Geometric Mechanics

Waldyr Muniz Oliva

Professor Emérito do Instituto de Matemática e Estatística
da Universidade de São Paulo-Brasil

Professor Catedrático Convidado do Instituto Superior Técnico
da Universidade Técnica de Lisboa- Portugal

Lisbon, June 2000

Contents

Preface

These notes were prepared as a guide for the lectures of one semester course on Geometric Mechanics. They were written inside the level of a master course. I started some years ago teaching them at the "Instituto de Matemática e Estatística" of the "Universidade de São Paulo", and, more recently, at the "Instituto Superior Técnico" of the "Universidade Técnica de Lisboa".

The spectrum of the participants of such a course ranges usually from young Master students to Phd students. So, it is always very difficult to decide how to organize all the material to be taught. I decided that the expositions should be self contained, so some material that we expect to be interesting for someone, results, often, tedious for others and frequently unreachable for a few ones.

In any case, for young researchers interested in differential geometry and or dynamical systems it is basic and fundamental to see the foundations and the development of a classical subject like Mechanics.

Geometric Mechanics in this monograph means Mechanics on a pseudo-riemannian manifold and the main goal is the study of some mechanical models and concepts, with emphasis on the intrinsic and geometric aspects arising in classical problems. Topics like calculus of variation and the theories of symplectic, hamiltonian and poissonian structures including reduction by symmetries, integrability etc., also related with most of the considered models, were avoided because they already appear in many modern books on the subject and are also contained in other courses of the majority of Master and PhD programs of many Institutions (see [AM],[Ma],[MR]). In Chapter 1 I started with Newtonian mechanics where it is described the galilean space-time structure and Newton equations; the chapter ends with some critical remarks in order to motivate the introduction of special and general relativity. Chapter 2 and 3 include the fundamental calculus on a differential manifold with a brief introduction on vector fields, differential forms and tensor fields. Chapter 4 starts with the concept of affine connections and special attention is given to the notion of curvature; E. Cartan structural equations of a connection are also derived in chapter 4. Chapter 5 deals with Mechanical systems on a riemannian manifold including some classical examples like the dynamics of rigid and pseudo-rigid bodies and notions derived from the dissipation in mechanical systems like Morse-Smale systems and structural stability; also some generic properties are discussed. Chapter 6 considers mechanical systems with non-holonomic constraints and describes d'Alembertian geometric mechanics including conservative and dissipative situations. In Chapter 7 one talks about hyperbolicity and Anosov systems arising in mechanics and it is also mentioned the so-called non-holonomic mechanics of vaconomic type. The last two chapters deal mainly with some topics on special and general relativity.

I wish to thank a number of colleagues from several different Institutions as well as Master and PhD students from São Paulo and Lisbon who motivated and helped us with comments and suggestions when I was writing this text. Among them we mention Jack Hale, Ivan Kupka, Giorgio Fusco, Paulo Cordaro, Carlos Rocha, Luis Magalhães, Esmeralda Dias, Zaqueu Coelho, Helena Castro, Marcelo Kobayashi, Sónia Garcia, Diogo Gomes and Manuel Natário. Thanks are also due to Ana Bordalo for her fine typing of this work.

Lisbon, June 2000.

Chapter 1

Newtonian mechanics

1.1 Galilean space-time structure and Newton equations

Let A be an affine space associated to a finite dimensional vector space V, that is, it is defined a map

$$A \times V \to A$$

called **sum** $(x+v) \in A$ of a point $x \in A$ with a vector $v \in V$, and the following axioms hold:

a_1) $x + 0 = x$, for all $x \in A$ and 0 the zero vector in V.

a_2) $x + (v_1 + v_2) = (x + v_1) + v_2$, for all $x \in A$ and $v_1, v_2 \in V$.

a_3) Given $x, y \in A$, there is just one vector $u \in V$ such that $x + u = y$; u is denoted by $(y - x)$.

Example 1.1 Any finite dimensional vector space can be considered as an affine space associated to itself. Note that the cartesian product $A_1 \times A_2$ of two affine spaces A_1 and A_2 is an affine space.

If the vector space V is euclidean (in V is defined an inner product $\langle , \rangle$), we say that any affine space associated to V is **euclidean**. In this last case one can talk about the distance between two points $x, y \in A$, by setting:

$$\rho(x,y) \stackrel{def}{=} \| x - y \| = \sqrt{\langle (x-y), (x-y) \rangle}.$$

The presentation of this section 1.1, follows closely [Ar] *"Mathematical Methods of Classical Mechanics"* by V.I. Arnold, Springer-Verlag, 1980, p.3 to 11.

A **galilean space-time structure** is a triple $(A^4, \tau, (,))$ where A^4 is a dimension four affine space associated to a vector space V^4, τ is a non-zero linear form

$$\tau : V^4 \to \mathbf{R}$$

and $(,)$ is an inner product defined on the three dimensional kernel $S = \tau^{-1}(0)$ of τ. The elements in A^4 are the **world points or events**, τ is the **absolute time** and $\tau(x-y)$ is the **time interval from event x to event y**. When $\tau(x-y) = 0$, x and y are said to be **simultaneous events** and then $(x-y) \in S$.

The set $S_x = x + S$ of all events simultaneous to x is a three dimensional euclidean affine space associated to S; in fact S is an euclidean vector space with the given inner product $(,)$. Then it makes sense to talk about the distance between two simultaneous events but does **not** makes sense to talk about the distance between two events with a positive time interval.

Let A_1, A_2 be two affine spaces associated to vector spaces V_1, V_2 respectively. An **affine transformation (affine isomorphism)** between A_1 and A_2 is a bijection $T : A_1 \to A_2$ such that there exists a bijective linear map $T^* : V_1 \to V_2$ and $T(x) - T(y) = T^*(x-y)$ for all $x, y \in A_1$. When $A_1 = A_2 = A$ and $V_1 = V_2 = V$ the affine transformations form a group called the **affine group** of A.

One defines the **galilean group** of a galilean structure $(A^4, \tau, (,))$ as the subgroup G_{A^4} of the affine group of A^4 whose elements preserve the time intervals of any pair of events and also preserve the distances between two simultaneous events.

So $T \in G_{A^4}$ means that T is an affine transformation of A^4 and, moreover:

$G_1)$ $\tau(x-y) = \tau(T(x) - T(y))$ for any $x, y \in A^4$;

$G_2)$ $x_1, x_2 \in A^4$ and $\tau(x_1 - x_2) = 0$ imply $\| x_1 - x_2 \| = \| T(x_1) - T(x_2) \|$.

It is clear that conditions $G_1)$ and $G_2)$ above are equivalent to the following:

$\bar{G}_1)$ $\tau = \tau \circ T^*$ (that, in particular, shows that T^* leaves invariant the subspace $S = \tau^{-1}(0)$).

$\bar{G}_2)$ The restriction of T^* to S is an orthogonal transformation on S, that is $(T^* v, T^* u) = (v, u)$ for all $v, u \in S$.

Example 2.1 Let us consider $\mathbf{R} \times \mathbf{R}^3$ as an affine space, $\tau : \mathbf{R} \times \mathbf{R}^3 \to \mathbf{R}$ be the projection $\tau(t, x) = t$ for all $(t, x) \in \mathbf{R} \times \mathbf{R}^3$, and $S = \tau^{-1}(0) = \{(0, x) | x \in \mathbf{R}^3)$

with the inner product $(,)$ induced by $\mathbf{R}^3$. The galilean space-time structure $(\mathbf{R} \times \mathbf{R}^3, \tau, (,))$is the so called **galilean coordinate space** and its galilean group $G_{\mathbf{R}\times\mathbf{R}^3}$ will be denoted by G.

Exercise 1.1 The following affine transformations of $\mathbf{R} \times \mathbf{R}^3$ belong to G:

g_1) **Uniform motion with velocity** v:

$$g_1((t,x)) = (t, x+tv), \quad (t,x) \in \mathbf{R} \times \mathbf{R}^3;$$

g_2) **Translation of the origin**$(0,0)$ **to** $(s,\omega) \in \mathbf{R} \times \mathbf{R}^3$:

$$g_2((t,x)) = (t+s, x+\omega), \quad (t,x) \in \mathbf{R} \times \mathbf{R}^3;$$

g_3) **Rotation** R **of the coordinate axes:**

$$g_3((t,x)) = (t, Rx), \quad (t,x) \in \mathbf{R} \times \mathbf{R}^3 \quad \text{and } R$$

is an orthogonal transformation of $\mathbf{R}^3$ (proper ($det \ \ R = 1$) or not).

Exercise 2.1 Show that any transformation $g \in G$ can be written in a unique way as a composition $g = g_1 \circ g_2 \circ g_3$; specify $g_1{}^*, g_2{}^*, g_3{}^*$.

Remark that the group G has dimension 10 and the affine group of $\mathbf{R} \times \mathbf{R}^3$ has dimension 20 (here **dimension** means the number of real parameters that one needs to determine a generic element of the group).

Two galilean space-time structures $(A_1, \tau_1, (,)_1)$ and $(A_2, \tau_2, (,)_2)$ are said to be **isomorphic** if there exists an affine isomorphism $T : A_1 \longrightarrow A_2$ such that

i) $\tau_1 = \tau_2 \circ T^*$ (and in particular T^* takes $\tau_1{}^{-1}(0)$ onto $\tau_2{}^{-1}(0)$);

ii) The restriction $T^*/\tau_1{}^{-1}(0) : \tau_1{}^{-1}(0) \rightarrow \tau_2{}^{-1}(0)$ preserves the euclidean structures, that is, $(T^*u, T^*v)_2 = (u,v)_1$ for all $u, v \in \tau_1{}^{-1}(0)$.

Exercise 3.1 Show that any two galilean space-time structures are isomorphic. Start by showing that any galilean space-time structure is isomorphic to the galilean coordinate space.

Let M be a set and $\varphi_1 : M \rightarrow \mathbf{R} \times \mathbf{R}^3$ a bijective map that it is called a **galilean coordinate system on** M. If φ_2 is another galilean system such that $\varphi_2 \circ \varphi_1{}^{-1} : \mathbf{R} \times \mathbf{R}^3 \rightarrow \mathbf{R} \times \mathbf{R}^3$ belongs to the galilean group G, one says that φ_2 **moves uniformly** with respect to φ_1.

Using a galilean coordinate system φ_1 on M and the galilean coordinate space $(\mathbf{R}\times\mathbf{R}^3, \tau, (,))$, one easily define a galilean space-time structure $(A_1, \tau_1, (,)_1)$.

In fact let $V_1 = {\varphi_1}^{-1}(\mathbf{R} \times \mathbf{R}^3)$ with the structure of a four dimensional vector space induced by the vector space $\mathbf{R} \times \mathbf{R}^3$, and let $A_1 = M = V_1$ be the four dimensional affine space associated to itself. The map $\tau_1 = \tau \circ \varphi_1$ is obviously a non zero linear map $\tau_1 : V_1 \to \mathbf{R}$ and on the three dimensional kernel ${\tau_1}^{-1}(0) = {\varphi_1}^{-1}(\{0\} \times \mathbf{R}^3)$ one defines the inner product $(,)_1$ induced by $(,)$.

It is clear that if φ_2 moves uniformly with respect to φ_1 (that is $\varphi_2 \circ {\varphi_1}^{-1} \in G$), the galilean space-time structure $(A_2, \tau_2, (,)_2)$ defined by φ_2, as above, is isomorphic to $(A_1, \tau_1, (,)_1)$ and, of course, isomorphic to the galilean coordinate space $(\mathbf{R} \times \mathbf{R}^3, \tau, (,))$.

A **motion** in $\mathbf{R}^N$ is a C^2 map $x : I \to \mathbf{R}^N$ where $I \subset \mathbf{R}$ is an open interval. The vectors $\dot{x}(t_o)$ and $\ddot{x}(t_o)$ in $\mathbf{R}^N$ are the **velocity** and the **acceleration at the point** $t_o \in I$. The image $x(I) \subset \mathbf{R}^N$ is called a **curve in** $\mathbf{R}^N$.

Let $\alpha : I \to \mathbf{R}^3$ be a motion in $\mathbf{R}^3$. The graph $\{(t, \alpha(t)) | t \in I\}$ is a curve in $\mathbf{R} \times \mathbf{R}^3$.

Let us come back to the case of a set M with a galilean system of coordinates $\varphi_1 : M \to \mathbf{R} \times \mathbf{R}^3$ and the corresponding galilean space-time structure induced by φ_1 on M. Consider also the atlas $\mathsf{a} = \{\varphi : M \to \mathbf{R} \times \mathbf{R}^3 | \varphi \circ {\varphi_1}^{-1} \in G\}$, that is, this atlas is the collection $\mathsf{a} = \{g \circ \varphi_1 | g \in G\}$.

A **world line on** M **relative to** a is the image $\gamma(J) \subset M$ of a map $\gamma : J \to M$ ($J \subset \mathbf{R}$ is an interval) such that $\varphi_1(\gamma(J))$ is the graph $\{(t, \alpha(t)) | t \in I\}$ of a motion $\alpha : I \to \mathbf{R}^3$.

Remark: If instead of φ_1 we use any $\varphi \in \mathsf{a}$, one can show that $\varphi(\gamma(J))$ is also the graph of a motion in $\mathbf{R}^3$. This fact follows from what can be proved in the next Exercise.

Exercise 4.1 Show that the maps g_1, g_2 and g_3 considered in Exercise 1.1 transform graphs of motions in $\mathbf{R}^3$ into graphs of motions in $\mathbf{R}^3$.

Example 3.1 Let E^3 be the affine space whose elements are the points of the euclidean geometry; E^3 is associated to the set V^3 of all translations of E^3 which is a three dimensional vector space. The set $M = \mathbf{R} \times E^3$ is an affine space associated to the four dimensional vector space $\mathbf{R} \times V^3$. $M = \mathbf{R} \times E^3$ is a model for the so called **physical space-time**; E^3 is said to be the **absolute space** and the first projection is the absolute time.

Any galilean system of coordinates (bijection) $\varphi_1 : M \to \mathbf{R} \times \mathbf{R}^3$ induces on M, as we saw, a galilean space-time structure and also defines the atlas $\mathsf{a} = \{g \circ \varphi_1 | g \in G\}$.

A motion of a **mechanical system of** n **points defined on** M, will give on M n world lines relative to a and correspondly n mappings $x_i : I \to \mathbf{R}^3$, $i = 1, \ldots, n$ that define one mapping $x : I \to \mathbf{R}^{3n}$ called a motion of a system of n points in the galilean coordinate space $\mathbf{R} \times \mathbf{R}^3$. The direct product $\mathbf{R}^3 \times \ldots \times \mathbf{R}^3 = \mathbf{R}^{3n}$ is called the **configuration space**.

According to the **Newton principle of determinacy** all motions of a mechanical system of n points are uniquely determined by their initial positions $x(t_o) \in \mathbf{R}^N$ and initial velocities $\dot{x}(t_o) \in \mathbf{R}^N$, $N = 3n$. In particular the accelerations are determined. So, there is a function $F : \mathbf{R}^N \times \mathbf{R}^N \times \mathbf{R} \to \mathbf{R}^N$ such that $\ddot{x} = F(x, \dot{x}, t)$, the **Newton equation**, is assumed to be of class C^1. This second order differential equation is determined experimentally for each specific mechanical system and constitutes a definition of it. By a classical theorem of existence and uniqueness of solutions, each motion is uniquely determined by $x(t_o)$ and $\dot{x}(t_o)$.

Galileo principle of relativity imposes strong constraints to Newton equations of a mechanical system. Its statement is the following: "*The physical space-time $\mathbf{R} \times E^3$ has a special galilean coordinate system φ_1 and its atlas $\mathfrak{a} = \{g \circ \varphi_1 | g \in G\}$ (the elements in $\mathfrak{a}$ are called the* **inertial coordinate systems**) *having the following property: If we subject the world lines relative to $\mathfrak{a}$ of all the n points of any mechanical system to one and the same galilean transformation, we obtain world lines relative to $\mathfrak{a}$ of the same mechanical system (with new initial conditions)*".

This imposes a series of restrictions on the form of the right-hand side F of Newton equations written in an inertial coordinate system.

Example 4.1 Since $g_2 \in G$ (see Exercise 4.1), if $x(t)$ is a solution of $\ddot{x} = F(x, \dot{x}, t)$ then $x(t+s)$ is also a solution for all $s \in \mathbf{R}$, so we have $\ddot{x}(t+s) = F(x(t+s), \dot{x}(t+s), t)$. As a consequence we have $F(x, \dot{x}, t) = F(x, \dot{x}, t-s)$ which shows that $\frac{\partial F}{\partial t} = 0$, so $F = F(x, \dot{x})$.

The invariance with respect to the translations $g_2 \in G$ means that "the space is homogeneous".

Exercise 5.1 Show that the second member of Newton equation depends only on the relative coordinates $x_j - x_k$ and $\dot{x}_j - \dot{x}_k$, that is $\ddot{x} = F(x, \dot{x})$ is written in its components $\ddot{x}_i$ as

$$\ddot{x}_i = F_i(\{x_j - x_k, \dot{x}_j - \dot{x}_k\}) \quad i, j, k = 1, \ldots, n.$$

Hint: First use g_2 (with $(s, \omega) = (0, -x_2)$) and see that the F_i depend on $x_j - x_2$ only; after use g_1 (with $v = -\dot{x}_2$) to show that the F_i depend on the relative

$\dot{x}_j - \dot{x}_2$ only.

Remark: The invariance under $g_3 \in G$ means that "the space is isotropic".

Exercise 6.1 Analyze the invariance under $g_3 \in G$ to see what one can say about the right hand side $F(x, \dot{x})$ of the Newton equation. After that, show that if a mechanical system consists of only one point, then its acceleration (in an inertial coordinate system) is equal to zero ("Newton's first law").

Hint: Use the results of Exercise 5.1 and the invariance under $g_3 \in G$.

Example 5.1 A mechanical system consists of two points. At the initial moment their velocities (in some inertial coordinate system) are equal to zero. Show that the points will stay on the line which connected them at the initial moment.

The two points satisfy $x_1(0) - x_2(0) = a \neq 0$, $\dot{x}_1(0) = \dot{x}_2(0) = 0$ and the system is

$$\begin{cases} \ddot{x}_1 = F_1(x_1 - x_2, \dot{x}_1 - \dot{x}_2) \\ \ddot{x}_2 = F_2(x_1 - x_2, \dot{x}_1 - \dot{x}_2) \end{cases} \tag{1.1}$$

where F_1 and F_2 are C^1-functions; by the invariance under $g_3 \in G$ we know that if (x_1, x_2) is a motion then $(\bar{x}_1 = Rx_1, \bar{x}_2 = Rx_2)$ is also a motion, that is,

$$\begin{aligned} \ddot{\bar{x}}_1 &= R\ddot{x}_1 = RF_1(x_1 - x_2, \dot{x}_1 - \dot{x}_2) = F_1(R(x_1 - x_2), R(\dot{x}_1 - \dot{x}_2)) \\ \ddot{\bar{x}}_2 &= R\ddot{x}_2 = RF_2(x_1 - x_2, \dot{x}_1 - \dot{x}_2) = F_2(R(x_1 - x_2), R(\dot{x}_1 - \dot{x}_2)). \end{aligned}$$

Assume, by contradiction, that or $x_1(t)$ or $x_2(t)$ does not remain on the line defined by $x_1(0)$ and $x_2(0)$. So, with a small rotation $R(\theta)$ (of angle θ) around that line (we may also assume that $0 \in \mathbf{R}^3$ is on the same line), one has $\bar{x}_1(t) \neq x_1(t)$ or $\bar{x}_2(t) \neq x_2(t)$. But

$$\begin{aligned} \dot{\bar{x}}_1(0) &= R\dot{x}_1(0) = 0 & \qquad & \dot{\bar{x}}_2(0) = R\dot{x}_2(0) = 0 \\ \bar{x}_1(0) &= Rx_1(0) = x_1(0) & \text{and} \qquad & \bar{x}_2(0) = Rx_2(0) = x_2(0) \end{aligned}$$

and, then, by uniqueness of solution of system (1.1) one has $\bar{x}_1(t) = x_1(t)$ and $\bar{x}_2(t) = x_2(t)$, which is a contradiction. So $x_1(t)$ and $x_2(t)$ remain on the line defined by $x_1(0), x_2(0)$, for all values of t.

1.2 Critical remarks on newtonian mechanics

By the end of the last century, the existence of an absolute space in the model $\mathbb{R} \times E^3$ of the physical space-time as an example of a galilean space-time structure, as well as the existence of a special galilean coordinate system that appears in the Galileo's principle of relativity, became dubious when highly accurate optical experiments were performed.

On a "human scale", the account of motion in Newtonian Mechanics is quite accurate but when it is pushed to extremes, some difficulties arise. For instance, no material object has been observed to travel faster than the (finite) speed c of light in a vacuum; but, in Newtonian theory, c plays no special role. Moreover, the light is propagated isotropically (with the same speed in all directions) in each supposed inertial system; if two inertial systems are passing one another (one inertial coordinate system is in uniform translation motion with respect to the other) and assuming a light pulse is emitted at their common origin at time zero, it is observed that both systems see their respective origins as the centers of the resulting spherical light pulse for all time. This phenomenon is known as the **light pulse paradox** and the observation was done, essentially, in the Michelson-Morley experiment.

This, together with other electromagnetic considerations, led Albert Einstein and other people to reject the notion of an absolute space. He still retained, however, the notion of a distinguished (but undefined) class of inertial systems. Einstein then showed that this rejection of an absolute space and the resulting notion of absolute motion of an inertial system forces us to abandon also the idea of an absolute time! (see [F], "Gravitational Curvature" by Theodore Frankel, W.H. Freeman and Co., San Francisco,1979).

