Mathematical Preparations

1 Introduction

The theory of relativity was developed to explain experiments which studied the propagation of electromagnetic radiation in moving coordinate systems. Within experimental error the velocity of EM radiation was found to be un-affected by the motion of source of radiation with respect to the coordinate system, *i.e.* the space through which the EM radiation travels.

On the other hand, Newton postulated that in the absence of forces, the spatial coordinates of a moving point are linear in time. Thus the position of a particle measured in a rest frame of reference, X', related to its position in a moving frame, X, is given by $X = X' + V_0 t$. Using this coordinate transformation, the velocities in two systems moving with constant velocity, V_0, relative to each other is;

$$\frac{dX}{dt} = \frac{dX'}{dt} + V_0$$

$$V = V' + V_0$$

The above equations represent a spatial transformation between coordinate frames moving with constant velocity, V_0, with respect to each other. It is presumed that there is a fixed, universal coordinate frame (inertial frame), and all frames moving with constant velocity with respect to this frame have the same acceleration;

$$\frac{d^2X}{dt^2} = a = a'$$

The invariance of the laws of physics in different coordiante frames is a symmetry called the "Principle of Relativity". In the above case, Newton's laws of motion are the same in all inertial frames, as the force (acceleration) $\vec{F} = M\vec{a}$ is independent of the inertial frame. However, the independence of the velocity of EM waves between different coordinate frames is not consistent with Neutonian physics.

2 Galilean transformation

To be consistent, the mathematical form of all physics laws cannot be changed by a coordinate transformation. In the case of Newtonian physics, a transformation between inertial frames preserves Newton's laws of mechanics, and is called a Galilean transformation. The Galilean transformation transformation is defined below.

Suppose 2 reference frames related to each other by a constant velocity along the Z axis, (X, Y, Z) and (X', Y', Z'). A system not subject to a force experiences no force in any inertial system. Thus if the force in one frame is given by $\vec{F} = M\vec{a}$ then the force in the other frame is; $\vec{F} = \vec{F'} = M\vec{a}'$ with $\vec{a} = \vec{a}'$;

$$\frac{d^2X'}{dt^2} = \frac{d^2X}{dt^2} = a$$

The Galilean transformation between inertial systems must take the form;

$$\begin{aligned} X' &= X \\ Y' &= Y \\ Z' &= Z + V_0 t \\ t' &= t \end{aligned}$$

Although Newton's laws of mechanics are invarient under a Galilean transformation, Maxwell's equations which describe electrodynamics are not, and this was recognized long before the theory of relativity. Thus the description of electromagnetic radiation was inconsistent with a Galelian transformation. It was originally thought that Maxwell's equations were incomplete, and theories were proposed to correct EM under the assumption that a Galelian transformation correctly described the coordiante transformation between moving bodies. We now know of course, that EM was correct and Newtonian mechanics required modification.

2.1 Generalized coordinates

Because 4 rather that 3 dimension (3 spatial and one time coordinate) are necessary to describe the relative motion of systems, it is important to first discuss geometry and transformations in a generalized set of coordinates. Most students have been minimaly exposed to this mathematics. However, only the parts of tensor analysis required for special relativity are developed here. General relativity requires more in-depth development which is not necessary for the study of classical electrodynamics.

All coordinate systems are defined relative to a Cartesian set of axes. For 3-D write (x_1, x_2, x_2), although extension to more spatial dimensions is trivial. Thus there is a 3-D function of the coordinates which locates some point in space. This point can also be located in a different coordinate frame, $\zeta_i \, (i = 1, 2, 3)$;

$$\zeta_i(x_1, x_2, x_3) \quad i = 1, 2, 3$$

There also exists a unique inverse of the transformation function between the coordinates. Mathematically, this is described by a one-to-one mapping of each point in one frane to one point in the other. This mapping must have a unique inverse so that each point has only one location in all frames of reference.

