MATSCIENCE REPORT 84
1 st JULY, 1975

AN INTRODUCTION TO VECTORS TENSORS AND RELATIVITY

K.H. MARIWALLA

THE INSTITUTE OF MATHEMATICAL SCIENCES, MADRAS-600 020, (INDIA)

MATSCIENCE REPORT 84

AN INTRODUCTION TO VECTORS TENSORS AND RELATIVITY

K.H. Mariwalla
MATSCIENCE, Madras-600020

THE INSTITUTE OF MATHEMATICAL SCIENCES
MADRAS-600 020

C O N T E N T S

FOREWARD

These notes are based on lectures* given by the author to the mathematics teachers of Madras University in connection with the prescribed syllabus for M.Sc. in respect of tensors and relativity. The concepts involving vectors ~~are~~ and tensors are common to three prescribed courses, viz. Modern Algebra, Mechanics and Relativity. The principal books recommended by the University were Fundamental Structures of Algebra by G.D.Mostow, J.H.Sampson, J.P.Meyer, Differential Geometry by Stouik, Introduction to Riemanian Geometry and Tensor Calculus by C.E.Weatherburn, Theory of Relativity by A.S.Eddington, Between Mostow et. al. and other books recommended, t~~T~~here is a wide gap in the approach to concepts of vectors and tensors. In ths present notes an attempt has been made to bridge the gap, since also the current literatures in physics requires increasing awareness of more general stand-point as presented in Mostow et. al. It must be emphasised that no attempt is made to develop the subject of tensor analysis ri~~ve~~gorously from a modern standp~~o~~oint. The object of these notes is merely to expose the reader to the modern terminology which enables one to have a clearer, more unified picture of the various concepts that arise in the discussion. The reader is recommended to consult the following books for further pursuing the subject of relativity in the spirit of the present introduction to the subject.

1. LECTURES ON GENERAL RELATIVITY, Brandeis Summer Institute in Theoretical Physics, Vol. 1, 1964, A.Trautman, F.A.E.Pirani, H.Bondi (Prentice Hall, New Jersey)

*The lectures were given in May 1973 at two centres: Matscience (to the Government College Teachers, under Inservice Training Programme) and Loyola College (Under the College Science Improvement Programme)

2. GRAVITATION (1973), C.W.Misner, K.S.Thorne, J.A.Wheeler (W.H.Freeman, San Francisco).

3. The Large Scale Structure of Space-time, S.W.Hawking and G.F.R.Ellis, Cambridge University Press (1973).

On the subject of special theory of relativity the following references may be found useful.

1. The theory of Relativity, by R.K.Pathria (Hindustan Publishing Corp. - Delhi, 1963), particulary for historical introduction.

2. The Special Theory of Relativity by J.Aharoni (2nd Ed., Oxford, 1965).

The plan of these notes is as follows: First four Chapters are devoted to introducing the concepts on vectors, matrices and tensors. In Chapter V Euclidean Spaces are discussed. It is useful to emphasise that whereas in a vector space all parallel vectors of the same length and same orientation are identified in a Euclidean space the basic notion is that of a vector together with its starting point. Consequently what was taken as a basis in a vector space is reinterpreted as 'frame vectors' and refer to a coordinate frame. This interpretation becomes particularly important when one introduces general coordinates in a Euclidean space or when one considers more general spaces; in this case the frame vectors are actually 'vector fields'. The basis is now given by $\partial/\partial x^i$ (and the coordinate differentials dx^i in the dual space) in a local coordinate system. This point is explained in sections (V.6, VI.1,2), and in Section VI.3 the general curved spaces involving a linear connection are discussed in this spirit. In section VI.6 the frame vector fields are emphasised and Riemannian spaces are discussed from this viewpoint. The discussion in sections V.9,10, which is also applicable mutatis mutandis to the general Riemannian spaces, is strongly coordinate dependent and is in the spirit of the formulation given in Weatherburn. It is hoped that these various treatments would enable the postgraduate teachers to broaden their outlook on the subject and thus further their understanding of the basic notions involved. Furthermore it should also help those who are interested in pursuing research to understand current literature in modern theoretical physics and Relativity. The last two chapters on special and general relativity, are developed in the same spirit.

Equations in each chapter are numbered afresh starting from 1. Except in chapters V and VI, the chapter number is not indicated on the equations.

MATSCIENCE
Madras-600 020
Date: June 1975.

K.H.Mariwalla

1. Some Definitions.

I.1. Let S denote a set of element $\{x, y, ---\}$. Let θ denote a Binary operation for every pair x, y of the set S such that $x \theta y$ is again an element of the set S (closure property). θ is said to be the internal operation of the set.

I.2. DEFINITION. The set S together with the internal operation is called a Groupoid.

I.3. DEFINITION. A groupoid with the following properties is called a group.

(a) the operation θ is associative

(b) there exists a 'neutral element' E such that for each X in S, $E \theta X = X \theta E = X$. E is also called the identity element.

(c) To every X in S there corresponds an inverse element (usually denoted) X^{-1} in S such that

$$X \theta X^{-1} = X^{-1} \theta X = E$$

I.4. DEFINITION. A group in which the operation θ is such that $X \theta Y = Y \theta x$ for all pairs x, y in S is called an Abelian group. The operation θ is said to be commutative; it is often denoted by +, and the neutral element by 0.

I.5. Other related notions that we might have occasion to use are Subgroup (subset of a group satisfying group axioms), representation of a group, discrete and continuous groups. We shall explain these notions in the relevant context.

I.6. One can define more than one type of internal binary operatio on a set. Such a set S together with two binary operations θ_1, θ_2 will be denoted by $S_{\theta_1 \theta_2}$

DEFINITION. A set $S_{\theta_1\theta_2} = S_{+\cdot}$ is said to be a Field if

(a) S_+ is an abelian group, 0 denotes the neutral element

(b) $S_\cdot \equiv \{S_+ - 0\}$. is a group, i.e. the set S from which the neutral element of is excluded is a group with respect to the operation

(c) If X, Y, Z are any elements of S, then the following holds: $(X+Y)Z = XZ + YZ$, and $X(Y+Z) = XY + XZ$.

I.7. DEFINITION. A field is said to be commutative if $S_\cdot$ is an abelian group. Well known examples of commutative fields are the fields of real and complex numbers. An example of a non-commutative field is the field of quarternions. The elements of these fields may be represented as linear combinations of 2 x 2 Pauli matrices. Thus

$$X = \alpha_0 I + i \sum_k \alpha_k \sigma_k$$

where I is the identity matrix, α_k are real numbers and $\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are Pauli matrices. In what follows we shall refer to real and complex numbers as scalars. The symbol of a general fiel will be K and for real numbers R. In this manner an n-tupple of fields will be denoted by $R^n (R_n)$ or $K^n (K_n$

II. VECTOR SPACE

II.1. Let us consider a system consisting of two sets together with the internal operations and inter-set operations (external operation). Let $K\{a, b, \cdots\}$ and $V\{\vec{x}, \vec{y}, \cdots\}$ be sets such that

(a) $K_{+\cdot}$ is a field; 0 is the neutral element in $+$ operation and 1 the identity element in the dot operation

(b) V_{+} is an abelian group with respect to $+$ operation and 0 is its neutral element

(c) there exists an interset operation called scalar multiplication obeying the following rules

$$\left.\begin{array}{ll} \text{(i)} & a(b\vec{x}) = a\cdot b(\vec{x}) \\ \text{(ii)} & (a+b)\vec{x} = a\vec{x} + b\vec{x} \\ \text{(iii)} & 1\vec{x} = \vec{x} \\ \text{(iv)} & a(\vec{x}+\vec{y}) = a\vec{x} + a\vec{y} \\ \text{(v)} & a\vec{x} = \vec{0},\ a \neq 0 \text{ implies } \vec{x} = \vec{0} \end{array}\right\} \begin{array}{c} \text{for all} \\ a, b \text{ in } K \\ \text{and} \\ x, y \text{ in } V \end{array}$$

DEFINITION. The set V together with the set K and with these properties is called a vector space or linear space over the field K. $\vec{x}, \vec{y}$ are called vectors and a, b scalars.

In what follows we shall generally take $K = R$.

II.2. Let $\vec{u}_1, \vec{u}_2, \ldots$ be elements of V, then they are linearly dependent if there exist scalars $a_1, a_2 \cdots a_r$ not all zero such that

$$a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_r\vec{u}_r = 0 . \tag{II.1}$$

If (1) is true only if all a_k are zero, then $u_1, u_2, \ldots, u_r$ are said to be linearly independent.

DEFINITION. A vector space V is said to be of dimension n if there exists a maximal set of n linearly independent vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ such that any vector $\vec{X}$ in V can be written as a linear combination of these vectors $\{\vec{e}_1, \ldots, \vec{e}_n\}$ with coefficients in K:

$$\vec{X} = x^j \vec{e}_j \tag{II.2}$$

where, and in what follows we shall assume summation whenever there is one upper and one lower index repeated (Einstein summation convention).

DEFINITION. An ordered pair of n linearly independent vectors is called a basis of the vector space and the n-tupple of scalars $\{x^j\}$ as in eqn. (2) are called the coordinates or components of the vector $\vec{X}$ in the basis $\{\vec{e}_j\}$.

II.3. In the above, by choosing a basis in V_n we have represented an arbitrary vector in V_n in terms of n-tuples of scalars. Since vectors are elements of V_n and scalars elements of K one expresses this by saying that one has mapped from V_n to K^n. If we denoted this map by T, then

$$T(\vec{X}) = (x^1, x^2, \ldots, x^n) \tag{II.3}$$

$$T(\vec{X}+\vec{Y}) = T(\vec{X}) + T(\vec{Y}) \tag{II.4}$$

$$= (x^1+y^1, x^2+y^2, \ldots, x^n+y^n)$$

$$T(a\vec{X}) = a\,T(\vec{X}) \tag{II.5}$$

i.e. there is a one to one correspondence that preserves vector operations. Thus K' behaves like a vector space of dimension n and is said to be isomorphic to V_n. (In fact all vector spaces of the same dimension are isomorphic to each other).

II.4. We have considered the linear map T from V_n to K^n. We now consider a linear map from V^n to K defined by

$$f(a\vec{X}+b\vec{Y}) = a f(\vec{X}) + b f(\vec{Y})$$

where $f(\vec{X}), f(\vec{Y})$ are elements of K and X, Y any two elements of V_n. f is called a linear form or linear function (or one-form) on V_n. The set of all linear forms $\{f, g, \cdots\}$ from V_n to K, with the following properties

$$\begin{aligned} (f+g)(\vec{X}) &= f(\vec{X}) + g(\vec{X}) \\ (af)(\vec{X}) &= a f(\vec{X}) \\ f(\vec{X}) = 0 \text{ for all } \vec{X} &\text{ in } V_n \text{ implies } f = 0 \end{aligned} \tag{II.7}$$

Clearly forms a vector space of dimension n. It is called the dual space of V_n and is denoted by V_n^*. If $\vec{e}_j$ is a basis in V_n than a linear form is completely determined by n scalars

$$f_j \stackrel{\text{def.}}{=} f(\vec{e}_j). \tag{II.8}$$

Conversely, a set of n scalars uniquely determines a linear form on V_n in a given basis of V_n.

Let us define n linear forms e^k as follows

$$e^k(\vec{e}_j) = \delta^k_{\ j} = \begin{cases} 0 & k \neq j \\ 1 & k = j \end{cases} \tag{II.9}$$

The symbol δ^k_j is called the Kronecker δ-symbol. For any f in V_n^* and $\vec{X} = X^j \vec{e}_j$ in V_n, it follows from (II.5) and (II.9) that

$$\left.\begin{aligned} e^i(\vec{X}) &= X^i \\ f(\vec{X}) = X^j f_j &= f_j e^j(\vec{X}) \end{aligned}\right\} \tag{II.10}$$

Since this is true for any $\vec{X}$ it further follows that

$$f = f_j e^j ; \tag{II.11}$$

the scalars $\{f_j\}$ are said to be the components of f in the basis $\{e^j\}$ of V_n^*. The basis $\{e^j\}$ is the dual basis of $\{\vec{e}_j\}$.

II.5. Let $\{\vec{e}_j\}, \{\vec{e}_{j'}\}$ be two basis in V_n. These must be linearly related.

$$\vec{e}_{j'} = A_{j'}{}^k \vec{e}_k \ , \ \vec{e}_k = A_k{}^{j'} \vec{e}_{j'} \tag{II.12}$$

For reasons of consistancy we must have

$$\left.\begin{aligned} A_{j'}{}^k A_k{}^{l'} &= \delta_{j'}{}^{l'} \\ A_k{}^{l'} A_{l'}{}^m &= \delta_k{}^m \end{aligned}\right\} \tag{II.13}$$

If $X^j, X^{j'}$ are the components of $\vec{X}$ in the two basis t then one must have

$$\vec{X} = \vec{e}_j X^j = \vec{e}_{l'} X^{l'} . \tag{II.14}$$

On substitution from (II.12) and (II.13) we find that

$$x^{j'} = A^{j}{}_{k} x^{k} \; ; \; x^{k} = A^{k}{}_{j'} x^{j'} . \tag{II.15}$$

Let $e^{l}, e^{l'}$ be the basis in V_n^* corresponding to $\vec{e}_l, \vec{e}_{l'}$ in V_n. Consider

$$e^{l'}(\vec{e}_j) = e^{l'}(A_j{}^{m'} \vec{e}_{m'}) = A_j{}^{m'} \delta^{l'}_{m'} \tag{II.16}$$

Since also $A^{l'}{}_{j} = A^{l'}{}_{k} e^{k}(\vec{e}_j)$, it follows that

$$e^{l'} = A^{l'}{}_{k} e^{k} . \tag{II.17}$$

Now, a vector in V_n^* must satisfy

$$f = f_j e^{j} = f_{j'} e^{j'} \tag{II.18}$$

so that the components of a one-form transform under the change of basis according to the formula

$$f_{l'} = f_k A^{k}{}_{l'} \tag{II.19}$$

II.6. We have found that V and V* have same dimension and are therefore isomorphic. But this does not mean that one can just abolish V*. The trouble can be traced to equation (II.9), which shows that the basis e^{k} and $\vec{e}_k$ transform in an inverse fashion so that the two can not be identified except in a special choice of basis (canonical basis). On the other hand one can show that V**, the vector space of linear forms on V* can be put into correspondence with V independent of the choice of a particular basis chosen. Thus there is a natural isomorphism between V and V**, and the two may be taken as identical.

III. MATRICES AND DETERMINANTS

III.1 In (II.5) we have considered transformations from one basis to the other in terms of n^2 scalars. According to the notation used there we denoted both the transformations for $\vec{e}_j \to \vec{e}_{j'}$ and $\vec{e}_{j'} \to \vec{e}_j$ by the same Kernel letter A. The difference between the nature of two transformations was indicated in terms of the place of primed and unprimed indices. An alternate notation is to use different kernel letters and if the use of primes is necessary then put these on kernel letters and not on the indices. Instead of putting primes on kernel letters and could also use different kernel letters. Thus the equations (15),(1[illegible]) and (13) of section II may be rewritten as

$$X'^j = A^j{}_k X^k \quad , \quad X^k = B^k{}_j X'^j \tag{1}$$

$$f'_l = f_k B^k{}_l \quad , \quad f_k = f'_l A^l{}_k \tag{2}$$

$$A^l{}_m B^m{}_k = \delta^l{}_k \quad , \quad B^m{}_k A^k{}_l = \delta^m{}_l \tag{3}$$

One can arrange the n^2 scalars $A^j{}_k$ in a square array called a "square matrix"

$$A^j{}_k = \begin{array}{c} j \\ \downarrow \end{array} \; k \rightarrow \begin{bmatrix} A^1{}_1 & A^1{}_2 & \cdots & A^1{}_n \\ \vdots & \vdots & & \vdots \\ A^n{}_1 & A^n{}_2 & \cdots & A^n{}_n \end{bmatrix} \tag{4}$$

and similarly for $B^j{}_k$. The letters j, l, k are called indices of the matrix, labelling its components (which are scalars) in a given basis. The matrix itself may just be written as A or B. The matrix components $\delta^m{}_j$ in (3) then correspond to the unit matrix denoted by I:

$$\begin{bmatrix} 1 & & & \\ & 1 & & 0 \\ & & \ddots & \\ & 0 & & 1 \end{bmatrix} \tag{5}$$

The equations (3) then collectively read

$$AB = BA = I \tag{6}$$

I follows that B is a matrix inverse to the matrix A and one writes $B = A^{-1}$. One may verify that the arrays thus introduced are required to satisfy the following rules of addition and multiplication for any $n \times n$ matrices A,B,C.

a) $C = A + B = B + A$; $C^l{}_j = A^l{}_j + B^l{}_j$. (7)

b) $A(BC) = (AB)C$: $A^l{}_j (B^j{}_m C^m{}_k) = (A^l{}_j B^j{}_m) C^m{}_k$ (8)

c) $A(B+C) = AB + AC$; $(B+C)A = BA + CA$ (9)

The null matrix, with all elements zero may be defined as the neutral element with respect to addition of matrices. It is clear however that given an arbitrary matrix M there does not necessarily exist a matrix inverse to it. In this important respect the set of all matrices differs from a field, and such a system is called a Ring.

III.2 : One can also introduce rectangular matrices by considering a mapping between two vector spaces of different dimensions but defined over the same field . These have the form

$$C^{\ell}{}_{\alpha} = {\scriptstyle\ell}\!\downarrow \alpha\!\rightarrow = \begin{bmatrix} c^1{}_1 & c^1{}_2 & \cdots & c^1{}_n \\ c^2{}_1 & \cdots & \cdots & c^2{}_n \\ \vdots & & & \vdots \\ c^m{}_1 & \cdots & \cdots & c^m{}_n \end{bmatrix} \tag{10}$$

such an arry is called $m \times n$ matrix $(\ell = 1 \cdots m, \alpha = 1 \cdots n)$. Of particular inter st are matrices in which either m or n is equal to 1. Thus $C^1{}_\alpha$ is a $1 \times n$ row matrix and $C^\ell{}_1$ is a $m \times 1$ colomn matrix. For instance the n-tripple of scalars $\{x^j\} \equiv \{x\}$ and $\{f_j\} \equiv [f]$ are respectively $n \times 1$ and $1 \times n$ matrices

$$\{x\}^j\!\downarrow = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix}, \quad [f\,]_j = \begin{bmatrix} f_1 & f_2 & \cdots & f_n \end{bmatrix}. \tag{11}$$

This notation is consistent with the rules for matrix multiplication, and the transform tions (1) and (2) now read

$$\{x'\} = A\{x\} \qquad \{x\} = A^{-1}\{x'\}$$

$$[f'] = [f]A^{-1} \qquad [f] = [f']A. \tag{12}$$

Also we see that

$$f(\vec{X}) = (f, X) = f_j X^j = \boxed{f_1 \cdots f_n} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [f]\{X\} \qquad (13)$$

It further follows that the "product" of (colomn matrix)$\times$(row matrix) is a rectangular matrix. Thus we get the square matrix

$$\begin{bmatrix} X^1 \\ X^2 \\ \vdots \\ X^n \end{bmatrix} \boxed{f_1 \; f_2 \; \cdots \; f_n} = \begin{bmatrix} x^1 f_1 & & X^1 f_n \\ X^2 f_1 & & \\ & & \\ X^n f_1 & \cdots & X^n f_n \end{bmatrix} \qquad (14)$$

III.3: Since the basis of V_n and V_n^* satisfy

$$e^k(\vec{e}_j) = \delta^k{}_j ,$$

it is easy to see that it is possible to choose the basis such that $\vec{e}_k$ are colomn vectors and e^j row vectors in conformity with the rules for the multiplication of row and colomn matrices. The simplest form of such a representation of the basis vectors is called canonical and is given by

$$(\vec{e}_k)^j = j\downarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{jth place } j=k \quad ; \quad (e^k)_j = j \rightarrow \boxed{0 \cdots\cdots 1 \cdots 0 \, 0} , \quad \text{jth place } j=k \qquad (15)$$

where j indicates the component of the colomn and row vectors.

According to this scheme, we introduce the following notation

$$\text{ket}\,`k' = |k\rangle = (\vec{e}_k)^j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} k = j\text{th place}$$

$$\text{Bra}\,`k' = \langle k| = (e^k)_j$$
$$= \underset{\uparrow\ k = j\text{th place}}{\boxed{0 \cdots 1 \cdots 0}} \tag{15A}$$

Then the various important formulae read

$$|x\rangle \stackrel{\text{def}}{=} \text{ket}\,`x' = \sum_j x^j |j\rangle = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \tag{16A}$$

$$\langle f| \stackrel{\text{def}}{=} \text{Bra}\,`f' = \sum_j \langle j| f_j = \boxed{f_1 \cdots f_n} \tag{16B}$$

$$\langle j|k\rangle = \delta^j{}_k \quad \text{- - - - orthogonality} \tag{16C}$$

$$|j\rangle\langle k| = \begin{array}{c} \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 1 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \leftarrow k\text{th row} \\ \uparrow\ j\text{th column} \end{array} \tag{16D}$$

$$\sum_k |k\rangle\langle k| = I \qquad \text{closure} \tag{16E}$$

$$f(\vec{x}) = \langle f|x\rangle = f_j x^j \tag{16F}$$

$$|x\rangle\langle f| = \begin{bmatrix} x^1 f_1 & \cdots & x^1 f_n \\ \vdots & & \vdots \\ x^n f_1 & \cdots & x^n f_n \end{bmatrix} \tag{16G}$$

It is further clear that if $a^j{}_k$ are components of a matrix 'a', then the matrix itself is given by

$$a = \sum_{j,k=1}^{n} a^j{}_k \, |j\rangle\langle k| . \tag{17}$$

Thus (16D) furnishes the canonical basis for the matrix ring.

III.4: We have seen that an arbitrary $n \times n$ matrix involves n^2 scalars, and in the canonical basis one can write a matrix in the form (17). Though the canonical basis appears rather natural, because of its highly singular nature it is never explicitly used. One may ask do there exist any other basis that are not singular. When the dimension of space is $n = e^{2m}$ (m=integer $\geqslant 1$) there exist n^2 elements of "clifford algebra" [An algebra is a vector space in which a multiplication is defined with the properties $(x+y)z = xz + yz$, $x(y+z) = xy + xz$, $a(xy) = (ax)y = x(ay)$. For an associative algebra $(xy)z = x(yz) = xyz$]. Which can be used as a basis. The n^2 elements of the algebra are determined in terms of (2m) matrices γ_μ ($\mu = 1, 2 \cdots 2m$) which satisfy

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\,\delta_{\mu\nu} \tag{18}$$

Actually (2m+1) matrices satisfy this relation. The Pauli matrices mentioned elsewhere are particular case of the algebra for m = 1.

For an arbitrary (finite) dimension there is a theorem of Alladi Ramakrishnan according to which one can expand an arbitrary

$n \times n$ matrix A as

$$A = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} B^{j} C^{k} \tag{19}$$

where α_{jk} are elements of R and matrices B,C arise as representations of nth roots of unity. In particular one can take

$$B = \begin{bmatrix} 1 & & & & \\ & \omega & & 0 & \\ & & \omega^2 & & \\ & 0 & & \ddots & \\ & & & & \omega^{n-1} \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & & & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & & 0 & 1 & \vdots \\ \vdots & & & \ddots & \ddots \\ 0 & & & 0 & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{20}$$

where $\omega = \exp(2\pi i / n)$, $i = \sqrt{-1}$, $B^n = C^n = 1$. Since also BC = ω CB it follows that in the algebra of these matrices there are exactly n^2 independant elements. The Alladi basis is apparently the only nonsingular matrix basis valid for any dimensionality.

<u>III.5</u>: In equations (16-17) we have seen how one may write a matrix in a given basis. But in general it is not necessary to worry about the basis as we are dealing with R^n (or more generally K^n) and entries in the matrix components is what matters. But if we change the basis: $A|j\rangle = |j'\rangle$, then an arbitrary matrix M also undergoes a linear transformation:

$$M' = A M A^{-1} \tag{21}$$

and is called a similarity transformation. The matrices M and M' are said to be similar. We now mention some important properties of matrices that are unchanged under similarity transformations.

(a) Let A be a matrix transformation in V_n. For every $\vec{X}$ in V_n, $A\dot{X} = Y$ is in V_n. The transposed matrix, denoted A^T, is defined by

$$(f, Ax) = (f A^T, x) \tag{22}$$

It is easy to verify that A^T is related to A with its rows and colomns interchanged. If $A = A^T$, it is called a symmetric matrix and if $A = -A^T$ it is called a ske -symmetric matrix; these properties are unchanged under similarity transformations.

(b) An important class of matrix equations is the eigenvalue equation for a symmetric matrix

$$M\vec{X} = \lambda\vec{X} \tag{23}$$

In n dimensions there are in eneral n solution and λ's are called the eigenvalues. Hence the is associated with n simultaneous equations (23) the nth degree equation in λ's:

$$\prod_{j=1}^{n} (\lambda - \lambda_j) = \sum_{r=1}^{n} \alpha_r \lambda^{n-r}, \quad \alpha_1 = 1. \tag{24}$$

The eigenvalues λ_j, as also the α_j which are functions of λ_j are characteristic of the matrix and are unchanged under a similarity transformation. In fact there always exists a similarity transformation that can diagonalize a symmetric matrix:

(25)

A little reflection shows that the coefficients α_k can be described completely in terms of two types of invariant functions on matrices called Trace or Spur (Tr. or Sp.) and determinant (det or $\|\cdot\|$). Thus

$$\alpha_2 = \sum_{i=1}^{n} \lambda_i = \text{sum of the diagonal elements of } M = \operatorname{Tr} M \tag{26}$$

$$\alpha_n = \prod_{i=1}^{n} \lambda_i = \text{product of diagonal elements of} = \operatorname{Det}(M) \tag{27}$$

Similarly other α's are determined in terms of "Trace" and determinant of powers of M. In general the determinant of a matrix may be expressed as

$$\operatorname{Det} M = \frac{1}{n!}\,\delta^{m_1 \cdots m_n}_{l_1 \cdots l_n} M^{l_1}_{m_1} \cdots M^{l_n}_{m_n}, \tag{28}$$

where the generalized Kronecker δ is defined as

$$\delta^{m_1 \cdots m_k}_{l_1 \cdots l_k} = \begin{cases} +1 \text{ if } m_1 \cdots m_k \text{ is an even permutation of } l_1 \neq l_2 \cdots \neq l_n \\ -1 \text{ if } m_1 \cdots m_k \text{ is an odd permutation of } l_1 \neq l_2 \cdots \neq l_n \\ 0 \text{ in all other cases} \end{cases} \qquad k \leq n \tag{29}$$

The equation of the type (24) for an arbitrary matrix A (whether symmetric or not) may be written as

$$\operatorname{Det}(A-\lambda I) = \det A - \frac{\lambda}{(n-1)!}\,\delta^{l_1\cdots l_{n-1}}_{k_1\cdots k_{n-1}} A^{k_1}_{l_1}\cdots A^{k_{n-1}}_{l_{n-1}} + \frac{\lambda^2}{(n-2)!}\cdots\cdots - - - - + \frac{(-1)^r\lambda^r}{(n-r)!}\,\delta^{l_1\cdots l_{n-r}}_{k_1\cdots k_{n-r}} A^{k_1}_{l_1}\cdots a^{k_{n-r}}_{l_{n-r}} + \frac{(-\lambda)^{n-1}}{1!}\,a^{k_1}_{k_1} + (-\lambda)^n \tag{30}$$

where we have repeatedly used the formula

$$\delta^{k_1\cdots k_r}_{l_1\cdots l_r}\,\delta^{l_r}_{k_r} = (n-k+1)\,\delta^{k_1\cdots k_{r-1}}_{l_1\cdots l_{r-1}}$$

in evaluating the expansion on the right, and I is the unit matrix.

From this definition of the determinant one can easily deduce the various familiar properties of a determinant, but we shall not do that. An important property to note is that determinant of a product of several matrices is equal to the product of their determinants. The operation of taking a determinant is not linear; thus if c is a scalar, det $(cA) = c^n(\det A)$ for the dimension n. It follows that determinant of a skew symmetric matrix in odd dimensions vanishes identically. Since the trace of a skew symmetric matrix vanishes for all dimensions, it follows from equation (30) that one root of λ for a skew symmetric matrix in odd dimensions is zero. Determinant of a symmetric matrix in even dimensions is positive semi-definite.

We define the cofactor of a determinant as

$$a^{l_1}{}_{k_1} \overset{\text{def.}}{=} \frac{1}{(n-1)!} \delta^{l_1 \cdots l_n}_{k_1 \cdots k_n} A^{k_2}{}_{l_2} \cdots A^{k_n}{}_{l_n}, \tag{31}$$

so that

$$(\det A)\, \delta^{l_1}{}_{p} = a^{l_1}{}_{k_1} A^{k_1}{}_{p} \tag{32}$$

From this it is easy to see that the cofactor $a^{l_1}{}_{k_1}$ of the element $A^{k_1}{}_{l_1}$ is $(-1)^{i+j}$ times the determinant of the matrix M when its jth row & ith colomn are deleted. If we define

$$\alpha^{i}{}_{j} = \frac{a^{i}{}_{j}}{\det A} \tag{33}$$

We see that the matrix α is inverse of matrix A

$$\alpha^{i}{}_{j} A^{j}{}_{l} = \delta^{i}{}_{l}, \quad \alpha A = A\alpha = I \tag{34}$$

It is also clear that $\det a = (\det A)^{n-1}$ and $\operatorname{Det} \alpha = (\det A)^{-1}$.

