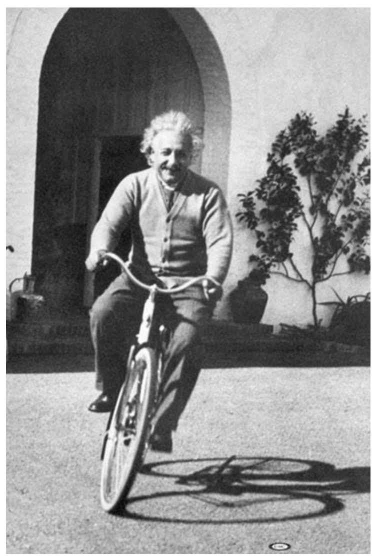
Universidad Simón Bolívar



Introduction to Special Relativity

Mario I. Caicedo
Departamento de Física

V 1.0: Bogotá, September, 2022

Purpose of these notes

These notes are to be used as a compass to guide the students to follow up USB's introductory course on *Relativity*.

We introduce Lorentz transformations from a group theoretical point of view. For a more conventional approach we refer to the references.

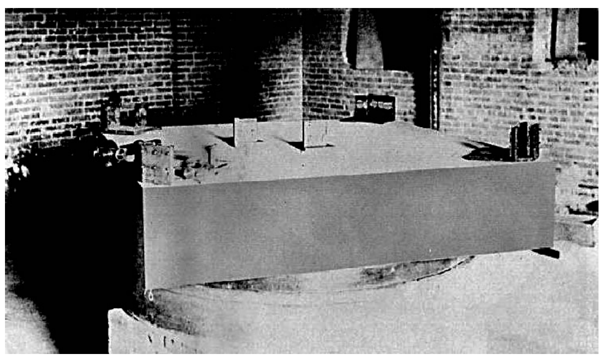
The student is supposed to read material from several sources. Reference [1] gives a detailed account of Einstein's work. References [2, 3, 4] comprise the material of the first third of the course. As always, Landau and Lifshitz's books cannot be forgotten, [5] gives a beautiful description of the dynamics of a particle. Chapter 11 of Jackson's [6] is a bit more advanced, we will use it for some interesting applications, references [7, 8] has a lot of interesting advanced material. Rindler's exposition [9] is remarkable. Finally, for a very modern perspective, we recommend [10]

Contents

List of Figures

Chapter 1

Physics at the end of the XIX Century



In this chapter we suggest some reading topics that may help the students to have a broad view of some of the physics before 1900. We also propose some actual problems that will later show to be helpful.

1.1 Suggested topics

- Criticism of Classical mechanics
- The galilean Group
- What did Leibnitz thought about absolute space-time?

- The speed of light
- Michelson-Morley experiment 1887
- In the frame of the *elastic-solid* theory of light. Voigt (1887) suggested that one could introduce a *local time* into a moving reference frame mantaining the time scale in such a a way that the wave equation remain unchanged in the moving frame.
- In 1892, 1895 H. A. Lorentz obtained new physical results. Introducing a local time t' in a moving frame he showed that all experimentally observed phenomena (to order $\frac{v}{c}$) could be explained if the motion of the electrons with respect to the ether was accounted for.
- Larmor was the first in formulating *modern* Lorentz's transformations.
- In 1904 H. A. Lorentz showed that (provided adequate rules for transforming the fields) Maxwell's equations were invariant under

$$y' = y,\ \ z' = z \tag{1.1}$$

$$x' = \kappa \frac{x - vt}{\sqrt{1 - (\frac{v}{c})^2}}, \tag{1.2}$$

$$t' = \kappa \frac{t - \frac{v}{c^2} t}{\sqrt{1 - (\frac{v}{c})^2}} \tag{1.3}$$

  Unfortunately the proof was rigurous for charge free space only. This late problem was definitely solved by Poincarè who showed (1905) that the full set of equations including charges and currents were invariant under Lorentz's transformations.

  Einstein (1905) did finally found the same results but based on purely physical reasoning.

1.2 Galilean Transformations

Inertial reference frames stand at the very foundation of ckassical mechanics. From a mathematical point of view, Newton's first law defines the equivalence class of all observers that can describe particle dynamics in terms of Newton's equations of motion (second law).

I would dare to say that there is an unstated axiom stating that it exists a privileged observer ($\mathcal{O}$) who describes the dynamics of an arbitrary system of discrete material points according to

$$\ddot{\mathbf{x}}_i = \frac{1}{m_i}\mathbf{F}_i(\mathbf{x}_i, \dot{\mathbf{x}}_i, t), \quad i = 1, 2, \dots, N \tag{1.4}$$

Newton's first law is the satement that all those observers connected with this hypotetical one can relate their coordinate systems with $\mathcal{O}$ through the following transformations

1. Space translations $(\mathbf{x}, t) \to (\mathbf{x} + \mathbf{a}, t)$
2. Change to a moving frame (boost): $(\mathbf{x}, t) \to (\mathbf{x} + \mathbf{u}t, t)$
3. Rigid rotations: $(\mathbf{x}, t) \to (\mathbf{R}\mathbf{x}, t)$
4. In the special case of all forces being time independent, $(\mathbf{x}, t) \to (\mathbf{x}, t + \tau_0)$

A subset of these set of transformations, constitute the Galilean symmetries, which are those reansformations that can be uniquely written as the composition of a rotation, a translation and a uniform motion of spacetime.

With patiente, reading and time you will learn that Galilean transformations constitute a 10 dimensional Lie Group.

In order to have a flavour of the meaning of the transformations we've talking about it is wise to state some problems

Problem 1 *Show that invariance under time shifts implies that the system is autonomous, i.e.*

$$\mathbf{F}_i(\mathbf{x}_i, \dot{\mathbf{x}}_i, t) = \mathbf{F}_i(\mathbf{x}_i, \dot{\mathbf{x}}_i) \tag{1.5}$$

Problem 2 *Show that invariance under global translations implies that the dependece of the forces on the positions of the particles can only be through their relative positions, i.e.*

$$\mathbf{F}_i = \mathbf{F}_i(\mathbf{x}_i - \mathbf{x}_j, \dot{\mathbf{x}}_i)$$

Problem 3 *Show that invariance under boosts imply*

$$\mathbf{F}_i = \mathbf{F}_i(\mathbf{x}_i - \mathbf{x}_j, \dot{\mathbf{x}}_i - \dot{\mathbf{x}}_j)$$

Problem 4 *Show that tha invariance of form of Newton's equations of motion is guaranteed if*

$$\mathbf{F}(\mathbf{R}\mathbf{x}, \mathbf{R}\dot{\mathbf{x}}) = \mathbf{R}\mathbf{F}(\mathbf{x}, \dot{\mathbf{x}})$$

Problem 5 *Use first principles to build the action for a free particle.*
What do you get from your findings?

1.3 Galilean transfromations and the wave equation

Consider the $1+1$ dimensional sourceless wave equation in its canonical form:

$$\frac{\partial^2 u(x,t)}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u(x,t)}{\partial t^2} = 0 \tag{1.6}$$

We want to discuss the behavior of the equation under the galilean transformations

$$\tilde{x} = x - vt \tag{1.7}$$

$$\tilde{t} = t \tag{1.8}$$

to that goal we look for the jacobain of the transformation

$$\begin{array}{ll} \dfrac{\partial \tilde{x}}{\partial x} = 1 & \dfrac{\partial \tilde{x}}{\partial t} = -v \\ \dfrac{\partial \tilde{t}}{\partial x} = 0 & \dfrac{\partial \tilde{t}}{\partial t} = 1 \end{array} \tag{1.9}$$

THe chain rule establishes that

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \tilde{x}}\frac{\partial \tilde{x}}{\partial x} + \frac{\partial u}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial x}, \tag{1.10}$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial \tilde{x}}\frac{\partial \tilde{x}}{\partial t} + \frac{\partial u}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial t}, \tag{1.11}$$

which upon substitution of the entries of the jacobian matrix imply

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \tilde{x}} \tag{1.12}$$

$$\frac{\partial\, u}{\partial t} = -v\frac{\partial u}{\partial \tilde{x}} + \frac{\partial u}{\partial \tilde{t}} \tag{1.13}$$

and for the second derivatives,

$$\frac{\partial^2\, u}{\partial^2 x} = \frac{\partial^2 u}{\partial^2 \tilde{x}} \tag{1.14}$$

$$\begin{aligned} \frac{\partial^2\, u}{\partial^2 t} &= \partial_t \left[-v\frac{\partial u}{\partial \tilde{x}} + \frac{\partial u}{\partial \tilde{t}}\right] = \\ &= -v\partial_{\tilde{x}}\left[-v\frac{\partial u}{\partial \tilde{x}} + \frac{\partial u}{\partial \tilde{t}}\right] + \partial_{\tilde{t}}\left[-v\frac{\partial u}{\partial \tilde{x}} + \frac{\partial u}{\partial \tilde{t}}\right] = \\ &= v^2\frac{\partial^2 u}{\partial \tilde{x}^2} - v\frac{\partial^2 u}{\partial \tilde{x}\tilde{t}} - v\frac{\partial^2 u}{\partial \tilde{t}\tilde{x}} + \frac{\partial^2 u}{\partial^2 \tilde{t}} = \\ &= v^2\frac{\partial^2 u}{\partial \tilde{x}^2} - 2v\frac{\partial^2 u}{\partial \tilde{x}\tilde{t}} + \frac{\partial^2 u}{\partial^2 \tilde{t}} \end{aligned} \tag{1.15}$$

from where we finally obtain,

$$\frac{\partial^2 u(x,t)}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u}{\partial \tilde{x}^2} - \frac{1}{c^2}\left[v^2\frac{\partial^2 u}{\partial \tilde{x}^2} - 2v\frac{\partial^2 u}{\partial \tilde{x}\tilde{t}} + \frac{\partial^2 u}{\partial^2 \tilde{t}}\right] \tag{1.16}$$

meaning that in the new reference frame, the wave equation looks like

$$\left(\frac{c^2 - v^2}{c^2}\right)\frac{\partial^2 u}{\partial \tilde{x}^2} + 2\frac{v}{c^2}\frac{\partial^2 u}{\partial \tilde{x}\tilde{t}} - \frac{1}{c^2}\frac{\partial^2 u}{\partial^2 \tilde{t}} = 0\,. \tag{1.17}$$

We have thus shown that the wave equation is not form invariant (covariant) under the galilean group.

Chapter 2

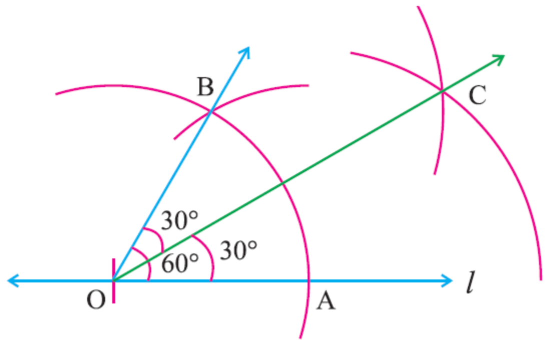
A little bit of Geometry

2.1 Euclidean geometry without rulers

Before entering the marvelous world of physics, we will take a sightseeing walk through some very old math (Euclidean geometry on the plane) as seen form a different perspective.

Our walk begins by asking ourselves a favor, namely, to make an effort in avoiding the use of drawings and of any previous geometrical notions unless explicitly asked for.

Now that we are in the right set of mind, we define $\Re^2$ as the set of $2 - tuples$, i.e. an element $\mathcal{P} \in \Re^2$ is just a pair of reals (x, y). We also consider the set $\mathcal{F}$ of functions from the



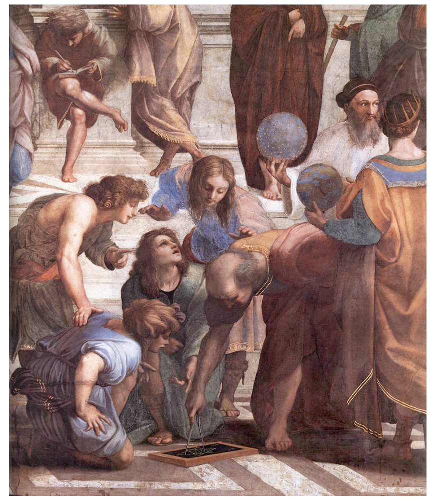


Figure 2.1: **Left**: We are all acquainted with the beautiful ruler and compass constructions of Euclidean gometry, here the bisection of an angle. **Right**:Detail from Raphael's The School of Athens featuring a Greek mathematician – perhaps representing Euclid or Archimedes – using a compass to draw a geometric construction.

reals to $\Re^2$, i.e. $\mathbf{x} \in \mathcal{F}$ iff $\mathbf{x} : \Re \to \Re^2$.

Let now S be the following functional that depends on a sufficiently differentiable element[1] $(x(\tau), y(\tau))$ of $\mathcal{F}$:

$$S[x,y] = \int_{t_0}^{t_f} d\tau \, [(\dot{x})^2 + (\dot{y})^2] \,, \tag{2.1}$$

where as usual, $\dot{x} = dx/d\tau$.

The Euler-Lagrange equations for S show that the critical points of S have the form:

$$x(t) = A_1 \, \tau + A_2 \qquad y(t) = B_1 \, \tau + B_2 \,, \tag{2.2}$$

where A_1, A_2, B_1 and B_2 are real constants.

This is a good point for a disgression. If we appeal to our previous knowledge we recogniz the critical points of S as lines in the cartesian plane. Angles between lines are characterized by the values of A_1 and B_1, the fact that different lines intersect either at one point or not at all comes form the solution of a system of equations, etc. In brief, all Euclidean geometry of

[1] $\tau \in \Re$

the plane can be built form the critical points of S. Several questions arise immediately, the most obvious being:

How did geometry came from a variational problem?

We will go back to this question at the end of this section, for now, let us retake our tour with the following

Definition 1 *Given two points*[2] $\mathcal{P}_1$, $\mathcal{P}_2$ *in* $\Re^2$ *the displacement* $\Delta\mathbf{x}$ *between those points is given by the pair*

$$(x(\mathcal{P}_2) - x(\mathcal{P}_1), y(\mathcal{P}_2) - y(\mathcal{P}_1)) \ . \tag{2.3}$$

we call $\mathcal{P}_1$ *the base of the displacement.*

The set $\Re^2_{\mathcal{P}}$ *of all displacements based at* $\mathcal{P}$ *obviously has the structure of a two dimensional real vector space*[3]. $\Re^2_{\mathcal{P}}$ *is usually referred to as* $T_{\mathcal{P}}\Re^2$ *(the tangent space to* $\Re^2$ *at* $\mathcal{P}$*). We aditionally endow* $T_{\mathcal{P}}\Re^2$ *with a structure called the arc length (or metric), defined as*

$$\Delta\ell^2 = \mathbf{g}(\Delta\mathbf{x}, \Delta\mathbf{x}) \equiv \Delta\mathbf{x}^2 = [x(\mathcal{Q}) - x(\mathcal{P})]^2 + [y(\mathcal{Q}) - y(\mathcal{P})]^2 \tag{2.4}$$

Let us go back to the question of geometry. In first place, it is clear that $\Re^2_{\mathcal{P}}$ is the set of arrows based on $\mathcal{P}$, we may also think of a new set, namely $T\Re^2 = \cup_{\text{all } \mathcal{P}\in\Re^2}\Re^2_{\mathcal{P}}$, the *tangent bundle*, which is the set of all arrows based at any point in $\Re^2$. A curve in $\Re^2$ $((x(t), y(t)))$ has a tangent vector (the velocity, $\dot{\mathbf{x}}$) at each point. The speed in its usual sense[4] is clearly $\mathbf{g}(\dot{\mathbf{x}}, \dot{\mathbf{x}})$, a little thinking leaves clear that the functional 2.1 is the arc length of he curve and so, its critical points are straight lines.

2.1.1 Homogeneous Isometries of the Plane

In the following, we devote some lines to study those transformations on vectors that leave their length invariant-.

[2]The 2-tuple corresponding to $\mathcal{P}$ is $(x(\mathcal{P}, y(\mathcal{P})$

[3]We think of $\Delta\mathbf{x}$ as the column vector: $\Delta\mathbf{x} = \begin{pmatrix} x(\mathcal{P}_2) - x(\mathcal{P}_1) \\ y(\mathcal{P}_2) - y(\mathcal{P}_1) \end{pmatrix}$

[4]Physics 101

Our intuition tells us that such transformations are, translations, rotations and reflections, just reflections and rotations leave the base point of vectors fixed, we well limit ourselves to think of these two last transformations.

It is a simple exercise to prove that the set $G\ell(2,\Re)$ of invertible 2×2 is a **group**[5] that acts naturally on $\Re^2_{\mathcal{P}}$. We are particularly interested in a subset of it, referred to as $O(2)$ defined as the set of matrices such that the transformation

$$\Delta\mathbf{x} \to \Delta\bar{\mathbf{x}} \equiv \boldsymbol{M}\,\Delta\mathbf{x}\,, \tag{2.5}$$

leaves the metric invariant, i.e.

$$\Delta\ell^2 = \Delta\bar{\mathbf{x}}^T\boldsymbol{I}\,\Delta\bar{\mathbf{x}} = \Delta\mathbf{x}^T\boldsymbol{I}\,\Delta\mathbf{x}\,. \tag{2.6}$$

We want to explore the set of infinitesimal $O(2)$ transformations, i.e. the set of all matrices $\boldsymbol{M}$ which are very close to the identity and which allow condition 2.6. To achieve this goal we write $\boldsymbol{M}$ as

$$\boldsymbol{M} = \mathbf{1} + \boldsymbol{\epsilon}\,. \tag{2.7}$$

where $\boldsymbol{\epsilon}$ is a matrix with small (infinitesimal) entries only. Condition 2.6 is equivalent to the identity

$$\boldsymbol{M}^T\boldsymbol{I}\,\boldsymbol{M} = \boldsymbol{I}\,, \tag{2.8}$$

which, upon substitution and retaining first order terms only, implies

$$\begin{aligned}\boldsymbol{M}^T\boldsymbol{I}\,\boldsymbol{M} &= [\mathbf{I}+\boldsymbol{\epsilon}]^T\,\boldsymbol{I}\,[\mathbf{1}+\boldsymbol{\epsilon}]\\ &\approx \boldsymbol{I} + \boldsymbol{\epsilon}^T\boldsymbol{I} + \boldsymbol{I}\boldsymbol{\epsilon}\,,\end{aligned} \tag{2.9}$$

therefore, the infinitesimal matrices must satisfy

$$\boldsymbol{\epsilon}^T + \boldsymbol{\epsilon} = 0\,, \tag{2.10}$$

which implies

$$\boldsymbol{\epsilon} = \begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix} \tag{2.11}$$

[5] under matrix multiplication

where $\epsilon \in \Re << 1$ and $\mathbf{T}$ is the 2×2 matrix

$$\mathbf{T} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.12}$$

There is an important remark to make here, the product of two of these (close to the identity) transformations, say, $\boldsymbol{M}(\varepsilon_1)$ and $\boldsymbol{M}(\varepsilon_2)$ is also an $O(2)$ transformation infinitessimally close to the identity, in fact,

$$\boldsymbol{M}(\varepsilon_1)\boldsymbol{M}(\varepsilon_1) = \boldsymbol{M}(\varepsilon_3) \tag{2.13}$$

with $\epsilon_3 = \epsilon_1 + \epsilon_2$. This in turn shows an interesting result:

$$[\boldsymbol{M}(\varepsilon)]^{-1} = \boldsymbol{M}(-\varepsilon) \tag{2.14}$$

In what follows we will need the following simple lemma

Lemma 1 $\mathbf{T}^2 = -\mathbf{1}$, $\mathbf{T}^3 = -\mathbf{T}$ *and* $\mathbf{T}^4 = \mathbf{T}$

The proof[6] *is left as an excercise*

Let us now go back to formula 2.13 and let us try the following idea, think of multiplying several small transformations to build a finite (non infinitesimal) transformation. To put things in action let $\theta \in \Re$ and $n \in N$ and build the product

$$\boldsymbol{M}(\theta, n) = \underbrace{(\mathbf{1} + \frac{\theta}{n}\mathbf{T}) \dots (\mathbf{1} + \frac{\theta}{n}\mathbf{T})}_{\text{n-times}} \tag{2.15}$$

clearly

$$\boldsymbol{M}(\theta, n) = \left[\mathbf{1} + \frac{\theta}{n}\mathbf{T}\right]^n , \tag{2.16}$$

and for extremely large n, θ/n is indeed an infinitesimal angle, meaning that in such limit we are implementing the original idea of a huge number of infinitesimal rotations..

We now recall that for real numbers,

$$\lim_{n\to\infty} \left[1 + \frac{x}{n}\right]^n = e^x . \tag{2.17}$$

[6]Notice that $\mathbf{T}$ behaves just like $i = \sqrt{-1}$, can you give a reason for this? Hint: think about the complex plane

Fortunately, life is full of joy and the mathematicians have shown that for some matrices (ours being one of those) we may also take the limit and get

$$\boldsymbol{M}(\theta) \equiv \lim_{n\to\infty} \boldsymbol{M}(\theta, n) = e^{\theta \mathbf{T}}\,, \tag{2.18}$$

The exponential may be calculated using the formal series:

$$e^{\theta \mathbf{T}} = \mathbf{1} + \theta \mathbf{T} + \frac{1}{2!}\left(\theta \mathbf{T}\right)^2 + \frac{1}{3!}\left(\theta \mathbf{T}\right)^3 + \frac{1}{4!}\left(\theta \mathbf{T}\right)^4 + \frac{1}{5!}\left(\theta \mathbf{T}\right)^5 + \dots \tag{2.19}$$

using Lemma 1 this expression simplifies to

$$e^{\theta \mathbf{T}} = \left[1 - \frac{1}{2!}\theta^2 + \frac{1}{4!}\theta^4 + \dots\right]\mathbf{1} + \left[\theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 + \dots\right]\mathbf{T} \tag{2.20}$$

which in turn is nothing but

$$e^{\theta \mathbf{T}} = cos\theta\,\mathbf{1} + sen\theta\,\mathbf{T} = \begin{pmatrix} cos\theta & sen\theta \\ -sen\theta & cos\theta \end{pmatrix} \tag{2.21}$$

i.e. a standard rotation -as expected.

The matrix $\mathbf{T}$ is referred to as the generator of rotations on the plane, and in the standar mathematical literature, we find

$$\mathbf{T} = \frac{d\mathbf{M}(\theta)}{d\theta}|_{\theta=0} \tag{2.22}$$

2.2 Thinking like a Physicist: What is a $3D$ Vector?

Let us now think of the gyro in figure 2.2, imagine its axis at the time that the photo was taken to be the vector $\mathbf{r} = \vec{OP}$ of figure 2.3, we want to find a formula to rotate $\mathbf{r}$ around the axis $\hat{\mathbf{n}}$ to get the final vector $\mathbf{r}' = \vec{OQ}$

From the figure, it is clear that the parallel ($\mathbf{r}_{||}$) and orthogonal (to $\hat{\mathbf{n}}$) ($\mathbf{r}_{\perp}$) elements of $\mathbf{r}$ are $\mathbf{r}_{||} = \vec{ON} = \hat{\mathbf{n}}(\hat{\mathbf{n}}.\mathbf{r})$ and $\mathbf{r}_{\perp} = \vec{NP} = \mathbf{r} - \hat{\mathbf{n}}(\hat{\mathbf{n}}.\mathbf{r})$

Obviously,

$$\mathbf{r}' = \mathbf{r}_{||} + \vec{NV} + \vec{VQ}\,, \tag{2.23}$$

but $\vec{VQ} = \mathbf{r} \times \hat{\mathbf{n}}\, sin\Phi$ and $\vec{OQ} = \mathbf{r}_{\perp}\, cos\Phi$, so after substitution

$$\mathbf{r}' = \hat{\mathbf{n}}\left(\hat{\mathbf{n}}.\mathbf{r}\right) + \left[\mathbf{r} - \hat{\mathbf{n}}\left(\hat{\mathbf{n}}.\mathbf{r}\right]\, cos\Phi + \mathbf{r} \times \hat{\mathbf{n}}\, sin\Phi\,. \tag{2.24}$$

Figure 2.2: The precession of a gyroscope

Finally, a litle rearrengement gives

$$\mathbf{r}' = \mathbf{r}\, cos\Phi + [1 - cos\Phi]\, \hat{\mathbf{n}}\, (\hat{\mathbf{n}}.\mathbf{r}) + \mathbf{r} \times \hat{\mathbf{n}}\, sin\Phi \tag{2.25}$$

A couple of remarks are worth at this point, the first is that neither $\mathbf{r}$ nor $\hat{\mathbf{n}}$ have anything in particular, wihich implies that formula 2.25 is completely general[7]. The second is that what we did to $\mathbf{r}$ can be thought of as a map $\mathbf{R}(\hat{\mathbf{n}}, \Theta)$ which acts on $\mathbf{r}$ as follows:

$$\mathbf{r} \to \mathbf{R}(\hat{\mathbf{n}}, \Theta)(\mathbf{r}) = \mathbf{r}\, cos\Phi + [1 - cos\Phi]\, \hat{\mathbf{n}}\, (\hat{\mathbf{n}}.\mathbf{r}) + \mathbf{r} \times \hat{\mathbf{n}}\, sin\Phi\, . \tag{2.26}$$

In fact, since $\hat{\mathbf{n}}$ is unitary, $\mathbf{R}(\hat{\mathbf{n}}, \Theta)$ is a map that depends in three real parameters. Even more, if the axis is held fixed, the set of all $\mathbf{R}(\hat{\mathbf{n}}, \Theta)$ with the obvious composition law that two consecutive rotations of angles Φ_1 and Φ_2 gives a third rotation of angle $\Phi_3 = \Phi_1 + \Phi_2$ is clearly an abelian group.

If we think of a tiny counterclockwise rotation $\Phi = \delta\phi \to 0$ and mantain first order only

[7] Any vector with its base point located at O rotates according to the same formula

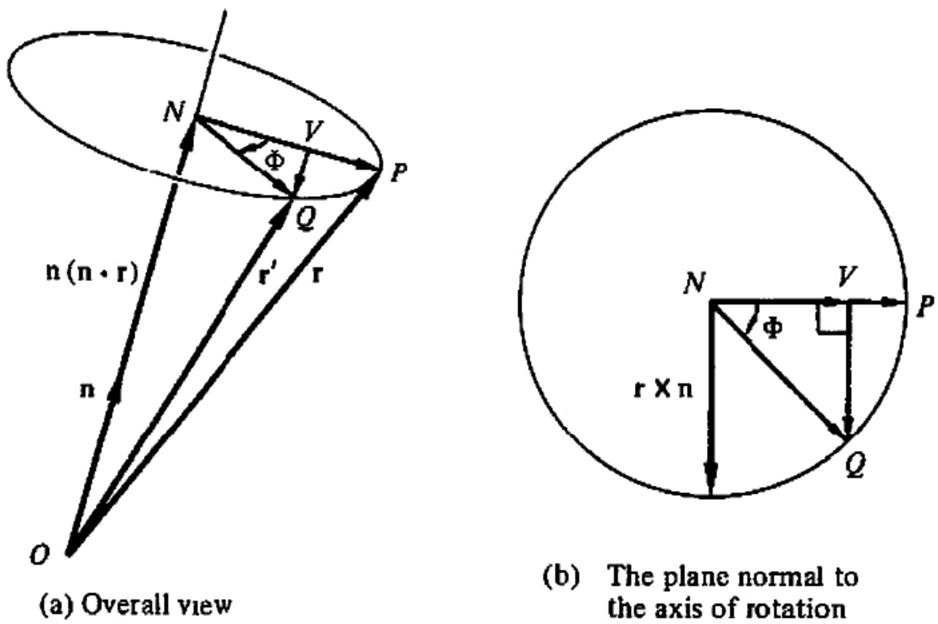


Figure 2.3: Rotating the vector $\mathbf{r}$ clockwise around $\hat{\mathbf{n}}$ gives $\mathbf{r}'$

$$\mathbf{r}' = \mathbf{r} + \delta\Phi\, \hat{\mathbf{n}} \times \mathbf{r}\,. \tag{2.27}$$

At this point we leave the discussion on rotations and turn our attention to something cute, the product, $\hat{\mathbf{n}} \times \mathbf{r}$ can be represented as a matrix product, indeed, if we identify $\mathbf{r}$ and $\hat{\mathbf{n}}$ with their cartesian coordinates and think a little bit, we find

$$\hat{\mathbf{n}} \times \mathbf{r} = \mathbf{M}\,\mathbf{r} \begin{pmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \tag{2.28}$$

The matrix $\mathbf{M}$ can be expressed as $\mathbf{M} = \sum_{i=1}^{3} n_i\, \mathbf{T}_i$ where the $\mathbf{T}_i$ are *rotation generators* that generalize what we did in section2.1.1.

In physics literature (awfully abusing the notation) the linear combination $\mathbf{M} = \sum_{i=1}^{3} n_i\, \mathbf{T}_i$

is usually writen as

$$\mathbf{M} = \sum_{i=1}^{3} n_i\, \mathbf{T}_i = \hat{\mathbf{n}}.\vec{\mathbf{T}} \tag{2.29}$$

Going back to rotations, we can now express formmula 2.27 as

$$\mathbf{r}' = \mathbf{r} + \delta\Phi\, \hat{\mathbf{n}}.\vec{\mathbf{T}}\,. \tag{2.30}$$

Which upon the definition

$$\delta\mathbf{r} \equiv \mathbf{r}' - \mathbf{r}\,, \tag{2.31}$$

reads,

$$\delta\mathbf{r} = \delta\Phi\, \hat{\mathbf{n}}.\vec{\mathbf{T}}\,. \tag{2.32}$$

In the next section we will explore $3D$ rotations in far more detail.

2.2.1 Guided excercise, a short visit to $O(3)$

The purpose of this section/excercise is to generalize what we just learned about rotations.

Part I

1. Define the set $\Re^3$
2. Define the set $\Re^3_{\mathcal{P}}$ of displacements based on one point
3. Describe the vector space structure of $\Re^3_{\mathcal{P}}$
4. Introduce the EUCLIDEAN metric in $\mathbf{g}(\ ,\ )$
5. This was not discussed in the previous section. Define the two entry map

$$\langle \Delta\mathbf{x}, \Delta\mathbf{y} \rangle \equiv \mathbf{g}(\Delta\mathbf{x}, \Delta\mathbf{y})$$

¿Is this an internal product?. Show that $\mathbf{g}$ so defined is invariant under the action of $O(3)$

Part II

We continue this excercise by studying the set $O(3)$ of all 3×3 matrices that preserve the metric close to the identity.

