Treball final de màster

MÀSTER DE MATEMÀTICA AVANÇADA

Facultat de Matemàtiques
Universitat de Barcelona

A MATHEMATICAL APPROACH TO GENERAL RELATIVITY THEORY

Autor: Pere Joan Falcó Solsona

Director:	**Dr. Carlos Currás Bosch**
Realitzat a:	**Departament de Àlgebra i Geometria**

Barcelona, 11 de gener de 2015

Acknowledgements

To my tutor, Dr. Carlos Currás, who gave me the idea of the project and has led me during the process of developing it, making me discover the outlines of Einstein's theory and helping me to learn much more than i had ever imagined when I started.

And, specially, to Jaume Salvó, who was my math's teacher in my last two years in high school; without him I would have never discovered the wonders of Mathematics nor considered the possibility of devote all these years to this science I love so much.

Contents

Introduction

When Albert Einstein presented in 1905 his paper *"Zur Elektrodynamic bewegter Körper"* ("On the Electrodynamics of Moving Bodies"), phycisists all over the world accepted with no hesitation the classical Physics mainly developed by Sir Isaac Newton back in the 17th century. However, some problems with this theory had rised due to the discovery that light is a kind of electromagnetic wave in the 19th century, since it was believed that such waves needed a medium through which be propagated. This fact forced the physicists of the moment to postulate the existence of the so-called ether, a theoretical medium filling the whole universe and acting as the propagation medium that the light electromagnetic wave needed, according to what was known at the time.

Nevertheless, this assumption led to more problems when Michelson and Morley performed an experiment trying to detect the effects of the postulated presence of this light-bearer, presence needed for the current theory to work, and failed. The results contradicted the theoretical existence of ether, and no one could give a satisfactory explanation to this fact.

It was in this moment when Einstein proposed his Special Relativity theory in 1905, which solved the problem of motion and speed of light and was in agreement with classical Physics regarding kinematics, dynamics and electromagnetism.However, Special Relativity was unable to give a good explanation about what has to do with gravitation as Newtonian theory did.

So in 1916, Einstein published his paper *"Die Grundlage der allgemeine Relativitätstheorie"* ("The foundations of General Relativity"), including the Laws of Universal Gravitation, developing a theory which has been proved to be more accurate to that of Newton.

The aim of this project is to outline the Special Relativity and the General Relativity theory, paying special attention to some mathematical concepts, such as Lorentz transformations and the Ricci tensor, which happen to be key points to this theory which revolutionated Physics in the beginning of last century. We assume that the reader have some previous mathematical knowledge, specially about Differential Geometry, and will therefore skip definitions and theorems regarding this area of Mathematics (such as the definition of connexion, Riemannian metrics or curvature). We shall give some notions of classical Physics in the first chapter, in order to look over them in Chapters 2 and 3 at light of Special Relativity and General Relativity theories, respectively.

Chapter 1

Physical preliminaries

The aim of this first chapter is to provide the reader with some basic concepts of classical Physics which shall be of further use in the next chapters to develop properly the goal of the project.

1.1 Kinematics and Newtonian dynamics of a material particle

Newtonian Kinematics

We shall consider that physical bodies are in $\mathbb{R}^3$, whose canonical basis and origin are considered fixed. Moreover, we shall consider that we can know the exact instant of time in which an event takes place, and that this time t is universal for every particle.

If a material particle P moves, it will be in a position $P(t) \in \mathbb{R}^3$ at every instant t in such a way that $P(t)$ is a curve on $\mathbb{R}^3$ parametrized by time. Moreover, we shall consider that every curve described by the motion of a particle is a differential curve (it will usually suffice to consider that $P(t)$ is $\mathcal{C}^2$).

If $P(t)$ describes the motion of a particle P depending on time, we can define the vectors velocity and acceleration at the instant t as $\vec{v}(t) = d\vec{P}(t)/dt$ and $\vec{a} = d^2\vec{P}/dt^2$. We shall say that the particle P has a uniform motion if the vector $\vec{v}$ is constant.

Moreover, having two particles P and O we can describe the relative velocity and the relative acceleration of P with respect to O as the vectors $d(P(t) - O(t))/dt$ and $d^2(P(t) - O(t))/dt^2$, respectively.

Fundamental law of Newtonian dynamics

It is assigned to every physical body a real positive number m, the so-called mass, which satisfes the following properties:

i) If a material body K_1 has a corresponding mass m_1 and another material body K_2 has a corresponding mass m_2, then the union body $K_1 \cup K_2$ has a corresponding mass $m_1 + m_2$.

ii) Any material body K can be divided in two bodies K_1 and K_2 with the same corresponding mass.

Let C denote a criterion such that, given two bodies K_1 and K_2, we could determine either which one of them has bigger mass or if both have the same mass.

Theorem 1.1.1. *If we were to have a criterion C as defined above, then the process of given a material body its mass would be uniquely determined by conditions i) and ii), and by the choice of an arbitrary material body U taken as unite of measure, i.e., whose mass is considered to be 1.*

Observe that the definition of mass depends on the criterion C. There exist several candidates to be the criterion C used in Theorem 1.1.1, and we shall give a couple of examples of valid criterions to compare masses.

If in a inertial region - that is, a region of space in which every body left in it either stays stopped or describes a straight path with constant velocity -, two bodies K_1 and K_2 whose masses want to be compared are left in such a way that their motions are on the same straight path and they collide, a candidate to be criterion C is giving the bigger mass to the body whose difference of velocities before and after the collision is smaller.

Definition 1.1.2. *When a material particle with mass m which is in repose or moves with uniform motion suffers an acceleration $\vec{a}$, we say that on this particle has acted a force, which is defined as $m\vec{a}$.*

Observe that this definition contains the inertia principle, by which a particle which is in repose or has a uniform motion remains in its state as long as no force acts on it.

Momentum, kinetic energy, power, work

Let P be a material particle with mass m which moves describing a path $\vec{P}(t)$. We define the momentum of P in the instant t as the vector $\vec{\mathcal{P}}(t) = m \cdot \vec{v}(t)$, where $\vec{v}(t)$ is the velocity of P at t. Observe that $F = m\vec{a} = md\vec{v}/dt = d\vec{\mathcal{P}}/dt$. Thus, if no force acts on a body, the momentum of this body remains constant.

With the same notation, we call kinetic energy of a particle P at the instant t to the scalar $\frac{1}{2}m\|\vec{v}(t)\|^2$, and we call power in the instant t to the scalar product $\vec{F}(t) \cdot \vec{v}(t)$, where $\vec{F}(t)$ is the force which acts on the particle at the instant t. Finally, we define the work between the instants t_0 and t_1 as

$$\int_{t_0}^{t_1} \vec{F}(t) \cdot \vec{v}(t)dt.$$

One has

$$work = \int_{t_0}^{t_1} \vec{F}(t) \cdot \vec{v}(t)dt = m \int_{t_0}^{t_1} d\vec{v}(t) \cdot \vec{v}(t)dt = \frac{1}{2}[m\vec{v}(t) \cdot \vec{v}(t)]_{t_0}^{t_1}.$$

Thus, the work between t_0 and t_1 coincides with the variation of kinetic energy between these two instants.

If a material body K is formed of n particles $P_1, ..., P_n$ we call momentum, kinetic energy, power and work of K to the sum of the momentums, kinetic energies, powers and works of the n particles which form K, respectively. If on a system of particles acts no force, the momentum of the system remains constant.

Forces resulting from a potential

Assume that, in a connected open set $U \subset \mathbb{R}^3$, for each $p \in U$, each scalar $m \in \mathbb{R}^+$ and each instant t, we have a vector $\vec{F}(m,p,t)$, and assume that every particle with mass m left in repose in a point p at the instant t suffers an acceleration

$$\vec{a} = \frac{1}{m}\vec{F}(m,p,t).$$

Then one says that there is a force field in U. One says that a force field results from a potential if it does not depend on t and there exists a smooth function $V(m,p)$, called potential, such that for $p = (x,y,z)$ and for any $p_0 \in U$ we have

$$\vec{F}(m,p_0) = -\nabla_{p_0} V,$$

where

$$\nabla_{p_0} V = \Big((\frac{\partial V}{\partial x})_{p_0}, (\frac{\partial V}{\partial y})_{p_0}, (\frac{\partial V}{\partial z})_{p_0} \Big).$$

A virtual trajectory between two points p_1 and p_2 in $\mathbb{R}^3$ is a parametrized curve in $\mathbb{R}^3$, piecewise $\mathcal{C}^2$, joining p_1 and p_2. Knowing the forces that act on a particle at an instant t (for instance, in a force field), the motion of this particle is determined; this would be a real trajectory. Among the curves which start in p_1, only a few can be pathed by material particles following the law of motion $\vec{F} = \vec{a}m$. Thus we call virtual trajectory to any curve which starts in p_1, whether it can be pathed by a particle following the law of motion or not. Moreover, if U is a connected open set in $\mathbb{R}^3$ with a force field $\vec{F}(m,p,t)$, p_1, p_2 two points of U and $C(t)$ a virtual path joining p_1 and p_2 with $C(t_1) = p_1$ and $C(t_2) = p_2$, we call virtual work of a particle P with mass m along $C(t)$ between p_1 and p_2 to the integral

$$\int_{t_1}^{t_2} \vec{F}(m, C(t), t) \cdot \vec{C}'(t) dt.$$

Thus, the virtual work is the work associated to a particle of mass m if it were to follow the path $C(t)$.

Theorem 1.1.3. *Let U be a connected open set in $\mathbb{R}^3$ and $\vec{F}(m,p)$ a force field over U not depending on t. Then, $\vec{F}$ results from a potential if, and only if, the virtual work between any pair of points p_1, p_2 of U along any virtual path connecting both points does not depend on the virtual path.*

We define the rotational of a vector field $\vec{F} = (F^1, F^2, F^3)$ as

$$\vec{rot}\vec{F} = \Big(\frac{\partial F^3}{\partial y} - \frac{\partial F^2}{\partial z}, \frac{\partial F^1}{\partial z} - \frac{\partial F^3}{\partial x}, \frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \Big).$$

Observe that if a force field $\vec{F}(m,p)$ results from a potential, then $\vec{rot}\vec{F} = 0$; and that if the region U is simply connected and $\vec{rot}\vec{F} = 0$, then F results from a potential over U.

Conservation of energy

Let a particle P with mass m moving in an open set $U \subset \mathbb{R}^3$ under the action of a force field $\vec{F}$ on U. Let F be resulting from a potential V and let $P(t)$ be the path of the particle. Then, the work between the instants t_0 and t_1 is

$$\int_{t_0}^{t_1} \vec{F} \cdot \frac{d\vec{P}}{dt} dt = -\int_{t_0}^{t_1} \nabla V \cdot \frac{d\vec{P}}{dt} dt =$$

$$= -\int_{t_0}^{t_1} \frac{d}{dt} V(P(t)) dt = V(P(t_0)) - V(P(t_1)).$$

Recall that the variation of work is equal to the variation of kinetic energy, denoting by $E_K(t)$ the kinetic energy at the instant t. Then

$$E_K(t_1) - E_K(t_0) = V(P(t_0)) - V(P(t_1)).$$

Thus,

$$E_K(t_0) + V(P(t_0)) = E_K(t_1)) + V(P(t_1)),$$

which means that the sum of kinetic energy and potential is constant. This fact is known as the conservation of energy Theorem.

1.2 Lagrange equations

Systems of particles with holonomic constraints non-depending on time

Given a body formed of n particles $P_1, ..., P_n$ of masses $m_1, ..., m_n$, let (x_i, y_i, z_i) be the coordinates of the position of P_i and let $f_1, ..., f_s$ be real funcions with variables in $\mathbb{R}^{3n}$. We say that $P_1, ..., P_n$ is a system of particles with holonomic constraints non-depending on time if they fullfil the equations

$$f_j(x_1, y_1, z_1, x_2, ..., z_{n-1}, x_n, y_n, z_n) = 0 \tag{1.1}$$

for $j = 1, ..., s$. From now on, we shall consider the constraints to be holonomic non-depending on time, unless we say otherwise.

We will usually call $x^1, ..., x^{3n}$ to the variables $x_1, y_1, z_1, ..., x_n, y_n, z_n$ and assume that $s < 3n$, that the functions f_j are smooth and that for any $x_0 \in \mathbb{R}^{3n}$ satisfying $f_j(x_0) = 0$ for all j, the Jacobian

$$\left(\frac{\partial f_j}{\partial x^i}\right)$$

has rank s. Then the equations 1.1 define a smooth submanifold M of dimension $r = 3n - s$ in $\mathbb{R}^{3n}$; we shall call this submanifold the manifold of positions of the system, since each point of M correponds to the coordinates of the position of each particles satisfying the constraints. Moreover, every x_0 has a neighbourhood in which M can be expressed as

$$\begin{cases} x^1 = & x^1(q_1, ..., q_r) \\ \dots & \dots\dots\dots\dots \\ x^{3n} = & x^{3n}(q_1, ..., q_r). \end{cases}$$

The parameters $q_1, ..., q_r$ are the so-called generalized coordinates of the system and the dimension r is the degree of freedom of the system. We shall now see some examples.

Example 1.2.1. [The simple pendulum]

The pendulum is a good example of a system of particles with constraints. Let P be a particle of mass m in $\mathbb{R}^2$ fixed at the end of a bar of lenght l without mass which is fixed by the other end to a fixed point O. Taking O as origin and taking x, y the coordinates of the position of P, such a position is determined by the angle φ which the bar forms with the $y-$axis. Thus, due to the time-dependance of the angle, the coordinates of the position of P are given by

$$\begin{cases} x(t) = l \sin \varphi(t) \\ y(t) = -l \cos \varphi(t) \end{cases}$$

In this case, the generalized coordinate is $\varphi(t)$.

Virtual velocities of the system: tangent bundle

Let $P_1, ..., P_n$ be material particles with constraints. For every $i \in \{1, ..., n\}$, we denote by π_i the canonical projection

$$\begin{array}{ccc} \mathbb{R}^{3n} & \xrightarrow{\pi_i} & \mathbb{R}^3 \\ (x_1, y_1, z_1, ..., x_n, y_n, z_n) & \longmapsto & (x_i, y_i, z_i). \end{array}$$

Observe that if $p \in M$, being M the manifold of positions of the system, then $p_i = \pi_i(p)$ corresponds to the position of the particle P_i. Assuming that the particles move describing curves $p_i(t) = (x_i(t), y_i(t), z_i(t))$, we have that only those curves satisfying $p(t) = (p_1(t), ..., p_n(t)) \in M$ for all t satisfy the constraints, and then are the only possible trajectories. Thus, we can define a virtual velocity of the system at a position p as any vector $v \in T_pM$, which is a possible velocity corresponding to a possible trajectory of the particles of the system. Moreover, observe that, given a virtual velocity $v \in T_pM$, $v_i = (\pi_*)v$ is the virtual velocity of the particle P_i corresponding to the virtual velocity v of the system. The set of all virtual velocities of the different points of M, $T(M) = \bigcup_{q \in M} T_qM$, is endowed with a smooth manifold structure by considering the tangent bundle structure.

Constraits forces and virtual powers principle

Let $P_1, ..., P_n$ be a set of particles of masses $m_1, ..., m_n$ which are submitted by constraints, and assume that these particles are in a connected open set $U \subset \mathbb{R}^3$. If on each P_i acts a force F_i, setting $F = (F_1, ..., F_n)$ and letting under the same notation $v = (v_1, ..., v_n)$ a vector of virtual velocities, we define the virtual power of the system as the sum of scalar products $\sum_{i=1}^{n} F_i \cdot v_i$. Assuming now that there is a force field $\vec{F}(m, p, t)$ in U, one can observe that the particles of the system will not move as if they were independent because of the constraints, and so, taking $\vec{a}_i(t)$ the acceleration suffered by the particle P_i at the instant t, the equations $\vec{F}(m_i, p_i, t) = m_i \cdot a_i(t)$ will not be real because they do not reflect the forces caused by the constraints. Therefore, one defines

$$\vec{\Phi}_i(t) = m_i \cdot \vec{a}_i(t) - \vec{F}(m_i, p_i, t),$$

where Φ_i are the so-called constraints forces which act on P_i.

We shall recall now a theorem, due to D'Alembert, which shall be taken as an axiom.

Axiom 1.2.2. *For every virtual velocity $\vec{v}$ of the system, the constraints forces are such that $\sum \vec{\Phi}_i \cdot \vec{v}_i = 0$, that is, the virtual power corresponding to $\vec{v}$ vanishes.*

We shall now check that the axiom is satisfied in the simple pendulum example that we gave above.

Example 1.2.3. [The simple pendulum]

Consider a pendulum as described in Example 1.2.1. The constraints force acting on P has direction PO since P is joined to the fixed point O through PO. Since any virtual velocity is tangent to the circle described by P in its motion, one has that the scalar product of the constraints force by any virtual velocity will be zero. Let us perform a further analysis of this case, since the constraints force, usually difficult to calculate, is easy to compute in it. Assume that the pendulum moves uniquely due to a given initial velocity, with no other external forces acting on it. Let $\vec{p}(t)$ be the so-called position vector of P, whic depends on the angle φ; let also l be the length of PO. Then, since φ depends on the time, we have

$$\vec{p}(t) = (l \sin \varphi(t), -l \cos \varphi(t)).$$

Differentiating with respect to time, one gets

$$\dot{p} = (l\dot{\varphi} \cos \varphi, l\dot{\varphi} \sin \varphi).$$

Now, setting $\vec{T}(t)$ the unit tangent vector

$$\vec{T}(t) = \frac{\dot{p}(t)}{\|p\dot{(}t)\|} = (\cos \varphi(t), \sin \varphi(t)),$$

and $\vec{N}(t)$ the unit normal vector

$$\vec{N}(t) = \frac{p(t)}{\|p(t)\|} = (\sin \varphi(t), -\cos \varphi(t)),$$

one has $\dot{p} = l\dot{\varphi}\vec{T}$. And differenciating again, one gets

$$\ddot{p} = l\ddot{\varphi}T + l\dot{\varphi}\frac{dT}{dt} = l\ddot{\varphi}T - l\dot{\varphi}^2 N.$$

Thus, since we have assumed that there are no external forces, the constraints force acting on P must be equal to $m\ddot{p}$; but the constraints force has direction PO, and so the tangent part of $m\ddot{p}$ has to be zero. Then, $\ddot{\varphi} = 0$, which implies that $\dot{\varphi}$ =constant=initiial angular velocity. Therefore, the constraints force is $m\ddot{p} = -ml\dot{\varphi}^2 N$ (the centripetal force).

On the other hand, if one assumes that the pendulum is under the action of a force field $\vec{F}$, at each instant of time one can decompose $\vec{F}$ in the normal part and the tangent one with respect to the circle described by the motion of P, that is, $\vec{F} = \vec{F}_1 + \vec{F}_2$. If we call $\vec{\Phi}$ to the constraint force acting on P, one has

$$\vec{F} + \vec{\Phi} = m\ddot{p} = ml\ddot{\varphi}T - ml\dot{\varphi}^2\vec{N},$$

where we have seen that Φ must be normal. Then the normal component of the expression above is
$$\vec{F}_1 + \vec{\Phi} = -ml\dot{\varphi}^2\vec{N},$$
and so we have $\vec{\Phi} = -ml\dot{\varphi}^2\vec{N} - \vec{F}_1$.

Motion equations in M

Let M be the manifold of positions of the system. We want to study now the differential equations of the motion of a system of particles $P_1, ..., P_n$ with constraints. Let $p_i(t)$ be the position of the particle P_i at the instant t. As we have seen, we have, for any $i \in \{1, ..., n\}$,

$$F_i + \Phi_i = m_i \frac{d^2 p_i(t)}{dt^2}. \tag{1.2}$$

These equations have the problem of the difficulty to compute the constraints forces in a general case. Thus, we shall find equations equivalent to equation 1.2 which do not contain the constraints forces; we shall do so by using the virtual powers principle. First of all, we set $x(t) = (p_1(t), ..., p_n(t))$, and so $\dot{x}(t) = (v_1(t), ..., v_n(t))$ where $v_i = \frac{dp_i}{dt}$. Fix $t = t_0$ and let w be any vector in $T_{x(t_0)}M$, which will be of the form $w = (w_1, ..., w_n)$ with $w_i \in \mathbb{R}^3$. Then, scalar-multiplying the equation 1.2 by w_i and applying the virtual powers principle, one gets
$$\sum_{i=1}^{n} \vec{F}_i \cdot \vec{w}_i = \sum_{i=1}^{n} m_i \frac{d^2 \vec{p}_i}{dt^2} \cdot \vec{w}_i. \tag{1.3}$$

Let us study the equation 1.3. Consider in $\mathbb{R}^{3n}$ two different scalar products. On one hand, the Euclidian product g and on the other one g_T, defined by

$$g_T(\vec{u}, \vec{w}) = \sum_{i=1}^{n} m_i \vec{u}_i \cdot \vec{w}_i.$$

Taking $\vec{F} = (\vec{F}_1, ..., \vec{F}_n)$, where F_i is the force acting on P_i, the equation 1.3 is written as

$$g(\vec{F}, \vec{w}) = g_T(\ddot{x}(t), w). \tag{1.4}$$

Observe now that $\ddot{x} = \nabla_{\dot{x}}\dot{x}$, being ∇ the usual covariant derivative and that, since the metrics g_T does not depend on the point, its associated covariant derivative will also be ∇. Let $\overline{\nabla}$ be the covariant derivative on M associated to g_T. Since $\nabla_{\dot{x}}\dot{x} = \overline{\nabla}_{\dot{x}}\dot{x}$+normal part in M, then, since $w \in T(M)$,

$$g_T(\ddot{x}, w) = g_T(\nabla_{\dot{x}}\dot{x}, w) = g_T(\overline{\nabla}_{\dot{x}}\dot{x}, w).$$

Then equation 1.4 can be written as

$$g(F, w) = g_T(\overline{\nabla}_{\dot{x}}\dot{x}, w). \tag{1.5}$$

Let $f \in T_{x(t_0)}M$ be the vector such that

$$g_T(f, w) = g(F, w) \ \forall w \in T_{x(t_0)}M.$$

Thus, the equation 1.5 is equivalent to

$$g_T(f,w) = g_T(\overline{\nabla}_{\dot{x}}\dot{x}, w)\ \forall w$$

and so

$$f = \overline{\nabla}_{\dot{x}}\dot{x}. \tag{1.6}$$

The vector f is the so-called generalized force. From now on we shall note ∇ the covariant derivative in M associated to g_T instead of $\overline{\nabla}$, unless those cases in which such notation could lead to confusion. Thus, the motion equation will be

$$f = \nabla_{\dot{x}}\dot{x}. \tag{1.7}$$

Lagrange form of the motion equations

We want to study the equation 1.7 in this subsection. To that purpose, we shall state the following proposition with the same notation as above and the one introduced right below.

Let (M,g) be a Riemannian manifold of dimension r and let $T(M)$ be its tangent bundle. We define the function related to the metrics g, which we denote also by g, as

$$\begin{array}{rccc} g: & T(M) & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & g_x(v,v). \end{array}$$

Taking $(U,x^1,...,x^r)$ a local chart on M, and $(U,x^1,...,x^r,v^1,...,v^r)$ a local chart on $T(M)$, the function g on $\pi^{-1}(U)$ is written as

$$g = g_{ij}v^iv^j,$$

where $g_{ij} = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. Given $x(t)$ a curve on M, we define the corresponding curve $\tilde{x}(t)$ on $T(M)$ as follows: given a local chart $(U,x^1,...,x^r)$ on M such that $x(t) = \big(x^1(t),...,x^r(t)\big)$, $\tilde{x}(t)$ is written on $\big(\pi^{-1}(U),(x^1,...,x^r,v^1,...,v^r)\big)$ as

$$\tilde{x}(t) = \big(x^1(t),...,x^r(t),\dot{x}^1(t),...,\dot{x}^r(t)\big).$$

We can now state the proposition.

Proposition 1.2.4. *Let $x(t)$ be a curve on M and $\tilde{x}(t)$ the curve on $T(M)$ corresponding to x. Let g be function on $T(M)$ corresponding to the metrics g of M. Let $(U,x^1,...,x^r)$ be a local chart of M and $(\pi^{-1}(U),x^1,...,x^r,v^1,...,v^r)$ the corresponding local chart in $T(M)$. For every $h = 1,...,r$, consider the function*

$$\frac{d}{dt}\Big(\frac{\partial g}{\partial v^h}\Big)_{\tilde{x}(t)} - \Big(\frac{\partial g}{\partial x^h}\Big)_{\tilde{x}(t)},$$

which is an ordinary function on t. Consider another function on t, $g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h})$. Then, it holds that

$$2g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h}) = \frac{d}{dt}\Big(\frac{\partial g}{\partial v^h}\Big)_{\tilde{x}(t)} - \Big(\frac{\partial g}{\partial x^h}\Big)_{\tilde{x}(t)}.$$

Proof. The components of $\nabla_{\dot{x}}\dot{x}$ on the basis $\left\{\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^1}\right\}$ are

$$\left(\nabla_{\dot{x}}\dot{x}\right)^k = \ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j.$$

Then we have

$$g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h}) = g_{hk}\ddot{x}^k + g_{hk}\Gamma^k_{ij}\dot{x}^i\dot{x}^j. \tag{1.8}$$

From the expression of the metrics with the Christoffel symbols we get, denoting $\partial_j = \frac{\partial}{\partial x^j}$,

$$g_{hk}\Gamma^k_{ij} = \frac{1}{2}\Big(\partial_j g_{hi} + \partial_i g_{jh} - \partial_h g_{ij}\Big).$$

Substituting these expressions in equation 1.8,

$$g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h}) = g_{hk}\ddot{x}^k + \frac{1}{2}\partial_j g_{hi}\dot{x}^i\dot{x}^j + \frac{1}{2}\partial_i g_{jh}\dot{x}^i\dot{x}^j - \frac{1}{2}\partial_h g_{ij}\dot{x}^i\dot{x}^j$$

and, since the second and the third terms are equal, we have

$$g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h}) = \frac{1}{2}g_{hk}\ddot{x}^k + \frac{1}{2}\partial_j g_{hi}\dot{x}^i\dot{x}^j - \frac{1}{2}\partial_h g_{ij}\dot{x}^i\dot{x}^j. \tag{1.9}$$

And, in $T(M)$

$$g = g_{sk}v^s v^k \Longrightarrow \frac{\partial g}{\partial v^k} = 2g_{hk}v^k \Longrightarrow$$

$$\Longrightarrow \left(\frac{\partial g}{\partial v^k}\right)_{\tilde{x}(t)} = 2g_{hk}\big(x^1(t), ..., x^n(t)\big)\dot{x}^k(t).$$

Then

$$\frac{d}{dt}\left(\frac{\partial g}{\partial v^k}\right)_{\tilde{x}(t)} = 2\partial_i g_{hk}\dot{x}^i\dot{x}^k + 2g_{hk}\ddot{x}^k$$

and, analogously,

$$\left(\frac{\partial g}{\partial x^h}\right)_{\tilde{x}(t)} = \partial_h g_{ij}\dot{x}^i\dot{x}^j.$$

Therefore,

$$\frac{d}{dt}\left(\frac{\partial g}{\partial v^k}\right)_{\tilde{x}(t)} - \left(\frac{\partial g}{\partial x^h}\right)_{\tilde{x}(t)} = 2g(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial x^h}).$$

□

Consider the metrics g_T and its corresponding function on $T(M)$, with $v = (v_1, ..., v_n) \in T_x M$, $v_i \in \mathbb{R}^3$, in the manifold of positions of the system M,

$$g_T(v, v) = \sum_{i=1}^{n} m_i \vec{v}_i \cdot \vec{v}_i.$$

Observe that, if we denote by T the kinetic energy, we have $g_T = 2T$. Thus, given a local chart $(U, q^1, ..., q^r)$ in M we have, applying Proposition 1.2.4 on the coordinates $(q^1, ..., q^r, \dot{q}^1, ..., \dot{q}^r)$ in $\pi^{-1}U$,

$$g_T(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial q^h}) = \frac{d}{dt}\frac{\partial T}{\partial \dot{q}^h} - \frac{\partial T}{\partial q^h}.$$

And, by equation 1.7,

$$g_T(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial q^h}) = g_T(f, \frac{\partial}{\partial q^h}) = g(F, \frac{\partial}{\partial q^h}),$$

where g is the Euclidean metrics in $\mathbb{R}^{3n}$. Parametrizing M by $\vec{p}_i = \vec{p}_i(q^1, ..., q^r)$, $\vec{p}_i \in \mathbb{R}^3$, the tangent vector $\partial/\partial q^h$ in $\mathbb{R}^{3n}$ is $(\partial\vec{p}_i/\partial q^h)$, and then $g(F, \frac{\partial \vec{p}_i}{\partial q^h}) = \sum_{i=1}^{n} \vec{F}_i \cdot \frac{\partial \vec{p}_i}{\partial q^h}$. Therefore,

$$\sum_{i=1}^{n} \vec{F}_i \cdot \frac{\partial \vec{p}_i}{\partial q^h} = g(F, \frac{\partial}{\partial q^h}) = g_T(\nabla_{\dot{x}}\dot{x}, \frac{\partial}{\partial q^h}) = \frac{d}{dt}\frac{\partial T}{\partial \dot{q}^h} - \frac{\partial T}{\partial q^h}, \tag{1.10}$$

and we have a system of equations in a local chart which is equivalent to equation 1.7

$$\sum_{i=1}^{n} \vec{F}_i \cdot \frac{\partial \vec{p}_i}{\partial q^h} = \frac{d}{dt}\frac{\partial T}{\partial \dot{q}^h} - \frac{\partial T}{\partial q^h}, h = 1, ..., r, \tag{1.11}$$

with T the kinetic energy on $T(M)$, $\vec{F}_i$ the force acting on paticle P_i and $\vec{p}_i$ the position of this particle.