$$x_i(\zeta_1, \zeta_2, \zeta_3) \quad i = 1, 2, 3$$

Now at the intersection of the planes;

$$\zeta_i \;=\; constant_i \quad i = 1, 2, 3$$

define a set of unit vectors, $\hat{a}_i$, perpendicular to each surface. If these vectors are mutually orthorgonal, an orthorgonal coordinate system is defined. A reference frame with orthogonal coordinates is not necessary in general, but orthogonal coordinates greately simplifies the mathematics. The direction cosines of the coordinates with respect to the set of Cartesian unit vectors are;

$$\hat{a}_1 \cdot \hat{x}_i \;=\; \alpha_i \;=\; \gamma_{1i}$$

$$\hat{a}_2 \cdot \hat{x}_i \;=\; \beta_i \;=\; \gamma_{2i}$$

$$\hat{a}_3 \cdot \hat{x}_i \;=\; \gamma_i \;=\; \gamma_{3i}$$

For an orthorgonal system, the 3 non-trivial direction cosines are related, as may be shown by calculating $\hat{a}_i \cdot \hat{a}_j$ for $i, j = 1, 2, 3$.

$$\sum_{s=1}^{3} \gamma_{ms}\gamma_{ns} \;=\; \sum_{s=1}^{3} \gamma_{sm}\gamma_{sn} \;=\; \delta_{mn}$$

Then;

$$\hat{a}_n \;=\; \sum_j \gamma_{nj}\, \hat{x}_j$$

$$\hat{x}_j \;=\; \sum_n \gamma_{nj}\, \hat{a}_n$$

Now consider the differential element of length, $|\vec{ds}|$. In the Cartesian system, the square of this element is;

$$\vec{ds} \cdot \vec{ds} \;=\; \sum_{i=1}^{3} dx_i^2$$

Suppose a general curvilinear set of coordinates is introduced as defined above.

$$dx_i \;=\; \sum_{j=1}^{3} \frac{\partial x_i}{\partial \zeta_j}\, d\zeta_j$$

The square of the length elements is then

$$ds^2 \;=\; \sum_{j,k=1}^{3} \sum_{i=1}^{3} \frac{\partial x_i}{\partial \zeta_j} \frac{\partial x_i}{\partial \zeta_k}\, d\zeta_j\, d\zeta_k$$

This is rewritten as ;

$$g_{jk} \;=\; \sum_{i=1}^{3} \frac{\partial x_i}{\partial \zeta_j} \frac{\partial x_i}{\partial \zeta_k}$$

where g_{jk} are the metric elements which define the space. Therefore;

$$ds^2 \;=\; \sum_{jk} g_{jk}\, d\zeta_j\, d\zeta_k$$

In the case of an orthorgonal system $g_{jk} = 0$ if $j \neq k$, so define a scale factor, $h_j^2 \;=\; \sum_i (\frac{\partial x_i}{\partial \zeta_j})^2$. Note that $h_j\, d\zeta_j$ is the length element for the j^{th} coordinate.

$$ds^2 \;=\; \sum_i (h_i d\zeta_i)^2$$

Using this, one can obtain the differential volume and area elements;

$$d\tau \;=\; (h_1 d\zeta_1)(h_2 d\zeta_2)(h_3 d\zeta_3)$$

$$d\sigma_k \;=\; (h_i d\zeta_i)(h_j d\zeta_j)$$

To obtain the various surface areas in the above, apply cyclic permentations of i, j, k. In general h_i varies at each point in the coordinate space. The direction cosines along the new coordinate axes (ONLY for an orthorgonal system) are;

$$\gamma_{ni} \;=\; (1/h_n)\, \frac{\partial x_i}{\partial \zeta_n} \;=\; h_n\, \frac{\partial \zeta_n}{\partial x_i} \text{ (no sum)}$$

Not only do the scale factors change with position, but also the unit vectors change directions, Fig. 1. For example,

$$\hat{a}_j \;=\; \sum_i \frac{\hat{x}_i}{h_j} \frac{\partial x_i}{\partial \zeta j}$$

$$\frac{\partial \hat{a}_j}{\partial \zeta_k} \;=\; \frac{\partial}{\partial \zeta_k} \sum_i \frac{\hat{x}_i}{h_j} \frac{\partial x_i}{\partial \zeta_j}$$

Which can be reduced to;