Let the elements of a matrix M be differentiable functions of a parameter t. Then using the above definitions for $\det A$. and $\alpha^{i}{}_{j}$ we get

$$\frac{\partial \alpha^{l_1}{}_{q_1}}{\partial t} = \frac{1}{n-2} \delta^{l_1 \cdots l_n}_{q_1 \cdots q_n} \frac{\partial a^{q_2}{}_{l_2}}{\partial t} a^{q_3}{}_{l_3} \cdots a^{q_n}{}_{l_n} \tag{35}$$

$$\frac{\partial \det A}{\partial t}\,\delta^{l_1}_{k_1} = \frac{1}{(n-2)!}\,\delta^{l_1 \cdots l_n}_{q_1 \cdots q_n} A^{q_1}_{k_1} \frac{\partial A^{q_2}_{l_2}}{\partial t} A^{q_3}_{l_3} \cdots A^{q_n}_{l_n}$$

$$+ \frac{1}{(n-1)!}\,\delta^{l_1 \cdots l_n}_{q_1 \cdots q_n} \frac{\partial A^{q_1}_{l_1}}{\partial t} A^{q_2}_{l_2} \cdots A^{q_n}_{l_n} \tag{36}$$

putting $l_1 = k_1$ we obtain

$$\frac{\det A}{\partial t} = \alpha^{l_1}{}_{q_1} \frac{\partial A^{q_1}{}_{l_1}}{\partial t} \tag{37}$$

IV TENSORS:

I. Let V W be vector spaces and V x W their cartesian product.

Def. A function $F : V \times W \longrightarrow R$ is bilinear if

$$F(a\vec{x}+b\vec{y},\vec{u}) = aF(\vec{x},\vec{u}) + bF(\vec{y},\vec{u})$$
$$F(\vec{x},a\vec{u}+b\vec{v}) = aF(\vec{x},\vec{u}) + bF(\vec{x},\vec{v}) \tag{1}$$

for all a,b in R, $\vec{x},\vec{y}$ in V and $\vec{u},\vec{v}$ in W.

The set of all bilinear functions F on V x W can be given the structure of a vector space by the procedure used earlier for linear forms on a vector space. Let $\vec{e}_i$, $\vec{I}_\alpha$ be the basis in V,W; Then

$$\begin{aligned} F(\vec{x},\vec{u}) &= F(x^i\vec{e}_i, u^\alpha\vec{I}_\alpha) \\ &= F(\vec{e}_i,\vec{I}_\alpha)\, x^i u^\alpha \end{aligned} \tag{2}$$

Now $x^i = e^i(\vec{x}),\ u^\alpha = I^\alpha(\vec{u})$ where e^i, I^α are the basis in V^*, W^*. If we rewrite

$$x^i u^\alpha = e^i(\vec{x})\, I^\alpha(\vec{u}) = (e^i \otimes I^\alpha)(\vec{x},\vec{u}), \tag{3}$$

Then we see that the functions $e^i \otimes I^\alpha$ are linear on (x,) (i.e. linear in each factor) and are linearly independ nt, since $K_{i\alpha}\, e^i \otimes e^\alpha = 0$ implies that

$$K_{i\alpha} = K_{j\beta}\, e^j \otimes I^\beta(\vec{e}_i, \vec{I}_\alpha) = 0,$$ for all couples of indices (i,α).

The symbol $\otimes$ is read tensor product and the vector space of bilinear functions on V x W is called the tensor product $V^* \otimes W^*$ of V^* and W^*. The dimension of $V^* \otimes W^*$ is the product of the dimensions of V^* and W^*; its elements are of the form

$$F = F_{i\alpha}\, e^i \otimes I^\alpha \tag{4}$$

and are called twice covariant tensors; $F_{i\alpha}$ are the components of F in the basis $e^i \otimes I^\alpha$

Since V,W may be identified with V^{**},W^{**} one can in the same fashion construct the vector space of "twice contravariant tensors" $V \otimes W$ with elements of the form $F = F^{i\alpha}\, \vec{e}_i \otimes \vec{e}_\alpha$. Similarly one can construct spaces of once covariant and once contravariant tensors: $V^* \otimes W$ of tensors $F_i{}^\alpha\, e^i \otimes \vec{I}_\alpha$ and $V \otimes W^*$ of tensors $F^i{}_\alpha\, \vec{e}_i \otimes I^\alpha$.

One can generalize the notion to consider tensor product of of several vector spaces as it is associative

$$\Sigma(\vec{\gamma}) \otimes [A \otimes f\ (\vec{u}, \vec{x})] = [\Sigma \otimes A\ (\vec{\gamma}, \vec{u})] \otimes f(\vec{x})$$

$$= (\Sigma \otimes A \otimes f)\ (\vec{\gamma}, \vec{u}, \vec{x}) \tag{5}$$

by the defining equation.

In what follows we shall confine ourselves to tensor products of a vector space with itself or with its dual. The elements of the repeated tensor product space $\otimes^k V \otimes (\otimes^\ell V^*)$ are called mixed tensors of valence (k, ℓ) and rank $k+\ell$ and have the form

$$F = F^{j_1 j_2 \cdots j_k}{}_{p_1 \cdots p_\ell}\, \vec{e}_{j_1} \otimes \cdots \otimes \vec{e}_{j_k} \otimes e^{p_1} \cdots \otimes e^{p_\ell}. \tag{6}$$

It is cle r that there can be several tensor spaces of the same rank and even of the same valence; they are isomorphic in the ~~same~~ sense of having the same dimension; however there is no natural isomorphism between them in the sense that there is no basis independent one to one correspondence between their elements in general. (In the particular case of a "Riemannian Space" the natural isomorphism does exist).

IV.2: Take any F in V $\otimes$ V; Then

$$F = F^{ij}\, \vec{e}_i \otimes \vec{e}_j \tag{7}$$

In the new basis

$$F = F^{i'j'}\, \vec{e}_{i'} \otimes \vec{e}_{j'} \tag{8}$$

On substituting for $\vec{e}_{i'}$, $\vec{e}_{j'}$ we find

$$F^{i'j'} = A^{i'}{}_{\ell}\, A^{j'}{}_{m}\, F^{\ell m}. \tag{9}$$

This is the usual definition of a contravariant tensor of second rank in terms of its components. Similarly the components of a covariant tensor of second rank transform as

$$F_{i'j'} = A_{i'}{}^{\ell}\, A_{j'}{}^{m}\, F_{\ell m}, \tag{10}$$

and those of the mixed tensor as

$$F^{i'}{}_{j'} = A^{i'}{}_{\ell}\, A_{j'}{}^{m}\, A^{\ell}{}_{m}$$

$$F_{i'}{}^{j'} = A_{i'}{}^{\ell}\, A^{j'}{}_{m}\, F_{\ell}{}^{m} \qquad (11)$$

For the kronecker symbol we have

$$\delta^{j'}_{k'} = e^{j'}(\vec{e}_{k'}) = A^{j'}{}_{j}\, A_{k'}{}^{k}\, e^{j}(\vec{e}_{k})$$

$$= A^{j'}{}_{j}\, A_{k'}{}^{k}\, \delta^{j}{}_{k}\,. \qquad (12)$$

So the Kroneckor symbol is a mixed tensor of the second rank.

We note that matrices of linear maps $V \rightarrow W$ are mixed tensors of $V^* \otimes W$. Thus square matrices give components of the elements of $V^* \otimes V$. If M is a square matrix acting on V, then the map $M\vec{X}$ is the value of the tensor $M = M_i{}^j\, e^i \otimes \vec{e}_j$ on the vector $\vec{X}$ in V :

$$M\vec{X} = M_i{}^j\, e^i \otimes \vec{e}_j(\vec{X}, \cdots) = M_i{}^j\, e^i(\vec{X})\, \vec{e}_j$$

$$= (M_i{}^j X^i)\, \vec{e}_j \qquad (13)$$

and is clearly an element of V. We call this map the contraction of an element M of $V^* \otimes V$ by the element $\vec{X}$ of V (or just <u>contraction of M by $\vec{X}$</u>; or in older texts inner product of M and $\vec{X}$).

It is clear that the Kronecker tensor $\delta_i{}^j$ gives the identity map δ of V into itself. The contraction of a mixed tensor T of valence (p,q) with δ is a tensor of valence (p-1,q-1) and is referred to as <u>"the contraction of T"</u>. For a mixed tensor $T^{j_1 j_2}{}_{k_1}$ we have contractions

$$(a)\; T^{j_1 j_2}{}_{k_1}\,\delta_{j_1}{}^{k_1} = T^{j_1 j_2}{}_{j_1} \rightarrow \tau^{j_2}$$

$$(b)\; T^{j_1 j_2}{}_{k_1}\,\delta_{j_2}{}^{k_1} = T^{j_1 j_2}{}_{j_2} \rightarrow t^{j_1} \qquad (14)$$

which shows that contraction is prescribed with respect to a pair of upper and lower indices. Contraction of an arbitrary tensor by another tensor can be similarly effected if there are suitable pairs of upper and lower indices. This latter type of contraction is also referred in older texts as inner product of two tensors.

The contraction of a matrix $M_i{}^j$ is its trace $M_i{}^i$; by the definition of contraction, it is a real number independent of the basis in V.

Suppose we have an object $M^{l_1 \cdots l_k}{}_{m_1 \cdots m_j}$ of unknown character under transformation of the basis; but given a covariant vector v_l (contravariant vector v^l) the set of quantities $v_{l_1} M^{l_1 \cdots l_k}{}_{m_1 \cdots m_j}$ $(v^{m_1} M^{l_1 \cdots l_k}{}_{m_1 \cdots m_j})$ transform as components of a tensor of valence $(k-1, j)$ $[(k, j-1)]$ then it would follow from the transformation character of the vector that $M^{l_1 \cdots l_k}{}_{m_1 \cdots m_j}$ is a tensor of valence (k, j) (<u>Quotient Law</u>)

<u>IV.3</u>: If we have a covariant (contravariant) tensor of second rank, we see that under the transformation of basis each index transforms linearly. Hence symmetry or antisymmetry with respect to interchange of indices is a property that is independent of the basis.

Let

$$F_{ij} = \frac{1}{2}\left(F_{ij} + F_{ji}\right) + \frac{1}{2}\left(F_{ij} - F_{ji}\right)$$
$$= F_{\underline{ij}} + F_{\underset{\vee}{ij}} \; ; \qquad (15)$$

one can express the tensor corresponding to the components $F_{\underset{\vee}{ij}}$ as

$$F_{\vee} = \frac{1}{2!}\, \delta^{lm}_{ij}\, F_{lm}\, e^i \otimes e^j = F_{\underset{\vee}{ij}}\, e^i \otimes e^j$$
$$\overset{def}{=} F_{ij}\, e^i \wedge e^j \, , \qquad (16)$$

where the symbol $\wedge$ is read as Wedge product or exterior product of vector spaces and

$$\delta^{lm}_{ij} = \delta^l_i \delta^m_j - \delta^l_j \delta^m_i = \det \begin{vmatrix} \delta^l{}_i & \delta^m{}_j \\ \delta^l{}_j & \delta^m{}_i \end{vmatrix} \qquad (17)$$

is called the generalized kronecker δ tensor of rank 4; its tensor character follows fom the tensor character of $\delta^i{}_j$;

In a space of n-dimensions one can define several generalized tensors of rank $2m \leq 2n$:

$$\delta^{j_1 \cdots j_m}_{k_1 \cdots k_m} = \begin{cases} +1 & \text{if } j_1 \cdots j_m \text{ is an even permutation of } k_1 \cdots k_m \\ -1 & \text{if } j_1 \cdots j_m \text{ is an odd permutation of } k_1 \cdots k_m \\ 0 & \text{in all other cases; } k_1 \neq k_2 \cdots \neq k_m \end{cases}$$

Each of these can be expressed in terms of the kronecker δ symbol as a determinant:

$$\delta^{j_1 \cdots j_m}_{k_1 \cdots k_m} = \begin{Vmatrix} \delta^{j_1}_{k_1} & \delta^{j_1}_{k_2} & \cdots & \delta^{j_1}_{k_m} \\ \delta^{j_2}_{k_1} & \cdots & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ \delta^{j_m}_{k_1} & \delta^{j_m}_{k_2} & \cdots & \delta^{j_m}_{k_m} \end{Vmatrix} \qquad (18)$$

The tensor character of generalized δ is also obvious. Given an arbitrary covariant (contravariant) tensor one can decompose it into tensors of different symmetry type by taking suitable linear combinations as we did for a second rank tensor. For an mth rank covariant tensor, the completely antisymmetric part is given by the formula which generalized (16):

$$F_{[j_1 \cdots j_m]} = \frac{1}{m!} \delta^{l_1 \cdots l_m}_{j_1 \cdots j_m} F_{l_1 \cdots l_m} . \tag{19}$$

The square bracket denotes that it is antisymmetric with respect to the interchange of any two indices. The vector space of antisymmetric tensors is an important subspace of tensor space. To further clarify it in the spirit of generalizing (16), we can write the tensor as

$$[F] = F_{[j_1 \cdots j_m]} e^{j_1} \otimes \cdots \otimes e^{j_m} = F_{j_1 \cdots j_m} e^{j_1} \wedge \cdots \wedge e^{j_m} \tag{20}$$

where

$$e^{j_1} \wedge e^{j_2} \cdots \wedge e^{j_m} = \frac{1}{m!} \delta^{j_1 \cdots j_m}_{k_1 \cdots k_m} e^{k_1} \otimes e^{k_2} \cdots \otimes e^{k_m} \tag{21}$$

$$\left.\begin{aligned} \vec{e}_{1'} \wedge \vec{e}_{2'} \cdots \wedge \vec{e}_{m'} &= (\det A^{a}{}_{b'}) \vec{e}_1 \wedge \vec{e}_2 \cdots \wedge \vec{e}_m \\ e^{(m+1)'} \wedge \cdots \wedge e^{n'} &= (\det A^{k'}{}_{e}) e^{m+1} \wedge \cdots \wedge e^{n} \end{aligned}\right\} \tag{22}$$

Thus the subspace V_m is determined completely (within a scalar factor) by an m-vector or a (m-n) form and **conversely.** In fact the number of components of a m-vector is same as that of an (m-n) form. Since

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!} = \binom{n}{n-m}. \qquad (23)$$

Geometrically m-dimensional surfaces in an n-dimensional space correspond to the components of an m-vector.

V. Euclidean and Pseudo-Euclidean Spaces.

V.1. Let V be a vector space, we can define a scalar product (dot product) in V as a mapping of $V \times V \longrightarrow K$ with the following properties

$$\vec{x}\cdot\vec{y} = \vec{y}\cdot\vec{x} \qquad \vec{x}\cdot(a\vec{y}+b\vec{z}) = a\vec{x}\cdot\vec{y} + b\vec{x}\cdot\vec{z} \tag{V.1}$$

From these also follows $(a\,x+by)\cdot\vec{z} = a\vec{x}\cdot\vec{z} + b\,\vec{y}\cdot\vec{z}$.

We note parenthatically that if $\vec{x}\cdot\vec{y} = 0$ for all $\vec{y}$ in V implies that $x = 0$, then the scalar product is called non-degenerate. As we shall see this is always so for Euclidean spaces. For Pseudo-Euclidean spaces the degenerate case arises.

Let $\vec{e}_j$ be a basis in V, define the set of scalars

$$g_{ij} = \vec{e}_i\cdot\vec{e}_j\,. \tag{V.2}$$

Consider the change in basis $\vec{e}_{\ell'} = A_{\ell'}{}^{j}\,\vec{e}_j$, then

$$g_{\ell'm'} = A_{\ell}{}^{j} A_{m'}{}^{k}\, g_{jk}\,; \tag{V.3}$$

this shows that the set of scalars g_{jk} are components of a co-variant tensor of the second rank (compare eqn. IV.10):

$$g = g_{ij}\,\vec{e}^{\,i} \otimes \vec{e}^{\,j}. \tag{V.4}$$

We shall now show that using this tensor one can establish a natural isomorphism between V and V* in the sense that for every element of V (or one of its tensor product spaces) there is an exact element in V* (or its tensor product space) and vice-versa. For this reason the tensor g_{ij} is called the <u>fundamental tensor</u>.

Recall that if F is an element of $V \otimes V$ and f an element of V^*, then the map

$$f(F) = (f_k F^{kl})\vec{e}_l \tag{V.5}$$

defines a unique element of V. Similarly one may consider the map

$$V^* \otimes V^* \xrightarrow{V} V^*$$

If $u = u^j \vec{e}_j$ is an element of V then clearly

$u = (g_{ij} u^j) e^i$ is an element of V^*.

Let us denote the components of u in the basis e^i by u_i, than

$$u = u_i e^i \quad , \quad u_i = g_{ij} u^j. \tag{V.6}$$

Let g^{ij} be components of the matrix which is inverse of the matrix with components g_{ij}, then

$$g_{ij} g^{jk} = \delta_i{}^k. \qquad (V.7($$

From (V.7) the tensor character of the components g^{ik} is obvious If $v = v_i e^i$ is an element of V^* then the corresponding element of V is given by

$$\vec{v} = (g^{ij} v_j)\vec{e}_i ; \quad v^j = g^{ij} v_i \tag{V.8}$$

In this manner g can be used to define a <u>natural isomorphism</u> between tensor spaces of the same dimension. We note that if e_i denotes the linear form corresponding to $\vec{e}_i$ in V and if we define the scalar product in V* by

$$\left.\begin{aligned} e_i \cdot e_j &\overset{def}{=\!=} \vec{e}_i \cdot \vec{e}_j \\ \text{then} \quad e^i \cdot e^j = g^{ij} \quad &, \quad e^i = g^{ij} e_j \end{aligned}\right\} \tag{V.9}$$

The tensor $g^{ij}\vec{e}_i \otimes \vec{e}_j$ is called the contravariant fundamental tensor.

V.2. Given an n dimensional (Pseudo) Euclidean space it is always possible to choose a basis $\vec{e}_i$ such that

$$\vec{e}_i \cdot \vec{e}_j = \eta_{ij} = \begin{cases} \mp 1 & \text{for } i = j \\ = 0 & 0 \quad i \neq j \end{cases} \tag{V.10}$$

$\sum_i \eta_{ii} = s$ is called the signature of V_n. Of the total number n of the components η_{ii}, $\frac{1}{2}(s+n)$ are $+1$ and $\frac{1}{2}(n-s)$ are -1. If $|s| = n$, the basis is orthonormal and we get the ordinary Euclidean space. If $|s| < n$, the space is called pseudo-Euclidean of signature s. The space in Newtonian mechanics is Euclidean and space-time of special relativity is pseudo-Euclidean of signature $\pm$ 2.

V.3. A simple basis in Euclidean space may be chosen as

$$\vec{e}_j = (\underset{1}{0}, \underset{2}{0} - - - - \underset{j-1}{0}, \underset{j}{1}, \underset{j+1}{0} - - - - \underset{n}{0}) \tag{V.11}$$

This basis is called the canonical basis. If the space is not pseudo-Euclidean then in this basis the distinction between V and V^* completely disappears. For every vector $\vec{X}$ in a Euclidean space, the norm of the vector is defined as an element of R,

$$|\vec{X}| = \sqrt{X_1^2 + - - - + X_n^2} = \sqrt{\vec{X} \cdot \vec{X}}. \tag{V.12}$$

The notion of inner product can be easily defined by means of the formula

$$|X+y|^2 = |X|^2 + |y|^2 + 2(X, y). \tag{V.13}$$

It follows from the Schwarz inequality, $(x,y)^2 \leq |x|^2 |y|^2$, that

$$-1 \leq \frac{(x,y)}{|x|\,|y|} \leq +1. \qquad \text{(V.14)}$$

and is a measure of the angle between two vectors and .
If we put $\cos\theta = (x,y)/|y||x|$, then

$$(x, \vec{e}_j) = x_j = |x| \cos\alpha_j, \qquad \text{(V.15)}$$

where α_j denotes the 'angle' between $\vec{x}$ and $\vec{e}_j$ and $\cos\alpha_j$ are the direction cosines in this basis. If $\vec{x}$, $\vec{y}$ are two vectors with components x_j, y_j in the canonical basis, then

$$S_{xy}^2 = |\vec{x} - \vec{y}|^2 = \sum_i (x_i - y_i)^2, \qquad \text{(V.16)}$$

and is called square of the distance between two vectors. We note that there is a whole class of basis related to the canonical basis for which the norm (V.12) and square of the distance (V.16) are given by the same formulas as above. The transformation matrices connecting these basis have the property that

$$A^T = A^{-1}, \quad A A^T = 1 \quad (A^T = \text{transposed of } A) \qquad \text{(V.17)}$$

Such matrices are called orthogonal, hence the transformations which leave the equations (V.12), (V.16) unchanged as also the corresponding class of basis are called orthogonal. These transformations form a group called the orthogonal group O(n). The subgroup of O(n) with $\det A = +1$ is denoted by SO(n).

V.4. For a Euclidean space, in the canonical basis, the matrices corresponding to the components of g_{ij} and g^{ij} are unit matrices. But if the space is pseudo-Euclidean then their components in the canonical basis, though again identical to each other are now given by

$$[\eta^{ij}] = [\eta_{ij}] = \begin{bmatrix} \epsilon_1 & & & 0 \\ & \epsilon_2 & & \\ & & \ddots & \\ 0 & & & \epsilon_n \end{bmatrix}, \quad \epsilon_j = \pm 1 \qquad \text{(V.18)}$$

For the norm of the vector we now write

$$\pm |X|^2 = \sum_{j=1}^{n} \epsilon_j \, X^{j\,2} \qquad \text{(V.12A)}$$

From the definition of the fundamental tensor and the 'distance formula' for vectors (V.12, V.16), we see that in general these may be written as

$$S_{xy}^2 = \eta_{ij} \, (x^i - y^i)(x^j - y^j). \qquad \text{(V.16a)}$$

For this reason the tensor $\eta = \eta_{ij} \, e^i \otimes e^j$ is also called the <u>Metric Tensor.</u> If X^i are components of a contravariant vector, then for a covariant vector the components are $X_j = g_{ij} X^i$. for the pseudo-Euclidean space these are $\{ \epsilon_{(i)} X^i \}$, in this canonical basis.

( ) denotes no summation

The transformations which leave (V.16A) unchanged in form are given by the matrices Λ which satisfy

$$\Lambda^T \eta \Lambda = \eta \quad \underline{\text{or}} \quad \Lambda_{lj} \Lambda^l{}_m = \eta_{jm} \qquad \text{(V.19)}$$

where,

$$(\Lambda^T)^l{}_j = \Lambda_j{}^l, \text{ and } \Lambda_{lj} = \Lambda^m{}_j \eta_{lm} = \Lambda_l{}^k \eta_{kj}$$

We consider the particular case when space is of the Lorentz type: $\epsilon_j = -1$ for $j = 1, \cdots, n-1$, and $+1$ for $j = n$. The transformations are of the following types

(a) Ordinary rotations R

$$-\pi \leq \theta \leq \pi \quad R_{12} = \left[\begin{array}{cccccc|c} \cos\alpha & \sin\alpha & 0 & \cdots & \cdots & 0 & 0 \\ -\sin\alpha & \cos\alpha & & & & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & \cdots & 0 & \vdots \\ \vdots & & & \ddots & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 1 & 0 \\ \hline 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{array}\right] \qquad \text{(V.20)}$$

This matrix represents a rotation in 1-2 plane. There are in all $\frac{1}{2}(n-1)(n-2)$ such matrices giving rotations in $x^1-x^2, x^2-x^3, \ldots, x^{n-1}-x^1$ planes. These matrices are orthogonal and so are their products. Hence the set of all these matrices and their products form a group - the subgroup of orthogonal group in $(n-1)$ dimensions: $SO(n-1)$ where the symbol S denotes that the determinant of these matrices is $+1$.

(b) Hyperbolic Rotations $\mathcal{H}$

$$-\infty < \theta < +\infty \qquad \mathcal{H}_{1n} = \left[\begin{array}{ccccc|c} \cosh\theta & 0 & - & - & 0 & \sinh\theta \\ & 1 & & & & 0 \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ 0 & 0 & - & - & 1 & 0 \\ \hline \sinh\theta & 0 & - & - & 0 & \cosh\theta \end{array}\right] \tag{V.21}$$

This matrix represents a hyperbolic rotation in $1-2$ plane. There are $(n-1)$ such hyperbolic rotations in the $1-n, 2-n, \ldots, (n-1)-n$ planes. The products of two such matrices in different planes is not of the same form (notice that $\mathcal{H}_{1n}$ is a symmetric matrix) as a pure hyperbolic rotation but also involves ordinary rotations. Hence hyperbolic rotations do not form a group though together with ordinary rotations R they do form a group, the <u>proper Lorentz</u> group in n dimensions

We note that the most general matrix Λ involving both ordinary and hyperbolic rotations may be written as

$$\Lambda = R_a H_k R_b \, . \tag{V.22}$$

To show this consider the expression $R_a^{-1} \Lambda R_b^{-1}$; it is of the form

$$\left[\begin{array}{c|c} a & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \; \cdots \; 0 & 1 \end{array}\right] \left[\begin{array}{c|c} \lambda & \vec{\Lambda}_0 \\ \hline \vec{\lambda}_0 & \Lambda_0 \end{array}\right] \left[\begin{array}{c|c} b & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \; \cdots \; 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{c|c} a\lambda b & a\vec{\Lambda}_0 \\ \hline \vec{\lambda}_0 & \Lambda_0 \end{array}\right] \tag{V.23}$$

For a suitable choice of $(n-1)\times(n-1)$ orthogonal matrices a and b we can put

$$a\vec{\Lambda}_0 = (\vec{\Lambda}^0 b)^T = (\underset{1}{0}, \underset{2}{\ldots}, 0, \underset{k}{\alpha}, 0, \ldots, \underset{n-1}{0}), \quad \alpha = \sqrt{\Lambda_0^{0\,2} - 1}$$

$$a\lambda b = \begin{array}{c} \begin{array}{cccccc} 1 & 2 & \cdots & k & \cdots & n-1 \end{array} \\ \left[\begin{array}{cccccc} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \cdots & \Lambda_0^0 & \cdots & 0 \\ \vdots & & & & 1 & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 1 \end{array}\right] \end{array} \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ k \\ \vdots \\ n-1 \end{array} \tag{V.24}$$

to obtain H_k as a hyperbolic rotation in the $k-n$ plane.

(c) Discrete transformations

space reflection

$$I_S = \left[\begin{array}{cccc|c} -1 & \cdots & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ \vdots & & -1 & \vdots & \vdots \\ 0 & \cdots & \cdots & -1 & 0 \\ \hline 0 & \cdots & \cdots & 0 & +1 \end{array}\right] \tag{V.25}$$

Time reversal

$$I_T = \left[\begin{array}{cccc|c} +1 & \cdots & \cdots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & \cdots & +1 & 0 & \vdots \\ 0 & \cdots & \cdots & +1 & 0 \\ \hline 0 & \cdots & \cdots & 0 & -1 \end{array}\right] \tag{V.26}$$

Space-time inversion $\quad I_{ST} = -I. \quad$ (V.27)

With these defining matrices we can now classify the transformations leaving (V.12A) unchanged when there is a Lorentz signature $\pm(n-2)$:

$$\left.\begin{array}{lll} L_+^{\uparrow}\ (\text{Transformations } L) & \det\Lambda = +1 & \Lambda^0{}_0 \geq +1 \\ L_-^{\uparrow} = I_s L_+^{\uparrow} & \det\Lambda = -1 & \Lambda^0{}_0 \geq +1 \\ L_+^{\downarrow} = I_{sT} L_+^{\uparrow} & \det\Lambda = +1 & \Lambda^0{}_0 \leq -1 \\ L_-^{\downarrow} = I_T L_+^{\uparrow} & \det\Lambda = -1 & \Lambda^0{}_0 \leq -1 \end{array}\right\}. \tag{V.28}$$

Except for $L_+^{\uparrow}$, the rest of the sets do not form a group. But some of their combination do form groups and may be classified as under $(\epsilon \stackrel{\text{def.}}{=} \Lambda^0{}_0 \det\Lambda)$

Lorentz group type	Symbol	Characterization
Restricted (connected)	$L_+^{\uparrow}$	$\Lambda^0{}_0 \geq 1,\ \det\Lambda = 1,\ \epsilon > 0$
Orthochronous	$L_0^{\uparrow} = L_+^{\uparrow} \cup I_s = L_+^{\uparrow} \cup L_-^{\uparrow}$	$\Lambda^0{}_0 \geq 1$ (V.29)
Orthochorous	$L_0 = L_+^{\uparrow} \cup I_T = L_+^{\uparrow} \cup L_-^{\downarrow}$	$\epsilon > 0$
Proper	$L_+ = L_+^{\uparrow} \cup I_{sT} = L_+^{\uparrow} \cup L_+^{\downarrow}$	$\det\Lambda = 1$

In the above the symbol $\cup$ denotes the union of sets.

