

# LECTURE 1: A BRIEF HISTORY OF SPACETIME

MAT LANGFORD

*Email:* mathew.langford@anu.edu.au

\*Most of the contents of these notes are adapted from a special relativity course given by Ian Benn at Newcastle University in, I think, 2007

*“He who attempts natural philosophy without geometry is lost” - Galileo<sup>1</sup>*

## 1. ARISTOTELIAN SPACETIME

It is very intuitive to think of events as occurring in ‘space’ at given ‘times’. Since only differences in times are meaningful, it makes sense mathematically to model time as a *one dimensional affine space*,  $T$ . Since ‘space’ seems to have three independent directions, it makes sense to model it with a *three dimensional affine space*,  $S$ .

**Definition 1.1.** A (real)  $k$ -dimensional affine space is a triple  $(A, -, \mathbb{A})$ , where  $A$  is a set,  $\mathbb{A}$  is a  $k$ -dimensional (real) linear space (the space of displacements), and

$$\begin{aligned} - : A \times A &\rightarrow \mathbb{A} \\ (p, q) &\mapsto p - q \end{aligned}$$

satisfies

- (1) for all  $q \in A$ , the map  $p \mapsto p - q$  is a bijection; and
- (2) for all  $p, q, r \in A$ ,

$$(p - q) + (q - r) = p - r.$$

This second condition is known as the triangle rule<sup>2</sup>.

So we could model ‘events’ as being points in the four dimensional affine space  $A := T \times S$ . This is the cartesian product of the underlying sets with the ‘obvious’ affine structure<sup>3</sup>

$$(t_2, p_2) - (t_1, p_1) := (t_2 - t_1, p_2 - p_1).$$

Moreover, we should equip  $T$  and  $S$  with Euclidean structures (i.e inner products on their spaces of displacements) corresponding to the existence of clocks and rulers respectively.

---

<sup>1</sup>Galileo Galilei, 1564 – 1642

<sup>2</sup>Of course, the rule dictates the notation, *not* the other way around: the affine structure ‘ $-$ ’ is not the additive inverse of any ‘ $+$ ’.

<sup>3</sup>The symbol ‘ $:=$ ’ means *equal to by definition*.

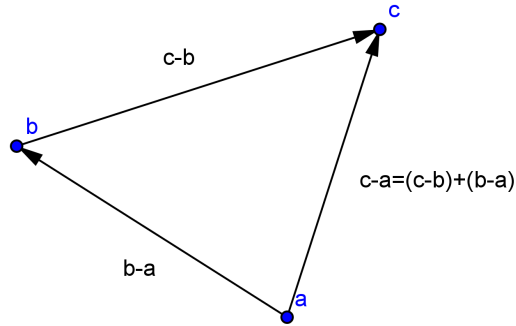


FIGURE 1. The triangle rule

One could model physical objects such as electrons, galaxies, or apples as a curve

$$C: T \rightarrow S,$$

assigning to each time the<sup>4</sup> point at which the electron, galaxy, or apple lies in space. We call such curves (Aristotelian) *observers*, or *particles*. The graph  $\{(t, C(t)) : t \in T\}$  is called the (Aristotelian) *worldline* of the observer.

The spacetime  $A$  inherits the following canonical projections:

$$\begin{aligned} \Pi_T: A &\rightarrow T \\ (t, p) &\mapsto t \end{aligned}$$

and

$$\begin{aligned} \Pi_S: A &\rightarrow S \\ (t, p) &\mapsto p. \end{aligned}$$

Thus, given  $t \in T$ , the set  $\Pi_T^{-1}(t) \subset A$ , can be thought of as a set of *simultaneous* events (since each  $p \in \Pi_T^{-1}(t)$  has the same ‘time coordinate’:  $t$ ). Since the set of simultaneous events is determined naturally by the structure of  $A$ , we say that  $A$  has an *absolute simultaneity* or has *absolute time*.

Similarly, given  $x \in S$ , we can think of the set of events  $\Pi_S^{-1}(x)$  as occurring at the same point,  $x$ , in space. We say that  $A$  is equipped with *absolute rest*, since an observer whose worldline coincides with  $\Pi_S^{-1}(t)$  would be considered, by everybody, to be stationary.

Now, Aristotle taught that objects not acted on by ‘forces’ remained at rest, whilst moving objects required external ‘forces’ to keep them in motion. So we refer to  $A$  as *Aristotelian spacetime*.

In fact, since Aristotle also believed that the Earth is at the centre of the universe, we should equip  $A$  with an origin. This identifies  $A$  with its space of displacements,  $\mathbb{R} \times \mathbb{R}^3$ .

<sup>4</sup>Of course, this means that we’re thinking of these objects as being ‘pointlike’

### 1.1. Problems with Aristotelian Spacetime.

*“E pur si muove”* - Galileo

In his *Dialogue Concerning the Two Chief World Systems*, Galileo (via his protagonist Salviati) poses the following thought experiment<sup>5</sup>:

*Shut yourself up below decks on some large ship... have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.*

This argument clearly refutes the idea of absolute rest, and hence, if we are to accept Salviati's point of view, we must reject Aristotelian spacetime as a model for our world.

Moreover, Galileo (and before Galileo, Copernicus) refuted the idea that the Earth is at the centre of the universe. More importantly, he used (in vain) experimental evidence to support his claims—his observations of the moons of Jupiter.

