NOTES on SPECIAL RELATIVITY

R. M. KIEHN 01/15/88 through 01/22/97

Introduction. (Dogma, mis-interpretations, what is a Signal?)

Much has been written about the theory of Special Relativity since its inception near the turn of the century. Remarkably enough, a certain amount of (not quite correct) dogma that was developed during the first 50 years of the theory has been propagated into current textbooks. For example, in our current text (KRANE) there is picture of a globe which is said to have the appearance of an ellipsoid to observers in highspeed relative motion. The statement is just not true, and is a hangover from G. Gamow's early popular book on relativity theory. Rods do not "shrink". Twins may or may not "age". Mass does not "change". -- so then, just what is this special relativity all about?

The point of departure from "classical" physics started with the Michelson experiments (well before 1900) where is was observed that measurements of transverse electromagnetic waves propagating in the vacuum did not agree with the behavior of longitudinal sound waves propagating in a fluid media. The relative translational motion of the media-and-source or media-and-receiver influenced the effects observed with sound waves, but the relative *translational* motions at constant velocity of the media did not seem to be of observable consequence for transverse electromagnetic waves when the "medium" was the "vacuum".

However note that in elementary discussions of special relativity it is usually **NOT** stated (or ignored) that constant relative *rotational* velocity effects (the Sagnac effect) are observable in the "vacuum", while constant relative translational velocity effects are not. Michelson did experiments on both phenomena.

In ordinary wave motion, it is presumed that the frequency is determined by the source, the phase propagation speed is determined by the medium (and its motion) , and the observable wavelength adjusts to preserve the fundamental equation, λf = phase velocity. In the Michelson experiment, the relative translational speed of the "vacuum" did not seem to give any electromagnetic wavelength changes, and hence no interference fringes were observed. (Slightly different wavelengths superimposed will produce spatial intensity modulations called fringes. Slightly different frequencies superimposed will produce temporal modulations called beats.)

To explain this phenomenon, Albert E. suggested that the speed of a light "signal" relative to the vacuum was constant for all observers - independent from constant relative translational velocities. He defined a light "signal" as something that moved along paths of "zero" differential length in the 4 dimensional, non-euclidean geometry of space time. Just how this idea of a signal relates to electromagnetism was never made clear by Albert.

It was not until some 20 years later that V. Fock demonstrated that Maxwell's equations of electromagnetism admitted **discontinuous** solutions, and that these discontinuities propagate with the speed $c = 1/\sqrt{em}$. Light Signals therefor can be defined as propagating discontinuities in **E** and **B** field strengths. Thank you Professor Fock for telling us what a signal is.

Albert also made the claim that the laws of "nature" are the same for all observers. This idea originally was interpreted as meaning that the equations of physics would have the same ("covariant") format in all inertial frames of reference. But what was an inertial frame? Newton had expressed the existence of a unique fundamental inertial reference system, but now-a-days an inertial system is one element of an equivalence class of reference systems which have the "properties" of the vacuum. These properties of an inertial system are :

1. The domain is isotropic
2. The domain is homogeneous
3. The speed of light is c, independent of direction and relative translational motion.

Prior to Albert E.'s 1905 paper, Lorentz had established that Maxwell's equations were preserved by a linear transformation now called the Lorentz transformation. But this result was obtained before the Tensor Calculus was invented. At the time, it was heralded as a great achievement, and the result is still adulated by most scientists. Albert used the Lorentz transformations as an example of defining an equivalence class of inertial reference systems which contain the vacuum as a component.

Later on Albert recognized that if the laws of nature - such as Maxwell's equations - were expressed as Tensor Equations, then the concept of covariance would be satisfied for all frames of reference which were related one unto another by means of a map which was continuous, differentiable, and had a continuous differentiable inverse. Such maps are defined as diffeomorphisms (the technical word for a coordinate transformation). The importance of the Lorentz transformation (which is one of many many diffeomorphisms) paled upon the recognition of the Tensor Calculus. Still it is often stated that Maxwell's equations do not preserve their form covariantly under a Galilean transformation, but do maintain their form under a Lorentz transformation. This statement is (without qualification) **not true**. In fact this mis-statement is often used to focus importance on the Lorentz transformations. The Galilean transformation is a diffeomorphism, and as Maxwell's equations can be written as tensor equations, they will transform in a covariant manner with respect to Galilean transformations, Lorentz transformations, or any other diffeomorphism.

So then, just what is the importance, what is the significance of the Lorentz transformations?

Again V. Fock gave the explanation. Now that he had established that a light signal is a propagating discontinuity (like a shock wave), the question to be asked is What are the equivalence classes of frames of reference such that a discontinuity to one observer appears as a discontinuity to another? Fock proved that the Singular, or Discontinuous solutions to Maxwell's equations (not the equations, but their singular solution sets) were preserved by **TWO** (2) transformations types. The first one, which is the only linear diffeomorphism, Fock proves is the class of Lorentz transformations. The importance of the Lorentz transformations therefore resides with the fact that a signal (Field discontinuity) to one observer in the equivalence class appears as a signal (Field discontinuity) to another observer in the equivalence class -- not with the fact that

Maxwell's equations are preserved in covariant form. It is possible to show that the propagation speed of the discontinuity is a finite constant for all elements of the equivalence class of reference systems related one to another by means of a Lorentz transformation.

HOWEVER, Fock also proved that there was one other class of transformations that will make two observers agree that they both are observing a propagating discontinuity. That (only) other transformation is the Moebius fractional transformation (a non-linear transformation of conformal projection) for which the speed of discontinuity propagation can be infinite!!! This result has been mostly ignored in current scientific literature.