Chapter 2

Differentiable manifolds

A **topological manifold** Q **of dimension** n is a topological Hausdorff space with a countable basis of open sets such that each $x \in Q$ has an open neighborhood homeomorphic to an open subset of the euclidean space $\mathbf{R}^n$. Each pair (U, φ) where U is open in $\mathbf{R}^n$ and φ is a homeomorphism of U onto the open set $\varphi(U)$ of Q is called a **local chart**, $\varphi(U)$ is a **coordinate neighborhood** and the inverse $\varphi^{-1} : \varphi(U) \longrightarrow U$, given by $y \in \varphi(U) \mapsto \varphi^{-1}(y) = (x^1(y), \ldots, x^n(y))$, is called a **local system of coordinates**. If a point $x \in Q$ is associated to two local charts $\varphi : U \longrightarrow Q$ and $\overline{\varphi} : \overline{U} \longrightarrow Q$, that is $x \in \varphi(U) \cap \overline{\varphi}(\overline{U})$, one obtains the bijection $\overline{\varphi}^{-1} \circ \varphi : W \longrightarrow \overline{W}$ where the open sets $W \subset U$ and $\overline{W} \subset \overline{U}$ are given by

$$W = \varphi^{-1}\left[\varphi(U) \cap \overline{\varphi}(\overline{U})\right] \quad \text{and} \quad \overline{W} = \overline{\varphi}^{-1}\left[\varphi(U) \cap \overline{\varphi}(\overline{U})\right]$$

The charts (φ, U) and $(\overline{\varphi}, \overline{U})$ are said to be C^k- **compatible** if $\overline{\varphi}^{-1} \circ \varphi : W \longrightarrow \overline{W}$ is a C^k-diffeomorphism, $k \geq 1$, $k = \infty$ or $k = \omega$.

A C^k**-atlas** is a set of C^k compatible charts covering Q. Two C^k-atlases are said to be **equivalent** if their union is a C^k-atlas. A C^k **differentiable manifold** is a topological manifold Q with a class of equivalence of C^k-atlases. A manifold is **connected** if it cannot be divided into two disjoint open subsets (if no mention is made, a manifold means a C^∞-differentiable manifold).

Examples of differentiable manifolds:

1.2 $\mathbf{R}^n$

2.2 The sphere $S^2 = \{(x, y, z) \in \mathbf{R}^3 | x^2 + y^2 + z^2 = 1\}$.

3.2 The configuration space S^1 of the planar pendulum.

4.2 The configuration space of the double planar pendulum, that is, the torus $T^2 = S^1 \times S^1$.

5.2 The configuration space of the double spherical pendulum, that is, the product $S^2 \times S^2$ of two spheres.

6.2 The configuration space of a "rigid" line segment in the plane, $\mathbf{R}^2 \times S^1$.

7.2 The configuration space of a "rigid" right triangle AOB, $\hat{O} = 90°$, that moves around O; it can be identified with the set $SO(3)$ of all 3×3 orthogonal matrices with determinant 1.

8.2 $P^n(\mathbf{R})$, the n-dimensional real projective space (set of lines passing through $o \in \mathbf{R}^{n+1}$), $n \geq 1$.

2.1 Imbedded manifolds in $\mathbf{R}^N$

We say that $Q^n \subset \mathbf{R}^N$ is a C^k **submanifold** of (**manifold imbedded** in) $\mathbf{R}^N$ with dimension $n \leq N$, if Q^n is covered by a finite or countable number of images $\varphi(U)$ of the so called **regular parametrizations** , that is, C^k-maps, $k \geq 1$,

$\varphi : U \subset \mathbf{R}^n \longrightarrow \mathbf{R}^N$, U open set of $\mathbf{R}^n$, such that:

i) $\varphi : U \longrightarrow \varphi(U)$ is a homeomorphism where $\varphi(U)$ is open in Q^n with the topology induced by $\mathbf{R}^N$;

ii) $\frac{\delta\varphi}{\delta x}(x_0) : \mathbf{R}^n \longrightarrow \mathbf{R}^N$ is injective for all $x_o \in U$.

Here $\varphi(x_1, \ldots, x_n) = (\varphi^1(x_1, \ldots, x_n), \ldots, \varphi^N(x_1, \ldots, x_n))$ and $\frac{\delta\varphi}{\delta x}(x_0)$ is the $N \times n$ matrix $\frac{\delta\varphi}{\delta x}(x_o) = (\frac{\delta\varphi^i}{\delta x_j}(x_0))$.

To show that Q^n is a C^k-manifold we prove the next two propositions:

Proposition 1.2 *Let Q^n be a C^k submanifold of $\mathbf{R}^N$ with dimension n and $\varphi : U \longrightarrow \mathbf{R}^N$ a regular parametrization in a neighborhood of $y_0 \in \varphi(U) \subset Q^n$. Then, there exist an open neighborhood Ω of y_0 in $\mathbf{R}^N$ and a C^k-map $\Psi : \Omega \longrightarrow \mathbf{R}^N$ such that*

$$\Psi | (Q^n \cap \Omega) = \varphi^{-1} | (Q^n \cap \Omega) .$$

Proof We may assume, without loss of generality, that the first determinant (n first lines and n columns) of $\frac{\delta\varphi}{\delta x}(x_0)$ does not vanish (here $y_0 = \varphi(x_0)$). Define the function $F : U \times \mathbf{R}^{N-n} \longrightarrow \mathbf{R}^N$ by $F(x; z) = (\varphi^1(x), \ldots, \varphi^n(x)); \varphi^{n+1}(x) +$

$z_1, \ldots, \varphi^N(x) + z_{N-n})$ which is of class C^k; we have, clearly, $F(x,0) = \varphi(x)$, for all $x \in U$, so $F(x_0, 0) = \varphi(x_0) = y_0$ and

$$\frac{\delta F}{\delta(x,z)}(x_0, 0) = \left[\begin{array}{ccc} \frac{\delta \varphi^i}{\delta x_j}(x_0) & \vdots & 0 \\ i,j = 1,\ldots,n & \vdots & \\ \cdots\cdots\cdots\cdots & \vdots & \cdots \\ * & \vdots & I \end{array}\right].$$

From this it follows that $det \frac{\delta F}{\delta(x,z)}(x_0, 0) \neq 0$. The result comes, using the inverse function theorem, that is, F is a (local) diffeomorphism onto an open neighborhood Ω of y_0 in $\mathbf{R}^N$ with an inverse Ψ defined in Ω. It is also clear that $\Psi|(Q^n \cap \Omega) = \varphi^{-1}|(Q^n \cap \Omega)$.

Remark: Denote by π_2 the second projection $\pi_2 : U \times \mathbf{R}^{N-n} \longrightarrow \mathbf{R}^{N-n}$ and let f be the composition $f = \pi_2 \circ \Psi : \Omega \longrightarrow \mathbf{R}^{N-n}$, so, to any $y \in Q^n$ that belongs to Ω, one associates $N - n$ functions $f_1, \ldots, f_{N-n} : \Omega \longrightarrow \mathbf{R}$ such that $f = (f_1, \ldots, f_{N-n})$ and $\Omega \cap Q^n$ is given by the equations $f_1 = \ldots = f_{N-n} = 0$, the differentials $df_1(y), \ldots, df_{N-n}(y)$ being linearly independent.

Conversely we have the following:

Proposition 2.2 *Let $Q \subset \mathbf{R}^N$ be a set such that any point $y \in Q$ has an open neighborhood Ω in $\mathbf{R}^N$ and $N - n$ C^k-differentiable functions, $k \geq 1$,*

$$f_1 : \Omega \longrightarrow \mathbf{R}, \ldots, f_{N-n} : \Omega \longrightarrow \mathbf{R}$$

such that $\Omega \cap Q$ is given by $f_1 = \ldots = f_{N-n} = 0$, with $df_1(y), \ldots, df_{N-n}(y)$ linearly independent. Then Q is a C^k submanifold of (manifold imbedded in) $\mathbf{R}^N$ with dimension n.

Proof The linear forms $df_i(y) : \mathbf{R}^N \longrightarrow \mathbf{R}$, $i = 1, \ldots, N-n$, define a surjective linear transformation

$$(df_1(y), \ldots, df_{N-n}(y)) : \mathbf{R}^N \longrightarrow \mathbf{R}^{N-n}$$

with a n-dimensional kernel $K \subset \mathbf{R}^N$. Let $L : \mathbf{R}^N \longrightarrow \mathbf{R}^n$ be any linear transformation such that the restriction $L|K$ is an isomorphism from K onto $\mathbf{R}^n$. Define $G : \Omega \subset \mathbf{R}^N \longrightarrow \mathbf{R}^N$ by

$$G(\xi) = (f_1(\xi), \ldots, f_{N-n}(\xi), L(\xi))$$

whose derivative at $y \in Q$ is given by

$$dG(y)v = (df_1(y)v, \ldots, df_{N-n}(y)v, L(v)).$$

Then $dG(y)v$ is non singular and so, by the inverse function theorem, G takes an open neighborhood $\tilde{\Omega}$ of y, diffeomorphically onto a neighborhood $G(\tilde{\Omega})$ of $(o, L(y))$. Note that if $f \stackrel{def}{=} (f_1, \ldots, f_{N-n})$, $f^{-1}(o) \cap \tilde{\Omega} = Q \cap \tilde{\Omega}$ corresponds, under the action of G, to points of the hyperplane $(o, \mathbf{R}^n)$ since G takes $f^{-1}(0) \cap \tilde{\Omega}$ onto $(0, \mathbf{R}^n) \cap G(\tilde{\Omega})$. The inverse φ of G restricted to $f^{-1}(o) \cap \tilde{\Omega}$ is a C^k-bijection:

$$\varphi : U \stackrel{def}{=} (o, \mathbf{R}^n) \cap G(\tilde{\Omega}) \longrightarrow \varphi(U) = Q \cap \ \tilde{\Omega}.$$

To the point $y \in Q$ then corresponds a local chart (φ, U), that is, Q is a C^k-submanifold of $\mathbf{R}^N$, with dimension n.

Exercise 1.2 The orthogonal matrices are obtained between the real 3×3 matrices (these are essentially $\mathbf{R}^9$) as the zeros of six functions (the orthogonality conditions). This way we obtain two connected components, since the determinant of an orthogonal matrix is equal to $+1$ or -1. The component with determinant $+1$ is the group $SO(3)$ of rotations of $\mathbf{R}^3$. Show that $SO(3)$ is a compact submanifold of $\mathbf{R}^9$ of dimension 3.

2.2 The tangent space

Let Q be a n-dimensional submanifold of $\mathbf{R}^N$. To any $y \in Q$ is associated a subspace T_yQ of dimension n; in the notation of Proposition 2.2, T_yQ is the kernel K of the linear map

$$(df_1(y), \ldots, df_{N-n}(y)) : \mathbf{R}^N \longrightarrow \mathbf{R}^{N-n}.$$

The vectors of $T_yQ = K$ are called the **tangent vectors** to Q at the point $y \in Q$ and the subspace T_yQ is the **tangent space** of Q at the point y. The tangent vectors at y can also be defined as the velocities $\dot{\gamma}(0)$ of all C^1-curves $\gamma : (-\varepsilon, +\varepsilon) \longrightarrow \mathbf{R}^N$ with values on Q and such that $\gamma(0) = y$.

In the general case of a manifold Q one defines an equivalence relation at $y \in Q$ between smooth curves. So, a continuous curve $\gamma : I \longrightarrow Q$ (I is any interval containing $0 \in \mathbf{R}$) is said to be smooth at zero if for any local chart $(U; \varphi), \gamma(0) = y \in \varphi(U)$, the curve $\varphi^{-1} \circ \gamma|_{\gamma^{-1}(\varphi(U))} : \gamma^{-1}(\varphi(U)) \longrightarrow \mathbf{R}^n$ is smooth. Two smooth curves $\gamma_1 : I_1 \longrightarrow Q$ and $\gamma_2 : I_2 \longrightarrow Q$ such that $\gamma_1(0) = \gamma_2(0) = y$ are equivalent if $\frac{d}{dt}(\varphi^{-1} \circ \gamma_1)|_{t=0} = \frac{d}{dt}(\varphi^{-1} \circ \gamma_2)|_{t=0}$. This concept does not depend on the local chart $(U;\varphi)$. A **tangent vector v_y at** $y \in Q$ is a class of equivalence of that equivalence relation. We write simply $v_y = \dot{\gamma}_1(0) = \dot{\gamma}_2(0)$. One defines sum of tangent vectors at y and product of a real number by a tangent vector. This way the set T_yQ of all tangent vectors to Q at $y \in Q$ is a vector space with dimension n. With a local chart $(U;\varphi)$ and the canonical basis $\{e_i\}(i = 1,\ldots,n)$ of $\mathbb{R}^n$, it is possible to construct a basis of T_yQ at $y \in \varphi(U)$; if we set $x_0 = \varphi^{-1}(y)$, consider the tangent vectors associated to the curves $\gamma_i : t \mapsto \varphi^{-1}(x_0 + te_i)$ and let $\frac{\partial}{\partial x_i}(y) \overset{def}{=} \dot{\gamma}_i(0), i = 1,\ldots,n$.

2.3 The derivative of a differentiable function

A continuous function $f : Q_1 \longrightarrow Q_2$ defined on a differentiable manifold Q_1 with values on a differentiable manifold Q_2 is said to be **C^r- differentiable** at $y \in Q_1$ if for any two charts (U,φ) and $(\overline{U},\overline{\varphi})$, $\varphi^{-1}(y) \in U$ and $\overline{\varphi}^{-1}(f(y)) \in \overline{U}$, the map $\overline{\varphi}^{-1}.f.\varphi : U \longrightarrow \overline{U}$ is C^r differentiable at $\varphi(y)$, $r > 0$; of course we are assuming (as we can) $f(\varphi(U)) \subset \overline{\varphi}(\overline{U})$ (reducing U if necessary), due to the continuity of f at $y \in Q_1$. The notion of differentiability does not depend on the used local charts. One uses to say **smooth** instead of C^∞. The derivative $df(y)$ or $f_*(y)$ of a C^1- differentiable function $f : Q_1 \longrightarrow Q_2$ at $y \in Q_1$ is a linear map

$$f_*(y) : T_yQ_1 \longrightarrow T_{f(y)}Q_2$$

that sends a tangent vector represented by a curve $\gamma : I \longrightarrow Q_1$, $\gamma(0) = y \in Q_1$, into the tangent vector at $f(y) \in Q_2$ represented by the curve $f_o\gamma : I \longrightarrow Q_2$. One can show that $f_*(y)$ is linear.

If $g : Q_2 \longrightarrow Q_3$ is another C^1-differentiable function one has:

$$T_yQ_1 \overset{f_*(y)}{\longrightarrow} T_{f(y)}Q_2 \overset{g_*(f(y))}{\longrightarrow} T_{g(f(y))}Q_3$$

and it can be proved that

$$(g \circ f)_*(y) = g_*(f(y)) \circ f_*(y) \quad \text{for all} \quad y \in Q_1.$$

A **C^r-diffeomorphism** $f : Q_1 \longrightarrow Q_2$ is a bijection such that f and f^{-1} are C^r-differentiable, $r \geq 1$.

2.4 Tangent and cotangent bundles of a manifold

Let Q be a C^k-differentiable manifold, $k \geq 2$. Consider the sets $TQ = \cup_{y\in Q}\ T_yQ$ and $T^*Q = \cup_{y\in Q}T_y{}^*Q$ where $T_y{}^*Q$, $y \in Q$, is the dual of T_yQ, that is, $T_y{}^*Q$ is the set of all linear forms defined on T_yQ.

Exercise 2.2 Show that TQ and T^*Q are C^{k-1}-manifolds if Q is a C^k-manifold, $k \geq 2$. Show also that the canonical projections:

$$\begin{array}{rcll} \tau & : & v_y \in TQ \mapsto y \in Q & \text{and} \\ \tau^* & : & \omega_y \in T^*Q \mapsto y \in Q & \end{array}$$

are C^{k-1} maps.

TQ and T^*Q are called the **tangent** and **cotangent bundles** of Q, respectively.

Exercise 3.2 The cartesian product of two manifolds is a manifold.

Exercise 4.2 (Inverse image of a regular value) Let $F : U \subset \mathbf{R}^n \longrightarrow \mathbf{R}^m$ be a differentiable map defined on an open set $U \subset \mathbf{R}^n$. A point $p \in U$ is a **critical point** of F if $dF(p) : \mathbf{R}^n \longrightarrow \mathbf{R}^m$ is not surjective. The image $F(p) \in \mathbf{R}^m$ of a critical point is said to be a **critical value** of F. A point $a \in \mathbf{R}^m$ is a **regular value** of F if it is not a critical value. Show that the inverse image $F^{-1}(a)$ of a regular value $a \in \mathbf{R}^m$ either is a submanifold of $\mathbf{R}^n$, contained in U, with dimension equal to $n - m$, or $F^{-1}(a) = \phi$.

Let Q be a differentiable manifold. Q is said to be **orientable** if Q has an atlas $\mathbf{a} = \{(U_\alpha, \varphi_\alpha)\}$ such that $(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) in $\mathbf{a}$ satisfying $\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta) \neq \phi$, the derivative of $\varphi_\beta{}^{-1} \circ \varphi_\alpha$ at any $x \in \varphi_\alpha{}^{-1}[\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)]$ has positive determinant. If one fix such an atlas, Q is said to be **oriented**. If Q is orientable and connected, it can be oriented in exactly to ways.

Exercise 5.2 Show that TQ is orientable (even if Q is not orientable). Show that a two-dimensional submanifold Q of $\mathbf{R}^3$ is orientable if, and only if, there is on Q a differentiable normal unitary vector field $N : Q \longrightarrow \mathbf{R}^3$, that is, for all $y \in Q, N(y)$ is orthogonal to T_yQ.

Exercise 6.2 Use the stereographic projections and show that the sphere $S^n = \{(x_1, \ldots, x_{n+1}) \in \mathbf{R}^{n+1} | \sum_{i=1}^{n+1} x^2{}_i = 1\}$ is orientable.

2.5 Discontinuous action of a group on a manifold

An **action** of a group G on a differentiable manifold M is a map

$$\varphi : G \times M \longrightarrow M$$

such that:

1) for any fixed $g \in G$, the map $\varphi_g : M \longrightarrow M$ given by $\varphi_g(p) = \varphi(g, p)$ is a diffeomorphism and $\varphi_e =$ Identity of M ($e \in G$ is the identity);

2) if g and h are in G then $\varphi_{gh} = \varphi_g \circ \varphi_h$ where gh is the product in G.

An action $\varphi : G \times M \longrightarrow M$ is said to be **properly discontinuous** if any $p \in M$ has a neighborhood U_p in M such that $U_p \cap \varphi_g(U_p) = \phi$ for all $g \neq e$, $g \in G$.

Any action of G on M defines an equivalence relation $\sim$ between elements of M; in fact, one says that $p_1 \sim p_2$ (p_1 equivalent to p_2) if there exists $g \in G$ such that $\varphi_g(p_1) = p_2$. The quotient space M/G under $\sim$ with the quotient topology is such that the canonical projection $\pi : M \longrightarrow M/G$ is continuous and open. ($\pi(p) \in M/G$ is the class of equivalence of $p \in M$).

The open sets in M/G are the images by π of open sets in M. Since M has a countable basis of open sets, M/G also has a countable basis of open sets.

Exercise 7.2 Show that the topology of M/G is Hausdorff if and only if given two non equivalent points p_1, p_2 in M, there exist neighborhoods U_1 and U_2 of p_1 and p_2 such that $U_1 \cap \varphi_g(U_2) = \phi$ for all $g \in G$.

Exercise 8.2 Show that if $\varphi : G \times M \longrightarrow M$ is properly discontinuous and M/G is Hausdorff then M/G is a differentiable manifold and $\pi : M \longrightarrow M/G$ is a local diffeomorphism, that is, any point of M has an open neighborhood Ω such that π sends Ω diffeomorphically onto the open set $\pi(\Omega)$ of M/G. Show also that M/G is orientable if and only if M is oriented by an atlas $\mathsf{a} = \{(U_\alpha, \varphi_\alpha)\}$ preserved by the diffeomorphisms $\varphi_g, g \in G$ (that is, $(U_\alpha, \varphi_g \circ \varphi_\alpha)$ belongs to a for all $(U_\alpha, \varphi_\alpha) \in \mathsf{a}$).

Examples: 9.2 Let $M = S^n \subset \mathbf{R}^{n+1}$ and G be the group of diffeomorphisms of S^n with two elements: the identity and the antipodal map $A : x \mapsto -x$. The quotient S^n/G can be identified with the projective space $P^n(\mathbf{R})$.

10.2 Let $M = \mathbf{R}^k$ and G be the group Z^k of all integer translations, that is, the action of $g = (n_1, \ldots, n_k) \in Z^k$ on $x = (x_1, \ldots, x_k) \in \mathbf{R}^k$ means to obtain $x + g \in \mathbf{R}^k$. The quotient $\mathbf{R}^k/Z^k$ is the **torus** T^k. The torus T^2 is diffeomorphic to the **torus of revolution** $\tilde{T}^2$, submanifold of $\mathbf{R}^3$ obtained as the inverse image of zero under the map $f(x,y,z) = z^2 + (\sqrt{x^2+y^2} - a)^2 - r^2$.

11.2 Let S be a submanifold of $\mathbf{R}^3$ symmetric with respect to the origin and $G = \{e, A\}$ be the group considered in example 9.2 above. The special case $S = \tilde{T}^2$ (torus of revolution in $\mathbf{R}^3$) gives us the quotient manifold $\tilde{T}^2/G \stackrel{def}{=} K$, the so called **Klein bottle.** When S is the manifold $S = \{(x,y,z) \in \mathbf{R}^3 | x^2 + y^2 = 1, -1 < z < 1\}$ then S/G is called the **Möbius band.**

Exercise 9.2 Show that the Klein bottle, the Möbius band and $P^2(\mathbf{R})$ are not orientable. Show also that $P^n(\mathbf{R})$ is orientable if and only if n is odd.

2.6 Immersions and imbeddings . Submanifolds

Let M and N be differentiable manifolds and $\varphi : M \longrightarrow N$ be a differentiable map. φ is said to be an **immersion** of M **into** N if $\varphi_*(p) : T_pM \longrightarrow T_{\varphi(p)}N$ is injective for all $p \in M$.

An **imbedding of** M **into** N is an immersion $\varphi : M \longrightarrow N$ such that φ is a homeomorphism of M onto $\varphi(M) \subset N$, $\varphi(M)$ with the topology induced by N. If $M \subset N$ and the inclusion $i : M \longrightarrow N$ is an imbedding of M into N, M is said to be a **submanifold of** N.