Show that setting

$$\mathbf{M} = \mathbf{I} + \boldsymbol{\epsilon}\,, \tag{2.33}$$

leads to

$$\boldsymbol{\epsilon} = \epsilon^1\,\mathbf{T}_1 + \epsilon^2\,\mathbf{T}_2 + \epsilon^3\,\mathbf{T}_3\,, \tag{2.34}$$

where $\epsilon^1, \epsilon^2, \epsilon^3$ are three real numbers (give a physical/heuristical reason that would had lead us to think there were just three numbers), and the *rotation generators* are

$$\boldsymbol{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\,,\; \boldsymbol{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}\,,\; \boldsymbol{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.35}$$

What would be the general form (we do not need an explicit matrix, just an expression in terms of exponentials) of a rotation of angle α around an axis?

Use what we learned in section 2.1.1 to explicitly build the matrix that represents a rotation of angle ϕ around the y axis.

For the final part of this long excercise we need to introduce a

Definition 2 *A (contravariant) vector* $\mathbf{A}$ *under* $O(3)$ *is a column of three numbers*

$$\mathbf{A} = \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} \tag{2.36}$$

which under rotations transforms as $\mathbf{x} \to \mathbf{M}\,\mathbf{x}$ *where* $\mathbf{M}$ *is a element of* $O(3)$ *(i.e. a* 3×3 *orthogonal matrix).*

Problem 6 *Connect definition 2 with what we learned in section 2.2*

Show that with this notion, the cartesian components (x, y, z) *of position vector of a point particle moving in* 3 *space constitute indeed a vector but the column*

$$\begin{pmatrix} x^2 \\ x\,y \\ z\,x \end{pmatrix}$$

does not.

Problem 7 *How does the following array*

$$\mathbf{T} = \begin{pmatrix} x^2 & x\,y & x\,z \\ y\,x & y^2 & y\,z \\ z\,z & z\,y & z^2 \end{pmatrix} \tag{2.37}$$

transform under rotations? HINT: There is an extremely simple way to answer this question

Chapter 3

SpaceTime

3.1 What does the word SIMULTANEOUS mean? A First Glimpse to Special Relativity.

Let us consider a very simple *gedankenexperiment*. Three starships of the United Federation of Planets, the Discovery (NCC-1031), the Enterprise (NCC-1701) and the Voyager (NCC-74656) are given a scientific mission. Their crews must observe two stars which are close to become supernovae, namely Antares and Betelgeuse. The three starships are to be stationed along the line that joins the two stars, Voyager's skipper, Captain Kathryn Janeway is to sit at mid distance ($D/2$) between the two stars, Captain Lorca (Discovery), shall be located close to Betelgeuse at one fourth of the distance between the two stars ($D/4$), while Captain Jean Luc Piccard will take measurements at $D/4$ of Antares.

Janeway reports that *both stars became supernovas simultaneously* on stardate 2364.

Let us now try to answer the following questions,

1. What would Captains Piccard and Lorca report? (forget dates, be qualitative)

2. Are there any discrepancies?, if the answer is yes, can you point out the source of them?

3. What does *simultaneously* mean?

4. Let us change our experiment and consider three regular persons in a bar, they see two

cups falling to the floor and for sheer luck the geometry is (up to rescaling) exactly the same of our treky experiment. What would the three observers report?, how could we reproduce some discrepancies just as before?

5. What would you say about causality?

What we have just realized is that human experience is extremely limited, to say the least. Most of our lives we deal with phenomena at a very particular scale (ours). The International Space Station, for example, is far from our common grasp, indeed, the ISS orbits the Earth each $92\,min$ at a speed of nearly $17150\,miles/h = 27594\,Km/h$, that's about $5\,miles\,s^{-1}$ or $8.04\,Km\,s^{-1}$,i.e. the ISS orbits the earth nearly 50 times faster that a F1 car, even so, the ISS travels at just $2.6819 \times 10^{-3}\,\%$ of the speed of light, ¡it is practically standing still!.

Light is fantastically fast, it travels at approximately $1\,ft/ns$, i.e. it takes nearly three nanoseconds for light to travel one meter and of course, we cannot access such tiny times unless we use (not too fast for today's standars) an electronic timer.

Things get a little more complicated when observers move with respect to each other. Nevertheless, the effect of a finite but very fast speed of light is with us even in the simple case of observers ocuppying different positions. c is therefore, the "source" of special relativity.

The Galilean/Newtonian notions of space and time are absolute. All observers have clocks that tick at the same rate and all of them interpret the events that they detect as being universally simultaneous. Such notions are not foolish at all, they come from our experience. Have Galileo, or Newton, or us had warp speed starships allowing us to boldly go where no one else has gone before, as a usual activity, or intuition would have not been clouded by an entire life of moving at very small speeds.

Hopefully we will use the language of four dimensional space time to develop a new intuition to become expert relativists.

3.2 Spacetime: what is it?

This section is strongly based on references [2, 3].

The following quotation comes directly from [3]:

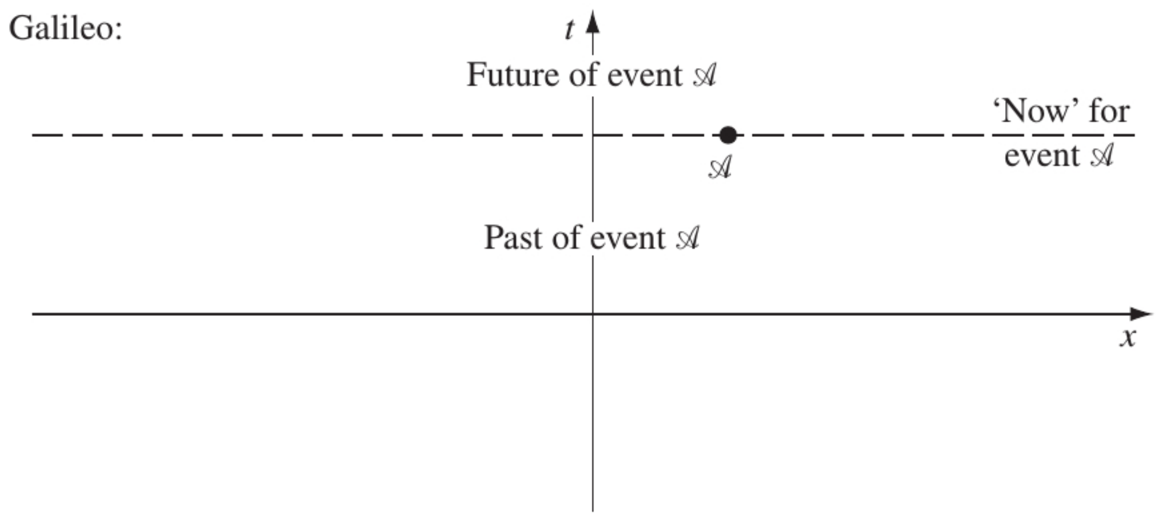


Figure 3.1: Galilean space time, all observers share the same simultaneity lines

"The fundamental concept in physics is an event. An event is specified not only by a place but also by a time of happening. Some examples of events are: collisions of particles or flashes of light (explosions), reflection or absorption of particles or light flashes, collisions, and near-collisions called coincidences."

To state it in the simplest possible way, space time is the place where and when physical phenomena take place (time and physical place). An example can be useful, Mount St Helen eruption (event) took place at 08:32 Pacific Day Time, May 18, 1980 at 46.2º N, 122.2º W (time and location in space).

From the mathematical point of view, space time is regarded as a manifold, a continuum on which experimenters may set up coordinates. We may say that Newtonian mechanics also uses a sace time description, and this brings a natural question, what is different between Newtonian and Einsteinian physics?

The answer is to be found in geometry. In Newtonian physics, time and space can be thought of as independent, in Einstenian physics time and space are **inextricably intertwined** through geometry, and geometry comes in motivated by a postulate. Indeed, Einstein's Special Relativity is mainly based on the following explicit postulate

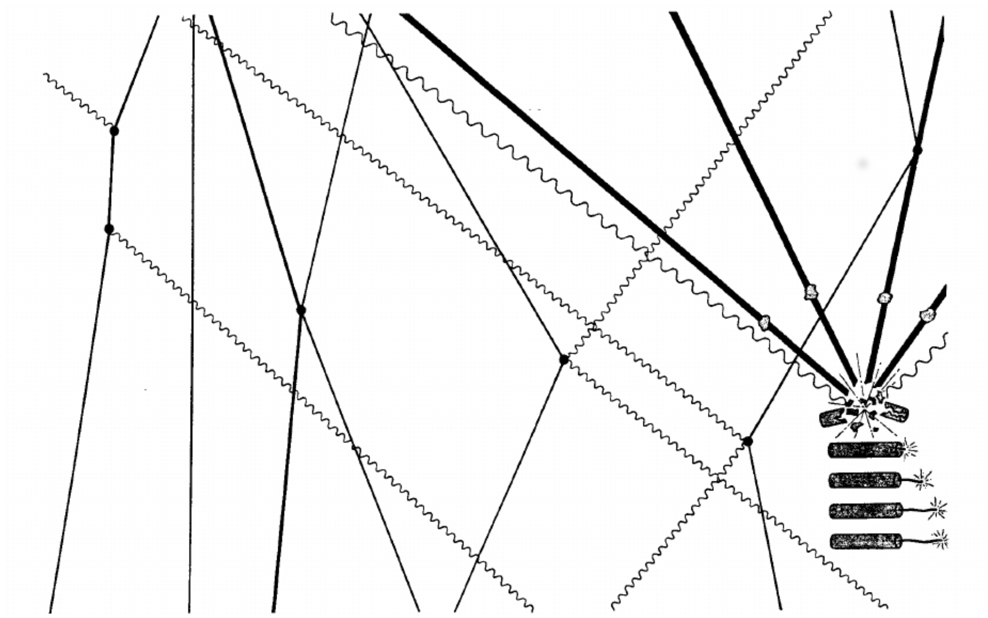


Figure 3.2: Particles and photons moving in space time. The motion traces a curve called the particle's world line. Space time does not need coordinates for its existence

Postulate 1 *The speed of light*[1] *($c = 299,792,458\,m\,s^{-1}$) is a universal constant.*

At first sight it does not seem clear that postulating c as a universal constant may have something to do with geometry, so we better do some physics.

Consider the transit of a light ray between two points (Q and P) on space, the two events ($\mathcal{Q}$ and $\mathcal{P}$) that we are interested in are the detections of the signal at the mentioned points.

For a certain observer, the coordinates of the two events are given[2] by

$$(t_{\mathcal{Q}}, x_{\mathcal{Q}}, y_{\mathcal{Q}}, z_{\mathcal{Q}}) \qquad \text{and} \qquad (t_{\mathcal{P}}, x_{\mathcal{P}}, y_{\mathcal{P}}, z_{\mathcal{P}}) \tag{3.1}$$

for another observer, the same two eventes have coordinates

$$(\bar{t}_{\mathcal{Q}}, \bar{x}_{\mathcal{Q}}, \bar{y}_{\mathcal{Q}}, \bar{z}_{\mathcal{Q}}) \qquad \text{and} \qquad (\bar{t}_{\mathcal{P}}, \bar{x}_{\mathcal{P}}, \bar{y}_{\mathcal{P}}, \bar{z}_{\mathcal{P}}) \tag{3.2}$$

[1] Its **exact value** is 299792458 metres per second (approximately $3.00 \times 10^8\, m/s$, or 300,000 km/s (186,000 mi/s))

[2] the space coordinates are regarded as cartesian

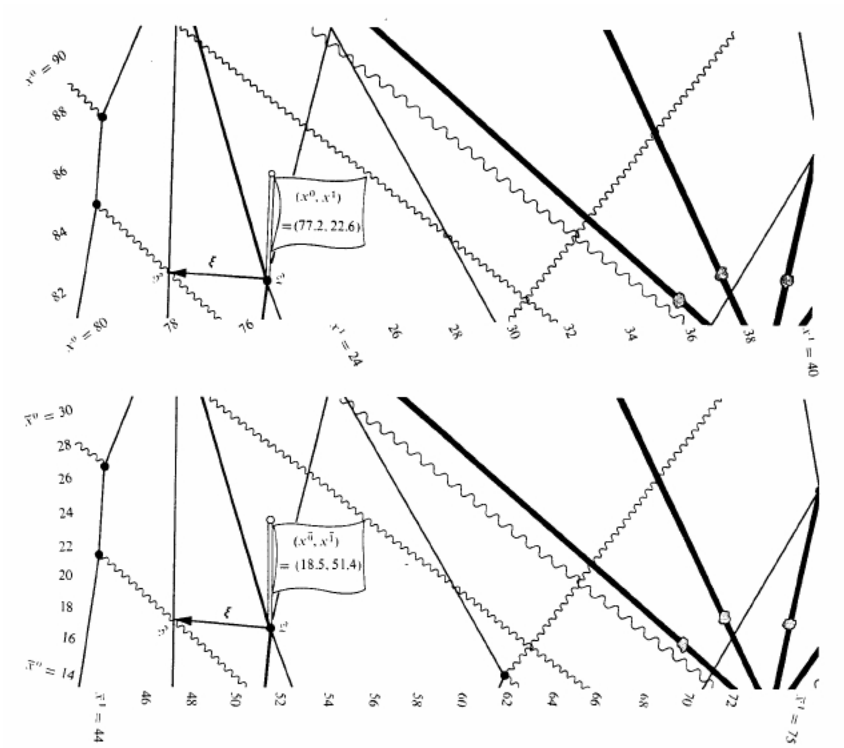


Figure 3.3: Coodinates in space time are numbers assigned to events. Different obsevers may assign different coordinates to tha same event, the question is: how do we relate coordinates given by different observers?

Since the signal is a light signal it obviously occurs that

$$c^2(t_{\mathcal{P}} - t_{\mathcal{Q}})^2 = (x_{\mathcal{P}} - x_{\mathcal{Q}})^2 + (y_{\mathcal{P}} - y_{\mathcal{Q}})^2 + (z_{\mathcal{P}} - z_{\mathcal{Q}})^2 \tag{3.3}$$

or

$$\Delta s^2 \equiv c^2(t_{\mathcal{P}} - t_{\mathcal{Q}})^2 - (x_{\mathcal{P}} - x_{\mathcal{Q}})^2 - (y_{\mathcal{P}} - y_{\mathcal{Q}})^2 - (z_{\mathcal{P}} - z_{\mathcal{Q}})^2 = 0 \tag{3.4}$$

Here is when things begin to get interesting, according to postulate 1 it **must** also happen that

$$\Delta \bar{s}^2 \equiv c^2(\bar{t}_{\mathcal{P}} - \bar{t}_{\mathcal{Q}})^2 - (\bar{x}_{\mathcal{P}} - \bar{x}_{\mathcal{Q}})^2 - (\bar{y}_{\mathcal{P}} - \bar{y}_{\mathcal{Q}})^2 - (\bar{z}_{\mathcal{P}} - \bar{z}_{\mathcal{Q}})^2 = 0 \tag{3.5}$$

Before going any further, we set a notation, the spatial coordinates (x, y, z) are to be called (x^1, x^2, x^3) or x^i, while the new time coordinate x^0 is defined as $x^0 = c\,t$, c being, of course, the speed of light. The four (space time) coordinates of any event are collectivy noted by a greek superscript (x^μ).

Going back to physics, and using the new notation we introduce the following definition

Definition 3 *Given two events $\mathcal{P}$ and $\mathcal{Q}$, the square of the 4-dimensional interval (Δs^2) between them is given by*

$$\Delta s^2 = \eta_{\mu\nu}\,\Delta x^\mu\,\Delta x^\nu \tag{3.6}$$

where $\Delta x^\mu \equiv x^\mu(\mathcal{P}) - x^\mu(\mathcal{Q})$, and $\eta_{00} = 1$, $\eta_{11} = \eta_{22} = \eta_{33} = -1$ all the other components of η being nill.

In accordance with this definition, our observation regarding the two events $\mathcal{P}$ and $\mathcal{Q}$ can be summed up by stating that the square of the 4 dimensional interval between any two events linked by a light ray equals zero for all observers.

This notion is generalized to any two events through the We may say that most of the geometric content of special relativity reduces to the following postulate:

Postulate 2 *The square of the 4-dimensional interval (Δs^2) is invariant for all inertial observers*

This is where geometry comes in, if we think about cartesian coordiantes, we immediatly realize that Eucidean geometry can be analytically derived from the invariant distance (arc length) between points as given by the Pythagorean theorem

$$\Delta\ell^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{3.7}$$

Indeed, lins are defined as curves minimizing the arc lenght, ... etc etc etc.

Postulate 2 states that space time has a geometry defined by the *pseudo euclidean metric*

$$\boldsymbol{\eta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.8}$$

Let us now see how this simple postulate brings in a notable physical consequence. Consider one clock wich is in straight uniform motion with respect an inertial observer (O) , and let x^μ be the coordinates of the clock according to O, we can set another inertial observer $(\bar{O})$ which is commoving with the watch and call the new coordinates of the watch $\bar{x}^\mu$, moreover, we will make the origin of $\bar{O}$ coincident with the clock, so at any instant, the coordinates of the clock in this system have the form $(\bar{x}^0, 0, 0, 0)$. Think now on two events, two consecutive ticks of the clock. The invariant interval between these two events is given by

$$\Delta s^2 = (\Delta\bar{x}^0)^2 = (\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 \tag{3.9}$$

it follows that

$$\Delta\bar{x}^0 = \sqrt{1 - \frac{(\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2}{(\Delta x^0)^2}}\ \Delta x^0 \tag{3.10}$$

Watching the above formula carefuly, one realizes three elementary but important and not art all obvious facts

1. $\Delta\bar{x}^0$ is c times the time that passes between two ticks of the clock as measured by $\bar{O}$
2. Δx^0 is c times the time that passes between two ticks of the clock as measured by O
3. The quotient in the square root is nothing but the square of the speed (v) of the clock as seen by O divided by c^2

Factoring out the $c's$ in front of the times, we finally get

$$\Delta\bar{t} = \sqrt{1 - \left(\frac{v}{c}\right)^2}\ \Delta t \tag{3.11}$$

The square root is smaller than 1 and thus, **¡the time runs differently for the two observers!**, in fact, it runs slower for the observer that measures the motion. This last fact may lead to paradoxes but they are just weaknesses in our understanding of special relativity that will fade away with practice.

Definition 4 *The motion of a particle is a curve in the space time, it is called the* **world line** *of the particle.*

Being a curve, the world line of a particle is parametrized by one and only one real number, i.e. $\tau \in \Re \to x^\mu(\tau)$

Remember our previous remark that in the comoving frame of a clock, the coordinates are given as $(ct, 0, 0, 0)$ consequently, its world line consists of a list of times and the invariant interval yields the time lapsed between any two events, that is the reason why the line element is also called proper time. We shall learn that the world line of massive particles can be always parametrized in proper time, the same cannot be done for massless particles such as photons, the reason for this is clear, the invariant lenght of a light ray is always $ds^2 = 0$ so the proper time is nill.

3.3 The notion of a reference frame

In this section I we will just quote two references

"The fundamental concept in physics is an event. ...

How can one determine the place and time at which an event occurs in a given inertial reference frame? Think of constructing a frame by assembling meter sticks into a cubical latticework similar to the "jungle gym" seen on playgrounds (see figure 3.4). At every intersection of this latticework fix a clock. These clocks can he constructed in any way. but are callibrated in meters of light-travel time." [3]

"It is important to realize that an 'observer' is in fact a huge information-gathering system, not simply one man with binoculars. In fact, we shall remove the human element entirely from our definition, and say that an inertial observer is simply a coordinate system for spacetime, which makes an observation simply by recording the location (x, y, z) and time (t) of any event. This coordinate system must satisfy the following three properties to be called inertial: (1) The distance between point $\mathcal{P}_1$ (coordinates (x_1, y_1, z_1) and point $\mathcal{P}_2$ (coordinates (x_2, y_2, z_2) is independent of time. (2) The clocks that sit at every point ticking off the time coordinate t are synchronized and all run at the same rate. (3) The geometry of space at any constant time t is Euclidean." [2]

3.3.1 How to syncrhonize the clocks

"Synchronizing clocks in Lattice

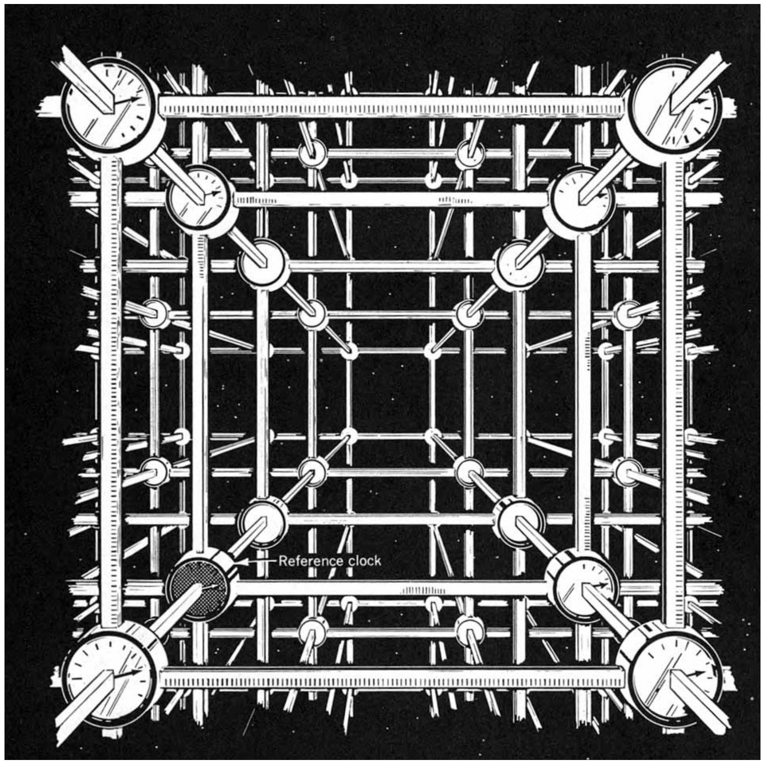


Figure 3.4: A reference frame consists of an inmense lattice of metersticks and clocks calibrated in light travel time. Remeber, $c \approx 1\, ft/ns$)

How are the different clocks in the lattice to be synchronized with one another? As follows: Pick one of the clocks in the lattice as the standard of time and take it to be the origin or an (x, y, z) *coordinate system*[3]*. Start this reference clock with its pointer at* $t = O$*. At this Instant let it send out a flash of light that spreads in all directions. Call this flash of light the* **reference flash**. *When the reference flash gets to a clock* 5 *meters away. we want that clock to read* 5 *meters of light-travel time. So an assistant* **sets** *that clock to* 5 *meters of time long before the experiment begins, holds it at* 5 *meters. and* **releases** *it only when the reference flash arrives. When assistants at all the clocks in the lattic ehave followed this procedure (each setting his clock to a time in meters equal to his own distance from the reference clock and starting it when the light flash arrives), the clocks in the lattice are said to be syncrhonized."* [3]

[3] My comment, notice that here we are talking about cartesian coordinates in a $3D$ euclidean space

3.4 Space Time Diagrams

We begin this section by saying that there is no way to overemphasize the importance and elegance of space time diagrams and thinking the following figure

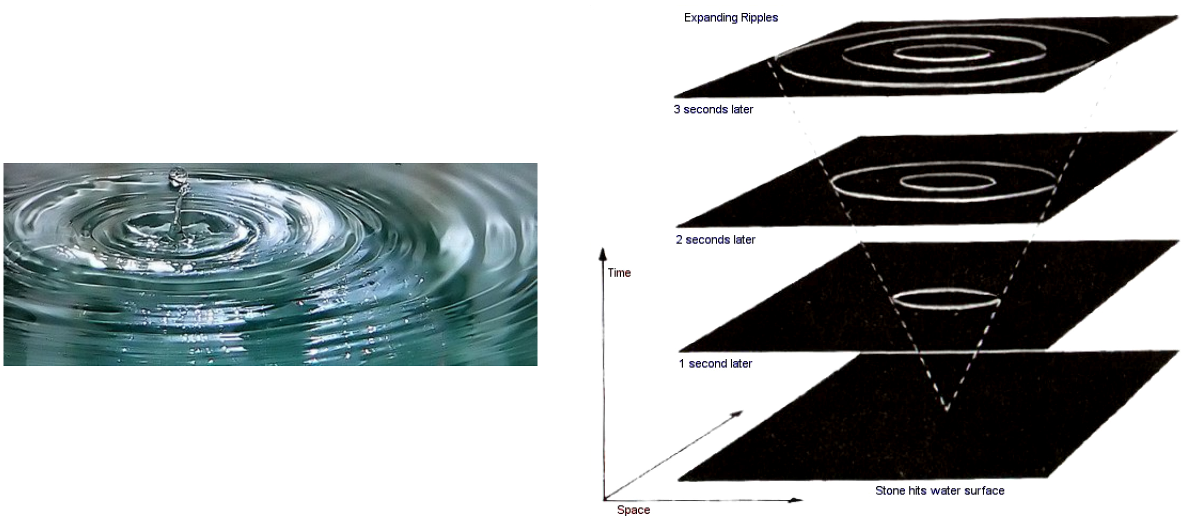


Figure 3.5: **Left**: A stone falls on a pond creating ripples, a pattern of circular wave front whose radii increase linearly with time.. **Right**: We set a vertical tme axis to catch a movie of the ripples. The planes of constant time represent the surface of the pond (the “universe”.

In most the same fashion as the ripples of figure 3.4, a light signal emanating from a point P in space at some instant of time creates a cone, the cone being the set of all events that will be illuminated by the light ray. Any object moving with speed less than that of light has a world line that remains in the interior of the light cone, whiile points that are outide the light cone cannot receive any signal sent at event[4] $\mathcal{P}$. There is also a cone coing from the past and collapsing at $\mathcal{P}$, it is the set of all light rays that wil converge at P at the same time that the flash at P is turned on.

Consider two events $\mathcal{P}$ and $\mathcal{Q}$ and the square of the interval between them

$$\Delta s^2 = c^2 \Delta t^2 - \Delta \vec{x}^2 \tag{3.12}$$

[4]recall that the event itself is the flash of light sent from P at some time

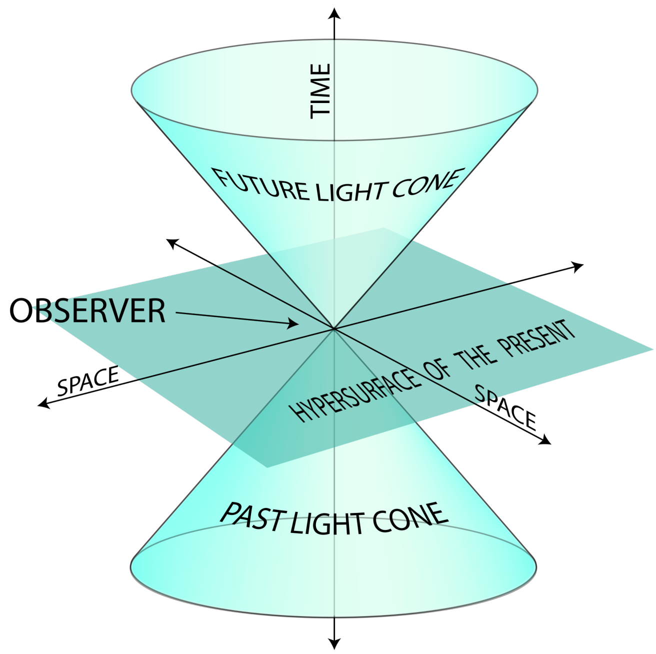


Figure 3.6: The light cone. Notice the inner and exterior regions.

If $\Delta s^2 = 0$ we say that the events are light-like separated, if $\Delta s^2 > 0$ we say their separation is time-elike, and if $\Delta s^2 < 0$ we say that they have a time-like separation. The light cone emanating from any of the to events enlightens this terminology, when the separation is time like both events can be connected by a light signal, when they are time like separated, a message sent with speed lower than that of light may reach one event from the other, finally, if the searaion is space like, no signal can be sent between both events since it would be super luminar.

A space time diagram is just a graphical representation of space time, which mimics figures 3.4 and 3.6, a typical diagram is two dimensional and shows time and one space coordinate (t, x) according to a particular observer.

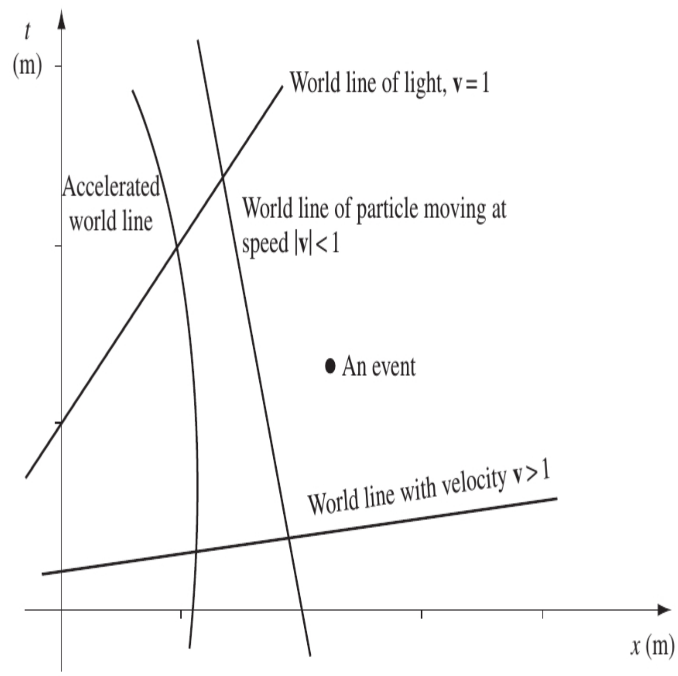


Figure 3.7: A space time diagram showing several world lines. Vertical lines are objects sitting in a fixed position, the time axis, in particular is the world line of the origin itself. Horizontal lines are lines of simultaneity. The axes are calibrated in natural units so any light signal looks as a line making a 45^0 angle with the time axis.

3.4.1 How to draw the coordinate lines of a "moving frame"

The problem at hand is this, we want to superimpose the coordinate lines of two frames, one, the lab frame (O) is at rest, while the other ($\bar{O}$), is moving with speed v, say along the positive x axis, for simplicity we add the hipothesis that that the origins of the two frames coincide at $t = \bar{t} = 0$.