2. Lorentz transformations

Special relativity may be described as the study of the equivalence class of Lorentz transformations acting on a four dimensional variety of space and time. The basic philosophy of the theory is that the Newtonian hypothesis of one absolute (3 dimensional) "inertial" frame of reference, with a universal time keeping system, is to be replaced by a large equivalence class of (4 dimensional) space time reference systems, with respect to which measurements made in terms of electromagnetic "signals" maintain a common significance for all observers. The Lorentz transformations, which supposedly link each of the members of this special equivalence class of reference systems, are one of many possible sets of transformations that generate equivalence classes. The Lorentz group is a subset of the most general linear group of transformations on a space of four variables or dimension. The general linear group (and each of its subgroups) may be given a matrix representation with n^2 independent parameters (or less). The action of the transformation group on a vector of the space is given by the matrix equation, where the n-by-n matrix representing the transformation acts on a column position vector, **x**, of n components transforming in into a new vector, **y**.

$$\begin{bmatrix} m11 & m12 & m13 & m14 \\ m21 & m22 & m23 & m24 \\ m31 & m32 & m33 & m34 \\ m41 & m42 & m43 & m44 \end{bmatrix} \left| \begin{matrix} x1 \\ x2 \\ x3 \\ x4 \end{matrix} \right\rangle = \left| \begin{matrix} y1 \\ y2 \\ y3 \\ y4 \end{matrix} \right\rangle$$

However, the Lorentz equivalence class is a subset of this most general linear group of transformations, for instead of requiring $n^2 = 4^2 = 16$ parameters it utilizes only 10 parameters. Sometimes it is argued physically these ten parameters correspond to the three parameters needed to describe a rotation about the origin of a local frame of reference, the three parameters needed to describe the relative velocity of this local frame of reference, and the four parameters required to define the "origin" of the space and time coordinates. Why this ten parameter group is important to the processes of physics will become evident below.

It may be shown abstractly that for an equivalence class of transformations to preserve "the property of maximum uniformity" then the number of independent parameters (or independent matrix elements) must be n(n+1)/2, which for n = 4 gives precisely the 10 parameters needed to describe the Lorentz transformations. The Lorentz transformations thus fit the bill of being elements of the equivalence class of "most uniformity" for the four dimensional arena of space time. "Uniformity" is interpreted to mean that the space-time variety is homogeneous, isotropic, and as will be seen below, has a characteristic "speed". The concept of a limiting speed can also be phrased by the statement that all finite points must be transformed into finite points. Note that these properties are topological properties independent from geometric scales. From the point of view of uniformity, the Lorentz equivalence class, and therefore the theory of special relativity, is in a certain sense a topological theory, not a geometrical theory.

Of course, when one looks at the non-uniformity of the stars in the night sky, one wonders why the "most uniform" Lorentz transformations have any utility at all in the real physical universe. The answer is, of course, that the Lorentz transformations of special relativity are *inadequate* to describe precisely the *global* features of the Cosmos (the presence of stars destroys the concept of global homogeneity, hence the Lorentz constraint of uniformity cannot be valid everywhere). In order to solve this paradox, the argument could be presented that the properties of homogeneity and isotropy and linearity associated with the Lorentz class are only *approximations* and these constraints are valid over only limited domains of space time, or for domains which are very dilute (called empty space or the vacuum). The argument would claim that although the domain of the universe is obviously not uniform and not strictly homogeneous on the scale of the galaxies and stars, it is very dilute and so on these scales, perhaps, it is possible to have faith in the "in approximation" status of the Lorentz equivalence class.

But philosophically speaking there is more to it than that. For in our local laboratories, the walls of the lab, and the boundaries of the experimental apparatus, certainly destroy any uniformity, and the "lab" is certainly not "dilute". Why, then, is there any significance to this Lorentz equivalence class of "most uniformity", when it seems impossible to construct practical local laboratory domains with these properties. It is astounding to think that the agreement between experiment and a theory based upon the Lorentz equivalent class should have any viability at all. But it does, and not for the usual reasons.

V. Fock recognized these difficulties with the theory of relativity, and offered a possible solution to the epistemologies paradox: the universe was to be considered as a domain in which there would be local modifications to the "most uniformity" principle, and the problem to be solved should be the one which insisted that at points "infinitely far" away from these interior non-Lorentzian regions (which was supposed to contain things like stars and galaxies) the space time domain should have (perhaps asymptotically) the "most uniform" properties of the Lorentz equivalence class. He proved that such problems led to a privileged class of coordinates, called harmonic coordinates, and would generated unique solutions. Such a space time might be described as "Asymptotically Lorentzian", where the elementary theory of special relativity would be described as "Globally or Everywhere Lorentzian"

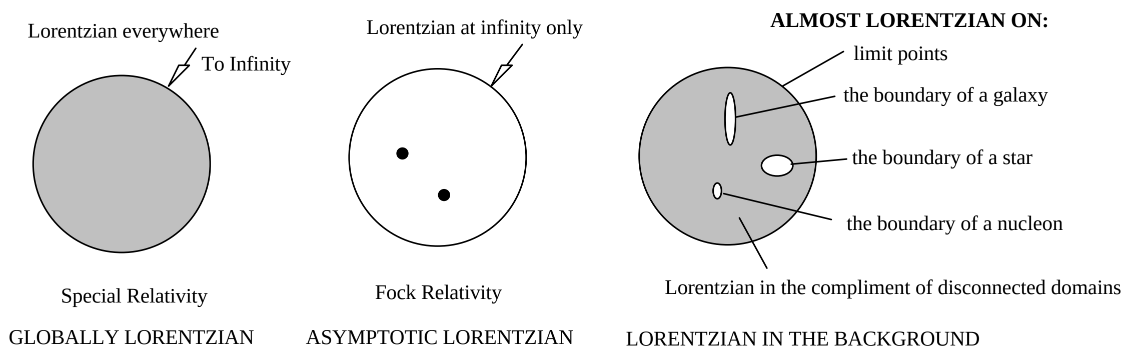
However, as mentioned above and shown below, and as noted by Fock, the fundamental significance of the Lorentz class of transformations is that this class leaves invariant those subsets of space-time upon which the solutions to the Maxwell equations of electromagnetism are NOT unique. He also noted that there were two transformation equivalence classes that had this property. The Lorentz group is the only linear group, and the 15 parameter projective (or Poincare) group of Moebius fractional linear transformations was the only non-linear group, that has the property of leaving the discontinuous subsets invariant. The Lorentz transformation was the only transformation that preserved the discontinuous subsets, **and** transformed finite points into finite points. It is this last condition that adds the concept of a limiting finite speed to the concepts of homogeneity and isotropy, to form the principle of "maximum uniformity". The fractional projective transformations have the property of homogeneity and isotropy, but do not necessarily map finite points into finite points. It is interesting to note, also, that from the point of view of differential geometry, the concept of maximum uniformity on a space of dimension 4 requires the same number of parameters (10) needed to construct an affine subgroup with a fixed point, a concept which is known to preserve "parallelism".