V.5. If we substitute the series expression for $\sin\alpha$ and $\cos\alpha$ in equation(V.20) we see on close examination that one can rewritte it as

$$R_{12} = \exp\{\alpha S_{12}\}, \tag{V.30}$$

where the matrix S_{12} is given by

$$S_{12} = \frac{\partial R_{12}}{\partial \alpha}\bigg|_{\alpha=0} = \left[\begin{array}{cccccc|c} 0 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ -1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \end{array}\right] \qquad \text{(V.31)}$$

Alternately we can also write

$$R_{12} = \exp\{\alpha L_{12}\}, \qquad \text{(V.32)}$$

$$L_{12} = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}. \qquad \text{(V.33)}$$

This latter form of R_{12} is in a certain sense more general in that one can use it for function of the coordinate vector also. For the particular case when R_{12} acts on a coordinate vector we get the expression (V.30-31). In general the rotation transformation in a k-l plane is given by

$$\left.\begin{aligned} R_{kl}(\theta) &= \exp\{\theta L_{kl}\} \\ L_{kl} &= x_k \partial_l - x_l \partial_k \end{aligned}\right\} \quad \partial_k = \frac{\partial}{\partial x_k} \qquad \text{(V.33)}$$

In the case of 3-dimensional Euvlidean space, the most general rotation may be expressed in terms of three Euler angles;

$$R(\alpha,\beta,\gamma) = R_z(\alpha)\, R_y(\beta)\, R_z(\gamma), \qquad \text{(V.34)}$$

where, e.g.

$$R_z = R_{12} = \exp(\alpha L_z) \;,$$

and L_x, L_y, L_z are defined by

$$L_k = x_i \partial_j - x_j \partial_i \quad i,j,k \text{ cyclic } \& = x,y,z \qquad \text{(V.35)}$$

For a pure Lorentz transformation $(n=4)$

$$\mathcal{H}_k(\theta) = \exp(\theta M_k)$$

$$M_k = x_k \partial_t + t\,\partial x_k \,, \quad \left[\begin{cases} t = x^4 \\ \partial_t = \dfrac{\partial}{\partial t} \,. \end{cases} \right] , \qquad \text{(V.36)}$$

L_k and M_k are called generators of the transformations and are 6 in number. When acting on the coordinate space it is possible to express L_k and M_k as 4×4 matrices.

If we had an n-dimensional space there will be $\frac{1}{2}n(n-1)$ generators which can be looked upon as components of an antisymmetric matrix. The significance of this is that in n-dimensional space there are exactly $\frac{1}{2}n(n-1)$ two-dimensional surfaces.

Under a general Lorentz transformation the basis transform as

$$\vec{e}\,' = \Lambda \vec{e} \qquad \text{(V.37)}$$

where $\vec{e}$ stands for the colmn vector of base vectors. Differentiating with respect to an arbitrary parameter we get

$$\frac{d\vec{e}\,'}{d\theta} = \frac{d\Lambda}{d\theta}\vec{e} \stackrel{\text{def}}{=\!=} C(\Lambda)\vec{e}\,' \qquad \text{(V.38)}$$

$$C(\Lambda) \stackrel{\text{def}}{=\!=} \frac{d\Lambda}{d\theta}\Lambda^{-1}. \qquad \text{(V.39)}$$

The matrix $C(\Lambda)$ is called the Cartan matrix of Λ. If $\omega_i{}^j$ denote the components of the Carton matrix C, then for the Euclidean space, since matrices Λ are orthogonal, we get $\omega_i{}^j = -\omega^j{}_i$. For a pseudo-Euclidean space if η_{ij} are components of the metric tensor, then

$$\eta_{ik}\omega^k{}_j = \omega_{ij} = -\omega_{ji} \qquad (V.40)$$

The equation (V.38) are called the frenet equations for an orthogonal frame. The Cartan matrix has the following important properties

$$C(k\Lambda) = kC(\Lambda) \quad , k \text{ a scalar}$$

$$C(\Lambda\mu) = C(\Lambda) + \Lambda C(\mu)\Lambda^{-1} \qquad (V.41)$$

These are of considerable importance in differential geometry.

V.6. We have seen that square of a vector takes a particularly simple form in the canonically adopted basis. All these basis are related to each other by orthogonal or pseudo orthogonal transformation which form the group $O(n-s, s)$. The corresponding co-ordinates are called cartesian. If we admit transformations of the general linear group, i.e. when transformation matrix has no particular restriction on it, then the matrix corresponding to η_{ik} is no longer diagonal and we denote it by g_{ik}. Here one has to clearly distinguish between the basis $\{\vec{e}_k\}$ and its dual $\{e^k\}$. Unlike Cartesian co-ordinates which are (pseudo) orthogonal, the coordinates in this case one in general oblique to each other in the pictorial sense.

One can further generalise the transformations such that the transforming matrix itself depends on the coordinates. If this case one can not speak of a coordinate vector and one has to consider instead a coordinate differential, i.e. we define a general coordinate transformation as a point transformation

$$x'^{i} = f^{i}(x^1, x^2, \ldots, x^n) \tag{V.42}$$

where f^i are assumed to be differentiable ——— sufficient number of times (at least four fold differentiable). The Jacobian matrix at p_0 , in neighbourhood of which the coordinate transformation is envisaged is assumed to be non-singular:

$$\det\left[\frac{\partial x'^{i}}{\partial x^{j}}\right] = \det\left[J^{i'}{}_{j}(x', x)\right] \neq 0 \tag{V.43}$$

In this case the coordinate differentials transform linearly and homogeneously

$$dx'^{i} = \frac{\partial x'^{i}}{\partial x^{j}}\, dx^{j} = A^{i'}{}_{j}\, dx^{j} \tag{V.44}$$

and are said to form components of a contravariant vector (compare eqn. II.15). Similarly

$$\frac{\partial}{\partial x^{j'}} = \frac{\partial x^{k}}{\partial x^{j'}}\, \frac{\partial}{\partial x^{k}} = A^{k}{}_{j'}\, \frac{\partial}{\partial x^{k}} \tag{V.45}$$

transform as components of a covariant vector.

Here it is useful to make a cautionary remark that the coordinate differentials or component of the gradient operator do not have any intrinsic meaning. It turns out that one has to interpret these in the sense $e^i \to dx^i$ and $\vec{e}_i \to \partial/\partial x^i$. Consider a curve $\sigma(t)$ in R^n, where t is a parameter along the curve. In terms of coordinates in R^n it is given by the function $f(x^1(t), x^2(t), \cdots x^n(t))$, which we assume to be differentiable; then

$$\frac{df}{dt} = \frac{dx^j}{dt} \frac{\partial}{\partial x^j} f \qquad \text{(V.46)}$$

Through a point p_0 we can have an infinity of such curves and dx^j/dt are the usual components of the tangent vector to the curve. The vector components $\partial/\partial x^j$ then given the basis of the tangent vector space T_{p_0} at p_0. It is then natural to take dx^j as components in the cotangent space which is dual to T_{p_0}.

It is possible to justify the identification $e^i \to dx^i$, $\vec{e}_i \to \partial/\partial x^i$, we have just made, in a negative sense which brings out the fact that the usual vector space basis can not be taken over us such and has to be reinterpreted. By the rules of calculus, the components of a tangent vector at p_0 transform as

$$\bar{V}^k(y) = \frac{\partial y^k}{\partial x^j} V^j(x); \qquad \text{(V.47)}$$

On differentiation we obtain

$$\frac{d\bar{V}^j}{dt} = \frac{\partial y^j}{\partial x^k}\left[\frac{dV^k}{dt} + \frac{\partial x^k}{\partial y^m}\frac{\partial^2 y^m}{\partial x^\ell \partial x^p}\frac{dx^\ell}{dt}V^p\right] \tag{V.48}$$

If y^k are the cartesian coordinates, it is clear that in transition to a general coordinate system the derivative of a vector does not transform as a vector. In the canonical basis the coordinate vector is given by $Y = y^j \vec{e}_{oj}$; since $\vec{e}_{oj}$ are constants,

$$dy = dy^j \vec{e}_{oj} = dx^k\left(\frac{\partial y^j}{\partial x^k}\vec{e}_{oj}\right) = dx^k \vec{e}_k \tag{V.49}$$

Can one consider $\vec{e}_k \overset{\text{def.}}{=} \vec{e}_{oj}\,\partial y^j/\partial x^k$ as components of the new basis ? To answer this we note that under the general change of coordinates (say from cartesian to polar coordinates $(y^1, y^2 \rightarrow$, $x^1 = r, x^2 = \theta)$ the relation $\vec{y} = y^j \vec{e}_{oj}$ has no meaning; it is however still possible to interpret $e_k = \partial y^j / \partial x^k \vec{e}_{oj}$ as components of a vector. It is clear that e_k are in general (differentiable) functions of coordinates. Since dy^j and dx^k are the basis in different coordinate systems the functions $\vec{e}_{oj}, \vec{e}_k$ prescribe the coordinate frame. We shall refer to $\{\vec{e}_k\}$ as frame vectors. An arbitrary contravariant vector at p_0 can be written down as a linear combination of its components with $\{\vec{e}_k\}$ as the coefficients which prescribe the coordinate frame. This set of vectors $\vec{e}_k$ is some times referred to as n-bien (Vierbiean for 4 dimensions). Since $\{\vec{e}_k\}$ are functions of coordinates.

they change from point to point and are therefore strictly speaking frame fields. The n vectors $\{e_k\}$ represent n^2 quantities which can be arranged in this form of a matrix with components $e^a{}_k$ By the method we obtained these (they are linearly independent at p_0) it is clear that the matrix $e^a{}_k$ is nonsingular. One can therefore define the reciprocal vectors e^l with components $(e^l)_a$ such that

$$e^l{}_a\, e^a{}_k = \delta^l{}_k. \tag{V.50}$$

We note that for an arbitrary set of linearly independent vectors $\vec{e}_k$ at p , the conditions

$$\partial_l e^a{}_k - \partial_k e^a{}_l = 0 \qquad (\partial_l = \partial/\partial x^l) \tag{V.51}$$

are the integrability conditions of the system

$$e^a{}_k = \frac{\partial y^a}{\partial x^k}, \quad y^a(x^1, x^2 \cdots x^n). \tag{V.52}$$

We shall see that the conditions (V.51) are an important set of conditions to be fulfilled for a Riemannian space.

V.7. Since the $\{dx^k\}$ correspond to the basis at p_0 in the cotangent space, the co-vectors (covariant vectors) are pfaffians or one forms:

$$\omega = f_j\, dx^j. \tag{V.53}$$

If the components of a covariant vector f_j are of the form $f_j = \partial f / \partial x^j$, then ω is a complete differential. In section (IV.3) we had introduced the concept of exterior product. We generalize this to the case of exterior differential forms on an n-dimensional pseudo-Euclidean space. If $\omega_{j_1 \cdots j_k}$ is a completely antisymmetric tensor field of the kth rank $(k < n)$ then a kth differential form called k-form is defined by

$$\omega^{(k)} = \omega_{j_1 \cdots j_k}(x^1 \cdots x^n)\, dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_k}.$$

If $\omega_{j_1 \cdots j_k}$ are differentiable functions of x^k then the formal 'exterior differential' of $\omega^{(k)}$ is a $k+1$-form given by

$$d\omega^{(k)} = \frac{\partial \omega_{j_1 \cdots j_k}(x^1 \cdots x^n)}{\partial x^l}\, dx^l \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}. \qquad \text{(V.54)}$$

It follows from this definition that

$$dd\omega^{(k)} = -dd\,\omega^{(k)} = 0.$$

This is called Poincares formula, it is basis to the Euclidean geometry and generalizes to the Riemannian geometry. An important point however is that one can use this method as a formal calculus: the calculus of differential forms.

In particular for the coordinate differential we must have

$$dd\vec{x} = \frac{\partial \vec{e}_k}{\partial x^l}\, dx^l \wedge dx^k = (\partial_l e_k - \partial_k e_l)\, dx^l \otimes dx^k$$

$$= 0 \qquad \text{(V.55)}$$

since the term in parenthesis is just integrability condition for the system (V.51). This condition also characterises a Riemannian geometry.

V.8. Now consider a vector in general coordinates given by

$$\vec{V} = \vec{e}_j v^j = \underset{\sim}{e}^j v_j \tag{V.56}$$

where $v^j = g^{ij} v_i$. We consider the case when

$$\vec{e}_k \cdot \vec{e}_j = \begin{cases} g_{jj} = (h_j)^2 & \text{for } k = j \\ 0 & \text{for } k \neq j \end{cases} \tag{V.57}$$

In this case by eqn. (V.7)

$$g^{kj} = \begin{cases} (g_{jj})^{-1} = h_j^{-2} & \text{for } k = j \\ 0 & \text{for } k \neq j \end{cases} \tag{V.57a}$$

It is then possible to define (no sum over j)

$$\begin{aligned} \vec{u}_j &= \underset{\sim}{e}^j h_{(j)} = \vec{e}_j / h_{(j)} \\ \vartheta_j &= v^j h_{(j)} = v_j / h_{(j)} \end{aligned} \tag{V.58}$$

so that

$$\vec{V} = \sum_j \vec{u}_j \vartheta_j \qquad \vec{u}_j \cdot \vec{u}_k = \delta_{kj} \tag{V.59}$$

The components ϑ_j are called the physical components of a vector in the basis $\vec{u}_j$ which is an orthogonal basis. Different 'orthogonal basis' are connected to each other by orthogonal transformations. If y^i denote the cartesian coordinates and X^i the generalized 'orthogonal curvilinear' coordinates (as oXe coordinates with property (V.57) somtimes called), then if follows from (V.49) and (V.57) that

$$g_{jj} = h_j^2 = \sum_k \left(\frac{\partial y_k}{\partial x^j}\right)^2 = \sum_k \left(\frac{\partial x_j}{\partial y_k}\right)^{-2} \qquad \text{(V.60)}$$

In particular the physical components of the 'coordinate differentials (Pfaffians)' are $dQ^j = h_j dx^j$ (no sum over j)) In the following table we summarize some of the well-known curvilinear orthogonal coordinate systems in three dimensions

Curvilinear orthogonal coordinate system as a function of cartesian coordinates	$dQ^1 = h_1 dx^1$	$dQ^2 = h_2 dx^2$	$dQ^3 = h_3 dx^3$
1. Cartesian $y^1, y^2, y^3 = x, y, z$	dx	dy	dz
2. Cylindrical $\rho = \sqrt{x^2+y^2}$ $\varphi = \tan^{-1} y/x$, z	$d\rho$	$\rho\, d\varphi$	dz
3. Spherical Polar $r = \sqrt{x^2+y^2+z^2}$ $\theta = \tan^{-1}\rho/z$, $\varphi = \tan^{-1}\frac{y}{x}$	dr	$r\, d\theta$	$r\sin\theta\, d\varphi$
4. Parabolic $\lambda = \sqrt{r+z}$ $\mu = \sqrt{r-z}$, $\varphi = \tan^{-1}\frac{y}{x}$ $\cos\frac{\theta}{2} = \lambda/\sqrt{\lambda^2+\mu^2}$	$\sqrt{\lambda^2+\mu^2}\, d\lambda$	$\sqrt{\lambda^2+\mu^2}\, d\mu$	$\lambda\mu\, d\varphi$

Just as different orthogonal basis are connected via orthogonal transformations the various physical components of 'coordinate differentials (Pfaffians)' are also similarly connected

$$dQ^j = h_j dx^{(j)} = A^j{}_k \, dy^k \tag{V.61}$$

For the cartesian coordinates the matrix A is identity matrix. For the other three cases listed above it is given by

$$\begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \end{bmatrix},$$

$$\begin{bmatrix} \cos\varphi\cos\frac{\theta}{2} & \sin\varphi\cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \cos\varphi\sin\frac{\theta}{2} & \sin\varphi\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \\ -\sin\varphi & \cos\varphi & 0 \end{bmatrix}. \tag{V.62}$$

The elements of these matrices are t the components of the scalar product matrix $\vec{u}_j(x)\cdot\vec{e}_{ok}$ according to the scheme

$$\vec{u}_j(x) = \sum_k (\vec{u}_j(x)\cdot\vec{e}_{ok})\,\vec{e}_{ok} \tag{V.63}$$

The surface tensor for 3-dimensional subspace has 'physical' components $h_{(j)} h_{(k)} dx^j \wedge dx^k = dS_{jk}$ and the value element is $h_{(j)} h_{(k)} h_{(l)} dx^j \wedge dx^k \wedge dx^l = dV_{jkl}$. Since in three dimensions a completely antisymmetric tensor has just one component the volume element may be written as $dV = h_1 h_2 h_3 \, dx^1 dx^2 dx^3$, where we have omitted the tensor product sign. Similarly the 'surface vector' is $S_\ell = h_{(j)} h_{(k)} \, dx^j dx^k$ with j, k, ℓ cyclic.

V.9. In a Euclidean geometry the distance between two points on a curve is by definition given by integrating $ds = \sqrt{d\vec{y} \cdot d\vec{y}}$ where $\vec{y} = y^i \vec{e}_{0i}$; y^i are the cartesian coordinates and $\vec{e}_{0i}$ is the canonical basis. It may also be written as

$$ds^2 = \vec{e}_{0i} \cdot \vec{e}_{0j} \, dy^i dy^j = \eta_{ij} \, dy^i dy^j \; . \qquad \text{(V.64)}$$

Under a general coordinate transformation $y^i \to x^i = f^i(y^1, \cdots y^n)$, we get

$$\left.\begin{aligned} ds^2 &= \sum_{l,m} \Big(\sum_{ij} \eta_{ij} \frac{\partial y^i}{\partial x^\ell} \frac{\partial y^j}{\partial x^m} \Big) dx^\ell dx^m \\ &= \sum_{l,m} g_{\ell m} \, dx^\ell dx^m \end{aligned}\right\} \qquad \text{(V.65)}$$

where $g_{\ell m}$ are components of this metric tensor

$$g_{\ell m} = \vec{e}_\ell \cdot \vec{e}_m = \sum_{i,j} \eta_{ij} \frac{\partial y^i}{\partial x^\ell} \frac{\partial y^j}{\partial x^m} , \qquad \text{(V.66)}$$

and $dx^l dx^m$ actually mean $dx^l \otimes dx^m$, but since there is no possibility of confusion we shall omit the tensor product sign. For a Euclidean space $\vec{e}_{oi} \cdot \vec{e}_{oj} = \delta_{ij}$ [$=1$ for $i=j$, and 0 for $i \neq j$], but in the pseudo-Euclidean space $\eta_{ij} = \pm 1$ for $i=j$ and 0 for $i \neq j$. The contravariant components of the metric tensor are

$$g^{lm} = e^l \cdot e^m = \sum \eta^{ij} \frac{\partial x^l}{\partial y^i} \frac{\partial x^m}{\partial y^j}, \tag{V.67}$$

where

$$\eta^{ij} = e_o^i \cdot e_o^j \quad \text{and} \quad g^{lm} g_{mk} = \delta^l_k \tag{V.67A}$$

The equation for a straight line in cartesian coordinate is given by

$$\frac{d^2 y^i}{ds^2} = 0 \tag{V.68}$$

We would like to find the corresponding equation in general coordinates. Under the change of coordinates a contravariant tensor transforms as

$$V^j(y) = \frac{\partial y^j}{\partial x^k} V^k(x) \tag{V.69}$$

Let y^j be the cartesian coordinates, then

$$\frac{dV^j(y)}{ds} = \frac{\partial y^j}{\partial x^k} \frac{dV^k(x)}{ds} + \frac{\partial^2 y^j}{\partial x^l \partial x^k} \frac{dx^l}{ds} V^k(x)$$

$$= \frac{\partial y^j}{\partial x^k} \left[\frac{dV^k(x)}{ds} + \frac{\partial x^k}{\partial y^m} \frac{\partial^2 y^m}{\partial x^l \partial x^n} \frac{dx^l}{ds} V^n(x) \right] \tag{V.70}$$

At this point it is convenient to introduce some new symbols, a straight forward computation shows that

$$[lm,n] \stackrel{\text{def}}{=} \frac{1}{2}(g_{nm,l}+g_{nl,m}-g_{lm,n}) \qquad \text{(V.71)}$$
$$= \eta_{ij}\,\partial_n y^i\,\partial_l\partial_m y^j ,$$

where $g_{lm,n} = \partial_n g_{lm} = \frac{\partial}{\partial x^m} g_{lm}$, etc. and

$$\left\{ {k \atop l\,m} \right\} \stackrel{\text{def.}}{=} g^{kn}[lm,n] = \frac{\partial x^k}{\partial y^n}\frac{\partial^2 y^n}{\partial x^l \partial x^m}. \qquad \text{(V.72)}$$

The symbols $[lm,n]$ and $\left\{ {k \atop lm} \right\}$ are called Christoffel symbols of the first and second kind respectively. They are both symmetric with respect to the interchange of indices l, m. Under a change of the coordinates $x^i \to \bar{x}^i = F^i(x^1, x^2 - - - - x^n)$, it is straightforward to show that

$$\overline{\left\{ {k \atop l\,m} \right\}} = \frac{\partial \bar{x}^k}{\partial x^n}\frac{\partial x^i}{\partial \bar{x}^l}\frac{\partial x^j}{\partial \bar{x}^m}\left\{ {n \atop i\,j} \right\} + \frac{\partial \bar{x}^k}{\partial x^r}\frac{\partial^2 x^r}{\partial \bar{x}^l \partial \bar{x}^m}, \qquad \text{(V.73)}$$

From this it is clear that the Christoffel symbols are not tensors, but can be made to vanish in a suitable coordinate system (a cartesian system for Euclidean space). In this notation eqn. (V.70) read

$$\frac{dV^j(y)}{ds} = \frac{\partial y^j}{\partial x^k}\left[\frac{dV^k(x)}{ds} + \left\{ {k \atop l\,m} \right\}\frac{dx^l}{ds} V^m(x)\right],$$
$$= \frac{\partial y^j}{\partial \bar{x}^k}\left[\frac{d\bar{V}^k(\bar{x})}{ds} + \overline{\left\{ {k \atop l\,m} \right\}}\frac{d\bar{x}^l}{ds}\bar{V}^m(\bar{x})\right]. \qquad \text{(V.70A)}$$

From this equation we can draw several conclusions.

(A) If we denote the expressions in the square brackets by $\frac{V^k(x)}{ds}$ and $\frac{\bar{V}^k(\bar{x})}{ds}$ then we see that under the coordinate change $x^j \rightarrow \bar{x}^j$,

$$\frac{DV^k(x)}{ds} = \frac{\partial x^k}{\partial \bar{x}^l} \frac{D\bar{V}^l(\bar{x})}{ds} \tag{V.71}$$

Since V^k are components of a vector field and s is the distance parameter along a curve, we call

$$\frac{DV^k}{ds} \stackrel{def}{=\!=} \frac{dV^k}{ds} + \left\{ {k \atop l\ m} \right\} \frac{dx^l}{ds} V^m \tag{V.72}$$

as the absolute derivative of V^k and transforms as a contra-variant vector. The generalization of the equation of a straight line, eqn. (V.68) in generalized coordinates is then given by

$$\frac{D}{ds}\left(\frac{dx^k}{ds}\right) \stackrel{def.}{=\!=} \frac{d^2x^k}{ds^2} + \left\{ {k \atop l\ m} \right\} \frac{dx^l}{ds}\frac{dx^m}{ds} = 0 \tag{V.68A}$$

and is called the equation of a geodesic or of a "geodetic line".

B) One can rewrite the absolute derivative as

$$\left.\begin{aligned} &\frac{DV^k}{ds} \stackrel{def}{=\!=} \Delta_l V^k \frac{dx^l}{ds}, \\ &\text{where,} \\ &\Delta_l V^k \stackrel{def}{=\!=} V^k_{;l} = v^k_{,l} + \left\{ {k \atop m\ l} \right\} V^m \\ &v^k_{,l} = \partial_l v^k = \partial v^k/\partial x^l \end{aligned}\right\} \tag{V.73}$$

The set of n^2 quanties $V^k{}_{;l}$ transform as components of mixed tensor of second rank and is called <u>covariant derivative</u> of a contravariant vector. Similarly one finds for the covariant derivative of a covariant vector

$$\Delta_l V_k = V_{k;l} = V_{k,l} - \left\{ {m \atop lk} \right\} V^m . \tag{V.74}$$

Covariant derivatives of higher rank tensors can also be written down by direct computation; thus e.g.

$$\begin{aligned}
\Delta_l V^{mn} &= V^{mn}{}_{;l} = V^{mn}{}_{,l} + \Gamma^m_{kl} V^{kn} + \Gamma^n_{kl} V^{mk} \\
\Delta_l V^m{}_n &= V^m{}_{n;l} = V^m{}_{n,l} + \Gamma^m_{kl} V^k{}_n - \Gamma^k_{nl} V^m{}_k \\
\Delta_l V_{mn} &= V_{mn;l} = V_{mn,l} - \Gamma^k_{ml} V_{kn} - \Gamma^k_{ln} V_{mk}
\end{aligned} \tag{V.75}$$

where we have put $\Gamma^k_{ml} \equiv \left\{ {k \atop m\,l} \right\}$ for convenience.

V.10. Using the equation (V.74 and 75) we want to evaluate the expression

$$V^k{}_{;l;m} - V^k{}_{;m;l} = +R^k{}_{pml} V^p . \tag{V.76}$$

Direct substitution yields,

$$\begin{aligned}
R^k{}_{plm} &= -\Gamma^k_{pl,m} + \Gamma^k_{pm,l} + \Gamma^k_{nl}\Gamma^n_{mp} - \Gamma^k_{nm}\Gamma^n_{lp} \\
&= \partial_{[l}\Gamma^k_{m]p} + \Gamma^k_{n[l}\Gamma^n_{m]p} .
\end{aligned} \tag{V.77}$$

where the $[l \quad m]$ stands for antisymmetrization:

$\partial_{[l}\Gamma^k_{m]p} = \partial_l \Gamma^k_{mp} - \partial_m \Gamma^k_{lp}$ etc. Since left hand side of (V.76) is a difference of two tensors of the same type (mixed of rank three) it is also a tensor. The right hand side is therefore also a tensor. By quotient rule it follows that since V^k are components of a vector, the set of quantities $R^k{}_{pml}$ are components of a fourth rank tensor. The tensor $R^k{}_{pml}$ is called the Riemann-Christoffel curvature tensor. If we lower one of its indices, we get

$$
\begin{aligned}
R_{hijk} &= \partial_j [ik,h] - \partial_k [ij,h] + [hk,a] \begin{Bmatrix} a \\ i\,j \end{Bmatrix} \\
&\qquad - [hj,a] \begin{Bmatrix} a \\ i\,k \end{Bmatrix} \\
&= \tfrac{1}{2}\left(\partial_i \partial_j g_{hk} + \partial_h \partial_k g_{ij} - \partial_i \partial_k g_{hj} - \partial_h \partial_j g_{ik}\right) \\
&\quad + g^{lm}\left([ij,m][hk,l] - [ik,m][hj,l]\right)
\end{aligned}
\tag{V.78}
$$

It is easy to see that it has the following symmetry properties

$$
\left.
\begin{aligned}
& R_{hijk} + R_{ihjk} = R_{(hi)jk} = 0 \\
& R_{hijk} + R_{hikj} = R_{hi(jk)} = 0 \\
& R_{hijk} = R_{jkhi}, \quad R_{hijk} + R_{hjki} + R_{hkij} = 0
\end{aligned}
\right\}
\tag{V.79}
$$

The first three conditions tell us that the Riemann tensor is equivalent to a symmetric tensor of the second rank in $d = \frac{1}{2}n(n-1)$ dimensions. Hence it's components are $\frac{1}{2}d(d+1)$ in number.

From this one must subtract the number of conditions due to the fourth identity. Since the last three indices are cyclically symmetrized in this equation, it is equivalent to a complete antisymmetric tensor of the dourth rank which has $\binom{n}{4}$ independent components. Hence the number of algebraically independent components of the Riemann tensor are

$$\frac{1}{8}n(n-1)(n^2-n+2)-\binom{n}{4}=\frac{1}{12}n^2(n^2-1). \qquad (V.80)$$

In addition to the algebraic identities discussed above one can show by straightforward computation that the Riemann tensor also satisfies the following differential identities

$$R^h{}_{ijk;m}+R^h{}_{ikm;j}+R^h{}_{imj;k} \quad \text{(Bianchi)} \qquad (V.81)$$

$$R^h{}_{ijk;m}+R^h{}_{jmi;k}+R^h{}_{mkj;i}+R^h{}_{kim;j} \quad \text{(Veblen)}. \qquad (V.82)$$

At this point it is useful to note that it follows from the definition of a covariant derivative and the Cristoffel symbol of the second kind that the covariant derivative of the metric tensor vanishes identically

$$g_{ij;k}=0\,, \quad g^{ij}{}_{;k}=0. \qquad (V.83)$$

We shall use the properties (V.81-83) to construct an identity for a second rank tensor which is of considerable importance in the theory of general relativity. Referring to (V.77) and contracting on k and m we get

$$R_{p\ell} = \partial_p \Gamma^k_{\ell k} - \partial_k \Gamma^k_{p\ell} + \Gamma^k_{m\ell}\Gamma^m_{pk} - \Gamma^k_{mk}\Gamma^m_{p\ell}$$

$$= \partial_p\partial_\ell \ln\sqrt{g} - \partial_k \Gamma^k_{p\ell} + \Gamma^k_{m\ell}\Gamma^m_{kp} - \Gamma^m_{p\ell}\,\partial_m \ln\sqrt{g} \qquad \text{(V.84)}$$

where $g = \det[g_{ij}]$ and $\Gamma^k_{\ell k} = \partial_\ell \ln\sqrt{g}$. ~~The~~ $R_{p\ell}$ is a symmetric tensor of the second rank and is called the Ricci tensor. We note that

$$R_{p\ell} = R^k{}_{p\ell k} = R_{pkm\ell}\, g^{km} = -R^k{}_{pk\ell} \,. \qquad \text{(V.85)}$$

Returning to (V.81) and contracting on h and m we get $R^h{}_{ijk;h} = R_{i[j;k]}$ on multiplying throughout by g^{jk} and using (V.83) we obtain

$$G^h{}_{j;h} \overset{\text{def.}}{=\!=} \left(R^h{}_j - \tfrac{1}{2} R\, \delta^h{}_j\right)_{;h} = 0 \,, \qquad \text{(V.86)}$$

where we have put $R = g^{ik} R_{ik}$ (scalar curvature), and $G^h{}_j$ are components of the Einstein tensor

$$G^h{}_j = R^h{}_j - \tfrac{1}{2} R\, \delta^h{}_j$$

The entire analysis of the detailed properties of the Riemann tensor has a glaring draw back. It is based on the initial assumption that the space is (pseudo-) Euclidean. For such a space we have seen that it is possible to choose a basis such that g_{ij} are constants everywhere and therefore the Riemann tensor vanishes identically. Conversely if the Riemann tensor vanishes, then

$$(\Delta_j\Delta_k - \Delta_k\Delta_j)V^i = 0$$

and the covariant derivatives commute as one would expect in a Euclidean space. Hence the vanishing of the Riemann tensor is necessary and sufficient condition for the space to be (pseudo) Euclidean or Flat.