Examples: 12.2 The map $\varphi : \mathbf{R} \longrightarrow \mathbf{R}^2$ given by $\varphi(t) = (t, |t|)$ is not differentiable at $t = 0$.

13.2 The map $\varphi : \mathbf{R} \longrightarrow \mathbf{R}^2$ defined by $\varphi(t) = (t^3, t^2)$ is differentiable but is not an immersion because $\varphi_*(0) : \mathbf{R} \longrightarrow \mathbf{R}^2$ is the zero map that is not injective.

14.2 The map $\varphi : (0, 2\pi) \longrightarrow \mathbf{R}^2$ defined by

$$\varphi(t) = (2\cos(t - \frac{\pi}{2}), \quad \sin 2(t - \frac{\pi}{2}))$$

is an immersion but is not a imbedding. The image $M = \varphi((0, 2\pi))$ is an "eight". Also, the inclusion $i : M \longrightarrow \mathbf{R}^2$ is not an imbedding, so $M = \varphi((0, 2\pi))$ is not a submanifold of $\mathbf{R}^2$.

15.2 The curve $\varphi : (-3, 0) \longrightarrow \mathbf{R}^2$ given by:

$$\varphi(t) = \begin{cases} (0, -(t+2)) & \text{if } \ t \in (-3, -1) \\ \text{a regular curve for } & t \in (-1, -\frac{1}{\pi}) \\ (-t, -\sin\frac{1}{t}) & \text{if } \ t \in (-\frac{1}{\pi}, 0) \end{cases}$$

is an immersion but is not an imbedding.

A neighborhood of $O = (0,0)$ has infinitely many connected components if one considers the induced topology for the set $\varphi(-3,0) \subset \mathbf{R}^2$.

16.2 $\varphi : \mathbf{R} \longrightarrow \mathbf{R}^3$ defined by $\varphi(t) = (\cos 2\pi t, \sin 2\pi t, t)$ is an imbedding. The image $\varphi(\mathbf{R})$ is homeomorphic to $\mathbf{R}$.

17.2 The image $\varphi(\mathbf{R})$ of the map $\varphi : \mathbf{R} \longrightarrow \mathbf{R}^2$ given by $\varphi(t) = (\cos 2\pi t, \sin 2\pi t)$ is $S^1 \subset \mathbf{R}^2$. The map φ is an immersion but not an imbedding since is not injective. But $\varphi(\mathbf{R}) = S^1$ is a submanifold of $\mathbf{R}^2$ if we consider the inclusion map $i : S^1 \longrightarrow \mathbf{R}^2$.

18.2 Analyze the maps:

$$\begin{aligned} \varphi_1(t) &= (\frac{1}{t}\cos 2\pi t, \frac{1}{t}\sin 2\pi t), \quad t \in (1,\infty); \\ \varphi_2(t) &= (\frac{t+1}{2t}\cos 2\pi t, \quad \frac{t+1}{2\pi}\sin 2\pi t), \quad t \in (1,\infty) \end{aligned}$$

2.7 Partition of unity

Let X be a topological space. A **covering** of X is a family $\{U_i\}$ of open sets U_i in X such that $\bigcup_i U_i = X$. A covering of X is said to be **locally finite** if any point of X has a neighborhood that intersects a finite number of elements in the covering, only. One says that a covering $\{V_k\}$ is **subordinated** to $\{U_i\}$ if each V_k is contained in some U_i. Let B_r be the ball of $\mathbf{R}^m$ centered at $0 \in \mathbf{R}^m$ and radius $r > 0$.

Proposition 3.2 *Let X be a differentiable manifold,* dim $X = m$. *Given a covering of X, there exists an atlas $\{(V_k, \varphi_k)\}$ where $\{V_k\}$ is a locally finite covering of X subordinated to the given covering, and such that $\varphi_k^{-1}(V_k)$ is the ball B_3 and, moreover, the open sets $W_k = \varphi_k(B_1)$ cover X.*

For a proof see the book [L] "Differential Manifolds" by S. Lang, Addison Wesley Pu. Co., p. 33, taking into account that in the last Proposition 3.2 X is Hausdorff, finite dimensional and has a countable basis.

The **support** *supp* (f) of a function $f : X \to \mathbf{R}$ is the closure of the set of points where f does not vanish. We say that a family $\{f_k\}$ of differentiable functions $f_k : X \to \mathbf{R}$ is **a differentiable partition of unity subordinated to a covering** $\{V_k\}$ **of** X if:

(1) For any k, $f_k \geq 0$ and *supp* (f_k) is contained in a coordinate neighborhood V_k of an atlas $\{(V_k, \varphi_k)\}$ of X.

(2) The family $\{V_k\}$ is a locally finite covering of X.

(3) $\sum_\alpha f_\alpha(p) = 1$ for any $p \in X$ (this condition makes sense since for each p, $f_\alpha(p) \neq 0$ for a finite number of indices, only).

Proposition 4.2 *Any connected differentiable manifold X has a differentiable partition of unity.*

Proof The idea is the following: by Proposition 3.2, for each k one defines a smooth "cut off" function $\psi_k : X \to \mathbf{R}$ of compact support contained in V_k such that ψ_k is identically 1 on W_k and $\psi_k \geq 0$ on X. From the fact that $\{V_k\}$ is a locally finite covering of X subordinated to the given initial covering $\{U_i\}$ of X, the sum $\sum_k \psi_k = \psi$ exists; moreover ψ is smooth and $\psi(p) > 0$ for any $p \in X$. Then the functions $f_k = \psi_k/\psi$ have the desired properties (1), (2) and (3) above.

For a complete proof see also the book [R] "Differentiable Manifolds" by G. de Rham, Springer-Verlag 1984, p.4.

Chapter 3

Vector fields, differential forms and tensor fields

We already saw that given a local chart $\varphi : U \longrightarrow \varphi(U) = V$ of a differentiable manifold Q, to each $x = (x^1, \ldots, x^n) \in U \subseteq \mathbf{R}^n$ corresponds the vectors $x + e_i \in \mathbf{R}^n$, (e_i) being the canonical basis. The curves $\varphi(x + te_i), i = 1, \ldots, n$, for $|t| < \epsilon$, ($\epsilon > 0$ small in order that $x + te_i \in U$) define the tangent vectors to Q at $\varphi(x)$ denoted by $\frac{\partial}{\partial x_i}(\varphi(x))$. We may also write

$$\varphi_*(x)e_i = \frac{\partial}{\partial x_i}(\varphi(x)), \quad i = 1, \ldots, n$$

and

$$(\frac{\partial}{\partial x_i}(\varphi(x))_{1 \leq i \leq n} \quad \text{span} \quad T_{\varphi(x)}Q.$$

A **vector field** X on a C^∞-manifold Q is a map $y \in Q \mapsto X(y) \in T_yQ \subset TQ$. It is clear that if $\varphi : U \mapsto V \subset Q$ is a local chart, the maps $\frac{\partial}{\partial x_i} : y \in V \longrightarrow \frac{\partial}{\partial x_i}(y)$ are vector fields on V. A vector field X on Q is said to be **of class** C^∞ (or **smooth**) if given any local chart $\varphi : U \longrightarrow V \subset Q, X$ is written as

$$X(y) = \sum_{i=1}^{n} a_i(y) \frac{\partial}{\partial x_i}(y)$$

with the functions $a_i : y \in V \mapsto a_i(y) \in \mathbf{R}$ being C^∞ -functions. This means that the map $X : Q \mapsto TQ$ satisfies $\tau \circ X = idQ$ and is a C^∞-differentiable map ($\tau : TQ \longrightarrow Q$ is the canonical projection and idQ is the identity map on Q). Let $D(Q)$ be the set of all C^∞-functions $f : Q \longrightarrow \mathbf{R}$ and $\mathcal{X}(Q)$ be the set of all C^∞-vector fields on Q. Any $X \in \mathcal{X}(Q)$ is a derivative of functions, in the sense that given a C^∞-differentiable function $f : Q \longrightarrow \mathbf{R}, f \in D(Q)$, then $X(f) \in D(Q)$ is the C^∞-differentiable function defined as

$$X(f)(y) = df(y)\,[X(y)] \quad \text{for any} \quad y \in Q.$$

In local coordinates, if $X = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i}$ then $X(f) = \sum_{i=1}^n a_i \frac{\partial f}{\partial x_i}$ (this equality holds in V).

Remark also that $\frac{\partial}{\partial x_i}(f \circ \varphi)(x) = \frac{\partial f}{\partial x_i}(\varphi(x))$, for all $x \in U$.

It is easy to see that if f and g are C^∞-differentiable functions and $\alpha, \beta \in \mathbf{R}$ we have, for any C^∞-vector field X, the equalities:

$$\begin{aligned} X(\alpha f + \beta g) &= \alpha X(f) + \beta X(g), \\ X(fg) &= f X(g) + g X(f). \end{aligned}$$

Given two C^∞-vector fields X and Y defined on a C^∞-manifold Q, they define the **Lie bracket** $[X, Y]$ as the unique C^∞-vector field Z such that, for any C^∞-differentiable function $f : Q \longrightarrow \mathbf{R}$, one has $Zf = [X, Y]f \overset{def}{=} X(Yf) - Y(Xf)$. If, in local coordinates, we have the expressions:

$$X = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i}, \qquad Y = \sum_{i=1}^n b_i \frac{\partial}{\partial x_i}$$

then, a simple computation shows that

$$[X, Y]f = \sum_{i,j=1}^n (a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i}) \frac{\partial f}{\partial x_j}$$

so, the uniqueness follows. To prove the existence of $[X, Y]$ we define, locally

$$[X, Y] = \sum_{i,j=1}^n (a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i}) \frac{\partial}{\partial x_j},$$

and show that the definition is coherent in the intersection of two coordinate neighborhoods.

Exercise 1.3 Complete the proof of the existence of $[X, Y]$.

Exercise 2.3 Show that if X, Y, Z are C^∞ vector fields, f, g C^∞-functions and $\alpha, \beta \in \mathbf{R}$ one has:

$$\begin{aligned}
[X,Y] &= -[Y,X]; \quad [\alpha X+\beta Y, Z] = \alpha[X,Z]+\beta[Y,Z]. \\
[fX, gY] &= fg[X,Y]+f(Xg)Y-g(Yf)X. \\
[[X,Y],Z] &+ [[Y,Z],X]+[[Z,X],Y]=0 \qquad \text{(Jacobi identity)}.
\end{aligned}$$

We want now to introduce some other machinery used on manifolds: differential forms, exterior derivative, interior product, tensor fields and Lie derivative.

Let $\Lambda^k Q$ be the manifold of all **exterior k-forms** on Q. This means that

$$\Lambda^k Q = \cup_{y\in Q} \Lambda^k T_y{}^* Q$$

where $\Lambda^k T_y{}^* Q$ is the space of all alternate k-linear forms on $T_y Q$; recall that $\Lambda^1 Q = T^* Q$ and $\Lambda^0 Q = D(Q)$. Denote by $\tau^k : \Lambda^k Q \longrightarrow Q$ the natural projection and by $\Gamma^k(Q)$ the set of all C^∞-differentiable k-forms on Q, that is, $\sigma \in \Gamma^k(Q)$ is a cross section, with respect to τ^k, of the vector bundle $(\Lambda^k Q, Q, \tau^k)$; so σ is a C^∞-map $\sigma : Q \longrightarrow \Lambda^k Q$ such that $\tau^k \circ \sigma = id \ Q$.

Any smooth map $f : Q_1 \longrightarrow Q_2$ from a manifold Q_1 into a manifold Q_2 has a natural extension f^* that acts on the k-forms $\sigma \in \Gamma^k(Q_2)$; in fact, $f^*\sigma \in \Gamma^k Q_1$ is defined as follows:

$$(f^*\sigma)(y)(X_1(y),\ldots,X_k(y)) \stackrel{def}{=} \sigma(f(y))(f_* X_1(y),\ldots,f_* X_k(y))$$

for all $y \in Q_1$ and $X_1,\ldots,X_k \in \mathcal{X}(Q_1)$.

We also write, for simplicity,

$$f^*\sigma(X_1,\ldots,X_k) = \sigma(f_* X_1,\ldots,f_* X_k) \tag{3.1}$$

It is clear that f^* is linear.

If σ^k and σ^l are in $\Gamma^k(Q)$ and $\Gamma^l(Q)$, respectively, $\sigma^k \wedge \sigma^l$ is the $(k+l)$-form in $\Gamma^{k+l}(Q)$ defined by

$$\sigma^k \wedge \sigma^l(X_1,\ldots,X_{k+l}) \stackrel{def}{=} \sum (-1)^\epsilon \sigma^k(X_{i_1},\ldots,X_{i_k})\sigma^l(X_{j_1},\ldots,X_{j_l}) \tag{3.2}$$

the $\sum$ being extended to all sequences $(i_1 < \ldots < i_k; j_1 < \ldots < j_l)$ where $(i_1,\ldots,i_k,j_1,\ldots,j_l)$ is a sign ϵ permutation of the indices $(1,\ldots,k+l)$ such that $i_1 < i_2 < \ldots < i_k$ and $j_1 < j_2 < \ldots < j_l$.

Given $\sigma^k \in \Gamma^k(Q)$, $\sigma^l \in \Gamma^l(Q)$ and $f : Q \longrightarrow Q$ differentiable, one has:

$$f^*(\sigma^k \wedge \sigma^l) = f^*\sigma^k \wedge f^*\sigma^l.$$

For $\sigma^k \in \Gamma^k(Q)$ and local coordinates $(V; x_1, \ldots, x_n)$ on Q, we have:

$$\sigma^k(y) = \sum_{i_1 < \ldots < i_k} S_I(y) dx_{i1}(y) \wedge \ldots \wedge dx_{ik}(y),$$

where $I = (i_1, \ldots, i_k)$ and $S_I(y) = S_I(x_1, \ldots, x_n)$ are differentiable functions on V. Omitting the point $y \in V$ we set, simply,

$$\sigma^k = \sigma^k / V = \sum_{i_1 < \ldots < i_k} S_I dx_{i1} \wedge \ldots \wedge dx_{ik}. \tag{3.3}$$

One can also show that:

$$\sigma^k \wedge (\sigma_1{}^l + \sigma_2{}^l) = \sigma^k \wedge \sigma_1{}^l + \sigma^k \wedge \sigma_2{}^l; \tag{3.4}$$

$$\sigma^k \wedge (\sigma^l \wedge \sigma^m) = (\sigma^k \wedge \sigma^l) \wedge \sigma^m \tag{3.5}$$

The **exterior derivative** $d : \Gamma(Q) \longrightarrow \Gamma(Q)$ is an operation on the algebra $\Gamma(Q) = \bigoplus_{k=o}^{n} \Gamma^k(Q)$, where $n = dim\ Q$ and $k \geq 0$, which is linear and

$$d\sigma^k \in \Gamma^{k+1} \quad (Q) \qquad \text{for} \quad \sigma^k \in \Gamma^k(Q);$$

$$d(\sigma^k \wedge \omega) = d\sigma^k \wedge \omega + (-1)^k \sigma^k \wedge d\omega; \tag{3.6}$$

$$d^2 = 0. \tag{3.7}$$

In local coordinates $(V; x_1, \ldots, x_n)$ on Q, if $f \in D(Q) = \Gamma^o(Q)$, on has $df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$ and by (3.6) if $\sigma^k = \sum_{i_1 < \ldots < i_k} S_I dx_{i1} \wedge \ldots \wedge dx_{in}$, then

$$d\sigma^k = \sum_{i_1 < \ldots < i_k} dS_I \wedge dx_{i1} \wedge \ldots \wedge dx_{ik}. \tag{3.8}$$

The properties of **d** imply that if $f : Q_1 \longrightarrow Q_2$ is differentiable, then

$$d(f^* \sigma) = f^*(d\sigma) \qquad \text{for all} \qquad \sigma \in \Gamma^k Q_2. \tag{3.9}$$

The **interior product** of $\sigma \in \Gamma^k(Q)$, $k \geq 0$, by a vector field $X \in \mathcal{X}(Q)$ is the $(k-1)$-form $i(X)\sigma$ (also denoted by $X \rfloor \sigma$) such that

$$i(X)f = 0 \quad \text{if} \quad k = 0 \;\; (f \in \mathcal{D}(Q)); \tag{3.10}$$

$$i(X)\sigma = \sigma(X) \quad \text{if} \;\; k = 1; \tag{3.11}$$

$$i(X)\sigma(X_1, \ldots, X_{k-1}) = \sigma(X, X_1, \ldots, X_{k-1}) \quad \text{if} \;\; k > 1. \tag{3.12}$$

It can be shown that

$$
\begin{aligned}
i(X)(a\sigma_1 + b\sigma_2) &= ai(X)\sigma_1 + bi(X)\sigma_2, \quad a, b \in \mathbf{R}, \quad \sigma_1, \sigma_2 \in \Gamma^k(Q) & (3.13)\\
i(X)(\sigma^k \wedge \omega) &= [i(X)\sigma^k] \wedge \omega + (-1)^k \sigma^k \wedge [i(X)\omega]. & (3.14)
\end{aligned}
$$

3.1 Lie derivative of tensor fields

A **covariant tensor field of order** r on Q is a multilinear map

$$\Phi : \mathcal{X}(Q)_{\times \ldots \times} \mathcal{X}(Q) \longrightarrow \mathcal{D}(Q)$$

that is, Φ is $\mathcal{D}(Q)$-linear in each one of the r factors:

$$\Phi(Y_1, \ldots, fX + gY, \ldots, Y_r) = f\Phi(Y_1, \ldots, X, \ldots, Y_r) + g\Phi(Y_1, \ldots, Y, \ldots, Y_r)$$

for all $X, Y \in \mathcal{X}(Q)$ and $f, g \in \mathcal{D}(Q)$.

In an analogous way one defines a **contravariant tensor field** using $\Gamma^1 Q$ instead of $\mathcal{X}(Q)$. Also one defines **mixed tensor field**of type (r, s) as a $\mathcal{D}(Q)$-multilinear map $\Phi : (\Gamma^1(Q))^r \times (\mathcal{X}(Q))^s \longrightarrow \mathcal{D}(Q)$, and so $\Phi(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s) \in \mathcal{D}(Q)$. One says that the one form σ^i occupies the ith contravariant slot and that the vector field Y_j occupies the jth covariant slot of Φ.

Exercise 3.3 Show that the covariant tensor fields of order 1 are naturally identified with the elements of $\Gamma^1(Q)$ and the contravariant tensor fields of order 1 are identified with vector fields.

Exercise 4.3 Define the **contraction** $C^i_j(\Phi)$ of a tensor field Φ of type (r, s) which is a tensor field of type $(r-1, s-1)$. **Hint:** Start defining the contraction $C(A)$ of a tensor field of type $(1, 1)$ as a function on Q, using local coordinates $x^1, \ldots, x^n$, by $C(A) = \sum_{i=1}^n A(dx^i, \frac{\partial}{\partial x^i})$, and show that the definition does not depend on the used coordinates. Follow defining $[C^i_j(\Phi)](\sigma^1, \ldots, \sigma^{r-1}, Y_1, \ldots, Y_{s-1})$ as the contraction $C(A)$ of the following tensor field A of type (1,1):

$$A : (\theta, X) \longmapsto \Phi(\sigma^1, \ldots, \theta, \ldots, \sigma^{r-1}, Y_1, \ldots, X, \ldots, Y_{s-1})$$

where θ occupies the ith contravariant slot and X occupies the jth covariant slot of Φ.

Let $X \in \mathcal{X}(Q)$ be a vector field on Q. The local flow X_t of X is a one-parameter group of diffeomorphisms that act in a neighborhood V of a point $y \in Q$, for $|t| < \epsilon$, $\epsilon > 0$ small. Given a **covariant tensor field** Φ one can compute the derivative of Φ along integral curves of X; in other words, the diffeomorphism X_t induces a map $X_t{}^*$ that acts on covariant tensor fields in V. So $X_t{}^*\Phi(X_t(y))$ is a tensor at y then $X_t{}^*\Phi(X_t(y)) - \Phi(y)$ makes sense. The **Lie derivative of** Φ is another tensor field $L_X\Phi$ of the same type (also denoted by $\theta(X)\Phi$) defined by

$$L_X\Phi(y) \stackrel{def}{=} \lim_{t\to 0} \frac{1}{t}\left[X_t{}^*\Phi(X_t(y)) - \Phi(y)\right] = \frac{d}{dt}X_t{}^*\Phi(X_t(y))|_{t=0}. \tag{3.15}$$

Let us see some properties of the Lie derivative:

$$L_X(a\Phi_1 + b\Phi_2) = aL_X\Phi_1 + bL_X\phi_2, \quad a, b \in \mathbb{R} \tag{3.16}$$

$$\text{If} \quad f \quad \text{is a diffeomorphism of} \quad Q, \quad L_X(f^*\Phi) = f^*L_X\Phi. \tag{3.17}$$

When B is a bilinear map of tensors, that is, if $B(\Phi_1, \Phi_2)$ is a tensor that depends linearly on Φ_1 and Φ_2, then

$$L_X B(\Phi_1, \Phi_2) = B(\Phi_1, L_X\Phi_2) + B(L_X\Phi_1, \Phi_2). \tag{3.18}$$

In particular:

$$L_X(\sigma^k \wedge \omega^\ell) = (L_X\sigma^k) \wedge \omega^\ell + \sigma^k \wedge L_X\omega^\ell. \tag{3.19}$$

For the case in which $\Phi = f$ is a function ($f \in \mathcal{D}(Q)$) one has:

$$L_X f = X(f). \tag{3.20}$$

Let $\Phi = \sigma^k$ be an exterior differential k-form and take a local system of coordinates $(V; y_1, \ldots, y_n)$ for Q where $X = \sum_{i=1}^n X^i \frac{\partial}{\partial y_i}$ and the local diffeomorphism X_t has the components

$$X_t(y_1, \ldots, y_n) = (X_t{}^1(y_1, \ldots, y_n), \ldots, X_t{}^n(y_1, \ldots, y_n)).$$

Each function $X_t{}^i(y_1, \ldots, y_n)$ depends, in a differentiable way, on t and $y_1, \ldots, y_n$, so $X_t{}^i(y_1, \ldots, y_n) = f^i(t, y_1, \ldots, y_n)$ and $\frac{\partial^2 f^i}{\partial t \partial y_j} = \frac{\partial^2 f^i}{\partial y_j \partial t}$. Thus, for $\sigma^k = dy_i$ on has:

$$X_t{}^*dy_i = dX_t{}^i = \sum_{j=1}^n \frac{\partial X_t^i}{\partial y_j} dy_j, \quad \text{and}$$

$$L_X dy_i = \frac{d}{dt}(X_t{}^*dy_i)\,|_{t=0} = \sum_{j=1}^n \left(\frac{d}{dt}\frac{\partial X_t{}^i}{\partial y_j}\right)|_{t=0}\; dy_j = \sum_{j=1}^n \frac{\partial X^i}{\partial y_j} dy_j. \tag{3.21}$$

Now, if $\Phi = \sigma^k$ is a one form σ, locally given by:

$$\sigma = \sum_{i=1}^{n} S_i dy_i = \sum_{i=1}^{n} S_i(y_1, \ldots, y_n) dy_i,$$

applying the properties of the Lie derivative and the fact that $(S, dy) \in \Gamma^o(Q) \times \Gamma^1(Q) \to S dy \in \Gamma^1(Q)$ is a bilinear map of tensors, then

$$\begin{aligned} L_X \sigma &= \sum_{i=1}^{n} L_X(S_i dy_i) = \sum_{i=1}^{n} (L_X S_i)\, dy_i + (S_i L_X dy_i) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X^j \frac{\partial S_i}{\partial y_j} \right) dy_i + \sum_{i=1}^{n} S_i \left(\sum_{j=1}^{n} \frac{\partial X^i}{\partial y_j} dy_j \right); \end{aligned}$$

so

$$L_X \sigma = \sum_{i,j} \left(X^j \frac{\partial S_i}{\partial y_j} + S_j \frac{\partial X^j}{\partial y_i} \right) dy_i. \tag{3.22}$$

Finally, $\phi = \sigma^k = \sum_{i_1 < \ldots < i_k} S_I dy_{i_1} \wedge \ldots \wedge dy_{i_k}$ being an exterior differential k-form, one uses (3.16) and (3.19) and $L_X \sigma^k$ is obtained under the usual rules.