The point of departure in this article is to focus attention on the requirement that "signals must be signals to all observers". Hence, the equivalence class must be homogeneous and isotropic, but the transformations need not be linear. Such a requirement is the equivalent of the statement: "subsets of electromagnetic discontinuities must be invariants of the equivalence class of transformations. One observer in one frame of reference observes an electromagnetic discontinuity, and another observer will observe the same discontinuity if the frames of reference are restricted to be members of either the :Lorentz or the Poincare equivalence class.". Such domains which are homogeneous and isotropic, but not uniformly continuous, will be defined as "Almost Lorentzian" domains. Using this idea it is then conceivable that through measurement of electromagnetic signals, which indeed are defined to be electromagnetic discontinuities, the universe may be divided, or partitioned, at all scales, into disconnected sets or domains, with one non-Lorentzian solution on one side of the domain boundary and another Lorentzian solution on the other side of the boundary. On the limit points that form the boundaries of the disconnected non-Lorentzian domains, or the portions of boundaries (called cycles) of their connected - but not simply connected - compliment, the solutions are harmonic, and discontinuous, but not necessarily of maximum uniformity. The solutions are defined to be "Almost Lorentzian" on these limits sets that form the boundaries of the non-Lorentzian domains.

Attention can be focused either on the discrete disconnected components, or on their non-simply connected compliment, defined as the "background". The model yields the analogue of a Swiss cheese universe full of defect "holes" and "lines of self-intersection" where the defect domains are domains of non-Lorentzian topology. The "holes" or defects will be topological obstructions of dimension 1, 2, and 3 in the 4 dimensional space. The "values" of these obstructions will be related to period integrals over the portions of the boundary called cycles. Three species of period integrals can be related to the quantized concepts of flux, charge and spin [6]. In the interior of the compliment (defined as the "background") the solutions are harmonic and Lorentzian of maximum uniformity. Upon the limit points that form the boundaries and cycles of the "background" only isotropy and homogeneity is required, but the constraint of maximal uniformity is dropped. The solutions are harmonic but need not unique on the limit points. As mentioned above, these limit point domains are "Almost Lorentzian". In the interior of the defect domains, the topology need not be Lorentzian at all.

This proposal implies that *on the surfaces of discontinuities, the requirement for a finite propagation speed is not mandatory!* The preservation of the sets of non-uniqueness is a topological requirement that can admit non-linear transformations which can transform finite points to infinity. A physical interpretation of such processes would be that they can occur at non-finite speed. A raison d'etre for the collapse of the wave front at infinite speed is thereby established by such a proposal. The topological collapse of a singularity surface of electromagnetism discontinuities can take place with "infinite" speed, but the propagation of the set of discontinuities into the background can occur only with finite speed.

The shape and size of these limit sets, which form pieces of the boundary of the "background", are geometric concepts of secondary (but often interesting) consequence. The more important idea is that on such limit sets, discontinuities are permissible, and for the special relativity theory, the Lorentz transformations are the only group of linear transformations (and the projective transformations of Moebius, the only non-linear group of transformations) that preserve the properties of electromagnetic discontinuities. These ideas lead to the concept of the Almost Lorentzian universe:



THREE MODELS OF THE UNIVERSE

In the compliment of the disconnected non-Lorentzian parts, the space time domain is presumed to be of "maximum" uniformity; i,.e., Lorentzian. In this background domain the rules of Lorentzian special relativity are accurately true. On the limit sets, however, the concepts of homogeneity and isotropy are assumed to be invariants, such that the discontinuity subsets will be preserved, but the concept of a finite speed is not required. These special subsets of limit points are "almost Lorentzian". The have the required properties of homogeneity and isotropy necessary to preserve the singular sets, but do not require linearity. The concepts of uniformity and homogeneity are presumed to be satisfied globally over the exterior of the disconnected domains and their boundaries. But the concept maximum uniformity and the "Lorentz laws of special relativity" are satisfied only on the exterior regions. *The topological boundaries of electromagnetic discontinuities form the boundaries of a set of disconnected Lorentzian defects.*

Such a concept is reminiscent to certain more readily observable features of hydrodynamics, where large scale turbulent structures, such as a hurricane, or a cloud, apparently exist within a non-turbulent placid or streamline domain. The topology inside the hurricane is different from the topology of the exterior domain, and a topological boundary separates the two domains. Solutions to the equations of motion describing the fluid dynamics need not be the same on the boundary. Discontinuities can exist. The cloud is white, the sky is blue. What are the qualities and properties of these boundaries? What is the equivalence class such that all observers recognize the same boundary?

For electromagnetic theory, it is apparent from experiment that the theory of special relativity and the Lorentz equivalence class must have merit. But some of the most impressive results of the special theory of relativity come from the application to problems of collisions between atoms, nuclei, or elementary particles, moving at high relative speeds. Here the rules of Newtonian mechanics conserving Newtonian momentum and energy do not work accurately, but the rules deduced under the constraints of the Lorentz group can be compared to experiment with dramatic accuracy. However, the application of special relativity theory to such collision problems does not involve interior domains of the interacting nuclei (where homogeneity and isotropy cannot be strictly true), but instead the rules of the theory are applied to "boundary" regions "far" from the actual objects themselves and their domain of interaction. In these experiments, far is measured in terms of the "size" of the objects. The laws of special relativity are applied to observations made on a "boundary" envelope of the collision event, many orders of magnitude removed from the collision domain.