We emphasise that the various formulas obtained for the absolute and covariant derivatives and for the Riemann tensor continue to be valid for an arbitrary space to which is assigned a twice covariant metric tensor field g_{ij} so as to provide a measure of the Pythagorean distance according to the eqn. (V.65); such a space is called a Riemannian space.

VI. Generalization of Euclidean Structure

VI.1 Suppose we are given a general space can one do things on it as we did in the case of a Euclidean space? Recall that if we have an abstract vector space V_n one may translate it into K^n or R^n by the mapping $V_n \to R^n$; then since R^n has some very natural properties familiar to us we could do many things. In fact in this case because of the linear structure of a vector space we end up with a Euclidean space. Similarly if we have an abstract group G one makes a homomorphic map of G into another group or a linear space H (many to one map from G to H such that not every element of H is necessarily an image of G). The homomorphic image then gives the representation of G . It is clear from these examples that one must prescribe some such map for our general space.

Now in case of the number space R, if $a < b$ are two numbers, we define an open interval as $a < x < b$. If we suitably devide the number space into a denumerable set of intervals then any open set of R (including R and the empty set) can be expressed as a union of such intervals. The set of all the open intervals and their unions is said to provide the natural topology of R and the basic set of open intervals considered gives the base of the topology. One can now discuss the questions of continuity etc. in terms of open sets. An important property of the real line is that if we define the neighbourhood of a point as an open set then for two distinct points there exist neighbourhoods whose intersection is empty (Housdorf topology). We assume all these properties for our general space $\mathcal{S}$ and call it a topological space $\mathcal{S}$ (this is not a definition)

DEFINITION. If there exists a homeomorphic mapping (one to one map such that the mapping and its inverse are both continuous) of a neighbourhood $U(p)$ of every point p of the topological space $\mathcal{S}$ into a neighbourhood V of a point P in a Euclidean space R^n, then $\mathcal{S}$ is called a manifold and n is the dimension of $\mathcal{S}$ (Recall here that a linear manifold is a vector subspace).

If we define a set $\mathcal{D}$ of real valued functions on $\mathcal{S}$ with values in $\mathbb{R}$ then via the homeomorphic mapping of a neighbourhood $U(p)$ in $\mathcal{S}$ to a corresponding neighbourhood of a point in R^n one may specify coordinates in $\mathcal{S}$; thus

$$f^a(p) = F^a(x^1(p), x^2(p), \ldots, x^n(p)); \quad \text{(VI-1)}$$

f^a, x^k are real valued functions on $\mathcal{S}$ (i.e. they are in D) and F^a are functions from $R^n \to R$.

In this manner we introduce coordinates for each neighbourhood of the points of $\mathcal{S}$ and the coordinates so introduced are elements of R^n. If the coordinates are n-fold differentiable functions we say that $\mathcal{S}_n$ is a differentiable manifold of class C^l. In general we shall assume it to be a C^∞ manifold.

For two different neighbourhoods of a point p in S' we get different coordinates in R^n. Since the corresponding neighbourhoods have points in common (for instance the point P) it is possible to introduce the concept of coordinate transformations. Since these are all assumed to be differentiable functions, the Jacobian matrices defined by

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

are non-singular and the transformation in R^n is defined by $x^a \rightarrow x^{a'} = F^{a'}(x^1, \cdots x^n)$, where $x^a, x^{a'}$ are in R^n and the transformation is defined only in a certain neighbourhood (an open subset in) of P.

<u>VI.2</u>. <u>Def</u>. Let $-\epsilon < t < +\epsilon$ be an open interval in R. Then the $R \rightarrow S'$ differentiable map $t \rightarrow p(t)$ for points p in S' is called a differentiable curve in $S' : \sigma(t)$. The tangent vector to $\sigma(t)$ at p_0 is a map from the set of all real valued differentiable functions $\{f\}$ defined in the neighbourhood of p_0 into R and is given by

$$\vec{u}_{p_0}(f) = \frac{d}{dt} f \cdot p(t) \Big|_{t=t_0} \qquad (2)$$

In the local coordinate system in R^n, this may be computed as

$$\vec{u}_{p_0}(f) = \frac{d}{dt} F(x^1, \cdots, x^n) \Big|_{p_0} = \frac{dx^a}{dt} \frac{\partial F}{\partial x^a} \Big|_{p_0}, \qquad (3)$$

The set of all tangent vectors at p_0 ($T_{p_0}(S^1)$ has a vector space structure, since

(1) $$(\vec{U}_{p_0}+\vec{V}_{p_0})f = \vec{U}_{p_0}(f) + \vec{V}_{p_0}(f) \qquad (4)$$

(2) $$(a\vec{U}_{p_0})f = a\vec{U}_{p_0}(f) \qquad (5)$$

(3) As (117) is true for a whole class of F we can write

$$\vec{U}_{p_0} = \frac{dx^j}{dt}\frac{\partial}{\partial x^j}\Big|_{p_0} = U^j\vec{e}_j\Big|_{p_0}, \qquad (6)$$

The dimension of this vector space is therefore clearly n and the — natural basis is given by

$$\vec{e}_j = \frac{\partial}{\partial x^j} \qquad (7)$$

We note that in addition to the properties (1),(2),(3) of the vector space the tangent vectors satisfy

$$\vec{U}(fg)\Big|_{p_0} = f_{p_0}U(g) + g_{p_0}U(f). \qquad (8)$$

We shall often refer to such action as a derivation.

In order to define the dual space $T^*_{p_0}(S^1)$ it has to be kept in mind that this property (V.8) is again satisfied. We call the elements of $T^*_{p_0}(S^1)\{\omega,\pi,\cdots\}$ as covectors or differential forms at p_0 and $T^*_{p_0}$ is itself called the <u>cotangent</u> vector <u>space</u>. If $\vec{U}_p$ is an element of T_p then the map $T_p \to R$ defined by $\vec{U}_p \to \vec{U}_p(f)$ is a linear map and is locally a map of $\mathcal{S}_n \to R$, since f is in $\mathcal{D}$. Hence it determines a differential form at p . If we denote it by df , then

$$df(\vec{U}_p) \overset{\text{def.}}{=\!=} \vec{U}_p(f) \tag{9}$$

and satisfies

$$d(fg)(\vec{u}) = g\, df(\vec{u}) + f\, dg(\vec{u}) \tag{10}$$

in virtue of (V.8). Since this holds for all $\vec{u}$,

$d(fg) = df \cdot g + f\, dg$ In particular if we choose for f the local coordinates $\{x^a\}$

$$dx^a(\vec{u}) = u(x^a) = u^a. \tag{11}$$

But $\vec{u} = u^a \vec{e}_a$, so that $dx^a(\vec{u}) = u^b dx^a(e_b) = u^a$; hence

$$dx^a(\vec{e}_b) = \delta^a_b \quad , dx^a = e^a \tag{12}$$

Thus $\{dx^a\}$ is a basis dual to $\{\partial/\partial x^a\}$ and any differential form at p can be written as

$$\omega = \omega_a dx^a \tag{13}$$

In particular if ω is a complete differential, ω_a is of the form $\partial f/\partial x^a$ with f a real valued function on $\mathcal{S} \longrightarrow R$ (ie. in $\mathcal{D}$).

The directional derivative of f in the direction of a vector $\vec{a}$ is said to be the 'value' of df on $\vec{a}$ and is written $df(\vec{a}) = a^j \partial f/\partial x^j$. Whereas the first derivative of f is $df = dx^j \partial f/\partial x^j$ the second derivative is written as

$$d^2 f = \frac{\partial f}{\partial x^i \partial x^j} dx^i \otimes dx^j \tag{14}$$

(see further section V.7

Now we can define a vector field on $\mathcal{S}$ as a correspondence which associates with every point p in $\mathcal{S}$ and every system of local coordinates $\{x^j\}$ around p a set of real numbers a^j in R^n which transform under the coordinates as components of a contravariant vector. In a similar fashion one can also define higher rank tensor fields. The covariant tensor fields then arise as differentiable form fields on $\mathcal{S}$.

VI.3 In the spirit of the above formulation, an affine connection is introduced as the map $T_p(\mathcal{S}) \times \mathcal{S}$ given by

$$(u, v) \longrightarrow \vec{u} \cdot \nabla \vec{v} = u^a \vec{e}_b \cdot \nabla_a v^b \tag{15}$$

where,

$$\vec{e}_a \cdot \nabla \vec{e}_b = \Gamma^c_{ab} \vec{e}_c, \quad \nabla_a v^b = \partial_a v^b + \Gamma^b_{ac} v^c. \tag{16}$$

$\nabla_a v^b$ the covariant derivative of a contravariant vector field. Similarly one can define a covariant derivative of a covariant vector field and generalize the notion to tensor fields of higher rank.

If $\vec{u}$ is a tangent vector field defined on a curve $\sigma(t)$ then the absolute derivative of a vector field V in $\mathcal{S}$ along the curve is given by

$$\frac{D\vec{v}}{dt} = \vec{u}\cdot\nabla\vec{v} = \vec{e}_a\left(\frac{dV^a}{dt} + \Gamma^a_{bc}V^b\frac{dx^c}{dt}\right). \qquad (17)$$

If the absolute derivative of V vanishes along σ we say that $\vec{V}$ is parallel along σ and the value of $\vec{V}$ at one point on σ determines its value at any other point on σ by parallel transport.

If σ is such that the absolute derivatives of $\vec{u}$ vanishes along it then σ is called a geodesic and t is called the affine parameter. In a local coordinate system is given by

$$\frac{d^2x^j}{dt^2} + \Gamma^j_{\ell m}\frac{dx^\ell}{dt}\frac{dx^m}{dt} = 0. \qquad (18)$$

The curvature tensor can be introduced through the mapping $\mathcal{S}' \times \mathcal{S}' \times \mathcal{S}' \to \mathcal{S}'$ given by

$$(\vec{u},\vec{v},\vec{w}) \to R_p(\vec{u},\vec{v},\vec{w})$$
$$= \vec{W}\cdot\nabla(\vec{V}\cdot\nabla\vec{u}) - \vec{V}\cdot(\vec{W}\cdot\nabla\vec{u}) + (\vec{V}\cdot\nabla\vec{W})\cdot\nabla\vec{u} - (\vec{W}\cdot\nabla\vec{V})\cdot\nabla\vec{u}$$
$$= \vec{e}_a R^a{}_{bcd}\, u^b v^c w^d \qquad (19)$$

In local coordinates the components of the curvature tensor are given by

$$R^a{}_{bcd} = \partial_{[d} \Gamma^a{}_{c]b} + \Gamma^a{}_{de} \Gamma^e{}_{cb} - \Gamma^a{}_{ce} \Gamma^e{}_{db}. \qquad (20)$$

We remark that the components Γ^a_{bc} of a general affine connection can be decompared into symmetric and antisymmetric parts

$$\Gamma^a{}_{bc} = \gamma^a{}_{bc} + L^a{}_{bc} \qquad (21)$$

where $L^a_{bc} = -L^a{}_{cb}$ are the components of a tensor of third rank.

VI.4 . The basic mathematical notions introduced above even though rather general enable us only to deal with local questions in the neighbourhood of a point p in S . Many relativists feel that it is necessary for a full understanding of relativity, (particularly in connection with quantization of relativity) to have more general structure that would enable one to treat questions of global nature. But here we are not interested in these. Our interest here is local Riemannian geometry. A Riemannian space is as an open set in a cartesian space in which is defined a symmetric metric

$$ds^2 = g_{ij}(x)\, dx^i \otimes dx^j \Rightarrow g_{ij}\, dx^i dx^j \qquad (22)$$

A Reimannian manifold is a manifold on which is defined a symmetric twice covariant tensor field $g = g_{ij}(x)$

In the terminology introduced above the metric tensor field g is given on a differentiable manifold S' if there is defined in the tangent space at every point p in S' a scalar product

$$\vec{u}\cdot\vec{v} = g(u,v) = g_{ij}\,u^i v^j. \tag{23}$$

From this difinition it is clear that in the tangent space the scalar product induced by the coordinates is Euclidean one. A metric affine connections exists if for any differentiable curve $\sigma(t)$ and any vector field $\vec{V}$ parallel along σ we have

$$\frac{d}{dt}g(\vec{v},\vec{v}) = \frac{dg_{ij}}{dt} + g_{ij}\left(\frac{dv_i}{dt}v_j + v_j\frac{dv_i}{dt}\right) = 0 \tag{24}$$

Since $\vec{v}$ is parallel along σ, $\frac{D\vec{v}}{dt} = 0$ and in the local coordinate system $\frac{dv^j}{dt} = -\Gamma^j_{\ell m}\, v^\ell u^m$, where $u^m = dx^m/dt$. Hence if we write $dg_{ij}/dt = u^\ell u^m\, \partial_\ell g_{ij}$, we get

$$\frac{d}{dt}g(\vec{v},\vec{v}) = v^i v^j u^\ell (\nabla_\ell g_{ij}) = 0. \tag{25}$$

It follows that

$$\nabla_\ell g_{ij} = \partial_\ell g_{ij} - \Gamma^m_{i\ell}\, g_{mj} - \Gamma^m_{j\ell}\, g_{im} = 0 \tag{26}$$

If we do not assume the symmetry of $\Gamma^{m}_{i\ell}$ then one has to be careful about the order of indices and we write in full

$$\nabla_i g_{j\ell} = \partial_i g_{j\ell} - \Gamma^{m}_{ji} g_{m\ell} - \Gamma^{m}_{\ell i} g_{jm} \quad , \tag{27A}$$

$$\nabla_j g_{\ell i} = \partial_j g_{\ell i} - \Gamma^{m}_{\ell j} g_{mi} - \Gamma^{m}_{ij} g_{\ell m} \; . \tag{27B}$$

If we subtract the first of these equations from the sum of the last two equations we obtain on multiplying throughbout by $\frac{1}{2}$,

$$\Gamma^{\ell}_{ij} = \left\{ {\ell \atop i\,j} \right\} + g^{\ell p}(g_{im} L^{m}_{\;\,jp} + g_{jm} L^{m}_{\;\,ip}) + L^{\ell}_{\;\,ij} \, . \tag{28}$$

Thus we see that the symmetric part is given by (28):

$$\gamma^{\ell}_{ij} = \left\{ {\ell \atop i\,j} \right\} + g^{\ell p}(g_{im} L^{m}_{\;\,jp} + g_{jm} L^{m}_{\;\,ip}) . \tag{29}$$

If $L^{\ell}_{\;\,ij}$ vanishes then the symmetric part is completely determined by the metric tensor field g_{ij}. Hence if the affine connection is symmetric and a metric tensor field is given then the latter completely determines the affine connection. The symbols

$$\begin{aligned} \left\{ {i \atop j\,k} \right\} &= g^{i\ell} \cdot \tfrac{1}{2}(g_{\ell j,k} + g_{\ell k,j} - g_{jk,\ell}) \\ &= g^{i\ell} \; [jk,\ell] \end{aligned} \tag{30}$$

are called christoffed symbols of the second kind and $[jk,l]$ of the first kind.

At this point it is useful to note that as far as the geodesic equation is concerned the nonsymmetric part is in any case not relevant. It is therefore amusing to consider a situation when the affine connection is given as

$$\gamma^{l}_{ij} = \left\{ {l \atop ij} \right\} + g^{lk} T_{lij} \tag{31}$$

where T_{ij} is a tensor symmetric in ij. Let the geodesic equation corresponding to this be given by

$$\frac{d^2 x^l}{dt'^2} + \gamma^{l}_{ij} \frac{dx^i}{dt'} \frac{dx^j}{dt'} = 0 \tag{32}$$

We ask under what restrictions on is the solution of (32) identical to that given by

$$dt^2 = g_{ij} \, dx^i dx^j \, . \tag{33}$$

If we multiply (144) by $g_{lm} dx^m / dt'$ it takes the form

$$\frac{d}{dt'}\left[g_{ij} \frac{dx^i}{dt'} \frac{dx^j}{dt'} \right] + T_{mij} \frac{dx^m}{dt'} \frac{dx^i}{dt'} \frac{dx^j}{dt'} = 0, \tag{34}$$

which shows that necessary condition that the parameter along a geodesic is given by (33) is that γ^{l}_{ij} is symmetric of the form (32) and such that the tensor T_{lij} satisfy

$$T_{\ell i j} + T_{i j \ell} + T_{j \ell i} = 0, \tag{35}$$

We make several remarks regarding the nature of the curvature tensor in various cases. For the general affine connections the formula (V.76) does not hold, we get instead

$$\nabla_i \nabla_j V^k - \nabla_j \nabla_i V^k = -B^k_{\ell j i} V^\ell - 2K^\ell_{ji} V^k_{;\ell} \tag{36}$$

when

$$B^k_{\ell(ji)} = B^k_{\ell ji} + B^k_{\ell ij} = 0 \tag{37}$$

There are $n^2\binom{n+1}{2}$ conditions and hence $B^k_{\ell ji}$ has $n^2\binom{n}{2}$ independent components. The curvature tensor can be written as a sum of two tensors one depending on the symmetric part and the other on the antisymmetric part. For the symmetric affinity, (V.76) type relation holds:

$$(\nabla_i \nabla_j - \nabla_j \nabla_i) V^k = -B^k_{\ell j i} V^\ell. \tag{38}$$

The $B^k_{\ell j i}$ now satisfy the following symmetry conditions

(a) $B^k_{\ell(ji)} = 0$, (37A)

(b) $B^k_{(\ell j i)} = B^k_{\ell j i} - B^k_{j i \ell} + B^k_{i \ell j} = 0$. (39A)

The first of these conditions impose $n^2\binom{n+1}{2}$ condition. The tensor $U^k_{\ell j i} \equiv B^k_{(\ell j i)}$ in light of the conditions (a) is completely antisymmetric in all the three indices $l, \hat{j}, \check{i}$; hence the conditions (b) are $n\binom{n}{3}$ in number. The number of independent components therefore are $n^4 - n^2\binom{n+1}{2} - n\binom{n}{3} = \frac{n^2(n^2-1)}{3}$. When the affinity is completely determined by the metric one denotes the curvature tensor by the symbol $R^k{}_{l\dot{j}\dot{i}}$ and is called Reimann-Christoffel tensor. Using the metric tensor for lowering indices one finds that $R_{kl\dot{j}\dot{i}}$ is also antisymmetric with respect to the interchange of k and l. Thus

(a) $B^k{}_{l(\dot{j}\dot{i})} = 0$ $\quad n^2\binom{n+1}{2}$ conditions (37B)

(b) $B^k{}_{(l\dot{j}\dot{i})} = 0$ $\quad n\binom{n}{3}$ conditions (39B)

(c) $B_{(kl)\dot{j}\dot{i}} = 0$ (40)

Since $B_{(kl)\dot{j}\dot{i}}$ is antisymmetric in $\dot{j}\dot{i}$ it has $\frac{n(n+1)}{2} \cdot \frac{n(n-1)}{2}$ independent components. Hence the number of independent components of the Riemann tensor is

$$n^4 - n^2\binom{n+1}{2} - n\binom{n}{3} - \binom{n+1}{2}\binom{n}{2} = \frac{n^2(n^2-1)}{12}.$$

From the condition (a) and (c) it follows the symmetry conditions

$$\text{(d)} \quad R_{kl\dot{i}\dot{j}} = R_{\dot{i}\dot{j}kl}. \tag{41}$$

Now according to (a) and (c) each pair of indices $\dot{i}\dot{j}$ and kl represent $\frac{1}{2}n(n-1) = \binom{n}{2}$ independent components. According to (d), $R_{\dot{i}\dot{j}kl}$ is symmetric with respect to the interchange of the pairs $\dot{i}\dot{j}$ and kl, hence by (a), (c), (d) the number independent components is

$$\frac{1}{2}\binom{n}{2}\left[\binom{n}{2}+1\right] = \frac{1}{8}n(n-1)\left[n(n-1)+2\right].$$

In the light of conditions (a),(c) and (d) we see that $U_{klji} = R_{k(lji)}$ is completely anti-symmetric, with respect to interchange of any two indices ; it therefore represents $\binom{n}{4}$ conditions. Hence the number of independent components of the Riemann tensor ~~an~~ are

$$\frac{1}{2}\binom{n}{2}\left[\binom{n}{2}+1\right]-\binom{n}{4} = \frac{n^2(n^2-1)}{12},$$

as before.

The curvature tensor in general has two contractions, one anti-symmetric and other nonsymmetric. When the connection is symmetric the sum of the nonsymmetric and antisymmetric contractions is symmetric. For the metric connection the antisymmetric contraction vanishes identically and we are left with the symmetric contraction R_{ik} which is called the Rieci tensor.

VI.6. In this section we want to establish connection with considerations of chapter V. We have defined a Riemannian space as an open set in R^n in which is defined a symmetric metric

$$ds^2 = g_{ij}(x)\, dx^i \otimes dx^j. \tag{42}$$

Since g_{ij} is a symmetric matrix there always exists a similarity transformation which can diagenalize it. In particular at some point P_0 there exists a coordinate system and a transformation matrix $J^{\ell}_{k'}(x, x') = \dfrac{\partial x^{\ell}}{\partial x^{k'}}$ such that

$$\left[\frac{\partial x^{\ell}}{\partial x^{k'}}\right] g_{\ell m}(p_0) \left[\frac{\partial x^{m}}{\partial x^{n'}}\right] = \eta_{k'n'}(p_0). \qquad (43)$$

Since $x^{k'} = f^{k}(x^1, \ldots x^n)$ (by implicit function theorem) it follows that this condition (43) may be satisfied in some neighbourhood of p_0 (though not outside it). In the tangent space at T_{p_0} the scalar product $\vec{u}\cdot\vec{v} = u^i \eta_{ij} v^j$ induced by the primed coordinates is therefore euclidean. Hence the covariant and contravariant tensors of the same rank may be identified as in (Pseudo) euclidean geometry and various considerations of sections V.9 and V.10 carry over with the appropriate proviso that the property (43) and hence the choice of the cartesian system as assumed in eqns. (V.64 - V.72) is valid only in the neighbourhood of p_0 and not outside it.

Considerations of chapter V utilized the concept of frame vectors. Here also in each tangent space T_{p_0} we can choose different frames and as before the scalar product is

$$g_{jk} = \vec{e}_j \cdot \vec{e}_k . \qquad (44)$$

We remarked that property (43) can be made to hold only in neighbourhood of a point. On the other hand if the transforming matrix J can not be written down in the form $\partial x^{\ell}/\partial x'^{k}$ there exists no coordinate system in which this can be done. Suppose there exists a matrix A such that there exists 'orthogonal frame vectors' $\vec{u}_j$, and

$$\vec{u}_a = A^j_a \vec{e}_j \tag{45}$$

$$\vec{u}_a \cdot \vec{u}_b = \eta_{ab} \tag{46}$$

holds everywhere. We refer to such frames $\vec{u}_a$ as non-holonomic frames; the reason for this terminology is that the integrability condition (V.51) is not satisfied so that the solution in the form of (V.52) is not possible. Consequently also the vectors $\vec{u}_a$ in general can not be chosen as tangents to parameter lines.

Consider now a differentiable curve $\vec{X}(t)$ through the point $p_0 \ (= X(t=0) = X_0)$. The tangent vector to our curve is

$$d\vec{X} = \vec{e}_j \, dx^j \, , \tag{47A}$$

$$d\vec{X} = \vec{u}_a \, \omega^a ; \tag{47B}$$

where ω^a is a pfaffian given in terms of coordinate differential as $(A^{-1})^a{}_j \, dx^j = \omega^a$. Equation (42) now takes the form

$$ds^2 = g_{ij} \, dx^i dx^j = \sum_a (\omega^a)^2 \tag{48}$$

Since $\vec{u}_a$ prescribes an orthogonal frame by considerations at the end of section V.5, the frenet equations for the curve will have the form (see eqns. V.38-40)

$$d\vec{u}_a = \Omega_a{}^b \vec{u}_b \,, \tag{49}$$

where the Cartan matrix $\Omega_a{}^b$ satisfies $\Omega_a{}^b \eta_{bc} = \Omega_{ac} = -\Omega_{ca}$. Differentiating (45) gives $[d\vec{u}] = dA[\vec{e}] + A\,d[\vec{e}]$ where $[\vec{e}]$ denotes a colomn vector with components $\vec{e}_a$ and A is the matrix in eqn.(45). One may rewrite this with the help of (46) and (49) as

$$de_j = \Gamma_j{}^k e_k \tag{50}$$

where the matrix Γ is given by

$$\Gamma = A^{-1}\left[\Omega - dA\,A^{-1}\right]A \tag{51}$$

If we recognise that $dA\,A^{-1}$ is the cartan matrix $C(A)$ of A, (51) may be rewritten as

$$\Omega = A\Gamma A^{-1} + C(A) \tag{51A}$$

It is clear from this and the formula (V.41) for the composition of cartan matrices that Γ transforms like a cartan matrix. Since ω^a are pfaffiams and so are $\Omega_a{}^b$, the matrix Γ may be written as

$$\Gamma^i{}_j = \Gamma^i{}_{jk}\,dx^k. \tag{52}$$

As an axiom of Riemanian geometry are now assume that the exterior derivative of (47) vanishes

$$dd\vec{X} = d\vec{e}_j \wedge dx^j = \vec{e}_k \Gamma^k_{jl} dx^l \wedge dx^j = 0 \quad (53)$$

From this it follows that Γ^k_{jl} is symmetric in j and l. From (50) and (52) we see that hence on differentiating (44) we obtain; $\partial_k \vec{e}_j = \Gamma^i_{jk} \vec{e}_i$;

$$\partial_k g_{ij} = \Gamma_{i,jk} + \Gamma_{j,ik} ; \; g_{il}\Gamma^l_{jk} = \Gamma_{i,jk} \quad (54)$$

as before.

Let $\vec{U}_{oa}$ be an orthonormal frame in the tangent space T_{p_0} at p_0. Introduce a standard euclidean space R^n taking $\vec{U}_{oa}$ as the basis. Then each tangent space T_p is the image of R^n under the map $\vec{U}_{o\alpha} \xrightarrow{M} \vec{U}_\alpha(p)$ and $\omega = C(M)$. Let $\vec{V}_0$ be a vector in R^n and Let $\vec{V}$, $\vec{\bar{V}}$ be the images of $\vec{V}_0$ in the tangent spaces T_p, $T_{\bar{p}}$; then it follows that

$$\vec{\bar{V}} = M(\bar{x}) M^{-1}(x) \vec{V} \quad (55)$$

More generally this relationship holds it $\vec{V}$, $\vec{\bar{V}}$ are parallel. If a vector $\vec{V}_0$ is transported along a curve $\vec{X}(t)$ such that it remains parallel to itself then clearly

$$\vec{V}[X(t)] = \vec{V}_0 + \int_{t_0}^{t} dV[X(t)] = M[X(t)]\, M^{-1}(x_0) \vec{V}_0 \; ; \quad (56)$$

differentiating along the curve (or one may expand $M[x(t)]$ in a Taylor series), we get

$$\left.\begin{aligned} dV[x(t)] &= dM[x(t)]\, M^{-1}[x(t)]\, \vec{V}[x(t)] \\ &= \omega[x(t)]\, \vec{V}[x(t)]. \end{aligned}\right\} \qquad (57)$$

If the vectors V and V_0 are not parallel then it follows that the operation

$$\Delta V = dV - \omega V \qquad (58)$$

gives the measure of deviation of a vector from parallelism. In a local coordinate system one may compute it formally as follows.

$$\vec{e}_j DV^j \overset{\text{def}}{=\!=} d(V^j \vec{e}_j) = dV^j \vec{e}_j + \Gamma^j{}_k e_j V^k, \qquad (59)$$

where we have used eqn. (50). On writing $DV^j = dx^\ell \Delta_\ell V^j$ we obtain the covariant derivative of a contravariant vector:

$$\Delta_\ell V^j = \partial_\ell V^j + \Gamma^j_{k\ell} V^k. \qquad (60)$$

In an analogous fashion other covariant derivatives may be computed. Although we have used $dd\vec{X} = 0$ to define in a unique fashion the matrix Γ it is not possible to choose Γ such that $dd\vec{e}_k = 0$ unless the space is flat; in fact

$$dd\vec{e}_k = (\partial_m \Gamma^j{}_{kl}\,\vec{e}_j + \Gamma^j{}_{kl}\,\Gamma^p{}_{jm}\,\vec{e}_p)\,dx^m \wedge dx^l$$
$$\stackrel{\text{def}}{=} R^j{}_{kml}\,\vec{e}_j\,dx^m \otimes dx^l \tag{61}$$

$$R^j{}_{kml} = \partial_m \Gamma^j{}_{lk} - \partial_l \Gamma^j{}_{mk} +$$
$$+\Gamma^j{}_{pm}\Gamma^p{}_{lk} - \Gamma^j{}_{pl}\Gamma^p{}_{mk}. \tag{62}$$

In terms of the language of cartan matrix if we write

$dd\vec{e}_k = \mathcal{R}(\Gamma)\,\vec{e}_k$, then using (45) and (51) we get

$$A\mathcal{R}(\Gamma)A^{-1} = \Omega,$$

With components $\Omega^i{}_{kml}\,\omega^m \wedge \omega^l$. If Γ is a cartan matrix, then $\Gamma = dA\,A^{-1}$; in this case the frame $\vec{u}_a = A^{-1}\vec{e}_j$ is constant over the whole space, so $d\vec{u}_a = 0$; the metric $u_a \cdot u_b$ is also constant and the space is flat.