Exercise 5.3 Try to define the Lie derivative of a contravariant tensor field.

Let now $\Phi = Y \in \mathcal{X}(Q)$ be a vector field on Q and let us show that $L_X Y = [X, Y]$, that is, $L_X Y$ is precisely the Lie bracket $[X, Y]$ introduced above. In fact, in local coordinates $(V, y_1, \ldots, y_n)$ one can write $X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial y_i}$ and $Y = \sum_{i=1}^{n} Y^i \frac{\partial}{\partial y_i}$. We start computing $L_X \left(\frac{\partial}{\partial y_i} \right) = \frac{d}{dt} X_{-t*} \left(\frac{\partial}{\partial y_i} \right) |_{t=0}$. The j-component of the vector field $X_{-t*} \frac{\partial}{\partial y_i}$ at $p \in Q$ is

$$\begin{aligned} & dy_j(p) \left[X_{-t*} \frac{\partial}{\partial y_i}(X_{-t}(p)) \right] = dX^j_{-t} \left[\frac{\partial}{\partial y_i}(X_{-t}(p)) \right] \\ = \quad & \frac{\partial X^j_{-t}}{\partial y_i}(X_{-t}(p)) \end{aligned}$$

so that,

$$L_X \left(\frac{\partial}{\partial y_i} \right) = \sum_{j=1}^{n} \frac{d}{dt} \left[\frac{\partial X^j_{-t}}{\partial y_i}(X_{-t}(p)) \right]_{t=0} \frac{\partial}{\partial y_j} = \sum_{j=1}^{n} - \frac{\partial X^j}{\partial y_i} \frac{\partial}{\partial y_j},$$

and

$$\begin{aligned} L_X(Y) &= L_X\left(\sum_{i=1}^{n} Y^i \frac{\partial}{\partial y_i}\right) = \sum_{i=1}^{n}\left((L_X Y^i)\left(\frac{\partial}{\partial y_i}\right) + Y^i L_X\left(\frac{\partial}{\partial y_i}\right)\right) = \\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} X^j \frac{\partial Y^i}{\partial y_j}\frac{\partial}{\partial y_i} + Y^i \sum_{j=1}^{n}\left(-\frac{\partial X^j}{\partial y_i}\right)\frac{\partial}{\partial y_j}\right), \end{aligned}$$

so,

$$L_X(Y) = \sum_{i,j=1}^{n}\left(X^j \frac{\partial Y^i}{\partial y_j} - Y^j \frac{\partial X^i}{\partial y_j}\right)\frac{\partial}{\partial y_i} = [X, Y]. \quad \blacksquare \tag{3.23}$$

The next formulae (3.24), (3.25), (3.26), (3.27) and (3.28) are useful and relate the notions of exterior derivative, interior product and Lie derivative.

$$L_X(d\sigma) = dL_X\sigma, \quad \sigma \in \Gamma(Q); \tag{3.24}$$

this follows because $X_t{}^* d\sigma = dX_t{}^*\sigma$.

$$L_X[i(Y)\sigma] = i[X, Y]\sigma + i(Y)L_X\sigma; \tag{3.25}$$

it is enough to observe that $i(Y)\sigma$ is bilinear in Y and σ, so by (3.18) we have

$$L_X[i(Y)\sigma] = i(L_X Y)\sigma + i(Y)L_X\sigma,$$

and by (3.23) formula (3.25) follows.

3.2 The H. Cartan formula

Consider now the so called **H. Cartan formula**:

$$L_X\sigma = i(X)d\sigma + d[i(X)\sigma], \quad \sigma \in \Gamma(Q). \tag{3.26}$$

To prove (3.26) one remarks that the second member of this last formula is a derivation on the algebra $\Gamma(Q)$ and then, it is enough to show that the equality holds when applied to functions and 1-forms.

If $f \in \Gamma^o(Q) = \mathcal{D}(Q)$ and since by (3.10) $i(X)f = 0$, formula (3.26) reduces, due to (3.11), to

$$L_X f = df(X) = i(X)df \tag{3.27}$$

If $\sigma \in \Gamma^1(Q)$, σ is locally the sum $\sigma = \sum_{i=1}^n S_i dx_i$ and so, it is enough to prove now formula (3.26) for the 1-forms of type $g.df$, $f, g \in \mathcal{D}(Q)$.

We have $L_X(g.df) = (L_X g)df + gL_X df$ by (3.19), and

$$L_X(g.df) = X(g)df + g.d(i(X)df) \tag{3.28}$$

by (3.20), (3.24) and (3.27). On the other hand, (3.6) and (3.7) imply

$$i(X)d(g.df) + d(i(X)(g.df)) = i(X)(dg \wedge df) + d(i(X)(g.df));$$

by (3.14), (3.11) and (3.10) one obtains

$$i(X)d(g.df) + d(i(X)(g.df)) = X(g)df - X(f).dg + d(g.df(X)),$$

so, by (3.6) one has

$$\begin{aligned} i(X)d(g.df) + d(i(X)(g.df)) &= X(g)df - X(f).dg + dg.X(f) + g.d(X(f)) \\ &= X(g)df + g.d(i(X)df), \end{aligned}$$

thus, by (3.28) we finally have

$$L_X(g.df) = i(X)d(g.df) + d(i(X)(g.df)),$$

and (3.26) is proved. ∎

Exercise 6.3 Prove that $L_{[X,Y]} = [L_X, L_Y]$ where, as usual, the right-hand side means $L_X L_Y - L_Y L_X$.

Exercise 7.3 In the following two examples, which one is a tensor field:

i) $(\sigma, X) \in \Gamma^1(Q) \times \mathcal{X}(Q) \longmapsto \sigma(X) \in \mathcal{D}(Q)$

ii) $(X, Y) \in \mathcal{X}(Q) \times \mathcal{X}(Q) \longmapsto X(\theta(Y)) \in \mathcal{D}(Q)$
where $\theta \in \Gamma^1(Q)$ is a fixed one form.

Exercise 8.3 Show that if σ is a k-form and $X_0, X_1, \ldots, X_k$, are vector-fields one has

$$\begin{aligned} d\sigma(X_0, X_1, \ldots, X_k) = \sum_{i=0}^{k} (-1)^i X_i(\sigma(X_0, \ldots, \hat{X}_i, \ldots, X_k)) + \\ + \sum_{i<j} (-1)^{i+j} \sigma([X_i X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k). \end{aligned} \tag{3.29}$$

$$L_X\sigma(X_1, \ldots, X_k) = X(\sigma(X_1, \ldots, X_k)) - \sum_{i=1}^{k} \sigma(X_1, \ldots, [X, X_i], \ldots, X_k). \tag{3.30}$$

Chapter 4

Pseudo-riemannian manifolds

A **pseudo-riemannian metric** on a differentiable manifold Q is a law that to each point $y \in Q$ associates a non degenerate symmetric bilinear form $\langle , \rangle_y$ on the tangent space T_yQ, varying smoothly, that is, given a local system of coordinates $(V; x_1, \ldots, x_n)$, $y \in V$, and considered the local vector fields $\frac{\partial}{\partial x_i}, i = 1, \ldots, n$, the functions $g_{ij} : V \to \mathbf{R}$ defined by $g_{ij} = \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle$ are differentiable. The $n \times n$ matrix (g_{ij}) is symmetric and $\langle , \rangle_y$ be non degenerate means *det* $g_{ij}(y) \neq 0$ for all $y \in V$. If the pseudo riemannian metric is such that $\langle , \rangle_y$ is positive definite for all $y \in Q$ we say that the law $\langle , \rangle : y \mapsto \langle , \rangle_y$ is a **riemannian metric** on Q. In both cases we use to say that $\langle , \rangle$ is simply a **metric**.

A **pseudo-riemannian (riemannian) manifold** is a pair $(Q, \langle , \rangle)$ where $\langle , \rangle$ is a pseudo-riemannian (riemannian) metric on a differentiable manifold Q. If one computes the composition $g_{ij} \circ \varphi$ of g_{ij} with the local chart $\varphi : U \to V$, one obtains $g_{ij} \circ \varphi(x_1, \ldots, x_n)$ or simply $g_{ij}(x_1, \ldots, x_n)$.

Given two pseudo-riemannian (riemannian) manifolds $(Q_1, \langle , \rangle_1), (Q_2, \langle , \rangle_2)$ and a diffeomorphism $f : Q_1 \to Q_2$ such that $\langle u_y, v_y \rangle_1 = \langle f_* u_y, f_* v_y \rangle_2$ for all $y \in Q_1$ and $u_y, v_y \in T_yQ_1$, then f is said to be a **pseudo-riemannian (riemannian) isometry**.

Exercise 1.4 Show that the product of two pseudo-riemannian (riemannian) manifolds is a pseudo-riemannian (riemannian) manifold.

Examples: 1.4 Any submanifold $Q \subset \mathbf{R}^N$ has a riemannian metric induced by $\mathbf{R}^N$ with its usual inner product. The **flat torus** is the manifold $S^1 \times \ldots \times S^1$

with the product metric, provided that $S^1 \subset \mathbf{R}^2$ has the induced metric.

2.4 A **Lie group** is a group G with a structure of differentiable manifold such that the map

$$(x, y) \in G \times G \mapsto xy^{-1} \in G$$

is differentiable. The left and right translations L_x, R_x by an element $x \in G$ are the diffeomorphisms of G given by $L_x(y) = xy$ and $R_x y = yx$, respectively. A pseudo-riemannian (or riemannian) metric on G is said to be **left invariant** if L_x is an isometry for all $x \in G$. An analogous definition for **right invariant** metric can be introduced using R_x instead L_x. The left invariant metrics on G are obtained if one introduces at T_eG (e is the identity of G) a non degenerate bilinear form $\langle , \rangle$ and defines $\langle , \rangle_x$ for any $x \in G$ using $L_{x^{-1}}$, that is

$$\langle u, v \rangle_x \stackrel{def}{=} \langle d(L_{x^{-1}})(x)u, d(L_{x^{-1}})(x)v \rangle$$

for all $u, v \in T_xG$.

3.4 - Immersed pseudo-riemannian (riemannian) manifolds

Let $f : N \to Q$ be an immersion that is, f is differentiable and $f_*(p) : T_pN \to T_{f(p)}Q$ is injective for all $p \in N$ (this implies that $\dim N \leq \dim Q$). If Q has a pseudo-riemannian (riemannian) structure $\langle , \rangle$, f induces on N a pseudo-riemannian (riemannian) structure by the formula

$$\ll u, v \gg_p \stackrel{def}{=} \langle f_*u, f_*v \rangle, \qquad \text{for all} \quad p \in N$$

and all $u, v \in T_pN$. It is easy to see that f_* injective implies that $\ll , \gg_p$ is non-degenerate, (positive definite), so $\ll , \gg$ is a pseudo-riemannian (riemannian) metric on N.

Let Q be a C^∞ manifold, $\tau : TQ \to Q$ the canonical differentiable projection and let $c : I \to Q$ ($I \subset \mathbf{R}$ an open interval) be a differentiable curve (not necessarily injective). A **vector field** V **along a differentiable curve** $c : I \to Q$ is a map $V : I \to TQ$ such that to each $t \in I$ corresponds $V(t) \in T_{c(t)}Q$, that is, $\tau \circ V = c$. V is said to be **differentiable** if the map $V : t \in I \to V(t) \in TQ$ is differentiable. This means that given (in Q) any coordinate neighborhood $(\Omega; x_1, \ldots, x_n)$ and any $t_o \in I$ such that $c(t_o) \in \Omega$, we have $V(t) = \sum_{i=1}^n a_i(t) \frac{\partial}{\partial x_i}(c(t))$ for t in a neighborhood of t_o with the $a_i(t)$ being differentiable functions. The vector field $\dot{c} = \frac{dc}{dt} \stackrel{def}{=} c_*(\frac{d}{dt})$ is called the **velocity field** or the **tangent field** of $c = c(t)$.

When $c : I \to Q$ is of class C^2, the velocity field $\dot{c}$ is of class C^1. A **segment** is the restriction of a C^1 curve $c : I \to Q$ to a closed interval $[a, b] \subset I$. It is possible to compute the **length of a segment**, provided that $(Q, \langle , \rangle)$ is a pseudo-riemannian structure:

$$\text{length of} \quad (c \mid [a,b]) = l_a{}^b(c) \stackrel{def}{=} \int_a^b \langle \frac{dc}{dt}, \frac{dc}{dt} \rangle^{1/2} dt.$$

Remark that the integral above makes sense because $t \in [a,b] \to \langle \frac{dc}{dt}, \frac{dc}{dt} \rangle^{1/2} \in \mathbf{R}$ is a continuous map.

Example 4.4 Recall that the torus T^k in Example 10.2 is the quotient $\mathbf{R}^k/Z^k$ and that $\mathbf{R}^k/Z^k$ is diffeomorphic to $S^1 \times \ldots \times S^1$; so, the quotient map corresponds to the natural projection $\pi(x_1, \ldots, x_k) = (e^{ix_1}, \ldots, e^{ix_k})$ which is a local isometry from $\mathbf{R}^k$ onto the manifold $\mathbf{R}^k/Z^k$ with a suitable riemannian structure. One can show that T^k with that structure and the flat torus $S^1 \times \ldots \times S^1$ are isometric riemannian manifolds.

Example 5.4 The flat torus $T^2 = S^1 \times S^1$ and the torus of revolution $\tilde{T}^2 \subset \mathbf{R}^3$ (see example 11.2) with the induced metric are **not** isometric riemannian manifolds. Why?

Example 6.4 Let $\mathbf{R}$ be considered as an affine space and G be the Lie group of all proper affine transformations, that is, $g \in G$ means that $g : \mathbf{R} \to \mathbf{R}$ is given by

$$g(t) = yt + x \qquad \text{for all} \quad t \in \mathbf{R},$$

with $y > 0$ and $x \in \mathbf{R}$ being fixed numbers. So G, as a differentiable manifold, can be identified with the set

$$\{(x,y) \in \mathbf{R}^2 \mid y > 0\},$$

with the differentiable structure induced by $\mathbf{R}^2$. The left invariant riemannian metric on G that at the identity $e(e(t) = t$ for all $t \in \mathbf{R}$ or $e = (0,1))$ of the group G is the usual metric (given by $\bar{g}_{11} = \bar{g}_{22} = 1, \bar{g}_{12} = 0$), is defined by $g_{11} = g_{22} = \frac{1}{y^2}$ and $g_{12} = 0$. That metric (g_{ij}) is the riemannian metric of the non euclidean geometry of Lobatchevski.

4.1 Affine connections

Let Q be a C^∞ differentiable manifold, $\mathcal{X}(Q)$ be the set of C^∞ vector fields on Q and $\mathcal{D}(Q)$ be the collection of all real valued C^∞ functions defined on Q. An **affine connection** on Q is a map

$$\nabla : \mathcal{X}(Q) \times \mathcal{X}(Q) \longrightarrow \mathcal{X}(Q)$$

(one denotes $\nabla(X,Y) \overset{def}{=} \nabla_X Y$) such that

$$\begin{aligned}
\nabla_{fX+gY} Z &= f\nabla_X Z + g\nabla_Y Z, && (4.1)\\
\nabla_X(Y+Z) &= \nabla_X Y + \nabla_X Z, && (4.2)\\
\nabla_X(fY) &= f\nabla_X Y + X(f)Y, && (4.3)
\end{aligned}$$

for all $X, Y, Z \in \mathcal{X}(Q)$ and all $f, g \in \mathcal{D}(D)$.

Proposition 1.4 *Let ∇ be an affine connection on a C^∞-manifold Q. Then:*

i) *If X or Y is zero on an open set Ω of Q then $\nabla_X Y = 0$ in Ω.*

ii) *If $X, Y \in \mathcal{X}(Q)$ and $p \in Q$, then $(\nabla_X Y)(p)$ depends on the value $X(p)$ and on the values of Y along a curve tangent to X at p, only.*

iii) *If $X(p) = 0$ then $(\nabla_X Y)(p) = 0$.*

Proof To prove i) when $X = 0$, one uses (4.1) making $f = g = 0$ and $Z = Y$; when $Y = 0$ one uses (4.3) making $f = 0$. (In particular ∇ defines on the manifold Ω an affine connection $\tilde{\nabla}$; if $X; Y$ are vector fields on Ω we extend them to $\bar{X}, \bar{Y} \in \mathcal{X}(Q)$ and define $\tilde{\nabla}_Y X$ as the restriction to Ω of $\nabla_{\bar{Y}} \bar{X}$. It follows from i) that $\tilde{\nabla}_Y X$ does not depend on the extensions chosen. To simplify the notation, $\tilde{\nabla}_Y X$ is also denoted by $\nabla_Y X$. To prove ii) we write in a coordinate neighborhood $(\Omega; x_1, \ldots, x_n)$:

$$X = \sum_{j=1}^{n} a_j \frac{\partial}{\partial x_j}, \qquad Y = \sum_{i=1}^{n} b_i \frac{\partial}{\partial x_i};$$

using (4.1), (4.2), and (4.3) one obtains locally:

$$\begin{aligned}
\nabla_X Y &= \nabla_{\sum_j a_j \frac{\partial}{\partial x_j}} \left(\sum_i b_i \frac{\partial}{\partial x_i} \right) = \sum_j a_j \left[\nabla_{\frac{\partial}{\partial x_j}} \left(\sum_i b_i \frac{\partial}{\partial x_i} \right) \right] \\
&= \sum_j a_j \left[\sum_i b_i \nabla_{\frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i} + \sum_i \frac{\partial b_i}{\partial x_j} \frac{\partial}{\partial x_i} \right];
\end{aligned}$$

One denotes,

$$\nabla_{\frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_i} \overset{def}{=} \sum_k \Gamma_{ji}^k \frac{\partial}{\partial x_k} \tag{4.4}$$

where the functions $\Gamma_{ji}^k(x_1, \ldots, x_n)$ are called the **Christoffel symbols** of the connection ∇, relative to the coordinate neighborhood $(\Omega; x_1, \ldots, x_n)$. So, we have:

$$\nabla_X Y = \sum_k \left[\sum_{i,j} a_j b_i \Gamma_{ji}^k + \sum_j a_j \frac{\partial b_k}{\partial x_j} \right] \frac{\partial}{\partial x_k}$$

$$= \sum_k \left[\sum_{i,j} a_j b_i \Gamma_{ji}^k + X(b_k) \right] \frac{\partial}{\partial x_k}; \tag{4.5}$$

Then:

$$(\nabla_X Y)(p) = \sum_k \left[\sum_{i,j} a_j(p) b_i(p) \Gamma_{ji}{}^k(p) + X(b_k)(p) \right] \frac{\partial}{\partial x_k}(p) \tag{4.6}$$

where $X(b_k)(p) = (X(p))(b_k) = db_k(p)[X(p)]$.

Formula (4.6) shows that $(\nabla_X Y)(p)$ depends on the values $a_j(p)$ (value of $X(p)$ in the chose coordinates) and on $(X(p))(b_k)$ only; but $(X(p))(b_k)$ depends on the values of Y along a curve tangent to the vector field X at p, only. That proves ii). The expression (4.6) also proves iii). ∎

Proposition 2.4 *Let Q be a C^∞ differentiable manifold with an affine connection ∇. So, there exists a unique law $\frac{D}{dt}$ that to each differentiable vector field V along a differentiable curve $c : I \to Q$ ($I \subset \mathbf{R}$ an open interval) associates another vector field $\frac{DV}{dt}$ along c, called the covariant derivative of V along c, such that:*

a_1) $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$.

a_2) $\frac{D}{dt}(fV) = (\frac{df}{dt})V + f\frac{DV}{dt}$, *where V and W are differentiable vector fields along c and $f \in \mathcal{D}(I)$.*

a_3) *If V is induced by a vector field $Y \in \mathcal{X}(Q)$, that is, $V(t) = Y(c(t))$, then $\frac{DV}{dt} = \nabla_{\dot c} Y$, where $\dot c$ is the velocity field of c.*

Remarks: In the last condition a_3), the expression $\nabla_{\dot c} Y$ makes sense by condition ii) of proposition 1.4; in fact $\nabla_{\dot c} Y$ is a tangent vector to the manifold Q at the point $c(t)$. When $c(t) \equiv c_o \in Q$, $\frac{DV}{dt}$ is the usual derivative on $T_{c_o}Q$.