Example 1.2.5. [The simple pendulum]

Take the same pendulum as given in example 1.2.1, assuming that on P acts an external force $\vec{F}$. Using the same notation as before, one has

$$\begin{cases} x = l\sin\varphi \\ y = -l\cos\varphi \end{cases} \Longrightarrow \begin{cases} \dot{x} = l\dot{\varphi}\cos\varphi \\ \dot{y} = l\dot{\varphi}\sin\varphi \end{cases} \Longrightarrow T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\varphi}^2$$

Then we get

$$\tfrac{\partial T}{\partial \dot{\varphi}} = ml^2\dot{\varphi},\ \tfrac{d}{dt}\tfrac{\partial T}{\partial \dot{\varphi}} - \tfrac{\partial T}{\partial \varphi} = ml^2\ddot{\varphi}.$$

Since $\vec{p} = (l\sin\varphi, -l\cos\varphi)$, then $\frac{\partial \vec{p}}{\partial \varphi} = (l\cos\varphi, l\sin\varphi)$ and taking the external force $\vec{F} = (F^1, F^2)$, the system 1.11 corresponds to the equation

$$F^1 l\cos\varphi + F^2 l\sin\varphi = ml^2\ddot{\varphi},$$

that is,

$$F^1\cos\varphi + F^2\sin\varphi = ml\ddot{\varphi}.$$

If, for example, we let the external force F to be the gravity, i.e., $\vec{F} = -mg\overrightarrow{(0,1)}$, then $F^1 = 0$ and $F^2 = -mg$ and the equation is

$$-g\sin\varphi = l\ddot{\varphi}.$$

Motion equations for forces resulting from a potential

In this section we shall see how the equations 1.11 can be written in a easier way if the forces $\vec{F}_i$ acting on the particles P_i result from a potential. We assume then that the system is under the action of a force field $\vec{F}(m, x, y, z)$ resulting from a potential $V(m, x, y, z)$. We shall denote $V(m_i, x, y, z)$ by $V_i(x, y, z)$ and the position of P_i by $\vec{p}_i = (x_i, y_i, z_i)$, and then we have

$$\vec{F}_i = \Big(-\frac{\partial V_i}{\partial x_i}, -\frac{\partial V_i}{\partial y_i}, -\frac{\partial V_i}{\partial z_i} \Big).$$

Therefore,

$$\vec{F}_i \cdot \frac{\partial \vec{p}_i}{\partial q^h} = -\frac{\partial V_i}{\partial x_i}\frac{\partial x_i}{\partial q^h} - \frac{\partial V_i}{\partial y_i}\frac{\partial y_i}{\partial q^h} - \frac{\partial V_i}{\partial z_i}\frac{\partial z_i}{\partial q^h} = -\frac{\partial V_i}{\partial q^h}$$

and so, taking $V = \sum\limits_{i=1}^{n} V_i$,

$$\sum_{i=1}^{n} \vec{F}_i \cdot \frac{\partial \vec{p}_i}{\partial q^h} = -\frac{\partial V_i}{\partial q^h}.$$

Then, the equations 1.11 can be written as

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}^h} - \frac{\partial (T - V)}{\partial q^h} = 0. \tag{1.12}$$

We define the so-called Lagrange function as $L = T - V$. Observe that L is a function on $T(M)$ determined but for additive constants, since V is so. And observe that, since V depends only on $q^1, ..., q^r$ and not on $\dot{q}^1, ..., \dot{q}^r$, we have

$$\frac{\partial T}{\partial \dot{q}^h} = \frac{\partial L}{\partial \dot{q}^h}.$$

Then the equations 1.12 can be written as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^h} - \frac{\partial L}{\partial q^h} = 0 \tag{1.13}$$

Example 1.2.6. [The simple pendulum] Take a simple pendulum which is under a force field generated by gravity as in Example 1.2.5. As we have seen in Example 1.2.5, we have $\vec{F} = -mg\vec{(0,1)}$ and then the potential function is $V(m, x, y) = mgy$. The Langrange function shall be

$$L(\varphi, \dot{\varphi}) = T(\varphi, \dot{\varphi}) - V(\varphi) = \frac{1}{2}ml^2\dot{\varphi}^2 + lmg\cos\varphi.$$

Then

$$\tfrac{d}{dt}\tfrac{\partial L}{\partial \dot{q}^h} = ml^2\ddot{\varphi},\ \tfrac{\partial L}{\partial q^h} = -lmg\sin\varphi.$$

Then, the equations 1.11 are in this case

$$l\ddot{\varphi} + g\sin\varphi = 0,$$

which are the same that we found in Example 1.2.5.

To end this section, we shall make an observation: in the case of a single particle moving freely in $\mathbb{R}^3$, with no constraints, the equations 1.12 are $\vec{F} = m\vec{a}$, the force definition. Let $\vec{F} = (F^1, F^2, F^3)$ and let $\vec{p} = (x, y, z)$ be the position vector of the particle; then, $\frac{\partial \vec{p}}{\partial x} = (1, 0, 0)$. Taking $q^h = x$, the first member of equation 1.11 is

$$\vec{F} \cdot \frac{\partial \vec{p}}{\partial x} = \vec{F} \cdot \vec{(1,0,0)} = F^1.$$

And, computing the second member, one gets

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\ \frac{\partial T}{\partial \dot{x}} = m\dot{x}$$

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}} - \frac{\partial T}{\partial x} = m\ddot{x}.$$

Then, the equation 1.11 for $q^h = x$, is

$$F^1 = m\ddot{x}.$$

By the same procedure for y and z, we have

$$F^2 = m\ddot{y}$$

$$F^3 = m\ddot{z}.$$

Then these three equations give

$$\vec{F} = m\vec{a}.$$

1.3 Conservation theorems. Collision of particles

Conservation theorems

Let us assume the very same hypothesis of the previous section, with a system of particles $P_1, ..., P_n$ with masses $m_1, ..., m_n$ and with holonomic constraints, denoting by M the manifold of positions of the system. We shall use the same notation also. Let φ_s be a one-parameter group of transformations in $\mathbb{R}^3$; then we shall denotbe also by φ_s the one-parameter group in $\mathbb{R}^{3n}$ defined by $\varphi_s(\vec{x}) = (\varphi_s(\vec{x}_1), ..., \varphi_s(\vec{x}_n))$, where $\vec{x} = (\vec{x}_1, ..., \vec{x}_n) \in \mathbb{R}^{3n}$, with $\vec{x}_i \in \mathbb{R}^3$. One says that the system of particles $P_1, ..., P_n$ is invariant for φ_s if $\varphi_s(x) \in M, \forall x \in M$, that is, for any position of the system x, $\varphi_s(x)$ is also a position of the system. We shall denote by $(T_v)_s$, for any $v \in \mathbb{R}^3$, the one-parameter group of translations,

$$(T_v)_s(P) = P + sv.$$

And, given a point $O \in \mathbb{R}^3$ and a vector $v \in \mathbb{R}^3$, we shall denote by $(R_{O,v})_s$ the one-parameter group of rotations around the straight line passing through O in the direction of v.

Theorem 1.3.1 (Momentum conservation)**.** *Assume that the system of particles $P_1, ..., P_n$ is invariant for a one-parameter group of translations in $\mathbb{R}^3$, $(T_v)_s$, and is under the action of a force field resulting from a potential V; assume also that the potential function V is invariant for $(T_v)_s$. Then, the scalar product $\vec{\mathcal{P}} \cdot \vec{v}$ remains constant along any trajectory of the system.*

Proof. Let $x(t) = (p_1(t), ..., p_n(t))$ a trajectory of the system and denote by X the field on M corresponding to $(T_v)_s$. Fix t_0, and the we can take, in a neighbourhood of $x(t_0)$ a system of coordinates $q_1, ..., q_r$ in M such that $X = \partial/\partial q_1$. Since V is invariant under $(T_v)_s$, we have $X(V) = 0$, and equivalently $\partial V/\partial q_1 = 0$. The corresponding Lagrange equation is

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_1} - \frac{\partial T}{\partial q_1} = -\frac{\partial V}{\partial q_1} = 0 \tag{1.14}$$

Since $T = \frac{1}{2}\sum m_i \vec{v}_i \cdot \vec{v}_i$, one has

$$\frac{\partial T}{\partial \dot{q}_1} = \sum m_i \frac{\partial \vec{v}_i}{\partial \dot{q}_1} \cdot \vec{v}_i.$$

And, since $v_i = \sum \dot{q}_k \frac{\partial p_i}{\partial q_k}$, then $\partial v_i / \partial \dot{q}_1 = \partial p_i / \partial q_1$. Thus

$$X = \Big(\frac{\partial p_1}{\partial q_1}, ..., \frac{\partial p_n}{\partial q_1}\Big) = \frac{\partial p}{\partial q_1}.$$

Since X has to be tangent to the curve in $\mathbb{R}^{3n}$, we have

$$s \longrightarrow (\vec{p}_1 + s\vec{v}, ..., \vec{p}_n + s\vec{v})$$

and $X = (v, ..., v)$ and $\partial p_i / \partial q_1 = v$. To sum up, we have

$$\frac{\partial T}{\partial \dot{q}_1} = \sum_i m_i \vec{v} \cdot \vec{v}_i = \big(\sum_i m_i \vec{v}_i\big) \cdot \vec{v} = \vec{\mathcal{P}} \cdot \vec{v}.$$

Observe now that

$$\frac{\partial T}{\partial q_1} = \sum m_i \frac{\partial v_i}{\partial q_1} \cdot v_i.$$

And we have

$$\vec{v}_i = \sum \dot{q}_k \frac{\partial p_i}{\partial q_k} \Longrightarrow \frac{\partial T}{\partial q_1} = \sum_{i,k} \dot{q}_k \frac{\partial^2 p_i}{\partial q_k \partial q_1} \cdot v_i.$$

But observe that derivating with respect to q_1 is derivating in the direction of the one-parameter group. Therefore,

$$\frac{\partial p_i}{\partial q_1} = \frac{d}{ds}(P + sv)|_{s=0} = v.$$

Thus, $\partial p_i / \partial q_1$ is costant and, then, $\partial^2 p_i / \partial q_k \partial q_1 = 0$, which implies that $\partial T / \partial q_1 = 0$. Thus, the equation 1.14 can be written as

$$\frac{d}{dt}(\vec{\mathcal{P}} \cdot \vec{v}) = 0,$$

which proves the Theorem. □

In order to state the following theorem, we shall give first some notation. Given $O \in \mathbb{R}^3$, we define the moment of momentum with respect to O in an instant t as the vector

$$\vec{M}_O = \sum (\vec{p}_i - \vec{O}) \wedge m_i \vec{v}_i.$$

Theorem 1.3.2 (Moment of momentum conservation)**.** *Let $(R_{O,v})_s$ be a one-parameter group under which the system of particles and its potential function V are invariant. Then $\vec{v} \cdot \vec{M}_O$ remains constant along any trajectory of the system.*

Proof. Let X be the vector field in M corresponding to the one-parameter group $(R_{O,v})_s$, let $x(t) = (p_1(t), ..., p_n(t))$ be a trajectory of the system, fix t_0 and take coordinates $q_1, ..., q_n$ in a neighbouhood of $x(t_0)$ in such a way that $X = \partial/\partial q_1$. The corresponding Lagrange equation is equation 1.14. Then we have, as in the previos proof,

$$\frac{\partial T}{\partial \dot{q}_1} = \sum m_i \frac{\partial \vec{p}_i}{\partial q_1} \cdot \vec{v}_i.$$

We have $X = (\partial p_1/\partial q_1, ..., \partial p_n/\partial q_1)$ and it has to be tangent to the curve

$$s \longrightarrow \big((R_{O,v})_s(p_1), ..., (R_{O,v})_s(p_n)\big).$$

Observe now that, taking orthonormal coordinates in $\mathbb{R}^3$ with O being the origin and z the rotation axis, we have $\vec{v} = (0, 0, \lambda)$. Denoting by (x, y, z) the coordinates of P and $(x(s), y(s), z(s))$ the coordinates of $(R_{O,v})_s)(P)$, we have

$$\begin{cases} x(s) = x\cos s - y\sin s \\ y(s) = x\sin s + y\cos s \\ z(s) = z. \end{cases}$$

Therefore,

$$\left(\frac{d}{ds}(R_{O,v})_s(P)\right)_{s=0} = (-y, x, 0) = (0, \vec{0}, 1) \wedge (x, \vec{y}, z) = \frac{1}{\lambda}\vec{v} \wedge (\vec{P} - \vec{O}),$$

that is,

$$\left(\frac{d}{ds}(R_{O,v})_s(P)\right)_{s=0} = K\vec{v} \wedge (\vec{P} - \vec{O})$$

where K is a constant. Then,

$$\frac{\partial T}{\partial \dot{q}_1} = K\sum m_i\vec{v}_i \cdot \big(\vec{v} \wedge (\vec{p}_i - \vec{O})\big) = K\sum v \cdot \big((\vec{p}_i - \vec{O}) \wedge m_i\vec{v}_i\big) = K\vec{v} \cdot \vec{M}_O.$$

Observe that

$$\frac{\partial T}{\partial q_1} = \frac{1}{2}\sum m_i \frac{\partial}{\partial q_1}(\vec{v}_i \cdot \vec{v}_i) = \sum m_i \frac{\partial \vec{v}_i}{\partial q_1}\vec{v}_i.$$

And observe that $v_i = \sum \dot{q}_k \frac{\partial p_i}{\partial q_k} \Longrightarrow \frac{\partial v_i}{\partial q_1} = \sum \dot{q}_k \frac{\partial^2 p_i}{\partial q_1 \partial q_k}$, and derivating with respect to q_1 is derivating in the direction of the one-parameter group,

$$\frac{\partial p_i}{\partial q_1} = K\vec{v} \wedge (\vec{p}_i - \vec{O}).$$

Hence,

$$\frac{\partial v_i}{\partial q_1} = \sum \dot{q}_k K v \wedge \frac{\partial p_i}{\partial q_k} = Kv \wedge v_i.$$

Then $\partial T/\partial q_1 = \sum m_i K(\vec{v} \wedge \vec{v}_i) \cdot \vec{v}_i = 0$. Then, the equation 1.14 can be written as $\frac{d}{dt}K(\vec{v} \cdot \vec{M}_O) = 0$, which proves the Theorem. □

Theorem 1.3.3 (Energy conservation)**.** *Assume that we have a system of particles with holonomic constraints on which acts a force field resulting from a potential. Denote by V the potential function on M the manifold of positions of the system and by T the kinetic energy. Then, the function $H = T + V$ remains constant on any trajectory of the system.*

Proof. Let $x(t_0) \in M$ correspond to the position of the system at the instant t_0. Take coordinates $q_1, ...q_r$ in a neighbourhood of $x(t_0)$. We write then

$$T = \sum_{k,h} T_{kh} \dot{q}_k \dot{q}_h.$$

Then

$$\frac{\partial T}{\partial \dot{q}_k} = 2 \sum_h T_{kh} \dot{q}_h.$$

Therefore,

$$\sum \dot{q}_k \frac{\partial T}{\partial \dot{q}_k} = 2T.$$

We have then

$$H = T + V = 2T - (T - V) = 2T - L = \sum \dot{q}_k \frac{\partial T}{\partial \dot{q}_k} - L = \sum \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L.$$

Restricting H over $x(t)$ on a neighbourhood $x(t_0)$, H is a function on t. Thus we have

$$\frac{dH}{dt} = \frac{d}{dt}\Big(\dot{q}_k \frac{\partial L}{\partial \dot{q}_k}\Big) - \frac{dL}{dt} =$$

$$\ddot{q}_k \frac{\partial L}{\partial \dot{q}_k} + \dot{q}_k \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k}\dot{q}_k - \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k = \Big(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k}\Big)\dot{q}_k = 0.$$

□

Collision of particles

Let P_1 and P_2 two physical bodies, which we will consider to be points, on which acts no force. Therefore, both move following a straight path with uniform motion. We want to study what happens when these two bodies collide, that is, when the two of them are in the same position at the same instant. Experience shows us:

1. The trajectory of the particles ceases to be differentiable when they collide. Then, mechanical laws given to this point turn out to be useless, since we have assumed that the trajectories were smooth.

2. We need further assumptions to get unique solutions, since the behaviour of the motion of the bodies after the collision depends not only on kinematics aspects. For instance, experience shows us that two billiard balls do not behave as two lead balls after the collision between them, since the former are not deformed by the collision and the latter are.

We shall take the following statement as an axiom for every sort of collision.

Axiom 1.3.4. *When two bodies P_1, P_2 with masses m_1, m_2 and moving in uniform motion at velocities $\vec{v}_1, \vec{v}_2$, respectively, collide, the momentum of the system remains constant before and after the collision. That is, denoting by $\vec{v}_1, \vec{v}_2$ the velocities of P_1 and P_2, repectively, before the collision, and $\vec{v}'_1, \vec{v}'_2$ the velocities of P_1 and P_2, respectively, after the collision, it holds that*

$$m_1\vec{v}_1 + m_2\vec{v}_2 = m_1\vec{v}'_1 + m_2\vec{v}'_2. \tag{1.15}$$

Observe that the equation given by the axiom is not enough to determine the velocities after the collision. But we can do so by studying the sort of collision. For instance, if we consider the case of the two lead balls, we have that they keep together after the collision. This is a so-called inelastic collision, and in this sort of collisions we have that $\vec{v}'_1 = \vec{v}'_2$, which, combined with the equation 1.15, determines the velocities after the collision whenever knowing the initial velocities. Consider now a case like the two billiard balls and consider that the two particles move along a straight line. Then we have a one-dimensional motion, and that implies that the velocities vectors are scalars indeed. This is a case of an elastic collision, which are defined as those collisions in which the kinetic energy remains constant before and after the collision. Thus, the scalar equations

$$\begin{cases} m_1v_1 + m_2v_2 = m_1v'_1 + m_2v'_2 \\ m_1v_1^2 + m_2v_2^2 = m_1v'^2_1 + m_2v'^2_2 \end{cases}$$

determine v'_1 and v'_2 depending on v_1 and v_2. To study the motion of two general particles, moving with uniform motion, after the two of them collide, we have two take as an axiom that the trajectories after the collision are contained in the same plane containing the trajectories before the collision. Thus, we have to study the collision of particles moving on a plane. Let a and b the trajectories of two particles P_1, P_2 before they collide, respectively. Let $O = a \cap b$, let t be the bisectrix of the angle between a and b, and let n be the straight line passing through O and perpendicular to t. Take as an axiom of the elastic collision the fact that the collision does not affect the motion of the projections on t, whereas the motion of the projections on n follow the laws stated in the one-dimensional case. This determines the motion after the collision.

1.4 Law of universal gravitation and the two-body problem

Equations of conics in polar coordinates

In order to deal with the two-body problem, we shall first briefly study some geometric properties of conics. Given $a, b \in \mathbb{R}_+ \cup 0$, $a \geq b$, consider in $\mathbb{R}^2$

$$\frac{x^2}{a^2} + \epsilon\frac{y^2}{b^2} = 1, \tag{1.16}$$

where $\epsilon \in \{1, -1\}$. This is the equation on an ellipse if $\epsilon = 1$ and it is this equation of an hyperbola if $\epsilon = -1$. We define the focus of the conics 1.16 as $(c, 0)$ and $(-c, 0)$, where $c = \sqrt{a^2 - b^2}$, and the eccentricity as $e = c/a = \sqrt{1 - \epsilon(b/a)^2}$. Remark that $e < 1$ if the conics is a an ellipse, and $e > 1$ if it is an hyperbola. We want to write the equation 1.16

with respect to the focus $(c, 0)$, taking in the case of the hyperbola the closest branch to the focus. Taking polar coordinates (r, φ), we have

$$\begin{cases} x - c = r\cos\varphi \\ y = r\sin\varphi. \end{cases}$$

Substituting those in equation 1.16, we get

$$\frac{r^2\cos^2\varphi + c^2 + 2rc\cos\varphi}{a^2} + \epsilon\frac{r^2\sin^2\varphi}{b^2} = 1$$

which can be written as

$$r^2\Big(\frac{\cos^2\varphi}{a^2} + \epsilon\frac{\sin^2\varphi}{b^2}\Big) + r\frac{2e}{a}\cos\varphi + (e^2 - 1) = 0, \tag{1.17}$$

which is a second degree equation on r, whose discriminant is

$$\nabla = \frac{4e^2}{a^2}\cos^2\varphi - 4(e^2-1)\Big(\frac{\cos^2\varphi}{a^2} + \epsilon\frac{\sin^2\varphi}{b^2}\Big) =$$

$$= \frac{4}{a^2}\cos^2\varphi - 4\epsilon(e^2-1)\frac{\sin^2\varphi}{b^2} = \frac{4}{a^2}(\cos^2\varphi + \sin^2\varphi) = \frac{4}{a^2}.$$

Thus, isolating r in the equation 1.17, one gets

$$r = \frac{-\frac{2e}{a}\cos\varphi \pm \frac{2}{a}}{2\Big(\frac{\cos^2\varphi}{a^2} + \epsilon\frac{\sin^2\varphi}{b^2}\Big)}. \tag{1.18}$$

Substituting $\sin^2\varphi$ by $1 - \cos^2\varphi$ and computing, we can write the equation 1.18 as

$$r = \frac{b^2\epsilon}{a}\cdot\frac{-e\cos\varphi \pm 1}{1 - e^2\cos^2\varphi}.$$

We have to study now the ellipse and the hyperbola separatedly. Let us start with the ellipse; then $\epsilon = 1, e < 1, 1 - e^2\cos^2\varphi > 0$, and the symbol $\pm$ in the equation 1.18 has to be a positive sign since $r < 0$ otherwise. Therefore, the equation of the ellipse in polar coordinates is

$$r = \frac{b^2}{a}\cdot\frac{1 - e\cos\varphi}{1 - e^2\cos^2\varphi} = \frac{b^2}{a}\cdot\frac{1 - e\cos\varphi}{(1 - e\cos\varphi)(1 + e\cos\varphi)},$$

that is,

$$r = \frac{b^2}{a}\cdot\frac{1}{1 + e\cos\varphi}. \tag{1.19}$$

In the case of the hyperbola, we have $\epsilon = -1$ and, since we are studying the right hand side branch, we want $\varphi = \pi/2$ to correpond to a point of this branch; thus, $\pm$ in equation 1.18 has to be a neative sign. Then we have

$$r = \frac{b^2}{a}\cdot\frac{1 + e\cos\varphi}{1 - e^2\cos^2\varphi} = \frac{b^2}{a}\cdot\frac{1}{1 - e\cos\varphi},$$

that is,

$$r = \frac{b^2}{a} \cdot \frac{1}{1 + e\cos(\varphi - \pi)}.$$

Consider finally the equation of the parabola

$$x + ay^2 = 0, a > 0. \tag{1.20}$$

We define the focus of the parabola 1.20 as the point $(-1/4a, 0)$. With an analogous calculation, the equation 1.20 can be written in polar coordinates as

$$r = \frac{1}{2a} \cdot \frac{1}{1 + \cos\varphi}.$$

Thus, we have the following theorem.

Theorem 1.4.1. *Let $P \in \mathbb{R}^2$ and s a line with origin in P. Take polar coordinates (r, φ) with origin P and angle φ starting at line s counter-clockwise. The equation*

$$r = \frac{p}{1 + e\cos(\varphi - \varpi)},$$

where $p, e, \varpi \in \mathbb{R}$ such that $p > 0, e \geq 0$, is the equation of an ellipse if $e < 1$, it is the equation of an hyperbola if $e > 1$, and it is the equation of a parabola if $e = 1$. In each case, the point P is the focus of the conics.

Universal gravitation law and the two-body problem

Newton, working on the previous investigation of Kepler, stated a law which explained the motion of planets and satellites in the Solar System; we shall take this law, which is stated below, as an axiom. Let P_1, P_2 be two particles with masses m_1, m_2, respectively, freely moving in $\mathbb{R}^3$. Then, for $i, j = 1, 2$, $i \neq j$, P_i exerts on P_j a force

$$\vec{F}_{ij} = \frac{K m_i m_j (\vec{p}_i - \vec{p}_j)}{\|\vec{p}_i - \vec{p}_j\|^3}, \tag{1.21}$$

where K is the so-called gravitational constant and $\vec{p}_i$ is the position vector of the particle P_i.