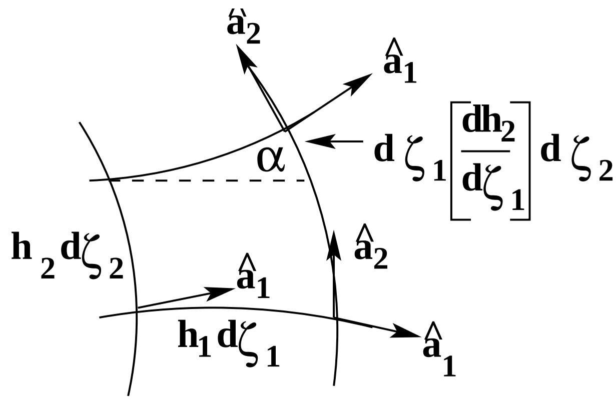

Figure 1: A cross section of an area element in a generalized coordinate system

$$\frac{\partial \hat{a}_j}{\partial \zeta_i} = \frac{\hat{a}_i}{h_i}\frac{\partial h_j}{\partial \zeta_j}$$

It is then interesting to apply these equations to a familiar coordinate system. Use spherical coordinates for this example.

$$x = rcos(\phi)sin(\theta) \quad \hat{a}_1 = sin(\theta)\,cos(\phi)\,\hat{x} + sin(\theta)\,sin(\phi)\,\hat{y} + cos(\theta)\,\hat{z}$$

$$y = rsin(\phi)sin(\theta) \quad \hat{a}_2 = cos(\theta)\,cos(\phi)\,\hat{x} + cos(\theta)\,sin(\phi)\,\hat{y} - sin(\theta)\,\hat{z}$$

$$x = rcos(\theta) \quad \hat{a}_3 = -sin(\phi)\,\hat{x} + cos(\phi)\,\hat{y}$$

Take the partial derivatives to show that an orthorgonal system is produced ($\sum_i \frac{\partial x_i}{\partial \zeta_j}\frac{\partial x_i}{\partial \zeta_k} = 0$). The square of the metric length is;

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 sin^2(\theta) d\phi^2$$

as expected. Unit vectors, volume/area elements, and the vector operations *gradient*, *div*, and *curl* can be obtained from the physical definition of these operators.

2.2 Tensors

Tensors are defined by considering the transformation properties of functions under a coordinate rotation and reflection. Thus a scalar function does not change value under rotation or reflection. As an example the function $f = \sum_{i=1}^{3} (x_i - x_{0,i})^2$ remains constant and for this example, is the magnitude of a vector. On the other hand if we consider;

$$\vec{\nabla} f = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} \hat{x}_i$$

then $\vec{\nabla} f$ transforms as a vector which preserves magnitude but changes direction. It also changes sign upon reflection $x_i \to -x_i$. All these properties are preserved when a coordinate transformation is applied so that the representation of a vector is independent of the coordinate frame. A true scalar remains the same under all coordinate transformations, including reflections. However, if a scalar function is constant under rotation but changes sign under reflection it is a pseudo-scalar. Similarly if a vector does not change sign under reflection it is a pseudo-vector. As an example, a pseudo-vector is the result of the cross product of 2 true vectors as will be observed below.

Now generalize this description of functions by defining a scalar function as a tensor of rank 0, and a vector function as a tensor of rank 1. This can be generalized by extending the transformation properties to higher rank. To help with notation, the summation convention is employed unless it leads to ambiguities. The summation convention suppresses the $\sum$ symbol and is represented by a repeated index on the variables. Thus the definition;

$$\sum_i \frac{\partial x_k}{\partial x'_i} \frac{\partial x'_i}{\partial x_j} = \frac{\partial x_k}{\partial x'_i} \frac{\partial x'_i}{\partial x_j}$$

Suppose an n-dinemsional space, with N independent variables x_i $i = 1, \cdots, n$. The set of x_i define a point in this space. Now define a set of n linearly independent functions $\zeta_i(x_1, \cdots, x_n)$ $i = 1, \cdots, n$. The Jacobian of a set of lineraly independent functions does not vanish.

$$J = \begin{vmatrix} \frac{\partial \zeta_1}{\partial x_1} & \cdots & \frac{\partial \zeta_n}{\partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial \zeta_1}{\partial x_n} & \cdots & \frac{\partial \zeta_n}{\partial x_n} \end{vmatrix} \neq 0$$

The functions, ζ_i, define a new coordinate system. Make the substitution $x'_i = \zeta_i$, and evaluate $\frac{\partial x_k}{\partial x_j}$ for future use.

$$\frac{\partial x_k}{\partial x_j} = \delta_{ij} = \frac{\partial x_k}{\partial x'_i} \frac{\partial x'_i}{\partial x_j}$$

