

ACT II. SCENE 1: SPACE, TIME & SPACE-TIME

T.J. BOGDAN

September 16, 2018

1. Introduction

This Scene is somewhat more mathematical and much more philosophical in nature than the previous five. It is predicated upon our assertion that it is *always* best to work out the microphysics in the rest, or comoving frame, of the fluid *and* the experimentally derived knowledge that observers in motion relative to one another agree on some things but have different opinions on, and measurements of, other things. Therefore it is essential that we know how various quantities measured in the comoving frame *transform* into the laboratory frame where we are solving our equations of RMHD.

If we get these frame or coordinate system transformations wrong, then, at best, we end up solving the wrong problem, correctly. A pretty useless endeavor.

The required transformations are, in fact, the *Lorentz Transformations* of special relativity. This Scene attempts to explain and motivate why in fact these are the correct and indeed the only transformations which relate quantities in the comoving and laboratory frames. To do so, it is necessary to develop some background in the mathematics of groups, algebraic fields, and vector spaces. With these tools and their logical consequences, the Lorentz Transformations emerge in a very straightforward and in some sense inevitable fashion.

Therefore, the reader who is quite content to accept that the Lorentz Transformations *are* the de facto means to connect physical quantities measured in the comoving and laboratory frames can simply pass over this Scene and move on to the next without serious repercussions.

2. Vector Spaces

The three space dimensions that we live in, up-down, north-south, east-west, say, is a familiar example of a *three-dimensional Euclidean space*. In this section we set about answering the question of what, precisely, is it about our three-dimensional world that makes it a Euclidean space (as opposed to some other space).

Our spatial surroundings is an example of a mathematical structure called a *vector space*. The *elements*, or *vectors*, of the vector space are simply all the possible locations, or points, in this space. In keeping with our notation, we will denote individual points by boldface Latin letters, say, \mathbf{x} , \mathbf{y} and so forth, but we need not insist that $\mathbf{x} = (x_1, x_2, x_3)$. This is simply a particular concrete *representation* of \mathbf{x} in some coordinate system derived from a specific *basis*. We'll get to these concepts shortly. But, in fact, vector spaces can be very flexible and quite abstract in terms of the things they represent and describe. A vector space is a set \mathcal{V} of elements, \mathbf{x} , \mathbf{y} , \mathbf{z} , etc, that comes equipped with a single binary operation, called *vector addition*, denoted by $+$, that associates with any pair of vectors \mathbf{x} and \mathbf{y} a third vector $\mathbf{z} = \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ which is also an element of the vector space \mathcal{V} .

Indeed, mathematically speaking, for any set \mathcal{S} , with elements x, y, z , etc., there are properties, or laws, that one would like any binary operation—denoted here generically by \circ —to possess. In order of increasing sophistication, we can list the following familiar laws:

$$I. \text{ Closure Law : } \quad \forall x, y \in \mathcal{S}, \quad x \circ y \in \mathcal{S} ,$$

$$II. \text{ Associative Law : } \quad \forall x, y, z \in \mathcal{S}, \quad x \circ (y \circ z) = (x \circ y) \circ z ,$$

$$III. \text{ Identity : } \quad \exists e_{\circ} \in \mathcal{S} \mid, \forall x \in \mathcal{S}, \quad e_{\circ} \circ x = x \circ e_{\circ} = x ,$$

$$IV. \text{ Inverse : } \quad \forall x \in \mathcal{S}, \exists x_{\circ} \in \mathcal{S} \mid \quad x \circ x_{\circ} = x_{\circ} \circ x = e_{\circ} .$$

$$V. \text{ Commutative Law : } \quad \forall x, y \in \mathcal{S}, \quad x \circ y = y \circ x .$$

Read the symbol \forall as “for every”, \exists as “there exists”, \in as “contained in”, and \mid as “such that”. A set which is endowed with a single binary operation satisfying the first four of these laws is called a *group*. And if the fifth law is also valid, it is called a *commutative*, or an *Abelian* group. You can convince yourself that the set of all real numbers, \mathbb{R} , is a group under both addition and multiplication separately (*provided* you set 0 aside when considering multiplication only), but that the set of all integers, \mathbb{Z} , is a group under addition only but not multiplication.

A vector space is therefore a group with respect to vector addition. This is a recurring theme in mathematics, viz., many mathematical structures can be viewed as simultaneously possessing within them other, generally less sophisticated mathematical structures. Vector addition is a *commutative* binary operation because the order in which the vectors are added does not matter. Vector addition is also *associative*, in the sense that $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$. There is a special unique element of the vector space, denoted by $\mathbf{0}$, and called the *identity vector*, with the property that $\mathbf{0} + \mathbf{x} = \mathbf{x}$, for every vector \mathbf{x} . The last property of vector addition is: for every $\mathbf{x} \neq \mathbf{0}$ there is a unique vector, which we will suggestively write as $\mathbf{y} = -\mathbf{x}$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$, called the *inverse* of \mathbf{x} . Obviously, \mathbf{x} is the inverse of $-\mathbf{x}$. These properties satisfied by vector addition are simply another way of saying that a vector space is, or has the structure of, an *Abelian* or *commutative group*.

A vector space, however, is richer in its algebraic structure than simply being an Abelian group. Every vector space comes equipped with a distinct set of auxiliary objects, called an *algebraic field*, of *scalars*. Here, scalars will be represented by Greek letters, α, β , and so forth to distinguish them from the elements of the vector space.