The key idea is that on boundaries separating different physical domains of perhaps different topology, the solutions to the laws of nature need not be unique, and it is the Lorentz equivalence and its extension to the class of fractional projective transformations that preserves this property of non-uniqueness from one observer to another.

Relativity, Geometry, and Topology

The concept of the boundary is a topological concept. It is important to realize from the outset that the ideas employed by the theory of special relativity are based upon geometrical and topological (global) concepts which are different from those employed by the conventional and more familiar geometry of Euclid. The topology of a euclidean space is a simply connected topology, and is without boundary. Euclidean space is presumed to be uniform and homogeneous, without holes or obstructions and singularities, and to extend to infinity. The topology of the universe in the presence of matter is not so simple. It is these different geometric and topological issues which generate many of the mis-interpretations associated with the theory. In order to convey the ideas incorporated in the theory of special relativity to new students, it has become fashionable to express certain "facts" about the theory in a dogmatic manner, employing catchy phrases that are easily remembered, and passed (often incorrectly) from one student generation to another.

For example, the theory of Special Relativity and the associated study of the Lorentz equivalence class is often given a high level of credence by a statement equivalent to the following:

"Maxwell's equations of electrodynamics are covariant in form with respect to Lorentz transformations, and are not covariant in form with respect to Galilean transformations".

The implication is that as every one knows the importance of electromagnetism, then the lack of "covariance" of Maxwell's equations with respect to the Galilean transformation, puts the Lorentz equivalence class on a pedestal. The concept of "Covariance" means that when the equations of a theory are transformed to a new functional representation in a new coordinate system, the equations have the same functional format. If in one reference system, the equations have only second order partial derivatives, then in the new frame of reference the equations must involve only second order partial derivatives. If the equations in the new frame of reference had both second and first order derivatives, then the equations would not be "Covariant" in form. The dogma of modern science assumes that the "laws" of science should be the same for all observers.

However, the above statement about the lack of Galilean covariance of Maxwell's equations is false. For in the notation of Maxwell (involving **E,B,D,H**), Maxwell's equations are *tensor* equations. The engineering format

$$\text{curl}\,\mathbf{E} + \partial\mathbf{B}/\partial t = 0 \quad \text{and} \quad \text{curl}\,\mathbf{H} - \partial\mathbf{D}/\partial t = 0 \tag{1}$$

for domains free of charge and currents, are tensorial statements (see A. Sommerfeld). The **E** and **B** fields are components of a second rank covariant tensor, and the **D** and **H** fields are components of a contravariant tensor density.

The beauty of a tensor statement is that it is naturally covariant in form with respect to **all** reference systems, related to one another by diffeomorphisms (continuous maps with differentiable inverses). But both the Lorentz transformation and the Galilean transformation are diffeomorphisms; hence, when the rules of tensor analysis are applied properly, Maxwell's equations are naturally covariant in form with respect to **both** the Lorentz and the Galilean transformation, as well as an infinity of other diffeomorphic coordinate transformations. QED. The mistake that is usually made in many text books is that **E** and **B** are transformed (correctly) as covariant tensor components, but **D** and **H** are not transformed as the components of a contravariant tensor density. Instead, based on the sophomoric and incorrect assumption that the constitutive relations

$$\mathbf{D} = \varepsilon\mathbf{E} \text{ and } \mathbf{B} = \mu\mathbf{H} \tag{2}$$

imply that **D** is that "same kind" of tensor as **E,** and **B** is the "same kind" of tensor as **H**, it is assumed that the components of **D** and **H** should be transformed as a covariant tensor, instead of a contravariant tensor density. Such an error of course leads to a lack of "covariance" relative to the Galilean (and all other) transformations, for the second set of Maxwell equations. The constitutive properties of matter are not deformation invariants, and hence are not true scalar quantities. The constitutive properties of matter may be interpreted as a set of geometrical constraints on the topological equations of Maxwell. As a matter of fact, the idea that Maxwell's theory of electromagnetism is a metric free theory - and therefore a theory describing topological and not geometrical concepts - seems to have started with the work of Van Dantzig about 1931.

The fact is that the above statement about the lack of general covariance of Maxwell's equations is false (Maxwell's equations are naturally covariant with respect to both the Lorentz transformation and the Galilean transformation, as well as all other diffeomorphisms if the transformations of the covariant and the contravariant parts are done properly) Yet many generations of physicists, and many textbooks, parrot the above dogmatic phrase from one generation to the next generation, without thinking, giving an incorrect motivation for the study of the Lorentz transformations of special relativity. The Lorentz transformations and therefore special relativity, indeed are important, but not for the dogmatic statement made above. Maxwell's electromagnetism goes far beyond the constraints of special relativity theory.

As mentioned above and discussed below, the importance of the Lorentz transformations, and their dominance in physical applications, is NOT due to the natural covariance of Maxwell's equations relative to such equivalence classes, but instead is related to the preservation, or invariance, of those solution subsets to Maxwell's equations (the solutions, not the equations) which are singular. The point of departure is to search for those solution subsets and domains for which the solutions to Maxwell's equations are *not unique*. It is these discontinuous subsets which represent the propagating discontinuities called electromagnetic "signals". In fact, following Fock, Luneberg, Klein and others, it is important to recognize that an electromagnetic wave is indeed a propagating discontinuity. To paraphrase Victor Fock:

> "The laws of propagation of light in empty space are thoroughly understood. They find their expression in the well-known equations of Maxwell ... (see equations 1 above with the constitutive assumptions 2 for empty space, where **E** and **B** are the vectors of electric and magnetic intensity, and **D** and **H** are the vectors of electric and magnetic excitation) ... However, we are not interested in the general case of light propagation, but only in the propagation of a signal advancing with maximum speed; i.e., the propagation of a wave front. Ahead of the front of the wave all components of the field vanish. Behind it some of them are different from zero. Therefore, some of the field components must be discontinuous at the front.
>
> On the other hand, given the field (a solution to the equations) on some surface moving in space, the derivatives of the field are, in general,

> determined by Maxwell's equations. Hence the value of the field at an infinitely near surface is also (uniquely) determined (by analytical continuation) and discontinuities are impossible. The only case when this is not so is when the form and the motion of the surface satisfies certain special conditions subject to which the values of the derivatives is not determined by the values of the field components themselves. Such a surface is called a characteristic surface, or briefly, a characteristic. Thus discontinuities of the field can occur only on a characteristic, but since there must certainly be discontinuities at a wave front (the signal), such a front is clearly a characteristic."