To begin the construction, we first notice that the $\bar{t}$ axis is nothing but the world line of the origin of $\bar{O}$ so it must be drawn as a straight line that passes through the origin of O makes an angle $\theta = arctan(v/c)$ with the t axis

To build the $\bar{x}$ axis we first consider a simple *gedankenexperiment*. Imagine a lamp located at the origin of the $\bar{x}$ axis, and a mirror located at some position $\bar{d}_M$ to the right of the origin

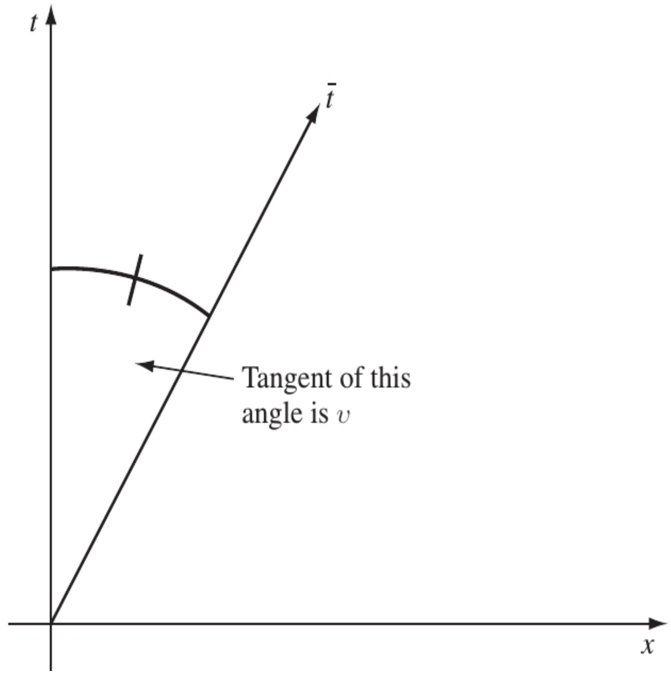


Figure 3.8: The $\bar{t}$ axis as seen from the rest frame. The $\bar{t}$ axis can be thought of as the worldline of $\bar{O}$

of $\bar{O}$. A light ray is sent form the origin at some instant in the past. Clearly, the light ray shall be reflected by the mirror and will reach the origin of the $\bar{x}$ axis again in the future. notice that both the emission and detection times are symmetrical with respect to the origin of $\bar{t}$.

In order to end the construction, i.e. build the $\bar{x}$ axis as seen by O we carefully think of the previous paragraph and its related figure 3.9. Take any point (P) at the negative $\bar{t}$ axis and look for its symmetric (with respect to $\bar{t} = 0$ point (P_S), draw a wavy (photon) line travelling to the future, from P_S draw a photon line travelling to the past. The point of intersection of these two world lines is just the event of th reflection (Q) and thus its spatial position (the position of the mirror) is along the $\bar{x}$ axis, the lone that joins Q with the origin is therefore the $\bar{x}$ axis along the simultaneity line $\bar{t} = 0$ see figure 3.10

The construction we have just finished captures the essence of relativity, simultaneity is no longer possible. The lines of constant time are completely different for observers in motion with respect with each other. Indeed, as seen by the O system (lab or rest system) the simultaneity

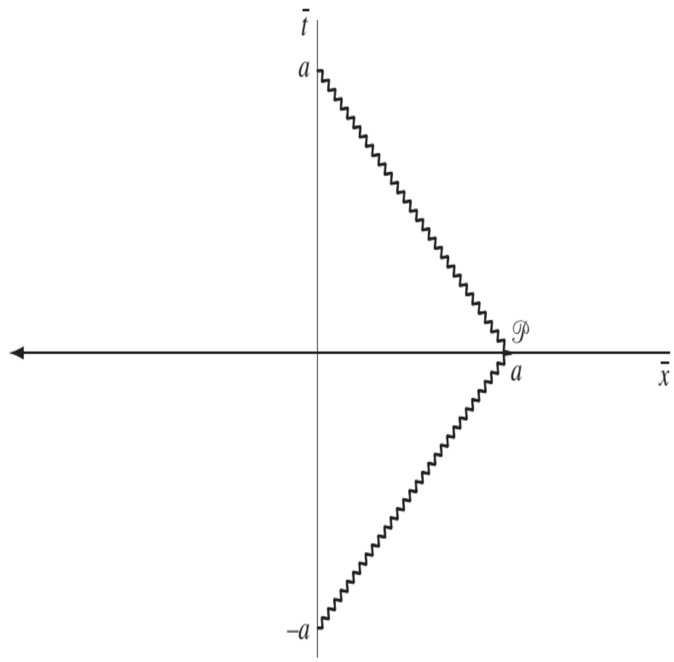


Figure 3.9: Emission and detection of a light ray in the $\bar{O}$ frame.

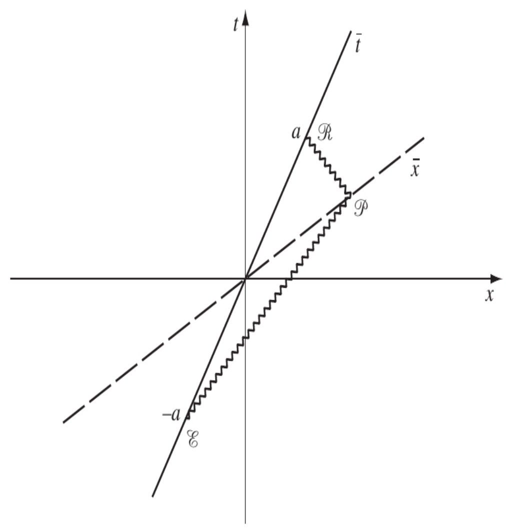


Figure 3.10: The coordinate axes of the $\bar{O}$ system as seen form O.

lines are horizontal while those of $\bar{O}$ are not since they are paralell to the $\bar{x}$ axis see figure 3.11

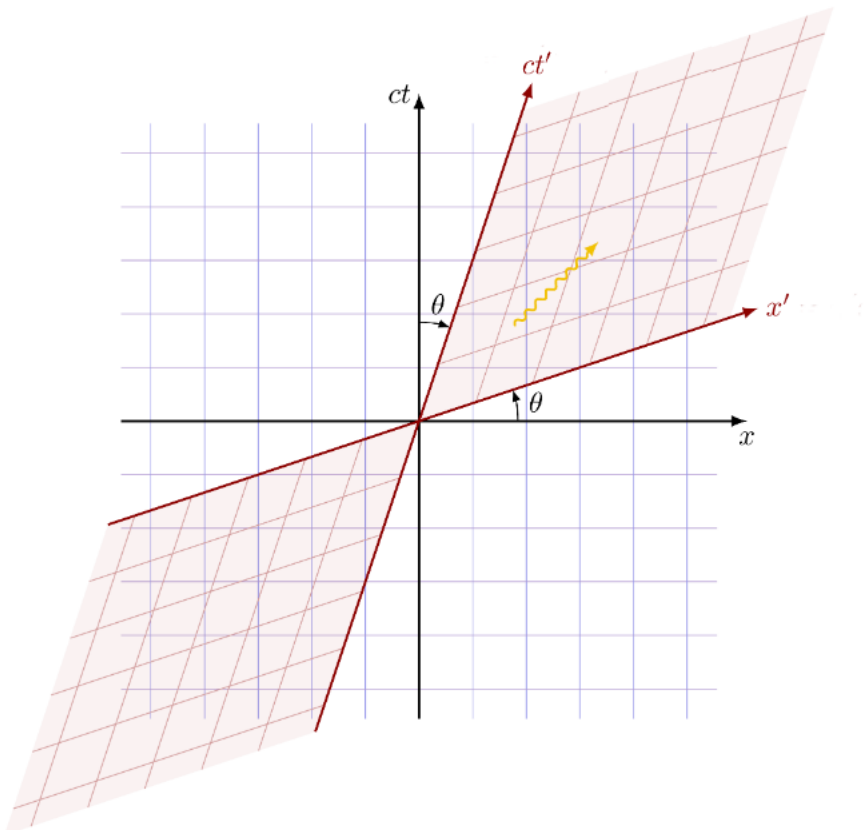


Figure 3.11: The coordinate lines of the LAB system are black, the red lines are coordinate lines for the moving frame. Simultaneity lines for the primed system are parallel to the x' axis. In relativity, simultaneity is completely lost as a concept to be shared between observers.

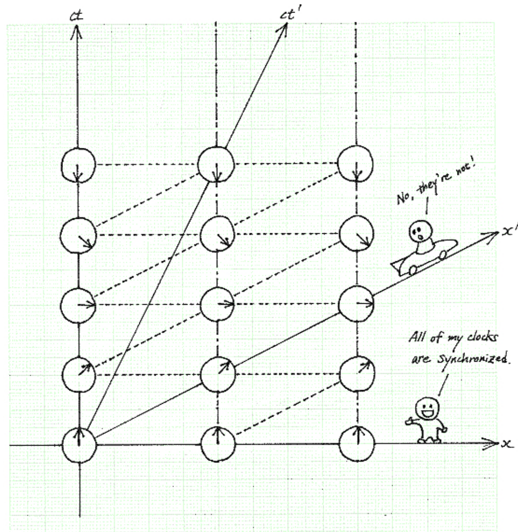


Figure 3.12: Clocks along horizontal lines are perfectly synchronized for the LAB system, they are not for the moving observer

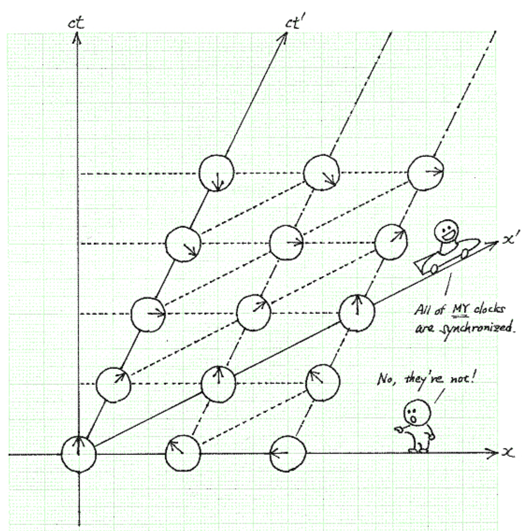


Figure 3.13: Clocks along lines parallel to the x' axis are perfectly synchronized as seen from the moving observer, the resting observer totally disagrees

Chapter 4

Lorentz Transformations

4.1 Building $2D$ lorentz transformations

This section closely follows the discussion presented in sections 2.1 and 2.1.1 We begin by working on what is usually referred o as $1+1$-*dimensional space-time*, i.e. one time and one space dimensions. In other words, we imagine doing experiments along a line and taking notes of the times at which the results are observed, in the usual language, we would say that we limi ourselves to think of events ocurring along the x axis, to make this a little more precise, we give the following

Definition 5 *The* $2D$ *"Minkowski" space time* ($\mathcal{M}^2$) *is the set of all events ocuuring along a line, we regard it as a two dimensional manifold to which an inertial observer assigns coordinates (sets of pairs) in such a way that event* $\mathcal{S}$ *is assigned the pair*

$$\mathbf{x}(\mathcal{S}) = \begin{pmatrix} x^0(\mathcal{S}) \\ x^1(\mathcal{S}) \end{pmatrix} \tag{4.1}$$

As with $\Re^2$, the tangent space $T_{\mathcal{P}}\mathcal{M}^2$ of $\mathcal{M}^2$ at $\mathcal{P}$ is the vector space of all dispacements based at $\mathcal{P}$ which consissts on the following, for each event $\mathcal{Q}$ a contravariant vector is built with coordinates:

$$\Delta\mathbf{x}^\mu = \mathbf{x}^\mu(\mathcal{Q}) - \mathbf{x}^\mu(\mathcal{P})\,, \tag{4.2}$$

The tangent bundle $T_{\mathcal{P}}\mathcal{M}^2$ is the union of all tangnet spaces.

The square of the interval (or proper time) between two events is given by the (pseudo)-metric

$$\Delta s^2 = (\Delta x^0)^2 - (\Delta x^1)^2 = \Delta \mathbf{x}^t \, \boldsymbol{\eta} \, \Delta \mathbf{x} \,, \tag{4.3}$$

Of particular interest are the Lorentz transformations, they are simply are matrices

$$\begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 \\ \Lambda^1{}_0 & \Lambda^1{}_1 \end{pmatrix} \tag{4.4}$$

such that the transformation

$$\Delta \mathbf{x} \rightarrow \Delta \bar{\mathbf{x}} \equiv \boldsymbol{\Lambda} \, \Delta \mathbf{x} \tag{4.5}$$

leaves the square of the proper time

$$\Delta s^2 = \Delta \mathbf{x}^T \boldsymbol{\eta} \, \Delta \mathbf{x} \,, \tag{4.6}$$

invariant, i.e.

$$\Delta s^2 = \Delta \bar{\mathbf{x}}^T \boldsymbol{\eta} \, \Delta \bar{\mathbf{x}} = \Delta \mathbf{x}^T \boldsymbol{\eta} \, \Delta \mathbf{x} \,. \tag{4.7}$$

We want to explore the set of infinitesimal Lorentz transformations, i.e. the set of all $\boldsymbol{\Lambda}$ matrices clode to the identity which allow condition 4.7, to achieve this goal we write $\boldsymbol{\Lambda}$ as

$$\boldsymbol{\Lambda} = \mathbf{1} + \boldsymbol{\epsilon} \,. \tag{4.8}$$

where $\boldsymbol{\epsilon}$ is a matrix with small (infinitesimal) entries only, namely

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon^0{}_0 & \epsilon^0{}_1 \\ \epsilon^1{}_0 & \epsilon^1{}_1 \end{pmatrix} \tag{4.9}$$

. Condition 4.7 is equvalent to the identity

$$\boldsymbol{\Lambda}^T \boldsymbol{\eta} \, \boldsymbol{\Lambda} = \boldsymbol{\eta} \,, \tag{4.10}$$

which, upon substitution and retaining first order terms only, implies

$$\begin{aligned} \boldsymbol{\Lambda}^T \boldsymbol{\eta} \, \boldsymbol{\Lambda} &= [\mathbf{1} + \boldsymbol{\epsilon}]^T \, \boldsymbol{\eta} \, [\mathbf{1} + \boldsymbol{\epsilon}] \\ &\approx \boldsymbol{\eta} + \boldsymbol{\epsilon}^T \boldsymbol{\eta} + \boldsymbol{\eta} \boldsymbol{\epsilon} \,, \end{aligned} \tag{4.11}$$

therefore, the infinitesimal matrices must satisfy

$$\boldsymbol{\epsilon}^T \boldsymbol{\eta} + \boldsymbol{\eta}\boldsymbol{\epsilon} = 0\,, \tag{4.12}$$

$$\begin{pmatrix} \epsilon^0{}_0 & \epsilon^1{}_0 \\ \epsilon^0{}_1 & \epsilon^1{}_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \epsilon^0{}_0 & \epsilon^0{}_1 \\ \epsilon^1{}_0 & \epsilon^1{}_1 \end{pmatrix} = \mathbf{0}\,, \tag{4.13}$$

or

$$\begin{pmatrix} \epsilon^0{}_0 & -\epsilon^1{}_0 \\ \epsilon^0{}_1 & -\epsilon^1{}_1 \end{pmatrix} + \begin{pmatrix} \epsilon^0{}_0 & \epsilon^0{}_1 \\ -\epsilon^1{}_0 & -\epsilon^1{}_1 \end{pmatrix} = \mathbf{0}\,, \tag{4.14}$$

Which in turn implies:

$$\boldsymbol{\epsilon} = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix} \tag{4.15}$$

or

$$\boldsymbol{\Lambda} = \mathbf{1} + \epsilon\, \mathbf{T}\,, \tag{4.16}$$

$\mathbf{T}$ being the 2×2 transposition matrix

$$\mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.17}$$

As happened with rotations, the product of two of two infinitesima Lorentz transformations is itself a Lorentz transformation ($\boldsymbol{\Lambda}(\varepsilon_1)\boldsymbol{\Lambda}(\varepsilon_1) = \boldsymbol{\Lambda}(\varepsilon_3)$, with $\epsilon_3 = \epsilon_1 + \epsilon_2$, this implies that $[\boldsymbol{\Lambda}(\varepsilon)]^{-1} = \boldsymbol{\Lambda}(-\varepsilon)$

We have a simple lemma for powers of the generator,

Lemma 2 $\mathbf{T}^2 = \mathbf{1}$, $\mathbf{T}^3 = \mathbf{T}$ *and* $\mathbf{T}^4 = \mathbf{T}$

The proof is left as an excercise

Just as with did with rotations on the plane, see section 2.1.1, we may build a finite Lorentz transformation (connected to the identity) by iterating a huge number of infinitesimal ones to get:

$$\boldsymbol{\Lambda}(\theta) \equiv \lim_{n\to\infty} \boldsymbol{\Lambda}(\theta, n) = e^{\theta \mathbf{T}} \tag{4.18}$$

The exponential is found using the power series

$$e^{\theta \mathbf{T}} = \mathbf{1} + \theta \mathbf{T} + \frac{1}{2!}\left(\theta \mathbf{T}\right)^2 + \frac{1}{3!}\left(\theta \mathbf{T}\right)^3 + \frac{1}{4!}\left(\theta \mathbf{T}\right)^4 + \frac{1}{5!}\left(\theta \mathbf{T}\right)^5 + \dots \tag{4.19}$$

which, by virtue of Lemma 2 finally yields

$$e^{\theta \mathbf{T}} = \left[1 + \frac{1}{2!}\theta^2 + \frac{1}{4!}\theta^4 + \dots\right]\mathbf{1} + \left[\theta + \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 + \dots\right]\mathbf{T} \tag{4.20}$$

which in turn is nothing but

$$e^{\theta \mathbf{T}} = cosh\theta\,\mathbf{1} + senh\theta\,\mathbf{T} = \begin{pmatrix} cosh\theta & senh\theta \\ senh\theta & cosh\theta \end{pmatrix} \tag{4.21}$$

4.1.1 The physical interpretation of Lorentz transformations

We frst recall the following useful formulas

$$cosh^2\theta - sinh^2\theta = 1 \tag{4.22}$$

$$tanh\theta = \frac{sinh\theta}{cosh\theta} \tag{4.23}$$

From which we find

$$cosh\theta = \frac{1}{\sqrt{1 - tanh^2\theta}}\,, \tag{4.24}$$

and

$$sinh\theta = \frac{tanh\theta}{\sqrt{1 - tanh^2\theta}} \tag{4.25}$$

Which allows us to reexpress Lorentz transformations as

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{1}{\sqrt{1-tanh^2\theta}} & \frac{tanh\theta}{\sqrt{1-tanh^2\theta}} \\ \frac{tanh\theta}{\sqrt{1-tanh^2\theta}} & \frac{1}{\sqrt{1-tanh^2\theta}} \end{pmatrix} \tag{4.26}$$

Or defining $tanh\theta \equiv -\beta$

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix} \tag{4.27}$$

Let us go back to small transformations, i.e. take the limit small θ which allows us to take, $tanh\theta \approx \theta$, undern such condition

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \tag{4.28}$$

According to this, when transforming the unbarred coordinates to the barred ones,

$$\begin{aligned} c\,\bar{t} &= c\,t - \beta\,x \\ \bar{x} &= -\beta\,c\,t + x \end{aligned} \tag{4.29}$$

If we look at these formulas with some care we will note that taking

$$\beta = \frac{v}{c}\,, \tag{4.30}$$

and interpret v appropriately we get that this transformation of coordinates is just a Galilean transformation between two observers, $\mathcal{O}$ and $\bar{\mathcal{O}}$, the latter moving along the positive x axis of former. This is the reason that justifies the name *velocity parameter* for the hyperbolic angle θ

In brief, define

$$tanh\theta = -\beta = -\frac{v}{c} \tag{4.31}$$

to recover that the Lorentz transformations between an observer moving to the right along the x axis with speed v with respect to an observer at rest is just

$$\mathbf{\Lambda} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix}\,, \tag{4.32}$$

with

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{4.33}$$

According to this, the explicit Lorentz boost is given by

$$\begin{aligned} \bar{t} &= \gamma(t - \beta\,x) \\ \bar{x} &= \gamma(x - v\,t) \end{aligned} \tag{4.34}$$

As a simple exercise, prove that the inverse Lorentz Boost is obtained by the change $v \to -v$.

The results of this section can be painlessly generalized, and we shall do it as a guided exercise.

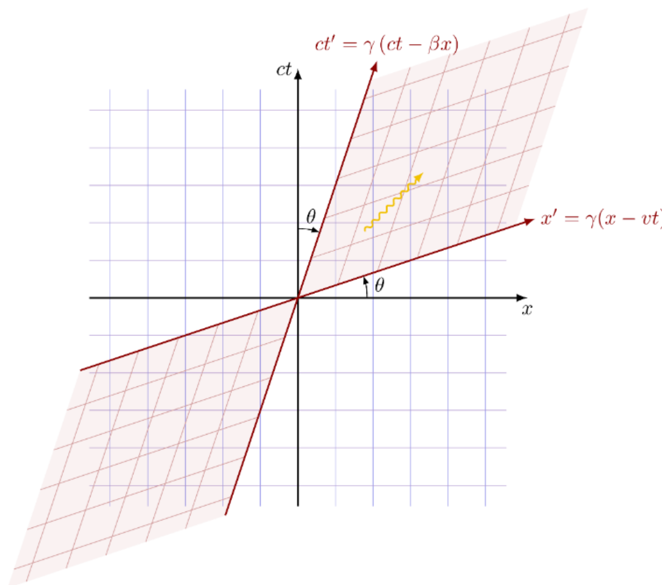


Figure 4.1: Coordinate lines of the right moving frame are built by boosts

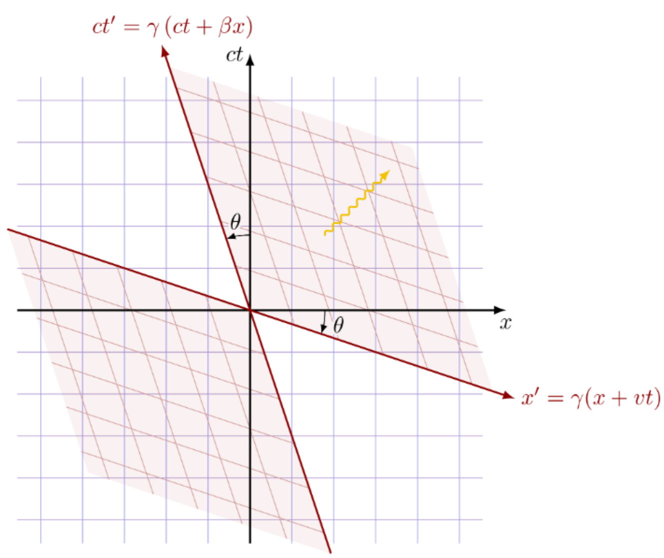


Figure 4.2: Coordinate lines of a left moving frame as seen from the LAB

4.1.2 Long problem: the $3+1$-space time

This section is thought of as a long guided problem on reading and summarizing..

1. Define $\mathcal{M}^4$ and assign coordinates to events.
2. Define the tangent space at an event $T_{\mathcal{P}}\mathcal{M}^4$
3. Define the tangent bundle $T\mathcal{M}^4$
4. Define the metric for $T_{\mathcal{P}}\mathcal{M}^4$
5. A curve on $\mathcal{M}^4$ is a map $x : I \in \Re \rightarrow \mathcal{M}^4$, what is a world lne?

4.2 Guided Problem, $4D$ Lorentz transformations, a short visit to $SO(1,3)$

In this section, You(the reader) are expected to investigate the group of tranformations that leave the Minkowski metric invariant, feel free to rely on section 2.2.1

1. Use what you have learned to guess the form of a Lorentz transformation changing from the lab system to a referential which travels at speed v along the positive y axis of the lab system.
2. What about the inverse of such Lorentz transformation?
3. Do the same for an observer travelling along the z axis.
4. Give a heuristic argument to support that a general boost might be written as

$$\bar{x}^0 = \gamma\left[x^0 - \vec{\beta}.\vec{x}\right] \tag{4.35}$$

$$\vec{\bar{x}} = \vec{x} + \frac{(\gamma - 1)}{\beta^2}\,(\vec{\beta}.\vec{x})\,\vec{\beta} - \gamma\,\vec{\beta}\,x^0 \tag{4.36}$$

5. Show that, if you think on infinitesimal transformations you get the following set of six generators[1]

$$\boldsymbol{S}_1 = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} , \boldsymbol{S}_2 = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} , \boldsymbol{S}_3 = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.37}$$

$$\boldsymbol{K}_1 = i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \boldsymbol{K}_2 = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \boldsymbol{K}_3 = i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{4.38}$$

6. Give an interpretation of the 3 **S** generators.

7. Do the same for the **K** generators.

8. Show that these six generators satisfy the following algebra:

$$[\boldsymbol{S}_i, \boldsymbol{S}_j] = i\, \epsilon_{ijk}\, \boldsymbol{S}_k \tag{4.39}$$

$$[\boldsymbol{S}_i, \boldsymbol{K}_j] = i\, \epsilon_{ijk}\, \boldsymbol{K}_k \tag{4.40}$$

$$[\boldsymbol{K}_i, \boldsymbol{K}_j] = -i\, \epsilon_{ijk}\, \boldsymbol{S}_k \tag{4.41}$$

9. Formula 4.39 has a clear meaning, can you state it?.

10. Formula 4.41 has an important physical interpretation, try to figure it out. Define[2]:

$$\boldsymbol{A}_i = \frac{\boldsymbol{S}_i + i\, \boldsymbol{K}_i}{2} \tag{4.42}$$

$$\boldsymbol{B}_i = \frac{\boldsymbol{S}_i - i\, \boldsymbol{K}_i}{2} \tag{4.43}$$

[1]Notice that the form of these generators is such that they become hermitian, this has interest in quantum mechanics and affects the formulas by changing the arguments in the exponentials of matrices by the introduction of an extra i we should writte, for example: $exp(i\, \theta^i\, \mathbf{K}_i)$

[2]this is just a new basis for the algebra

Show that in this new basis, the Lorentz algebra takes the form

$$
\begin{aligned}
[\boldsymbol{A}_i, \boldsymbol{A}_j] &= i\,\epsilon_{ijk}\,\boldsymbol{A}_k \\
[\boldsymbol{B}_i, \boldsymbol{B}_j] &= i\,\epsilon_{ijk}\,\boldsymbol{B}_k \\
[\boldsymbol{A}_i, \boldsymbol{B}_j] &= 0
\end{aligned}
$$

Interpret this result

Chapter 5

Simple Consequences

While reading this chapter keep in mind the fundamental postulate of special relativity:

The speed of light: c = 299792458 m/s is the same for all inertial observers

5.1 Simultaneity, a Relative Concept

We have already discussed this topic, but such weird consequences of relativity as the lost of a "very old and well known concept" as simultaneity must be readdressed several times. We will do it graphically through our favourite tools, space time diagrams.

In figure 5.1 events A and B lie on the same horizontal line of the lab system and are therefore simultaneous for the "resting observer" $\mathcal{O}$. In the moving or boosted frame, A lies along the green dashed line while B does along the black one, therefore, there is a time difference $\Delta t' > 0$ between both events which makes them non simultaneous for $\mathcal{O}'$. In figure 5.2 the situation is reversed, events A and B lie along the thick red dashed line which is a line of simultaneity for the boosted frame, and both events are obviously nos simultaneous for the LAB frame, in fact, in the LAB frame B happens with a delay Δt.

The scenario shown in figure 5.3 is mind-boggling to say the least, the time ordering of events A and B is reversed for each observer.

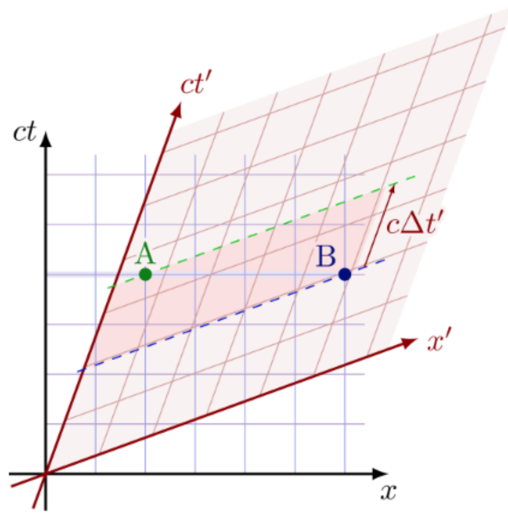


Figure 5.1: Events A and B are simultaneous in the rest frame S, but in the boosted frame S', B happens before A.

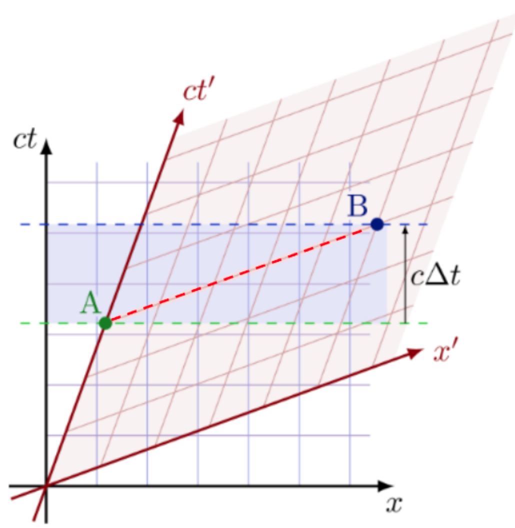


Figure 5.2: Events A and B are simultaneous in the boosted frame S', but in the rest frame S, A happens before B..

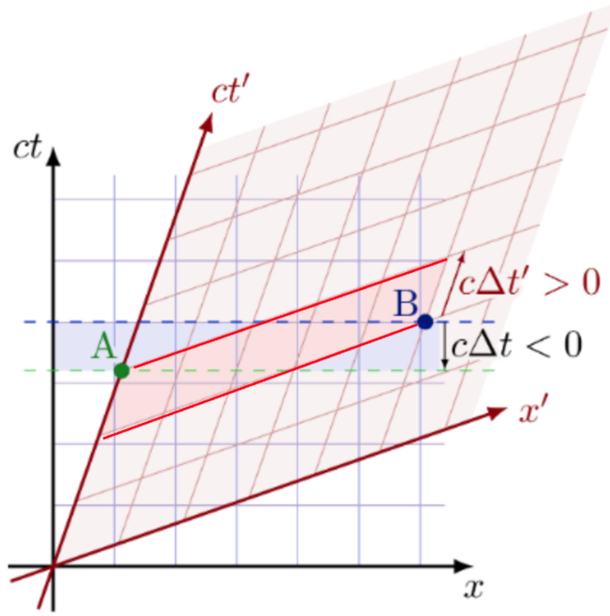


Figure 5.3: Relativity of simultaneity. The Two observers experience events at in different order: In frame S, A happens before B, while in frame S', B happens before A.

5.2 Time Dilation

5.2.1 Spacetime Diagram Approach

Consider an object moving in straight line with speed v with respect to the LAB frame $\mathcal{O}$ certain reference frame $\bar{\mathcal{O}}$. Think of two events $\mathcal{A}$ and $\mathcal{B}$, the coordinates of which in the comoving frame of the object are $\mathcal{A} = (\bar{\tau}_{\mathcal{A}}, 0, 0, 0)$, $\mathcal{B} = (\bar{\tau}_{\mathcal{B}}, 0, 0, 0)$.