## 2. GALILEAN SPACETIME

To incorporate the indistinguishability of rest and uniform motion into our spacetime model, we see that we need to remove the space projection  $\Pi_S$  from the above construction. So let's consider some four-dimensional ‘space’ of events  $G$  equipped with an absolute time projection

$$\Pi_T: G \rightarrow T.$$

So we retain the spaces  $\Pi_T^{-1}(t)$  of simultaneous events. Such a construction can be modelled mathematically by a fibre bundle. Roughly speaking:

**Definition 2.1.** A fibre bundle is a quadruple  $(G, \Pi_T, T, F)$ , where<sup>6</sup>

- (1) the base space  $T$ , total space  $G$ , and fibre  $F$  are topological spaces;
- (2) the bundle projection  $\Pi_T: G \rightarrow T$  is a continuous surjection;
- (3) for each  $p \in G$ ,  $\Pi_T^{-1}(p)$  ‘looks like’<sup>7</sup>  $F$ ; and
- (4) locally,  $G$  ‘looks like’<sup>8</sup>  $T \times F$ .

We will often denote the fibre bundle  $(G, \Pi_T, T, F)$  by  $\Pi_T: G \rightarrow T$ , or simply by referring to the base space  $G$ .

This construction looks an awful lot like a fancy way of defining  $T \times F$ ! The main point is that, whilst we can identify (locally)  $G$  with  $T \times F$ , there is no longer a *canonical* way to do so. We still want to be able to measure relative displacements and distances

<sup>5</sup>[http://en.wikipedia.org/wiki/Galileo's\\_ship](http://en.wikipedia.org/wiki/Galileo's_ship)

<sup>6</sup>The precise definition is not necessary for our discussion. Nor is it overly difficult to grasp; a good reference is *Differential Topology* by Morris Hirsch.

<sup>7</sup>Via a morphism. E.g. if  $F$  is a vector space, or an affine space, the right meaning for ‘looks like’ is isomorphism.

<sup>8</sup>This time, by ‘looks like’, we mean via a ‘local trivialisation’.

between simultaneous events, so the fibres  $F_t := \Pi_T^{-1}(t)$  should be 3-dimensional affine spaces equipped with a Euclidean structure on their spaces of displacements,  $\mathbb{F}_t$ .

A (Galilean) observer is a *section* of the bundle  $G$ , that is, a continuous map

$$O: T \rightarrow G$$

such that

$$\Pi_T(O(t)) = t.$$

The requirement  $\Pi_T(O(t)) = t$  can be thought of as ensuring that all observers have ‘synchronised their watches’, in the sense that their ‘personal time’ (the curve parameter) agrees with the ‘absolute time’ singled out by the time projection.

Given two observers  $O$  and  $P$ , we can use the affine structure on the fibres (the 3-spaces of simultaneous events) to define the relative displacement of  $P$  with respect to  $O$ , that is,  $\vec{r}(t) := P(t) - O(t)$ . The Euclidean structure then allows us to define their relative distance:  $r(t) := |\vec{r}(t)|$ . We could try to define the relative velocity of  $P$  with respect to  $O$  in the obvious way:

$$\vec{v}(t_0) := \lim_{t \rightarrow t_0} \frac{\vec{r}(t) - \vec{r}(t_0)}{t - t_0}.$$

However, whilst the denominator in the fraction on the right hand side of the definition makes sense, the numerator does not, since we are trying to take the difference of two vectors in different spaces! Thus we need to add a little more structure to our model: an ability to compare displacement vectors at different times. Let’s call such a structure a *parallelism* for  $G$ , since, physically, it tells us when directions at different times are parallel. In order that ‘metre sticks remain metre sticks’, we should require that the parallelism consist of isometries (they should preserve the Euclidean structure).

**Definition 2.2.** A parallelism for  $G$  is a collection of isometries  $\tau_{t_2, t_1} : \mathbb{F}_{t_2} \rightarrow \mathbb{F}_{t_1}$  for each pair  $t_1, t_2 \in T$ .

**Example.** Let  $O$  be a Galilean observer. Then  $O$  can define a parallelism (at least locally) by setting up a family  $\mathcal{O} = \{O_x\}_{x \in \mathbb{F}^3}$  of ‘co-moving’ (equidistant) observers that foliate  $G$ . That is,

- (1) each pair  $O_x, O_y$  of observers remains equidistant:

$$|O_x(t_2) - O_y(t_2)| = |O_x(t_1) - O_y(t_1)| \quad \text{for all } t_1, t_2 \in T;$$

and

- (2) every event in  $G$  lies on exactly one worldline from the family.

The observer  $O$  can now compare displacement vectors by ‘transporting along the foliation’. That is,

$$\tau_{t_2, t_1}(O_x(t_2) - O_y(t_2)) := O_x(t_1) - O_y(t_1).$$

If we assume that  $G$  is equipped with a parallelism  $\tau$ , then we can define the relative velocity of an observer  $P$  with respect to another observer  $O$  by:

$$\vec{v}(t_0) := \lim_{t \rightarrow t_0} \frac{\tau_{t,t_0}(\vec{r}(t)) - \vec{r}(t_0)}{t - t_0}$$

We refer to the collection of observers whose relative velocities with respect to  $O$  are uniform (constant up to identification at different times via the parallelism) as the *inertial class* of  $O$ . The (strong) Galilean principle of relativity asserts that there is a special inertial class<sup>9</sup>, with respect to which “physical laws hold good in their simplest form”<sup>10</sup>. Observers in this special inertial class are called *inertial observers*.