We can now state the two-body problem: let P_1, P_2 be two particles, with masses m_1 and m_2, in $\mathbb{R}^3$ with free motion, provided that the only force acting on each one is the attraction force due to the other particle by means of the gravitational law; the problem consists on studying the motion of these particles. Since we have two particles freely moving in $\mathbb{R}^3$, we will have a system with six degrees of freedom. Let

$$\vec{r} = \frac{m_1\vec{p}_1 + m_2\vec{p}_2}{m_1 + m_2}$$

the center of mass of the system. Then

$$\vec{p}_2 = \vec{r} + \frac{m_1}{m_1 + m_2}(\vec{p}_2 - \vec{p}_1).$$

Take (X^1, X^2, X^3) the coordinates of the center of mass with respect to fixed rectangular axis and (x^1, x^2, x^3) the coordinates of $\vec{p}_2 - \vec{p}_1$ with respect to the same axis as generalized coordinates. Then the coordinates of $\vec{p}_2$ are

$$\left(X^i + \frac{m_1}{m_1+m_2}x^i\right)$$

and the coordinates of $\vec{p}_1$ are

$$\left(X^i - \frac{m_2}{m_1+m_2}x^i\right).$$

We can calculate the kinetic energy of the system, which we denote by T:

$$T = \frac{1}{2}m_1 \sum \left(\dot{X}^i - \frac{m_2}{m_1+m_2}\dot{x}^i\right)^2 + \frac{1}{2}m_2 \sum \left(\dot{X}^i + \frac{m_1}{m_1+m_2}\dot{x}^i\right)^2 =$$

$$= \frac{1}{2}(m_1+m_2)\sum \left(\dot{X}^i\right)^2 + \frac{1}{2}\frac{m_1 m_2}{m_1+m_2}\sum(\dot{x}^i)^2.$$

By the universal gravitation law, the force that P_2 exerts on P_1 is

$$\vec{F}_{21} = \frac{Km_1m_2(\vec{p}_2 - \vec{p}_1)}{\|\vec{p}_2 - \vec{p}_1\|^3} = \frac{Km_1m_2}{\sqrt{\sum(x^i)^2}^3}(x^1, x^2, x^3). \tag{1.22}$$

Analogously, the force that P_1 exerts on P_2 is $\vec{F}_{12} = -\vec{F}_{21}$. Using the given generalized coordinates and equation 1.22, we have

$$\vec{F}_{21}\frac{\partial \vec{p}_1}{\partial x^i} + \vec{F}_{12}\frac{\partial \vec{p}_2}{\partial x^i} = -\frac{Km_1m_2x^i}{\sqrt{\sum(x^i)^2}^3}$$

and

$$\vec{F}_{21}\frac{\partial \vec{p}_1}{\partial X^i} + \vec{F}_{12}\frac{\partial \vec{p}_2}{\partial X^i} = 0.$$

Thus, taking the function

$$V(x^1, x^2, x^3) = -\frac{Km_1m_2x^i}{\sqrt{\sum(x^i)^2}^3}$$

we have

$$\begin{cases} -\frac{\partial V}{\partial x^i} = \vec{F}_{21}\frac{\partial \vec{p}_1}{\partial x^i} + \vec{F}_{12}\frac{\partial \vec{p}_2}{\partial x^i} \\ -\frac{\partial V}{\partial X^i} = \vec{F}_{21}\frac{\partial \vec{p}_1}{\partial X^i} + \vec{F}_{12}\frac{\partial \vec{p}_2}{\partial X^i}. \end{cases}$$

Then V is the potential function of the system. Therefore, the Lagrange funtion L is going to be

$$L = \frac{1}{2}(m_1+m_2)\sum(\dot{X}^i)^2 + \frac{1}{2}\frac{m_1m_2}{m_1+m_2}\sum(\dot{x}^i)^2 + \frac{Km_1m_2}{\sqrt{\sum(x^i)^2}}$$

and the Lagrange equations of the system are

$$\begin{cases} \ddot{X}^i = 0 \\ \frac{m_1m_2}{m_1+m_2}\ddot{x}^i = -\frac{\partial V}{\partial x^i}. \end{cases}$$

Therefore, the center of mass either moves with uniform straight motion or stays in repose, and the motion of P_2 with respect to P_1 is such as if P_1 was the fixed origin and P_2 moved under the action of a force field resulting from the potential

$$V'(m_2, x^1, x^2, x^3) = -\frac{K(m_1 + m_2)m_2}{\sqrt{\sum(x^i)^2}}. \tag{1.23}$$

Therefore, we shall study now the motion of a particle with mass m_2 under the action of a force field resulting from the potential 1.23. This is obviously a problem symmetric under the rotation on any axis containing the origin, and that implies that the momentum with respect to the origin is constant over time. Then, if $\vec{p}$ is the position vector of the particle and $\vec{\mathcal{P}}$ is the momentum, the vector $\vec{M} = \vec{p} \wedge \vec{\mathcal{P}}$ perpendicular to $\vec{p}$ is constant. Then, $\vec{p}$ is always perpendicular to a fixed vector of $\mathbb{R}^3$, and therefore the trajectory of the particle is contained in a plane. Take then polar coordinates (r, φ) in this plane. The kinetic energy will be

$$T = \frac{1}{2}m_2(\dot{r}^2 + r^2\dot{\varphi}^2)$$

and the potential function

$$V = -\frac{K(m_1 + m_2)m_2}{r}.$$

Then, the Lagrange equations are

$$\begin{cases} \frac{d}{dt}(r^2\dot{\varphi}) = 0 \\ \ddot{r} = r\dot{\varphi}^2 - \frac{K(m_1+m_2)}{r^2}. \end{cases} \tag{1.24}$$

Integrating the first equation, one gets

$$r^2\dot{\varphi} = h,$$

with h being a constant. This is the so-called areas law, due to Kepler. If $h = 0$, then $\dot{\varphi} = 0$ and the trajectory is a straight line. Assume then that $h \neq 0$. We substitute now d/dt by $(h/r^2)d/d\varphi$ (which is deduce from the equation above) in the second equation of 1.24, and write $u = 1/r$ to simplify the result. One gets

$$\frac{d^2u}{d\varphi^2} + u = \frac{K(m_1 + m_2)}{h^2},$$

which has linear coefficients. Its general solution is

$$u = \frac{K(m_1 + m_2)}{h^2}\big(1 + e\cos(\varphi - \varpi)\big),$$

with e, ϖ integration constants. One can assume $e \geq 0$ modifying, if necessary, ϖ, and we obtain

$$r = \frac{h^2}{K(m_1 + m_2)(1 + e\cos(\varphi - \varpi))}. \tag{1.25}$$

By Theorem 1.4.1 this is the equation in polar coordinates of a conics with focus in the origin and, as we have seen above, this conics will be an ellipse if $e < 1$, a parabola if

$e = 1$ and an hyperbola if $e > 1$. The constants e, h and ϖ can be computed if we know the initial position and velocity if the particle. In the case that the trajectory is an ellipse, we will compute the period of the trajectory, that is, the time taken to complete an orbit. Let O be the origin, which is the focus of the ellipse, and let P be the particle describing the trajectory. Let t_0 be the origin of time and denote by $A(t)$ the area of the ellipse between $\overline{OP(t_0)}$ and $\overline{OP(t)}$. Then

$$\frac{dA}{dt} = \frac{1}{2}r^2\dot{\varphi} = \frac{h}{2}. \tag{1.26}$$

Denote by $\mathcal{T}$ the period of the trajectory and let t_1 such that $t_1 - t_0 = \mathcal{T}$. Integrating the equation 1.26 we get

$$\pi ab = \frac{h}{2}\mathcal{T}. \tag{1.27}$$

Comparing the equations 1.19 and 1.25 one gets

$$\frac{b^2}{a} = \frac{h^2}{K(m_1+m_2)} \Longrightarrow h = \frac{b\sqrt{K(m_1+m_2)}}{\sqrt{a}}. \tag{1.28}$$

And, substituting in equation 1.27, we get

$$\mathcal{T} = \frac{2\pi ab}{h} = \frac{2\pi a^{\frac{3}{2}}}{\sqrt{K(m_1+m_2)}}$$

which is the first Kepler law. He stated it as

$$\frac{a^3}{\mathcal{T}^2} = \frac{K(m_1+m_2)}{4\pi^2}, \tag{1.29}$$

that is, $\mathcal{T}$ depends only on the masses. We shall determine now r and φ depending on time in order to determine the position of the particle at any instant of time. To do so, we shall use that the energy $E = T + V$ is constant. Then

$$\begin{aligned} E &= \frac{1}{2}m_2(\dot{r}^2 + r^2\dot{\varphi}^2) - \frac{K(m_1+m_2)m_2}{r} \overset{r^2\dot{\varphi}=h}{=} \\ &= \frac{1}{2}m_2\Big(\dot{r}^2 + \frac{h^2}{r^2}\Big) - \frac{K(m_1+m_2)m_2}{r}. \end{aligned} \tag{1.30}$$

Assuming that we have an elliptic orbit, when the particle is as close as possible to the origin, $\dot{r} = 0$. Then $r = a - c = a(1-e)$ and so, according to the equation 1.30,

$$E = m_2\Big(\frac{1}{2}\frac{h^2}{a^2(1-e)^2} - \frac{K(m_1+m_2)}{a(1-e)}\Big)$$

and, substituting h according to equation 1.28 and knowing that $b = a\sqrt{1-e^2}$, we get

$$E = -\frac{K(m_1+m_2)m_2}{2a}$$

which, substituted in equation 1.30, gives us the differential equation

$$\frac{\dot{r}^2}{2} + \frac{a(1-e^2)K(m_1+m_2)}{2r^2} - \frac{K(m_1+m_2)}{r} = -\frac{K(m_1+m_2)}{2a}.$$

Denoting $K(m_1+m_2)$ by μ, this equation can be written as

$$\dot{r}=\sqrt{\frac{2\mu}{r}-\frac{\mu}{a}-\frac{a(1-e^2)}{r^2}\mu}\Longrightarrow r\dot{r}=\sqrt{2\mu r-\frac{\mu}{a}r^2-a(1-e^2)\mu}.$$

Doing the change of variables

$$r=a(1-e\cos\alpha), \tag{1.31}$$

we get

$$a^2(1-e\cos\alpha)e\sin\alpha\dot{\alpha}=\sqrt{\mu a(2(1-e\cos\alpha)-(1-e\cos\alpha)^2-(1-e^2))}=\cdots=\sqrt{\mu a}e\sin\alpha$$

which can be written

$$(1-e\cos\alpha)\dot{\alpha}=\sqrt{\frac{\mu}{a^3}}$$

and, since by equation 1.29 $\sqrt{\frac{\mu}{a^3}}=\frac{2\pi}{\mathcal{T}}$, the solution of this equation is

$$\alpha-e\sin\alpha=\frac{2\pi}{\mathcal{T}}(t-t_0) \tag{1.32}$$

This equation gives us α depending on t, and by using the equation 1.31, we can get r depending on t as wanted. The equation 1.32, known as the Kepler equation, was obtained by the German mathematician and astronomer Johannes Kepler, and it is the first transcendental equation in the history of science.

1.5 Electromagnetism

The existance of electromagnetic properties is known since ancient times. The French physicist Charles-Augustin de Coulomb stated in the 17th century that the attraction or repulsion electrostatic force between two punctual electrically charged bodies, with positions P and P', and charges q and q', respectively, is given by

$$\vec{F}=k\frac{qq'}{r^3}\vec{r},$$

where $\vec{r}=\vec{P}'-\vec{P}$, $r=\|\vec{r}\|$ and k is a universal constant depending only on the environment. We can take unities in such a way that $k=1$ in vacuum. We define the electric field $\vec{E}$ generated by an electric charge q as the force which the electric charge q exerts, at every point, on the unit electric charge located at that point, that is,

$$\vec{E}=\frac{q}{r^3}\vec{r}. \tag{1.33}$$

Assuming that the electric charge q is located in a point ξ, we have that the function

$$V(x)=\frac{q}{r(x,\xi)}, \tag{1.34}$$

where $r(x,\xi)$:= distance between x and ξ, is the potential function. Observe that it is defined at every point but $x=\xi$.

Electric field and potential generated by a continuum charge

The equations 1.33 and 1.34 refer to the electric field and potential generated by a punctual charge. Assume now that we have a non-punctual body K with a charge distribution given by a density function σ such that the electric charge of an arbitrary region $D \subset K$ is given by

$$Q(D) = \int_D \sigma(x)dx, \tag{1.35}$$

where $dx = dx^1 \wedge dx^2 \wedge dx^3$ is the volume element. The electric field generated by K on a point x will be

$$\vec{E}(x) = \int_K \sigma(\xi)\frac{\vec{x}-\vec{\xi}}{r(x,\xi)^3}d\xi, \tag{1.36}$$

where $r(x\xi) = \|\vec{x}-\vec{\xi}\|$. Since the function under the integral sign has a singularity in $x = \xi$, we shall now give some results regarding the convergence of this integral.

Proposition 1.5.1. *Let $x \in \mathbb{R}^3$ and $B(x,\rho)$ the open ball of radious ρ centered at x. Denote the Euclidean distance between x and ξ by $r(x,\xi)$. Let $\lambda \in \mathbb{R}_+$. Then the integral*

$$I = \int_{B(x,\rho)} \frac{d\xi}{r(x,\xi)^\lambda}$$

is convergent whenever $\lambda < 3$ and divergent otherwise.

As a corollary of this Proposition, we have that the integral which defines the electric field in equation 1.36 is convergent. The i-th component of $\vec{E}$ is indeed

$$E^i(x) = \int_K \sigma(\xi)\frac{x^i-\xi^i}{r(x,\xi)}\frac{1}{r(x,\xi)^2}d\xi.$$

The function $F(x,\xi) = \frac{x^i-\xi^i}{r(x,\xi)}$ is bounded and, letting $\rho \in \mathbb{R}_+$, we have

$$|E^i(x)| \le C\Big(\int_{K\cap B(x,\rho)} \frac{d\xi}{r(x,\xi)^2} + \int_{K\cap(B(x,\rho))^C} \frac{d\xi}{r(x,\xi)^2}\Big),$$

where C is a boundary of the function $\sigma(\xi)F(x,\xi)$ on K. The convergence of the first integral follows from Proposition 1.5.1 and the second one is bounded by $(C/\rho^2)\mathrm{vol}(K)$ since, when $\xi \notin B(x,\rho)$, one has

$$\frac{1}{r(x,\xi)^2} \le \frac{1}{\rho^2}.$$

For practical reasons, we shall assume from now on that the charge-density function σ is well-defined in $\mathbb{R}^3$, of class $\mathcal{C}^\infty$ and compactly supported. Therefore the equation 1.36 of the electric field can be written as

$$\vec{E}(x) = \int_{\mathbb{R}^3} \sigma(\xi)\frac{\vec{x}-\vec{\xi}}{r(x,\xi)^3}d\xi, \tag{1.37}$$

Then, we have

Proposition 1.5.2. *Let*

$$V(x) = \int\limits_{\mathbb{R}^3} \frac{\sigma(\xi)}{r(x,\xi)^3} d\xi. \tag{1.38}$$

This integral converges, the obtained function is smooth and its derivatives are obtained using derivation under the integral sign:

$$\frac{\partial V}{\partial x^\alpha} = \int\limits_{\mathbb{R}^3} \frac{\partial}{\partial x^\alpha}\Big(\frac{1}{r(x,\xi)}\Big)\sigma(\xi) d\xi.$$

Corollary 1.5.3. *It holds that $\vec{E} = -\vec{\nabla} V$. Then, the function V defined by the equation 1.38 is the potential function of the field $\vec{E}$.*

Proposition 1.5.4. *It holds that $\partial V/\partial x^\alpha$ is differentiable with respect to x^α and one has*

$$\frac{\partial^2 V(x)}{\partial (x^\alpha)^2} = \int\limits_{\mathbb{R}^3} \frac{\partial}{\partial x^\alpha}\Big(\frac{\xi^\alpha - x^\alpha}{r(x,\xi)} \cdot \frac{\sigma(\xi)}{r(x,\xi)^2}\Big) d\xi - \sigma(x) \int\limits_{\varpi \in S^2} (\varpi^\alpha)^2 d\varpi,$$

and the first integral is convergent.

Corollary 1.5.5 (Poisson's identity)**.** *It holds that*

$$\triangle V = -4\pi\sigma.$$

Magnetostatics

Oersted observed in 1819 that an electric current of constant intensity I flowing through a straight electrical conductor produces a magnetic field in a neighbourhood of it. Biot and Savart observed that the modulus of this magnetic field at a distance r from its source, denoted by B, satisfies

$$\|B\| = K\frac{I}{r}, \tag{1.39}$$

while if the electrical network producing the magnetic field B is circular, it holds

$$|B| = \pi K \frac{I}{R'}, \tag{1.40}$$

where R' is the radious of the electrical network. It was Laplace who found the general law under these phenomenon. Let $\vec{x}(l) \in \mathbb{R}^3$ a parametrized curve by the arc length, with

$l_0 \leq l \leq l_1$ and through which flows an electric current with constant intensity I in the sense of the curve. Given this notation, Laplace stated that

$$\vec{B}(x) = \int_{l_0}^{l_1} kI \frac{\vec{\dot{x}}(l) \wedge (\vec{x} - \vec{x}(l))}{\|x - x(l)\|^3} dl, \tag{1.41}$$

where k is a universal constant. Assume now that we have a magnetic field created by a magnet and that we put an electrical network through which flows an electric current with intensity I in this magnetic field. A force acts on this network, and, as Ampére discovered, it is given by

$$\vec{F} = \int_{l_0}^{l_1} kI\vec{\dot{x}}(l) \wedge \vec{B}(x(l))dl, \tag{1.42}$$

assuming as above that the electric current is given by a curve $x(l)$ parametrized by arc length. The constant k in equations 1.41 and 1.42 is the same and depends on the unities taken for the intensity, the magnetic field and the distances; thus, we can take those unities in such a way that $k = 1$. But if we measure the intensity in coulombs per second, the magnetic field in oesterds and the distances in centimeters, one gets the constant $k = 1/c$, being c the light velocity.

Equations 1.41 and 1.42 refer to electric currents of intensity I flowing through wires. Let us delete "wires" from our theory. Assume now that we have electric currents generated by motion of electric charges in $\mathbb{R}^3$. We define the vector density of current as $\vec{J} = \sigma\vec{v}$, where $\vec{v}$ is the particles velocity field and σ is, as above, the charge density. For any piece of surface orthogonal to $\vec{J}$, the intensity crossing S is given by

$$I = \| \int_S \vec{J}(x)ds \|,$$

where ds is the area element of the surface. We assume that J does not depend on time, and so does not I. The generalized form of the equations 1.41 and 1.42 for an electric current of current density $\vec{J}$ will be

$$\vec{B}(x) = \frac{1}{c} \int_{\mathbb{R}^3} \vec{J}(x') \wedge \frac{\vec{x} - \vec{x'}}{\|x - x'\|^3} dx', \tag{1.43}$$

$$\vec{F} = \frac{1}{c} \int_{\mathbb{R}^3} \vec{J}(x) \wedge \vec{B}(x)dx. \tag{1.44}$$

To simplify the notation, we write $\vec{r}(x, x') = \vec{x'} - \vec{x}$, $r(x, x') = \|\vec{r}(x, x')\|$. Let x' fixed and x variable, we have

$$\frac{\vec{r}(x, x'}{r^3} = \nabla_x \frac{1}{r(x, x')},$$

Thus the equation 1.43 can be written as

$$B(x) = \frac{1}{c} \int_{\mathbb{R}^3} \vec{J}(x') \wedge \frac{\left(- \vec{r}(x, x')\right)}{r(x, x')^3} dx' = \frac{1}{c} \int_{\mathbb{R}^3} \nabla_x \left(\frac{1}{r(x, x')} \right) \wedge J(x')dx'.$$

Recall that if $\vec{A}(x)$ is a vector field in $\mathbb{R}^3$, we define rot$A$ as the vector field

$$\Big(\frac{\partial A^3}{\partial x^2}-\frac{\partial A^2}{\partial x^3},\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1},\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\Big).$$

Then

$$\nabla_x\Big(\frac{1}{r(x,x')}\Big)\wedge J(x')=\operatorname*{rot}_x\Big(\frac{J(x')}{r(x,x')}\Big)\Longrightarrow \vec{B}(x)=\frac{1}{c}\int_{\mathbb{R}^3}\operatorname*{rot}_x\Big(\frac{\vec{J}(x')}{r(x,x')}\Big)dx'.$$

Using Proposition 1.5.2, we get

$$B(x)=\frac{1}{c}\operatorname*{rot}_x\int_{\mathbb{R}^3}\frac{J(x')}{r(x,x')}dx'\Longrightarrow \operatorname{div}B=0, \tag{1.45}$$

since the divergence of a rotational vanishes. Moreover, it is easy to check that, for any A vector field in $\mathbb{R}^3$, it holds that

$$\operatorname{rotrot}A=\operatorname{graddiv}A-\triangle A,$$

where $\triangle A=(\triangle A^1,\triangle A^2,\triangle A^3)$. Taking rotationals in equation 1.45, we then obtain

$$\operatorname{rot}B=\frac{1}{c}\operatorname*{graddiv}_x\int_{\mathbb{R}^3}\frac{J(x')}{r(x,x')}dx'-\frac{1}{c}\triangle_x\int_{\mathbb{R}^3}\frac{J(x')}{r(x,x')}dx'.$$

Call I_1 and I_2 the two members of the second term of this equality. One has

$$\operatorname*{div}_x\Big(\frac{1}{r(x,x')}\Big)=-\operatorname*{div}_{x'}\Big(\frac{1}{r(x,x')}\Big)\Longrightarrow I_1=-\frac{1}{c}\operatorname*{grad}_x\int_{\mathbb{R}^3}J(x')\operatorname*{div}_{x'}\Big(\frac{1}{r(x,x')}\Big)dx'.$$

And, by Stoke's theorem,

$$\int_{\mathbb{R}^3}\operatorname*{div}_{x'}\Big(\frac{J(x')}{r(x,x')}\Big)dx'=0,$$

since $\frac{J(x')}{r(x,x')}$ vanishes at infinity. Therefore,

$$I_1=\frac{1}{c}\operatorname*{grad}_x\int_{\mathbb{R}^3}\Big(\operatorname*{div}_{x'}J(x')\Big)\frac{1}{r(x,x')}dx'.$$

And, since the electric charge is constant, we have that

$$0=\int_{\partial V}(\vec{J}\cdot\vec{n})ds=\int_V\operatorname{div}J dx\Longrightarrow \operatorname{div}J=0\Longrightarrow I_1=0.$$

On the other hand, with an argument analogous to the one in Proposition 1.5.4 and its Corollary, one has

$$\triangle_x\int\frac{J(x')}{r(x,x')}dx'=-4\pi J(x).$$

Then,

$$\operatorname{rot}B=\frac{4\pi}{c}J. \tag{1.46}$$

Magnetic fields depending on time. Maxwell equations

Assume that we have an electrical network given by a curve C in $\mathbb{R}^3$ parametrized by $\vec{x}(l)$, $l_0 \leq l \leq l_1$. If an electric current flows through C, the electromotive force ϵ is defined as

$$\epsilon = \int_{l_0}^{l_1} \left(\vec{E}(x(l)) \cdot \dot{\vec{x}}(l)\right) dl, \tag{1.47}$$

being $\vec{E}$ the electric field. Assume now that there is no current flowing through C, and that we put C in a magnetic field depeding on time. This variation of the magnetic field produces an electric current in C; Faraday found the law which describes this phenomenon. Let S be a piece of oriented surface in a magnetic field $\vec{B}$. We define the magnetic flux Φ through S as

$$\Phi = \int_S (\vec{B} \cdot \vec{n}) ds,$$

where $\vec{n}$ is the unitary normal vector with respect to S and ds is the area element of S. The Faraday law states that, given a surface S with boundary C, the taken parametrization on C gives an orientation on S. Then,

$$\epsilon = -K \frac{d\Phi}{dt}. \tag{1.48}$$

Using Stokes theorem, one can easily check that the equation 1.48 does not depend on the surface S with boundary C. Constant K in equation 1.48 depends on the units used for B, E, time and distance. If B is measured in oersterds, E in coulombs/cm^2, distance in cm and time in seconds, then $K = 1/c$, being c the light velocity as before. Then,

$$\epsilon = -\frac{1}{c} \frac{d\Phi}{dt}. \tag{1.49}$$

Applying Stokes theorem on equation1.47, one gets

$$\epsilon = \int_S \left((\mathrm{rot} E) \cdot n\right) ds.$$

Using this equality and the definition of magnetic flow, the equation 1.49 can be written as

$$\int_S \left((\mathrm{rot} E) \cdot n\right) ds = -\frac{1}{c} \frac{d}{dt} \int_S (B \cdot n) ds,$$

which holds for any surface S, and then

$$\mathrm{rot} E = -\frac{1}{c} \frac{\partial B}{\partial t}, \tag{1.50}$$

which is equivalent to the Faraday law (equation 1.49). Observe also that if the electric charge of a region depends on time, then

$$\frac{d}{dt} \int_V \sigma dv = -\int_{\partial V} \vec{J} \cdot \vec{n} ds \overset{Stokes}{=} -\int_V \mathrm{div} J dv. \tag{1.51}$$

Then follows the continuity law

$$\frac{\partial \sigma}{\partial t} + \mathrm{div} J = 0. \tag{1.52}$$

Observe now that, taking divergences in equation 1.46, then

$$\mathrm{div} B = 0,$$

which is not compatible with equation 1.52. That is because for equation 1.46 we assumed that E, B, σ and J did not depend on time. Observe the other equations relating E and B if both depend on time. The Poisson equation given in Corollary 1.5.5 works also for electric fields depending on time. So does the equation 1.45, since the arguments used to obtain the result do not depend on time. From equation 1.52 and the Poisson equation in Corollary 1.5.5, we obtain

$$\mathrm{div} J = -\frac{\partial \sigma}{\partial t} = -\frac{1}{4\pi} \mathrm{div} \frac{\partial E}{\partial t} \Longrightarrow \mathrm{div}\left(J + \frac{1}{4\pi}\frac{\partial E}{\partial t}\right) = 0.$$

Hence, we shall take the equation

$$\mathrm{rot} B = \frac{4\pi}{c} J + \frac{1}{c}\frac{\partial E}{\partial t}.$$

We have found now the equations that Maxwell proposed to characterize electromagnetism, those equations which every electromagnetic field fulfills. These are the so-called Maxwell equations

$$\begin{cases} \mathrm{div} E = 4\pi\sigma \\ \mathrm{rot} B = \frac{4\pi}{c} J + \frac{1}{c}\frac{\partial E}{\partial t} \\ \mathrm{rot} E = -\frac{1}{c}\frac{\partial B}{\partial t} \\ \mathrm{div} B = 0 \\ \frac{\partial \sigma}{\partial t} + \mathrm{div} J = 0. \end{cases} \tag{1.53}$$

We shall take these equations as basic physical laws. Assume now that we have a particle with charge q moving within a magnetic field B and an electric field E depending on time. Thus the force acting on this particle will be the sum of the forces acting on the particle due to the electric field and the magnetic field, that is,

$$\vec{F} = q\vec{E} + \frac{q}{c}\vec{v} \wedge \vec{B}. \tag{1.54}$$

This equation, due to Lorentz, must be add to the Maxwell equations. Observe that the equations 1.53 express the relations between E, B, σ and J, while the equation 1.54 describes the motin of any particle with charge in the electromagnetic field. Thus, we finally have

$$\begin{cases} \mathrm{div} E = 4\pi\sigma \\ \mathrm{rot} B = \frac{4\pi}{c} J + \frac{1}{c}\frac{\partial E}{\partial t} \\ \mathrm{rot} E = -\frac{1}{c}\frac{\partial B}{\partial t} \\ \mathrm{div} B = 0 \\ \frac{\partial \sigma}{\partial t} + \mathrm{div} J = 0 \\ \vec{F} = q\vec{E} + \frac{q}{c}\vec{v} \wedge \vec{B}. \end{cases} \tag{1.55}$$

Chapter 2

Special relativity

2.1 Frames of reference

In classical Physics, Newton considered the existence of an absolute immobile space, in which every physical body is located. Assume that there is another observer $O'(t)$ moving with respect to the origin O in $\mathbb{R}^3$ with uniform motion, and let P be a particle moving with respect to the origin. Its acceleration with respect to O' and its acceleration with respect to O are te same, since

$$\frac{d^2(P(t) - O'(t))}{dt^2} = \frac{d^2P(t)}{dt^2} - \frac{d^2O'(t)}{dt^2} = \frac{d^2P(t)}{dt^2}.$$

Thus, two observers located at the origin O an at O' will attribute the motion of P to the very same force. Therefore, the concepts of acceleration and force are the same for an observer located at the origin and an observer moving with uniform motion with respect to the origin.

The same occurs when we observe the universal gravitation law:

$$\vec{F}_{ij} = \frac{Km_im_j(\vec{p}_i - \vec{p}_j)}{\|p_i - p_j\|^3},$$

being p_i, p_j the position of the particles P_i and P_j with masses m_i, m_j, respectively, and $\vec{F}_{ij}$ the force which P_i exerts on P_j. It is easy to observe that, if we consider the motion with respect to an observer located at O' such as considered above, the expression remains unchanged, since we have to substitute the positions p_i and p_j by $p_i - O'$ and $p_j - O'$, and so its difference is again $p_i - p_j$. The same argument holds with the collision of particles. Thus, classical mechanics states when referring to observers moving with uniform motion as when referring to stationary observers. This implies that one cannot know whether an observer is stationary in the absolute space considered by Newton, or otherwise moves with uniform motion with respect to the stationary reference considered by Newton. But it also implies that we can consider a class of "preferred" observers: those moving with uniform motion with respect to the other observers of the class. Such an observer is called an inertial observer.