A *field* is an algebraic structure, like a group, but one which has *two* binary operations. A set \mathcal{S} , with elements x, y, z , etc., and two binary operations, denoted generically by \circ and \square , is an algebraic field, if \mathcal{S} is an Abelian group with respect to \circ , and \mathcal{S} with the identity e_{\circ} omitted, is also an Abelian group with respect to \square , and, moreover

$$VI. \text{ Distributive Law : } \quad \forall x, y, z \in \mathcal{S}, \quad x \square (y \circ z) = (x \square y) \circ (x \square z) ,$$

VII. *Cancellation Law* : $\forall x, y \in \mathcal{S}, x \boxdot y = e_{\circ} \iff x = e_{\circ} \wedge y = e_{\circ}$.

Notice that $e_{\boxdot} \neq e_{\circ}$, so that the two identity elements are distinct.

The real numbers—with the associations $+$ $=$ \circ , \times $=$ \boxdot , $e_{\circ} = 0$, and $e_{\boxdot} = 1$ —are an algebraic field, and they are usually the scalar field of choice for most vector spaces, and especially for Euclidean spaces. For some set of vectors, say \mathcal{V} , to be a vector space over the field of scalars \mathbb{F} , the following must be true

$$\text{for all } \mathbf{x}, \mathbf{y} \in \mathcal{V} \text{ and } \alpha, \beta \in \mathbb{F}, \alpha\mathbf{x} + \beta\mathbf{y} = \beta\mathbf{y} + \alpha\mathbf{x} \in \mathcal{V} .$$

That is, not only can we add vectors together, but we can also multiply them by scalars from our field \mathbb{F} and the outcome of these operations has to be another vector in our vector space. This ability to “stretch” a vector out, or turn it around by multiplying it by a positive or negative scalar is the essential aspect of a vector space that sets it apart from other algebraic structures.

The identity vector $\mathbf{0} \in \mathcal{V}$ must not be confused with the scalar 0 that lives in \mathbb{F} , however, it certainly can be created by using 0, as in

$$\mathbf{0} = 0\mathbf{x} , \text{ for all } \mathbf{x} \in \mathcal{V} .$$

Likewise, for every $\mathbf{x} \in \mathcal{V}$, the unique *additive inverse* $\mathbf{y} = -\mathbf{x} = -1\mathbf{x}$, in terms of scalar multiplication by the additive inverse of the multiplicative identity from the field \mathbb{F} . The final property necessary for the definition of a vector space is the distributive law

$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} , \quad (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x} .$$

If the points around us in our three-dimensional world are the vectors, and the real numbers are the scalars, and the usual addition of vectors is the binary operation of addition, then our three-dimensional space is clearly a vector space.

A *basis* for a vector space, $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, is any *maximal* set of elements selected from \mathcal{V} such that the *only* solution to the equation

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \dots + \alpha_n\mathbf{e}_n = \mathbf{0},$$

with the α_i selected from \mathbb{F} , is $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, and, if any other vector \mathbf{e}_{n+1} is appended to \mathcal{B} then this equation has solutions for some $\alpha_i \neq 0 \in \mathbb{F}$. The integer, n , is the *dimension* of the space. Every basis of an n -dimensional vector space must have n elements. It follows that every vector $\mathbf{b} \in \mathcal{V}$ has a unique representation in terms of n scalars in each basis, \mathcal{B} :

$$\mathbf{b} = \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \dots + \beta_n\mathbf{e}_n ,$$

and the scalars $(\beta_1, \beta_2, \dots, \beta_n)$ are the *coordinates* of \mathbf{b} reckoned in this particular basis.

For every basis, \mathcal{B} , we can associate with each element of $\mathbf{b} \in \mathcal{V}$ a unique (ordered) set of n -scalars from \mathbb{F} , and so every n -dimensional vector space over \mathbb{F} is equivalent to the cartesian product of $\mathbb{F} \otimes \mathbb{F} \otimes \dots \equiv \mathbb{F}^n$. Thus, we often designate

our three-dimensional Euclidean Space as \mathbb{R}^3 , and as a preferred (selected from the large number of) basis we take the \mathbf{e}_i to be (unit) vectors in the three orthogonal (Cartesian) directions (up, east, north). However, it is worth noting that *any* basis is as good as any other, and there are infinitely many to choose from. The concept of many independent bases is a fundamental aspect of vector spaces that distinguishes them from other algebraic structures like groups and fields.

To make a vector space, \mathcal{V} , a *Euclidean vector space*, it is necessary for us to take $\mathbb{F} = \mathbb{R}$, and to create some additional geometric structure by defining an *inner product*, or equivalently, a *bilinear functional* from $\mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$. We'll use a dot between the two boldface vectors to denote this inner product. We require that the inner product satisfy:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$$

$$(\alpha \mathbf{x} + \beta \mathbf{y}) \cdot \mathbf{z} = \alpha \mathbf{x} \cdot \mathbf{z} + \beta \mathbf{y} \cdot \mathbf{z}$$

$$\mathbf{x} \cdot \mathbf{x} \geq 0, \quad \mathbf{x} \cdot \mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0},$$

for all $\alpha, \beta \in \mathbb{R}$ and all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$.