Equipped with a parallelism  $\tau$ , and a class  $I$  of inertial observers, we refer to  $\Pi_T : G \rightarrow T$  as a *Galilean spacetime*.

## 2.1. Inertial Coordinates.

Any observer  $O$  can define coordinates on  $G$  adapted to his worldline. Choosing a ‘time origin’  $t_0 \in T$ , he defines a time coordinate  $t : G \rightarrow \mathbb{R}$  using the projection map, and the affine structure on  $T$ :

$$t(p) = \Pi_T(p) - t_0.$$

Choosing an orthonormal basis  $\{e_i\}_{i=1}^3$  for  $\mathbb{F}_{t_0}$ , the space of displacements of the fibre over  $t_0$ , he defines three space coordinates  $x^i : G \rightarrow \mathbb{R}$  using the Euclidean affine structure:

$$x^i(p) := (p - O(\Pi_T(p))) \cdot e_i(t),$$

where  $e_i(t) := \tau_{t_0,t}(e_i)$  is the ‘parallel translate’ of  $e_i$ , and  $\cdot$  is the inner product on the fibres.

Of course, these coordinates depend on the choice of observer. If  $\hat{O}$  is a second observer, then she may define her own ‘hat’ coordinates by choosing her own time origin  $\hat{t}_0 \in T$  and orthonormal basis  $\{\hat{e}_i\}_{i=1}^3$  for  $\mathbb{F}_{\hat{t}_0}$ :

$$\begin{aligned} \hat{t}(p) &= \Pi_T(p) - \hat{t}_0 \\ \hat{x}^i(p) &= (p - \hat{O}(\Pi_T(p))) \cdot \hat{e}_i(\Pi_T(p)). \end{aligned}$$

where  $\hat{e}_i(t)$  is the parallel translate of  $\hat{e}_i$ .

Let’s assume that the relative velocity of  $\hat{O}$  with respect to  $O$  is uniform, that is,  $\vec{v}(t) = \tau_{t,t_0}(\vec{v}_0)$  for some  $\vec{v}_0 \in \mathbb{F}_{t_0}$ .

We can relate the two time coordinates as follows:

$$\hat{t}(p) = \Pi_T(p) - \hat{t}_0 = \Pi_T(p) - t_0 + (t_0 - \hat{t}_0) = t(p) - a,$$

where we’ve defined  $a := \hat{t}_0 - t_0$ .

---

<sup>9</sup>According to Mach, the inertial class is determined by the distribution of *all* the matter in the universe. Mach justifies this point of view with the following question: If there is but a single particle in the universe, how can it distinguish rotation from non-rotation? Linear acceleration from rest?

<sup>10</sup>Albert Einstein, *The Foundation of the General Theory of Relativity*.

Since any two orthonormal bases are related by an orthogonal transformation, we have  $\hat{e}_i(t_0) = B(e_i(t_0))$  for some orthogonal transformation  $B \in O(\mathbb{F}_{t_0})$ .

Now, since the relative velocity of the two observers is uniform, we have,

$$\begin{aligned} (t_0 - t_1)\vec{v}_0 &= \tau_{t_0, t_1} \left( \hat{O}(t_0) - O(t_0) \right) - \left( \hat{O}(t_1) - O(t_1) \right) \\ \Rightarrow \left( \hat{O}(t_1) - O(t_1) \right) \cdot e_i(t_1) &= \left( \hat{O}(t_0) - O(t_0) \right) \cdot e_i(t_0) + (t_1 - t_0)\vec{v}_0 \cdot e_i, \end{aligned}$$

which gives the component equations

$$\left( O(t_1) - \hat{O}(t_1) \right)^i = \left( \hat{O}(t_0) - O(t_0) \right)^i + (t_1 - t_0)v_0^i.$$

Now, putting this together, and writing  $\Pi_T(p) = \tau$ , we obtain

$$\begin{aligned} \hat{x}^i(p) &= \sum_{j=1}^3 B_{ij} \left( x^j(\tau) - \left( \hat{O}(t_0) - O(t_0) \right)^j - (\tau - t_0)v_0^j \right) \\ &= \sum_{j=1}^3 B_{ij} \left( x^j(\tau) - A^j - t(p)v_0^j \right), \end{aligned}$$

where we have defined  $A := \hat{O}(t_0) - O(t_0)$ . Therefore, observers in the same inertial class may relate coordinates via a transformation of the form:

$$\begin{cases} \hat{t} = & t - a \\ \hat{x}^i = & \sum_{j=1}^3 B_{ij}(x^j - A^j - tv_0^j). \end{cases}$$

These transformations form a 10-parameter group, which we call the *Galilean group*. The assertion that the laws of physics should be the same for inertially related observers can be interpreted as a requirement that any formulation of these laws should be invariant under the action of the Galilean group.

**Exercise.** Show that, if  $G \cong T \times F$ , then, using the parallelism and inertial structure, it is possible to define a natural affine structure on  $G$ , making it a four dimensional affine space.

## 2.2. Acceleration and Force.

**Exercise.** Galilean velocities add in the obvious way: For any three Galilean observers,  $O, P, Q$ , the velocity of  $Q$  with respect to  $O$  is the velocity of  $Q$  with respect to  $P$  plus the velocity of  $P$  with respect to  $O$ . That is,

$$\dot{Q}_O = \dot{Q}_P + \dot{P}_O,$$

where, for example,

$$\dot{Q}_O(t_0) := \lim_{t \rightarrow t_0} \frac{\tau_{t, t_0} (Q_O(t)) - Q_O(t_0)}{t - t_0},$$

and

$$Q_O(t) := Q(t) - O(t)$$

are the relative velocity and displacement vectors of  $Q$  with respect to  $O$ .