VII. FOUNDATIONS OF NEWTONIAN MECHANICS AND SPECIAL RELATIVITY

VII.1. Basic Assumptions of Newton on Space-time

Issac Newton made many contributions to a wide ranging branches of science - what was in those days called natural philosophy. In physics proper, among his well known contributions are in the subjects of heat, optics, electricity and magnetism. However he is most known for his formulation of mechanics and his law of gravitational attraction[1]. Before going on to state laws of motion, Newton discussed at length the significance of various terms that he was using and carefully stated his basic assumptions on space and time in the context of the prevalent notions and terminology. According to Einstein[2] "Newton himself was better aware of the weaknesses inherent in his intellectual edifice than the generations which followed him. This fact has always roused my admiration".

We shall therefore start here by stating the basic assumptions made by Newton and in the spirit of modern trends in physics deduce Newton's laws from there. This approach would enable us to see clearly the "weaknesses" that Einstein has spoken of, and would otherwise help us in developing the course. These assumptions of Newton are:

1. Space is absolute[3], has dimension 3 and is Euclidean[4]
2. Time is absolute and flows uniformaly[3]

By absolute in the above is meant that any physical phenomena whatsoever occurring in nature (in space time) have no effect on the properties of space time (and vice versa), which continue to be the same as stated in (1) and (2). By the Euclidean nature of space one means that the space is homogeneous and isotropic; in other words the physical phenomena are unchanged under the following coordinate changes

$$x^i \to x^i + a^i \qquad \text{homogeneity of space} \qquad (1)$$

$$x^i \to A^i{}_j x^j \qquad AA^T = A^T A = I \text{ Isotropy of space} \qquad (2)$$

$$t \to t + b \qquad \text{uniformity of time} \qquad (3)$$

The coordinate changes clearly leave unchanged the Euclidean distance between two points a,b:

$$S_{ab}^2 = |\vec{X}_a - \vec{X}_b|^2. \qquad (4)$$

In addition the following transformation also leaves it unchanged

$$x^i \to x^i + v^i t \qquad \text{Galilean transformations} \qquad (5)$$

where v^i are the parameters of the transformation and can be interpreted as velocities. The concept of velocity arises because of the concept of flow of time. All physical phenomena have to be studied in the background of the concept of flow of time.

<u>Definition</u>. A point (particle) in space is said to have a non-vanishing velocity if its position (coordinates x^i) changes in time. It is clear that if the flow is not uniform then the

concept of velocity will not be very useful.

Collecting the transformations (1-3,5) we find that the most general transformation that leaves (4) unchanged is

$$x'^i = A^i{}_j x^j + b^i + v^i t \tag{6A}$$

$$t' = t + c \tag{6B}$$

It might appear that if we replace t by f(t) in (5), the form (4) is again left unchanged. However then the form of the most general transformation will not be (6) and would lead to severe complications due to (6B).

VII.2. First Law of Motion and "Conserved" Objects

We now ask what is the simplest law of motion for a free particle consistent with this transformation. The answer is clearly the Newton's first law:

$$\frac{d^2x^i}{dt^2} = 0 . \tag{7A}$$

Notice that (7A) is also unchanged under the affine change of t:

$$t \to at + b . \tag{8A}$$

On the other hand, if we replace eqns. (7A) by

$$M \frac{d^2x^i}{dt^2} = 0 \tag{7B}$$

and assume that when t undergoes an affine change (8A), M undergoes the change of scale

$$M \rightarrow a^2 M, \tag{8B}$$

then the left hand side of (7B) remains unchanged under the combined transformations (8,A,B).

Now consider a closed system of several particles. This system as a whole will satisfy the first law; however, what is to be the interpretation of M? If we consider M to be "property of a point particle", then the analogue of (7B) for a closed system of particles will be

$$\sum M_\alpha \frac{d^2 \vec{X}_\alpha}{dt^2} = 0 \tag{9}$$

As t undergoes the transformation $t \rightarrow at$, M_α undergoes the change $M_\alpha \rightarrow a^2 M_\alpha$. It is clear that if we interpret M_α as a number characteristic of a <u>property</u> of the particle, then the numerical value assigned to this property is arbitrary upto <u>a constant</u>. If we therefore fix this constant, it fixes not only the relative values of M_α but also the parameter associated with the affine transformation (8) of t. In particular for a=1, we obtain precisely the group of transformations that leaves (4) unchanged together with the concept of "property of a particle" which is characterised by the numbers M_α which we shall refer to as "<u>Mass</u>" of the αth particle. From the basic assumptions of space-time it is clear that M is independent of $\vec{x}$ and t.

If we integrate (7B) over a time interval $[t_0, t]$ we obtain the conservation law:

$$\vec{p} \overset{\text{def.}}{=\!=} M \frac{d\vec{x}}{dt}\Big|_t = M \frac{d\vec{x}}{dt}\Big|_{t_0} \overset{\text{def.}}{=\!=} \vec{p}_0 . \tag{10}$$

The vector $\vec{p}$ is called momentum ("Quantity of Motion"-Newton) of the particle. Taking the scalar product of (7B) with $\vec{v}$ and integrating over the interval $[t_0, t]$ gives

$$T \overset{\text{def}}{=\!=} \frac{1}{2} M \vec{v}^{\,2}\Big|_t = \frac{1}{2} M \vec{v}^{\,2}\Big|_{t_0} = T_0 \tag{11}$$

We call T the kinetic energy ("Vis Viva"-Newton) of the particle.

Multiply (7B) by x^j and subtract from it the expression obtained by interchanging i and j:-

$$\frac{d}{dt}\left[p^i x^j - p^j x^i\right] \overset{\text{def.}}{=\!=} \frac{d}{dt} L_{ij} = 0 \tag{12}$$

Components of the antisymmetric tensor are the familiar components of the vector product

$$L_k = (\vec{x} \wedge \vec{p})_k = \frac{1}{2} \sum_{i,j} \epsilon_{ijk} L_{ij} . \tag{13}$$

(density!)

where ϵ_{ijk} is the Levi-Civita, completely antisymmetric tensor;

it is +1 for ijk an ever permutation of 1,2,3; and -1 for odd permutation of 1,2,3 and zero in all other cases. P,T and L are said to be constants of the motion. It further follows from (10) that the particle moves with uniform velocity V, so that $\vec{V}$ a further constant of the motion - we state without proof that these constants are connected respectively with the symmetries defined by (1), (3), (2) and (5).

Formula (10) can be immediately generalised for the case of a system of particles, eqn.(9):

$$\sum_\alpha M_\alpha \vec{V}_\alpha \Big|_t \stackrel{\text{def}}{=} \vec{P} = \vec{P}_0 \stackrel{\text{def}}{=} \sum_\alpha M_\alpha \vec{v}_\alpha \Big|_{t_0}. \tag{14}$$

A further integration of (14) yields

$$\sum M_\alpha \vec{X}_\alpha = \vec{P}_0 t - \sum M_\alpha \vec{X}_\alpha \Big|_{t=t_0} \tag{15}$$

We define the "centre of mass" of a closed system of particles as

$$X = \frac{\sum_\alpha M_\alpha \vec{X}_\alpha}{\sum_\alpha M_\alpha} = \frac{\vec{P}}{\sum_\alpha M_\alpha}. \tag{16}$$

It is then clear that the centre of mass of a closed system of particles moves with a constant velocity $\vec{V}_0 = \vec{P}_0 / \sum M_\alpha$ as if it was a free particle. Constancy of the velocity of centre of mass, as pointed out earlier arises from the invariance of equations of motion under Galilean transformations.

VII.3. Concept of Force: the IInd and IIIrd laws

Next let us consider a closed system of two particles. This system satisfies the first law:

$$m_A \frac{d^2\vec{X}_A}{dt^2} + m_B \frac{d^2 X_B}{dt^2} = 0. \tag{17}$$

If we call $\frac{d^2\vec{X}}{dt^2} = \vec{a}$ as the acceleration vector and denote it by $\vec{a}$, then we see from (16) that one can also define ratio of the masses of two particle in a closed system as the **inverse** ratio of the magnitudes of their accelerations (provided of course that their accelerations are nonvanishing): imparted, say by ~~the~~ same force.

$$\frac{m_A}{m_B} = \frac{|\vec{a}_B|}{|\vec{a}_A|}. \tag{18}$$

Further we can rewrite (17) as

$$m_A \frac{d^2\vec{X}_A}{dt^2} \overset{def.}{=} \vec{F}_{AB}, \quad m_B \frac{d^2\vec{X}_B}{dt^2} \overset{def}{=} \vec{F}_{BA}. \tag{19A}$$

$$\vec{F}_{AB}(x,v,t) + \vec{F}_{BA}(x,v,t) = 0; \tag{19B}$$

and we call the vector function $\vec{F}_{AB}$ as the force vector: the force due to B on A; similarly $\vec{F}_{BA}$ is the force due to A and B. Equation (19B) then says that these forces that A and B exert on each other are equal and opposite.

In this manner we have thus arrived at the concept of force (second law) and the third law which says that action and reaction are equal and opposite.

VII.4A. Symmetry and Conservation Laws

Let us multiply the first of equations (19A) with m_B and the second with m_A and subtract:

reduced mass

$$\mu \frac{d^2(\vec{X}_A - \vec{X}_B)}{dt^2} = F_{AB}, \quad \mu \stackrel{\text{def}}{=} \frac{m_A m_B}{m_A + m_B} \tag{20}$$

It follows from this equation and equations (19) that $\vec{F}$ has the form

$$\vec{F}_{AB} = \vec{F}(\vec{X}_A - \vec{X}_B, \dot{\vec{X}}_A - \dot{\vec{X}}_B; t), \tag{20A}$$

where $\dot{\vec{X}} = d\vec{X}/dt$. We have thus reduced the two particle system (19) to a single particle equation

$$\mu \frac{d^2\vec{X}}{dt^2} = f(\vec{X}, \dot{\vec{X}}; t), \tag{21A}$$

together with the conservation law

$$m_A \vec{V}_A + m_B \vec{V}_B = \text{constant} \tag{22}$$

where μ is the reduced mass, $\vec{X} = \vec{X}_A - \vec{X}_B$ the relative position, and $\vec{v} = \dot{\vec{X}}$ is the relative velocity. We leave as an ~~exercises~~, to prove the following statements on conservation laws arising from (21A) as integrals of the motion. From (21A) we obtain the following formal integrals whose existence is determined by certain symmetry properties

$$p^{j}(x) = \mu \vec{V}^{j} - \int f^{j}\, dt \Big|_{t} \tag{23}$$

$$\begin{aligned} H(t) &= \tfrac{1}{2}\mu \vec{V}^{2} - \int f^{k}\, dx_{k} \\ &= \tfrac{1}{2}\mu \vec{V}^{2} - \int f^{k} v_{k}\, dt. \end{aligned} \tag{24}$$

where the line integral is over an arbitrary path of the particle.

$$J^{k}(\alpha,\beta,\gamma) = L^{k} - \int \Omega^{k}\, dt \tag{25}$$

where α, β, γ are the Euler angles, $L = \vec{X} \wedge \vec{P} = \mu \vec{X} \wedge \vec{v}$ is "orbital" angular momentum and Ω^{k} are the components of the vector $\vec{X} \wedge \vec{F}$.

The 'integrals' p^{j} exist and are constants of the motion if $p^{j}(x) = p^{j}(x + \delta x)$, i.e. if p^{j} is independent of $\vec{X}$; and in this case p^{j} are called generalized momenta. The condition for this is

$$\frac{\partial f^{i}}{\partial x^{j}} = 0\,; \tag{23A}$$

and the general form is $[\vec{A} = \vec{A}(\vec{v}, t)]$

$$\vec{f} = -\frac{\partial \vec{A}}{\partial t}, \quad p^{j} = \mu v^{j} + A^{j}(\vec{v}, t). \tag{23B}$$

When H(t) exists it is called the energy integral; the condition is that $H(t) = H(t+\delta t)$ which yields

$$\vec{f} = \vec{v} \wedge \vec{B} - \nabla \phi(\vec{x}) \tag{24B}$$

where B is a function of x and t. Similarly for conservation of angular momentum one has to consider an infinitesimal change in α, β, γ. For $\vec{F} \wedge \vec{X} = 0$, orbital angular momentum $\vec{L} = \vec{x} \wedge \vec{p}$ is conserved. The generalized momentum defined by

$$J = \vec{L} + \vec{\omega}, \quad \alpha = \frac{d\omega(t)}{dt} \tag{25A}$$

is conserved if

$$\vec{\alpha} \cdot \vec{x} = 0, \quad F = \frac{\vec{x} \wedge \vec{\alpha}}{|\vec{x}|^2}. \tag{25B}$$

The vectors $\vec{x}, \vec{\alpha}, \vec{F}$ form an orthogonal triad and may be taken along the spherical polar unit vectors $\hat{r}, \hat{\theta}, \hat{\varphi}$. Even if $\vec{L}$ is not conserved, it could happen that L^2 is constant of the motion; this is the case if

$$\vec{F} = \vec{A} + \vec{B} = \vec{A} + \vec{v} \wedge \vec{C} \tag{25C}$$

where A and B are parallel and perpendicular to $\vec{X}$; $\vec{C}$ is normal to $\vec{v}$ and is in the $\vec{x}$-$\vec{v}$ plane; so that $\vec{v}$, $\vec{c}$ and $\vec{L}$ form an orthogonal triad.

To summarise, one can deduce the conserved quantities by studying the behaviour of $\vec{F}$ under space-time changes as under

infinitesimal change	conserved quantity
$\vec{x} \rightarrow \vec{x} + \delta\vec{x}$	linear momentum
$t \rightarrow t + \delta t$	energy
$\vec{x} \rightarrow A\vec{x}$, $AA^T = I$	angular momentum
$x \rightarrow x + t\,\delta\vec{v}$	centre of mass

In addition, <u>mass</u> is always strictly conserved.

VII.4B. Using the results of section V.8 and V.9 in particular equations (V.68, 68A) we can rewrite the Newtons law of motion, eqn. (21A) in general coordinates as

$$\mu\left[\frac{d^2x^j}{dt^2} + \left\{{j \atop k\ell}\right\}\frac{dx^k}{dt}\frac{dx^\ell}{dt}\right. = f^j(\vec{x}, \dot{\vec{x}}\,; t) \tag{21B}$$

$$\left\{{j \atop k\ell}\right\} = \frac{\partial x^j}{\partial y^m}\frac{\partial^2 y^m}{\partial x^k \partial x^\ell} = g^{jm}\frac{1}{2}\left(g_{mk,\ell} + g_{m\ell,k} - g_{k\ell,m}\right)$$

where y^j are the cartesian and x^k the general coordinates; comma in $g_{mk,\ell}$ denotes differentiation with respect to x^ℓ, and

$$ds^2 = d\vec{y}\cdot d\vec{y} = g_{k\ell}\,dx^k dx^\ell\,; \quad g_{k\ell} = \frac{\partial\vec{y}}{\partial x^k}\cdot\frac{\partial\vec{y}}{\partial x^\ell}\,. \tag{4A}$$

If we multiply (21B) throughout by g_{jm} and sum over j we find on using (V.7,7A)

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}^k} - \frac{\partial T}{\partial x^k} = g_{jk} f^j = f_k \tag{21C}$$

where T is the expression for kinetic energy in generalized coordinates:

$$T \stackrel{\text{def}}{=} \frac{1}{2}\mu g_{\ell k}\dot{x}^\ell \dot{x}^k, \quad \dot{x}^\ell \stackrel{\text{def.}}{=} \frac{dx^\ell}{dt} \stackrel{\text{def}}{=} v^\ell \tag{11A}$$

Now let us consider the case when force is of the type [combining (23B) and (24B)]

$$\vec{f} = \vec{V} \wedge \vec{B} - \nabla\phi(x) - \frac{\partial \vec{A}}{\partial t}, \quad B = \nabla \wedge \vec{A} \tag{26}$$

In component form one may write

$$f_k = \epsilon_{k\ell m} v^\ell B^m - \partial_k \phi - \frac{\partial A_k}{\partial t}$$

where $B^k = \epsilon^{k\ell m}\partial_\ell A^m \stackrel{\text{def.}}{=} \frac{1}{2}\epsilon^{k\ell m} B_{\ell m}$. Using the definition $\epsilon^{mij}\epsilon_{mk\ell} = \delta^{ij}_{k\ell} = \delta^i_k \delta^j_\ell - \delta^i_\ell \delta^j_k$, we get

$$f^k = -\partial_k \phi + v^\ell B_{\ell k} - \frac{\partial A_k}{\partial t} \tag{26A}$$
$$= \left[\frac{d}{dt}\frac{\partial}{\partial v^k} - \frac{\partial}{\partial x^k}\right](\phi - \vec{v}\cdot\vec{A}).$$

Eqn. (21C) then takes the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^k} - \frac{\partial L}{\partial x^k} = 0 \tag{27}$$

where $L = \frac{1}{2} g_{jk} v^j v^k - \phi + \vec{v}\cdot\vec{A}$ is called Lagrangian of the system and (27) are called the Euler-Lagrange equations. They may also be obtained by the Hamilton's variational principle :

$$\begin{aligned} 0 &= \delta \int_{t_1}^{t_2} L(\vec{x}, \dot{\vec{x}}; t) \\ &= \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial \dot{x}^j} \delta \dot{x}^j + \frac{\partial L}{\partial x^j} \delta x^j \right) \qquad (28) \\ &= \left[\frac{\partial L}{\partial \dot{x}^j} \delta x^j \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \delta x^j \left(\frac{\partial L}{\partial x^j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^j} \right) dt ; \end{aligned}$$

where it is assumed that the variation vanishes at the end points $[\delta x_j(t_2) = \delta x_j(t_1) = 0]$ and that the variations are otherwise arbitrary and independent of each other.

With force of the type (26) eqns. (21B) may also be written as

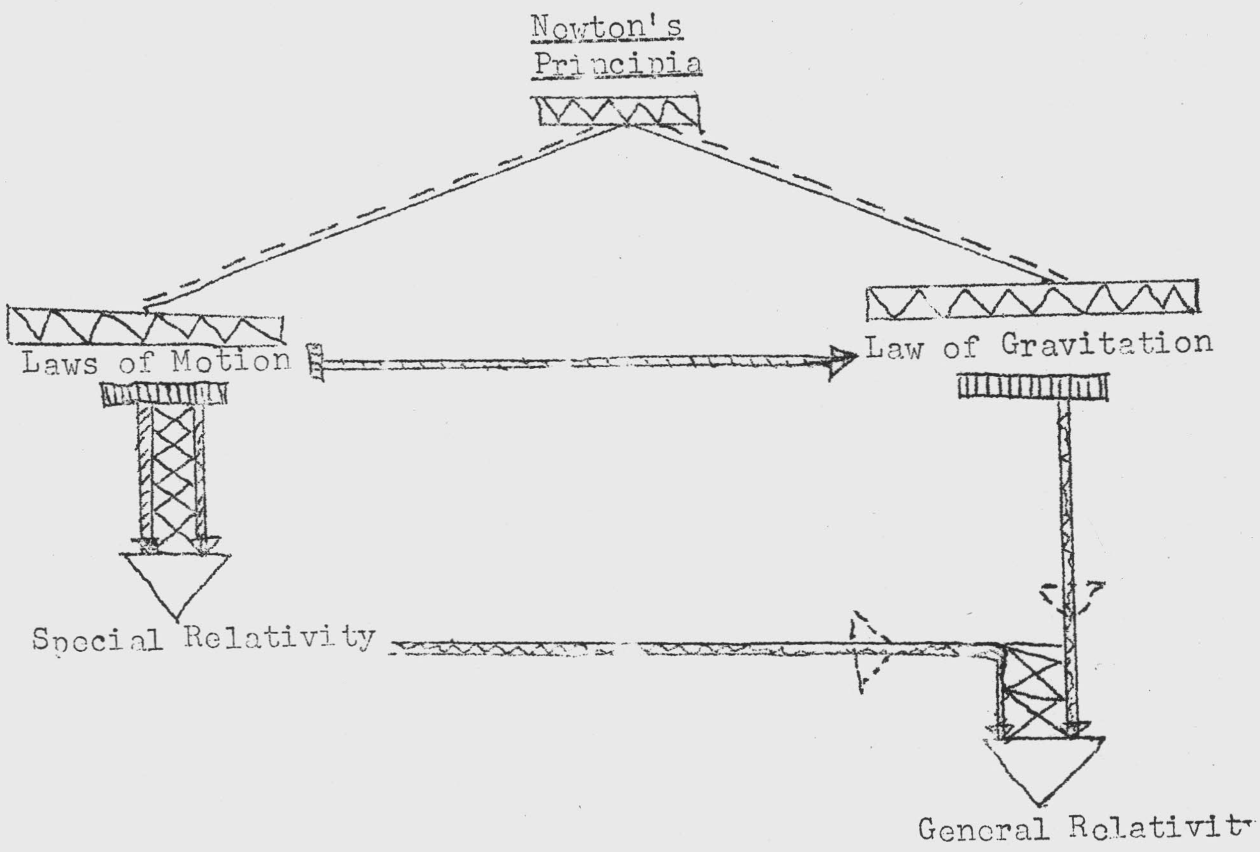
$$\mu\left[\frac{d^2x^j}{dt^2} + \Gamma^j_{kl} v^k v^l\right] = \left(\frac{-\partial\phi}{\partial x^k} - \frac{\partial A_k}{\partial t}\right) g^{kj} \tag{27A}$$

$$\Gamma^j_{kl} = \left\{ {j \atop kl} \right\} + T^j_{kl} = \left\{ {j \atop kl} \right\} + \frac{B^j_k v_l + B^j_l v_k}{2v^2}$$
$$(B_{lk} = \partial_l A_k - \partial_k A_l) \tag{29}$$

where $B^j_k = g^{jl} B_{lk}$ and T^j_{kl} is symmetric in k, l; also $T_{(jkl)} = 0$. Thus we see that Γ^j_{kl} here is precisely of the form considered in equations (VI.31,35). Hence if right hand side of (27A) vanishes, i.e. force is of the type $\vec{f} = \vec{V} \wedge \vec{B}$, $\vec{B} = \nabla \wedge \vec{A}(x)$, then the solution of the geodesic equation is same as in the absence of force and similarly kinetic energy is also the same as if there is no force. Such forces are said to do no work. Examples of these forces are: Gyroscopic forces and forces due to a magnetic field.

VII.5. Weaknesses in the Newtonian Theory and Special Relativity as its Natural Completion

Newton's two contributions on laws of motion and law of gravitation have been so to say given logical completion in the theories of special and general relativity of Einstein as pictorialy represented in the following diagram



Double arrow in the figure above, denotes logical completion. Single arrow denotes just further development consistent with the previous hypothesis.

Let us clarify what we mean by logical completion. We confine ourselves here to special relativity. The important conservation laws of Newtonian mechanics are

1. Conservation of mass
2. Conservation of energy
3. Conservation of linear momentum
4. Conservation of angular momentum
5. Conservation of centre of mass

In all problems in Newtonian mechanics mass is assumed to be strictly conserved. Let us consider the scattering process between two equal mass particles. First, let the scattering be elastic; then to treat the problem one can either use the principle of conservation of linear momentum or of the conservation of energy. However if the scattering is inelastic one gets into some problems as represented in the following table.

Quantity	Initial state	Final state	Conservation or not
Mass	$m+m$	$M=2m$	yes
Momentum	$m\vec{v}+0$	$M\vec{V}$	yes
Kinetic energy	$\frac{1}{2}m\vec{v}^2+0$	$\frac{1}{2}M\vec{V}^2$	no

The table shows that if we assume the strict conservation of mass and momentum than the kinetic energy is not conserved.

The loss in kinetic energy is explained by saying that it has gone into the production of heat and sound produced during the collision and into the elastic energy of the combined mass. However for a point particle such an explanation is far too complicated and in any case involves extra-Newtonian assumptions.

At this point we make parenthetic remarks of historical nature. At the time of Newton there was considerable controversy regarding the nature of force in the sense whether it is associated with the <u>vis viva</u> $m\vec{v}^2$ or with the <u>quantity of motion</u> $m\vec{v}$. In a competition organised by the Royal Society, London in 1668, to discuss the then burning problem of the relative importance of $m\vec{v}$ and $m\vec{v}^2$ in connection with the concept of force, Cristian Huygens pointed out that $m\vec{v}$ was preferable as having to do with force, since one could use it to treat both elastic and inelastic processes, whereas $m\vec{v}^2$ was a useful concept only for elastic scattering. We should perhaps add that had the loss of kinetic energy in inelastic scattering been recognised as part of the potential energy of the combined system one would have seen the early rise of Hamiltonian system of equations which has proved highly useful in the study of quantum phenomena.

Notwithstanding what we have said in the preceding paragraph, the criticism against the law of conservation of kinetic energy in inelastic scattering voiced in an earlier paragraph stands. This difficulty is removed in the theory of special relativity where the principles of conservation of mass and energy are replaced by a single law - one defines

$$\text{Linear Momentum} = \vec{P} = m\gamma\vec{v}$$
$$\text{Mass-Energy} = E = m\gamma c^2 \qquad (30)$$

where C is the velocity of light and $\gamma = (1-\frac{v^2}{c^2})^{-\frac{1}{2}}$,
so that

$$E^2 - \vec{P}^2 c^2 = m^2 c^4 . \qquad (29A)$$

Whether the scattering is elastic or inelastic we get the <u>same</u> set of basic equations: (Same problem as considered in the above table)

$$\left.\begin{aligned} m\vec{v}\gamma &= M\vec{V}\Gamma \\ m\gamma + m &= M\Gamma \end{aligned}\right\} \quad \Gamma = (1-V^2/c^2)^{-\frac{1}{2}} \qquad (30)$$

For elastic scattering M=2m. Thus we see that in special relativity no extra assumptions are required to deal with inelastic scattering. In this sense one may consider <u>special relativity as a logical completion of Newtonian mechanics</u>.

Special relativity is a <u>natural completion of Newtonian ideas also in another sense</u>. We saw that the equations of motion of a closed system of particles (eqn.9) are left unchanged under the Galilean transformations (5). The physical significance of this transformation is that <u>the laws of physics are the same in all frames of reference that move relative to each other with uniform velocity (principle of Galilean Relativity)</u>. The frames of reference that move relative to each other with uniform velocity are called Inertial Frames. From the study of optical phenomena in moving bodies and from the Maxwell's equations which describe this phenomena Einstein[5] found using the two postulates of Galilean relativity and constancy of the velocity of light that the transformations relating Inertial frames of reference are not given by the Galilean transformations (5). It follows from the principle of Galilean Relativity that in particular the velocity of light has the same constant value in all inertial frames. This fact was verified in the classic experiments of Michelson and Morley and others[6]. A careful analysis shows that the principle of Galilean Relativity together with the principle of constancy of the velocity of light leads to the transformation group which leaves Maxwell's equations unchanged. Since Lorentz[7] first discovered the transformation group that leaves Maxwell's equations unchanged these transformations are called Lorentz transformations. An analysis of the concept of simultaneity of two events in different inertial frames assuming (1) the principle of Galilean relativity (2) the principle of the constancy of the velocity of light shows that one cannot reconcile these principles with the concept of absolute time and absolute space in Newtonian mechanics[8], instead one obtains:

Space-time is absolute and pseudo-euclidean

In this sense we see again that special relativity is a natural completion of Newtonian Mechanics. Finally we mention one more viewpoint which brings out how special relativity is a natural completion of Newtonian ideas. This viewpoint was much elaborated by Professor Alladi Ramakrishnan in his lectures and I shall therefore only briefly review it.

Starting point of Alladi's argument is that if $\vec{v}_1$ and $\vec{v}_2$ are velocities in Newtonian mechanics then $\vec{v}_1 + \vec{v}_2$ is also a velocity, so that there is no upper limit to the numerical value of velocity that a particle can attain. In Alladi's viewpoint, such a world is considered to be chaotic. Hence it is desirable to put a suitable limit to the velocity that a particle can attain. The claim now is that one can obtain the Einsteinian law of addition of velocities (and also Lorentz transformations) as generalization of the Newtonian law when the assumption regarding upper limit of velocity is made. Precisely stated, the assumptions are

(1) The upper limit to a realizable velocity is 1 in suitable units.

(2) For velocities much smaller than 1 the Newtonian law results

(3) If v_1 and $\bar{v}_2$ are realisable velocities then their composition $\vec{v}_{12}$ is also realisable.