Proof Assume there exists such a law verifying $a_1), a_2)$ and $a_3)$. Let us assume also that in a coordinate neighborhood $(\Omega; x_1, \ldots, x_n)$ of Q, the local expressions of $V = V(t)$ and $c(t)$ are

$$V(t) = \sum_i v_i(t) \frac{\partial}{\partial x_i}(c(t)) \quad \text{and} \quad c(t) = (x_1(t), \ldots, x_n(t)),$$

for all t in a suitable interval contained in I where $v_i(t)$ and $x_i(t)$ are differentiable functions. Using a_1) and a_2) we may write:

$$\begin{aligned}\frac{DV}{dt} &= \frac{D}{dt}\left(\sum_i v_i(t)\frac{\partial}{\partial x_i}(c(t))\right)\\ &= \sum_i \frac{D}{dt}\left[v_i(t)\frac{\partial}{\partial x_i}(c(t))\right]\\ &= \sum_i \left[\frac{dv_i(t)}{dt}\frac{\partial}{\partial x_i}(c(t)) + v_i(t)\frac{D}{dt}\frac{\partial}{\partial x_i}(c(t))\right];\end{aligned}$$

using a_3) we have that

$$\frac{D}{dt}\frac{\partial}{\partial x_i}(c(t)) = \nabla_{\dot{c}}\frac{\partial}{\partial x_i} = \nabla_{\sum_j \dot{x}_j(t)\frac{\partial}{\partial x_j}(c(t))}\frac{\partial}{\partial x_i} = \sum_j \dot{x}_j(t)\left(\nabla_{\frac{\partial}{\partial x_j}}\frac{\partial}{\partial x_i}\right)c(t),$$

and so

$$\frac{DV}{dt} = \sum_i \left[\frac{dv_i(t)}{dt}\frac{\partial}{\partial x_i}(c(t)) + v_i(t)\sum_j \dot{x}_j(t)\left(\nabla_{\frac{\partial}{\partial x_j}}\frac{\partial}{\partial x_i}\right)(c(t))\right] \tag{4.7}$$

Last formula (4.7) shows that $\frac{DV}{dt}$ is uniquely determined because the right hand side depends on the curve $c = c(t)$, on $V = V(t)$ and on ∇, through $(\nabla_{\frac{\partial}{\partial x_j}}\frac{\partial}{\partial x_i})(c(t))$, only. To show the existence of the law, one uses, in the same coordinate neighborhood, the expression (4.7) to define $\frac{D}{dt}$ and verify that $\frac{D}{dt}$ has the desired properties a_1), a_2) and a_3). If we take another coordinate neighborhood $(\tilde{\Omega}; \tilde{x}_1, \ldots, \tilde{x}_n)$ on Q such that $\Omega \cap \tilde{\Omega} \neq \phi$, one defines, analogously, $\frac{DV}{dt}$ on $\tilde{\Omega}$ using (4.7); clearly the two definitions coincide on $\Omega \cap \tilde{\Omega}$ due to the uniqueness of $\frac{DV}{dt}$ on Ω. This way $\frac{DV}{dt}$ can be extended to the entire manifold Q, using an atlas. ∎

Given an affine connection ∇ on a differentiable manifold and a differentiable vector field $V = V(t)$ along a differentiable curve $c : t \in I \mapsto c(t) \in Q$, one says that V is **parallel along** c if $\frac{DV}{dt} = 0$.

Proposition 3.4 (Parallel translation) *Let Q be a C^∞ differentiable manifold with an affine connection ∇, $c = c(t)$ a differentiable curve on Q and $V_o \in T_{c(t_o)}Q$ a tangent vector to Q at the point $c(t_o)$ of the curve. Then there exists a unique parallel vector field V along c such that $V(t_o) = V_o$.*

Proof If $c : I \subset \mathbf{R} \longrightarrow Q$ is the given differentiable curve, let $[t_o, t_1] \subset I$ be a closed interval (then compact) and assume that the compact image $c([t_o, t_1]) \subset$

Q is covered by a finite number of coordinate neighborhoods $(\Omega; x_1, \ldots, x_n)$. For simplicity let us suppose that $c([t_o, t_1]) \subset \Omega$. Using (4.4) and (4.7) we have that $\frac{DV}{dt} = 0$ on $[t_o, t_1]$ if, and only if,

$$\frac{dv_k(t)}{dt} + \sum_{i,j} v_i(t)\dot{x}_j(t)\Gamma^k_{ji}(c(t)) = 0, \qquad k = 1, \ldots, n. \tag{4.8}$$

The last equations (4.8) is a system of ordinary differential equations in the unknowns $v_k(t), k = 1, \ldots, n$ ($\dot{x}_j(t)$ and $\Gamma^k_{ji}(c(t))$ are given functions of t). Since that system is linear with coefficients given by continuous functions defined on the interval $[t_o, t_1]$, it is well known that it has a unique solution $(v_k(t))$ defined on $[t_o, t_1]$ provided that $(v_k(t_o))$ are given. In the present case one can make $v_k(t_o)$ equal to the k-component of V_o, that is,

$$V_o = \sum_k v_k(t_o)(\frac{\partial}{\partial x_k})(c(t_o)).$$

So, the vector field along c, defined by

$$V = V(t) \stackrel{def}{=} \sum_k v_k(t)\frac{\partial}{\partial x_k}(c(t)), \qquad t \in [t_o, t_1]$$

is parallel along c and it is clearly unique. The general case in which $c([t_o, t_1])$ has to be covered by a finite number of local coordinate neighborhoods can be easily formalized. ∎

A **geodesic** of an affine connection ∇ on Q is a differentiable curve $c = c(t)$ on Q such that the corresponding velocity field $V = \dot{c}(t)$ is parallel along c, i.e. $\frac{D\dot{c}}{dt} = 0$ for all t. In local coordinates, $c(t) = (x_1(t), \ldots, x_n(t))$, the system of ordinary differential equations giving the geodesics is obtained from (4.8) making $v_k(t) = \dot{x}_k(t)$:

$$\ddot{x}_k(t) + \sum_{i,j} \dot{x}_i\dot{x}_j\Gamma^k{}_{ji}(x_1(t), \ldots, x_n(t)) = 0, \quad k = 1, \ldots, n. \tag{4.9}$$

Equations (4.9) show that the geodesics are at least of class C^2.

4.2 The Levi-Civita connection

Assume it is given a C^∞-pseudo-riemannian manifold $(Q, \langle,\rangle)$ with an affine connection ∇ on Q. We say that ∇ **is compatible with the metric** $\langle,\rangle$ if for any differentiable curve $c = c(t)$ on Q and all pair of parallel vector-fields $E_1(t), E_2(t)$ along c we have that

$$\langle E_1(t), E_2(t)\rangle = k \tag{4.10}$$

where k does not depend on t.

Proposition 4.4 *Let $(Q, \langle , \rangle)$ be a C^∞-pseudo-riemannian manifold with an affine connection ∇ on Q. Then ∇ is compatible with the metric $\langle , \rangle$ if, and only if, for any differentiable curve $c = c(t)$ and any two differentiable vector fields V and W along c we have:*

$$\frac{d}{dt}\langle V, W\rangle = \langle \frac{DV}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle. \tag{4.11}$$

Proof To see that (4.11) implies (4.10) it is enough to choose $V = E_1(t)$ and $W = E_2(t)$, both parallel along c so $\frac{DE_1(t)}{dt} = \frac{DE_2(t)}{dt} = 0$ and then $\frac{d}{dt}\langle E_1(t), E_2(t)\rangle = 0$ for all t, which implies (4.10). Conversely, assume (4.10) is true and consider an orthonormal basis $(E_1(t_o), \ldots, E_n(t_o))$ for $T_{c(t_o)}Q$ (see Exercise 2.4 below). Using Proposition 3.4 we obtain by parallel translation an orthonormal basis $(E_1(t), \ldots, E_n(t))$ for all t because $\langle E_i(t_o), E_j(t_o)\rangle = \epsilon_{ij}$ ($\epsilon_{ii} = +1$ or -1 and $\epsilon_{ij} = 0$ if $i \neq j$), and by (4.10) we also have $\langle E_i(t), E_j(t)\rangle = \epsilon_{ij}$. In particular V and W can be written as

$$V(t) = \sum_i v_i(t)E_i(t), \qquad W(t) = \sum_j w_j(t)E_j(t) \tag{4.12}$$

and then

$$\begin{aligned}\frac{d}{dt}\langle V, W\rangle &= \frac{d}{dt}\langle \sum_i v_i(t)E_i(t), \sum_j w_j(t)E_j(t)\rangle \\ &= \frac{d}{dt}\left(\sum_{i,j} v_i(t)w_j(t)\epsilon_{ij}\right) = \frac{d}{dt}\sum_i \epsilon_{ii}v_i(t)w_i(t).\end{aligned} \tag{4.13}$$

But

$$\begin{aligned}\frac{DV}{dt} &= \frac{D}{dt}\left(\sum_i v_i(t)E_i(t)\right) = \sum_i \dot{v}_i(t)E_i(t) + \sum_i v_i(t)\frac{DE_i}{dt}(t) \\ &= \sum_i \dot{v}_i E_i(t),\end{aligned} \tag{4.14}$$

because $\frac{DE_i(t)}{dt} = 0$; analogously,

$$\frac{DW}{dt} = \sum_j \dot{w}_j(t)E_j(t). \tag{4.15}$$

Substituting (4.12), (4.14) and (4.15) in the right hand side of (4.11) one obtains

$$\langle \frac{DV}{dt}, W\rangle + \langle V, \frac{DW}{dt}\rangle = \langle \sum_i \dot{v}_i(t)E_i(t), \sum_j w_j(t)E_j(t)\rangle +$$

$$+ \langle \sum_i v_i(t)E_i(t), \sum_j \dot{w}_j(t)E_j(t)\rangle = \frac{d}{dt}\sum_i \epsilon_{ii} v_i(t) w_i(t).$$

The last equality and (4.13) prove that (4.11) holds. ∎

Exercise 2.4 Show that any finite dimensional vector space with a non degenerate and symmetric bilinear form has an orthonormal basis. Give a counter-example showing that, in this case, is not true, in general, the Gram-Schmidt method used to obtain orthonormal basis relative to a positive definite symmetric bilinear form.

Exercise 3.4 Show that an affine connection ∇ on a pseudo-riemannian manifold $(Q, \langle,\rangle)$ is compatible with the metric $\langle,\rangle$ if, and only if, for any $X, Y, Z \in \mathcal{X}(Q)$ we have

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle. \tag{4.16}$$

Given two vector fields $X, Y \in \mathcal{X}(Q)$ one can construct $[X, Y]$ depending on Q only, and $\nabla_X Y - \nabla_Y X$ that depends on a given affine connection ∇ on a C^∞ differentiable manifold Q. We say that ∇ is **symmetric** if

$$\nabla_X Y - \nabla_Y X = [X, Y] \qquad \text{for all } X, Y \in \mathcal{X}(Q). \tag{4.17}$$

Exercise 4.4 Show that ∇ is symmetric if and only if for any coordinate neighborhood $(\Omega; x_1, \ldots, x_n)$ the corresponding Christoffel symbols (see 4.4) are symmetric, that is,

$$\Gamma^k_{ij} = \Gamma^k_{ji} \qquad i, j, k = 1, \ldots, n. \tag{4.18}$$

Proposition 5.4 (Levi-Civita) *Given a pseudo-riemannian metric $\langle,\rangle$ on a C^∞ differentiable manifold Q, there exists a unique affine connection ∇ on Q such that*

a) *∇ is symmetric;*

b) *∇ is compatible with the metric $\langle,\rangle$.*

Proof Let us define ∇ by the formula:

$$\begin{aligned} 2\langle\nabla_Y X, Z\rangle &= X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle Y,X\rangle \\ &- \langle[X,Z],Y\rangle - \langle[Y,Z],X\rangle - \langle[X,Y],Z\rangle \end{aligned} \tag{4.19}$$

for all $X, Y, Z \in \mathcal{X}(Q)$. Since $\langle,\rangle$ is non degenerate, $\nabla_Y X$ is well defined. Now it is a simple computation to show that ∇ is an affine connection and that (4.16) and (4.17) hold, so, there exists such a ∇. But conversely, given any affine connection ∇ satisfying *a*) and *b*) one can compute $X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle Y,X\rangle$ using (4.16) for the three terms of that expression; after this one uses (4.17) and see that $\nabla_Y X$ satisfies (4.19), that proves uniqueness. ∎

The affine connection given by Proposition 5.4 is called the **Levi-Civita connection associated to the pseudo-riemannian metric $\langle,\rangle$ on Q.**

Exercise 5.4 If ∇ is the Levi-Civita connection associated to the pseudo-riemannian metric $\langle,\rangle$ on a manifold Q and $(\Omega; x_1, \ldots, x_n)$ is a coordinate neighborhood, show that:

$$\Gamma_{ij}^m = \frac{1}{2}\sum_k \left[\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ki}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k}\right] g^{km} \tag{4.20}$$

where Γ_{ij}^m are the Christoffel symbols of ∇ relative to $(\Omega; x_1, \ldots, x_n)$, (see 4.4), $g_{ij} = \langle\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\rangle$ and (g^{km}) is the inverse matrix of the matrix (g_{km}).

Let $(Q, \langle,\rangle)$ be a smooth pseudo-riemannian manifold and ∇ be the Levi-Civita connection associated to the pseudo-riemannian metric $\langle,\rangle$. We saw that the geodesics of ∇ are the curves $c = c(t)$ such that the vector field $V = \dot{c}(t)$ is parallel along c, i.e. $\frac{D\dot{c}}{dt} = 0$ for all t. Locally $c(t) = (x_1(t), \ldots, x_n(t))$ and the $x_i(t)$ must satisfy (4.9), that is the system:

$$\ddot{x}_k(t) + \sum_{i,j} \dot{x}_i(t)\dot{x}_j(t){\Gamma^k}_{ij}(x_1(t), \ldots, x_n(t)) = 0,$$

$k = 1, \ldots, n$. We know that

$$\frac{d}{dt}\langle\dot{c}(t), \dot{c}(t)\rangle = 2\langle\frac{D}{dt}\dot{c}, \ \dot{c}\rangle = 0$$

so, the norm $|\dot{c}(t)|$ is constant. We will assume that $|\dot{c}(t)| = c_o \neq 0$, that is, we will exclude the geodesics given by constant functions (cannot be reduced to a point). The length of $c = c(t)$ from $\bar{t}$ to t is given by $s(t) = \int_{\bar{t}}^{t} |\dot{c}(u)|du = c_o(t - \bar{t})$; that shows that the parameter $(t - \bar{t})$ of any geodesic is proportional to the length from $\bar{t}$ to t.

The second order system of ordinary differential equations (4.9) defining the geodesics, can be written as a first order system:

$$\begin{cases} \dot{x}_k = v_k \\ \dot{v}_k = -\sum_{ij} \Gamma^k{}_{ij}(x_1, \ldots, x_n) v_i v_j. \end{cases} \tag{4.21}$$

So, in natural coordinates $(x_k, v_k), k = 1, \ldots, n$, of TQ, corresponding to the local system of coordinates $(x_1, \ldots, x_n)$ in Q, equations (4.21) describe the intrinsic condition $\frac{D\dot{c}}{dt} = 0$ and it is defined a vector-field on TQ

$$S : v_p \in TQ \mapsto S(v_p) \in T_{v_p}(TQ)$$

called the **geodesic flow** of the pseudo-riemannian metric $\langle , \rangle$; in the coordinates above we have

$$v_p = (x_k, v_k) \qquad \text{and} \qquad S(v_p) = ((x_k, v_k), (v_k, -\sum_{ij} \Gamma^k{}_{ij} v_i v_j)).$$

The trajectories of S are projected onto the geodesics by the canonical projection $\tau : TQ \to Q$; the condition $\dot{x}_k = v_k$ shows that the trajectories of S are, precisely, the curves $t \mapsto (c(t), \dot{c}(t)) \in TQ$, derivatives of the geodesics. By Exercise 5.4 we see that $\langle , \rangle \in C^k$, $k \geq 2$, implies that S is of class C^{k-1}.

The vector $S(v_p)$ can be also obtained through the **horizontal lifting operator** $H_{v_p} : w_p \in T_pQ \longrightarrow T_{v_p}(TQ)$ defined as follows. Take the geodesic $c(t)$ characterized by the conditions $c(0) = p, \dot{c}(0) = w_p$, and consider the curve $V(t)$ as the parallel transport of v_p along $c(t)$, that is, such that $V(0) = v_p$ and $\frac{DV(t)}{dt} = 0$. So, $H_{v_p}(w_p)$ is, by definition, the tangent vector at $t = 0$ to the curve $(c(t), V(t)) \in TQ$. We easily see that $d\tau(v_p)(H_{v_p} w_p) = w_p$ and that H_{v_p} is linear and injective. On the other hand $S(v_p) = H_{v_p}(v_p)$. The elements of $H_{v_p}(T_pQ) \subset T_{v_p}(TQ)$ are said to be **horizontal vectors at** v_p.

From the theory of ordinary differential equations applied to (4.21), one can state the following result:

Proposition 6.4 *Given a point $p \in Q$ one can find: an open set $\mathcal{U}$ in TQ, $\mathcal{U} \subset TV$, where $(V, x_1, \ldots, x_n)$ is a local system of coordinates around $p \in V$, $\mathcal{U}$ containing $O_p = (p, o) \in TV$, a number $\delta > 0$ and a differentiable map*

$$\Phi : (-\delta, \delta) \times \mathcal{U} \to TV,$$

such that $t \mapsto \Phi(t, q, v)$ is the unique trajectory of the geodesic flow S that verifies the initial condition $\Phi(o, q, v) = (q, v)$ for all $(q, v) \in \mathcal{U}$.

If we call $c = \tau \circ \Phi$, Proposition 6.4 implies the following

Proposition 7.4 *Given a point $p \in Q$, there exist an open set $V \subset Q$, $p \in V$, real numbers $\delta, \bar{\varepsilon} > 0$ and a differentiable map*

$$c : (-\delta, +\delta) \times \mathcal{U} \to Q,$$

$\mathcal{U}$ being the set $\mathcal{U} = \{(q,v) | q \in V, v \in T_qQ, |v| < \bar{\varepsilon}\}$, such that the curve $t \mapsto c(t,q,v)$, $t \in (-\delta, +\delta)$, is the unique geodesic of $(Q, \langle , \rangle)$ that passes through $q \in Q$ at the time $t = 0$ with velocity v, for all $q \in V$ and all $v \in T_qQ$ such that $|v| < \bar{\varepsilon}$.

It can be seen that it is possible to increase the initial velocity of a geodesic if one decreases, properly, the interval of definition, and conversely. In fact we have

Proposition 8.4 *If the geodesic $c = c(t,q,v)$ is defined for $t \in (-\delta, +\delta)$, then the geodesic $c(t,q,av), a > 0$, is defined in the interval $(-\delta/a, +\delta/a)$ and $c(t,q,av) = c(at,q,v)$.*

Proof Let $h : (-\delta/a, +\delta/a) \longrightarrow Q$ be the curve defined by $h(t) = c(at,q,v)$. It is clear that $h(0) = c(0,q,v) = q$ and that $\dot{h}(0) = a\dot{c}(0,q,v) = av$. Moreover, h is a geodesic because

$$\frac{D\dot{h}}{dt} = \nabla_{\frac{d}{dt}c(at,q,v)} \frac{d}{dt} c(at,q,v) = a^2 \nabla_{\dot{c}(at,q,v)} \dot{c}(at,q,v) = 0$$

where in ∇, $\frac{d}{dt}c(at,q,v)$ represents an extension of $\dot{h}$ to a neighborhood of $c(at,q,v)$, in Q. The uniqueness of geodesics gives finally:

$$h(t) = c(at,q,v) = c(t,q,av) \quad \text{for} \quad t \in (-\delta/a, +\delta/a).$$

From what was said above, one can define a local exponential map:

$$\exp : \mathcal{U} \longrightarrow Q, \qquad \text{given by}$$

$$\exp(q,v) = c(1,q,v) = c\left(|v|, q, \frac{v}{|v|}\right) \tag{4.22}$$

called the **exponential map** in $\mathcal{U}$, which is a differentiable map. If we fix $q \in Q$, one may consider $B_{\bar{\varepsilon}}(0) \subset \mathcal{U} \cap T_qQ$ where $B_{\bar{\varepsilon}}(0)$ is a ball centered at $0 \in T_qQ$ with a suitable radius $\bar{\varepsilon} > 0$, and define

$$\exp_q : B_{\bar{\varepsilon}}(0) \longrightarrow Q$$

by $exp_q(v) = exp(q,v), v \in B_{\bar{\varepsilon}}(0)$. One can see that $d(\exp_q)(0)$ is non singular because

$$\begin{aligned} d(\exp_q)(0)[v] &= \frac{d}{dt}\exp_q(tv) \mid_{t=0} = \frac{d}{dt} c(1,q,tv) \mid_{t=0} \\ &= \frac{d}{dt} c(t,q,v)|_{t=0} = v \end{aligned}$$

that is, $d(\exp_q)(0) = id\ T_qQ$. So, by the inverse function theorem, $\exp_q$ is a local diffeomorphism, that is, there exists $\varepsilon > 0$ and a ball $B_\varepsilon(0)$ centered at $0 \in T_qQ$ with radius $\varepsilon > 0$ such that the **exponential map at** $q \in Q$.

$$\exp_q : B_\varepsilon(0) \subset T_qQ \longrightarrow Q$$

is a diffeomorphism from $B_\varepsilon(0)$ onto an open neighborhood of q in Q. Denote by $B_\varepsilon(q)$ the set $B_\varepsilon(q) = \exp_q(B_\varepsilon(0))$ called a **normal ballor a geodesic ball** of center q and radius $\varepsilon > 0$. "Geometrically speaking", $\exp_q(v)$ is the point of Q obtained on the geodesic passing through $q \in Q$ at the time $t = 0$ with velocity $v/|v|$, after "walking" a length equal to $|v|$.