In particular, relative velocities are observer dependent. But this should be obvious. On the other hand, within an inertial class, acceleration (of any observer) does not depend on the choice of observer (this should also be obvious). That is,

**Exercise.** Let  $O, P$ , and  $C$  be observers, such that  $O$  and  $P$  are in the same inertial class. Define the relative acceleration of  $C$  with respect to  $O$  by

$$\ddot{C}_O(t_0) = \lim_{t \rightarrow t_0} \frac{\tau_{t,t_0} \left( \dot{C}'_O(t) \right) - \dot{C}'_O(t_0)}{t - t_0},$$

and similarly for  $P$ . Then  $\ddot{C}_P = \ddot{C}_O$ .

We define the *absolute acceleration* (or, simply, the *acceleration*) of an observer to be its acceleration with respect to the inertial observers.

Now consider a *mechanical particle*, that is, a pair  $(C, m)$ , where  $C$  is an observer, and  $m > 0$  is a number called the *inertial mass*. Then, according to Newton, we should attribute any non-inertial behaviour of the particle to the existence of a *force*  $f$ ;

$$m\ddot{C}(t) = f(t), \tag{2.1}$$

where  $\ddot{C}$  is the absolute acceleration of  $C$ . We remark that this is simply the definition of the concept of a ‘force’. The picture is completed by specifying  $f$  and some initial conditions, solving for  $C$ , and aiming your artillery accordingly.

The preceding discussion should sound very familiar (albeit written, perhaps, in an unnecessarily complicated way!), and the Newtonian-Galilean formalism we have set up has been a very successful description of ‘mechanical forces’. However, as it stands, it has two major problems, the first of which involves gravity, and the second of which involves electromagnetism.

**Exercise.** Show that, for an arbitrary observer  $x$ , and a particle  $(C, m)$ , we have

$$m \left( \ddot{C}_x(t) + \ddot{x}(t) \right) = f(t),$$

where  $\ddot{C}_x$  is the acceleration of  $C$  with respect to  $x$ .

### 2.3. Newtonian Gravity.

It can be argued that Newton’s greatest achievement was his “universal law of gravitation”, which unified the celestial motions of stars with the terrestrial motions of apples by asserting that the force exerted on a particle,  $C$  say, with *gravitational mass*  $\mu > 0$  by a second particle,  $D$  say, with gravitational mass  $\Omega > 0$  is given by the famous law

$$f(t) = -G \frac{\mu, \Omega}{r(t)^2} \vec{r}(t),$$

where  $\vec{r} := C - D$  is the relative displacement of  $C$  from  $D$ ,  $r(t) = |\vec{r}(t)|$  is their relative separation, and  $G$  is the gravitational constant, which we henceforth set to 1 (or, later, some multiple of  $\pi$ ). More generally, we may write

$$f(t) = -\mu g(t),$$

where  $g$  is some ‘gravitational field’. According to Poisson, gravitational fields are generated as follows: If we assume that  $g = -\text{grad } \phi$  for some function  $\phi : G \rightarrow \mathbb{R}$  (interpretation: ‘gravitational potential’) and  $\rho := \text{div } g$  (up to factors of  $\pi$ ; interpretation: ‘mass density’), we obtain the equation:

$$\Delta\phi = \rho,$$

which is a convenient mathematical way of saying that gravitational fields are determined by mass distributions.

Now, suppose that a particle  $C$  is moving in a gravitational field as above. A passer-by,  $x$ , notes a relative acceleration,  $\ddot{C}_x$  of  $C$  with respect to herself. Assuming that gravity is the only force present, in the form of a gravitational field (which she denotes  $g_x$ ), she deduces that

$$m(\ddot{C}_x(t) - \ddot{x}(t)) = \mu g_x(t),$$

where  $m$  is the inertial mass of  $C$ . That is,

$$\ddot{C}_x(t) = \ddot{x}(t) + \frac{\mu}{m} g_x(t). \quad (2.2)$$

Now  $x$  might hope to use (2.2) to measure her own absolute acceleration as follows: she takes two cannon balls,  $C$  and  $D$ , of unit inertial mass, and respective gravitational masses  $\lambda$  and  $\mu$ . Then, by dropping the cannon balls (from a conveniently oblique local tower perhaps),  $x$  may measure the gravitational field  $g_x$  by eliminating her acceleration from the two equations derived as in (2.2). She obtains

$$\ddot{C}_x - \ddot{D}_x = (\lambda - \mu)g_x.$$

Therefore, if  $x$  can measure the relative acceleration of  $C$  and  $D$  with respect to herself, as well as the difference between their respective gravitational to inertial mass ratio, then she may determine the value of  $g_x$ , and hence, by (2.2), her own absolute acceleration (so long as the difference in mass ratios is non-zero).

Unfortunately, her hopes are dashed by her observation that all her cannon balls fall with the same acceleration, so that they all apparently have the same mass ratios. Moreover, the Eötvös experiments<sup>11</sup> showed that mass ratio of any ‘ordinary matter’ is always, in appropriate units, equal to unity (to one part in 20 million!). We are forced to conclude that

$$\ddot{C}_x(t) = \ddot{x}(t) + g_x(t). \quad (2.3)$$

---

<sup>11</sup>Eötvös (1885, 1889).



