$

Since this is all the relative momentum in the system at time t, and it has no relative momentum at all at time $t + dt$, we get from (16), multiplied by dt,

$$-\mu^{**} v_\varrho v_\sigma n_\sigma dt dS = P_\varrho dt, \tag{18}$$

where P_ϱ is the external relative force applied to the target to keep it at rest. Therefore

$$P_\varrho = -\mu^{**} v_\varrho v_\sigma n_\sigma dS. \tag{19}$$

This gives a stress on the target; in the case of normal impact, we get a normal pressure

$$p = \mu^{**} v^2. \tag{20}$$

This differs from the corresponding Newtonian formula only through the replacement of Newtonian density by the relative density of relative mass.

We turn now to (17). Before impact, the moving particles have a total relative energy

$$\mu^{**} c^2 v_\sigma n_\sigma dt dS.$$

This is reduced by impact, but not entirely lost, since the particles retain their proper masses (we shall assume these unchanged on impact). Now the total proper mass of these particles is

$$\mu^* v_\sigma n_\sigma dt dS,$$

where μ^* is their relative density of proper mass. We get their relative energy after absorption in the target by multiplying this by c^2.

If the target is held at constant temperature, its proper energy does not change. It is at rest, and so its relative energy does not change. Thus the only change in relative energy comes from the particles striking the target in the interval $(t, t + dt)$, and so, by (17) and (7),

$$\begin{aligned}&\text{energy supplied to system by external agencies in time } dt\\ &= (\mu^* - \mu^{**})c^2 v_\sigma n_\sigma dt dS\\ &= -(1 - \gamma^{-1})\mu^{**} c^2 v_\sigma n_\sigma dt dS. \qquad (21)\end{aligned}$$

Thus, to keep the target at constant temperature, heat-energy must be conducted off at a rate

$$(1 - \gamma^{-1})\mu^{**} c^2 v_\sigma n_\sigma, \qquad (22)$$

per unit time per unit area. For small v/c, this is approximately

$$\tfrac{1}{2}\mu^{**} v^2 v_\sigma n_\sigma, \qquad (23)$$

which we can compare with the Newtonian formula for the loss of kinetic energy of the particles.

Let us now suppose that, instead of being absorbed, the particles rebound elastically from the target, the tangential component of velocity being unchanged and the normal component being reversed. In this case it is easy to see that the relative force required to hold the target fixed is

$$P_\varrho = -2\mu^{**}(v_\sigma n_\sigma)^2 n_\varrho dS, \qquad (24)$$

which gives a normal pressure

$$p = 2\mu^{**}(v_\sigma n_\sigma)^2. \qquad (25)$$

Here $v_\sigma n_\sigma$ is the component of velocity normal to the target, and μ^{**} is the relative density of relative mass of the incident particles. Thus $2\mu^{**}$ is the relative density of relative mass in the region in which there are both incident and reflected particles.

In this case the relative energy of each particle is unchanged on collision with the target, and we do not have to conduct away any heat-energy in order to keep the temperature of the target constant.

§ 4. PRESSURE IN A RELATIVISTIC GAS

Let us now consider a more general experiment in which one side of the fixed target is bombarded elastically by particles of various proper masses and various speeds, approaching it from all possible directions. We may think of these particles as forming a relativistic gas.

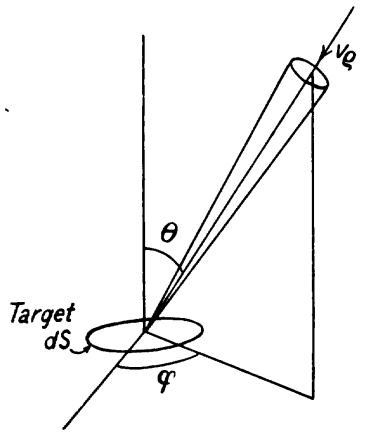


Fig. 2 – Particles incident on a target from all directions

Let us introduce spherical polar coordinates θ, ϕ as in Fig. 2, and denote by

$$f(\theta, \phi, v, m)d\omega dv dm \qquad (26)$$

the relative numerical density of particles approaching the target from directions lying in the solid angle $d\omega = \sin\theta\, d\theta\, d\phi$ enclosing the direction θ, ϕ, with speeds in the range $(v, v + dv)$, and proper masses in the range $(m, m + dm)$. Then the relative density of relative mass of particles in this category is

$$m\gamma f(\theta, \phi, v, m)d\omega dv dm. \qquad (27)$$

If we superimpose elementary pressures, as given by (25) for the various elementary streams involved, we get for the total pressure due to the gas

$$p = 2\int_{\theta=0}^{\frac{1}{2}\pi}\int_{\phi=0}^{2\pi}\int_{v=0}^{c}\int_{m=0}^{\infty} m\gamma v^2 \cos^2\theta\, f(\theta, \phi, v, m) \sin\theta\, d\theta\, d\phi\, dv\, dm. \qquad (28)$$

Let us simplify this by assuming isotropy, so that f is independent of θ and ϕ; we can write it $f(v, m)$. Then (28) gives

$$p = \frac{4\pi}{3}\int_{v=0}^{c}\int_{m=0}^{\infty} m\gamma v^2 f(v, m)dv dm. \qquad (29)$$

Multiplying (27) by c^2, integrating, and multiplying by 2 to allow for particles receding from the target, we get for the relative density of relative energy

$$4\pi\int_{v=0}^{c}\int_{m=0}^{\infty} m\gamma c^2 f(v, m)dv dm. \qquad (30)$$

The relative density of relative kinetic energy is then

$$4\pi\int_{v=0}^{c}\int_{m=0}^{\infty} m(\gamma - 1)c^2 f(v, m)dv dm. \qquad (31)$$

From (26), integrated and multiplied by 2 to allow for particles receding from the target, the relative numerical density is

$$N = 4\pi \int_{v=0}^{c} \int_{m=0}^{\infty} f(v, m)dvdm, \tag{32}$$

and so the average relative kinetic energy per particle is

$$\begin{aligned} \bar{T} &= \frac{4\pi}{N} \int_{v=0}^{c} \int_{m=0}^{\infty} m(\gamma - 1)c^2 f(v, m)dvdm \\ &= \frac{\int_{v=0}^{c} \int_{m=0}^{\infty} m(\gamma - 1)c^2 f(v, m)\, dvdm}{\int_{v=0}^{c} \int_{m=0}^{\infty} f(v, m)dvdm}. \end{aligned} \tag{33}$$

To link these formulae with Newtonian theory, let us suppose that the particles are moving slowly, or equivalently that the function $f(v, m)$ falls rapidly to zero as we increase v/c from zero. Then we may replace γ by unity in (29) and $(\gamma - 1)$ by $\frac{1}{2}v^2/c^2$ in (33). In this approximation, we have

$$\begin{aligned} p &= \frac{4\pi}{3} \int_{v=0}^{c} \int_{m=0}^{\infty} mv^2 f(v, m)dvdm, \\ \bar{T} &= \frac{2\pi}{N} \int_{v=0}^{c} \int_{m=0}^{\infty} mv^2 f(v, m)dvdm, \end{aligned} \tag{34}$$

and hence

$$p = \tfrac{2}{3}N\bar{T}, \tag{35}$$

a familiar formula in ordinary kinetic theory (cf. JEANS [1925, p. 116]).

Returning to the case where v/c is not small, let us now suppose that all the particles have the same proper mass m. We approach this case as a limit by supposing that $f(v, m)$ vanishes for all m except in a short range, so that in the limit as this range tends to zero we have

$$\int_0^{\infty} f(v, x)dx = F(v), \quad \int_0^{\infty} xf(v, x)dx = mF(v). \tag{36}$$

This function $F(v)$ is identified physically by the fact that $4\pi F(v)dv$ is the numerical density of particles (approaching and receding) with speeds in the range $(v, v + dv)$. The total relative numerical density is,

by (32),

$$N = 4\pi \int_{v=0}^{c} F(v)dv, \tag{37}$$

and (29) and (33) give

$$p = \frac{4\pi m}{3} \int_{v=0}^{c} \gamma v^2 F(v)dv,$$

$$\bar{T} = \frac{4\pi m}{N} \int_{v=0}^{c} (\gamma - 1)c^2 F(v)dv = \frac{m \int_{v=0}^{c} (\gamma - 1)c^2 F(v)dv}{\int_{v=0}^{c} F(v)dv}. \tag{38}$$

The relative density of relative mass is

$$\mu^{**} = 4\pi m \int_{v=0}^{c} \gamma F(v)dv, \tag{39}$$

and thus we have the relation

$$\bar{T} = \frac{\mu^{**}c^2}{N} - mc^2. \tag{40}$$

We may define the absolute temperature Θ of the gas by the usual gas law,

$$p = NR\Theta, \tag{41}$$

where R is a constant. Then by (37) and (38) we have

$$R\Theta = \tfrac{1}{3}m \frac{\int_{v=0}^{c} \gamma v^2 F(v)dv}{\int_{v=0}^{c} F(v)dv} \tag{42}$$

For small velocities this gives approximately

$$R\Theta = \tfrac{1}{3}m \frac{\int_{v=0}^{c} v^2 F(v)dv}{\int_{v=0}^{c} F(v)dv} = \tfrac{2}{3}m\bar{T}, \tag{43}$$

another familiar formula of kinetic theory (cf. JEANS [1925, p. 113]).

§ 5. PRESSURE DUE TO THE IMPACT OF PHOTONS

Let us make some calculations as in §§ 3, 4, but now for photons instead of for material particles. We must remember that a photon has

zero proper mass and that it travels with the speed of light. We recall also from p. 173 that the relative momentum of a photon is hvl_ϱ/c and its relative energy hv, where h is Planck's constant, v the frequency of the photon, and l_ϱ the direction cosines of the direction of its motion.

Consider first the impact of a stream of photons on a target of area dS (Fig. 1), all with the same frequency v and the same direction l_ϱ; we suppose them to be absorbed by the target. Let N be their relative numerical density.

The volume occupied by the photons which strike the target in time dt is $cl_\sigma n_\sigma dtdS$, and so the total relative momentum lost in time dt is

$$Nhvl_\varrho l_\sigma n_\sigma dtdS. \tag{44}$$

The total relative energy of the photons here involved is

$$cNhvl_\sigma n_\sigma dtdS. \tag{45}$$

It follows from (16) that the relative force which must be applied to the target to hold it at rest is

$$P_\varrho = -Nhvl_\varrho l_\sigma n_\sigma dS. \tag{46}$$

In the case of normal incidence ($l_\varrho = n_\varrho$) this means that the photons exert on the target a pressure

$$p = Nhv. \tag{47}$$

The relative energy absorbed per unit time per unit area of target is $cNhvl_\sigma n_\sigma$, and for normal incidence this is

$$cNhv = cp. \tag{48}$$

To keep the target at a constant temperature, heat energy must be conducted away at this rate per unit area of target.

Now suppose that the conductor is a perfect reflector, the photons rebounding with unchanged frequency in a direction given by the usual law of reflection. Then it is easy to see that the target experiences a pressure

$$p = 2Nhv\cos^2\theta, \tag{49}$$

where θ is the angle of incidence ($\cos\theta = l_\sigma n_\sigma$) and N the relative numerical density of the incident photons. (The numerical density is $2N$ if we include both incident and reflected photons.)

Finally, let us pass to the case of reflection from a target when the photons have all frequencies and approach the target from all possible directions. Again we use spherical polar angles as in Fig. 2, and write

$$f(\theta, \phi, v)d\omega\, dv \tag{50}$$

for the relative numerical density of photons approaching from directions lying in the solid angle $d\omega$, with frequencies in the range $(\nu, \nu + d\nu)$. Superimposing elementary pressures as in (49), we get a total pressure

$$p = 2h \int_{\theta=0}^{\frac{1}{2}\pi} \int_{\varphi=0}^{2\pi} \int_{\nu=0}^{\infty} \nu \cos^2 \theta f(\theta, \phi, \nu) \sin \theta \, d\theta \, d\phi \, d\nu. \tag{51}$$

Let us now assume an isotropic distribution of photons, so that f is independent of θ and ϕ. Then (51) gives

$$p = \frac{4\pi}{3} h \int_0^{\infty} \nu f(\nu) d\nu. \tag{52}$$

The relative density of relative energy is (allowing for both incident and reflected photons)

$$W = 4\pi h \int_0^{\infty} \nu f(\nu) d\nu, \tag{53}$$

and so we have

$$p = \tfrac{1}{3} W. \tag{54}$$

The pressure exerted by the photons on the target is one third of their relative density of relative energy.

In terms of the function f, the numerical density of photons, coming and going, is

$$N = 4\pi \int_0^{\infty} f(\nu) d\nu. \tag{55}$$

If we use the gas law (41) to define the absolute temperature Θ of our "photonic gas", we get

$$R\Theta = \frac{p}{N} = \tfrac{1}{3} \frac{W}{N} = \frac{h}{3} \frac{\int_0^{\infty} \nu f(\nu) d\nu}{\int_0^{\infty} f(\nu) d\nu}. \tag{56}$$

We note that W/N is the average relative energy per photon.

If we wished to calculate the pressure on a target moving with uniform velocity, the best plan would be to reduce the target to rest by changing the frame of reference. But then we would have to remember that a distribution of photons, isotropic with respect to the moving target, would not be isotropic with respect to the target at rest, and so the substitution of $f(\nu)$ for $f(\theta, \phi, \nu)$ would not be permissible.

This completes our preliminary treatment of the statistical passage from a discrete system to a continuum, showing how the macroscopic quantities (density, pressure and temperature) emerge from the microscopic picture. We shall now pass on to more powerful space-time methods, treating a general continuous medium.

§ 6. WORLD TUBES AND THEIR CROSS-SECTIONS

The history of a single particle is a world line in space-time, and the history of a set of particles is a set of world lines. If the particles are numerous, close together, and all moving with approximately the same velocity, then their world lines form a thin tube in space-time — a *world tube.* In fact, a world tube is the history of a small cloud of particles. In the case of material particles the world tube is timelike; in the case of photons it is null.

The concept of world tubes is fundamental in the transition from a discrete system to a continuum, and we shall now develop formulae describing the 3-volumes of their cross-sections.

In Minkowskian space-time two adjacent events define a separation

$$ds = (\varepsilon\, dx_r dx_r)^{\frac{1}{2}}, \quad \varepsilon = \pm 1. \tag{57}$$

This is the measure of Minkowskian length, and we shall use the word *length* in that sense in what follows (M-geometry as on p. 65). We recall the meaning of orthogonality: 4-vectors X_r, Y_r are orthogonal if

$$X_r Y_r = 0. \tag{58}$$

We can construct *cells* of two, three and four dimensions in space-time as follows. Draw two elements of lengths ds_1, ds_2, orthogonal to one another, and complete the elementary rectangle. This is a *2-cell,* and its *area* we define to be simply $ds_1 ds_2$, just as if we were in a Euclidean plane. Similarly we can construct a *3-cell,* starting from three mutually orthogonal edges ds_1, ds_2, ds_3, and define its *3-volume* to be $ds_1 ds_2 ds_3$. And we can construct a *4-cell* from four mutually orthogonal edges ds_1, ds_2, ds_3, ds_4, and define its *4-volume* to be $ds_1 ds_2 ds_3 ds_4$. That is of course as far as we can go, since space-time is four-dimensional, and we cannot have more than four mutually orthogonal directions. These cells are illustrated in Fig. 3.

When we speak of the area of a portion of a 2-flat, we understand the limit of the sum of the areas of small 2-cells into which it may be divided. More accurately, it is the upper bound of the sum of the areas

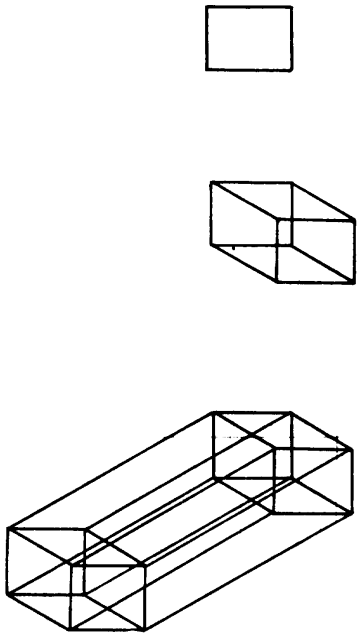


Fig. 3 – Space-time diagrams of 2-cell, 3-cell and 4-cell

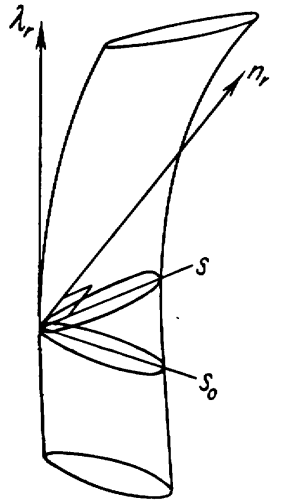


Fig. 4 – Section of a world tube by a 3-space S

of non-overlapping 2-cells contained within its boundary. By making use of orthogonal projections on 2-flats, we can define in this way the area of a curved 2-space immersed in space-time — a procedure which we apply in ordinary Euclidean space. In the same way we can define the 3-volume of a curved 3-space in space-time, or (more simply) the 4-volume of a portion of space-time itself. These areas, 3-volumes and 4-volumes then present themselves as double, triple and quadruple integrals. But it is good to realise how these concepts rest fundamentally on the cells shown in Fig. 3, and these in turn on the *metric* of space-time.

Let us now consider a thin world tube (Fig. 4), which may be either timelike or spacelike, the latter being included because we shall need it later, although it does not present itself in the case of either material particles or photons. The null world tube we omit; we shall later approach it as a limit.

Let λ_r be the unit tangent vector drawn along one of the curves which make up the tube. Then

$$\varepsilon(\lambda)\lambda_r\lambda_r = 1,$$

$$\varepsilon(\lambda) = -1 \text{ for timelike tube}, \tag{59}$$

$$\varepsilon(\lambda) = +1 \text{ for spacelike tube}.$$

We cut the tube across by a 3-space with unit normal n_r. Then

$$\varepsilon(n) n_r n_r = 1,$$
$$\varepsilon(n) = -1 \text{ for timelike } n_r, \tag{60}$$
$$\varepsilon(n) = +1 \text{ for spacelike } n_r.$$

Our object is to find how the 3-volume of the cross-section of the tube depends on n_r. We shall denote this 3-volume by S.

To do this, we take a normal section of the tube, i.e. we cut it with the 3-space having λ_r for normal. Let this normal cross-section have a 3-volume S_0. We shall obtain a simple formula connecting S and S_0.

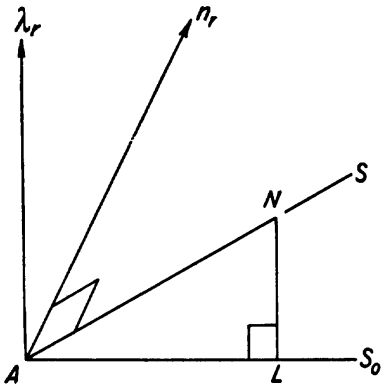


Fig. 5 – Normal (S_0) and oblique (S) cross-sections of a world tube

Fig. 5 shows the 2-flat containing the vectors λ_r and n_r, and the traces on that 2-flat of the two cross-sections; AL is the trace of S_0 and AN the trace of S. In order to estimate their 3-volumes, we think of the two cross-sections divided into 3-cells, each 3-cell having two of its edges parallel respectively to two directions which are orthogonal to one another and to the 2-flat shown in Fig. 5; the third edge of a 3-cell in S_0 is parallel to AL, and the third edge of a 3-cell in S is parallel to AN. We may think of the 3-cells in S_0 as generated by first taking 3-cells in S, and then projecting them on S_0 by drawing lines parallel to λ_r, this direction of projection being NL in Fig. 5. Since this projection does not change the lengths of edges which are orthogonal to the 2-flat of λ_r and n_r, it follows that the ratio of the 3-volumes of corresponding 3-cells (and hence the ratio S_0/S) is simply the ratio AL/AN.

We have then

$$S_0/S = AL/AN, \tag{61}$$

and we proceed to calculate this ratio. Since the 4-vector AL_r lies in the 2-flat of λ_r and n_r, we may write

$$AL_r = a\lambda_r + bn_r, \tag{62}$$

and since $AN_r = AL_r + LN_r$, and LN_r is parallel to λ_r, we have

$$AN_r = a'\lambda_r + bn_r. \tag{63}$$

In the above equations, the coefficients a, a', b are scalars, at present

unknown. Now AL_r and AN_r are orthogonal to λ_r and n_r respectively and so

$$\begin{aligned} 0 &= \lambda_r AL_r = a\varepsilon(\lambda) + b\lambda_r n_r, \\ 0 &= n_r AN_r = a'\lambda_r n_r + b\varepsilon(n), \end{aligned} \tag{64}$$

and hence

$$a = -\varepsilon(\lambda) b\lambda_r n_r, \quad a' = -\varepsilon(n) b/(\lambda_r n_r). \tag{65}$$

Then from (62) and (63) we have (with uncertain sign), using (65),

$$\begin{aligned} \pm\frac{AL^2}{AN^2} &= \frac{a^2\varepsilon(\lambda) + 2ab\lambda_r n_r + b^2\varepsilon(n)}{a'^2\varepsilon(\lambda) + 2a'b\lambda_r n_r + b^2\varepsilon(n)} \\ &= \frac{\varepsilon(\lambda)\,(\lambda_r n_r)^2 - 2\varepsilon(\lambda)\,(\lambda_r n_r)^2 + \varepsilon(n)}{\varepsilon(\lambda)\,(\lambda_r n_r)^{-2} - 2\varepsilon(n) + \varepsilon(n)} \\ &= \frac{\varepsilon(n) - \varepsilon(\lambda)\,(\lambda_r n_r)^2}{\varepsilon(\lambda) - \varepsilon(n)\,(\lambda_r n_r)^2}\,(\lambda_s n_s)^2. \end{aligned} \tag{66}$$

But

$$\varepsilon(\lambda)\varepsilon(n)[\varepsilon(\lambda) - \varepsilon(n)\,(\lambda_r n_r)^2] = \varepsilon(n) - \varepsilon(\lambda)\,(\lambda_r n_r)^2, \tag{67}$$

and so the absolute value of the fraction in the last line of (66) is unity. Therefore

$$AL^2/AN^2 = (\lambda_r n_r)^2, \tag{68}$$

and so, by (61),

$$\boxed{S_0 = S\,|\,\lambda_r n_r\,|.} \tag{69}$$

This important projection formula expresses the connection between the 3-volume S_0 of the normal cross-section of a world tube and the 3-volume S of an oblique cross-section with unit normal n_r. The formula is easy to remember, for it is the direct analogue of the Euclidean formula for the projection of an area: $S_0 = S\cos\theta$.

If the tube has a null direction, there is no unit tangent vector λ_r, and indeed there is no normal section. The formula (69) fails. To get round this difficulty, we take first a tube for which the direction is not null, and of that tube we take two oblique sections with unit normal vectors n_r and $\bar{n}_r$; let their 3-volumes be S and $\bar{S}$, respectively. Then, by (69),

$$S_0 = S\,|\,\lambda_r n_r\,|, \quad S_0 = \bar{S}\,|\,\lambda_r \bar{n}_r\,|, \tag{70}$$

and hence

$$\frac{S}{\bar{S}} = \frac{|\,\lambda_r \bar{n}_r\,|}{|\,\lambda_r n_r\,|}. \tag{71}$$

Let t_r be any vector in the direction of λ_r; then (71) may be written

$$\frac{S}{\bar{S}} = \frac{|\, t_r \bar{n}_r \,|}{|\, t_r n_r \,|}. \tag{72}$$

Now, holding n_r and $\bar{n}_r$ fixed, let us change the direction of the tube into a null direction. We cannot use (70) in this process, since the components of λ_r tend to infinity. But we can use (72), and so obtain the formula

$$\boxed{S\,|\, t_r n_r \,| = \bar{S}\,|\, t_r \bar{n}_r \,|.} \tag{73}$$

This formula connects the 3-volumes S, $\bar{S}$ of cross-sections normal to n_r, $\bar{n}_r$ respectively, of a world tube, t_r being any tangent vector to the tube. This formula is valid for all tubes, whether null or not. It is the analogue of the Euclidean formula $S \cos\theta = \bar{S} \cos\bar{\theta}$.

§ 7. GREEN'S THEOREM AND THE EXPANSION OF WORLD TUBES

In Euclidean 3-space we have Green's theorem:[1]

$$\int \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) dxdydz = \int (lu + mv + nw) dS. \tag{74}$$

Here the integral on the left is taken through a volume (for simplicity a single integral sign is used here and later for multiple integrals, the multiplicity being otherwise obvious), and the integral on the right is taken over the surface bounding the volume; l, m, n are direction cosines of the normal, drawn outwards.

The usual proofs of this formula involve a certain amount of geometrical intuition, but it is not hard to generalise the argument from three to four or more dimensions. However, the indefinite character of the Minkowskian metric introduces a certain complication when we seek to generalise to space-time; since Green's theorem is fundamental in the mechanics of a continuum, it will be well to give a proof here.[2]

Let R be a region of space-time, bounded by a closed 3-space S (Fig. 6). Let n_r be the unit normal to S, drawn outwards. There is a subspace S' of S, formed of events on S where the normal has a null direction, and where in consequence a *unit* normal n_r does not exist; but in

[1] Also known as Gauss's theorem; cf. COURANT [1936, p. 384].

[2] For a more general treatment, see SYNGE and SCHILD [1952, p. 275].

general S' is two-dimensional[1], and may be neglected in taking an integral over S.

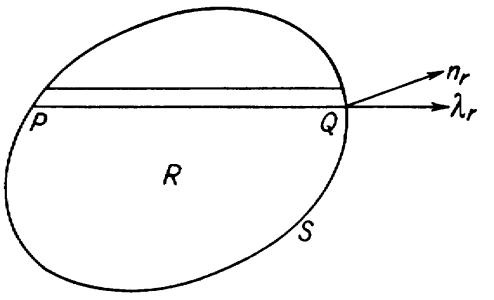


Fig. 6 – A region R of space-time, bounded by a closed 3-space S

Let λ_r be a unit vector at any event on S. What is the condition that λ_r should point *out of* R? By *out of* R we mean "on the same side of S as n_r points", because we stipulated that n_r should point outwards. To answer this question, we examine the scalar product $\lambda_r n_r$ as λ_r changes direction continuously. If λ_r coincides with n_r, then

$$\lambda_r n_r = n_r n_r = \varepsilon(n), \tag{75}$$

where $\varepsilon(n)$ is the indicator of n_r, chosen equal to $+1$ or -1 so as to make

$$\varepsilon(n) n_r n_r = 1. \tag{76}$$

Now $\lambda_r n_r$ vanishes when λ_r lies in the tangent 3-flat to S, and only then. Hence the condition (necessary and sufficient) that λ_r should point out of R is that $\lambda_r n_r$ should have the same sign as it has when λ_r coincides with n_r, i.e. the sign of $\varepsilon(n)$. So we have this result: λ_r *points out of* R *if*

$$\varepsilon(n) \lambda_r n_r > 0, \tag{77}$$

and λ_r *points into* R *if*

$$\varepsilon(n) \lambda_r n_r < 0. \tag{78}$$

Consider now the integral

$$I = \int_R \frac{\partial X}{\partial x_1} \, d\tau, \tag{79}$$

taken throughout the region R bounded by S; here X is any function of the coordinates, continuous and with continuous first derivatives, and $d\tau$ is the element of 4-volume:

$$d\tau = | \, dx_1 dx_2 dx_3 dx_4 \, |. \tag{80}$$

We want to change I into an integral over S, and to do this we split up R into thin tubes parallel to the x_1-axis. In Fig. 6 PQ is one such tube, starting at P and ending at Q, so that x_1 increases from P to Q.

[1] For the case where S' is three-dimensional, see IX–(386).

Integrating first along each tube, we obtain

$$I = \int | dx_2 dx_3 dx_4 | \int_P^Q \frac{\partial X}{\partial x_1} dx_1 = \int_S [X]_P^Q | dx_2 dx_3 dx_4 | . \tag{81}$$

Here $| dx_2 dx_3 dx_4 |$ is the 3-volume of the normal section of a tube, and so by (69) can be written in the following forms:

$$| dx_2 dx_3 dx_4 | = (| \lambda_r n_r | dS)_Q = (| \lambda_r n_r | dS)_P, \tag{82}$$

where $(dS)_P$ and $(dS)_Q$ are the 3-volumes cut off from S by the tube at P and Q respectively, and λ_r is the unit vector pointing in the direction of the x_1-axis, so that

$$\lambda_1 = 1, \quad \lambda_2 = \lambda_3 = \lambda_4 = 0, \quad \lambda_r n_r = n_1. \tag{83}$$

Since λ_r points out of R at Q, we have there, by (77),

$$0 < \varepsilon(n)\lambda_r n_r = \varepsilon(n) n_1, \quad | \lambda_r n_r | = \varepsilon(n) n_1, \tag{84}$$

and since λ_r points in at P, we have there, by (78),

$$0 > \varepsilon(n)\lambda_r n_r = \varepsilon(n) n_1, \quad | \lambda_r n_r | = -\varepsilon(n) n_1. \tag{85}$$

Thus (82) may be written

$$| dx_2 dx_3 dx_4 | = (\varepsilon(n) n_1 dS)_Q = -(\varepsilon(n) n_1 dS)_P, \tag{86}$$

and when we insert these values in (81) we get

$$I = \int_R \frac{\partial X}{\partial x_1} d\tau = \int_S \varepsilon(n) n_1 X dS, \tag{87}$$

dS being an element of 3-volume of S. We have thus converted the integral through R into an integral over S. Similarly, we have

$$\left.\begin{aligned} \int_R \frac{\partial X}{\partial x_2} d\tau &= \int_S \varepsilon(n) n_2 X dS, \\ \int_R \frac{\partial X}{\partial x_3} d\tau &= \int_S \varepsilon(n) n_3 X dS. \end{aligned}\right\} \tag{88}$$

Since the time-coordinate is different from the three space-coordinates, it would be rash to assume that there is a fourth member of the set (87), (88) of the same form with the suffix 4. It is in fact so, but let us prove it.

Remembering that $x_4 = ict$, we have

$$\int_R \frac{\partial X}{\partial x_4} d\tau = \int_R \frac{\partial X}{\partial x_4} \mid dx_1 dx_2 dx_3 dx_4 \mid = \int \mid dx_1 dx_2 dx_3 \mid \int_P^Q \frac{\partial X}{\partial x_4} \frac{dx_4}{i}, \quad (89)$$

where P and Q are the events at which a parallel to the t-axis, drawn in the sense of t increasing, enters and leaves R. Hence

$$\int_R \frac{\partial X}{\partial x_4} d\tau = -i \int_S [X]_P^Q \mid dx_1 dx_2 dx_3 \mid . \quad (90)$$

Now as in (82)

$$\mid dx_1 dx_2 dx_3 \mid = (\mid \lambda_r n_r \mid dS)_Q = (\mid \lambda_r n_r \mid dS)_P, \quad (91)$$

where λ_r is a unit vector drawn in the positive sense of the t-axis, so that

$$\lambda_1 = \lambda_2 = \lambda_3 = 0, \quad \lambda_4 = i, \quad \lambda_r n_r = i n_4. \quad (92)$$

Then by (77) we have at Q

$$0 < \varepsilon(n) \lambda_r n_r = i\varepsilon(n) n_4, \quad \mid \lambda_r n_r \mid = i\varepsilon(n) n_4, \quad (93)$$

and by (78) at P

$$0 > \varepsilon(n) \lambda_r n_r = i\varepsilon(n) n_4, \quad \mid \lambda_r n_r \mid = -i\varepsilon(n) n_4. \quad (94)$$

Therefore, if S_1 is that part of S where parallels to the t-axis (in the positive sense) emerge from R, and S_2 that part where they enter, we have from (90)

$$\begin{aligned} \int_R \frac{\partial X}{\partial x_4} d\tau &= -i \int_{S_1} X i\varepsilon(n) n_4 dS + i \int_{S_2} X(-i)\varepsilon(n) n_4 dS \\ &= \int_S \varepsilon(n) n_4 X dS. \end{aligned} \quad (95)$$

Collecting our results, we have *Green's theorem* in the form[1]

$$\boxed{\int_R \frac{\partial X}{\partial x_r} d\tau = \int_S \varepsilon(n) n_r X dS.} \quad (96)$$

Since no transformation of coordinates has been used in the above work, the result holds irrespective of the transformation properties of

[1] See IX–(386) for a form valid when S is a null 3-space.

X; it may be an invariant, the component of a vector, the component of a tensor, or none of these. We note the particularly important forms, corresponding to vector and tensor respectively, with a summation involved in each case:

$$\int_R \frac{\partial X_r}{\partial x_r} d\tau = \int_S \varepsilon(n) n_r X_r dS, \tag{97}$$

$$\int_R \frac{\partial X_{rs}}{\partial x_s} d\tau = \int_S \varepsilon(n) X_{rs} n_s dS. \tag{98}$$

We are to remember that n_r is the *outward* normal.

It may be remarked in passing that, since we cannot add vectors or tensors at different points in a curved space (in the sense that the sums so obtained are not themselves tensors), it is only the form (97) that can be taken over into the curved space-time of general relativity. In (97) the integrands are invariants with respect to the Lorentz transformation; the corresponding invariants in general relativity are $X^r{}_{|r}$ and $n_r X^r$, the stroke indicating covariant differentiation.

We shall now establish the following formula for the *expansion* of a tube of timelike world lines:

$$\frac{dS_0}{ds} = \frac{\partial \lambda_r}{\partial x_r} S_0. \tag{99}$$

Here S_0 is the 3-volume of the normal cross-section of the tube (assumed to be of infinitesimal section), ds is an element of separation along the world lines forming the tube, and λ_r the unit vector tangent to these world lines.

Consider first a tube of finite size; take a portion R of it, bounded by cross-sections, S_1 and S_2, not necessarily normal sections (Fig. 7). Applying Green's theorem in the form (97), we have

$$\int \varepsilon(n) \lambda_r n_r dS = \int_R \frac{\partial \lambda_r}{\partial x_r} d\tau. \tag{100}$$

The integral on the left is taken over the bounding 3-space of R, and n_r is the unit vector normal to this 3-space, drawn outwards. The sides of the tube give no contribution, since there we have $\lambda_r n_r = 0$.

Now make the section of the tube infinitesimal, and take S_1 and S_2 to be normal sections. Then $n_r = \lambda_r$ on S_2 and $n_r = -\lambda_r$ on S_1;

$\varepsilon(n) = -1$ on both S_1 and S_2, the tube being timelike. Hence (100) gives

$$S_2 - S_1 = \int_R \frac{\partial \lambda_r}{\partial x_r} d\tau. \qquad (101)$$

Let us now take S_2 close to S_1, and denote by ds the separation between them, measured along the tube. Put $S_2 - S_1 = dS_0$, and let S_0 stand for S_1 or S_2. Then (101) gives

$$dS_0 = \frac{\partial \lambda_r}{\partial x_r} S_0 ds. \qquad (102)$$

We have thus established the formula (99), which tells us how rapidly the cross-section of a thin tube is increasing. In kinematical language, it tells us how fast the proper 3-volume of a cloud of particles is increasing.

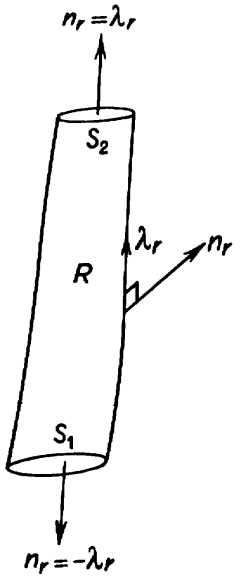


Fig. 7 – Expansion of a world tube

The formula (99) belongs mathematically to the geometry of a congruence of curves in space-time; this in turn is a special case of the geometry of a congruence of curves in Riemannian N-space (cf. EISENHART [1926], MCCONNELL [1928–29]).

§ 8. THE ENERGY TENSOR OF A CONTINUOUS MEDIUM

We have now prepared the mathematical tools needed for the transition from a discrete system to a continuum. To make this transition, we consider again a discrete system as discussed in VII–§ 4, but, to get the maximum of generality, let us include intrinsic angular momentum as treated in VII–§ 7. This means that we have before us an assembly of material particles, photons and internal impulses; we shall use the word *particle* to cover them all. Each has a straight world line between collisions, and each possesses (on each straight stretch of world line) a constant 4-momentum and a constant intrinsic angular momentum (skew-symmetric tensor). The 4-momentum points along the world line — that is important. At each collision there is conservation of 4-momentum and intrinsic angular momentum unless external impulses or impulsive torques act; in what follows we shall assume that there are no external impulses or impulsive torques. It is not assumed that the collisions are elastic.

In order to deal statistically with the system, we suppose the

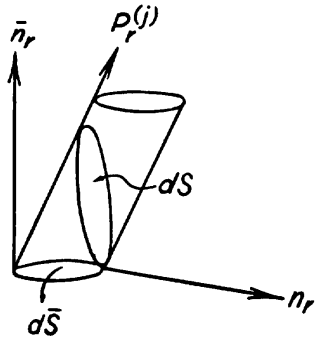


Fig. 8 – Calculation of the flux of 4-momentum across the 3-volume $d\bar{S}$

number of particles to be very great. Take a constant vector field defined by $\bar{n}_r$, this being a unit timelike vector. We shall consider those world lines which intersect an element of 3-volume $d\bar{S}$, normal to $\bar{n}_r$ (Fig. 8).

In the statistical assembly which we are substituting for the discrete system, there are particles with all possible 4-momenta. There are a certain number of particles with world lines intersecting $d\bar{S}$, each with a 4-momentum lying in a 4-cell $dM_1dM_2dM_3dM_4$ containing some specified M_r; this number might be written

$$f(M, x, \bar{n})d\bar{S} \mid dM_1dM_2dM_3dM_4 \mid, \tag{103}$$

f being a type of numerical density. It is, however, easier to write our formulae (and it is essentially equivalent) if we assume that all the particles may be put into classes 1, 2, 3, ... such that all those in class j have precisely the 4-momentum $M_r^{(j)}$; the number of classes is of course very great. Then the number of particles in class j with world lines intersecting $d\bar{S}$ may be written in the form

$$\bar{N}^{(j)}d\bar{S}, \tag{104}$$

$\bar{N}^{(j)}$ being a function of $\bar{n}_r$ and x_r. Here $\bar{N}^{(j)}$ is the numerical density of class j (cf. p. 261) relative to an observer who takes $\bar{n}_r$ for time-axis. Then, since all particles are propagated from past to future, the flux of 4-momentum of class j across $d\bar{S}$ is [cf. VII–(14)]

$$M_r^{(j)}\,\bar{N}^{(j)}\,d\bar{S}. \tag{105}$$

(The summation convention does not apply to the bracketted suffix j).

Let S be a closed 3-space in space-time, with outward unit normal n_r. Take a 3-element dS of S, and through it draw the tube with the direction of $M_r^{(j)}$. Let $d\bar{S}$ be the 3-volume of the section of this tube normal to $\bar{n}_r$ (Fig. 8). All the world lines of particles of class j which intersect $d\bar{S}$ also intersect dS, except those which escape from the tube on account of collisions occurring in the 4-volume between dS and $d\bar{S}$. There are also particles which leak in through the sides of the tube, and are converted into class j by collisions occurring in that 4-volume. But since the 4-volume is an infinitesimal of the fourth order, while dS and $d\bar{S}$ are of the third order for a thin tube, collisions in the

4-volume may be neglected, and so we have

$$\begin{aligned}&\text{Flux of 4-momentum of class } j \text{ across } dS\\ &= \eta\,(\text{flux of 4-momentum of class } j \text{ across } d\bar{S}),\end{aligned} \tag{106}$$

where η is $+1$ or -1; this factor η is essential because, by VII–(14), a contribution to the flux across dS is to be added or subtracted according as the sense of propagation across dS is in the sense of n_r or in the reverse sense.

Let $P_r^{(j)}$ be a vector drawn in the sense of propagation, i.e. into the future. Its direction is that of $M_r^{(j)}$, except for internal impulses of the attractive type, when the directions are opposite. In any case we have, by (77) and (78),

$$\left.\begin{aligned}\eta &= 1 \quad \text{if } \varepsilon(n)P_r^{(j)}n_r > 0,\\ \eta &= -1 \text{ if } \varepsilon(n)P_r^{(j)}n_r < 0,\end{aligned}\right\} \tag{107}$$

and so

$$\eta\varepsilon(n)P_r^{(j)}n_r > 0. \tag{108}$$

By (73) we have

$$d\bar{S}\,|\,P_r^{(j)}\bar{n}_r\,| = dS\,|\,P_r^{(j)}n_r\,|, \tag{109}$$

and so (105) and (106) give

$$\begin{aligned}&\text{Flux of 4-momentum of class } j \text{ across } dS\\ &= \eta M_r^{(j)}\bar{N}^{(j)}dS\,|\,P_s^{(j)}n_s\,|\,|\,P_t^{(j)}\bar{n}_t\,|^{-1}\\ &= \varepsilon(n)M_r^{(j)}\bar{N}^{(j)}P_s^{(j)}n_s dS\,|\,P_t^{(j)}\bar{n}_t\,|^{-1}.\end{aligned} \tag{110}$$

If we now add together the fluxes for all classes across dS, we get

$$\boxed{\begin{aligned}&\text{Total flux of 4-momentum across } dS\\ &= \varepsilon(n)T_{rs}n_s dS,\end{aligned}} \tag{111}$$

where

$$T_{rs} = \sum_j \frac{\bar{N}^{(j)}M_r^{(j)}P_s^{(j)}}{|\,P_t^{(j)}\bar{n}_t\,|}. \tag{112}$$

It might appear that T_{rs} depended on the choice of $\bar{n}_r$, but that is not so; for the number (104) is clearly independent of the way in which we slice the tube across, and so, by (73), the quotient

$$\frac{\bar{N}^{(j)}}{|\,P_t^{(j)}\bar{n}_t\,|} \tag{113}$$

is independent of the choice of $\bar{n}_r$, provided it is timelike. We have avoided taking a normal section of the tube in order to include the case where $P_r^{(j)}$ is a null vector, for then there is no normal section.[1]

We note that since $M_r^{(j)}$ and $P_r^{(j)}$ have either the same or opposite directions, the tensor T_{rs} is *symmetric*:

$$T_{rs} = T_{sr}. \tag{114}$$

This symmetric tensor T_{rs} is fundamental in the theory of a continuous medium; it is called the *energy tensor* for reasons which will appear later. The derivation of the energy tensor in the above statistical way from the case of a discrete system is unconventional. The reader who does not wish to bother with this derivation can get on very well without it, simply accepting the existence of a symmetric energy tensor T_{rs}. For purposes of physical interpretation he must accept the statement (111), and he must also accept the equation (117) below as a hypothesis. We shall however proceed to *prove* (117) on the basis of hypotheses already made for a discrete system.

Since we suppose that no external impulses act on the system, we have the closed law of conservation of 4-momentum, as stated on p. 215. By (111) this gives

$$\int_S \varepsilon(n) T_{rs} n_s dS = 0, \tag{115}$$

for any closed 3-space S in space-time, and hence, by Green's theorem in the form (98),

$$\int_R \frac{\partial T_{rs}}{\partial x_s} \, d\tau = 0, \tag{116}$$

this integral being taken over *any* portion of space-time. From this it follows that

$$\boxed{\frac{\partial T_{rs}}{\partial x_s} = 0} \tag{117}$$

[1] For some purposes it is more useful not to employ a single constant timelike vector field $\bar{n}_t$, but rather a set of timelike vector fields $\bar{n}_t^{(j)}$, one for each class, these fields varying from event to event. For example, for material particles, one might choose $\bar{n}_t^{(j)}$ in the direction of the 4-momentum $M_t^{(j)}$. This generalisation of the procedure makes little change, the formula (112) being modified only by writing $\bar{n}_t^{(j)}$ instead of $\bar{n}_t$. We shall use this modification on pp. 288 and 299.

everywhere. For if there were any event P at which this is not true, then there would be one component of $\partial T_{rs}/\partial x_s$ which is positive or negative throughout some small domain enclosing P, and so the integral (116) taken throughout this domain would not vanish, as it must. We express (117) in words by saying that *the divergence of the energy tensor vanishes*; we may call (117) the *equation of conservation of 4-momentum.*

Intrinsic angular momentum does not figure in the above argument, but we can discuss it in the same way. We may speak of a class j having a propagation vector $P_r^{(j)}$ and an intrinsic angular momentum $H_{rs}^{(j)}$. Then instead of (111) we obtain

$$\text{Total flux of intrinsic angular momentum across } dS$$
$$= \varepsilon(n) U_{rst} n_t dS, \tag{118}$$

where

$$U_{rst} = \sum_j \frac{\bar{N}^{*(j)} H_{rs}^{(j)} P_t^{(j)}}{| P_m^{(j)} \bar{n}_m |}, \tag{119}$$

$\bar{N}^{*(j)} d\bar{S}$ being the number of particles in class j with world lines intersecting $d\bar{S}$. We may call this tensor U_{rst} the *intrinsic angular momentum flux tensor.* It is skew-symmetric in rs:

$$U_{rst} = - U_{srt}. \tag{120}$$

By the closed law of intrinsic angular momentum VII–(232), we have, in the absence of external impulsive torques,

$$\int_S \varepsilon(n) U_{rst} n_t dS = 0, \tag{121}$$

and hence

$$\frac{\partial U_{rst}}{\partial x_t} = 0. \tag{122}$$

§ 9. THE PHYSICAL MEANING OF THE ENERGY TENSOR

The energy tensor T_{rs} has 16 components, but on account of symmetry only 10 of these are independent. We shall now use (111) to interpret these ten independent components physically; in other words, we shall describe ideal experiments by which they could be measured, if such ideal experiments could be actually carried out.

First, let us take the unit vector n_r of (111) in the direction of the

observer's time-axis, so that

$$n_\varrho = 0, \quad n_4 = i, \quad \varepsilon(n) = -1. \tag{123}$$

The 3-element dS is then a portion of the observer's instantaneous space; we may think of it as a little box of volume dS which is suddenly created. The flux of 4-momentum across dS is then simply the total 4-momentum of all those particles which happen to find themselves inside the little box at a certain time t, and we can express (111) as follows:

Total 4-momentum in box of volume dS at time t

$$= -iT_{r4}\,dS. \tag{124}$$

If we now use the relations (9) connecting 4-momentum M_r with relative momentum and relative energy, we can split the above statement into two:

$$\text{Relative density of relative momentum} = -icT_{\varrho 4}. \tag{125}$$

$$\text{Relative density of relative energy} = -c^2T_{44}. \tag{126}$$

We have here interpreted four of the ten independent components of T_{rs} in terms of the momentum and energy which the observer traps inside a little box which he suddenly creates, assuming that he has at his disposal some way of measuring the momentum and energy so trapped.

We note that the relative density of relative energy appears as the 44-component of a tensor of the second order. This explains the factor γ^2 in equation (7), or the corresponding equations relating to energy instead of mass; a Lorentz transformation applied to the component A_4 of a vector A_r brings in a factor γ, but when applied to the component A_{44} of a tensor A_{rs} it brings in a factor γ^2.

Now let the observer take a thin rigid loop enclosing an area dS_2 and hold it fixed in his frame of reference for a time dt. The history of dS_2 for time dt is a 3-element of space-time with 3-volume $dS = dS_2 c\,dt$, and the unit normal n_r to dS in space-time has for its first three components (n_ϱ) the components of the unit normal in space to the plane of the loop; we have further $n_4 = 0$, $\varepsilon(n) = 1$, since it is spacelike. Application of (111) then gives

Flux of 4-momentum through loop of area dS_2 in time dt

$$= T_{r\sigma}n_\sigma dS_2 c\,dt. \tag{127}$$

Using (9), we can split this into the two statements:

$$\text{Flux of relative momentum per unit time per unit area normal to } n_\sigma = c^2 T_{\varrho\sigma} n_\sigma. \tag{128}$$

$$\text{Flux of relative energy per unit time per unit area normal to } n_\sigma = -ic^3 T_{4\sigma} n_\sigma. \tag{129}$$

Just as in the ordinary kinetic theory of gases, flux of momentum may be interpreted as *stress*; on that account T_{rs} is sometimes called the *stress-momentum-energy tensor*, but this name is clumsy and we shall continue to call it the *energy tensor*.

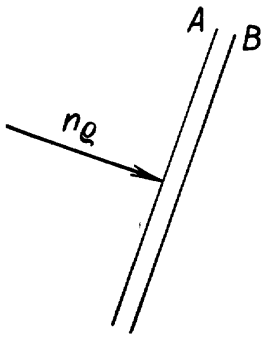


Fig. 9 – Double target for the measurement of $T_{r\sigma} n_\sigma$

It may be justly objected that an observer has no way of measuring fluxes of momentum and energy through a loop. To meet this objection, we may replace the loop by a double target, consisting of two small parallel plates, A and B, close to one another (Fig. 9). Let n_ϱ be the normal in the sense from A to B. Then the flux, in the sense used above, whether of momentum or energy, is in fact

$$(\text{Amount received by } A) - (\text{amount received by } B).$$

If the plates absorb the particles which fall on them, these amounts can be measured according to the methods of §§ 3 and 5, and so in this way $T_{\varrho\sigma}$ and $T_{4\sigma}$ can be measured mechanically and thermally. However, the interpretation in terms of flux through a loop is easier to think about, and so we shall retain that interpretation.

Let us now collect our results:

$$T_{\varrho\sigma} n_\sigma = c^{-2} \times \text{flux of relative momentum per unit time per unit area normal to } n_\sigma, \tag{130}$$

or

$$\left.\begin{matrix} T_{11}, T_{21}, T_{31} \\ T_{12}, T_{22}, T_{32} \\ T_{13}, T_{23}, T_{33} \end{matrix}\right\} = c^{-2} \times \text{flux of relative momentum per unit time per unit area normal to the axis of } \begin{cases} x_1, \\ x_2, \\ x_3. \end{cases} \tag{131}$$

$$T_{4\sigma} n_\sigma = ic^{-3} \times \text{flux of relative energy per unit time per unit area normal to } n_\sigma, \tag{132}$$

or

$$\left.\begin{matrix} T_{41} \\ T_{42} \\ T_{43} \end{matrix}\right\} = ic^{-3} \times \text{flux of relative energy per unit time per unit area normal to the axis of} \left\{\begin{matrix} x_1, \\ x_2, \\ x_3. \end{matrix}\right. \quad (133)$$

$$T_{\varrho 4} = ic^{-1} \times \text{relative density of relative momentum.} \quad (134)$$

$$T_{44} = -c^{-2} \times \text{relative density of relative energy.} \quad (135)$$

We note the two interpretations of $T_{\varrho 4}$ $(= T_{4\varrho})$. These components of the energy tensor may be regarded either as flux of energy or density of momentum; we have in fact

$$\text{Flux of energy} = c^2 \times \text{density of momentum.} \quad (136)$$

This is one of the surprises of relativity, for this formula links together quantities which are independent in Newtonian physics. (We check that each side of (136) has the dimensions $[MT^{-3}]$.)

Had we accepted (111), but assumed an unsymmetric energy tensor $(T_{rs} \neq T_{sr})$, the interpretations (133) and (134) might have been made, but we would not have got the relationship (136).

§ 10. THE ENERGY TENSOR FOR AN INCOHERENT STREAM OF MATERIAL PARTICLES

Consider now the case of a stream of material particles. We assume that they do not collide with one another, for if they did they would move in many directions and so cease to constitute a single stream. We use the word *incoherent* to imply the absence of interactions. For an incoherent stream the energy tensor takes a particularly simple form, as we shall now show.

It was pointed out that the quotient (113), occurring in the expression (112) for the energy tensor of a general medium, is independent of the choice of the timelike unit vector $\bar{n}_r$. In § 8 we made $\bar{n}_r$ a constant vector field, but that was a matter of convenience and by no means essential. Let us now choose $\bar{n}_r = \lambda_r$, where λ_r is the 4-velocity of a particle, i.e. the unit tangent to the world line of the stream passing through the event in which we are interested. For the propagation vector we have $P_r = \lambda_r$, and so

$$| P_t \bar{n}_t | = 1. \quad (137)$$

Let us suppose for simplicity that all the particles have the same proper mass, so that we have only one class to consider $(j = 1)$ and

$\bar{N}^{(j)}$ becomes (for our choice of $\bar{n}_r$) the proper numerical density N_0 of the particles. Then (112) gives

$$T_{rs} = N_0 M_r \lambda_s. \tag{138}$$

Here M_r is the 4-momentum of a particle, so that $M_r = m\lambda_r$ where m is its proper mass. Thus $N\,m$ is the proper density of proper mass μ as in (6), and so (138) gives $T_{rs} = \mu\lambda_r\lambda_s$.

It is easy to extend this to the case where the particles have different proper masses, by dividing them into classes, each class with the same proper mass. We may therefore state this result: *For an incoherent stream of material particles the energy tensor is*

$$T_{rs} = \mu\lambda_r\lambda_s, \tag{139}$$

where μ is the proper density of proper mass and λ_r the 4-velocity.

Let us apply to (139) the equation of conservation (117). We get

$$\frac{\partial}{\partial x_s}(\mu\lambda_r\lambda_s) = \lambda_r \frac{\partial}{\partial x_s}(\mu\lambda_s) + \mu \frac{\partial\lambda_r}{\partial x_s}\lambda_s = 0. \tag{140}$$

Since λ_r is a unit vector, we have

$$\lambda_r \frac{\partial\lambda_r}{\partial x_s} = 0, \tag{141}$$

and so multiplication of (140) by λ_r gives

$$\frac{\partial}{\partial x_s}(\mu\lambda_s) = 0. \tag{142}$$

On account of this, (140) reduces to

$$\frac{\partial\lambda_r}{\partial x_s}\lambda_s = 0, \tag{143}$$

or, equivalently,

$$\frac{d\lambda_r}{ds} = 0, \quad \lambda_r = \text{const.} \tag{144}$$

This last result merely confirms what we know already — in an incoherent stream the world lines of the particles are straight. The equation (142) is more interesting. If we introduce the 3-velocity u_ϱ, writing as in v–(59)

$$\lambda_\varrho = \gamma u_\varrho/c, \quad \lambda_4 = i\gamma, \quad \gamma = (1 - u^2/c^2)^{-\frac{1}{2}}, \tag{145}$$

then (142) becomes

$$\frac{\partial}{\partial t}(\gamma\mu) + \frac{\partial}{\partial x_\varrho}(\gamma\mu u_\varrho) = 0. \tag{146}$$

This is immediately recognised as the equation which, in hydrodynamics, is commonly called the *equation of continuity*; the only change from Newtonian form consists in substituting the relative density of proper mass $\gamma\mu = \mu^*$ [cf. (6)] for Newtonian density.