4.3 Tubular neighborhood

Let N be a manifold imbedded in Q, $n = \dim\ N < \dim\ Q$. Let E be the **normal bundle over** N, that is, E is the union $\cup_{x\in N} T_x{}^\perp N$ where $T_x{}^\perp N$ is the subspace of T_xQ orthogonal to T_xN in the metric $\langle,\rangle$ of the pseudo-riemannian manifold $(Q, \langle,\rangle)$. So we have a direct sum $T_x{}^\perp N \oplus T_xN = T_xQ$ for each point $x \in N$. The fiber bundle E is a submanifold of TQ. Let $\pi : E \to N$ be the restriction $\tau|E$. Note that E and Q have the same dimension.

A tubular neighborhood of N in Q is a diffeomorphism $f : Z \longrightarrow \Omega$ from an open neighborhood Z of the zero section in E onto an open set Ω in Q containing N and such that $f(0_x) = x$ for any zero vector $0_x \in E$, $x \in N$. The neighborhood Z is said to be a **tube** in E while $\Omega = f(Z)$ is a **tube** in Q. The composition $p = \pi \circ f^{-1} : \Omega \to N$ is a projection ($p^2 = p$) from the tube Ω onto N. In fact, given $y \in \Omega$, $y = f(\tilde{y}_x)$ with $\tilde{y}_x \in E$. So, $p(y) = \pi f^{-1}(y) = \pi\tilde{y}_x = x$ and $p^2(y) = p(x) = \pi f^{-1}(x) = \pi(0_x) = x$. It is also usual to call the pair (Ω, p) a tubular neighborhood of N in Q.

Proposition 9.4 (Tubular neighborhood) *Given a pseudo-riemannian manifold $(Q, \langle,\rangle)$, $Q \in C^\infty, \langle,\rangle \in C^k, k \geq 2$ and a submanifold $N \subset Q$ (N is imbedded in Q), $n = \dim\ N < \dim\ Q$, then there exists a tubular neighborhood $f : Z \longrightarrow \Omega$ of class C^{k-1} of N in Q.*

Proof To each $x \in Q$ one associates an open neighborhood $\mathcal{U} = \mathcal{U}(x)$ in TQ and a local exponential map $exp : \mathcal{U} \longrightarrow Q$ of class C^{k-1} (the class of the geodesic flow S). It is clear that in $\mathcal{D} = \cup_{x\in Q}\mathcal{U}(x)$ it is defined a **global exponential map**:

$$\exp : \ \mathcal{D} \longrightarrow Q$$

extending all the local exponential maps. Let $\tilde{Z} = \mathcal{D} \cap E$ which is an open neighborhood (inE) of the zero section of E. To each 0_x in this zero section, $f = exp|\tilde{Z}$ is a local diffeomorphism because its "vertical" derivative is the identity of $T_x{}^\perp N$ (restriction of $d\ exp_x(0) = id(T_xQ)$) and the "horizontal" derivative is also an isomorphism since the restriction of f to the zero section of E satisfies $f(0_x) = x$. The images in Q of all these local diffeomorphisms (that are restrictions of f) define a covering of N by open sets of the manifold Q. So, one can apply the results of Proposition 3.2 to the manifold N that has the induced topology of Q (N is a submanifold of Q) and obtain a covering of N by open sets V_i in Q such that to each i we have diffeomorphisms

$$f_i : Z_i \longrightarrow V_i \qquad \text{and} \qquad g_i : V_i \longrightarrow Z_i$$

(one is the inverse of the other) between V_i and open sets Z_i in $\tilde{Z}$, such that each Z_i contains a point 0_x of the zero section of E, with $x \in N$; moreover, the f_i act like identities when restricted to the zero section of E while the $g_i|N$ are also identities; but the f_i are restrictions to Z_i of the same map f. One can also obtain a locally finite covering $\{W_i\}$ of N by open sets in Q such that $\bar{W}_i \subset V_i$. Define $W = \cup_i W_i$ and denote by $\tilde{W}$ the set of all elements $y \in W$ such that, if y belongs to an intersection $\bar{W}_i \cap \bar{W}_j$, one has $g_i(y) = g_j(y)$. It is clear that $\tilde{W}$ contains N. Let us show that $\tilde{W}$ contains an open set of Q containing N. Take $x \in N$; there exists an open neighborhood G_x of x in Q that meets a finite number of the $\bar{W}_i$, only, say $\bar{W}_{i_1} \cap \ldots \cap \bar{W}_{i_r}$. Choosing G_x sufficiently small one can (not only) assume that x is in $\bar{W}_{i_1} \cap \ldots \cap \bar{W}_{i_r}$ and (also) that G_x is contained in each one of the sets $\bar{V}_{i_1}, \ldots, \bar{V}_{i_r}$. Since $x \in \bar{W}_{i_1} \cap \ldots \cap \bar{W}_{i_r}$ we have $G_x \subset [V_{i1} \cap \ldots \cap V_{i_r}]$ (because $\bar{W}_i \subset V_i$) and then the maps $g_{i1}, \ldots, g_{i_r}$ take the same value o_x at x. Since the $f_{i_1}, \ldots, f_{i_r}$ are restrictions of f, one concludes that the corresponding $g_{i_1}, \ldots, g_{i_r}$, have to agree at the points of G_x which can be reduced again to obtain $G_x \subset [\bar{W}_{i_1} \cap \ldots, \cap \bar{W}_{i_r}]$ that is, G_x is open and is contained in $\tilde{W}$. So we have

$$\tilde{W} \supset G \stackrel{def}{=} \cup_{x \in N} G_x.$$

The set G is open and one can define $g : G \longrightarrow g(G) \subset \tilde{Z}$ taking $g = g_i$ over $G \cap W_i$. The set $g(G)$ is open in $\tilde{Z}$ and the restriction of f to $f(G)$ is an inverse for g. We get, this way, a tubular neighborhood for N, $f : Z \longrightarrow \Omega$, where $Z = g(G)$ and $\Omega = G$. ∎.

Exercise 6.4 Show that in a normal ball $B_\epsilon(q) = \exp_q(B_\epsilon(o))$ there are coordinates $(x_1, \ldots, x_n)$ determined by an orthonormal basis $(e_1, \ldots, e_n)$ at T_qQ; that is, to each $\xi \in B_\epsilon(q)$ the coordinates $(x_1(\xi), \ldots, x_n(\xi))$ are given by $\exp_q^{-1}(\xi) = \sum_{i=1}^n x_i(\xi) e_i$. Prove that in these coordinates we have

$$g_{ij}(q) = \delta_{ij}\epsilon_j \qquad \text{where} \qquad \langle e_i, e_j \rangle = \delta_{ij}\epsilon_j, \quad \epsilon_j = +1 \text{ or } -1$$

and $\Gamma_{ij}^k(q) = 0$. $(B_\epsilon(q), x_1, \ldots, x_n)$ is called a **normal coordinate system**.

4.4 Curvature

Let $(Q, \langle , \rangle)$ be a pseudo-riemannian manifold and ∇ an affine connection. The function $R : \mathcal{X}(Q) \times \mathcal{X}(Q) \times \mathcal{X}(Q) \longrightarrow \mathcal{X}(Q)$ given by

$$
\begin{aligned}
R_{X,Y}Z &= \nabla_{[X,Y]}Z - [\nabla_X, \nabla_Y]Z \\
&= \nabla_{[X,Y]}Z - \nabla_X(\nabla_Y Z) + \nabla_Y(\nabla_X Z)
\end{aligned}
\tag{4.23}
$$

is called the **curvature tensor** on Q of the connection ∇. If ∇ is the Levi-Civita connection, R is said to be the **riemannian curvature tensor.**

In fact, to the map R corresponds the map

$$\bar{R} : (\sigma, X, Y, Z) \in \Gamma^1(Q) \times \mathcal{X}^3(Q) \longmapsto \sigma(R_{X,Y}Z) \in \mathcal{D}(Q)$$

and one has the following:

Proposition 11.4 *The map $\bar{R}$ is a mixed tensor field of type (1,3).*

Proof It is enough to show that the map $\bar{R}$ is $\mathcal{D}(Q)$-multilinear. But since the linearity in each variable is quite obvious, we only need to check that one can factor out functions. For instance:

$$
\begin{aligned}
R_{X,fY}Z &= \nabla_{[X,fY]}Z - \nabla_X(\nabla_{fY}Z) + \nabla_{fY}(\nabla_X Z) \\
&= (Xf)\nabla_Y Z + f\nabla_{[X,Y]}Z - \nabla_X(f\nabla_Y Z) + f\nabla_Y(\nabla_X Z) \\
&= fR_{X,Y}Z. \quad \blacksquare
\end{aligned}
$$

If we fix a point $p \in Q$ and take $x, y \in T_pQ$ one can also consider the so called **curvature operator**, the linear operator:

$$R_{xy} : T_pQ \longrightarrow T_pQ$$

sending $z \in T_pQ$ to $R_{xy}z \in T_pQ$.

The reason of this is the fact that any tensor field Φ on Q, and in particular the tensor field $\bar{R}$ associated to R, is a field on Q, assigning a value Φ_p at each point $p \in Q$. The main point is that, when Φ is computed on one-forms and vector fields to give a real valued function

$$\Phi(\sigma^1, \ldots, \sigma^r, X_1, \ldots, X_s),$$

the value of this function at $p \in Q$ depends on the values of the arguments at p, only.

Exercise 7.4 Prove this last fact.

The **tensor product** $A \otimes B$ of a mixed tensor field A of type (r, s) by a mixed tensor field B of type (r', s') is a mixed tensor field of type $(r + r', s + s')$ defined as

$$
\begin{aligned}
&(A \otimes B)(\sigma^1, \ldots, \sigma^{r+r'}, Y_1, \ldots, Y_{s+s'}) = \\
&= A(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s).B(\sigma^{r+1}, \ldots, \sigma^{r+r'}, Y_{s+1}, \ldots, Y_{s+s'}).
\end{aligned}
$$

The case $r' = s' = 0$ (B is a function $f \in \mathcal{D}(Q)$) can be also included in this definition and get

$$A \otimes f = f \otimes A = fA.$$

(the same for A of type $(r,s) = (0,0)$.)

Remark The tensor product is an associative (but not commutative) operation. In fact, in a local system of coordinates $x^1, \ldots, x^n$, we have

$$(dx^1 \otimes dx^2)(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}) = 1$$

$$(dx^2 \otimes dx^1)(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}) = 0.$$

Using Exercise 7.4 it is an easy matter to show that in a local system of coordinates $(U; x^1, \ldots, x^n)$ a mixed tensor field A of type (r,s) has, uniquely defined, its (local) components $A^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$, that are real-valued functions defined in U, by:

$$A^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} = A(dx^{i_1}, \ldots, dx^{i_r}, \frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_s}}),$$

where all the indices run from 1 to $n = \dim Q$. One can see also that the tensor fields

$$\frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots, \otimes dx^{j_s}$$

generate all mixed tensor fields of type (r,s) in the sense that

$$A = \Sigma A^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}$$

where each index is summed from 1 to n. In particular, for a $(0,1)$ tensor, that is, a one-form σ, we have

$$\sigma = \sum_{i=1}^{n} [\sigma(\frac{\partial}{\partial x^i})] dx^i,$$

and, for a $(1,0)$ tensor, that is, a vector field Y, one can write

$$Y = \sum_{i=1}^{n} [dx_i(Y)] \frac{\partial}{\partial x^i}.$$

One can extend the notion of components $\Phi^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ of a tensor field Φ of a type (r,s) with respect to any (local) basis $X_1, \ldots, X_n$ of vector fields defined in U and to its dual basis $\omega^1, \ldots, \omega^n$; they are the coefficients of the expression of Φ when it is written in terms of the local basis for the tensor fields in U of type (r,s), that is, in terms of the family of tensor fields

$$X_{i_1} \otimes \ldots \otimes X_{i_r} \otimes \sigma^{j_1} \otimes \ldots \otimes \sigma^{j_s}.$$

Explicitly we write

$$\Phi = \sum \Phi^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} X_{i_r} \otimes \ldots \otimes X_{i_r} \otimes \sigma^{j_1} \otimes \ldots \otimes \sigma^{j^s}$$

where each index is summed from 1 to n.

Example 7.4 If A is a type $(2,3)$ tensor field, we know that the contraction $C^1_3(A)$, or simply $C^1_3 A$, is the type $(1,2)$ tensor field given by (see Exercise 4.3):

$$(C^1_3 A)(\sigma, X, Y) = C\{A(\cdot, \sigma, X, Y, \cdot)\};$$

then, relative to a given system of coordinates, the components of $C^1_3 A$ are:

$$\begin{aligned}(C^1_3 A)^k_{ij} &= (C^1_3 A)(dx^k, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}) = C\{A(\cdot, dx^k, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \cdot)\} = \\ &= \sum_{l=1}^{n} A(dx^l, dx^k, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^l}) = \sum_{l=1}^{n} A^{lk}_{ijl}.\end{aligned}$$

In generalizing that example, if A is a mixed tensor field of type (r,s), and for fixed i, j, $1 \le i \le r$ and $1 \le j \le s$, the local components of $C^i_j A$ are

$$\sum_{l=1}^{n} A^{i_1,\ldots,l,\ldots,i_r}_{j_1,\ldots,l,\ldots,j_s}$$

(the l "up" is the ith index, the l "down" is the jth index and $A^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$ are the local components of A).

If A is a mixed tensor field of type (r,s) and when we fix two integers a, b, $1 \le a \le r$ and $1 \le b \le s$ the tensor field A can be identified with a mixed tensor field $\bar{A} \stackrel{def}{=} D^a_b A$ of type $(r-1, s+1)$ using the isomorphism

$$\mu : V \in \mathcal{X}(Q) \longrightarrow \mu(V) \in \Gamma^1(Q)$$

where $\mu(V) \in \Gamma^1(Q)$ is given by

$$\mu(V)(X) = \langle V, X \rangle, \quad \text{for all} \quad X \in \mathcal{X}(Q). \tag{4.24}$$

(The inverse isomorphism $\mu^{-1} : \sigma \in \Gamma^1(Q) \longrightarrow \mu^{-1}(\sigma) = V \in \mathcal{X}(Q)$ is given by $\sigma(\cdot) = \langle V, \cdot \rangle = \langle \mu^{-1}(\sigma), \cdot \rangle$). More precisely, $D^a_b A$ is defined by

$$\begin{gathered} D^a_b A(\theta_1, \ldots, \theta_{r-1}, Y_1, \ldots, Y_{s+1}) = \\ A(\theta_1, \ldots, \sigma, \ldots, \theta_{r-1}, Y_1, \ldots, Y_{b-1}, Y_{b+1}, \ldots, Y_{s+1}) \end{gathered}$$

where is the right hand side we lose the bth covariant slot and in the ath contravariant slot appears the 1-form $\sigma = \mu(Y_b)$ given by (4.24) with $V = Y_b$. For example, let A be a (2,2) tensor field and $\bar{A}$ be the (1,3) tensor field given as $\bar{A} = D_2^1 A$; so $\bar{A}(\theta, Y_1, Y_2, Y_3) = A(\sigma, \theta, Y_1, Y_3)$ for all $\theta \in \Gamma^1(Q)$, and all $Y_i \in \mathcal{X}(Q)$, $i = 1,2,3, \sigma$ being obtained from Y_2 through (4.24), that is, $\sigma = \mu(Y_2)$.

In a local system of coordinates $x^1, \dots, x^n$, (4.24) makes $\frac{\partial}{\partial x^i}$ corresponding to the one form $\Sigma_j g_{ij} dx^j$, so, the (local) components of $\bar{A}$ are

$$\begin{aligned} \bar{A}^i_{jkl} &= \bar{A}(dx^i, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}) = A(\sum_{m=1}^n g_{km} dx^m, dx^i, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^l}) = \\ &= \sum_{m=1}^n g_{km} A^{mi}_{jl}. \end{aligned}$$

The operation D_2^1 uses $\langle,\rangle$ to turn first superscripts into second subscripts. We have an isomorphism inverse for the operation D_b^a, denoted as the operation U_b^a, that, analogously, takes the ath one-form and inserts its corresponding vector field given by (4.24) in the bth slot among the vector fields; D_b^a acts **lowering an index** and U_b^a acts **raising an index**; they are **type-changing** operations.

In local coordinates, $\Sigma_{j=1}^n g^{ij} \frac{\partial}{\partial x^j}$ is the vector field that corresponds to the one-form dx^i ((g^{ij}) is the inverse of the matrix $g_{ij} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$).

For example, as above, if $\bar{A} = D_2^1 A$ is the type (1,3) tensor field with local components $\bar{A}^j_{krl}$, then $(U_2^1 \bar{A})$ is the corresponding (2,2) tensor field with local components

$$(U_2^1 \bar{A})^{ij}_{kl} = \sum_{r=1}^n g^{ir} \bar{A}^j_{krl}.$$

So, $[U_2^1 \circ D_2^1 \bar{A}]^{ij}_{kl} = \Sigma_{r=1}^n g^{ir} \bar{A}^j_{krl} = \Sigma_{r=1}^n g^{ir} \Sigma_{m=1}^n g_{rm} \bar{A}^{mj}_{kl} = \bar{A}^{ij}_{kl}$, that is, $U_2^1 \circ D_2^1 = id$. In general $U_b^a \circ D_b^a = id$.

Using the operations D_b^a and U_b^a we can also define contractions either between two covariant slots or between two contravariant slots; these are the so called **metric contractions**. In fact, for instance, if A is a mixed tensor field of type (1,3) we define the contraction between the 2nd and 3rd covariant slots; from the components A^i_{jkr} of A one obtains $\Sigma_{k,r=1}^n g^{kr} A^i_{jkr}$ for the components of the contraction $C_{23}A$. In terms of the operations D_b^a and U_b^a, $\quad C_{23}A = C_2^2 U_3^2 A$, and we obtain

$$(U_3^2 A)^{ir}_{jk} = \sum_{l=1}^n g^{rl} A^i_{jkl} \quad \text{and} \quad (C_2^2 U_3^2 A)^i_j = \sum_{l,r=1}^n g^{rl} A^i_{jrl}.$$

Analogously, it is possible to define a (metric) contraction between contravariant slots. For instance if A is of type (3,1) and has components A^{ijk}_l one can obtain the C^{23} contraction between the 2nd and 3rd contravariant slots: $(C^{23}A)^i_l = \Sigma_{k,j=1}^n g_{jk} A^{ijk}_l$, or equivalently, $C^{23}A = C_2^2 D_2^3 A$, that is, $(D_2^3 A)^{ij}_{lk} = g_{kr} A^{ijr}_l$ and so $(C_2^2 D_2^3 A)^i_l = \Sigma_{r,k=1}^n g_{kr} A^{ikr}_l$.

Proposition 12.4 *If $x, y, z, v, w \in T_pQ$ and if ∇ is the Levi-Civita connection,* *then*

(a) $R_{xy} = -R_{yx}$.

(b) $\langle R_{xy}v, w\rangle = -\langle v, R_{xy}w\rangle$.

(c) $R_{xy}z + R_{yz}x + R_{zx}y = 0$.

(d) $\langle R_{xy}v, w\rangle = \langle R_{vw}x, y\rangle$.

Proof Since the operations ∇_X and bracket on vector fields are local operations, we only need to work locally, that is, on a coordinate neighborhood, and since the equalities to be proved are equivalent to tensor equations, the vectors x, y, z, v, w can be extended to vector fields X, Y, Z, V, W with constant components, so their brackets are zero and, in particular, $R_{X,Y}Z$ reduces to $\nabla_Y(\nabla_X Z) - \nabla_X(\nabla_Y Z)$. Then:

(a) is immediate.

(b) By polarization of bilinear forms it is enough to show that $\langle R_{xy}v, v\rangle = 0$, and this follows from the fact that the connection ∇ is compatible with $\langle,\rangle$, that is, from (4.16).

(c) follows from the fact that ∇ is symmetric, that is from (4.17), and (d) is just an algebraic exercise that uses (a), (b) and (c). In fact to prove (d) we use (c) and write

$$\langle R_{YW}V, X\rangle + \langle R_{VY}W, X\rangle + \langle R_{WV}Y, X\rangle = 0$$
$$\langle R_{YX}V, W\rangle + \langle R_{VY}X, W\rangle + \langle R_{XV}Y, W\rangle = 0$$
$$\langle R_{YW}X, V\rangle + \langle R_{XY}W, V\rangle + \langle R_{WX}Y, V\rangle = 0$$
$$\langle R_{VW}X, Y\rangle + \langle R_{XV}W, Y\rangle + \langle R_{WX}V, Y\rangle = 0.$$

Using (a) and (b) one obtains, after summation of the four equations:

$$2\langle R_{VW}X, Y\rangle + 2\langle R_{XY}W, V\rangle = 0 \qquad \text{or}$$

$$\langle R_{VW}X, Y\rangle = \langle R_{XY}V, W\rangle. \quad \blacksquare$$

The last proposition showed the symmetries of the curvature operator, and also, the considerable skew-symmetry it has. Property (b) says that R_{xy} is a skew-symmetric linear operator; (a) and (c) hold for any symmetric connection

Δ; (c) is called the **first Bianchi identity** and (d) is said to be the **symmetry by pairs**.

Exercise 8.4 Show that, in local coordinates $(x^1, \ldots, x^n)$, we have

$$R_{\frac{\partial}{\partial x^k}\frac{\partial}{\partial x^l}}\frac{\partial}{\partial x^j} = \sum_i R^i_{jkl}\frac{\partial}{\partial x^i},$$

where

$$R^i_{jkl} = \frac{\partial}{\partial x_l}\Gamma^i_{kj} - \frac{\partial}{\partial x_k}\Gamma^i_{lj} + \Sigma_m \Gamma^i_{lm}\Gamma^m_{kj} - \Sigma_m \Gamma^i_{km}\Gamma^m_{lj}.$$

Remark that $R^i_{jkl} = -R^i_{jlk}$.