However, these arguments do not hold for the electromagnetic laws that we stated in the previous chapter. Those laws need the reference origin to be stationary to hold. Consider an electric field $\vec{E}$ and a magnetic field $\vec{B}$ generated by a distribution of electric charges in motion. Let $\vec{J}$ be the corresponding current density vector field. Taking rotational in the second equation 1.53 in Chapter 1, one gets

$$\mathrm{rotrot}B - \frac{1}{c}\frac{\partial}{\partial t}\mathrm{rot}E = \frac{4\pi}{c}\mathrm{rot}J.$$

Taking derivatives with respect to time in the third equation of 1.53, one gets

$$\frac{\partial}{\partial t}\mathrm{rot}E = -\frac{1}{c}\frac{\partial^2 B}{\partial t^2}.$$

Substituting in the previous equation, and knowing that $\mathrm{rotrot} = \nabla\mathrm{div} - \Delta$,

$$-\Delta B + \frac{1}{c^2}\frac{\partial^2 B}{\partial t^2} + \nabla\mathrm{div}B = \frac{4\pi}{c}\mathrm{rot}J.$$

Since we know that $\mathrm{div}B = 0$ (4th equation in 1.53), one observes that whenever $J = 0$ we have

$$\Delta B - \frac{1}{c^2}\frac{\partial^2 B}{\partial t^2} = 0.$$

Consider that the inertial frame S' consists of 3 orthogonal axes moving relative to the "stationary" axes with constant speed $\vec{u} = (u, 0, 0)$. Let E', B' be the electric and magnetic fields, respectively, seen by the observers traveling on the axes of S'. Suppose that one of these observers leaves a particle P with electric charge q in his system. According to the electromagnetism formuli, there exists a force $\vec{F}' = q\vec{E}'$ acting on the particle P. But this particle moves with speed $\vec{u}$ according to a "stationary" observer, and thus the force acting on this particle is $\vec{F} = \vec{E}q + \frac{q}{c}\vec{u} \wedge \vec{B}$. And, since we have seen that forces are equal for inertial observers, we have

$$F' = F \Longrightarrow E' = E + \frac{1}{c}u \wedge B.$$

Assume now that the observers in S' see the particle P with initial speed $\vec{v}$; thus, the force acting on P that they will observe is $F' = qE' + \frac{q}{c}v \wedge B$, whereas the force acting on P that the observers in S will measure will be $F = qE + \frac{q}{c}(v + u) \wedge B$. We have again that $F = F'$ and we know already that $E' = E + \frac{1}{c}u \wedge B$; from these we deduce $v \wedge B = v \wedge B'$. Since this argument holds for any initial speed v, we have that $B = B'$. Therefore, the equations of electromagnetism are invariant under the so-called Galileo transformations (changes from stationary reference to references in uniform motion), and we have

$$\vec{B}' = \vec{B}$$
$$\vec{E}' = \vec{E} + \frac{1}{c}\vec{u} \wedge \vec{B}.$$

The equations of the change are

$$\begin{cases} x_1 = x_1' + ut' \\ x_2 = x_2' \\ x_3 = x_3' \\ t = t' \end{cases}$$

From these one obtains

$$\begin{aligned}\frac{\partial}{\partial t'} &= u\frac{\partial}{\partial x_1} + \frac{\partial}{\partial t}\\ \frac{\partial}{\partial x_i} &= \frac{\partial}{\partial x_i'}.\end{aligned} \tag{2.1}$$

In those regions of space where there are no electric charges, we have

$$\frac{\partial^2 B}{\partial x_1^2} + \frac{\partial^2 B}{\partial x_2^2}\frac{\partial^2 B}{\partial x_3^2} - \frac{1}{c^2}\frac{\partial^2 B}{\partial t^2} = 0. \tag{2.2}$$

Since $B = B'$, it also holds that

$$\frac{\partial^2 B}{\partial x_1'^2} + \frac{\partial^2 B}{\partial x_2'^2}\frac{\partial^2 B}{\partial x_3'^2} - \frac{1}{c^2}\frac{\partial^2 B}{\partial t'^2} = 0.$$

This last equation can be written, in virtue of equation 2.1, as

$$\sum_{i=1}^{3}\frac{\partial^2 B}{\partial x_i^2} - \frac{u^2}{c^2}\frac{\partial^2 B}{\partial x_1^2} - \frac{2u}{c^2}\frac{\partial^2 B}{\partial x_1 \partial t} - \frac{1}{c^2}\frac{\partial^2 B}{\partial t^2} = 0.$$

And, since B satisfies the equation 2.2, we have, for any $u \in \mathbb{R}$ and any magnetic field B generated by electrical charges in motion,

$$u\frac{\partial^2 B}{\partial x_1^2} = -2\frac{\partial^2 B}{\partial x_1 \partial t} \Longrightarrow \frac{\partial^2 B}{\partial x_1^2} = \frac{\partial^2 B}{\partial x_1 \partial t} = 0.$$

This holds only if B is lineal, that is, if

$$\vec{B}(x_1, x_2, x_3, t) = x_1\vec{a} + x_2\vec{b} + x_3\vec{c} + t\vec{d} + \vec{e},$$

where $\vec{a}, \vec{b}, \vec{c}, \vec{d}, \vec{e}$ are constant vectors. Imposing that B vanishes at infinity, which is logic by experience, we have that $B \equiv 0$. Therefore, if the magnetic field equations are invariant under Galileo transformations, any electromagnetic field generated by electrically charged particles in motion have to vanish outside the region where the generator particles are, and this is absurd.

2.2 Postulates of special relativity

After this brief introduction on inertial frames, we can now introduce the postulates of special relativity.

The main point, regarding the Newtonian assumption that there exists an immobile space, is the fact that there is no reason to choose an origin and three coordinate axes over rather than another choice of origin and axes. Every observer has the right, so to speak, to consider that he remains stationary and that is the of the universe which moves, so that he can choose consequently three coordinate axes whose origin is the very same observer. He could as well choose a "universal clock" relative to him, in such a way that the whole frame relative to this observer has synchronized clocks; thus, we could say that a body is in the point (x, y, z) at the instant t if an observer, stationary in (x, y, z) relative to the observer at the origin, sees this body at time t of his clock.

In order to clarify what "synchronized clock" means assume that we have three orthogonal axes as defined above and imagine that there are tiny clocks attached to every point. We say that a clock is synchronized with the clock at the origin if, being this clock at distance L from the origin, a light flash emmited from the origin at $t = 0$ (according to the origin's clock) reaches the point where the clock is located when the point's clock reads time $t = L/c$.

Definition 2.2.1. *We call frame to the set of 3 orthogonal oriented axes with a common origin, with a chosen distance unit and with a clock in every point stationary relative to the origin, in such a way that all these clocks are synchronized.*

The existence of these synchronized clocks is, in fact, a very strong condition, and one might think at first glance that there is no such frames. However, we shall see that, in fact, their existence is stated in the postulates of special relativity.

Definition 2.2.2. *Let S, S' be frames. We say that S' moves with uniform motion (at constant speed and within a straight trajectory) relative to S if $\exists \vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ such that every particle A stationary relative to an observer in S' is, at any instant t of the time in S, in the point $\big(A_1(t), A_2(t), A_3(t)\big)$ of the frame S, where A_1, A_2, A_3 are such that $\frac{dA_i(t)}{dt} = v_i$, $i = 1, 2, 3$.*

Definition 2.2.3. *We say that two frames S, S' are equivalent if either they are stationary relative to each other, and have the same distance unit and time scale, or they satisfy:*

1) they move with uniform motion relative to each other

2) two events that take place in a stationary point of S are seen by an observer in S' in the same temporal order as does an observer in S, and viceversa

3) If two events observed from S' happen in a point A stationary relative to S and with a time separation of 1 unit in S, then the distance between the two events observed in S' (recall that S' moves relative to S and so the two events happen at two different points according to the frame S) is the very same as if we interchanged S and S' in this statement.

Special relativity deals with the observation of inertial observers, that is, observers at rest in an inertial frame, and the comparison of two different observers in uniform motion relative to each other.

Remark 2.2.4. Observe that inertial frames are ideal, since the effects of gravity cannot be completely eliminated and hence there always exist forces acting on the systems. However, when experimentig with particles traveling near the speed of light, the Earth may be considered a inertial frame, since the deviation of the straight-line path of such particles caused by the gravity is negligible.

When talking about the historical development of the special relativity theory, it is worth talking briefly about the ether and the Michelson-Morley experiment.

In 1864, the Scottish physiscist James C. Maxwell showed that light is an electromagnetic phenomenon. According to the knowledge on electromagnetic waves of that time, all forms of electromagnetic waves needed a medium trough which the wave could be propagated; just as it happens with the sound wave. Hence, the ether was hypothesized as the light-bearer medium, which filled all the space but, according to experience, neither acted or was affected by other matter.

In 1887, the American physiscists Albert A. Michelson and Edward W. Morley performed an experiment, named the Michelson-Morley experiment after them, to study the motion of light through the ether and how this light-bearer affected its motion. Surprisingly, the result of the experiment was not the expected one: the shift caused by the presence of the ether was not observed.

The result suggested that the speed of light was isotropic (i.e., the same speed in all directions) in every inertial frame taken as a reference and that the ether wind, if it truly existed, cannot be observed. One possible cause of the result of the experiment was that the speed of light was an invariant, that is, that the speed of light is the same in all inertial frames regardless of the state of motion of its source. But this would contradict the then-assumed principle of addition of velocities. Several theories were proposed to explain the failure in the expected result of the Michelson-Morley experiment retaining the idea of ether.

It was not until 1905 that the German physiscist Albert Einstein (1879-1955), while working as a patent examiner in Bern, Switzerland, proposed a theory which solved the problem discovered regarding relative motion and the speed of light. The postulates of this theory, known as the Special Relativity Theory, are the follwing:

Postulates 2.2.5 (of Special Relativity and invariance of the speed of light)**.** *There exists a class of preferred frames of reference, which is precisely the class of inertial frames, and a positive constant c, with speed dimensions, satisfying:*

A) Is S is an inertial frame, so is any frame S' stationary relative to S which is obtained from S by a displacement of axes and a change of time origin.

B) If S is an inertial frame and P is a point moving with uniform motion relative to S, with speed $\vec{v}$, $\|\vec{v}\| < c$, then there exists an inertial frame S' with origin at P.

C) Two inertial frames S and S' are equivalent.

D) The observers in an inertial frame observe the physical laws equally, regardless of the inertial frame where they are.

E) If an observer stationary relative to an inertial frame S observes the propagation in vacuum of an instantaneous flash of light emmited by any source, he will observe that it is propagated with uniform motion relative to S.

F) The modulus of the speed of such flash of light is c, regardless of the inertial frame where the observer is.

Albert Einstein stated, in fact, two postulates in his famous paper "*Zur Elektrodynamik bewegter Körper*" (On the Electrodynamics of Moving Bodies), published in 1905. These two postulates are:

Postulates 2.2.6 (of Special Relativity and invariance of the speed of light. Shorter version)**.** *P1 All physical laws valid in one frame of reference are equally valid in any other frame moving uniformly relative to the first.*

P2 The speed of light (in vacuum) is the same in all inertial frames, regardless of the motion of the light source.

Observe two immediate consequences of these two postulates. The first one, called the "*Principle of Relativity*", implies that no physical experiment carried out entirely within one inertial frame can ever reveal the state of motion of that system with respect to any other. Therefore, if the ether exists, it cannot be detected, since there is no preferred frame of referece nor such a thing as absolute state of rest. Hence, the idea of ether was finally dismissed, since it is unnecessary for the theory.

The second postulate contradicts the classical principle of addition of velocities. This principle had always been accepted with no hesitation and it appears to be logical according to our intuition. As we shall see below, the reason of this contradiction lies on our classical conception of time, in the assumption that absolute time, just as absolute space, is the very same to every one. This is, actually, a deep-rooted idea in our intuition: Newton himself said in his "Principia Mathematica": "Absolute, true, and mathematical time, of itself, and from its own nature, flows equably, without relation to anything external".

Einstein realized that these classical ideas concerning space and time were fundamentally wrong. Thus, classical Physics laws and the principle of addition of velocities are mere approximations of the real true laws, which work well for "daylife" speeds but turn out to be mistaken when dealing with speeds near the speed of light.

Relativity of simultaneity

Let us briefly examine now some consequences of the postulates; the reader will see that we will find conceptual difficulties rather than mathematical. To overcome those difficulties, the reader is asked to dismiss preconceived ideas about space and time.

We shall present, throughout an example, the new conception of time arisen from the postulates: consider two trains T and T' in relative motion along two adjacent railroads. Assume that T' is moving to the right with constant speed v relative to T.

Consider two fixely located observers O and O' which are in T and T', respectively. Suppose that two bolts of lightning happen to strike the road bed, leaving marks on both trains. One of them leaves marks at point A in T and at point A' in T'. Thus, A and A' have to be in front of each other at the time and space locus when the lightning bolt happens. Similarly, we call B and B' to the points of T and T', repectively, where the flash leaves marks; again, these points have to be in front of each other when the flash occurs.

Assume also that O is situated at the midpoint between A and B, and O' at the midpoint between A' and B'. Without lost of generality, we can also assume that A and A' are situated at the left of O and O', respectively, and B and B' at the right.

Suppose that O sees both flashes simultaneously, that is, O receives the flash signals of A and B simultaneously. That is logical, sinc the distances AO and BO are the same for being O the midpoint between A and B, and light travels at constant speed c regardless of the diretion. And, since O' is the midpoint between A' and B' and, when the flashes occurred A was opposite A' and B opposite B', then O was opposite B' at that instant. Thus, according to O, the follwing events happened simultaneously:

a) a lightning bolt struck at A and A',

b) a lightning bont struck at B and B',

c) O' was opposite O.

However, what does O' observe? Recall that as light signals from the flashes travel with speed c, O' moves to the right relative to O, away from A and towards B. As a result, O' will meet the BB' signal before the AA' signal; and, since O' is equidistant from A' and B', the observer at O' will logically conclude that the flash at B' occurred before that the flash at A'. O' will explain that both flashes meet simultaneously O by the fact that O moves to the left relative to O', away from the earlier flash at B and towards the later one at A.

Note that it does not matter if T is at rest and T' moves to the right, or T' is at rest and T moves to the left, or if they two move along the track in opposite directions. According to Einstein's postulates, these points of view are equivalent whenever the relative speed of the trains remains constant. Anyway, the speed of light is c for any observer.

This is an example of the so-called relativity of simultaneity, that is, that events which are simultaneous for one observer need not to be simultaneous from the point of view of another observer.

Coordinates

In Euclidean geometry, we talk about points: once we have chosen a distance unit and a set of orthogonal coordinate axes, we can express every point as a triple of Cartesian coordinates (x, y, z).

In Physics, however, we work with events rather than points. An event determines both time and space, and thus we assign them four coordinates (x, y, z, t) depending upon the oberver (as we have seen, time as well as space can have different values for different observers). The collection of all possible events is called spacetime.

It would be mathematically interesting, for geometric purposes, refer to both space and time in terms of the same unit. This is easier that what seems to be, since the speed of light, by being an invariant, provides us with a conversion factor.

For instance, if we use cm to measure the spacial coordinates, we can measure time with cm too, understanding $1cm$ of time as one cm of light-travel time, i.e., the time which light needs to travel one centimeter. Thus, $1cm \approx 3.33 \cdot 10^{-11} sec$. Using this convertions, analogous to the light-year distance-measuring unit which the reader might be more familiarized with, we can express the four coordinates of an event in terms of geometric units.

Observe that the speed of light is 1 in geometric units, since light travels exactly $1cm$ (equivalently $1m$, $1km$, etc.) in one cm (equivalently $1m$, $1km$, etc.) of light-travel time. Thus, for any velocity v measured in classical units, say cm/sec, the equivalent value in geometric units will be

$$v\text{cm}/\text{sec} \cdot \frac{1\text{cm}/\text{cm}}{c\text{cm}/\text{sec}} = \frac{v}{c}.$$

Since v/c is a ratio of two velocities, it is a dimensionless quantity, i.e., it is independent of the conventional units chosen (whenever, obviously, v and c are measured using the same conventional units). We shall use this ratio when studying when studying below the so-called Lorentz transformations.

2.3 Lorentz transformations

Let S, S' be two inertial frames. An event in S is determined by (x, y, z, t) the spacetime coordinates in S; the very same event has corresponding spacetime coordinates (x', y', z', t') in S'. Thus, given two inertial frames S and S', there is a map $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ sending the spacetime coordinates (x, y, z, t) of an event observed in S to the spacetime coordinates (x', y', z', t') of the very same event observed in S'. Such a map is called Lorentz transformation and we are about to perform a study of these kind of maps.

Since we have assumed that any event can be observed from both S and S', and these two are inertial frames, then the map sending the spacetime coordinates (x', y', z', t') of an event observed in S' to the spacetime coordinates (x, y, z, t) of the same event observed in S will be the inverse of f. Therefore, such an f is bijective.

Lemma 2.3.1. *f transforms straight lines in $\mathbb{R}^4$ not contained in an hyperplane of the form t =constant in straight lines in $\mathbb{R}^4$ not contained in an hyperplane of the form t' =constant.*

Proof. Let r be a straight line in $\mathbb{R}^4$ not contained in an hyperplane of the form t =constant. Then, r can be parametrized by:

$$\begin{cases} x = x(t) = v_1 t + b_1 \\ y = y(t) = v_2 t + b_2 \\ z = z(t) = v_3 t + b_3 \\ t = t' \end{cases}$$

Such a straight line r represents the motion of a particle $P(t) = \big(x(t), y(t), z(t)\big) \in \mathbb{R}^3$ moving with speed $\vec{v} = (v_1, v_2, v_3)$ relative to S. By condition B) in the Postulates 2.2.5, there exists an inertial frame S'' whose origin is P; and, by condition C) in the same Postulates 2.2.5, S'' moves with uniform motion relative to S'. Hence, any observer in S' will observe the motion of P as a straight line in $\mathbb{R}^4$, which will be $f(r)$. $\square$

Lemma 2.3.2. *If V is a plane of $\mathbb{R}^4$ not contained in any hyperplane of the form t =constant, then $f(V)$ is again a plane in $\mathbb{R}^4$ not contained in any hyperplane of the form t' =constant.*

Proof. V is of the form $a + H$, where $H \subset \mathbb{R}^4$ is 2-dimensional such that $H = \langle v_1, v_2 \rangle$. By Lemma 2.3.1, the straight lines $a + \langle v_1 \rangle$ and $a + \langle v_2 \rangle$ are sent to two straight lines r_1 and r_2, respectively, which intersect at $f(a)$. Let W be the plane containing r_1 and r_2. Let $x \in V$ such that $x \notin a + \langle v_1 \rangle, a + \langle v_2 \rangle$, and let β be a straight line in V passing through x with vector $\vec{u}$ (with fourth component $\neq 0$), $\vec{u} \notin \langle v_1 \rangle$, $\vec{u} \notin \langle v_2 \rangle$. β will cut $a + \langle v_1 \rangle$ and $a + \langle v_2 \rangle$. By Lemma 2.3.1, $f(\beta)$ will be a straight line not contained in an hyperplane of the form $t \neq$constant cutting r_1 and r_2, and such that $f(x) \in f(\beta)$. Therefore, $f(V) \subset W$. An analogous argument on f^{-1} proves $W \subset f(V)$, and hence $f(V) = W$, which proves the Lemma. $\square$

A similar argument proves the following Lemma.

Lemma 2.3.3. *f sends hyperplanes of the form $t \neq$constant to hyperplanes of the form $t' \neq$constant.*

Theorem 2.3.4. *f is an affine trasformation, i.e., f is of the form*

$$f(\vec{x}) = a(\vec{x}) + \vec{b},$$

where $\vec{x} = (x, y, z, t)$, $\vec{b} = (b_1, b_2, b_3, b_4)$ and a is a linear transformation in $\mathbb{R}^4$.

Proof. Since any straight line can be seen as the intersection of three non-horitzonal hyperplanes (that is, of the form $t \neq$constant), f has to send straight lines to straight lines. Let us check that f preserves also parallelism. Let r, r' be two parallel straight lines and let H be the plane determined by r, r'. Then $f(r), f(r')$ are two straight lines contained in $f(H)$. Assume that $f(r), f(r')$ are not parallel; then, since they are contained in the plane $f(H)$, $f(r)$ and $f(r')$ intersect at a point A. Since we know that f is bijective, then $f^{-1}(A)$ has to belong to r and r'. But this is a contradiction, since we assumed r and r' to be parallel. Therefore, f sends straight lines to straight lines and preserves parallelism. By a classical geometric theorem, f is then an affine transformation and we are done. $\square$

Let us seek an explicit expression of a Lorentz transformation f. Assume that $f : S \longrightarrow S'$, being S and S' two inertial frames with coordinates (x_1, x_2, x_3, t) and (x'_1, x'_2, x'_3, t'), respectively, and such that S' moves with uniform motion relative to S in the direction of a common $x_1x'_1-$axis, $x_2x'_2-$axis and $x_3x'_3-$axis remain parallel, and the origins O and O' coincide when $t = t' = 0$. Then, since f is an affine trasnformation by Theorem 2.3.4, it has to have the form

$$\begin{cases} x'_i = \sum\limits_{j=1}^{3} a_{ij}x_j + a_{i4}t + k_i, \quad i = 1,2,3 \\ t' = \sum\limits_{j=1}^{3} a_{4j}x_j + a_{44}t + k_4 \end{cases}$$

Since the origins coincide at the time origins, we have that $k_i = 0$, $i = 1,2,3,4$. And, since S and S' have the same x_1-axis, whenever $x_2 = x_3 = 0$ we have $x'_2 = x'_3 = 0$, and therefore $a_{21} = a_{24} = a_{31} = a_{34} = 0$. Due to condition 2) of Definition 2.2.3 (about equivalent frames), $a_{44} > 0$. Write $v = -a_{14}/a_{44}$ (we will study the physical meaning of v later). The we have

$$\begin{cases} x'_1 = & a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & - & va_{44}t \\ x'_2 = & & & a_{22}x_2 & + & a_{23}x_3 & & \\ x'_3 = & & & a_{32}x_2 & + & a_{33}x_3 & & \\ t' = & a_{41}x_1 & + & a_{42}x_2 & + & a_{43}x_3 & + & a_{44}t \end{cases} \tag{2.3}$$

We impose the principle of invariance of the speed of light. Suppose that at instant $t = 0$ in S, a flash of light is emmited from the origin spreading through all directions. Therefore, this flash of light will arrive at time t to the points $(x_1, x_2, x_3) \in \mathbb{R}^3$ satisfying

$$x_1^2 + x_2^2 + x_3^2 = c^2t^2. \tag{2.4}$$

Moreover, since at $t = 0$ in S the origins of S and S' coincided at $t' = 0$, this flash of light would have source at the origin at time $t' = 0$ for an observer of S'. Thus, this flash of light will arrive at time t' to the points $(x'_1, x'_2, x'_3) \in \mathbb{R}^3$ satisfying

$$(x'_1)^2 + (x'_2)^2 + (x'_3)^2 = c^2(t')^2. \tag{2.5}$$

Hence, f has to send points (x_1, x_2, x_3) satisfying the equation 2.4 to points (x'_1, x'_2, x'_3) satisfying the equation 2.5.

Let $Q(x_1, x_2, x_3, t)$ the expression obtained from $(x'_1)^2 + (x'_2)^2 + (x'_3)^2 - c^2(t')^2$ substituting x'_1, x'_2, x'_3, t' by x_1, x_2, x_3, t according to equations 2.3. Thus $Q(x_1, x_2, x_3, t)$ will be a quadratic expression on x_1, x_2, x_3, t. Whenever (x_1, x_2, x_3, t) satisfy equation 2.4, we have $Q(x_1, x_2, x_3, t) = 0$ and conversely. Then the cone in $\mathbb{R}^4$ generated by equation 2.4 must also be the quadric $Q(x_1, x_2, x_3, t) = 0$. Therefore, there exists a constant ρ such that

$$Q(x_1, x_2, x_3, t) = \rho(x_1^2 + x_2^2 + x_3^2 - c^2t^2). \tag{2.6}$$

Let us study now the action of f on the plane of equations $x_2 = 0$ and $x_3 = 0$. Since, by equation 2.3

$$f : (x_1, 0, 0, t) \longrightarrow (x'_1, 0, 0, t'),$$

we study only

$$\begin{cases} x'_1 = a_{11}x_1 - va_{44}t \\ t' = a_{41}x_1 + a_{44}t. \end{cases} \tag{2.7}$$

By equation 2.6 we deduce

$$Q(x_1, 0, 0, t) = \rho(x_1^2 - c^2t^2)$$

and then

$$(x'_1)^2 - c^2(t')^2 = \rho(x_1^2 - c^2t^2).$$

Computing using equation 2.7, we have

$$\begin{aligned} (x'_1)^2 - c^2(t')^2 &= (a_{11}x_1 - va_{44}t)^2 - c^2(a_{41}x_1 + a_{44}t)^2 = \\ &= (a_{11}^2 - a_{41}^2c^2)x_1^2 + (v^2a_{44}^2 - a_{44}^2c^2)t^2 - 2(a_{41}a_{44}c^2 + va_{11}a_{44}) = \\ &= \rho(x_1^2 - c^2t^2). \end{aligned}$$

Then,

$$a_{11}^2 - a_{41}^2c^2 = \rho, \tag{2.8}$$

$$a_{44}^2v^2 - a_{44}^2c^2 = -\rho c^2, \tag{2.9}$$

$$-va_{11}a_{44} - c^2a_{41}a_{44} = 0. \tag{2.10}$$

From equation 2.10 we obtain $a_{41} = -(va_{11})/c^2$ and, substituting on equation 2.8, we obtain

$$\rho = (1 - \frac{v^2}{c^2})a_{11}^2.$$

From equation 2.9 we obtain

$$\rho = (1 - \frac{v^2}{c^2})a_{44}^2.$$

Therefore, $a_{11}^2 = a_{44}^2$. Without loss of generality, we can assume $a_{11} > 0$ (otherwise we could consider a different frame S' with the same axes but converse orientation on the x_1-axis) and, since $a_{44} > 0$, we have $a_{11} = a_{44}$. Then, equation 2.7 is written

$$\begin{cases} x_1' = a_{44}x_1 - a_{44}vt \\ t' = -\frac{va_{44}}{c^2}x_1 + a_{44}t. \end{cases} \tag{2.11}$$

Since we know that f is bijective, we want equation 2.11 to have an inverse. Then, the determinant of its matrix must be nonzero, which implies $\|\vec{v}\| \neq c$. The inverse transformation of equation 2.11 is

$$\begin{cases} x_1 = \frac{1}{a_{44}(1-\frac{v^2}{c^2})}x_1' + \frac{v}{a_{44}(1-\frac{v^2}{c^2})}t' \\ t = \frac{v}{c^2 a_{44}(1-\frac{v^2}{c^2})}x_1' + \frac{1}{a_{44}(1-\frac{v^2}{c^2})}t'. \end{cases} \tag{2.12}$$

Let us study the effects of condition 3) of Definition 2.2.3. Consider two events which happen at the origin of S with a time interval of length 1 between them (for instance, the first one at $t = 0$ and the second one at $t = 1$). The firste event is seen in S' at the origin, and the second one, according to equations 2.11, at the point $(-a_{44}v, 0, 0)$ and time $t' = a_{44}$. The distance in S' between the two events is $a_{44}|v|$. Doing the same experiment interchanging the roles of S and S', we obtain that the distance is $\frac{|v|}{a_{44}(1-\frac{v^2}{c^2})}$. Since, by Definition 2.2.3 of equivalent frames, the result must be the same, we have

$$a_{44} = \frac{1}{a_{44}(1-\frac{v^2}{c^2})}.$$

Therefore, $-v^2/c^2 > 0$, which implies that $|v| < c$, and

$$a_{44} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Moreover,

$$\rho = (1 - v^2/c^2)a_{44}^2 = 1.$$

Thus, going back to equation 2.3, we have

$$\begin{cases} x_1' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}x_1 + a_{12}x_2 + a_{13}x_3 - \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t \\ x_2' = a_{22}x_2 + a_{23}x_3 \\ x_3' = a_{32}x_2 + a_{33}x_3 \\ t' = -\frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}}x_1 + a_{42}x_2 + a_{43}x_3 + \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}t. \end{cases} \tag{2.13}$$

Moreover, we have seen that

$$(x_1')^2 + (x_2')^2 + (x_3')^2 - c^2(t')^2 = Q(x_1, x_2, x_3, t) = x_1^2 + x_2^2 + x_3^2 - c^2t^2. \tag{2.14}$$

Computing and equalizing the two x_1x_2-terms in equation 2.14, we get

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}a_{12}+\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}a_{42}=0\Longrightarrow a_{12}=-va_{42}.$$

Doing the same for x_2t-terms, we get

$$-\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}a_{12}-\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}c^2a_{42}=0.$$

Substituting $a_{12}=-va_{42}$, we get

$$v^2a_{42}-c^2a_{42}=0\overset{v^2<c^2}{\Longrightarrow}a_{42}=0\Longrightarrow a_{12}=0.$$

Using the same argument with x_1x_3-terms and x_1t-terms, in equation 2.14, we get $a_{13}=a_{43}=0$. Thus equation 2.13 can be written as

$$\begin{cases} x_1'=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}x_1-\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t \\ x_2'=a_{22}x_2+a_{23}x_3 \\ x_3'=a_{32}x_2+a_{33}x_3 \\ t'=-\frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}}x_1+\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}t. \end{cases} \tag{2.15}$$

Moreover, equation 2.14 has to be satisfied. From equation 2.15 one can see that $(x_1')^2-c^2(t')^2=x_1^2-c^2t^2$ is fulfilled. So we just need $(x_2')^2+(x_3')^2=x_2^2+x_3^2$ for equation 2.14 to be satisfied. Therefore,

$$\begin{cases} x_2'=a_{22}x_2+a_{23}x_3 \\ x_3'=a_{32}x_2+a_{33}x_3 \end{cases} \tag{2.16}$$

has to be an isometry. We change the axes in the the plane x_2x_3 in S by means of the chane givn by equations 2.16. Naming S again the obtained frame, then $S\longrightarrow S'$ is given by

$$\begin{cases} x_1'=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}x_1-\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t \\ x_2'=x_2 \\ x_3'=x_3 \\ t'=-\frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}}x_1+\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}t. \end{cases} \tag{2.17}$$

To sum up, we proved:

Theorem 2.3.5. *Let S, S' be two inertial frames. We can make orthogonal changes between S and S' (i.e., changes preserving distances and angles) and a change in the time origin whenever such a change is given by equations 2.17.*

Let u study now the physiscal meaning of v. Observe that the inverse of equations 2.17 is given by

$$\begin{cases} x_1 = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} x_1' + \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} t' \\ x_2 = x_2' \\ x_3 = x_3' \\ t = \frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}} x_1' + \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} t'. \end{cases} \tag{2.18}$$

At the origin of S', we have that $t = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} t'$. The origin of S' is seen from S at the point $\left(\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} t', 0, 0\right) = (vt, 0, 0)$. Thus, the origin of S' moves relative to S with uniform motion at constant speed $\vec{v} = (v, 0, 0)$.

Let us give now an important and useful defition and rewrite the Lemmas given at the beginning of this section in terms of this definition.