A vector space equipped with an inner product, and indeed there can, like bases, be many inner products to choose from, is called an *inner product space*. An inner product space is necessarily a *normed space*, because the inner product can be used to define a *norm*, which is a mapping from $\mathcal{V} \rightarrow \mathbb{R}^+$ according to

$$|\mathbf{x}| \equiv \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

Like inner products, there are many additional choices for norms besides this one. Also, it is possible to have vector spaces that are normed but which do not have an inner product. A norm must satisfy three conditions,

$$|\mathbf{x}| = 0 \iff \mathbf{x} = \mathbf{0},$$

$$|\alpha \mathbf{x}| = |\alpha| |\mathbf{x}|,$$

and the triangle inequality

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|.$$

A Euclidean vector space is therefore a normed inner product space where the norm is defined in terms of the inner product as above. A norm constructed from an inner product is a *Euclidean Norm*, and this is precisely what makes our three-dimensional world Euclidean. The Cauchy-Schwarz inequality follows directly from the definition of our norm and inner product,

$$|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}| |\mathbf{y}|$$

This enables us to define an *angle* θ between two vectors \mathbf{x} and \mathbf{y} by

$$\cos \theta \equiv \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}| |\mathbf{y}|}.$$

We say two vectors are *orthogonal* if their inner product is precisely 0. A collection of vectors is said to be *orthonormal* if they are all mutually orthogonal and have a norm equal to 1. Any basis can be converted into an orthonormal basis in any number of ways. Orthonormal bases are particularly useful because they give rise to orthogonal coordinate systems for vector spaces equipped with an inner product.

A normed space is necessarily also a *metric space*, because the norm can be used to define a measure of *distance*, $d(\mathbf{x}, \mathbf{y}) \equiv |\mathbf{x} - \mathbf{y}|$, between any two vectors in the vector space. The distance between two vectors, is a mapping from $\mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$, which satisfies:

$$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \geq 0 ,$$

$$d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y} ,$$

and

$$d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) .$$

With the concepts of norms, distances and angles, we now have rudiments of geometry and geometrical constructions in place as well as the ability to define topologies and topological concepts essential for doing calculus and analysis.

3. Euclidean Geometry

The mathematician Felix Klein felt that *geometry* was essentially the idea of what sorts of things are invariant, or preserved, under a specified set of linear symmetry transformations from a given space to itself. The type of space you have is specified once you tell me what all the linear transformations—or symmetries—are, and what are the aspects of the space that these transformations preserve.

A *linear transformation* from an n -dimensional vector space $\mathcal{V}_n(\mathbb{R})$ to another m -dimensional vector space $\mathcal{U}_m(\mathbb{R})$, say A, B, C , etc., which we can write as

$$A : \mathcal{V}_n(\mathbb{R}) \rightarrow \mathcal{U}(\mathbb{R})$$

or

$$A : \mathbf{x} \mapsto \mathbf{y} \text{ or } A\mathbf{x} = \mathbf{y}, \text{ for } \mathbf{x} \in \mathcal{V}_n(\mathbb{R}) \text{ and } \mathbf{y} \in \mathcal{U}_m(\mathbb{R}),$$

satisfy the property

$$A(\alpha\mathbf{x} + \beta\mathbf{z}) = \alpha A\mathbf{x} + \beta A\mathbf{z} ,$$

for all \mathbf{x} and \mathbf{z} in \mathcal{V} .

So first, it is essential that the two vector space share the same auxiliary set of scalars, here \mathbb{R} , but as usual, any field \mathbb{F} will do. Second, remember that both $A\mathbf{x}$ and $A\mathbf{z}$ are two vectors that live in the vector space \mathcal{U} . So this definition is a statement about what must always be true in the image vector space.

Of the many linear transformations available to us, the ones from a vector space \mathcal{V} to itself are of particular interest. And of these, the subset of all linear transformations which are *one-to-one*,

$$A\mathbf{x} = A\mathbf{y} \iff \mathbf{x} = \mathbf{y} ,$$

and *onto*

$$\forall \mathbf{x}, \exists \mathbf{y} \text{ such that } \mathbf{x} = A\mathbf{y} ,$$

form a group, $GL_n[\mathbb{F}]$, called the *general linear group* of dimension n over the vector space $\mathcal{V}_n(\mathbb{F})$. Since, as we have demonstrated, $\mathcal{V}_n(\mathbb{F})$ and \mathbb{F}^n are connected by an isomorphism, the group applies to all $\mathcal{V}_n(\mathbb{F})$ and the reference to any particular $\mathcal{V}_n(\mathbb{F})$ can be omitted.

The binary operation for the group $GL_n[\mathbb{F}]$, whose elements we will continue to denote by italicized upper case Latin letters, A, B, C , etc., is *composition*, or successive application of linear transformations:

$$C\mathbf{x} = B(A\mathbf{x}) = (AB)\mathbf{x} .$$

The unique identity linear transformation is

$$I\mathbf{x} = \mathbf{x}$$

and the one-to-one and onto properties imply that if

$$\mathbf{x} = A\mathbf{y}$$

then there is another unique element in $GL_n[\mathbb{F}]$, call it A^{-1} such that

$$\mathbf{y} = A^{-1}\mathbf{x} .$$

Which completes our demonstration that $GL_n[\mathbb{F}]$ is in fact a group.

Although

$$AA^{-1} = A^{-1}A , \text{ and } AI = IA$$

for all A , the General Linear group is *not* commutative. Finally, we can deduce that for all $A \in GL_n[\mathbb{F}]$,

$$A\mathbf{0} = \mathbf{0}$$

[hint: use the fact that $\mathbf{0} = 0\mathbf{x}$ for any \mathbf{x}].

If \mathcal{B} is any basis for $\mathcal{V}_n(\mathbb{F})$, then we can associate with each $A \in GL_n[\mathbb{F}]$ a unique set (or array) of scalars

$$\{A_{ij}\}_{i,j=1}^n$$

drawn from \mathbb{F} such that, if

$$\mathbf{b} = \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \cdots \beta_n\mathbf{e}_n \in \mathcal{V}(\mathbb{F}),$$

then

$$\mathbf{g} = A\mathbf{b}$$

implies

$$\mathbf{g} = \gamma_1\mathbf{e}_1 + \gamma_2\mathbf{e}_2 + \cdots \gamma_n\mathbf{e}_n \in \mathcal{V}(\mathbb{F}),$$

where

$$\gamma_i = A_{ij}\beta_j \text{ OR } \gamma_i = A_{ji}\beta_j .$$

Different authors will take one or the other of these two depending upon whether they want to think of the A_{ji} as elements of a *matrix* with j indexing rows and i indexing columns—corresponding to the second of the two choices above. We’ll go with the second choice to remain consistent with matrix theory.