In the Newtonian case the set of all realisable velocities satisfy the following relations (we consider the case of one dimension)

$$u_1+(u_2+u_3)=(u_1+u_2)+u_3$$ distribution

$$u_1+u_2 = u_2+u_1$$ commutative

$$u+(-u)=0$$ has unique inverse (i)

$$u+0=u$$ identity element

$$\lim_{v\to\infty}(u+v)=\infty$$ on the extended real line

At this point we observe that all the transformations we have considered in Newtonian theory, so far, are continuous transformations that form groups. If we confine ourselves to such continuous transformation groups then on quite general grounds one can show that if we perform two successive transformations involving a single parameter (e.g. two successive translations along x-axis, or two successive rotations in a plane); $X \to X' = f_1(X, a_1)$ $X' \to X'' = f_2(X, a_2)$, then $X \to X'' = f_{12}(X, a_1+a_2)$. In technical language we say that a one parameter group is an abelian group - this statement is realy a tautology.

From these remarks about one parameter groups it follows that any "relativistic" law of composition of velocities will also have these properties, but with one difference: whereas the Newtonian velocities span the entire real line, the Einsteinian velocities span a bounded interval between -1 and +1. Let us denote by * the sign of composition of relativistic velocities.

Then from what we have said above, we must have

$$
\begin{aligned}
&v_1 * (v_2 * v_3) = (v_1 * v_2) * v_3 \\
&v_1 * v_2 = v_2 * v_1 = v_{12} \leq 1 \quad \text{if } v_1, v_2 \leq 1 \\
&v * (-v) = 0 \\
&v * (0) = 0 * (v) = v \\
&v * (\pm 1) = (\pm 1) * v = \pm 1
\end{aligned}
\tag{ii}
$$

If we compare (i) and (ii) we see that "0" of relativistic formula behaves like "0" of the Newtonian case, but ± 1 of the relativistic case behaves as $\pm \infty$ of the Newtonian case. Since the Newtonian case corresponds to the real line it means that there exists a one-to-one continuous map from the open interval $-1 < v < +1$ to the real line $-\infty < \theta < +\infty$ and vice versa. One can of course also take the closed interval $-1 \leq v \leq +1$ for the relativistic velocities, in which case the mapping is to the extended real line (which includes infinity as a "number"). Any such map is necessarily unique. It is given by

$$
v = \tanh\theta, \quad \theta = \tanh^{-1} v = \frac{1}{2}\ln\frac{1-v}{1+v}. \tag{A}
$$

Now observe that on the real line, θ would satisfy all the composition rules of Newtonian velocities: (i). In particular

$$
\theta_{12} = \theta_1 + \theta_2 = \theta_2 + \theta_1 .
$$

Therefore if we define $\tanh\theta_1 = v_1$, $\tanh\theta_2 \, v_2$, $\tanh\theta_{12} = v_{12}$, then

$$v_{12} = v_1 * v_2 = \frac{v_1 + v_2}{1 + v_1 v_2}. \qquad \text{(B)}$$

From uniqueness of the mapping (A) follows the uniqueness of the formula (B).

Now let v_1 and v_2 be the velocities of an object as observed by two observers moving relative to each other with velocity u; then

$$v_2 = \frac{v_1 + u}{1 + v_1 u}.$$

Let us introduce the homogeneous coordinates $v_2 = \frac{y}{y_0}$, $v_1 = \frac{x}{x_0}$, $u = a / a_0 = \tanh\theta$, then

$$y = K(a_0 x + a x_0), \; y_0 = K(a_0 x_0 + a x)$$

For K=1, we see that since $a / a_0 = \tanh\theta$ we will have

$$y^2 - y_0^2 = x^2 - x_0^2. \qquad \text{(C)}$$

From this it would follow that x_0, y_0 are to be interpreted as time so that $v_1 = x / x_0$ and $v_2 = y / y_0$ are velocities. Hence we obtain

$$x' = y = \frac{x + v x_0}{\sqrt{1 - v^2}}, \quad t' = y_0 = \frac{t + v x}{\sqrt{1 - v^2}}, \quad (x_0 = t)$$

which are Lorentz transformations. Similarly one can also consider the 3-dimensional case, which however we shall not do here.

VII.6. Lorentz Transformations

From the principle of the constancy of the velocity of light in all inertial frames it follows that,

$$c = \left|\frac{d\vec{x}}{dt}\right| = \text{ has the same constant value in all inertial frames;}$$

hence for a light signal we must have

$$dt^2 - \frac{d\vec{x}^2}{c^2} = 0. \tag{31}$$

Since other particles will travel slower than light, for these $dt^2 - d\vec{x}^2/c^2 > 0$ and we write

$$d\tau^2 = c^2 dt^2 - dx^2. \tag{32}$$

We add that in principle, it is possible that there are particles which also travel faster than light; for these $c^2dt^2 - d\vec{x}^2 < 0$; however we reject this possibility for several reasons. Firstly such particles have no Newtonian analogue: i.e. in no limit is their mechanics governed by Newtonian mechanics; secondly the concept of flow of time, so essential to enable one to define this concept of "velocity", is no longer meaningful - hence it is meaningless to talk about velocities greater than that of light. In this connection we note that the concept of flow of time has also another important use in physical phenomena in that one can order with its help any sequence of physical phenomena: what is earlier is the "cause" and what follows is the "effect". This is called the principle of causality. Obviously this also breaks down as was already pointed out by Einstein[9,10] long ago. Still another reason for rejecting the existence of these particles is that they can have positive as well as negative energies: the existence of negative energies will create an anomalous situation. Further more such particles can not be charged as contribution of the electromagnetic field to its self mass square (see eqn. 25A) is always $\geqslant 0$; this would make detection of such particles almost impossible.

On the other hand we note that one may interpret the inequality $d\vec{X}^2 - c^2 dt^2 > 0$ as a rigid rod whose maximum possible measured length square is $dX^2 - c^2 dt^2$ and this happens in a frame of reference which is at rest relative to the rod. In this rest frame of the rod $dt = 0$, whereas in all other inertial frames $dt \neq 0$. One may interpret this to mean that if two events (e.g. flashes of light) at the ends of a rod in its rest frame are observed to be simultaneous $(dt = 0)$ then in any other inertial frame they are not simultaneous $dt \neq 0$.. In this view-point $\frac{d\vec{X}}{dt}$ does not have the interpretation of a velocity, but rather $\frac{d\vec{X}}{dt} \to \frac{c^2}{v}$, where v is the velocity of the inertial frame relative to the rest frame so that $ds = |d\vec{X}| \sqrt{1 - v^2/c^2}$. Hence a rod of rest-length $|d\vec{X}|$ appears to be contracted when observed from a frame moving relative to it with velocity v - the contraction being only in the dimension of the direction of motion.

Returning to equations (31),(32), we see that among the transformations which leave these unchanged are

$$\vec{x} \to \vec{x} + \vec{a} \quad \text{space translation}, \tag{1}$$

$$t \to t + b \quad \text{Time Translation}, \tag{3}$$

$$\vec{x} \to A\vec{x}, \quad AA^T = 1, \quad \text{Spatial Rotations}, \tag{2}$$

The most general transformations that leave (31) unchanged are clearly those which would lead to multiplying the left side of (31) by at most a factor k. If we require that (1) the set of transformations leaving (31) unchanged form a group (2) space time translations and spatial rotations form a subgroup of this group, then it turns out that the factor k is necessarily unity. We show this in two steps.

Consider a transformation in x_1-t plane, that leaves the left hand side of (31) unchanged; it has the obvious form

$$x_1' = x_1 \cosh\theta + ct \sinh\theta, \quad x_2' = x_2, \quad x_3' = x_3$$
$$ct' = ct \cosh\theta + x_1 \sinh\theta. \tag{33}$$

where $(\vec{x}, t)$ are coordinates of an arbitrary event in frame Σ and $(\vec{x}', t')$ are coordinates of the same event as seen in Σ'. In particular we consider the event to refer to the origin of the system in Σ'; in that case $x_1' = 0$ and so

$$-\frac{v_1}{c} \stackrel{\text{def.}}{=\!=} -\frac{x_1}{ct}\bigg|_{x_1'=0} = \tanh\theta \tag{34}$$

eliminating θ gives

$$x_1' = \gamma(x_1 - v_1 t), \quad x_2' = x_2, \quad x_3' = x_3$$

$$ct' = \gamma\left(ct - \frac{v_1 x_1}{c}\right), \quad \gamma \stackrel{\text{def.}}{=} \left(1 - \frac{v_1^2}{c^2}\right)^{-\frac{1}{2}}. \tag{35}$$

These are called pure Lorentz transformations along x_1 .

Alternately, if we divide $\vec{x}'$ by ct in (33) we obtain

$$\frac{u_1'}{c} = \frac{x_1'}{ct'} = \frac{\frac{u}{c} + \tanh\theta}{1 + \frac{u_1}{c}\tanh\theta} \tag{36A}$$

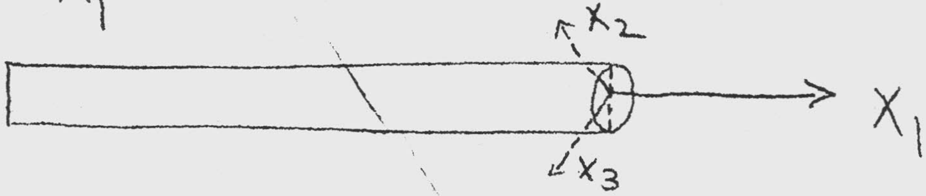
$$\frac{u_2'}{c} = \frac{x_2'}{ct'} = \frac{u_2/c}{\left(1 + \frac{u_1}{c}\tanh\theta\right)\cosh\theta}, \text{ etc.} \tag{36B}$$

where we have put $u_1' = x_1'/t'$, is velocity of the event as seen in Σ' and $u_1 = \frac{x_1}{t}$, its velocity, as seen in Σ . We see that $c\tanh\theta$ has to be identified with the relative velocity of two frames, since in the Newtonian limit $c \to \infty$, and we get $u_1' = u_1 + c\tanh\theta$. In both these derivations of eqn. (34) from which the expressions (35) for pure Lorentz transformations follow immediately, we have made use of the condition that in the limit of small u_1 (compared to C) we get the Newtonian expressions; stated differently, we have assumed that there exist inertial frames that are relatively at rest.

Now suppose we do not assume that the left hand side of (37) is left unchanged, but rather it takes on a factor k; then in place of (35) we get

$$\left.\begin{aligned} x_1' &= \gamma\left(x_1 - \frac{v_1}{c}t\right)k\,, \quad x_3' = k x_3\,, \quad x_2' = k x_2 \\ ct' &= \gamma\left(ct - \frac{v_1 x_1}{c}\right)k\,, \quad k \overset{\text{def}}{=\!=} k(\vec{v}). \end{aligned}\right\} \quad (35)'$$

Since all directions in space are equivalent (isotropy of space— invariance under spatial rotations), it follows that k depends only on the magnitude of $\vec{v}$. Consider a cylinder moving along the X_1-axis such that the axis



of the cylinder coincides with the X_1-axis in the frame Σ . Envisage now a rotation of 180° in $X_1 - X_2$ plane in Σ and in $X_1' - X_2'$ plane in Σ' :

$$\begin{aligned} &X_1 \to -X_1\,, \quad X_2 \to -X_2 \\ &X_1' \to -X_1'\,, \quad X_2' \to -X_2' \\ &v_1 \to -v_1\,, \quad \gamma \to \gamma\,, \quad k(v) \to k(-v). \end{aligned}$$

We get then

$$X_1' = \gamma\left(x_1 - \frac{v_1}{c}t\right)k(-v) \quad ct' = \gamma\left(ct - \frac{v x_1}{c}\right)k(-v)$$

$$X_2' = X_2\, k(-v)\,, \quad X_3' = X_3\, k(-v).$$

Since rotations in $X_1 - X_2$ plane should have no effect on X_3 , we must have

$$k(v) = k(-v) = k(|v|).$$

Now consider a frame Σ'' moving relatively to Σ' along the direction $-X_1$ with the same velocity v. Applying (35) twice gives

$$X''_\mu = k^2 X_\mu \;, \quad \{X_\mu = X_1, X_2, X_3, ct\}$$

Since we are interested only in those transformations that are connected continuously to the identity we must have $k = +1$, as we wanted to show.

Lorentz transformations in an arbitrary direction making an angle α_1 with the X_1 -axis may be obtained from (31) as follows. Resolve the vector $\vec{X}$ along $(\vec{X}_{\parallel})$ and perpendicular $(\vec{X}_{\perp})$ to velocity vector $\vec{v}$; (31) then takes the form

$$\left.\begin{aligned} \vec{X}'_{\parallel} &= \gamma\left(\vec{X}_{\parallel} - \frac{\vec{v}}{c}t\right), \quad \vec{X}'_{\perp} = \vec{X}_{\perp} \\ t' &= \gamma\left(t - \frac{\vec{v}\cdot\vec{X}}{c^2}\right), \quad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \end{aligned}\right\} \quad (34A)$$

Since, $\vec{v}\cdot\vec{X} = \vec{v}_1\cdot\vec{X}_{\parallel} = v_1 X_1$,

$\vec{X}_{\parallel} = \dfrac{\vec{X}\cdot\vec{v}}{v^2}\,\vec{v}$, $\vec{X}_{\perp} = \vec{X} - \vec{X}_{\parallel}$. Substituting these in the above we get

$$\left.\begin{aligned} \vec{X}' &= \vec{X} + (\gamma - 1)\frac{\vec{X}\cdot\vec{v}}{v^2}\vec{v} - \gamma t \vec{v} \\ &= \gamma\left[\tilde{\vec{X}} - \vec{v}t\right] \\ t' &= \gamma\left(t - \vec{v}\cdot\vec{X}/c^2\right), \end{aligned}\right\} \quad (35)$$

where we have put

$$\tilde{\vec{X}} = \frac{1}{\gamma}\vec{X} - \left(\frac{1}{\gamma} - 1\right)\frac{\vec{v}\cdot\vec{X}}{v^2}\vec{v} = \frac{\vec{X}_{\perp}}{\gamma} + \vec{X}_{\parallel}. \quad (35A)$$

If we denote by R the matrix for a general 3-dimensional rotation, then the most general Lorentz transformation that is connected continuously to the identity may be written as

$$\vec{x}' = \gamma R(\vec{x} - \vec{v}t), \quad t' = \gamma\left(t - \frac{\vec{v}\cdot\vec{x}}{c^2}\right) \tag{36}$$

From remarks following eqn.(31) it is clear that velocity of light has been assumed to be an upper limit; on the other hand in Newtonian mechanics there is no such upper limit. Hence to obtain an interpretation of Lorentz transformations we consider the limit $c \to \infty$. We obtain from (35)

$$x_1' = x_1 - v_1 t, \quad x_2' = x_2, \quad x_3' = x_3, \quad t' = t,$$

which are just the Galilean transformations. Thus pure Lorentz transformations are transformation between different inertial frames. From (32) we see that

$$d\tau = dt\sqrt{1 - v^2/c^2}. \tag{32A}$$

For $v = 0$, $dt = d\tau$; hence $d\tau$ is to be interpreted as the time interval in a clock at rest with respect to the observer. Equation (32A) then shows that if $d\tau$ is a time interval in Σ_0 and the frame Σ moves relative to Σ_0 with velocity $\vec{v}$, then the time interval observed in Σ is longer in the time scale of Σ,

$$dt = \frac{d\tau}{\sqrt{1 - v^2/c^2}}. \tag{32 }$$

Therefore as observed in Σ, the clock in Σ_0 which is moving relative to it with velocity $\vec{v}$ will lag behind the clock at rest in Σ.

As a cautionary remark we note that in obtaining the law of addition of velocities we have put $v = \frac{x}{t}$; there is nothing wrong with this mathematically; but if we want to interpret v as velocity then one must use $v = \frac{\Delta x}{\Delta t}$ meaning that a point has changed it's position from $\vec{x}_1$ to $\vec{x}_2$ $(\Delta\vec{x} = \vec{x}_2 - \vec{x}_1)$ during the passage of time $\Delta t = t_2 - t_1$.

It is clear from (32) that the proper time interval is unchanged under Lorentz transformations; hence if we consider the space-time difference 4-vector dx^μ with components $(d\vec{x}; c\,dt)$, we see that $\frac{dx^\mu}{d\tau}$ is also a four vector— in fact $dx^\mu/d\tau$ are the components of the tangent vector to a curve $\tau(x)$ and τ is the geodetic parameter. In terms of x, t we have

$$v^\mu = \frac{dx^\mu}{d\tau} = \gamma\,\frac{dx^\mu}{dt} = \{\gamma\vec{v}; c\gamma\}, \quad \vec{v} = \frac{d\vec{x}}{dt}; \tag{37}$$

for $v \ll c$, $v^\mu = \{\vec{v}; 0\}$, so we interpret v^μ as components of the velocity 4-vector. If $\vec{e}_\mu$ denotes the frame base, the vector itself is $v^\mu \vec{e}_\mu$. Similarly $\underset{\sim}{dx} = dx^\mu e_\mu$ and therefore (32) may be rewritten as

$$d\tau^2 = \underset{\sim}{dx}\cdot\underset{\sim}{dx} = \vec{e}_\mu\cdot\vec{e}_\nu\, dx^\mu dx^\nu = \eta_{\mu\nu}\,dx^\mu dx^\nu,$$

$$\begin{aligned} \eta_{\mu\nu} &= +1 \quad \text{for } \mu = \nu = 0,\ x_0 = ct \\ &= -1 \quad \text{for } \mu = \nu = 1, 2, 3 \\ &= 0 \quad \text{in all other cases.} \end{aligned} \tag{32A}$$

Corresponding to the contravariant components v^{μ}, we have the covariant components $v_{\mu} = \{-\gamma \vec{v} ; \gamma\}$ so that

$$v^{\mu} v_{\mu} = -\gamma^2 \vec{v}^{\,2} + \gamma^2 = +1 \tag{37A}$$

and v^{μ} is called a time like four-vector; in general for an arbitrary vector A^{μ}

$$A^{\mu} = \{A^1, A^2, A^3 ; A^0\} = \{\vec{A} ; A_0\}$$

$$A_{\mu} = \{-\vec{A} ; A_0 = A^0\}$$

The vector is said to be time-like, space like or null according as

$$A^{\mu} A_{\mu} > 0, \quad A^{\mu} A_{\mu} < 0, \quad A^{\mu} A_{\mu} = 0$$

The frame base for the covariant components may be denoted by e^{μ} and then $d\tau^2 = \eta^{\mu\nu} dX_{\mu} dX_{\nu}$; e^{μ} is the base dual to $\vec{e}_{\mu}$ such that

$$e^{\mu} \cdot \vec{e}_{\nu} = \delta^{\mu}{}_{\nu} = \eta^{\mu\nu} \eta_{\nu\alpha}$$

More precisely we define a "Minkowski space" M as a 4-dimensional pseudo-euclidean space over the field of reals and of signature -2. It is clearly 4 dimensional, since any event in M can be described by 3 space coordinates and one time parameter. By Euclidean we mean that there is homogeneity and isotropy property and the scalar product of a vector with itself is always positive (or zero only if the vector itself is zero). For a pseudo-euclidean space the scalar product can be positive, negative or zero. By signature we mean the between the number of positive and the number of negative squares.

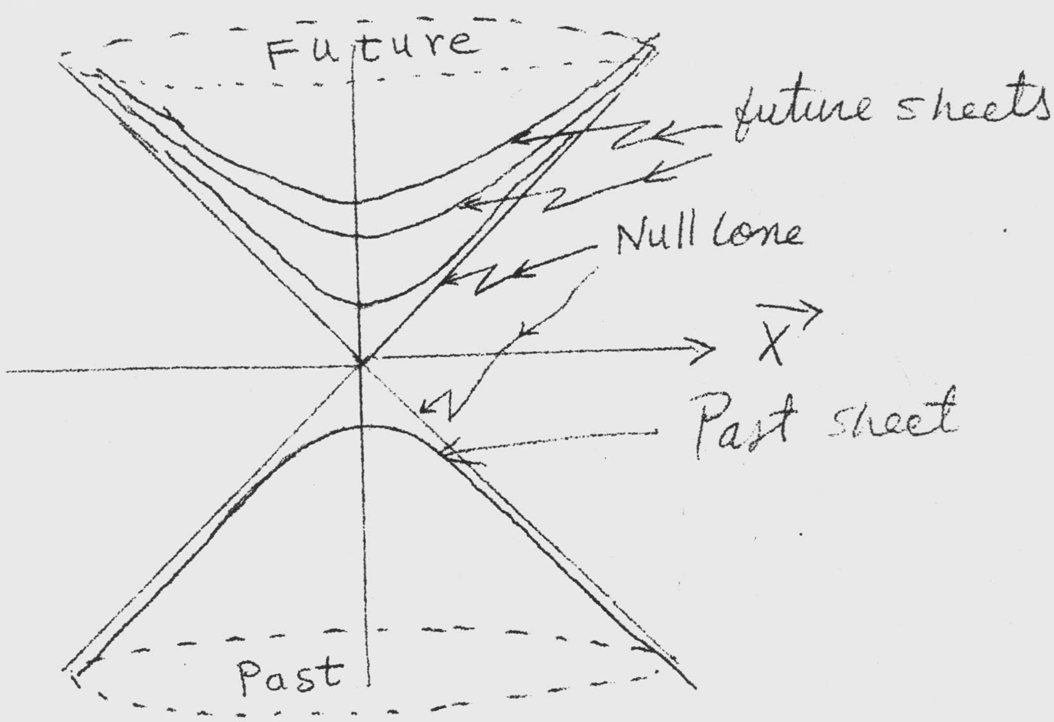
It is clear that M is a linear space and, one can consider its subspaces; these are

(i) A subspace is space-like if <u>all</u> its vectors are space-like

(ii) A subspace is time-like if it contains a time-like vector

(iii) A subspace is null if it contains a null vector N such that if A be any other vector in the subspace, then $N.A = 0$.

It is clear that a null vector can be orthogonal only to itself, or to some space-like vectors; the only vector that can be orthogonal to a time like vector is a space-like vector. The set of all null vectors form a cone called the null cone.



It is a 3 dimensional hypersurface in the four-dimensional Minkowski space. The set of all time-like vectors lie inside the cone. The zero vector, since it is orthogonal to all other vectors in Minkowski space forms a seperate subspace and divides the cone in two parts: upper and the lower cones: the set of all time like vectors lying in the upper cone are called the future sheets and those in the lower cone are referred to as the past sheets. All other vectors lying outside the cone are space-like vectors. Because of the invariant nature of the zero vector, under Lorentz transformations, points on and inside the future (part) cone are transformed respectively into points on and inside the future (part) cones. The set of all transformations that preserve this past future relationship are called orthochronous. If we further restrict the transformations such that inner orientation[11] of a space-like hypersurfaces is also preserved then we get precisely the continuous transformations of the Lorentz group we have considered above. For further details consult section V.4. We only note here that when space-time translations are included as space-time symmetry then the enlarged group is called Inhomogeneous Lorentz group or the Poincare group .

References

1. Isaac Newton, Principia: Mathematical principles of natural philosophy and His system of the World; translated by Andrew Motte, 1729; translation revised and annotated by Florian Cajori (Uni. of California Press, Berkeley (1947)).

2. A.Einstein, *Essays in Science*, Philosophical Library, N.Y. 1934
3. *Principia*, Chapter I
4. The Euclidean nature of space is emphasised by Newton in his introduction in principia, loc. cit.
5. A.Einstein: "On the electrodynamics of moving bodies" (in *Principles of Relativity*, collection of translated articles, Dover Publications, Inc) Translated from Annalen der Physik, 17 (1905) 891
6. A.A.Michelson and E.W.Morley, Am.J.Sci. 34 (1887) 333; Phil. Mag. 24 (1887) 449.

   This experiment shows using light source from the earth that the velocity of light along the direction of the motion of the earth and in the direction perpendicular to it have the same numerical value.

   D.C.Miller, Proc. Nat. Acad. Sci. 11 (1925) 311 showed that the results are the same if source of light is the Sun.

   R.Tomashek, Ann. Physik 73 (1924) 105 showed the same using starlight.

   De Sitter, Proc. Amsterdam Acad. 15 (1913) 1297; 16 (1913) 395 showed from study of the binary stars that velocity of light is independent of the source. A related experiment is that of Trouton and Noble, Phil. Trans. A202 (1903) 165, Proc.Roy.Soc. (Lond.) 72 (1903) 132 which reports negative result of attempts to detect torque on a charged condenser due to the absolute motion of the earth.

There is a possibility that medium is dragged along with the source; this is ruled out by the occurence of aberation when observing fixed stars; and also Fizean's experiments (H. Fizeau Comptes Rendus, 33 (1851) 349) on the velocity of light in a moving medium (water) suggest that empty space (ether) does not exhibit this property. Whereas these experiments can also be explained by Lorentz contraction hypothesis, the experiment of Kennedy and Thorndyke, Phys. Rev. 42 (1932) 400 using unequal arms in the interferrometer (Michelson uses equal arms) rule out that explanation.

7. H.A.Lorentz, Proceedings Academy of Sciences of Amsterdam, 6 (1904) 809. Lorentz had made a slight error which was later corrected by H.Poincare who used the term "Lorentz group" for the first time. H.Poincare' Rend. Circ. Mat. Palermo, 21 (1906) 129; C.R.Acad. Acad. Sci., Paris 140 (1905) 1504 Poincare' also for the first time considered the principle of the constancy of the velocity of light. However Einstein's formulation was entirely different. He started from the principle of Galilean relativity and applied it to the study of optical phenamena in moving bodies: see ref.(5). The results of Michelson and Morley's experiments were actually not used by Einstein in his derivation and apparently played no role in his original considerations, though he explicitly mentions the postulate of constancy of the velocity of light.

8. For a very lucid analysis, see A.Einstein, Meaning of Relativity.

9. A.Einstein: Ann. Phys., Lpz., 23 (1907) 371

10. See also W.Pauli: Theory of Relativity, Pergamon Press (1958) p.16.

11. Recall that if A and B are two vectors in 3-dimensional space then their "vector product" $A \wedge B$ is a vector normal to the plane determined by A, B; its direction may be chosen outwards or inwards according as we choose a right-handed or a left handed system. In either case, whether we count the "angle" between A and B from A to B or from B to A the two dimensional plane determined by A and B is assigned an inner orientation by the arrow of "counting" the positive angle in the plane. The "direction" of the vector $A \wedge B$ which is normal to the plane then gives the outer orientation. In case of 3 dimensional space it is clear that we cannot give an outer orientation. Inner orientation may be defined as the sign of $\vec{e_1} \cdot (\vec{e_2} \wedge \vec{e_3})$ where $\vec{e_j}$ are the base vector of the vector space.

VIII. GENERAL THEORY OF RELATIVITY

VIII.1. Newtonian Gravitation.

We mentioned two important contributions of Newton: to mechanics and gravitation which find their way as natural completion in the theories of special and general relativity.

In Newtonian mechanics the basic assumptions are

(a) space is absolute, 3 dimensional and Euclidean (Homogeneous and isotropic).

(b) Time is absolute and flows uniformly.

These, as we found in the last chapter imply that

$$S_{ab} = |\vec{X}_a - \vec{X}_b| = \text{invariant} , \qquad \text{(VIII.1)}$$

and therefore the allowed continuous symmetry transformation of Newtonian theory is

$$x^k \to a^k{}_\ell (x^\ell + v^\ell t + b^\ell) , \quad aa^T = 1 ; \qquad \text{(VIII.2)}$$

accordingly the equations of motion for a system of n particles is

$$\sum_{j=1}^{n} m_j \frac{d^2 \vec{X}^j}{dt^2} = \sum_{\substack{k=1 \\ j=1}}^{n} \vec{F}_{jk} = 0 ; \quad \vec{F}_{jk} = -\vec{F}_{kj} ; \qquad \text{(VIII.3)}$$

where m_j is mass of the jth particle and $\vec{F}_{jk}$ is the force that the kth particle exerts on the jth particle. We shall refer to m_j as defined here as the 'inertial mass' of the jth particle. For a closed system of two particles $m_{1I}\vec{v}_1 + m_{2I}\vec{v}_2 = 0,$

so that

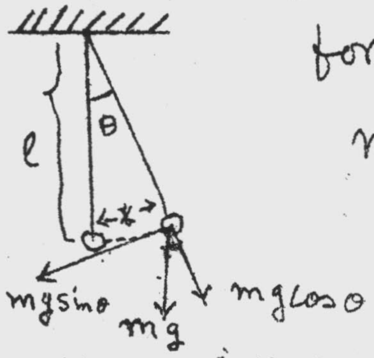
$$\frac{m_{1I}}{m_{2I}} = \frac{|\vec{\dot{v}}_2|}{|\vec{\dot{v}}_1|} : \left\{ \begin{array}{l} \text{Ratio of the masses = inverse ratio} \\ \text{of accelerations ina closed system =} \\ \text{inverse ratio of accelerations imported} \\ \text{by the same 'force' or forces of} \\ \text{the same magnitude} \end{array} \right. \quad \text{(VIII.4)}$$

In the above, the subscript I on m denotes that we are considering inertial masses.