§ 11. EIGEN VALUES OF THE ENERGY TENSOR

There is no difficulty about representing a 4-vector in a space-time diagram; it looks just like an ordinary vector. But a tensor is a different matter, and we have now to discuss how the energy tensor T_{rs} is to be represented.

In Newtonian physics we are accustomed to symmetric tensors of the second order; they are 3×3 symmetric matrices, the elements of which transform tensorially under orthogonal transformations of the coordinates. Two familiar examples are (i) the components of stress in an elastic body, and (ii) moments and products of inertia (the products being prefixed with minus signs). In both these cases we accustomed to think of the quantities involved in terms of quadric surfaces (the stress-quadric and the momental ellipsoid). The reason for this is that, if we write the equation of a central quadric in the form

$$a_{\varrho\sigma}x_\varrho x_\sigma = 1 \quad (a_{\varrho\sigma} = a_{\sigma\varrho}),$$

then the coefficients are the components of a symmetric tensor. Thus a quadric surface is a geometrical representation of a symmetric tensor, and hence a geometrical representation of stress and of moments and products of inertia.

But it is still more important that a quadric surface may be represented by an orthogonal triad (its principal axes) and three numbers associated with them (the squares of its semi-axes or their reciprocals). In this way we are led to the concepts of principal stresses and principal moments of inertia; indeed, in dynamics we rarely use products of inertia at all — they are eliminated by the use of principal axes.

This suggests that we should seek similar representations of the energy tensor T_{rs} in space-time. Let T_{rs} be evaluated at some event P, and let ξ_r be Minkowskian coordinates measured relative to P as origin.

Then the invariant equations

$$T_{rs}\xi_r\xi_s = \varepsilon, \quad \varepsilon = \pm 1, \tag{147}$$

give a pair of quadric 3-spaces in space-time. It may be that one choice of ε gives no real 3-space; that will be the case if the left hand side of (147) is positive for all choices of the vector ξ_r, or if it is negative for all such choices. For that reason it is well to include the ambiguous sign. The quadrics (or quadric) (147) represent the energy tensor at the event P, in the same way that the stress-quadric represents stress, or the momental ellipsoid represents moments and products of inertia.

We now seek an orthogonal tetrad of principal directions and four numbers associated with them. The indefinite character of the Minkowskian metric introduces complications, so that we are faced with something more than a mere generalisation from three to four dimensions.

Consider the equations

$$T_{rs}V_s = XV_r; \tag{148}$$

here T_{rs} is the energy tensor at some given event, and V_r, X are to be found to satisfy these equations, V_r being a vector and X an invariant, so that the equations are invariant under Lorentz transformation. If the equations possess a solution V_r (other than the trivial solution $V_r = 0$), then X must satisfy the determinantal equation

$$\det(T_{rs} - X\delta_{rs}) = 0. \tag{149}$$

This is an algebraic equation of the fourth degree, and so has four roots $X^{(1)}$, $X^{(2)}$, $X^{(3)}$, $X^{(4)}$, but we have no guarantee that they are real. If they are real, we call them the *eigen values* of the energy tensor, and the four corresponding vectors given by (148) the *prinicpal directions* of the energy tensor, any solution for V_r being of course undetermined to within a scalar factor.

It is easy to establish *the orthogonality of the principal directions*. Let $X^{(1)}$ and $X^{(2)}$ be two roots of (149), and let $V_r^{(1)}$ and $V_r^{(2)}$ be principal vectors corresponding to them, so that

$$T_{rs}V_s^{(1)} = X^{(1)}V_r^{(1)}, \quad T_{rs}V_s^{(2)} = X^{(2)}V_r^{(2)}. \tag{150}$$

Hence, on account of the symmetry of T_{rs}, we have

$$(X^{(1)} - X^{(2)})V_r^{(1)}V_r^{(2)} = 0. \tag{151}$$

If the eigen values are distinct ($X^{(1)} \neq X^{(2)}$), this implies the orthogonality of $V_r^{(1)}$ and $V_r^{(2)}$.

In the case of equal eigen values (degeneracy) there is a partial indeterminacy in the principal directions, corresponding to the partial indeterminacy in ordinary geometry of the principal axes of a quadric surface when the lengths of two of these axes are equal. Such cases may be approached as limits; it is simpler in general arguments to suppose the eigen values distinct.

We see then, associated with a given symmetric tensor T_{rs}, an orthogonal tetrad of vectors (the *principal tetrad*), an eigen value being associated with each of them. The space-time picture so suggested is analogous to the picture of an elastic solid in 3-space, where we have three principal axes of stress with three principal stresses associated with them. But the Minkowskian character of space-time makes a difference; of the four principal directions, one is timelike and so is essentially different from the other three.

We now turn to the important question of the *reality of the eigen values*, the roots of the determinantal equation (149). Here it is well to be as general as possible and not tie the argument to the approach so far adopted. Let us not for the present think of T_{rs} as given by the formula (112) in terms of elementary particles, but rather as any symmetric tensor subject only to the conditions we now see fit to impose on it.

The following theorem will be proved: *If* T_{rs} *is such that*

$$T_{rs}p_r p_s > 0 \tag{152}$$

for all vectors p_r *which are timelike or null, then* T_{rs} *has four real eigen values.*

Before proceeding with the proof of this, we note that if we take p_r for time-axis, so that $p_\varrho = 0$, $p_4 = i$, then (152) gives

$$-T_{44} > 0. \tag{153}$$

By (126) this reads: *the relative density of relative energy is positive.* But note that (152) is asserted to hold for null vectors as well as timelike vectors, and (153) does not imply (152) on that account. In fact, the positive character of energy density, although implied by (152), is not in itself a sufficient assumption to establish the reality of the eigen values.

To prove the theorem, we proceed as follows. First take p_r to be a null vector. Then

$$l_\varrho = ip_\varrho/p_4 \tag{154}$$

is a unit 3-vector in the observer's space, for we have

$$l_\varrho l_\varrho - 1 = - p_\varrho p_\varrho/(p_4)^2 - 1 = - p_r p_r/(p_4)^2 = 0, \tag{155}$$

since $p_r p_r = 0$. Now

$$\begin{aligned} T_{rs} p_r p_s &= T_{\varrho\sigma} p_\varrho p_\sigma + 2T_{\varrho 4} p_\varrho p_4 + T_{44} p_4^2 \\ &= - p_4^2 (T_{\varrho\sigma} l_\varrho l_\sigma + 2iT_{\varrho 4} l_\varrho - T_{44}). \end{aligned} \tag{156}$$

Here p_4^2 is negative, since p_4 is a pure imaginary, and so (152) implies

$$T_{\varrho\sigma} l_\varrho l_\sigma + 2iT_{\varrho 4} l_\varrho - T_{44} > 0 \tag{157}$$

for all unit 3-vectors l_ϱ.

Next we take p_r timelike and of unit magnitude. It is then the 4-velocity of a particle, and in terms of the 3-velocity u_ϱ we have

$$p_\varrho = \gamma u_\varrho/c, \quad p_4 = i\gamma, \quad \gamma = (1 - u^2/c^2)^{-\frac{1}{2}}. \tag{158}$$

Then

$$T_{rs} p_r p_s = \gamma^2 (T_{\varrho\sigma} u_\varrho u_\sigma/c^2 + 2iT_{\varrho 4} u_\varrho/c - T_{44}) \tag{159}$$

We introduce a unit 3-vector l_ϱ in the direction of the 3-velocity u_ϱ or in the reverse direction, so that

$$u_\varrho = u' l_\varrho, \quad u' = \pm u, \tag{160}$$

and (159) takes the form

$$\begin{aligned} T_{rs} p_r p_s &= \gamma^2 (T_{\varrho\sigma} l_\varrho l_\sigma u'^2/c^2 + 2iT_{\varrho 4} l_\varrho u'/c - T_{44}), \\ \gamma^2 &= (1 - u'^2/c^2)^{-1}. \end{aligned} \tag{161}$$

We have here a function of l_ϱ and u', with u' in the range $-c < u' < c$, and the condition (152) tells us that it is positive for all choices of l_ϱ and u'. Let us write (161) in the form (with $u'/c = x$)

$$T_{rs} p_r p_s = \frac{Ax^2 + 2Bx + C}{1 - x^2}, \tag{162}$$

with A, B and C functions of l_ϱ. By (152) this fraction is positive for $-1 < x < 1$. What happens when $x \to 1$? The denominator tends to zero but the numerator does not — it approaches the expression in (157). At this limit then, the fraction (162) tends to infinity, and the same is easily seen to be true for $x \to -1$.

We have then a function of x which is positive in the range $-1 < x < 1$ and tends to infinity at the ends of the range, being finite elsewhere in the range. It therefore has a *minimum* corresponding to some value of x, i.e. some value of u'. This means that having chosen

a direction l_ϱ in the observer's space, we find a speed u (with velocity along l_ϱ or in the reverse direction) which minimises $T_{rs}p_r p_s$. We can plot these values of u on a unit sphere. There will exist some direction giving a minimum value to this minimum u. The corresponding velocity u_ϱ *defines a 4-velocity* p_r *which, in fact, minimises* $T_{rs}p_r p_s$. Since p_r is a timelike unit vector, we have $p_r p_r = -1$, and so that 4-vector p_r which minimises $T_{rs}p_r p_s$ satisfies[1]

$$T_{rs}V_s = X^{(4)}V_r, \tag{163}$$

where $X^{(4)}$ is a real scalar (we easily see that it is negative). *Thus* (152) *implies the existence of a real negative eigen value with a timelike principal direction associated with it.* The essence of the rather long argument by which we have established this is, of course, that there exists some timelike unit vector p_r minimising $T_{rs}p_r p_s$.

To complete the proof of the theorem connected with (152) we have to speak of the other eigen values. This is easy. We use a frame of reference having for time axis the vector V_r satisfying (163). Then $V_\varrho = 0$, $V_4 = i$, and (163) gives

$$T_{\varrho 4} = 0, \quad T_{44} = X^{(4)}. \tag{164}$$

The matrix of T_{rs} is now in the form

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} & 0 \\ T_{21} & T_{22} & T_{23} & 0 \\ T_{31} & T_{32} & T_{33} & 0 \\ 0 & 0 & 0 & T_{44} \end{pmatrix} \tag{165}$$

Here the 3×3 block of $T_{\varrho\sigma}$ behaves, under rotation of space axes, like the stress tensor in elasticity, and by a suitable rotation of space axes we can change it to diagonal form. Under such a rotation $T_{\varrho 4}$ behaves as a 3-vector, and so this rotation leaves in the matrix the zeros already present in (165). After applying this rotation, the equation (149) for eigen values reads

$$\begin{vmatrix} T_{11} - X & 0 & 0 & 0 \\ 0 & T_{22} - X & 0 & 0 \\ 0 & 0 & T_{33} - X & 0 \\ 0 & 0 & 0 & T_{44} - X \end{vmatrix} = 0. \tag{166}$$

It is clear that there are four real eigen values, and so our theorem is proved.

[1] We are in fact minimising the quotient $-T_{rs}p_r p_s/(p_n p_n)$.

Having established this result without imposing on T_{rs} any conditions except (152), and the condition $T_{rs} = T_{sr}$, let us turn back to the expression (112) for T_{rs} and see whether the condition (152) is actually satisfied by it. We introduce a factor $k^{(j)}$ for the class j such that $k^{(j)} = 1$ for material particles, photons and repulsive internal impulses, and $k^{(j)} = -1$ for attractive internal impulses. Then we may write for the propagation vector

$$P_r^{(j)} = k^{(j)} M_r^{(j)}, \tag{167}$$

and (112) reads

$$T_{rs} = \sum_j k^{(j)} \bar{N}^{(j)} \frac{M_r^{(j)} M_s^{(j)}}{| M_t^{(j)} \bar{n}_t |}. \tag{168}$$

Then for any vector p_r, and in particular for a timelike one,

$$T_{rs} p_r p_s = \sum_j k^{(j)} \bar{N}^{(j)} \frac{(M_r^{(j)} p_r)^2}{| M_t^{(j)} \bar{n}_t |}. \tag{169}$$

This is certainly positive if $k^{(j)} = 1$ for all j, and it may be positive otherwise, provided the terms with positive $k^{(j)}$ outbalance those with negative $k^{(j)}$. *Thus the condition* (152) ***leading to real eigenvalues is certainly satisfied by the energy tensor*** (112) ***if only material particles, photons and repulsive internal impulses are present, and will be satisfied even if attractive internal impulse are present, provided their negative contribution to*** (169) *is outbalanced by the other terms.*

We mention here for future reference (§ 16) an obvious fact: *If* T_{rs} *is such that*

$$T_{rs} T_{rt} p_s p_t < 0 \tag{170}$$

for all timelike vectors p_r, *then the 4-vector*

$$F_r = -T_{rs} p_s \tag{171}$$

is timelike for all timelike vectors p_r. This is evident, since (170) makes $F_r F_r < 0$.

With T_{rs} as in (168), we find

$$T_{rs} T_{rt} p_s p_t = \sum_j \sum_h k^{(j)} k^{(h)} \bar{N}^{(j)} \bar{N}^{(h)} \frac{M_r^{(j)} M_r^{(h)} (M_s^{(j)} p_s)(M_t^{(h)} p_t)}{| M_u^{(j)} \bar{n}_u | \, | M_v^{(h)} \bar{n}_v |}, \tag{172}$$

the summations for j and h being over the classes. Confining our attention to the case where $k^{(j)} = 1$ for all classes, we see that the condition (170) is satisfied, because $M_r^{(j)} M_r^{(h)}$ is negative (scalar product of two future-pointing vectors) and $M_s^{(j)} p_s$, $M_t^{(h)} p_t$ have the same sign, negative if p_r points into the future and positive if it points

into the past. Thus the vector F_r of (171) is timelike if attractive internal impulses are absent; it will still be timelike in the presence of attractive impulses if their effect is outbalanced in (172).

As the simplest of all illustrative cases, consider an incoherent stream of particles with T_{rs} as in (139). Then

$$T_{rs}p_r p_s = \mu(p_r\lambda_r)^2,$$

and so (152) is satisfied provided μ is positive. The eigen values are $(0, 0, 0, -\mu)$, as we see at once on taking the time axis along λ_r in (139). As for (170), we have

$$T_{rs}T_{rt}p_s p_t = \mu^2\lambda_r\lambda_r(p_s\lambda_s)^2 = -\mu^2(p_s\lambda_s)^2,$$

and so (170) is satisfied. The vector F_r has the direction of λ_r:

$$F_r = -\mu\lambda_r(p_s\lambda_s).$$

§ 12. MEAN DENSITY, MEAN VELOCITY AND STRESS

We shall henceforth assume that the energy tensor T_{rs} has four real eigen values, and that it can be represented in space-time by an orthogonal tetrad of unit vectors, with each of which an invariant is associated. Let us write $V_r^{(\sigma)}$ for the three spacelike vectors ($\sigma = 1, 2, 3$) and λ_r for the timelike one. Let us write $X^{(\sigma)}$ for the three roots of (149) associated with $V_r^{(\sigma)}$, and $X^{(4)}$ for the root associated with λ_r (Fig. 10). Then $X^{(\tau)}$ are the four roots of

$$\det(T_{rs} - X\delta_{rs}) = 0, \tag{173}$$

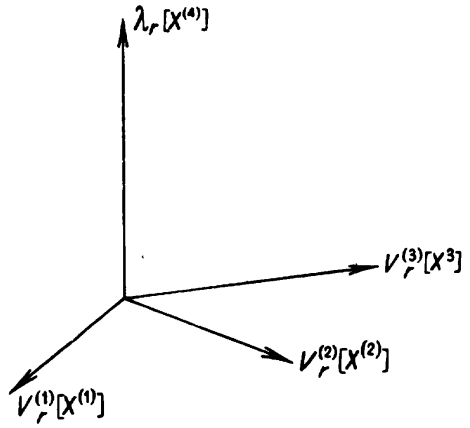


Fig. 10 – Representation in space-time of the tensor T_{rs} by the unit orthogonal tetrad $(V_r^{(\sigma)}, \lambda_r)$ and the four invariants $X^{(s)}$

and we have also the equations

$$\begin{aligned} T_{rs}V_s^{(\tau)} &= X^{(\tau)}V_r^{(\tau)}, \\ T_{rs}\lambda_s &= X^{(4)}\lambda_r, \end{aligned} \tag{174}$$

there being no summation for the bracketted suffix.

On the basis of Newtonian analogy, we expect the medium under discussion to possess a density, a velocity, and a stress tensor; they appear in the relativistic theory as an invariant, a 4-vector and a symmetric tensor of the second order, re-

spectively. We proceed to define these things in terms of the eigen values and principal directions of T_{rs}, justifying the definitions a little later.

We define the *mean mass-density* μ of the medium (or its *macroscopic density*, or briefly its *density*) by

$$\mu = -X^{(4)}, \tag{175}$$

$X^{(4)}$ being, as indicated above, that eigen value of T_{rs} which corresponds to a timelike principal direction. The *mean energy-density* is then μc^2.

We define the *mean velocity 4-vector* of the medium (or, briefly, its *4-velocity*) to be λ_r, that is, the unit vector (pointing into the future, of course) along the timelike principal direction. Then, as in v–(59), the 3-velocity u_ϱ satisfies

$$\gamma u_\varrho/c = \lambda_\varrho, \quad i\gamma = \lambda_4, \quad \gamma = (1 - u^2/c^2)^{-\frac{1}{2}}. \tag{176}$$

Finally, we define the *stress tensor* S_{rs} by

$$S_{rs} = -c^2(X^{(1)} V_r^{(1)} V_s^{(1)} + X^{(2)} V_r^{(2)} V_s^{(2)} + X^{(3)} V_r^{(3)} V_s^{(3)}). \tag{177}$$

We shall now prove that the energy tensor may be expressed in terms of the quantities just defined by the formula

$$\boxed{T_{rs} = \mu\lambda_r\lambda_s - c^{-2}S_{rs}.} \tag{178}$$

To establish this formula, we note that by (175) and (177) it is equivalent to

$$T_{rs} = X^{(1)} V_r^{(1)} V_s^{(1)} + X^{(2)} V_r^{(2)} V_s^{(2)} + X^{(3)} V_r^{(3)} V_s^{(3)} - X^{(4)} \lambda_r \lambda_s. \tag{179}$$

This is a tensor equation, and so it holds in all Galileian frames of reference if it holds in one. We introduce, then, the *principal frame of reference* at the event with which we are concerned, i.e. the frame of reference having λ_r for time-axis, and we take the space-axes in the principal directions $V_r^{(\sigma)}$. Then we have, at the event,

$$\left.\begin{aligned} &V_1^{(1)} = 1, \quad V_2^{(2)} = 1, \quad V_3^{(3)} = 1, \quad \lambda_4 = i, \\ &\text{all other components of } V_r^{(\sigma)}, \lambda_r \text{ vanish.} \end{aligned}\right\} \tag{180}$$

It follows from (174) that

$$\left.\begin{aligned} &T_{11} = X^{(1)}, \quad T_{22} = X^{(2)}, \quad T_{33} = X^{(3)}, \quad T_{44} = X^{(4)}, \\ &T_{rs} = 0 \text{ if } r \neq s. \end{aligned}\right\} \tag{181}$$

When we substitute from (180) and (181) in (179), we find that it is satisfied. Being true in the principal frame, it is true in general, and so (178) is established. It is an important formula; it breaks the energy tensor into a part of the form already obtained in (139) for an incoherent stream of particles, and a part due to stress in the medium.

However, the use of the words *density* for μ and *stress* for S_{rs} remains to be justified, and so we turn to (130)–(134) in order to interpret these quantities in terms of momentum and energy. Again we take the principal frame, so that (180) holds, and thus by (177) and (181) we have

$$\left.\begin{aligned} S_{11} &= -c^2X^{(1)} = -c^2T_{11}, \\ S_{22} &= -c^2X^{(2)} = -c^2T_{22}, \\ S_{33} &= -c^2X^{(3)} = -c^2T_{33}, \\ &\text{all other components of } S_{rs} \\ &\text{vanish.} \end{aligned}\right\} \tag{182}$$

By (178) we have then

$$\begin{gathered} T_{11} = -c^{-2}S_{11}, \quad T_{22} = -c^{-2}S_{22}, \quad T_{33} = -c^{-2}S_{33}, \\ T_{44} = -\mu, \quad T_{rs} = 0 \text{ if } r \neq s. \end{gathered} \tag{183}$$

In view of (130)–(134) we may make these statements:

i) The fluxes of relative momentum across areas normal to the axes of x_1, x_2, x_3 are in the directions of these axes and of amounts $-S_{11}$, $-S_{22}$, $-S_{33}$ per unit area per unit time. Thus S_{11}, S_{22}, S_{33} are principal stresses (as stress is understood in the kinetic theory of gases), the signs being in accordance with the convention which makes tension positive and pressure negative. In a gas, we think of the momentum as carried by migrating particles; in a solid, we may think of it as carried by internal impulses.

ii) The relative density of relative momentum vanishes, and so does the flux of relative energy.

iii) The relative density of relative energy is μc^2; thus it is appropriate to regard μ as mass-density in a macroscopic sense.

Further, in the principal frame the 3-velocity of the medium vanishes, by (176). This completes the justification of the terms used.

To sum up, given the energy tensor T_{rs}, we are in a position to speak of the density of the medium, its 4-velocity, and the stress in it,

these quantities being defined in a Lorentz-invariant way in terms of the eigen values and principal directions of T_{rs}. These are macroscopic quantities, and it is unnecessary to make further reference to the microscopic picture, although it may on occasion be interesting to do so.

We note here another form for (178) which is easy to verify:

$$T_{rs} = \mu\lambda_r\lambda_s - c^{-2}S_{(\varrho\sigma)}V_r^{(\varrho)}V_s^{(\sigma)}, \tag{184}$$

where the summation convention for repeated suffixes now operates in spite of the brackets (the range is 1, 2, 3), and $S_{(\varrho\sigma)}$ are the components of the stress tensor in the principal frame of reference as used above.

We are to remember that a principal frame of reference is a local thing, in general valid for one event only, since the principal directions at different events are in general not parallel. Thus while such frames are very useful for local interpretations, they are not suited to general arguments. That is why (178) is a much more useful formula than (183), for example.

The energy tensor yields, as we have seen, four invariants, viz. its eigen values $X^{(r)}$. But their determination involves the solution of an equation of the fourth degree. Their sum, however, is easily found, for it is the *trace* of T_{rs}, the trace being defined as the sum of the diagonal elements of the matrix, viz. the invariant T_{rr}. The invariant relation

$$T_{rr} = X^{(1)} + X^{(2)} + X^{(3)} + X^{(4)} \tag{185}$$

follows from (181). By (178) we can also write the trace in the form

$$T_{rr} = -\mu - c^{-2}S_{rr}, \tag{186}$$

where, by (177),

$$S_{rr} = -c^2(X^{(1)} + X^{(2)} + X^{(3)}); \tag{187}$$

this is the sum of the three principal stresses.

It is interesting to see what the trace is in terms of the microscopic picture. By (112), modified as explained in the footnote on p. 284, we have

$$T_{rr} = \sum_j \frac{\bar{N}^{(j)} M_r^{(j)} P_r^{(j)}}{|P_t^{(j)}\tilde{n}_t^{(j)}|}. \tag{188}$$

But $P_r^{(j)}$ and $M_r^{(j)}$ have the same direction or opposite directions. Hence the contribution to (188) due to null vectors vanishes; that wipes out the contributions due to photons and internal impulses, and

we have only to consider material particles. For them we may put

$$\bar{n}_r^{(j)} = \lambda_r^{(j)}, \quad P_r^{(j)} = \lambda_r^{(j)}, \quad M_r^{(j)} = m^{(j)}\lambda_r^{(j)}, \tag{189}$$

where $\lambda_r^{(j)}$ is the 4-velocity and $m^{(j)}$ the proper mass for particles in class j. Then (188) reduces to

$$T_{rr} = -\sum_j \bar{N}^{(j)} m^{(j)} = -\sum_j \mu^{(j)}, \tag{190}$$

where $\mu^{(j)}$ is the proper density of proper mass for particles in class j. When this is compared with (186), we get

$$\mu = \sum_j \mu^{(j)} - c^{-2} S_{rr}. \tag{191}$$

Thus the macroscopic density equals the sum of all the microscopic proper densities of proper mass, reduced by the sum of the three principal stresses (divided by c^2). This gives a further justification for our definition of density, since in most applications the last term in (191) is small.

§ 13. EQUATIONS OF MOTION OF A CONTINUOUS MEDIUM

In (117) we have the equation of conservation

$$\frac{\partial T_{rs}}{\partial x_s} = 0, \tag{192}$$

and in (178) we have

$$T_{rs} = \mu\lambda_r\lambda_s - c^{-2} S_{rs}. \tag{193}$$

Substitution from (193) in (192) gives us the equations of motion of a continuous medium.

These equations can be written in two forms. First, we have

$$\lambda_r \frac{\partial}{\partial x_s}(\mu\lambda_s) + \mu \frac{\partial \lambda_r}{\partial x_s}\lambda_s = c^{-2} \frac{\partial S_{rs}}{\partial x_s}. \tag{194}$$

Multiplication by λ_r gives [cf. (141)]

$$\frac{\partial}{\partial x_s}(\mu\lambda_s) = -c^{-2}\lambda_r \frac{\partial S_{rs}}{\partial x_s}, \tag{195}$$

and when this is substituted in (194) we get

$$\mu \frac{d\lambda_r}{ds} = c^{-2}(\delta_{rt} + \lambda_r\lambda_t) \frac{\partial S_{st}}{\partial x_s}. \tag{196}$$

This is the first form of the equations of motion; we note that the left hand side is

$$\text{density} \times \text{4-acceleration},$$

since d/ds indicates differentiation along the world line with tangent λ_r.

To get the second form of the equations, we write (194) in the equivalent form

$$\frac{\partial}{\partial x_s}(\mu\lambda_r)\lambda_s + \mu\lambda_r\frac{\partial\lambda_s}{\partial x_s} = c^{-2}\frac{\partial S_{rs}}{\partial x_s}. \tag{197}$$

Let σ be the 3-volume of the normal section of a thin tube of world lines of mean motion. Then, by (99),

$$\frac{d\sigma}{ds} = \frac{\partial\lambda_s}{\partial x_s}\sigma, \tag{198}$$

and (197) becomes, on multiplication by σ,

$$\sigma\frac{d}{ds}(\mu\lambda_r) + \mu\lambda_r\frac{d\sigma}{ds} = \sigma c^{-2}\frac{\partial S_{rs}}{\partial x_s}, \tag{199}$$

or

$$\frac{d}{ds}(\sigma\mu\lambda_r) = \sigma c^{-2}\frac{\partial S_{rs}}{\partial x_s}. \tag{200}$$

This is the second form of the equations of motion.

If we regard the thin world tube as the history of a "particle", then the proper mass of that particle is naturally taken to be $m = \sigma\mu$ (the product of 3-volume and density), and its 4-momentum to be $M_r = m\lambda_r = \sigma\mu\lambda_r$. Then (200) takes the form

$$\frac{dM_r}{ds} = \sigma F_r. \tag{201}$$

This agrees with the basic equation of motion of a particle [VI–(4)], if we interpret F_r as 4-force per unit volume. Thus we may say that the stress produces a 4-force per unit volume of amount

$$F_r = c^{-2}\frac{\partial S_{rs}}{\partial x_s}, \tag{202}$$

a formula similar to that which gives the resultant force due to stress on an elementary volume of an elastic body or a fluid, in the Newtonian mechanics of continua.

It is curious how the concept of a particle bobs up again in (201), for

we might have thought that the transition from the discrete to the continuous would have obliterated it. The sequence of ideas here is parallel to that by which, in Newtonian mechanics, we first think of a gas microscopically, with particles running about in all directions; then we smooth this microscopic chaos into the macroscopic theory of hydrodynamics; but in hydrodynamics we recognise that a fluid element moves as if it were a particle subject to a force due to pressure-gradient, or (in the case of a viscous fluid) to the gradient of a stress tensor; so the concept of "particle" is reborn in the macroscopic continuous theory. We may call it a *macroscopic particle.*

§ 14. THE PERFECT FLUID IN RELATIVITY

In Newtonian physics a perfect fluid is characterised by the condition that at each point in it the three principal stresses are equal to one another. We define a *perfect fluid* in relativity by the condition that the three eigen values of the energy tensor corresponding to spacelike principal directions are equal to one another. We write

$$X^{(1)} = X^{(2)} = X^{(3)} = p/c^2, \tag{203}$$

and call p the *pressure* in the fluid. The conditions (203) imply a spatial isotropy in the principal frame of reference at each event, but there is no suggestion of homogeneity — in general p varies from event to event. By (182) the principal stresses at an event have the common value $-p$.

We have now to find, by inserting the appropriate S_{rs} in (178), the energy tensor of a perfect fluid. By (177) we have

$$S_{rs} = -p(V_r^{(1)}V_s^{(1)} + V_r^{(2)}V_s^{(2)} + V_r^{(3)}V_s^{(3)}). \tag{204}$$

The vectors occurring here are not determined, because we introduced a degeneracy in making three of the eigen values equal. We may take $V_r^{(\sigma)}$ to be any triad of unit vectors, orthogonal to λ_r and to one another. Then

$$V_r^{(1)}V_s^{(1)} + V_r^{(2)}V_s^{(2)} + V_s^{(3)}V_s^{(3)} = \delta_{rs} + \lambda_r\lambda_s. \tag{205}$$

This tensor equation is immediately verified by using the principal frame of reference, so that (180) holds.

If we put (205) into (204), the stress tensor for a perfect fluid becomes (for any frame of reference)

$$S_{rs} = -p(\delta_{rs} + \lambda_r\lambda_s), \tag{206}$$

and (178) gives for *the energy tensor of a perfect fluid*

$$\boxed{T_{rs} = (\mu + c^{-2}p)\lambda_r\lambda_s + c^{-2}p\delta_{rs}.} \tag{207}$$

We recall that μ is the density and λ_r the 4-velocity.

By direct calculation from (206) we have

$$\left.\begin{aligned} &\frac{\partial S_{rs}}{\partial x_s} = -\frac{\partial p}{\partial x_s}(\delta_{rs} + \lambda_r\lambda_s) - p\lambda_r\frac{\partial \lambda_s}{\partial x_s} - p\lambda_s\frac{\partial \lambda_r}{\partial x_s}, \\ &\lambda_r\frac{\partial S_{rs}}{\partial x_s} = p\frac{\partial \lambda_s}{\partial x_s}, \\ &(\delta_{rt} + \lambda_r\lambda_t)\frac{\partial S_{st}}{\partial x_s} = -\frac{\partial p}{\partial x_s}(\delta_{rs} + \lambda_r\lambda_s) - p\frac{d\lambda_r}{ds}, \end{aligned}\right\} \tag{208}$$

and so (196) gives as equations of motion

$$(\mu + c^{-2}p)\frac{d\lambda_r}{ds} = -c^{-2}\frac{\partial p}{\partial x_s}(\delta_{rs} + \lambda_r\lambda_s). \tag{209}$$

However, if we multiply by λ_r, we get the identity $0 = 0$, and so we have only three independent equations here. The fourth equation is (195), which now reads

$$\frac{\partial}{\partial x_s}(\mu\lambda_s) = -c^{-2}p\frac{\partial \lambda_s}{\partial x_s}, \tag{210}$$

or

$$\frac{d\mu}{ds} + (\mu + c^{-2}p)\frac{\partial \lambda_s}{\partial x_s} = 0. \tag{211}$$

Any four independent equations contained in the set (209), (211) *are to be regarded as the equations of motion of a perfect fluid.*

Let us see what these equations look like when we introduce the 3-velocity u_ϱ by (176), and put $ds = cdt/\gamma$. The first three of (209) give

$$\gamma(\mu + c^{-2}p)\frac{d}{dt}(\gamma u_\varrho) = -\frac{\partial p}{\partial x_\varrho} - c^{-1}\gamma u_\varrho\lambda_s\frac{\partial p}{\partial x_s}, \tag{212}$$

and the fourth gives

$$\gamma(\mu + c^{-2}p)\frac{d\gamma}{dt} = c^{-2}\frac{\partial p}{\partial t} - c^{-1}\gamma\lambda_s\frac{\partial p}{\partial x_s}. \tag{213}$$

If we multiply (213) by u_ϱ and substract it from (212), we get

$$\gamma^2(\mu + c^{-2}p)\frac{du_\varrho}{dt} = -\frac{\partial p}{\partial x_\varrho} - c^{-2}u_\varrho\frac{\partial p}{\partial t}. \tag{214}$$

Further

$$c\frac{\partial \lambda_s}{\partial x_s} = \frac{\partial}{\partial x_\sigma}(\gamma u_\sigma) + \frac{\partial \gamma}{\partial t}, \tag{215}$$

and so (211) gives

$$\gamma\frac{d\mu}{dt} + (\mu + c^{-2}p)\left[\frac{\partial \gamma}{\partial t} + \frac{\partial}{\partial x_\sigma}(\gamma u_\sigma)\right] = 0. \tag{216}$$

The four equations (214), (216) *are the equations of motion of a perfect fluid in terms of the 3-velocity.* In these equations d/dt means "differentiation following the fluid" in the usual hydrodynamical sense, viz.

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u_\sigma\frac{\partial}{\partial x_\sigma}. \tag{217}$$

In exploring the limiting form of these equations of motion for large c, there are two dimensionless quantities which may be taken to be small independently:

$$\text{(i) } u/c, \qquad \text{(ii) } p/(\mu c^2). \tag{218}$$

If u/c is small, we may say that the motion is *slow* (in a technical sense, of course, since the velocity might still be quite large in cm $\sec^{-1}$); if $p/(\mu c^2)$ is small, we may say that the pressure is *low* (again in a technical sense). The limiting forms are as follows (we assume $u\partial p/\partial x_\varrho$ and $\partial p/\partial t$ to be of the same order of magnitude).

Fluid in slow motion:

$$\left.\begin{aligned}(\mu + c^{-2}p)\frac{du_\varrho}{dt} &= -\frac{\partial p}{\partial x_\varrho},\\ \frac{d\mu}{dt} + (\mu + c^{-2}p)\frac{\partial u_\sigma}{\partial x_\sigma} &= 0.\end{aligned}\right\} \tag{219}$$

Here we recognise the usual hydrodynamical equations of motion and the equation of continuity, with a change of the density μ into $(\mu + c^{-2}p)$ in two places.

Fluid with low pressure:

$$\left.\begin{aligned} \gamma^2\mu\frac{du_\varrho}{dt} &= -\frac{\partial p}{\partial x_\varrho} - c^{-2}u_\varrho\frac{\partial p}{\partial t}, \\ \gamma\frac{d\mu}{dt} + \mu\left[\frac{\partial\gamma}{\partial t} + \frac{\partial}{\partial x_\sigma}(\gamma u_\sigma)\right] &= 0. \end{aligned}\right\} \tag{220}$$

Here the usual hydrodynamical equations are modified by the presence of γ and an additional term in the first equation.

For slow motion combined with low pressure, we drop $c^{-2}p$ from (219), and get the classical hydrodynamical equations

$$\left.\begin{aligned} \mu\frac{du_\varrho}{dt} &= -\frac{\partial p}{\partial x_\varrho} \\ \frac{d\mu}{dt} + \mu\frac{\partial u_\sigma}{\partial x_\sigma} &= 0. \end{aligned}\right\} \tag{221}$$

Returning to the general equations (214), (216), we note that we have only four equations for the five unknowns (u_ϱ, μ, p), and so we meet the indeterminacy familiar in Newtonian hydrodynamics. We could get determinacy by adding another unknown, the *temperature* Θ, and two more equations, viz. an *equation of state* connecting μ, p and Θ, and an *equation of conductivity*. But, in view of consequent complications, one may prefer to cut the knot, as one often does in Newtonian hydrodynamics, by simply imposing an equation between density and pressure,

$$\mu = f(p). \tag{222}$$

We now have in (214), (216), (222) five equations for the five unknowns (u_ϱ, μ, p).

These equations represent a definite relativistic advance over the Newtonian equations (221), because they are Lorentz-invariant whereas the Newtonian equations are not. A relativitist could, for that reason, never accept (221) as "correct" equations. But if he tries to produce experimental evidence in favour of the relativistic equations, he meets with a common difficulty: the motions with which he can experiment are (in our technical sense) slow and at low pressure, and experimental techniques cannot distinguish between the rival equations.

In the general theory of a perfect fluid with energy tensor (207) we need not commit ourselves as to the signs of μ and p. But in the case of a fluid regarded as the statistical limit of a microscopic system composed of material particles, it is possible to assert that

$$\mu > 0, \quad p \geqslant 0. \tag{223}$$

This is seen as follows. Insert (189) in (112); then

$$T_{rs} = \sum_j \bar{N}^{(j)} m^{(j)} \lambda_r^{(j)} \lambda_s^{(j)}, \tag{224}$$

where $\lambda_\varrho^{(j)}$ are real and $\lambda_4^{(j)}$ pure imaginary and not zero. Hence

$$T_{11} \geqslant 0, \quad T_{44} < 0, \tag{225}$$

for all Galileian frames. But if we use a principal tetrad, (207) gives

$$T_{11} = c^{-2} p, \quad T_{44} = -\mu, \tag{226}$$

and so (223) follows.

In the next section the concept of the incompressible fluid will be discussed, but a full treatment of relativistic hydrodynamics will not be attempted in this book. For some developments, including discussion of a stationary principle and vorticity, reference may be made to EISENHART [1924], SYNGE [1937a], and LICHNEROWICZ [1940, 1941, 1944].

§ 15. INCOMPRESSIBLE FLUIDS

Most of classical hydrodynamics deals with incompressible fluids. That is partly because water and other liquids are only slightly compressible, but probably more because the assumption of incompressibility simplifies the mathematics and makes it possible to obtain results. Strangely enough, the theory of incompressible fluids is successfully applied in aerodynamics, where the fluid (air) is far from being incompressible.

For similar reasons of mathematical simplicity, we seek in relativity a definition of *incompressibility* as applied to a perfect fluid.

The most obvious definition of incompressibility is kinematical. We demand that an element of the fluid shall retain the same proper volume; in other words, we demand that *the expansion of an elementary world-tube shall be zero*, or by (99)

$$\frac{\partial \lambda_s}{\partial x_s} = 0. \tag{227}$$

Then (211) gives $d\mu/ds = 0$, so that the density is constant along each world line. For simplicity we add an assumption of homogeneity, so that μ is constant, not merely along each world line, but throughout the fluid. Thus we have, as our first definition of an incompressible homogeneous fluid,

$$\mu = \text{const.} \tag{228}$$

Under this definition, (214) retains its form and (216) simplifies by the vanishing of the first term, so that the equations of motion read

$$\begin{gathered}\gamma^2(\mu + c^{-2}p)\frac{du_\varrho}{dt} = -\frac{\partial p}{\partial x_\varrho} - c^{-2}u_\varrho\frac{\partial p}{\partial t}, \\ \frac{\partial \gamma}{\partial t} + \frac{\partial}{\partial x_\sigma}(\gamma u_\sigma) = 0.\end{gathered} \tag{229}$$

Here we have four equations for the four unknowns (u_ϱ, p). But this definition is open to objection, as we shall see later.

To form a second definition of incompressibility, we go back to (191) which reads (for a perfect fluid)

$$\mu = \sum_j \mu^{(j)} + 3c^{-2}p. \tag{230}$$

We take as our definition the condition that in an incompressible fluid *the proper density $\mu^{(j)}$ of each class j shall be constant.* By (190) this implies that the trace of the energy tensor shall be constant:

$$T_{rr} = \text{const.}, \tag{231}$$

and by (230) we have

$$\mu - 3c^{-2}p = \text{const.} \tag{232}$$

a relationship of the form (222). If we substitute for μ from (232) in (214), (216), we get four equations for the four unknowns (u_ϱ, p).

We note that, according to (54), a photonic gas satisfies $\mu - 3c^{-2}p = 0$, and therefore qualifies as an incompressible fluid under this second definition. This does not seem an appropriate use of the word "incompressible". Nevertheless, the definition (231) or (232) was recommended by EDDINGTON [1924, pp. 122, 168–170] and used by SEN [1934a] in discussing the equilibrium of an incompressible fluid sphere under its gravitational attraction, the definition being put into a form suitable to the general theory of relativity.

There is however a third way of looking at the question of incompressibility. In a compressible fluid (pressure p, density μ) the speed

of sound is $(dp/d\mu)^{\frac{1}{2}}$ according to Newtonian physics. As the fluid tends to incompressibility, this speed tends to infinity. This limiting infinite speed of sound violates no principle of Newtonian physics, but in relativity the existence of any signal-velocity greater than c (the speed of light) would force us to revise our most basic assumptions. In fact, we should so frame our definition of incompressibility that the speed of sound in an incompressible fluid does not exceed c.

Let us then examine the speed of sound in a fluid which moves according to the equations (214), (216), with density and pressure satisfying a general relationship of the form (222). To deal with vibrations of the fluid, we shall assume u_ϱ small and put

$$\mu = \bar{\mu} + \mu', \quad p = \bar{p} + p', \tag{233}$$

where $\bar{\mu}$, $\bar{p}$ are constants and μ', p' small. We shall make no approximation based on the largeness of c. Then, to the first order, (214) and (216) give

$$\left.\begin{aligned} (\bar{\mu} + c^{-2}\bar{p})\,\frac{\partial u_\varrho}{\partial t} &= -\frac{\partial p'}{\partial x_\varrho}, \\ \frac{\partial \mu'}{\partial t} + (\bar{\mu} + c^{-2}\bar{p})\,\frac{\partial u_\sigma}{\partial x_\sigma} &= 0. \end{aligned}\right\} \tag{234}$$

By (222) we have

$$\mu' = F(\bar{p})p', \tag{235}$$

where $F(\bar{p}) = d\bar{\mu}/d\bar{p}$, and when we eliminate μ' from (234) we get

$$\frac{\partial^2 p'}{\partial x_\varrho \partial x_\varrho} - F(\bar{p})\,\frac{\partial^2 p'}{\partial t^2} = 0. \tag{236}$$

This is a wave equation of the usual form, and it gives for the speed of sound

$$v = [F(\bar{p})]^{-\frac{1}{2}} = (d\bar{p}/d\bar{\mu})^{\frac{1}{2}}, \tag{237}$$

which is, in fact, the Newtonian formula.

The demand $v \leqslant c$ for an incompressible fluid, being only an inequality, does not provide us with the form of the (μ, p) relationship (222). But it provides a test, and rules out the definition $\mu = \text{const.}$ as in (228); that gives $v = \infty$. It does not rule out (232), for that gives $v = c/\sqrt{3}$.

But if we go further and demand that the highest possible speed of

sound ($v = c$) is attained in an incompressible fluid, no matter what the pressure is, then (237) gives us the relationship

$$p = (\mu - \mu_0)c^2, \tag{238}$$

where μ_0 is a constant, viz. the density when $p = 0$. This may be regarded as a third definition of incompressibility. The corresponding equations of motion are, by (214) and (216),

$$\left.\begin{aligned} \gamma^2(\mu_0 + 2c^{-2}p)\frac{du_\varrho}{dt} &= -\frac{\partial p}{\partial x_\varrho} - c^{-2}u_\varrho\frac{\partial p}{\partial t}, \\ \gamma c^{-2}\frac{dp}{dt} + (\mu_0 + 2c^{-2}p)\left[\frac{\partial \gamma}{\partial t} + \frac{\partial}{\partial x_\sigma}(\gamma u_\sigma)\right] &= 0, \end{aligned}\right\} \tag{239}$$

four equations for the four unknowns (u_ϱ, p).

To sum up, three different definitions of incompressibility have been discussed. They are as follows, v being the speed of sound:

$$\left.\begin{aligned} \mu &= \text{const.}, & v &= \infty, \\ \mu - 3c^{-2}p &= \text{const.}, & v &= c/\sqrt{3}, \\ p &= (\mu - \mu_0)c^2, & v &= c. \end{aligned}\right\} \tag{240}$$

§ 16. ISOLATED SYSTEMS AND THE ENERGY TENSOR

To repeat what was said in VII–§ 1, it is quite meaningless to ask whether matter is *really* discrete or continuous, for the concepts of continuity and discontinuity belong to mathematics, not to nature. We use now this mathematical representation, now that, seeking those representations which are convenient to work with and which yield at least some correct physical predictions. In Ch. VII we constructed a model of a discrete system, and passed on in the present chapter to a continuum, considered as a limit of a discrete system. But that order of ideas is by no means essential, and indeed the reverse order is more popular in physics today. We can start with a continuum (field theory), and derive "particles" as singularities in the continuum by a limiting process. In the present section we shall follow the method of field theory, but without passing to the limit of singularities.

Laying aside our previous ideas about the energy tensor, we think of a domain in space-time in which there exists a tensor T_{rs} which we

call the *energy tensor* and on which there is imposed the *conservation condition*

$$\frac{\partial T_{rs}}{\partial x_s} = 0. \tag{241}$$

(This is of course our equation (117), but for present purposes (241) is to be thought of as a basic hypothesis, not a deduction.) To link T_{rs} with terms familiar in physics, we say as in (111) that the flux of 4-momentum across a 3-element dS orthogonal to a unit vector n_r is

$$\varepsilon(n) T_{rs} n_s dS, \tag{242}$$

or, what is really equivalent, we may interpret the individual components as in § 9:

$$\begin{aligned} T_{\varrho\sigma} &= -c^{-2} \times \text{stress}, \\ T_{\varrho 4} &= ic^{-1} \times \text{momentum-density}, \\ T_{4\varrho} &= ic^{-3} \times \text{energy-flux}, \\ T_{44} &= -c^{-2} \times \text{energy-density}. \end{aligned} \tag{243}$$

An advantage of the present abstract approach is that we are not compelled to make T_{rs} symmetric. We shall impose symmetry later as a further hypothesis, but for the moment we shall proceed without assuming it. We note that (241) implies, by Green's theorem,

$$\int \varepsilon(n) T_{rs} n_s dS = 0, \tag{244}$$

the integral being taken over any closed 3-space, and this may be regarded as a *closed law of 4-momentum* in the sense of VII–§ 4.

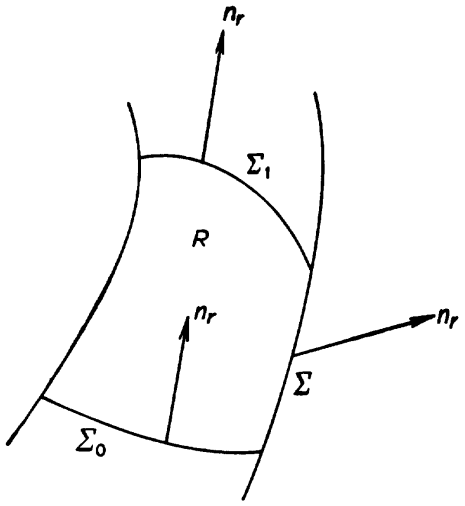


Fig. 11 – An isolated system

We shall consider an *isolated system.* By this we mean that our system occupies the interior of a timelike tube in space-time bounded by a 3-space Σ across which there is no flux of 4-momentum; thus on Σ we have

$$T_{rs} n_s = 0 \tag{245}$$

where n_r is the unit normal to Σ, drawn outwards (Fig. 11). This is the *condition of isolation.* We do not assume that the

tube is necessarily of small section; in fact the theory is free of all approximation.[1]

Let Σ_0 and Σ_1 be two sections of the tube. Their forms are completely arbitrary, but we shall suppose that Σ_0 is before Σ_1 in time and that they do not intersect, so that a 4-dimensional region R is enclosed by Σ, Σ_0 and Σ_1. (The non-intersection of Σ_0 and Σ_1 is not really essential, but it simplifies the description.) Let n_r be the unit normal to Σ_0 and Σ_1, pointing into R on Σ_0 and out of R on Σ_1.

Then by (241) and Green's theorem (98) we have

$$\int_{\Sigma} T_{rs} n_s dS + \int_{\Sigma_1} \varepsilon(n) T_{rs} n_s dS - \int_{\Sigma_0} \varepsilon(n) T_{rs} n_s dS = 0. \tag{246}$$

The integral over Σ vanishes by the isolation condition (245). Therefore if we hold Σ_0 fixed and take all possible sections Σ_1, the vector

$$M_r = \int_{\Sigma_1} \varepsilon(n) T_{rs} n_s dS \tag{247}$$

is independent of Σ_1, and so is a constant of the tube. We call M_r the *4-momentum of the isolated system*, and note that this 4-momentum is a constant; by (242) M_r is in fact the flux of 4-momentum across any section of the tube.

On account of (241) we have

$$\frac{\partial}{\partial x_p}(x_r T_{sp} - x_s T_{rp}) = T_{sr} - T_{rs}, \tag{248}$$

and so, by Green's theorem and (245),

$$\int_{\Sigma_1} \varepsilon(n)(x_r T_{sp} - x_s T_{rp}) n_p dS = \int_{\Sigma_0} \varepsilon(n)(x_r T_{sp} - x_s T_{rp}) dS + \int_R (T_{sr} - T_{rs}) d\tau. \tag{249}$$

If we hold Σ_0 fixed and change Σ_1, the first integral on the right remains fixed, but in general the second integral changes, and so we get nothing interesting unless we impose the *symmetry condition*

$$T_{rs} = T_{sr}, \tag{250}$$

[1] If the components T_{rs} tend to zero sufficiently rapidly at infinity in the space of an observer, the theory can be applied to a system filling the whole of space-time. The essential requirements are (i) that the various integrals converge when Σ is removed to infinity, and (ii) that the integrals $\int T_{rs} n_s dS$ and $\int x_r T_{sp} n_p dS$, taken over Σ, vanish when Σ is removed to infinity. The theory is used in that way in IX–§ 14.

and we shall now impose it. Then (249) tells us that the tensor

$$H_{rs} = \int_{\Sigma_1} \varepsilon(n)\ (x_r T_{sp} - x_s T_{rp}) n_p\, dS \tag{251}$$

is a constant of the isolated system. We call it the *angular momentum* (cf. VII–§ 5 for the discrete approach).

We now write down the four equations [cf. VII–(30)]

$$(\bar{x}_r M_s - \bar{x}_s M_r) M_s = H_{rs} M_s. \tag{252}$$

The solution of these equations gives events $\bar{x}_r$ lying on a straight line in space-time which has the direction of M_r. We call this straight line the *history of the mass-centre.*

Thus, using nothing but the conservation condition (241), the isolation condition (245), and the symmetry condition (250), we have shown that an isolated system has a constant 4-momentum M_r, a constant angular momentum H_{rs}, and a mass-centre as in (252).

We have however no assurance that M_r is timelike or that it points into the future. This demands further hypotheses, and we shall impose on T_{rs} the two inequalities (152), (170), already considered in connection with eigen values: we assume that

$$T_{rs} p_r p_s > 0 \tag{253}$$

for all timelike or null vectors p_r, and that

$$T_{rs} T_{rt} p_s p_t < 0 \tag{254}$$

for all timelike p_r. We are then assured of the reality of the eigen values of T_{rs}, but that does not concern us here. What matters is that by (254) the vector

$$F_r = - T_{rs} p_s \tag{255}$$

is timelike for any timelike p_r, and (253) tells us that if p_r points into the future, then so does F_r, as is not hard to see.

Now if we take n_r timelike in (247) (and necessarily pointing into the future, since it points out of R in Fig. 11), we have

$$M_r = - \int_{\Sigma_1} T_{rs} n_s\, dS. \tag{256}$$

This may be regarded as the sum of infinitesimal vectors, each timelike and future-pointing, being like F_r in (255). Hence M_r itself is timelike and points into the future. It is certainly well to have it so; if M_r pointed into the past the system would have negative energy, and if it were spacelike, the mass-centre would travel faster than light. We

note that the inequalities (253) and (254), if imposed on the energy tensor, eliminate these embarrassments.

We can now choose M_r for time axis, so that $M_\varrho = 0$. Then the equations (252) for the history of the mass-centre reduce to

$$\bar{x}_\varrho M_4^2 = H_{\varrho 4} M_4. \tag{257}$$

If we choose Σ_1 to be $x_4 = \text{const.}$, this gives

$$\bar{x}_\varrho = \frac{H_{\varrho 4}}{M_4} = \frac{\int_{\Sigma_1} x_\varrho T_{44} dS}{\int_{\Sigma_1} T_{44} dS}. \tag{258}$$

This is the (fixed) position of the mass-centre in the rest-frame of the isolated system (time-axis parallel to M_r). The integrals in (258) are taken over the instantaneous position of the body, and are independent of the time.

Since by (243) we have $T_{44} = -\mu$ where μc^2 is energy-density, (258) may also be written

$$\bar{x}_\varrho = \frac{\int_{\Sigma_1} x_\varrho \mu dS}{\int_{\Sigma_1} \mu dS}; \tag{259}$$

this has the usual form of an expression for mass-centre [cf. VII–(37) for a discrete system]. By (253) we have

$$T_{44} < 0, \quad \mu > 0. \tag{260}$$

If we write (259) in the form

$$\int_{\Sigma_1} (x_\varrho - \bar{x}_\varrho) \mu dS = 0, \tag{261}$$

we see that $(x_\varrho - \bar{x}_\varrho)$ must change sign in Σ_1. This implies that any sphere which contains the section Σ_1 at time t must contain also the (fixed) mass-centre. Or, if Σ_1 is convex, then it contains the mass-centre. *In space-time language, this last statement tells us that the world tube of the system contains the straight world line of the mass-centre.* In view of the high indeterminacy of the conditions of conservation (241) — four equations for ten unknowns — this result is rather remarkable. It shows that, although the world tube may spread out in section, it cannot (in the case of positive energy-density) wander *as a whole* away from the mass-centre.