In addition;

$$dx'_i = \frac{\partial x'_i}{\partial x_j} dx_j$$

3 Tensor contraction and direct product

In the following, use the results of the differential operations between the primed and unprimed frame which were obtained in the last section. The differential quantities dx_i and dx'_j are related by a linear transformation, $\dfrac{\partial x'_i}{\partial x_j}$. A tensor function is defined by the linear transformation of its differential form between two coordinate frames. Thus a tensor, A, of rank 1 (a vector) has the transformation properties;

$$A'^{i} = \sum_j \frac{\partial x'_i}{\partial x_j} A^j$$

For the record, this is a contravarient tensor indicated by the super-scripted index. A sub-scripted index indicates a covarient tensor, and higher order tensors with both super- and sub-scripts are called a mixed tensor.

$$A'_i = \sum_j \frac{\partial x_j}{\partial x'_i} A_j$$

For Cartesian coordinates covarient and contravarient tensors are identical since;

$$\sum_i \frac{\partial x'_i}{\partial x_j} \frac{\partial x_k}{\partial x'_i} = \delta_{jk}$$

The contraction of any tensor by a vector for example (for 2 vectors this is the dot product) reduces the order of the tensor by one (the rank of the tensor less the rank of the vector). In the case of contracting 2 vectors a scalar is produced.

$$\sum_i A'^{i} B'_i = \sum_{ijk} \frac{\partial x_j}{\partial x'_i} \frac{\partial x'_i}{\partial x_k} A^j_k = \sum_j A^j B_j$$

On the other hand, the direct product of 2 tensors multiplies each element of a tensor by the elements of the other tensor. This increases the rank of the tensor by the sum of the ranks of each tensor. Thus the direct product of a tensor of rank 1 (a vector) with another tensor of rank 1, produces a tensor of rank 2 (a matrix).

$$A'^{i} B'_l = \sum_{j,m} \frac{\partial x_j}{\partial x'_i} \frac{\partial x'_l}{\partial x_m} A^j B_m$$

Note the above form transforms like a tensor of rank 2.

4 The metric tensor

As previously, the square of the length element is;

$$ds^2 \; = \; dx^i dx_i \; = \; g_{jk} dx'^j dx^{k\prime}$$

The g_{ij} form a tensor of 2^{nd} rank called the metric tensor of the space. The determinant is $g \; = \; |g_{ij}| \neq 0$. It is possible in general to have $ds^2 < 0$, however, this would not be consistent with length, so the measure of the space is taken as the absolute value of ds^2. Note that ds^2 is a tensor of rank 0, *ie* a scalar quantity.

5 Levi-Civita tensor

It is useful to define the following tensor of rank 3 or higher.

$$\epsilon_{ijk} = d$$

The constant d takes on the following values. $\epsilon_{ijk} \; = \; 1$ when $i, j, k \; = \; 1, 2, 3$ if $i \neq j \neq k$ and with an even permutation of $1, 2, 3$. The tensor equals -1 if the indicies are an odd permutation of $1, 2, 3$ and the tensor is 0 if any of the indicies have the same value. This tensor is a pseudo-tensor, *ie* a tensor with inverted symmetry upon interchange of indicies. A conjugate tensor with the same properties can also be defined. The contracton of a pseudo-scalar tensor tensor with another tensor produces another pseudo-tensor, perhaps a pseudo-scalar or a pseudo-vector. This leads to the definition of dual tensors to be defined below. The vector cross product (vector product) is a tensor of rank 2 but it has a dual representation as a pseudo-vector.

6 Contraction with the Levi-Civita tensor

Suppose an anti-symmetric tensor of rank 2, $A_{ij} \; = \; -A_{ji}$. We contract this tensor with the Levi-Civita tensor of rank 3, ϵ_{ijk}

$$[A_{ij}] \; = \; \begin{pmatrix} 0 & A_{12} & A_{13} \\ -A_{21} & 0 & A_{23} \\ -A_{31} & -A_{32} & 0 \end{pmatrix}$$

Thus;

$$\sum_{ij} A_{ij} \, \epsilon_{ijk} \; = \; [A_{ij} - A_{ji}]_k$$

It is obvious that this results in a form which transforms like a tensor of rank 1, but does not change sign under a coordinate inversion. It is then a pseudo-vector and a dual of the tensor of second rank. It is also clear that this represents the cross product of two vectors.