It is of historical interest that both in the formulation of Newtonian mechanics and Newtonian gravitation a beginning was already made in the work of Galileo. Thus Galileo essentially found the principle underlying the first law and pointed out that force had to do with acceleration. In the problems involving gravitation he gave laws of simple pendulum relating period of the pendulum with its length and showed that it was independent of the nature of the material and mass of the pendulum bob. He also pointed out that two objects, irrespective of their nature and inertial masses fall in earth's gravitational field with the <u>same</u> acceleration. It follows that the force acting on each particle is somehow dependent on mass of the particle. If $\vec{g}$ denotes <u>intensity</u> of the gravitational field, the force acting on a particle of intertial m_I is

$$m_I \vec{\ddot{x}} = m_g \vec{g} \, , \quad m_g = k \, m_I \, , \qquad \text{(VIII.5)}$$

where k is a constant of proportionality to be determined and m_g is called the gravitational mass. For a simple pendulum, from the following diagram and corresponding Newtonian equations



for small θ $\sin\theta \simeq \theta \simeq \frac{x}{\ell}$

$$m_I \ddot{x} = \left(m_g \frac{g}{\ell}\right) x$$

$$T = \frac{2\pi}{\sqrt{k}} \sqrt{\frac{\ell}{g}} \; ; \; k^{-1} = \frac{m_I}{m_g} . \qquad \text{(VIII.6)}$$

of motions and their solution, it is clear that relation between inertial and gravitational masses may be determined in terms of k. By performing experiments with different materials of the bob and using different inertial masses of the bob one finds the same numerical value for $\sqrt{kg}$. One can therefore conveniently but $k = 1$ and conclude that

Inertial mass = 'Passive' gravitational mass (VIII.7)

We note that Newton originally performed such pendulum experiments before he identified the 'passive' gravitational and inertial masses. Many experiments with various kinds of pendulum were repeated which established with considerable accuracy that the period of a pendulum was independent of the composition of the pendulum bob. The accuracy of such experiments is however limited by the accuracy with which one can determine the period. In 1889 Baron Roland V.Eotvos performed an experiment using torsion balance which showed with considerable accuracy (a few parts in 10^9) that all bodies fall with the same acceleration. He employed a horizontal torsion beam, 40 cm. long, suspended by a fine wire. Ends of the beam carried two masses of different composition, one slightly lowe-r than the other. The component of earth's gravitational pull acting on each mass was balanced by the centrifugal force field of the earth acting on it. A lack of strict propertionality between the inertial and gravitational masses of the two bodies will lead to a torque tending to rotate the balance. R.V.Eotvos, D. Pekar, E.Fekete, Ann. Phys. 68, (1922), 11]. Recent experiments by the Princeton group (P.G.Roll, R.Kratkov and R.H.Dicke), who employ Sun's gravitational field rather than the field due to the earth's rotation, have improved the figure by $2\frac{1}{2}$ magnitudes. Hence one may consider equation (VIII.7) to be exactly true. This assumption, of the exact equality of gravitational and inertial masses which implies that all objects fall with the same gravitational acceleration is a given gravitational field is called the 'Weak Principle of Equivalence'. We shall come back to the importance of formulating this principle in relation to the general theory of relativity.

Newton's own contribution to Gravitation theory was first to realise the truth of equation (VIII.7) and then to give a formula for gravitational intensity of a massive object. According to Newton, if m and M are masses of two objects, then between them there is a force of magnitude mMG/r^2, and is along the line joining their centres of mass, i.e.

$$\vec{F} = m\frac{MG}{r^3}\vec{r} = m\vec{g}(\vec{r}) \qquad \text{(VIII.8)}$$

where G is called the Newtonian gravitational constant [dimensions $(M/L^3)^{-1} T^{-2}$ = (mass density sec^2)$^{-1}$; numerical value 6.7×10^{-8} $cm^3/gm/sec^2$]. The mass M which gives rise to the gravitational field intensity $\vec{g}$ is sometimes referred to as Active Gravitational mass, in contrast to m which we called passive gravitational mass. With these distinctions equation (VIII.8) may rewritten as

$$m_I \ddot{\vec{X}} = m_g \frac{m_A G}{r^3} \vec{r}, \tag{VIII.8A}$$

where m_I = inertial mass, m_g = passive gravitational mass, m_A = active gravitational mass. Newton did not consider any distinction between active and passive gravitational masses due to the symmetry of equation (VIII.8) and we shall also do likewise. One may write equations (VIII.8A) more generally as

$$m_I \ddot{\vec{X}} = - m_g \operatorname{grad} \phi \tag{VIII.8B}$$

Where the gravitational potential ϕ is given by the Poisson's equation

$$\nabla^2 \phi = 4\pi G \rho \tag{VIII.9A}$$

inside the region Γ where the body has density ρ ; and outside the region Γ where $\rho = 0$, it satisfies the Laplace's equation

$$\nabla^2 \phi = 0 . \tag{VIII.9B}$$

The gravitational field intensity is then given by

The gravitational field intensity is then given by $\vec{g} = -\nabla\phi$, and the formal solution is

$$\phi = -G\int \frac{\rho(X')}{|\vec{X} - \vec{X}'|}\, d^3X'. \qquad \text{(VIII.10)}$$

For a point particle equation (VIII.10) immediately yields the expression $\phi = \frac{-MG}{r}$ consistent with (VIII.8) and (VIII.8A). We note that $M_g\phi$ is the potential energy of M_g in the gravitational field of M and can therefore be expressed as the work done in bringing a particle of mass M_g from $-\infty$ to $r = \int_{-\infty}^{r} \vec{F}\cdot d\vec{r} = M_g\phi$.

The constants of the motion associated with the problem $(m_I = m_g)$ are the total energy

$$E = \frac{1}{2} m v^2 + m\phi, \qquad \text{(VIII.11)}$$

the orbital angular momentum $\vec{L} = \vec{r} \wedge \vec{P}$, and the Runge-Lenz vector

$$\vec{\Pi} = \vec{L} \wedge \vec{P} - m\phi\vec{X}. \qquad \text{(VIII.12)}$$

In spherical polar coordinates, the energy integral may be rewritten as

$$dt^2 = \frac{1}{2\left(\frac{E}{m} - \phi\right)}\left(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\varphi^2\right). \qquad \text{(VIII.11A)}$$

Since the potential ϕ is independent of angles and orbital angular momentum is conserved, it is clear that one can reduce the motion to a two-dimensional planar motion by choosing, say, $\theta = \frac{\pi}{2}$

θ = constant = $\frac{\pi}{2}$. In this case we obtain

$$dt^2 = \frac{1}{2(\frac{E}{m} - \phi)} (dr^2 + r^2 d\varphi^2), \qquad \text{(VIII.11B)}$$

where r, φ are the plane polar coordinates $(x = r\cos\varphi, y = r\sin\varphi)$. In terms of these the equation (VIII.8A) may be expanded as

$$(\ddot{r} - r\dot{\varphi}^2)\hat{r} + (r\ddot{\varphi} + 2\dot{r}\dot{\varphi})\hat{\varphi} = -\frac{MG}{r^2}\hat{r},$$

and yields the two equations

$$\ddot{r} - r\dot{\varphi}^2 = \frac{MG}{r^2}, \quad \frac{d}{dt}(r^2\dot{\varphi}) = 0 \qquad \text{(VIII.8C)}$$

These equations may be also be obtained from Newton's equations in generalized coordinates as considered in the last chapter ($r\dot{\varphi}^2$ and $\frac{2}{r}\dot{r}\dot{\varphi}$ are just components of $\left\{ {i \atop jk} \right\} \dot{x}^j \dot{x}^k$).). If we put $r = u^{-1}$ and write $\frac{d}{dt} = \dot{\varphi}\frac{d}{d\varphi}$, then on using

$$r^2\dot{\varphi} = l = \text{constant} \qquad \text{(VIII.13)}$$

We obtain for the radial equation

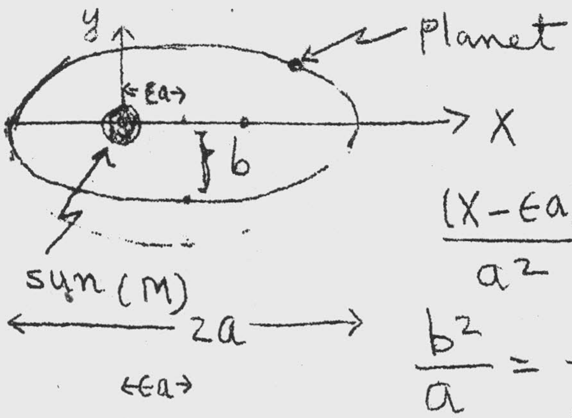
$$\frac{d^2u}{d\varphi^2} + u = \frac{\mu}{l^2}, \quad \mu = MG, \qquad \text{(VIII.14)}$$

which has the solution

$$u = \frac{\mu}{l^2}(1 + \epsilon\cos\varphi); \qquad \text{(VIII.14A)}$$

where the eccentricity $\epsilon = 1 + \frac{2E}{m} \cdot \frac{l^2}{\mu}$; the orbit is elliptical, parabalic or hyperbolic according as

$$\epsilon < 1 \ (E < 0), \ \epsilon = 1 \ (E = 0), \ \epsilon > 1 \ (E > 0)$$



$$\frac{(x-\epsilon a)^2}{a^2} + \frac{y^2}{b^2} = 1 ,$$

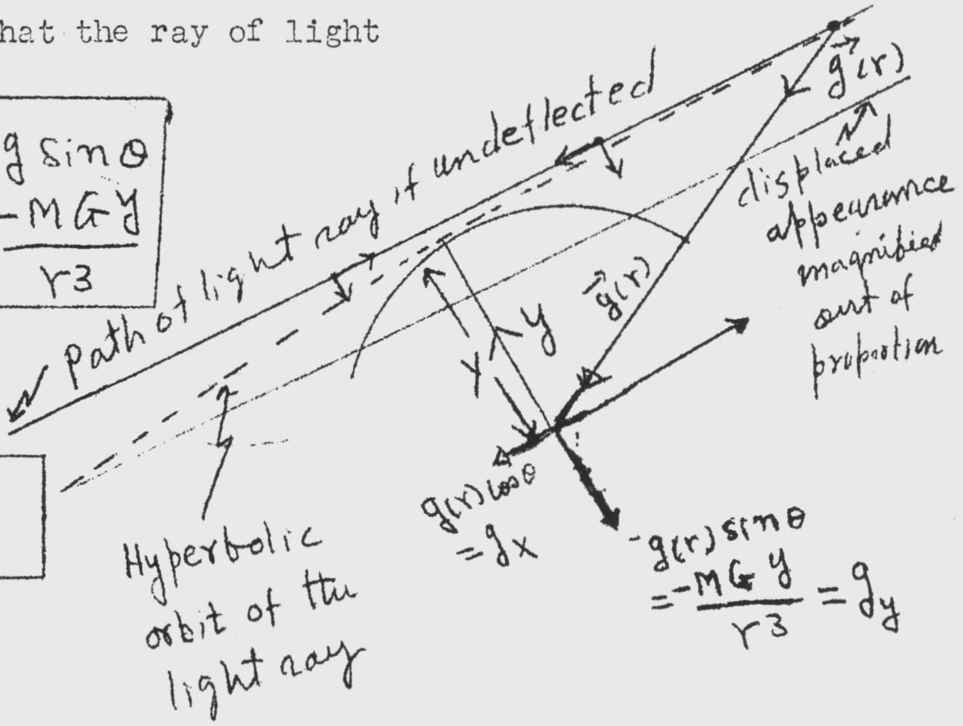
$$\frac{b^2}{a} = \frac{\ell^2}{\mu} = a(1-\epsilon^2). \qquad \text{(VIII.14B)}$$

Eilliptical orbital of a planet in Newtonian Theory.

We consider two further applications of the Newtonian theory relevant here; these were first treated in the work of A. Einstein[1)] in his attempts to develop the general theory of relativity. According to the principle of special relativity the energy equivalent of mass is mass times (velocity of light)^2. conversely radiant energy has a mass equivalent and therefore also weight due to the gravitational attraction of this equivalent mass. Consider a ray of light coming from $-\infty$. An observer at $+\infty$ will see it slightly deflected if the ray passes near a large mass during its journey because of gravitational attraction. This deflection was actually predicted much before the advent of relativity by Soldner in the eighteenth century. He was able to derive this expression without using special relativity because if one considers a light ray to consist of particles then due to the equivalence of gravitational and inertial masses the actual 'mass of the ray' is irrelevant (equivalence principle!). It is clear from the following figure that the ray of light

$$\boxed{g_y = \frac{dC_y}{dt} = g \sin\theta = \frac{-MGy}{r^3}}$$

$$\boxed{\Delta C_y = \frac{g_y}{c}\Delta x}$$



appears defelected towards a large mass by an amount (in radians)

$$\theta = \int_0^c \frac{dC_y}{C(r)} = \int_{-\infty}^{+\infty} \frac{1}{c^2}\frac{dC_y}{dt}dx = \frac{-2MG}{c^2 y} = \frac{+2\phi}{c^2} \qquad \text{(VIII.15)}$$

We can now compare the rate of a clock at rest in the gravitational field of a massive body of mass M with a clock not under the effect of the gravitational field (infinitely far away from M). Suppose a light say goes to and from between two mirrors a distance x apart. Such a system clealy constitutes a clock G. In the system at ∞ (K_∞) the time taken to traverse a distance X in the clock G at K_∞ is $t = \frac{X}{c}$. Near the massive body (system K) according to above considerations the velocity of light is (of the order) $c-\delta c$ as observed from K_∞, so that the time taken to traverse a distance X near the massive body is in the same order of approximation given by

$$\left.\begin{aligned} t_0 &\simeq \frac{X}{c-\delta c} \simeq t\left(1+\frac{\delta c}{c}\right) \simeq t\left(1+\frac{gX}{c^2}\right) \\ \underline{\text{or}}\quad t &\simeq t_0\left(1-\frac{gX}{c^2}\right) \end{aligned}\right\} \qquad \text{(VIII.16)}$$

We now put gX = negative of the gravitational potential of the sun = $-\phi$, that of earth being considered zero; then if for the clocks C_1 situated on the sun the time interval is t_0 the corresponding time interval observed on the earth would appear as

$$t \simeq t_0\left(1+\frac{\phi}{c^2}\right) \qquad \text{(VIII.17)}$$

or in terms of frequencies of the light waves

$$\nu = \nu_0\left(1-\frac{\phi}{c^2}\right), \quad \nu = \frac{2\pi}{t}. \qquad \text{(VIII.17a)}$$

If therefore a spectral line is produced on the sun, when observed on the earth its frequency will be shifted towards the red (smaller frequency) by an amount

$$\frac{\Delta\nu}{\nu_0} = \frac{\nu-\nu_0}{\nu_0} = -\frac{\phi}{c^2} \quad \left(\text{for the sun}\right) \qquad \text{(VIII.17b)}$$

Alternate way of deriving these results is to note that since $\vec{g} = d\vec{c}/dt$, therefore $c^2(r) = c_\infty^2 + 2\phi$, which gives

$$\frac{\Delta x}{\Delta t} = c(r) \simeq \left(1 + \frac{\phi}{c^2}\right) c \qquad \text{(VIII.18)}$$

From this we find $\Delta c = \frac{\phi}{c}$; and if we assume that Δx is unchanged from frame to frame, then (VIII.18) yields the results of equations (VIII.17, 17a, 17B).

Newton in his discussion of distinction between absolute and relative motion gave the example of rotational motion to illustrate absolute motion. In a certain sense absolute motion of rotation is equivalent to gravitational field. For instance in a massive rotating sphere there are two forces, viz. the gravitational centripetal force and the centrifugal force-field due to rotation; for a large enough angular velocity of rotation the two could even be balanced. We shall return to consider the significance of these remarks in the next section.

Newtonian gravitation can also be applied to consider global problems of space-time structure of the universe as a whole - Newtonian cosmology. Here again there are several problems involved For instance whether the universe is finite. What is its mass density or total mass ? Since there appear to be so many stars radiating, an interesting problem is that of density of radiation in this universe. One however gets into serious difficulties if one tries to answer these questions on the basis of Newtonian theory. With these remarks we end this section.

VIII.2 Principles of Equivalence and Foundations of Relativistic Gravitation Theory.

The exact equality of (passive) gravitational and inertial masses (equation 7), as we had remarked earlier, implies that all bodies irrespective of their composition fall in a given homogeneous gravitational field with the same acceleration. Consider a small room in free space, away from any gravitational field; if the room is accelerated, (say in the direction z, with constant acceleration g,), then an observer inside the room will find that objects when thrown in the room follow the following equations of motion

$$\frac{d^2x}{dt^2} = 0 = \frac{d^2y}{dt^2} \quad ; \quad \frac{d^2z}{dt^2} = -g \qquad (19)$$

These are precisely the equations of motion of an object in a homogeneous gravitational field in the direction $-z$. Summarising then: A gravitational field can be locally stimulated by a frame of reference in uniform acceleration; in other words a (homogeneous) gravitational field (K) is equivalent a uniformly accelerated frame of reference (K'). This is called the weak principle of equivalence.

It is clear that the results of the last section on the deflection of light and the red shift can also be obtained if instead of the homogeneous gravitational field we consider a uniformly accelerated frame and compare it with an inertial frame. In fact Einstein had originally demonstrated[1] these results by comparing K and K'. It would then follow that one

could represent the motion of a test body in any gravitational field at least in some small region by a suitable force free geodesic equation. But two things need to be settled
(1) dimensionally of space: for instance whether it is four-dimensional space-time as in special relativity; and
(2) the nature of the linear connection Γ , since it would clearly represent the gravitational field.

We had remarked that special relativity is from several viewpoints a natural completion of Newtonian mechanics. Now Newtonian gravitation also assumes Newtonian mechanics; hence any generalization of Newtonian gravitation must involve as a first step transition to special relativity. For instance let us consider the absolute motion of rotation in special relativity. In an inertial frame we have

$$\begin{aligned} d\tau^2 &= c^2 dt^2 - dx^2 - dy^2 - dz^2 \\ &= c^2 dt^2 - d\rho^2 - \rho^2 d\varphi^2 - dz^2. \end{aligned} \tag{20}$$

If the rotating frame has cylindrical symmetry about the z-axis we must replace $\varphi \rightarrow \varphi + \omega t$ where ω is the angular velocity of the rotating frame. We obtain

$$\left. \begin{aligned} d\tau^2 = {} & \left(1 - \frac{\omega^2 \rho^2}{c^2}\right) c^2 dt^2 - (2\omega\rho^2)\, d\varphi\, dt \\ & - \rho^2 d\varphi^2 - d\rho^2 - dz^2. \end{aligned} \right\} \tag{21}$$

For small angular velocities and for $v = \omega\rho \ll c$ we get the Newtonian approximation:

In the inertial frame $d\tau \approx c\,dt$

In the rotating frame $d\tau \approx \sqrt{1-\frac{\omega^2 \rho^2}{c^2}}\, c\,dt$

Let K_0 denote the inertial frame and K the rotating frame; then as observed in K_0, the clock in K shows time dilatation given by

$$c\,dt = \frac{d\tau}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{d\tau}{\sqrt{1-\frac{\omega^2\rho^2}{c^2}}} \simeq d\tau\left(1-\frac{\Phi}{c^2}\right) \tag{22}$$

where $\Phi = \frac{-1}{2}\omega^2\rho^2$ is the 'potential' of the centrifugal field of force which may be considered as a gravitational field according to the ideas discussed at the and of the last section.

Alternately, let us consider the equation of a geodesic corresponding to the metric (21); we obtain[3] on simplifying

$$\begin{aligned}
&\ddot{r} - r\dot{\varphi}^2 = -2\left\{{1 \atop 24}\right\}\dot{\varphi}\dot{t} - \left\{{1 \atop 44}\right\}\dot{t}^2 \\
&\frac{1}{r^2}\frac{d}{dt}(r^2\dot{\varphi}) = -2\left\{{2 \atop 14}\right\}\dot{\rho}\dot{t} = -2\omega\dot{\rho}\dot{t}\Big/\rho\left(1+\frac{3\omega^2\rho^2}{c^2}\right) \\
&\ddot{z} = 0 \\
&\ddot{t} = -2\left\{{4 \atop 14}\right\}\dot{\rho}\dot{t} = \frac{-6\omega^2\rho}{c^2}\left(1+\frac{3\omega^2\rho^2}{c^2}\right)\dot{\rho}\dot{t}
\end{aligned} \tag{23}$$

In the Newtonian approximation $d\tau \sim c\,dt$ and therefore $\dot{\varphi}$ and $\dot{\varrho}$ are much smaller as compared to $\dot{t}$; in this approximation we get

$$\ddot{r} - r\dot{\varphi}^2 = -\left\{ {1 \atop 44} \right\} = \text{centrifugal force field}$$

$$\ddot{z} = 0 \ , \quad \ddot{t} = 0$$

$$\frac{1}{r^2}\frac{d}{dt}(r^2\dot{\varphi}) \sim -2\left\{ {2 \atop 14} \right\} \dot{\varrho}\,\dot{t} = \text{Coriolis force field} \qquad (23A)$$

It is clear from these considerations that the generalization of Newtonian gravitation has to be such that in the absence of gravitation (weak field limit) we should get special relativity and when the velocities are small it should yield the Newtonian gravitation. Hence we have to consider the 4-dimensional space-time and motion of a test particle in a gravitational field is given by

$$\frac{d^2x^j}{ds^2} + \Gamma^j_{lm}\frac{dx^l}{ds}\frac{dx^m}{ds} = 0 \qquad (24)$$

where Γ^j_{lm} can have the general form discussed in Chapter VI

$$\Gamma^j_{lm} = \left\{ {j \atop lm} \right\} + T^j_{lm} + K^j{}_{lm} \qquad (25)$$

$$T^j_{lm} = T^j_{ml} \ , \quad T_{(jlm)} = 0 ; \quad K^j{}_{(lm)} = 0$$

where T and K are both tensors; the antisymmetric part of Γ (i.e.k.) has no effect on the system of geodesics of (24). On the other hand irrespective of T and K, the first integral of (24) may always be written

$$ds^2 = g_{ij}\, dx^i\, dx^j \, , \tag{26}$$

so that $\left\{ {j \atop l\, m} \right\}$ are completely determined in terms of metric g_{ij}. In this sense the systems (24,25) and (26) are compatible with each other. It may be pointed out here that (24) only determines the parameter S along a geodesic, whereas the parameter S in (26) may refer to any other curve also. This is as far as one can proceed on the basis of the weak principle of equivalence. To obtain general relativity we need a stronger assumption.

The experiments of Eötvös and Dicke give sensitive evidence in support of the weak principle of equivalence; we ask whether these experiments say anything more. Since any material body involves a complex of various types of 'forces' in nature (viz. electromagnetic interactions, strong interactions weak interactions; gravitational interactions), and bodies of different composition involve these various interactions in different degrees, the experiments of Eötvös-Dicke actually say a great deal more. For instance one can conclude that unto a certain degree of accuracy the various constants that arise in treating, strong, weak and electromagnetic interactions via quantum theory and special relativity are universal constants. This would imply that, locally (in a small neighbourhood), there always exists an inertial frame so that in this local frame special relativity holds. This assumption is called the <u>strong principle of equivalence</u>. It may be expressed in terms of the following thought experiment. Consider a small room falling freely in earth's gravitational field. All objects in this room would also be falling freely. Hence an observer, experimenting inside the room would feel as if he was in an inertial frame. We conclude that a freely falling frame is locally equivalent to an inertial frame. In other words, locally the space is Euclidean. It follows that

$$T^{j}_{\ lm} = 0 \ , \quad K^{j}_{\ lm} = 0 \ , \tag{24}$$

$$\frac{d^2 x^j}{ds^2} + \left\{ {j \atop lm} \right\} \frac{dx^j}{ds} \frac{dx^m}{ds} = 0 \ , \tag{24A}$$

and the space-time of general relativity is Riemannian.

In Newtonian theory, we had the equation of motion and the field equation for the potentials. We would like to know what is the corresponding field equation here. The immediate generalization of the Poisson's equation to special relativity is

$$\Box \phi = -\nabla^2 \phi + \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -4\pi G \rho \tag{28}$$

This would mean that the gravitational field is determined by a scalar potential. Another generalization is to consider the gravitational potentials to form a vector and equations may be written as

$$\left.\begin{aligned} &\Box A^{\mu} = j^{\mu} = \rho v^{\mu} (4\pi) \\ &\text{or} \\ &F^{\mu\nu}{}_{;\nu} = j^{\mu} \ , \ F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} \end{aligned}\right\} \quad (29)$$

We recall that equations (29) are equations for the electromagnetic field; in this case we know that there are two types of charges positive and negative. On the other hand there is only one type. of mass; therefore the gravitational potential cannot be a vector. At this point we note that if we expand the solution of Poisson's equation:

$$\varphi = \int \frac{G\rho(x')}{|\vec{x} - \vec{x}'|} d\tau' = \frac{Q}{r} + \frac{\vec{Q}\cdot\vec{x}}{r^3} + \frac{1}{2!}\frac{Q_{ij}x^i x^j}{r^5} + \cdots ; \quad (30)$$

where $Q = G\int \rho(x')\, d^3x'$ is a scalar, $\vec{Q} = G\int \rho(x')\, \vec{x}'\, d^3x'$ is a vector, $Q_{ij} = G\int \rho(x')(3x'_i x'_j - \vec{x}'^2 \delta_{ij})\, d^3x'$ is a symmetric traceless tensor of the second rank, etc. They are respectively proportional to the spherical harmonics $Y_0^0, Y_1^m, Y_2^m, \ldots Y_\ell^m \ldots$ etc. These are also referred to as monopole, dipole, quaderpole... 2^{ℓ}-pole moments. In the case of an electric charge distribution each of these moments are in general nonvanishing even for a system in equilibrium; hence the first nonvanishing moment of importance is the dipole moment. This being a vector, one expects the electromagnetic potentials to form a vector. On the other hand if $\rho(x)$ represents a mass distribution, then the moment corresponding to $\vec{Q}$ is the linear momentum. In conditions of equilibrium, this can always be made to vanish. The first nonvanishing moment thus corresponds to Q_{ij} and is the moment of Inertia; we therefore expect on anology with the electromagnetic case that the gravitational potentials form a symmetric tensor. This conclusion, based purely on physical considerations, can be supported by considerations based on equivalence principle and the Newtonian limit; i.e. we consider the Newtonian limit of (24A). In anology with considerations on the rotational motion, we get in the limit of small velocities (i.e. terms of order v^2/c^2 can be neglected) and weak field $(\mu = 1, 2, 3)$

$$\frac{d^2x^\mu}{ds^2} + \left\{ {\mu \atop 44} \right\} \left(\frac{dt}{ds}\right)^2 = 0 \quad , \quad \frac{d^2t}{ds^2} = 0. \tag{31}$$

Taking $ct = s$, and assuming that the field is static we get

$$\left.\begin{aligned} \frac{d^2x^\mu}{d\tau^2} = -\left\{ {\mu \atop 44} \right\} = -\frac{\partial\phi}{\partial x^\mu} \ , \ \phi = \frac{1-g_{44}}{2}c^2 \\ ds = c\,d\tau \end{aligned}\right\} \tag{31A}$$

where the additive constant in ϕ has been so chosen that in absence of the gravitational field $\phi = 0$ and g_{44} has the normal value +1. We note that actually in this approximation the other components of g_{ik} have not been assumed to be small and in fact could be of the same order of magnitude as g_{44}. But in spite this it turns out that g_{44} effectively determines the gravitational field It is for this circumstance that gravitational potential can be taken approximately as a scalar potential. However, it is clear that the complete set of potentials is given by the ten components of the metric tensor in agreement with our analysis on moments of a mass distribution.

Having determined the nature of the gravitational potentials, it is clear that the gravitational field components are given by the Christoffel symbols. Since Cristoffel symbols do not transform as tensor components they can be transformed to zero in accodence with the weak principle of equivalence. However for physical purposes it is desirable to have tensors to describe the properties of a gravitational field; the reason is that if a tensor vanishes in one coordinate system it will vanish in all other coordinate systems. We therefore look for field equations which are tensorial. Since Poisson's equation is of the second order in potential, it follows that field equations for the gravitational field should be of the second order in the metric tensor. It is known that all the tensors that are of the second order in g_{ij} can be constructed algebraically from the Riemann tensor. The various possibilities are

Riemann tensor R_{ijkl}

Ricci tensor $R_{ij} = R^{l}{}_{ijl}$

Curvature scalar $R = g^{ij} R_{ij}$

and their combinations.

In the absence of matter, we should get in analogy with Laplace's equation

$$R_{ijkl} = 0 , \tag{32A}$$

or

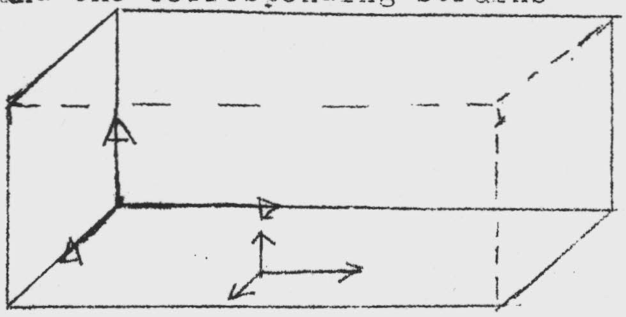
$$R_{ij} = 0 , \tag{32B}$$

or

$$R = 0 . \tag{32C}$$

The first of these would imply that space is flat in the absence of matter. This would mean that for instance outside the sun there is no curvature and hence no gravitational field! We therefore reject this possibility. The third possibility being only one condition does not appear very interesting since R_{ijkl} are 20 nonvanishing components; this would mean that there are 19 undetermined quantities whereas there are ten potentials to be determined. Hence (32B) appears to be the best candidate. We shall now give further arguments in support of this choice.