In the lab frame the corrdinates are $\mathcal{A} = (\tau_{\mathcal{A}}, x_{\mathcal{A}}, y_{\mathcal{A}}, z_{\mathcal{A}})$ and $\mathcal{B} = (\tau_{\mathcal{B}}, x_{\mathcal{B}}, y_{\mathcal{B}}, z_{\mathcal{B}})$ where without fearof expressing the obvious, since the particle is moving, at least one the space coordinates of the events are necessarily different.

The proper time interval is invariant and so

$$\Delta s^2 = c^2 \Delta\bar{\tau}^2 = c^2 \left(\bar{\tau}_{\mathcal{A}} - \bar{\tau}_{\mathcal{B}}\right)^2 = c^2 \left(\tau_{\mathcal{A}} - \tau_{\mathcal{B}}\right)^2 - \left(x_{\mathcal{A}} - x_{\mathcal{B}}\right)^2 - \left(y_{\mathcal{A}} - y_{\mathcal{B}}\right)^2 - \left(z_{\mathcal{A}} - z_{\mathcal{B}}\right)^2 \qquad (5.1)$$

or

$$\Delta\bar{\tau}^2 = \left[1 - \frac{\Delta x^2}{c^2 \Delta\tau^2}\right] \Delta\tau^2 \qquad (5.2)$$

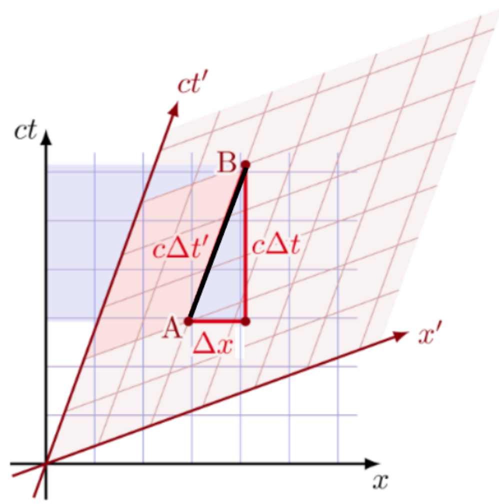


Figure 5.4: Events A and B occur at the same space point in the moving frame, where the they are separated by a time interval $\Delta t'$, as seen by the LAB system both events happen at different space points ($\Delta x > 0$) and are separated by a time interval Δt. The time separation between A and B is longer in S' than in S; $\Delta t' > \Delta t$.

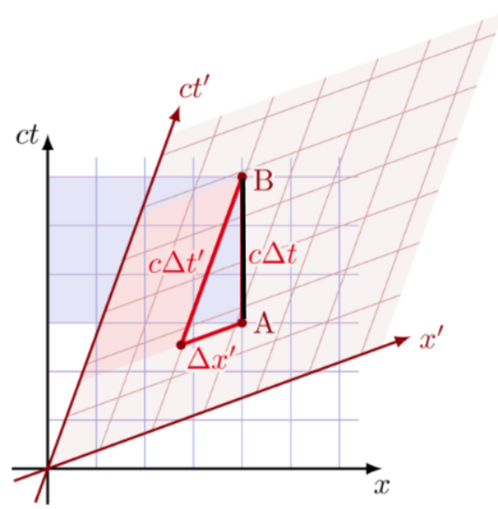


Figure 5.5: Time dilation in the boosted frame S' for events at a fixed point of space in frame S. The time separation between A and B is longer in S' than in S; $\Delta t' > \Delta t$.

where

$$
\begin{aligned}
\Delta\bar{\tau} &= \bar{\tau}_{\mathcal{A}} - \bar{\tau}_{\mathcal{B}} \\
\Delta\tau &= \tau_{\mathcal{A}} - \tau_{\mathcal{B}} \\
\Delta x^2 &= (x_{\mathcal{A}} - x_{\mathcal{B}})^2 - (y_{\mathcal{A}} - y_{\mathcal{B}})^2 - (z_{\mathcal{A}} - z_{\mathcal{B}})^2 \ .
\end{aligned}
\tag{5.3}
$$

By definition, the quotient

$$\frac{\Delta x^2}{c^2 \Delta\tau^2}$$

is the square of the velocity of the object as measured by the LAB observer $\mathcal{O}$, in brief

$$\Delta\bar{\tau} = \sqrt{1 - \left(\frac{v}{c}\right)^2}\, \Delta\tau \tag{5.4}$$

the meaning of this identity is clear, since the value of the square root is always less than 1, the quantity $\delta\tau$ which is the time lapse between the two events according to $\mathcal{O}$ must be larger than $\delta\bar{\tau}$ the time lapse in the comoving system, i.e. the proper time.

The situation depicted in Figure 5.4 exactly reproduces the situation we have just described. Indeed, events A and B are located at the same space point in the moving system S' and are separated by a time interval $\Delta t'$. In the lab system both events are separated in space by an amount Δx and Δt in time. Here is where the geometry of space time comes in, the relation between those intervals is

$$(\Delta t')^2 = (\Delta t)^2 - \frac{(\Delta x)^2}{c^2}\, , \tag{5.5}$$

which clearly implies $\Delta t > \Delta t'$, i.e. the LAB frame reads a dilated time interval.There is something confusing, though, our natural tendency is to say the opposite, this is so because we are using euclidean geometry and see a right triangle where the hypotenuse is $c\Delta t'$ which implies that this time lag is bigger than $c\Delta t'$, but, and this is a big and fundamental BUT, the geometry we are dealing with is not euclidean, and in this geometry, the Pythagorean theorem is to be substituted by the interval length also referred to as the metric.

Figure 5.5 is very important indeed. Reference frames should be inmaterial, and therefore the predictions of special relativity should be symmetrical[1]. Events A and B are located at the same x coordinates in the LAB system where the time delay Δt between them is clear. Since

[1]from the point of view of the comoving frame, it is the lab system what is actually moving.

the simultaneity lines in the x' axis and the moving frame time runs along lines parallel to the t' axis, we must draw the triangle shown in the figure where it is absolutely clear that $\Delta t' > \Delta t$, i.e. time runs slowly as seen by S'.

5.2.2 Standard Approach

The usual approach to explain the time dilation effect consists in explaining what happens in figure 5.6.

A light pulse is emitted from a point in one wall of the train car, reflected in the mirror located in the other wall of the train car and received back at the detector.

For an observer fixed in the train car (comoving observer), the two way travel time of the signal is obviously

$$\Delta t = 2\frac{d}{c} \tag{5.6}$$

where d is the width of the rail car.

On the other hand, since the train car is moving, the path followed by light signal as seen by an observer fixed on the train station (LAB observer) is an isosceles triangle of "height" d whose base is the distance traveled by the emitter-receiver system during the time ($\Delta t'$) taken by the signal to reach the receiver. If we call D one of the two equal sides of the triangle then,

$$\Delta t' = 2D/c\,. \tag{5.7}$$

Now, during the two way travel time of the light signal, the emitter-receiver system moves a distance $s = v\Delta t'$ (the base of the isosceles triangle).

Given the geometry, we can think of half the isosceles triangle which is clearly a right triangle with catheti d and $s/2$ and hypotenuse $D = \sqrt{d^2 + (s/2)^2}$.

Consequently, the two way travel time of the signal according to the LAB observer is

$$\Delta t' = 2\frac{\sqrt{d^2 + (v\Delta t'/2)^2}}{c}\,. \tag{5.8}$$

Equation 5.6 implies

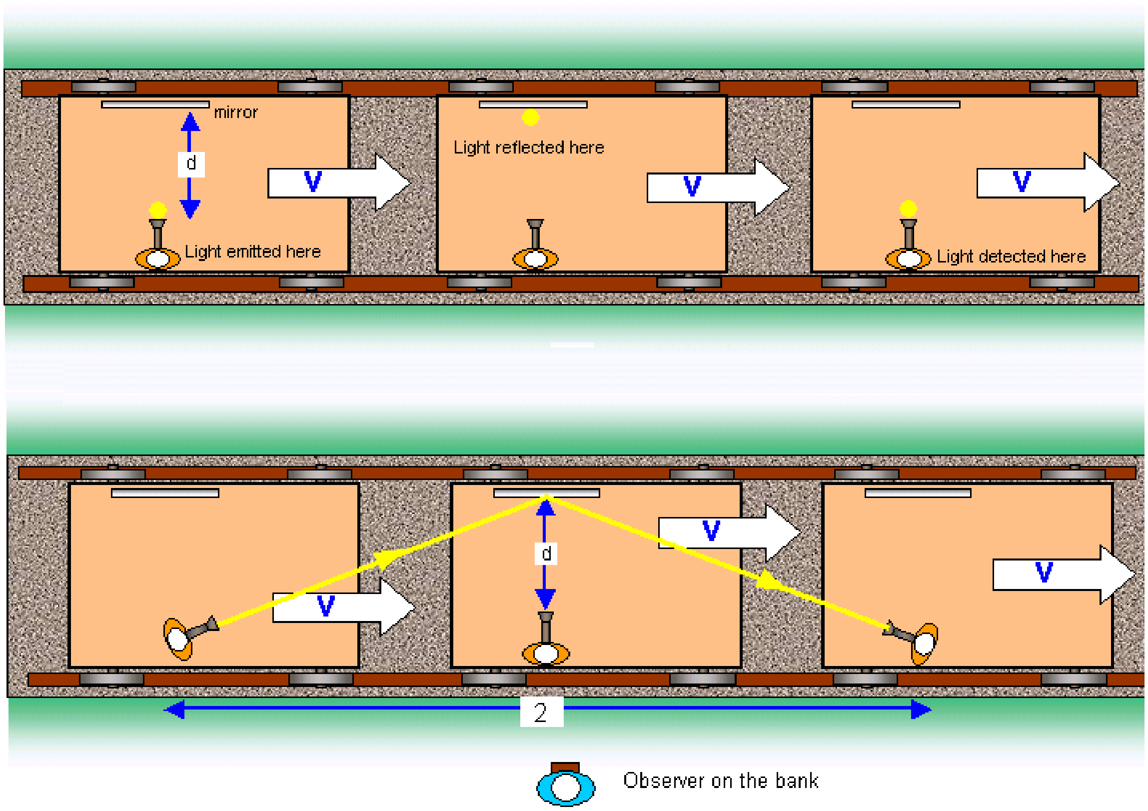


Figure 5.6: An emitter receiver system is located inside a train car moving with velocity **v** along the rails. train car

$$
\begin{aligned}
\Delta t' &= 2\frac{\sqrt{(c\Delta t/2)^2 + (v\Delta t'/2)^2}}{c} = \\
&= 2\sqrt{(\Delta t/2)^2 + (v\Delta t'/2c)^2} = \\
&= \sqrt{(\Delta t)^2 + (v\Delta t'/c)^2}\,,
\end{aligned}
\tag{5.9}
$$

which after a little bit of algebra gives us

$$
\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}{}^2}}\,, \tag{5.10}
$$

the formula for time dilation.

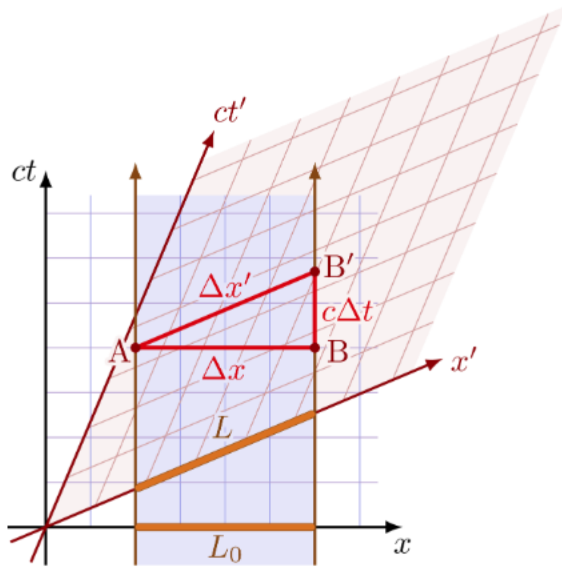


Figure 5.7: A rod at rest in the LAB system, has lenght L_0 with respect to S, its lenght with respecto to the movimg frame is $L < L_0$

5.3 Length Contraction

To talk about length contraction we need to address the concept of length measurement. In order to measure a length we need to place a -say- ruler in such a way that the extremes of the object to be measured, a rod for example. Then we should made sure that photons coming from those extreme points reach our eyes simultaneously.

Consider a rod positioned[2] along the x axis and let x_1 and x_2 be the coordinates of its ends in the lab system Figure 5.7. Consider two observers in a moving frame whom after careful synchronization will be located at the rod's ends at a prepared time t', let the positions of those two observers be x'_1 and x'_2.

Given the nature of the experiment being carried out by the moving observers, they may assure that the length of the rod in the moving frame is given by

$$L = x'_2 - x'_1 \,. \tag{5.11}$$

[2]The rod is at rest in the LAB system

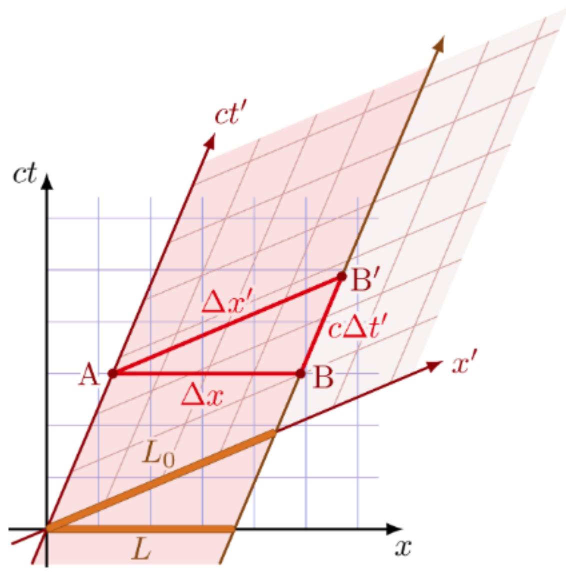


Figure 5.8: Length contraction of a rod moving in frame S. The rod is shorter in the frame S', because it is moving relative to this frame.

Now, the length of the rod in the lab system is clearly

$$L_0 = x_2 - x_1 \,, \tag{5.12}$$

a simple Lorentz transformation tells us that

$$x_2 - x_1 = \gamma(x_2' + \beta t') - \gamma(x_1' + \beta t') = \gamma(x_2' - x_1')\,, \tag{5.13}$$

meaning that

$$L_0 = \gamma L\,, \tag{5.14}$$

i.e.

$$L = \sqrt{1-\beta^2}\, L_0 \tag{5.15}$$

looks shorter in the moving frame.

Notice that the length of the $\Delta x'$ (or L if you will) side of the triangle appear longer than side Δx between points A and B. Once more we are been deceived by our euclidean experience.

Let us now turn our attention to the case that the road moves with respect to the LAB frame (Fig 5.8). As the rod moves, we have a hard time lining up a meter stick next to the rod to make any measurements. Instead, we watch as the rod passes a fixed point and record the time interval from the time the first end passes the point to the time the back end does. The time obtained is

$$\Delta t = \frac{L}{v}\,, \tag{5.16}$$

where L is the length of the moving rod as recorded by the stationary observer. The observer in the rest frame of the rod would record a time of

$$\Delta t' = \frac{L_0}{v}\,. \tag{5.17}$$

However, we can use time dilation to relate the times. The time measured by the stationary observer is measured by focusing on a fixed point in space. So, Δt is the proper time in system S. It is shorter than that measured in S'. Thus,

$$\Delta t' = \gamma \Delta t\,. \tag{5.18}$$

Note that the time measured in frame S' cannot be a proper time as the observer has to measure two different times in two different locations, as we will see using the Minkowski diagrams.

Continuing with the computation, we have so far

$$\Delta t' = \gamma \Delta t\,, \quad \Delta t = \frac{L}{v}\,, \quad \Delta t' = \frac{L_0}{v}\,. \tag{5.19}$$

Eliminating the time variables, we are left with the length contraction equation:

$$L = \frac{L_0}{\gamma}\,. \tag{5.20}$$

This indicates that the proper length is larger than the length measured in other inertial frames. So, moving rods contract.

Let's address the Euclidean confusion, in fig 5.8 the Euclidean length of the $\Delta x'$ side of the triangle appear longer than side $\Delta x'$ between points A and B. How can this be?. The answer comes in when we remember, that these increments in spacetime are given by the invariant length Δs in both systems.

Consider the situation in Figure 5.8. The squared minkowskian lenght (Δs^2) between events A and B is given by

$$\Delta s^2 = -(\Delta x^2) = (-\Delta t')^2 - (\Delta x')^2\,. \tag{5.21}$$

From the Lorentz Boost, and using the fact $\Delta t = 0$, we get

$$c\Delta t' = -\beta\gamma\Delta x\,, \tag{5.22}$$

So,

$$-(\Delta x)^2 = (-\beta\gamma\Delta x)^2 - (\Delta x')^2\,. \tag{5.23}$$

Rearranging,

$$(\Delta x')^2 = (1+\beta^2\gamma^2)(\Delta x)^2\,. \tag{5.24}$$

But,

$$1+\beta^2\gamma^2 = 1+\frac{\beta^2}{1-\beta^2} = \frac{1-\beta^2+\beta^2}{1-\beta^2} = \gamma^2\,, \tag{5.25}$$

from where we conclude the Fitzgerald-Lorentz contraction formula

$$L = \frac{L_0}{\gamma} = \sqrt{1-\beta^2}\,L_0\,, \tag{5.26}$$

5.4 Velocity Addition

Let $\mathcal{O}$, $\mathcal{O}'$ two inertial frames, the former is assumed to be the lab or resting system while the latter is moving along the positive x axis of $\mathcal{O}$ with speed $v_{\mathcal{O}'}$. $\mathcal{A}$ be an event, the speed of the moving observer which is supposed to move along the x axis.

Unprimed and primed coordinates are assigned to events according to the observer that describe them, accordingly, two close events have separations that obey the relations then

$$\begin{aligned}\frac{dx}{dt} &= \frac{\left[dx' + v_{\mathcal{O}'}dt'\right]}{\left[dt' + \frac{v_{\mathcal{O}'}}{c^2}dx'\right]} = \\ &= \frac{\left[\frac{dx'}{dt'} + v_{\mathcal{O}'}\right]}{\left[1 + \frac{v_{\mathcal{O}'}}{c^2}\frac{dx'}{dt'}\right]}\,,\end{aligned} \tag{5.27}$$

where

$$\gamma = \frac{1}{\sqrt{1+(\frac{v_{\mathcal{O}'}}{c})^2}} \tag{5.28}$$

If the two close events correspond to the worldline of a particle, dx/dt and dx'/dt' are the "horizontal components" of the velocity of the particle as measured by each of the two frames, and so, the velocity composition law is given by

$$\boxed{v = \frac{v' + v_{\mathcal{O}'}}{1 + \frac{v_{\mathcal{O}'} v'}{c^2}}\,.} \tag{5.29}$$

According to this formula, the velocity as seen form the lab system can never be larger than c. Indeed, even in the limit case $v_{\mathcal{O}'} = c$, v equals c.

In the limit $c \to \infty$ one gets

$$v = \lim_{c\to\infty} \frac{v' + v_{\mathcal{O}'}}{1 + \frac{v_{\mathcal{O}'} v'}{c^2}} = v' + v_{\mathcal{O}'}\,. \tag{5.30}$$

which is nothing but the Galilean formula for velocity composition.

Problem 8 *Show that, under the conditions just discussed, the composition velocity formulas are*

$$\begin{aligned}
\frac{dx}{dt} &= \frac{\frac{dx'}{dt'} + v_{\mathcal{O}'}}{1 + \frac{v_{\mathcal{O}'}}{c^2}\frac{dx'}{dt'}}\,, \\
\frac{dy}{dt} &= \frac{\frac{dy'}{dt'} + v_{\mathcal{O}'}}{\gamma\left[1 + \frac{v_{\mathcal{O}'}}{c^2}\frac{dx'}{dt'}\right]}\,, \\
\frac{dz}{dt} &= \frac{\frac{dz'}{dt'} + v_{\mathcal{O}'}}{\gamma\left[1 + \frac{v_{\mathcal{O}'}}{c^2}\frac{dx'}{dt'}\right]}
\end{aligned} \tag{5.31}$$

Chapter 6

Four Velocity

We have "learned" that displacements ($\Delta\mathbf{x}$) are contravariant vectors, i.e. if $\mathbf{\Lambda}$ is a Lorentz transformation, the components of a displacement transform as

$$\Delta\mathbf{x}' = \mathbf{\Lambda}\,\mathbf{\Delta x}\,, \tag{6.1}$$

or:

$$\Delta x^{\mu\,\prime} = \Lambda^{\mu}_{\nu}\,\Delta x^{\nu}\,. \tag{6.2}$$

This is true in particular for displacements between very close events $dx^{\mu\,\prime} = \Lambda^{\mu}_{\nu}\,dx^{\nu}$, the invariant interval or proper time is given by

$$d\tau^2 = \eta_{\mu\nu}\,dx^{\mu}\,x^{\nu}\,. \tag{6.3}$$

6.1 What is velocity?

The next concept we need to discuss is, obviously, *velocity.* In the usual sense, velocity is the derivative of the position with respect to time, this is, the limit of the quotient between space displacements and intervals of time when the later become infinitesimal. Let us examine the concept closely,

To find the velocity of a moving particle, an observer O measures succesive positions of the moving object along the space part of its world line and uses her watch to measure the time intervals (Δt) between those positions. With such measurements at hand, O defines the velocity as the limit of the incremental coeficients (displacements/time intervals), when the time intervals become very small, in formulas,

$$\vec{v} = \lim_{\Delta t \to 0} \frac{\Delta \vec{x}}{\Delta t}\,, \tag{6.4}$$

Here we must notice that **space displacements** ($\Delta \vec{x}$) are ordinary 3-vectors[1], and therefore, so is the velocity. Unfortunately, we have already learned that the quantities involved in the calculation are not universally defined for all observers.

From the point of view of what be have been studying, what is really going on is that O is observing the world line of the particle and parametrizing it according to his time i.e. O parametrizes the worldline as the following map

$$t \to (t, x^1(t), x^2(t), x^3(t))\,, \tag{6.5}$$

and the standard definition yields the velocity of the particle as been given by

$$\vec{v} = \frac{d\vec{x}(t)}{dt}\,. \tag{6.6}$$

We might improve things a litle bit by recalling that the proper time length

$$d\tau^2 = \eta_{\mu\nu} dx^\mu\, dx^\nu\,, \tag{6.7}$$

is a scalar (a relativistic invariant), this implies that the parametrization of the world line given by

$$\tau \to (t(\tau), \vec{x}(\tau))\,, \tag{6.8}$$

is universally defined, indeed, displacements along the world line are well defined four vectors.

[1]i.e. they transform under $O(3)$

We may now define a new kind of 3−velocity

$$\vec{v}_\tau \equiv \frac{d\vec{x}(t)}{d\tau} \tag{6.9}$$

and get

$$\vec{v}_\tau = \frac{d\vec{x}(t)}{d\,t}\,\frac{d\,t}{d\tau} = \frac{\vec{v}}{\sqrt{1-\beta^2}} \tag{6.10}$$

6.2 Four velocity

Formula 6.10 contains the factor $\sqrt{1-\beta^2}$, which obviously comes from comes from the difference between proper time and time as measured by the LAB observer.

Let us now recall that the components of the 4−displacements dx^μ are good 4-vectors, while $d\tau^2$ is a scalar. According to this, the quotients

$$\frac{dx^\mu}{d\tau}$$

constitute the components of a good contravariant vector, namely, the 4-velocity.

$$U^\mu \equiv \frac{d\,x^\mu}{d\tau} \tag{6.11}$$

$$\mathbf{U} = \left(\frac{1}{\sqrt{1-\beta^2}}, \frac{\vec{v}}{\sqrt{1-\beta^2}}\right)\,. \tag{6.12}$$

Let us now state two remarks about $\mathbf{U}$. The first one, -though trivial- is that formula 6.10 just defines the space components of the four velocity. The second which is also very simple is that $\mathbf{g}(\mathbf{U},\mathbf{U})$ the length squared of $\mathbf{U}$ is constant and equal to 1, indeed,

$$\mathbf{g}(\mathbf{U},\mathbf{U}) = \eta_{\mu\nu}\,U^\mu\,U^\nu = \left(\frac{1}{\sqrt{1-\beta^2}}\right)^2 - \frac{\vec{v}.\vec{v}}{(\sqrt{1-\beta^2})^2} = \frac{1}{1-\beta^2} - \frac{\beta^2}{1-\beta^2} = 1\,. \tag{6.13}$$

There is a hidden trick in what we have done so far. The fact that the length of the four velocity is 1 follows directly from the definition of the proper time interval $d\tau^2 = (dx^0)^2 - (d\vec{x})^2$, but the fact follows only if the interval is not light like. For light like intervals, i.e. light signals, $d\tau^2 = 0$ and the proper time interval cannot be used to define the four velocity.

6.2.1 Geometric definition of four velocity

Our definition of the four velocity lacks the elegance of a geometric (coordinate free definition), it is therefore interesting to turn our attention to such a definition (at least for time like world lines), to achieve this goal we begin by recalling that in Galilean/Newtonian physics, velocity is a 3-vector tangent to a particle's path. The obvious generalization of this concept is given by the following

Definition 6 *Given a world line whose points have time like separation, the four velocity* $\mathbf{U}$ *is the tangent to the world line characterized by having unit minkowskian length (*$\sqrt{\mathbf{g}(\mathbf{U},\mathbf{U})} = 1$*).*

The name four velocity is justified by the close relation between its space components and those of the ordinary velocity. There is a nice interpretation of four velocity, it is not only a tangent vector to the world line, but a tangent such that its length stretches one unit of time in the particle's frame [2]

Let us consider the geometric definition of four velocity for a uniformly moving particle, we do it in the inertial frame in which the particle is at rest. Then, according to definition 6 it should be clear that $\mathbf{U}$ points parallel to the time axis and is one unit of time long. That is, it is identical with the basis vector $\hat{\mathbf{e}}_0$ of that frame. Thus we could also use as our definition of the four-velocity of a uniformly moving particle that $\mathbf{U}$ is the vector $\hat{\mathbf{e}}_0$ of the particle's rest frame.

When dealing with a particle in non uniform motion we find the situation, there is no inertial frame in which the particle is always at rest. Fortunatley, there is always an inertial frame which momentarily has the same velocity as the particle, but which a moment later is of course no longer comoving with it. This frame is referred to as the momentarily comoving reference frame (MCRF), and is a concept os such great importance that we should not fear repeating it as follows:

Definition 7 *At any instant of time (*t*), an instantaneous (momentarily) comoving reference frame (MCRF) of a particle is an inertial reference frame which has the same velocity as the particle at* t*.*

At any instant of time, the world line of a particle in a MCRF is given by

$$x^\mu_{MCRF}(\tau) \rightarrow (\tau, 0, 0, 0)\,, \tag{6.14}$$

then, by definition 6, the velocity in such MCRF is just

$$\mathbf{U}_{MCRF} = (1, 0, 0, 0) \tag{6.15}$$

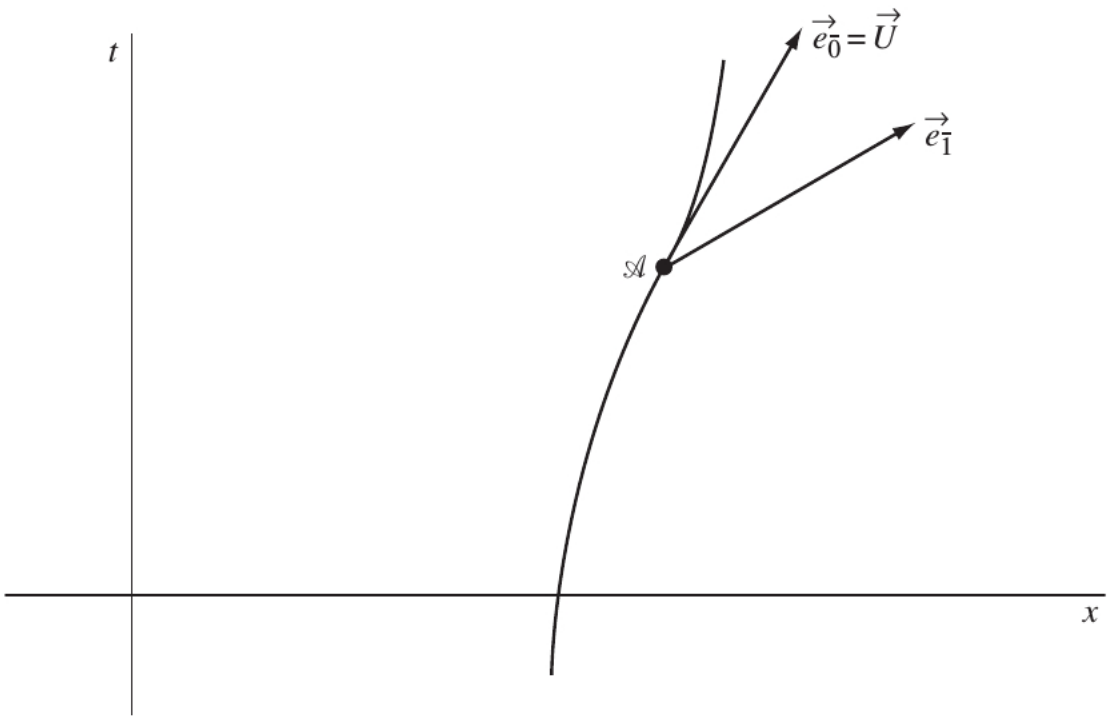


Figure 6.1: The world line of a particle as seen by the lab observer O, at event $\mathcal{A}$ we draw the MCRM $\bar{O}$ where the four velocity is identical with the basis vector $\hat{\mathbf{e}}_{\bar{0}}$

The four velocity of the particle as described by the LAB system is therefore $U^\mu = \Lambda^\mu_\nu\, U^\nu_{MCRF}$, where Λ^μ_ν is the Lorentz transformation with velocity $-v$, v being the velocity of the particle as seen from the LAB.

$$\mathbf{U} = \begin{pmatrix} \gamma & \beta\,\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{U}_{MCRF} = \begin{pmatrix} \gamma \\ \gamma\beta \\ 0 \\ 0 \end{pmatrix} \tag{6.16}$$

6.3 4 acceleration

Four acceleration is defined in an obvios way,

$$\mathbf{a} = \frac{d\mathbf{U}}{d\tau}\,, \tag{6.17}$$

given the fact that τ is a scalar, the 4-acceleration is a good contravariant vector, besides, and because $\mathbf{U}$ has constant unit length, the four acceleration is always orthogonal (in the minkowskian sense) to the acceleration, i.e.

$$\mathbf{g}(\mathbf{U}, \mathbf{a}) = 0\,. \tag{6.18}$$

In the next chapter we will explore some propierties of $\mathbf{a}$ in more detail.