Fock then goes on "..to determine the equation of the characteristic for the system of Maxwell's equations" subject to the constitutive equations given by (2). Note that homogeneous system of Maxwell's equations involve 12 functions (3 components each of **E,D,B,H**) and 6 equations. The constitutive equations give 6 more equations, hence the system is homogeneously soluble when a constitutive format is assumed. Over four dimensions, Cramer's solubility criteria for homogeneous systems requires a certain determinantal condition be true, and leads to a polynomial characteristic equation of fourth degree as a necessary condition for solubility. In the simplest cases all four roots of this characteristic polynomial turn out to be of the same magnitude: these roots are related to the speeds of propagation of the characteristic, and in the simplest situations (the conventional Lorentz vacuum), the speeds are degenerately the same. This result leads to the often used statement that the speed of light is the same for all observers.

For geometrical constraints that admit a center of symmetry, this characteristic equation for the " characteristic surface" turns out to be a polynomial equation consisting of *two* quadratic forms with signature {+,-,-,-} or { -,-,-,+}:

$$\{\pm (\ )^2 \mp (\ )^2 \mp (\ )^2 \mp (\ )^2 \} = 0.$$

For a "surface" defined by the equation

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

the explicit form of the characteristic constraint is

$$\pm(\partial\phi/c\partial t)^2 \mp(\partial\phi/\partial x)^2 \mp (\partial\phi/\partial y)^2 \mp (\partial\phi/\partial z)^2 = 0. \qquad (3)$$

This equation is called the Eikonal equation.

From this perspective, the issue of physical importance is to develop that domain of logical intersection where by measurements made in terms of electromagnetic "signals" (or propagating discontinuities) by one observer, will agree with measurements made in terms of electromagnetic signals by a second observer. Both observers agree that the laws of physics are tensorial, so that when written and transformed properly they will have the same format. But now the comparative domain of different observers is that of the measured signals, or propagating discontinuities. These signals, or wave fronts, are solution subsets of the governing equations (Maxwell's equations) upon which the solutions are not unique.

It is no accident that such propagating discontinuities are called wave fronts. For example, consider the standard form of the wave equation:

$$\{\nabla^2\phi \;-\; (1/c^2)\partial^2\phi/\partial t^2\} = 0 \tag{4}$$

The partial differential equation called the wave equation admits characteristics (as do all hyperbolic PDE's including Maxwell's system of partial differential equations). It is remarkable but not accidental that the equation for the characteristics of the wave equation is the *same* Eikonal expression given above (3). What is even more remarkable about the eikonal solutions (functions that satisfy both (3) and (4)) is that then $F(\phi)$ is also a solution to the wave equation, where F is any twice differentiable function of ϕ, *linear or non-linear*. By substituting $F(\phi)$ into the wave equation, it is an easy exercise to obtain:

$$\nabla^2 F \;-\; (1/c^2)\partial^2 F/\partial t^2 \;=\; (\partial F/\partial\phi)\;\{\nabla^2\phi \;-\; (1/c^2)\partial^2\phi/\partial t^2\} \;+$$
$$(\partial^2 F/\partial\phi^2)\{+(\partial\phi/\partial x)^2 +(\partial\phi/\partial y)^2 +(\partial\phi/\partial z)^2 - (\partial\phi/c\partial t)^2\}$$

Hence if ϕ satisfies both the wave equation

$$\{\nabla^2\phi \;-\; (1/c^2)\partial^2\phi/\partial t^2\} = 0$$

and the "eikonal" equation

$$\{+(\partial\phi/\partial x)^2 +(\partial\phi/\partial y)^2 +(\partial\phi/\partial z)^2 - (\partial\phi/c\partial t)^2\} = 0,$$

even for ***non-linear*** functions, F, then $F(\phi)$ is a solution to the wave equation. Note that if the function F is an anti-symmetric function, or linear function, of the phase function, ϕ , then $F(\phi)$ is a solution to the wave equation if ϕ is a solution, independent from whether or not ϕ satisfies the eikonal equation. The chief feature of the characteristics is that $F(\phi)$ need not be anti-symmetric in ϕ, such that $(\partial^2F/\partial\phi^2) \neq 0$, and if ϕ satisfies the eikonal equation, then $F(\phi)$ is a still a wave function. It is those cases where $(\partial^2F/\partial\phi^2) \neq 0$ that generate diffraction patterns and the bright caustics or webs of light intensity seen in a swimming pool with a "wavy" surface. The classic example of an eikonal solution is the linear function

$$\phi = \pm i(\mathbf{k} \bullet \mathbf{r} \;\pm\; \omega\, t),$$

which satisfies the wave equation and the eikonal equation as long as

$$\mathbf{k} \bullet \mathbf{k} - (\omega/c)^2 = 0.$$

This last equation is called a dispersion relation (or characteristic polynomial), and constrains the values of the reciprocal periods in space and time, called the wave vector, **k**, and the angular frequency ω. In the electromagnetic case, the expression for the dispersion relation is a bit more complex, and depends upon the symmetries imposed by the constitutive equations. The functions $\phi(x,y,z,t)$ that satisfy both the wave equation and the eikonal equation are often called "phase" functions, and on the surfaces of zero phase, the solutions to the Wave Equation are not necessarily unique. Note that any linear function of a wave function is also a solution to a wave equation, but any non-linear function of a ***phase function*** is also a solution to a wave equation. There are three situations:

1. ϕ is a solution to both the Eikonal and the Wave equation.
   Then any $F(\phi)$ is a solution to the Wave equation.