If we have an extended body, in addition to its rotational and translational motion one has also to take into account stresses acting on the body which give rise to deformation of the body called strain. Since the space is of three dimensions there are in all $\left.\frac{n(n+1)}{2}\right|_{n=3} = 6$ independent stresses in general and the corresponding strains



If $\vec{\eta}$ is a displacement vector then under stresses it may be deformed into the vector $\vec{\eta}'$. The deformation $\delta\vec{\eta}$ may be expressed in terms of the original displacement vector as

$$\delta\eta_\mu = \varepsilon_{\mu\nu}\eta^\nu \qquad \mu,\nu = 1,2,3 \tag{33}$$

where the coefficients $\varepsilon_{\mu\nu}$ contain all the information on deformation in various directions; since there can be only $\frac{1}{2}n(n+1)$ deformations the $\varepsilon_{\mu\nu}$ form a symmetric tensor and is called <u>strain tensor</u>. The stresses which cause these deformations also form a symmetric stress tensor. If $\overrightarrow{d\sigma} \sim \xi^\mu d\sigma_\mu$ denotes an infinitesimal surface and $\vec{\xi}$ the unit positive normal to it, the force acting on $d\sigma$ may be expressed as

$$F^\mu d\sigma = S^{\mu\nu}\xi_\nu d\sigma = S^{\mu\nu} d\sigma_\nu \tag{34}$$

where $S^{\mu\nu}$ are components of the stress tensor. Let j^μ be the force per unit mass and a^μ the accelerations; then

$$\int \rho j^\mu d\tau - \int \rho a^\mu d\tau = \int S^{\mu\nu} d\sigma_\nu$$

$$(d\tau = \text{volume element} = d^3x) \tag{35}$$

where ρ is the mass per unit volume; using Green's theorem (assuming suitable boundary conditions), we get for the right hand side $-\int \partial_\mu S^{\mu\nu} \, d\tau$, so that

$$\rho a^\mu = \partial_\nu S^{\mu\nu} + \rho \dot{j}^\mu . \tag{36}$$

In generalized coordinates $\partial_\mu S^{\mu\nu} \rightarrow S^{\mu\nu}{}_{;\mu}$.

In the absence of external acceleration, we have

$$\rho \dot{j}^\mu = f^\mu = -\partial_\mu S^{\mu\nu} \tag{36A}$$

for a closed system. For a closed system of particles one can thus express any force as divergence of a symmetric tensor called the stress tensor. In transition to special relativity we get a four-dimensional stress-energy tensor whose divergence is the four-force:

$$\rho \dot{j}^k = f^k = -\partial_j S^{kj} \quad , \quad j, k = 0, 1, 2, 3 . \tag{37}$$

The components of the four force are

$$f^k = \left(f^0 = \frac{\vec{f} \cdot \vec{v}}{c} \, ; -\vec{f} \right) \tag{38}$$

where the 4th component f^0 gives the measure of rate of doing work.

We note that in the case of electromagnetic forces, one can again express all forces in terms of a symmetric second rank tensor T^{jk}. If we write the left hand side of (37) as $\partial(\rho u^j u^k)/\partial x^j$ and assume that the equation of continuity (conservation of momentum density)

$$\frac{\partial}{\partial x^j}(\rho u^j) = 0 \tag{39}$$

is satisfied then equation (37) takes the form

$$\partial_j T^{jk} = \partial_j(\rho u^j u^k + S^{jk}) = 0. \tag{37A}$$

In the absence of stress the components of T^{jk} are

$$T_0^{jk} = \left[\begin{matrix} [c\rho u^\mu u^\nu \gamma^2] & \rho\vec{u}\gamma^2 c \\ \rho\vec{u}\gamma^2 c & \rho\gamma^2 c^2 \end{matrix}\right] \quad \left.\begin{matrix} u^\mu = \{\gamma\vec{u}\,;\,\gamma c\} \\ \gamma = (1-\frac{u^2}{c^2})^{-\frac{1}{2}} \end{matrix}\right\} \tag{40}$$

The components $T_0^{\mu\nu}$ correspond to a 'momentum current', $T^{0\mu}$ and $T^{\mu 0}$ to moment density and T^{00} to energy density. For this reason the tensor T^{jk} is called the energy momentum tensor. One can show from very general considerations that for a closed system of particles the force may be expressed in terms of a second rank tensor which is symmetric if angular momentum is conserved. For a closed system it is clear that $\partial_j T^{jk} = 0$. In generalized coordinates this may be written as

$$\nabla_j T^{jk} = T^{jk}{}_{;j} = 0. \tag{41}$$

From these remarks we see that the field equations of general relativity should be of the form

$$K_{ij} = T_{ij} \tag{42}$$

where K_{ij} is a tensor constructed from R_{ij} such that its divergence vanishes identically. In section VI we have already seen how this may be done: the required tensor was the Einstein tensor $G_{ij} = R_{ij} - \frac{1}{2} g_{ij} R$. Hence the field equations of general relativity may be written as

$$G_{ij} \equiv R_{ij} - \frac{1}{2} g_{ij} R = -k T_{ij} \tag{43}$$

where k is a constant to be determined by comparison with Newtonian theory.

In the Newtonian approximation we neglect all components of T_{ij} of the order $\frac{v}{c}$; the only nonvanishing component then is $T_{00} = \rho c^2$. Hence, also

$$\left.\begin{aligned} T = g^{ik} T_{ik} = g^{00} T_{00} = \rho c^2 \\ R_{00} = -k T_{00} + \frac{1}{2} g_{00} R = -\frac{1}{2} k \rho c^2 \end{aligned}\right\} \tag{44}$$

In the Newtonian approximation we also consider the gravitational field to be weak and static. Then in evaluating R_{00} , the only nonvanishing derivatives of g_{ij} are the space derivatives,

so that

$$\left.\begin{aligned} R_{oo} &= \partial_o^2 \ln\sqrt{g} - \partial_j \Gamma^j_{oo} + \Gamma^j_{ko}\Gamma^k_{jo} - \Gamma^j_{oo}\partial_o \ln\sqrt{g} \\ &\simeq -\partial_j \Gamma^j_{oo} = \frac{1}{2}\partial_j \left(g^{j\alpha}\partial_\alpha g_{oo}\right) \\ &\simeq \frac{1}{2}\nabla^2 g_{oo}, \end{aligned}\right\} \quad (45)$$

since $\partial_j = (-\nabla, \partial_o)$ and $g^{j\alpha}$ is approximately diagonal $(1, -1, -1, -1)$. If we put $\phi = \frac{1 - g_{oo}}{2}$ as in (31A), we obtain the Poisson's equation

$$\left.\begin{aligned} \nabla^2 \phi &= \frac{1}{2} k c^2 \rho = 4\pi G \rho \\ k &= \frac{8\pi G}{c^2} = 1.87 \times 10^{-27} \, cm/gm \end{aligned}\right\} \quad (4$$

This completes the derivation of Einstein equations. We make some pertinent remarks. The covariant divergence of G^i_j vanishes identically so that

$$G^i_{j;i} = -k T^i_{j;i} = 0 \qquad (47)$$

This vanishing of the divergence may be compared to equation (37A) for a closed system of particles in free space. However, there are some essential differences. In equations (43) the left hand side represents geometrical properties of space and hence the gravitational field; the right hand side represents the material world: the material energy momentum tensor. Thus there is a certain dichotomy in the definition of the energy-momentum complex: the gravitational field energy-momentum is treated on a different footing than the energy-momentum arising from matter. Another way of looking at it is that equation (47) is a covariant divergence so that one cannot apply to it the usual divergence theorems to obtain conservation laws[4] even though formally one can construct the ordinary divergence

$$T^{jk}{}_{;k} \longrightarrow \partial_k (T^{jk} + t^{jk}) = 0. \qquad (48)$$

where the quantities do not transform as components of a tensor except under linear transformations. However such a t_{jk} is not well defined in the sense that there are several arbitrary choices for it. We shall not further pursue the question of conservation laws, except to state that attempts are being made to extend the method of space-time symmetries we used in the case of Newtonian equations of motion to obtain an invariant description[5] of conservation laws in general relativity.

In arriving at the concept of stress tensor and energy momentum tensor we started with an extended body. We could have equally well started with a liquid or a gas which is very suggestive from equation (39). For an incompressible fluid the energy-momentum tensor may be written as

$$T_{ik} = \left(\rho + \frac{p}{c^2}\right) u_i u_k + p\, g_{ik}, \qquad (49)$$

where p vanishes on the boundary enclosing the liquid and thus represents pressure.

VIII.3 Some Applications of General Theory of Relativity

As an application we consider the motion of a planet around the sun. Since mass of the earth or of the other planets is much smaller than that of the sun, the problem is essentially a one particle problem. In this case we consider a planet to be a test particle and its equation of motion is given by the geodesic equation. From the nature of the problem it is clear that there is spherical symmetry in three dimensional space. Thus in the Newtonian case the first integral of motion has the form (11-A,B). Further we assume as in the Newtonian case that the motion does not depend explicitly on time; i.e. the gravitational field is static. In this case there exists a coordinate system such that $g_{0k} = 0$ and the metric is time independent $\left(\frac{\partial}{\partial t} g_{ik} = 0\right)$ It then follows (from spherical symmetry and static nature of the potentials) that metric has the form

$$d\tau^2 = A\,(c\,dt)^2 - B\,dr^2 - C\,r^2(d\theta^2 + r^2\sin^2\theta\, d\varphi^2),$$

where A,B,C are functions of r alone. By suitable choice of coordinates, one can choose $C = 1$ and obtain

$$d\tau^2 = A\,(c\,dt)^2 - B\,dr^2 - r^2(d\theta^2 + r^2\sin^2\theta\, d\varphi^2). \quad (50)$$

If we put $A=e^{\nu}$ and $B=e^{\lambda}$ we obtain for the non-vanishing components of the Christoffel symbols

$$\left\{ {1 \atop 1\,1} \right\} = \frac{1}{2}\lambda', \quad \left\{ {1 \atop 2\,2} \right\} = -re^{-\lambda}$$

$$\left\{ {1 \atop 3\,3} \right\} = -r\sin^2\theta, \quad \left\{ {1 \atop 4\,4} \right\} = \frac{1}{2}\nu' e^{\nu-\lambda} \tag{51}$$

$$\left\{ {2 \atop 1\,2} \right\} = \left\{ {3 \atop 1\,3} \right\} = \frac{1}{r}, \quad \left\{ {2 \atop 3\,3} \right\} = -\sin\theta\cos\theta$$

$$\left\{ {3 \atop 2\,3} \right\} = \cot\theta, \text{ where } \lambda' = \frac{\partial\lambda}{\partial r}.$$

It is obvious that just as in Newtonian case the motion here is again two dimensional; we can therefore take $\theta = \frac{\pi}{2}$ as the plane in which the motion takes place. Substituting, we then obtain for the geodesic, the equations:

$$\left.\begin{aligned} &\ddot{r} + 2\lambda'\dot{r}^2 - re^{-\lambda}\dot{\varphi}^2 + \frac{1}{2}e^{(\nu-\lambda)}\nu'\dot{t}^2 = 0, \\ &\ddot{\varphi} + \frac{2}{r}\dot{r}\dot{\varphi} = 0, \\ &\ddot{t} + \nu'\dot{r}\dot{t} = 0, \end{aligned}\right\} \tag{52}$$

where dot and prime respectively denote differentiation with respect to τ and r. The second and third of the equations may be immediately integrated to give

$$r^2\dot{\varphi} = \text{constant} = h \tag{53}$$

$$\dot{t} = ae^{-\nu}, \quad a = \text{constant}. \tag{54}$$

Equation (53) may be compared with the corresponding Newtonian form (13). In the Newtonian case dot refers to absolute time whereas here the dot refers to proper time. Instead of solving the first of the differential equations in (52) we consider the equation (50) as a first integral of the geodesic equation (we follow here the same procedure as in the Newtonian case 1). But before we do that we must evaluate A and B in terms of r. To this end we note that outside the body the field equations are analogous to the Laplace's equation: $\nu - ij$.

$$G_{ij} = 0 \Rightarrow R_{ij} = 0.$$

Substituting from (51) and making use of the expression for R_{ij} in terms of Christoffel symbols as given in Chapter V, we obtain

$$0 = R^1_{\ 1} = \frac{1}{r^2} - e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) \quad \text{(a)}$$

$$0 = R^2_{\ 2} = -e^{-\lambda}\left(\frac{\nu''}{2} - \frac{\lambda'\nu'}{4} + \frac{\nu'^2}{4} + \frac{\nu' - \lambda'}{2r}\right) \quad \text{(b)} \qquad (55)$$

$$0 = G^4_{\ 4} = \frac{1}{r^2} + e^{-\lambda}\left(\frac{\lambda'}{r} - \frac{1}{r^2}\right) \quad \text{(c)}$$

Combining (a) and (c) we get

$$\lambda' + \nu' = 0; \qquad (56A)$$

and on further using (b) we get

$$\nu'' + \nu'^2 + \frac{2\nu'}{r} = 0 \qquad (56B)$$

This may be solved to give

$$A = e^{\nu} = K_1 + \frac{K_2}{r} = e^{-\lambda}$$

We now use the condition that for flat space time A = 1; this gives $K = 1$ (since as $r \to \infty$ $\Longrightarrow$, the second term vanishes). Now we note that in the Newtonian approximation we found (eqn.31A) that $g_{44} = 1 - \frac{2\Phi}{c^2}$; hence comparing we find that

$$A = e^{\nu} = e^{-\lambda} = 1 - \frac{2MG}{rc^2} = 1 - \frac{2\Phi}{c^2} . \qquad (57)$$

If we combine (50),(57) together with the condition $\theta = \frac{\pi}{2}$ we obtain

$$A^{-1}\dot{r}^2 + r^2\dot{\varphi}^2 - A\dot{t}^2 = -1. \qquad (58A)$$

To eliminate and we use (53), (54) which gives

$$\left(\frac{h}{r^2}\frac{dr}{d\varphi}\right)^2 + \frac{h^2}{r^2} - \frac{a^2}{A} = -1 ; \qquad (58B)$$

and on substituting for A $(\mu = MG)$

$$\left(\frac{h}{r^2}\frac{dr}{d\varphi}\right)^2 + \frac{h^2}{r^2} = a^2 - 1 + \frac{2\mu}{r}\left(1 + \frac{h^2}{r^2}\right) . \qquad (58C)$$

We now put $u = r^{-1}$ and differentiate the resulting expression with respect to φ to obtain

$$\frac{d^2u}{d\varphi^2} + u = \frac{\mu}{h^2} + 3\mu u^2. \qquad (59)$$

From (58C) and (59) it is clear on comparing with the corresponding Newtonian expression that the effective gravitational potential in Newtonian terms is

$$\phi_{effective} = \phi_{Newtonian}\left(1 + \frac{h^2}{r^2}\right); \qquad (60)$$

i.e. there is an additional $\frac{1}{r^3}$ term that depends on h^2 and hence on angular momentum (which is a constant). On account of this we should expect precision of the orbit. This may be shown by calculating precision of the perhelion of the orbit. We shall first calculate the orbit equation by the method of successive approximations.

In the first approximation let us neglect the non-Newtonian terms; then

$$\frac{d^2u_1}{d\phi^2} + u_1 = \frac{\mu}{h^2}, \qquad (59A)$$

which has the solution

$$u_1 = \frac{\mu}{h^2}(1 + \epsilon \cos\varphi) \qquad (61A)$$

In the second approximation

$$\frac{d^2u_2}{d\varphi^2} + u_2 - \frac{\mu}{h^2} \simeq 3\mu u_1^2 \simeq \frac{3\mu^3}{h^4}(1 + 2\epsilon\cos\varphi + \cdots) \qquad (59B)$$

It's formal solution is

$$u_2 = u_1 + \frac{1}{1+D^2}\, 3\mu u_1^2 \;; \quad \left(D = \frac{d}{d\varphi}\right) \qquad (61B)$$

If in the expansion u_1^2 we retain only terms up to the first power of ϵ and neglect terms in ϵ and recall that the 'particular integral' $\frac{1}{1+D^2}\cos\varphi = \frac{1}{2}\varphi\sin\varphi$; we obtain

$$u_2 = u_1 + \frac{6\mu^3}{h^4}\,\epsilon\,\frac{1}{2}\,\varphi\sin\varphi. \qquad (61C)$$

On simplification, this gives

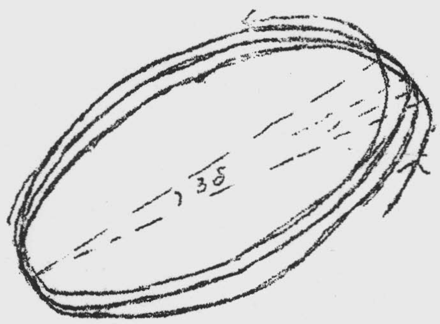
$$\begin{aligned} u_2 &= \frac{\mu}{h^2}\left[1 + \epsilon\cos\varphi + \frac{3\mu^2}{h^2}\,\epsilon\,\varphi\sin\varphi\right] \\ &\approx \frac{\mu}{h^2}\left[1 + \epsilon\cos(\varphi - \delta)\right] \end{aligned} \qquad (61D)$$

where $\delta = 3\mu^2\varphi/h^2$ is small and we have taken $\cos\delta = 1$, $\sin\delta = \delta$.

Since δ itself involves φ, it is not a mere phase difference but causes the orbit to precese. Thus δ gives a measure of the rotation of the Newtonian ellipse and hence of the perhelion and of the aphelion which advance in time. We now recall that $h^2/\mu = \sqrt{a^2+b^2} = a(1-\epsilon^2)$

so that

$$\delta = \frac{3\mu c l}{a(1-\epsilon^2)} = \frac{6\pi M G}{a(1-\epsilon^2)c^2} \text{ per revolution of the planet.} \qquad (62)$$



This rotation of the orbit has been known for a long time in the case of planet mercury since Leverier.[6] Even after applying several corrections arising from the close vicinity of the planet to the sun as also the corrections due to special relativity[7] one finds a descripancy in the precision of the perhelion of 43 seconds of an arc per century. Using the formula (62) not only has this precision been verified but also the much smaller precisions of the planets venus and earth.

We mention two further predictions of deflection of light and the red shift. According to the formulas (54) and (57), if we put a = 1,

$$ds = dt\left(1-\frac{2\phi}{c^2}\right)^{\frac{1}{2}} \approx dt\left(1-\frac{\phi}{c^2}\right). \qquad (63)$$

where dt is the measure of time interval in the gravitational potential ϕ and dS its measure as seen by an observer for whowm $\phi = 0$. In terms of frequency of spectral lines

$$\nu = \nu_0 \left(1 - \frac{\phi}{c^2}\right)$$

$$\frac{\delta \nu}{\nu_0} = \frac{-\phi}{c^2} = \frac{MG}{c^2 r} . \qquad (64)$$

This is the same as for the Newtonian theory. We mention in passing that this effect has been verified by experiments on earth by using the fact that at different heights gravitational potential is different.

In order to consider the 'deflection of light' effect we make a change in coordinate system

$$r \to \left(1 + \frac{\mu}{2R}\right)^2 R \qquad (65)$$

in equations (50), (57); this yields the isotropic line element

$$dS^2 = \left(\frac{1 - \frac{\mu}{2R}}{1 + \frac{\mu}{2R}}\right)^2 dt^2 - \left(1 + \frac{\mu}{2R}\right)^4 (dx^2 + dy^2 + dz^2) . \qquad (66)$$

A ray of light will propagate along a null geodesic and we must have for this $dS = 0$. Since in a local inertial frame we must have

$$\left| \frac{d\vec{X}}{dT} \right| = c = \text{velocity of light in an inertial frame.} \qquad (67)$$

Therefore in a general frame

$$0 \simeq \left(1-\frac{\mu}{R}\right)^2 c^2 dt^2 - \left(1+\frac{\mu}{R}\right)^2 |d\vec{x}|^2, \qquad (68A)$$

so that in this frame the velocity of light is given by.

$$c' = \left|\frac{d\vec{x}}{dt}\right| = c\,\frac{1-\mu/R}{1+\mu/R} = c\left(1-\frac{2\mu}{R}\right). \qquad (68B)$$

Comparing with the Newtonian result, eqn.(18), we see that this is smaller; hence we shall get a larger deflection:

$$\Delta\theta = \int \frac{dc'}{c} = \int_{-\infty}^{+\infty} \frac{1}{c}\frac{\partial c'}{\partial y}\,dx = \frac{4MG}{c^2 R_0} \qquad (69)$$

for R_0 as radius of the body of mass M in whose vicinity the ray of light is passing and hence suffers a deflection. This deflection, as we see is twice the Newtonian value.

There are several other experiments conducted more recently or are in the process of implementation with the advance of technology but we shall not discuss these here.

References

1) Annals of Physics 26 (1964) 442

2) On the influence of gravitation on the propagation of Light, A Einstein in Annalen der Physik, 35, 1911. Translation in The Principle of Relativity with notes by A. Sammerfeld, Dover, Publications (1923)

3)

$$g_{11}=g_{33}=-1\ ,\ g_{22}=-\rho^2,\ g_{44}=c^2-\omega^2\rho^2$$

$$g_{24}=-2\omega\rho^2\ ;\ g^{11}=g^{33}=-1,\ g^{22}=g_{44}/g$$

$$g^{44}=g_{22}/g,\ g^{24}=-g_{24}/g\ ;\ g=-\rho^2c^2\left(1+\frac{3\omega^2\rho^2}{c^2}\right)$$

Nonvanishing components of Christoffel symbols are

$$\left\{ {1 \atop 22} \right\}=-\{22,1\}=-r,\ \left\{ {1 \atop 44} \right\}=-\{44,1\}=-\omega^2 r$$

$$\left\{ {1 \atop 24} \right\}=-\{24,1\}=-2\omega\rho$$

$$\left\{ {2 \atop 12} \right\}=\frac{1}{r}\ ,\ \left\{ {2 \atop 14} \right\}=\frac{2\omega}{r(1+3\omega^2r^2/c^2)}\ ,$$

$$\left\{ {4 \atop 14} \right\}=3\omega^2 r\Big/c^2\left(1+\frac{3\omega^2r^2}{c^2}\right)\ ;$$

where dot denoted differentiation with respect to τ . Reader is highly recommended to work out these expressions to gain practice without which it is difficult to obtain any feeling for the subject.

4) If $\partial j^k/\partial x^k=0$ then by integrating over a four-dimensional volume bounded by two hypersurfaces t = constant one finds that $q=\int j^0\,dx\,dy\,dz$ has the same value on the two hypersurfaces and is therefore a constant of the motion. Involved in this derivation is the assumption that j^k rapidly decrease outside a certain region and may therefore be considered to vanish outside some region. Under similar

assumptions one finds that if $\partial_j T^{jk} = 0$, then
$P^k = \int T^{0k}\, dx\, dy\, dz$ are constants of the motion: these are the components of energy and momentum.

5) K.H.Mariwalla 'Coordinate transformations that form groups in the Large' in <u>Lectures in Theoretical Physics</u> Vol.XIII Ed. Barut and Brittin, Colorado Associated University Press (1971). Also see conservation laws, etc. in 'Proc. of the conference on Cosmology, Gravitation and its applications to Particle Physics', Matscience Report 76.

6) U.J.Leverier, Ann. Obs. Paris, Vol. 5 (1859).

7) According to special relativity the effective mass increases with velocity $m \rightarrow m / \sqrt{1 - v^2/c^2}$; this effect in the case of a classical electronic orbit around the nucleus was first predicted by Sommerfeld. The effect of this is to cause a similar rotation of the orbit and is called Rosseti motion.

ERRATA

Page 2 : line 8 the netrual element of S_+ is .....

line 9 with respect to the ' ' operation

line 10 If x, y, z are any .....

line 14, 7th word, commutative (abelian)

Page 4 : line 12 .. An ordered pair N linearly ...

line 20 read 'denote' in place of 'denoted'

Page 5 : line 2 K^n

Page 8 : line 11 letters <u>one</u> could <u>also use</u> different ...

Page 9 : line 7 first letter reads: It

line 15 ~~matrix~~ neutral element with .....

Page 13 : line 10 read $n = 2^m$ for $n = e^{2m}$

Page 15 : line 14 ... are in general n.....

line 15 .... Hence there is associated with n ....

Page 16 : line 1 $SMS^{-1} = \Lambda$, $\Lambda^i_j = \lambda_{(i)}\,\delta^{(i)}_j$

Page 17 : eq. (36) replace a everywhere by A.

line 7 $\delta^{k_1\ldots k_r}_{l_1\ldots l_r}\,\delta^{l_r}_{k_r} = (n-r+1)\,\delta^{k_1\ldots k_{r-1}}_{l_1\ldots l_{r-1}}$

line 14 ... is a Scalar, det ....

line 18 ....(30) that <u>at least</u> one root of λ

line 19 symmetric → skew-symmetric

Page 18 : line 12 read matrix A, for matrix M

Eq.(35) replace a everywhere by A

Bage 19 : Eq. (36) Ind. term reads

$$\frac{1}{(n-1)} \delta^{l_1 \cdots l_n}_{q_1 \cdots q_n} \frac{\partial A^{q_1}{}_{k_1}}{\partial t} A^{q_2}_{l_2} \cdots A^{q_n}_{l_n}$$

Page 19 : Eq.(37) replace α by a on the right side

Page 20. : line 16, read $(\vec{x}, \vec{u})$ for $(x, \ )$

Page 21 : line 15, ... spaces as the tensor product is ....

Page 22 : line 3 ... clear that there can be several tensor...

Page 23 .: line 5, eq.(12) ... $= A^{i'}_{j} A^{k}_{k'} e^{i}(\vec{e}_k)$

Page 25 : lin2 12, ... generalized δ

Page 31 : line 3 ..... two vectors $\vec{x}$ and $\vec{y}$

line 17 cross out: (V.16)

Page 34 : Eq.V.22 $\Lambda = R_a H_k R_b$; R = Ordinary rotation

H = Hyperbolic rotation

Eq.V.23

$$\left[\begin{array}{c|c} a & \vec{0} \\ \hline \vec{0} & 1 \end{array}\right] \left[\begin{array}{c|c} \lambda & \vec{\Lambda}_0 \\ \hline \vec{\Lambda}^0 & \Lambda^0_0 \end{array}\right] \left[\begin{array}{c|c} b & \vec{0} \\ \hline \vec{0} & 1 \end{array}\right] = \left[\begin{array}{c|c} a\lambda b & a\vec{\Lambda}_0 \\ \hline \vec{\Lambda}^0 b & \Lambda^0_0 \end{array}\right]$$

Page 38 : Line 17 column in place of colmn

Page 39 : last line one $\Rightarrow$ are

Page 40 : line 3, ... If this $\longrightarrow$ In this

Page 41 : line 13 .... $\partial/\partial x^j$ then give the ....

Page 43 : line 3 .... arranged in the form of ....

Page 44 : line 14, ... are an important analogue of a set ...

Page 44 : line 14, it is basic to the Euclidean geometry

Page 48 : line 15, this $\longrightarrow$ the

Page 54 : line 4, dourth $\rightarrow$ fourth

Eq.(V.81) $R^h{}_{ijk;m} + R^h{}_{ikm;j} + R^h{}_{imj;k} = 0$

Eq.(V.82) $R^h_{ijk;m} + R^h{}_{jmi;k} + \cdots + \cdots = 0$

Page 59 : line 11, (an open subset in R^n) of P.

Page 60 : line 5, (3) As (3) is true ....

Page 62 : line 11, ... under the coordinate change

line 16, $T_p(S) \times S \Rightarrow S$

Page 64 : line 5, decompared $\Rightarrow$ decomposed

Page 65 : line 17, $u^l u^m \partial_l g_{ij} \Rightarrow u^l \partial_l g_{ij}$

Page 67 .: line 1, Christoffed $\longrightarrow$ Christoffel

line 11, ... restruction on γ^l_{ij} is the ...

line 16, (32) $\longrightarrow$ (31)

Page 68 .: line 11, antisymmetic part of the affinity

Page 70 : line 4, It therefore represents $\binom{n}{4}$ conditions, equivalent $R_{(klji)} = 0$

Page 70 : line 13, Rieci $\rightarrow$ Ricci

Page 74 : line 1, are $\longrightarrow$ we

line 11, ... Thes each $\longrightarrow$ Then each ....

Page 81 : line 14, ... tensor $\underline{L_{ij}}$ are ....

Page 82 : line 1, ... an even permutation of 1,2,3,-1, for odd

line 5, $\overrightarrow{V}$ is a ....

Page 88	:	line 2 (V.7.7A) $\longrightarrow$ (V.7,8)
Page 96	:	line 20, ... -1 and +1
Page 101	:	line 2, ... equations (31), (32), ~~are~~ we
		line 15, ... side of (31) unchanged;....
Page 104	:	line 9, 11, .. (31) $\longrightarrow$ (35)
Page 105	:	line 4 (36) $\longrightarrow$ (35C)
		line 5, (37) $\longrightarrow$ (31)
		last line (32) $\longrightarrow$ (32A)
Page 106	:	last equation (32A) $\longrightarrow$ (32B)
Page 108	:	line 3... mean the difference between the number...
Page 114	:	line 18, ... of inertial mass M_I is
Page 115	:	line 8 , ... conveniently put
Page 116	:	line 8 , (P. G. Roll, ...)[1] see ref.1.
Page 117	:	last line, delete this line.
Page 120	:	line 4, A.Einstein[1].... $\longrightarrow$ A.Einstein[2] ....
Page 125	:	line 4, (1) dimensionality ....
Page 139	:	line 9, ... the quantities t_{ik} do not ....
Page 43	.:	line 14, To eliminate $\dot{q}$ and $\dot{t}$ we .....