But, in either case, every basis \mathcal{B} generates a unique *representation* of A in terms of a square array—or “matrix”—of scalars drawn from \mathbb{F} . Because A is one-to-one and onto, the determinant of this matrix $\det[A] \neq 0$. Therefore the matrix array is invertible and the inverse is the representation of A^{-1} in this basis! In Act I Scene 1, we provided an explicit formula for the determinant of A when $n = 3$.

The *Special Linear Group* $\text{SL}_n[\mathbb{F}]$ is a subset, and indeed a subgroup (meaning to say it can stand along on its own as a group), of $\text{GL}_n[\mathbb{F}]$ with the property that in any basis, $A \in \text{SL}_n[\mathbb{F}] \iff \det[A] = 1$. These linear transformations preserve the volume and orientation of any parallelepiped formed by three non-colinear vectors \mathbf{x} , \mathbf{y} and \mathbf{z} .

They do not, however, preserve the *angles* between these vectors or *distances*. The subset (and again, subgroup) of $\text{SL}_n[\mathbb{F}]$ which achieves this desired result is the *Special Orthogonal Group* $\text{SO}_n[\mathbb{F}]$, with the additional restriction that the transpose of the matrix A , written A^T is also the inverse of A . That is if A is associated with the array of scalars $\{A_{ij}\}$ then A^{-1} is associated with the array $\{A_{ji}\}$.

$\text{SO}_n(\mathbb{R})$ is also known as the *n-dimensional rotation group*, because it describes rigid body rotations about the origin $\mathbf{0}$. It’s larger sibling, the *Orthogonal Group* $\text{O}_n[\mathbb{R}]$ relaxes the requirement $\det[A] = 1$ to $\det[A] = \pm 1$. The orthogonal group continues to preserve angles, and distances, but not orientation, as it includes reflections through various 2-dimensional planes that pass through the origin.

What makes our space Euclidean, of course, is the notion of angle, distance and orientation, which everyone, no matter how they set up their local coordinate system, should agree upon. And so returning to Klein’s program, the Special Orthogonal Group $\text{SO}_3[\mathbb{R}]$ of rotations in three dimensions is (part of!) the symmetry group of Euclidean three-space—it is the largest subset of all the general linear transformations from our Euclidean three-space to itself, which preserve angles, distances and orientation. Stated another way, individuals can orient their coordinate axes, or choose their bases, however they wish, and they will all agree on things like the sum of the angles of a triangle adding up to 180 degrees, parallel lines never meeting, and so on. This is a very useful and powerful property, because we do not want to be slaved to using someone else’s coordinate system and axes to carry out our investigations.

We said “part of” above, because despite its immense size, all the linear transformations of $\text{GL}_n[\mathbb{R}]$ leave the origin $\mathbf{0}$ fixed. Our space is not just isotropic—in the sense that we can orient our axes any way we like—but it is also homogeneous in that it extends indefinitely in all directions and so there really is no privileged spot that all can agree deserves to be called the origin.

An *affine* transformation, such as,

$$\varphi_{\mathbf{y}} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{y}$$

for any fixed $\mathbf{y} \in \mathcal{V} = \mathbb{R}^n$ also preserves distances, angles and orientation, but simply shifts the $\mathbf{0}$ by an amount \mathbf{y} . The collection of all $\phi_{\mathbf{y}}$, if we call this $T_n[\mathbb{R}] = \mathbb{R}^n$ the n -dimensional group of affine translations, form a group under the operation of composition of translations. The identity affine transformation is

$$\varphi_{\mathbf{0}} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{0} = \mathbf{x} = I\mathbf{x}$$

and has the same effect on \mathbf{x} (specifically, *no effect*) as the identity element $I \in SO_n[\mathbb{R}]$. But otherwise these two groups $SO_n[\mathbb{R}]$ and $T_n[\mathbb{R}]$ have no other elements in common.

To construct the complete symmetry group of n -dimensional Euclidean Space, the *Euclidean Group* E_n , we simply “join” these two groups together:

$$E_n = SO_n[\mathbb{R}] \bigcup T_n[\mathbb{R}] .$$

Voilà!

4. Galilean Space-Time

Time, of course, adds another dimension to our experience. And so we must add an orthogonal dimension to our Euclidean three-space, to create a four-dimensional space-time. In so doing, we do not want to destroy or adversely impact the Euclidean geometry of three-space which was so carefully constructed in the previous section. This suggests that we will want to seek some additional stand-alone group(s) that we can “join” to the rotations and translations. The 4-dimensional Euclidean Group, E_4 , by the way, cannot be the symmetry group of Galilean Space-Time! (Why?)

So now we confine our attention to three spatial Euclidean dimensions, and one time dimension. Our vector space for the Galilean space-time can be written as the direct Cartesian product $\mathbb{R}^3 \otimes \mathbb{R}$. With time come the notions of motion, velocity, acceleration, jerk and so forth. In addition to coordinate axes and origins, we will all have to work with “clocks” to measure time.