Thus, *the motions of freely falling<sup>12</sup> particles with respect to a uniformly accelerated frame are indistinguishable from the motions of freely falling particles in a corresponding gravitational field.* This is called the *weak equivalence principle*.

We could brush the problem aside by including  $\ddot{x}$  in the gravitational field,  $g_x$ . This is known as the *strong equivalence principle*: “If it looks like gravity, and it quacks like gravity, then it *is* gravity”. Einstein, in typical Einsteinian fashion, demonstrated the idea with an elegant Gedanken (involving an elevator, that presumably you have all heard at one time or another).

The assertion of the strong equivalence principle suggests that the preferred class of observers should be the freely falling observers. That is, the laws of physics should look the same to all freely falling observers (henceforth ‘freefallers’).

Recall that the inertial observers were precisely those curves satisfying  $\ddot{C} = 0$  with respect to the chosen parallelism. Thus, since the relative acceleration of two inertial observers is zero, Newton’s law of motion looks the same to all inertial observers. This is no longer the case for freefallers, since the difference in acceleration of two freefallers is in general non-zero. Later, we will modify the Galilean parallelism so that all of the freely falling observers satisfy  $C'' = 0$  (with the dashes appropriately interpreted). But first, we’ll show how Galilean relativity *cannot* be adjusted to accommodate the electromagnetic force.

## 2.4. Electromagnetism.

In the 19th century, the nature electromagnetic interaction came to be rather well understood, with the discovery and formulation of Maxwell’s equations and the Lorentz force law.

Observe that a force, as described above, is at each point  $t \in T$ , a vector in  $\mathbb{F}_t$ , the space of displacements of the fibre at  $t$ . We think of  $f$  as a ‘time-dependent vector field’ along the observer curve. More generally, we could define a time dependent vector field to be a map  $V : G \rightarrow \mathbb{F} := \sqcup_{t \in T} \mathbb{F}_t$  such that  $V(p) \in \mathbb{F}_{\Pi_T(p)}$  (at each point  $p \in G$ , we attach a displacement vector  $V(p)$  from the fibre containing  $p$ ).

The electromagnetic force is defined by electric and magnetic fields  $E$  and  $B$  via the Lorentz force law (see below). These fields are time-dependent vector fields that solve Maxwell’s equations:

$$\left\{ \begin{array}{l} \operatorname{div} E = \rho \\ \frac{\partial B}{\partial t} + \operatorname{curl} E = j \\ \operatorname{div} B = 0 \\ \frac{\partial E}{\partial t} - \operatorname{curl} B = 0, \end{array} \right.$$

where  $\rho : G \rightarrow \mathbb{R}$  is some specified function (called the *charge density*), and  $j$  is some specified time-dependent vector field (called the *current density*). Now, we can make sense of the operators  $\operatorname{div}$  and  $\operatorname{curl}$ , by using the divergence and curl operators on the space slices  $\mathbb{F}_t$  at each time  $t \in T$ . However, in order to define the time derivative  $\partial_t$ , we need to

<sup>12</sup>That is, not undergoing non-gravitational forces.

choose an origin for the fibres  $F_t$  at each time, and this amounts to choosing an observer, or equivalently, some inertial coordinates  $(t, x^1, x^2, x^3)$  (with respect to some inertial basis  $\{e_1, e_2, e_3\}$ ). In that case, we have

$$\frac{\partial B}{\partial t} := \frac{\partial B^i}{\partial t} e_i(t).$$

The Lorentz force law ties the electromagnetic fields to the Newtonian formalism by asserting that the electromagnetic force on a particle  $C$  with ‘charge’  $q \in \mathbb{R}$  is given by the expression

$$f(C(t)) = q(E + \dot{C} \times B),$$

where  $\times$  is the cross product on the fibres.

Now, observe that the Lorentz force  $f$  depends on the velocity of the particle, which is inertial observer dependent. When we equate the force to the acceleration via Newton’s law, we appear to get an (inertial) observer dependent acceleration! Clearly something is amiss. In fact, the problem is really that

**Exercise.** *Maxwell’s equations are not invariant under Galilean transformations.*

Moreover,

**Exercise.** *Maxwell’s equations ‘predict the existence of electromagnetic waves’, which have constant velocity in any inertial frame.*

Finally, and most importantly, the experiments of Michelson-Morley and others provided evidence that light is described by these electromagnetic waves; in particular, its velocity appears to be the same with respect to any inertial frame. These matters are difficult to reconcile with the Galilean description of space and time as outlined above.

### 3. MINKOWSKIAN SPACETIME

*The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality. – Hermann Minkowski, 1908*

To measure the speed of light, in units of metres per second say, we would have to time light over a set distance. So we assume that we know how to measure time in seconds and distances in metres. Then the anomaly of the constancy of the speed of light must correspond to some inconsistency in this presupposed method for measuring velocities.

Now suppose that you are working as a patent clerk in Geneva. Then you *know* there is nothing wrong with your measurement of seconds: you are wearing a Swiss watch! Supposing further that Maxwell and Michelson-Morley are correct in saying that light travels with constant speed  $c \text{ ms}^{-1}$ , then there must be some problem with our measurement

of distances. The solution to the problem is ‘obvious’:  $1\text{m} = \frac{1}{c}\text{ls}$ , where one light second (ls) is the distance traversed by light in one second. In these units, the speed of light is equal to  $c\text{ms}^{-1}$ , or, in more natural units, one light second per second!