2. ϕ is a solution to the wave equation but is not a solution
   to the Eikonal equation. Then any odd function $F(\phi)$
   in the sense that $(\partial^2 F/\partial\phi^2) = 0$ is a solution to the wave equation.

3. ϕ is a solution to Eikonal equation but is not a solution
   to the wave equation. Then any even function $F(\phi)$
   such that $(\partial F/\partial\phi) = 0$ is a solution to the wave equation.

(In the statistical mechanics of "particles" - which are disconnected objects with some form of "boundaries" - there are three types of statistics. Maxwell-Boltzmann, Fermi-Dirac, and Bose-Einstein. It would be a beautiful Ph.D. thesis to show the connection between the three forms of boundary types given above and the three forms of statistics)

The question arises as to what classes of transformations preserve the Eikonal equation of singularities (not the original Maxwell's equations but the singular solution sets). But the eikonal is a sum of squares with the signature {+,+,+, - } or {-,-,-,+ } and it has been shown that the only linear transformation group that preserves such an algebraic structure is the Lorentz group (Fock). There is also a *non-linear* transformation, the Moebius or Poincare projective transformation (of 15 parameters, not 10) that also preserves the signature, but these are the only two groups of transformations that will. In order for propagating discontinuities called wave signals in one frame of reference to be interpreted as propagating discontinuities called wave signals in another frame of reference, it follows that the two frames of reference must be connected by means of a Lorentz transformation. (Otherwise the Eikonal condition is not necessarily satisfied in both systems) These ideas form the foundation of special relativity, which is then the study of systems for which signals to one experimenter appear as signals to another. The significance of the Lorentz transformation is that is establishes those frames of reference in which each observer will agree to the detection of a signal, or propagating surface of electromagnetic discontinuity..

Another mis-interpretation of the special relativity theory is the statement that a physical rod undergoes a real euclidean decrease in length when it is moving at a constant speed v relative to another observer. An equivalent statement would be to say that a moving sphere undergoes a relativistic distortion into the shape of an ellipsoid, with the diameter contracted in the direction of motion. Such statements are false. James Terrel was the first to show that the appearance of a spherical object remains a spherical object when observed with electromagnetic signals. After some fifty years of continued mis-statements in text books and dogmatic lectures, Terrel's work met with several years of rejection by the physics community; finally, recognition came after V. Weiskopf presented Terrel's paper at a meeting of the APS. Shortly thereafter R. Penrose came out with a formal proof, based on the conformal qualities of the Poincare transformation, that a spherical shape is transformed into a spherical shape. The apparent contraction of a rod is due to a optical illusion of apparent rotation, and for the sphere, the diameter does not change, but part of the back surface of the sphere at rest rotates around and becomes visible by means of signals under the Lorentz transformation at high speeds.

Another mis-statement is that the speed of light is the same for all observers. The more correct statement is that for frames of reference equivalent to the vacuum by means of a Lorentz transformation, the speed of light (the propagation speed of the discontinuities) is the same. Such a result is due to the assumed form of the constitutive relations, which have certain geometric symmetries that cause the fourth order characteristic polynomial of the Maxwell system to have four equal roots. But such a statement is not meaningful in quartz, where waves with left-handed polarization propagate with the speeds different form those waves with right handed polarization.. In quartz, the constitutive relationships have more complicated symmetry relations than those associated with the vacuum. The fourth order characteristic polynomial for the Maxwell system now separates into two quadratic polynomials with *different* pairs of roots, for each mode of polarization. But it is indeed true that in quartz, as well as in vacuum, the inbound signals propagate with same speed as the outbound signals, as long as they have the same polarization. In fact this assumption that the speed of light in the outbound direction is the same as the speed of light in the inbound direction is crucial to Einstein's derivation of the special theory of relativity.

However a simple example shows that this reciprocity in direction is certainly violated at a discontinuous surface of glass and air. But it is even more strange that there are *exact* solutions to Maxwell's equations for which the characteristic speeds of propagation in the interior of a domain are such that the characteristic polynomial has four distinctly different roots. The speed of light in different directions and for different states of polarization are distinct in such systems. Such systems are associated with constitutive equations that do not provide a center of symmetry. Such domains can occur only in spaces of four dimensions or more, and involve topological domains and defects of the third rank called torsion cycles. The phenomenon can not occur in three dimensional space, for in the theory of finite groups, the first occurrence of irreducible subgroups occurs for groups of order (dimension) four. The exact solutions to Maxwell's are in the form of Quaternions, whose non-abelian structure can be adapted to the lack of a center of symmetry. The first experiments conducted to demonstrate this fact were accomplished at the University of Houston by V. Sanders [2] in 1977 who experimented with dual polarized ring lasers. The first exact solutions demonstrating the four speeds of wave front propagation were created by a another University of Houston student, A. Schultz [3], in 1979. A more detailed theory of singular surfaces was published by the present author in 1991 [4].

The moral of this story is that it is important to understand the limitations and philosophy behind special relativity, and not to accept dogmatic phrases without understanding their domain of applicability.

Victor Fock pointed out the need to study the field equations and their boundary conditions together. He championed the idea that the important systems to study were where space time was "uniform in the sense of Lorentz at infinity". The assumption led to a preferred system of coordinates, called by Fock, :"harmonic coordinates". This physical existence of any preferred system of coordinates was denied by Einstein. But Fock's successes and contributions are not to be denied.