There is another way (LANCZOS [1930]) of looking at these questions; it is more direct, but conceals the question of Lorentz invariance. We assume the conservation condition (241), the isolation condition (245), and the symmetry condition (250). At any instant t the system occupies a portion of space bounded by some surface B, the equation of which may be written[1]

$$f(x_1, x_2, x_3) = ct \text{ or } f(x_1, x_2, x_3) + ix_4 = 0. \tag{262}$$

The outward unit normal to B is

$$\nu_\varrho = \eta f_{,\varrho}(f_{,\sigma}f_{,\sigma})^{-\frac{1}{2}}, \quad \eta = \pm 1, \tag{263}$$

the commas indicating partial derivatives, and the normal velocity of B (counted positive if in the sense of f increasing) is easily seen to be

$$u_n = \eta c(f_{,\sigma}f_{,\sigma})^{-\frac{1}{2}}. \tag{264}$$

Considered as a 3-space in space-time, (262) has a normal vector n_r with components in the ratios $f_{,1} : f_{,2} : f_{,3} : i$, and so the isolation condition (245) gives

$$T_{r\sigma}f_{,\sigma} + iT_{r4} = 0, \tag{265}$$

or

$$T_{r\sigma}\nu_\sigma + iT_{r4}u_n/c = 0. \tag{266}$$

Then dS being an element of 3-volume and dB an element of surface area, we have, by (241) and Green's theorem for 3-space,

$$\begin{aligned}\frac{d}{dt}\int T_{r4}\,dS &= ic\int \frac{\partial T_{r4}}{\partial x_4}\,dS + \int T_{r4}u_n\,dB \\ &= -ic\int \frac{\partial T_{r\sigma}}{\partial x_\sigma}\,dS + \int T_{r4}u_n\,dB \\ &= -ic\int (T_{r\sigma}\nu_\sigma + iT_{r4}u_n/c)\,dB \\ &= 0 \end{aligned} \tag{267}$$

by (266). Thus

$$M_r = -i\int T_{r4}\,dS \tag{268}$$

[1] If B is fixed ($u_n = 0$), we cannot write its history as in (262); but in that case it is obvious that the isolation condition is given by (266), with $u_n = 0$. For a discussion of moving surfaces, see Appendix A, p. 419.

is a constant of the motion; it is in fact a particular case of (247). Further

$$\begin{aligned}\frac{d}{dt}\int x_\varrho T_{44}\,dS &= ic\int x_\varrho \frac{\partial T_{44}}{\partial x_4}\,dS + \int x_\varrho T_{44}u_n\,dB\\ &= -ic\int x_\varrho \frac{\partial T_{4\sigma}}{\partial x_\sigma}\,dS + \int x_\varrho T_{44}u_n\,dB\\ &= ic\int T_{\varrho\varrho}\,dS - ic\int x_\varrho(T_{4\sigma}v_\sigma + iT_{44}u_n/c)\,dB\\ &= -cM_\varrho \qquad (269)\end{aligned}$$

which is constant. (We used the symmetry of T_{rs}.)

We now define the *centroid* of the system by

$$\tilde{x}_\varrho = \frac{\int x_\varrho T_{44}\,dS}{\int T_{44}\,dS}, \qquad (270)$$

avoiding the word *mass-centre* because this is actually not the mass-centre. Differentiation gives

$$\frac{d\tilde{x}_\varrho}{dt} = -\frac{cM_\varrho}{iM_4}, \quad \frac{d^2\tilde{x}_\varrho}{dt^2} = 0. \qquad (271)$$

Thus the centroid $\tilde{x}_\varrho$ of an isolated system moves with constant velocity; its world line is parallel to M_r.

If we assume that energy-density is positive (T_{44} negative), then (270) implies that, for a convex body, the centroid lies inside it.

The relationship between mass-centre (252) and centroid (270) is interesting. The mass-centre is certainly relativistically invariant; the centroid is not. If we slice the world tube of the system by sections orthogonal to some selected vector n_r, we get a centroid on each section, and these centroids lie on a straight line in space-time parallel to M_r. But if we change n_r to n'_r, we get a different locus of centroids. To see this, we write (270) in the form [cf. (251)]

$$\tilde{x}_\varrho = \frac{iH_{\varrho 4} + \tilde{x}_4\int T_{\varrho 4}\,dS}{iM_4} = \frac{H_{\varrho 4} + \tilde{x}_4 M_\varrho}{M_4}, \qquad (272)$$

or

$$\tilde{x}_\varrho M_4 - \tilde{x}_4 M_\varrho = H_{\varrho 4}. \qquad (273)$$

Consider now the Lorentz-invariant equations

$$(\tilde{x}_r M_s - \tilde{x}_s M_r)n_s = H_{rs}n_s. \qquad (274)$$

These equations determine, for any n_r, a straight world line parallel to M_r. If we take n_r for time axis, so that $n_\varrho = 0$, they reduce to (273). *In fact,* (274) *is the invariant form of equation for the world line of centroids of sections orthogonal to* n_r. It is *not* the same as (252) (n_s has replaced M_s), and of course the straight line (274) depends on the choice of n_r.

We conclude that mass-centre and centroid, as defined above, are different concepts, the former Lorentz-invariant and the latter not so; they coincide only when the slicing for the centroid is orthogonal to the 4-momentum M_r of the system.

Let the time-axis be along M_r and the origin at the mass-centre; then

$$M_\varrho = 0, \quad \bar{x}_\varrho = 0, \quad H_{\varrho 4} = 0, \tag{275}$$

and (274) gives for the centroid

$$\tilde{x}_\varrho = H_{\varrho\sigma} l_\sigma / (iM_4), \quad l_\sigma = in_\sigma / n_4, \quad l_\sigma l_\sigma = 1. \tag{276}$$

As on p. 225, $H_{\varrho\sigma}$ defines a 3-vector; taking the x_1-axis in its direction, we have

$$-M_4^2 \tilde{x}_\varrho \tilde{x}_\varrho = H_{23}^2 (l_2^2 + l_3^2) = H_{23}^2 (1 - l_1^2). \tag{277}$$

It follows that a convex body with positive energy-density contains a sphere of radius r, where

$$r^2 = -\frac{H_{23}^2}{M_4^2} = -\tfrac{1}{2} \frac{H_{rs} H_{rs}}{M_n M_n}, \tag{278}$$

the last expression being an invariant.[1]

The point-particle with intrinsic angular momentum (p. 221) should therefore be regarded as an independent concept, suggested by the consideration of an extended system, but not necessarily the limit of such a system.

[1] Cf. MØLLER [1949]. I am much indebted to Dr. F. A. E. Pirani for discussions on these matters.

CHAPTER IX

THE ELECTROMAGNETIC FIELD IN VACUO

§ 1. THE ELECTROMAGNETIC TENSOR F_{rs}

When we think of an electromagnetic field in vacuo in pre-relativistic physics, we have in mind the fields of electric and magnetic force; there is an electric 3-vector E_ϱ and a magnetic 3-vector H_ϱ, both functions of position and time, satisfying Maxwell's equations, which read

$$\begin{aligned} \frac{1}{c}\frac{\partial E_\varrho}{\partial t} &= \varepsilon_{\varrho\mu\nu}\frac{\partial H_\nu}{\partial x_\mu}, \quad \frac{\partial E_\varrho}{\partial x_\varrho} = 0, \\ -\frac{1}{c}\frac{\partial H_\varrho}{\partial t} &= \varepsilon_{\varrho\mu\nu}\frac{\partial E_\nu}{\partial x_\mu}, \quad \frac{\partial H_\varrho}{\partial x_\varrho} = 0, \end{aligned} \tag{1}$$

if we use the 3-dimensional permutation symbol instead of the more usual curl or rot. In order to absorb these concepts into relativity, we ask: Can we assign a law of transformation for E_ϱ and H_ϱ so that these Maxwellian equations retain their form when the space-time coordinates undergo a Lorentz transformation?

This question is best answered by making a fresh start, in which we forget E_ϱ and H_ϱ for the moment and take, as the quantities which define an electromagnetic field in vacuo, the six independent components of a skew-symmetric tensor F_{rs} $(= -F_{sr})$ — we are using Minkowskian coordinates with $x_4 = \mp ict$. This tensor we call the *electromagnetic tensor*, and in terms of it we *define* two 3-vectors [cf. VII–§ 8] by

$$E_\varrho = iF_{\varrho 4}, \quad H_\varrho = -iF^*_{\varrho 4}, \tag{2}$$

where F^*_{rs} is the dual of F_{rs}, given by

$$F^*_{rs} = \tfrac{1}{2}i\varepsilon_{rsmn}F_{mn}, \quad F_{rs} = -\tfrac{1}{2}i\varepsilon_{rsmn}F^*_{mn}, \tag{3}$$

the second equation here being a consequence of the first.

We call E_ϱ the *electric 3-vector* and H_ϱ the *magnetic 3-vector*. The

formulae (2) are fundamental, and we shall write them out in other equivalent forms:

$$E_\varrho = iF_{\varrho 4}, \quad H_\varrho = \tfrac{1}{2}\varepsilon_{\varrho\mu\nu}F_{\mu\nu}, \tag{4}$$

$$E_\varrho = \tfrac{1}{2}\varepsilon_{\varrho\mu\nu}F^*_{\mu\nu}, \quad H_\varrho = -iF^*_{\varrho 4}, \tag{5}$$

or explicitly

$$\boxed{\begin{aligned} &E_1 = iF_{14}, \quad E_2 = iF_{24}, \quad E_3 = iF_{34}, \\ &H_1 = F_{23}, \quad H_2 = F_{31}, \quad H_3 = F_{12}; \end{aligned}} \tag{6}$$

$$\begin{aligned} &E_1 = F^*_{23}, \qquad E_2 = F^*_{31}, \qquad E_3 = F^*_{12}, \\ &H_1 = -iF^*_{14}, \quad H_2 = -iF^*_{24}, \quad H_3 = -iF^*_{34}. \end{aligned} \tag{7}$$

For F_{rs}, F^*_{rs} in terms of E_ϱ, H_ϱ, we have

$$F_{\varrho 4} = -iE_\varrho, \quad F_{\varrho\sigma} = \varepsilon_{\varrho\sigma\mu}H_\mu, \tag{8}$$

$$F^*_{\varrho 4} = iH_\varrho, \qquad F^*_{\varrho\sigma} = \varepsilon_{\varrho\sigma\mu}E_\mu. \tag{9}$$

We recall that $F_{\varrho 4}$ and $F^*_{\varrho 4}$ are pure imaginaries, all the other components of F_{rs} and F^*_{rs} being real, so that E_ϱ and H_ϱ are real 3-vectors.

The definitions of E_ϱ and H_ϱ in (2) tell us how these 3-vectors transform under a Lorentz transformation, because we have chosen F_{rs} to be a tensor. We shall presently consider these transformations in detail, but our immediate concern is with Maxwell's equations (1). To show that these equations do in fact retain their form under a Lorentz transformation, the simplest plan is to forget (1) for the moment and take as our *relativistic Maxwell equations in vacuo*

$$\frac{\partial F_{rs}}{\partial x_s} = 0, \quad \frac{\partial F^*_{rs}}{\partial x_s} = 0. \tag{10}$$

These equations are obviously tensor equations, and so they retain their form under Lorentz transformation. To show that (10) are in fact the same as (1), we note that, by (8),

$$\begin{aligned} &\frac{\partial F_{\varrho s}}{\partial x_s} = \frac{\partial F_{\varrho\sigma}}{\partial x_\sigma} + \frac{\partial F_{\varrho 4}}{\partial x_4} = \varepsilon_{\varrho\sigma\mu}\frac{\partial H_\mu}{\partial x_\sigma} - \frac{1}{c}\frac{\partial E_\varrho}{\partial t}, \\ &\frac{\partial F_{4s}}{\partial x_s} = \frac{\partial F_{4\sigma}}{\partial x_\sigma} = i\frac{\partial E_\sigma}{\partial x_\sigma}, \end{aligned} \tag{11}$$

and so the first of (10) is equivalent to the first set of the equations (1); similarly, the second of (10) is equivalent to the second set of (1).

To sum up, we lay the foundation of relativistic electromagnetic theory in vacuo by stating that *the electromagnetic field is specified by a skew-symmetric tensor F_{rs} which satisfies Maxwell's equations in the form* (10).

It should be noted that, unlike Newtonian mechanics which is modified essentially by relativity, Maxwellian electromagnetism passes over into relativity without change, in the sense that any Galileian observer can use the familiar equations (1). It is only when we change from one such observer to another that the Lorentz transformation makes itself felt. The transformation formulae for F_{rs} are simpler than those for E_ϱ and H_ϱ (we shall discuss these in the next section), and for that reason it is usually more convenient in relativity to use F_{rs} and the equations (10). This change in notation makes relativistic electromagnetism look more different from the familiar theory than it really is.

The second equation in (10) may be put in another form. By (3) we may write it (using a comma for partial differentiation)

$$\varepsilon_{rsmn} F_{mn,s} = 0. \tag{12}$$

If $r = 1$, then smn must be 234 in some order, if the term is to survive, and so for $r = 1$ we get

$$F_{34,2} + F_{42,3} + F_{23,4} = 0. \tag{13}$$

Giving r in turn the values 1, 2, 3, 4, we see that (12) may be written

$$F_{mn,r} + F_{nr,m} + F_{rm,n} = 0. \tag{14}$$

This form is interesting because, although we have in mind at present only Lorentz transformations, for which the tensor character of (14) is obvious, this equation is actually a tensor equation for *any* transformation of coordinates if F_{rs} is taken to be a covariant tensor. This is not true of the first of (10).

Since it is usually most convenient to use F_{rs} only (instead of F_{rs} and F^*_{rs}), we shall write down here for reference the complete set of Maxwell's equations in vacuo in the form

$$\boxed{F_{rs,s} = 0, \quad F_{mn,r} + F_{nr,m} + F_{rm,n} = 0.} \tag{15}$$

Note that in the second equation the suffixes are permuted cyclically.

We have actually eight equations in (15), but of these only six are to be regarded as independent. This is most easily seen by referring to the

equivalent equations (10), in which the left hand sides satisfy the identities

$$\frac{\partial}{\partial x_r}\left(\frac{\partial F_{rs}}{\partial x_s}\right) = 0, \quad \frac{\partial}{\partial x_r}\left(\frac{\partial F^*_{rs}}{\partial x_s}\right) = 0, \tag{16}$$

on account of the skew-symmetry of F_{rs} and F^*_{rs}.

§ 2. LORENTZ TRANSFORMATIONS OF THE ELECTRIC AND MAGNETIC 3-VECTORS

Let S and S' be two Galileian reference systems with coordinates x_r and x'_r respectively, connected by a Lorentz transformation as in IV–(24):

$$x'_r = A_{rs}x_s, \quad A_{rs}A_{rt} = \delta_{st}. \tag{17}$$

If the same electromagnetic field is represented by F_{rs} in S and by F'_{rs} in S', then (since F_{rs} is a tensor) we have

$$F'_{rs} = A_{rm}A_{sn}F_{mn} \tag{18}$$

as the formula of transformation of the electromagnetic tensor.

We can use (18) to express (E'_ϱ, H'_ϱ) in terms of (E_ϱ, H_ϱ) for any Lorentz transformation, but it is more interesting to work with the special Lorentz transformation of IV–(202):

$$\begin{gathered} x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma(t - vx/c^2), \\ \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}. \end{gathered} \tag{19}$$

It will be recalled that this transformation is quite general in the sense that the transformation from S to S' can always be put in this form by the cooperation of the observers in their choice of space axes; v is the velocity of S' relative to S, this relative velocity being along Ox (or Ox_1 in the indicial notation). For (19) the surviving elements of the matrix A_{mn} are, as in IV–(201),

$$\begin{gathered} A_{11} = \gamma, \qquad A_{14} = i\gamma v/c, \\ A_{22} = A_{33} = 1, \\ A_{41} = -i\gamma v/c, \quad A_{44} = \gamma, \\ A_{11}A_{44} - A_{14}A_{41} = 1. \end{gathered} \tag{20}$$

Then (18) gives explicitly

$$\left.\begin{aligned} F'_{14} &= A_{1m}A_{4n}F_{mn} = (A_{11}A_{44} - A_{14}A_{41})F_{14} = F_{14}, \\ F'_{24} &= A_{2m}A_{4n}F_{mn} = A_{41}F_{21} + A_{44}F_{24}, \\ F'_{34} &= A_{3m}A_{4n}F_{mn} = A_{41}F_{31} + A_{44}F_{34}, \\ F'_{23} &= A_{2m}A_{3n}F_{mn} = F_{23}, \\ F'_{31} &= A_{3m}A_{1n}F_{mn} = A_{11}F_{31} + A_{14}F_{34}, \\ F'_{12} &= A_{1m}A_{2n}F_{mn} = A_{11}F_{12} + A_{14}F_{42}. \end{aligned}\right\} \tag{21}$$

When we substitute from (6) and (20), we get the following formulae for *the transformation of the electric and magnetic 3-vectors under the simple Lorentz transformation* (19):

$$\left.\begin{aligned} E'_1 = E_1, \quad E'_2 = \gamma(E_2 - H_3 v/c), \quad E'_3 = \gamma(E_3 + H_2 v/c), \\ H'_1 = H_1, \quad H'_2 = \gamma(H_2 + E_3 v/c), \quad H'_3 = \gamma(H_3 - E_2 v/c). \end{aligned}\right\} \tag{22}$$

An interesting feature of this transformation is the mixing of the electric field with the magnetic field, and the generation of the one from the other as a result of the motion of the observer. If S observes a purely electric field E_ϱ (with $H_\varrho = 0$), then S' observes a mixed field with

$$\left.\begin{aligned} E'_1 = E_1, \quad E'_2 = \gamma E_2, \quad E'_3 = \gamma E_3, \\ H'_1 = 0, \quad H'_2 = \gamma E_3 v/c, \quad H'_3 = -\gamma E_2 v/c. \end{aligned}\right\} \tag{23}$$

Thus a purely electric field (or a purely magnetic field) is not a relativistic concept, since this pure character is not preserved under a Lorentz transformation.

It should be noted that the transformation given above is a consequence of the tensor character of F_{rs}, and depends in no way on the fact that F_{rs} satisfies Maxwell's equations.

In the special theory of relativity we attach great importance to quantities which are invariant under Lorentz transformations. Now it is easy to verify from (22) that

$$E'^2 - H'^2 = E^2 - H^2, \quad E'_\varrho H'_\varrho = E_\varrho H_\varrho, \tag{24}$$

and therefore the quantities

$$E^2 - H^2 \text{ and } E_\varrho H_\varrho = EH\cos\theta, \tag{25}$$

where θ is the angle between E_ϱ and H_ϱ, are *Lorentz invariants of an*

electromagnetic field. (There are no other invariants, except of course combinations of these two.)

We note that, in consequence, the relation $E = H$ is true for all Galileian observers if true for one, and that the electric and magnetic 3-vectors are orthogonal for all Galileian observers if orthogonal for one.

There is however a more fundamental way of obtaining these invariants [cf. VII–(61)]. It is obvious that $F_{rs}F_{rs}$ and $F_{rs}F^*_{rs}$ are invariants, and we have by (6)

$$\begin{aligned} -\tfrac{1}{2}F_{rs}F_{rs} &= -(F_{14}^2 + F_{24}^2 + F_{34}^2 + F_{23}^2 + F_{31}^2 + F_{12}^2) \\ &= E^2 - H^2, \end{aligned} \tag{26}$$

and

$$\begin{aligned} \tfrac{1}{4}F_{rs}F^*_{rs} &= \tfrac{1}{2}(F_{14}F^*_{14}+F_{24}F^*_{24}+F_{34}F^*_{34}+F_{23}F^*_{23}+F_{31}F^*_{31}+F_{12}F^*_{12}) \\ &= i(F_{14}F_{23} + F_{24}F_{31} + F_{34}F_{12}) = E_\varrho H_\varrho, \end{aligned} \tag{27}$$

and so the invariance of the two quantities in (25) follows directly.

The invariant combination I of VII–(62) reads

$$\begin{aligned} I &= \tfrac{1}{2}[(F_{rs}F_{rs})^2 + (F_{rs}F^*_{rs})^2]^{\frac{1}{2}} \\ &= (E^4 + H^4 + 2E^2H^2\cos 2\theta)^{\frac{1}{2}}. \end{aligned} \tag{28}$$

This is always real, and it vanishes if, and only if, $E = H$ and $\theta = \frac{1}{2}\pi$, i.e. if the electric and magnetic 3-vectors are equal and perpendicular to one another. We shall call such a field an electromagnetic *null field*; it is characterised by

$$F_{rs}F_{rs} = 0, \quad F_{rs}F^*_{rs} = 0, \tag{29}$$

or equivalently by

$$E = H, \quad E_\varrho H_\varrho = 0. \tag{30}$$

The expression *self-conjugate* field is also used.

§ 3. THE ENERGY TENSOR

In pre-relativistic physics we are familiar with the idea that an electromagnetic field possesses stress, momentum and energy. In VIII–§ 8 we studied these concepts for a continuous mechanical medium (regarded as the limit of a discrete system), using an energy tensor T_{rs}. The details of that work do not concern us here; the

salient points are that T_{rs} was symmetric ($T_{rs} = T_{sr}$), that it satisfied the conservation equations VIII–(117), viz.

$$\frac{\partial T_{rs}}{\partial x_s} = 0, \tag{31}$$

and that the physical (mechanical) meanings of its ten components were as set out in the equations VIII–(130) to (136).

We now write down, as *the energy tensor of an electromagnetic field in vacuo*,

$$T_{rs} = \tfrac{1}{2}c^{-2}(F_{mr}F_{ms} + F^*_{mr}F^*_{ms}). \tag{32}$$

The factor c^{-2} is inserted here to give the components of T_{rs} the dimensions of a mass; this will become clear when we consider later material charged particles in connection with the field, and the reason for the factor $\frac{1}{2}$ will then become evident [X–§9].

We observe that T_{rs} as given by (32) is symmetric. To show that it satisfies the conservation equations (31), we note that, on account of Maxwell's equations (10),

$$2c^2T_{rs,s} = F_{mr,s}F_{ms} + F^*_{mr,s}F^*_{ms}. \tag{33}$$

But by (14), and the similar starred equation equivalent to the first of (10),

$$F_{mr,s} = -F_{rs,m} - F_{sm,r}, \quad F^*_{mr,s} = -F^*_{rs,m} - F^*_{sm,r}. \tag{34}$$

Substituting these in (33), we obtain

$$\begin{aligned} 2c^2T_{rs,s} = &-F_{rs,m}F_{ms} + \tfrac{1}{2}(F_{ms}F_{ms})_{,r} \\ &- F^*_{rs,m}F^*_{ms} + \tfrac{1}{2}(F^*_{ms}F^*_{ms})_{,r}. \end{aligned} \tag{35}$$

But, as is easily verified,

$$F_{ms}F_{ms} + F^*_{ms}F^*_{ms} = 0, \tag{36}$$

and so the terms with the factor $\frac{1}{2}$ disappear from (35); with further use of (10), (35) may be written

$$\begin{aligned} 2c^2T_{rs,s} &= -(F_{rs}F_{ms} + F^*_{rs}F^*_{ms})_{,m} \\ &= -2c^2T_{rm,m}, \end{aligned} \tag{37}$$

and so

$$T_{rs,s} = 0. \tag{38}$$

Thus T_{rs}, as given by (32), *satisfies the conservation equations* (31).

To express the energy tensor in terms of F_{rs} alone, we note the identity[1]

$$F_{mr}F_{ms} - F^*_{mr}F^*_{ms} = \tfrac{1}{2}\delta_{rs}F_{mn}F_{mn}, \tag{39}$$

which gives, as an alternative expression to (32),

$$T_{rs} = c^{-2}(F_{mr}F_{ms} - \tfrac{1}{4}\delta_{rs}F_{mn}F_{mn}). \tag{40}$$

Using (6), we now write out the ten components of T_{rs} in terms of the electric and magnetic 3-vectors:

$$\left.\begin{aligned}
T_{11} &= \tfrac{1}{2}c^{-2}(-E_1^2 + E_2^2 + E_3^2 - H_1^2 + H_2^2 + H_3^2),\\
T_{22} &= \tfrac{1}{2}c^{-2}(E_1^2 - E_2^2 + E_3^2 + H_1^2 - H_2^2 + H_3^2),\\
T_{33} &= \tfrac{1}{2}c^{-2}(E_1^2 + E_2^2 - E_3^2 + H_1^2 + H_2^2 - H_3^2),\\
T_{44} &= -\tfrac{1}{2}c^{-2}(E^2 + H^2),\\
T_{23} &= -c^{-2}(E_2E_3 + H_2H_3), \quad T_{31} = -c^{-2}(E_3E_1 + H_3H_1),\\
T_{12} &= -c^{-2}(E_1E_2 + H_1H_2),\\
T_{14} &= ic^{-2}(E_2H_3 - E_3H_2), \quad T_{24} = ic^{-2}(E_3H_1 - E_1H_3),\\
T_{34} &= ic^{-2}(E_1H_2 - E_2H_1).
\end{aligned}\right\} \tag{41}$$

We assign to the components of T_{rs} mechanical meanings in accordance with VIII–(130) to (136). It is useful to introduce the *Poynting vector* defined by

$$P_\varrho = c\varepsilon_{\varrho\mu\nu}E_\mu H_\nu = c(E \times H)_\varrho, \tag{42}$$

this last symbol being the usual vector product. Then by VIII–(132)

$$\begin{aligned}&\text{flux of relative energy per unit time}\\ &\text{per unit area normal to } n_\varrho = -ic^3T_{4\varrho}n_\varrho = P_\varrho n_\varrho;\end{aligned} \tag{43}$$

by VIII–(134)

$$\text{relative density of relative momentum} = -icT_{\varrho 4} = c^{-2}P_\varrho; \tag{44}$$

and by VIII–(135)

$$\text{relative density of relative energy} = -c^2T_{44} = \tfrac{1}{2}(E^2 + H^2). \tag{45}$$

The components $T_{\varrho\sigma}$ are interpreted in terms of flux of relative momentum as in VIII–(130), or as stress — what used to be called "stress

[1] This is equivalent to VII–(142), as we see by substituting Y^*_{rs} for Y_{rs} and noting that, by VII–(141), $(Y^*_{rs})^* = -Y_{rs}$. The following interesting identity, due to RUSE [1936], is easily established:

$$16c^4T_{mr}T_{ms} = \delta_{rs}[(F_{mn}F_{mn})^2 + (F_{mn}F^*_{mn})^2].$$

in the ether", but it is best to avoid that dangerous word in relativity.

To sum up, we regard an electromagnetic field in vacuo as having mechanical properties (stress, momentum and energy) as described above. In that way an electromagnetic field is similar to the mechanical continuum discussed in Ch. VIII, but there are important differences arising from the peculiar form of the electromagnetic energy tensor. These will be discussed below. We note here in particular that *the trace of the electromagnetic energy tensor vanishes*:

$$T_{rr} = 0. \tag{46}$$

This follows immediately from (32) or (40); it may be verified directly from (41).

We recall that, by VIII–(188), $T_{rr} = 0$ for a mechanical system composed of photons only, but that $T_{rr} < 0$ if material particles are present.

§ 4. EIGEN VALUES AND PRINCIPAL DIRECTIONS FOR THE ELECTROMAGNETIC ENERGY TENSOR

In the theory of a continuous mechanical system the eigen values of the energy tensor and its principal directions played an important role; in particular, in VIII–§ 12 we defined the mean motion by the timelike principal direction. Being now in possession of the energy tensor of an electromagnetic field, as in (32) or (40), it is interesting to investigate its eigen values and principal directions.

The equations satisfied by any principal vector ξ_r and the eigen value X associated with it are

$$T_{rs}\xi_s = X\xi_r, \tag{47}$$

the determinantal equation for X being

$$\det(T_{rs} - X\delta_{rs}) = 0. \tag{48}$$

We have under consideration some definite event and some Galileian observer. Without changing the frame of reference, the space axes in it can be rotated so that the Ox_3 axis is perpendicular to both the vectors E_ϱ, H_ϱ at the event in question. We have then

$$E_3 = H_3 = 0, \tag{49}$$

and this leads to considerable simplification in (41); we have

$$\left.\begin{aligned} T_{11} &= -T_{22} = -\tfrac{1}{2}c^{-2}(E_1^2 - E_2^2 + H_1^2 - H_2^2), \\ T_{33} &= -T_{44} = \tfrac{1}{2}c^{-2}(E^2 + H^2), \\ T_{23} &= T_{31} = 0, \quad T_{12} = -c^{-2}(E_1E_2 + H_1H_2), \\ T_{14} &= T_{24} = 0, \quad T_{34} = ic^{-2}(E_1H_2 - E_2H_1). \end{aligned}\right\} \tag{50}$$

The equation (48) now reads

$$\begin{vmatrix} T_{11} - X & T_{12} & 0 & 0 \\ T_{21} & T_{22} - X & 0 & 0 \\ 0 & 0 & T_{33} - X & T_{34} \\ 0 & 0 & T_{43} & T_{44} - X \end{vmatrix} = 0, \tag{51}$$

and this gives the two quadratic equations

$$\begin{aligned} (X - T_{11})(X - T_{22}) - T_{12}^2 &= 0, \\ (X - T_{33})(X - T_{44}) - T_{34}^2 &= 0. \end{aligned} \tag{52}$$

By (50) these reduce to

$$X^2 = T_{11}^2 + T_{12}^2 = \tfrac{1}{4}c^{-4}[(E_1^2 - E_2^2 + H_1^2 - H_2^2)^2 + 4(E_1E_2 + H_1H_2)^2],$$
$$X^2 = T_{33}^2 + T_{34}^2 = \tfrac{1}{4}c^{-4}[(E^2 + H^2)^2 - 4(E_1H_2 - E_2H_1)^2].$$

By direct calculation it is easy to see that these two right hand sides are equal to one another, and that their common value may be written

$$\tfrac{1}{4}c^{-4}[(E^2 - H^2)^2 + 4(E_1H_1 + E_2H_2)^2],$$

which is, by (26) and (27),

$$\tfrac{1}{16}c^{-4}[(F_{rs}F_{rs})^2 + (F_{rs}F_{rs}^*)^2].$$

Therefore *the four eigen values of the electromagnetic energy tensor are real and equal in pairs, being*

$$X = \pm k, \quad X = \pm k, \tag{53}$$

where

$$\begin{aligned} k &= \tfrac{1}{4}c^{-2}[(F_{rs}F_{rs})^2 + (F_{rs}F_{rs}^*)^2]^{\frac{1}{2}} \\ &= \tfrac{1}{2}c^{-2}(E^4 + H^4 + 2E^2H^2\cos 2\theta)^{\frac{1}{2}}, \end{aligned} \tag{54}$$

θ being the angle between E_ϱ and H_ϱ. Although we used special space axes to simplify the algebra, the result (53) is obviously invariant and true for all axes.

The invariant k (differing from I as in (28) only by the factor $\frac{1}{2}c^{-2}$) has a simple interpretation in terms of minimum energy-density. We have the elementary inequality

$$E^4 + H^4 + 2E^2H^2 \cos 2\theta \leqslant (E^2 + H^2)^2,$$

equality occurring if, and only if, $\theta = 0$ or π (i.e. if E_ϱ and H_ϱ have the same direction or opposite directions). This is equivalent to

$$kc^2 \leqslant \tfrac{1}{2}(E^2 + H^2). \tag{55}$$

If $k = 0$, this tells us nothing. If $k \neq 0$, it tells us that kc^2 *is the minimum of all the energy-densities of the field when observed by all possible Galileian observers*; the minimum is attained when the observer uses a canonical frame, as described in § 5, for then $\theta = 0$ or π.

To find the principal vectors ξ_r, we turn back to (47) and again use the space axes for which (49) holds. Then (47) read explicitly

$$\left.\begin{aligned} T_{11}\xi_1 + T_{12}\xi_2 - X\xi_1 &= 0, \\ T_{21}\xi_1 + T_{22}\xi_2 - X\xi_2 &= 0, \\ T_{33}\xi_3 + T_{34}\xi_4 - X\xi_3 &= 0, \\ T_{43}\xi_3 + T_{44}\xi_4 - X\xi_4 &= 0, \end{aligned}\right\} \tag{56}$$

where $X = \pm k$. It is necessary to treat separately the cases of a null field and a general field; for a null field $k = 0$, by (29).

Case (i): *Null field* ($k = 0$).

Since the electric and magnetic 3-vectors are equal and perpendicular in this case, we may further specialise our space axes by choosing Ox_1 in the direction of E_ϱ and Ox_2 in the direction of H_ϱ, the correct orientation of the axes being taken care of by the sense of Ox_3. Then we have

$$E = E_1 = H_2 = H, \quad E_2 = H_1 = 0, \tag{57}$$

and so by (50) the only surviving components of the energy tensor are

$$T_{33} = -T_{44} = c^{-2}E^2, \quad T_{34} = T_{43} = ic^{-2}E^2. \tag{58}$$

Remembering that $X = \pm k = 0$ in the present case, we see that the first two equations of (56) disappear, and we are left with only two equations,

$$\xi_3 + i\xi_4 = 0, \quad i\xi_3 - \xi_4 = 0, \tag{59}$$

which are in fact only a single equation. This high degree of degeneracy

results of course from the equality of all four eigen values. We are to regard as principal vectors the totality of vectors satisfying (59); we note that

$$\xi_3^2 + \xi_4^2 = 0, \quad \xi_r\xi_r = \xi_1^2 + \xi_2^2 \geqslant 0, \tag{60}$$

so that all the principal directions are spacelike or null.

Among the principal directions there is just one null direction; it is given by

$$\xi_1 = 0, \quad \xi_2 = 0, \quad \xi_3 = 1, \quad \xi_4 = i. \tag{61}$$

We could of course reverse this principal null direction, and so get another, but the distinction between these two is not significant; we shall think only of (61), which points into the future.

To see the relationship between the principal null direction (61) and the totality of principal directions, let us write ξ'_r for the vector (61) and ξ_r for any other principal vector. Then we have

$$\xi_r\xi'_r = \xi_3\xi'_3 + \xi_4\xi'_4 = 0, \tag{62}$$

since (59) holds for both ξ_r and ξ'_r. Now, if ξ_r are regarded as running coordinates, $\xi_r\xi'_r = 0$ is the equation of the 3-flat tangent to the null cone $\xi_r\xi_r = 0$ along the generator ξ'_r. Hence we may sum up by stating that *in the case of a null field (for which the electric and magnetic 3-vectors are equal and perpendicular) the energy tensor has one and only one principal null direction, and the totality of all principal directions form the 3-flat tangent to the null cone along the principal null direction.* This is indicated in Fig. 1.

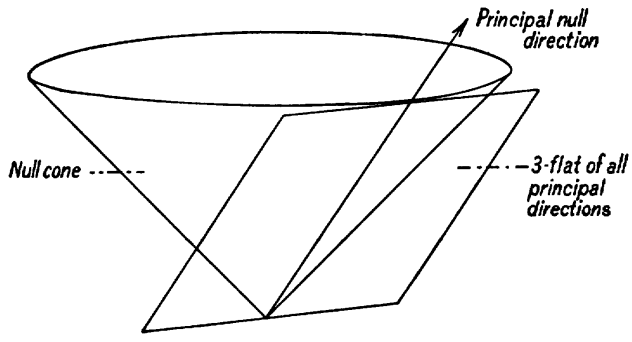


Fig. 1 – Principal directions of T_{rs} for a null field ($k = 0$)

To interpret the principal null direction kinematically, we recall that any direction in space-time corresponds to the motion of a point in the observer's space, a null direction corresponding to motion with the speed of light. The 3-velocity corresponding to a vector ξ_r is

$$u_\varrho = \frac{dx_\varrho}{dt} = ic\frac{\xi_\varrho}{\xi_4}. \tag{63}$$

Thus (61) corresponds to motion in the direction of Ox_3 with the speed

of light. But Ox_3 has the direction of the vector product $(E \times H)_\varrho$, which has the direction of the Poynting vector (42), and so we may say that *in the case of a null field the single principal null direction corresponds to the motion of a point in the direction of the Poynting vector with the speed of light.*

To get a kinematical interpretation of the totality of principal directions, we note that a 3-flat corresponds to the history of the plane (2-flat) which is the intersection of the 3-flat in question with $t = \text{const}$. Now for the principal directions, ξ_1 and ξ_2 are arbitrary, whereas the relation (59) between ξ_3 and ξ_4 is the same for them all. But this relation determines the component of velocity in the direction of the Poynting vector, and hence it follows that *for a null field the totality of principal directions correspond to the motion, in the space of any Galileian observer, of an elementary plane wave travelling with the speed of light in the direction of the Poynting vector.*

Case (ii): *General field* $(k \neq 0)$.

We pass on now to consider the principal directions defined by the energy tensor of a general electromagnetic field $(k \neq 0)$. The argument applies if the electric and magnetic 3-vectors are equal in magnitude, or if they are perpendicular; only the case where they are both equal in magnitude and perpendicular is ruled out (Case (i) above).

With Ox_3 perpendicular to E_ϱ and H_ϱ, we have the equations (56). Now the determinantal equation of the first two of (56) is the first equation of (52), and its roots are $\pm k$; and the determinantal equation of the last two of (56) is the second of (52), also with roots $\pm k$. Consequently, if we put $X = k$, the first of (56) implies the second and the third implies the fourth; the same is true if we put $X = -k$. Thus any direction in the 2-flat

$$(X = k) \quad T_{11}\xi_1 + T_{12}\xi_2 - k\xi_1 = 0, \quad T_{33}\xi_3 + T_{34}\xi_4 - k\xi_3 = 0, \quad (64)$$

is a principal direction, and any direction in the 2-flat

$$(X = -k) \quad T_{11}\xi_1 + T_{12}\xi_2 + k\xi_1 = 0, \quad T_{33}\xi_3 + T_{34}\xi_4 + k\xi_3 = 0, \quad (65)$$

is a principal direction.[1] In fact, *the principal directions are confined to*

[1] If an equation in (64) or (65) should degenerate into an identity, we must replace it by a non-degenerate equation from (56). Note that the two 2-flats intersect only at $\xi_r = 0$; no other values of ξ_r satisfy (56) for $X = k$ and for $X = -k$.

two 2-flats, but are otherwise indeterminate. Since T_{rs} is symmetric, it follows from VIII–(151) that these 2-flats are orthogonal, in the sense that every vector in one is orthogonal to every vector in the other.

Let us now investigate *principal null directions* for this case ($k \neq 0$). They will lie in one of the above 2-flats, but not in both, since the 2-flats are orthogonal. To see which they lie in, we write, as in (47),

$$T_{rs}\xi_s = X\xi_r, \quad T_{rs}\eta_s = X\eta_r, \tag{66}$$

ξ_r and η_r being any two principal vectors, both corresponding to the same value of X. To explore the 2-flat containing ξ_r and η_r, we put

$$\zeta_r = \alpha\xi_r + \beta\eta_r. \tag{67}$$

Then (66) implies

$$T_{rs}\zeta_s = X\zeta_r, \tag{68}$$

and hence

$$T_{rs}\zeta_r\zeta_s = X\zeta_r\zeta_r. \tag{69}$$

Now if the 2-flat of ξ_r and η_r contains null directions (i.e. cuts the null cone), then there exist values of α and β such that ζ_r is timelike. Choose it for time-axis, so that $\zeta_\varrho = 0$, $\zeta_4 = i$. Then (69) becomes

$$-X = -T_{44} = \tfrac{1}{2}c^{-2}(E^2 + H^2) > 0 \tag{70}$$

by (41). Therefore *the principal 2-flat containing null directions is* (65), *corresponding to a negative* X ($X = -k$). The arrangement of principal directions is indicated in Fig. 2.

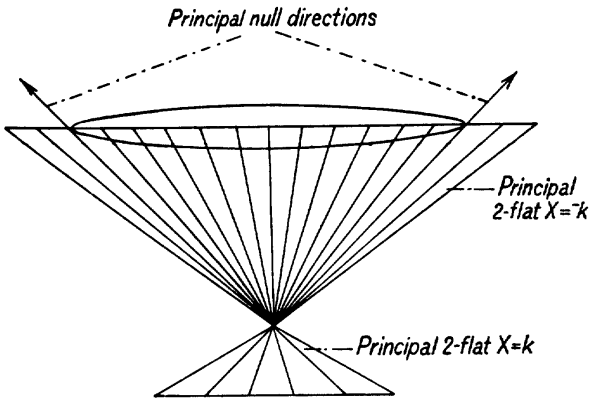


Fig. 2 – Principal directions of T_{rs} for a general field ($k \neq 0$)

The principal null directions are easily found explicitly by putting $X = -k$ in (56) and solving these equations together with $\xi_r\xi_r = 0$. Omitting the details of the algebra [cf. SYGNE (1935b)], we shall state the results, expressing the principal null vectors in the form

$$\xi_1 = l_1, \quad \xi_2 = l_2, \quad \xi_3 = l_3, \quad \xi_4 = i,$$
$$l_1^2 + l_2^2 + l_3^2 = 1, \tag{71}$$

so that each corresponds to the motion of a point with the speed of

light in a direction with direction cosines l_ϱ. We find, if $T_{11} + k > 0$,

$$l_1 = -\, l_2 \frac{T_{12}}{T_{11} + k}, \quad l_2 = \pm \left(\frac{T_{11} + k}{T_{33} + k} \right)^{\frac{1}{2}}, \quad l_3 = -\, \frac{iT_{34}}{T_{33} + k}, \tag{72}$$

and, if $T_{22} + k > 0$,

$$l_1 = \pm \left(\frac{T_{22} + k}{T_{33} + k} \right)^{\frac{1}{2}}, \quad l_2 = -\, l_1 \frac{T_{12}}{T_{22} + k}, \quad l_3 = -\, \frac{iT_{34}}{T_{33} + k}. \tag{73}$$

Here we are using space axes with Ox_3 perpendicular to E_ϱ and H_ϱ, so that (50) hold. Thus T_{11} and T_{22} have opposite signs, and so, since k is positive, either $T_{11} + k > 0$ or $T_{22} + k > 0$. Further, by (50), $T_{33} + k$ is positive, so that we are assured of the reality of either (72) or (73).

In particular, if $E = H$ (but $E_\varrho H_\varrho \neq 0$) and if we take Ox_1 and Ox_2 along $H_\varrho + E_\varrho$ and $H_\varrho - E_\varrho$ respectively, then, if θ is acute,

$$l_1 = \pm\, (1 - \tan^2 \tfrac{1}{2}\theta)^{\frac{1}{2}}, \quad l_2 = 0, \quad l_3 = \tan \tfrac{1}{2}\theta, \tag{74}$$

and, if θ is obtuse,

$$l_1 = 0, \quad l_2 = \pm\, (1 - \cot^2 \tfrac{1}{2}\theta)^{\frac{1}{2}}, \quad l_3 = \cot \tfrac{1}{2}\theta. \tag{75}$$

On the other hand, if $E_\varrho H_\varrho = 0$ (but $E \neq H$), and if we take Ox_1, Ox_2 in the directions of E_ϱ, H_ϱ respectively, then, if $E > H$,

$$l_1 = \pm\, (E^2 - H^2)^{\frac{1}{2}}/E, \quad l_2 = 0, \quad l_3 = H/E, \tag{76}$$

and, if $E < H$,

$$l_1 = 0, \quad l_2 = \pm\, (H^2 - E^2)^{\frac{1}{2}}/H, \quad l_3 = E/H. \tag{77}$$

§ 5. THE CANONICAL FORMS FOR AN ELECTROMAGNETIC FIELD AT AN EVENT

In § 4 we used any assigned Galileian frame of reference, rotating the space-axes to suit our needs. Now we ask: What frame of reference gives the simplest possible appearance to a field specified by a skew-symmetric tensor F_{rs}? To answer this question we make use of the principal null directions discussed in the preceding section, again separating the cases of (i) a null field ($k = 0$) and (ii) a general field ($k \neq 0$). It must be emphasised that the following reductions hold at one event only, unless we happen to be dealing with a constant field; that event may be taken as the origin of space-time for convenience.

Case (i): *Null field* ($k = 0$).

For $k = 0$ we have a single principal null direction (Fig. 1); taking any timelike vector for Ox_4, let us choose Ox_1 so that this principal null vector ξ_r is contained in the 2-flat Ox_1x_4 with $\xi_1 > 0$. The axes of Ox_2 and Ox_3 are chosen arbitrarily, except of course for orthogonality conditions. The resulting axes we call *canonical axes* (Fig. 3).

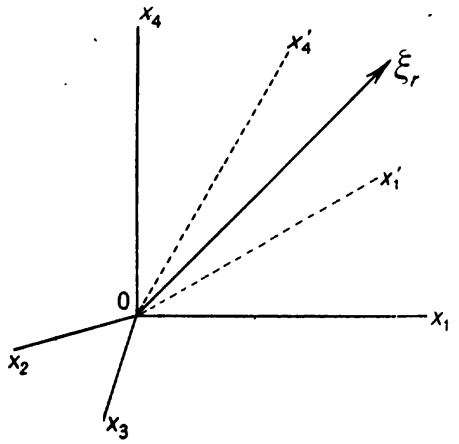


Fig. 3 – Canonical axes for a null field ($k = 0$)

We may then write for the components of the principal null vector

$$\xi_1 = 1, \quad \xi_2 = 0, \quad \xi_3 = 0, \quad \xi_4 = i. \tag{78}$$

These values (with $X = 0$ since we are dealing with the case $k = 0$) are to satisfy (47), and so we get

$$T_{11} + iT_{14} = 0, \quad T_{21} + iT_{24} = 0, \quad T_{31} + iT_{34} = 0, \quad T_{41} + iT_{44} = 0. \tag{79}$$

By (41) the first and last of these give

$$\tfrac{1}{2}(-E_1^2 + E_2^2 + E_3^2 - H_1^2 + H_2^2 + H_3^2) - (E_2H_3 - E_3H_2) = 0, \quad (E_2H_3 - E_3H_2) - \tfrac{1}{2}(E^2 + H^2) = 0; \tag{80}$$

adding these, we get $E_1^2 + H_1^2 = 0$ and hence

$$E_1 = H_1 = 0; \tag{81}$$

either of (80) then gives

$$(E_2 - H_3)^2 + (E_3 + H_2)^2 = 0, \tag{82}$$

which implies

$$E_2 = H_3, \quad E_3 = -H_2. \tag{83}$$

Thus the observer who uses the canonical axes finds the electric and magnetic 3-vectors in the plane Ox_2x_3. The fact that these vectors are equal and perpendicular is verified by (83), but of course it is a known consequence of the condition $k = 0$ for all frames of reference.

The simplification produced in a null field by the use of canonical axes is really very slight; we have merely got rid of E_1 and H_1. Let us now study the effect of changing the axes, keeping Ox_2 and Ox_3 fixed and rotating Ox_1 and Ox_4 to Ox_1' and Ox_4' in the 2-flat containing ξ_r. The new axes form another canonical set, and the transformation is a

simple Lorentz transformation as in (19). By (22) the field changes as follows:

$$\begin{aligned} E_1' &= E_1, \quad E_2' = \gamma(E_2 - H_3 v/c), \quad E_3' = \gamma(E_3 + H_2 v/c), \\ H_1' &= H_1, \quad H_2' = \gamma(H_2 + E_3 v/c), \quad H_3' = \gamma(H_3 - E_2 v/c), \end{aligned} \tag{84}$$

and so, by (81) and (83),

$$\begin{aligned} E_1' &= 0, \quad E_2' = E_2\left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}}, \quad E_3' = E_3\left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}}, \\ H_1' &= 0, \quad H_2' = H_2\left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}}, \quad H_3' = H_3\left(\frac{1 - v/c}{1 + v/c}\right)^{\frac{1}{2}}. \end{aligned} \tag{85}$$

Thus the effect of this transformation is to multiply all the components of E_ϱ and H_ϱ by the same factor $(1 - v/c)^{\frac{1}{2}}/(1 + v/c)^{\frac{1}{2}}$.

We note that if we let $v \to c$ in (85), then the components of the field all tend to zero; and if we let $v \to -c$, they tend to infinity, except of course E_1' and H_1', which remain zero. This means that *a null field can be made as small as we please, or as large as we please, by changing the frame of reference*; for that reason we should be cautious in relativity of speaking of a *weak* null field or a *strong* null field, since these concepts are not Lorentz invariant.

Case (ii): *General field* ($k \neq 0$).

The canonical reduction of a general field ($k \neq 0$) is more interesting than that of a null field ($k = 0$). Now we have two principal null vectors, say ξ_r and η_r, and these define a 2-flat (Fig. 2). As canonical axes, we choose Ox_1 and Ox_4 in this 2-flat (Fig. 4), there being one degree of freedom in this choice (there were three degrees of freedom in the case $k = 0$, for then Ox_4 was arbitrary).

Then the components of the principal null vectors may be written

$$\begin{aligned} \xi_1 &= 1, \quad \eta_1 = -1, \\ \xi_2 &= 0, \quad \eta_2 = 0, \\ \xi_3 &= 0, \quad \eta_3 = 0, \\ \xi_4 &= i, \quad \eta_4 = i. \end{aligned} \tag{86}$$

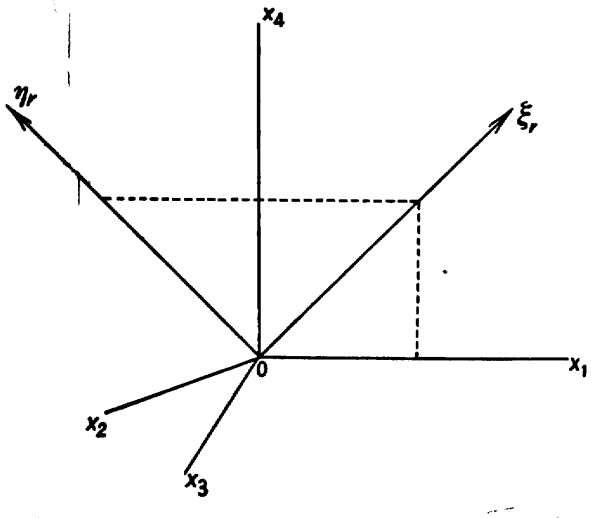


Fig. 4 – Canonical axes for a general field ($k \neq 0$)

These components are to satisfy (47) with $X = -k$, and so we have

$$\begin{aligned} \varepsilon T_{11} + iT_{14} + \varepsilon k &= 0, \\ \varepsilon T_{21} + iT_{24} &= 0, \\ \varepsilon T_{31} + iT_{34} &= 0, \\ \varepsilon T_{41} + iT_{44} + ik &= 0, \end{aligned} \tag{87}$$

for $\varepsilon = 1$ and $\varepsilon = -1$. Hence

$$T_{31} = T_{12} = T_{14} = T_{24} = T_{34} = 0, \quad T_{11} = T_{44} = -k. \tag{88}$$

But by (41) we have

$$c^2(T_{11} - T_{44}) = E_2^2 + E_3^2 + H_2^2 + H_3^2, \tag{89}$$

and, since this vanishes by (88), we have

$$E_2 = E_3 = H_2 = H_3 = 0. \tag{90}$$

Thus the field has been greatly simplified by the use of canonical axes; *we have electric and magnetic 3-vectors with a common line* (Ox_1); their senses on this line may be the same or opposite.

By analogy with statics, we may call a field with E_ϱ and H_ϱ on a single line an *electromagnetic wrench*; the corresponding formula is

$$pE_\varrho = qH_\varrho, \tag{91}$$

where p and q are scalars, and we may call the ratio p/q the *pitch* of the wrench. We have established that *at any event where an electromagnetic field is not null, it can be reduced to an electromagnetic wrench by the use of canonical axes. This reduction is not possible in the case of a null field.*

As mentioned above, there is a single degree of freedom in the choice of canonical axes Ox_1, Ox_4, and two sets of canonical axes are related by the simple Lorentz transformation (19). But this produces no change at all in the field, since (84) and (90) give

$$E_1' = E_1, \quad H_1' = H_1, \quad E_2' = E_3' = H_2' = H_3' = 0. \tag{92}$$

Thus, for a field which is not null, we may speak of *absolute electric and magnetic strengths* at an event, since E_1 and H_1 are not changed in (92), being the same for all canonical axes; and we can express these

absolute strengths in terms of the invariants, for (26), (27) and (90) give

$$E_1^2 - H_1^2 = -\tfrac{1}{2}F_{rs}F_{rs}, \quad E_1H_1 = \tfrac{1}{4}F_{rs}F_{rs}^*, \tag{93}$$

and so, by simple algebra with k as in (54), we get for the squares of the absolute strengths

$$E_1^2 = kc^2 - \tfrac{1}{4}F_{rs}F_{rs}, \quad H_1^2 = kc^2 + \tfrac{1}{4}F_{rs}F_{rs}. \tag{94}$$

Or, if we use the second form of k in (54), we can express the absolute electric and magnetic strengths (the magnitudes of the vectors forming the electromagnetic wrench) as follows in terms of the field strengths E, H as observed by an arbitrary Galileian observer, and the angle θ between the vectors:

$$\begin{aligned}(E_1^2)_{abs} &= \tfrac{1}{2}(E^4 + H^4 + 2E^2H^2\cos 2\theta)^{\frac{1}{2}} + \tfrac{1}{2}(E^2 - H^2),\\ (H_1^2)_{abs} &= \tfrac{1}{2}(E^4 + H^4 + 2E^2H^2\cos 2\theta)^{\frac{1}{2}} - \tfrac{1}{2}(E^2 - H^2).\end{aligned} \tag{95}$$

We get a simple check on these formulae by putting in turn $E = 0$ and $H = 0$, in each of which cases the field is an electromagnetic wrench for the observer.

Let us sum up with two statements:

(*a*) A null field cannot be reduced to an electromagnetic wrench; in all Galileian frames it has electric and magnetic 3-vectors which are equal in magnitude and perpendicular, the common magnitude being anything from zero to infinity according to the frame chosen.

(*b*) A field which is not null can always be reduced to an electromagnetic wrench, with E_ϱ and H_ϱ on the same line, their magnitudes being definite quantities as shown in (95) in terms of the values of E, H, θ for an arbitrary frame.