So what does it mean to measure distance in light seconds? In the 70s, NASA sent Apollo missions to the moon to, among other things, place a mirror on its surface. From Earth, a laser was fired at the moon, and the time taken for its return trip measured (with Swiss watches situated on the Earth’s surface). Those watches counted approximately 2.5 seconds for the return journey, putting the moon 1.25 light seconds away.

There can be no mystery about the constancy of the speed of light if we measure distances this way. If we instead try to define distances using, say, rigid measuring rods, then are we measuring the same thing (albeit in different units) or is there some other sort of distance to be measured? One can interpret experiments such as the Michelson-Morley experiment by saying that these two notions of distance (radar ranging and rigid measuring rods) are equivalent. Surely this is what we should have expected—for steel rules are really held together by electromagnetic interactions.

If we agree to measure distances by radar ranging, then we reduce distance measurements to time measurements. As we will see, it follows from the finiteness of the speed of light that these time measurements cannot accord with Galileo-Newtonian ideas. In particular, we are forced to give up the existence of absolute simultaneity measured by absolute time.

### 3.1. Eine Relativitätsgedanken.

Suppose that we observe a duel, fought with pistols, and we wish to determine which duellist fired first. We could set up delicate instruments to exactly time the arrival of the muzzle flashes from the pistols. We could then determine who fired first according to which flash we observed first. But that would not be quite right. Since we know that light travels at a finite speed, the time of arrival of the photons from the muzzle flashes will depend upon the distance travelled. So if, for example, both flashes arrive together, we would deem the person furthest away from us to have fired first. Thus we cannot decide who fired first unless we know how to measure distances. We have agreed to do this by radar ranging.

Suppose that some observer, a martial at the duel, fires a radar ‘starting gun’ pulse which travels to both duellists. The pulse triggers both weapons and the muzzle flashes are subsequently seen by the martial at the same time. The martial concludes that the duellists are the same distance apart and that both weapons were fired at the same instance. (The martial might not be standing in a very clever place, but we assume he sacrifices personal safety to help simplify the discussion.) We can depict this sequence of events on a space-time diagram (see figure 3.1). The martial fires his radar pulse at the point  $O$ , a definite time and place. The electromagnetic pulse meets one duellist at point  $A$ , and the other at point  $B$ . The instant at which the muzzle flashes are detected by the martial is  $C$ . Thus points  $O$  and  $C$  lie on the martial’s worldline. The lines  $OA$  and  $AC$  represent the worldlines of photons, as do  $OB$  and  $BC$ . The martial concludes that both weapons were fired simultaneously and so the duel was fair. Notice that we assume nothing about the

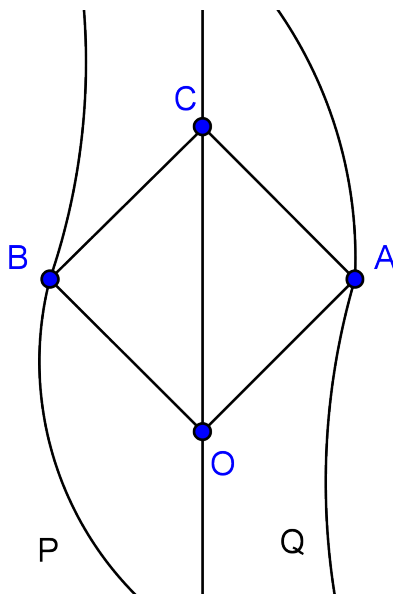


FIGURE 2. A ‘fair’ duel

motion of the duellists. They could be moving relative to each other. All we need to know is the point on their respective worldlines at which they fired their weapon. What they did before or after does not affect our determination of fairness (even if it might determine the outcome of the duel).

Now suppose (to ensure impartiality of the martial) that each duellist has his second monitor the fight. Suppose that one of these seconds is present when the ‘starting gun’ is fired, but that he moves at uniform speed with respect to the martial (see figure 3.1). We suppose that he receives one muzzle flash at point  $D$  with the other arriving later at  $E$ . Thus  $O$ ,  $D$  and  $E$  lie on the worldline of the second. He must conclude that the pistol fired at  $A$  was closest, as the electromagnetic signal returns first, and therefore that duellist received the starting signal first and subsequently fired first. So he deems the duel unfair! Clearly both points of view are equally valid. We must conclude that the ordering of the events  $A$  and  $B$  depends upon your point of view. The martial deems them to be simultaneous whereas our second deems  $A$  to precede  $B$ . If we measure distances by radar then the finiteness of the speed of light forces us to abandon absolute simultaneity.

We need to assume that each observe can measure time, but we are forced to conclude that this time is a personal matter. On our diagram, the ratio of the lengths of  $OD$  and  $OE$  represents the ratio of the times at which the second deems  $A$  and  $B$  to occur. However, as we will see, we cannot just take a ruler and measure on our diagram the lengths of  $OD$  and  $OC$  to measure the ratio of the times measured by the different observers. To see this, suppose that the second travels as fast as he can away from the martial in the direction of  $A$ . The point  $D$  will become arbitrarily close to  $A$  as his velocity approaches the speed of light. So, since  $OA$  and  $AD$  are traversed by the flashes in the same period of time (as determined by the second), the time between firing the starting gun and receiving