Definition 2.3.6. *A vector $\vec{u} = (u_1, u_2, u_3, u_4) \in \mathbb{R}^4$ is said to be timelike if it satisfies*

$$u_1^2 + u_2^2 + u_3^2 < c^2 u_4^2.$$

Considering the cone $C = \{(x_1, x_2, x_3, t) \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 - c^2t^2 < 0\}$, we have

$$\vec{u} \text{ timelike} \Longleftrightarrow \vec{u} \in C.$$

We call $C^+ = \{(x_1, x_2, x_3, t) \in C : t > 0\}$. We call positive timelike vectors to those $\vec{u} \in C^+$.

Let S and S' be two inertial frames and let $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ sending S to S'. Let r be a straight line in $\mathbb{R}^4$ with timelike direction vector $\vec{u}$. Then $\vec{u}/u_4$ is again a timelike direction vector of r, and then r can be parametrized as

$$\begin{cases} x_i = x_i(t) = v_i t + b_i, \quad i = 1, 2, 3 \\ t = t \end{cases}$$

where $v_1^2 + v_2^2 + v_3^2 < c^2$.

The previous given results can be rewritten then as follow.

Lemma 2.3.7. *If r is a straight line in $\mathbb{R}^4$ with timelike direction vector, then $f(r)$ is a straight line in $\mathbb{R}^4$ with timelike direction vector.*

Definition 2.3.8. *We say that a plane (respectively, an hyperplane) in $\mathbb{R}^4$ is admissible if it is of the form $a + H$, where H is a vector subspace of dimension 2 (respectively, of dimension 3) whose basis is in C^+.*

Lemma 2.3.9. *If V is an admissible plane in $\mathbb{R}^4$, then $f(V)$ is again an admissible plane in $\mathbb{R}^4$.*

Lemma 2.3.10. *f sends admissible hyperplanes to admissible hyperplanes.*

Theorem 2.3.11. *f is an affine transformation.*

Consequences of Lorentz transformations

Let S and S' be two inertial frames such that the change of coordinates from S to S' is given by equations 2.17. Let A be a stationary point relative to S, located at the $x-$axis of S (then, $A = (A_1, 0, 0)$). This point A is seen by an observer in S' with coordinates (x', y', z', t'_0) satisfying

$$\begin{cases} x' = \frac{A_1 - vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ y' = 0 \\ z' = 0 \\ t'_0 = -\frac{(v/c^2)A_1 + t}{\sqrt{1-\frac{v^2}{c^2}}}. \end{cases}$$

Isolating t in the last equation and substituting, we get

$$x' = A_1\sqrt{1-\frac{v^2}{c^2}} - vt_0.$$

Thus, A is seen in S' at instant t'_0 as

$$(A_1\sqrt{1-\frac{v^2}{c^2}} - vt_0, 0, 0).$$

Let B be another point $B = (B_1, 0, 0)$ sucht that $B_1 - A_1 = 1$. Thus, $\text{dist}(A, B) = 1$ in S. B is seen in S' at the instant t'_0 with coordinates

$$(B_1\sqrt{1-\frac{v^2}{c^2}} - vt_0, 0, 0).$$

Thus, an observer in S' will measure a distance between A and B as $\sqrt{1-\frac{v^2}{c^2}}$. Then, one meter seen from S has length $\sqrt{1-\frac{v^2}{c^2}}$ meters from S'. Hence, lengths shorten in the direction of motion.

Consider now two events which happen in the same point A of S at different instants t_0 and t_1. These events are seen in S' at instants t'_0 and t'_1 such that

$$t_i = -\frac{(v(c^2)A_1 + t_i}{\sqrt{1-\frac{v^2}{c^2}}}, \, i = 0, 1.$$

Therefore,

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

Hence, time intervals are enlarged by motion!

Consider now two events which occur in S in two different points $A = (A_1, A_2, A_3)$ and $B = (B_1, B_2, B_3)$ at the same instant t_0. These events are going to be seen by an observer in S' at instants t'_1 and t'_2 given by

$$t'_1 = -\frac{(v/c^2)A_1 + t_0}{\sqrt{1-\frac{v^2}{c^2}}}, \, t'_2 = -\frac{(v/c^2)B_1 + t_0}{\sqrt{1-\frac{v^2}{c^2}}},$$

which is a new example of the yet-explained relativity of simultaneity.

Vector form of Lorentz transformations

We want to seek now a simple form of equations 2.17 sending S to S', being S and S' two inertial frames, without need of making a good choice of axes in S and S'.

Assume that the axes of S and S' are "good", so that equations 2.17 hold. Let $\vec{r} = (x, y, z)$ and $\vec{r}' = (x', y', z')$. Our goal is to write equation 2.17 as $(\vec{r}, t) \longmapsto (\vec{r}', t')$. Equations 2.17 can be written as

$$\begin{cases} x' = x + \Big(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\Big)x - \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t \\ y' = y \\ z' = z \\ t' = -\frac{v}{c^2\sqrt{1-\frac{v^2}{c^2}}}x + \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}t. \end{cases} \tag{2.19}$$

From the first three equations, we deduce

$$\vec{r}' = \vec{r} + \Big((\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)x - \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t, 0, 0\Big).$$

Being $\vec{v} = (v, 0, 0)$ the speed vector of S' relative to S, we have $\vec{e}_1 := (1, 0, 0) = \vec{v}/v$ and then

$$\Big((\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)x - \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t, 0, 0\Big) = (\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)x\vec{e}_1 - \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}t\vec{e}_1$$

and

$$x = \vec{r} \cdot \vec{e}_1 = \frac{\vec{r} \cdot \vec{v}}{v}.$$

Therefore,

$$\vec{r}' = \vec{r} + (\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)\frac{\vec{r} \cdot \vec{v}}{v}\vec{v} - \frac{t}{\sqrt{1-\frac{v^2}{c^2}}}\vec{v}.$$

The last equation in equations 2.19 is written as

$$t' = -\frac{1}{c^2\sqrt{1-\frac{v^2}{c^2}}}\vec{r} \cdot \vec{v} + \frac{t}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Then, equations 2.19 can be written as

$$\begin{cases} \vec{r}' = \vec{r} + \Big((\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)\frac{\vec{r}\cdot\vec{v}}{v^2} - \frac{t}{\sqrt{1-\frac{v^2}{c^2}}}\Big)\vec{v} \\ t' = -\frac{1}{c^2\sqrt{1-\frac{v^2}{c^2}}}\vec{r} \cdot \vec{v} + \frac{t}{\sqrt{1-\frac{v^2}{c^2}}}. \end{cases} \tag{2.20}$$

and the inverses (changing $\vec{v}$ by $-\vec{v}$) are

$$\begin{cases} \vec{r} = \vec{r}' + \Big((\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)\frac{\vec{r}'\cdot\vec{v}}{v^2} + \frac{t'}{\sqrt{1-\frac{v^2}{c^2}}}\Big)\vec{v} \\ t = -\frac{1}{c^2\sqrt{1-\frac{v^2}{c^2}}}\vec{r}' \cdot \vec{v} + \frac{t'}{\sqrt{1-\frac{v^2}{c^2}}}. \end{cases} \tag{2.21}$$

Equations 2.20 are equivalent to equations 2.17 and, since any orthogonal change of axes preserves the scalar product in $\mathbb{R}^3$, equations 2.20 hold regardless of the position of axes of S and S' if the origins of S and S' coincide at their time origins. And analogously for equations 2.21 and the inverses of equations 2.17.

Law of addition of velocities

Let S and S' be two inertial frames and let P be a particle in motion, being $\vec{v}$ the speed of S' relative to S. Let $\vec{r}(t)$ be the position vector of P at instant t in S and $\vec{r}'(t)$ the position vector of P at instant t' in S'. The speed of P at time t in S will be

$$\vec{u} = \frac{d\vec{r}(t)}{dt},$$

whereas the speed of P at time t' in S' will be

$$\vec{u}' = \frac{d\vec{r}'(t')}{dt'}.$$

Derivating equations 2.21, one gets

$$\vec{u} = \frac{d\vec{r}}{dt} = \Big(\frac{d\vec{r}'}{dt'} + ((\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1)\frac{1}{v^2}\frac{d\vec{r}'}{dt'}\cdot\vec{v} + \frac{1}{\sqrt{1-\frac{v^2}{c^2}}})\vec{v}\Big)\frac{dt'}{dt},$$

$$\frac{dt}{dt'} = \frac{1+(\vec{u}'\cdot\vec{v})/c^2}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Then,

$$\vec{u} = \frac{1}{1+(\vec{u}'\cdot\vec{v})/c^2}\Big((\frac{\vec{u}'\cdot\vec{v}}{v^2}+1)\vec{v} + \sqrt{1-\frac{v^2}{c^2}}(\vec{u}' - \frac{\vec{u}'\cdot\vec{v}}{v^2}\vec{v})\Big).$$

Therefore,

$$\vec{u} = \frac{1}{1+(\vec{u}'\cdot\vec{v})/c^2}\Big((\frac{\vec{u}'\cdot\vec{v}}{v^2}+1)\vec{v} + \sqrt{1-\frac{v^2}{c^2}}(\vec{u}' - \frac{\vec{u}'\cdot\vec{v}}{v^2}\vec{v})\Big)$$

which is the relativistic law of addition of velocities. Observe that if v/c^2 is small, we have $\vec{u} \approx \vec{u}' + \vec{v}$, the classical law of addition of velocities.

2.4 The Minkowski space

We have seen that Lorentz transformations sending the origin to the origin preserve the quadratic form $x^2+y^2+z^2-c^2t^2$. We define the inner product g in $\mathbb{R}^4$, given $X = (x_1, x_2, x_3, x_4)$ and $Y = (y_1, y_2, y_3, y_4)$, as

$$g(X,Y) = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4.$$

Proposition 2.4.1. *Lorentz transformations sending the origin to the origin preserve the inner product g, that is,*

$$g(X,Y) = g(f(X), f(Y)).$$

Proof. Observe that

$$g(X,Y) = \frac{1}{2}\big(g(X+Y, X+Y) - g(X,X) - g(Y,Y)\big),$$

and we know that f preserves the quadratic form

$$x^2 + y^2 + z^2 - c^2t^2 = g\big((x_1, x_2, x_3, t), (x_1, x_2, x_3, t)\big).$$

Therefore, f preserves $g(X,X)$, $g(Y,Y)$ and $g(X+Y, X+Y)$, and hence preserves also $g(X,Y)$. □

Let C be the cone in $\mathbb{R}^4$ defined previously; observe that $C = \{X = (x,y,z,t) \in \mathbb{R}^4 : g(X,X) < 0\}$. Let C^+ be the connected component of C such that $t > 0$, and define a vector to be timelike as above.

Proposition 2.4.2. *Let f be a Lorentz transformation sending the origin to the origin. Then $f(C^+) = C^+$.*

Proof. By Proposition 2.4.1 we have that $f(C) = C$. Since f is continuous (by linearity), it sends the connected component $C^+ \subset C$ to a connected component of C. Let us study the image of the point $(0,0,0,1) \in C^+$. By Theorem 2.3.5, $f = h' \circ F \circ h$ where F is a transformation given by equations 2.17 and h, h' are transformations in $\mathbb{R}^4$ of the form

$$\begin{cases} \mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R} & \longrightarrow \quad \mathbb{R}^3 \times \mathbb{R} = \mathbb{R}^4 \\ (\vec{x}, t) & \longmapsto \qquad (H(\vec{x}), t), \end{cases}$$

being H a linear isometry in $\mathbb{R}^3$ with the Euclidean metrics.

Then

$$f(0,0,0,1) = h'Fh(0,0,0,1) = hF(0,0,0,1),$$

since $h(0,0,0,1) = (0,0,0,1)$. Since F is of the form of the equations 2.17, we see that $F(0,0,0,1) \in C^+$.

And, moreover, h' sends C^+ to C^+. Therefore, $f(0,0,0,1) \in C^+$ and we are done. □

Definition 2.4.3. *We say that a basis $\{e_1, e_2, e_3, e_4\}$ of $\mathbb{R}^4$ is positive ortho-c-normal is it satisfies*

$$\begin{cases} g(e_i, e_i) = 1, & i = 1,2,3 \\ g(e_4, e_4) = c^2 & \\ g(e_i, e_j) = 0, & i \neq j \\ e_4 \in C^+. & \end{cases}$$

Let S and S' be two inertial frames and f the Lorentz transformation sending S to S'. Thus, $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ and then an event corresponding in S to $x \in \mathbb{R}^4$ has coordinates $f(x) \in \mathbb{R}^4$ in S', and we know that $f(\vec{x}) = F(\vec{x}) + \vec{a}$, being F linear. We call $\{e_1, ..., e_4\}$ the canonical basis in $\mathbb{R}^4$ and $\epsilon_i = F^{-1}(e_i)$. Consider the four axes determined by $f^{-1}(0)$ and direction vectors ϵ_i.

If $f(x)$ has coordinates (y_1, y_2, y_3, y_4) in the canonical axes of $\mathbb{R}^4$, then x has also coordinates (y_1, y_2, y_3, y_4) in the basis given by $f^{-1}(0)$ and ϵ_i.

We have

$$\begin{cases} f(x) = y = \sum_{i=1}^{4} y_i e_i \\ \\ f(x) = F(x) + a \end{cases} \Longrightarrow F(x) + a = \sum_{i=1}^{4} y_i e_i$$

Then

$$x = F^{-1}(\sum_{i=1}^{4} y_i e_i - a) = \sum_{i=1}^{4} y_i \epsilon_i + F^{-1}(-a).$$

And observe that

$$f(F^{-1}(-a)) = FF^{-1}(-a) + a = 0 \Longrightarrow F^{-1}(-a) = f^{-1}(0).$$

Therefore

$$x = \sum_{i=1}^{4} y_i \epsilon_i + f^{-1}(0) \Longrightarrow x - f^{-1}(0) = \sum_{i=1}^{4} y_i \epsilon_i.$$

Thus, the following are equivalent:

1) Observers in S' use, as observers in S, the canonical axes in $\mathbb{R}^4$. But an event given in S by coordinates $x = (x_1, x_2, x_3, x_4)$ is given in S' by coordinates $f(x) = (y_1, y_2, y_3, y_4)$.

2) Both observers in S and S' represent the same event as a same point in $\mathbb{R}^4$. But observers in S use the canonical axes in $\mathbb{R}^4$ and assign to x coordinates (x_1, x_2, x_3, x_4) whereas observers in S' use the axes given by $f^{-1}(0)$ and $\vec{\epsilon}_i$, and therefore assign coordinates (y_1, y_2, y_3, y_4) to the event x.

We define the Minkowski space as $\mathbb{R}^4$ endowed with the metrics g defined above and the choice of the connected component C^+. Phenomena of special relativity's kinematics are geometrical problems in the Minkowski space.

Lemma 2.4.4. *In the 2-dimensional Minkowski space ($\mathbb{R}^2$ endowed with the inner product $g(v, w) = v_1 w_1 - c^2 v_2 w_2$),*

$$(v, w)^2 \geq g(v, v) g(w, w), \forall v, w.$$

Proof. We want to check

$$v_1 w_1 - c^2 v_2 w_2 \geq (v_1^2 - c^2 v_2^2)(w_1^2 - c^2 w_2^2) \Longleftrightarrow$$

$$\Longleftrightarrow -2 v_1 w_1 v_2 w_2 c^2 \geq -c^2 (v_2 w_1)^2 - c^2 (v_1 w_2)^2 \Longleftrightarrow$$

$$\Longleftrightarrow (v_2 w_1)^2 + (w_2 v_1)^2 - 2 v_1 w_1 v_2 w_2 \geq 0 \Longleftrightarrow$$

$$(v_2 w_1 - v_1 w_2 \geq 0,$$

which always holds. $\square$

Proposition 2.4.5. *Let $\vec{v}$ be a timelike vector in the Minkowski space, and let $\vec{w}$ be any vector. Then*

$$g(v,w)^2 \geq g(v,v)g(w,w).$$

Proof. Consider in $\mathbb{R}^4$ the subspace $\langle \vec{v}, \vec{w} \rangle$. If $\vec{v} = \lambda \vec{w}$, the inequality is obvious (it is an equality). Otherwise, $\langle \vec{v}, \vec{w} \rangle$ is two-dimensional. Consider in $\langle \vec{v}, \vec{w} \rangle$ the metrics induced by the metrics g in $\mathbb{R}^4$. Then, $(\langle \vec{v}, \vec{w} \rangle, g)$ is a two-dimensional Minkowski space (since $\vec{v}$ is timelike the metrics has index≥ 1; and, since the original metrics has index 1, so has the induced one). Therefore, the result follows from Lemma 2.4.4. $\square$

Proposition 2.4.6. *Let $\vec{v}, \vec{w} \in C^+$ in the Minkowski space. Then*

$$\frac{1}{i}\sqrt{g(v+w,v+w)} \geq \frac{1}{i}\sqrt{g(v,v)} + \frac{1}{i}\sqrt{g(w,w)}.$$

Proof. Note that if both $\vec{v}, \vec{w} \in C^+$, then $g(v+w,+w), g(v,v), g(w,w) < 0$ and they have imaginary square roots. Dividing by i we get positive real numbers. Making squares, the inequality to be proved is equivalent to

$$-g(v+w,v+w) \geq -g(v,v) - g(w,w) - 2\sqrt{g(v,v)}\sqrt{g(w,w)} \Longrightarrow$$

$$\Longrightarrow g(v+w,v+w) \leq g(v,v) + g(w,w) + 2\sqrt{g(v,v)}\sqrt{g(w,w)}.$$

Observe now that, since

$$g(v+w,v+w) = g(v,v) + g(w,w) + 2g(v,w),$$

we need to prove

$$g(v,w) \leq \sqrt{g(v,v)}\sqrt{g(w,w)}.$$

Observe that the inner product of two vectors of C^+ is negative and then $g(v,w), g(v,v), g(w,w) < 0$. Thus, the previous inequality is equivalent to

$$g(v,w)^2 \geq g(v,v)g(w,w),$$

which is true in virtue of Proposition 2.4.5. $\square$

Definition 2.4.7. *Let P_1, P_2 be two points in the Minkowski space in $\mathbb{R}^4$. We say that P_1 is previous to P_2 if the vector $\overrightarrow{P_1P_2} \in C^+$.*

P_1 and P_2 are, thus, two events such that, being observed from any inertial frame, happen always in the same order: P_1 first and P_2 later.

2.5 Proper time

Let S be an inertial frame; we have seen above that this is equivalent to choosing axes in $\mathbb{R}^4$ with directions $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$, where $\{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4\}$ is a positive ortho-c-normal basis in $\mathbb{R}^4$.

Let r be a straight line in $\mathbb{R}^4$ whose direction vector is timelike. Then r can be parametrized by

$$\begin{cases} x_i = w_i t + b_i, & i = 1,2,3 \\ t = t \end{cases}$$

where $w_1^2 + w_2^2 + w_3^2 < c^2$. In S, r represents the uniform motion of a particle P at speed $\vec{w} = (w_1, w_2, w_3)$ (in fact, every straight line with timelike direction vector represents the uniform motion of a particle in an inertial frame).

We know, though, that there is an inertial frame S' whose origin is P (being the point S' stationary in S').

Let P_1, P_2 be two points in r. Without loss of generality, we consider P_1 to be previous to P_2 (which can be done since r has timelike direction vector). In S', both P_1 and P_2 correspond to the origin, but observed at instants t_1' and t_2', where $t_1' < t_2'$. Thus their coordinates in S' are $(0,0,0,t_1')$ and $(0,0,0,t_2')$, respectively. Then the distance in the Minkowski space is

$$\mathrm{d}(P_1, P_2) = \sqrt{g(P_2 - P_1, P_2 - P_1)} = ic(t_2' - t_1').$$

Therefore, the time interval $t_2' - t_1'$ in the frame where P is stationary is $\mathrm{d}(P_1, P_2)/ic$. And, since distace is an invariant in the Minkowski space, we have that calculating $\mathrm{d}(P_1, P_2)/ic$ gives always the length of the time interval between P_1 and P_2 measured in the frame in which P is stationary.

It is easy to observe the following: let us calculate the coordinates of P_1 and P_2 using the equations 2.17 of the Lorentz transformation. Thus one obtains

$$P_1 = \Big(\frac{vt_1'}{\sqrt{1-\frac{v^2}{c^2}}}, 0, 0, \frac{t_1'}{\sqrt{1-\frac{v^2}{c^2}}}\Big) \text{ and } P_2 = \Big(\frac{vt_2'}{\sqrt{1-\frac{v^2}{c^2}}}, 0, 0, \frac{t_2'}{\sqrt{1-\frac{v^2}{c^2}}}\Big).$$

Then

$$P_2 - P_1 = \Big(\frac{v(t_2' - t_1')}{\sqrt{1-\frac{v^2}{c^2}}}, 0, 0, \frac{(t_2' - t_1')}{\sqrt{1-\frac{v^2}{c^2}}}\Big) \Longrightarrow$$

$$\Longrightarrow \sqrt{g(P_2 - P_1, P_2 - P_1)} = \sqrt{\frac{v(t_2' - t_1')}{\sqrt{1-\frac{v^2}{c^2}}} - \frac{c^2(t_2' - t_1')}{\sqrt{1-\frac{v^2}{c^2}}}} =$$

$$= \sqrt{-c^2(t_2' - t_1')\frac{\cancel{v^2 - c^2}}{\cancel{v^2 - c^2}}} = ic(t_2' - t_1'),$$

and so $\mathrm{d}(P_1, P_2) = ic(t_2' - t_1')$ no matter the chosen inertial frame. Note also that any observer sees the events P_1 and P_2 in the same order, since $\overrightarrow{P_1P_2} \in C^+$.

Assume now that an observer travels with the particle P_1 moving from a point A to B, and then changing it direction towards a third point C. This is not a uniform motion but the sum of two different uniform motions. Events P_A, P_B and P_C are seen according to then inertial frame as (A_i, t_A), (B_i, t_B) and (C_i, t_C), respectively, and $\overrightarrow{P_A P_B}, \overrightarrow{P_B P_C} \in C^+$. Observe that the observer traveling on P will measure a time interval $\mathrm{d}(P_A, P_B)/ic$ from A to B, and $\mathrm{d}(P_B, P_C)/ic$ from B to C, since both of these motions are uniform. We state that the time of this observer suffers no change when changing direction at point B, and so that the time measured by him for the whole motion from A to C is

$$\frac{1}{ic}\big(\mathrm{d}(P_A, P_B) + \mathrm{d}(P_B, P_C)\big).$$

Assume now that an observer travels with the particle P at speed $v < c$ describing a smooth curve γ in $\mathbb{R}^4$ relative to an inertial frame S, whose tangent vector at every point is timelike. Let P_A and P_B be the first and the last point of γ, respectively. The question arisen is: which will be the time recorded by the observer traveling with P along γ between the events P_A and P_B?

Take n points on γ such that $P_1 = P_A$, $P_n = P_B$ and $\overrightarrow{P_i P_{i+1}} \in C^+$ for $i = 1, ..., n-1$. This concatenation of straight lines corresponds to the motion of a particle P' approximating γ. The time recorded by an observer traveling with P' between P_A and P_B will be

$$\frac{1}{ic}\sum_{i=1}^{n-1} \mathrm{d}(P_i, P_{i+1}).$$

We state then that, tending proper times correspond to trending trajectories, and so taking $n \longrightarrow +\infty$ we get that the time interval recorded by an observer traveling with P is

$$\tfrac{1}{ic} \text{ length of } \gamma \text{ between } P_A \text{ and } P_B.$$

Hence, the time interval between two events P_A and P_B recorded by an observer moving along a piecewise smooth curve with positive timelike tangent vector at every point is the length of γ between P_A and P_B divided by ic. This is the so-called proper time of this observer.

Example 2.5.1. [The twin paradox] Jack and Joe are two twin brothers living in an inertial frame S. One day, Joe takes a rocket and starts traveling away from Jack, who remains stationary in S, with uniform motion following the $x-$axis. Suddenly, he stops, turns back and meets his brother again. Let us study this case and its shocking consequences.

Jack is always in S, and so represents events with coordinates (x, t). In fact, since he can be considered to remain stationary at the origin of S, his coordinates are $(0, t)$. Joe's motion will be represented, however, by two straight lines $t = x/v$ and $t = -x/v + a$, being a the point where Joe turns back.

Let A be the event where Joe leaves Jack, B the event where Joe turns back and C the event where Joe meets Jack again, in the Minkowski space.

Thus, Joe's proper time is $\frac{1}{ic}\big(\mathrm{d}(P_A, P_B) + \mathrm{d}(P_B, P_C)\big)$, whereas Jack's proper time is $\frac{1}{ic}\big(\mathrm{d}(P_A, P_C)\big)$. Then, by Proposition 2.4.6, we have

$$\frac{1}{ic}\big(\mathrm{d}(P_A, P_B) + \mathrm{d}(P_B, P_C)\big) \leq \frac{1}{ic}\big(\mathrm{d}(P_A, P_C)\big)$$

and hence Joe is now younger than Jack: they are twins no more!

2.6 Dynamics of special relativity

We shall now check the concepts and properties given in the first chapter in terms of special relativity. We shall keep everything as unchanged as possible, as long as no contradiction with the Postulates of Special Relativity arises. We assign a mass to every body (the mass of the body when it stays stationary). We could thus define $\vec{\mathcal{P}}$ the momentum relative to an inertial frame as $\vec{\mathcal{P}} = m\vec{v}$, being $\vec{v}$ the speed of the particle relative to such inertial frame. One could think, as in Classical Mechanics, that the momentum remains constant before and after a collision. But we shall observe that this does not hold. Consider two particles with masses m and $2m$, and speeds $\vec{u}$ and $-\vec{u}$, respectively, which collide frontally. Since their motions follow the same straight-lined path, we can consider the case to be one-dimensional and take $u, -u$ scalars. In a classical collision, the particles will have speeds $-5u/3$ and $u/3$ after colliding. Thus, we have, before the collision, that

$$\mathcal{P} = mu - 2mu = -mu,$$

and, after it, that

$$\mathcal{P} = -\frac{5mu}{3} + \frac{2mu}{3} = -mu.$$

We observe now the same event from another inertial frame traveling on a common $xx'-$axis at speed v relative to the first. Denote by S the first inertial frame and S' the one just defined. Applying the formuli for speed transformation, we get that a particle traveling at speed u in S will have speed u' in S', given by

$$u' = \frac{1}{1 - \frac{uv}{c^2}}(u - v).$$

Thus, the momentum before the collision in S' is

$$\frac{m}{1 - \frac{uv}{c^2}}(u - v) + \frac{2m}{1 + \frac{uv}{c^2}}(-u - v)$$

and the momentum after the collision is

$$\frac{m}{1 + \frac{5uv}{3c^2}}\big(-\frac{5}{3}u - v\big) + \frac{2m}{1 - \frac{uv}{3c^2}}\big(\frac{u}{3}\big),$$

which is different from the one in S, unless $v = 0$. This is a problem, since we have postulated that all physical laws are equally valid regardless of the inertial frame we are in. We have to work out a smarter statement of the collision laws then. Recall that the motion of any particle can be interpreted as a curve in the Minkowski space, with timelike tangent vectors, which can be parametrized in terms of the proper time of the particle. Let then $\vec{\gamma}(\tau)$ be such a curve in the Minkowski space. We define the Minkowski speed and the Minkowski acceleration of the particle as $\frac{d\vec{\gamma}(\tau)}{d\tau}$ and $\frac{d^2\vec{\gamma}(\tau)}{d\tau^2}$, respectively. Note that these are four-dimensional vectors, unlike the classical three-dimensional speed and acceleration. If a particle with mass m moves corresponding to a curve $\vec{\gamma}(\tau)$ in the Minkowski space, we define its Minkowski momentum at an instant τ of the particle's proper timeas $\vec{\mathcal{P}} = m\frac{d\vec{\gamma}(\tau)}{d\tau}$. And we define the Minkowski force as the vector $\frac{d\vec{\mathcal{P}}}{d\tau} =$

$m\frac{d^2\vec{\gamma}(\tau)}{d\tau^2}$. We postulate that, whenever two particles collide, the Minkowski momentum of the system (the sum of the Minkowski momentums of the particles of the system) remains constant before and after the collision. Observe that both the curve $\vec{\gamma}(\tau)$ and the proper time are geometric concepts independent of the chosen inertial frame: different inertial frames give different coordinate references so that the coordinates of $\vec{\gamma}(\tau)$ change, but the points in the Minkowski space (recall that every point in the Minkowski space represents an event) are the same. Thus, the Minkowski speed does not depend on the choice of the inertial frame, and so neither does the postulate we just stated. We shall now study the relation between the classical speed and momentum, and the new ones just defined. Let S be an inertial frame and consider the motion of a particle in S, given by a curve in $\mathbb{R}^4$ parametrized by

$$\begin{cases} x_1 = \gamma_1(t) \\ x_2 = \gamma_2(t) \\ x_3 = \gamma_3(t) \\ x_4 = t. \end{cases}$$

Since the time t of S can be written in terms of the proper time τ of the particle, we can parametrize the motion curve as

$$\begin{cases} x_1 = \gamma_1(\tau) \\ x_2 = \gamma_2(\tau) \\ x_3 = \gamma_3(\tau) \\ x_4 = t(\tau). \end{cases}$$

The Minkowski speed will be then

$$\vec{u} = \frac{d\vec{\gamma}}{d\tau} = \Big(\frac{d\gamma_1}{d\tau}, \frac{d\gamma_2}{d\tau}, \frac{d\gamma_3}{d\tau}, \frac{dt}{d\tau}\Big) =$$

$$= \Big(\frac{d\gamma_1}{dt}, \frac{d\gamma_2}{dt}, \frac{d\gamma_3}{dt}, 1\Big)\frac{dt}{d\tau} = (v_1, v_2, v_3, 1)\frac{dt}{d\tau},$$

where (v_1, v_2, v_3) is the usual speed of the particle relative to S. By definition of proper time, we have

$$\tau(t) = \frac{1}{ic}\int_0^t \sqrt{g\big(\frac{d\vec{\gamma}}{d\xi}, \frac{d\vec{\gamma}}{d\xi}\big)}d\xi.$$

Then

$$\frac{d\tau}{dt} = \frac{1}{ic}\sqrt{g\big(\frac{d\vec{\gamma}}{dt}, \frac{d\vec{\gamma}}{dt}\big)} = \frac{1}{ic}\sqrt{g\big((v_1, v_2, v_3, 1), (v_1, v_2, v_3, 1)\big)} = \frac{1}{c}\sqrt{c^2 - v^2} = \sqrt{1 - \frac{v^2}{c^2}},$$

where $v = \|\vec{v}\|$. Hence,

$$\vec{u} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}(v_1, v_2, v_3, 1).$$

Then, the Minkowski momentum will be

$$\vec{\mathcal{P}} = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}}(v_1, v_2, v_3, 1).$$

Observe that the first three components of $\vec{\mathcal{P}}$ define the vector $\vec{p} = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\vec{v}$, which depends on the inertial frame and tends to the classical momentum vector for speeds $\vec{v}$ small compared with the speed of light. This vector $\vec{p}$ would be the classical momentuam of a particle with mass

$$M = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}.$$

On the other hand, the fourth component of $\vec{\mathcal{P}}$ is precisely M. Expanding by Taylor the function $\frac{1}{\sqrt{1-x^2}}$ at the origin, we have

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{x^2}{2} + \cdots,$$

and then

$$M = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}} \simeq m + \frac{1}{c^2}\Big(\frac{1}{2}mv^2\Big) + \cdots,$$

that is, the mass enlarged by the classical kinetic energy divided by c^2.