Newton’s formulation of dynamics is centered on the concept of *acceleration*. Remember the mantra: objects remain in motion unless acted upon by a force, then they are subject to acceleration, etc. And so acceleration needs to be something that is left invariant by the symmetry group that describes our Galilean space-time. Distance already is an invariant because of embedded 3-dimensional Euclidean geometry, and so this implies that time must be an invariant as well. We can choose the origin of time however we like, so, we can immediately add in the group $T_1[\mathbb{R}]$ of affine time transformations

$$\varphi_{\tau} : t \mapsto t + \tau ,$$

for $t, \tau \in \mathbb{R}$ by incrementing $T_3[\mathbb{R}] \rightarrow T_4[\mathbb{R}]$ or joining through a union, take your pick.

Velocities, however, need not be invariant under the transformation group. Indeed, in the comoving frame, by definition, the fluid is at rest, but in the laboratory frame the same fluid appears to be in motion. Consider, therefore, the linear transformation

$$\psi_{\mathbf{u}} : (\mathbf{x}, t) \mapsto (\mathbf{x} - \mathbf{u}t, t)$$

where $\mathbf{u} \in \mathbb{R}^3$. This linear transformation from $\mathbb{R}^3 \otimes \mathbb{R}$ to itself preserves distances, angles, orientation, and *accelerations*. The set of all $\psi_{\mathbf{u}}$ also form a group. The inverse of $\psi_{\mathbf{u}}$ is $\psi_{-\mathbf{u}}$, and the identity is

$$\psi_{\mathbf{0}} : (\mathbf{x}, t) \mapsto (\mathbf{x} - \mathbf{0}t, t) = (\mathbf{x}, t) .$$

Notice that this is the same *identity* as each of the following

$$I : (\mathbf{x}, t) \mapsto (I\mathbf{x}, t) = (\mathbf{x}, t) ,$$

$$\varphi_{\mathbf{0}} : (\mathbf{x}, t) \mapsto (\mathbf{x} + \mathbf{0}, t) = (\mathbf{x}, t) ,$$

$$\varphi_0 : (\mathbf{x}, t) \mapsto (\mathbf{x}, t + 0) = (\mathbf{x}, t) .$$

The set of all $\psi_{\mathbf{u}}$ is sometimes referred to as the *boost* group. We'll give it the label $\text{GB}_3[\mathbb{R}]$ which stands for Galilean Boosts, in three directions—this is not a standard notation but it serves our purposes here.

Therefore the *full* Galilean Group of symmetry transformations on a 3+1 dimensional space-time is

$$G = \text{SO}_3[\mathbb{R}] \bigcup \text{T}_3[\mathbb{R}] \bigcup \text{GB}_3[\mathbb{R}] \bigcup \text{T}_1[\mathbb{R}] ,$$

where we have added an additional copy of the translation group to account for the fact that time is homogeneous and we can set the origin of time arbitrarily as we did for the origin of our three-dimensional Euclidean space. The Galilean Group is the disjoint union of four subgroups: rotations in 3-space, translations in 3-space, acceleration preserving boosts in 3(+1) space-time, and translations in time.

The G group is called a *continuous* group, or a *Lie Group*, because it takes (count 'em) ten continuous real variables to index its vast (uncountable in fact) number of elements: 4 translational parameters, 3 boost parameters, and 3 rotation angles. Because the subgroup $\text{SO}_3[\mathbb{R}]$ is not commutative, the overall Galilean group is not commutative either, even though the other subgroups are commutative within themselves.

The final step in the Klein program is the assertion, which we leave unproved, that any two *viable* sets of space-time coordinates (\mathbf{x}, t) and (\mathbf{x}', t') , say, for doing Newtonian physics in this 3+1 space-time are connected by a unique element (and its inverse) from the group G .

Nice!

5. Minkowski Space-Time

Maxwell's Equations, unlike Newton's laws of motion and gravity, are *not* invariant under the action of the Galilean Group. The simplest way to see this is to recall from Scene 3 that the wave equations for the vector and scalar potentials require that light propagate at the velocity c . This is true in *any* set of viable spacetime coordinates one cares to employ. In otherwords, light travels at the same speed c in every viable set of spacetime coordinates! To belabor the point to perhaps tedium suppose that instead of working in the unprimed laboratory frame of Scene 3, I choose to ride along at a uniform velocity on an Air Canada flight from Denver to Montreal and use light to read a magazine on board the plane. I don't use your laboratory coordinates in Montreal (unprimed) but my own (\mathbf{x}', t') . This is fine since there is an element of the Galilean Group of the previous section that tells us how to translate my coordinates to yours (being the unprimed ones). My Maxwell Equations are:

$$\begin{aligned}\nabla' \cdot \mathbf{E}' &= 4\pi\delta' , & c\nabla' \times \mathbf{E}' &= -\frac{\partial \mathbf{B}'}{\partial t'} , \\ \nabla' \cdot \mathbf{B}' &= 0 , & c\nabla' \times \mathbf{B}' &= 4\pi\mathbf{J}' + \frac{\partial \mathbf{E}'}{\partial t'} ,\end{aligned}$$

and I follow through all the mathematical developments of Scene 3 with primes on all my quantities (not c however!) and I end up with a wave equation for photons

$$\begin{aligned}\left(\nabla'^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t'^2}\right)\phi' &= -4\pi\delta'(\mathbf{x}', t') , \\ \left(\nabla'^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t'^2}\right)\mathbf{A}' &= -\frac{4\pi}{c}\mathbf{J}'(\mathbf{x}', t') ,\end{aligned}$$

which travel at c in my frame of reference. But this can't be! Because velocities are *not* invariant under Galilean boosts and you, in Montreal, have written out all these equations without primes to describe the light in my airplane cabin and you know the photons are travelling at c in *your* frame of reference, not mine.