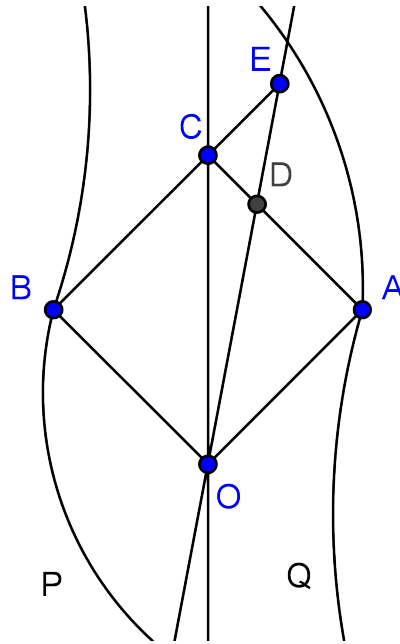


FIGURE 3. A 'fair' duel?

the muzzle flash will get arbitrarily small. Thus, as  $D$  gets closer to  $A$ , the length of  $OD$  represents a decreasing time. In the limit in which  $D$  approaches  $A$  this time must become zero. So a non-zero line segment represents a vanishingly small time. Then we cannot compare times measured by different observers by using a ruler!

Note that, in addition to assuming that the speed of light is constant, we are implicitly using the Galilean assertion that *all inertial observers are created equal*.

### 3.2. Lorentzian Geometry.

We now recall some geometry. Unsubstantiated claims should be considered exercises. Most of what follows is analogous to what you learned in linear algebra 101.

**Definition 3.1.** A pseudo-orthogonal structure on a real-linear space  $V$  is a map

$$g : V \times V \rightarrow \mathbb{R}$$

that

- (1) is bilinear,
- (2) symmetric, and
- (3) non-degenerate.

The final statement means the map  $Y \mapsto g(X, Y)$  is the zero map only if  $X = 0$ . That is, "the only vector orthogonal to everything is the zero vector".

A pair  $(V, g)$  is called a pseudo-orthogonal space.

**Theorem 3.2.** *There exists a basis  $\{X_i\}_{i=1}^n$  for every  $n$ -dimensional pseudo-orthogonal space such that*

$$g(X_i, X_j) = \begin{cases} \pm 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

*Such a basis is called a (pseudo)-orthonormal basis. Moreover, for every orthonormal basis, the respective number  $(p, m)$  of plus and minus signs is the same.*

The pair  $(p, m)$  is called the *signature* of the pseudo-orthogonal structure,  $g$ . A pseudo-orthogonal structure of signature  $(n - 1, 1)$  is called a *Lorentzian* structure. We shall consider  $\mathbb{R}^{3,1}$ , the 4-dimensional Lorentzian space.

**Definition 3.3.** *A vector  $X \in V$  is called*

- (1) *timelike if  $g(X, X) < 0$ ,*
- (2) *spacelike if  $g(X, X) > 0$ , and*
- (3) *lightlike, or null, if  $g(X, X) = 0$ .*

Recall that a *cone* is a subset of a linear space that is closed under positive scalar multiplication.

**Proposition 3.4.** *The set  $J := \{X \in \mathbb{R}^{3,1} : g(X, X) < 0\}$  of timelike vectors in  $\mathbb{R}^{3,1}$  has two connected components,  $J_{\pm}$ , each of which is an open, convex cone. If  $X \in J_+$  ( $J_-$ ) and  $Y \in J$  is such that  $g(X, Y) < 0$ , then  $Y \in J_+$  ( $J_-$ ).*

The set  $J$  is called the *lightcone*. We orient  $J$  by fixing some  $j \in J$ , and asserting  $j \in J_+$ . We call  $J_+$  the *future lightcone*, and  $J_-$  the *past lightcone*.

We define the *length* of a vector  $X \in \mathbb{R}^{3,1}$  by  $|X| := \sqrt{|g(X, X)|}$ .

**Theorem 3.5** (Reverse Cauchy-Schwarz inequality). *For all  $X, Y \in J_+$ , it holds that*

$$\frac{g(X, Y)}{|X||Y|} \leq -1.$$

*Moreover, equality holds if and only if  $X$  and  $Y$  are colinear.*

**Corollary 3.6** (Reverse triangle inequality). *For all  $X, Y \in J_+$*

$$|X + Y| \geq |X| + |Y|,$$

*with equality if and only if  $X$  and  $Y$  are colinear.*

Now recall from high school that the ‘dot product’ is defined by

$$X \cdot Y := |X||Y| \cos \theta,$$

where  $\theta$  is ‘the angle between  $X$  and  $Y$ ’. Of course, like most of your teenage years, you were lied to: rather, the angle is actually defined by the dot product, as allowed by the standard Cauchy-Schwarz inequality. Since in the Lorentzian case the inequality is backwards:

$$\frac{|g(X, Y)|}{|X||Y|} \geq 1$$

we can no longer use the cosine to define the angle between future pointing timelike vectors. Instead, a well-defined ‘angle’ is determined by the hyperbolic cosine:

$$\cosh \theta := \frac{|g(X, Y)|}{|X||Y|}.$$

In Lorentzian geometry, hyperbolic trigonometry plays an analogous role to the elliptic trigonometry of Euclidean geometry.

### 3.3. Minkowskian Spacetime.

Recall that a spacetime in which all inertial observers measure the same speed for light cannot support absolute simultaneity, and, therefore, cannot have a time projection  $\Pi_T$ . Since we are keeping the Galilean (inertial) postulate, we take (bare) spacetime to be a 4 dimensional affine space,  $M$ . Furthermore, we saw that the constancy of the speed of light imparts a natural cone structure on  $M$ , such that the length of the worldline of light is zero. We are led to equip the space of displacements  $\mathbb{M}$  of  $M$  with a Lorentzian structure.