Equivalence mass-energy

Consider two particles of mass m moving along the $x-$ axis of an inertial frame with speeds v and $-v$. Let us study the collision of these particles. Their Minkowski momentum before the collision will be

$$\vec{\mathcal{P}} = \Big(\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}, 0, 0, \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\Big) + \Big(-\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}, 0, 0, \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}\Big) =$$

$$\Big(0, 0, 0, \frac{2m}{\sqrt{1-\frac{v^2}{c^2}}}\Big).$$

Name v_1' and v_2' the speeds of the two particles after the collision. Then the Minkowski momentum after the collision is

$$\Big(\frac{mv_1'}{\sqrt{1-\frac{(v_1')^2}{c^2}}}, 0, 0, \frac{m}{\sqrt{1-\frac{(v_1')^2}{c^2}}}\Big) + \Big(\frac{mv_2'}{\sqrt{1-\frac{(v_2')^2}{c^2}}}, 0, 0, \frac{m}{\sqrt{1-\frac{(v_2')^2}{c^2}}}\Big).$$

Thus, we get

$$\begin{cases} \frac{mv_1'}{\sqrt{1-\frac{(v_1')^2}{c^2}}} + \frac{mv_2'}{\sqrt{1-\frac{(v_2')^2}{c^2}}} = 0 \\ \frac{m}{\sqrt{1-\frac{(v_1')^2}{c^2}}} + \frac{m}{\sqrt{1-\frac{(v_2')^2}{c^2}}} = \frac{2m}{\sqrt{1-\frac{v^2}{c^2}}}. \end{cases}$$

The solutions are $v_2' = -v_1' = \pm v$. A simple examination shows that the only solution is $v_2' = v_1' = v$, which coincides with the classical elastic collision. But one observes that the inellastic case seems not to be a possibility in the previous example, and it should be according to experience. If the two particles remained together after the collision, we would have a Minkowski momentum $(0,0,0,m')$ and, by the principle of the momentum conservation, we would have

$$m' = \frac{2m}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Then, we could have that the mass of the system after the collision would be bigger than before, since we have $m' > 2m$. In the previus system we had implicitely assumed that the mass remained constant. The mass difference will be

$$m' - 2m = 2m\Big(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\Big) \simeq \frac{\cancel{2}mv^2}{\cancel{2}c^2} = \frac{1}{c^2}mv^2.$$

Thus, $m' - 2m$ is similar to the kinetic energy divided by c^2, which leads us to thiking that the kinetic energy is somehow used to increase the mass of the system. Hence, if the particles remain together after the collision, either the mass increases in relation to the kinetic energy, or either this energy becomes heat. Anyway, we need to come into terms with the idea that energy is equivalent to mass multiplied by c^2; otherwise, we ought to dismiss the principle of conservation of the Minkowski momentum. From now on, we call stationary energy of a particle or system of particles with mass m to

$$E = mc^2.$$

And we call energy of a particle or system of particles with mass m relative to an inertial frame to

$$\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} = mc^2 + \frac{1}{2}mv^2 + \cdots,$$

which is the stationary energy plus a term $T = \frac{1}{2}mv^2$ which coincides with the kinetic energy for small velocities. Thus, the energy is the fourth component of the Minkowski momentum vector multiplied by c^2, whenever it is written using the reference coordinates of the Minkowski space of the inertial frame in which we measure the energy. To sum up, $\vec{\mathcal{P}}$ is an intrinsic vector of the Minkowski space, but given an inertial frame (with its coordinates), the first three components of $\vec{\mathcal{P}}$ relative to this frame are

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}},$$

the momentum relative to the inertial frame. The fourth component is

$$\frac{m}{\sqrt{1-\frac{v^2}{c^2}}},$$

which is E/c^2, where E is the energy, depending on the inertial frame. The sum of the Minkowski momentum remains constant before and after any collision and, whenever the sum of masses remains constant too, we say that the collision is elastic.

The Minkowski force vector

Let S be an inertial frame (then we have an ortho-c-nrmal reference of the Minkowski space). Thus,

$$\vec{\mathcal{P}} = (p_1, p_2, p_3, E/c^2).$$

The Minkowski force vector $\vec{F} = \frac{d\vec{\mathcal{P}}}{d\tau}$ is

$$\vec{F} = (F_1, F_2, F_3, F_4).$$

The first three components generate a three-dimesional vector denoted by $\vec{f}$:

$$\vec{f} = \frac{d\vec{p}}{d\tau} = \frac{\vec{p}}{dt}\frac{dt}{d\tau} = \frac{d}{dt}\Big(\frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\Big)\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}.$$

And

$$\frac{d}{dt}\Big(\frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}}\Big) = \frac{m\frac{d\vec{v}}{dt}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{m}{c^2}\Big(1-\frac{v^2}{c^2}\Big)^{-\frac{3}{2}}\vec{v}\frac{dv}{dt}.$$

For small speeds and accelerations, the terms v^2/c^2 and $\frac{\vec{v}}{c^2}\frac{dv}{dt}$ are negligible and $\vec{f}$ coincides with the classical Newtonian force. Observe now that

$$F_4 = \frac{d}{d\tau}\Big(\frac{E}{c^2}\Big),$$

which is the variation of energy with respect to the proper time, divided by c^2. Let us find another expression for F_4. Observe that $\vec{F}$ and the Minkowski speed vector $\vec{u}$ are orthogonal: let $\vec{\gamma}(t)$ be the curve in the Minkowski space corresponding to the particle's life. Then

$$\frac{d\vec{\gamma}}{d\tau} = \frac{d\vec{\gamma}}{dt}\frac{dt}{d\tau} = \frac{d\vec{\gamma}}{dt}\frac{ic}{\sqrt{g(d\vec{\gamma}/dt, d\vec{\gamma}/dt)}}.$$

Then

$$g(d\vec{\gamma}/d\tau, d\vec{\gamma}/d\tau) = -\frac{c^2}{g(d\vec{\gamma}/dt, d\vec{\gamma}/dt)}g(d\vec{\gamma}/dt, d\vec{\gamma}/dt) = -c^2.$$

Derivating with respect to τ, we get

$$2g\Big(\frac{d^2\vec{\gamma}}{d\tau^2}, \frac{d\vec{\gamma}}{d\tau}\Big) = 0.$$

Since $\vec{u} = \frac{d\vec{\gamma}}{d\tau}$, multiplying by m, we get

$$g(\vec{F}, \vec{u}) = 0.$$

Recall $\vec{u} = (v_1, v_2, v_3, 1)\frac{dt}{d\tau}$, where $v_i = \frac{d\gamma_i}{dt}$. Then,

$$0 = g(\vec{F}, \vec{u}) = F_1u_1 + F_2u_2 + F_3u_3 - c^2F_4u_4 =$$

$$= \vec{f}\cdot\vec{v}\frac{dt}{d\tau} - c^2\frac{d}{d\tau}\Big(\frac{E}{c^2}\Big)\frac{dt}{d\tau},$$

which implies
$$\frac{dE}{d\tau} = \vec{f} \cdot \vec{v}.$$
Hence,
$$\frac{dE}{dt} = \frac{dE}{d\tau}\frac{d\tau}{dt} = \vec{f} \cdot \vec{v}\sqrt{1 - \frac{v^2}{c^2}}.$$
The second term of this expression coincides with the classical power for small speeds compared with the speed of light. We shall call relativistic power of the particle relative to the chosen inertial frame to
$$W = \vec{f} \cdot \vec{v}\sqrt{1 - \frac{v^2}{c^2}}.$$
Then, power is the variation of energy as happens in Classical Mechanics. Therefore,

$$F_4 = \frac{d}{d\tau}\left(\frac{E}{c^2}\right) = \frac{1}{c^2}\frac{dE}{dt}\frac{dt}{d\tau} = \frac{1}{c^2\sqrt{1 - \frac{v^2}{c^2}}}W.$$

2.7 Electromagnetism and special relativity

Recall that we characterized continuum electrical charges in motion by means of the charge density ρ and the current density vector $\vec{J}$, tangent to the trajectories in such a way that the charge density of a region of space V is

$$Q(V) = \int_V \rho(x)dx,$$

and $\vec{J} = \rho\vec{v}$, where $\vec{v}$ is the field of speeds of the charged particles. Such a system generates an electric field $\vec{E}$ and a magnetic field $\vec{B}$ fulfilling the Maxwell equations

$$\begin{cases} \mathrm{div}E = 4\pi\rho \\ \mathrm{rot}B = \frac{4\pi}{c}J + \frac{1}{c}\frac{\partial E}{\partial t} \\ \mathrm{rot}E = -\frac{1}{c}\frac{\partial B}{\partial t} \\ \mathrm{div}B = 0 \\ \frac{\partial \rho}{\partial t} + \mathrm{div}J = 0. \end{cases}$$

Recall also that on a particle with charge q with speed $\vec{v}$ acts a force caused by these fields, being such force

$$\vec{F} = q\vec{E} + \frac{q}{c}\vec{v} \wedge \vec{B} \ (\textit{Law of Lorentz}).$$

This needed an absolute space and an absolute time, since it was not invariant under changes of inertial frames. Thus we want to find now intrinsic laws keeping the spirit of the classical ones. Let P be a particle with charge q in the moving system. At any instant τ_0 of its proper time, we consider an inertial frame S' such that, at $t' = 0$, the origin of S' coincides with P and the speed of P relative to S' is zero at that instant; this is the so-called proper frame of P at τ_0. If we measure in S' the charge density at the origin at $t' = 0$, we get a number σ, the so-called proper charge density of P at τ_0, and depends only on P

and its proper time. Let $\vec{\gamma}(\tau)$ be the curve in the Minkowski space describing the life of P, parametrized by the proper time τ of P. Let $\vec{u} = d\vec{\gamma}/d\tau$ be the intrinsic Minkowski speed vector. Denote by $\vec{j} = \sigma\vec{u}$, an intrinsic vector in the Minkowski space. Consider another inertial frame S from which we observe the moving system of electrical charges. Such a frame will give us a coordinate reference in the Minkowski space. Let $\vec{E}, \vec{B}$ be the electric and magnetic fields, respectively, generated by the system and observed from S. Let α be the differential 2-form in the Minkowski space given, in terms of coordinates of S, by

$$\alpha = 2c\sum_j E_j dx_j \wedge dt + 2B_1 dx_2 \wedge dx_3 + 2B_2 dx_3 \wedge dx_1 + 2B_3 dx_1 \wedge dx_2.$$

Observe that $\mathrm{div}B = 0$ and $\frac{\partial B}{\partial t} + c\,\mathrm{rot}E = 0$ are equivalent to the only equation $d\alpha = 0$. Furthermore,

$$\vec{j} = \sigma\vec{u} = \sigma\frac{d\vec{\gamma}}{dt}\frac{dt}{d\tau} = \sigma(v_1, v_2, v_3, 1)\frac{dt}{d\tau}.$$

For small speeds compared with the speed of light, the first three components are approximately the classical vector $\vec{J}$, whereas the fourth component will be the charge density relative to S. Let us compute the divergence of α. We have that the components α_{ij} of α are given by the antisymmetric matrix

$$(\alpha_{ij}) = \begin{pmatrix} 0 & -B_3 & B_2 & -cE_1 \\ B_3 & 0 & -B_1 & -cE_2 \\ -B_2 & B_1 & 0 & -cE_3 \\ cE_1 & cE_2 & cE_3 & 0 \end{pmatrix}.$$

From now on, we shall use the first index for columns and the second one for rows, and thus we have

$$(\mathrm{div}\alpha)^i = \bigtriangledown_j \alpha^{ij},$$

where

$$(\alpha^{ij}) = \begin{pmatrix} 0 & -B_3 & B_2 & E_1/c \\ B_3 & 0 & -B_1 & E_2/c \\ -B_2 & B_1 & 0 & E_3/c \\ -E_1/c & -E_2/c & -E_3/c & 0 \end{pmatrix}.$$

We get then

$$\mathrm{div}\alpha = \left(\mathrm{rot}B + \frac{1}{c}\frac{\partial E}{\partial t}, \frac{1}{c}\mathrm{div}E\right) \overset{Maxwell}{=} \frac{4\pi}{c}(\vec{J}, \rho).$$

Thus we get the instrinsic equations equivalent to the Maxwell equations for small speeds compared with the speed of light, and small accelerations. These equations are given by

$$\left\{ \begin{array}{l} \mathrm{div}\alpha = \frac{4\pi}{c}\vec{j} \\ d\alpha = 0 \end{array} \right).$$

Observe that the differential 2-form α is postulated to be intrinsic; that means that, despite the fact that E and B change under changes of inertial frames, they will always generate the same α. Observe also that $\mathrm{div}(\mathrm{div}\alpha) = 0$ holds in the Minkowski space: we know that

$$\mathrm{div}(\mathrm{div}\alpha) = \bigtriangledown_j \bigtriangledown_i \alpha^{ij}.$$

Since the Christoffel symbols relative to linear frames are zero in the Minkowski space, then

$$\mathrm{div}(\mathrm{div}\alpha) = \partial_i \partial_j \alpha^{ij},$$

where $\partial_i = \partial/\partial x_i$. And observe that, interchanging the index names, we get

$$\partial_i \partial_j \alpha^{ij} = \partial_j \partial_i \alpha^{ji} \overset{Schwarz}{=} \partial_i \partial_j \alpha^{ji} \overset{antisymmetry}{=} -\partial_i \partial_j \alpha^{ij} \Longrightarrow \partial_i \partial_j \alpha^{ij} = 0.$$

Then $\mathrm{div} j = \frac{c}{4\pi} \mathrm{div}(\mathrm{div}\alpha) = 0$, and

$$0 = \mathrm{div} j = \frac{\partial j_1}{\partial x_1} + \frac{\partial j_2}{\partial x_2} + \frac{\partial j_3}{\partial x_3} + \frac{\partial j_4}{\partial t} \simeq\simeq \mathrm{div}\vec{J} + \frac{\partial \rho}{\partial t}.$$

Thus, the equation $\mathrm{div} j$ substitutes the classical equation of continuity

$$\mathrm{div}\vec{J} + \frac{\partial \rho}{\partial t} = 0,$$

but we do not need to impose it since it is deduced frm

$$\mathrm{div}\alpha = \frac{4\pi}{c} j.$$

Force acting on a charged particle in an electromagnetic field

Consider a particle with charge q in an electromagnetic field (as before, $\vec{E}$ denotes the electric field and $\vec{B}$ denotes the magnetic field). Let $\vec{\gamma}(\tau)$ be the corresponding curve parametrized by the proper time of the particle in the Minkowski space and let $\vec{u} = d\vec{\gamma}/d\tau$ be the corresponding Minkowski speed. Let F be the only vector in the Minkowski space such that

$$g(X, F) = \tfrac{q}{c} \alpha(X, u), \forall X,$$

where g is the Minkowski metrics and α is the electromagnetic field. Fix an inertial frame S, which gives us a coordinate reference in the Minkowski space. Using these coordinates, we get

$$F_j = \frac{q}{c} u_i \alpha_i^j,$$

where $\alpha_i^j = g^{jk} \alpha_{ki}$. Thus,

$$(\alpha_i^j) = \begin{pmatrix} 0 & -B_3 & B_2 & E_1/c \\ B_3 & 0 & -B_1 & E_2/c \\ -B_2 & B_1 & 0 & E_3/c \\ cE_1 & cE_2 & cE_3 & 0 \end{pmatrix}.$$

Since $\vec{u} = (v_1, v_2, v_3, 1) \frac{dt}{d\tau}$, we have

$$\begin{aligned} F_1 &= \tfrac{q}{c}(B_3 v_2 - B_2 v_3) \tfrac{dt}{d\tau} + q E_1 \tfrac{dt}{d\tau} \\ F_2 &= \tfrac{q}{c}(-B_3 v_1 + B_1 v_3) \tfrac{dt}{d\tau} + q E_2 \tfrac{dt}{d\tau} \\ F_3 &= \tfrac{q}{c}(B_2 v_1 - B_1 v_2) \tfrac{dt}{d\tau} + q E_3 \tfrac{dt}{d\tau} \\ F_4 &= \tfrac{q}{c}(E_1 u_1 + E_2 u_2 + E_3 u_3). \end{aligned}$$

Let $\vec{f}$ denote the vector $\vec{f} = (F_1, F_2, F_3)$, then

$$\vec{f} = \left(\frac{q}{c}\vec{v} \wedge \vec{B} + q\vec{E}\right)\frac{dt}{d\tau};$$

recall that $\frac{q}{c}\vec{v} \wedge \vec{B} + q\vec{E}$ is the classical expression of the force caused by an electromagnetic field acting on a particle with charge q. And the fourth component F_4 is the power of the force $\vec{f}$. Hence, the law of Lorentz has to be substituted by the intrinsic law stated as follows: if a particle with electrical charge q is in an electromagnetic field, the such a particle is affected by a Minkowski force $\vec{F}$ related to the electromagnetic field α by

$$g(\vec{F}, \vec{X}) = \tfrac{q}{c}\alpha(\vec{X}, \vec{u}),\ \forall X,$$

where $\vec{u}$ is the speed of the particle in the Minkowski space.

Chapter 3

General Relativity

As we have seen in the previous chapter, the Special Relativity Theory compares the events and physical laws seen by inertial observers in uniform motion relative to each other; the motion of such inertial frames can only be described as motion with respect to another inertial frame. And, as we have already said, no experiment carried out entirely within one inertial frame can ever reveal the state of motion of that frame.

In this chapter, we shall deal with frames with accelerated motion, whose effects are rapidly felt. The differences between uniform and accelerated frames were a problem to Einstein, who wanted his theory to embrace all kinds of frames regardless of their state of motion.

He made it in his 1916 paper *"Die Grundlage der allgemeinen Relativitätstheorie"* ("The foundations of General Relativity"), in which he extended his relativistic theory to arbitrary frames. In the following sections, we shall develop the principal ideas of General Relativity and its geometry.

3.1 The Principle of Equivalence

Since the time of Galileo Galilei (1564-1642) it was known that bodies under the influence of gravity fall with same acceleration (under negligible resistance); this is also an immediate consequence of the laws formulated by Sir Isaac Newton. This property was confirmed by Neil Armstrong carrying out an experiment on the Moon's surface in 1969.

According to Newton's Second Law of Motion, the force needed to impart a given acceleration to a body is proportional to the body's mass. And, according to Newton's Law of Gravitation, the force exerted by the gravity of the Earth on an object is proportional to the object's mass, in such a way that the acceleration caused is the same for any object. One could somehow think that the Earth, as gravitation source, "senses" the object mass and modules the gravitational force that needs to be exerted on the object in order to cause the same gravity acceleration for any body.

Einstein realized that this feature of gravity exhibits two aspects of mass which had never been satisfactorily explained. He therefore talked about inertial mass (the mass appearing in Newton's Second Law) as a way of measuring an object's resistance to a change in its motion, that is, its resistance to experiment an acceleration; and the gravitational mass (the mass appearing in the Law of Gravitation) as a measure of an object's response to gravitational attraction. Einstein realized that physicists had been using the word "mass" to express two different attributions of matter, and solved the problem by postulating the so-called Principle of Equivalence.

We can express his idea by means of a "thought experiment".

Consider thus a tall building and an elevator in it, which is suddenly cut loose from its cables and begins to fall freely (neglect air resistance or friction). Observe that, since the elevator and the objects within fall with the very same acceleration relative to the ground, they do not experiment acceleration relative to each other. That is, if a passenger had been holding a package and releases it, such a package would appear to be floating. Similarly, if he pushes it away, it would move away from him moving in a straight-lined path with constant speed relative to him and the elevator.

Therefore, the elevator in free fall is, apparently, an inertial frame. If the passengers were not to be aware of where they are, they might believe that they are in a space station floating in outer space.

As we shall soon see, Einstein's Principle of Equivalence implies that the falling elevator and the floating station are in fact equivalent, in the sense that if one were suddenly placed in one of them, he would have no way of telling in which one he is; there is no internal experiment you could perform which tell in which frame you are located.

Observe that this equivalence of systems, this impossibility of determine whether your weightless state is due to free fall in a gravitational field or by uniform motion in absence of a gravitational field, implies that gravity and acceleration cease to have an absolute meaning, but depends on the observer. We shall now perform a second thought experiment to clarify this.

Consider a group of observers in a windowless rocket in a region of space without gravitational influence of any celestial bodies. Such a rocket moves with constant acceleration (relative to a nearby inertial observer). Due to the acceleration, the observers inside the rocket will be pressed against the wall opposite to the direction of the acceleration; this wall will then become the "floor", and the opposite one, the "ceiling", according to the observers. If one of the inhabitants of the rocket had been holding a package and releases it, it will fall to the "floor". This is caused by the Law of Inertia, by which the package remains with the same state of motion it had when released, relative to an inertial observer; but the rocket gains speed, and therefore the package "falls" to te "floor". And all the objects fall with the same acceleration regardless of its mass. Therefore, if the rocket had an acceleration of $9{,}8m/s^2$, the inhabitants of the rocket might be fooled into thinking that they are in a stationary windowless laboratory on the Earth. Again, no experiment carried out entirely inside the rocket can reveal if it is a rocket or a stationry laboratory as described above.

This to last examples are given by Faber in [2], and with them the author introduces the Principle of Equivalence. Let us now see how the same concept is introduced by Girbau in [1]; he does so with an example. Let S be an inertial frame and let S' be a disk of radious R located in the plane $z = 0$ of S. Assume that S' can spin around his center O and that O is the origin of S, and that S' is now at rest. Assume that we take polar coordinates (r, θ), being r the distance from any point $P \in S'$ to O, and θ the angle between $\bar{PO}$ and the $x-$axis of S.

Assume that at the instant $t = 0$ of S, the disk starts spinning with constant angular speed ω, assuming $\omega R < c$. Any point $P \in S'$ with polar cordinates (r, θ) will describe the curve relative to S,

$$\begin{cases} x(t) = r\cos(\omega + \theta) \\ y(t) = r\sin(\omega + \theta) \\ z = 0, \end{cases} \tag{3.1}$$

being (x, y, z) the coordinates of P in S, and t the time in S. Since $z = 0$, we shall not take it in account. Once the disk S' is in motion, the observers in S' will use the coordinates (r, θ) for every point, whereas the observers in S will describe the motion of a point P with coordinates (r, θ) as a curve in the 3-dimensional Minkowski space

$$\gamma(t) = \big(x(t), y(t), t\big),$$

being $x(t)$ and $y(t)$ given by equation 3.1. Observe that $\|\dot{\gamma}(t)\| = \sqrt{\omega^2 r^2 - c^2}$, non depending on time. Therefore, the proper time of P is

$$\tau = \frac{1}{ic}\big(\sqrt{\omega^2 r^2 - c^2}\big)t = \sqrt{1 - \frac{r^2\omega^2}{c^2}}t. \tag{3.2}$$

Thus, the proper time depends on r, and so, in S' the bigger the radious is, the slower the clocks are. The observers in S' do not have a common time coinciding with the proper time; however, we can take as common time for S' the time at the origin $r = 0$ by using equation 3.2. Thus, any observer in S' can express any event by terms of the coordinates (r, θ, t), related with the coordinates in S by

$$\begin{cases} x(t) = r\cos(\omega + \theta) \\ y(t) = r\sin(\omega + \theta) \\ t = t. \end{cases} \tag{3.3}$$

Any stationary point in S' with coordinates (r_0, θ_0) describes a curve $t \longrightarrow (r_0, \theta_0, t)$ in spacetime which, using coordinates in S will be the helice

$$\gamma(t) : \begin{cases} x(t) = r_0\cos(\omega + \theta_0) \\ y(t) = r_0\sin(\omega + \theta_0) \\ t = t \end{cases}$$

which, being parametrized in terms of proper time, is

$$\gamma(t) : \begin{cases} x(\tau) = r_0\cos(\frac{\omega}{\alpha_0}\tau + \omega_0) \\ y(\tau) = r_0\sin(\frac{\omega}{\alpha_0}\tau + \omega_0) \\ t(\tau) = \frac{\tau}{\alpha_0}, \end{cases}$$

where $\alpha_0 = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}}$. The acceleration vector will be

$$\frac{d^2\gamma(\tau)}{d\tau^2} = -\frac{\omega^2 r_0}{\alpha_0^2}\big(\cos(\frac{\omega}{\alpha_0}\tau + \omega_0), \sin(\frac{\omega}{\alpha_0}\tau + \omega_0), 0\big).$$

Then, a stationary point of mass m relative to S' with coordinates (r_0, θ_0) experiments a Minkowski force $md^2\frac{\gamma(\tau)}{d\tau^2}$ with modulus $\frac{m\omega^2 r_0}{\alpha_0^2}$ and pointing out to the origin of S'; this is the so-called centripetal force.

Let us calculate the Minkowski force in terms of S'; if our theory is consistent, it should give the same result, since the Minkowski force is intrinsic.

The observers in S' use the coordinates (r, θ, t); we have to start by writing the Minkowski metric in terms of (r, θ, t). From equation 3.3, we get

$$\begin{cases} dx = \cos(\omega t + \theta)dr - r\omega \sin(\omega t + \theta)dt - r\sin(\omega t + \theta)d\theta \\ dy = \sin(\omega t + \theta)dr + r\omega \cos(\omega t + \theta)dt + r\cos(\omega t + \theta)d\theta \\ dt = dt. \end{cases}$$

Therefore,

$$g = dx^2 + dy^2 - c^2dt^2 = dr^2 + r^2d\theta^2 + 2r^2\omega dt d\theta - (c^2 - r^2\omega^2)dt^2. \tag{3.4}$$

A stationary point in S' with coordinates (r_0, θ_0) will follow the curve $\gamma(t) = (r_0, \theta_0, t)$ with tangent vector $\dot\gamma(t) = (0, 0, 1)$. Thus, according to equation 3.4 we have $\|\dot\gamma(t)\| = \sqrt{-(c^2 - r_0^2\omega^2)}$. The proper time will then be

$$\tau = \frac{\sqrt{-(c^2 - r_0^2\omega^2)}}{ic} t = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}} t.$$

Parametrizing γ in terms of proper time, we get

$$\gamma(\tau) = \big(r_0, \theta_0, \frac{\tau}{\alpha_0}\big),$$

where $\alpha_0 = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}}$.