The resolution of this paradox is that Maxwell's Equations as they are written above, are *not* invariant in form under the action of the Galilean Group. If we replace them by

$$\begin{aligned}\nabla' \cdot \mathbf{E}' &= 4\pi\delta' , & c\nabla' \times \mathbf{E}' &= -\frac{\partial \mathbf{B}'}{\partial t'} , \\ \nabla' \cdot \mathbf{B}' &= 0 , & c\nabla' \times \mathbf{B}' &= 4\pi\mathbf{J}' ,\end{aligned}$$

or

$$\begin{aligned}\nabla' \cdot \mathbf{E}' &= 4\pi\delta' , & c\nabla' \times \mathbf{E}' &= 0 , \\ \nabla' \cdot \mathbf{B}' &= 0 , & c\nabla' \times \mathbf{B}' &= 4\pi\mathbf{J}' + \frac{\partial \mathbf{E}'}{\partial t'} ,\end{aligned}$$

then we *do* have invariance under the Galilean Group! (Try it!) But, alas, we don't have any photons. So we have several options here. We get rid of photons

(not good). We agree that in Montreal, for example, the Maxwell Equations are what they are and all of us moving with respect to Montreal will need to use different electromagnetic equations depending upon our velocity relative to that city (not good either). We give up on the Galilean Group and go looking for something else (not great, but better than the other two choices).

A closer look reveals that Maxwell's Equations *are* in fact invariant under the action of the transformations in the Galilean subgroups SO_3 , T_3 , and T_1 —it is the boost group GB_3 which causes the inconsistencies. It would be nice to find a different subgroup for boosts that preserves both Newton and Maxwell, but since we are restricted in our program to *linear transformations* there is insufficient freedom to accomplish this. Therefore, you can have one, or you can have the other, but, alas, you cannot have both.

The (3 dimensional) *Lorentz "Group"* of boosts, L , ensures that light travels at the same speed c , in all viable coordinate systems, but, it causes accelerations to appear to be different from coordinate system to system. The "group" is unique up to an isomorphism. In otherwords there is one and only one set of transformations which is a "group" and which accomplishes this desired objective. (The reason I am putting quotes around "group" will become clear shortly.) The elements of the one-dimensional subgroup (no quotes are needed here!) where all the action takes place are

$$\psi_u : (x, t) \mapsto (\gamma[x - ut], \gamma[t - ux/c^2]) ,$$

where

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}} .$$

Although it is slightly less obvious at first sight than it was for the Galilean boosts, GB_3 , this set of ψ_u for $u \in \mathbb{R}$ is also a group, with ψ_{-u} again being the inverse of ψ_u and ψ_0 being the identity element. Frequently these mappings are called the *Lorentz Transformations* between the two coordinate (x, t) and $(x', t') \equiv (\gamma[x - ut], \gamma[t - ux/c^2])$.

The Lorentz Transformations leave the two components of \mathbf{x} transverse to \mathbf{u} unaffected. So we can take account of this fact, by enlarging the scope and dimension of the "group" elements:

$$\psi_{\mathbf{u}} : (\mathbf{x}, t) \mapsto (\mathbf{x} + \mathbf{u}\{[(\gamma - 1)(\mathbf{u} \cdot \mathbf{x})/u^2] - \gamma t\}, \gamma[t - \mathbf{u} \cdot \mathbf{x}/c^2]) ,$$

where

$$\gamma = \frac{1}{\sqrt{1 - u^2/c^2}} , \quad u = |\mathbf{u}| ,$$

and $\mathbf{u} \in \mathbb{R}^3$, now.

We now find a very interesting thing—the collection of all linear transformations $\psi_{\mathbf{u}}$, L , is not actually a group. There is an identity, $\psi_{\mathbf{0}}$ and $\psi_{-\mathbf{u}}$ is the inverse of $\psi_{\mathbf{u}}$, but when \mathbf{u} and \mathbf{v} are not parallel to one another, then

$$\psi_{\mathbf{v}}\psi_{\mathbf{u}}, \psi_{\mathbf{u}}\psi_{\mathbf{v}} \notin L ,$$

however,

$$\psi_{\mathbf{v}}\psi_{\mathbf{u}}, \psi_{\mathbf{u}}\psi_{\mathbf{v}} \in L \bigcup \text{SO}_3[\mathbb{R}] .$$

The successive application of two *nonparallel* Lorentz Transformations is equivalent to the combination of another Lorentz Transformation *and* a rotation from $\text{SO}_3[\mathbb{R}]$.

The full symmetry group of the Minkowski space-time, is the *Poincaré Group*,

$$P = \text{SO}_3[\mathbb{R}] \bigcup L \bigcup T_3[\mathbb{R}] \bigcup T_1[\mathbb{R}] ,$$

where L by itself, while a subset, is not a subgroup—all the other subsets are subgroups as well. For comparison

$$G = \text{SO}_3[\mathbb{R}] \bigcup \text{GB}_3[\mathbb{R}] \bigcup T_3[\mathbb{R}] \bigcup T_1[\mathbb{R}] .$$

Notice, too, that I did not give L a field argument \mathbb{R} , because, after all \mathbf{u} is restricted to the interior of an open ball in \mathbb{R}^3 with a radius of c .

Like the Galilean Group, the Poincaré group also requires 10 continuous parameters to uniquely determine each element. Finally, notice that in the limit $u \ll c$, one recovers the elements of GB_3 as the leading order behavior of the elements of L.