Chapter 7

Acceleration

7.1 Review of Fundamental Notions

Let $\mathcal{W}$ be the worldline of a particle as described from an iretial reference frame. The four velocity ($\mathbf{U}$) is tangent to $\mathcal{W}$ and satisfies $\mathbf{g}(\mathbf{U}, \mathbf{U}) = c^2$, the first derivative $\dot{\mathbf{U}} = d\mathbf{U}/d\tau$ is the four acceleration ($\mathbf{a}$) of the particle.

Since $u_\mu u^\mu = c^2$, the four velocity and the four acceleration must be orthogonal, indeed,

$$\frac{1}{2}\frac{d\mathbf{g}(\mathbf{u},\mathbf{u})}{d\tau} = \mathbf{g}(\mathbf{u},\mathbf{a}) = 0\,, \tag{7.1}$$

If we think of a coordinate frame and let τ be the proper time of a particle, each event $\mathcal{A}$ along along the worldline is located by a four vector $\mathbf{x}(\mathcal{A}) = \mathbf{x}(\tau)$ with entries

$$(x^0(\tau), x^1(\tau), x^2(\tau), x^3(\tau))\,.$$

At each event of the world line of a free particle we can always set an inertial reference frame usually referred to as the **instantaneous rest frame**[1] or **comoving frame**. In such frame, the coordinates of an event at proper time τ and its correspondent four velocity are given by

$$\mathbf{x}(\mathcal{A}) = (c\tau, 0, 0, 0) \qquad \mathbf{U} = (c, 0, 0, 0).$$

An **accelerated** particle has no **inertial** frame in which it is always at rest. However, there is always an inertial frame which momentarily has the same velocity as the particle, but which a moment later is of course no longer comoving with the particle. Such a frame see definition 7 -which is conceptually very important- is called **momentarily comoving reference frame (MCRF)** [2].

Actually, at any event along the world line of an accelerated particle there is an infinite number of MCRFs; all of which move at the same velocity with respect a given inertial frame, but have different orientations, i.e. their spatial axes are obtained from one another by spatial rotations.

Landau gives far more credit to the reader and speaks of the comoving frame in the understanding of it being instantaneous and inertial.

When described by a general inertial frame, the four velocity is ($c\,d\tau = c\,dt\,\sqrt{1 - v^2/c^2}$)

$$\mathbf{U} \equiv \frac{d\mathbf{x}}{d\tau} = (\frac{c}{\sqrt{1 - v^2/c^2}}, \frac{\vec{v}}{\sqrt{1 - v^2/c^2}})\,, \quad \vec{v} = \frac{d\vec{x}}{dt}\,. \tag{7.2}$$

In Landau's insightful notation,

$$u^i = \frac{dx^i}{d(c\tau)} = \frac{dx^i}{c\,dt\sqrt{1 - v^2/c^2}} = \frac{v^i}{c\sqrt{1 - v^2/c^2}} \tag{7.3}$$

[1] Woodhouse.

7.2 Constant Acceleration

We will now present the study of the constant accelerated motion. We will carry on the analysis in three different sections closely following references [5], [4] and [7].

For pedagogical reasons, an inertial reference frame is chosen so the orientation of its axis ensure that the motion is along x. Accordingly, for all values of the proper time τ, the world line of the particle as described for have the form

$$\mathbf{x}(\tau) = (ct(\tau), x(\tau), 0, 0)\,, \tag{7.4}$$

consequently, the four velocity and four acceleration are

$$\begin{aligned} \mathbf{U}(\tau) &= (c\dot{t}(\tau), \dot{x}(\tau), 0, 0) \\ \mathbf{a}(\tau) &= (c\ddot{t}(\tau), \ddot{x}(\tau), 0, 0) \end{aligned} \tag{7.5}$$

and satisfy,

$$\mathbf{g}(\mathbf{U}, \mathbf{U}) \quad = c^2 \tag{7.6}$$

$$\mathbf{g}(\mathbf{U}, \mathbf{a}) \quad = 0 \tag{7.7}$$

$$\mathbf{g}(\mathbf{a}, \mathbf{a}) \quad = -\alpha^2\,, \tag{7.8}$$

in the particular case of constant α we talk of constant acceleration[2].

xxxxx

and setting the initial condition $v(t=0) = 0$ A second integration with $x = 0$ for $t = 0$ gives

---

2

$$\begin{aligned} c^2\dot{t}^2 - \dot{x}^2 &= c^2 \\ c^2\ddot{t}^2 - \ddot{x}^2 &= -\alpha^2 \\ c^2\dot{t}\ddot{t} - \ddot{x}\dot{x} &= 0\,, \end{aligned}$$

7.2.1 Landau -Lifschitz [5] Presentation

In the "lab" system[3] the four acceleration has the form,

$$\frac{d\mathbf{U}}{d\tau} = \frac{d}{d\tau}(\frac{1}{\sqrt{1-v^2/c^2}}, \frac{\vec{v}}{\sqrt{1-v^2/c^2}}), \tag{7.9}$$

given the choice of coordinates, this condition implies $\vec{v} = (v, 0, 0)$. Under these conditions, any experimenter in the lab system will declare the acceleration to be constant if

$$\frac{d}{dt}\left[\frac{v}{\sqrt{1-v^2/c^2}}\right] = a = constant\,,$$

or

$$\frac{v}{\sqrt{1-v^2/c^2}} = a\,t + b\,.$$

setting the initial condition $v(t=0) = 0$ the integration constant is zero $(b = 0)$ and then

$$v(t) = \frac{dx}{dt} = \frac{at}{\sqrt{1+a^2t^2/c^2}}\,. \tag{7.10}$$

A second integration with $x = 0$ for $t = 0$ gives

$$x(t) = \frac{c^2}{a}\left(\sqrt{1+a^2t^2/c^2} - 1\right) \tag{7.11}$$

Notice that for $c \to \infty$

$$\begin{aligned} v(t) &= \frac{at}{\sqrt{1+a^2t^2/c^2}} \approx at \\ x(t) &= \frac{c^2}{a}\left(\sqrt{1+a^2t^2/c^2} - 1\right) \approx \frac{1}{2}at^2\,, \end{aligned} \tag{7.12}$$

we recover the standard formulas for uniform acceleration.

The proper time can be calculated from

$$\tau = \int \sqrt{1-v^2/c^2}\,dt$$

[3]any inertial not comoving frame

Substitution of eq. 7.10 yields

$$
\begin{aligned}
\tau &= \int \sqrt{1 - v^2/c^2}\, dt = \\
&= \int dt \sqrt{1 - \frac{a^2t^2}{c^2(1 + a^2t^2/c^2)}} \\
&= \int dt \sqrt{\frac{1 + a^2t^2/c^2 - a^2t^2/c^2}{1 + a^2t^2/c^2}} = \\
&= \int dt \frac{1}{\sqrt{1 + a^2t^2/c^2}} = \frac{c}{a}\, senh^{-1}\left(\frac{at}{c}\right),
\end{aligned} \tag{7.13}
$$

frome where

$$
\tau = \frac{c}{a}\, senh^{-1}\left(\frac{at}{c}\right) . \tag{7.14}
$$

Which in turn implies

$$
t = \frac{c}{a}\, senh\left(\frac{a\tau}{c}\right)
$$

Subsitution of t in $x(t)$ gives,

$$
\begin{aligned}
x(t) &= \frac{c^2}{a}\left(\sqrt{1 + a^2\left(\frac{c}{a}\, senh\left(\frac{a\tau}{c}\right)\right)^2/c^2} - 1\right) = \\
&= \frac{c^2}{a}\left(\sqrt{1 + senh^2\left(\frac{a\tau}{c}\right)} - 1\right) = \\
&= \frac{c^2}{a}\left[cosh\left(\frac{a\tau}{c}\right) - 1\right]
\end{aligned} \tag{7.15}
$$

The motion (world line) is therefore described by the two parametric equations

$$
\boxed{\begin{aligned}
t(\tau) &= \frac{c}{a}\, sinh\left(\frac{a\tau}{c}\right) \\
x(\tau) &= \frac{c^2}{a}\left[cosh\left(\frac{a\tau}{c}\right) - 1\right]
\end{aligned}} \tag{7.16}
$$

Once again, and for the sake of completeness, we take the large c limit,

$$\boxed{\begin{aligned} t(\tau) &= \frac{c}{a}\, sinh\left(\frac{a\tau}{c}\right) = \frac{c}{a}\left(\frac{a\tau}{c} + \dots\right) \approx \tau \\ x(\tau) &= \frac{c^2}{a}\left[cosh\left(\frac{a\tau}{c}\right) - 1\right] = \\ &= \frac{c^2}{a}\left[1 + \frac{1}{2!}\left(\frac{a\tau}{c}\right)^2 + \dots - 1\right] = \\ &\approx \frac{c^2}{a}\,\frac{1}{2}\left(\frac{a\tau}{c}\right)^2 = \frac{a\tau^2}{2} \end{aligned}} \tag{7.17}$$

7.2.2 Woodhouse

The presentation of reference [4] is completely geometric giving us new insights about the theory.

Let $\mathcal{A}$ be an event along the worldline. As we already know, in any MCRF the components of the four velocity $\mathbf{U}$ at $\mathcal{A}$ are $(c, 0, 0, 0)$, since $g(\mathbf{U}, \dot{\mathbf{U}}) = 0$, the entries of the four acceleration of the particle at event $\mathcal{A}$ may always be set as $(0, \vec{a})$.

If we let τ be the proper time of event $\mathcal{A}$, at proper time $\delta\tau$ after $\mathcal{A}$, the 4-velocity is

$$\mathbf{U}(\tau + \delta\tau) = (c, \delta\vec{v}) = (c, 0) + (0, \vec{a})\,\delta\tau + O(\delta\tau^2)\,. \tag{7.18}$$

Where we have used that to the first order in $\delta\tau$: $\gamma(0) = 1$, $\delta\tau = \delta t$ $(t = x^0/c)$. Therefore

$$\vec{a} = \frac{d\vec{v}}{d\tau} = \frac{d\vec{v}}{dt}\frac{dt}{d\tau} = \frac{d\vec{v}}{dt}\,, \tag{7.19}$$

meaning that in the instantaneous rest frame, the spatial part of $\dot{\mathbf{U}}$ is the acceleration in the ordinary sense. Accordingly, the scalar

$$a \equiv \sqrt{-\mathbf{g}(\dot{\mathbf{U}}, \dot{\mathbf{U}})} \tag{7.20}$$

measures the acceleration felt by an accelerating observer with worldline $\mathcal{W}$. a is called the proper acceleration of the particle.

We want to find general expressions for t and x for the case $a = constant$. Some dimensional analysis, intuition and formula 7.8 suggest the following ansäntz[4]

$$\begin{aligned} \ddot{t} &= \frac{a}{c}\, sinh\left(\frac{a\tau}{c}\right) \\ \ddot{x} &= a\, cosh\left(\frac{a\tau}{c}\right) . \end{aligned} \tag{7.21}$$

An easy integration with initial conditions $\dot{t}(0) = 1$ and $\dot{x}(0) = 0$ yields

$$\begin{aligned} \dot{t} &= cosh\left(\frac{a\tau}{c}\right) \\ \dot{x} &= c\, sinh\left(\frac{a\tau}{c}\right) \end{aligned} \tag{7.22}$$

which indeed satisfy identities 7.6,7.7 and 7.8.

Another integration with initial conditions $t(0) = 0$ and $x(0) = 0$ gives the final result (see formulas 7.16)

$$\begin{aligned} t(\tau) &= \frac{c}{a}\, sinh\left(\frac{a\tau}{c}\right) \\ x(\tau) &= \frac{c^2}{a}\, \left[cosh\left(\frac{a\tau}{c}\right) - 1\right] \end{aligned} \tag{7.23}$$

7.2.3 Misner-Thorne-Wheeler

In chapt 6 of reference [7] the same problem is studied with the condition,

$$\mathbf{g}(\mathbf{a}, \mathbf{a}) = -g^2 , \tag{7.24}$$

where it is clear that they are using *geometrodynamic* units ($c = \hbar = \kappa_B = 1$).

The following system defines the world line.

$$\begin{aligned} \frac{dt}{d\tau} &= u^0, \quad \frac{dx}{d\tau} = u^1 , \\ \frac{du^0}{d\tau} &= a^0, \quad \frac{du^1}{d\tau} = a^1 . \end{aligned} \tag{7.25}$$

The orthogonality between the four velocity and four acceleration

$$u_\mu a^\mu = u^0 a^0 - u^1 a^1 = 0 , \tag{7.26}$$

[4] $cosh^2(x) - sinh^2(x) = 1$

imply

$$a^0 = \alpha u^1\,, \quad a^1 = \alpha u^0\,, \tag{7.27}$$

condition 7.24 uniquely sets choosing (uniquely) $\alpha = g$.

Which implies the following ODEs for the four velocity

$$\frac{du^0}{d\tau} = g u^1\,, \qquad \frac{du^1}{d\tau} = g u^0\,, \tag{7.28}$$

which upon differentiation yield

$$\begin{aligned} \frac{d^2u^0}{d\tau^2} - g u^0 = 0\,, \\ \frac{d^2u^1}{d\tau^2} - g u^1 = 0\,, \end{aligned} \tag{7.29}$$

with general solution

$$\begin{aligned} u^0(\tau) = u^1(0) sinh(g\tau) + u^0(0) cosh(g\tau)\,, \\ u^1(\tau) = u^0(0) sinh(g\tau) + u^1(0) cosh(g\tau)\,. \end{aligned} \tag{7.30}$$

The same result is found by writing the first order system in matrix form

$$\frac{d}{d\tau}\mathbf{u} = \mathbf{M}\mathbf{u} \quad \text{with} \quad \mathbf{u} = \begin{pmatrix} u^0(\tau) \\ u^1(\tau) \end{pmatrix} \text{ and } \mathbf{M} = g \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \tag{7.31}$$

indeed, an elementary exercise shows that

$$e^{\mathbf{t}\tau} = \begin{pmatrix} cosh(g\tau) & sinh(g\tau) \\ sinh(g\tau) & cosh(g\tau) \end{pmatrix}\,, \tag{7.32}$$

which in turns implies

$$\begin{pmatrix} u^0(\tau) \\ u^1(\tau) \end{pmatrix} = \begin{pmatrix} cosh(g\tau) & sinh(g\tau) \\ sinh(g\tau) & cosh(g\tau) \end{pmatrix} \begin{pmatrix} u^0(0) \\ u^1(0) \end{pmatrix} \tag{7.33}$$

The initial conditions $\dot{t}(0) = 1$ and $\dot{x}(0) = 0$ translate to $u^0(0) = 1$ and $u^1(0) = 0$, from where we get

$$\begin{aligned} u^0(\tau) = cosh(g\tau)\,, \\ u^1(\tau) = sinh(g\tau)\,. \end{aligned} \tag{7.34}$$

Another integration gives

$$\begin{aligned} x^0(\tau) &= \frac{1}{g}sinh(g\tau) + C_1\,, \\ x^1(\tau) &= \frac{1}{g}cosh(g\tau) + C_2\,. \end{aligned} \tag{7.35}$$

The initial conditions $t(0) = 0$ and $x(0) = 0$ set $C_1 = 0$ and $C_2 = g^{-1}$, and we reach the final form of the parametric equations for the world line in geometrodynamic units

$$\boxed{\begin{aligned} t(\tau) &= g^{-1}sinh(g\tau) \\ x^1(\tau) &= g^{-1}\left[cosh(g\tau) - 1\right] \end{aligned}} \tag{7.36}$$

We finally reinstall c to obtain the formulas in more conventional units

$$\begin{aligned} t(\tau) &= cg^{-1}sinh\left(\frac{g\tau}{c}\right) \\ x^1(\tau) &= c^2g^{-1}\left[cosh\left(\frac{g\tau}{c}\right) - 1\right] \end{aligned} \tag{7.37}$$

7.3 Space Travel with Constant Aceleration

Imagine that you take a spaceship and travel with acceleration g, the acceleration due to gravity, for 35 years, then

$$\frac{c}{g} \sim 3\times 10^7 seg \approx 1\,\mathrm{yr} \tag{7.38}$$

and so

$$t(35\,\mathrm{yr}) = \frac{c}{g}sinh(g\tau/c)|_{\tau=35\,\mathrm{yr}} \sim e^{35}\,\mathrm{yr} = 1.6\times 10^{15}\,\mathrm{yr}\,, \tag{7.39}$$

put in words after you travel out from earth with a constant acceleration g for a period of 35 yrs in your watch, to the observers in the earth you will have aged for nearly $1.6\times 10^{15}\,\mathrm{yr}$ a number that you should comare with the current estimations of the life span of the universe.

7.4 Four acceleration Formulas in 3+1 notation

As a general rule, formulas in $3+1$ notation have an usually unpleasant look and acceleration is not an exception, the purpose of this section is to explore this claim.

Problem 9 *Consider the four velocity of a particle moving along the x axis, of an inertial frame, namely*

$$U = \gamma(v)(c, v\mathbf{e}_x)\,, \qquad v = \frac{dx}{dt}\,.$$

Show that

$$\begin{aligned}\frac{d\mathbf{U}}{d\tau} &= \frac{v}{c^2(1-v^2/c^2)^{3/2}}\frac{dv}{d\tau}\,(c, v\,\hat{\mathbf{e}}_x) + \frac{1}{(1-v^2/c^2)^{1/2}}\,(0, \hat{\mathbf{e}}_x)\frac{dv}{d\tau} = \\ &= \frac{1}{c(1-v^2/c^2)^{3/2}}\,(v, c\,\hat{\mathbf{e}}_x)\,\frac{dv}{d\tau}\end{aligned} \tag{7.40}$$

Problem 10 *Use problem 9 and the definition of proper acceleration ($a = \sqrt{\mathbf{g}}(\dot{\mathbf{U}}, \dot{\mathbf{U}})$) to show that*

$$a = \frac{1}{1-v^2/c^2}\,\frac{dv}{d\tau} = \frac{1}{(1-v^2/c^2)^{3/2}}\,\frac{dv}{dt} \tag{7.41}$$

Chapter 8

Charged Particles in Electromagnetic Fields I

In this chapter we will introduce some results without justification.

To avoid any source of confusion, we begin by recalling some basic definitions and formulas [6].

The coordinates of an event are $(x^0, x^1, x^2 x^3) = (ct, x, y, z)$. The square of arc length between two very close events is given the scalar

$$ds^2 = c^2 d\tau^2 = c^2 (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \tag{8.1}$$

the velocity (or better said, what everyone usually refers as velocity) is the 3-vector of entries

$$\vec{v}^i = \frac{dx^i}{dt} \,. \tag{8.2}$$

The four velocity, a real Lorentz vector- is the natural tangent to the world line of a particle and therefore its entries are

$$U^\mu = \frac{dx^\mu}{d\tau} \tag{8.3}$$

$$\mathbf{U} = (\gamma c, \gamma \vec{v}) \tag{8.4}$$

and the corresponding contravariant 4-momentum is just $p^\mu = m\, U^\mu$ which en $3+1$ notation is

$$\mathbf{p} = \Big(\frac{mc}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\Big) = \Big(\frac{\mathcal{E}}{c}, \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\Big)\,, \tag{8.5}$$

the energy being defined as

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \,, \tag{8.6}$$

8.1 The Faraday Tensor

We begin with the following, relativity will show us that the correct way to think of usual electromagnetic fields

$$\begin{aligned} \vec{E} &= E_x \hat{\mathbf{e}}_x + E_y \hat{\mathbf{e}}_y + E_z \hat{\mathbf{e}}_z \\ \vec{B} &= B_x \hat{\mathbf{e}}_x + B_y \hat{\mathbf{e}}_y + B_z \hat{\mathbf{e}}_z \,, \end{aligned} \tag{8.7}$$

is by **unifying** them into an antisymmetric covariant 2-tensor (the Faraday tensor) as

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \,,$$

to which we may easily associate a contravariant 2-tensor

$$F^{\mu\nu} = \eta^{\mu\alpha} \eta^{\nu\beta} F_{\alpha\beta} \tag{8.8}$$

a simple excercise (please do it) shows that

$$[F^{\mu\nu}] = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}$$

8.2 Equation of Motion

motionThe second notion states that the dynamics of charged a particle moving in an electromagnetic field is determined by the following relativistic equation of motion

$$\frac{dp^\mu}{d\tau} = \frac{q}{c} F^{\mu\nu} U_\mu \tag{8.9}$$

This is obviously a Lorentz covariant formula.

In order to get a better understanding of equation 8.9 we begin with a simple exercise, finding the contraction $F^{\mu\nu}U_\nu$, we can think of the contraction as a matrix product

$$[F^{\mu\nu}U_\nu] = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} U_o \\ U_1 \\ U_2 \\ U_3 \end{pmatrix} ,$$

which yields,

$$[F^{\mu\nu}U_\nu] = \begin{pmatrix} -E_xU_1 - E_yU_2 - E_zU_3 \\ E_xU_0 - B_zU_2 + B_yU_3 \\ E_yU_0 + B_zU_1 - B_xU_3 \\ E_zU_0 - B_yU_1 + B_xU_2 \end{pmatrix} .$$

The natural (geometric) object is U^μ, given the metric, we end up with

$$[F^{\mu\nu}U_\nu] = \begin{pmatrix} E_xU^1 + E_yU^2 + E_zU^3 \\ E_xU^0 + B_zU^2 - B_yU^3 \\ E_yU^0 - B_zU^1 + B_xU^3 \\ E_zU^0 + B_yU^1 - B_xU^2 \end{pmatrix} .$$

we now change to the time-space split and use standard vector notation to get

$$[F^{\mu\nu}U_\nu] = \begin{pmatrix} \vec{U}.\vec{E} \\ U^0\vec{E} + \vec{U} \times \vec{B} \end{pmatrix} .$$

The relativistic form of the equation of motion ends being

$$\begin{aligned} \frac{dp^0}{d\tau} &= \frac{q}{c}\,\vec{U}.\vec{E} \\ \frac{d\vec{p}}{d\tau} &= \frac{q}{c}\left[\vec{E} + \vec{U} \times \vec{B}\right] . \end{aligned} \tag{8.10}$$

$$\begin{aligned} \frac{dp^0}{d\tau} &= \frac{q}{c}\gamma\,\vec{v}.\vec{E} \\ \frac{d\vec{p}}{d\tau} &= \frac{q}{c}\left[\gamma c\vec{E} + \gamma\vec{v} \times \vec{B}\right] , \end{aligned} \tag{8.11}$$

The first of these equations is the power that the electromagnetic field gives to the particle, while the second is just the usual Lorentz force.

8.3 Motion in an Uniform Magnetic Field

In this case, equation 8.10 takes the form

$$\begin{aligned} \frac{dp^0}{d\tau} &= 0 \\ \frac{d\vec{p}}{d\tau} &= \frac{q}{c}\gamma\vec{v}\times\vec{B}\,, \end{aligned} \tag{8.12}$$

Reinserting the γ's and the identity

$$d\tau = dt/\gamma\,, \tag{8.13}$$

we get

$$\begin{aligned} \gamma\frac{d(\gamma mc^2)}{dt} &= 0 \\ \gamma\frac{d\vec{p}}{dt} &= \frac{q}{c}\gamma\vec{v}\times\vec{B}\,, \end{aligned} \tag{8.14}$$

or in full lab system coordinates,

$$\begin{aligned} \gamma\frac{d(\gamma mc^2)}{dt} &= 0 \\ \gamma\frac{d\vec{p}}{dt} &= \frac{q}{c}\vec{v}\times\vec{B}\,, \end{aligned} \tag{8.15}$$

The firts of these equations (equation for p^0) implies that γ is constant which in turns shows that the speed is constant as well.

Let now $(\hat{\boldsymbol{\epsilon}}_1, \hat{\boldsymbol{\epsilon}}_2, \hat{\boldsymbol{\epsilon}}_3)$ a righthanded basis with $\hat{\boldsymbol{\epsilon}}_3$ parallel to the magnetic induction- In such basis the space equation of motion can be written as

$$\frac{d\vec{v}}{dt} = \omega_B\,\vec{v}\times\hat{\boldsymbol{\epsilon}}_3 \tag{8.16}$$

where

$$\omega_B = ccc \tag{8.17}$$

8.4 Motion in an Uniform Electric Field

Instead of using the space.time split equations ... las 3+1 wee will go back to tensor notation and take $\vec{E} = E_0\,\hat{\mathbf{e}}_x$ so

$$[F^{\mu\nu}] = \begin{pmatrix} 0 & -E & 0 & 0 \\ E & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The only nonzero elements of $F^{\mu\nu}\,U_\nu$ are

$$\begin{aligned} F^{0\nu}\,U_\nu = F^{01}U_1 = -EU_1 = EU^1 \\ F^{1\nu}\,U_\nu = F^{10}U_0 = EU^0 \end{aligned} \tag{8.18}$$

Therefore,

$$\begin{aligned} \dot{p}^0 &= \frac{q}{c}EU^1 \\ \dot{p}^1 &= \frac{q}{c}EU^0 \end{aligned} \tag{8.19}$$

while $\dot{p}^2 = \dot{p}^3 = 0$. From here we get

$$\begin{aligned} \gamma\frac{d\mathcal{E}/c}{dt} &= \frac{q}{c}E\gamma v_x \\ \gamma\frac{dp^1}{dt} &= \frac{q}{c}\gamma c\,E \end{aligned} \tag{8.20}$$

$$\begin{aligned} \frac{d\mathcal{E}}{dt} &= qE\,v_x \\ \frac{dp^1}{dt} &= q\,E \end{aligned} \tag{8.21}$$

The second of these equations is easily integrated,

$$p^1(t) = qE\,t + p^1_0 \tag{8.22}$$

the other two components of the momentum being constant, $p^2(t) = p^2_0$ and $p^3(t) = p^5_0$. Now, since

$$p^i = \frac{mv^i}{\sqrt{1 - v^2/c^2}} \tag{8.23}$$

further integration is complicated by the coupling between the equations.

Let us assume that the initial momentum is perpendicular to the electric field, i.e. $p(o)^1 = 0$. in that case,

$$\mathcal{E}^2 = (mc^2)^2 + c^2\vec{p}^2 = \mathcal{E}_0^2 + c^2(qEt)^2 \tag{8.24}$$

with

$$\mathcal{E}_0^2 = (mc^2)^2 + p_0^2\,, \qquad p_o^2 = (p_0^2)^2 + (p_0^3)^2\,. \tag{8.25}$$

But

$$\frac{d\mathcal{E}}{dt} = qE\,v_x \tag{8.26}$$

or

$$\frac{d\mathcal{E}}{dt} = qE\,\frac{dx}{dt}\mathcal{E}_0^2 + c^2(qEt)^2 \tag{8.27}$$

which implies

$$\mathcal{E}(t) = qE\,x(t) + \mathcal{E}_0\,. \tag{8.28}$$

Consistency implies the identity

$$\begin{aligned}[qE\,x(t) + \mathcal{E}_0]^2 = c^2(qEt)^2 + \mathcal{E}_0^2 = \\ = \mathcal{E}_0^2\left[\left(\frac{cqEt}{\mathcal{E}_0}\right)^2 + 1]\right]\end{aligned} \tag{8.29}$$

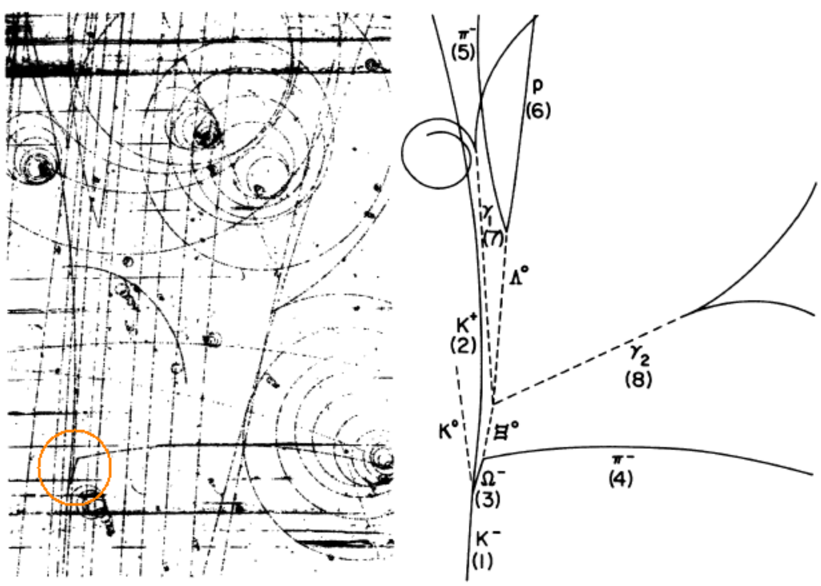
$$qE\,x(t) = \pm\mathcal{E}_0\left[\left(\frac{cqEt}{\mathcal{E}_0}\right)^2 + 1]\right]^{1/2} - \mathcal{E}_0 \tag{8.30}$$

choosing the + which means .... we end up with

$$x(t) = \frac{\mathcal{E}_0}{qE}\left\{\left[\left(\frac{cqEt}{\mathcal{E}_0}\right)^2 + 1]\right]^{1/2} - 1\right\} \tag{8.31}$$

Chapter 9

Action for a free particle



We want to formulate the dynamics of a free paticle moving in Minkowski space-time. We know it is trivial in the sense that the possible world lines of the particle should be straight lines.

The action must be a scalar built from the coordinates of the particle, for symmetry[1] reasons there is only one scalar at our disposal, namely the proper time interval ds, so it is reasonable to try the following form for an action[2]

$$S = -\alpha \int_{\mathcal{I}}^{\mathcal{F}} \sqrt{\eta_{\mu\nu}\, dx^\mu\, dx^\nu}\,, \tag{9.1}$$

where $\mathcal{I}$ and $\mathcal{F}$ are the initial and final points of the world line of the particle.

Action 9.1 is reffered to as *manifestly covariant*, this expression has the following particular meaning, we say that a formula or equation is manifestly covariant when it is evidently invariant under some transformations. For example, Newton's second law

$$\frac{d\vec{p}}{d\,t} = \vec{F}\,, \tag{9.2}$$

is manifestly covariant under galilean transformations, but we simply do not have a sociological tendency to say so. The proper time interval, is manifestly covariant under Translations and Lorentz transformations and people just say manifestly covariant and under the Poincaré group. In fact, the correct way to express that some equations are manifestly covariant is to include the group of transformations that leave them invariant, as heppens with Einstein's fields equations

$$\mathbf{G} + \lambda\,\mathbf{g} = \frac{8\pi\,G}{c^4}\mathbf{T}\,, \tag{9.3}$$

which are manifestly covariant under general coordinate transformations (also called diffeomorphisms).