The acceletarion $\nabla_{\dot\gamma}\dot\gamma$ of γ has coordinates

$$(\nabla_{\dot\gamma}\dot\gamma)^i = \ddot\gamma^i + \Gamma^i_{jk}\dot\gamma^j\dot\gamma^k.$$

Since $\dot\gamma^1 = \dot\gamma^2 = 0$ and $\dot\gamma^3 = \frac{1}{\alpha_0}$, we will have $\ddot\gamma^i = 0$. Then,

$$(\nabla_{\dot\gamma}\dot\gamma)^i = \frac{1}{\alpha_0^2}\Gamma^i_{33}.$$

Let us compute te Christoffel symbols Γ^i_{33}. We know that

$$\Gamma_{jkl} = g_{il}\Gamma^i_{jk} = \frac{1}{2}\Big(\frac{\partial g_{jl}}{\partial x_k} + \frac{\partial g_{kl}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_l}\Big).$$

The metric g is given, at the point (r_0, θ_0, t), by

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r_0^2 & r_0^2\omega \\ 0 & r_0^2\omega & r_0^2\omega^2 - c^2 \end{pmatrix}. \tag{3.5}$$

Then,

$$\Gamma_{33l} = \frac{\partial g_{3l}}{\partial x_3} - \frac{1}{2}\frac{\partial g_{33}}{\partial x_l}.$$

The derivatives of g vanish with respect to all coordinates but r (x_1) and, then, the only $\Gamma_{33l} \neq 0$ is

$$\Gamma_{331} = -\frac{1}{2}\frac{\partial g_{33}}{\partial x_1} = -r_0\omega^2.$$

Thus, $\nabla_{\dot\gamma}\dot\gamma$ have covariant components $\frac{1}{\alpha_0^2}(-r_0\omega^2, 0, 0)$, being these covariant components

$$(\nabla_{\dot\gamma}\dot\gamma)_i = g_{ij}(\nabla_{\dot\gamma}\dot\gamma)^j.$$

To get the contravariant components $(\nabla_{\dot\gamma}\dot\gamma)^i$, we have to apply the matrix g^{-1}, the inverse of g. The inverse of equation 3.5 is

$$g^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r_0^2} - \frac{\omega^2}{c^2} & \frac{\omega}{c^2} \\ 0 & \frac{\omega}{c^2} & -\frac{1}{c^2} \end{pmatrix}. \tag{3.6}$$

Therefore, the contravariant components of $\nabla_{\dot\gamma}\dot\gamma$ are the very same we got above, that is,

$$\nabla_{\dot\gamma}\dot\gamma = \frac{1}{\alpha_0^2}(-r_0\omega^2, 0, 0).$$

Assuming that the particle we are working with has mass m, the Minkowski force will be

$$\vec{F} = \frac{m r_0 \omega^2}{\alpha_0^2}(-1, 0, 0), \tag{3.7}$$

which is the same we got above.

To sum up, a particle of mass m stationary relative to S' with coordinates (r_0, θ_0) experiments a centripetal force given by equation 3.7. Assume now that another particle Q of mass m is freely released at (r_0, θ_0) at the instant $t = 0$ with no initial speed; for observers in S', such a particle Q is in motion. If Q moves following the curve $\gamma(t) = \big(r(t), \theta(t), t\big)$ in the Minkowski space, with $\gamma(0) = (r_0, \theta_0, 0)$, then the condition for no force acting on Q is $\nabla_{\dot\gamma}\dot\gamma = 0$. Using $(r, \theta, t)-$coordinates, the equation $\nabla_{\dot\gamma}\dot\gamma = 0$ can be written

$$\ddot\gamma^i + \Gamma^i_{jk}\dot\gamma^j\dot\gamma^k = 0. \tag{3.8}$$

At $t = 0$, equation 3.8 is written as

$$\ddot\gamma^i = -\Gamma^i_{33}.$$

If the observer in S' expresses the acceleration in terms of its proper time $\tau = \alpha_0 t$ instead of using the official common time, we get an acceleration

$$\frac{\partial^2\gamma^i}{d\tau^2} = \frac{1}{\alpha^i}\ddot\gamma^i = -\frac{1}{\alpha_0^i}\Gamma^i_{33},$$

and that induces him into thinking that a force of components $-\frac{1}{\alpha_0^2}m\Gamma^i_{33}$ is acting on Q, a force with the same modulus but the opposite sense relative to the centrifugal force.

However, an observer of the inertial frame S will think that there is no force acting on the particle Q.

Meanwhile, an observer in S' will attribute the force which acts on the particle Q to a gravitational field. Observe that if we had freely released a particle with mass M instead of m in the same point, it would have experimented the same acceleration; this property is a characteristic from gravitational fields.

Summing up, there is no way to distinguish whether the apparent force acting on freely released particles is an actual force due to a gravitational field or it is a fake force due to the motion relative to inertial observers.

Let us go on studying this example.

The Principle of Equivalence states that acceleration and gravity are equivalent, in the sense that there is no way to distinguish between their effects. This equivalence implies the identification between inertial and gravitational mass. Thus, given an observer seeing particles accelerating relative to him, he could equally attribute their motion either to an acceleration of his frame or to the influence of a gravitational field. As a consequence, motion (uniform or not) has no absolute meaning but exists relative to a frame of reference.

An almost immediate consequance of the Principle of Equivalence is that light rays are bent by gravitational fields. Let us reconsider the rocket example: assume that the rocket has a window through which a ray of light enters the rocket, with a direction perpendicular to the rocket's acceleration.

In the point of view of an outsider inertial observer, the light ray follows a straight-lined path to the opposite wall of the cabin. However, while the light ray travels inside the cabin from the window and towards the opposite wall, the rocket moves a short but non-zero distance, in such a way that it will strike the opposite wall of the cabin a small distance below where it should have struck had the rocket not been accelerating.

So that the inhabitants of the rocket observe the light ray following a curved path. And, according the Principle of Equivalence, a light ray passing through a laboratory located on the Earth has to follow a curved path too, but the deviation from the straight-lined path is too tiny to be naked-eye measured; that is why this feature had never been observed before Einstein proposed his theory.

Let us now mathematically express the Principle of Equivalence by means of a simple thought experiment, in which the relative acceleration between the two obervers (let us call them A and B) is constant in modulus and direction. Consider a system of n particles with masses $m_1, ..., m_n$ such that every one of them exerts a gravitational (or electrostatic) force on each other, which we shall denote by F_{ij} (the force that the ith particle exerts on the jth particle). To make it easier, we shall consider that these forces follow the straight line between the two involved particles, and that $F_{ij} = -F_{ji}$ depending on the distance between particles (by Newton's Third Law of Motion).

Denote by x, y, z, t the coordinates used by observer A and x', y', z', t' the coordinates used by observer B, and assume that the speeds involved are small enough to make a non-relativistic study of the case suffice. Therefore, we take $t = t'$, and we use the notation

$$X = (x, y, z, t),\ X' = (x', y', z', t')$$

for spacial vectors. Denote the location of the ith particle by X_i for observer A, and X_i' for observer B.

Let observer A assume that he is under the influence of a gravitational field; so that, under the point of view of observer A, all particles fall with a same acceleration g (a vector with constant components) and the motion equations are, according to Newton's Second Law of Motion,

$$m_i \frac{d^2 X_i}{dt^2} = m_i g + \sum_{j \neq i} F_{ji}, \tag{3.9}$$

for $i = 1, ..., n$. Assume that observer B moves relative to A in such a way that

$$X' = X - \frac{1}{2} g t^2,$$

that is, frame A accelerates relative to frame B and vice versa. Then, differenciating, we get

$$\frac{d^2 X_i}{dt^2} = \frac{d^2 X_i'}{dt^2} + g$$

and, by equation 3.9, the motion equations according to observer B are

$$m_i \frac{d^2 X_i'}{dt^2} = \sum_{i \neq j} F_{ji}$$

(since F_{ij} depends on $\|X_i - X_j\|$, and $\|X_i - X_j\| = \|X_i' - X_j'\|$, F_{ij} is the same for both observers).

Thus, we see that for observer B there is no gravitational field! Observer A will explain the situation by considering that observer B is in free fall and does not realize about the gravitational field because of the weightless state created by this free fall. On the other hand, observer B will disagree and consider that there is no gravitational field present and that observer A's view is distorted by his acceleration relative to him.

The next and natural question should be: who is right? They both are! That is so since both points of view are equally valid for describing physical events; and so, the Principle of Equivalence is an externsion of the Principle of Relativity, placing all frames on an equal footing regardless of their state of motion.

3.2 Gravity as spacetime curvature. Tidal effects of gravity

In actual world, the thought experiments we have performed in the last section are not entirely possible; that is because of our assumption of an uniform gravitational field (that is, all falling objects have an acceleration constant in modulus and direction relative to an inertial observer), which is an idealization: in reality, the acceleration caused by gravity varies from point to point.

For instance, consider two test particles which are at rest relative to a free falling capsule to the Earth. Suppose that these two capsules are horizontally separated, that is, perpendicular to the line between the capsule and the Earth's surface. Both particles are falling heading towards the center of the Earth, and therefore the distance between them will decrease at an increasing speed.

Similarly, if the two particles are vertically separated (the line between them is the line of the direction of the free falling of the particles), the one which is closer to the Earth's surface will suffer a slightly greater acceleration, due to the fact that the effects of gravity are slightly stronger for being closer to the Earth's center, and hence both particles will draw farther apart.

These are the so-called tidal effects of gravity, due to the nonuniformity in the gravitational field of the Earth. The tidal effects are the cause of oceanic tides, for example; but, regarding our capsule, the effects are so small that, if both particles are initially close and we study their motion for a short interval of time, the acceleration relative to each other (and thus the presence of a gravitational field) will not be detectable, which implies that one would think to be working in an inertial frame.

At the light of these new observations, let us rewrite the Principle of Equivalence: for each event and for a given degree of accuracy, there exists a frame in which, in a sufficiently small spacetime neighbourhood of the event, the effects of gravity are negligible and the frame is inertial to the given degree of accuracy.

This is called a locally inertial frame (at the event), the frame of a free falling observer, and the free falling observer at rest in this frame is called a locally inertial observer. Moreover, a frame moving uniformly relative to a locally inertial frame is also a locally inertial frame (at the same event).

We shall study now a geometric analogue to the relative acceleration of free falling test particles. Consider two travelers setting out from two different points of the equator and travelling northwards along two meridians; despite their paths are initially parallels, they will meet at the north pole because of the curvature of the sphere. Let us write this down in terms of geodesics: let $\alpha(s)$ be the parametrization of one of the traveler's path in terms of arc length ($\alpha(0)$ on the equator) and let $v(s)$ be the distance from $\alpha(s)$ to the other geodesic measured along the parallel passing through $\alpha(s)$. It can be proved that

$$\frac{d^2v}{ds^2} + \frac{v}{R^2} = 0,$$

being R the sphere radious. Recalling that $\frac{1}{R^2}$ is the Gauβ curvature of the sphere and denoting it by K, we have

$$\frac{d^2v}{ds^2} + Kv = 0. \tag{3.10}$$

This equation, called Jacobi's equation or equation of geodesic deviation, turns out to be true regardless of the surface, being the curvature able to vary pointwise and being K a function of arc length along α ($K = K(\alpha(s))$).

These resemblances between the *relative acceleration of free particles in a nonuniform gravitational field* and the *variation of separation of nearby geodesics on a curved surface* are not mere coincidence, but Einstein arrived to the conclusion that gravity is not a force, as believed up to then, but rather the curvature of spacetime whose source is matter itself.

Einstein generalized by Law of Inertia, that is, free particles follow geodesics (straight lines) in the flat spacetime of Special Relativity (without gravity and so without curvature), by making the hypothesis that free particles follow instead geodesics in the curved spacetime of a gravitational field. Thus, all material objects as well as light rays move along geodesics in spacetime (geodesics which are curves in four dimensions).

Let us properly mathematically define what a locally inertial frame is. To this purpose, we shall introduce the symbols

$$\eta_{ij} = \begin{cases} 1 & \text{if} \quad i = j = 0 \\ -1 & \text{if} \quad i = j = 1, 2, 3 \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Thus (η_{ij}) is the matrix of the Lorentz metric.

Given that the effects of gravity are negligible in a locally inertial frame, a locally inertial observer may set up a coordinate system in which the interval is given by

$$d\tau^2 = \eta_{ij} dx_i dx_j = (dx_0)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2$$

as taken in Special Relativity (remark that $t = x0$, $x = x_1$, $y = x_2$, $z = x_3$ and we are using geometric units, in which the speed of light takes value 1). That is, in a locally inertial frame at an event P, one can set up a coordinate system (x_i) such that

$$d\tau^2 = g_{ij} dx_i dx_j$$

and the functions g_{ij} satisfy

$$g_{ij}(P) = \eta_{ij} \tag{3.11}$$

and

$$\frac{\partial g_{ij}}{\partial x_k}(P) = 0 \tag{3.12}$$

for all $i, j, k = 0, ..., 3$. Equation 3.11 implies that the metric is Lorentzian at P exactly, whereas equation 3.12 implies that the g_{ij}'s have small rates of change in small neighbourhoods of the event P and so differ very little from η_{ij}; therefore,

$$d\tau^2 \approx \eta_{ij} dx_i dx_j$$

with the difference being able to be made arbitrarily small by restricting the size of the taken sacetime neighbourhood. A coordinate system fulfilling both equation 3.11 and 3.12 is called a locally Lorentz coordinate system at P.

Light bending

We have already seen how gravity affects a light ray by bending it. For example, if a light ray emitted from a far star passes close to the sun its path to the Earth, the gravitational field of the sun will bend it towards the sun; hence, the apparent position of the star seen from the Earth will be different (displaced outward from the sun) from the position of the star seen from somewhere else.

Einstein computed the spacetime around the sun and predicted the curvature that stars near the edge of the sun observed during a solar eclipse would be of 1,74''. His theory was supported by the measurements made by the Royal Society and the Royal Astronomical Society during the eclipse of May 29, 1919. Later, other experiments during eclipses and using radio waves emitted from the so-called quasars ("quasi-stellar" radio sources, about which are no satisfactory explanation on what they are) have come to agree with Einstein's theory. Experimental results have shown Einstein's gravitational theory to be more accurate than Newton's.

3.3 The Universal Law of Gravitation

When studying them, Sir Isaac Newton realized that the force which makes a cannon ball follow a parabollic trajectory or his famous apple fall from the tree, and he force which keeps the Moon in its orbit about the Earth, are of the same nature.

When it came to mathematically formulating his theory, Newton was influenced by the laws of planetary motion of the physicist Johannes Kepler (1571-1630), who stated that only circles or spheres could describe astronomical motions. Kepler's laws are:

A) A planet revolves about the sun in a planar elliptical orbit with the sun at one focus (whereas the other focus is a mathematical point, physically meaningless).

B) The radious vector from the sun to the planet sweeps out area at constant rate.

C) The period of the orbit of the planet is proportional to the $\frac{3}{2}$ power of the orbit's major axis.

Thus Newton stated his Law of Universal Gravitation in his *Principia* as follows: "every particle in the universe attracts every other particle in such a way that the force between them is directed along the line between them and its magnitude is proportional to the product of their masses and inversely proportional to the square of the distance between them".

Thus the force between two particles of masses M and m at distance r from each other is

$$F = G\frac{Mm}{r^2}, \tag{3.13}$$

where G is a constant of proportionality called gravitational constant. Taken as units of distance, mass and time cm, g and *seconds*, respectively, its value is, approximately, $6,67 \cdot 10^{-8} cm^3/(g \cdot sec^2)$.

Before moving on, we shall make some observations on units. As we saw in Chapter 2, time can be written in terms of ditance units by using $c \approx 2,997925 \cdot 10^{10} cm/sec$ as a conversion fator; so can be done with mass, this time by using G. Observe the units of equation 3.13:

$$G\frac{Mm}{r^2}\left(\frac{cm^3}{g \cdot sec^2}\frac{g^2}{cm^2}\right) = F\left(\frac{g \cdot cm}{sec^2}\right)$$

Unit symbols can be cancelled as in algebra, and thus, one has

$$\frac{G}{c^2}\left(\frac{cm^3}{g \cdot sec^2}\frac{sec^2}{cm^2}\right) = \frac{G}{c^2}\left(\frac{cm}{g}\right),$$

which can be used as a conversion factor! We have $\frac{G}{c^2} \approx 7,425 \cdot 10^{-29} cm/g$.

For instance, the mass of the sun in conventional units is $M_{conv} \approx 1,99 \cdot 10^{33} g$ whereas $M = M_{conv} \cdot \frac{G}{c^2} \approx 1,48 \cdot 10^5 cm$.

From now on, we shall always use geometric units. We can use the following table for transformations:

Time	$t_{conv}(sec) \cdot c(cm/sec) = t(cm)$
Velocity	$v_{conv}(cm/sec) \cdot \frac{1}{c}(sec/cm) = v(dimesionless)$
Acceleration	$a_{conv}(cm/sec^2)\frac{1}{c^2}(sec^2/cm^2) = a(cm^{-1})$
Mass	$m_{conv}(g) \cdot \frac{G}{c^2}(cm/g) = m(cm)$
Force	$F_{conv}(g \cdot cm/sec^2) \cdot \frac{G}{c^4}(sec^2/cm \cdot g) = F(dimensionless)$
Energy	$E_{conv}(cm^2 \cdot g \cdot sec^{-2}) \cdot \frac{G}{c^4}(\frac{sec^2}{cm \cdot g} = E(cm)$

Thus, equation 3.13 can be written in terms of geometric units:

$$F_{conv} = \frac{G \cdot M_{conv} \cdot m_{conv}}{r^2} \Longleftrightarrow \frac{F\cancel{c^4}}{\cancel{G}} = \cancel{G}\frac{\cancel{c^2}}{\cancel{G}}M\frac{\cancel{c^2}}{\cancel{G}}m\frac{1}{r^2} \Longleftrightarrow$$

$$F = \frac{M \cdot m}{r^2}.$$

Thus, by using geometric units, the gravitational constant and the speed of light take value 1.

3.4 Geodesics

As we saw in Section 3.2, General Relativity understands motion as natural motion alongside a 4-dimensional geodesic, and not as a response to a force. Spacetime is (viewed as) a semi-Riemannian 4-manifold, whose metric has the form

$$d\tau^2 = g_{ij} dx_i dx_j$$

in each coordinate system (x_0, x_1, x_2, x_3). We say that a vector $v = v_i \frac{\partial}{\partial x_i}$ is timelike, light-like or spacelike if the value

$$\langle v, v \rangle = g_{ij} v_i v_j$$

is positive, zero or negative, respectively.

Let us give now a definition of geodesic of further use to our purposes:

Definition 3.4.1. *A spacetime curve α is said to be a geodesic if it has a parametrization $x_i(\rho)$ satisfying*

$$\frac{d^2 x_i}{d\rho^2} = \Gamma^i_{jk} \frac{dx_j}{d\rho} \frac{dx_k}{d\rho} = 0, \tag{3.14}$$

for $i = 0, 1, 2, 3$.

The definition is independent of the choice of the coordinate system. By equation 3.14, one has that

$$\langle \alpha', \alpha' \rangle = \left(\frac{d\tau}{d\rho}\right)^2 = g_{ij}\frac{dx_i}{d\rho}\frac{dx_j}{d\rho}$$

is a constant function. Let us call its value C^2. We shall say that α is timelike, lightlike or spacelike if C^2 is positive, zero or negative, respectively.

Let α be timelike. Thus one might integrate $\frac{d\tau}{d\rho} = \pm C$ for ρ, and get $\rho = a\tau + b$ for some $a, b \in \mathbb{R}$. We should have to assume $a > 0$ so that α is future-directed (otherwise a particle following α would trvel backwards in time). Reparametrizing now equation 3.14 in terms of proper time, it holds for ρ replaced by τ.

Assuming now α to be lightlike, then τ is constant along α and we have $\langle \alpha', \alpha' \rangle = 0$; thus, proper time cannot be used as a parameter in such a case.

Finally, assuming α to be spacelike, one has that $\frac{d\tau}{d\rho}$ is imaginary and the curve might be reparametrized in terms of the so-called proper distance: $\rho = a\sigma + b$. We shall skip this case for being meaningless to our purposes.

Remark that a curve α (not necessarily a geodesic) is said to be timelike if $\langle \alpha', \alpha' \rangle > 0$ at every point. Let us now write down a couple of theorems dealing geodesics in spacetime, analogous to classical theorems on geodesics in Differential Geometry.

Theorem 3.4.2. *If α is a timelike curve which extremizes the spacetime distance between its two end points, then α is a geodesic.*

Theorem 3.4.3. *Let P be an event and let v a nonzero vector at P. Then, there exists a unique geodesic $\alpha = \alpha(\rho)$ such that $\alpha(0) = P$ and $\alpha'(0) = v$.*

Observe that if v is timelike, Theorem 3.4.3 implies that all particles fall with the same acceleration in the same gravitational field, since the spacetime path followed depends only on the initial conditions $\alpha(0)$ and $\alpha'(0)$.

Einstein postulated that the world-line of a free particle is a timelike geodesic in spacetime, and that the world-line of a photon is a lightlike geodesic. Observe that the coefficients Γ^i_{jk} in equation 3.14 express the deviation from the straight path $\frac{d^2x_i}{d\rho^2} = 0$ caused by the presence of a gravitational field. Thus, we may interpret the Christoffel symbols as the components of the gravitational field.

3.5 The field equations

As we have seen, particles follow spacetime geodesics, that is, the geometry of spacetime influences matter. But this is not the whole story, and so to complete the General Relativity Theory we shall study how matter determines geometry, i.e., we shall find a set of equations relating the metric coefficients g_{ij} to the distribution of matter. These are the so-called field equations, given by Einstein in 1916 in his paper *"Die Grundlage der allgemeine Relativitätstheorie"* ("The foundations of General Relativity").

Thus, the aim of this section is to outline the reasoning which led Einstein to the field equations. We shall consider in detail, however, only the case of the gravitational field produced by an isolated spherically symmetric mass like, for instance, the sun. Let us start by taking another look at Newton's Law of Gravitation: consider a point of mass M locaed at the origin of a 3-dimensional Cartesian system (which we shall denote by x_1, x_2, x_3). Let $X = (x_1, x_2, x_3)$ and

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2} = \|X\|.$$

Let $u_r = \frac{1}{r}X$ the unit radial vector. According to Newton's Law, the force F acting on a particle with mass m located at X is (taking geometric units only)

$$F = -\frac{Mm}{r^2}u_r.$$

Combining this equation with Newton's Second Law,

$$F = m\frac{d^2X}{dt^2},$$

one gets

$$\frac{d^2X}{dt^2} = -\frac{M}{r^2}u_r.$$

Define now the potential function $\Phi = \Phi(r)$ as

$$\Phi = -\tfrac{M}{r},\ r > 0.$$

And observe that, since

$$\frac{\partial r}{\partial x_i} = \frac{\partial}{\partial x_i}(X \cdot X)^{\frac{1}{2}} = \frac{\cancel{2}x_i}{\cancel{2}(X \cdot X)^{\frac{1}{2}}} = \frac{x_i}{r}$$

for $i = 1, 2, 3$ and $\frac{\partial \Phi}{\partial x_i} = \frac{\partial \Phi}{\partial r}\frac{\partial r}{\partial x_i}$, we have

$$-\nabla\Phi = -\Big(\frac{\partial\Phi}{\partial x_1}, \frac{\partial\Phi}{\partial x_2}, \frac{\partial\Phi}{\partial x_3}\Big) = -\frac{M}{r^2}\Big(\frac{x_1}{r}, \frac{x_2}{r}, \frac{x_3}{r}\Big)$$

$$-\frac{M}{r^2}u_r = \frac{d^2X}{dt^2}.$$

Writing it in cordinates,

$$\frac{d^2x_i}{dt^2} = -\frac{\partial\Phi}{\partial x_i} \tag{3.15}$$

for $i = 1, 2, 3$. Differentiating the obtained equation

$$\frac{\partial\Phi}{\partial x_i} = \frac{Mx_i}{r^3},$$

one gets

$$\frac{\partial^2\Phi}{\partial x_i^2} = M\Big(\frac{r^3 - 3x_i^2 r}{r^6}\Big) = \frac{M}{r^5}(r^2 - 3x_i^2).$$

Summing this last equation for $i = 1, 2, 3$, one obtains the Laplace equation

$$\Delta\Phi = \frac{\partial^2\Phi}{\partial x_1^2} + \frac{\partial^2\Phi}{\partial x_2^2} + \frac{\partial^2\Phi}{\partial x_3^2} = 0, \tag{3.16}$$

valid everywhere but at the origin (where the mass is located). Equation 3.16 is the equation for the potential function in the empty space surrounding an isolated point mass (equivalently, a spherically symmetric homogeneous mass distribution).

If we had the case of a finite number of point masses, equations 3.15 and 3.16 would still be valid in the region between masses, being the potential function Φ a sum of terms, each one contributed by one of the masses.

When considering the case of a continuous distribution of mass throughout a region of space, it can be shown that the Laplace equation (equation 3.16) is replaced by Poisson's equation

$$\Delta\Phi = 4\pi\rho, \tag{3.17}$$

where $\rho = \rho(X)$ is the density of matter at the point X. Solving equation 3.17 for the potential Φ (difficult but for some special cases), the acceleration of any free particle (and hence the gravitational force) can be obtained by equation 3.15.

In General Relativity, equation 3.15 containing the first partial derivatives of the potential funtion Φ are replaced by

$$\frac{d^2x_i}{d\tau^2} + \Gamma^i_{jk}\frac{dx_j}{d\tau}\frac{dx_k}{d\tau} = 0 \tag{3.18}$$

for $i = 0, 1, 2, 3$, which contain the first partial derivatives of the metric coefficients g_{ij}, since

$$\Gamma^i_{jk} = \frac{1}{2}g^{il}\Big(\frac{\partial g_{jl}}{\partial x_k} + \frac{\partial g_{kl}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_l}\Big). \tag{3.19}$$

Thus, in a sense, the metric coefficients play the role of gravitational potential functions in Einstein's General Relativity. The field equations relating these potential functions to the distribution of matter correspond to equation 3.17, or to equation 3.16 in the case of the empty space outside a mass distribution.

To keep the analogy, one might expect the field equations in empty space to be a system of the form

$$G = 0, \tag{3.20}$$

being G an expression involving the second partial derivatives of the "potentials" g_{ij}. One would also expect G to have some relation with curvature, in virtue of what we saw in Section 3.2. And, as happens with equation 3.18, equation 3.20 should have the same form regardless of the choice of coordinate system (that is, it has to be a tensor equation).

This observations made, it turns out that the only tensors being constructible from the metrics coefficients g_{ij}, and from their first and second derivatives, are those which are functions of the g_{ij}'s and the components R^l_{ijk} of the curvature tensor,

$$R^l_{ijk} = \frac{\partial\Gamma^l_{ik}}{\partial x_j} - \frac{\partial\Gamma^l_{ij}}{\partial x_k} + \Gamma^m_{ik}\Gamma^l_{jm} - \Gamma^m_{ij}\Gamma^l_{mk}. \tag{3.21}$$

Moreover, one possible solution of the field equations should be the flat spacetime of Special Relativity. Thus, in Lorentz coordinates (in fact, in every coordinate system), all the g_{ij}'s are constant, and hence from equation 3.19 and equation 3.21 one would have

$$R^l_{ijk} = 0 \tag{3.22}$$

for $i, j, k, l = 0, 1, 2, 3$.

Was it not for the fact that it is a too restrictive condition, equation 3.22 would be a good candidate. However, equation 3.22 implies the existence of a coordinate system with g_{ij}'s constants, which leads to the flat spacetime of Special Relativity. Then, in order to allow the existence of curved spacetime, we need a less restrictive condition. Setting $l = k$ in equation 3.22 and summing over l, one gets the so-called Ricci tensor:

$$R_{ij} = R^k_{ijk} = \frac{\partial \Gamma^k_{ik}}{\partial x_j} - \frac{\partial \Gamma^k_{ij}}{\partial x_k} + \Gamma^m_{ik}\Gamma^k_{jm} - \Gamma^m_{ij}\Gamma^k_{mk}. \tag{3.23}$$

Einstein finally chose, as his field equations for the gravitational field in empty spacetime,

$$R_{ij} = 0 \tag{3.24}$$

for $i, j = 0, 1, 2, 3$. These are second order partial differential equations to be solved for g_{ij}. Despite seeming that when he chose equations 3.24 as his field equations, Einstein was mainly led by guess, he had in fact several clues, both physical and mathematical, which shortened the list of possible candidates for being a suitable field equation. For instance, he know that G in equation 3.20 had two subscripts and was a symmetric tensor (as satisfies the Ricci tensor); he also got a clue from the fact that the field equations must reduce to equation 3.16 in the limit of a very weak field and low velocities, in the case of a continuous distribution of mass. Finally, the great agreement between theoretical predictions and observed phenomena stands for justificaton for equation 3.24.