In the case $k = 0$, we can simplify (81), (82) to

$$E_1 = E_3 = H_1 = H_2 = 0, \quad E_2 = H_3 \tag{96}$$

by merely rotating the axes Ox_2, Ox_3; indeed the possibility of this simplification is obvious, because $k = 0$ implies that E_ϱ and H_ϱ are equal in magnitude and perpendicular, and any Galileian observer can obtain (96) by choosing Ox_2 along E_ϱ and Ox_3 along H_ϱ. With the values (96) inserted in (6) and (7), the matrices F_{rs} and F_{rs}^* take the following *canonical forms* for $k = 0$ (the first suffix indicates the row, the second the column):

$$\left.\begin{aligned} F_{rs} &= \begin{pmatrix} 0 & E_2 & 0 & 0 \\ -E_2 & 0 & 0 & -iE_2 \\ 0 & 0 & 0 & 0 \\ 0 & iE_2 & 0 & 0 \end{pmatrix} \\ F^*_{rs} &= \begin{pmatrix} 0 & 0 & -E_2 & 0 \\ 0 & 0 & 0 & 0 \\ E_2 & 0 & 0 & iE_2 \\ 0 & 0 & -iE_2 & 0 \end{pmatrix} \end{aligned}\right\} \tag{97}$$

By (90) we have the following *canonical forms* for $k \neq 0$:

$$\left.\begin{aligned} F_{rs} &= \begin{pmatrix} 0 & 0 & 0 & -iE_1 \\ 0 & 0 & H_1 & 0 \\ 0 & -H_1 & 0 & 0 \\ iE_1 & 0 & 0 & 0 \end{pmatrix} \\ F^*_{rs} &= \begin{pmatrix} 0 & 0 & 0 & iH_1 \\ 0 & 0 & E_1 & 0 \\ 0 & -E_1 & 0 & 0 \\ -iH_1 & 0 & 0 & 0 \end{pmatrix} \end{aligned}\right\} \tag{98}$$

We note that in (97) E_2 is not determined absolutely by the field; it depends on the frame of reference, and may take any value from $-\infty$ to $+\infty$. On the other hand, E_1 and H_1 in (98) are determined, except for an ambiguity of sign. To make this question of sign clearer, let us introduce the two real non-negative invariants ω, ω^* defined by

$$\left.\begin{aligned} \omega &= (kc^2 - \tfrac{1}{4}F_{rs}F_{rs})^{\frac{1}{2}} = (kc^2 + \tfrac{1}{4}F^*_{rs}F^*_{rs})^{\frac{1}{2}}, \\ \omega^* &= (kc^2 + \tfrac{1}{4}F_{rs}F_{rs})^{\frac{1}{2}} = (kc^2 - \tfrac{1}{4}F^*_{rs}F^*_{rs})^{\frac{1}{2}}, \end{aligned}\right\} \tag{99}$$

the positive square roots being understood, as always. Then, by (93) and (94), we have

$$E_1 = \varepsilon\omega, \quad H_1 = \varepsilon^*\omega^*, \tag{100}$$

where ε and ε^* may take the values ± 1, subject to

$$\varepsilon\varepsilon^* F_{rs}F^*_{rs} \geqslant 0. \tag{101}$$

§ 6. EIGEN PROPERTIES OF THE TENSORS F_{rs} AND F^*_{rs}

Consider the two sets of equations:

$$F_{rs}\eta_s = Y\eta_r \tag{102}$$

and

$$F^*_{rs}\zeta_s = Z\zeta_r. \tag{103}$$

The eigen properties associated with them are most easily investigated by using canonical axes and the canonical matrices (97), (98). Accordingly we separate the two cases.

Case (i): *Null field* ($k = 0$).

When we substitute from (97), (102) and (103) read

$$\begin{aligned} E_2\eta_2 &= Y\eta_1, & -E_2\zeta_3 &= Z\zeta_1, \\ -E_2\eta_1 - iE_2\eta_4 &= Y\eta_2, & 0 &= Z\zeta_2, \\ 0 &= Y\eta_3, & E_2\zeta_1 + iE_2\zeta_4 &= Z\zeta_3, \\ iE_2\eta_2 &= Y\eta_4; & -iE_2\zeta_3 &= Z\zeta_4. \end{aligned} \tag{104}$$

From these equations we find $Y = 0$ and $Z = 0$; *for a null field the eigen values of* F_{rs} *and* F^*_{rs} *are all zero.*

Substituting $Y = 0$, $Z = 0$ in (104), we see that the principal vectors η_r of F_{rs} are to satisfy

$$\eta_2 = 0, \quad \eta_1 + i\eta_4 = 0, \tag{105}$$

and those of F^*_{rs} are to satisfy

$$\zeta_3 = 0, \quad \zeta_1 + i\zeta_4 = 0. \tag{106}$$

We note that (105) defines a 2-flat, and so does (106); these 2-flats cut along the vector $(1, 0, 0, i)$, which is in fact the principal null vector of T_{rs} [cf. (78)]. The principal 2-flats of F_{rs} and F^*_{rs} are contained in the principal 3-flat of T_{rs} (Fig. 1). Note that there is a single direction (a null direction) which ~~is principal for F_{rs}, F^*_{rs} and T_{rs}~~.

The principal 2-flat of F_{rs} is represented kinematically by the motion, with the speed of light, of a linear element having the direction of the magnetic 3-vector; the principal 2-flat of F^*_{rs} has the same interpretation, with "electric" substituted for "magnetic".

Case (ii): *General field* ($k \neq 0$).

For a general field we substitute the canonical matrices (98) in (102) and (103); this gives

$$\begin{aligned} -iE_1\eta_4 &= Y\eta_1, & iH_1\zeta_4 &= Z\zeta_1, \\ H_1\eta_3 &= Y\eta_2, & E_1\zeta_3 &= Z\zeta_2, \\ -H_1\eta_2 &= Y\eta_3, & -E_1\zeta_2 &= Z\zeta_3, \\ iE_1\eta_1 &= Y\eta_4; & -iH_1\zeta_1 &= Z\zeta_4. \end{aligned} \tag{107}$$

Elimination of the vectors gives the eigen values

$$Y = \pm E_1, \quad \pm iH_1; \quad Z = \pm H_1, \quad \pm iE_1. \tag{108}$$

Although imaginary eigen values are significant mathematically, in exploring the geometry of the electromagnetic field we are interested only in vectors η_r, ζ_r which are real for real coordinate systems, or equivalently have their first three components real and the fourth a pure imaginary in Minkowskian coordinates. We therefore reject the imaginary eigen values in (108), so that, by (100), we have

$$Y = \pm \omega, \quad Z = \pm \omega^*. \tag{109}$$

We now put these values into (107) and solve, obtaining the following principal vectors:

$$\text{for } F_{rs}\begin{cases} Y = \omega: & \eta_1 = \varepsilon, \quad \eta_2 = 0, \quad \eta_3 = 0, \quad \eta_4 = i, \\ Y = -\omega: & \eta_1 = -\varepsilon, \quad \eta_2 = 0, \quad \eta_3 = 0, \quad \eta_4 = i; \end{cases} \tag{110}$$

$$\text{for } F^*_{rs}\begin{cases} Z = \omega^*: & \zeta_1 = -\varepsilon^*, \quad \zeta_2 = 0, \quad \zeta_3 = 0, \quad \zeta_4 = i, \\ Z = -\omega^*: & \zeta_1 = \varepsilon^*, \quad \zeta_2 = 0, \quad \zeta_3 = 0, \quad \zeta_4 = i; \end{cases} \tag{111}$$

here ε and ε^* are as in (100).

The ambiguous signs are confusing, but the result is really very simple:

(*a*) The eigen values of F_{rs} are $\pm \omega$, where ω is the magnitude of the electric vector in the electromagnetic wrench.

(*b*) The eigen values of F^*_{rs} are $\pm \omega^*$, where ω^* is the magnitude of the magnetic vector in the electromagnetic wrench.

(*c*) F_{rs} has precisely two principal directions, which are at the same time principal directions of F^*_{rs}. These are null directions, and they coincide with the principal null directions of the energy tensor T_{rs} (cf. Fig. 2).[1]

The connection between the eigen properties of T_{rs} and those of F_{rs} and F^*_{rs} depends essentially on the fact that (40) may be written in matrix language as

$$c^2\mathbf{T} = -\mathbf{F}^2 - \tfrac{1}{4}\mathbf{I}F_{mn}F_{mn}, \tag{112}$$

[1] For simplicity we have taken the general case ($\omega \neq 0$, $\omega^* \neq 0$). But if $F_{rs}F^*_{rs} = 0$ (equivalently $\det F_{rs} = 0$), then either E_1 or H_1 vanishes in (107), and (110), (111) need modification. If $E_1 = 0$, $H_1 \neq 0$, then $Y = 0$, and the principal direct ions of F_{rs} have merely to satisfy $\eta_2 = \eta_3 = 0$, which means that they form a timelike 2-flat containing the two (null) principal directions of F^*_{rs}. If $E_1 \neq 0$, $H_1 = 0$, the roles of F_{rs} and F^*_{rs} are interchanged.

where **T**, **F** and **I** are matrices, **I** being the unit matrix. We shall however use indicial notation here to explore this connection. For the eigen properties of F_{rs} we have, as in (102),

$$F_{rs}\eta_s = Y\eta_r, \tag{113}$$

and if we multiply this by F_{rt} we get

$$F_{rt}F_{rs}\eta_s = YF_{rt}\eta_r = -YF_{tr}\eta_r = -Y^2\eta_t. \tag{114}$$

Subtracting

$$\tfrac{1}{4}\delta_{ts}F_{mn}F_{mn}\eta_s$$

from both sides and using (40), we get, on division by c^2,

$$T_{ts}\eta_s = -c^{-2}(Y^2 + \tfrac{1}{4}F_{mn}F_{mn})\eta_t. \tag{115}$$

But this is of the form

$$T_{ts}\eta_s = X\eta_t, \tag{116}$$

and so it follows that *every principal direction of* F_{rs} *is also a principal direction of* T_{rs}, *and the eigen values* X *of* T_{rs} *are related to the eigen values* Y *of* F_{rs} *by*

$$Y^2 = -Xc^2 - \tfrac{1}{4}F_{mn}F_{mn}. \tag{117}$$

A similar result holds for F^*_{rs}. Since we already know that $X = \pm k$ by (54), (117) tells us at once that $Y = 0$ if $k = 0$ (since then $F_{mn}F_{mn} = 0$) and $Y = \pm\omega$ if $k \neq 0$. Similarly $Z = 0$ if $k = 0$ and $Z = \pm\omega^*$ if $k \neq 0$.

§ 7. THE TENSORS F_{rs} AND F^*_{rs} EXPRESSED IN TERMS OF INVARIANTS AND PRINCIPAL NULL VECTORS

Consider first a field which is not null ($k \neq 0$). Then there exists a pair of null vectors (say ξ_r and η_r) which are principal for both F_{rs} and F^*_{rs}. Consider the formulae

$$F_{rs} = \alpha\frac{\xi_r\eta_s - \xi_s\eta_r}{\xi_t\eta_t} + i\beta\frac{\varepsilon_{rsmn}\xi_m\eta_n}{\xi_t\eta_t}, \tag{118}$$

and

$$F^*_{rs} = \alpha^*\frac{\xi_r\eta_s - \xi_s\eta_r}{\xi_t\eta_t} + i\beta^*\frac{\varepsilon_{rsmn}\xi_m\eta_n}{\xi_t\eta_t}, \tag{119}$$

where α, β, α^*, β^* are invariants. We are going to show that these invariants can be chosen so that (118) and (119) are true.

We note that these are tensor formulae, and therefore true in all Galileian frames if true in one. Let us examine them for a canonical frame for which F_{rs} and F^*_{rs} are as in (98) and the components of the principal vectors are [cf. (110), (111), the notation being slightly changed]

$$\begin{aligned} \xi_1 &= 1, \quad \xi_2 = 0, \quad \xi_3 = 0, \quad \xi_4 = i, \\ \eta_1 &= -1, \quad \eta_2 = 0, \quad \eta_3 = 0, \quad \eta_4 = i. \end{aligned} \tag{120}$$

Then the matrix of the right hand side of (118) reads

$$\begin{pmatrix} 0 & 0 & 0 & -i\alpha \\ 0 & 0 & \beta & 0 \\ 0 & -\beta & 0 & 0 \\ i\alpha & 0 & 0 & 0 \end{pmatrix} \tag{121}$$

and the matrix of the right hand side of (119) is the same, with α^*, β^* written for α, β. Comparison with (98) shows that (118) and (119) are in fact true, provided

$$\begin{aligned} \alpha &= E_1 = \varepsilon\omega, & \beta &= H_1 = \varepsilon^*\omega^*, \\ \alpha^* &= -H_1 = -\varepsilon^*\omega^*, & \beta^* &= E_1 = \varepsilon\omega, \end{aligned} \tag{122}$$

E_1 and H_1 being the vectors of the electromagnetic wrench and ε, ε^* as in (100).

Thus we can express F_{rs} and F^*_{rs} in terms of the principal null vectors ξ_r, η_r and the invariants ω, ω^* by the formulae

$$\begin{aligned} F_{rs} &= \varepsilon\omega \frac{\xi_r\eta_s - \xi_s\eta_r}{\xi_t\eta_t} + i\varepsilon^*\omega^* \frac{\varepsilon_{rsmn}\xi_m\eta_n}{\xi_t\eta_t}, \\ F^*_{rs} &= -\varepsilon^*\omega^* \frac{\xi_r\eta_s - \xi_s\eta_r}{\xi_t\eta_t} + i\varepsilon\omega \frac{\varepsilon_{rsmn}\xi_m\eta_n}{\xi_t\eta_t}. \end{aligned} \tag{123}$$

The frame is now any Galileian frame. We note that $\varepsilon = \pm 1$, $\varepsilon^* = \pm 1$, the sign of their product being given by (101). The ambiguity of sign in (123) is natural, because it is not decided a priori which of the two principal vectors should be called ξ_r and which η_r, and an interchange of them reverses the sign of every term on the right hand side of (123).

We shall now express the energy tensor T_{rs} in terms of ξ_r, η_r and the invariants. Let us define

$$A_{rs} = (\xi_r\eta_s - \xi_s\eta_r)/(\xi_p\eta_p). \tag{124}$$

Then by the general definition (3) of the dual of a skew-symmetric tensor,

$$A^*_{rs} = \tfrac{1}{2}i\varepsilon_{rsmn}A_{mn} = i\varepsilon_{rsmn}\xi_m\eta_n/(\xi_p\eta_p). \tag{125}$$

Remembering that ξ_r and η_r are null vectors, we have

$$A_{rs}A_{rt} = -(\xi_s\eta_t + \xi_t\eta_s)/(\xi_p\eta_p), \quad A_{rs}A_{rs} = -2, \quad A_{rs}A^*_{rt} = 0, \tag{126}$$

and further, by the identity (39) which holds for any skew-symmetric tensor,

$$A^*_{rs}A^*_{rt} = A_{rs}A_{rt} - \tfrac{1}{2}\delta_{st}A_{mn}A_{mn}. \tag{127}$$

We can write (123) in the form

$$\begin{gathered} F_{rs} = \alpha A_{rs} + \beta A^*_{rs}, \quad F^*_{rs} = \alpha^* A_{rs} + \beta^* A^*_{rs}, \\ \alpha = \varepsilon\omega, \quad \beta = \varepsilon^*\omega^*, \quad \alpha^* = -\varepsilon^*\omega^*, \quad \beta^* = \varepsilon\omega, \end{gathered} \tag{128}$$

so that, by (99),

$$\alpha^2 + \alpha^{*2} = \beta^2 + \beta^{*2} = \omega^2 + \omega^{*2} = 2kc^2. \tag{129}$$

By (32) the energy tensor is given by

$$2c^2T_{rs} = F_{mr}F_{ms} + F^*_{mr}F^*_{ms}, \tag{130}$$

and when we substitute from (128) and use (129) and the last of (126) this becomes

$$2c^2T_{rs} = 2kc^2(A_{mr}A_{ms} + A^*_{mr}A^*_{ms}). \tag{131}$$

By (127) this may be written

$$T_{rs} = 2k(A_{mr}A_{ms} - \tfrac{1}{4}\delta_{rs}A_{mn}A_{mn}), \tag{132}$$

or, on substitution from (126),

$$T_{rs} = k\left(\delta_{rs} - 2\frac{\xi_r\eta_s + \xi_s\eta_r}{\xi_p\eta_p}\right). \tag{133}$$

This is the required expression for the energy tensor in terms of the principal null vectors ξ_r, η_r and the invariant k for the case $k \neq 0$. As a check, we verify that $T_{rr} = 0$ since $\delta_{rr} = 4$.

We turn now to the case of a null field ($k = 0$). We might derive a formula from (123) by letting the invariants tend to zero and at the same time letting the null vector η_r tend to coincidence with ξ_r, but it is clearer and simpler to state the following result and prove it: *Given a null field F_{rs}, there exists a certain null vector ξ_r lying along the*

principal null direction of F_{rs} *and pointing into the future, and spacelike unit vectors* p_r, q_r *orthogonal to* ξ_r, *such that*

$$F_{rs} = \xi_r p_s - \xi_s p_r, \quad F^*_{rs} = i\varepsilon_{rsmn}\xi_m p_n, \tag{134}$$

and

$$F_{rs} = i\varepsilon_{rsmn}\xi_m q_n, \quad F^*_{rs} = \xi_r q_s - \xi_s q_r. \tag{135}$$

It is clear that the second of (134) follows from the first, and that (135) will follow if we have proved (134). Thus the formula we have to establish is

$$F_{rs} = \xi_r p_s - \xi_s p_r, \tag{136}$$

the prescribed conditions on ξ_r and p_r being

$$F_{rs}\xi_s = 0, \quad \xi_r\xi_r = 0, \quad \xi_r p_r = 0, \quad p_r p_r = 1, \quad \xi_4/i > 0. \tag{137}$$

Given F_{rs} satisfying the null conditions

$$F_{rs}F_{rs} = 0, \quad F_{rs}F^*_{rs} = 0, \tag{138}$$

we have to prove the existence of ξ_r, p_r satisfying (136) and (137).

Since (136) is a tensor equation, and (137) also, it suffices to establish existence in some one frame of reference. Choose, then, a frame for which F_{rs} takes the canonical form (97), the principal null direction for this frame being given by a vector of the form

$$\xi_1 = a, \quad \xi_2 = 0, \quad \xi_3 = 0, \quad \xi_4 = ia,$$

with a positive to make it point into the future. Choose for p_r

$$p_1 = 0, \quad p_2 = 1, \quad p_3 = 0, \quad p_4 = 0.$$

Then (137) are satisfied. The matrix on the right hand side of (136) reads

$$\begin{pmatrix} 0 & a & 0 & 0 \\ -a & 0 & 0 & -ia \\ 0 & 0 & 0 & 0 \\ 0 & ia & 0 & 0 \end{pmatrix}$$

This agrees with (97) if we choose $a = E_2$, and so the result is established.

Note that p_r is by no means uniquely determined, for we can add to it any multiple of ξ_r without disturbing (136) or (137). As regards ξ_r, it differs from principal null vectors previously considered by having definite components; in other cases only the ratios of the components

mattered. For any canonical axes the fourth component ξ_4 is iE_2, where E_2 $(= H_3)$ is the electric vector.

We now calculate the energy tensor from (134). We find by (137)

$$F_{rs}F_{rt} = (\xi_r p_s - \xi_s p_r)(\xi_r p_t - \xi_t p_r) = \xi_s \xi_t, \tag{139}$$

and so, by (40) and (138), *the energy tensor of a null field* $(k = 0)$ *is expressed in terms of the principal null vector* ξ_r *by the very simple formula*[1]

$$T_{rs} = c^{-2}\xi_r \xi_s. \tag{140}$$

Note that T_{rs} has the dimensions [mass] and the components ξ_r the dimensions [energy]$^{\frac{1}{2}}$.

From (140) it follows that the energy tensor of a null field is the same as that of a certain distribution of photons. Consider a stream of photons, each with a (null) 4-momentum M_r, the number of their world lines cutting a 3-element dS normal to a timelike unit vector $\bar{n}_r$ being $\bar{N}dS$. Then, by VIII–(112), the energy tensor of this stream is

$$T_{rs} = \frac{\bar{N}M_r M_s}{|M_t \bar{n}_t|}. \tag{141}$$

Let us choose $M_r = b\xi_r$, where b is some invariant factor and ξ_r the definite principal null vector occurring in (140); then (141) reads

$$T_{rs} = \frac{b\bar{N}\xi_r \xi_s}{|\xi_t \bar{n}_t|}, \tag{142}$$

and this agrees with the T_{rs} of (140) provided b and $\bar{N}$ are chosen to satisfy

$$b\bar{N} = c^{-2}|\xi_t \bar{n}_t|. \tag{143}$$

The formulae (134) and (135) may be looked at in a different way. Let F_{rs} be a null field and ξ_r a given vector lying in the principal null direction of F_{rs}, so that

$$F_{rs}\xi_s = 0, \quad F^*_{rs}\xi_s = 0.$$

Consider the equations

$$F_{rs} = \xi_r p_s - \xi_s p_r, \quad F^*_{rs} = \xi_r q_s - \xi_s q_r \tag{144}$$

for the unknown vectors p_r, q_r. Using a canonical frame, so that F_{rs}

[1] Cf. RUSE [1936]. MARIOT [1954] has shown that, for a null Maxwellian field in general relativity, the world lines defined by ξ_r are null geodesics, and that if the field is null on some spacelike 3-space S, it is also null off S.

and F^*_{rs} are as in (97) and $\xi_2 = \xi_3 = 0$, $\xi_4 = i\xi_1$, we find

$$\begin{aligned} p_1/p_4 &= \xi_1/\xi_4 = -i, \quad p_2 = E_2/\xi_1, \quad p_3 = 0, \\ q_1/q_4 &= \xi_1/\xi_4 = -i, \quad q_2 = 0, \quad q_3 = -E_2/\xi_1. \end{aligned} \tag{145}$$

Thus p_r is determined to within an additive part proportional to ξ_r and moreover it is orthogonal to ξ_r; the same is true of q_r. Now $p_r p_r$ is an invariant, and it does not change when we add a vector proportional to ξ_r, for

$$(p_r + \theta\xi_r)(p_r + \theta\xi_r) = p_r p_r + 2\theta p_r \xi_r + \theta^2 \xi_r \xi_r = p_r p_r.$$

There is also an invariant $q_r q_r$, but it is the same; by (145)

$$p_r p_r = q_r q_r = (E_2/\xi_1)^2. \tag{146}$$

Thus, although the two invariants (FF), (FF^*) of (29) vanish for a null field, we see that *a null field F_{rs} together with a vector ξ_r lying in its principal null direction define a single invariant $p_r p_r$, where p_r is any vector satisfying*

$$F_{rs} = \xi_r p_s - \xi_s p_r. \tag{147}$$

Note that by (145) p_r and q_r are orthogonal.

The results established in this section and the preceding ones belong to what one might call *the geometry of the electromagnetic field.* In pre-relativistic days mathematical physicists were thoroughly familiar with the two 3-vectors (**E**, **H**), and as long as we have only one frame of reference in mind (perhaps that determined by a piece of electromagnetic or optical apparatus) these two vectors are very satisfactory. But as soon as we consider Lorentz transformations (and they are the core of the special theory of relativity), the pair of 3-vectors become very hard to think about, and we naturally turn to the absolute geometry of Minkowskian space-time, using instead the skew-symmetric tensor F_{rs}. But a tensor is hard to visualise, and that is why we have troubled to express F_{rs} and F^*_{rs} (and also T_{rs}) in terms of 4-vectors, which are easy to visualise. These formulae help us to understand the structure of an electromagnetic field.

As regards method in such work, we have on the one hand tensor technique in a general Galileian frame (symmetrical and elegant, but often leading to complicated algebra) and on the other hand methods based on the use of specially chosen frames of reference (unsymmetrical but with simple algebra). Neither method has been used consistently above; the reader may find satisfaction in changing the method used

to that of his own choice. The theory considered here has been treated along the same general lines, but with rather more detail and with a slightly different notation, in SYNGE [1935b]. Other references are S. R. MILNER [1922], E. T. WHITTAKER [1922], G. Y. RAINICH [1925], N. R. SEN [1934b] and H. S. RUSE [1936].

§ 8. THE 4-POTENTIAL

We return now to Maxwell's equations (15); these played no part in the work of §§ 4–7, for we were there concerned only with the field at a single event. And indeed we shall not at present concern ourselves with the full set (15), but only with the equations

$$F_{mn,r} + F_{nr,m} + F_{rm,n} = 0. \tag{148}$$

There are actually only four distinct equations here, obtained by choosing for *mnr* a set of three distinct numbers from the set 1234, without regard to order. Thus (148) may be exhibited in the form

$$F_{\mu\nu,\varrho} + F_{\nu\varrho,\mu} + F_{\varrho\mu,\nu} = 0, \tag{149}$$

$$F_{\mu\nu,4} + F_{\nu 4,\mu} + F_{4\mu,\nu} = 0, \tag{150}$$

where, as always, Greek suffixes have the values 1, 2, 3.

We note that if F_{mn} is of the form

$$F_{mn} = \phi_{n,m} - \phi_{m,n}, \tag{151}$$

then (148) is certainly satisfied; this follows from the fact that $\phi_{r,mn} = \phi_{r,nm}$.

Our object is now to prove the converse, viz. to prove the existence of a 4-vector ϕ_r (the 4-*potential*) satisfying (151), F_{mn} being any given skew-symmetric tensor field satisfying (148), or equivalently (149) and (150).

Before embarking on this proof, a word should be said about the *domain of existence*. Suppose that F_{mn}, $F_{mn,r}$ exist, are continuous, and satisfy (148) throughout some domain D of space-time. Then our proof will not establish the existence of ϕ_r throughout D in general, but only in a *box* B given by inequalities of the form

$$x_\varrho^{(0)} < x_\varrho < x_\varrho^{(1)}, \quad t^{(0)} < t < t^{(1)}, \tag{152}$$

contained in D. If D is the whole of space-time, then we may take $B = D$, so that the bounds in (152) are infinite, but if D is a finite

domain, bounded by curved 3-spaces and perhaps multiply connected,[1] we make no claim to establish the existence of ϕ_r throughout D, but only in any box (152) contained in D. In fact, our proof of existence will be a proof "in the small" and not "in the large".

We take the origin inside the box to simplify the writing, and write (151) in the equivalent form

$$\phi_{\nu,4} - \phi_{4,\nu} = F_{4\nu}, \tag{153}$$

$$\phi_{\nu,\mu} - \phi_{\mu,\nu} = F_{\mu\nu}. \tag{154}$$

We proceed to solve these equations for ϕ_r. Choosing ϕ_4 arbitrarily, we satisfy (153) by taking

$$\phi_\nu = f_\nu(x_1, x_2, x_3) + \int_0^{x_4} \phi_{4,\nu} dx_4 + \int_0^{x_4} F_{4\nu} dx_4, \tag{155}$$

where the three functions f_ν are (for the moment) arbitrary. The limits of integration, here and later, are not to carry us out of the box.

Now (154) is satisfied by (155) provided f_ϱ satisfy

$$f_{\nu,\mu} - f_{\mu,\nu} + \int_0^{x_4} (F_{4\nu,\mu} - F_{4\mu,\nu}) dx_4 = F_{\mu\nu}. \tag{156}$$

But by (150)

$$F_{4\nu,\mu} - F_{4\mu,\nu} = F_{\mu\nu,4}, \tag{157}$$

and so (156) may be changed to read

$$f_{\nu,\mu} - f_{\mu,\nu} + F_{\mu\nu} - (F_{\mu\nu})_0 = F_{\mu\nu}, \tag{158}$$

the subscript 0 indicating evaluation for $x_4 = 0$. Thus (151) is satisfied by ϕ_r (ϕ_4 arbitrary, ϕ_ν as in (155)) provided that the three functions $f_\varrho(x_1, x_2, x_3)$ satisfy

$$f_{\nu,\mu} - f_{\mu,\nu} = (F_{\mu\nu})_0, \tag{159}$$

or, explicitly,

$$\left.\begin{aligned} f_{2,3} - f_{3,2} &= (F_{32})_0, \\ f_{3,1} - f_{1,3} &= (F_{13})_0, \\ f_{1,2} - f_{2,1} &= (F_{21})_0. \end{aligned}\right\} \tag{160}$$

We choose f_3 arbitrarily. The first two of (160) are then satisfied if

[1] In the analogous 3-dimensional problem, the condition div $\mathbf{f} = 0$, analogous to (148), is not enough to ensure the existence of a solution $\mathbf{v}$ of the vector equation curl $\mathbf{v} = \mathbf{f}$; we need also $\int f_n dS = 0$ for every closed surface in D; cf. McMahon [1953], Stevenson [1954].

we take

$$f_2 = g_2(x_1, x_2) + \int_0^{x_3} f_{3,2}dx_3 + \int_0^{x_3} (F_{32})_0 dx_3,$$
$$f_1 = g_1(x_1, x_2) + \int_0^{x_3} f_{3,1}dx_3 - \int_0^{x_3} (F_{13})_0 dx_3, \tag{161}$$

the functions g_1 and g_2 being arbitrary for the moment; and then the last of (160) is satisfied if these two functions satisfy

$$g_{1,2} - g_{2,1} - \int_0^{x_3} (F_{13,2} + F_{32,1})_0 dx_3 = (F_{21})_0. \tag{162}$$

By virtue of (149) this may be written

$$g_{1,2} - g_{2,1} = -(F_{12})_{00}, \tag{163}$$

where the subscript 00 means evaluation for $x_3 = x_4 = 0$. We have now only this equation to satisfy, and this we do by choosing g_2 arbitarily and then putting

$$g_1 = h(x_1) + \int_0^{x_2} g_{2,1}dx_2 - \int_0^{x_2} (F_{12})_{00}dx_2, \tag{164}$$

the function h being arbitrary. *Thus we have shown that, if F_{mn} satisfies* (148), *the equations* (151) *possess a solution ϕ_r, the solution having a high degree of arbitrariness on account of the arbitrary choice of the functions ϕ_4, f_3, g_2 and h.*

We may express (151) in words by saying that F_{mn} is the *curl* or *rot* of ϕ_r. For the dual we have

$$F^*_{rs} = i\varepsilon_{rsmn}\phi_{n,m}. \tag{165}$$

So far our discussion of the 4-potential ϕ_r has been based entirely on the second set of Maxwell's equations in (15). We now turn to the first set ($F_{rs,s} = 0$); they impose on ϕ_r the four conditions

$$\phi_{r,ss} - (\phi_{s,s})_{,r} = 0. \tag{166}$$

Any 4-vector ϕ_r satisfying these partial differential equations generates a Maxwellian field by (151); *conversely, except for the limitation on domain of existence mentioned earlier, any Maxwellian field possesses a 4-potential satisfying these equations.*

Mathematically, the 4-potential ϕ_r is much easier to use than the tensor F_{mn}. There are only four components to deal with instead of six, and the four partial differential equations (166) (although of the second order) are more straightforward than the eight equations (10)

or (15), which are interconnected by (16). The unsatisfactory thing about ϕ_r is its lack of determinacy; this we shall now examine further.

Let ϕ_r and ψ_r be two different 4-potentials for the same electromagnetic field F_{mn}, so that

$$F_{mn} = \phi_{n,m} - \phi_{m,n} = \psi_{n,m} - \psi_{m,n}. \tag{167}$$

If we define χ_r by

$$\chi_r = \phi_r - \psi_r, \tag{168}$$

it follows from (167) that

$$\chi_{n,m} = \chi_{m,n}, \tag{169}$$

and this implies that χ_r is a gradient, so that we may write

$$\chi_r = \chi_{,r}, \tag{170}$$

where χ is a scalar. Therefore, *given any one 4-potential ψ_r, all other 4-potentials of the same field are found by adding a gradient to ψ_r, and so are of the form*

$$\phi_r = \psi_r + \chi_{,r}. \tag{171}$$

We can use this result to simplify the partial differential equations (166) satisfied by the 4-potential. Let ψ_r be any given 4-potential of some assigned field and let χ be a scalar; then ϕ_r, as given in (171), is another 4-potential of the same field, and we have

$$\phi_{r,r} = \psi_{r,r} + \chi_{,rr}. \tag{172}$$

Choose χ to satisfy the partial differential equation

$$\chi_{,rr} = -\psi_{r,r} \text{ or } \Box\chi = -\psi_{r,r}, \tag{173}$$

where $\Box$ is the *d'Alembertian operator*

$$\Box = \frac{\partial^2}{\partial x_s \partial x_s} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{174}$$

This operator is the analogue in Minkowskian space-time of the Laplace operator in Euclidean 3-space, and plays as fundamental a role in relativity as does the Laplace operator in Newtonian physics, or nearly so. But there is actually a far greater difference between the two than a mere increase in dimensionality; the Laplace operator leads to partial differential equations of the *elliptical type*, whereas with the d'Alembertian we get equations of the *hyperbolic type*, this hyperbolic character arising from the imaginary coordinate x_4, or (what is the same thing) the minus sign in (174).

Now, with χ chosen to satisfy (173), the equation (172) becomes

$$\phi_{r,r} = 0, \tag{175}$$

and in consequence (166) simplifies to

$$\phi_{r,ss} = 0 \text{ or } \Box\phi_r = 0. \tag{176}$$

A change of 4-potential as in (171) (addition of a gradient) is called a *gauge transformation*; when we work with a 4-potential instead of the tensor F_{mn}, we must make sure that our work is *gauge invariant* in the sense that the results are unchanged by a gauge transformation.

We have used a gauge transformation to simplify the equations satisfied by the 4-potential from the form (166) to the form (176), at the same time imposing on the 4-potential the *normalising condition* (175) (vanishing of the divergence). We shall not discuss the regularity conditions on $\psi_{r,r}$ sufficient to ensure the existence of a solution χ of (173); but, supposing such conditions satisfied, we may be sure that the solution of (173) is far from unique, since no boundary conditions are associated with the partial differential equation. Thus, although the normalising condition (175) reduces the indeterminacy of the 4-potential, it does not destroy it, and we must therefore beware of regarding the normalised 4-potential ϕ_r as a unique vector field associated with a given Maxwellian field F_{mn}.

Normalisation based on (175) is not the only possible one. For example, in the work from (153) to (164) we might choose

$$\phi_4 = 0, \quad g_3 = 0, \quad f_2 = 0, \quad h = 0. \tag{177}$$

This gives a uniquely determined 4-potential with only three components (ϕ_1, ϕ_2, ϕ_3). But this procedure (since it depends on the axes chosen) does not commend itself in relativity; the invariant equation (175) is preferable, even though it does not give uniqueness.

The standard procedure is to use a 4-potential subject to the partial differential equations

$$\boxed{\phi_{r,r} = 0, \quad \Box\phi_r = 0.} \tag{178}$$

We know that the curl of any solution of these equations is a Maxwellian field. That is the important thing. As for the converse, we can be confident that (unusual cases excepted) a Maxwellian field possesses a set of 4-potentials satisfying these equations.

It might appear that we have too much in (178) — five equations

for four unknowns. But they are not independent, since

$$\Box(\phi_{r,r}) = (\Box\phi_r)_{,r}. \tag{179}$$

§ 9. PLANE ELECTROMAGNETIC WAVES

The simplest of all electromagnetic fields is the *uniform* field in which all the components of F_{rs} are constant. We can contemplate such a field in a domain D of space-time, which may in particular be the whole of space-time. Maxwell's equations (10) are certainly satisfied in D, since all the derivatives vanish. The reductions of § 5 can be made throughout D, and in this way the field reduced to either (*a*) equal orthogonal vectors E_ϱ, H_ϱ (in the case of a null field), or (*b*) an electromagnetic wrench with E_ϱ, H_ϱ on a common line, this latter being the more general case.

When we pass on to variable fields, we are led to think of electromagnetic waves. It is, however, difficult to say when a variable field may be regarded as composed of waves and when it may not be so regarded. This is due to the free use of the word *wave* in physics, and it would be unwise to attempt to tie down such a useful word by any arbitrary definition. In what follows we shall use the word *wave* in ways which seem in accordance with the rather vague concepts lying behind the word.

Consider the electromagnetic field given by

$$F_{rs} = K_{rs}P(a_n x_n); \tag{180}$$

here K_{rs} is a constant skew-symmetric *amplitude tensor* and a_r a constant 4-vector; P is any *phase function.*

In general such a field does not satisfy Maxwell's equations (10). Before examining the conditions to be imposed on K_{rs} and a_r in order that it may do so, let us see what (180) implies. We recognise a *system of plane waves* in the sense that the components of F_{rs} are constant over each of the 3-flats

$$a_n x_n = \text{const.} \tag{181}$$

This is also true of a more general type of field given by

$$F_{rs} = K_{rs}P(a_n x_n) + L_{rs}Q(a_n x_n), \tag{182}$$

or indeed of one obtained by adding on more terms of this type. We may call (180) a *simple* system of plane waves and (182) and its

generalisations *compound* systems. We shall for the present confine our attention to simple systems.[1]

Each of the 3-waves (181) is the history of a 2-wave moving the *wave velocity* [cf. Appendix A, p. 419]

$$w_\varrho = -ic\frac{a_\varrho a_4}{a_\sigma a_\sigma}, \tag{183}$$

and the unit normal to the 2-wave has components

$$n_\varrho = \frac{a_\varrho}{(a_\sigma a_\sigma)^{\frac{1}{2}}}. \tag{184}$$

The concepts of frequency and wave length do not appear; they cannot, until we impose a periodicity on the function P.

Noting that the dual of (180) is

$$F^*_{rs} = K^*_{rs}P(a_n x_n), \quad K^*_{rs} = \tfrac{1}{2}i\varepsilon_{rsmn}K_{mn}, \tag{185}$$

we now apply Maxwell's equations (10) to (180) and (185), assuming the function P to possess a derivative in the domain D considered. (This means that there are no electromagnetic shock waves in D; see §§ 10, 20.) We have

$$F_{rs,s} = K_{rs}a_s P'(a_n x_n), \quad F^*_{rs,s} = K^*_{rs}a_s P'(a_n x_n), \tag{186}$$

and so the necessary and sufficient conditions on K_{rs} and a_r for the satisfaction of Maxwell's equations (10) are

$$K_{rs}a_s \neq 0, \quad K^*_{rs}a_s = 0; \tag{187}$$

these give 6 independent conditions on 10 constants.

To explore the meaning of (187), we note that the theory of §§ 2–7 applies to any skew-symmetric tensor and so to K_{rs}. It possesses an invariant [cf. (54)]

$$k = \tfrac{1}{4}c^{-2}[(K_{rs}K_{rs})^2 + (K_{rs}K^*_{rs})^2]^{\frac{1}{2}}. \tag{188}$$

Let us suppose $k \neq 0$. Then K_{rs} and K^*_{rs} can be reduced to the canonical forms (98), in which we must assume either E_1 or H_1 different from zero if the field (180) is not to disappear. When we put these canonical forms into (187) we get $a_r = 0$, and so the field (180) becomes uniform. Thus if (180) is not to reduce to a uniform field (and that gives us no

[1] We might also use the term *compound* system of plane waves to refer to a more general expression than (182), with $Q(a_n x_n)$ replaced by $Q(b_n x_n)$.

waves), we must take $k = 0$, and hence

$$K_{rs}K_{rs} = 0, \quad K_{rs}K^*_{rs} = 0. \tag{189}$$

In fact the field K_{rs} must be a null field, with canonical forms as in (97),

$$\left.\begin{aligned} K_{rs} &= \begin{pmatrix} 0 & K & 0 & 0 \\ -K & 0 & 0 & -iK \\ 0 & 0 & 0 & 0 \\ 0 & iK & 0 & 0 \end{pmatrix} \\ K^*_{rs} &= \begin{pmatrix} 0 & 0 & -K & 0 \\ 0 & 0 & 0 & 0 \\ K & 0 & 0 & iK \\ 0 & 0 & -iK & 0 \end{pmatrix} \end{aligned}\right\} \tag{190}$$

with $K \neq 0$.[1] Then (187) read explicitly, for the canonical axes corresponding to (190),

$$\begin{gathered} Ka_2 = 0, \quad -Ka_1 - iKa_4 = 0, \quad 0 = 0, \quad iKa_2 = 0, \\ -Ka_3 = 0, \quad 0 = 0, \quad Ka_1 + iKa_4 = 0, \quad -iKa_3 = 0, \end{gathered} \tag{191}$$

so that

$$a_1 + ia_4 = 0, \quad a_2 = a_3 = 0. \tag{192}$$

Thus a_r is a null vector, and it is in fact the principal vector common to K_{rs} and K^*_{rs}.

To sum up: *in order that a wave system* (180) *may be Maxwellian, it is necessary and sufficient that* K_{rs} *be a null field (i.e. satisfying* (189) *and so representable by two equal orthogonal 3-vectors), and further that* a_r *lie along the unique null direction principal for both* K_{rs} *and* K^*_{rs}.

Since a_r is a null vector, it follows at once from (183) that *plane electromagnetic waves are propagated with the speed c*, a most fundamental result. Indeed, since light is now universally regarded as an electromagnetic phenomenon, it is only with the establishment of this result that the hypothesis of I–(60) (that the history of a photon is a null line) loses its ad hoc character.

[1] It is interesting to note that if $\mathbf{M}$, $\mathbf{M}^*$ denote the two matrices in (190), then

$$\mathbf{M}^2 = \mathbf{M}^{*2} = \begin{pmatrix} -K^2 & 0 & 0 & -iK^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -iK^2 & 0 & 0 & K^2 \end{pmatrix}$$

and $\mathbf{M}^3 = \mathbf{M}^{*3} = \mathbf{0}$. This last result follows from (39), (189):

$$K_{rm}K_{mn}K_{ns} = K^*_{rm}K^*_{mn}K_{ns} = 0.$$

If we use canonical axes and substitute from (192) into (183), we get

$$w_1 = c, \quad w_2 = 0, \quad w_3 = 0, \tag{193}$$

showing that the wave travels in the direction of the Ox_1 axis.

It is remarkable that the form of the function P in (180) plays no part at all in our geometrical analysis of the wave system; the derivative P' occurs in (186) but divides out of the homogeneous Maxwell equations. The geometrical structure of all Maxwellian wave systems of the form (180) is extremely simple, all such systems being capable of construction in the following steps:

(i) Choose any phase function P.
(ii) Choose any null vector a_r.
(iii) Take an orthogonal tetrad of axes with a_r contained in the 2-flat Ox_1x_4.
(iv) Assign a value to K and define K_{rs} for these axes by (190).
(v) Transform away from these axes to a general set, so obtaining (180).

There is another way of expressing the plane wave system for general axes. Taking a_r to be any null vector, write

$$K_{rs} = a_r b_s - a_s b_r, \quad K^*_{rs} = i\varepsilon_{rsmn} a_m b_n. \tag{194}$$

Then

$$K_{rs} a_s = a_r a_s b_s, \quad K^*_{rs} a_s = 0, \tag{195}$$

and so the Maxwellian conditions (187) are satisfied provided

$$a_s b_s = 0, \tag{196}$$

i.e. if b_r is orthogonal to a_r or, equivalently, is any vector lying in the 3-flat tangent to the null cone along a_r. It is easy to see that (194) is the most general solution for K_{rs} of (187), a_r being assigned. This means that *any simple plane Maxwellian wave system may be written*

$$F_{rs} = (a_r b_s - a_s b_r) P(a_n x_n), \quad F^*_{rs} = i\varepsilon_{rsmn} a_m b_n P(a_t x_t), \tag{197}$$

where a_r is an arbitrary null vector and b_r any vector orthogonal to it.

If we view the wave system (180), not from the space-time point of view, but through the eyes of some Galileian observer, we see space, at any instant, sliced by the planes

$$a_\varrho x_\varrho = \text{const.} \tag{198}$$

At every point of space the vectors E_ϱ and H_ϱ are equal to one another

in magnitude and perpendicular to one another. Their common magnitude is constant over each of the planes (198). The direction of E_ϱ (and of H_ϱ) is the same everywhere, except for reversal of sense where P changes sign. As time passes, this whole pattern moves unchanged with speed c in a direction normal to (198).

What happens if we change from one Galileian observer to another? Let S and S' be two Galileian observers, S' appearing to S to be moving in the direction of the propagation of the waves with velocity v. (This is of course not the most general case, but the easiest to think about.) Then it follows from (85) that, if v is positive, the field appears weaker to S' than it does to S, in the ratio

$$\frac{(1 - v/c)^{\frac{1}{2}}}{(1 + v/c)^{\frac{1}{2}}}; \tag{199}$$

and if v is negative, then the field appears stronger to S'. Thus by pursuing the waves the strength is reduced, the limit being zero as the velocity of S' relative to S approaches c; and by going against the waves the strength is increased, the limit in this case being infinity.

§ 10. SOME SPECIAL SYSTEMS OF PLANE WAVES

We pass from the general plane waves of (180) to sinusoidal waves by taking a sine or cosine for the function P; let us put

$$F_{rs} = K_{rs} \cos (a_n x_n + \alpha), \tag{200}$$

or equivalently

$$F_{rs} = K_{rs} \cos (2\pi f_n x_n/c + \alpha), \tag{201}$$

so that we have a constant amplitude tensor K_{rs}, a constant 4-vector f_r, and a phase constant α; these may be chosen arbitrarily except for the conditions

$$K_{rs} f_s = 0, \quad K^*_{rs} f_s = 0, \tag{202}$$

as in (187). By the argument of p. 352 we know that f_r is a null vector; without loss of generality we take it pointing into the future.

Referring to v–(91), we recognise f_r as the frequency 4-vector, its components being

$$f_\varrho = n_\varrho \nu, \quad f_4 = i\nu, \tag{203}$$

where ν is the frequency and n_ϱ the direction cosines of the normal to the waves. The transformation of frequency, as the fourth component of a 4-vector, was discussed in v–§ 6.

The sinusoidal waves (201) may also be written in the form (197) as

$$F_{rs} = (f_r b_s - f_s b_r) \cos (2\pi f_n x_n/c + \alpha), \tag{204}$$

b_r being any vector orthogonal to the frequency vector f_r, so that $b_r f_r = 0$.

For various reasons (mathematical simplicity and the use of Fourier transforms) sinusoidal waves play a dominant part in mathematical physics. But actually the function P in (180) and (197) is quite arbitrary, and other interesting choices can be made. Let us choose

$$P(\xi) = \exp [- q^2(\xi - \alpha)^2], \tag{205}$$

where q and α are constants; then our plane wave system is

$$F_{rs} = (a_r b_s - a_s b_r) \exp [- q^2(a_n x_n - \alpha)^2], \quad a_r b_r = 0. \tag{206}$$

On the 3-flat Π with equation

$$a_n x_n = \alpha \tag{207}$$

the field is

$$F_{rs} = a_r b_s - a_s b_r, \tag{208}$$

and as we recede from Π in space-time the field falls off, rapidly if q is large. As we let q become larger and larger, the field becomes more and more concentrated in the neighbourhood of Π, and we are led to contemplate the limiting field

$$F_{rs} = (a_r b_s - a_s b_r) I(a, \alpha), \tag{209}$$

where

$$I(a, \alpha) \begin{cases} = 1 \text{ on } a_n x_n = \alpha, \\ = 0 \text{ elsewhere.} \end{cases} \tag{210}$$

Regarded mathematically, (209) defines a tensor field throughout space-time with a discontinuity on the 3-flat Π. It would be wrong to say that this field satisfies Maxwell's equations throughout space-time, since the partial derivatives of F_{rs} do not exist on Π. But any limit which can be approached as closely as we like within Maxwellian theory is acceptable as a model of reality; that applies to the field (209), which we may call an *electromagnetic sheet*. Such a sheet appears to any Galileian observer as a plane which moves perpendicularly to itself with speed c; there is no field except on the sheet, and on it the vectors E_ϱ and H_ϱ lie in it, with equal magnitudes ($E = H$) and

perpendicular ($E_\varrho H_\varrho = 0$), each giving a constant vector field over the sheet. The common magnitude (E or H) may have any value from zero to infinity according to the observer.

Other interesting cases are those in which the phase function P has the step forms shown in Fig. 5, approached as limits of forms in which the corners are rounded off to admit the derivative P'. These give *electromagnetic shock waves* (the electromagnetic sheet may be regarded as a particular case). In the first case shown, the Galileian observer sees two parallel planes advancing with speed c, the field being constant between them and zero outside them. In the second case he sees an advancing plane, with a constant field behind it and zero field in front of it. We shall return to shock waves in § 20.

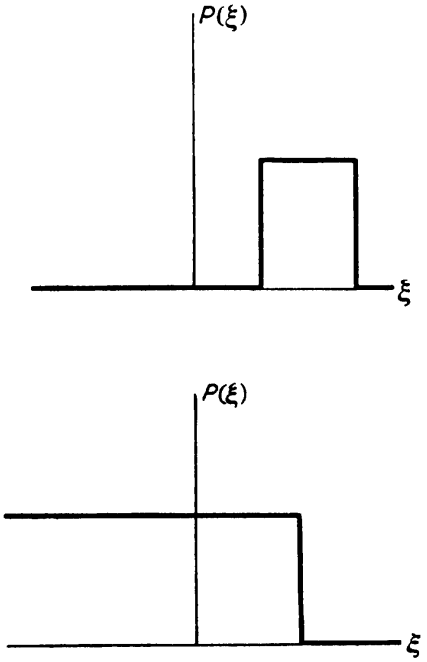


Fig. 5 – Step forms of the phase function for electromagnetic shock waves

There are of course infinitely many special forms which may be regarded as legitimate limits of Maxwellian fields in the sense that they can be approached as closely as we like by means of Maxwellian fields. Note that the concept of *frequency* is not at all inherent in the idea of plane electromagnetic waves, being introduced only by imposing a periodicity on the phase function P, as we did in (200) for sinusoidal waves. It is of course true that any function P (subject to certain restrictions) can be represented by a Fourier integral, and in that sense any plane wave system may be regarded as a superposition of sinusoidal waves of all frequencies; but, useful as that idea is, it is in the nature of a mathematical device, and we can often get on better without using it.

§ 11. REDUCTION OF A PAIR OF SINUSOIDAL PLANE WAVE SYSTEMS. INTERFERENCE

Consider the pair of sinusoidal plane wave systems,

$$F'_{rs} = K'_{rs} \cos(2\pi f'_n x_n/c + \alpha'), \quad F''_{rs} = K''_{rs} \cos(2\pi f''_n x_n/c + \alpha''), \tag{211}$$

where

$$K'_{rs} f'_s = K'^*_{rs} f'_s = K''_{rs} f''_s = K''^*_{rs} f''_s = 0, \tag{212}$$

so that the fields are Maxwellian. Since Maxwell's equations (10) are linear and homogeneous, the field

$$F_{rs} = F'_{rs} + F''_{rs}, \tag{213}$$

obtained by superimposing the two fields (211), is itself Maxwellian. Although the fields F'_{rs} and F''_{rs} are null, it is easy to see that F_{rs} is not null in general.

The question now is this: can we, by suitable choice of frame of reference, give to F_{rs} a particularly simple form?

We might think that the answer is to be found in the canonical axes of F_{rs}. But these canonical axes change from event to event, and so do not yield a frame of reference.

There are two cases to consider, depending on the properties of the frequency vectors, f'_r and f''_r, which are null vectors. There is the special case (i) in which these vectors have the same direction in space-time, so that we can write

$$f'_r = \phi f''_r, \tag{214}$$

where ϕ is an invariant, and the general case (ii) in which the directions are different.

In case (i) no significant simplification in F_{rs} is possible. Every Galileian observer see two systems of plane waves (all waves parallel to one another) advancing in the same spatial direction with speed c. In general the frequencies are different ($\nu' \neq \nu''$). By changing his frame, the observer can give these frequencies all values from zero to infinity, *but he cannot change their ratio*; that fixed ratio is $\nu'/\nu'' = \phi$, an invariant.

Case (ii) is more amenable. Now f'_r, f''_r have different directions; $(f'_r + f''_r)$ is timelike, and $(f'_r - f''_r)$ is spacelike, since it is orthogonal to $(f'_r + f''_r)$, for we have

$$(f'_r + f''_r)(f'_r - f''_r) = f'_r f'_r - f''_r f''_r = 0. \tag{215}$$

We shall use these vectors to define a simplifying frame. But first let us move the origin of space-time on to the 2-flat with equations

$$2\pi f'_n x_n/c + \alpha' = 0, \quad 2\pi f''_n x_n/c + \alpha'' = 0. \tag{216}$$

This 2-flat exists, since the two 3-flats in (216) are not parallel. By this device we get rid of the phase constants in (211), obtaining

$$F'_{rs} = K'_{rs} \cos(2\pi f'_n x_n/c), \quad F''_{rs} = K''_{rs} \cos(2\pi f''_n x_n/c). \tag{217}$$

It is easy to see that we can in general eliminate phase constants in this way from *four* waves, but not from more.

Now we take $(f'_r - f''_r)$ for Ox_1 axis and $(f'_r + f''_r)$ for Ox_4 axis. In such a frame we have

$$\begin{aligned} f'_2 - f''_2 = 0, \quad f'_1 + f''_1 = 0, \\ f'_3 - f''_3 = 0, \quad f'_2 + f''_2 = 0, \\ f'_4 - f''_4 = 0, \quad f'_3 + f''_3 = 0. \end{aligned} \tag{218}$$

We observe that the frequencies are now the same (say ν), and we have

$$f'_1 = -f''_1 = \nu, \quad f'_2 = f''_2 = f'_3 = f''_3 = 0, \quad f'_4 = f''_4 = i\nu. \tag{219}$$

In this frame the field F_{rs} becomes

$$F_{rs} = K'_{rs} \cos 2\pi\nu(x_1/c - t) + K''_{rs} \cos 2\pi\nu(x_1/c + t). \tag{220}$$

This means that (provided f'_r and f''_r have not the same direction) *it is possible, by suitable choice of frame of reference, to reduce any pair of sinusoidal plane wave systems to waves of the same frequency travelling in opposite directions* [cf. p. 190].

We can also show F_{rs} as a superposition of two standing wave systems:

$$\begin{aligned} F_{rs} = (K'_{rs} + K''_{rs}) \cos(2\pi\nu x_1/c) \cos 2\pi\nu t \\ + (K'_{rs} - K''_{rs}) \sin(2\pi\nu x_1/c) \sin 2\pi\nu t. \end{aligned} \tag{221}$$

The coefficients occurring in (220) and (221) are conditioned by (212) and (219); we have

$$\begin{aligned} K'_{14} = K'_{23} = 0, \quad K'_{12} = iK'_{24}, \quad K'_{31} = -iK'_{34}, \\ K''_{14} = K''_{23} = 0, \quad K''_{12} = -iK''_{24}, \quad K''_{31} = iK''_{34}. \end{aligned} \tag{222}$$

We might discuss *interference* on the basis of (221), but it is more interesting to revert to a general frame, except that we take the origin on the 2-flat (216) so that our field is

$$F_{rs} = K'_{rs} \cos(2\pi f'_n x_n/c) + K''_{rs} \cos(2\pi f''_n x_n/c). \tag{223}$$

This may be written

$$\begin{aligned} F_{rs} &= K'_{rs} \cos[\pi(f'_n + f''_n)x_n/c + \pi(f'_n - f''_n)x_n/c] \\ &\quad + K''_{rs} \cos[\pi(f'_n + f''_n)x_n/c - \pi(f'_n - f''_n)x_n/c] \\ &= (K'_{rs} + K''_{rs}) \cos[\pi(f'_n + f''_n)x_n/c] \cos[\pi(f'_n - f''_n)x_n/c] \\ &\quad - (K'_{rs} - K''_{rs}) \sin[\pi(f'_n + f''_n)x_n/c] \sin[\pi(f'_n - f''_n)x_n/c]. \end{aligned} \tag{224}$$

Interference may be said to occur whenever we superimpose two wave systems, but it is of interest chiefly in the sinusoidal case when the amplitude tensors K'_{rs} and K''_{rs} are nearly equal or nearly opposite. These two cases are quite similar, and we shall suppose that all the components of $(K'_{rs} - K''_{rs})$ are small. This is not a relativistically invariant condition, and so what follows must be regarded as applying only to some restricted set of Galileian observers.