This is an absolutely remarkable and important result, because it indicates that our conventional notion of space-time, being heavily steeped in Newton and Galileo, is an accurate portrayal of Minkowskian physics that moves very slowly. But, if Maxwell's Equations are indeed correct as they stand, then there is a very different notion of our space-time for things that move rapidly and for photons which travel at the speed of light in every viable coordinate system. It also represents a symmetry breaking, because GB_3 is a subgroup of G, while L is not a subgroup of P.

6. Summary

It is curious that at a very basic level, the essential four-dimensional space-time that underlies both Galilean and Minkowski space-times is essentially the very same vector space! What is different is the symmetry group of transformations which is employed to connect the allowable bases and associated coordinate systems in which the laws of physics are deemed to be operative and invariant. For Newton (and, so to speak Galileo), everyone needed to agree on the same forces and accelerations no matter where they set up their experiments and coordinate systems (so long as they were *inertial coordinate systems*). But in this case, light travels at different speeds for everyone moving relative to each other, which is inconsistent with Maxwell's Equations being the same in everyone's coordinate system. We get Minkowski and Poincaré if we instead require that everyone uses the same Maxwell Equations in their set ups and coordinate systems. They will not, then, agree on accelerations, and forces, with these discrepancies (between inertial coordinate systems) increasing with the relative rectilinear speeds between the two systems.

Since, as far as anyone has been able to ascertain by careful and devious experiments, light does travel at the same speed in all inertial reference frames, *and*, discrepancies in forces and accelerations increasing with relative velocity have been verified, it seems that given a choice, we must choose the Minkowski space-time as the better representation of the universe we inhabit.

In Minkowski space-time, we agree on the invariance of Maxwell's Equations—which requires, a revision of Newton's laws of motion that becomes more pronounced with increasing relative speed—but we do not in particular need to agree upon things like the actual magnitudes and orientations of electric and magnetic fields, charge, and current densities that enter into these equations! The Poincaré Group is set up *only* to ensure the invariance of Maxwell's Equations and the fact that light travels at the same speed in every inertial frame. Nothing more can be asked of it, as there are no additional degrees of freedom left in the specification of the 10-parameter Poincaré Group. So if electric fields are invariant, that is nice, but there is no guarantee they will be. And in fact, they are not.

Stepping back and thinking a little more generally, if observers in different inertial frames will observe the same phenomenon and measure different electromagnetic fields and sources, they may also measure different, densities, pressures, gravitational potentials, temperatures, frequencies and wavelengths, etc. Already in Galilean space-time different inertial observers observed the same object in motion and concluded it was traveling at different velocities. In Minkowski space-time the same is true but with the interesting twist that now they conclude that for something, light, they actually measure the same speed (but *not* necessarily the same velocity).

What it remains to do, and we take this up in the next Scene, is to deduce the relationships between physical quantities measured in two different inertial reference frames which are connected by an element of the Poincaré Group.

7. Exercises

Exercise 1: THE SMALLEST THREE DIMENSIONAL VECTOR SPACE!

The three-dimensional Euclidean vector space is a very familiar, and also a very *large* vector space, because, the field of real numbers \mathbb{R} is itself a very big set of scalars. So big in fact, that it is not possible to make a listing of all the real numbers, which, at least we could do for the integers, another set, that we denote by $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$ (which is *not* an algebraic field. Why?). To construct the *smallest* three dimensional vector space, we should try to use the smallest possible field for our scalars, which turns out to be $\mathbb{Z}_2 = \{0, 1\}$, containing only two scalars, 0 and 1. All the usual ways in which you multiply and add zero and one work just as you expect with the one exception that because there is no “2”, we require that $1 + 1 = 0$. That is, one is its own inverse under addition, and, cunningly enough, also its own inverse under multiplication since $1 \cdot 1 = 1$.

(A) Convince yourself that \mathbb{Z}_2 and the obvious generalization $\mathbb{Z}_3 = \{0, 1, 2\}$ are in fact algebraic fields, but that as it stands $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ is not a field! For bonus points look at \mathbb{Z}_5 (which *is* a field) and \mathbb{Z}_6 (which is *not* a field) and see

if you can guess why \mathbb{Z}_p is an algebraic field if and only if p is a prime number.

(B) By definition, a three dimensional vector space has a basis with three linearly independent elements, $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Therefore convince yourself that the three dimensional vector space over the field \mathbb{Z}_2 has just $2^3 = 8$ unique elements! Then list them in terms of their coordinates with respect to the basis \mathcal{B} . Remember that one of them has to be the identity $\mathbf{0}$ element. [Hint: every element must be its own inverse!] Therefore the smallest possible vector space in three dimensions must have 8 unique elements—you cannot make anything smaller. It is equivalent to one of the five Platonic solid with 6 faces, 12 edges and 8 vertices (being the elements of the vector space)—also known as a cube. The other 4 Platonic solids are the tetrahedron (4 vertices), octahedron (6 vertices), icosahedron (12 vertices) and the dodecahedron (20 vertices). For bonus points convince yourself that *only one* of these can be equivalent to a three dimensional vector space over some finite field [hint: see the discussion in part (D) below, and note that the octahedron is the dual of the cube under the exchange of faces with vertices.]

(C) If we retain our three dimensional basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, but use *all* of the integers \mathbb{Z} as our scalar field, then we have shown that a three dimensional lattice of points, like we might find in a crystal or structured solid, is, in fact, another (bigger) three-dimensional vector space \mathbb{Z}^3 —but still only a very tiny part of our full Euclidean space! Lattices and variables that can take only two states, on = 1, and off = 0, are the essential ingredients of Boolean Algebra and communication theory.