9.1 Non covariant form of the action

If one chooses the time ($x^0 = c\,t$) as parameter for the world line, the action takes the form

$$S = -c\alpha \int_{t_{in}}^{t_{fin}} \sqrt{1 - \left(\frac{\vec{v}}{c}\right)^2}\, dt\,, \tag{9.4}$$

[1]the world symmetry in this phrase means invariant under combined translations and Lorentz transformations, i.e. the Poincaré Group

[2]an argument along this same line of reasonign is given in Landau & Lifschitz for the case of a nonrelativistic particle

this form of the action is not manifestly covariant, we have made an explicit galilean-like split between space and time which, as we have already learned, forgets about the correct geometry of spacetime, this lack of a proper geometry is the source of some light unpleasanties which deserve some remarks.

To begin with, velocity squared means:

$$(\vec{v})^2 = \sum_{i=1}^{3} \frac{dx^i}{dt}\frac{dx^i}{dt}\,, \tag{9.5}$$

i.e. it is a 3-D euclidean scalar product. The heart of the matter lies in how we treated the proper time interval,

$$c^2\, d\tau^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu = \eta_{00}\, dx^0\, dx^0 + \eta_{ij}\, dx^i\, dx^j = (dx^0)^2 + \eta_{ij}\, dx^i\, dx^j\,, \tag{9.6}$$

here comes the trick, η_{ij} is nothing but $-\delta_{ij}$ which is from where we get the euclidean metric. Now, the entries of the inverse euclidean metric are identical to those of the metric, i.e. $\delta^{ij} = \delta_{ji}$, this means that $\eta_{ij}x^j = -x^i$ which is something that will jump to our faces very soon.

Indeed, according to the comments just made,

$$c\, d\tau = \sqrt{1 - \frac{\sum_{i=1}^{3} dx^i\, dx^j}{(dx^0)^2}}\, dx^0 = c\sqrt{1 - \left(\frac{\vec{v}}{c}\right)^2}\, dt\,, \tag{9.7}$$

Going back to the action 9.4 we easily read the lagrangian,

$$L = -c\alpha\sqrt{1 - \left(\frac{\vec{v}}{c}\right)^2}\,, \tag{9.8}$$

for speeds far below c,

$$L \approx\; -\alpha\, c\left(1 - \frac{v^2}{2\, c^2}\right) = \alpha\,\frac{v^2}{2\, c} - \alpha\, c \tag{9.9}$$

the constant term can be dropped out for the dynamics leaving us with

$$L' = \alpha\,\frac{v^2}{2\, c}\,, \tag{9.10}$$

which, if we choose, $\alpha = mc$, coincides with the nonrelativistic action for a free particle.

We have then obtained the non manifestly covariant form of the action for a free relativistic particle

$$S = -m\, c^2 \int \sqrt{1 - \left(\frac{v}{c}\right)^2}\; dt \tag{9.11}$$

9.2 Introducing the four momentum

Having gotten an action, we can easily calculate the canonical momenta as

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = -\frac{m\,\delta_{ij}\,\dot{x}^j}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\,. \tag{9.12}$$

At this point is where not using a manifestly covariant formulation gets unpleasant, indeed, a litle observation shows that, what we have just obtained is proportional to the space components of the covariant four velocity (U_μ), to see this we just need to remember that the four velocity is defined as a contravariant four vector with components

$$U^\mu \rightarrow (\frac{1}{\sqrt{1-\beta^2}}, \frac{\dot{x}^1}{\sqrt{1-\beta^2}}, \frac{\dot{x}^2}{\sqrt{1-\beta^2}}, \frac{\dot{x}^3}{\sqrt{1-\beta^2}}) = (\frac{1}{\sqrt{1-\beta^2}}, \frac{\vec{v}}{\sqrt{1-\beta^2}})\,, \tag{9.13}$$

which clearly implies

$$U_\mu \rightarrow (\frac{1}{\sqrt{1-\beta^2}}, -\frac{\vec{v}}{\sqrt{1-\beta^2}})\,, \tag{9.14}$$

Formula 9.12 is thus, -apparently- telling us that the canonical momenta conjugate to the position coordinates (x^i) are

$$p_i = -m\,U_i = m\,\eta_{i\mu}U^\mu \tag{9.15}$$

Since we are using cartesian coordinates for the space part of the space time, the canonical momenta correspond to the components of the mechanical (Newtonian) momentum[3]. To make the connection, we need the contravariant part of the momenta, i.e. p^i which should just be

$$p^i = m\,U^i\,, \tag{9.16}$$

or, in vector notation,

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\left(\frac{v}{c}\right)^2}}\,, \tag{9.17}$$

which clearly goes to the newtonian momentum in the low speed limit.

[3]remember, for instance, that the conjugate momenta of an angle is a component of the angular momentum

Since the Lagrangian is independent of time, the quantity $\vec{p}.\vec{v} - L$ is the energy of the particle, performing the calculation yields:

$$E = \frac{m\vec{v}.\vec{v}}{\sqrt{1-\left(\frac{v}{c}\right)^2}} + mc^2\sqrt{1-\left(\frac{v}{c}\right)^2} = \frac{m\,c^2}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \tag{9.18}$$

In the low speed limit we find

$$E = mc^2 + \frac{m\,v^2}{2} \tag{9.19}$$

which shows Einstein's *rest energy*.

It is tempting to unify the energy and the 3−momentum to build a proper four vector as follows

$$p^\mu \equiv (E/c, \vec{p})\,, \qquad p_\mu \equiv (E/c, -\vec{p}) \tag{9.20}$$

nevertheless, at this point we are lacking a rigurous proof of the transformation propierties of this object showing that this construction is indeed a four vector. The proof is not hard but it is better to forget about it and just accept -for the moment that the combination just proposed is indeed a contravariant four vector. According to this assumption, the invariant "length" of the momentum is given by

$$\eta_{\mu\nu}\, p^\mu\, p^\nu = E^2/c^2 - \mathbf{p}^2 = \frac{m^2\,c^2}{1-\left(\frac{v}{c}\right)^2} - \frac{m^2\,c^2 v^2/c^2}{1-\left(\frac{v}{c}\right)^2} = m^2\,c^2\,, \tag{9.21}$$

a result impying that the mass of a particle is invariant.

Let us briefly review what we have found in this section,

1. The non covariant formulation was instrumental in finding tha mass of the particle through comparission with Newton's theory

2. We found an extremely good candidate for four momentum.

3. We are not sure about our finding because our formulation is not manifestly covariant.

4. The definition of p^μ as $m\,U^\mu$ is supported by intuition but does certainly gives us a well defined four vector. Neverhtelles it is not what we get directly form the action.

Why am I so picky about manifest covariance?. In the case at hand things are prettty simple and there is no real necesity of being so rigurous. Nevertheless in more interesting problems, such as when discussing certain field theories, proving the Lorentz covariance can be not only uncomfortable but an imperative nuisance.

The case of string theory is emblematic. The interest is to build a consistent quantum theory. In the light cone gauge formalism, the manifest lorentz covariance of bosonic string theory is broken on purpose, consequently, the final steps before trying to extract some interesting physics is to check for Lorentz covariance, which turns to be a titanic tour the force rendering the space time to be a $D = 26$ manifold.

The obvious way to avoid the headaches of trying to prove Lorentz covariance is to work in a manifest covariant way, which is the topic of the next section.

9.3 Covariant formulation of the free particle

Taking our starting point in the covariant form of the action

$$S = -m\,c \int_{\mathcal{I}}^{\mathcal{F}} \sqrt{\eta_{\mu\nu}\, dx^\mu\, dx^\nu}\,, \tag{9.22}$$

We parametrize by proper time interval

$$S = -m\,c \int_{\mathcal{I}}^{\mathcal{F}} \sqrt{\eta_{\mu\nu}\, \dot{x}^\mu\, \dot{x}^\nu}\, ds\,, \qquad \dot{x}^\mu = \frac{dx^\mu}{ds} \tag{9.23}$$

and obtain the canonical momentum as

$$p_\mu = \frac{\partial\, L}{\partial \dot{x}^\mu}\,, \tag{9.24}$$

Since the $\dot{x}^\mu$ are the components of the contravariant four velocity, the p_μ's constitute the components of a covariant vector, or as we might say, the indices are where they belong. Here we see the first advantage of using a manifestly covariant formulation, there is no need to prove anything. Neverhtelles, let us write a litle bit more,

$$p_\mu = \frac{mc\,\eta_{\mu\nu}\dot{x}^\mu}{\sqrt{\dot{x}_\mu\, \dot{x}^\mu}}\,, \tag{9.25}$$

and so, parametrizing by propertime lenght $\dot{x}_\mu\,\dot{x}^\mu = 1$ and recalling that $\dot{x}^\mu = U^\mu$ we get, as we suspected,

$$p_\mu = m\,c\,U_\mu\,, \tag{9.26}$$

in passing, we insist in that the canonical momentum is naturally a covariant four vector.

This is a formula to recall, in SI units:

$$p_\mu\,p^\mu = m^2c^2 \tag{9.27}$$

Chapter 10

Relativistic Correction to Kepler Orbit

This chapter, which closely follows reference [11], is important for both, historical and physical reasons.

10.1 Equations of Motion

We begin with the Lagrangian,

$$L = -mc^2\sqrt{1 - \left(\frac{v}{c}\right)^2} - U(r)\,, \tag{10.1}$$

and assume that planetary orbits are confined to a plane, so we can use polar coordinates to write $v^2 = \dot{r}^2 + r^2\dot{\theta}^2$. Consequently the equations of motion are

$$\frac{d}{dt}\left[\gamma r^2\dot{\theta}\right] = 0 \tag{10.2}$$

$$\gamma\ddot{r} + \dot{\gamma}\dot{r} + \frac{GM}{r^2} - \gamma r\dot{\theta}^2 = 0 \tag{10.3}$$

Equation 10.2 defines the constant angular momentum

$$\ell = \gamma r^2\dot{\theta} = constant\,, \tag{10.4}$$

impliying that the last term in eq 10.3 canbe cast as

$$\gamma r\dot{\theta}^2 = \frac{\ell^2}{\gamma r^3}\,. \tag{10.5}$$

At this point we stop to notice that the chain rule applied to an arbitrary function of time $f(t)$ yields

$$\frac{df}{dt} = \frac{df}{d\theta}\frac{d\theta}{dt} = \frac{\ell}{\gamma r^2}\frac{df}{d\theta}\,, \tag{10.6}$$

this is equivalent to the operator identity

$$\frac{d}{dt} = \frac{\ell}{\gamma r^2}\frac{d}{d\theta} \tag{10.7}$$

Accordingly,

$$\dot{r} = \frac{\ell}{\gamma r^2}\frac{dr}{d\theta} = -\frac{\ell}{\gamma}\frac{d}{d\theta}\frac{1}{r}; \tag{10.8}$$

Now, since γ is independent of θ,

$$\frac{d(\gamma\dot{r})}{dt} = \frac{\ell}{\gamma r^2}\frac{d\gamma\dot{r}}{d\theta} = \frac{\ell}{r^2}\frac{d\dot{r}}{d\theta} = -\frac{\ell^2}{\gamma r^2}\frac{d}{d\theta^2}\frac{1}{r}\,, \tag{10.9}$$

but also,

$$\frac{d(\gamma\dot{r})}{dt} = \dot{\gamma}\dot{r} + \gamma\ddot{r}\,, \tag{10.10}$$

consequently

$$\gamma\ddot{r} = -\dot{\gamma}\dot{r} - \frac{\ell^2}{\gamma r^2}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\frac{1}{r}. \tag{10.11}$$

substitution of formulas 10.5, 10.11 in equation 10.3 gives us the relativistic form of the orbit equation

$$\boxed{\ell^2\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\frac{1}{r} + \frac{\ell^2}{r} = GM\gamma\,.} \tag{10.12}$$

Defining

$$r_c = \frac{\ell^2}{GM} \quad \mathit{and} \quad \lambda = \gamma - 1 \tag{10.13}$$

the orbit equation takes the form

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\frac{r_\mathrm{c}}{r} + \frac{r_\mathrm{c}}{r} = 1 + \lambda\,. \tag{10.14}$$

In the limit $c \to \infty$ we recover the usual Newtonian orbit equation

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\frac{r_\mathrm{c}}{r} + \frac{r_\mathrm{c}}{r} = 1\,. \tag{10.15}$$

with the well known conic solutions solution

$$\frac{r_c}{r} = 1 + e\cos\theta, \tag{10.16}$$

where e is the eccentricity of the orbit.

10.2 Lowest Order Corrections to the Orbits

The series of $\lambda = \gamma - 1$ in terms of powers of c^{-1} has the leading term

$$\lambda \approx \tfrac{1}{2}(r\dot{\theta}/c)^2\,, \tag{10.17}$$

or

$$\lambda \approx \tfrac{1}{2}(\ell/rc)^2. \tag{10.18}$$

To this order, the equation of motion (10.14) is approximately expressed as

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\frac{r_c}{r} + \frac{r_c}{r} \approx 1 + \frac{1}{2}\epsilon\left(\frac{r_c}{r}\right)^2. \tag{10.19}$$

where

$$\epsilon \equiv (GM/\ell c)^2\,. \tag{10.20}$$

Equation 10.19 is clearly non linear and therefore quite difficult to solve. Nevertheless, and as stated in the previous section, setting $\epsilon = 0$ ($c \to \infty$) recovers the conic-sections of Newtonian mechanics (equations (10.15) and (10.16)). The solution of eq. (10.19) for $\epsilon \neq 0$ gives the first -special- relativistic corrections to Keplerian orbits.

Under the assumption of small values of ϵ, it is convenient to make the change of variable

$$1/s \equiv r_c/r - 1\,, \tag{10.21}$$

and think of the case $1/s \ll 1$. Then

$$(r_c/r)^2 \approx 1 + 2/s\,, \tag{10.22}$$

resulting in the following elementary equation for $1/s(\theta)$:

$$\frac{2}{\epsilon}\frac{d^2}{d\theta^2}\frac{1}{s} + \frac{2(1-\epsilon)}{\epsilon}\frac{1}{s} \approx 1\,. \tag{10.23}$$

The change of variable $\alpha \equiv \sqrt{1-\epsilon}\,\theta$ transforms the ODE to the familiar form

$$\frac{d^2}{d\alpha^2}\frac{s_c}{s} + \frac{s_c}{s} \approx 1, \tag{10.24}$$

where $s_c \equiv 2(1-\epsilon)/\epsilon$. The solution is similar to that of Eq. (10.15):

$$\frac{s_c}{s} \approx 1 + A\cos\alpha, \tag{10.25}$$

where A is an arbitrary constant of integration. In terms of the original coordinates, Eq. (10.25) becomes

$$\frac{\tilde{r}_c}{r} \approx 1 + \tilde{e}\cos\tilde{\kappa}\theta, \tag{10.26}$$

where

$$\tilde{r}_c \equiv \frac{1-\epsilon}{1-\frac{1}{2}\epsilon} r_c\,, \tag{10.27}$$

$$\tilde{e} \equiv \frac{\frac{1}{2}\epsilon A}{1-\frac{1}{2}\epsilon}, \tag{10.28}$$

$$\tilde{\kappa} \equiv (1-\epsilon)^{\frac{1}{2}}. \tag{10.29}$$

The "correspondence principle", establishes that in the limit $\epsilon \to 0$ $(c \to \infty)$ we must recover Kepler's orbits, Eq. (10.16), therefore, $\frac{1}{2}\epsilon A \equiv e$ is the eccentricity of Newtonian mechanics.

To first order in ϵ, Eqs. (10.27)–(10.29) become

$$\tilde{r}_c \approx (1-\tfrac{1}{2}\epsilon)\, r_c, \tag{10.30}$$

$$\tilde{e} \approx (1+\tfrac{1}{2}\epsilon)e, \tag{10.31}$$

$$\tilde{\kappa} \approx 1-\tfrac{1}{2}\epsilon, \tag{10.32}$$

so that relativistic orbits in this limit are described concisely by

$$\frac{r_c(1-\frac{1}{2}\epsilon)}{r} \approx 1 + e(1+\tfrac{1}{2}\epsilon)\cos(1-\tfrac{1}{2}\epsilon)\theta\,. \tag{10.33}$$

This approximate orbit equation clearly displays three characteristics: precession of perihelion; reduced radius of circular orbit; and increased eccentricity.

Besides, eq. 10.33 has the same form as that derived from general relativity [12],

$$\frac{r_c(1-3\epsilon)}{r} \approx 1 + e(1+3\epsilon)\cos(1-3\epsilon)\theta, \tag{10.34}$$

In the next section we will comment on the real observational parameters.

10.3 Characteristics of Near-Keplerian Orbits

The famous procession of Mercury's perihelion problem is quite simple to understand. In a perfectly conic orbit, a planet as seen from the sun returns to its initial position at the beginning of the planet' s year. In trms of the coordinates used to solve the orbits equation, after a complete theta revolution (a change of 2π rad in theta), the planet goes to its original position.

In equation 10.33 the cosine changes as

$$cos(\tilde{\kappa}\theta)\,, \tag{10.35}$$

showing that the angular period (Δ) of the orbit is not 2π.

Let's us ask what is the value ($\Delta\theta$) that added to $\theta = 2\pi$ makes the cosine argument change in exactly 2π, i.e. we are looking for $\Delta\theta$ (the so called shift) such that

$$2\pi = \tilde{\kappa}(2\pi + \Delta\theta)\,, \tag{10.36}$$

a little bit of elementary algebra shows that the prediction for the shift of the perihelion is

$$\Delta\theta \equiv 2\pi(\tilde{\kappa}^{-1} - 1) \approx \pi\epsilon \tag{10.37}$$

per revolution. This prediction is identical to that derived in reference [13] incorporating special relativity into the Kepler problem. It is also comparable to observations assuming that the relativistic and Keplerian angular momenta are approximately equal. For a Keplerian orbit $\ell^2 = GMa(1-e^2)$, where $G = 6.670 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$, $M = 1.989 \times 10^{30}\,\mathrm{kg}$ is the mass of the Sun, and a and e are the semi-major axis and eccentricity of the orbit, respectively. Therefore, the relativistic correction defined after Eq. (10.19),

$$\epsilon \approx \frac{GM}{c^2 a(1-e^2)}, \tag{10.38}$$

is largest for planets closest to the Sun and for planets with very eccentric orbits. For Mercury [14] $a = 5.79 \times 10^{10}\,\mathrm{m}$ and $e = 0.2056$, so that $\epsilon \approx 2.66 \times 10^{-8}$. (The speed of light is taken to be $c^2 = 8.987554 \times 10^{16}\,\mathrm{m^2/s^2}$.) According to Eq. (10.37), Mercury precesses through an angle

$$\Delta\theta \approx \frac{\pi GM}{c^2 a(1-e^2)} = 8.36 \times 10^{-8}\,\mathrm{rad} \tag{10.39}$$

per revolution. This angle is very small and is usually expressed cumulatively in arc seconds per century. The orbital period of Mercury is 0.24085 terrestrial years, so that

$$\Delta\Theta \equiv \frac{100\,\mathrm{yr}}{0.24085\,\mathrm{yr}} \times \frac{360 \times 60 \times 60}{2\pi} \times \Delta\theta \tag{10.40a}$$

$$\approx 7.16\,\mathrm{arcsec/century}. \tag{10.40b}$$

The general relativistic (GR) treatment results in a prediction of 43.0 arcsec/century [15], and agrees with the observed precession of perihelia of the inner planets [15, 16]. Historically, this contribution to the precession of perihelion of Mercury's orbit precisely accounted for the observed discrepancy, serving as the first triumph of the general theory of relativity [1][16]. The present approach, using only special relativity, accounts for about one-sixth of the observed discrepancy, Eq. (10.40b). Precession due to special relativity is illustrated in Fig. **??**.

The approximate relativistic orbit equation Eq. (10.33) [or Eq. (10.26) together with Eqs. (10.30)–(10.32)] predicts that a relativistic orbit in this limit has a reduced radius of circular orbit ($e = 0$). This characteristic is not discussed in the standard approach to incorporating special relativity into the Kepler problem, but is consistent with the GR description. An effective potential naturally arises in the GR treatment of the central-mass problem [14],

$$V_{\mathrm{eff}} \equiv -\frac{GM}{r} + \frac{\ell^2}{2r^2} - \frac{GM\ell^2}{c^2 r^3}, \tag{10.41}$$

that reduces to the Newtonian effective potential in the limit $c \to \infty$. In the Keplerian limit, the GR angular momentum per unit mass ℓ is also taken to be approximately equal to that for a Keplerian orbit [16, 14, 15]. Minimizing V_{eff} with respect to r results in the radius of a stable circular orbit,

$$R_{\mathrm{c}} = \frac{1}{2} r_{\mathrm{c}} + \frac{1}{2} r_{\mathrm{c}} \sqrt{1 - 12\epsilon} \approx r_{\mathrm{c}}(1 - 3\epsilon), \tag{10.42}$$

so that the radius of circular orbit is predicted to be reduced: $R_{\mathrm{c}} - r_{\mathrm{c}} \approx -3\epsilon r_{\mathrm{c}}$. (There is also an unstable circular orbit. See Fig. **??**.) This reduction in radius of circular orbit is six times that predicted by the present treatment using only special relativity, for which $\tilde{r}_{\mathrm{c}} - r_{\mathrm{c}} \approx -\frac{1}{2}\epsilon r_{\mathrm{c}}$. [See Eq. (10.30).]

Most discussions of the GR effective potential Eq. (10.41) emphasize relativistic capture, rather than reduced radius of circular orbit. The $1/r^3$ term in Eq. (10.41) contributes negatively to the effective potential, resulting in a finite centrifugal barrier and affecting orbits very near the central mass (large-velocity orbits). (See Fig. **??**.) This purely GR effect is not expected to be described by the approximate orbit equation Eq. (10.33), which is derived using only special relativity and assumes orbits very far from the central mass (small-velocity orbits).

An additional characteristic of relativistic orbits is that of increased eccentricity. Equation (10.33) predicts that a relativistic orbit will have increased eccentricity, when compared to a Keplerian orbit with the same angular momentum: $\tilde{e} - e \approx \frac{1}{2}\epsilon e$. [See Eq. 10.31.] This characteristic of relativistic orbits also is not discussed in the standard approach to incorporating special relativity into the Kepler problem, but is consistent with the GR description. The GR orbit equation in this Keplerian limit Eq. (10.34) predicts an increase in eccentricity $\tilde{e} - e \approx 3\epsilon e$, which is six times that predicted by the present treatment using only special relativity.

Chapter 11

Charged Particles in Electromagnetic Fields II

11.1 Another Visit to Faraday

In any intermediate course on electromagnetism we learn that a first integration of Maxwell's equations leads naturally to the introduction of the electric and magnetic potentials ϕ and $\vec{A}$ from which

$$\begin{aligned} \vec{E} &= -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \nabla \times \vec{A} \end{aligned} \tag{11.1}$$

As we will learn in section 13.4, the Faraday tensor can be obtained as the four dimensional generalization of the curl of the covariant potential

$$A_\mu \equiv (\phi, -\vec{A})\,. \tag{11.2}$$

as follows

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial\nu A_\mu\,. \tag{11.3}$$

Clearly there is a lot to be proven with these statements, and that will be done in due time.

11.2 Non Relativistic Action of a Charged Particle In An Electromagnetic Field

This will be a very short section that begins with the following

Teorema 1 *The nonrelativistic lagrangian*

$$L = \frac{1}{2}mv^2 - U \tag{11.4}$$

where U is the velocity dependent potential

$$U = q\phi - q\vec{A}.\vec{v}\,, \tag{11.5}$$

yields the correct equations of motion the force for a particle moving in an electromagnetic field.

The proof is left as an exercise. The physical interpretation of the theorem is very simple, meaning U gives rise to the electromagnetic force

$$\vec{F}_{EM} = q\vec{E} + q\vec{v}\times\vec{B} \tag{11.6}$$

The action associated with L is obviously

$$S = \int_{t_{in}}^{t_{fin}} dt\, \left[\frac{1}{2}mv^2 - q\phi + q\vec{A}.\vec{v}\right] \tag{11.7}$$

11.3 Covariant Form of The Action

we note that the last term can be written as a line integral

$$I = q\int_{t_{in}}^{t_{fin}} dt\,\vec{A}.\vec{v} = q\int_{t_{in}}^{t_{fin}} \vec{A}.d\vec{r} \tag{11.8}$$

The goal of this chapter is to look for a covariant form for the interaction.

The first interesting and tempting observation is that

$$I = -q\phi + \vec{A}.d\vec{v} \tag{11.9}$$

reminds us of a four product, it looks in fact similar to

$$A_\mu \, U^\mu$$

This really suggest that a covariant looking coupling might be something like

$$\int d\,\tau A_\mu(x(\tau))\, U^\mu(\tau) \tag{11.10}$$

where -up to constants- **A** should be the covariant vector potential for the Faraday tensor.

Let us explore this possibility by studying the variation of the coupling, clearly

$$\delta(A_\mu \, U^\mu) = \partial_\nu A_\mu \, U^\mu \delta \, x^\nu + A_\mu \, \delta U^\mu = \partial_\nu A_\mu \, U^\mu \delta \, x^\nu + A_\mu \, \delta \dot{x}^\mu \,,$$

integration by parts gives

$$\delta(A_\mu \, U^\mu) = \partial_\nu A_\mu \, \delta x^\nu \, U^\mu + A_\mu \, \delta \frac{dx^\mu}{d\tau} = ... + \frac{d}{d\tau} \left[\, A_\mu \, \delta x^\mu \right] - \frac{dA_\mu}{d\tau} \, \delta x^\mu \,,$$

or

$$\delta(A_\mu \, U^\mu) = \partial_\nu A_\mu \, U^\mu \delta \, x^\nu - \frac{dA_\mu}{d\tau} \, \delta x^\mu + bd$$

some little manipulation gives

$$\delta(A_\mu \, U^\mu) = \partial_\nu A_\mu \, U^\mu \delta \, x^\nu - \partial_\nu A_\mu \, U^\nu \delta x^\mu + bd \,,$$

A little rearrangement of indices shows that

$$\delta(A_\mu \, U^\mu) = F_{\mu\nu} \, U^\nu \delta x^\mu + bd$$

At the end, if we define the full action of the particle as

$$S = -m\,c \int_{\mathcal{I}}^{\mathcal{F}} \sqrt{\eta_{\mu\nu} \, \dot{x}^\mu \, \dot{x}^\nu} \, ds + \alpha \int_{\mathcal{I}}^{\mathcal{F}} A_\mu U^\mu \, ds \tag{11.11}$$

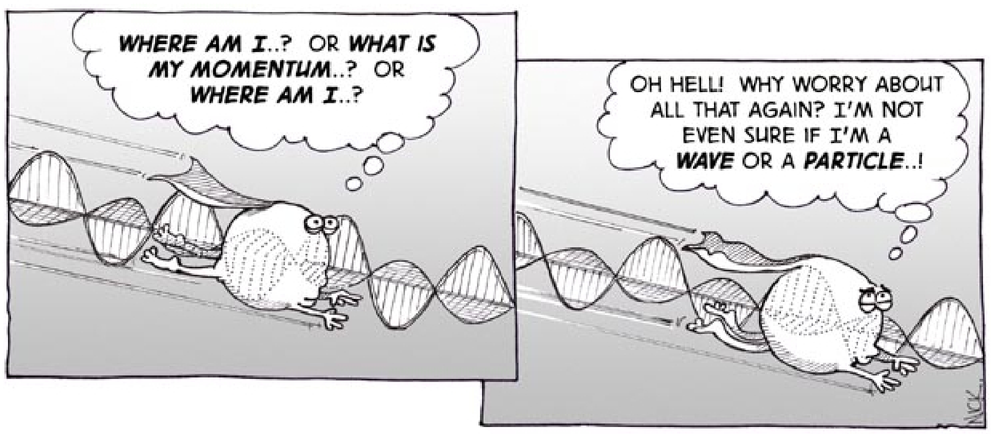
the equation of motion will be

$$\dot{p}^\mu = \alpha F^{\mu\nu} U_\nu \tag{11.12}$$

All that's missing is the value of α which clearly depends on the system of units and which im te iniys of regerence [6] is $\alpha = q/c$.

Chapter 12

Photons



We have already learned how to treat (free) massive particles, the next logical step is to carry on the same kind of study for massless particles, but in doing so we find a clear difficulty, the separation between any two events along the world line of a massless particle is a light like vector, therefore, we cannot associate an instantaenous comoving frame to massless particles, making the definition of their momenta somehow tricky.

There is another problem, the prototype of a massless particle is a photon, and as anyone is perfectly aware, photons are intrinsically quantum, i.e. there is no sucha thing as a classical model for a photon. Consequently, we must look for quantum mechanical tools to bring in any model (even symplistic) of a photon, just as Einstein did in 1905.

12.1 De Broglie's relations and Schrödinger equation

We begin by showing some heuristics that allow a construction of Scrhödinger equation. To such goal we recall the De Broglie's relations for a free particle,

$$\vec{p} = \hbar\,\vec{k}\,, \qquad E = \hbar\,\omega\,, \tag{12.1}$$

and think of the old quantum mechanical idea of the guiding wave, which for a free particle is a plane wave, say, a right moving wave along the x axis

$$\psi(x,t) = e^{i(k\,x-\omega\,t)}\,. \tag{12.2}$$

Given this wave, the question of interest is this:

Can a wave equation be built that respects De Broglie's relations?

From our knowledge about wave motion, we expect the angular frequency and the wave vector to be connected by a dispersion relation

$$\omega = \omega(k)\,. \tag{12.3}$$

Thinking about the plane wavee, it is clear that $\partial_x\psi$ brings a factor ik in front of ψ, so, if we want to get rid of the i and introduce $\hbar$ we may use use the operator

$$-i\hbar\,\partial_x\,, \tag{12.4}$$

where we must note that the minus sign has been engineered to cancel the i^2 that will be obtained when the operator acts on ψ.