We now have approximately

$$F_{rs} = 2K'_{rs} \cos [\pi(f'_n + f''_n)x_n/c] \cos [\pi(f'_n - f''_n)x_n/c]. \tag{225}$$

Hence, to this approximation, $F_{rs} = 0$ on the two discrete sets of 3-flats $\Pi_+(m)$, $\Pi_-(m)$ defined as follows:

$$\begin{aligned} \Pi_+(m)&: \quad (f'_n + f''_n)x_n = (m + \tfrac{1}{2})c, \\ \Pi_-(m)&: \quad (f'_n - f''_n)x_n = (m + \tfrac{1}{2})c, \\ &(m = 0, \pm 1, \pm 2, \ldots). \end{aligned} \tag{226}$$

These we may call *interference 3-waves*, each of them being the history of an *interference 2-wave*. The speeds of propagation of these 2-waves are obtained as in Appendix A, p. 419; if ν', ν'' are the frequencies of the two wave systems in (223) and n'_ϱ, n''_ϱ the unit normals of their instantaneous positions, then for $\Pi_+(m)$ the speed is given by

$$w_+^2 = -c^2 \frac{(f'_4 + f''_4)^2}{(f'_\varrho + f''_\varrho)(f'_\varrho + f''_\varrho)} = c^2 \frac{(\nu' + \nu'')^2}{\nu'^2 + \nu''^2 + 2\nu'\nu'' n'_\varrho n''_\varrho}, \tag{227}$$

and for $\Pi_-(m)$ by

$$w_-^2 = -c^2 \frac{(f'_4 - f''_4)^2}{(f'_\varrho - f''_\varrho)(f'_\varrho - f''_\varrho)} = c^2 \frac{(\nu' - \nu'')^2}{\nu'^2 + \nu''^2 - 2\nu'\nu'' n'_\varrho n''_\varrho}. \tag{228}$$

Since $(f'_r + f''_r)$ is timelike and $(f'_r - f''_r)$ spacelike, we know that $w_+ > c$ and $w_- < c$; this can be verified immediately from the above formulae.

In the special frame used in (221), for which the frequencies are equal and the waves travel in opposite directions, we have $\nu' = \nu''$ and $n'_\varrho n''_\varrho = -1$; hence $w_+ = \infty$ and $w_- = 0$.

§ 12. SOME SCALAR WAVE FUNCTIONS

In (178) we saw that the 4-potential ϕ_r is to satisfy the equations

$$\phi_{r,r} = 0, \quad \Box \phi_r = 0, \tag{229}$$

the second of these being the wave equation. Let us now consider scalar wave functions W (invariant under Lorentz transformations) satisfying the equation

$$\Box W = 0 \quad (\Box W = W_{,rr}). \tag{230}$$

One of the simplest solutions (we shall call it a *fundamental* or *elementary* solution) of this equation is

$$W = 1/S, \quad S = (x_r - \alpha_r)\,(x_r - \alpha_r), \tag{231}$$

where α_r is some fixed event. Let us verify that this satisfies (230). We have

$$S_{,r} = 2(x_r - \alpha_r), \quad S_{,rs} = 2\delta_{rs}, \quad S_{,r}S_{,r} = 4S, \tag{232}$$

and

$$W_{,r} = -S^{-2}S_{,r}, \quad W_{,rs} = -S^{-2}S_{,rs} + 2S^{-3}S_{,r}S_{,s}, \tag{233}$$

so that

$$\Box W = W_{,rr} = -2S^{-2}\delta_{rr} + 8S^{-3}S = 0, \tag{234}$$

since $\delta_{rr} = 4$.

The wave equation (230) is the space-time analogue of Laplace's equation in Euclidean 3-space,

$$\Delta V = 0 \quad (\Delta V = V_{,\varrho\varrho}), \tag{235}$$

and the fundamental wave function (231) is the analogue of the fundamental harmonic function

$$V = 1/r, \quad r^2 = (x_\varrho - \alpha_\varrho)\,(x_\varrho - \alpha_\varrho). \tag{236}$$

In Euclidean n-space the fundamental harmonic function is r^{2-n} for $n > 2$ and $\log r$ for $n = 2$.

But we must remember that space-time is not Euclidean 4-space. This makes an important difference with respect to singularities. The fundamental harmonic function (236) (or indeed the fundamental harmonic function in Euclidean space of any dimensionality) has a singularity at one point only; but the wave function W of (231) has singularities not only at the event $x_r = \alpha_r$, but all over the null cone

$$(x_r - \alpha_r)\,(x_r - \alpha_r) = 0. \tag{237}$$

For some purposes singularities are convenient, but often they are not. We proceed now, by a simple generalisation of (231), to obtain wave functions which vanish at infinity and have no singularities (GÅRDING [1951]).

Let α_r be a fixed event and β_r a *timelike* vector, so that

$$\beta_r\beta_r = -\beta^2 < 0. \tag{238}$$

Consider the function

$$W = 1/S, \quad S = (x_r - \alpha_r - i\beta_r)\,(x_r - \alpha_r - i\beta_r), \tag{239}$$

which is the same as (231) except that we have replaced α_r by $\alpha_r + i\beta_r$. This change makes no significant difference in the formal work (232)–(234), and we get $\Box W = 0$; we have in (239) a *complex* scalar wave function.

We now break W into its real and imaginary parts. We write

$$\begin{gathered} S = A - iB, \\ A = (x_r - \alpha_r)\,(x_r - \alpha_r) + \beta^2, \quad B = 2\beta_r(x_r - \alpha_r), \end{gathered} \tag{240}$$

both A and B being real. Then

$$W = \frac{1}{S} = \frac{1}{A - iB} = \frac{A + iB}{A^2 + B^2} = U + iV \tag{241}$$

where

$$U = \frac{A}{A^2 + B^2}, \qquad V = \frac{B}{A^2 + B^2}. \tag{242}$$

Now

$$\Box U + i\Box V = \Box W = 0, \tag{243}$$

and therefore

$$\Box U = 0, \quad \Box V = 0. \tag{244}$$

We have in U and V two real scalar wave functions. Note that if we put $\beta_r = 0$, U reduces to the wave function (231) and V vanishes.

It may seem a little confusing to introduce complex quantities in this way, with the imaginary element already present in the fourth components of x_r, α_r and β_r. To avoid this, we might momentarily introduce real coordinates x^r with the diagonal fundamental tensor g_{rs} (1, 1, 1, − 1). Then instead of (239) and (240) we would have

$$\begin{gathered} S = g_{rs}(x^r - \alpha^r - i\beta^r)\,(x^s - \alpha^s - i\beta^s), \\ A = g_{rs}(x^r - \alpha^r)\,(x^s - \alpha^s) + \beta^2, \quad B = 2g_{rs}\beta^r(x^s - \alpha^s), \\ \beta^2 = -g_{rs}\beta^r\beta^s > 0, \end{gathered} \tag{245}$$

and the complex $W = 1/S$ would satisfy the wave equation in the form

$$g_{rs}\frac{\partial^2 W}{\partial x^r \partial x^s} = \frac{\partial^2 W}{\partial x_\varrho \partial x_\varrho} - \frac{1}{c^2}\frac{\partial W}{\partial t^2} = 0. \tag{246}$$

But actually this recourse to real coordinates is unnecessary, the argument given earlier being valid at every step.

The wave functions U and V of (242) *have no singularities anywhere in space-time.* For the denominator $A^2 + B^2$ can vanish only if $A = B = 0$, i.e. if the two equations

$$(x_r - \alpha_r)\,(x_r - \alpha_r) + \beta^2 = 0, \quad \beta_r(x_r - \alpha_r) = 0 \tag{247}$$

are satisfied simultaneously. The first of these equations confines the event x_r to the pseudosphere with centre α_r and radius β, lying inside the null cone; the second equation confines x_r to the 3-flat through α_r orthogonal to β_r. This 3-flat is spacelike since β_r is timelike, and so it cannot cut the pseudosphere. Hence there are no singularities.

Further, U and V *tend to zero as we go to infinity in any space-time direction*; this will be evident later when we write them in the form (286).

If we take β_r to be a null vector (so that $\beta = 0$), we still have wave functions as in (242), but now they have singularities along the locus

$$(x_r - \alpha_r)\,(x_r - \alpha_r) = 0, \quad \beta_r(x_r - \alpha_r) = 0; \tag{248}$$

this is easily seen to be the null line drawn through α_r along β_r, with equations

$$x_r = \alpha_r + \theta\beta_r, \quad -\infty < \theta < \infty. \tag{249}$$

Thus for $\beta = 0$ we get scalar wave functions with no singularities except on a certain straight null line[1].

It is interesting to note that we owe our success in generating the singularity-free wave functions (242) to the indefinite character of the space-time metric. If we try the same trick with the potential $1/r$ of (236), replacing α_ϱ by $\alpha_\varrho + i\beta_\varrho$, we find that the complex potential has singularities on a certain circle.

We shall later use U and V to construct electromagnetic fields which may be regarded as models of (*a*) an electromagnetic material particle when β_r is timelike, and (*b*) of a photon when β_r is null. But before passing to tensor fields, we should mention some other simple scalar wave functions.

[1] It is easy to see [cf. (339)] that, if $\beta = 0$, the most general wave function which depends only on A and B is of the form $W = B^{-1} f(A/B)$; for (239) we have $f(x) = (x - i)^{-1}$.

Consider

$$W = f(a_n x_n), \tag{250}$$

where f is any function and a_r any null vector, so that

$$a_r a_r = 0. \tag{251}$$

We have

$$W_{,rs} = a_r a_s f''(a_n x_n), \quad W_{,rr} = 0, \tag{252}$$

so that W is a wave function. If f is free from singularities, then so is W. Note that a wave function of the form (250) cannot tend to zero at infinity in all space-time directions (as we sometimes wish a wave function to do), because it is of course constant over each of the 3-flats $a_n x_n = \text{const.}$

Other scalar wave functions are

$$W = r^{-1} f(r + ct), \quad W = r^{-1} f(r - ct), \tag{253}$$

r being the spatial distance from a fixed point and the function f being arbitrary. It is easy to verify that these are wave functions; W being a function of r and t only, we have

$$\begin{aligned} \Box W &= \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial W}{\partial r}\right) - \frac{1}{c^2} \frac{\partial^2 W}{\partial t^2} \\ &= \frac{1}{r^2} \frac{\partial}{\partial r}(rf' - f) - \frac{1}{r} f'' = 0. \end{aligned} \tag{254}$$

If we take in particular $f(\xi) = \sin k\xi$ where k is any constant, and add the wave functions (253), we get the wave function

$$W = r^{-1} \sin k(r + ct) + r^{-1} \sin k(r - ct); \tag{255}$$

this may also be written

$$W = 2r^{-1} \sin kr \cos kct. \tag{256}$$

This wave function has no singularity; it goes to zero at infinity for all space-time directions except two, viz. $r = 0$, $t \to \pm \infty$.

§ 13. GENERATION OF A MAXWELLIAN FIELD FROM A SCALAR WAVE FUNCTION

Suppose we are given some scalar wave function W. Let K_{rs} be any *constant* skew-symmetric tensor. Write

$$\phi_r = -K_{rm} W_{,m}. \tag{257}$$

Then

$$\phi_{r,r} = -K_{rm}W_{,mr} = 0 \tag{258}$$

since $W_{,mr}$ is symmetric, and

$$\Box\phi_r = -K_{rm}(\Box W)_{,m} = 0. \tag{259}$$

Thus ϕ_r satisfies the conditions (229), and so the field

$$F_{rs} = \phi_{s,r} - \phi_{r,s} \tag{260}$$

is Maxwellian; explicitly we have

$$\boxed{F_{rs} = K_{rm}W_{,ms} - K_{sm}W_{,mr}.} \tag{261}$$

By this device we generate a Maxwellian field from any scalar wave function W with the aid of any constant skew-symmetric tensor K_{rs}.

It is not suggested that any arbitrary Maxwellian field can be put in this form. Rather we have here a way of generating a certain set of Maxwellian fields; if W is assigned, the set contains six parameters corresponding to the six independent components of K_{rs}.

Let us find the dual of (261). We have

$$F^*_{rs} = \tfrac{1}{2}i\varepsilon_{rspq}F_{pq} = i\varepsilon_{rspq}K_{pm}W_{,mq}, \tag{262}$$

and so, since

$$K_{pm} = -\tfrac{1}{2}i\varepsilon_{pmab}K^*_{ab}, \tag{263}$$

we get

$$F^*_{rs} = \tfrac{1}{2}\varepsilon_{prsq}\varepsilon_{pmab}K^*_{ab}W_{,mq}. \tag{264}$$

Now, as is easily verified,

$$\varepsilon_{prsq}\varepsilon_{pmab} = \begin{vmatrix} \delta_{rm} & \delta_{ra} & \delta_{rb} \\ \delta_{sm} & \delta_{sa} & \delta_{sb} \\ \delta_{qm} & \delta_{qa} & \delta_{qb} \end{vmatrix}, \tag{265}$$

and so

$$F^*_{rs} = K^*_{sq}W_{,rq} - K^*_{rq}W_{,sq} + K^*_{rs}W_{,qq}. \tag{266}$$

The last term vanishes because W is a wave function; *thus we have as the dual of* (261)

$$F^*_{rs} = -K^*_{rm}W_{,ms} + K^*_{sm}W_{,mr}. \tag{267}$$

We can put these formulae in matrix notation, denoting a tensor A_{rs} by $\mathbf{A}$ and its transpose by $\mathbf{A}_t$; for the product of two matrices we have

$$\mathbf{AB} = (A_{rm}B_{ms}). \tag{268}$$

The first suffix indicates the row, the second the column. We put

$$\mathbf{F} = (F_{rs}),\ \mathbf{F}^* = (F^*_{rs}),\ \mathbf{K} = (K_{rs}),\ \mathbf{K}^* = (K^*_{rs}),\ \mathbf{W} = (W_{,rs}), \quad (269)$$

so that

$$\mathbf{F}_t = -\mathbf{F},\ \mathbf{F}^*_t = -\mathbf{F}^*,\ \mathbf{K}_t = -\mathbf{K},\ \mathbf{K}^*_t = -\mathbf{K}^*,\ \mathbf{W}_t = \mathbf{W}. \quad (270)$$

Then the field (261), or equivalently (267), is

$$\mathbf{F} = \mathbf{KW} + \mathbf{WK},\quad \mathbf{F}^* = -\mathbf{K}^*\mathbf{W} - \mathbf{WK}^*. \quad (271)$$

As in (32), the energy tensor for any field is given by

$$2c^2 T_{rs} = F_{mr}F_{ms} + F^*_{mr}F^*_{ms}, \quad (272)$$

or in matrix notation

$$2c^2\mathbf{T} = -\mathbf{F}^2 - \mathbf{F}^{*2}. \quad (273)$$

To get the energy tensor for the field (271) we are to substitute

$$\begin{aligned}\mathbf{F}^2 &= \mathbf{KWKW} + \mathbf{KW}^2\mathbf{K} + \mathbf{WK}^2\mathbf{W} + \mathbf{WKWK},\\ \mathbf{F}^{*2} &= \mathbf{K}^*\mathbf{WK}^*\mathbf{W} + \mathbf{K}^*\mathbf{W}^2\mathbf{K}^* + \mathbf{WK}^{*2}\mathbf{W} + \mathbf{WK}^*\mathbf{WK}^*.\end{aligned} \quad (274)$$

Provided that the derivatives of W have no singularities and vanish sufficiently rapidly as we go to infinity, the integrals

$$M_r = -\int T_{rs}n_s\,dS,\quad H_{rs}(O) = -\int (x_r T_{sp} - x_s T_{rp})n_p\,dS \quad (275)$$

exist, the integrations being taken over a spacelike 3-space with unit normal n_s pointing into the future. By virtue of the conservation equations (31), these integrals are independent of the particular 3-space chosen; they are respectively the 4-momentum of the field [cf. VIII-(247)] and the angular momentum relative to the origin [cf. VIII–(251)]. If we substitute from (272), we can compute M_r and $H_{rs}(O)$ for the field (261). To ease the calculation, we may choose the 3-space to be $x_4 = \text{const.}$, so that $n_\varrho = 0$, $n_4 = i$, and then we have

$$M_r = -i\int T_{r4}\,dS,\quad H_{rs}(O) = -i\int (x_r T_{s4} - x_s T_{r4})dS. \quad (276)$$

To illustrate the above method of generating a Maxwellian field, let us take the scalar wave function (250); from (257) and (261) we get

$$\begin{aligned}\phi_r &= -K_{rm}a_m f'(a_n x_n),\\ F_{rs} &= (K_{rm}a_m a_s - K_{sm}a_m a_r)f''(a_n x_n).\end{aligned} \quad (277)$$

This is in fact the plane wave system (180) in another form; in (180) K_{rs} had to satisfy (187), whereas in (277) it is arbitrary. We reconcile

the two forms by writing L_{rs} for the coefficient in (277) and noting that $L_{rs}a_s = L^*_{rs}a_s = 0$ on account of (251) and (267).

Next consider the wave function

$$W = 1/S, \quad S = x_r x_r, \tag{278}$$

obtained from (231) by changing the origin. As in (233)

$$W_{,rs} = -2\delta_{rs}S^{-2} + 8x_r x_s S^{-3}. \tag{279}$$

By (261) this gives the field

$$F_{rs} = -4K_{rs}S^{-2} + 8S^{-3}(K_{rm}x_m x_s - K_{sm}x_m x_r). \tag{280}$$

On account of the singularity at the origin, the 4-momentum M_r does not exist.

§ 14. AN ELECTROMAGNETIC MODEL OF A MATERIAL PARTICLE

We shall now apply the generating formula (261) to the scalar wave functions U and V of (242):

$$U = \frac{A}{A^2 + B^2}, \quad V = \frac{B}{A^2 + B^2}, \tag{281}$$
$$A = (x_r - \alpha_r)(x_r - \alpha_r) + \beta^2, \quad B = 2\beta_r(x_r - \alpha_r),$$
$$\beta^2 = -\beta_r\beta_r.$$

This gives the Maxwellian fields

$$F_{rs} = K_{rm}W_{,ms} - K_{sm}W_{,mr}, \tag{282}$$

where W stands for U or V. It is clear that there are no singularities because there are none in U or V, and differentiation does not produce them. It is this absence of singularities, combined with a finite 4-momentum (as we shall see below), that makes these fields worth investigating as electromagnetic models of a material particle.

In spite of the simple definition of U and V, the calculation of the field is rather complicated. To shorten the writing, we put

$$A^2 + B^2 = Q, \quad 3A^2 - B^2 = Q_A, \quad 3B^2 - A^2 = Q_B. \tag{283}$$

Then, regardless of the forms of the functions A and B, we have from (281)

$$\begin{aligned} U_{,rs} &= -Q^{-2}[A_{,rs}(A^2 - B^2) + 2B_{,rs}AB] \\ &\quad + 2Q^{-3}[Q_A B(A_{,r}B_{,s} + A_{,s}B_{,r}) - Q_B A(A_{,r}A_{,s} - B_{,r}B_{,s})], \\ V_{,rs} &= -Q^{-2}[B_{,rs}(B^2 - A^2) + 2A_{,rs}AB] \\ &\quad + 2Q^{-3}[Q_B A(A_{,r}B_{,s} + A_{,s}B_{,r}) + Q_A B(A_{,r}A_{,s} - B_{,r}B_{,s})]. \end{aligned} \tag{284}$$

To simplify the work, we shall now use a special frame obtained by shifting the origin to α_r, and taking the time-axis along β_r (we assume that it points into the future). We have then

$$\begin{gathered}\beta_\varrho = 0; \quad \beta_4 = i\beta, \\ A = x_r x_r + \beta^2, \quad B = 2i\beta x_4 = -2\beta ct, \\ A_{,r} = 2x_r, \quad B_{,\varrho} = 0, \quad B_{,4} = 2i\beta, \\ A_{,rs} = 2\delta_{rs}, \quad B_{,rs} = 0.\end{gathered} \tag{285}$$

For such a frame the wave functions U and V are

$$\begin{aligned} U &= \frac{r^2 - c^2t^2 + \beta^2}{(r^2 - c^2t^2 + \beta^2)^2 + (2\beta ct)^2}, \\ V &= \frac{-2\beta ct}{(r^2 - c^2t^2 + \beta^2)^2 + (2\beta ct)^2}, \end{aligned} \tag{286}$$

where r is the spatial distance from the origin ($r^2 = x_\varrho x_\varrho$). Now (284) reduce to

$$\begin{aligned} U_{,\varrho\sigma} &= C\delta_{\varrho\sigma} + Dx_\varrho x_\sigma, & V_{,\varrho\sigma} &= C'\delta_{\varrho\sigma} + D'x_\varrho x_\sigma, \\ U_{,\varrho 4} &= iEx_\varrho, & V_{,\varrho 4} &= iE'x_\varrho, \\ U_{,44} &= F, & V_{,44} &= F', \end{aligned} \tag{287}$$

where the coefficients are functions of r and t, as follows:

$$\begin{aligned} C &= -2Q^{-2}(A^2 - B^2), \\ D &= -8Q^{-3}Q_B A, \\ E &= 8Q^{-3}(Q_A B\beta - Q_B Act), \\ F &= -2Q^{-2}(A^2 - B^2) + 8Q^{-3}[Q_B A(c^2t^2 - \beta^2) - 2Q_A B\beta ct], \\ C' &= -4Q^{-2}AB, \\ D' &= 8Q^{-3}Q_A B, \\ E' &= 8Q^{-3}(Q_B A\beta + Q_A Bct), \\ F' &= -4Q^{-2}AB - 8Q^{-3}[Q_A B(c^2t^2 - \beta^2) + 2Q_B A\beta ct]. \end{aligned} \tag{288}$$

We note the identity

$$0 = U_{,ss} = 3C + Dr^2 + F = 3C + D(A + c^2t^2 - \beta^2) + F, \tag{289}$$

and a similar primed formula.

Putting $W = U$ in (282) and substituting from (287), we get the

U-field as follows:

$$\begin{aligned} F_{\varrho\sigma} &= 2CK_{\varrho\sigma} + D(K_{\varrho\mu}x_\mu x_\sigma - K_{\sigma\mu}x_\mu x_\varrho) + iE(K_{\varrho 4}x_\sigma - K_{\sigma 4}x_\varrho), \\ F_{\varrho 4} &= (F + C)K_{\varrho 4} + DK_{\mu 4}x_\mu x_\varrho + iEK_{\varrho\mu}x_\mu. \end{aligned} \tag{290}$$

By (267) we get F^*_{rs} for the U-field by changing K_{rs} into $-K^*_{rs}$. The V-field is obtained from the U-field by changing C, D, E, F into C', D', E', F'.

To calculate the energy tensor, we have merely to substitute from (290) in (272), but we shall be concerned only with the components T_{r4} involved in the integrals (276) for 4-momentum and angular momentum; these are

$$\begin{aligned} 2c^2T_{\varrho 4} &= F_{\mu\varrho}F_{\mu 4} + F^*_{\mu\varrho}F^*_{\mu 4}, \\ 2c^2T_{44} &= F_{\mu 4}F_{\mu 4} + F^*_{\mu 4}F^*_{\mu 4}. \end{aligned} \tag{291}$$

Moreover we calculate the integrals over $x_4 = 0$, and this simplifies the quantities (288); we have, for $x_4 = 0$,

$$\begin{gathered} A = r^2 + \beta^2, \quad B = 0, \\ Q = A^2, \quad Q_A = 3A^2, \quad Q_B = -A^2, \end{gathered} \tag{292}$$

and so (288) become

$$\begin{gathered} C = -2A^{-2}, \quad D = 8A^{-3}, \quad E = 0, \quad F = -2A^{-2} + 8\beta^2A^{-3}, \\ C' = 0, \quad D' = 0, \quad E' = -8\beta A^{-3}, \quad F' = 0. \end{gathered} \tag{293}$$

Thus the U-field (290) and the corresponding V-field become on $x_4 = 0$

$$U\text{-field:} \begin{cases} F_{\varrho\sigma} = 2CK_{\varrho\sigma} + D(K_{\varrho\mu}x_\mu x_\sigma - K_{\sigma\mu}x_\mu x_\varrho) \\ F_{\varrho 4} = (C + F)K_{\varrho 4} + DK_{\mu 4}x_\mu x_\varrho, \end{cases} \tag{294}$$

$$V\text{-field:} \begin{cases} F_{\varrho\sigma} = iE'(K_{\varrho 4}x_\sigma - K_{\sigma 4}x_\varrho), \\ F_{\varrho 4} = iE'K_{\varrho\mu}x_\mu. \end{cases} \tag{295}$$

Without any further calculation, it is evident that *the angular momentum about the origin is zero, both for the U-field and for the V-field.* For T_{r4} is clearly of even degree in the spatial coordinates x_ϱ, and in the second integral of (276) it is multiplied either by x_4 (which is zero) or by a spatial coordinate, so that the integrand is odd and the integral vanishes.

We could now calculate the 4-momentum M_r, but the result is not interesting in the general case we have so far considered, in which the

tensor K_{rs} and the vector β_r are independent of one another. *Let us take β_r in the timelike principal 2-flat of K_{rs}* (cf. p. 330), *K_{rs} being any constant skew-symmetric tensor which is not null, i.e. its invariants*

$$(KK) = K_{rs}K_{rs}, \quad (KK^*) = K_{rs}K^*_{rs} \tag{296}$$

do not both vanish. By (47) this means that β_r satisfies

$$\tfrac{1}{2}c^{-2}(K_{mr}K_{ms} + K^*_{mr}K^*_{ms})\beta_s = -\varkappa\beta_r, \tag{297}$$

where $\varkappa$ is the invariant [cf. (54)]

$$\varkappa = \tfrac{1}{4}c^{-2}[(KK)^2 + (KK^*)^2]^{\frac{1}{2}}. \tag{298}$$

For the special axes employed above, we have $\beta_\varrho = 0$, $\beta_4 = i\beta$, and so

$$\begin{aligned} K_{\mu 4}K_{\mu 4} + K^*_{\mu 4}K^*_{\mu 4} &= -2c^2\varkappa, \\ K_{\mu\varrho}K_{\mu 4} + K^*_{\mu\varrho}K^*_{\mu 4} &= 0. \end{aligned} \tag{299}$$

Consider now the calculation of the integral

$$M_\varrho = -i\int T_{\varrho 4}dS \tag{300}$$

where $T_{\varrho 4}$ is as in (291), with F_{rs} taken from (294) or (295). In computing the integral we may use the formulae

$$\int f(r)dS = 4\pi\int_0^\infty f(r)\, r^2dr, \quad \int x_\varrho x_\sigma f(r)dS = \frac{4\pi}{3}\delta_{\varrho\sigma}\int_0^\infty f(r)r^4dr. \tag{301}$$

(In these formulae dS is of course the element of 3-volume.) It is not hard then to see that the second expression (299) comes out as a factor in (300), and since it vanishes by (299), we get

$$M_\varrho = 0. \tag{302}$$

Thus by connecting K_{rs} and β_r as we have done, we have made the 4-momentum M_r point in the direction of β_r.

It remains to compute M_4 from

$$M_4 = -i\int T_{44}dS. \tag{303}$$

If we regard the whole field as constituting a material particle, its proper mass m will then be given by $M_4 = im$.

First take the U-field. We have by (294)

$$\begin{aligned} F_{\varrho 4}F_{\varrho 4} = (C+F)^2K_{\varrho 4}K_{\varrho 4} + 2D(C+F)K_{\varrho 4}K_{\mu 4}x_\mu x_\varrho \\ + D^2r^2K_{\mu 4}K_{\nu 4}x_\mu x_\nu, \end{aligned} \tag{304}$$

and so by (301)

$$\int F_{\varrho 4} F_{\varrho 4}\, dS = 4\pi K_{\varrho 4} K_{\varrho 4} \int_0^\infty [(C+F)^2 r^2 + \tfrac{2}{3} D(C+F) r^4 + \tfrac{1}{3} D^2 r^6] dr. \quad (305)$$

When we substitute for the coefficients from (293), remembering that $A = r^2 + \beta^2$, we get an elementary integral which we evaluate by the general formula (for m even)

$$\int_0^\infty \frac{x^m dx}{(x^2+1)^n} = \frac{1.3.5\ldots(m-1).1.3.5\ldots(2n-m-3)}{2.4.6\ldots(2n-2)} \cdot \frac{\pi}{2}; \quad (306)$$

this yields in particular

$$\int_0^\infty \frac{x^4 dx}{(x^2+1)^6} = \tfrac{1}{8} \int_0^\infty \frac{x^2(3x^4 - 2x^2 + 3) dx}{(x^2+1)^6} = \frac{3\pi}{512}. \quad (307)$$

It is the second of these integrals which occurs in (305), and we get the simple result

$$\int F_{\varrho 4} F_{\varrho 4}\, dS = \pi^2 \beta^{-5} K_{\varrho 4} K_{\varrho 4}. \quad (308)$$

Hence

$$\begin{aligned} M_4 &= -\tfrac{1}{2} i c^{-2} \int (F_{\varrho 4} F_{\varrho 4} + F^*_{\varrho 4} F^*_{\varrho 4}) dS \\ &= -\tfrac{1}{2} \pi^2 i c^{-2} \beta^{-5} (K_{\varrho 4} K_{\varrho 4} + K^*_{\varrho 4} K^*_{\varrho 4}). \end{aligned} \quad (309)$$

Using (298), (299) we can pass back to a general frame and make the following statement: *The Maxwellian U-field given by* (281) *has the finite 4-momentum*

$$M_r = \frac{\pi^2 \beta_r}{4c^2 \beta^6} [(KK)^2 + (KK^*)^2]^{\frac{1}{2}}, \quad (310)$$

provided β_r is a timelike vector in the timelike principal 2-flat of the energy tensor of K_{rs} (Fig. 2, p. 330).

The proper mass of the "particle" is then

$$m = \frac{\pi^2}{4c^2 \beta^5} [(KK)^2 + (KK^*)^2]^{\frac{1}{2}}. \quad (311)$$

As we have already seen, the angular momentum about the origin is zero. It follows from VIII–(252) that *the history of the mass-centre is the straight world line drawn through α_r in the direction of β_r*. Regarded as a particle, the system has no intrinsic angular momentum.

We now calculate M_4 for the V-field. By (295)

$$F_{\varrho 4}F_{\varrho 4} = -E'^2K_{\varrho\mu}K_{\varrho\nu}x_\mu x_\nu, \tag{312}$$

and so by (301)

$$\int F_{\varrho 4}F_{\varrho 4}dS = -(4\pi/3)K_{\varrho\mu}K_{\varrho\mu}\int_0^\infty E'^2r^4dr. \tag{313}$$

Substituting for E' from (293) and using the first integral in (307), we get

$$\int F_{\varrho 4}F_{\varrho 4}dS = -\tfrac{1}{2}\pi^2\beta^{-5}K_{\varrho\mu}K_{\varrho\mu}, \tag{314}$$

and hence

$$M_4 = \tfrac{1}{4}\pi^2ic^{-2}\beta^{-5}(K_{\varrho\mu}K_{\varrho\mu} + K^*_{\varrho\mu}K^*_{\varrho\mu}). \tag{315}$$

But for any skew-symmetric tensor

$$K_{\varrho\mu}K_{\varrho\mu} + K^*_{\varrho\mu}K^*_{\varrho\mu} = -2(K_{\varrho 4}K_{\varrho 4} + K^*_{\varrho 4}K^*_{\varrho 4}), \tag{316}$$

and so (315) is the same as (309). *Thus we have the remarkable result that the U-field and the V-field give precisely the same 4-momentum, as in* (310).

We see then that out of a skew-symmetric tensor K_{rs} and a timelike vector β_r, linked together by (297), we can construct in a simple unique manner two Maxwellian fields (U-field and V-field) with the same finite momentum and energy. Each such field may be regarded as a model particle, with proper mass m as in (311). For the rest frame of the particle, the fields are given by (290) and its primed counterpart. The structure of these fields is not very complicated, but space does not permit a description of them here; suffice it to say that the fields go to zero at infinity for every space-time direction, but less rapidly for null directions than for others, so that we have what may be roughly described as a spherical electromagnetic sheet which contracts, dissolves, and expands again. In the infinite past and in the infinite future, the field at any fixed point is zero, but its total energy maintains the constant value mc^2.

The multiplicity of fields obtained in this way is not as great as one might suppose at first sight. If we choose K_{rs} and β_r subject to (297) and fixed values of the three invariants

$$\beta^2, \quad (KK), \quad (KK^*), \tag{317}$$

then it is not hard to see that any two U-fields obtained in this way are transformable into one another by a Lorentz transformation; *this*

means that all such U-fields are intrinsically the same, and this is true for V-fields also. Thus the totality of fields depend essentially only on the values of the three invariants (317).

With regard to the distribution of energy at time $t = 0$, it follows from (305) and (307) that half the energy of the U-field is contained in a sphere of radius r_1 determined by

$$\tfrac{1}{8}\int_0^{r_1/\beta} \frac{x^2(3x^4 - 2x^2 + 3)}{(x^2+1)^6}\,dx = \frac{3\pi}{1024}, \tag{318}$$

and half the energy of the V-field within a sphere of radius r_2 determined by

$$\int_0^{r_2/\beta} \frac{x^4 dx}{(x^2+1)^6} = \frac{3\pi}{1024}. \tag{319}$$

Calculation gives the approximate numerical values

$$r_1/\beta = 0.62, \quad r_2/\beta = 0.85. \tag{320}$$

§ 15. SUPERPOSITION OF ELEMENTARY WAVE FUNCTIONS

The real elementary wave functions U and V of (281) (or the complex combination $W = U + iV$) are analogues of the fundamental potential function $1/r$, but with an essential difference — they have no poles, provided $\beta \neq 0$. Following the analogy with ordinary potential theory, one is led to superimpose elementary wave functions as one superimposes the fields of positive and negative point charges, or of sources and sinks in hydrodynamics.

Let us call the event α_r in (281) the *centre* of the function and the vector β_r its *phase vector*. We can construct wave functions by superimposing elementary wave functions with different centres and phase vectors, each being weighted as we please by a scalar factor. A finite sum obtained in this way can have no singularity provided $\beta \neq 0$ for each of the elementary functions; but singularities may be induced if we superimpose by integration. In carrying out such work, it is simplest to use the complex elementary wave function

$$W = [(x_r - \alpha_r - i\beta_r)\,(x_r - \alpha_r - i\beta_r)]^{-1} \tag{321}$$

until one has found the field F_{rs}, but then to separate the real and

imaginary parts before proceeding to calculate the energy tensor T_{rs}, since this is quadratic.

The "particle" of § 14 seems a rather transitory thing, the field rushing in to the centre and then out again, all with the speed c. That is an oversimplified picture, but it suggests that we might get a particle of longer life by superimposing wave functions with centres along the line $\alpha_r = u\beta_r$ $(-\infty < u < \infty)$. With a weighting factor $f(u)$, we get in this way the wave function

$$W = \int_{-\infty}^{\infty} \frac{f(u)du}{g(u)}, \tag{322}$$

where

$$g(u) = (x_r - u\beta_r - i\beta_r)\,(x_r - u\beta_r - i\beta_r). \tag{323}$$

This may be written

$$g(u) = -\beta^2(u - u_1)\,(u - u_2), \tag{324}$$

where u_1, u_2 are the roots of the quadratic equation $g(u) = 0$; if we take β_r for time axis, we find

$$\begin{aligned} u_1, u_2 &= -i + (ct \pm r)/\beta, \\ u_1 - u_2 &= 2r/\beta, \quad r^2 = x_\varrho x_\varrho. \end{aligned} \tag{325}$$

We note that u_1 and u_2 lie in the lower half of the complex plane of u.

If we choose $f(u) = \text{const.}$, we get $W = 0$, as may be seen immediately on integrating round a contour (say C_1) formed by the real axis and the infinite semi-circle in the upper half-plane of u. *This means that a uniform distribution of elementary wave functions along the line of β_r cancels out.* We would not expect this from the analogy of the potential $1/r$, for there such a superposition gives the logarithmic potential; but the complex character of W introduces phase effects and destructive interference of the fields.

Let us now take for $f(u)$ some analytic function bounded at infinity with a single simple pole (residue a) at $u = u_0$ in the upper half-plane (it must have a pole somewhere if it is not a constant). Integrating round C_1 we get

$$W = \frac{2\pi i a}{g(u_0)}. \tag{326}$$

On the other hand, if we integrate clockwise round C_2, consisting of the real axis and the infinite semi-circle in the lower half-plane, we get

$$W = \frac{2\pi i}{\beta^2} \frac{f(u_1) - f(u_2)}{u_1 - u_2}. \tag{327}$$

We have in (326) and (327) two alternative expressions for the same complex wave function.

We can write $g(u_0)$ in the form

$$g(u_0) = (x_r - \alpha'_r - i\beta'_r)\,(x_r - \alpha'_r - i\beta'_r), \tag{328}$$

where α'_r is some event and β'_r some vector in the (timelike) direction of β_r. Thus, according to (326), *the effect of the superposition* (322) *is to generate a new elementary wave function with centre* α'_r *and phase vector* β'_r.

On the other hand, (327) with (325) gives

$$W = \frac{\pi i}{\beta r}\{f(-i + [ct + r]/\beta) - f(-i + [ct - r]/\beta\}, \tag{329}$$

a sum of functions of the form (253). We conclude that *any wave function of the form*

$$W = \frac{1}{r}[F(ct + r) - F(ct - r)] \tag{330}$$

is that due to some single elementary wave function, provided the function $f(u) = F[\beta(u + i)]$ *is bounded at infinity and has a single simple pole in the upper half plane.*

§ 16. A NEARLY STATIC ELECTROMAGNETIC PARTICLE (β LARGE)

We recall that the model "particle" we have been considering is a Maxwellian field in vacuo with no singularity and with finite momentum and energy. We cannot get a statical model of this type, because there is no statical Maxwellian field satisfying the requirements. But we can approach a statical condition in finite portions of space-time by making β large.

In (286) the parameter β appears homogeneously with r and ct, and hence, if we change β to β', say, the new field will be the same as the old one if expressed in new units of space and time, changed by a suitable factor. But two such fields would differ physically from one another, and it is best to regard the units as fixed, either as c.g.s. units or as those absolute units in which $m = h = c = 1$, where m is the

proper mass of an electron, h is Planck's constant, and c (as always) the speed of light; in terms of c.g.s. units, these absolute units are

$$\begin{aligned} &\text{unit of mass } (= m) = 9.1 \times 10^{-28} \text{ gm}, \\ &\text{unit of length } (= h/mc) = 2.4 \times 10^{-10} \text{ cm}, \\ &\text{unit of time } (= h/mc^2) = 8.0 \times 10^{-21} \text{ sec}. \end{aligned} \tag{331}$$

Largeness and smallness of quantities are best interpreted relative to these absolute units. The unit of length is the Compton wave length (p. 197).

By increasing β we prolong the life of the particle, but we do this by increasing its size. Taking the time-axis in the direction of β_r, the field for $t = 0$ is given accurately by (294) or (295). Half the energy is contained within a sphere of radius r_1 or r_2 as in (320), and since these radii are proportional to β we may say that *for large β the energy is spread over a large volume of order β^3 at $t = 0$.*

To investigate the building up of the field and its decay with time, we go to the origin in (290), putting $x_\varrho = 0$. The U-field is given by

$$F_{\varrho\sigma} = 2CK_{\varrho\sigma}, \quad F^*_{\varrho\sigma} = -2CK^*_{\varrho\sigma}, \tag{332}$$

the second following from the first by (267), so that we have only to consider the coefficient C, which by (288) is

$$C = -2\frac{A^2 - B^2}{(A^2 + B^2)^2} = -2\frac{\beta^4 - 6\beta^2c^2t^2 + c^4t^4}{(\beta^2 + c^2t^2)^4}. \tag{333}$$

The field is an electromagnetic wrench by (299), the vectors E_ϱ and H_ϱ having the same direction or opposite directions, and their magnitudes being proportional to C; they fall to half their initial values when $t = t_1$ where t_1 satisfies

$$\frac{\beta^4 - 6\beta^2c^2t_1^2 + c^4t_1^4}{(\beta^2 + c^2t_1^2)^4} = \frac{1}{2\beta^4}. \tag{334}$$

Clearly t_1 is proportional to β and so we may say that *the life of the particle is long of order β.* A similar argument applies to the V-field.

To make the proper mass finite, we choose the components K_{rs} large of order $\beta^{5/2}$, as required by (311); then by (332) the field at the origin is small of order $\beta^{-3/2}$ and the energy density small of order β^{-3}.

§ 17. MODEL OF A PHOTON WITH $\beta = 0$

The most natural way to get a photon-model from the wave functions (281) is to make β_r a null vector and K_{rs} null [cf. (29)] with β_r pointing

along its principal null direction. We have then

$$\beta = 0, \quad (KK) = 0, \quad (KK^*) = 0, \quad K_{rs}\beta_s = 0, \quad K^*_{rs}\beta_s = 0. \tag{335}$$

This gives a Maxwellian field with singularities along the infinite null line $x_r = \alpha_r + \theta\beta_r$ $(-\infty < \theta < \infty)$.

This model is worthy of attention on account of its mathematical simplicity, but it suffers from a serious defect — it has infinite energy, as we shall now prove.

Using (144), we write

$$K_{rs} = \beta_r p_s - \beta_s p_r, \quad K^*_{rs} = \beta_r q_s - \beta_s q_r, \tag{336}$$

where p_r and q_r are spacelike vectors satisfying

$$p_r\beta_r = q_r\beta_r = p_r q_r = 0, \quad p^2 = q^2. \tag{337}$$

The question of finiteness of energy is so important that we shall not limit ourselves to the wave functions of (281), but take any wave function of the form $W(A, B)$ where $A = x_r x_r$ and $B = 2\beta_r x_r$ with $\beta = 0$. Any such wave function satisfies

$$2\frac{\partial W}{\partial A} + A\frac{\partial^2 W}{\partial A^2} + B\frac{\partial^2 W}{\partial A \partial B} = 0, \tag{338}$$

so that $\partial W/\partial A$ is a homogeneous function of A, B of degree -2, and hence, as remarked earlier, W is of the form

$$W = \frac{1}{B} f\left(\frac{A}{B}\right). \tag{339}$$

When we substitute (336) in (282) and use (337) and (338), we find for the field

$$F_{rs} = 2\frac{\partial^2 W}{\partial A^2}\,[-A(\beta_r p_s - \beta_s p_r) + 2p_n x_n(\beta_r x_s - \beta_s x_r) - B(p_r x_s - p_s x_r)], \tag{340}$$

$-F^*_{rs}$ being given by the same expression with q written for p [cf. (267)]. From (272) and (340) we find the energy tensor to be

$$T_{rs} = \frac{4p^2}{c^2}\left(\frac{\partial^2 W}{\partial A^2}\right)^2 (A\beta_r - Bx_r)(A\beta_s - Bx_s), \quad A = x_n x_n, \quad B = 2\beta_n x_n, \tag{341}$$

a remarkably simple expression; here p^2 is the invariant defined by K_{rs} and β_r as in (147) for F_{rs} and ξ_r. Hence the energy density is

$$-c^2T_{44} = -4p^2\beta_4^2\left[B\frac{\partial^2 W}{\partial A^2}\left(\frac{A}{B}-\frac{x_4}{\beta_4}\right)\right]^2. \tag{342}$$

Differentiation of (339) gives

$$\frac{\partial^2 W}{\partial A^2} = \frac{1}{B^3}f''\left(\frac{A}{B}\right),$$

and thus, defining a function $F(X)$ by

$$F(X) = \left[\left(X-\frac{x_4}{\beta_4}\right)f''(X)\right]^2 \tag{343}$$

(here the ratio x_4/β_4 is a real constant), we can write (342) as

$$-c^2T_{44} = -\frac{4p^2\beta_4^2}{B^4}F\left(\frac{A}{B}\right).$$

To evaluate the total energy of our model photon we have to calculate the integral

$$\begin{aligned} E &= -c^2\int T_{44}\,dS = -4p^2\beta_4^2 I, \\ I &= \int\frac{1}{B^4}F\left(\frac{A}{B}\right)dS, \end{aligned} \tag{344}$$

the integral being taken over the whole of instantaneous space ($x_4 = \text{const.}$).

We shall now show that the integral I diverges. Since the integrand is positive, it is enough to show that the integral diverges when taken through the interior of the null cone $x_r x_r = 0$, i.e. throughout the sphere $r < ct$. Taking our axes so that $\beta_1 = b$, $\beta_2 = \beta_3 = 0$, $\beta_4 = ib$ and introducing cylindrical coordinates x, ϱ, ϕ so that

$$r^2 = \varrho^2 + x^2, \quad dS = \varrho\,dx\,d\varrho\,d\phi, \tag{345}$$

we have

$$A = \varrho^2 + x^2 - c^2t^2, \quad B = 2b(x-ct), \tag{346}$$

and the contribution to I from the interior of the null cone is

$$I_1 = \frac{\pi}{8b^4}\int_{x=-ct}^{ct}\frac{dx}{(x-ct)^4}\int_{\varrho=0}^{(c^2t^2-x^2)^{\frac{1}{2}}}F\left(\frac{\varrho^2+x^2-c^2t^2}{2b(x-ct)}\right)\varrho\,d\varrho, \tag{347}$$

or, with $\xi = ct - x$,

$$I_1 = \frac{\pi}{8b^4}\int_0^{2ct}\frac{d\xi}{\xi^4}\int_{\varrho=0}^{[\xi(2ct-\xi)]^{\frac{1}{2}}} F\left(\frac{\xi(2ct-\xi)-\varrho^2}{2b\xi}\right)\varrho d\varrho. \tag{348}$$

Putting

$$y = \frac{\xi(2ct-\xi)-\varrho^2}{2b\xi}, \tag{349}$$

we get

$$I_1 = \frac{\pi}{8b^3}\int_{\xi=0}^{2ct}\frac{d\xi}{\xi^3}\int_{y=0}^{(2ct-\xi)/2b} F(y)dy. \tag{350}$$

Since the second integral is finite for $\xi = 0$, it follows that I_1 diverges. *Thus the photon-model with $\beta = 0$ has infinite energy*, and in this respect suffers from the same disability as the point charge in electrostatics.

Nevertheless the model with $\beta = 0$ is so simple and fundamental that it is interesting to examine the field, in particular on the 3-flat $\beta_r x_r = 0$, which is the history of a plane wave advancing with speed c and always containing the singularity. We shall now be less general, taking the wave functions U and V of (281) with $\alpha_r = 0$, and instead of using (336) we shall proceed from (282), putting

$$\beta = 0, \quad B = 0, \quad Q = A^2, \quad Q_A = 3A^2, \quad Q_B = -A^2, \qquad A = r^2 - c^2t^2, \tag{351}$$

in (284), so that

$$\begin{aligned} U_{,rs} &= -2A^{-2}\delta_{rs} + 8A^{-3}(x_r x_s - \beta_r\beta_s), \\ V_{,rs} &= -8A^{-3}(x_r\beta_s + x_s\beta_r). \end{aligned} \tag{352}$$

Then by (335)

$$\begin{aligned} K_{rm}U_{,ms} &= -2A^{-2}K_{rs} + 8A^{-3}K_{rm}x_m x_s, \\ K_{rm}V_{,ms} &= -8A^{-3}K_{rm}x_m\beta_s, \end{aligned} \tag{353}$$

and so the fields are (on $\beta_r x_r = 0$) by (282)

$$\begin{aligned} &U\text{-field:}\ F_{rs} = -4A^{-2}K_{rs} + 8A^{-3}(K_{rm}x_m x_s - K_{sm}x_m x_r), \\ &V\text{-field:}\ F_{rs} = -8A^{-3}(K_{rm}x_m\beta_s - K_{sm}x_m\beta_r). \end{aligned} \tag{354}$$

Let us now take a canonical frame so that [cf. (190)]

$$\begin{gathered}\beta_1 = b, \quad \beta_2 = \beta_3 = 0, \quad \beta_4 = ib, \\ K_{12} = -K_{21} = K, \quad K_{24} = -K_{42} = -iK,\end{gathered} \tag{355}$$

the other components of K_{rs} vanishing. Then on $\beta_r x_r = 0$ we have $x_1 + ix_4 = 0$ and $A = x_2^2 + x_3^2 = \varrho^2$, say, this ϱ being distance from the space-track of the singularity. The V-field vanishes, and for the U-field we find

$$\begin{gathered}E_1 = iF_{14} = 0, \quad E_2 = iF_{24} = -\frac{4K}{\varrho^4} + \frac{8Kx_2^2}{\varrho^6}, \\ E_3 = iF_{34} = \frac{8Kx_2x_3}{\varrho^6}, \quad H_1 = F_{23} = 0, \\ H_2 = F_{31} = -\frac{8Kx_2x_3}{\varrho^6}, \quad H_3 = F_{12} = -\frac{4K}{\varrho^4} + \frac{8Kx_2^2}{\varrho^6},\end{gathered} \tag{356}$$

so that

$$E = H = 4K/\varrho^4. \tag{357}$$

The lines of force form two orthogonal systems of circles, as shown in Fig. 6.

§ 18. MODEL OF A PHOTON WITH β SMALL

Fundamental as the above model is mathematically, the divergence of the integral for energy makes us turn to another model in which β is not zero but very small.

Choosing K_{rs} small of order β^3, so that the invariants may be written

$$\begin{aligned}(KK) &= p\beta^6, \\ (KK^*) &= q\beta^6\end{aligned} \tag{358}$$

with p and q finite, we see from (310) and (311) that the 4-mo-

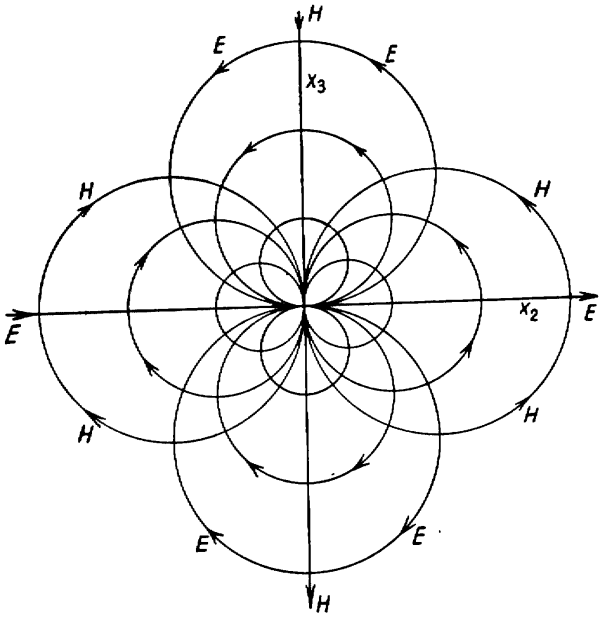


Fig. 6 – Model of photon with $\beta = 0$, showing lines of force in the plane wave advancing with the singularity, viewed from in front

mentum and proper mass are

$$M_r = m\lambda_r, \quad m = \frac{\pi^2\beta}{4c^2}(p^2 + q^2)^{\frac{1}{2}}, \tag{359}$$

where λ_r is the unit vector $\lambda_r = \beta_r/\beta$.

We have then a system with very small proper mass, or, equivalently, with very small proper energy. Such a system is mechanically insignificant when viewed in its rest-frame (time-axis along β_r), or any frame adjacent thereto; it becomes significant only when viewed from a frame relative to which it is moving fast, or equivalently when the components of the unit vector λ_r in (359) become sufficiently large to balance the smallness of the factor m. This means that the individual components of β_r are finite although the magnitude β is small. Thus the model photon now under consideration involves two limiting processes: (i) the Lorentz-invariant condition that β shall be small, and (ii) the choice of a frame to make β_r finite. The relative energy of the system being

$$E = mc^2\lambda_4/i, \tag{360}$$

(359) gives

$$\beta_4 = \frac{4iE}{\pi^2}(p^2 + q^2)^{-\frac{1}{2}}, \tag{361}$$

a formula which we shall use later.

The U-field and the V-field are determined by (282) with $U_{,rs}$ and $V_{,rs}$ as in (284); examining (284), we see that $U_{,rs}$ and $V_{,rs}$ are finite at a general event, and so the field is small of order β^3 since K_{rs} is small of that order. But there is a region in which the field is not so small, viz. that region in which A and B (and hence Q) are small. Making $\alpha_r = 0$ by choice of origin, we have

$$A = x_r x_r + \beta^2, \quad B = 2\beta_r x_r, \tag{362}$$

and so this region of concentrated field consists of those events which are close to both $A = 0$ (a pseudosphere) and $B = 0$ (a 3-flat). That such events exist is clear, for if we put $x_r = \theta\beta_r$ (θ any finite factor) we have

$$A = (1 - \theta^2)\beta^2, \quad B = -2\theta\beta^2, \quad Q = (1 + \theta^2)^2\beta^4, \tag{363}$$

which values are certainly small.

As we saw in § 14, the infinite line $x_r = \theta\beta_r$ $(-\infty < \theta < \infty)$ is the history of the mass-centre; it is moreover the locus of singularities when $\beta = 0$. Let us calculate the energy tensor on this line; it is easiest to do so by an exact argument, valid for any value of β (except zero).

On this line we have by (284)

$$\begin{aligned} U_{,rs} &= -2Q^{-2}(A^2 - B^2)\delta_{rs} + 8Q^{-3}[2Q_A B\theta - Q_B A(\theta^2 - 1)]\beta_r\beta_s, \\ V_{,rs} &= -4Q^{-2}AB\delta_{rs} + 8Q^{-3}[2Q_B A\theta + Q_A B(\theta^2 - 1)]\beta_r\beta_s, \end{aligned} \tag{364}$$

or, when we substitute for θ and $\theta^2 - 1$ from (363),

$$\begin{aligned} U_{,rs} &= -2Q^{-2}(A^2 - B^2)\delta_{rs} + 8Q^{-3}(-Q_A B^2 + Q_B A^2)\lambda_r\lambda_s, \\ V_{,rs} &= -4Q^{-2}AB\delta_{rs} + 8Q^{-3}(-Q_A - Q_B)AB\lambda_r\lambda_s, \end{aligned} \tag{365}$$

where $\lambda_r = \beta_r/\beta$. With Q_A, Q_B as in (283), these expressions reduce to

$$U_{,rs} = C(\delta_{rs} + 4\lambda_r\lambda_s), \quad V_{,rs} = C'(\delta_{rs} + 4\lambda_r\lambda_s), \tag{366}$$

where C and C' are the invariants which occur in (288), and now have the values

$$\begin{aligned} C &= -2\frac{A^2 - B^2}{Q^2} = -\frac{2}{\beta^4}\frac{(1-\theta^2)^2 - 4\theta^2}{(1+\theta^2)^4}, \\ C' &= -4\frac{AB}{Q^2} = \frac{8}{\beta^4}\frac{\theta(1-\theta^2)}{(1+\theta^2)^4}. \end{aligned} \tag{367}$$

Proceeding with the U-field, (282) gives

$$F_{rs} = 2C[K_{rs} + 2(K_{rm}\lambda_m\lambda_s - K_{sm}\lambda_m\lambda_r)]; \tag{368}$$

F^*_{rs} is given by changing K_{pq} into $-K^*_{pq}$ by (267). The energy tensor is given by

$$2c^2 T_{rs} = F_{mr}F_{ms} + F^*_{mr}F^*_{ms}, \tag{369}$$

and we substitute (368) in this, but in making the calculation we are to remember the linkage (297) between K_{rs} and β_r, which gives

$$(K_{mr}K_{ms} + K^*_{mr}K^*_{ms})\lambda_s = -2\varkappa c^2\lambda_r; \tag{370}$$

we have also $\lambda_r\lambda_r = -1$. *There results the following very simple exact expression for the energy tensor on the history of the mass-centre*:

$$c^2 T_{rs} = 2C^2(K_{mr}K_{ms} + K^*_{mr}K^*_{ms}). \tag{371}$$

This is for the U-field; for the V-field, read C' for C.