In differential geometry, the idea of harmonic coordinates is linked directly to the theory of "minimal surfaces" (think soap films and their analogs in higher dimensions). Hence, Fock's constraint of looking at harmonic coordinates is related to the idea that the fabric of space time becomes a minimal surface at infinity, and in the interior, local imperfections lead to an understanding of mass and gravity. The point of departure presented herein, is to recognize that the surfaces of discontinuity that separate physical domains, one from another, are such harmonic sets. They are the equivalent to minimal surfaces, act as membranes of tangential discontinuities impervious to mass, and are limit sets upon which dissipation is minimal if not zero. For the electromagnetic case, such limit sets are presumed to be governed by the Eikonal expression, and hence form equivalence classes relative to the Lorentz transformation. The fundamental hypothesis formulated by this author is that the theory of relativity and the study of Lorentz equivalence classes is important because:

The persistent and non-diffusive limit sets of tangential discontinuities that form boundaries between different physical domains are related to one another by means of the Lorentz transformations, or the more general projective transformations, both of which preserve the eikonal criteria necessary for the construction of a set of tangential discontinuities.

This idea goes beyond that of Fock who demands only the "uniformity" of the Lorentz condition at infinity. In the statement above, the Almost Lorentz condition is assumed to be applicable on the cycles of a non-simply connected domain as well as to the boundary points at "infinity". The limit points are where discontinuities are permissible.

A more standard treatment of Special Relativity

A more standard treatment of special relativity, starts with the notion of euclidean differential distance, or arc length, in 3 - dimensions. The concept of metric or measure is introduced in conventional way using the Pythagorean theorem of euclidean geometry"

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2 .$$

For a moving object described by a velocity vector, $d\mathbf{R} - \mathbf{V}(x,y,z)\ dt = 0$, Picard's theorem states that over a limited domain there exists a (unique) solution trajectory, and any object starting anywhere on one of these solution trajectories continues down the trajectory. The solution certainly cannot extend beyond a point where all components of the Velocity vanish (the stagnation points), These solution trajectories are said to form a set of streamlines in 3 dimensional space. The differential distance admits of an easy interpretation in terms of the euclidean norm of the velocity.

$$(ds)^2 = (\mathbf{V} \bullet \mathbf{V})\ (dt)^2$$

However, if the functions that make up the components of the velocity vector, $\mathbf{V}$, are dependent upon time as well as space, then the flow will not necessarily generate a set of streamlines. In fact the streamlines in 3 dimensions occur only when $\mathbf{V}$ is factorizable: $\mathbf{V}(x,y,z,t) = \mathbf{v}(x,y,z)\ T(t)$. But it is always possible to consider the three dimensional vector field as an element of a 4 dimensional space with an arbitrary parametrization, τ. Then the assumption is that the four components of $\mathbf{U}(x,y,z,t) = \lambda(x,y,z,t)(\mathbf{V}(x,y,z,t)/c, \mathbf{1})$ are independent of the "new" parameter, τ. Then again using Picard's theorem, it is possible to generate over a limited domain a set of "world lines" which are the four dimensional equivalent to "streamlines" in 3 dimensions. The worldlines never intersect within the domain of Picard solutions. Hence there must exist a function, ρ, such that

$$\mathrm{div}(\rho\mathbf{V}) + \partial\rho/\partial t = 0,$$

over that domain. This equation is called the equation of continuity.

Now the problem is to define a differential distance function, this time measuring an arclength along the "world lines". One choice would be to extend the euclidean 3 space to a euclidean 4 space. The norm of the vector would then be defined as

$$(\mathbf{U} \bullet \mathbf{U}) = \lambda^2 \{+(\mathbf{V}/c \bullet \mathbf{V}/c) + 1\}.$$

Note that this euclidean norm is a sum of squares with signature {+,+,+,+}. It is not an invariant of the Lorentz transformations. Another possibility, would be to use the Eikonal expression with signature {+,+,+,-} to guide in the construction of a 4 dimensional norm:

$$(\mathbf{U} \quad \mathbf{U}) = \lambda^2 \{+(\mathbf{V}/c \bullet \mathbf{V}/c) - 1\}.$$

This is the construction that is used in the theory of Special Relativity, and is given stature by the extraordinary experimental interpretations and implications of the resulting theory. A four dimensional space with the signature {+,+,+,-}for its line element is called a Minkowski space. It follows that the distance function becomes

$$(ds)^2 = \lambda^2 ((\mathbf{V}/c \bullet \mathbf{V}/c) - 1)(dt)^2 \qquad (6)$$

It is conventional to normalize the square of the differential arclength to a constant value, by choosing the value, $\lambda^2 = -c^2$, and to redefine the parameter of "world" time, τ, such that $(ds)^2 = (cd\tau)^2$. These conventions lead to the algebraic result:

$$\gamma\, d\tau = dt, \quad \text{where } \gamma = 1/\sqrt{1-(v^2/c^2)}$$

For speeds v < c, the factor γ is always a number greater than 1. The equation is often interpreted in the following way: A world time interval, $d\tau$, measured by a fixed observer, must be dilated by the factor γ to agree with the time interval as measured by a moving observer, dt. The dogma is "Moving clocks run slow".

It should be noted that dt is always considered to be a perfect differential, but as $d\tau$ is defined by a square root operation, it may not be a *perfect* differential. There may not exist a unique function whose differential is equal to the increment symbolized by $d\tau$. It perhaps would be best to put a slash through the $d\tau$ symbol, as is done in thermodynamics, to signify an imperfect differential. However, the quadratic form,

$$-(cd\tau)^2 = (dx)^2 + (dy)^2 + (dz)^2 - (cdt)^2 \;,$$

although constructed from a imperfect differential, is an *invariant* of the Lorentz transformations, and is the same for all electromagnetic signal observers. The idea of an imperfect differential is related to whether of not the integral of the differential around a closed path integration yields a zero result. The imperfect differential of work and the imperfect differential of heat are classic physical examples where such integrations give none null results.