At first sight, then, we might accept that wrtitting the momentum as

$$p = -i\hbar\partial_x\,. \tag{12.5}$$

Thinking along similar lines, it does not seem too odd to propose the following

$$E = i\hbar\,\partial_t\,. \tag{12.6}$$

The really bold steps are four, the first, to think of the classical relation between the energy and the momentum ($E = p^2/(2m)$), and the realization, that for the system under consideration, the energy equals the hamiltonian function $H(x, p)$ and the third, to promote the formula $H = p^2/(2m)$ to an operator acting on the wave function, this third step is nothing more than writing

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m}\,, \tag{12.7}$$

where $\mathbf{p}$ means

$$\mathbf{p} = -i\hbar\partial_x\,. \tag{12.8}$$

The fourth and last step is to change the equality between the hamiltonian and the energy as an operator equation actong on the wave function, i.e. stating that ψ satisfies the equation

$$\frac{\mathbf{p^2}}{2m}\psi(x,t) = -i\hbar\partial_t\psi(x,t)\,, \tag{12.9}$$

which in explicitly developed form reads

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} = -i\hbar\partial_t\psi(x,t)\,, \tag{12.10}$$

Applying this equation on the guiding wave, we find that quantum mechanics predicts the following dispersion relation for the frequency and wave number[1]

$$\omega(\vec{k}) = \frac{\hbar\,k^2}{2m}\,. \tag{12.11}$$

If the particle interacts with a potential, we just include it in the hamiltonian as usual and a Schrödinger's equation gets modified to

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t) = -i\hbar\partial_t\psi(x,t)\,, \tag{12.12}$$

At this point we must stop to think that we have just done pure math, ths equation might not mean anything at all, and there is one and only one judge of this equation as having some physical content, the predictions it yields and their comparission to experiments. As we are all well aware, the success is astonishing, most of our technology rests on the predictions of Schrödinger's equation and therefore we may proceed with -somehow good support- our exploration of theoretical physics.

[1]Here we note the following semyclassical reasoning, the group velocityi is defined as $\vec{v}_g = \nabla_{\vec{k}}\omega(\vec{k})$, performing the calculation we get $\vec{v}_g = \hbar\vec{k}/m = \vec{p}/m$ which is the velocity of th particle (if a uantum defintion of velocity were possible)

12.2 The momentum of the Photon

Let us know think about photons. We begin by recalling two things, the first is that light obeys the usual wave equation with dispersion relation $\omega^2/c^2 - \vec{k}^2 = 0$ and the second, that photons' world lines must be light like.

The dispersion relation of light comes from the D'Alambertian operator which is relativistic covariant. Now, think of a right moving (along the z axis) electromagnetic wave:

$$\mathbf{E}_\pm = \mathcal{E}\,\hat{\mathbf{e}}_\pm\, e^{i(kz-\omega t)}\,, \tag{12.13}$$

where $\mathcal{E}$ is the complex amplitude and $\hat{\mathbf{e}}_\pm$ is one of the two possible circular polarization vectors

$$\hat{\mathbf{e}}_\pm = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm\, i \\ 0 \end{pmatrix}\,. \tag{12.14}$$

What we know about relativity clearly establishes that the phase is a scalar, we also know that the coordinates transform as the components of a contravariant Lorentz vector, and so, under Lorentz transformations, ω and k must transform as a covariant vector. Therefore, if we follow the heuristics that we used to build Schrödinger's equation, we are compelled to introduce the following

Definition 8 *The four momentum for a photon moving along the positive z axis is given by the covariant vector*

$$p_\mu = \hbar\, k_\mu\,, \tag{12.15}$$

where

$$k_\mu \to (k, 0, 0, k)\,, \tag{12.16}$$

and $k = \omega/c$

12.3 The Relativistic Doppler effect

Let us now see (once more) the enormous advantage of using the geometric formalism. Consider two physicists O fixed in the LAB and $\bar{O}$ who moves along the x axis of O with velocity v, $\bar{O}$ sends a monochromatic light signal to O.

The Lorentz transformation that changes barred to unbarred coordinates is given by

$$\mathbf{\Lambda} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix}, \tag{12.17}$$

Now let $\bar{k}_\mu$ be the four wave vector (momentum of the photon) as seen by $\bar{O}$, clearly

$$k_0 = \gamma\bar{k}_0 - \gamma\beta\bar{k}_0 = \gamma(-\beta)\bar{k}_0\,, \tag{12.18}$$

or

$$k_0 = \gamma\,[1-\beta]\,\bar{k}_0 = \sqrt{\frac{1-\beta}{1+\beta}}\,\bar{k}_0\,. \tag{12.19}$$

this is all there is to the relativistic Doppler shift, a simple Lorentz transformation.

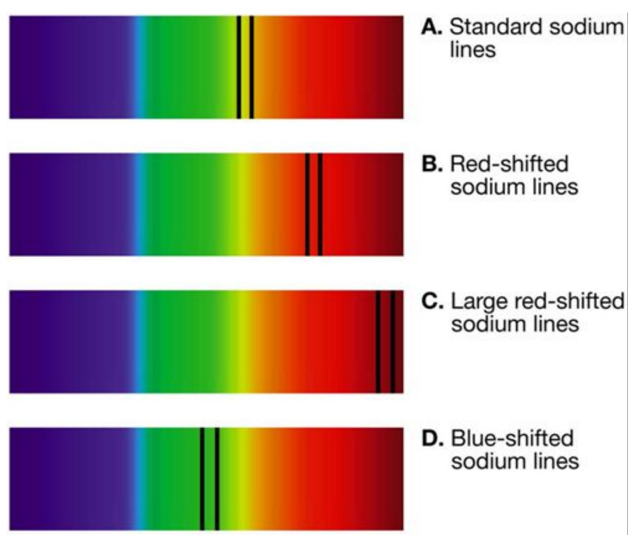


Figure 12.1: Doppler shift illustrated

In astrophysics it is customary to define the ratio

$$\frac{\text{signal frequency}}{\text{observd frequency}} = \frac{\bar{k}_0}{k_0} = \frac{\lambda}{\bar{\lambda}} = \sqrt{\frac{1+\beta}{1-\beta}}\,, \tag{12.20}$$

if the source is receding from the observer, the quotient is bigger than one implying that the observed wave length is greater than the emted one (red shift), if the sources are coming closer the opposite happens and we see a blue shift.

A very important astropysical quantity is the red shift factor z defined as

$$z \equiv \frac{\lambda - \bar{\lambda}}{\bar{\lambda}} = \sqrt{\frac{1+\beta}{1-\beta}} - 1 \tag{12.21}$$

12.4 Compton Effect

From a hystorical perspective, the Compton effect is a cornerstone showing the particle nature of photons. Compton scattering is nothing more that a collision between a photon and an electron.

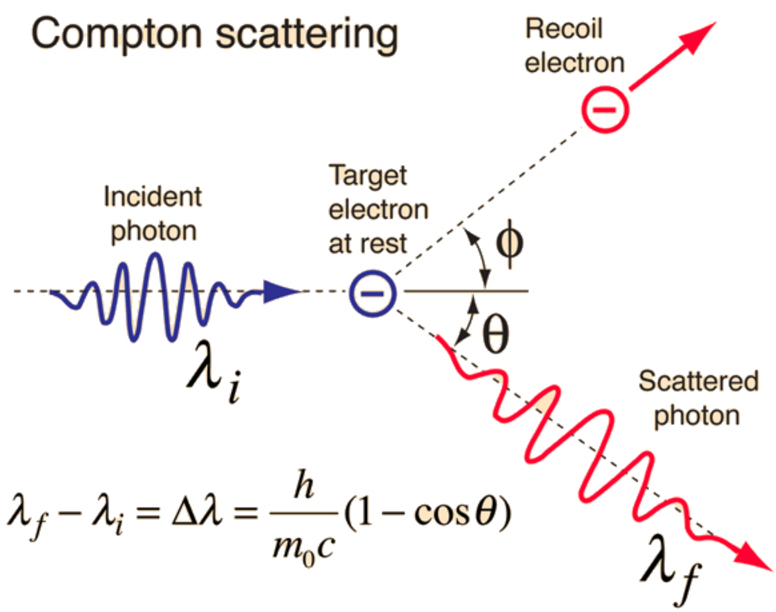


Figure 12.2: Kinematics of the Compton effect

The kinematics of Compton scatering reduces to momentum conservation. If we call $\mathbf{k}_i$ and $\mathbf{p}_i$ the initial momenta of the photon and the electron and $\mathbf{k}_f$ and $\mathbf{p}_f$ their final counterparts, all we have to writte is

$$\mathbf{k}_i + \mathbf{p}_i = \mathbf{k}_f + \mathbf{p}_f \,. \tag{12.22}$$

This implies

$$\mathbf{g}(\mathbf{k}_i + \mathbf{p}_i, \mathbf{k}_i + \mathbf{p}_i) = \mathbf{g}(\mathbf{k}_f + \mathbf{p}_f, \mathbf{k}_f + \mathbf{p}_f) \,. \tag{12.23}$$

Since $\mathbf{g}(\mathbf{k}_i, \mathbf{k}_i) = \mathbf{g}(\mathbf{k}_f, \mathbf{k}_f) = 0$, and $\mathbf{g}(\mathbf{p}_i, \mathbf{p}_i) = \mathbf{g}(\mathbf{p}_f, \mathbf{p}_f) = m^2$ the equation gets simplified to

$$\mathbf{g}(\mathbf{k}_i, \mathbf{p}_i) = \mathbf{g}(\mathbf{k}_f, \mathbf{p}_f) \,. \tag{12.24}$$

In the LAB system the electron is initially at rest ($\mathbf{p}_i = (m_e, 0, 0, 0)$), consequently

$$\mathbf{g}(\mathbf{k}_i, \mathbf{p}_i) = m\, k_i \,, \tag{12.25}$$

so

$$m\, k_i = \mathbf{g}(\mathbf{k}_f, \mathbf{p}_f) \,. \tag{12.26}$$

we now develop the rhs of this result to get

$$m\, k_i = k_{0\,f} p_{0\,f} - \vec{k}_f . \vec{p}_f \,. \tag{12.27}$$

Up to this point we have not used vector components, let us now set the coordinates of lab system in such a way that the incoming photon moves along the x axis, and let us choose the $x - y$ plane to be coincident with the plane containing the scattered photon, then:

$$\begin{aligned} \mathbf{k}_i &= (k_i, k_i, 0, 0) & (12.28)\\ \mathbf{k}_f &= (k_f, k_{x\,f}, k_{y\,f}, 0) & (12.29)\\ \mathbf{p}_f &= (p_{0\,f}, p_{x\,f}, p_{y\,f}, 0) & (12.30) \end{aligned}$$

and $\mathbf{g}(\mathbf{p}_i, \mathbf{p}_i) = \mathbf{g}(\mathbf{p}_f, \mathbf{p}_f) = m^2$ In the LAB system the electron is initialy at rest, therefore, the above equation can be writen as

$$\begin{aligned} k_i + m &= k_{0\,f} + p_{0\,f} & (12.31)\\ k_i &= k_{x\,f} + p_{x\,f} & (12.32)\\ 0 &= k_{y\,f} + p_{y\,f} \,. & (12.33) \end{aligned}$$

Using formulas 12.31, 12.32 and 12.33, in equation 12.27 we get

$$m\,k_i = k_{0\,f}\,[k_i - k_{0\,f} + m] - k_{x\,f}\,[k_i - k_{x\,f}] - k_{y\,f}(-k_{y\,f})\,, \tag{12.34}$$

here we do a litle reordering

$$m\,k_i = k_i\,k_{0\,f} - k_{0\,f}\,k_{0\,f} + k_{0\,f}\,m - k_i\,k_{x\,f} + k_{x\,f}\,k_{x\,f} + k_{y\,f}\,k_{y\,f}\,, \tag{12.35}$$

and still a litle more

$$m\,\left[k_i - k_{0\,f}\right] = k_i\,[k_{0\,f} - k_{x\,f}] - k_{0\,f}\,k_{0\,f} + k_{x\,f}\,k_{x\,f} + k_{y\,f}\,k_{y\,f}\,. \tag{12.36}$$

To find our final expression, we recall that in Compton scattering experiments one uses a goniometer to measure the photon's scattering angle, figure 12.2 clearly shows that

$$\mathbf{k}_f = k_f\,(1, cos\theta, -sen\theta, 0)\,, \tag{12.37}$$

we also need to recall (again!) that the photon's momentum is light like and that experimental physicists do not measure wave numbers but wavelengths in reasonable units (so we must restore $\hbar$ and c. Performing these substitutions we get

$$m\,c\,\hbar\,\left[\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\right] = \frac{\hbar^2}{\lambda_i}\frac{1}{\lambda_f}\,[1 - cos\theta] \tag{12.38}$$

And from here we finally find our well known Compton scattering formula

$$\lambda_f - \lambda_i = \frac{\hbar}{m\,c}\,[1 - cos\theta] \tag{12.39}$$

We would like to say a few closing words about Compton scattering. First of all, figure 12.3 was taken from this web site which has some nice and enlightening explanations. Second, formula 12.39 has kinematic meaning only, if one wants dynamical results, such as an explanation of the origin of the relative amplitudes of the scattered signals in figure 12.3 one must learn some QED, such knowledge allows a calculation of the differential scattering cross section ($d\sigma/d\Omega$), which is given by the famous Klein-Nishina formula. As a last remark, we want to stress that all the reasoning that led us to formula 12.39 is not really dependent of the scattering center

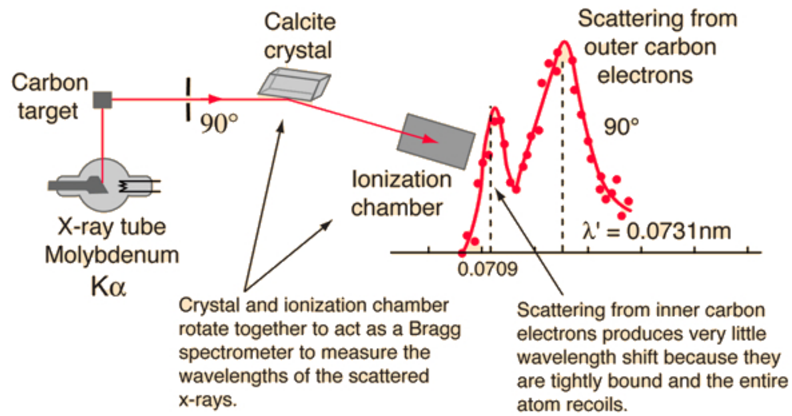
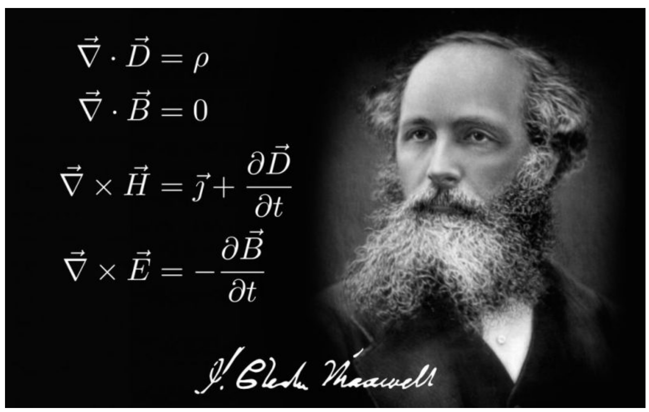


Figure 12.3: Compton's experiment explained, notice the shift

to be an electron, i.e. there is nothing in the reasoning that forbids Compton scattering from ocurring as the at might interact with it (charged particle). The interesting thing to note is the mass (m) which suppresses the effect, you may find interesting data of Compton scattering by protons (nearly tow thousand times more massive than the electron) and will notice that the effect is very small due to such big mass.

Chapter 13

Covariant Form of Electrodynamics



13.1 Maxwell's Equations and Charge Conservation

In this section we will use $3D$ euclidean notation (upper and lower indices are indistiguishable), using this, Maxwell's equations with sources in vacuum have the form

$$\partial_i E_i = 4\pi\rho \tag{13.1}$$

$$\partial_i B_i = 0 \tag{13.2}$$

$$\epsilon_{ijk}\partial_j E_k = -\frac{1}{c}\partial_t B_i \tag{13.3}$$

$$\epsilon_{ijk}\partial_j B_k = \frac{4\pi}{c}J_i + \frac{1}{c}\partial_t E_i \tag{13.4}$$

Contracting Ampere-Maxwell's law with[1] ∂_i and taking the time derivative[2] of Gauss' law for the electric field,

$$\frac{1}{c}\partial_t\partial_i E_i = \frac{4\pi}{c}\,\partial_t\rho \tag{13.5}$$

$$\epsilon_{ijk}\partial_i\partial_j B_k = \frac{4\pi}{c}\,\partial_i J_i + \frac{1}{c}\,\partial_i\partial_t E_i\,. \tag{13.6}$$

In the last equation we use that, since ∂^2_{ij} is symmetric in its indices[3], $\epsilon_{ijk}\partial_i\partial_j = 0$ so,

$$\partial_t\partial_i E_i = 4\pi\,\partial_t\rho \tag{13.7}$$

$$\tfrac{1}{c}\,\partial_t\partial_i E_i + \tfrac{4\pi}{c}\,\partial_i J_i = 0 \tag{13.8}$$

upon substitution of Gauss Law, we finally get[4]

$$\partial_t\rho + \partial_i J_i = 0\,, \tag{13.9}$$

which is the continuity equation, i.e. the mathematical expression of the law of conservation of electric charge.

For the rest of this discussion, we will take charge conservation as being an absolutely fundamental law of nature. Such hypotesis has an amusing consequence. Consider a volume containing certain charge

$$Q = \int_V \rho\, d^3x\,, \tag{13.10}$$

we require Q to be the same for all inertial observers (what would happen to say, a proton, if this were not true?), we perfectly know that under Lorentz transformations, volumes change as

$$d^3x \to \gamma^{-1}\, d^3x\,, \tag{13.11}$$

implying that the charge density must change as the time component of a four vector, i.e.

$$\rho \to \gamma\,\rho\,. \tag{13.12}$$

---

[1] i.e. taking the divergence

[2] we include the factor $1/c$ for obvious dimensional reasons

[3] this is a way to express that $div(curl(\vec{V})) = 0$

[4] $\frac{1}{c}\partial_t(c\rho) + \partial_i J_i = \partial_0(c\rho) + \partial_i J_i$

Even more, Since charge must be conserved, the continuity equation must be the same for all observers, if we think a litle we will realize that the operators

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \tag{13.13}$$

thransform as a covariant vector, and therefore, if we could show that the cureent density vector constitutes the space part of a contravariant four vector current[5],

$$J^\mu \equiv (c\rho, \vec{J}) \tag{13.14}$$

the continuity equation becomes a scalar expression ($\partial_\mu \, J^\mu = 0$) and we will find urselves walking along a good ttrack to establish a manifestly covariant version of Maxwell's equations.

13.2 Potentials for the electromagnetic field

Let us now turn to the magnetic gauss Law, it certainly implies that ther exists a vector field A_i such that[6] the magnetic induction can be expressed as

$$B_i = \epsilon_{ijk}\partial_j A_k \tag{13.15}$$

we can substitute this result in Faraday's law of induction (eq. 13.3)

$$\epsilon_{ijk}\partial_j E_k = -\frac{1}{c}\partial_t \epsilon_{ijk}\partial_j A_k \,, \tag{13.16}$$

which can therefore be cast in the form

$$\epsilon_{ijk}\partial_j \left[E_k + \frac{1}{c}\partial_t A_k\right] = 0 \,, \tag{13.17}$$

this in turn implies the existence of a scalar potential ϕ such that

$$E_k + \frac{1}{c}\partial_t A_k = -\partial_k \phi \,, \tag{13.18}$$

[5]the factor c is introduced for obvious dimensional reasons

[6]in ∇ notation, $\vec{B} = \nabla \times \vec{A}$, $B_x = \partial_2 A_3 - \partial_3 A_2$, $B_y = -\partial_1 A_3 + \partial_3 A_1$, $B_z = \partial_1 A_2 - \partial_2 A_1$

from where we get the following formula for the components of the electric field

$$E_k = -\partial_k \phi - \frac{1}{c}\partial_t A_k \tag{13.19}$$

At this point we must pay attention to a fundamental fact, the electric field and the magnetic induction are invariant under the substitution (gauge transformation)

$$A_i \to A_i + \partial_i \psi \qquad \phi \to \phi - \frac{1}{c}\partial_t \psi \tag{13.20}$$

Indeed, under such transformations the magnetic induction goes into itself as is shown below

$$B_i \to \epsilon_{ijk}\partial_j \left[A_k + \partial_k \psi\right] = B_i + \epsilon_{ijk}\partial^2_{jk}\psi = B_i\,, \tag{13.21}$$

similarly, the electric field is also unchanged

$$E_i \to E_k = -\partial_k \left[\phi - \frac{1}{c}\partial_t \psi\right] - \frac{1}{c}\partial_t \left[A_k + \partial_k \psi\right] = E_k - \frac{1}{c}\partial^2_{tk}\psi + \frac{1}{c}\partial^2_{kt}\psi = E_k \tag{13.22}$$

13.3 Maxwell's equations in terms of the potentials

The two sourceless Maxwell's equations allowed us to express the electric field and magnetic induction in terms of the vector and scalr potentials ($\vec{A}$ and ϕ) as

$$B_i = \epsilon_{ijk}\partial_j A_k \tag{13.23}$$

$$E_k = -\partial_k \phi - \frac{1}{c}\partial_t A_k \tag{13.24}$$

The next logical step consists in expressing the equations containing the sources in terms of the potentials, a simple substitution in Gauss' and Ampere-Maxwell's laws yield

$$\partial_k(\partial_k \phi + \frac{1}{c}\partial_t A_k) = -4\pi\rho \tag{13.25}$$

$$\epsilon_{ijk}\partial_j\left[\epsilon_{k\ell m}\partial_\ell A_m\right] + \frac{1}{c}\partial_t\left[\partial_i \phi + \frac{1}{c}\partial_t A_i\right] = \frac{4\pi}{c}J_i\,, \tag{13.26}$$

using the identity

$$\epsilon_{ijk}\,\epsilon_{k\ell m} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell}$$

We get a more convenient form of the equations

$$\partial^2_{kk}\phi + \frac{1}{c}\partial_t\partial_k A_k = -4\pi\rho \tag{13.27}$$

$$\partial_i\left[\frac{1}{c}\partial_t\phi + \partial_j A_j\right] + \frac{1}{c^2}\partial^2_{tt}A_i - \partial^2_{\ell\ell}A_i = \frac{4\pi}{c}J_i\,, \tag{13.28}$$

in this form, the D'Alambert operator

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\,, \tag{13.29}$$

becomes notoriously conspicuos

We may now use the gauge freedom to choose ϕ and A_i to satisfy

$$\frac{1}{c}\partial_t\phi + \partial_j A_j = 0 \tag{13.30}$$

Doing so the equations adopt the form

$$\frac{1}{c^2}\partial^2_{tt}\phi - \partial^2_{kk}\phi = 4\pi\rho \tag{13.31}$$

$$\frac{1}{c^2}\partial^2_{tt}A_i - \partial^2_{\ell\ell}A_i = \frac{4\pi}{c}J_i \tag{13.32}$$

We already introduced the current four vector

$$J^\mu \to (c\rho, \vec{J})\,, \tag{13.33}$$

If in addition we build a contravariant four vector potential

$$A^\mu \to (\phi, \vec{A})\,, \tag{13.34}$$

the entire set of equations get unified as[7],

$$\partial_\alpha\partial^\alpha\, A^\mu = \frac{4\pi}{c}\, J^\mu \tag{13.35}$$

[7] even the Lorentz gauge choice becomes manifestly covariant,

$$\partial_\mu\, A^\mu = 0$$

13.4 The Field strenght tensor

It may seem odd at this point, but the natural geometric object is not the contravariant potential four vector A^μ, but its covariant counterpart

$$A_\mu = (\phi, -\vec{A})\,. \tag{13.36}$$

With this object at hand, we can define a skewsymmetric covariant tensor (**F**) of two indices, namey, the electromagnetic field strength (or as Missner et all call it, Faraday) tensor with components:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{13.37}$$

Let us examine the entries of **F**. Since it is obvioulsy antysymmeeric,

$$F_{00} = F_{ii} = 0 \tag{13.38}$$

The space - time components coincide with the electric field, indeed

$$F_{0i} = \partial_0(-A_i) - \partial_i A_0 = -\left[\frac{1}{c}\partial_t A_i + \partial_i \phi\right] = E_i \tag{13.39}$$

And, finally, the pure space components[8]

$$F_{ij} = \partial_i(-A_j) - \partial_j(-A_i) = \partial_j A_i - \partial_i A_j\,, \tag{13.40}$$

are related to the magnetic induction as follows

$$F_{12} = \partial_2 A_1 - \partial_1 A_2 = -B_z \tag{13.41}$$

$$F_{13} = \partial_3 A_1 - \partial_1 A_3 = B_y \tag{13.42}$$

$$F_{23} = \partial_3 A_2 - \partial_2 A_3 = -B_x \tag{13.43}$$

There are two more tensors related to Faraday, namely, its contravariant version

$$F^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}\,F_{\alpha\beta}\,, \tag{13.44}$$

and the dual ${}^*\mathbf{F}$ with entries

$${}^*F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}\,F^{\alpha\beta}\,, \tag{13.45}$$

where as usual, $\epsilon_{\mu\nu\alpha\beta}$ is the completely antisymmetric tensor density of weight 1 with $\epsilon_{0123} = +1$

[8] Recall that $B_i = \epsilon_{ijk}\partial_j A_k$, so, for instance, $B_1 = \epsilon_{123}\partial_2 A_3 + \epsilon_{132}\partial_3 A_2 = \partial_2 A_3 - \partial_3 A_2 = -F_{23}$

Teorema 2 *The Faraday tensor* **F** *satisfies the following identity*

$$\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0 \tag{13.46}$$

the proof is a simple excercise,

$$\partial_\alpha \left[\partial_\beta A_\gamma - \partial_\gamma A_\beta\right] + \partial_\gamma \left[\partial_\alpha A_\beta - \partial_\beta A_\alpha\right] + \partial_\beta \left[\partial_\gamma A_\alpha - \partial_\alpha A_\gamma\right] = 0 \tag{13.47}$$

13.5 Maxwell's Equations

To unravel wether theorem 2 has a physical meaning we examine a particular case and leave the others s an excercise in index gymnastics

$$\partial_0 F_{12} + \partial_2 F_{01} + \partial_1 F_{20} = 0\,, \tag{13.48}$$

by virtue of the relation between the components of the Faraday tensor and the usual components of the electromagnetic field, the above formula traslates to

$$\frac{1}{c}\partial_t(-B_z) + \partial_2 E_x - \partial_1 E_y = 0 \tag{13.49}$$

which is just one of the components of faraday's law of induction.

In fact, all homogeneous Maxwell' s equations are encoded theorem 2.

What about the nonhomogeneous equations?, well it happens that they may be writen as

$$\partial_\nu\, F^{\nu\mu} = \frac{4\pi}{c} j^\mu\,. \tag{13.50}$$

or, in developed form

$$\partial_\nu \partial^\nu\, A^\mu - \partial_\mu \partial_\nu\, A^\nu = \frac{4\pi}{c} j^\mu\,, \tag{13.51}$$

which, upon imposing the Lorentz gauge choice, reduces to

$$\partial_\nu \partial^\nu\, A^\mu = \frac{4\pi}{c} j^\mu\,, \tag{13.52}$$

as desired.

To make contact with the usual form of Maxwell's equations, we check a particuular case (the zro component) to convince ourselves. We begin by calculating

$$\partial_\nu F^{\nu 0}\,, \tag{13.53}$$

once again we use the connection between the electrmagnetic fields and the Faraday tensor to get

$$\partial_i F^{i0} = -\partial_i F^{0i} = -\partial_i E^i = 4\pi\rho = \frac{4\pi}{c}(c\,\rho)\,, \tag{13.54}$$

where we have used Gauss' law,

$$div(\vec{E}) = -4\pi\,\rho\,, \tag{13.55}$$

to complete the right hand side

To summarize, Maxwell's equations are equivalent to the following two sets of equations

$$\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0 \tag{13.56}$$

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu \tag{13.57}$$

Equation 13.56 implies the existence of the potential, i.e., it yields $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. While equation 13.57 allows to calculate **F** from the current distribution, not only that, it guarantees that the charged is conserved. Indeed,

$$\partial_\mu J^\mu \propto \partial_\mu \partial_\nu F^{\mu\nu} = 0 \tag{13.58}$$

13.6 Closing remarks

We want to finish this lecture with several remarks, the first one concerns the gauge transformations

$$A_i \to A_i + \partial_i \psi \qquad \phi \to \phi - \frac{1}{c}\partial_t \psi\,, \tag{13.59}$$

they can be easily unified as

$$A_\mu \to A_\mu + \partial_\mu \Lambda\,, \tag{13.60}$$

in this form, it is evident that the Faraday tensor is gauge invariant.

There is a lot more to talk about Lorentz transformations of the 4 potential, and we will leave this discussion aside. Nevertheless, if we accept that A_μ is indeed a covariant vector, we automatically get that the Faraday tensor is indeed a two ndex covariant tensor implying that under Lorentz transformations it transforms as

$$\bar{F}_{\mu\nu} = \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}\, F_{\alpha\beta}\,. \tag{13.61}$$

From here one may read the transformation formulas for the electric ield and magnetic induction.

There is an important question to ask,

Is there an action from where we can get Maxwells' theory?

The answer is: indeed, there is, and not only that, there are differtn forms of the action and very general properties of the electromagnetic field such as energy momen tum conservation may me gotten form them.

Chapter 14

Fields of a Constant Velocity Moving Charge

Now that we know Maxwell's theory to be Lorentz covarian, it is interesting to deal with the simplest system of all, a classical pointlike particle carrying an electric charge Q.