The predictions of General Relativity must agree with those of classical physics, approximately, when studying the case of a slow moving particle in a weak static gravitational field. Let us study the metric coefficients in such a case.

Let (t, x_1, x_2, x_3) be a locally Lorentz coordinate system in a neighbourhood of an event on a moving particle's world-line. Provided that the particle's speed is small enough, i.e., $|dx_i/dt| \ll 1$ for $i = 1, 2, 3$), $dx_i/d\tau$ is negligible compared to $dt/d\tau$ and we can write the geodesic equations as

$$\frac{d^2 x_k}{d\tau^2} = -\Gamma^k_{ij}\frac{dx_i}{d\tau}\frac{dx_j}{d\tau} \approx -\Gamma^k_{00}\left(\frac{dt}{d\tau}\right)^2 \tag{3.25}$$

We have made the assumption of an static gravitational field, that is, time derivatives of the metric coefficient vanish, and then the computation of the Christoffel symbols (by using equation 3.19) gives

$$\Gamma^k_{00} = -\frac{1}{2} g^{mk} \frac{\partial g_{00}}{\partial x_m}.$$

Since the field is supposed to be weak, we may choose Lorentz coordinates satisfying

$$g_{ij} = \eta_{ij} + h_{ij}$$

where

$$(\eta_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and h_{ij} are small compared to 1. Then the inverse matrix of (g_{ij}) has entries of the form $g^{ij} = \eta^{ij} + k^{ij}$, where $\eta^{ij} = \eta_{ij}$ (since the matrix (η_{ij}) is its own inverse) and k^{ij} are small compared to 1. According to this, we have

$$\Gamma^k_{00} \approx -\frac{1}{2}\eta^{mk}\frac{\partial h_{00}}{\partial x_m}, \tag{3.26}$$

to first order in the h_{ij}'s. Substitute equation 3.26 into equation 3.25, and we approximately obtain

$$\frac{d^2x_k}{d\tau^2} = \frac{1}{2}\left(\frac{dt}{d\tau}\right)^2 \eta^{mk}\frac{\partial h_{00}}{\partial x_m}. \tag{3.27}$$

For $k = 0$, this yields to $\frac{d^2t}{d\tau^2} = 0$ or constant. Hence,

$$\frac{dx_i}{dt} = \frac{\frac{dx_i}{d\tau}}{\frac{dt}{d\tau}}, \frac{d^2x_i}{dt^2} = \frac{\frac{d^2x_i}{d\tau^2}}{\left(\frac{dt}{d\tau}\right)^2}$$

and we get, from equation 3.27 with $k = i \neq 0$,

$$\frac{d^2x_i}{dt^2} = \frac{1}{2}\eta^{mi}\frac{\partial h_{00}}{\partial x_m} = -\frac{1}{2}\frac{\partial h_{00}}{\partial x_i}, i = 1, 2, 3.$$

Comparing this latter equation with the Newtonian

$$\frac{d^2x_i}{dt^2} = -\frac{\partial \Phi}{\partial x_i},$$

we can see that both theories agree given that

$$\frac{\partial h_{00}}{\partial x_i} = 2\frac{\partial \Phi}{\partial x_i}, i = 1, 2, 3$$

or

$$h_{00} = 2\Phi + C, \text{ for some } C \text{ constant.}$$

Thus, the Newtonian potential function Φ is determined up to a constant. Since Φ has to vanish at infinity and so has h_{ij}, then C must be 0, and hence $h_{00} = 2\Phi$ and

$$g_{00} = 1 + 2\Phi. \tag{3.28}$$

Briefly speaking, we have seen that the low speed, weak field agreement between Newtonian theory and General Relativity needs equation 3.27 to hold in locally Lorentz coordinates.

3.6 The Schwarzschild solution

In this section, we shall solve the Einstein's field equations

$$R_{ij} = 0, \tag{3.29}$$

$i, j = 0, 1, 2, 3$, for the gravitational field outside an isolated spherically symmetric mass M, assumed to be at rest at the origin of the coordinate system we are within.

Assuming that the mass is a typical star, the departure of flatness is small and we may assume a coordinate system differing only slightly from the Special Relativity form

$$d\tau^2 = dt^2 - dx_1^2 - dx_2^2 - dx_3^2. \tag{3.30}$$

For convenience, we shall take spherical coordinates by using the transformation

$$\begin{cases} x_1 = \rho \sin\Phi \cos\theta \\ x_2 = \rho \sin\Phi \sin\theta \\ x_3 = \rho \cos\Phi. \end{cases} \tag{3.31}$$

Computing dx_1, dx_2 and dx_3 from equation 3.31 and substituting into equation 3.30, we get

$$d\tau^2 = dt^2 - d\rho^2 - \rho^2 d\Phi^2 - \rho^2 \sin^2\Phi d\theta^2. \tag{3.32}$$

This is still the metric of flat spacetime. We shall modify equation 3.32 in order to produce the metric of curved spacetime surrounding mass M. The metric we are looking for should be static and spherically symmetric, so that its expression is of the form

$$d\tau^2 = U(\rho)dt^2 - V(\rho)d\rho^2 - W(\rho)(\rho^2 d\Phi^2 + \rho^2 \sin^2\Phi d\theta^2), \tag{3.33}$$

being U, V and W functions of ρ and differ slightly from unity. Introducing $r = \rho W(\rho)^{\frac{1}{2}}$, we can write the metric as

$$d\tau^2 = A(r)dt^2 - B(r)dr^2 - r^2 d\Phi^2 - r^2 \sin^2\Phi d\theta^2. \tag{3.34}$$

This form of equation 3.34 contains two unknown functions to be determined. Observe that since functions U, V, W, A and B differ only slightly from unity, both r in equation 3.34 and ρ in equation 3.33 are approximate counterparts of radial distance ρ in equation 3.32. But, since we consider curved spacetime, neiter ρ nor r can have exactly the properties of distance from the origin in Euclidian space. Hence, we may choose r rather than ρ as the analog of radial distance in Newton's theory.

Define now the functions $m = m(r)$ and $n = n(r)$ by

$$e^{2m} = A, e^{2n} = B.$$

Thus, we have

$$d\tau^2 = e^{2m}dt^2 - e^{2n}dr^2 - r^2 d\Phi^2 - r^2 \sin^2\Phi d\theta^2. \tag{3.35}$$

Comparing with the general expression

$$d\tau^2 = g_{ij}dx_i dx_j$$

we have

$$(g_{ij}) = \begin{pmatrix} e^{2m} & 0 & 0 & 0 \\ 0 & -e^{2n} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \Phi \end{pmatrix}, \tag{3.36}$$

whose determinant is $g = -e^{2m+2n} r^4 \sin^2 \Phi$.

We shall determine functions m and n. To that purpose, we shall compute the Christoffel symbols (equation 3.19) in terms of m and n. Observe that, since $g_{ij} = 0$ for $i \neq j$, we have $g^{ii} = \frac{1}{g_{ii}}$ for all i, and $g^{ij} = 0$ for $i \neq j$. Equation 3.19 the reduces to

$$\Gamma^k_{ij} = \frac{1}{2g_{kk}} \Big(\frac{\partial g_{ik}}{\partial x_j} + \frac{\partial g_{jk}}{\partial x_i} - \frac{\partial g_{ij}}{\partial x_k} \Big). \tag{3.37}$$

There are now three cases to consider:

Case 1: For $k = j$,

$$\Gamma^j_{ij} = \frac{1}{2} \frac{\partial}{\partial x_i} \ln |g_{jj}|. \tag{3.38}$$

Case 2: For $i = j \neq k$,

$$\Gamma^k_{ii} = -\frac{1}{2g_{kk}} \frac{\partial g_{ii}}{\partial x_k}. \tag{3.39}$$

Case 3: For i, j, k distinct,

$$\Gamma^k_{ij} = 0. \tag{3.40}$$

Thus, by computation one gets the following Christoffel symbols:

$$\begin{array}{ll} \Gamma^0_{10} = \Gamma^0_{01} = m', & \Gamma^1_{00} = m' e^{2m} - 2n \\ \Gamma^1_{11} = n', & \Gamma^1_{22} = -re^{-2n} \\ \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}, & \Gamma^1_{33} = -re^{-2n} \sin^2 \Phi \\ \Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r}, & \Gamma^3_{23} = \Gamma^3_{32} = \cot \Phi, \\ \Gamma^2_{33} = -\sin \Phi \cos \Phi. & \end{array} \tag{3.41}$$

These expressions, together with the formula

$$\ln |g|^{\frac{1}{2}} = \frac{1}{2} \ln(e^{2m+2n} r^4 \sin^2 \Phi) = m + n + 2 \ln r + \ln \sin \Phi$$

are to be substituted into Einstein's field equations:

$$R_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} \ln |g|^{\frac{1}{2}} - \frac{\partial \Gamma^k_{ij}}{\partial x_k} + \Gamma^m_{ik} \Gamma^k_{jm} - \Gamma^m_{ij} \frac{\partial}{\partial x_m} \ln |g|^{\frac{1}{2}} = 0. \tag{3.42}$$

For instance, for $i = j = 0$, we have

$$R_{00} = \frac{\partial^2}{\partial t^2} \ln |g|^{\frac{1}{2}} - \frac{\partial \Gamma^k_{00}}{\partial x_k} + \Gamma^m_{k0} \Gamma^k_{m0} - \Gamma^m_{00} \frac{\partial}{\partial x_m} \ln |g|^{\frac{1}{2}} =$$

$$= 0 - \frac{\partial}{\partial r} (m' e^{2m-2n}) + (\Gamma^0_{10} \Gamma^1_{00} + \Gamma^1_{00} \Gamma^0_{10}) - \Gamma^1_{00} \frac{\partial}{\partial r} \ln |g|^{\frac{1}{2}} =$$

$$= \big(- m'' - m'(2m' - 2n') \big) e^{2m-2n} + 2(m')^2 e^{2m-2n} - m' e^{2m-2n} (m' + n' + \frac{2}{r}) =$$

$$= \big(- m'' + m'n' - (m')^2 - \frac{2m'}{r} \big) = 0.$$

With similar calculations we get

$$\frac{R_{00}}{e^{2m-2n}} = -m'' + m'n' + (m')^2 - \frac{2m'}{r} \tag{3.43}$$

$$R_{11} = m'' - m'n' + (m')^2 - \frac{2n'}{r} = 0 \tag{3.44}$$

$$R_{22} = e^{-2n}(1 + rm' - rn') - 1 = 0 \tag{3.45}$$

$$R_{33} = R_{22} \sin^2 \Phi = 0 \tag{3.46}$$

and other $R_{ij} = 0$.

Adding equations 3.43 and 3.44, we get $m' + n' = 0$, and so $m + n = b$ a constant. However, m and n must vanish as $r \longrightarrow \infty$, since the metric given in equation 3.35 must approach the Lorentz metric at great distances from mass M. This implies $b = 0$ and $n = -m$. Then equation 3.45 becomes

$$1 = (1 + 2rm')e^{2m} = (re^{2m})'.$$

Hence,

$$re^{2m} = r + C$$

for some constant C, or

$$g_{00} = e^{2m} = 1 + \frac{C}{r}.$$

But we saw in the previous section (Section 3.5) that far from mass, where the field is weak and Newtonian theory holds, we must have $g_{00} = 1 - \frac{2M}{r}$. Therefore we have $C = -2M$. Finally, we arrive to the solution

$$d\tau^2 = \Big(1 - \frac{2M}{r}\Big)dt^2 - \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 - r^2 d\Phi^2 - r^2 \sin^2 \Phi d\theta^2. \tag{3.47}$$

This equation was obtained by the German physicist Karl Schwarzschild (1873-1916) in 1916, only a few months after the publication of the General Relativity theory of Einstein.

3.7 Orbits in General Relativity

Our goal in this section is to calculate the orbit of a planet under the General Relativity theory. We begin with the Schwarzschild metric

$$d\tau^2 = \Big(1 - \frac{2M}{r}\Big)dt^2 - \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 - r^2 d\Phi^2 - r^2 \sin^2 \Phi d\theta^2 \tag{3.48}$$

for the spacetime geometry outside a spherically symmetric mass M. We shall consider the sun as a point mass and neglect the effects on the orbits by other planets. According to

Einstein's General Relativity theory, as we have seen, a planet follows a timelike geodesic in this geometry:

$$\frac{d^2x_k}{d\tau^2} + \Gamma^k_{ij}\frac{dx_i}{d\tau}\frac{dx_j}{d\tau} = 0 \tag{3.49}$$

for $k = 0, 1, 2, 3$, being $x_0 = t$, $x_1 = r$, $x_2 = \Phi$ and $x_3 = \theta$. The Christoffel symbols are the ones obtained from equation 3.41, where primes denote differentiation with respect to r, $n = -m$, and

$$e^{2m} = 1 - \frac{2M}{r}.$$

We choose first $k = 2$, and obtain

$$\frac{d^2\Phi}{d\tau^2} + \frac{2}{r}\frac{dr}{d\tau}\frac{d\Phi}{d\tau} - \sin\Phi\cos\Phi\left(\frac{d\Phi}{d\tau}\right)^2 = 0.$$

Performing a spacial rotation if necessary we may assume that, whenever $\tau = 0$, $\Phi = \pi/2$ and $d\Phi/d\tau = 0$, that is, the planet is initially in the plane $\Phi = \pi/2$. Due to the spherical symmetry of the field, the planet must remain in this plane: otherwise it would be preference to one of the plane's half-spaces over the other. We therefore take $\Phi = \pi/2$, which, substituted into the remaining geodesics equations, we get, as in orbital equations

$$\frac{d^2t}{d\tau^2} + 2m'\frac{dr}{d\tau}\frac{dt}{d\tau} = 0, \tag{3.50}$$

$$\frac{d^2r}{d\tau^2} + m'e^{2m-2n}\left(\frac{dt}{d\tau}\right)^2 + n'\left(\frac{dr}{d\tau}\right)^2 - re^{-2m}\left(\frac{d\theta}{d\tau}\right)^2 = 0, \tag{3.51}$$

$$\frac{d^2\theta}{d\tau^2} + \frac{2}{r}\frac{dr}{d\tau}\frac{d\theta}{d\tau} = 0. \tag{3.52}$$

For convenience, define

$$\gamma = 1 - \frac{2M}{r} = e^{2m}.$$

Since $m'\frac{dr}{d\tau} = \frac{dm}{dr}\frac{dr}{d\tau} = \frac{dm}{d\tau}$, we divide equation 3.50 by $dt/d\tau$ and set

$$\frac{d}{d\tau}\left(\ln\frac{dt}{d\tau}\right) = -2\frac{dm}{d\tau}.$$

Integrating and then exponentiating we set

$$\frac{dt}{d\tau} = be^{-2m} = \frac{b}{\gamma} \tag{3.53}$$

for some constant $b > 0$.

Similarly, we can integrate equation 3.52 and get

$$r^2\frac{d\theta}{d\tau} = h, \tag{3.54}$$

with $h > 0$ a constant.

Instead of attempt the integration of equation 3.51, we can get an alternative equation from the metric. From equation 3.49 with $\Phi = \pi/2$, one gets

$$1 = \gamma\left(\frac{dt}{d\tau}\right)^2 - \gamma^{-1}\left(\frac{dr}{d\tau}\right)^2 - r^2\left(\frac{d\theta}{d\tau}\right)^2 = \gamma\left(\frac{b}{\gamma}\right)^2 - \gamma^{-1}\left(\frac{dr}{d\theta}\frac{h}{r^2}\right)^2 - r^2\left(\frac{h}{r^2}\right)^2 \tag{3.55}$$

by using equations 3.52 and 3.54. Now, if we multiply by $\gamma = 1 - \frac{2M}{r}$ and rearrange, we get

$$\left(\frac{h}{r^2}\frac{dr}{d\theta}\right)^2 + \frac{h^2}{r^2} = b^2 - 1 + \frac{2M}{r} + \frac{2M}{r}\frac{h^2}{r^2}.$$

Let $u = 1/r$, which implies that $\frac{du}{d\theta} = -\frac{1}{r^2}\frac{dr}{d\theta}$. Dividing the previous equation by h^2, we get

$$\left(\frac{du}{d\theta}\right)^2 + u^2 = \frac{b^2-1}{h^2} + \frac{2Mu}{h^2} + 2Mu^3.$$

Finally, we differentiate with respect to θ and divide by $2\frac{du}{d\theta}$, and we get

$$\frac{d^2u}{d\theta^2} + u = \frac{M}{h^2} + 3Mu^2, \tag{3.56}$$

where $u = \frac{1}{r}$ and $h = r^2\frac{d\theta}{d\tau}$ a constant. This equation 3.56 is similar to the Newtonian equation

$$\frac{d^2u}{d\theta^2} + u = \frac{M}{h^2}$$

but for the relativistic term $3Mu^2$, usually quite small compared to $\frac{M}{h^2}$. As an example, for the planet Mercury the ratio of these terms is $7,4 \cdot 10^{-8}$ approximately. Therefore, the orbit should differ slightly from what Newton's theory predicted; the discrepancy only appears after many revolutions. As a first approximation to a solution to equation 3.56, we get the function $u_1 = u_1(\theta)$ given by

$$u_1 = \frac{M}{h^2}(1 + e\cos\theta). \tag{3.57}$$

We can substitute u_1 into the relativistic term $3Mu^2$, and thus we get a better approximation:

$$\frac{d^2u}{d\theta^2} + u = \frac{M}{h^2} + \frac{3M^3}{h^4}(1 + 2e\cos\theta + e^2\cos^2\theta)$$

or

$$\frac{d^2u}{d\theta^2} + u = \frac{M}{h^2} + \frac{3M^3}{h^4} + \frac{6M^3e}{h^4}\cos\theta + \frac{3M^3e^2}{2h^4}\cos(2\theta), \tag{3.58}$$

the last term obtained from the equation $\cos^2\theta = \frac{1}{2}\big(1 + \cos(2\theta)\big)$.

Let us perform now a Lemma which will help solving equation 3.58:

Lemma 3.7.1. *Let A be a real number. Then*

$$u = A \textit{ is a solution of } \frac{d^2u}{d\theta^2} + u = A,$$

$$u = \tfrac{1}{2}\theta\sin\theta \textit{ is a solution of } \frac{d^2u}{d\theta^2} + u = A\cos\theta,$$

$$u = -\tfrac{A}{3}\cos(2\theta) \textit{ is a solution of } \frac{d^2u}{d\theta^2} + u = A\cos(2\theta).$$

Apply this Lemma to the four additional terms of equation 3.58, that is, for each term T, solve $\frac{d^2u}{d\theta^2} + u = T$ and add the sum of the solutions to u_1. After factoring out M/h^2 and collecting terms, we get

$$u = \frac{M}{h^2}\Big(1 + \frac{3M^2}{h^2}(1 + \frac{e^2}{2}) + e\cos\theta + \frac{3M^2}{h^2}e\theta\sin\theta - \frac{M^2e^2}{2h^2}\cos(2\theta)\Big) \tag{3.59}$$

where e is a constant.

We shall determine how the extra terms in brackets modify the Newtonian elliptical orbit equation

$$u = \frac{M}{h^2}(1 + e\cos\theta).$$

The term $3M^2(1 + \frac{e^2}{2})/h^2$ which we have added is small compared to unity (for example, $8 \cdot 10^{-8}$ for Mercury), the effect of this term is negligible on the planet's orbit. In fact, if we where to set

$$\alpha = 1 + \frac{3M^2}{h^2}\Big(1 + \frac{e^2}{2}\Big)$$

and define $e' = e/\alpha$, we can rewrite the first terms of equation 3.59 as

$$\frac{M\alpha}{h^2}(1 + e'\cos\theta),$$

a function whose form is similar to equation 3.57. Since $\alpha \approx 1$, the effect of the added constant term is negligible on the orbit.

The final term involving $\cos(2\theta)$ is also very small and produces a negligible variation on the orbit. The crucial term is the one involving $\theta\sin\theta$, since θ appears outside the trigonometrical operator instead of being the argument of a periodic function, and therefore it has an accumulative effect over many revolutions, as θ increases. This effect is the perihelial advance. We can thus write (to a close approximation) the relativistic orbital equation as

$$u \approx \frac{M}{h^2}\big(1 + e\cos\theta + \frac{3M^2e}{h^2}\theta\sin\theta\big).$$

By using the approximations given by $\frac{M^2}{h^2} \approx 10^{-8}$,

$$\cos\left(\frac{3M^2\theta}{h^2}\right) \approx 1, \sin\left(\frac{3M^2\theta}{h^2}\right) \approx \frac{3M^2\theta}{h^2},$$

we get the orbital equation

$$u \approx \frac{M}{h^2}\Big(1 + e\cos\big(\theta - \frac{3M^2}{h^2}\theta\big)\Big). \tag{3.60}$$

When $u = 1/r$ reaches its maximum (when r reaches its minimum), the perihelion occurs. Taking equation 3.60, we see that two successive perihelia occur at $\theta = 0$ and at

$$\theta = \frac{2\pi}{1 - \frac{3M^2}{h^2}} \approx 2\pi\Big(1 + \frac{3M^2}{h^2}\Big).$$

Thus the direction of the perihelion advances $\Delta\theta \approx \frac{6\pi M^2}{h^2}$ radians per revolution, and the precession is in the direction of the planet's revolution.

Define by n the number of revolutions of the planet per (Earth's) century. Hence, the amount of precession or perihelion advance in 100 years, denoted by $\Delta_{\theta cent}$, is

$$\Delta_{\theta cent} = n\Delta_\theta = \frac{6\pi M^2 n}{h^2} = \frac{6\pi M n}{a(1-e^2)}$$

radians, since $h^2/M = ed = a(1-e^2)$.

3.8 The bending of light

As we ave seen in Section 3.2, according to Einstein, a ray of light follows a geodesic a matter does, but now τ is no suitable parameter since the lapse of proper time along a lightlike curve is zero. The equations for a light ray have the form

$$\tfrac{d^2x_k}{d\rho^2} + \Gamma^k_{ij}\tfrac{dx_i}{d\rho}\tfrac{dx_j}{d\rho} = 0,\ k = 0,1,2,3,$$

where $d\tau/d\rho = 0$. We orient the coordinate system we are in so that the path lies in the plane $\Phi = \pi/2$, and the equations are of the form of equations 3.50, 3.51 and 3.52, but with τ instead of ρ. The derivatio which follows is like the one done for planetary orbits in Section 3.7, but for the left hand side of equation 3.55 being zero instead of one, since we replace $dt/d\tau = 1$ by $dt/d\tau = 0$. The calculations yield

$$\frac{d^2u}{d\theta^2} + u = 3Mu^2, \tag{3.61}$$

where $u = 1/r$.

Let R be the minimum distance from mass M to the path of the light ray. We shall assume the coordinate system to be oriented so that the point of closest approach is on the ray $\theta = 0$. When there is no mass ($M = 0$), the solution to equation 3.61 is

$$u = \frac{1}{r} = \frac{1}{R}\cos\theta \tag{3.62}$$

or

$$x = r\cos\theta = R$$

a straight line. Thus, the term $3Mu^2$ represents the deviation from the classical rectilinear path, and in usual speeds (small compared to the speed of light) is small compared to R. For instance, consider a light ray grazing the sun: the mass of the sun is $M_\odot \approx 1,48 \cdot 10^5$cm, the radious of the sun is $R_\odot \approx 6,69 \cdot 10^{10}$cm, and so $3Mu^2 \approx 9,2 \cdot 10^{-17}$. Hence, we expect the solution to equation 3.61 to be

$$u \approx \frac{1}{R}\cos\theta.$$

If we substitute this last equation back into the relativisic term $3Mu^2$ in equation 3.61, we get

$$\frac{d^2u}{d\theta^2} + u \approx \frac{3M}{R^2}\cos^2\theta = \frac{3M}{2R^2}\big(1+\cos(2\theta)\big).$$

By Lemma 3.7.1, a particular solution of the last equation is

$$u_1 = sin\frac{3M}{2R^2}\Big(1 - \frac{1}{3}\cos(2\theta)\Big).$$

Using the identity $\cos(2\theta) = 2\cos^2(\theta) - 1$, we can rewrite this as

$$u_1 = \frac{M}{R^2}(2 - \cos^2\theta).$$

Now, adding u_1 to the homogeneous solution 3.62, we have that the solution to equation 3.61 is, closely approximated,

$$u = \frac{1}{r} = \frac{1}{R}\cos\theta + \frac{M}{R^2}(2 - \cos^2\theta). \tag{3.63}$$

Observe that when $r \longrightarrow \infty$, far from M redarless of the direction, spacetime is practically flat and the path of the light ray becomes almost straight. Thus the path has its asymptotic lines, and the angle $\Delta\theta$ between them is the amount of light deflection which general relativity predicts. As $r \longrightarrow \infty$, θ approaches $\pm(\pi/2 + \Delta\theta/2)$, and since $\Delta\theta$ is very small, the $\cos^2\theta$ term in equation 3.63 is negligible in comparison to the other terms. Doing $r \longrightarrow \infty$ in equation 3.63, we have

$$0 = \frac{1}{R}\cos\Big(\frac{\pi}{2} + \frac{\Delta\theta}{2}\Big) + \frac{2M}{R^2}.$$

Hence,

$$\frac{2M}{R} = -\cos\Big(\frac{\pi}{2} + \frac{\Delta\theta}{2}\Big) = \sin\frac{\Delta\theta}{2} \approx \frac{\Delta\theta}{2}$$

or

$$\Delta\theta \approx \frac{4M}{R}.$$

Taking the radious and mass of the Earth, we get $\Delta\theta \approx 8,51 \cdot 10^{-6}$ rad$\approx 1,75''$.

In fact, Newtonian mechanics also predicted a bending of light if one takes the assumption that light protons are influenced by gravity as are material particles. Thus, the path of a proton is a branch of an hyperbola and is given by

$$u = \frac{M}{h^2}(1 + e\cos\theta)$$

with $e > 1$:

$$\frac{1}{r} = \frac{M}{h^2}(1 + e\cos\theta), \tag{3.64}$$

where

$$h = r^2\frac{d\theta}{dt}. \tag{3.65}$$

Consider now the case of a ray just grazing the sun at $\theta = 0$. At the instant of closest approach, the radial component of the speed of photon is zero, and the speed, using the Newton equation

$$(ru_r)' = \frac{dr}{dt}u_r + r\frac{d\theta}{dt}u_\theta,$$

where $u_r = (\cos\theta, \sin\theta)$ and $u_\theta = (-\sin\theta, \cos\theta)$, the speed is $r\frac{d\theta}{dt} = 1$ (the speed of light). Equation 3.65 implies $h = r\big(r\frac{d\theta}{dt}\big) = R_\odot$. If we substitute $\theta = 0$, $M = M_\odot$, $r = R_\odot = h$ into equation 3.64 and solve for e, we get

$$e = \frac{R_\odot}{M_\odot} - 1 \approx \frac{R_\odot}{M_\odot} \approx 4{,}7 \cdot 10^5.$$

Let $\Delta\theta_N$ be the (accute) angle between the light path's asymptotes; thus, as $r \longrightarrow \infty$ in equation 3.64, $\theta \longrightarrow \pm\big(\frac{\pi}{2} + \frac{\Delta\theta_N}{2}\big)$. Hence we get, computing the limit,

$$0 = \frac{M}{R^2}\Big(1 + e\cos\big(\frac{\pi}{2} + \frac{\Delta\theta_N}{2}\big)\Big)$$

or

$$\frac{1}{e} = -\cos\big(\frac{\pi}{2} + \frac{\Delta\theta_N}{2}\big) = \sin\frac{\Delta\theta_N}{2} \approx \frac{\Delta\theta_N}{2}$$

since, the fact that e is large implies that $\Delta\theta$ is small. Hence, the Newtonian deflection is

$$\Delta\theta_N \approx \frac{2}{e} = \frac{2M_\odot}{R_\odot},$$

which is exactly half of what predicted General Relativity. Measurements of apparent star positions during solar eclipses since 1919 and recent observations of radio sources confirmed Einstein's theory over Newton's.

Bibliography

[1] Girbau, J., *Geometria diferencial i relativitat*, Manuals de la UAB (1993), Servei de Publicacions de la Universitat Autònoma de Barcelona, Bellaterra, pp. xii+375.

[2] Faber, R. L., *Differential Geometry and Relativity Theory: An Introduction*, Monographs and Textbooks in Pure and Applied Mathematics 76 (1983), Marcel Dekker, Inc., New York, pp. x+255.

[3] Einstein, A., *The Special and the General Relativity Theory. A Popular Exposition*, Methuen& Co. LTD. (1920), London, pp. xi+132.

[4] Hooft, G., *Introduction to General Relativity*, Utrecht, pp. 68.