One of the most direct ways of dealing with relativistic kinematics is to start with the 4-vector of differential position, and form the derivative with respect to world time, τ.

$$\begin{aligned}\mathbf{U} = d\mathbf{R}/d\tau &= \{dx/d\tau, dy/d\tau, dz/d\tau, cdt/d\tau\}\\ &= \{dx/dt, dy/dt, dz/dt, c\}\, dt/d\tau\\ &= \{u, v, w, c\}/\sqrt{1-(v^2/c^2)}\;.\end{aligned}$$

If the quadratic form is constructed using the {+,+,+,-} signature, then

$$\mathbf{U} \quad \mathbf{U} = -c^2.$$

The four vector of velocity has a Minkowski norm equal to $- c^2$. This value is an invariant of all Lorentz transformations! Now follow the lead of Newton and define a four vector of momentum by multiplying the four velocity by parameter of constant mass

$$\mathbf{P} = m_o\mathbf{U} = m_o \{ \mathbf{V} , c \}/\sqrt{1-(v^2/c^2)} \ .$$

The Minkowski norm of the four momentum is an invariant of the Lorentz transformations, and equal to :

$$\mathbf{P} \quad \mathbf{P} = - \quad m_o\mathbf{U} \quad m_o\mathbf{U} = \{(m_o\mathbf{v} \quad m_o\mathbf{v}) - (m_o c)^2\}/\sqrt{1-(v^2/c^2)} \ . = - (m_o c)^2 .$$

Multiplying by c^2 leads to the Lorentz invariant square of the "rest energy", $m_o c^2$. Redefining the spatial components of the relativistic momentum as

$$\mathbf{p} = m_o\mathbf{v}/\sqrt{1-(v^2/c^2)}$$

and the "total energy" as

$$E = m_o c^2/\sqrt{1-(v^2/c^2)}$$

permits the Lorentz invariant expression for the norm of the four momentum vector to be written as

$$\mathbf{p} \bullet \mathbf{p}\, c^2 - E^2 = - (m_o c^2)^2$$

Rearranging terms leads to the famous Einstein result

$$\mathbf{p} \bullet \mathbf{p}\, c^2 + (m_o c^2)^2 = E^2.$$

The first term is interpreted as the square of the "kinetic" energy, the second term is interpreted as the square of the "rest mass" energy, and the term on the RHS is interpreted as the square of the "total energy". The total energy is velocity dependent, a fact which originated with the idea that time, t, is not absolute and admits a "dilation" for a moving observer. Often this foundation is masked by the interpretation that the mass, m, is not a constant, but is indeed velocity dependent:

$$m = m_o/\sqrt{1-(v^2/c^2)} \ .$$

Now all of this may seem somewhat contrived. But, assume the validity of the above definitions, and construct the equivalent of the billiard ball collision problem. Instead of conserving as invariant the three components of the linear momentum, $m_o\mathbf{v}$, establish the equivalent collision rules for a relativistic two body collision by assuming that the fundamental rule over space time is the conservation of four momentum. Before the collision, construct the total relativistic four momentum vector by adding the individual four momentum vectors of each individual particle in the initial state. Then perform the same construction on the four momentum vectors of the individual particles of the final state. By the rule of conservation of relativistic momentum, the two vectors must be the same. This conservation law of relativistic mechanics puts severe constraints on the angular distribution and speeds of the particles after collision, similar to the billiard ball constraints of classical mechanics. But now, the momentum of any particle tends to infinity as the speed of the particle approaches the limiting speed c (the speed of light), a phenomena not within the domain of non-relativistic mechanics.

It is the almost uncanny accuracy by which these methods of relativistic mechanics agree with the angular distributions found in experiments of high speed collisions between electrons, nuclei and photons that establish the validity of the special theory of relativity.

The same techniques can be used to study the concepts of force:

The four vector of relativistic force is given by the expression:

$$\mathbf{F} = d\mathbf{P}/d\tau$$

and the relativistic power is given by the expression

Power is **F** **U.**

Other important invariants of the Lorentz transformation are the Maxwell field expressions defined as the first and second Poincare invariants. The are equal to the difference between the squares of the electric and magnetic fields, and the "dot" product (or the orthogonality) of **E** and **B**.

Poincare invariant I $\{\mathbf{E} \bullet \mathbf{E} - c^2\,\mathbf{B} \bullet \mathbf{B}\}$

Poincare invariant II $\{\mathbf{E} \bullet \mathbf{B}\}$

If the second Poincare invariant vanishes, and the first is greater than zero, it always possible to find a Lorentz transformation such in the new frame of reference the Magnetic field , **B**, vanishes. Such domains are called electric like, and the system looks a collection of electrostatic charges.

If the second Poincare invariant vanishes and the first is less than zero, than it is possible to find a frame of reference where the **E** field vanishes. Such a situation would correspond to the magnetic fields produced by an array of constant currents.

If the second Poincare invariant does NOT vanish, then the solutions to the problem are irreducibly four dimensional (Quaternionic).

[1] V. Fock , "Space Time and Gravitation" , Macmillan, N.Y. (1964)
[2] V. E. Sanders and R. M. Kiehn IEEE J Quant Elec QE-13 739 (1977) "Dual Polarized Ring Lasers"
[3] A. K. Schultz, R. M. Kiehn, E. J. Post and J. B. Roberds Physics Letters, **74A** 384 (1979) " Lifting of the Four-fold EM Degeneracy and PT Symmetry"
[4] R. M. Kiehn "Parity and Time reversal Symmetry breaking, Singular solutions, and Fresnel surfaces, Phys Rev A, **43** 5665 (1991)
[5] R. M. Kiehn, "Topological defects, coherent structures and turbulence in terms of Cartan's theory of differential systems" in "Developments in Theoretical and Applied Mechanics" XVI, B.N.Antar, R.Engels, A.A.Prinaris, T.H. Moulden, Editors, (UTSI Tullahoma, TN p.III.IV.02, 1992)
[6] R. M. Kiehn, "Periods on manifolds, Quantization, and Gauge" , J. Math Phys.**18**, 614 (1977)