We consider two reference frames, the comoving frame in which the particle is at rest and the lab (L) system from where the particle is seen as moving with constant velocity $\mathbf{v}$. For simplicity we assume that both frames have the same orientation and that $\mathbf{v} = v\hat{\mathbf{e}}_x$

14.1 Fields from the Potentials

The physics as described in the comoving frame is exceedingly simple. Indeed, in such frame all there is is a coulombian field, so the potentials are

$$\phi_{Co} = \frac{Q}{r_{Co}}\,, \qquad \mathbf{A}_{Co} = 0 \tag{14.1}$$

the corresponding fields are clearly

$$\mathbf{E}_{Co} = \frac{Q}{r_{Co}^3}\mathbf{r}_{Co}\,, \qquad \mathbf{H}_{Co} = 0 \tag{14.2}$$

Assuming that the origins of both reference frames are coincident at $t = 0$, and using[1]

$$\begin{aligned} cosh\theta &= \frac{1}{\sqrt{1-\beta^2}} \\ sinh\theta &= \frac{\beta}{\sqrt{1-\beta^2}}\,, \end{aligned} \tag{14.3}$$

the coordinate transformations relating the lab (L) and the Co frames are

$$\begin{aligned} ct_{Co} &= -x_L\, sinh\theta + ct_L\, cosh\theta \\ x_{Co} &= x_L\, cosh\theta - ct_L\, sinh\theta \\ y_{Co} &= y_L \\ z_{Co} &= z_L\,. \end{aligned} \tag{14.4}$$

These imply

$$r_{Co}^2 = (x_L - vt_L)^2 cosh^2\theta + y_L^2 + z_L^2 = s^2 cosh^2\theta \tag{14.5}$$

with

$$s^2 = (x_L - vt_L)^2 + (y_L^2 + z_L^2)[1-\beta^2] \tag{14.6}$$

We now recall that ϕ and $\mathbf{A}$ constitute a 4-vector, so its components in both frames are simply related by exactly the same Lorentz transformations. Therefore,

$$\begin{aligned} \phi_L &= A_{xCo}\, cosh\beta + c\phi_{Co}\, sinh\beta = c\phi_L\, sinh\beta \\ A_{LM} &= A_{LCo} = 0 \end{aligned} \tag{14.7}$$

But we know that,

$$\phi_{Co} = \frac{Q}{r_{Co}} = \frac{Q}{scosh\beta} \tag{14.8}$$

which substituted in the above gives

$$\begin{aligned} \phi_L &= c\phi_{Co}\, cosh\beta = \frac{Q}{s} \\ A_{xL} &= c\phi_{Co}\, sinh\beta = \frac{vQ}{cs} \\ A_{yL} &= A_{yCo} = 0 \\ A_{zL} &= A_{zCo} = 0\,. \end{aligned} \tag{14.9}$$

[1] $\beta = \frac{v}{c} = tanh\theta$

Summarizing

$$\begin{aligned} \phi_L &= c\phi_{Co}\, cosh\beta = \frac{Q}{s} \\ \mathbf{A}_{xL} &= \frac{Q}{c\,s}\mathbf{v}\,, \quad \mathbf{v} = v\,\hat{\mathbf{i}} \end{aligned} \tag{14.10}$$

We note that,

$$div\mathbf{A}_L = \partial_{x_L}\frac{vQ}{cs} = -\frac{vQ\,(x_L - vt_L)}{c\,s^3} \tag{14.11}$$

and

$$\partial_{ct_L}\frac{Q}{s} = \frac{vQ(x_L - vt_L)}{c\,s^3} \tag{14.12}$$

Formulas 14.11 and 14.12 express that the four potential $\mathbf{A}_L^{(4)}$ is in the covariant or Lorentz gauge[2].

To calculate the electric and magnetic fields from the potentials, all we need is to remember that

$$\mathbf{E} = -grad(\phi) - \partial_{ct}\mathbf{A} \quad \text{and} \quad \mathbf{H} = -curl\mathbf{A} \tag{14.13}$$

Taking the derivatives

$$\begin{aligned} \partial_{x_L}\phi_L &= -\frac{(x_L - vt_L)\,Q}{s^3} \\ \partial_{y_L}\phi_L &= -\frac{[1 - v^2/c^2]y_L\,Q}{s^3}\,. \quad \partial_{z_L}\phi_L = -\frac{[1 - v^2/c^2]z_L\,Q}{s^3} \\ \partial_{ct_L}\mathbf{A}_L &= \frac{1}{c}\frac{v^2Q\,(x_L - vt_L)}{c\,s^3}\,\hat{\mathbf{e}}_x\,. \end{aligned} \tag{14.14}$$

Direct substitution yields

$$\mathbf{E}_L = Q\left(1 - v^2/c^2\right)\frac{\mathbf{r}_L}{s^3}\,; \qquad \mathbf{H}_L = \frac{1}{c}\mathbf{v}\times\mathbf{E}_L\,, \tag{14.15}$$

where,

$$\mathbf{r}_L = (x_L - vt_L)\hat{\mathbf{e}}_x + y_L\hat{\mathbf{e}}_y + z_L\hat{\mathbf{e}}_z \tag{14.16}$$

in the limit of very large speed of light,

$$\mathbf{E}_M \approx Q\frac{\mathbf{r}_M}{r_M^3}\,; \qquad \mathbf{H} \approx \frac{Q}{c}\frac{\mathbf{v}\times\mathbf{r}_M}{r_M^3} \tag{14.17}$$

[2] $div\mathbf{A} + \frac{1}{c}\partial_t\phi = 0$

14.1.1 A different approach

Another way to find the fields in the lab frame is to calculate them in the comoving frame and then Lorentz transforming them to the L system. This approach is a bit different since we are now dealing with the entries of the Faraday tensor and must therefore use the right transformation rules.

The only nonzero field is

$$\begin{aligned}\mathbf{E}_{Co} &= \frac{Q}{r_{Co}^3}\mathbf{r}_{Co} = \\ &= \frac{Q}{s^3 cosh^3\beta}\left[(x_L - vt_L)cosh\beta\,\hat{\mathbf{e}}_x + y_L\,\hat{\mathbf{e}}_y + z_L\,\hat{\mathbf{e}}_z\right]\end{aligned} \tag{14.18}$$

For zero magnetic field, the transformed components of the fields are

$$\begin{array}{lll} E_{x_L} = E_{x_{Co}} & E_{y_L} = E_{y_{Co}}\,cosh\alpha & E_{Co}cosh\alpha \\ H_{x_L} = H_{x_{Co}} = 0 & H_{y_L} = -E_{z_{Co}}\,sinh\alpha & H_{z_L} = E_{y_{Co}}sinh\alpha \end{array}. \tag{14.19}$$

Explicitly

$$\begin{aligned}\mathbf{E}_L &= \frac{Q\,cosh\alpha}{s^3cosh^3\alpha}\left[(x_L - vt_L)\,\hat{\mathbf{e}}_x + y_L\,\hat{\mathbf{e}}_y + z_L\,\hat{\mathbf{e}}_z\right] = \\ &= \frac{Q}{s^3cosh^2\alpha}\left[(x_L - vt_L)\,\hat{\mathbf{e}}_x + y_L\,\hat{\mathbf{e}}_y + z_L\,\hat{\mathbf{e}}_z\right] = \\ &= Q\left[1 - v^2/c^2\right]\frac{(x_L - vt_L)\,\hat{\mathbf{e}}_x + y_L\,\hat{\mathbf{e}}_y + z_L\,\hat{\mathbf{e}}_z}{s^3}\end{aligned} \tag{14.20}$$

$$\begin{aligned}\mathbf{H}_L &= sinh\alpha\,\hat{\mathbf{e}}_x \times \mathbf{E}_{Co} = \\ &= \frac{Q\,sinh\alpha}{s^3cosh^3\alpha}\left[y_L\,\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_y + z_L\,\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_z\right] = \\ &= Q\,v(1 - v^2/c^2)\frac{y_L\,\hat{\mathbf{e}}_z - z_L\,\hat{\mathbf{e}}_y}{s^3} = \\ &= \frac{1}{c}\mathbf{v} \times \mathbf{E}_L\end{aligned} \tag{14.21}$$

Chapter 15

Action for Electrodynamics

15.1 Brief introduction to classical field theory

Let us consider a set of N identical masses connected by ideal springs of equilibrium length ℓ. Masses number 1 and number N are also connected to a fixed wall by an extra pair of springs.

We call $q_i(t)$ the displacement of the i-eth mass from its equilibrium position. According to this, the stretch undergone by the spring joining masses m and $m+1$ is

$$q_{m+1} - q_m - \ell\,. \tag{15.1}$$

The lagrangian of such a system is (we must suplement L with the boundary conditions $q_0 = q_{N+1} = 0$)

$$L = \sum_{p=0}^{N} \left\{ \frac{M}{2}\dot{q}_p^2 - \frac{\kappa}{2}\left(q_{p+1} - q_p - \ell\right)^2 \right\} \tag{15.2}$$

There is an easy way to get rid of the unpleasant ℓ in the lagrangian, we just make a shift[1]

$$q_i \to q_i - \frac{\ell}{2}\,, \tag{15.3}$$

Leaving us with the lagrangian

$$L = \sum_{p=0}^{N} \left\{ \frac{M}{2}\dot{q}_p^2 - \frac{\kappa}{2}\left(q_{p+1} - q_p\right)^2 \right\} \tag{15.4}$$

[1] a litle thought (excercise) justifies this.

from L we read the equations of motion

$$M\,\ddot{q}_p - \kappa\,(q_{p+1} - q_p) = 0 \tag{15.5}$$

The action of the system is

$$S = \int_{t_i}^{t_f} dt \sum_{p=0}^{N} \left\{ \frac{M}{2}\dot{q}_p^2 - \frac{\kappa}{2}\,(q_{p+1} - q_p)^2 \right\} \tag{15.6}$$

Let us now introduce ℓ in the following fashion

$$S = \int_{t_i}^{t_f} dt \sum_{p=0}^{N} \ell \left\{ \frac{M}{2\ell}\dot{q}_p^2 - \frac{\kappa\ell}{2}\left(\frac{q_{p+1} - q_p}{\ell}\right)^2 \right\} \tag{15.7}$$

If we now consider the following limiting procedure,

$$\begin{aligned}
N &\rightarrow \infty \\
\ell &\rightarrow 0 \\
q_i(t) &\rightarrow u(x,t) \\
\sum_{p=0}^{N} \ell &\rightarrow \int_{x_1}^{x_2} dx \\
M &\rightarrow 0 \quad M/\ell = \mu \\
\kappa &\rightarrow \infty \quad \kappa\ell = Y \\
\frac{q_{p+1} - q_p}{\ell} &\rightarrow \partial_x\, u(x,t)\,,
\end{aligned} \tag{15.8}$$

the action shall end up being writen as

$$S = \int_{t_i}^{t_f} \int_{x_1}^{x_2} dt\,dx\, \left[\frac{\mu}{2}\,(\partial_t u(x,t))^2 - \frac{Y}{2}\,(\partial_x u(x,t))^2 \right] \tag{15.9}$$

This is our first contact with classical field theory. The N finite degrees of freedom have been substituted by an infinite number of degrees of freedom $u(x,t)$, describing the longitudinal oscillations of a bar made of some material of Young modulus Y and linear mass density μ.

The Lagrangian is the space integral of the lagrangin density

$$\mathcal{L} = \frac{\mu}{2}\,(\partial_t u(x,t))^2 - \frac{Y}{2}\,(\partial_x u(x,t))^2 \tag{15.10}$$

For this kind of systems the equations of motion follow from the variational principle just as happended with a finite number of degrees of freedom. There is only one detail that needs mention, now there are spatial boundaries.

The variational principle always asks for initial and final times fixed boundaries, i.e.

$$\delta u(x,t_i) = \delta u(x,t_f)\,, \quad \forall x \in [x_1,x_2]\,, \tag{15.11}$$

but it says nothing about space boundaries.

Let us see how this works. The variation of the action is

$$\delta S = \int_{t_i}^{t_f}\int_{x_1}^{x_2} dt\,dx\;[\mu\partial_t u(x,t)\delta\partial_t u(x,t) - Y\partial_x u(x,t)\delta\partial_x u(x,t)] \tag{15.12}$$

Integration by parts yields

$$\delta S = -\int_{t_i}^{t_f}\int_{x_1}^{x_2} dt\,dx\,\delta u(x,t)\left[\mu\partial^2_{tt}u(x,t) - Y\partial^2_{xx}u(x,t)\right] + \text{boundary terms} \tag{15.13}$$

And, as anticipated, the boundary terms are two, namely

$$B_1 = \mu\int_{x_1}^{x_2} dx\;[\delta u(x,t)\partial_t u(x,t)]\,|_{t_i}^{t_f}\,, \tag{15.14}$$

and

$$B_2 = Y\int_{t_i}^{t_f} dt\;[\delta u(x,t)\partial_x u(x,t)]\,|_{x_1}^{x_2}\,. \tag{15.15}$$

B_1 is nill by virtue of the standar time boundary conditions 15.11. B_2 must be delicately examined.

The variational principle implies the wave equation form the bulk integral, but does also imply $B_2 = 0$. And there are several ways in which this may happen. Let us examine just two,

Fixed boundary conditions, these correspond to imposing some all time fixed values ($\delta u(x_1,t) = \delta u(x_2,t) = 0$) to $u(x_1,t)$ and $u(x_2,t)$.

Free boundary conditions, these corrrespond to imposing $\partial u(x,t)|_{x_1} = \partial u(x,t)|_{x_2} = 0$ for all times.

THe lesson to be learned here is that, in field theory, the variational principle does also lead to space boundary conditions.

Excercise: What other boundary conditions can you think of?

15.2 Actions for relativistic fields

When dealing with relativistic fields, we postulate the following requisites for the action.

1. We want it to be manifestly covariant under the Poincaré group.

2. The kinetic term should contain first derivatives at most, the reason being that we want second order differential equations at most.

3. We would like the kinetic term to contain the lowest possible powers of the fields or its derivatives (we want to avooid nonlinearities as much as possible)

Being the above said, the action has the following general look

$$S = \int_V d^4x\, \mathcal{L}(\phi, \partial_\mu \phi) \tag{15.16}$$

The equatins of motion follow from the variational principle

$$\delta_\phi S = \text{boundary term} + \int_V d^4x\, \delta\phi \left[\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right] \tag{15.17}$$

The bounddary term is an exact divergence, namely

$$\text{boundary term} = \int_V d^4x\, \partial_\mu \left[\delta\phi \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right] , \tag{15.18}$$

Which implies

$$\text{boundary term} = \int_{\partial V} ds\, n_\mu \left[\delta\phi \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right] , \tag{15.19}$$

where n_μ is the unitary vector pointing outside the four dimensional volume V, and ds is the three dimensional volume wich constitutes the boundary of V.

From the variational principle $\delta S = 0$, the boundary term fixes the boundary conditions for the fields, and the differential equations giving their dynamiics are just

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0 \tag{15.20}$$

15.3 A simple Action for a vector field

Our first approach, known as second order formalism consists in regarding a covariant vector field ($A_\mu(x)$) as the only independent field, from which we construct the second order tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \tag{15.21}$$

this object, containing derivatives of the dynamical field must appear in the kinetic part of the action.

There are at least to scalars that can be built form F, namely

$$F^{\mu\nu} F_{\mu\nu} \quad \text{and} \quad \epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu} \,, \tag{15.22}$$

for pedagogical reasons I will leave the latter out form the discussion (we will turn back to it in the future)

Considering the requisites to build an action we give a try to the following

$$S = \langle\, \alpha\, F_{\mu\nu} F^{\mu\nu} + \beta A_\mu\, J^\mu \rangle \,. \tag{15.23}$$

Notice the prescence of the vector field $J^\mu(x)$, which is to be considered as *external* (given), for reasons that will be clear very soon, we will call it, the *external current*.

We might have tried other possibilities for the action, but, for the moment, I prefer to stick with this one.

The equations of motion (including boundary conditions) are gotten from the variational principle

$$\delta_A S = 0 \,. \tag{15.24}$$

The variation of the kinetic term yields

$$\begin{aligned} \delta_A\, F_{\mu\nu} F^{\mu\nu} &= 2\, [F^{\mu\nu} \delta\; (\partial_\mu\, A_\nu - \partial_\nu\, A_\mu)] = \\ &= 4\partial_\mu\, [F^{\mu\nu} \delta A_\nu] - 4\, [\partial_\mu F^{\mu\nu}]\, \delta A_\nu \end{aligned} \tag{15.25}$$

After getting rid of the boundary term by impossing boundary conditions on A_μ, the variational principle reads

$$\delta_A S = \langle\, -4\alpha\; [\partial_\mu F^{\mu\nu}]\, \delta A_\nu + \beta\; J^\mu \delta A_\mu \rangle = 0 \,, \tag{15.26}$$

which implies

$$\partial_\mu F^{\mu\nu} = \frac{\beta}{4\alpha}\, J^\mu \tag{15.27}$$

We know that the nonhomogeneous Maxwell's equations are

$$\partial_\mu F^{\mu\nu} = \frac{\beta}{4\alpha}\, J^\mu\,, \tag{15.28}$$

we then realize that if we impose

$$\frac{\beta}{4\alpha} = \frac{4\pi}{c}\,, \tag{15.29}$$

the equations of motion we havee just obtained resemble the set of nonhomogeneous Maxwellś equations.

It is clear that the above observation regarding the constants is up to a common sign. At this point we recall that the interaction term between a charged particle and an electromagnetic feld is given by the following formula in the respective Lagrangian:

$$L_{int} = -\frac{1}{c} A_\mu\, eU^\mu\,, \tag{15.30}$$

besides, a point particle, gives rise to the following current vector

$$J^\mu = eU^\mu\,. \tag{15.31}$$

This last formula suggests that we choose $\beta = -$ and $\alpha = -\frac{1}{16\pi}$ to reproduce the set of inhomogeneous Maxwells equations form the action

$$S = -\langle\, \frac{1}{16\pi}\, F_{\mu\nu} F^{\mu\nu} + \frac{1}{c} A_\mu\, J^\mu \rangle\,, \tag{15.32}$$

At first sight, it seems we have done a nice job finding an action for Maxwellś electrodynamis, but is this really true?

In first place, one might ask about the homogeneous Maxwellś equations, and the answer is simple, we already satisfied them identically when we defined $F_{\mu\nu}$.

In second place, we may look at the equation of motion

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c}\, J^\mu\,, \tag{15.33}$$

and take its derivative with respect to x^μ

$$\partial_\mu \partial_\mu F^{\mu\nu} = \frac{4\pi}{c}\, \partial_\mu J^\mu \,, \tag{15.34}$$

The lhs vanishes identically, since there we have the contraction of an antysimmetric tensor (F) with a symmetric one $\partial^2_{\mu\nu}$, this guarantees that the current is conserved $\partial_\mu J^\mu$.

There is a delicate point here, recall that J^μ is an external current, and we obiously want to think of it as an electromangentic current. Consequently, the current must be identically conserved, wether F is the Faraday tensor or not!, what we have just seen is that we can guarantee current conservation when the equations of motion are satisfied (this is called ON-SHELL), but the current should be conserved always, even off shell.

Why am I making such a big fuss about this?, well, it is because of quantum mechanics. We must go quantum at some point, and the sum over histories forces us to add over all field configurations even if they are not solutiond of the classical field equations, i.e. we must go off shell in the sum over histories, but, since the current is external, it must be concerved always!.

The solution to this impasse comes in from one of the most beautiful principles, gauge invariance!. The action must be gauge invariant, i.e. tha action must be invariant under $A_\mu \to A_\mu + \partial_\mu \Lambda$

Let us give this a try

$$S_\Lambda = -\langle \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{c}\left(A_\mu - \frac{1}{c}\partial_\mu \Lambda\right) J^\mu \rangle = S - \frac{1}{c}\langle \Lambda\, \partial_\mu J^\mu \rangle - \frac{1}{c}\langle \partial_\mu \left(\Lambda\, J^\mu\right)\rangle \,, \tag{15.35}$$

the first term (F^2) is gauge invariant by construction, the last term, mamely $\langle \partial_\mu \left(\Lambda\, J^\mu\right)\rangle$ vanishes due to the boundary conditions on the currents, leaving us with the identity

$$S_\Lambda = S - \frac{1}{c}\langle \Lambda\, \partial_\mu J^\mu \rangle \,, \tag{15.36}$$

which implies that the action is gauge invariant iff the current is conserved off shell (which is because in the case under consideration, it is an external current).

We have therefore shown, that the action

$$S = -\langle \frac{1}{16\pi}\, F_{\mu\nu} F^{\mu\nu} + \frac{1}{c} A_\mu\, J^\mu \rangle \,, \tag{15.37}$$

Seems (under the most rigid tests we have at hand) perfectly well suited to represent the action for Maxwell's theory.

We must keep this in mind. When dealong with quantum theories, gauge symmetries can be broken due to quantum corrections, such breakings on gauge invariance that come form quanum corrections are referred to as anomalies and represent the kiss of dead of any theory. The only way aound to save a theory showing abomalies is their cancellation due to some kind of *magic*.

15.4 The first order Maxwell action

We now consider an external current, and two independent dynamical fields, namely, a vector field A_μ and a second order anystmmetric tensor field $F_{\mu\nu}$. From what we have already learned, the follwing acton seems natural

$$S = \langle\, \alpha\, F^{\mu\nu} F_{\mu\nu} + \gamma\, F^{\mu\nu} \partial_\mu A_\nu + \beta A_\mu\, J^\mu \rangle \tag{15.38}$$

We notice that F appears only algebraically, its derivatives never enter the action. We now have to impose the variational principle for variations on all indpentdent fields. A simple calculation yields (dropping of boundary terms is assumed)

$$\delta_F\, S = \langle\, 2\alpha\, \delta F^{\mu\nu} F_{\mu\nu} + \frac{\gamma}{2}\, \delta F^{\mu\nu}\, (\partial_\mu A_\nu - \partial_\nu A_\mu) \rangle = 0\,, \tag{15.39}$$

and

$$\delta_A\, S = \langle\, -\gamma\, \partial_\nu F^{\nu\mu} \delta A_\mu + \beta J^\mu \delta A_\mu \rangle = 0 \tag{15.40}$$

Equation 15.39 gives, F in terms of A, in fact, if we choose the normalization to ensure that

$$4\alpha = -\gamma \tag{15.41}$$

we find

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{15.42}$$

On the other hand, equation 15.40 yields

$$\frac{\beta}{\gamma} = \frac{4\pi}{c}\,, \tag{15.43}$$

Choosing $\beta = -c^{-1}$ imposes $\gamma = -\frac{1}{4\pi}$ and therefore, $\alpha = \frac{1}{16\pi}$.

We have thus found the following action that seems to be a suitable action for electrodynamics

$$S = \langle \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} - \frac{1}{4\pi} F^{\mu\nu} \partial_\mu A_\nu - \frac{1}{c} A_\mu J^\mu \rangle \tag{15.44}$$

Once again, the off shell gauge invariance of the action depends on the current to be conserved.

This is an excellent point for a discussion. 15.44 is **classically equivalent** to action 15.32, this means that the solution of the field equations of one action are in biyection with the soluitons of the other.

A nice excercise is the following, reexpress the second term in action 15.44 as

$$\langle -\frac{1}{4\pi} F^{\mu\nu} \partial_\mu A_\nu \rangle = \langle -\frac{1}{8\pi} F^{\mu\nu} \left(\partial_\mu A_\nu - \partial_\nu \right) \rangle \tag{15.45}$$

which due to the equations of motion is just

$$\langle -\frac{1}{4\pi} F^{\mu\nu} \partial_\mu A_\nu \rangle = \langle -\frac{1}{8\pi} F^{\mu\nu} F_{\mu\nu} \rangle \tag{15.46}$$

Substitution on the action converts it in to action 15.32.

Once again, this is an on shell equivalence and there is no reason a priori to think that the equivalence of both formulations will survive the quantization procedure.

15.5 A litle bit of dimensional analysis

Let us consider the structure of the kinetic term of the second order Maxwell action, namely

$$d^4x\, FF \sim d^4x (\partial A)^2 \,. \tag{15.47}$$

These quantities must have dimensons of action (units of $\hbar$. In clever units, $\hbar] = 1$ so

$$1 = L^4 \times L^{-2} \times [A]^2 \,. \tag{15.48}$$

Or

$$M \times L^2 \times T^{-1} = L^2 \times [A]^2 \,, \tag{15.49}$$

Meaning that the vector potential has units of reciprocal length (mass).

15.6 The Proca Lagrangian

having learned the little bit of dimensional analysis we did in the previous section, it i clear that an action for a covariant vector field might inlcude a new term as follows

$$S = \langle -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{m^2}{8\pi}\, A^\mu A_\mu - \frac{1}{c} A_\mu J^\mu \,\rangle \,, \tag{15.50}$$

where $F_{\mu\nu}$ is given as the "curl" of A. and m has units of mass, (reinsating units $m \to mc/\hbar$. Given the interpretation of the first and last terms as the satandard Maxwell action, this new action, called the Proca action should be interpreted as the action of a massive photon.

We begin by noticing that the mass term completely spoils gauge invariance. Indeed, under $A_\mu \to A_\mu + \partial_\mu \Lambda$ the mass term behaves as

$$\langle A^\mu\, \partial_\mu \Lambda \rangle = \text{boundary term} + \langle \Lambda \partial^\mu A_\mu \rangle \,, \tag{15.51}$$

which vanishes iff $\partial_\mu A^\mu = 0$ which, by no means is a gauge fixing condition. And it is in fact a necessary condition for current conservation.

A simple excercise shows that the nonhomogeneous equations of motion are

$$\partial_\mu\, F^{\mu\nu} + m^2\, A^\mu = \frac{4\pi}{c}\, J^\mu \,, \tag{15.52}$$

here we notice that even on shell, current conservation is valid iff $\partial_\mu\, A^\mu = 0$

The differences beteween the Proca and Maxwell lagrangian are huge. I will comment on something that might sound a little strange.

The dynamical field contents of the two theories are completely different, the Maxwell Lagrangian gives rise to two propagating degrees of freedom only. Indeed, if we think of waves in vacuum, if one gives a propagation vector, the electric field must be orthogonal to it, meaning that the polarization vector $(\vec{\mathcal{E}})$ is two dimensional, and the magnetic induction is completely constrained to be essentially $\vec{k} \times \vec{\mathcal{E}}$.

The massive vector field, on the other hand, has three degrees of freedom and clearly, it is not possible to change the number of polarizations continuosly.

15.7 A final excercise

In this section, we just want to ask a question, what dynamics cones from the simple action

$$S =< \epsilon^{\mu\nu\rho\alpha} F_{\mu\nu} F_{\alpha\rho} > , \tag{15.53}$$

where $F \sim \partial A$

The answer is simple, variations of S yield

$$\delta_A S =< 2\epsilon^{\mu\nu\rho\alpha} F_{\mu\nu} \delta F_{\alpha\rho} > , \tag{15.54}$$

$$\delta_A S =< 4\epsilon^{\mu\nu\rho\alpha} F_{\mu\nu} \partial_\rho \delta A_\alpha >=< 4\left(\epsilon^{\mu\nu\rho\alpha} \partial_\rho F_{\mu\nu}\right) \delta A_\alpha > + boundary\,, \tag{15.55}$$

And now we just notice that $\epsilon^{\mu\nu\rho\alpha}\partial_\rho F_{\mu\nu}$ vanishes identically, meaning that the equations of motion are trivial, there is no dynamics here.

This short exercise can be approached in a slightly different form by noticing that the action is an exact divergence.

15.8 Closing remarks

In this discussion the current has always been thought as given (external current) this is not very reallistic.

A reasonable way to improve on this would be to find a field representing charged matter (particles) and fnding a way to couple it to the Maxwell Lagrangian. This is exactly what is done in QED.

For the sake of completeness we will briefly discuss scalar electrodynamics.

One begins by considering a complex scalar field $\phi(x)$, and propose the simplest possible action, which by dimensional considerations has the form

$$S = \langle \frac{1}{2} \partial_\mu \phi^* \partial^\mu \phi + \frac{m^2}{2} \phi^* \phi \rangle \tag{15.56}$$

The equation of motion for this action is the Klein Gordon equation

It is clear that S in invariant under the field transformations[2]

$$\phi \to e^{i\alpha}\phi \tag{15.57}$$

where α is a real constant[3]

Noether's theorem shows that the current

$$J^\mu = \partial^\mu \phi^* \, \phi - \phi^* \partial^\mu \phi \tag{15.58}$$

is conserved.

The next step in the standard field theoretic approach to interactions is to ask about what happens if one considers *local* $U(1)$ transformations, i.e. transformatons where the phase is given by

$$e^{i\alpha(x)} \tag{15.59}$$

where now, $\alpha(x)$ is a function taking real values.

The result is that the invariance is spoiled due to the derivatives acting on the phases and one tries to restore it by introducing a *covariant derivative operator* defined as

$$\nabla_\mu = \partial_\mu - ie\, A_\mu(x) \tag{15.60}$$

where e is a constant (not euler's) to be identified later on.

Having done this, one realizes that the invarriance if the action under local phase transformations is restored provided the vector transforms as

$$A_\mu \to A_\mu + \partial_\mu \alpha \tag{15.61}$$

i.e. a gauge trnasformation of the elctrodynamic kind.

Expanding the action in full one finds a term $e\, A_\mu \, J^\mu$. Reabsorbing the constant in J^μ, we may identify the latter as the electric current. We have thus identified our action as the action of a charged fiedl in interaction with an external electromagnetic field A_μ, to have something more interesting, we must introduce a dynmaics for the *Gauge field* (A_μ), and this is achieved by nothng less than Maxwell's action.

---

[2]Notice that these transformations do not affect the space-time, this kind of transformations actong on the fields only are, historically, referred to as internal transformations

[3]$e^{i\alpha}$ is an element of $U(1)$ and therefore we say that the action is invariant under $U(1)$

BLAH BLAH no da dinamica ... es topológica .....
Chern Symmons ... en $3D$ ...

$$S = < A \wedge dA > \tag{15.62}$$

Bibliography

[1] A. Pais. *Subtle is the Lord : The Science and the Life of Albert Einstein: The Science and the Life of Albert Einstein*. Oxford University Press, USA, 1982.

[2] Bernard Schutz. *A First Course in General Relativity*. Cambridge University Press, 2 edition, 2009.

[3] E. F. Taylor and J.A. Wheeler. *Spacetime Physics*. W. H. Freeman and Co., 2 edition, 1992.

[4] N.M.J. Woodhouse. *Special Relativity*. Springer Undergraduate Mathematics Series. Springer London, 2012.

[5] L. Landau and Evgeny Lifshitz. *The Classical Theory of Fields: Volume 2*. Course of theoretical physics. Elsevier Science, 1975.

[6] John David Jackson. *Classical electrodynamics*. Wiley, New York, NY, 3rd ed. edition, 1999.

[7] Charles W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

[8] J.L. Anderson and J.L. Anderson. *Principles of Relativity Physics*. Academic Press, 1967.

[9] W. Rindler. *Relativity: Special, General, and Cosmological*. Relativity: Special, General, and Cosmological. Oxford University Press, 2001.

[10] T. Padmanabhan. *Gravitation: Foundations and Frontiers*. Cambridge University Press, 2010.

[11] Tyler J. Lemmon and Antonio R. Mondragon. Kepler's orbits and special relativity in introductory classical mechanics. arXiv:1012.5438 [astro-ph.EP].

[12] Tyler J. Lemmon and Antonio R. Mondragon. Alternative derivation of the relativistic contribution to perihelic precession. *American Journal of Physics*, 77(10):890–893, oct 2009.

[13] Jr. T. Phipps. Mercury's precession according to special relativity. *American Journal of Physics*, 54(3):245, 1986.

[14] S.M. Carroll. *Spacetime and Geometry*. Cambridge University Press, 2019.

[15] R. d'Inverno and J. Vickers. *Introducing Einstein's Relativity: A Deeper Understanding*. Oxford University Press, 2022.

[16] S. Weinberg and W. Steven. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley, 1972.