(D) See if you can modify slightly the rules of addition and multiplication to find a way to make the set of four distinct elements $\{0, 1, a, a^{-1}\}$ into a field [hint: assume $a a^{-1} = 1$]. Your success is consistent with a theorem that states that finite fields exist only if the number of elements in the field is an integer power of a prime number. Therefore, the success you had in making a set of $4 = 2^2$ elements into a field cannot be repeated for a set with 6 elements. There are no finite fields with 6, 10, 12, ..., etc, elements.

8. Further Reading

I agonized again over how little or how much to put in this Scene or even whether to bother with it all, period. Still, I persisted since I have often been uncomfortable the way the Lorentz Transformations are pulled out of thin air in MHD and to a lesser extent in radiative transfer. I've tried to strike a balance between difficult mathematical concepts and practical implications. Too often what is left unsaid in most books and monographs on these topics is exactly what I was looking for. I tried to state some of these things here.

Mihalas & Mihalas [MM 1] does quite a nice job of developing a completely relativistic approach to RHD (*Radiation Hydrodynamics*—remember, no MHD or electromagnetism is present in their treatment). It's systematic and builds nicely. It repays careful study many times over. To add in the MHD to their treatment, I suggest

[L 3] André Lichnerowicz, *Relativistic Hydrodynamics and Magnetohydrodynamics*, (New York, NY: W.A. Benjamin; 1967), ix+196.

Also useful, particularly if you like his style, is

[L 4] A. Lichnerowicz, Elements of Tensor Calculus, (New York, NY: John Wiley & Sons; 1962), vii+164,

which picks up where we leave off at the end of §2 and §3 and introduces the machinery of calculus on manifolds.

An all-around reference for many of the things covered here and later in the mathematical Appendix to this *Opera* is the encyclopedic and often accessible (and often *not* accessible in my opinion) magnum opus by Roger Penrose

[P 8] Roger Penrose, The Road to Reality. The Complete Guide to the Laws of the Universe, (New York, NY: Alfred A. Knopf; 2005), xxviii+1099.

The following papers provide interesting perspectives on Galilean non-invariance of Maxwell's Equations:

[BL-L 1] M. Le Bellac & J.-M. Lévy-Leblond, “Galilean electromagnetism”, *Il Nuovo Cimento*, **14B**(2), 217-34, 1973,

[PFM 1] Giovanni Preti, Fernando de Felice & Luca Masiero, “On the Galilean non-invariance of classical electromagnetism”, *European Journal of Physics*, **30**, 381-91, 2009,

[R 4] Germain Rousseaux, “Forty years of Galilean electromagnetism (1973-2013)”, *European Journal of Physics Plus*, **128**, 81, 2013.

Of the many many technical books on symmetry and Lie groups, perhaps a reasonable starting place is

[GMS 1] I.M. Gelfand, R.A. Minlos & Z. Ya. Shapiro, Representations of the Rotation Group and Lorentz Groups and Their Applications, (Mineola, NY: Dover Publications; 2018), xviii+384.

9. Appendix A: Relativistic RMHD

Under the action of the Poincaré Group, Maxwell's Equations are invariant in form. That is, if in Montreal you are solving

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 4\pi\delta, & c\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, & c\nabla \times \mathbf{B} &= 4\pi\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t},\end{aligned}$$

and on my Air Canada flight I am solving

$$\begin{aligned}\nabla' \cdot \mathbf{E}' &= 4\pi\delta', & c\nabla' \times \mathbf{E}' &= -\frac{\partial \mathbf{B}'}{\partial t'}, \\ \nabla' \cdot \mathbf{B}' &= 0, & c\nabla' \times \mathbf{B}' &= 4\pi\mathbf{J}' + \frac{\partial \mathbf{E}'}{\partial t'},\end{aligned}$$

we'll both get the right answer! Similarly if I am using

$$\frac{1}{c} \frac{\partial I'_{\nu'}}{\partial t'} + \mathbf{n}' \cdot \nabla' I'_{\nu'} = \eta'_{\nu'} - \chi'_{\nu'} I'_{\nu'}.$$

and you are using

$$\frac{1}{c} \frac{\partial I_{\nu}}{\partial t} + \mathbf{n} \cdot \nabla I_{\nu} = \eta_{\nu} - \chi_{\nu} I_{\nu}.$$

we'll again get the “same” results.

But we'll get “different” results, if, for example you use

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla p - \frac{1}{\rho} \nabla \cdot \boldsymbol{\sigma} = \mathbf{g} ,$$

and I try to use

$$\frac{\partial \mathbf{u}'}{\partial t'} + \mathbf{u}' \cdot \nabla' \mathbf{u}' + \frac{1}{\rho'} \nabla' p' - \frac{1}{\rho'} \nabla' \cdot \boldsymbol{\sigma}' = \mathbf{g}' ,$$

to describe the air flow in the Air Canada cabin. Because the Air Canada flight is at best travelling at a speed of $600 \text{ mph} = 2.6822 \times 10^4 \text{ cm sec}^{-1}$ and $c \approx 3 \times 10^{10} \text{ cm sec}^{-1}$ we can live happily with these differences unless we are keeping more than 6 significant digits in our calculations. The Parker Solar Probe will reach speeds approaching 700 times that of my Air Canada flight, so if I was on board the spacecraft we would start to get concerned if we were keeping more than 3-4 significant digits. Actually, I'd be a lot more concerned about a whole lot of other things, but that is besides the point.

When typical flows speed in an astrophysical problem begin to approach a percent or so of the speed of light, then one must use the relativistically correct equations for the material and the gravitational field instead of their nonrelativistic limits, like

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{1}{\rho} \nabla p - \frac{1}{\rho} \nabla \cdot \boldsymbol{\sigma} = \mathbf{g} ,$$

for example.