As we saw above, to the curvature tensor R corresponds the mixed tensor field $\bar{R}$ of type (1,3) that is, $\bar{R} : \Gamma^1(Q) \times \mathcal{X}^3(Q) \longrightarrow \mathcal{D}(Q)$, defined by

$$\bar{R}(\sigma, X, Y, Z) = \sigma(R_{X,Y}Z). \tag{4.25}$$

In order to introduce the notion of covariant derivative of tensors, we start defining the **covariant differential** $\nabla\sigma^1$ of a one-form σ^1, that is, of a tensor field of type $(0,1)$; $\nabla\sigma^1$ is a $(0,2)$ tensor field defined as

$$\nabla\sigma^1(Y, W) = W(\sigma^1(Y)) - \sigma^1(\nabla_W Y)$$

and the **covariant derivative** $\nabla_W\sigma^1$ of σ^1 with respect to W is the $(0,1)$ tensor field given by

$$(\nabla_W\sigma^1)(Y) \stackrel{def}{=} (\nabla\sigma^1)(Y, W) = W(\sigma^1(Y)) - \sigma^1(\nabla_W Y) \tag{4.26}$$

Given any mixed tensor field Φ of type (r, s):

$$\Phi : (\Gamma^1(Q))^r \times (\mathcal{X}(Q))^s \longrightarrow \mathcal{D}(Q),$$

one can define its **covariant differential** , which is a mixed tensor field $\nabla\Phi$ of type $(r, s+1)$, by the equality:

$$\begin{aligned}
&(\nabla\Phi)(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s, W) = W(\Phi(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s)) - \\
- \quad &\Phi(\sigma^1, \ldots, \sigma^r, \nabla_W Y_1, \ldots, Y_s) - \ldots - \Phi(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, \nabla_W Y_s) - \\
- \quad &\Phi(\nabla_W\sigma^1, \sigma^2, \ldots, \sigma^r, Y_1, \ldots, Y_s) - \ldots - \Phi(\sigma^1, \ldots, \nabla_W\sigma^r, Y_1, \ldots, Y_s),
\end{aligned}$$

where $\nabla_W\sigma^i, i = 1, \ldots, r$, is the covariant derivative introduced in (4.26).

It is a trivial matter to show that one can factor out functions and so $\nabla\Phi$ is really a mixed tensor field of type $(r, s+1)$.

We also define $\nabla_f = df$, for any $f \in \mathcal{D}(Q)$.

The **covariant derivative** $\nabla_W \Phi$ of Φ by the vector field W is the tensor field defined by

$$(\nabla_W \Phi)(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s) = \nabla\Phi(\sigma^1, \ldots, \sigma^r, Y_1, \ldots, Y_s, W). \qquad (4.27)$$

Exercise 9.4 Covariant derivative ∇_W and covariant differential ∇ of a mixed tensor field, commute with both contraction and type changing operations.

To the curvature tensor R, or to the associated mixed tensor field $\bar{R}$ of type (1,3), there correspond the covariant differential $\nabla\bar{R}$, which is a type (1,4) tensor field, as well as $\nabla_W \bar{R}$, a type (1.3) tensor field.

More precisely, from (4.23), (4.25) and (4.27) one has

$$\begin{aligned}(\nabla_W \bar{R})(\sigma, X, Y, Z) &= W(\bar{R}(\sigma, X, Y, Z)) - \bar{R}(\sigma, \nabla_W X, Y, Z) - \\ &- \bar{R}(\sigma, X, \nabla_W Y, Z) - \bar{R}(\sigma, X, Y, \nabla_W Z) \\ &- \bar{R}(\nabla_W \sigma, X, Y, Z),\end{aligned}$$

where $\nabla_W \sigma$ is defined in (4.26), so

$$\begin{aligned}&(\nabla_W \bar{R})(\sigma, X, Y, Z) = W(\sigma(R_{X,Y} Z)) - \\ - \;&\sigma(R_X, \nabla_{R_X, \nabla_W Y} Z + R_{\nabla_W X, Y} Z + R_{X,Y}(\nabla_W Z)) - (\nabla_W \sigma)(R_{X,Y} Z).\end{aligned}$$

Using (4.24) we identify σ with the vector field V given by $\sigma(\cdot) = \langle V, \cdot\rangle$, then $(\nabla_W \sigma)(\bar{X}) = \langle \nabla_W V, \bar{X}\rangle$ for all $\bar{X} \in \mathcal{X}(Q)$. We can give an interpretation to the last equality in the following way: for fixed $W, X, Y \in \mathcal{X}(Q)$, to each $Z \in \mathcal{X}(Q)$ one associates the vector field $(\nabla_W R)(X, Y)Z$ defined by

$$\begin{aligned}&\langle V, (\nabla_W R)(X, Y)Z\rangle = W(\langle V, R_{X,Y} Z\rangle) - \\ - \;&\langle V, R_{\nabla_W X, Y} Z + R_{X, \nabla_W Y} Z + R_{X,Y}(\nabla_W Z)\rangle - \langle \nabla_W V, R_{XY} Z\rangle \\ &\text{for all } V \in \mathcal{X}(Q). \qquad (4.28)\end{aligned}$$

As before, (4.28) makes sense for individual tangent vectors of $T_p Q$, say v, w, x, y, z, and then $(\nabla_w R)(x, y)$ is considered as a linear operator acting on $T_p Q$.

Proposition 13.4 (second Bianchi identity) *For any $x, y, z \in T_p Q$ one has* $(\nabla_z R)(x, y) + (\nabla_x R)(y, z) + (\nabla_y R)(z, x) = 0$, provided that ∇ is compatible with the metric.

Proof Apply the first member of the second Bianchi identity to a general vector $w \in T_p Q$ and compute the scalar product of the result with a vector $v \in T_p Q$. We have to extend v, x, y, z, w, to vector fields V, X, Y, Z, W, respectively, defined on a neighborhood of $p \in Q$. We choose a normal coordinate system (see Exercise 6.4) and let these extensions have constant components; then all the brackets $[\ ,\ ]$ vanish and, for instance, $R_{X,Y}W$ reduces to $\nabla_Y(\nabla_X W) - \nabla_X(\nabla_Y W)$ in (4.23); moreover, $\Gamma^k_{ij}(p) = 0$ and also all the covariant derivatives involving only V, X, Y, Z, W are equal to zero at $p \in Q$ (see (4.5)).

From (4.23) and (4.28) and the fact that ∇ is compatible with the metric we have at the point $p \in Q$:

$$\begin{aligned}
&\langle V, [(\nabla_Z R)(X,Y)W + (\nabla_X R)(Y,Z)W + (\nabla_Y R)(Z,X)W]\rangle = \\
= \;& Z\langle V, R_{X,Y}W\rangle + X\langle V, R_{Y,Z}W\rangle + Y\langle V, R_{Z,X}W\rangle \\
= \;& \langle \nabla_Z V, R_{X,Y}W\rangle + \langle V, \nabla_Z(R_{X,Y}W)\rangle + \\
+ \;& \langle \nabla_X V, R_{Y,Z}W\rangle + \langle V, \nabla_X(R_{Y,Z}W)\rangle + \\
+ \;& \langle \nabla_Y V, R_{Z,X}W\rangle + \langle V, \nabla_Y(R_{Z,X}W)\rangle = 0.
\end{aligned}$$

Then, the last equation evaluated at $p \in Q$ gives $\langle v, [(\nabla_z R)(x,y) + (\nabla_x R)(y,z) + (\nabla_y R)(z,x)]w\rangle = 0$ and since v and w are generic vectors of T_pQ the second Bianchi Identity follows. ∎

The covariant derivative law $\frac{D}{dt}$ of a vector field V along a differentiable curve $C : I \to Q$, introduced in Proposition 2.4, can be extended to any tensor field Φ of type (r,s), by the use of the definitions (4.26) and (4.27). Assume that the vector field W and the curve c satisfy $c(0) = p \in Q$ and $c'(t) = W(c(t))$ for all t. From (4.26) we define

$$\left(\frac{D\sigma^1}{dt}\right)Y(c(t)) \stackrel{def}{=} (W(\sigma^1(Y)))(c(t)) - \sigma^1\left(\frac{DY}{dt}(c(t))\right)$$

and, analogously, from (4.27) we set

$$\begin{aligned}
&\left(\frac{D\Phi}{dt}\right)(\sigma^1,\ldots,\sigma^r,Y_1,\ldots,Y_s)(c(t)) \stackrel{def}{=} (W(\Phi(\sigma^1,\ldots,\sigma^r,Y_1,\ldots,Y_s)))(c(t)) \\
&-\Phi(\sigma^1,\ldots,\sigma^r,\frac{DY_1}{dt},\ldots,Y_s)(c(t)) - \ldots - \Phi(\sigma^1,\ldots,\sigma^r,Y_1,\ldots,\frac{DY_s}{dt})(c(t)) \\
&-\Phi\left(\frac{D\sigma^1}{dt},\ldots,\sigma^r,Y_1,\ldots,Y_s\right)(c(t)) - \Phi(\sigma^1,\ldots,\frac{D\sigma^r}{dt},Y_1,\ldots,Y_s)(c(t)) \qquad (4.29)
\end{aligned}$$

If we start with an orthonormal basis $(e_1,\ldots,e_n)$ at the point p of a pseudo-riemannian manifold $(Q,\langle,\rangle)$, and if we work with the parallel transport of ∇ to construct a basis $(e_1(t),\ldots,e_n(t))$ along $c = c(t)$, and if $(\omega^1(t),\ldots,\omega^n(t))$ is the corresponding dual basis, the restriction $\Phi(c(t))$ of the tensor field Φ to the curve $c = c(t)$ has components $\Phi^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(t)$ relative to $(e_1(t),\ldots,e_n(t))$. And it is easy to see that the components of $\frac{D\Phi}{dt}$ at $c(t)$ relative to $(e_1(t),\ldots,e_n(t))$ are precisely the usual derivatives $\dot\Phi^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}(t)$ with respect to the real variable t.

Another notion to be considered is the **sectional curvature** that will be a simpler real-valued function K which completely determines the riemannian tensor field R. This function K is defined on the set of all non-degenerate tangent planes; recall that a tangent plane at $p \in Q$ is a two-dimensional subspace P of T_pQ and to be **non degenerate** means that

$$q(v,w) \stackrel{def}{=} \langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2 \neq 0$$

for one (hence every) basis $\{v,w\}$ of P. In fact if $\{x,y\}$ is another basis of P we have

$$\begin{aligned} v &= ax+by \\ w &= cx+dy \end{aligned}$$

with $ad-bc \neq 0$, and so, $q(v,w) = (ad-bc)^2 q(x,y)$. Since $\langle R_{v,w}v, w\rangle = (ad-bc)^2\langle R_{x,y}x,y\rangle$, the value

$$K(P) \stackrel{def}{=} \langle R_{v,w}v,w\rangle / q(v,w) \tag{4.30}$$

depends only on the non-degenerate tangent plane P and not on the basis $\{v,w\}$ used in the definition (4.30) of the **sectional curvature** $K(P)$ of P.

Proposition 14.4 *If the sectional curvature satisfies $K(P) = 0$ for all non degenerate tangent planes P at $p \in Q$, then the tensor field R is zero at p.*

We need, for the proof, the following result:

Lemma *If u, v are vectors of a vector space endowed with a non degenerate bilinear form $\langle,\rangle$, there exist vectors $\bar{u}, \bar{v}$ arbitrarily close to u, v, respectively, such that*

$$q(\bar{u},\bar{v}) = \langle\bar{u},\bar{u}\rangle\langle\bar{v},\bar{v}\rangle - \langle\bar{u},\bar{v}\rangle^2 \neq 0.$$

Proof (of the Lemma) Assume u, v linearly independent (because any two vectors can be approximated by independent ones) such that $q(u,v) = 0$. If there is a neighborhood of (u,v) such that $q(\bar{u},\bar{v}) = 0$ for all $(\bar{u},\bar{v})$ in that neighborhood, the analyticity implies that q is identically zero and this is a contradiction. In fact if $\langle,\rangle$ is indefinite there exists a vector $w \neq 0$ such that $\langle w,w\rangle = 0$ and also x such that $\langle w,x\rangle \neq 0$ (otherwise $\langle,\rangle$ is degenerate), then $q(w,x) = -\langle w,x\rangle^2 \langle 0$; if $\langle,\rangle$ is definite, we choose non zero orthogonal vectors a, b; then $q \equiv 0$ gives $\langle a,b\rangle^2 = \langle a,a\rangle.\langle b,b\rangle = 0$, so $\langle a,a\rangle = 0$ with $a \neq 0$ that cannot be. ∎

Proof (of Proposition 14.4). The first step is to see that $\langle R_{v,w}v,w\rangle = 0$ for all $v, w \in T_pQ$; the hypothesis implies that this is true if v, w span a non degenerate plane. If otherwise v, w span a degenerate plan, the last lemma together with the continuity of the function $(x, y) \longmapsto \langle R_{x,y}x, y\rangle$ imply that $\langle R_{v,w}v, w\rangle = 0$ for all $v, w \in T_pQ$. Now, for $v, w \in T_p(Q)$ and arbitrary $x \in T_pQ$ we have $0 = \langle R_{v,w+x}v, w+x\rangle = \langle R_{v,x}v, w\rangle + \langle R_{v,w}v, x\rangle$; the symmetry by pairs (Proposition 12.4 (d)) implies $\langle R_{v,w}v, x\rangle + \langle R_{v,w}v, x\rangle = 0$ and so $R_{v,w}v = 0$ for all $v, w \in T_pQ$. In particular we also have $0 = R_{v+x,w}(v+x) = R_{x,w}v + R_{v,w}x$ that together with Proposition 12.4 (c) imply

$$0 = R_{x,w}v + R_{w,v}x + R_{v,x}w = -R_{v,w}x + R_{w,v}x + R_{v,w}x$$

or $R_{w,v}x = 0$ for all $x \in T_pQ$, that means $R_{w,v} = 0$. Since v, w are arbitrary one obtains $R = 0$ at p. ∎

Given a tensor field Φ on Q of type (r, s), one considers its covariant differential $\nabla\Phi$; the contraction $C^i_{s+1}(\nabla\Phi)$ of the $(s+1)th$ covariant slot with the ith contravariant slot is a tensor field of type $(r-1, s)$ called the **ith-divergence of** Φ, denoted by $div_i\Phi$, that is,

$$div_i\Phi = C^i_{s+1}(\nabla\Phi). \tag{4.31}$$

Remark that Φ has r divergences (see Exercise 4.3).

Example 8.4 A vector field V on Q can be considered as a tensor field of type $(1,0)$ (see Exercise 3.3) so ∇V is a tensor field of type (1,1). The **divergence of** V (in this case ∇V has one only divergence) is the contraction

$$divV = C^1_1(\nabla V). \tag{4.32}$$

From the definition of ∇V we have $(\nabla V)(\sigma, W) = W(V(\sigma)) - V(\nabla_W\sigma) = W(\sigma(V)) - (\nabla_W\sigma)(V) = \sigma(\nabla_W V)$; in local coordinates $(x^1, \ldots, x^n)$, $V = \Sigma_m V^m\frac{\partial}{\partial x^m}$ and then $(\nabla V)^i_j = (\nabla V)(dx^i, \frac{\partial}{\partial x^j}) = dx^i(\nabla_{\frac{\partial}{\partial x^j}}V) = dx^i(\nabla_{\frac{\partial}{\partial x^j}}(\Sigma_m V^m\frac{\partial}{\partial x^m})) = dx^i[\Sigma_m V^m\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^m} + \Sigma_m\frac{\partial V^m}{\partial x^j}\frac{\partial}{\partial x^m}]$. Then $divV = \Sigma_i(\frac{\partial V^i}{\partial x^i} + \Sigma_k\Gamma^i_{ik}V^k)$.

When $Q = \mathbf{R}^3$ with the natural metric, one obtains the usual formula for the divergence of a vector field on $\mathbf{R}^3$.

Example 9.4 The **Hessian** $H(f)$ **of a function** $f \in \mathcal{D}(Q)$ is the covariant differential of df:

$$H(f) = \nabla(df). \tag{4.33}$$

Since $df \in \Gamma^1(Q)$, it can be considered as a tensor field of type $(0,1)$ (see Exercise 3.3) then $H(f)$ is a tensor field of type (0,2). Moreover, $H(f)$ is a symmetric tensor. In fact

$$\begin{aligned}(H(f))(X,Y) &= (\nabla(df))(X,Y) = Y(df(X)) - df(\nabla_Y X)\\ &= Y(X(f)) - (\nabla_Y X)(f);\end{aligned}$$

But since $XY - YX = [X,Y]$ and the Levi-Civita connection is symmetric, we have

$$XY - YX = \nabla_X Y - \nabla_Y X;$$

this last equality, allows us to write

$$(H(f))(X,Y) = (YX - \nabla_Y X)(f) = (XY - \nabla_X Y)(f) = (H(f))(Y,X).$$

The **gradient of a smooth function** $f : Q \longrightarrow \mathbb{R}$, characterized by $\langle grad\ f, X\rangle = df(X)$ for all $X \in \mathcal{X}(Q)$, makes sense in any pseudo-riemannian manifold $(Q,\langle,\rangle)$. In local coordinates $(x^1,\ldots,x^n)$ we have $f = f(x^1,\ldots,x^n)$ and then $df = \Sigma_{i=1}^n \frac{\partial f}{\partial x^i}dx^i$ and so $grad\ f = \Sigma_{i,j} g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j}$.

The **Laplacian** ∇f **of a function** $f \in \mathcal{D}(Q)$ is the divergence of its gradient:

$$\nabla f = div\ (grad\ f) = C_1^1(\nabla(grad\ f)). \tag{4.34}$$

To the riemannian curvature tensor R of $(Q,\langle,\rangle)$ there corresponds a tensor field $\bar{R}$ of type (1,3), (see (4.25)). The contraction $C_3^1(\bar{R})$, also denoted $C_3^1(R)$, is a mixed tensor field of type $(0,2)$ called the **Ricci curvature tensor of** $(Q,\langle,\rangle)$, that is

$$Ric\ (X,Y) = (C_3^1 R)(X,Y). \tag{4.35}$$

Proposition 15.4 *The Ricci curvature tensor of* $(Q,\langle,\rangle)$ *is a type* $(0,2)$ *symmetric tensor.*

Proof In local coordinates $(x^1,\ldots,x^n)$, the components of $\bar{R}$ are denoted by $R^i_{jkl} = \bar{R}(dx^i, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l})$ and the components of $C_3^1\bar{R}$ are $(Ric)_{ij} = (C_3^1\bar{R})(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}) = C\{\bar{R}(\cdot, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \cdot)\} = \Sigma_{l=1}^n \bar{R}(dx^l, \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^l}) = \Sigma_{l=1}^n R^l_{ijl}$, so

$$(Ric)_{ij} = Ric(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}) = \Sigma_{l=1}^n R^l_{ijl}. \tag{4.36}$$

On the other hand, from Proposition 12.4 (d), the symmetry by pairs implies

$$\langle R_{\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^r}\rangle = \langle R_{\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^r}} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\rangle$$

that is,

$$\langle R^l_{kij}\frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^r}\rangle = \langle R^l_{ikr}\frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^j}\rangle$$

where repetition of indices means summation of that index from 1 to n. Equivalently, we have:

$$g_{lr}R^l_{kij} = g_{lj}R^l_{ikr} \qquad \text{for all} \qquad r,k,i,j = 1,\dots,n.$$

From the last expression we get

$$g^{sr}g_{lr}R^l_{kij} = g^{sr}g_{lj}R^l_{ikr},$$

or

$$R^s_{kij} = g^{sr}R^l_{ikr}g_{lj}. \tag{4.37}$$

Contracting s and j one obtains $R^j_{kij} = R^r_{ikr}$ and, using (4.36), we show that

$$(Ric)_{ki} = (Ric)_{ik}. \quad \blacksquare$$

The **scalar curvature** of $(Q,\langle,\rangle)$ is the (metric) contraction S of the Ricci curvature tensor, that is,

$$S = C^1_1(U^1_1 Ric).$$

In local coordinates $(x^1,\dots,x^n)$, one can write: $(U^1_1 Ric)^i_r = g^{ij}(Ric)_{jr}$, and so, $S = g^{ij}(Ric)_{ij} = g^{ij}R^l_{ijl}$; from this it follows that $dS = \frac{\partial S}{\partial x^m}dx^m$ implies

$$\frac{\partial S}{\partial x^m} = \frac{\partial}{\partial x^m}(g^{ij}R^l_{ijl}). \tag{4.38}$$

Remark By definition $\bar{Ric} = U^1_1 Ric$ and we have $div\bar{Ric} = C^1_2(\nabla\bar{Ric})$. Since Ric is a symmetric tensor field of type $(0,2)$, the tensor fields $U^1_1 Ric$ and $U^1_2 Ric$ coincide as type (1,1) tensors; then $div\bar{Ric}$ depends on Ric, only.