Moreover, we saw that it shouldn’t make sense for ‘physical objects’ to travel faster than light. We define a *Minkowskian observer* as a smooth curve<sup>13</sup>  $C: \mathbb{R} \rightarrow M$  that is future-pointing timelike, that is, for all  $t \in \mathbb{R}$ , the tangent vector of  $C$ ,

$$C'(t) := \lim_{h \rightarrow 0} \frac{C(t+h) - C(t)}{h} \in \mathbb{M}, \quad (3.1)$$

satisfies  $C'(t) \in J_+$ .

Observe that the arc-length of  $C$ ,

$$L_{(a,b)}(C) := \int_a^b |C'(t)| dt, \quad (3.2)$$

is a geometric invariant: it does not depend on any choice of coordinates. Therefore, it should also have a physical interpretation: we assign to  $L_{(a,b)}(C)$  the concept of *personal time* as measured by  $C$  between the events  $C(a)$  and  $C(b)$ . In order to synchronise watches, we also assert that all observers are *arc-length parametrised* or *unit speed*. That is,  $|C'(t)| = 1$  for all  $t \in \mathbb{R}$ .

We refer to an observer  $C$  whose velocity  $C'$  is uniform as an *inertial observer*. If we define the acceleration of  $C$  by

$$C''(t) = \lim_{h \rightarrow 0} \frac{C'(t+h) - C'(t)}{h} \quad (3.3)$$

then it follows that inertial observers satisfy  $C'' = 0$ .

Now, at any moment  $t \in \mathbb{R}$ , the orthogonal complement of  $C'(t)$ ,  $\Sigma_t := \{X \in \mathbb{M} : g(C'(t), X) = 0\}$ , is also a geometric invariant. We refer to  $\Sigma_t$  as the *instantaneous rest space of displacements*<sup>14</sup> of  $C$  at the moment  $t$ . The set  $\sigma_t := \{P \in M : P - C(t) \in \Sigma_t\}$  is called the *instantaneous rest space* of  $C$  at  $t$ . We interpret  $\sigma_t$  as the set of points  $C$

<sup>13</sup>Note that we have sacrificed the freedom of choice of time-origin for clarity.

<sup>14</sup>I will refer to both the instantaneous rest space, and the instantaneous rest space of displacements as the (instantaneous) rest space.

deems simultaneous to the event  $C(t)$ . The motivation for this interpretation is clear if we consider an inertial observer measuring distances with radar as in the examples above.

**Exercise.** Show that, the acceleration  $C''$  of an observer  $C$  lies in his instantaneous rest space. That is,  $g(C', C'') = 0$ .

### 3.4. Rapidity and Velocity.

Consider now two observers,  $O$  and  $P$ , such that  $P(\tau)$  lies in the instantaneous rest space of  $O$  at time  $t$ . Denote  $X = O'(t)$  and  $Y = P'(\tau)$ . Then how does  $O$  interpret  $Y$ ? She splits it into components parallel and orthogonal to her own tangent vector:

$$Y = Y^{\parallel} + Y^{\perp},$$

where

$$Y^{\parallel} := \frac{g(X, Y)X}{g(X, X)} = -g(X, Y)X,$$

and

$$Y^{\perp} := Y - Y^{\parallel}.$$

Since  $Y^{\perp}$  lies in  $O$ 's instantaneous rest space of displacements at  $t$ ,  $|Y^{\perp}|$  is interpreted by  $O$  as a distance. Furthermore, if we consider the inertial observer  $I$  such that  $I(t) = O(t)$ , and  $I' = X$ , then a point  $I(|Y^{\parallel}|)$  is such that  $I(|Y^{\parallel}|) - I(t) = Y^{\parallel}$ . Since  $|Y^{\parallel}| = L_{(t, t+|Y^{\parallel}|)}(I)$ ,  $|Y^{\parallel}|$  is interpreted by  $O$  as a 'time'. Hence  $O$  views the ratio  $|Y^{\perp}|/|Y^{\parallel}|$  as a *speed*. By the definition of  $\theta$ , we have that

$$\begin{aligned} |Y^{\perp}| &= \cosh \theta \\ |Y^{\parallel}| &= \sinh \theta, \end{aligned}$$

therefore  $\tanh \theta$  is interpreted as the relative speed between the observers  $O$  and  $P$ , as observed by  $O$  at time  $t$ . The hyperbolic angle between two vectors often attains the name *rapidity*. Now, it's easy to see that  $P$  observes the same relative speed between  $P$  and  $O$  at his personal time  $\tau$ . However, if we choose to write  $\vec{v} := \tanh \theta \vec{s}$ , where  $\vec{s}$  is a unit vector in the direction of  $Y^{\perp}$ , then  $\vec{v}$  is  $P$ 's velocity vector as observed by  $O$ , however,  $-\vec{v}$  cannot (in general) be the velocity vector of  $O$  as observed by  $P$ , since  $\vec{v}$  lies in the rest space of  $O$ , and not (in general)  $P$ .

### REFERENCES

- (1) Ian M. Benn, special relativity lecture notes;
- (2) Herman Bondi, *Relativity and Common Sense*;
- (3) Peter Rowe, *Geometrical Physics in Minkowski Spacetime*;
- (4) Morris Hirsch *Differential Topology*;
- (5) [www.math.ucr.edu/home/baez/gr](http://www.math.ucr.edu/home/baez/gr);
- (6) John Stewart, *Advanced General Relativity*.