We now reassume the smallness of β and note that C and C' are large of order β^{-4} and K_{rs} small of order β^3, so that, *on the history of the mass-centre, T_{rs} is large of order β^{-2}, provided θ is finite.*

For large θ, (367) gives approximately

$$C = -\frac{2}{\beta^4\theta^4}, \quad C' = -\frac{8}{\beta^4\theta^5}. \tag{372}$$

Now at a general event in space-time, T_{rs} is small of order β^6 since F_{rs} is small of order β^3. Thus, as we follow the history of the mass-centre (θ increasing), the concentration of energy at the mass-centre disappears as soon as T_{rs}, as given in (371) or the corresponding expression with C', becomes small of order β^6, i.e. when C or C' becomes finite. (We are not of course thinking of a superposition of the U-field and the V-field, but of two distinct photon-models.) It follows then from (372) that the disappearance or diffusion of the concentration corresponds to

$$\begin{aligned} \theta &= a/\beta \qquad \text{for the } U\text{-field}, \\ \theta &= (a/\beta)^{4/5} \text{ for the } V\text{-field}, \end{aligned} \tag{373}$$

a being some finite constant with the dimensions of a length. But the event $x_r = \theta\beta_r$ occurs at time t given by

$$ct = \theta\beta_4/i = 4\pi^{-2}E\theta(p^2 + q^2)^{-\frac{1}{2}} \tag{374}$$

by (361), and so, by (359), the time to disappearance of the concentration is given by

$$ct = \frac{4E}{\pi^2}\,\frac{a}{\beta}\,(p^2 + q^2)^{-\frac{1}{2}} = \frac{aE}{mc^2} \tag{375}$$

for the U-field, and by

$$ct = \frac{4E}{\pi^2}\left(\frac{a}{\beta}\right)^{4/5}(p^2 + q^2)^{-\frac{1}{2}} = \frac{4E}{\pi^2}\left(\frac{a\pi^2}{4mc^2}\right)^{4/5}(p^2 + q^2)^{-1/10} \tag{376}$$

for the V-field.

To sum up, we had in § 17 a photon-model with $\beta = 0$, consisting of a Maxwellian field with a null line of singularities, and with infinite energy. Now we have a photon-model with β small, consisting of a Maxwellian field without singularities, the proper mass being small and the relative energy finite when the system is viewed by suitable observers. For such observers the mass-centre moves with a speed

only slightly less than c, and at it there is a high concentration of energy which diffuses away after a long time, this time being of order m^{-1} for the U-field and $m^{-4/5}$ for the V-field.

§ 19. NULL 3-SPACES AND GREEN'S THEOREM

At an event x_r on a 3-space $f(x) = 0$ there exists a vector $f_{,r}$ $(= \partial f/\partial x_r)$ which is normal to the 3-space in the sense that

$$df = f_{,r}dx_r = 0 \tag{377}$$

for every infinitesimal displacement dx_r in the 3-space. If $f_{,r}$ is a null vector, so that

$$f_{,r}f_{,r} = 0 \tag{378}$$

at some event on the 3-space, we say that the event is a *null point*, and if (378) holds over the 3-space we say that $f(x) = 0$ is a *null 3-space*.

In fact (378) is *the partial differential equation of all null 3-spaces*; written out explicitly, it reads

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 - \frac{1}{c^2}\left(\frac{\partial f}{\partial t}\right)^2 = 0. \tag{379}$$

Just as the wave equation $\Box W = 0$ is the space-time analogue of the harmonic equation $\triangle V = 0$, (379) is analogous to Hamilton's equation [cf. SYNGE 1937b, p. 19]

$$\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2 = \mu^2. \tag{380}$$

At a null point the normal vector $f_{,r}$ lies in the 3-space, since (377) is satisfied if we take dx_r proportional to $f_{,r}$. There is no unit normal vector at a null point; if the 3-space were not null, the unit normal vector would be $n_r = f_{,r} \mid f_{,s}f_{,s} \mid^{-\frac{1}{2}}$, but at a null point this is meaningless.

Null 3-spaces have an intimate connection with Maxwellian fields, but before discussing that we must return to Green's theorem [VIII–(96)]

$$\int_R \frac{\partial X}{\partial x_r}\, d\tau = \int_S \varepsilon(n) n_r X dS, \tag{381}$$

n_r being the unit vector pointing out of R, and ask what happens to

the right hand side when the 3-space S, which bounds the domain R of space-time, is null, at least in part. We shall approach this as a limit, taking first a 3-space S which is not null.

Let λ_r be a unit vector pointing out of R. Through the element dS of (381) draw a tube of world lines parallel to λ_r and let dS_0 be its normal section. Then by VIII–(69) and VIII–(77) we have

$$dS_0 = \varepsilon(n)\lambda_s n_s dS, \tag{382}$$

and so

$$\varepsilon(n)n_r dS = \frac{n_r}{\lambda_s n_s} dS_0. \tag{383}$$

The right hand side is unchanged by either of the following operations: (i) change of n_r to any codirectional vector N_r, and (ii) change of λ_r and consequently dS_0, this invariance being evident from (382) since dS is not changed when we change λ_r.

Let us then write

$$dS_r = \frac{N_r}{\lambda_s N_s} dS_0, \tag{384}$$

recognising this as an infinitesimal vector defined by an element of S, independent of the magnitude and sense of the normal vector N_r and independent of the choice of λ_r, provided it is a unit vector pointing out of R. If S is not null, we have by (383)

$$\varepsilon(n)n_r dS = dS_r, \tag{385}$$

and so Green's formula (381) may also be written

$$\boxed{\int_R \frac{\partial X}{\partial x_r} d\tau = \int_S X dS_r.} \tag{386}$$

But dS_r remains meaningful when S becomes null by continuous transformation, because it depends only on the ratios of N_r, and *so we have in* (386) *a form of Green's theorem valid when S is a null 3-space.* N_r is now any normal vector, and we can take $N_r = f_{,r}$ if S is $f(x) = 0$; we can use any unit vector λ_r we like (and indeed different vectors at different points on S), the only requirement being that it should point out of R.

§ 20. ELECTROMAGNETIC SHOCK WAVES

We have already seen in § 10 that plane electromagnetic shock waves (across which the field is discontinuous) exist as valid limits in Maxwellian theory. We shall now consider shock waves which are not plane.

For a Maxwellian field, satisfying (10), (386) gives

$$\int_S F_{rs}\,dS_s = 0, \quad \int_S F^*_{rs}\,dS_s = 0 \tag{387}$$

for any closed 3-space S. It is clear from the arbitrary character of S that (387) implies (10), and indeed we might accept (387) as the basic equations of the electromagnetic field, being more general than (10), for they do not involve the derivatives and can be applied directly to piecewise continuous fields.

If we take a 3-space Σ and apply (387) to a small flat cylinder enclosing a portion of Σ, we get, as the *junction conditions* across Σ,

$$(F^{(1)}_{rs} - F^{(2)}_{rs})dS_s = 0, \quad (F^{*(1)}_{rs} - F^{*(2)}_{rs})dS_s = 0, \tag{388}$$

where (1) and (2) refer to the two sides of Σ; by (384) we can also write these conditions as

$$(F^{(1)}_{rs} - F^{(2)}_{rs})N_s = 0, \quad (F^{*(1)}_{rs} - F^{*(2)}_{rs})N_s = 0, \tag{389}$$

N_r being a vector normal to Σ. Applying now the argument used at (187), we see that $(F^{(1)}_{rs} - F^{(2)}_{rs})$ must be null, and (more important) N_r must be a null vector. *Thus electromagnetic shock waves are null surfaces; equivalently, they are propagated with speed c.*

Consider now the propagation of light into darkness, or (what is the same thing) consider a region R of space-time with a Maxwellian field inside it, but with no field outside a 3-space Σ which bounds R. Then Σ must be a null 3-space, and on it we must have

$$F_{rs}N_s = 0, \quad F^*_{rs}N_s = 0, \tag{390}$$

N_r being normal to Σ.

As for the energy tensor, the conservation equations $T_{rs,s} = 0$ imply

$$\int T_{rs}\,dS_s = 0, \tag{391}$$

for any closed 3-space S, and so we are led to the junction conditions

$$(T^{(1)}_{rs} - T^{(2)}_{rs})N_s = 0, \tag{392}$$

or, in the case of transition from field to no-field,

$$T_{rs}N_s = 0. \tag{393}$$

Actually (392) is an algebraic consequence of (389); a proof based on (97) is easy, and is left to the reader.

It is interesting to consider the conditions (390) in connection with the model photon of § 17, for which $\beta = 0$. On the null cone ($x_r x_r = 0$) we have $N_r = x_r$, and (340) and the similar equation for F^*_{rs} lead to (390). *This means that the field of the model photon with* $\beta = 0$ *may be broken off at the null cone, and the part outside put equal to zero.* The model then appears to any observer as an expanding sphere of electromagnetic field, with a singularity on the sphere and no field outside it. But the energy of the field is not made finite by cutting it off like this [cf. (350)].

CHAPTER X

FIELDS AND CHARGES

§ 1. THE DISCRETE AND CONTINUOUS METHODS

Chapter IX dealt with the electromagnetic field in vacuo; we have now to consider electric charges, the fields due to them and the influence of fields on them.

There are two approaches. In the first method (*discrete method*) we regard the charges as world lines in space-time (equivalently, moving points) on which the field has singularities, the field elsewhere being Maxwellian in the sense of Ch. IX, i.e. the partial differential equations IX–(10) are satisfied. In the second method (*continuous method* or method of pure field theory) we think of the charges as smeared over space-time and represented at each event by a 4-vector J_r; Maxwell's equations IX–(10) are modified by inserting J_r instead of zero on the right hand side of the first equation, and other equations are supplied to make the problem of finding F_{rs} and J_r mathematically determinate.

The two methods come close to one another in limiting cases. If, in the continuous method, we think of J_r as concentrated in thin world tubes, then we approach the discrete method. On the other hand, if we take a great number of singular world lines and apply statistical ideas, we can bring the discrete method close to the continuous. But these transitions are not always easy, and the best plan is to develope the discrete and continuous methods separately and later establish the connections between them. We shall treat the discrete method in §§ 2–6 and then proceed with the continuous method.

The central problem of fields and charges may be broken down into three questions:

(i) How does an electromagnetic field behave in the absence of charges?

(ii) What is the electromagnetic field due to moving charges?

(iii) How do charges move in a given field?

The first question has been answered in Ch. IX; there are, it is true, other theories of the electromagnetic field in vacuo, but the Maxwellian theory is the commonly accepted one, and we shall confine our attention to it. The present chapter is devoted to the second and third questions, with a brief account in § 11 of field theory in moving matter.

§ 2. THE COULOMB FIELD OF AN ELECTRIC CHARGE

Regarding an electric charge, or charged particle, as a world line of singularities, we start with the case where that world line is straight, so that an observer S_0 who chooses his time-axis parallel to the world line sees a point charge at rest. *We shall assume that the field of a point charge e at rest is the familiar Coulomb*[1] *field,*

$$E_\varrho = \frac{e}{4\pi}\,\frac{x_\varrho - x'_\varrho}{r^3}, \qquad H_\varrho = 0, \tag{1}$$
$$r^2 = (x'_\sigma - x_\sigma)\,(x'_\sigma - x_\sigma).$$

Here x_ϱ are the coordinates of the point at which the field is calculated, and x'_ϱ those of the point charge.

Although we are at present concerned only with the field due to a charge, deferring to § 4 the question of the force exerted by the field on a charge, the *rational electrostatic units* used in (1) and throughout are best described by saying that the force (in dynes if we use c.g.s. units) between charges e, e' at a distance r apart is

$$\frac{ee'}{4\pi r^2}; \tag{2}$$

the inclusion of the factor 4π here removes it from Maxwell's equations, and that is why the rational units are to be preferred in general theory.[2]

By (1) and IX--(6) the electromagnetic tensor due to a charge at rest is

$$F_{\varrho 4} = -\frac{ie}{4\pi}\,\frac{x_\varrho - x'_\varrho}{r^3} = \frac{\partial}{\partial x_\varrho}\left(\frac{ie}{4\pi r}\right), \qquad F_{\varrho\sigma} = 0, \tag{3}$$

[1] The electrostatic inverse square law is commonly attributed to C. A. Coulomb; actually he was anticipated by J. Priestley (WHITTAKER [1951, p. 53]).

[2] These rational units were introduced by O. Heaviside; if we use Gaussian units, we delete the 4π in (1) and (2) and throughout §§ 2–5, but we put $4\pi J_r$ instead of J_r in the Maxwell equations of § 7.

and this is derivable [cf. IX–(151)] from a 4-potential

$$\phi_\varrho = 0, \quad \phi_4 = \frac{ie}{4\pi r}. \tag{4}$$

We verify at once that $\Box\, \phi_r = 0$, $\phi_{n,n} = 0$, so that (3) is a Maxwellian field in empty space-time with a line of singularities on $x_\varrho = x'_\varrho$.

Let us now pass from the special observer S_0, for whom the charge is at rest, to a general Galileian observer S, for whom the charge has a constant velocity. In Fig. 1, L is the straight world-line of the charge and $P(x)$ an event at which the field is to be calculated. We draw the complete null cone from P, the past sheet N' cutting L at $Q'(x')$ and the future sheet N'' cutting L at $Q''(x'')$. Let λ_r be the unit vector on L, pointing into the future, and let w', w'' be the invariants

$$w' = \lambda_n(x'_n - x_n), \quad w'' = \lambda_n(x''_n - x_n). \tag{5}$$

Consider the 4-vector

$$\Phi_r = \lambda_r\left(\frac{A'}{w'} - \frac{A''}{w''}\right), \tag{6}$$

where A' and A'' are any two constants satisfying

$$A' + A'' = 1. \tag{7}$$

We shall show that the 4-vector Φ_r of (6) is in fact the 4-vector ϕ_r of (4) to within a constant factor. To do this, we take L for time-axis, so that S becomes S_0; we have then

$$\begin{aligned} \lambda_\varrho = 0, \quad \lambda_4 = i, \\ w' = i(x'_4 - x_4) = c(t - t') = r, \\ w'' = i(x''_4 - x_4) = c(t - t'') = -r, \end{aligned} \tag{8}$$

where r is the spatial distance of P from L when measured in S_0 (the space-time perpendicular MP in Fig. 1). Thus in S_0 the components of Φ_r are

$$\Phi_\varrho = 0, \quad \Phi_4 = i\left(\frac{A'}{r} + \frac{A''}{r}\right) = \frac{i}{r}, \tag{9}$$

agreeing with (4) to within a constant factor. We conclude that *a charge e with*

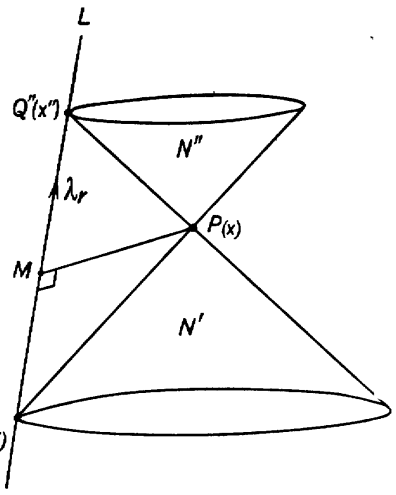


Fig. 1 – Space-time diagram for the Coulomb field

straight world line in the direction of a unit vector λ_r *produces an electromagnetic field* F_{rs} *given by*

$$F_{rs} = \phi_{s,r} - \phi_{r,s}, \quad \phi_r = \frac{e\lambda_r}{4\pi}\left(\frac{A'}{w'} - \frac{A''}{w''}\right), \tag{10}$$

with w', w'' *as in* (5) *and* A', A'' *any constants satisfying* (7).

Any choice of A', A'' consistent with (7) gives the same field. With $A' = 1$, $A'' = 0$, we get

$$\phi_r = \frac{e\lambda_r}{4\pi w'}, \tag{11}$$

and, since this depends on Q' and not on Q'', we think of this potential as "due to" Q' and speak of it as a *retarded potential*, P being later than Q' in time. Similarly with $A' = 0$, $A'' = 1$ we get

$$\phi_r = -\frac{e\lambda_r}{4\pi w''}, \tag{12}$$

and we speak of this as an *advanced potential*, P being earlier than Q''. Thus *the potential of the Coulomb field is a superposition of a retarded potential and an advanced potential, with weighting factors arbitrary except for* (7). The significance of this splitting into retarded and advanced potentials will appear in the next section.

To find explicitly the electric and magnetic 3-vectors due to a charge in *uniform* motion, we have a choice of two methods. We can transform the Coulomb field for the charge at rest by the method of IX–§ 2, or we can proceed directly from (10), taking $A' = 1$, $A'' = 0$ for simplicity. If we take the charge e moving parallel to the axis Ox_1 with velocity v and apply the first method, the field at the event (x_ϱ, t) comes out to be

$$\begin{aligned} &E_1 = \Lambda(x_1 - x_1'), \quad E_2 = \Lambda(x_2 - x_2'), \quad E_3 = \Lambda(x_3 - x_3'), \\ &H_1 = 0, \quad H_2 = -\Lambda(x_3 - x_3')v/c, \quad H_3 = \Lambda(x_2 - x_2')v/c, \end{aligned} \tag{13}$$

where

$$\begin{aligned} \Lambda = \frac{e}{4\pi r^3}\,\frac{\gamma}{[1 + (\gamma^2 - 1)\cos^2\theta]^{3/2}}, \qquad \gamma = (1 - v^2/c^2)^{-\frac{1}{2}}, \\ r^2 = (x_\varrho - x_\varrho')(x_\varrho - x_\varrho'), \quad \cos\theta = (x_1 - x_1')/r; \end{aligned} \tag{14}$$

here x_ϱ' are the coordinates of the moving charge at time t. In the limit $v \to c$, we have $\gamma \to \infty$ and the field disappears except for $\theta = \frac{1}{2}\pi$,

i.e. on the moving plane $x_1 = x_1'$, and there it becomes infinite, so that we have an electromagnetic sheet of infinite strength [cf. IX–§ 10].

From (13) we can calculate the field due to a linear current, obtaining the usual formula (cf. McCREA [1949, p. 53]).

§ 3. THE FIELD OF AN ACCELERATED CHARGE

We shall now generalise the field (10) to the case where the charge has a motion with acceleration, or, equivalently, where its world line L is curved, as shown in Fig. 2, which is constructed in the same way as Fig. 1. The unit tangent vector to L ($\lambda_r = dx_r/ds$) is no longer a constant; let its values at Q', Q'' be λ_r', λ_r'', respectively.

An obvious formal generalisation of (10) is

$$\begin{aligned} F_{rs} &= \phi_{s,r} - \phi_{r,s}, \\ \phi_r &= \frac{e}{4\pi}\left(\frac{A'\lambda_r'}{w'} - \frac{A''\lambda_r''}{w''}\right), \\ w' &= \lambda_n'(x_n' - x_n), \\ w'' &= \lambda_n''(x_n'' - x_n), \\ A' + A'' &= 1. \end{aligned} \tag{15}$$

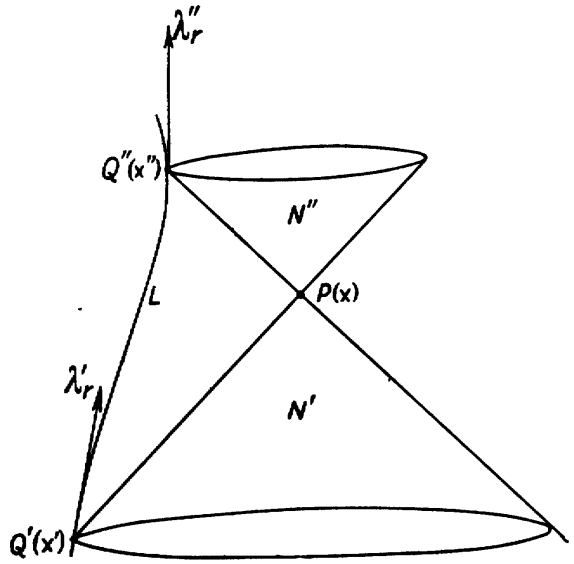


Fig. 2 – Space-time diagram for the field of an accelerated charge

This certainly reduces to (10) when L is straight ($\lambda_r = \lambda_r' = \lambda_r''$), but it is by no means obvious that it is a Maxwellian field, i.e. that the equations

$$\Box \phi_r = 0, \quad \phi_{r,r} = 0 \tag{16}$$

are satisfied. Such is in fact the case, as may verified by differentiation of (15) in the manner of Appendix B, p. 422, but the calculations are tedious and we shall use instead a method due to CONWAY [1903], [1953, p. 13], based on the theory of residues.

Let L have the equations $x_r = f_r(u)$, u being any monotonic parameter. We suppose the functions $f_r(u)$ defined, not only for real values of u, but for complex values also. Consider the equation

$$S(x, u) \equiv [x_n - f_n(u)]\,[x_n - f_n(u)] = 0, \tag{17}$$

x_r being the coordinates of P (Fig. 2). This equation has real roots

($u = u'$, $u = u''$, say) corresponding to Q', Q'' respectively. In the complex plane of u draw a contour C' enclosing u' but no other root of (17), and consider the integral

$$\Phi_r(x) = -i\int_{C'} \frac{df_r}{du}\,\frac{du}{S(x,u)}. \tag{18}$$

We know that $1/S$ is a wave function [cf. IX–(231)], and hence $\Box\Phi_r = 0$, since we may differentiate under the integral sign. Further

$$\frac{\partial\Phi_r}{\partial x_r} = 2i\int_{C'} \frac{df_r}{du}\,\frac{x_r - f_r(u)}{S^2}\,du = i\int_{C'} \frac{\partial}{\partial u}\left(\frac{1}{S}\right)du = 0. \tag{19}$$

Thus Φ_r satisfies the equations (16), and is therefore the 4-potential of a Maxwellian field in vacuo. It has no singularity except on L.

We now evaluate Φ_r by the method of residues. Expanding at the root u', we have

$$S(x,u) = -2(u-u')\,[x_n - f_n(u')]\,\frac{df_n(u')}{du'} + \ldots, \tag{20}$$

and so the residue at this pole of the integrand in (18) is

$$-\frac{df_r(u')/du'}{2[x_n - f_n(u')]df_n(u')/du'} = -\frac{\lambda'_r}{2(x_n - x'_n)\lambda'_n} = \tfrac{1}{2}\,\frac{\lambda'_r}{w'}. \tag{21}$$

Therefore

$$\Phi_r(x) = -i(2\pi i)\,\tfrac{1}{2}\,\frac{\lambda'_r}{w'} = \frac{\pi\lambda'_r}{w'}. \tag{22}$$

Since Φ_r satisfies (16), it follows that ϕ_r as in (15), with $A'' = 0$, satisfies (16). A similar argument applies to ϕ_r with $A' = 0$, and thus we see that *the field* (15) *is Maxwellian and reduces to the Coulomb field if L is straight, no matter how the weights A', A'' are assigned, subject to* $A' + A'' = 1$.

We shall obtain (15) again in § 8 by the continuous method.

The Coulomb field (10) was independent of the choice of A' and A'', but that is not true of (15), which in fact gives, not a unique field for an accelerating charge, but a single infinity of fields which satisfy the two conditions: (i) they are Maxwellian, with no singularities except on L, and (ii) they reduce to (10) when L is straight.

The field (15) may be regarded as due to the superposition of

retarded and advanced potentials. The idea of causality might lead us to reject the advanced potential, since it makes the field at P depend on an event later than P in the time-scales of all Galileian observers; in that case we put $A' = 1$, $A'' = 0$ and obtain

$$\phi_\tau = \frac{e\lambda'_\tau}{4\pi w'}, \qquad w' = \lambda'_n(x'_n - x_n), \tag{23}$$

which is the *Liénard-Wiechert retarded potential* (LIÉNARD [1898], WIECHERT [1900]). This is commonly accepted as defining the field due to an accelerated charge, but it is better (because more general) to keep an open mind about the rejection of the advanced potential and regard (15) as the fundamental formula; we can always pass to (23) by putting $A' = 1$, $A'' = 0$.

The explicit expressions for the electric and magnetic 3-vectors for the field (15) are complicated, but if we use the retarded potential (23), assume the velocity of the charge small compared with c, and calculate the field at a great distance from the charge, we get the following simple approximate expressions:

$$E_\varrho = \frac{e}{4\pi c^2 r}\left(\frac{x_\varrho f'_\sigma x_\sigma}{r^2} - f'_\varrho\right), \qquad H_\varrho = \frac{e}{4\pi c^2 r^2}\varepsilon_{\varrho\mu\nu} f'_\mu x_\nu, \tag{24}$$

where the origin is taken at Q' (Fig. 2) and f'_ϱ is the acceleration there ($f'_\varrho = d^2x'_\varrho/dt^2$ on L at Q'). From (24) we can calculate the Poynting vector and hence the flux of energy outwards across a large sphere; it comes out to be[1]

$$F = \frac{1}{6\pi}\frac{e^2}{c^3} f'^2. \tag{25}$$

This formula for *the radiation of energy from an accelerated charge* is due to LARMOR [1897]. (To convert to Gaussian units, replace $1/6\pi$ by $2/3$).

If we carry out the rather more complicated calculations in which both the advanced and retarded potentials are retained, the flux of energy comes out to be

$$F = \frac{1}{6\pi}\frac{e^2}{c^3}(A'^2 f'^2 - A''^2 f''^2); \tag{26}$$

[1] See Appendix B, p. 422, for an exact calculation.

here f' and f'' are the magnitudes of the accelerations at Q' and Q'' respectively.

These are the essential formulae to which reference must be made in any discussion of the question: "Does an accelerated charge radiate energy?" If we accept the retarded potential (23), then energy is radiated out according to the Larmor formula (25). But if we mix the advanced and retarded potentials, it is not necessarily so; if the charge moves in a circle with constant speed, then $f' = f''$, and we get $F = 0$ by (26) if we put $A' = A'' = \frac{1}{2}$, i.e. if we give equal weights to the advanced and retarded potentials. We shall return to this question in § 6.

§ 4. THE PONDEROMOTIVE FORCE

Consider a charged particle moving in a given Maxwellian field F_{rs}. The total field at any event is the sum of F_{rs} and the field due to the charge (say F'_{rs}), calculated as in § 3. Our question is: How does the particle move?

The most natural approach to this question would be to apply the law of conservation of 4-momentum, writing down a statement to the effect that the total flux of 4-momentum across any closed 3-space is zero [cf. VII–§ 4, VIII–§ 16]; this flux would be due in part to the 4-momentum of the particle ($m\lambda_r$ where $\lambda_r = dx_r/ds$) and in part to the energy tensor of the field $F_{rs} + F'_{rs}$. But this method fails on account of the divergence, at the world line of the particle, of the integral for the flux of 4-momentum of the field. We shall come back to this approach in dealing with the continuous method in § 9, but, as far as the discrete method is concerned, the best plan is to make a simple ad hoc hypothesis.

Accordingly we assume that *a charged particle in a field F_{rs} experiences a 4-force*

$$X_r = \frac{e}{c^2} F_{rs}\lambda_s, \tag{27}$$

where e is the charge and λ_r its 4-velocity ($\lambda_r = dx_r/ds$). This is the *law of ponderomotive force* due to O. Heaviside and H. A. Lorentz (see WHITTAKER [1951, pp. 392–396]). Note that only the field F_{rs} is involved; the field F'_{rs} due to the particle is assumed not to influence the motion.

The equations of motion are then

$$\frac{d}{ds}\left(m\frac{dx_r}{ds}\right) = \frac{e}{c^2}F_{rs}\frac{dx_s}{ds}, \tag{28}$$

where m is the proper mass of the particle: since $X_r\lambda_r = 0$, m necessarily remains constant [cf. VI–(9)], and so we may write the equations of motion in the form

$$m\frac{d\lambda_r}{ds} = \frac{e}{c^2}F_{rs}\lambda_s. \tag{29}$$

These equations are equivalent to

$$\frac{d}{dt}(m\gamma u_\varrho) = e(E_\varrho + c^{-1}\varepsilon_{\varrho\mu\nu}u_\mu H_\nu), \quad \frac{d}{dt}(m\gamma c^2) = eE_\varrho u_\varrho, \tag{30}$$

where $u_\varrho = dx_\varrho/dt$, $\gamma = (1 - u^2/c^2)^{-\frac{1}{2}}$; the last of (30) is the equation of energy, and is a consequence of the other equations. These are the basic equations of *electron optics*, telling us how an electron moves in a given electromagnetic field. For speeds small compared with c, we may approximate by putting $\gamma = 1$ in the first equation and replacing $m\gamma c^2$ by $\frac{1}{2}mu^2$ in the second [cf. VI–(20)].[1]

The equations of motion (29) may be derived from a variational principle,

$$\delta\int(mcds - ec^{-1}\phi_n dx_n) = 0, \tag{31}$$

where ϕ_r is a 4-potential of the field F_{rs}, so that $F_{rs} = \phi_{s,r} - \phi_{r,s}$, and the end events A, B are held fixed in the variation; m is assumed constant. To prove this, we introduce any monotone parameter σ for the set of timelike world lines joining two events A, B, with $\sigma = \sigma_1$ at A and $\sigma = \sigma_2$ at B; then (31) may be written

$$\delta\int_{\sigma_1}^{\sigma_2} f\,d\sigma = 0, \quad f = mc(-x'_n x'_n)^{\frac{1}{2}} - ec^{-1}\phi_n x'_n, \quad x'_r = dx_r/d\sigma. \tag{32}$$

The extremals satisfy the Euler-Lagrange equations

$$\frac{d}{d\sigma}\frac{\partial f}{\partial x'_r} - \frac{\partial f}{\partial x_r} = 0, \tag{33}$$

[1] For some details of motions in this approximation, see SYNGE and GRIFFITH [1949, Chap. 13].

and when we calculate these expressions, choosing $\sigma = s$ on the extremal, we obtain the equations of motion (29).

The fact that (31), which contains ϕ_r, leads to (29), which contains F_{rs}, assures us that the variational principle (31) is gauge invariant, although it might not appear so at first sight. The integrand is not gauge invariant, nor is the integral, for if we add to ϕ_n a gradient $\partial\psi/\partial x_n$, we increase it by

$$-\int_A^B ec^{-1}d\psi = ec^{-1}[(\psi)_A - (\psi)_B]. \tag{34}$$

But this quantity does not change under a variation which leaves A, B fixed, and so we confirm that the variational principle is gauge invariant, although the integral is not.

The theory of de Broglie waves for a charged particle is based on the variational principle (31) (cf. SYNGE [1954a, Ch. III]), the integral which is varied being the integral of action.

§ 5. THE ELECTROMAGNETIC KEPLER PROBLEM

The electromagnetic two-body problem is the problem of finding the motion of two charged particles, each moving in the field of the other. This problem is formulated mathematically once we have decided what the field due to a charge is (we might decide on the retarded potential (23)), and laid done the equations of motion (29). However, the retardation (or advancement) of the potential makes the force on one of the particles depend, not on the position, velocity and acceleration of the other particle at the *same* time, but on these quantities at a *different* time. The result of this is a serious complication in the equations of motion, which are not differential equations as we meet them in Newtonian mechanics, but a mixture of differential equations and finite difference equations. The only hope of handling the problem is to introduce some approximation at once, such as the smallness of the relative velocity (cf. DARWIN [1920]) or the smallness of the ratio of the masses (cf. SYNGE [1940]).

If we proceed to the limit in which the mass of one of the particles is infinite, we are to think of that particle as fixed in some Galileian frame, with the other particle moving around it in its Coulomb field. This constitutes the *electromagnetic Kepler problem*, analogous to the problem of the motion of a planet about a fixed sun under Newtonian

attraction, and we shall now solve this problem, using the equations (30) with the Coulomb field (1).

Let m be the proper mass of the moving particle, e its charge and e' the charge on the fixed particle; then the equations of motion read, if the origin is taken at the fixed particle,

$$\frac{d}{dt}(m\gamma u_\varrho) = -\frac{mkx_\varrho}{r^3}, \qquad \frac{d}{dt}(m\gamma c^2) = -\frac{mk}{r^3}x_\varrho u_\varrho, \tag{35}$$

where

$$k = -\frac{ee'}{4\pi m}. \tag{36}$$

We are interested chiefly in the case where e and e' have opposite signs, so that $k > 0$. From the second of (35) we get the *integral of energy*,

$$\gamma c^2 - k/r = W, \tag{37}$$

where W is a constant (energy/mass), and from the first of (35) we have

$$x_\sigma \frac{d}{dt}(\gamma u_\varrho) - x_\varrho \frac{d}{dt}(\gamma u_\sigma) = 0, \tag{38}$$

and hence three *integrals of angular momentum*,

$$\gamma(x_\sigma u_\varrho - x_\varrho u_\sigma) = A_{\varrho\sigma}, \tag{39}$$

where $A_{\varrho\sigma}$ are constants. It follows that the orbit lies in a plane, and if we take polar coordinates (r, θ) in that plane, (39) gives

$$\gamma r^2\dot{\theta} = A, \tag{40}$$

where A is a constant (angular momentum/mass); the dot indicates d/dt.

The solution is contained in (37) and (40). We eliminate t. With $\varrho = 1/r$, we have

$$\begin{aligned} \dot{r} &= -\frac{1}{\varrho^2}\dot{\varrho} = -\frac{1}{\varrho^2}\frac{d\varrho}{d\theta}\dot{\theta} = -\frac{A}{\gamma}\frac{d\varrho}{d\theta}, \\ 1 - \frac{1}{\gamma^2} &= \frac{u^2}{c^2} = \frac{1}{c^2}(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{A^2}{c^2\gamma^2}\left[\left(\frac{d\varrho}{d\theta}\right)^2 + \varrho^2\right]. \end{aligned} \tag{41}$$

Therefore

$$\gamma^2 = 1 + \frac{A^2}{c^2}\left[\left(\frac{d\varrho}{d\theta}\right)^2 + \varrho^2\right], \tag{42}$$

and substitution in (37) gives the following differential equation for the orbit:

$$1 + \frac{A^2}{c^2}\left[\left(\frac{d\varrho}{d\theta}\right)^2 + \varrho^2\right] = \frac{1}{c^4}(W + k\varrho)^2. \tag{43}$$

To integrate, we first differentiate, obtaining

$$\frac{d^2\varrho}{d\theta^2} + \left(1 - \frac{k^2}{A^2c^2}\right)\varrho = \frac{Wk}{A^2c^2}. \tag{44}$$

We see that ϱ is a sinusoidal, quadratic or hyperbolic function of θ according as the coefficient of ϱ is positive, zero or negative.

Let us assume the initial conditions such that

$$1 - \frac{k^2}{A^2c^2} > 0; \tag{45}$$

then the solution of (44) is

$$\varrho = B\cos(p\theta + C) + \frac{Wk}{p^2A^2c^2}, \tag{46}$$

where

$$p = \left(1 - \frac{k^2}{A^2c^2}\right)^{\frac{1}{2}}. \tag{47}$$

The constant C is arbitrary, but B is a function of W and A; substitution of (46) in (43) gives

$$B^2 = \frac{c^2}{A^2p^4}\left(\frac{W^2}{c^4} - p^2\right). \tag{48}$$

Thus the equation of the orbit is

$$\frac{1}{r} = \varrho = \frac{c}{Ap^2}\left[\left(\frac{W^2}{c^4} - p^2\right)^{\frac{1}{2}}\cos(p\theta + C) + \frac{kW}{Ac^3}\right]. \tag{49}$$

To obtain a finite orbit (i.e. one which does not go to infinity, where $\varrho = 0$), it is necessary and sufficient that

$$\left(\frac{W^2}{c^4} - p^2\right)^{\frac{1}{2}} < \left|\frac{kW}{Ac^3}\right|; \tag{50}$$

this is equivalent to $W^2 < c^4$, or to

$$W < c^2, \tag{51}$$

if we assume $W + c^2 > 0$. For small velocity, we have $\gamma - 1 = \frac{1}{2}u^2/c^2$ approximately, and the condition (51) becomes, by (37),

$$\tfrac{1}{2}u^2 - \frac{k}{r} < 0, \tag{52}$$

which is the familar condition in Newtonian mechanics for an elliptical orbit — the total energy must be negative.

The apsidal angle for the orbit (49), satisfying (51), is

$$\alpha = \pi/p, \tag{53}$$

and, if p is near to unity, we may regard the orbit as a *rotating ellipse*, the advance of perihelion per revolution being

$$2\alpha - 2\pi = 2\pi(p^{-1} - 1) = \frac{\pi k^2}{A^2c^2}, \tag{54}$$

approximately.

For scattering and capture by a fixed nucleus, see Appendix C, p. 426.

§ 6. RADIATION OF ENERGY AND THIRD-ORDER EQUATIONS OF MOTION

We note that by (37) the energy of the charged particle in the Kepler problem remains constant throughout the motion. On the other hand, the Larmor formula (25) indicates a radiation of energy to infinity, and on this account it is often said that the charge must actually spiral in towards the fixed centre. This is very confusing, since the orbit (49) shows no such tendency.

In the early years of this century, the Kepler problem was fundamental in atomic theory, the Rutherford-Bohr atom being pictured as such a system. Advances in quantum theory, notably the Schrödinger wave equation, have changed this situation radically and to such an extent that anomalies presented by the Kepler problem might now be dismissed as of no physical significance. Nevertheless we should try to remove the confusion presented by a system which radiates energy and yet loses none.

In the first place, we might ask ourselves whether that which we have called "energy" in (37) should really be so called, or indeed whether there is any such thing as "total energy" in the Kepler problem, because the integral which we should use to obtain the energy of the field does not converge; this divergence arises from the necessity of

taking the moving particle as a mathematical point, any attempt to give it finite size (a small sphere, for example) being fraught with insurmountable difficulties with respect to Lorentz invariance. Now we cannot speak of the conservation of something that does not exist, and any apparent contradiction between the Larmor formula (25) and the energy integral (37) is due to the misconception that infinity is a number. One may be tempted to say: "The energy of the system is infinite, but it is conserved, and so that energy which flows away at infinity must be supplied by the mechanical energy of the moving particle and its electrostatic energy in the field of the fixed particle." But such an argument is spurious for the reason stated: infinity is not a number, and we cannot speak of the constancy of an infinite sum.

M. von Laue and M. Abraham (cf. PAULI [1920, p. 654]) modified the Lorentz equations (29) so as to make the electron supply the energy of the Larmor radiation. The modified equations, obtained later by DIRAC [1938] in a different way, read[1]

$$m\lambda_r' = \frac{e}{c^2} F_{rs}\lambda_s + \frac{1}{6\pi}\frac{e^2}{c^2}(\lambda_r'' - \lambda_r \lambda_n' \lambda_n'), \tag{55}$$

where $\lambda_r = dx_r/ds$, so that

$$\lambda_r \lambda_r = -1; \tag{56}$$

the prime indicates d/ds. These equations are quite different from the usual equations of mechanics, because they involve the rate of change of acceleration ($\lambda_r'' = d^3x_r/ds^3$), and so a motion is not determined by initial position and velocity; initial acceleration must also be given.

Let us apply these equations to the Kepler problem of § 5. Instead of the last of (35) we now have

$$m\lambda_4' = \frac{e}{c^2} F_{4\sigma}\lambda_\sigma + \frac{1}{6\pi}\frac{e^2}{c^2}(\lambda_4'' - \lambda_4 \lambda_n' \lambda_n'), \tag{57}$$

and, with $ds = cdt/\gamma$, this can also be written

$$\frac{d}{dt}\left(m\gamma c^2 - \frac{mk}{r}\right) = \frac{1}{6\pi}\frac{e^2}{c}\left(\gamma\frac{d^2\gamma}{dt^2} + \frac{d\gamma}{dt}\frac{d\gamma}{dt} - \gamma^2\frac{d\lambda_n}{dt}\frac{d\lambda_n}{dt}\right), \tag{58}$$

since $\lambda_\varrho = \gamma u_\varrho/c$, $\lambda_4 = i\gamma$. For slow motion, $\gamma = 1 + \frac{1}{2}u^2/c^2$ approxi-

[1] To convert to Gaussian units, for $1/6\pi$ read $2/3$.

mately, and

$$\frac{d\lambda_\varrho}{dt} = \frac{1}{c} f_\varrho, \qquad \frac{d\lambda_4}{dt} = i\frac{d\gamma}{dt} = i\frac{u}{c^2}\frac{du}{dt}, \tag{59}$$

so that the principal part of (58) gives

$$\frac{d}{dt}\left(m\gamma c^2 - \frac{mk}{r}\right) = \frac{1}{6\pi}\frac{e^2}{c^3}\frac{d}{dt}\left(u\frac{du}{dt}\right) - \frac{1}{6\pi}\frac{e^2}{c^3}f^2, \tag{60}$$

f_ϱ being the acceleration (du_ϱ/dt), and f its magnitude.

The last term in (60) is the Larmor expression (25) with sign reversed (we replace f' by f in the present notation). Thus if we neglect the first term on the right of (60) (as we may in the case of a nearly circular orbit), we see that *the equations of motion* (55) *give a rate of decrease of the energy of the moving particle equal to the Larmor rate of radiation of energy*. For that reason we may call the last term in (55) the *radiation 4-force*.

In (55) and (56) we have five equations for the four quantities λ_r, m being assumed constant. However these five equations are identically related, as we see on multiplying (55) by λ_r and using the two following identities implied by (56):

$$\lambda_r\lambda_r' = 0, \quad \lambda_r\lambda_r'' + \lambda_r'\lambda_r' = 0. \tag{61}$$

Application of (55) to the case where the field F_{rs} is absent leads to a rather strange result. We have to solve

$$\lambda_r'' - \lambda_r\lambda_n'\lambda_n' = \mu\lambda_r', \quad \mu = 6\pi mc^2/e^2, \tag{62}$$

and from these equations we obtain

$$\lambda_r''\lambda_r' = \mu\lambda_r'\lambda_r', \quad \lambda_r''\lambda_s - \lambda_s''\lambda_r = \mu(\lambda_r'\lambda_s - \lambda_s'\lambda_r), \tag{63}$$

and hence the integrals

$$\lambda_r'\lambda_r' = B^2\exp(2\mu s), \tag{64}$$

$$\lambda_r'\lambda_s - \lambda_s'\lambda_r = K_{rs}\exp(\mu s), \tag{65}$$

where B and K_{rs} $(= -K_{sr})$ are constants. Choose the frame of reference so that for $s = 0$ we have

$$\begin{aligned} &\lambda_1 = \lambda_2 = \lambda_3 = 0, \quad \lambda_4 = i, \\ &\lambda_1' = B, \quad \lambda_2' = \lambda_3' = \lambda_4' = 0. \end{aligned} \tag{66}$$

Then $K_{14} = -K_{41} = iB$, the other components of K_{rs} vanishing, and so by (65) we have throughout the motion

$$\begin{aligned} &\lambda_2'\lambda_3 - \lambda_3'\lambda_2 = 0, \quad \lambda_1'\lambda_4 - \lambda_4'\lambda_1 = iB\exp(\mu s) \\ &\lambda_3'\lambda_1 - \lambda_1'\lambda_3 = 0, \quad \lambda_2'\lambda_4 - \lambda_4'\lambda_2 = 0, \\ &\lambda_1'\lambda_2 - \lambda_2'\lambda_1 = 0, \quad \lambda_3'\lambda_4 - \lambda_4'\lambda_3 = 0. \end{aligned} \tag{67}$$

It follows that $\lambda_2 = \lambda_3 = \lambda_2' = \lambda_3' = 0$, and (64) and (67) give the two equations

$$\lambda_1'^2 + \lambda_4'^2 = B^2\exp(2\mu s), \quad \lambda_1'\lambda_4 - \lambda_4'\lambda_1 = iB\exp(\mu s), \tag{68}$$

which we solve by putting, as we may by (56),

$$\lambda_1 = \sinh\theta, \quad \lambda_4 = i\cosh\theta; \tag{69}$$

we find

$$\theta = \frac{B}{\mu}[\exp(\mu s) - 1], \tag{70}$$

and thus *the motion of a charged particle, according to the equations* (55) *and in the absence of field, is given by*

$$\begin{aligned} \lambda_1 &= \sinh\left\{\frac{B}{\mu}[\exp(\mu s) - 1]\right\}, \\ \lambda_2 &= \lambda_3 = 0, \\ \lambda_4 &= i\cosh\left\{\frac{B}{\mu}[\exp(\mu s) - 1]\right\}, \end{aligned} \tag{71}$$

from which equations we obtain $x_r(s)$ by quadratures. We see that unless B is zero, λ_1 and λ_4/i tend to infinity with s, which means that *the speed of the particle tends ultimately to c.*

The fact is that the radiation term in (55) causes the curvature $\varkappa$ of the world line ($\varkappa^2 = \lambda_n'\lambda_n'$ by definition) to increase as we pass along the world line into the future, and that is what we want in the Kepler problem, for the smaller the energy, the closer the orbit to the centre, and the greater $\varkappa$. But this increasing curvature leads to the "run-away" motion (71) in the absence of field.

It is however easy to avoid the anomalous behaviour described above, and at the same time secure the energy balance (60) for slow motions, by altering the radiation 4-force in (55) so that the equations

of motion are

$$m\lambda_r' = \frac{e}{c^2} F_{rs}\lambda_s + \frac{1}{6\pi}\frac{e^2}{c^2}(\lambda_r'' - \lambda_r\lambda_n'\lambda_n')\Phi, \tag{72}$$

where

$$\Phi = \frac{e^2}{m^2c^4}\frac{F_{mp}\lambda_p F_{mq}\lambda_q}{\lambda_n'\lambda_n'}; \tag{73}$$

equivalently

$$mc^2\lambda_r'\lambda_n'\lambda_n' = eF_{rs}\lambda_s\lambda_n'\lambda_n' + \frac{1}{6\pi}\frac{e^4}{m^2c^4}(\lambda_r'' - \lambda_r\lambda_n'\lambda_n')F_{mp}\lambda_p F_{mq}\lambda_q. \tag{74}$$

If $F_{rs} = 0$, the only solution is $\lambda_r = \text{const}$. If $F_{rs} \neq 0$ and the motion is slow, the last term in (72) is small and the motion is given approximately by the first term; this makes $\Phi = 1$. Thus we may replace Φ by unity in (72), and so obtain the energy balance (60) just as we did for (55).

§ 7. MAXWELL'S EQUATIONS WITH CURRENT

Consider an incoherent stream of material particles, as described in VIII–§ 10. We shall suppose for simplicity that all the particles have the same proper mass m and the same constant charge e, although that is not essential; it is essential that they form a stream, specified at each event by a timelike unit vector λ_r, and do not wander off in all directions like the molecules of a gas.

We have then three invariants,

$$\begin{aligned} N &= \text{proper numerical density,} \\ \mu &= \text{proper density of proper mass,} \\ \varrho &= \text{proper density of charge,} \end{aligned} \tag{75}$$

with the following meanings: an observer who takes the local λ_r for time-axis (thus reducing the particles to rest) finds $N dS$ particles in a box of volume dS, with a total proper mass μdS and a total charge ϱdS. We have

$$N = \mu/m = \varrho/e. \tag{76}$$

In any Galileian frame, the fluxes of number, of proper mass, and of charge, across a target dS with unit normal n_r, are respectively

$$N\varepsilon(n)\lambda_r n_r dS, \quad \mu\varepsilon(n)\lambda_r n_r dS, \quad \varrho\varepsilon(n)\lambda_r n_r dS, \tag{77}$$

these quantities being in fact the number of world lines cutting dS, and that number multiplied by m and e respectively.

If, as we shall suppose, no particles are created or destroyed, and the mass and charge of each particle remain constant, then for any closed 3-space we have

$$\int N\varepsilon(n)\lambda_r n_r dS = 0, \tag{78}$$

and two similar equations, and hence, by Green's theorem, *the equations of conservation of number, mass, and charge*

$$\frac{\partial}{\partial x_r}(N\lambda_r) = 0, \quad \frac{\partial}{\partial x_r}(\mu\lambda_r) = 0, \quad \frac{\partial}{\partial x_r}(\varrho\lambda_r) = 0. \tag{79}$$

Of these it is only the last that interests us here. We define the *4-current* J_r by

$$J_r = \varrho\lambda_r, \tag{80}$$

and note the *conservation equation*

$$J_{r,r} = 0, \tag{81}$$

the comma denoting partial derivation.

The 3-vector cJ_σ may be called the *relative 3-current* and J_4/i ($= \varrho^*$ say) the *relative electrical density* or *relative charge density*; $\varrho^* dS$ is the charge found by an observer in a box of volume dS.

We have built up J_r from a picture of particles. But we may very well abandon that picture, and accept J_r as a physical 4-vector in its own right, satisfying the conservation equation (81) and with its four components named as above.

We now assume *Maxwell's equations with electric current*[1] in the form

$$F_{rs,s} = J_r, \quad F^*_{rs,s} = 0, \tag{82}$$

these equations reducing to the vacuum equations IX–(10) when $J_r = 0$. Note that the conservation equation (81) is a consequence of the first of (82). These equations may also be written [cf. IX–(1)] in the pre-relativistic form

$$\begin{gathered} \frac{1}{c}\frac{\partial E_\varrho}{\partial t} + J_\varrho = \varepsilon_{\varrho\mu\nu}\frac{\partial H_\nu}{\partial x_\mu}, \quad \frac{\partial E_\sigma}{\partial x_\sigma} = J_4/i = \varrho^*, \\ -\frac{1}{c}\frac{\partial H_\varrho}{\partial t} = \varepsilon_{\varrho\mu\nu}\frac{\partial E_\nu}{\partial x}, \quad \frac{\partial H_\sigma}{\partial x_\sigma} = 0. \end{gathered} \tag{83}$$

[1] Zero remains in the second equation because there is no magnetic current.

The Maxwellian equations IX–(10) for a field in vacuo form a determinate mathematical system (six independent equations for six independent unknowns), but the new equations (82) are not adequate to determine the ten quantities F_{rs}, J_r; other equations must be added, and to these we shall come later.

The second equation in (82) is equivalent to [cf. IX–(14)]

$$F_{mn,r} + F_{nr,m} + F_{rm,n} = 0, \tag{84}$$

and the existence of a 4-potential depends on this equation only. Hence, as in IX–(151), we may write

$$F_{rs} = \phi_{s,r} - \phi_{r,s}. \tag{85}$$

Substituting this in the first of (82), we see that the equations IX–(178) now become

$$\boxed{\phi_{r,r} = 0, \quad \Box \phi_r = -J_r.} \tag{86}$$

These equations may be regarded as equivalent to (82), although we are to remember that even if F_{rs} and J_r are known, ϕ_r is undetermined to within the gradient of an arbitrary wave function added to it.

§ 8. EXPLICIT FORMULA FOR THE 4-POTENTIAL

Suppose that the 4-current J_r is given. Then ϕ_r is to satisfy (86), and the most general solution is the sum of any particular solution of (86) and the general solution of the vacuum equations, i.e. (86) with $J_r = 0$. We shall prove that the expression (104) is a particular solution of (86). There are two steps, the first easy and the second harder. First, we show that the vector field (91) has zero divergence; secondly, we establish the key-formula (101), which is the space-time analogue of Poisson's equation in potential theory.

In order to express the solution neatly in Lorentz-invariant form, we shall make use of the absolute 2-content of an element of the null cone, as explained in Appendix D (p. 430). Let $P(x)$ be the vertex of a null cone $(N' + N'')$, N' being the past sheet and N'' the future sheet. Then the absolute 2-contents of 3-cells at $Q'(x')$ on N' and at $Q''(x'')$ on N'' are respectively

$$d\omega' = \frac{dS'}{n_r(x'_r - x_r)}, \quad d\omega'' = \frac{dS''}{n_r(x_r - x''_r)}, \tag{87}$$

where n_r is a unit vector pointing into the future and dS', dS'' the projections of the 3-cells on a 3-flat orthogonal to n_r; the denominators in (87) are positive. The quantities $d\omega'$, $d\omega''$ are Lorentz invariant and also independent of the choice of n_r. If we choose n_r along the time-axis, we get

$$d\omega' = \frac{dS'}{i(x_4' - x_4)} = \frac{dS'}{c(t - t')} = \frac{dS'}{r'},$$
$$d\omega'' = \frac{dS''}{i(x_4 - x_4'')} = \frac{dS''}{c(t'' - t)} = \frac{dS''}{r''}, \tag{88}$$

where r', r'' are the spatial distances of Q', Q'' from P; now dS', dS'' are 3-volumes in the observer's space,

$$dS' = dx_1' dx_2' dx_3', \quad dS'' = dx_1'' dx_2'' dx_3''. \tag{89}$$

Let J_m be a given vector field, satisfying the conservation equation

$$J_{m,m} = 0. \tag{90}$$

We shall suppose that J_m and its first derivatives vanish at infinity on the null cone $(N' + N'')$ as fast as $r^{-1-\alpha}$ $(\alpha > 0)$, and to simplify the work we shall assume that the second derivatives exist and are continuous.

Consider the integral

$$I_m(x, x_4) = \int_{M'} J_m(x', x_4') d\omega' \tag{91}$$

where x stands for (x_1, x_2, x_3), the space coordinates of P, and x' for (x_1', x_2', x_3'), the space coordinates of an event Q' on N'; the integral is taken over the portion M' of the null sheet N' down to a 3-flat Π with equation $t = \text{const}$. (Fig. 3).

If we displace the null cone as a whole parallel to one of the space-axes, we can put the 3-cells of M' in the old and new positions into correspondence, so that the elements $d\omega'$ are unchanged, and we get

$$\frac{\partial I_m}{\partial x_\sigma} = \int_{M'} \frac{\partial J_m'}{\partial x_\sigma'} d\omega', \tag{92}$$

where $J_m' = J_m(x', x_4')$. On the other hand, if we displace the null cone parallel to the time-axis and establish correspondence between elements by

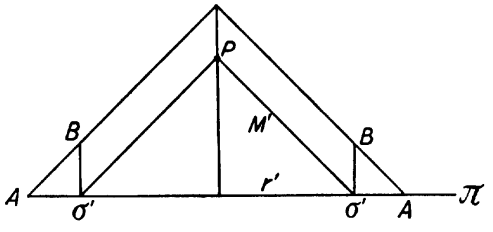


Fig. 3 – Null cone displaced parallel to the time-axis

lines drawn parallel to that axis, we see that the new M' is not covered by this correspondence, because a new piece of cone has been generated at Π (AB in Fig. 3). Hence we get

$$\frac{\partial I_m}{\partial x_4} = \int_{M'} \frac{\partial J'_m}{\partial x'_4} d\omega' - i \int_{\sigma'} J'_m \frac{d\sigma'}{r'}, \tag{93}$$

the last integral being taken over the intersection σ' of M' and Π; σ' is a sphere in the observer's space. The form of the last term in (93) is obtained by observing that, from (88), the elementary 2-content added at AB is $d\omega' = dS'/r'$, where dS' is the elementary 3-volume of a spherical shell ($dS' = dr'd\sigma'$); further $dx_4 = icdt = idr'$.

By reason of the condition imposed on J_m at infinity, the last term disappears from (93) when we remove Π to the infinite past, thus making $M' = N'$. Hence we have

$$I_m = \int_{N'} J'_m d\omega', \qquad \frac{\partial I_m}{\partial x_s} = \int_{N'} \frac{\partial J'_m}{\partial x'_s} d\omega', \tag{94}$$

or, putting $s = m$ and using (90),

$$\frac{\partial I_m}{\partial x_m} = \int_{N'} \frac{\partial J'_m}{\partial x'_m} d\omega' = 0. \tag{95}$$

This completes the first step in solving (86); we have a vector I_m with zero divergence.

Returning to the cone M' cut off by Π, we can differentiate (92) and (93), obtaining in the latter case an integral over M' and integrals over σ'. But the integrals over σ' disappear in the limit $M' = N'$, and so we can in fact differentiate (94) by shifting the null cone without paying attention to contributions at infinity on N'. Thus we get

$$\Box I_m = \int_{N'} \Box' J'_m d\omega', \tag{96}$$

where $\Box$, $\Box'$ are the d'Alembertian operators for x_r, x'_r respectively.

Let us write, with x'_r on N',

$$J'_m = J_m(x', x'_4) = J_m(x', x_4 - ir') = f_m(x', x, x_4), \tag{97}$$

using the fact that

$$x'_4 = x_4 - ir', \quad r'^2 = (x'_\varrho - x_\varrho)(x'_\varrho - x_\varrho). \tag{98}$$

In (97) x stands for (x_1, x_2, x_3) and x' for (x'_1, x'_2, x'_3). By calculating the partial derivatives of f_m with respect to x'_ϱ and x_4, we find

$$\Box' J'_m = \Delta' f_m + 2i \frac{\partial}{\partial x_4} \left(\frac{\partial f_m}{\partial x'_\varrho} \frac{\partial r'}{\partial x'_\varrho} + \frac{f_m}{r'} \right), \tag{99}$$

where $\Delta' = \partial^2/\partial x'_\varrho \partial x'_\varrho$. We now multiply (99) by $d\omega' = dS'/r'$ and integrate over N'. It is easy to see that the last term in (99) contributes nothing, and by a well known result in potential theory

$$\int_{N'} \Delta' f_m \, d\omega' = \int_{N'} \Delta' f_m \frac{dS'}{r'} = -4\pi f_m = -4\pi J_m, \tag{100}$$

evaluated at P. Thus, by (96),

$$\Box \int_{N'} J'_m \, d\omega' = \int_{N'} \Box' J'_m \, d\omega' = -4\pi J_m. \tag{101}$$

Details of the above calculations are given in Appendix E, p. 432.

From (91) on, we have worked with the past sheet N' of the null cone. But the same results follow for N'', and so, as in (95) and (101), we have the following formulae:

$$\begin{gathered} \frac{\partial}{\partial x_r} \int_{N'} J'_r \, d\omega' = 0, \quad \frac{\partial}{\partial x_r} \int_{N''} J''_r \, d\omega'' = 0, \\ \Box \int_{N'} J'_r \, d\omega' = -4\pi J_r, \quad \Box \int_{N''} J''_r \, d\omega'' = -4\pi J_r. \end{gathered} \tag{102}$$

It is then clear that the following theorem is true: *The equations*

$$\phi_{r,r} = 0, \quad \Box \phi_r = -J_r \tag{103}$$

are satisfied by

$$\phi_r = \frac{1}{4\pi} \left(A' \int_{N'} J'_r \, d\omega' + A'' \int_{N''} J''_r \, d\omega'' \right), \tag{104}$$

the integrations being over the past and future sheets of the null cone with vertex at the event where ϕ_r is to be evaluated, and A', A'' being any constants satisfying $A' + A'' = 1$, provided that J_r satisfies the conservation equation (90) *and the other conditions mentioned there.* We note that the term with factor A' is a retarded potential, and that with A'' an advanced potential

As commonly used, the formula (104), or its equivalent in different notation, is given with $A' = 1$, $A'' = 0$ (retarded potential only). It is usually approached in a different way, viz. as a consequence of Kirchhoff's formula (cf. BATEMAN [1944, p. 185], BAKER and COPSON

[1950, p. 38]), but it is not quite the same thing. Kirchhoff's formula expresses any solution of $\Box\phi_r = -J_r$ in terms of a volume integral (our $\int_{M'} J'_r d\omega'$) and a surface integral involving ϕ_r. If this surface integral vanishes, then Kirchhoff's formula gives (104) (with $A' = 1$, $A'' = 0$) as the unique solution of the partial differential equation and certain conditions at infinity. As we have approached the matter, there is nothing unique about (104); it is a solution of (103), that is all, the most general solution being obtained by adding to (104) a 4-potential of *any* Maxwellian field in vacuo, such as a system of plane waves, or the field discussed in IX-§ 14.

To link up this continuous theory with the discrete, let us now suppose that J_r vanishes outside a thin world-tube surrounding a world line L with unit tangent vector λ_r, not in general constant. We can then write (104) as

$$\phi_r = \frac{1}{4\pi}(A' J'_r d\omega' + A'' J''_r d\omega''). \tag{105}$$

In (87) let us take $n_r = \lambda'_r$ at Q' and $n_r = \lambda''_r$ at Q'' (these being the events where L cuts N' and N''); then

$$d\omega' = dS'/w', \quad d\omega'' = -dS''/w'' \tag{106}$$

in the notation of (15), dS' and dS'' being the normal sections of the tube. By (80) we have

$$J'_r = \varrho'\lambda'_r, \quad J''_r = \varrho''\lambda''_r, \tag{107}$$

where ϱ is the proper density of charge. The total charge (conserved by (79) or (90)) is

$$e = \varrho' dS' = \varrho'' dS''. \tag{108}$$

Substitution from (106)–(108) in (105) gives

$$\phi_r = \frac{e}{4\pi}\left(\frac{A'\lambda'_r}{w'} - \frac{A''\lambda''_r}{w''}\right), \quad (A' + A'' = 1). \tag{109}$$

Thus, by application of the continuous theory with Maxwell's equations in the form (82), *we have recovered the discrete formula* (15) *for the field due to a moving charge.*

We are therefore led to accept (104) as the 4-potential of the field due to a continuous distribution of current, putting $A' = 1$, $A'' = 0$ if we wish to have a definite result and at the same time avoid the awkward question of causality connected with the advanced potential.

§ 9. THE ENERGY TENSOR OF A FIELD WITH CURRENT

We have now to discuss the question of associating additional equations with (82) in order to make F_{rs} and J_r mathematically determinate, in the sense that the number of independent equations is equal to the number of independent unknowns. (Of course complete determinacy requires the specification of boundary conditions as well, but with this we are not concerned.) We shall describe two ways of doing this.

The first way is to write down the energy tensor

$$T_{rs} = \mu\lambda_r\lambda_s + c^{-2}(F_{mr}F_{ms} - \tfrac{1}{4}\delta_{rs}F_{mn}F_{mn}), \tag{110}$$

this being the sum of the energy tensor of an incoherent stream of particles of proper mass density μ [VIII–(139)] and the energy tensor of an electromagnetic field in vacuo [IX–(40)]. Differentiating (110) and using the conservation of mass (79) and the first of the Maxwell equations (82), we get

$$T_{rs,s} = \mu\lambda_{r,s}\lambda_s + c^{-2}(F_{mr}J_m + F_{mr,s}F_{ms} - \tfrac{1}{2}F_{mn}F_{mn,r}). \tag{111}$$

The last terms cancel by reason of the other Maxwell equations (84), and so we have

$$T_{rs,s} = \mu\lambda_{r,s}\lambda_s - c^{-2}F_{rs}J_s. \tag{112}$$

Let us now accept the conservation equation $T_{rs,s} = 0$. Changing the first term on the right of (112) into an equivalent form, we have

$$\mu\frac{d\lambda_r}{ds} = c^{-2}F_{rs}J_s = c^{-2}\varrho F_{rs}\lambda_s. \tag{113}$$

But $\mu/\varrho = m/e$ (the mass/charge ratio for a particle), and so *for each particle of the discrete system, which we smeared statistically into a continuum, we have the equations of motion* (29)

$$m\frac{d\lambda_r}{ds} = \frac{e}{c^2}F_{rs}\lambda_s. \tag{114}$$

As for determinacy, we have now the following equations:

$$\begin{gathered} F_{rs,s} = \varrho\lambda_r, \quad F^*_{rs,s} = 0, \\ m\lambda_{r,s}\lambda_s = ec^{-2}F_{rs}\lambda_s, \quad \lambda_r\lambda_r = -1. \end{gathered} \tag{115}$$

We do not include the conservation of charge $[(\varrho\lambda_r)_{,r} = 0]$ because it is implied by the first of (115). The third of (115) gives $0 = 0$ when multi-

plied by λ_r, on account of the fourth, and so we count $4+3+3+1=11$ independent equations for the 11 unknowns contained in F_{rs} (6), λ_r (4) and ϱ(1). *Thus* (115) *is a determinate system of equations.*

This approach is satisfactory if we regard the theory as a macroscopic one, applied statistically to a large number of charged particles. But we can adopt a different view, better suited to possible application to the ultimate structure of matter, by writing down Maxwell's equations

$$F_{rs,s} = J_r, \quad F^*_{rs,s} = 0 \tag{116}$$

and regarding the mechanical properties of the system (mass and momentum) as contained in the vacuum energy tensor [IX–(40)]

$$T_{rs} = c^{-2}(F_{mr}F_{ms} - \tfrac{1}{4}\delta_{rs}F_{mn}F_{mn}). \tag{117}$$

By a simple calculation as in (111) we find

$$T_{rs,s} = -c^{-2}F_{rs}J_s, \tag{118}$$

and so, if we combine the Maxwell equations (116) with the conservation equations $T_{rs,s} = 0$, we get *the following system of equations as a basis for electromagnetic theory with current*:

$$F_{rs,s} = J_r, \quad F^*_{rs,s} = 0, \quad F_{rs}J_s = 0. \tag{119}$$

This is a determinate system, for we count $4 + 3 + 3 = 10$ independent equations for the 10 unknowns in F_{rs} (6) and J_r (4).

That (119) forms a consistent system of equations is shown by the existence of the following special solution:

$$\begin{aligned} &E_1 = iF_{14} = 0, \quad E_2 = iF_{24} = x_2 f(\xi), \quad E_3 = iF_{34} = x_3 f(\xi), \\ &H_1 = F_{23} = 0, \quad H_2 = F_{31} = -x_3 f(\xi), \quad H_3 = F_{12} = x_2 f(\xi), \\ &J_1 = -iJ_4 = 2f(\xi) + \xi f'(\xi), \quad J_2 = J_3 = 0, \end{aligned} \tag{120}$$

where $\xi = (x_2^2 + x_3^2)^{\frac{1}{2}}$ and the function f is arbitrary. Note that J_r here is a null vector, although it need not be so in general. Unfortunately the system (120) is independent of x_1, and so it cannot be used as an illustrative model in connection with what follows below, since the integrals involved do not converge.

The system (119) may be regarded as a particular case of (115) corresponding to $m = 0$. Note that in those regions where J_r does not vanish, the last of (119) implies $\det F_{rs} = 0$, equivalent to

$(FF^*) = 0$, so that the electric and magnetic 3-vectors are perpendicular to one another in such regions, as we see is the case in (120).

We shall now show that (119) implies equations for the rate of change of 4-momentum of a system analogous to the equations of motion (114).

Let F'_{rs} be some Maxwellian field in vacuo, so that

$$F'_{rs,s} = 0, \quad F'^*_{rs,s} = 0; \tag{121}$$

we shall call this the external field. Let (F_{rs}, J_r) be a field and 4-current such that we get a field satisfying (119) when we superimpose them on F'_{rs}. Thus we have

$$F_{rs,s} + F'_{rs,s} = J_r, \quad F^*_{rs,s} + F'^*_{rs,s} = 0, \quad (F_{rs} + F'_{rs})J_s = 0, \tag{122}$$

and hence by (121)

$$F_{rs,s} = J_r, \quad F^*_{rs,s} = 0, \tag{123}$$

so that (F_{rs}, J_r) is a Maxwellian field and current, but with $F_{rs}J_s \neq 0$ in general.

Consider now the mechanical properties of the system (F_{rs}, J_r) as contained in the 4-momentum [VIII–(247)]

$$M_r = -i\int T_{r4}dS \tag{124}$$

where

$$c^2T_{r4} = F_{\mu r}F_{\mu 4} - \tfrac{1}{4}\delta_{r4}F_{mn}F_{mn}, \tag{125}$$

the integration being taken over the 3-flat $t = \text{const.}$, dS being the element of volume. We suppose that the field goes to zero fast enough at infinity so that the integral (124) exists, and so that the operations required below can be carried out. Differentiation of (124) gives

$$\frac{dM_r}{dx_4} = -i\int T_{r4,4}dS, \tag{126}$$

or, by (118),

$$\frac{dM_r}{dx_4} = i\int T_{r\sigma,\sigma}dS + ic^{-2}\int F_{rs}J_s dS. \tag{127}$$

By Green's theorem in 3-space we can replace the first integral by an integral over an infinite sphere, and this vanishes if the field goes to zero fast enough. We shall suppose that it does; then, using the last of

(122), we can write (127) in the form

$$\frac{dM_r}{cdt} = c^{-2}\int F'_{rs}J_s dS. \tag{128}$$

This expresses the rate of change of 4-momentum of the system (F_{rs}, J_r) in terms of the *external* field and the current, and may be regarded as an integral analogue of the differential formula (113), which is the same as (114).

If we regard the system (F_{rs}, J_r) as a model of an electromagnetic particle smeared out over space-time, then (128) tells us how the 4-momentum of this particle reacts to the imposition of an external field F'_{rs}; the interesting thing about (128) is that the field of the particle itself (F_{rs}) does not appear.

§ 10. MAXWELL'S EQUATIONS DERIVED FROM A VARIATIONAL PRINCIPLE

In modern mathematical notation a system of differential equations derived from a variational principle is often as easy to write down as the variational equation itself, and this robs variational principles of some of their old charm. For example, the explicit equations of motion (29) are at least as simple as the variational equation (31). Nevertheless, interest in variational principles remains strong, partly because the integral which we vary may itself possess an important meaning (energy or action, for example), and partly because, by changing the integrand in a known variational principle, we can generalise known physical laws into new forms.

We shall now derive Maxwell's equations from a variational principle. Sometimes this is done by assuming half the equations and deriving the other half, but we shall follow a more interesting method due to B. FINZI [1952]. The key to this method is the following fact: *Any skew-symmetric tensor F_{rs} can be split in the form*

$$F_{rs} = H_{rs} + K_{rs}, \tag{129}$$

where H_{rs} and K_{rs} are skew-symmetric tensors satisfying

$$H_{rs,s} = 0, \quad K^*_{rs,s} = 0, \tag{130}$$

provided that the vector $F_{rs,s}$ satisfies certain conditions at infinity, viz. the conditions imposed on J_r following (90).

To prove this, we take any 4-vector ϕ_r satisfying $\phi_{r,r} = 0$, and

define K_{rs} by

$$K_{rs} = \phi_{s,r} - \phi_{r,s}. \tag{131}$$

Then the second of (130) is satisfied. Define H_{rs} by

$$H_{rs} \doteq F_{rs} - K_{rs}. \tag{132}$$

Then

$$H_{rs,s} = F_{rs,s} + \Box\phi_r, \tag{133}$$

and so the first of (130) is satisfied provided ϕ_r satisfies

$$\phi_{r,r} = 0, \quad \Box\phi_r = -F_{rs,s}. \tag{134}$$

By the skew-symmetry of F_{rs} we have $(F_{rs,s})_{,r} = 0$, and so the problem of solving (134) is the same as that of solving (86); but we saw how to do this in § 8, and thus the validity of splitting F_{rs} as in (129) is established.

The formulae (129), (130) enable us to express any skew-symmetric tensor in terms of *two* 4-potentials. For $H_{rs,s} = 0$ implies

$$H^*_{mn,r} + H^*_{nr,m} + H^*_{rm,n} = 0, \tag{135}$$

and so [cf. IX–§ 8] there exists a 4-vector ψ_r such that

$$H^*_{rs} = \psi_{s,r} - \psi_{r,s}, \quad H_{rs} = -\tfrac{1}{2}i\varepsilon_{rsmn}H^*_{mn} = i\varepsilon_{rsmn}\psi_{m,n}. \tag{136}$$

This means that, subject only to certain conditions as to vanishing at infinity, *we can write any skew-symmetric tensor in the form*

$$F_{rs} = \phi_{s,r} - \phi_{r,s} + i\varepsilon_{rsmn}\psi_{m,n}. \tag{137}$$

We come now to the variational principle. Consider the integral

$$I = \int (\tfrac{1}{4}F_{rs}F_{rs} - J_r\phi_r)d\tau, \tag{138}$$

integrated over a domain R of space-time with boundary B; F_{rs} means the expression (137), and J_r, ϕ_r and ψ_r are arbitrary 4-vectors. We regard J_r as assigned, and seek those 4-vectors ϕ_r, ψ_r which give a stationary value to I when ϕ_r, ψ_r are assigned on B.

Variation gives

$$\begin{aligned}\delta I &= \int (\tfrac{1}{2}F_{rs}\delta F_{rs} - J_r\delta\phi_r)d\tau \\ &= \int (-F_{rs}\delta\phi_{r,s} + \tfrac{1}{2}iF_{rs}\varepsilon_{rsmn}\delta\psi_{m,n} - J_r\delta\phi_r)d\tau \\ &= \int (-F_{rs}\delta\phi_{r,s} + F^*_{rs}\delta\psi_{r,s} - J_r\delta\phi_r)d\tau. \end{aligned} \tag{139}$$

Integrating by parts with Green's theorem and rejecting the integrals

on B because $\delta\phi_r = \delta\psi_r = 0$ there, we get

$$\delta I = \int (F_{rs,s}\delta\phi_r - F^*_{rs,s}\delta\psi_r - J_r\delta\phi_r)d\tau. \tag{140}$$

Thus the variational equation $\delta I = 0$, *for arbitrary variations* $\delta\phi_r$, $\delta\psi_r$ *in* R, *implies Maxwell's equations*

$$F_{rs,s} = J_r, \quad F^*_{rs,s} = 0. \tag{141}$$

By taking in (138) other invariant quadratic combinations of F_{rs}, J_r, ϕ_r, ψ_r, FINZI [1952] has obtained modified forms of Maxwell's equations.[1]

§ 11. MAXWELL'S EQUATIONS IN MOVING MATTER

We consider now electromagnetic phenomena in a solid body moving with constant velocity, so that the world lines of its particles form a congruence of parallel straight lines. Let λ_r be the unit vector drawn in the common direction of these lines, i.e. the 4-velocity of the body. For a body at rest, pre-relativistic physics provides a satisfactory system of equations, and our task is to write down a set of tensorial equations which reduce to the accepted equations in the rest frame of the body, i.e. when we put $\lambda_\varrho = 0$, $\lambda_4 = i$.

We introduce two skew-symmetric tensors, F_{rs} and H_{rs}, and a 4-vector J_r, with the following descriptions to link them with physical experience:

$$\begin{aligned}
&\text{electric intensity} = E_\varrho = iF_{\varrho 4},\\
&\text{magnetic induction} = B_\varrho = -iF^*_{\varrho 4},\\
&\text{electric displacement} = D_\varrho = iH_{\varrho 4},\\
&\text{magnetic intensity} = H_\varrho = -iH^*_{\varrho 4},\\
&\text{electric current} = cJ_\varrho,\\
&\text{electric density} = J_4/i.
\end{aligned} \tag{142}$$

Note the analogy with the vacuum formulae IX–(6), (7); in vacuo we have $H_{rs} = F_{rs}$, $D_\varrho = E_\varrho$, $B_\varrho = H_\varrho$.

On the above quantities we impose the partial differential equations

$$H_{rs,s} = J_r, \quad F^*_{rs,s} = 0, \tag{143}$$

[1] See also UDESCHINI [1952] and STORCHI [1953]. Electromagnetic field equations obtained in this way are linear; for non-linear field equations, see BORN and INFELD [1934].

analogous to (82). When we substitute from (142), we get equations as in (83), but with E changed to D in the first line and H changed to B in the second; these are the familiar *Maxwellian equations* for a body at rest. By the scheme (142) we have made them valid for a body in uniform motion.

With (143) we associate *constitutive equations*, dependent on the nature of the body. For a homogeneous isotropic body we take

$$\begin{gathered} H_{rs}\lambda_s = \varepsilon F_{rs}\lambda_s, \quad F^*_{rs}\lambda_s = \mu H^*_{rs}\lambda_s, \\ J_r + \lambda_r J_s \lambda_s = c^{-1}\sigma F_{rs}\lambda_s. \end{gathered} \tag{144}$$

Here ε is the dielectric constant, μ the magnetic permeability, and σ the conductivity, all assumed to be constants, invariant with respect to Lorentz transformations. In the rest frame of the body these reduce to the familiar equations

$$D_\varrho = \varepsilon E_\varrho, \quad B_\varrho = \mu H_\varrho, \quad cJ_\varrho = \sigma E_\varrho. \tag{145}$$

We have then in (143) *and* (144) *a set of equations suitable for the description of electromagnetic phenomena in a body in uniform motion.* We make no attempt here to discuss the far more difficult problem of electromagnetism in a body in accelerated motion, e.g. in rotation.

The system (143), (144) is mathematically determinate; we count $4 + 3 + 3 + 3 + 3 = 16$ independent equations for the 16 independent unknowns contained in F_{rs}, H_{rs}, J_r. We are therefore not obliged, as we were in § 9, to introduce an energy tensor in order to secure determinacy. If we are to have an energy tensor T_{rs} satisfying the conservation equations $T_{rs,s} = 0$, then either these equations must be deducible from (143), (144), or they must involve new unknowns to compensate for the introduction of new equations. The second alternative is forced on us by the fact that an electric current generates heat, a form of energy not taken care of by the electromagnetic variables. Accordingly we write, for the energy tensor in the body,

$$T_{rs} = R_{rs} + S_{rs}, \quad T_{rs,s} = 0, \tag{146}$$

where R_{rs} is regarded as due to the motion of the body as a whole (with proper mass density changed by heat) and to elastic stresses, and S_{rs} to the field alone.

Following MINKOWSKI[1] [1908], we shall take the electromagnetic

[1] For a fuller treatment of these matters, with a discussion of Abraham's alternative form for the energy tensor, see PAULI [1920, p. 665] and MØLLER [1952, p. 202].

energy tensor to be

$$S_{rs} = c^{-2}(F_{mr}H_{ms} - \tfrac{1}{4}\delta_{rs}F_{mn}H_{mn}). \tag{147}$$

Noting that $(Y^*_{rs})^* = -Y_{rs}$, the general formula VII–(142) gives, on replacing Y by Y^*, and changing the suffixes,

$$X_{mr}Y_{ms} - X^*_{ms}Y^*_{mr} = -\tfrac{1}{2}\delta_{rs}(X^*Y^*) = \tfrac{1}{2}\delta_{rs}(XY); \tag{148}$$

thus (147) may also be written

$$S_{rs} = \tfrac{1}{2}c^{-2}(F_{mr}H_{ms} + F^*_{ms}H^*_{mr}). \tag{149}$$

Note the analogy between the above formulae and IX–(32), (40). But there is an important difference; hitherto we had symmetric energy tensors, but $S_{rs} \neq S_{sr}$.

Differentiation of (147) gives, with the first of (143) and rearrangement,

$$\begin{aligned} c^2S_{rs,s} &= F_{mr}J_m + F_{mr,s}H_{ms} - \tfrac{1}{2}F_{mn,r}H_{mn} \\ &\quad + \tfrac{1}{4}(F_{mn,r}H_{mn} - F_{mn}H_{mn,r}). \end{aligned} \tag{150}$$

The right hand side vanishes except for the first term. For, by use of skew-symmetries,

$$\begin{aligned} F_{mr,s}H_{ms} - \tfrac{1}{2}F_{mn,r}H_{mn} &= \tfrac{1}{2}H_{mn}(F_{mr,n} - F_{nr,m} - F_{mn,r}) \\ &= \tfrac{1}{2}H_{mn}(F_{mr,n} + F_{rn,m} + F_{nm,r}) \end{aligned} \tag{151}$$

and this vanishes by the second of (143). The last term in (150) is a vector; we prove it zero in all frames by proving it zero in the rest frame of the body, for which (144) gives

$$H_{\varrho 4} = \varepsilon F_{\varrho 4}, \quad F^*_{\varrho\mu} = \mu H^*_{\varrho\mu}, \quad F_{\varrho\sigma} = \mu H_{\varrho\sigma}, \tag{152}$$

and hence

$$\begin{aligned} F_{mn,r}H_{mn} &= F_{\mu\nu,r}H_{\mu\nu} + 2F_{\mu 4,r}H_{\mu 4} \\ &= \mu H_{\mu\nu,r}H_{\mu\nu} + 2\varepsilon F_{\mu 4,r}F_{\mu 4}, \\ F_{mn}H_{mn,r} &= F_{\mu\nu}H_{\mu\nu,r} + 2F_{\mu 4}H_{\mu 4,r} \\ &= \mu H_{\mu\nu}H_{\mu\nu,r} + 2\varepsilon F_{\mu 4}F_{\mu 4,r}. \end{aligned} \tag{153}$$

We have then

$$c^2S_{rs,s} = -F_{rs}J_s, \tag{154}$$

and (146) gives

$$R_{rs,s} = c^{-2}F_{rs}J_s. \tag{155}$$

Let us use the word *material* in connection with R_{rs}, recognising however that heat is included in it. Then, if we integrate (155) through the interior R of a closed 3-space S, the total flux of material 4-momentum out across S is given by

$$\int_S \varepsilon(n) R_{rs} n_s \, dS = c^{-2} \int_R F_{rs} J_s \, d\tau, \tag{156}$$

n_r being the unit outward normal of S.

Let us suppose the body isolated with respect to the material energy tensor, in the sense of VIII–(245), across a 3-space Σ which is the history of the bounding surface of the body, so that on Σ we have

$$R_{rs} n_s = 0. \tag{157}$$

This means that no heat flows across Σ. Let Σ_0, Σ_1 be two 3-spaces and R the region of space-time bounded by Σ, Σ_0, Σ_1, as in Fig. VIII–11 (p. 310). Replacing S in (156) by $\Sigma + \Sigma_0 + \Sigma_1$ and noting (157), we have, with n_r as in Fig. VIII–11,

$$\int_{\Sigma_1} \varepsilon(n) R_{rs} n_s \, dS - \int_{\Sigma_0} \varepsilon(n) R_{rs} n_s \, dS = c^{-2} \int_R F_{rs} J_s \, d\tau. \tag{158}$$

Now use the rest frame of the body and take Σ_0, Σ_1 to be the 3-flats $t = t_0$, $t = t_1$; the integrals on the left now represent the material 4-momentum of the system, say M_r [cf. VIII–(247)], and we have

$$M_r(t_1) - M_r(t_0) = c^{-2} \int F_{rs} J_s \, d\tau. \tag{159}$$

Taking $r = 4$ and multiplying by c^2/i, we get

$$\begin{aligned} &\text{increase in material energy in time interval } (t_0, t_1) \\ &= \int_{t_0}^{t_1} c\,dt \int (-iF_{4\sigma} J_\sigma \, dS) \\ &= \int_{t_0}^{t_1} dt \int c E_\sigma J_\sigma \, dS, \end{aligned} \tag{160}$$

the last integral being taken over the instantaneous space $t = \text{const.}$ with 3-element dS. We recognise a standard pre-relativistic formula, involving the scalar product of the electric intensity E_σ and the current cJ_σ, for the amount of *Joulean heat* generated in the interval (t_0, t_1).

APPENDIX A

3-WAVES AND 2-WAVES

[pp. 314, 351, 359]

Consider an equation

$$F(x_1, x_2, x_3, x_4) = 0. \tag{1}$$

It defines a 3-space, say Σ_3, in space-time; we call Σ_3 a 3-*wave*. If we put $x_4 = \text{const.}$, (1) gives a surface (2-space), say Σ_2, in the observer's space; we call Σ_2 a 2-*wave*. Thus *a 3-wave is the history of a 2-wave.*

Note that a 3-wave is an invariant concept — a geometrical object in space-time — whereas the form of the 2-waves which make up its history depends on the particular observer who slices across Σ_3 with his instantaneous 3-flats $x_4 = \text{const}$. For this reason it is better as a rule to think in terms of 3-waves rather than in terms of 2-waves; but we have to use 2-waves to make contact with ordinary physical thought.

The equation (1) covers both Σ_3 and Σ_2; x_4 is variable when we think of Σ_3, but constant when we think of Σ_2. The normal to Σ_3 is a 4-vector with components proportional to $F_{,r}$ and the normal to Σ_2 (i.e. the normal to a surface in the ordinary sense) has direction ratios $F_{,\varrho}$. Here the comma indicates partial differentiation: $F_{,r} = \partial F/\partial x_r$, $F_{,\varrho} = \partial F/\partial x_\varrho$; the reader is reminded of the convention that Latin suffixes have the range 1, 2, 3, 4 and Greek suffixes the range 1, 2, 3.

To find the *wave velocity*, as ordinarily understood, we proceed along the normal to Σ_2, drawn for $x_4 = k$, say, until we arrive at the Σ_2 drawn for $x_4 = k + icdt$, where dt is an infinitesimal time. Then the wave velocity is the 3-vector w_ϱ with components

$$w_\varrho = dx_\varrho/dt, \tag{2}$$

where dx_ϱ is the infinitesimal displacement along the normal. But since we are keeping on Σ_3 (the history of Σ_2), (1) gives

$$F_{,\varrho}\,dx_\varrho + F_{,4}\,dx_4 = 0, \tag{3}$$

and so, by (2),

$$w_\varrho F_{,\varrho} + icF_{,4} = 0. \tag{4}$$

We obtain w_ϱ from this equation and the condition expressing that w_ϱ is normal to Σ_2, viz.

$$w_\varrho = \theta F_{,\varrho}, \tag{5}$$

where θ is a factor of proportionality. Substitution in (4) gives

$$\theta = -ic\frac{F_{,4}}{F_{,\varrho}F_{,\varrho}}, \tag{6}$$

and so, by (5), *the wave velocity for the 3-wave* $F = 0$ *is the 3-vector*

$$w_\varrho = -ic\frac{F_{,\varrho}F_{,4}}{F_{,\sigma}F_{,\sigma}}. \tag{7}$$

We can also write this as

$$w_\varrho = -ic\frac{W_\varrho W_4}{W_\sigma W_\sigma}, \tag{8}$$

$$W_r = \phi F_{,r}, \tag{9}$$

ϕ being a factor of proportionality; it is sometimes convenient to choose ϕ so that W_r is a unit 4-vector (the *unit 3-wave normal*), but of course this cannot be done if $F_{,r}$ happens to be a null vector.

The simplest case is that in which F is linear, so that the 3-wave is

$$F = a_r x_r + a = 0. \tag{10}$$

Then Σ_3 is a 3-flat in space-time and the 2-waves Σ_2 are plane waves with the constant wave velocity

$$w_\varrho = -ic\frac{a_\varrho a_4}{a_\sigma a_\sigma}. \tag{11}$$

In general the speed w of a 2-wave (the magnitude of the wave velocity) is given by

$$w^2 = -c^2\frac{W_4^2}{W_\varrho W_\varrho}. \tag{12}$$

Therefore

$$1 - \frac{w^2}{c^2} = \frac{W_r W_r}{W_\varrho W_\varrho}. \tag{13}$$

Here the denominator is positive, being the sum of the squares of

three real quantities, and we have the following important result: *The speed of a wave is less than, equal to, or greater than that of light, according as the 3-wave normal W_r is spacelike* ($W_r W_r > 0$), *null* ($W_r W_r = 0$), *or timelike* ($W_r W_r < 0$).

There is nothing in the above mathematical argument to exclude waves which travel faster than light. Indeed the de Broglie waves associated with material particles do travel faster than light (cf. SYNGE [1954a] for a treatment of de Broglie waves along the above lines). This does not violate the basic relativistic ideas about time, as developed in Ch. I, because these ideas demand only that no *signal* can travel faster than light, and a signal (a transmission of energy) travels with group speed, not with phase speed; the group speed of de Broglie waves is in fact the speed of a particle, and that always remains less than c.

APPENDIX B

RADIATION OF ENERGY FROM AN ACCELERATED CHARGE

[pp. 391, 393]

The approximate calculation of the rate of radiation of energy from an accelerated charge (cf. JEANS [1927, p. 577]) conceals the invariant nature of the result. Here we shall make an exact calculation.

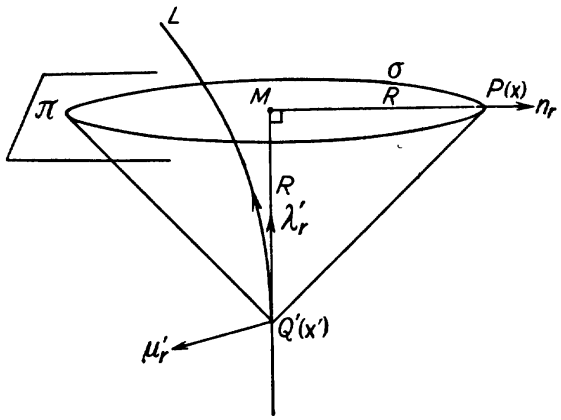


Fig. 1 – Radiation of energy

In Fig. 1, $P(x)$ is any event; $Q'(x')$ is the event where the world line L of a charge e cuts the null cone drawn from P into the past. Then P lies on the null cone drawn from Q' into the future.

We shall use the retarded potential, so that, as in x–(23),

$$\phi_r = \frac{e}{4\pi}\frac{\lambda'_r}{w'}, \qquad w' = \lambda'_n(x'_n - x_n), \tag{1}$$

and the electromagnetic field at P is

$$F_{mr} = \phi_{r,m} - \phi_{m,r}. \tag{2}$$

To calculate this, we need certain partial derivatives. The coordinates x'_r are determined by x_r and the equations $x'_r = x'_r(s')$ of L. Varying x_r in the equation

$$(x_n - x'_n)\,(x_n - x'_n) = 0, \tag{3}$$

and writing $dx'_n = \lambda'_n ds'$, we have

$$(x_n - x'_n)\,(dx_n - \lambda'_n ds') = 0, \tag{4}$$

and hence

$$\frac{\partial s'}{\partial x_n} = -\frac{x_n - x'_n}{w'}. \tag{5}$$

Therefore

$$\frac{\partial x'_n}{\partial x_p} = \lambda'_n \frac{\partial s'}{\partial x_p} = -\lambda'_n(x_p - x'_p)/w', \tag{6}$$

and

$$\frac{\partial w'}{\partial x_p} = \mu'_n \frac{\partial s'}{\partial x_p}(x'_n - x_n) + \lambda'_n\left(\frac{\partial x'_n}{\partial x_p} - \delta_{np}\right)$$
$$= -W'(x_p - x'_p)/w' + (x_p - x'_p)/w' - \lambda'_p, \tag{7}$$

where

$$\mu'_r = d\lambda'_r/ds', \quad W' = \mu'_n(x'_n - x_n), \tag{8}$$

μ'_r being in fact the 4-acceleration of the charge.

We now differentiate ϕ_r as given in (1), obtaining

$$\phi_{r,m} = \frac{e}{4\pi}\left(\frac{1}{w'}\mu'_r\frac{\partial s'}{\partial x_m} - \frac{1}{w'^2}\lambda'_r\frac{\partial w'}{\partial x_m}\right)$$
$$= \frac{e}{4\pi}\frac{1}{w'^2}\left[-\mu'_r(x_m - x'_m) + \frac{W'}{w'}\lambda'_r(x_m - x'_m)\right.$$
$$\left. - \frac{1}{w'}\lambda'_r(x_m - x'_m) + \lambda'_r\lambda'_m\right], \tag{9}$$

and thus by (2) the field at P is

$$F_{mr} = \frac{e}{4\pi}\frac{1}{w'^2}\left[\frac{W' - 1}{w'}\{\lambda'_r(x_m - x'_m) - \lambda'_m(x_r - x'_r)\}\right.$$
$$\left. - \mu'_r(x_m - x'_m) + \mu'_m(x_r - x'_r)\right]. \tag{10}$$

We note that

$$\lambda'_r\lambda'_r = -1, \quad \lambda'_r\mu'_r = 0, \tag{11}$$

and therefore

$$F_{mr}\lambda'_r = \frac{e}{4\pi}\frac{1}{w'^2}\left[\frac{W' - 1}{w'}\{-(x_m - x'_m) + w'\lambda'_m\} - w'\mu'_m\right]. \tag{12}$$

Now draw the 3-flat Π through P and orthogonal to λ'_r, cutting at the event M the tangent drawn to L at Q'; the coordinates of M may be written $x'_r + R\lambda'_r$. Let n_r be the unit vector along MP. Then

$$x_r - x'_r = R(\lambda'_r + n_r),$$
$$(x_r - x'_r)\lambda'_r = -R = -w', \quad (x_r - x'_r)n_r = R, \tag{13}$$
$$W' = -Rn_r\mu'_r.$$

Thus (12) may be written

$$F_{mr}\lambda'_r = \frac{e}{4\pi}\frac{1}{R^2}[(1 - W')n_m - R\mu'_m]. \tag{14}$$

Also

$$F_{mr}n_r = \frac{e}{4\pi}\frac{1}{R^2}[(1 - W')\lambda'_m - R\mu'_r n_r(\lambda'_m + n_m) + R\mu'_m]$$

$$= \frac{e}{4\pi}\frac{1}{R^2}(\lambda'_m + W'n_m + R\mu'_m), \tag{15}$$

and therefore

$$F_{mr}\lambda'_r F_{ms}n_s = \frac{e^2}{16\pi^2}\frac{1}{R^4}(W'^2 - R^2\mu'^2), \tag{16}$$

where $\mu'^2 = \mu'_r\mu'_r$, this being a spacelike vector.

By IX–(40) the energy tensor at P is

$$T_{rs} = c^{-2}(F_{mr}F_{ms} - \tfrac{1}{4}\delta_{rs}F_{mn}F_{mn}), \tag{17}$$

and so

$$c^2T_{rs}\lambda'_r n_s = F_{mr}\lambda'_r F_{ms}n_s, \tag{18}$$

the invariant evaluated in (16).

The 3-flat Π cuts the null cone with vertex Q' in a 2-sphere σ of radius R, and n_r is the unit normal to σ and to λ'_r. Let us evaluate the integral

$$I = c^2\int T_{rs}\lambda'_r n_s d\sigma. \tag{19}$$

This is an invariant, and so we can calculate it by using that frame for which λ'_r is time-axis; then

$$\lambda'_\varrho = 0, \quad \lambda'_4 = i, \quad \mu'_4 = 0, \quad n_4 = 0, \tag{20}$$

and by (16) and (18)

$$c^2T_{rs}\lambda'_r n_s = \frac{e^2}{16\pi^2}\frac{1}{R^2}[(\mu'_\varrho n_\varrho)^2 - \mu'^2]. \tag{21}$$

If dS is an element of 3-volume of Π inside σ, we have

$$\int(\mu'_\varrho n_\varrho)^2 d\sigma = R^{-1}\mu'_\varrho\mu'_\tau\int x_\varrho n_\tau d\sigma$$

$$= R^{-1}\mu'_\varrho\mu'_\tau\int\delta_{\varrho\tau}dS = \frac{4\pi}{3}R^2\mu'^2, \tag{22}$$

$$\int\mu'^2 d\sigma = 4\pi R^2\mu'^2,$$

and so

$$I = -\frac{e^2}{16\pi^2}\frac{8\pi}{3}\mu'^2 = -\frac{e^2}{6\pi}\mu'^2. \tag{23}$$

In the frame of reference we are employing we have by VIII–(132)

$$T_{rs}\lambda'_r n_s = iT_{4\varrho}n_\varrho = -c^{-3} \times \text{flux of energy per unit time per unit area normal to } n_\varrho, \tag{24}$$

and therefore the flux of energy per unit time out across the 2-sphere σ is

$$\begin{aligned} F &= -c^3 \int T_{rs}\lambda'_r n_s d\sigma \\ &= -cI = e^2 c\mu'^2/(6\pi). \end{aligned} \tag{25}$$

We can express this in terms of 3-acceleration f'_ϱ. We have, with $\gamma' = cdt'/ds'$,

$$\begin{aligned} \lambda'_r &= \frac{dx'_r}{ds'} = \frac{\gamma'}{c}\frac{dx'_r}{dt'}, \\ \mu'_\varrho &= \frac{\gamma'}{c}\frac{d}{dt'}\left(\frac{\gamma'}{c}\frac{dx'_\varrho}{dt'}\right) = \frac{1}{c^2}\frac{d^2x'_\varrho}{dt'^2} = \frac{1}{c^2}f'_\varrho, \end{aligned} \tag{26}$$

since $dx'_\varrho/dt = 0$, $\gamma' = 1$ at Q'; hence

$$\mu'^2 = f'^2/c^4, \tag{27}$$

and so the outward flux of energy per unit time is

$$F = \frac{e^2 f'^2}{6\pi c^3}, \tag{28}$$

as in X–(25).

The invariant result which we have established is this:

$$c^2 \int T_{rs}\lambda'_r n_s d\sigma = -\frac{e^2\mu'^2}{6\pi}, \tag{29}$$

where μ' is the magnitude of the 4-acceleration at Q'. The remarkable fact is that the value of this integral is independent of R; it is the same for all 2-spheres obtained by cutting the null cone with vertex Q' by 3-flats orthogonal to λ'_r. But the interpretation of this result as a relativistic conservation law is not apparent, since it involves integration over a 2-space and not over a 3-space.

APPENDIX C

SCATTERING AND CAPTURE BY A FIXED NUCLEUS

[p. 399]

Let O be a fixed charge e' (a nucleus, regarded as a point). A particle of proper mass m and charge e, of opposite sign to e', comes from infinity with speed v_0. We want to find out what happens, using the equations of motion of the relativistic Kepler problem (X–§ 5) (cf. DARWIN [1913]).

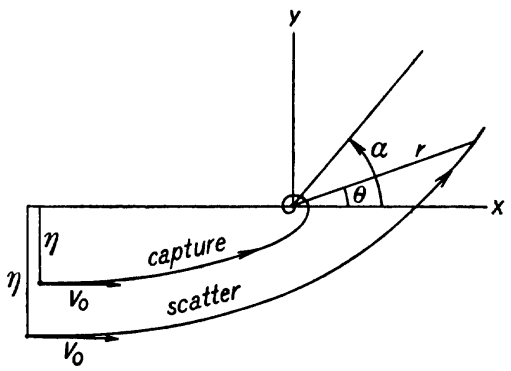


Fig. 1 – Scattering and capture by a fixed nucleus

In polar coordinates (r, θ) (Fig. 1) with $\varrho = 1/r$, the differential equation of the orbit is, as in X–(43),

$$\left(\frac{d\varrho}{d\theta}\right)^2 = f(\varrho), \tag{1}$$

where

$$f(\varrho) = -\varrho^2 - \frac{c^2}{A^2} + \frac{1}{A^2c^2}(W + k\varrho)^2. \tag{2}$$

Here A and W are constants of the motion:

$$A = \gamma r^2\dot{\theta}, \quad W = \gamma c^2 - k/r, \quad k = -ee'/m > 0. \tag{3}$$

We have made a slight change from Chap. X; e and e' are now measured in Gaussian electrostatic units.

To find the initial value of $d\varrho/d\theta$, we assume first that the particle starts from the finite point $x = x_0$, $y = -\eta$ with velocity $\dot{x} = v_0$, $\dot{y} = 0$, so that we have initially

$$\dot{r} = v_0 \cos\theta, \quad r\dot{\theta} = -v_0 \sin\theta, \quad r^2\dot{\theta} = v_0\eta, \tag{4}$$

and therefore

$$\frac{d\varrho}{d\theta} = -\frac{\dot{r}}{r^2\dot{\theta}} = -\frac{\cos\theta}{\eta}. \tag{5}$$

Here η is the distance by which the particle would miss the nucleus if there were no Coulomb field; it is called the *impact parameter*. Holding η and v_0 fixed in value and letting $x_0 \to -\infty$, we get as the appropriate initial conditions

$$\varrho = 0, \quad \frac{d\varrho}{d\theta} = \frac{1}{\eta}, \quad \theta = -\pi. \tag{6}$$

Substitution of these values in (1) and (2) gives

$$\frac{1}{\eta^2} = -\frac{c^2}{A^2} + \frac{W^2}{A^2c^2}, \quad \frac{c^2}{A^2} = \frac{c^4}{\eta^2(W^2 - c^4)}, \tag{7}$$

and so we can write $f(\varrho)$ in the form

$$\begin{aligned} f(\varrho) &= -\varrho^2 + \frac{(W + k\varrho)^2 - c^4}{\eta^2(W^2 - c^4)} \\ &= \frac{1}{\eta^2(W^2 - c^4)} \{\varrho^2[k^2 - \eta^2(W^2 - c^4)] + 2kW\varrho + W^2 - c^4\}. \end{aligned} \tag{8}$$

We note that

$$W = \gamma_0 c^2 > c^2, \quad \gamma_0 = (1 - v_0^2/c^2)^{-\frac{1}{2}}. \tag{9}$$

In (8) we have $f(\varrho)$ expressed in terms of the fixed constant k and the two parameters η, v_0 corresponding to initial position and velocity. Since f is quadratic in ϱ, we can solve (1) in terms of trigonometric or hyperbolic functions, but it is simpler to proceed as follows.

We see that $f(\varrho)$ has two real zeros:

$$\begin{aligned} \varrho_1 &= \frac{kW + [k^2c^4 + \eta^2(W^2 - c^4)^2]^{\frac{1}{2}}}{\eta^2(W^2 - c^4) - k^2}, \\ \varrho_2 &= \frac{kW - [k^2c^4 + \eta^2(W^2 - c^4)^2]^{\frac{1}{2}}}{\eta^2(W^2 - c^4) - k^2}. \end{aligned} \tag{10}$$

Case (i):

$$\eta^2 > \frac{k^2}{W^2 - c^4}. \tag{11}$$

In this case we have $\varrho_1\varrho_2 < 0$, with ϱ_1 positive and ϱ_2 negative. The graph of ϱ against θ is then as in Fig. 2 — a sine curve. The particle makes its closest approach to the nucleus at B, where $\varrho = \varrho_1$; ϱ decreases to zero at C, the particle having then receded to infinity.

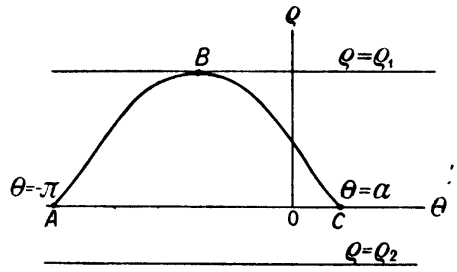


Fig. 2. – Graph of ϱ against θ for scattering without capture (Case (i))

Scattering occurs, and the angle of scattering α (Fig. 1) is given by

$$\alpha + \pi = \int_A^B d\theta + \int_B^C d\theta, \tag{12}$$

or

$$\alpha = -\pi + 2\int_0^{\varrho_1} \frac{d\varrho}{\sqrt{f(\varrho)}}. \tag{13}$$

We can write (8) in the form

$$f(\varrho) = K^2(\varrho_1 - \varrho)\,(\varrho - \varrho_2) \tag{14}$$

with

$$K = \sqrt{1 - \frac{k^2}{\eta^2(W^2 - c^4)}}, \tag{15}$$

and hence

$$\alpha = -\pi + \frac{2}{K}\int_0^{\varrho_1} \frac{d\varrho}{\sqrt{(\varrho_1 - \varrho)\,(\varrho - \varrho_2)}}. \tag{16}$$

Putting

$$\varrho = \varrho_1 \cos^2\phi + \varrho_2 \sin^2\phi, \tag{17}$$

we obtain

$$\alpha = -\pi + \frac{4}{K}\phi_1, \tag{18}$$

where ϕ_1 is defined by

$$\tan^2\phi_1 = -\frac{\varrho_1}{\varrho_2} = \frac{[k^2c^4 + \eta^2(W^2 - c^4)^2]^{\frac{1}{2}} + kW}{[k^2c^4 + \eta^2(W^2 - c^4)^2]^{\frac{1}{2}} - kW}, \tag{19}$$

$$0 < \phi_1 < \tfrac{1}{2}\pi.$$

From (18) we can calculate the angle of scattering for any value of v_0 and any value of η satisfying (11).

Case (ii):

$$\eta^2 < \frac{k^2}{W^2 - c^4}. \tag{20}$$

In this case it is clear from (8) that $\varrho_1\varrho_2 > 0$ and $\varrho_1 + \varrho_2 < 0$, so that ϱ_1 and ϱ_2 are both negative. Thus ϱ increases steadily to infinity, since

$d\varrho/d\theta$ cannot vanish; the particle spirals in to the nucleus and is captured ($r = 0$). This means that if a particle with energy $E = mW$ is projected in from a point at infinity near enough to the x-axis, it will be captured. To be captured, it must initially lie in a circular region on the plane at infinity, *the cross-section of capture*; the area of this cross-section is

$$\sigma = \frac{\pi k^2}{W^2 - c^4} = \frac{\pi(ee')^2}{E^2 - m^2c^4} . \tag{21}$$

APPENDIX D

THE ABSOLUTE 2-CONTENT OF A 3-CELL ON A NULL CONE

[p. 405]

Consider a pseudosphere with centre a_r and radius R (Fig. 1), its equation being

$$(x_r - a_r)\,(x_r - a_r) = -R^2. \tag{1}$$

Let x_r be any event on the pseudosphere, n_r any timelike unit vector pointing into the future, T a thin tube parallel to n_r, dS_p the 3-volume cut from the pseudosphere by T, and dS_n the normal 3-volume of T. Then, by the projection formula VIII–(69),

$$dS_n = dS_p \mid N_r n_r \mid, \tag{2}$$

where N_r is the unit normal to the pseudosphere; its components are

$$N_r = (x_r - a_r)/R, \tag{3}$$

as we see on differentiating (1). Hence

$$dS_n = \frac{dS_p}{R} \mid n_r(x_r - a_r) \mid, \tag{4}$$

or

$$d\omega = \frac{dS_p}{R} = \frac{dS_n}{\mid n_r(x_r - a_r) \mid}, \tag{5}$$

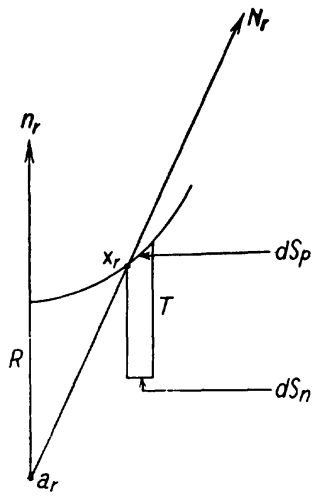


Fig. 1 – Pseudosphere used to define absolute 2-content on null cone

$d\omega$ being thus defined. We call $d\omega$ the *absolute 2-content* of the 3-element of the pseudosphere with 3-volume dS_p (note that $d\omega$ has the dimensions of an area).

Starting with a given 3-element on the pseudosphere, and drawing through it a tube T in any direction n_r, we can calculate $d\omega$ by evaluating the last fraction in (5), dS_n being the 3-volume of the normal section of T. Thus, although the numerator and denominator of this fraction depend on the choice of n_r, the value of the fraction is independent of n_r.

Holding n_r fixed and also the tube T, let the radius R tend to zero, so that the pseudosphere becomes a null cone in the limit. In this process, x_r is the only thing which changes in the last fraction in (5), and this fraction attains a definite limit. Therefore, as R tends to zero, dS_p tends to zero also, which means that *the 3-volume of a 3-element of a null cone is zero.* Nevertheless such a 3-element possesses an *absolute 2-content* given by

$$d\omega = \lim_{R\to 0} \frac{dS_p}{R} = \frac{dS_n}{|\, n_r(x_r - a_r)\,|}. \tag{6}$$

This 2-content is Lorentz-invariant. To calculate it, one must calculate the last fraction in (6), dS_n being the 3-volume of the normal section of a tube T, drawn through the 3-element of the null cone in the direction of some unit vector n_r; *the value of the fraction is independent of the choice of* n_r.

We always take $d\omega$ to be positive as in (6). Thus, if x_r is on the past sheet of the null cone, we have

$$d\omega = \frac{dS_n}{n_r(x_r - a_r)}, \tag{7}$$

and if x_r is on the future sheet, as indicated in Fig. 1,

$$d\omega = \frac{dS_n}{-\, n_r(x_r - a_r)}. \tag{8}$$

Note that in both cases n_r points into the future.

The important thing to remember is that, though the right hand sides in (6), (7), (8) can be evaluated only by making some choice of n_r, the values are independent of that choice. The invariant $|\, n_r(x_r - a_r)\,|$ has a very simple geometrical meaning: it is the length of the perpendicular dropped from x_r on the straight line through a_r with direction n_r. It is the spatial distance of the event x_r on the null cone from a_r (the vertex of the null cone), as measured by an observer who takes n_r for time axis.

APPENDIX E

CALCULATIONS FOR RETARDED POTENTIAL

[p. 408]

Here are some details which the reader may find useful in establishing formula x–(101).

Let x_r, x'_r be any two events subject only to the condition that the straight line joining them is a null line with x_r after x'_r (Fig. 1). Then

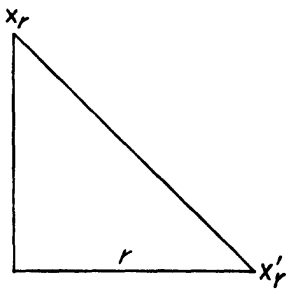


Fig. 1 – Space-time diagram for retarded potential

$$x'_4 = x_4 - ir, \quad r^2 = (x'_\varrho - x_\varrho)(x'_\varrho - x_\varrho) \tag{1}$$

(for simplicity we have written r instead of r' as in x–(98)). Then, if J is any function of position in space-time, we may write, with x', x standing for x'_ϱ, x_ϱ respectively,

$$J' = J(x', x'_4) = J(x', x_4 - ir) = f(x', x, x_4), \tag{2}$$

a function of the seven variables x', x, x_4, which are clearly independent. Giving arbitrary variations to the independent variables and denoting partial derivatives by subscripts, we have

$$J'_{\varrho'}\,dx'_\varrho + J'_{4'}\,dx'_4 = f_{\varrho'}\,dx'_\varrho + f_\varrho\,dx_\varrho + f_4 dx_4. \tag{3}$$

By (1)

$$\begin{aligned} dx_4 &= dx'_4 + ir_{\varrho'}(dx'_\varrho - dx_\varrho), \\ r_{\varrho'} &= (x'_\varrho - x_\varrho)/r = -r_\varrho. \end{aligned} \tag{4}$$

This we substitute in (3) so that only the differentials dx'_ϱ, dx_ϱ, dx'_4 appear. Although it was x_4 and not x'_4 that was previously named as one of the seven independent variables, we may equally well use x'_4. From the coefficients of dx'_ϱ and dx'_4 in (3), after substitution from (4), we get

$$J'_{\varrho'} = f_{\varrho'} + if_4 r_{\varrho'}, \quad J'_{4'} = f_4; \tag{5}$$

we do not bother about the coefficients of dx_ϱ.

Now each of the four equations contained in (5) is an equation of the same general form as (2), and so we can repeat the process, obtaining

$$\begin{aligned} J'_{\varrho'\sigma'} &= (f_{\varrho'} + if_4 r_{\varrho'})_{\sigma'} + i(f_{\varrho'} + if_4 r_{\varrho'})_4 r_{\sigma'} \\ &= f_{\varrho'\sigma'} + if_{4\sigma'} r_{\varrho'} + if_4 r_{\varrho'\sigma'} + if_{4\varrho'} r_{\sigma'} - f_{44} r_{\varrho'} r_{\sigma'}, \\ J'_{4'4'} &= f_{44}. \end{aligned} \tag{6}$$

Referring to (4), we see that

$$\begin{aligned} r_{\varrho'} r_{\varrho'} = 1, \quad r_{\varrho'\sigma'} &= \frac{\delta_{\varrho\sigma}}{r} - \frac{(x'_\varrho - x_\varrho)(x'_\sigma - x_\sigma)}{r^3}, \\ \Delta' r = r_{\varrho'\varrho'} &= \frac{3}{r} - \frac{1}{r} = \frac{2}{r}, \end{aligned} \tag{7}$$

and so we obtain from (6)

$$\begin{aligned} \Delta' J' = J'_{\varrho'\varrho'} &= \Delta' f + 2if_{4\varrho'} r_{\varrho'} + 2if_4/r - f_{44}, \\ \Box' J' &= \Delta' f + 2if_{4\varrho'} r_{\varrho'} + 2if_4/r, \end{aligned} \tag{8}$$

the term f_{44} disappearing; this is the same as x–(99).

Holding x_r fixed and integrating over the past sheet N' of the null cone with vertex x_r, we get

$$\int_{N'} \Box' J' d\omega' = \int \Delta' f \frac{dS'}{r} + 2iA, \tag{9}$$

where

$$A = \int (f_{4\varrho'} r_{\varrho'} + f_4/r) dS'/r. \tag{10}$$

Integration by parts gives, with use of (7), and rejection of an integral at infinity,

$$\begin{aligned} \int f_{4\varrho'} r_{\varrho'} \frac{dS'}{r} &= -\int f_4 \left(\frac{r_{\varrho'}}{r}\right)_{\varrho'} dS' \\ &= -\int f_4 \left(\frac{2}{r^2} - \frac{1}{r^2}\right) dS' = -\int f_4 \frac{dS'}{r^2}, \end{aligned} \tag{11}$$

and therefore $A = 0$. Thus (9) gives

$$\int_{N'} \Box' J' d\omega' = \int \Delta' f \frac{dS'}{r} = -4\pi f(x, x, x_4), \tag{12}$$

by the well known formula of potential theory. But if $x'_\varrho = x_\varrho$, then by (1) we have $r = 0$, $x'_4 = x_4$, and so

$$f(x, x, x_4) = J(x, x_4); \tag{13}$$

therefore

$$\int_{N'} \Box' J' \, d\omega' = -4\pi J, \tag{14}$$

which is the formula x–(101). By the device of shifting the null cone as in x–(92), this gives

$$\Box \int_{N'} J' d\omega' = -4\pi J. \tag{15}$$

We have used a single function J here instead of a 4-vector J_m as in x–(101) for notational simplicity, the vector character of J_m actually being unimportant as far as x–(101) is concerned. The several components are linked in the simpler formula x–(95).

REFERENCES

BADE, W. L. and H. JEHLE, Rev. Mod. Phys. **25** (1953) 714.

BAKER, B. B. and E. T. COPSON, 1950, The Mathematical Theory of Huygen's Principle (Clarendon Press, Oxford).

BATEMAN, H., 1944, Partial Differential Equations of Mathematical Physics (Dover Publications, New York).

BETHE, H. A., 1947, Elementary Nuclear Theory (Wiley, New York).

BONDI, H., 1952, Cosmology (Cambridge University Press).

BORN, M., Ann. Physik (4) **30** (1909) 1.

—— and L. INFELD, Proc. Roy. Soc. A **144** (1934), 425.

BRIDGMAN, P. W., 1927, The Logic of Modern Physics (Macmillan, New York).

—— 1952, The Nature of some of our Physical Concepts (Philosophical Library, New York).

BUCHWALD, E., Naturwissenschaften **38** (1951) 519.

CONWAY, A. W., Proc. London Math. Soc. (2) **1** (1903) 154.

—— 1953, Selected Papers (Dublin Institute for Advanced Studies).

COURANT, R., 1936, Differential and Integral Calculus, Vol. II (Blackie, London and Glasgow).

CUNNINGHAM, E., 1921, Relativity, the Electron Theory, and Gravitation (Longmans Green, London).

DARWIN, C. G., Phil. Mag. (6) **25** (1913) 201.

—— Phil. Mag. (6) **39** (1920) 537.

DIRAC, P. A. M., 1930, The Principles of Quantum Mechanics (Clarendon Press, Oxford).

—— Proc. Roy. Soc. A **167** (1938) 148.

DITCHBURN, R. W. and O. S. HEAVENS, Nature **170** (1952), 705.

DUGAS, R., 1950, Histoire de la Mécanique (Éditions du Griffon, Neuchatel).

EDDINGTON, A. S., 1924, The Mathematical Theory of Relativity (Cambridge University Press).

EINSTEIN, A., Ann. Physik (4) **17** (1905) 891.

—— Sitz. Preuss. Akad. Wiss. (1915) 778, 799, 831.

—— 1916, Die Grundlage der allgemeinen Relativitätstheorie (Barth, Leipzig) [Ann. Phys. (4) **49** (1916) 769].

—— 1920, Relativity (Methuen, London).

—— 1949, Albert Einstein: Philosopher-Scientist (Library of Living Philosophers Inc., Evanston, Illinois) [contains a bibliography of the writings of Einstein up to 1949].

EISENHART, L. P., Trans. Amer. Math. Soc. **26** (1924) 205.

—— 1926, Riemannian Geometry (Princeton University Press).

FINZI, B., Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat. (8) **12** (1952) 378, 477 and **13** (1952) 211.

FRENKEL, J., 1926, Lehrbuch der Elektrodynamik, Bd. 1 (Springer, Berlin).
GÅRDING, L., Acta Mathematica **85** (1951), 1.
GARDNER, G. H. F., Nature **170** (1952) 243.
—— 1953, The concept of a rigid body in special relativity (Thesis, Princeton University).
HERGLOTZ, G., Ann. Physik (4) **31** (1910) 393.
HUBBLE, E. P., Proc. Amer. Philosophical Soc. **95** (1951) 461.
—— Observatory **73** (1953) 99.
ILLINGWORTH, K. K., Phys. Rev. **30** (1927) 692.
INFELD, L., Phys. Zeitsch. **33** (1932) 475.
IVES, H. E., Sci. Proc. Roy. Dublin Soc. **26** NS (1952) 9.
—— and G. R. STILWELL, J. Opt. Soc. Amer. **28** (1938) 215 and **31** (1941) 369.
JEANS, J. H., 1925, The Dynamical Theory of Gases (Cambridge University Press).
—— 1927, The Mathematical Theory of Electricity and Magnetism (Cambridge University Press).
JEFFREYS, H. and B. S. JEFFREYS, 1946, Methods of Mathematical Physics (Cambridge University Press).
JOOS, G., Ann. Physik (5) **7** (1930) 385.
KENNEDY, R. J., Proc. Nat. Acad. Sci. **12** (1926) 621.
KERMACK, W. O. and W. H. MCCREA, Monthly Notices Roy. Astronomical Soc. **93** (1933) 519.
LANCZOS, C., Zeitschr. f. Physik **59** (1930) 514.
LARMOR, J., Phil. Mag. (5) **44** (1897) 503.
LAUB, J., Ann. Physik (4) **23** (1907) 741.
—— Ann. Physik (4) **25** (1908) 183.
LAUE, M. von, Ann. Physik (4) **23** (1907) 989.
LICHNEROWICZ, A., Comptes rendus Acad. Sci. **211** (1940) 117.
—— Ann. Sci. École Normale Sup. (3) **58** (1941) 285.
—— Comptes rendus Acad. Sci. **219** (1944) 270.
LIÉNARD, A., L'Éclairage électrique **16** (1898) 5, 53, 106.
LORENTZ, H. A., 1895, Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern (Brill, Leiden).
—— 1923, The Principle of Relativity (Methuen, London) [a collection of papers by Lorentz, Einstein, Minkowski and Weyl, reprinted by Dover Publications Inc., New York].
MARIOT, L., Comptes rendus Acad. Sci. **238** (1954) 2055 and **239** (1954) 1189.
MATRAI, T., Nature **172** (1953), 858.
MCCONNELL, A. J., Annali di Matematica (4) **6** (1928–29) 207.
MCCREA, W. H., 1949, Relativity Physics (Methuen, London).
MCMAHON, J., Proc. Roy. Irish Acad. A **55** (1953) 133.
MICHELSON, A. A. and E. W. MORLEY, Amer. J. Sci. (3) **34** (1887) 333.
—— F. G. PEASE and F. PEARSON, J. Opt. Soc. Amer. **18** (1929) 181.
MILLER, D. C., Rev. Mod. Phys. **5** (1933) 203.
MILNE, E. A., 1935, Relativity, Gravitation and World-Structure (Clarendon Press, Oxford).
MILNER, S. R., Phil. Mag. (6) **44** (1922) 705.

MINKOWSKI, H., Göttinger Nachr. (1908) 53.
—— Phys. Zeitschr. **10** (1909) 104 (translated in LORENTZ [1923]).
MØLLER, C., Comm. Dublin Institute for Advanced Studies A **5** (1949).
—— 1952, The Theory of Relativity (Clarendon Press, Oxford).
MURNAGHAN, F. D., 1938, The Theory of Group Representations (The Johns Hopkins Press, Baltimore).
NOETHER, F., Ann. Physik (4) **31** (1910) 919.
PAULI, W., 1920, Relativitätstheorie. Encyklopädie der mathematischen Wissenschaften V 19 (Teubner, Leipzig and Berlin).
PIRANI, F. A. E., Proc. Roy. Soc. A**228** (1955) 455.
POUNDER, J. R., Comm. Dublin Institute for Advanced Studies A **11** (1954).
PRESTON, T., 1912, Theory of Light (Macmillan, London).
RAINICH, G. Y., Trans. Amer. Math. Soc. **27** (1925) 106.
RAYNER, C. B., 1953, Foundations and Applications of Whitehead's Theory of Relativity (Thesis, Univ. of London).
—— Proc. Roy. Soc. A **222** (1954) 509.
RISCO, M., J. de Physique et le Radium (8) **13** (1952) 441.
RUSE, H. S., Proc. London Math. Soc. (2) **41** (1936) 302.
—— Proc. Roy. Soc. Edinburgh **57** (1937) 97.
RUSSELL, H. N., R. S. Dugan, and J. Q. Stewart, 1926, Astronomy, Vol. I (Ginn, Boston).
SALZMAN, G. and A. H. TAUB, Phys. Rev. **95** (1954) 1659.
SEN, N. R., Monthly Notices Roy. Astronomical Soc. **94** (1934a) 550.
—— Bull. Calcutta Math. Soc. **25** (1934b) 191.
SHANKLAND, R. S., S. W. MCCUSKEY, F. C. LEONE and G. KUERTI, Rev. Mod. Phys. **27** (1955) 167.
SHARPLESS, S., Astronomical Soc. of the Pacific, Leaflet 290 (1953).
SILBERSTEIN, L., 1924, The Theory of Relativity (Macmillan, London).
STEPHENSON, G. and C. W. KILMISTER, Nuovo Cimento (9) **10** (1953) 230.
STEVENSON, A. F., Quarterly App. Math. **12** (1954), 194.
STORCHI, E., Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat. (8) **14** (1953) 261.
STRUVE, O., Sky and Telescope **12** (1953) 203, 238.
SYNGE, J. L., Trans. Roy. Soc. Canada, Ser. 3, Sect. 3, **28** (1934) 127.
—— Phys. Rev. **47** (1935a) 760.
—— Univ. of Toronto Studies, Applied Math. Ser. **1** (1935b).
—— Proc. London Math. Soc. (2) **43** (1937a) 376.
—— 1937b, Geometrical Optics (Cambridge Univ. Press).
—— Proc. Roy. Soc. A **177** (1940) 118.
—— Institute for Fluid Dynamics and Applied Mathematics, Univ. of Maryland, Lecture Ser. **5** (1951).
—— Sci. Proc. Roy. Dublin Soc. **26** NS (1952a) 45.
—— Nature **170** (1952b) 243.
—— Proc. Roy. Soc. A **211** (1952c) 303.
—— 1954a, Geometrical Mechanics and de Broglie Waves (Cambridge Univ. Press).
—— 1954b, Studies in Mathematics and Mechanics presented to Richard von Mises (Academic Press, New York).

SYNGE, J. L., and B. A. GRIFFITH, 1949, Principles of Mechanics (McGraw-Hill, New York).
—— and A. SCHILD, 1952, Tensor Calculus (Univ. of Toronto Press).
TEMPLE, G., Proc. Physical Society **36** (1923), 176.
THACKERAY, A. D. and A. J. WESSELINK, Nature **171** (1953) 693.
TOLMAN, R. C., 1934, Relativity, Thermodynamics and Cosmology (Clarendon Press, Oxford).
UDESCHINI, P., Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat. (8) **13** (1952) 246.
WEYL, H., 1931, The Theory of Groups and Quantum Mechanics (Methuen, London) [reprinted by Dover Publications Inc., New York].
WEYSSENHOFF, J., Acta Physica Polonica **9** (1947) 26, 34, 46.
—— and A. RAABE, Acta Physica Polonica **9** (1947) 7, 19.
WHITEHEAD, A. N., 1922, The Principle of Relativity (Cambridge University Press).
WHITTAKER, Sir EDMUND, Proc. Roy. Soc. Edinburgh **42** (1922) 1.
—— 1937a, A Treatise on the Analytical Dynamics of Particles and Rigid Bodies (Cambridge University Press).
—— Proc. Roy. Soc. A **158** (1937b), 38.
—— 1951, A History of the Theories of Aether and Electricity (The Classical Theories) (Nelson, London).
—— 1953, A History of the Theories of Aether and Electricity (The Modern Theories, 1900–1926) (Nelson, London).
WIECHERT, E., Archives Néerlandaises **5** (1900) 549.
ZEEMAN, P., Proc. Roy. Acad. Amsterdam **17** (1914) 445; **18** (1916) 398, 1240.

INDEX

aberration 146 ff
Abraham, M. 400, 416
absolute 2-content of 3-cell 405, 430
  electric and magnetic strengths 334
  units 375
  world 122
absorption of photon 191
accelerated charge, radiation of energy from 393, 399 ff, 422 ff
acceleration 4-vector (4-acceleration) 130 ff, 301, 423
action and reaction 163, 255
  see also internal impulse
activity of relative force 169
advanced potential 390, 393, 408
aether *see* ether
after and before 12
age of earth and universe 157
Airy, G. B. 149
amplitude 133, 350
angle 29
  Eulerian 79 ff
  imaginary 74
  of aberration 147
  of scattering 428
angular momentum 216ff, 251ff, 312, 397
  of electromagnetic field 365, 368
  of continuous system 312
  open and closed laws of 218, 253, 254
  tensor 216 ff, 221 ff
  3-vector 225
  see also intrinsic angular momentum
annihilation of matter 199 ff
α-particle 182, 185
apparent distance 153 ff
Archimedes 2
area of 2-cell 272
atom, disintegration of 185
  emission and absorption of photon by 187 ff, 191 ff
  nucleus of 176, 181
  proper mass of 166
  Rutherford-Bohr 399
atomic clock 14, 120, 152
attractive internal impulses 212, 213, 238, 255 ff, 295
axes, canonical 332, 333
  of coordinates, apparent inclination of 117
  in space-time 75, 76, 123
axial null rays of 4-screw 88 ff
axioms 1 ff, 165
axis of rotation in space-time 79

Baade, W. 157
Bade, W. L. 103
Baker, B. B. 408
Balazs, N.L. viii
Balmer series 182, 193
Bateman, H. 408
before and after 11
Bethe, H. A. 181
binding energy 181
biquadratic equation in collision problems 233, 245, 249, 250
Bondi, H. 155
Born, M. 36, 415
Bridgman, P. W. 7
Brück, H. A. 157
Buchwald, E. 146

c (speed of light), *see* light
canonical axes and forms for electromagnetic field 331 ff, 336
capture by fixed nucleus 426 ff
Cartan, E. 75
Cartesian axes, oblique, in space-time
  diagram 124
  tensor 57
causality 393
cell in space-time 272, 273
  on null cone 430
centre of wave function 372
centroid of system 315
charge, electric 202, 388
  density of 403, 404
charged particle, accelerated 391 ff, 399 ff, 422
  equations of motion of 171, 394 ff, 400 ff, 410 ff
  scattering and capture of 426 ff
Chasles' theorem 85, 94
Christoffel symbols 33
classical *see* Newtonian
clock, apparent retardation of 118 ff, 122 ff, 189
  arbitrary 7
  paradox 16
  standard or atomic 14 ff, 38, 119, 152
closed law of angular momentum 218, 254

of 4-momentum 215, 253, 310
closed system *see* isolated system
coalescence of particles 183
collision 9, 173 ff
elastic 174, 205 ff, 227 ff
inelastic 174, 183 ff
Newtonian 173, 184, 227
of material particles 173 ff, 183 ff, 200, 202, 205 ff
with intrinsic angular momentum 227 ff, 235 ff, 237 ff, 246 ff
of photon and material particle 193 ff, 249
of two photons 200 ff, 250
semi-elastic 174
space-time diagrams of 9, 176, 194, 200, 202, 205
with high energy 237
with target 265 ff
comma denoting derivative xv
commutation of rotations 87
complex numbers 56
wave function 361, 366 ff
composition of 4-momenta 176 ff
of velocities 126 ff
Compton effect 193 ff, 202, 205, 249
wave length 197, 204, 375
concepts from Newtonian physics 2 ff
conductivity 305, 416
cone, null *see* null cone
congruence of curves 52, 281
conjugate vectors 66
conservation, equation of 285, 310, 323, 404
of angular momentum 218, 238, 248
of charge 202, 404
of energy 175
of 3-momentum 175
of 4-momentum 174 ff, 213 ff, 248, 285, 310
of number 404
of proper mass 167, 210, 238, 404
constant curvature 35
relative force 171
constitutive equations 416
2-content of 3-cell 405, 430
continuity, equation of 290
continuous medium or system 208
see also continuum, density, 4-velocity, etc.
continuum 261 ff
energy tensor of 281 ff, 309 ff
equations of motion of 300 ff
contraction of moving body 118 ff
contravariant tensors 57
convection coefficient 142 ff
convention, summation xv, 8
Conway, A. W. 391
coordinate vectors 78
coordinates in space-time, general 7 ff
Minkowskian 56 ff, 69 ff
oblique, in space-time diagram 124
real and imaginary xv
special 38, 47 ff
see also transformation
Copson, E. T. 408
cosmic radiation 176
cosmology 150 ff
Coulomb field 388 ff, 396 ff, 426
Courant, R. 276
covariant tensors 57
creation of matter 199 ff, 212
cross-section of capture 429
of world tube 273 ff
Cunningham, E. 146
curl 317
3-current and 4-current 404, 415
curvature, constant 35
tensor 34
curve in space-time 9, 64
curved space-time 35

δ_{mn} (Kronecker delta) 53
d'Alembertian operator 348
darkness, propagation of light into 385
Darwin, C. G. 396, 426
de Broglie waves 136, 197, 396, 421
decrease of proper mass in disintegration 180
degeneracy 292
density, macroscopic and microscopic 297, 300
mean 297
minimum 294, 327
numerical 261, 267, 270, 282, 289, 403
of charge 403, 404
of energy 262, 286 ff, 292, 324
of proper mass 262, 289, 403
of relative mass 262
of relative momentum 263, 286 ff, 324
positive character of 292, 306, 313
relative 261 ff, 404
Descartes, R. 75
determinacy of collision with angular momentum 234, 247 ff
deuteron 181
deviation of straight lines 45
diagonal energy tensor 294
diagram, *see* snapshot, space-time diagram
dielectric constant 416
Dirac, P. A. M. 103, 400
directions, principal, of electromagnetic tensor 336 ff
of energy tensor 290 ff, 325 ff

discrete and continuous electromagnetism 387
system 208 ff
disintegration 173 ff, 180 ff, 185 ff
energy released in 181
distance 23, 29 ff, 40, 45
apparent, of star 153 ff
luminosity 155
Ditchburn, R. W. 36
divergence of energy tensor 285
Doppler effect 135, 152, 189
drag coefficient 142 ff
dual tensor 222, 239
Dugan, R. S. 149
Dugas, R. 4
dummy indices *see* summation convention
dynamics *see* mechanics

E (relative energy) 169
E_o (proper energy) 170
ε_{mnrs} (permutation symbol) 73, 222, 239, 364
earth, age of 157
Eddington, A. S. 307
E-geometry 65, 123
eigen properties of electromagnetic tensor 336 ff
of energy tensor 290 ff, 325 ff
Einstein, A. 4, 5, 7, 22, 35, 164
Eisenhart, L. P. 75, 281, 306
elastic collision *see* collision
electric and magnetic strengths, absolute 334
electric charge 202, 388
current and density 404, 415
displacement 415
intensity (3-vector) 317, 320, 388, 415
electromagnetic field 317 ff, 387 ff
canonical forms 331 ff, 336
due to charge 387 ff, 409, 422
eigen properties 290 ff, 325 ff, 336 ff
energy tensor 323 ff, 365, 376, 385, 410 ff, 416 ff, 424
generated by scalar wave function 363 ff
geometry of 344
invariants 321, 322, 336 ff, 371
null (self-conjugate) 322, 327, 332, 335, 337, 341 ff, 352, 375
transformation of 320 ff, 333
electromagnetic fields intrinsically the same 372
model of material particle 366 ff, 374, 413
of photon 375 ff, 386
4-potential 345 ff, 363, 389 ff, 405 ff, 414, 422, 432
sheet 355, 391
shock waves 356, 385
tensor F_{rs} 317 ff, 387 ff, 415
H_{rs} 415
3-vectors 317, 415
waves 137, 350 ff
wrench 334, 335, 338
electron 166, 181, 182, 185, 375, 400
colliding with photon 193 ff
positron 202
optics 395
electrostatic units 388
elementary wave function 360, 361, 366 ff
emission of photon 187 ff, 191 ff
energy, binding 181
conservation of, 175
density 262, 286 ff, 292, 324
minimum of 294, 327
dissipated as heat 174, 184, 266, 270, 416, 418
flux of 287, 288, 324
identified with mass 165, 170
integral of 397
kinetic 169, 182
levels of atom 187
-momentum vector *see* 4-momentum
positive and negative 212, 292, 306, 313, 315
proper 170, 262
decreased in disintegration 180
increased in inelastic collision 184
numerical values for 182
radiation of 393, 399 ff, 422
relative 169, 173, 212
law of 175, 263 ff
numerical values of 182
of photon 173
released in disintegration 180
energy tensor 284 ff, 309 ff
eigen properties of 290 ff, 325 ff
electromagnetic 323 ff, 365, 376, 385, 410 ff 416, ff, 424
flux of 4-momentum in terms of 283, 310
for accelerated charge 424
in model of photon 376, 381
in terms of density and stress 297
in terms of null vectors 341, 343
inequalities for 292, 295, 312
Minkowskian 416, 417
of continuous medium 281 ff, 309 ff
of incoherent stream 288
of isolated system 309 ff
of perfect fluid 303
physical meaning of 285 ff, 310
principal directions of *see* eigen properties

symmetry of 284, 311, 417
trace of 299, 307, 325
equation of conservation 285, 310, 323, 404
of continuity 290
of motion 10, 33, 39, 166 ff
of charged particle 171, 394 ff, 400 ff, 410 ff
of continuous medium 300 ff
of particle with spin 222
of perfect fluid 303 ff
of state 305
ether 140, 145, 148, 325
Euclid 2, 75
Euclidean geometry *see* E-geometry
Euler, L. 84, 163
Eulerian angles and pseudoangles 79 ff
Euler-Lagrange equations *see* Lagrangian equations
events 5 ff
ordering of 11 ff
separation between 15 ff, 23 ff, 38 ff, 57
expanding universe 150 ff
expansion of world tube 280, 306
experiment, ideal 7
explosion 175
external impulse and impulsive torque 251 ff

F_{rs} (electromagnetic tensor) 317 ff, 387 ff, 415
field, electromagnetic *see* electromagnetic field
gravitational 5, 35, 36, 38, 53
theory 309, 387
finite separations 39 ff, 58, 84
Finsler space-time 19 ff
Finzi, B. 413, 415
FitzGerald, G. F. 119
FitzGerald-Lorentz contraction 118 ff
Fizeau, A. H. L. 146
1-flat 60
2-flat 60
principal 330, 337, 369
rotation in 79, 80
spacelike and timelike 61, 330
3-flat 59
principal 328, 337
flat space-time 34 ff, 38 ff
geometry of 59 ff
fluid, incompressible 306 ff
perfect 302 ff
flux of angular momentum 218, 285
of 4-momentum 215, 283, 286, 310
of relative energy 287, 288, 324
of relative momentum 287
force, Newtonian 163
relative 168, 169
3-force 168
4-force 167
on charged particle 394, 399 ff
on element of continuous medium 301
four-vector *see* vector
frame of reference, Galileian 52 ff, 118
principal 297, 331 ff
free particle 32 ff, 39, 58, 401
Frenkel, J. 222
frequency 133 ff, 142, 202 ff, 354
apparent and proper 152
changed by collision 196
in annihilation of matter 204
in emission 189
numerical values 182
transformation of 135, 136, 138 ff, 153
4-vector 134, 136, 202, 354
Fresnel's convection coefficient 142 ff
fundamental form and tensor xv, 16 ff, 22 ff, 32 ff, 42, 56, 69
wave function 360, 361, 366 ff
future and past 11 ff, 17 ff, 38, 40, 41, 57

g_{mn} (fundamental tensor) xv, 16 ff, 22 ff, 32 ff, 42, 56, 69
γ (coefficient in Lorentz transformation) 113, 118
galaxy, evolution of 157
Galileian frame of reference 52 ff, 118
transformation *see* Newtonian transformation
Galileo 52
Gårding, L. 360
Gardner, G. H. F. viii, 36, 239, 247
gas law 269
photonic 269 ff, 307
relativistic 265 ff
gauge invariance 349, 396
transformation 349
Gaussian units 388, 393, 400, 426
Gauss's theorem 276
general theory of relativity 35 ff, 38
geodesic 20
hypothesis 32 ff, 39
null 33, 34, 39
geometrisation of physics 24
geometry 4, 165
Euclidean and Minkowskian 65 ff, 123, 153
of electromagnetic field 344
of flat space-time 59 ff
of Lorentz transformation 76 ff
of skew-symmetric tensor 223 ff
of spinors 104
Riemannian, 19, 21, 22 ff, 165

gravitation 5, 35, 36, 38, 53
gravitino 212
γ-rays 137
Greek suffixes xv, 8
Green's theorem 276 ff, 383, 384
Griffith, B. A. 79, 395
group of Lorentz transformations 85
  speed 421
gyrocompass 54

h (Planck's constant) 154, 173, 182, 375
H_{rs} (angular momentum tensor) 216 ff
H_{rs} (electromagnetic tensor) 415
Hamiltonian methods 19 ff, 163
heat 174, 184, 266, 270
  Joulean 416, 418
Heavens, O. S. 36
Heaviside, O. 388, 394
helium 182
Herglotz, G. 36
Hermitian matrix 114 ff, 129
history of mass-centre 219 ff, 312 ff
Hoek, M. 149
Hubble, E. P. 157
Hubble's constant 156, 157
hydrodynamics *see* fluid
hydrogen 166, 182, 191, 193
hyperbolae and pseudosphere 67
hyperplane 60

ideal experiment 7
Illingworth, K. K. 158
imaginary angle 74
  components of vector or tensor 57
  time xv, 56, 69
impact of particles on target 265 ff
  parameter 248, 427
improper Lorentz transformation 97, 99
impulse, external 251 ff
  internal 210 ff, 238, 255 ff, 295
  4-vector 209, 252
impulsive torque 251 ff
incoherent stream of particles 261 ff, 288 ff, 296, 403
incompressible fluid 306 ff
indefinite form 59
indeterminacy in collision problems 174, 194, 227
  in fluid motion 305
index, refractive 142, 149
indicator 23
induction, magnetic 415
inelastic collision 174, 183 ff
inequality, polygonal 179, 204
  triangle 177 ff
inertial frame *see* Galileian frame
Infeld, L. viii, 103, 415
infinite energy 378, 399, 400
infinitesimal Lorentz transformation 85, 86
infra-red waves 137
integral of angular momentum 312, 397
  of energy 397
  of 4-momentum 311
intensity, electric and magnetic 317, 320, 388, 415
interference 46, 158 ff, 358, 359, 373
interferometer 158
internal impulse 210 ff, 238
  attractive 212, 213, 238, 255 ff, 295
  in two-body problem 255 ff
  repulsive 212, 213, 238, 295
interval *see* separation
intrinsic angular momentum 220 ff, 316
  flux of 285
  invariant of 222 ff, 227, 235, 238
  open and closed laws of 254
invariance of Maxwell's equations 317, 416
invariant of angular momentum 222 ff, 227, 235, 238
  of electromagnetic field 321, 322, 336 ff, 371
  of skew-symmetric tensor 222
isolated system 309 ff
isolation, condition of 310, 314
Ives, H. E. 120, 189

J_r (4-current) 404, 415
Jeans, J. H. 268, 269, 422
Jeffreys, B. S. and H. 73
Jehle, H. 103
Joos, G. 158
Joulean heat 416, 418
junction conditions across shock wave 385, 386

Kennedy, R. J. 158
Kepler problem 396 ff, 426
Kermack, W. O. 150
Kilmister, C. W. 19
kinetic energy 169
  numerical values 182
Kirchhoff's formula 408, 409
Kronecker delta 53
Kuerti, G. 162

laboratory, terrestrial 53, 54
Lagrangian equations 20, 33, 163, 395
Lanczos, C. viii, 314
Larmor, J. 393
Latin suffixes xv, 8

Laub, J. 146
Laue, M. von 144, 400
law of angular momentum 218, 253, 254
of 4-momentum 213 ff, 253, 284, 310
of relative energy and momentum 175, 263 ff
left-handed tetrad 97
length 23, 30, 31
Leone, F. C. 162
Lichnerowicz, A. 306
Liénard, A. 393
light 1, 5, 14, 23
frequency vector 136
in moving medium 142 ff
particle of *see* photon
ray 137
reflected by moving mirror 138 ff
speed of 54, 55, 182, 183, 352, 385
visible 137, 138, 145, 182
wave 133, 136 ff, 385
see also electromagnetic field
line element 16 ff, 22 ff, 34, 42, 56, 69
line breadth 191
linearity of Lorentz transformation 70
Lorentz, H. A. 4, 5, 24, 119, 146, 394
Lorentz contraction *see* FitzGerald-Lorentz contraction
Lorentz transformation 69 ff
applications of 118 ff
as orthogonal transformation 71
as projective transformation 102
as rigid displacement 84 ff
as rotation and translation 71, 72, 78 ff
as transformation of null rays or points on sphere 98 ff
degrees of freedom of 72
Eulerian form of 82
future-preserving 73, 99, 106, 109
generated by spin transformation 103 ff
geometrical meaning of 76 ff
group property of 85
infinitesimal 85, 86
linearity of 70
matrix of 72, 78, 82, 86, 114 ff
Maxwell's equations invariant under 317, 416
null rays unchanged by 88 ff
of electromagnetic tensor and vectors 320 ff, 333
of frequency and phase velocity 133 ff, 138 ff, 153
proper and improper 73, 97, 99, 106, 109
reduced to 4-screw 85, 86 ff
reduced to timelike rotation 112
simple 74, 107, 110 ff, 118 ff
luminosity 153, 154
distance 155

m (proper mass) 166
m^* (relative mass) 169
M_r (4-momentum) 166
μ (density) 262, 297
macroscopic density 297
particle 302
magnetic field, constant 171
induction 415
intensity (3-vector) 317, 320, 415
map of space-time 62, 121
Mariot, L. 343
mass-centre 218 ff, 312 ff, 370
reference system 180, 191, 220
in Compton effect 197
mass, proper 165 ff, 170, 172, 289
constant or variable 167, 181, 210
decreased in disintegration 180
numerical values 182
of electromagnetic particle 370, 380
of photon 173, 380
mass, relative 165, 169, 170
see also density
material energy tensor 416 ff
particle *see* particle
mathematical models 165
structure of nature 164
mathematics, pure and applied 3
Matrai, T. 36
matrix 72
canonical, of electromagnetic field 336, 352
Hermitian 114 ff, 129
of energy tensor 338, 365
of Lorentz transformation 72, 78, 82, 86, 114 ff
of 4-screw 94
orthogonal 73
singular 223, 240
skew-symmetric 223
symmetric 114, 116
matter, annihilation and creation of 199 ff, 212
Maxwell's equations in 415 ff
Maxwell's equations 1
derived from variational principle 413 ff
in moving matter 415 ff
in vacuo 317 ff
with current 403 ff
McConnell, A. J. 281
McCrea, W. H. 150, 391
McCuskey, S. W. 162
McMahon, J. 346
mean density and velocity 297

measurement of length 31
of spacelike separation 24 ff
of time 14 ff, 38, 49 ff, 119, 152
mechanics of continuum 261 ff, 309 ff
of discrete system 208 ff
of particle 163 ff
meson 176, 181
meteorite 157
metric in space-time 16 ff, 22 ff, 34, 42, 56, 69
M-geometry 65 ff, 123, 153
Michelson, A. A. 146, 158
Michelson-Morley experiment 158 ff
Miller, D. C. 158, 161
Milne, E. A. 121, 150
Milner, S. R. 345
minimum of energy density 294, 327
Minkowski, H. 4, 24, 56, 122, 416
Minkowskian coordinates 56 ff, 69 ff
energy tensor 416, 417
mirror, moving 138 ff
model, mathematical 165
of material particle 366 ff, 374 ff, 413
of photon 375 ff, 386
molecule 166
Møller, C. 16, 146, 219, 316, 416
momentum, angular *see* angular momentum
relative (3-momentum) 169, 173, 175, 212
conservation of 175
density of 263, 286 ff, 324
flux of 287
law of 263 ff
of internal impulse 212
of material particle 169, 212
of photon 173, 212
4-momentum 166
closed law of 215, 253, 310
conservation of 174 ff, 213 ff, 248, 285, 310
flux of 215, 283, 286, 310
of material particle 166, 173, 175
of photon 173, 175
of system 214 ff, 311 ff
open law of 214, 253
transmitted by internal impulse 211
momentum-energy tensor *see* energy tensor
4-vector *see* 4-momentum
Morley, F. W. 146, 158
motion in constant magnetic field 171
in Coulomb field 396 ff, 426
of free particle 32 ff, 39, 58, 401
of photon 23, 33, 34, 39, 58
relative, of frames of reference 113 ff, 118 ff
rigid 36, 37
under constant relative force 171
see also equation of motion, rest, velocity, etc.
moving medium 142 ff, 415 ff
see also continuum
Murnaghan, F. D. 103

ν (frequency) 133
nebulae 150, 153
negative energy 212
photon 212
neutron 181, 185
Newtonian angular momentum 218
collision 173, 184, 227
hydrodynamics 305, 308
mass-centre 220
mechanics, laws of 163, 164, 165, 169, 255
physics, concepts taken from 2 ff
time 3, 13, 16
transformation 114
Noether, F. 36
non-Euclidean geometry 165
normal to 3-wave 420
normalisation of null vectors 95
of 4-potential 349
notation xv, 8
nucleus, disintegration of 176, 181
scattering and capture by 426
null cone 13, 17, 23, 57, 386
absolute 2-content of 3-cell 405, 430
and spinors 103
cut by 2-flat 61, 330
diagram of 41, 65, 137
finite 40
3-wave tangent to 137
null directions, principal 328 ff, 352, 376
electromagnetic field in terms of 339 ff
null field, electromagnetic 322, 327, 332, 335, 337, 341 ff, 352, 375
geodesic 33, 34, 39
point on 3-space 383
rays, axial, of 4-screw 88 ff
triad of 94 ff
3-space 383
vector 19, 23
determined by spin triangle 104
orthogonal to itself 29
vectors, normalisation of 95
scalar product of 89
sum of 177, 202
world tube 275, 276
numerical density 261, 267, 270, 282, 289, 403
values 182
Ω (intrinsic angular momentum invariant) 222

oblique axes in space-time diagram 124
observer 11
  Galileian 49 ff
open law of angular momentum 218, 253, 254
  of 4-momentum 214, 253
operational method 7
optics *see* light
  electron 395
orbital angular momentum 221, 397
ordering of events 11 ff
orientation of tetrad 73, 97
orthogonal 2-flats 60, 86, 88, 330
  matrix 73
  projection 61 ff, 274
  tetrad 76 ff
    and triad of null rays 94 ff
    orientation of 73, 97
  transformation 71
orthogonality of principal directions 291
  of spacelike displacements 27
  of vectors 26 ff, 41, 58, 65, 66, 76
  physical meaning of 26 ff

p (pressure) 302
$P\rho$ (relative 3-force) 168
parallel plane mirrors 141
  straight lines 44 ff
particle 5, 213
  annihilation and creation of 199 ff
  charged *see* charged particle
  collision of, with photon 193 ff, 249
  disintegration of 173 ff, 180 ff, 185 ff
  electromagnetic model of 366 ff, 374 ff, 413
  free 32 ff, 39, 58, 401
  history confined by null cone 11 ff
  intrinsic angular momentum of 220 ff, 316
  macroscopic 302
  mass of 165 ff
  material 5, 238
  mechanics of 163 ff
  4-momentum of 166, 173, 175
  of light *see* photon
  ultimate 181
particles, colliding with target 265 ff
  elastic collision of 174, 205 ff, 227 ff 248
  incoherent stream of 261 ff, 288 ff, 296, 403
  inelastic collision of 174, 183 ff
past, present and future 11 ff, 17 ff, 38, 40, 41, 57
Pauli, W. 4, 146, 400, 416
Pearson, F. 158
Pease, F. G. 158
pendulum 35, 54
perfect fluid 302 ff
permeability 416
permutation symbol 73, 222, 239, 317, 364
phase function 350
  plane 134
  vector 372
  velocity 133, 135, 421
philosophy of physics 1 ff, 163 ff, 208
photon 5
  angular momentum of 218, 238, 249 ff
  collision of, with material particle 193 ff, 249
  electromagnetic model of 375 ff, 386
  emission and absorption of 187 ff, 191 ff
  in expanding universe 154
  4-momentum of 173, 175
  negative 212
  numerical values for 182
  proper mass of 173, 380
  relative energy and momentum 173, 212
  world line of 23, 33, 34, 39, 58
photons colliding with target 269 ff
  elastic collision of two 250
  gas of 269 ff, 307
  in annihilation and creation of matter 199 ff
  Lorentz invariant of two 201
  mass-centre reference system for two 191
  related to electromagnetic null field 343
  sameness of 137, 189
  stream of 269 ff, 343
physical meaning of coordinates 7, 47 ff
  of energy tensor 285 ff, 310
physicisation of geometry 24
physics, philosophy of 1 ff, 163 ff, 208
Pirani, F. A. E. viii, 212, 316
pitch of electromagnetic wrench 334
Planck's constant 154, 173, 182, 375
plane in space-time 60
  waves 133 ff, 350 ff
Poincaré, H. 4
point-event 5
points of space 52
polar coordinates in space-time 46
polarisation 190
polygonal inequality 179, 204
ponderomotive force 394
positive-definite form 59
positive nature of energy-density 292, 306, 313
positron-electron pair 202
potential 4-vector (4-potential) 345 ff, 405 ff
  advanced 390, 393, 408

due to 4-current 405 ff
due to moving charge 389 ff, 422
gauge transformation of 349
in variational principle 395, 414
obtained from scalar wave function 363, 364
retarded 390, 393, 408, 432
Pounder, J. R. 36
Poynting vector 324
present, past and future 11 ff, 17 ff, 38, 40, 41, 57
pressure due to impact 265 ff, 269 ff
in fluid 302
in relativistic gas 267 ff, 271
positive character of 306
Preston, T. 149
Priestley, J. 388
principal directions of electromagnetic tensor 336 ff
of energy tensor 290 ff, 325 ff
orthogonality of 291
principal 2-flat 330, 337, 369
3-flat 328, 337
frame of reference for continuous medium 297
null directions 328 ff, 352, 376
tetrad of energy tensor 292
principle, variational *see* variational principle
product, scalar *see* scalar product
projection, orthogonal 61 ff, 274
projective transformation 102
propagation, sense of 215
velocity of 133
proper density 261, 262, 289, 403
energy 170, 262
Lorentz transformation 73, 97, 99, 106, 109
mass *see* mass
time 54
proton 166, 181, 182, 185
pseudoangle 29, 74, 126
Eulerian 79 ff
in atomic disintegration 185
in composition of velocities 127
pseudosphere 67
pseudotrigonometry 186
Pythagoras, theorem of 28

Q (energy released) 181
quadratic form, fundamental xv, 16 ff, 22 ff, 34, 42, 56, 69
quantum conditions 187, 191
theory 34, 57, 175
see also photon

Raabe, A. 222
radar waves 182
radiation 137
cosmic 176
4-force 401
of energy from accelerated charge 393, 399 ff, 422
radio waves 137, 138, 182
Rainich, G. Y. 345
rank of skew-symmetric matrix 223
rational electrostatic units 388
ray, null 88 ff, 94 ff
of light 137
Rayner, C. B. 35
reaction 163, 255
reality of eigen values 292 ff, 338
recession of nebulae 150
red-shift, cosmological 150 ff
reference, frame of 52 ff, 118, 297, 331 ff
system of mass-centre 180, 191, 197, 220
reflection at moving mirror 138 ff
at fixed target 266, 270
refractive index 142, 149
relative 3-current 404, 415
density 261 ff, 404
energy 169, 173, 212
law of 175, 263 ff
numerical values 182
of photon 173
released in disintegration 180
force 168, 169
kinetic energy 169
mass 165, 169, 170
momentum *see* momentum
velocity 113 ff, 118 ff, 126 ff, 132
relativistic gas 265 ff
repulsive internal impulse 212, 213, 238, 295
rest and motion 53
retardation of moving clock 118 ff, 122 ff, 189
retarded potential 390, 393, 408, 432
Riemannian geometry 19, 21, 22 ff, 165
rigid body 36, 161
displacements and Lorentz transformations 84 ff
rod 30 ff
motion 36, 37
right angle 29
right-handed tetrad 97
Risco, M. 191
Robertson, H. P. 146
rod, moving, contraction of 118 ff
rigid 30 ff
rot 317
rotation in space-time 71, 78 ff
commutation of 87
spacelike 79
timelike 80, 112
run-away motion 402
Ruse, H. S. 103, 239, 324, 343, 345

Russell, H. N. 149
Rutherford-Bohr atom 399

Salzman, G. 36
scalar disturbance 133 ff
  product 26
    of null vectors 89
    of timelike vectors 27, 46, 178
  wave function 360 ff
scattering by fixed nucleus 426 ff
  see also collision
Schild, A. 8, 17, 276
Schrödinger, E. VIII, 399
Schwarz inequality 27
4-screw 85 ff
  axial null rays of 88 ff
  matrix of 94
  reduction of Lorentz transformation to 85, 86 ff
self-conjugate field 322
Sen, N. R. 307, 345
separation between events 15 ff, 23 ff, 38 ff, 57
  finite 39 ff, 58, 84
  Finslerian 19
  measurement of 24 ff
Shankland, R. S. 162
Sharpless, S. 157
sheet, electromagnetic 355, 391
shock wave 356, 385
sidereal year 157
signal 421
signature of fundamental form XV, 17
Silberstein, L. VIII, 146, 158
simultaneity not absolute 118, 152, 163, 173
singular matrix 223, 240
sinusoidal wave 133 ff, 354, 356 ff
six-vector 216, 224
skew straight lines 44 ff
skew-symmetric tensor, canonical forms for 331 ff, 336
  dual of 222, 239
  equivalent to pair of 3-vectors 224
  expressed by two 4-potentials 414
  geometrical representation of 223 ff
  invariants of 222
  null 322
  singular 223, 240
  spinors connected with 103
  see also angular momentum, electromagnetic tensor
slowing-down of clock *see* retardation
Smithsonian Tables 183
snapshot 120 ff
  of aberration 148
  of annihilation of particle 201
  of clock-retardation 122
  of collision 176, 206
  of Compton effect 195, 198, 199
  of expanding universe 151
  of FitzGerald-Lorentz contraction 121
  of intrinsic angular momentum 226, 235
  of light in moving medium 143
  of moving mirror 139
  of relative velocity 127
sodium 182
sound 133, 308, 309
space and time 49 ff
3-space, null 383
spacelike displacements, orthogonality of 27
  2-flat 61
  geodesic 39
  rotation 79
  separation, measurement of 24 ff
  vector 19, 23, 27
space-time 6 ff
  coordinates 7 ff, 38, 47 ff
space-time diagram 9, 63 ff
  of annihilation and creation 200, 202
  of axes of coordinates 123, 124
  of axial null rays 88
  of canonical axes 332, 333
  of cells and tubes 273, 274, 277, 281, 282
  of collision 9, 176, 205
  of composition of velocities 128
  of Compton effect 194
  of conservation law 176
  of contraction and retardation 124
  of disintegration 179
  of distance between world lines 29, 31, 45
  of emission of photon 188
  of Eulerian angles and pseudo-angles 83
  of events 64
  of expanding universe 150
  of frequency 4-vectors 138, 190
  of Galileian coordinates 48, 50
  of impulse 209, 210, 213
  of isolated system 310
  of light in moving medium 143
  of Michelson-Morley experiment 158
  of 4-momentum and angular momentum 215, 226, 253, 254, 282
  of null cone 12, 41, 65, 137, 406
  of orthogonal displacements or vectors 27, 28, 66
  of orthogonal projection 62, 274
  of past, present and future 12, 41
  of polygonal inequality 179

of 4-potential, advanced and retarded 389, 391, 422, 432
of principal directions 296, 328. 330
of pseudosphere 67, 430
of red-shift 153
of simple Lorentz transformation 111
of spacelike separation 25
of straight line 44
of triad of null rays 98
of triangle inequality 178
of two-body problem 256
of 4-velocity 131, 132
of 3-wave 137
space-time, Finsler 19 ff
flat 34 ff, 38 ff, 59 ff
of constant curvature 35
Riemannian 19, 21, 22 ff, 165
split into space and time 49 ff
special theory of relativity 4, 34, 38 ff
spectra of nebulae 150, 152
speed, group 421
of light 54, 55, 182, 183, 352, 385
of sound 308, 309
see also velocity
spin angular momentum 221
invariant 223
shape 108
tensor 223
transformation 105 ff
triangle 104
spinor 103 ff
splitting space-time into space and time 49 ff
spur *see* trace
standard clock 14 ff, 38, 119, 152
standing waves 141, 191, 358
stars, evolution of 157
state, equation of 305
statics 163
stationary principle *see* variational principle
Stephenson, G. 19
Stevenson, A. F. 346
Stewart, J. Q. 149
Stilwell, G. R. 120, 189
Storchi, E. 415
straight line 39, 42 ff, 60, 64
stream of particles or photons 261 ff, 267 ff, 288 ff, 343, 403
stress, electromagnetic 324, 325
in continuous medium 296 ff, 416
-momentum-energy tensor 287
see energy tensor
Struve, O. 157
summation convention xv, 8
superposition of plane waves 357 ff
of wave functions 372 ff
symmetry of energy tensor 284, 311, 417
of fundamental tensor 17
of matrix of Lorentz transformation 114, 116
Synge, J. L. 8, 17, 21, 35, 36, 79, 177, 197, 210, 219, 260, 276, 306, 330, 345, 383, 395, 396, 421
system, angular momentum of 217 ff, 251 ff, 312
centroid of 315
closed *see* isolated system
continuous 208, 309 ff
discrete 208 ff
field theory of 309
isolated 309 ff
mass-centre of 218 ff, 312 ff, 370
mechanics of 208 ff, 261 ff, 309 ff
4-momentum of 214 ff, 311 ff
with intrinsic angular momenta 251 ff, 281

T (relative kinetic energy) 169, 182
T_{rs} *see* energy tensor
target in relativistic gas 265 ff
Taub, A. H. 36
telescope filled with water 149
temperature 265, 266, 269, 271, 305
Temple, G. 35
tensor, angular momentum 216 ff, 221 ff
Cartesian 57
electromagnetic 317 ff, 387 ff, 415
energy *see* energy tensor
fundamental xv, 16 ff, 22 ff, 32 ff, 42, 56, 69
real and imaginary components of 57
skew-symmetric *see* skew-symmetric tensor
terrestrial laboratory 53, 54
tetrad, orthogonal 73, 76 ff, 94 ff, 97
principal 292
Thackeray, A. D. 157
three-vector *see* vector
time 11 ff
and space 49 ff
imaginary xv, 56, 69
Newtonian 3, 13, 16
proper 54
timelike 2-flat 61, 330
rotation 80, 112
separation *see* separation
vector 19, 23, 177
vectors, scalar product of 27, 46, 178
Tolman, R. C. 157
torque, impulsive 251 ff
trace of energy tensor 299, 307, 325

transformation of coordinates, general 8
two interpretations of 75
see also Lorentz transformation, spin transformation
translation of space-time 71
transparent medium 142 ff
triad of null rays 94 ff
of points on unit sphere 100
triangle inequality 177 ff
trivial solutions in collision problem 229, 233 ff, 247 ff
tube of world lines *see* world tube
two-body problem 254 ff

Udeschini, P. 415
ultimate particle 181
ultra-violet light 137
unimodular condition 105
unit vector 29, 46, 58, 178
units 54, 55
absolute 375
electrostatic 388
Gaussian 388, 393, 400, 426
universe, age of 156, 157
expanding 150 ff
uranium 182

variational principle for geodesic 20, 32 ff, 39
for Maxwell's equations 413 ff
for motion of charged particle 395
Veblen, O. 75
3-vector, angular momentum 225
electromagnetic 317, 320, 388, 415
relative force 168
relative momentum 169, 173, 175 212
relative velocity 129 ff
4-vector, acceleration 130 ff, 301, 423
coordinate 78
force 167
frequency 134, 136, 202, 354
null 19, 23
orthogonality 26 ff, 41, 58, 65, 66, 76
principal *see* principal direction
real and imaginary components of 57
spacelike and timelike 19, 23, 27, 177
unit 29, 46, 58, 178
velocity 130
velocity, composition of 126 ff
group 421
mean, in continuum 297
of light 54, 55, 182, 183, 352, 385
of recession of nebulae 151
of sound 308, 309
of wave 133 ff, 385, 419 ff
relative 113 ff, 118 ff, 126 ff, 132
3-vector 129 ff
4-vector 130
3-volume 272
contracted in moving system 119
of cross-section of world tube 273 ff
of element of null cone 431
relative and proper 261
4-volume 272
vorticity 306

wave, de Broglie 136, 197, 396, 421
electromagnetic 137, 350 ff, 385
equation 348, 360 ff
function 360 ff
elementary complex 361, 366 ff
superposition 372 ff
yielding Maxwellian field 363 ff
length, Compton 197, 204, 375
in moving medium 142 ff
numerical values 182
light 133, 136 ff, 385
shock 356, 385
sound 133
standing 141, 191, 358
transformation of 133 ff
velocity 133 ff, 385, 419 ff
2-wave and 3-wave 134 ff, 419
Wesselink, A. J. 157
Weyl, H. 103
Weyssenhoff, J. 222
Whitehead, A. N. 35
Whitrow, G. J. 157
Whittaker, Sir Edmund 4, 79, 103, 170, 345, 388, 394
Wiechert, E. 393
world line 9
of free particle 32 ff, 39, 58, 401
of photon 23, 33, 34, 39, 58
world lines, distance between 29 ff, 45
map 121
tube 272 ff
expansion of 280, 306
null 275, 276
wrench, electromagnetic 334, 335, 338

X_r (4-force) 167
X-rays 137, 182

Y_r (4-impulse) 209

Zeeman, P. 146