Proposition 16.4 $\quad dS = 2div\bar{Ric} \quad$ where $\quad \bar{Ric} = U^1_1 Ric.$

Proof We will fix a point $p \in Q$ and choose a normal coordinate system in a neighborhood of p as we did in Proposition 13.4. Since $\Gamma^l_{ij}(p) = 0$ for all $i,j,l = 1,\dots,n$, we have $\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^i}(p) = 0$ for all $i,j = 1,\dots,n$ and

$$[\frac{\partial}{\partial x^i}\langle\frac{\partial}{\partial x^j},\frac{\partial}{\partial x^k}\rangle](p) = \frac{\partial g_{jk}}{\partial x^i}(p) = 0 \ \ (\text{also } \frac{\partial g^{jk}}{\partial x^i}(p) = 0), \quad \text{for all} \quad i,j,k = 1,\dots,n.$$

On the other hand (4.27) implies

$$\begin{aligned}(\nabla\bar{Ric})^i_{jk} &= (\nabla\bar{Ric})(dx^i,\frac{\partial}{\partial x^j},\frac{\partial}{\partial x^k}) = \\ &= \frac{\partial}{\partial x^k}(\bar{Ric})^i_j - Ric(dx^i,\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j}) - \bar{Ric}(\nabla_{\frac{\partial}{\partial x^k}}dx^i,\frac{\partial}{\partial x^j}).\end{aligned}$$

From (4.26), (4.36) and the definition of U_1^1 we obtain

$$(\nabla \bar{R}ic)^i_{jk} = \frac{\partial}{\partial x^k}(g^{ir} R^l_{rjl}) - \Gamma^l_{jk} g^{ir} R^m_{rlm} + \Gamma^i_{kr} g^{rm} R^l_{mjl},$$

so we have

$$\begin{aligned} [div\bar{R}ic]_j = [C^1_2(\nabla \bar{R}ic)]_j &= (\nabla \bar{R}ic)^i_{ji} = \\ &= [\frac{\partial}{\partial x^i}(g^{ir} R^l_{rjl}) - \Gamma^l_{ji} g^{ir} R^m_{rlm} + \Gamma^i_{ir} g^{rm} R^l_{mjl}]. \end{aligned}$$

The choice of the normal coordinate system implies at $p \in Q$:

$$[div\bar{R}ic]_j(p) = [\frac{\partial}{\partial x^i}(g^{ir} R^l_{rjl})](p) = g^{ir}(p)\frac{\partial R^l_{rjl}}{\partial x^i}(p). \tag{4.39}$$

The second Bianchi identity, for all $r, m, s = 1, \ldots, n$, gives us

$$(\nabla_{\frac{\partial}{\partial x^r}} R)(\frac{\partial}{\partial x^m}, \frac{\partial}{\partial x^s}) + (\nabla_{\frac{\partial}{\partial x^s}} R)(\frac{\partial}{\partial x^r}, \frac{\partial}{\partial x^m}) + \nabla_{\frac{\partial}{\partial x^m}} R(\frac{\partial}{\partial x^s}, \frac{\partial}{\partial x^r}) = 0, \tag{4.40}$$

and, if we use (4.28) and introduce the notation $R^i_{jmr};s$ by the condition

$$\left[(\nabla_{\frac{\partial}{\partial x^r}} R)(\frac{\partial}{\partial x^m}, \frac{\partial}{\partial x^n})\right] \frac{\partial}{\partial x^j} = R^i_{jmr;s} \frac{\partial}{\partial x^i} \tag{4.41}$$

one obtains, at the point $p \in Q$, $R^i_{jmr;s} = \frac{\partial}{\partial x^s}(R^i_{jmr})$ and so

$$[R^i_{jms;r} + R^i_{jrm;s} + R^i_{jsr;m}](p) = 0. \tag{4.42}$$

From Exercise 8.4, we see that reversing r with s in the last parcel (so, with change of sign) and making the contraction of indices i and s one gets $(R^i_{jmi;r} + R^i_{jrm;i} - R^i_{jri;m})(p) = 0$, and then, by (4.36) we arrive to

$$\{\frac{\partial}{\partial x^r}[(Ric)_{jm}] + R^i_{jrm;i}\}(p) = R^i_{jri;m}(p). \tag{4.43}$$

Contracting (metrically) the covariant slots j and r, (4.43) gives us

$$g^{jr}(p)\{\frac{\partial}{\partial x^r}[(Ric)_{jm}]\}(p) + (g^{jr} R^i_{jrm;i})(p) = (g^{jr} R_{jri;m})(p). \tag{4.44}$$

Using (4.38) and (4.44) we can write

$$[\frac{\partial S}{\partial x^m}](p) = [g^{jr} R^i_{jmi;r}](p) + [g^{jr} R^i_{jrm;i}](p). \tag{4.45}$$

From (4.39) we have

$$2[div\bar{R}ic]_m(p) = 2[g^{sr} R^i_{rmi;s}](p). \tag{4.46}$$

Our point now is to show that the second members of (4.45) and (4.46) coincide; for that we use (4.37) and write

$$g^{tj}R^s_{kij} = g^{sr}R^t_{ikr}, \tag{4.47}$$

that, after derivative, gives us at $p \in Q$:

$$[g^{tj}R^s_{kij;m}](p) = [g^{ir}R^t_{ikr;t}](p). \tag{4.48}$$

By contracting indices (t, m) and (s, i) one obtains

$$[g^{tj}R^i_{kij;t}](p) = [g^{ir}R^t_{ikr;t}](p) \tag{4.49}$$

or equivalently

$$[g^{jr}R^i_{kij;r}](p) = [g^{jr}R^i_{jkr;i}](p). \tag{4.50}$$

The last equation shows that the following permutation between the covariant indices hold:

$$(kijr) \longrightarrow (jkri).$$

Now, using (4.50) and the symmetry of the Ricci tensor, we have the equalities:

$$[g^{jr}R^i_{jmi;r}](p) = [g^{jr}R^i_{mji;r}](p) = [g^{jr}R^i_{mri;j}](p) = [g^{jr}R^i_{rmi;j}](p)$$

$$[g^{jr}R_{jrm;i}](p) = [g^{jr}R_{rjm;i}](p) = [g^{jr}R^i_{mri;j}](p) = [g^{jr}R^i_{rmi;j}](p);$$

with the last two equalities, (4.45) and (4.46) imply

$$dS(p) = [2div\bar{R}ic](p)$$

and the proof of Proposition (16.4) is complete. ∎

4.5 E. Cartan structural equations of a connection

Given an affine connection ∇, we put

$$T(X, Y) = \nabla_Y X - \nabla_X Y + [X, Y]. \tag{4.51}$$

The mapping

$$(\sigma, X, Y) \in \Gamma^1(Q) \times \mathcal{X}^2(Q) \to \sigma(T(X, Y)) \in \mathcal{D}(Q)$$

is a mixed tensor field of type (1,2) called the **torsion tensor field** of ∇.

From (3.2) and (3.29), ω, σ being two one differential forms on Q we have

$$(\omega \wedge \sigma)(X, Y) = \omega(X)\sigma(Y) - \omega(Y)\sigma(X), \tag{4.52}$$

$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) \tag{4.53}$$

where $X, Y \in \mathcal{X}(Q)$. The covariant derivative of 1-forms is given in (4.26) by $(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y)$. This last equality together with (4.51) and (4.52) imply

$$d\omega(X,Y) = (\nabla_X \omega)(Y) - (\nabla_Y \omega)(X) - \omega(T(X,Y)). \tag{4.54}$$

Let $p \in Q$ and $(X_1, X_2, \ldots, X_n)$ a basis for the vector fields in some neighborhood N_p of p, that is, any vector field X on N_p can be written as $X = \sum_{i=1}^n f_i X_i$ where $f_i \in \mathcal{D}(N_p)$.

Let ω^i, $\omega_j^k (1 \leq i,j,k \leq n)$ be the one-differential forms in N_p characterized by the equalities

$$w^i(X_j) = \delta_j^i \qquad \text{and} \qquad \omega_j^k = \sum_{i=1}^n \Gamma_{ij}^k \omega^i,$$

the Γ_{ij}^k being smooth functions on N_p defined by the formula $\nabla_{X_i} X_j = \sum_{k=1}^n \Gamma_{ij}^k X_k$.

It is easy to see that the 1-forms ω_j^k determine the connection ∇ on N_p. The **structural equations of E. Cartan** (see the next equations (4.55) and (4.56)) relate the differentials $d\omega^i$ with special 2-forms $\omega^j(T)$ and Ω_j^k associated with the torsion $T(X,Y)$ and with the curvature tensor field $R_{X,Y}Z$, defined in (4.23) and (4.51), respectively.

Proposition 17.4 (E. Cartan) *The following structural equations hold:*

$$d\omega^j = \sum_{k=1}^n \omega^k \wedge \omega_k^j - \omega^j(T) \tag{4.55}$$

$$d\omega_j^k = \sum_{l=1}^n \omega_j^l \wedge \omega_l^k - \Omega_j^k \tag{4.56}$$

where $\omega^j(T)$ and Ω_j^k are 2-differential forms given by $\omega^j(T))(X,Y) = \omega^j(T(X,Y))$ and $\Omega_j^k(X,Y) = \omega^k(R_{X,Y}X_j)$.

Proof We start by observing that if $Z \in \mathcal{X}(N_p)$ we have

$$\nabla_Z X_j = \sum_{k=1}^n (\omega_j^k(Z)) X_k. \tag{4.57}$$

From the equalities

$$(\nabla_Z \omega^l)(X_j) = Z(\omega^l(X_j)) - \omega^l(\nabla_Z X_j) = -\omega^l(\nabla_Z X_j) = -\omega^l(\sum_{k=1}^n \omega_j^k(Z) X_k) = -\omega_j^l(Z)$$

we get

$$\nabla_Z \omega^l = -\sum_{j=1}^{n} (\omega_j^l(Z))\omega^j. \tag{4.58}$$

From (4.52) and (4.53) we have

$$\begin{gathered} d\omega^j(X,Y) - \sum_{k=1}^{n} (\omega^k \wedge \omega_k^j)(X,Y) = \\ X(\omega^j(Y)) - Y(\omega^j(X)) - \omega^j([X,Y]) - \\ -\sum_{k=1}^{n} \omega^k(X)\omega_k^j(Y) + \sum_{k=1}^{n} \omega^k(Y)\omega_k^j(X), \end{gathered}$$

and using (4.58) we arrive to

$$d\omega^j(X,Y) - \sum_{k=1}^{n} (\omega^k \wedge \omega_k^j)(X,Y) = \omega^j(\nabla_X Y - \nabla_Y X - [X.Y]).$$

Taking into account the definition (4.51) of $T(X,Y)$ we obtain (4.55). We now consider (4.53) applied to the 2-form $d\omega_j^k$ and use (4.57) twice to get:

$$\begin{aligned} \nabla_Y(\nabla_X X_j) &= \nabla_Y(\sum_{k=1}^{n} (\omega_j^k(X))X_k) \\ &= \sum_{l=1}^{n} [Y(\omega_j^l(X))X_k + \sum_{k=1}^{n} (\omega_j^k(X))(\nabla_y X_k)] \\ &= \sum_{l=1}^{n} [Y(\omega_j^l(X)) + \sum_{k=1}^{n} \omega_j^k(X)\omega_k^l(X)]X_l. \end{aligned}$$

With the last equality one can write the expression of $R_{X,Y}X_j$ given in (4.23) and obtain from (4.57):

$$\begin{aligned} R_{X,Y}X_j &= \nabla_Y(\nabla_X X_j) - \nabla_X(\nabla_Y X_j) + \nabla_{[X,Y]}X_j = \\ &= \sum_{l=1}^{n} [-d\omega_j^l(X,Y) + \sum_{k=1}^{n} (\omega_j^k \wedge \omega_k^l)(X,Y]X_l. \end{aligned}$$

Finally, the definition of Ω_j^l gives

$$\Omega_j^l(X,Y) = \omega^l(R_{X,Y}X_j) = -d\omega_j^l(X,Y) + \sum_{k=1}^{n} (\omega_j^k \wedge \omega_k^l)(X,Y)$$

and the proof is complete. ∎

As a consequence of Proposition 17.4 one can analyze the case of a riemannian manifold $(Q, \langle,\rangle)$ with the Levi-Civita connection ∇. If we assume that $(X_1, \ldots, X_n)$ is an orthonormal basis, that is $\langle X_r, X_s\rangle = \delta^r_s, r, s = 1, \ldots, n$, we obtain $T(X, Y) = 0$ for all $X, Y \in \mathcal{X}(N_p)$ and then the **structural equations of E. Cartan** for the riemannian case reduce to

$$d\omega^j = \sum_{k=1}^{n} \omega^k \wedge \omega^j_k \tag{4.59}$$

$$d\omega^k_j = \sum_{l=1}^{n} \omega^l_j \wedge \omega^k_l - \Omega^k_j \tag{4.60}$$

and the forms ω^j_k and Ω^k_j satisfy

$$\omega^k_j + \omega^j_k = 0 \tag{4.61}$$

$$\Omega^k_j + \Omega^j_k = 0. \tag{4.62}$$

In fact, since (4.57) holds we obtain

$$\omega^k_j(Z) = \langle \nabla_Z X_j, X_k\rangle.$$

But $Z\langle X_j, X_k\rangle = Z(\delta^j_k) = 0$ and then

$$\langle \nabla_Z X_j, X_k\rangle + \langle X_j, \nabla_Z X_k\rangle = 0$$

is true, that is, $\omega^j_k + \omega^k_j = 0$. But, moreover, (4.60) and (4.61) imply (4.62).

Chapter 5

Mechanical systems on riemannian manifolds

5.1 The generalized Newton law

Let $(Q, \langle , \rangle)$ be a riemannian manifold, $q = q(t)$ be a C^2-curve on Q and ∇ be the Levi-Civita connection associated to the given riemannian metric $<,>$. The acceleration of $q(t)$ is the covariant derivative of the velocity field $\dot{q} = \dot{q}(t)$, that is,

$$\textbf{acceleration of} \quad q(t) \stackrel{def}{=} \frac{D\dot{q}}{dt}. \tag{5.1}$$

If V is any (local) vector field extending $\dot{q} = \dot{q}(t)$, we also write, for simplicity, $\frac{D\dot{q}}{dt} = \nabla_{\dot{q}}\dot{q} = \nabla_{\dot{q}}V$. When $\dot{q}(t) \neq 0$, there exists such a V in a neighborhood of $q(t)$.

In local coordinates $(\Omega; q_1, \ldots, q_n)$ of Q, the functions $g_{ij} =< \frac{\partial}{\partial q_i}, \frac{\partial}{\partial q_j} >$ and the $\Gamma^k{}_{ij}$ given by $\nabla_{\frac{\partial}{\partial q_j}} \frac{\partial}{\partial q_i} = \sum_{k=1}^n \Gamma_{ji}{}^k \frac{\partial}{\partial q_k}$, are well known C^1-functions on Ω and the expressions (4.20) give each $\Gamma^k{}_{ij}$ as a function of the $g_{ij}(q_1, \ldots, q_n)$, so as a function of $q_1, \ldots, q_n$. If $(q_i, \dot{q}_i)$ are the corresponding natural coordinates of TQ on $\tau^{-1}(\Omega)$, one can write:

$$\dot{q} = \sum_{i=1}^n \dot{q}_i \frac{\partial}{\partial q_i} \tag{5.2}$$

and so, we have along $q = q(t)$ (see (4.7)):

$$\frac{D\dot{q}}{dt} = \sum_{k=1}^{n} \left[\ddot{q}_k + \sum_{i,j} \dot{q}_i \dot{q}_j \Gamma_{ij}^k \right] \frac{\partial}{\partial q_k} \tag{5.3}$$

The **kinetic energy** associated to the riemannian metric $<,>$ is the C^k-function $K : TQ \to \mathbf{R}$ given by $K(v_p) = \frac{1}{2} < v_p, v_p >$.

As we will see in some examples, the masses appear in the definition of the metric $<,>$; the **Legendre transformation** or **mass operator** μ is a diffeomorphism from TQ onto T^*Q,

$$\mu : TQ \to T^*Q \tag{5.4}$$

given by $\mu(v_p)(.) = \langle v_p, . \rangle$ for all $v_p \in TQ$. TQ is also called the phase space of velocities and T^*Q is called the phase space of momenta. Since $\langle , \rangle_p$ is non degenerate, we see easily that μ takes the fiber T_pQ onto the fiber $T_p{}^*Q$ and μ identifies, diffeomorphically, TQ with T^*Q. A **field of (external) forces** is a C^1-differentiable map

$$\mathcal{F} : TQ \to T^*Q \tag{5.5}$$

that sends the fiber T_pQ into the fiber $T^*{}_pQ$, for all $p \in Q$.

Remark that, by definition, $\mathcal{F}$ is not necessarily surjective but sends fibers into fibers. When $\mathcal{F}(v_p)$ is constant (for all $p \in Q$ and $v_p \in T_pQ$) the field of forces is said to be **positional**. For an example of a positional field of forces one defines

$$\mathcal{F}_U(v_p) = -dU(p) \qquad \forall v_p \in T_pQ, p \in Q,$$

where $U : Q \to \mathbf{R}$, the **potential energy**, is a given C^2-differentiable function. In that case one says that $\mathcal{F}_U$ is a **conservative** field of forces. It is clear that $\mathcal{F}_U$ is a positional field of forces. The map $\mu^{-1} \circ \mathcal{F}_U : TQ \to TQ$ defines, in this case, a vector field $\mathcal{X}$ on the manifold Q:

$$\mathcal{X} : p \in Q \longmapsto \mu^{-1} \circ \mathcal{F}_U(v_p) \in T_pQ,$$

that does not depend on $v_p \in T_pQ$, but on U and $p \in Q$, only. In fact $\mathcal{X}$ is equal to $-grad\ U$ (**- gradient** of U); take $w_p \in T_pQ$ and so:

$$\begin{aligned} \langle \mathcal{X}(p), w_p \rangle &= \langle \mu^{-1}\mathcal{F}_U(v_p), w_p \rangle = \mu(\mu^{-1}\mathcal{F}_U(v_p))(w_p) \\ &= \mathcal{F}_U(v_p)(w_p) = -dU(p)(w_p), \quad \text{that is } \mathcal{X}(p) = -(grad\ U)(p). \end{aligned}$$

Exercise 1.5 Show that in local coordinates we have

$$\mu\left(\frac{D\dot{q}}{dt}\right) = \sum_{j=1}^{n} \left(\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} - \frac{\partial K}{\partial q_j} \right) dq_j. \tag{5.6}$$

A **mechanical system on a riemannian manifold** $(Q, \langle,\rangle)$ is a triplet $(Q, \langle,\rangle, \mathcal{F})$ where $\mathcal{F}$ is an (external) field of forces. The manifold Q is said to be the **configuration space** and the corresponding **generalized Newton law** is the relation

$$\mu(\frac{D\dot{q}}{dt}) = \mathcal{F}(\dot{q}). \tag{5.7}$$

A **motion** $q = q(t)$ is a C^2-curve, with values on Q, that satisfies the Newton law (5.7). A **conservative mechanical system** is a triplet $(Q, \langle,\rangle, \mathcal{F} = -dU)$ where $U : Q \to \mathbf{R}$ is its potential energy. The function $E_m = K + U \circ \tau$ is the **mechanical energy**.

Proposition 1.5 (Conservation of energy)

In any conservative mechanical system $(Q, <,>, -dU)$ the mechanical energy $E_m = K + U \circ \tau$ is constant along a given motion $q = q(t)$.

Proof:

$$\begin{aligned}\tfrac{d}{dt}E_m(\dot{q}) &= \tfrac{d}{dt}\left[K(\dot{q}) + U \circ \tau(\dot{q})\right] = \tfrac{d}{dt}\left[\tfrac{1}{2} < \dot{q}, \dot{q} > + U(q)\right] = \\ &= \langle (\tfrac{D\dot{q}}{dt}), \dot{q}\rangle + (dU(q))\dot{q} = \langle \mu^{-1}[-dU(q)], \dot{q}\rangle + (dU(q))\dot{q} \\ &= -(dU(q))\dot{q} + (dU(q))\dot{q} = 0. \quad \blacksquare\end{aligned}$$

5.2 The Jacobi riemannian metric

Let $(Q, \langle,\rangle, -dU)$ be a conservative mechanical system on a riemannian manifold $(Q, <,>)$ and U be a C^2-potential energy. Let $v_p \in TQ$ be a critical point of the mechanical energy $E_m = K + U \circ \tau : TQ \to \mathbf{R}$, that is, $dE_m(v_p) = 0$. In local coordinates we have $v_p = (q_i, \dot{q}_i)$ and $E_m(v_p) = \frac{1}{2}\sum_{ij} g_{ij}(p)\dot{q}_i\dot{q}_j + U(q_1(p), \ldots, q_n(p))$, so

$$dE_m(q_i, \dot{q}_i) = \sum_{k=1}^{n}\left[\frac{1}{2}\sum_{ij}\frac{\partial g_{ij}}{\partial q_k}\dot{q}_i\dot{q}_j + \frac{\partial U}{\partial q_k}\right] dq_k + \sum_{k=1}^{n}\left[\sum_i g_{ik}\dot{q}_i\right] d\dot{q}_k = 0$$

and that implies the following equations:

$$\sum_i g_{ik}\dot{q}_i = 0, \qquad k = 1, \ldots, n, \tag{5.8}$$

$$\left[\frac{1}{2}\sum_{ij}\frac{\partial g_{ij}}{\partial q_k}\dot{q}_i\dot{q}_j + \frac{\partial U}{\partial q_k}\right] = 0, \qquad k = 1, \ldots, n. \tag{5.9}$$

By (5.8) and (5.9), and since $det(g_{ij}) \neq 0$, $v_p \in TQ$ is a critical point of E_m if, and only if:

$$\dot{q}_i = 0, \quad i = 1, \dots, n, \quad \text{and} \quad \frac{\partial U}{\partial q_k}(p) = 0.$$

This means that v_p is a **critical point** of E_m if, and only if, $p \in Q$ is a critical point of U and $v_p = o_p \in T_pQ$.

Let $h \in \mathbb{R}$ be a regular value of the mechanical energy E_m with $E_m^{-1}(h) \neq \phi$; so the points in $E_m^{-1}(h)$ are not critical points. Consider the open set of Q:

$$Q_h = \{p \in Q \mid U(p) < h\}. \tag{5.10}$$

On the manifold Q_h one can define the so called **Jacobi metric** g_h **associated to** $\langle , \rangle$; for each $p \in Q_h$ define $g_h(p)$ by

$$g_h(p)(u_p, v_p) \overset{def}{=} 2(h - U(p))\langle u_p, v_p\rangle, \tag{5.11}$$

Since $(h - U(p)) > 0$ for $p \in Q_h$, one sees that g_h is a riemannian metric on Q_h.

Proposition 2.5 (Jacobi) *The motions of a conservative mechanical system $(Q, \langle , \rangle, -dU)$ with mechanical energy h are, up to reparametrization, geodesics of the open manifold Q_h with the Jacobi metric associated to $\langle , \rangle$.*

Before proving the proposition 2.5 one goes to show the following:

Proposition 3.5 *Let $(Q, \langle , \rangle)$ be a riemannian manifold, $\rho : Q \to \mathbb{R}$ to be a C^2 function and grad ρ denote a vector field on Q, the gradient corresponding to $<,>$ of the function ρ. Let ∇ and $\tilde{\nabla}$ be the Levi-Civita connections associated to $<,>$ and $e^{2\rho}\langle , \rangle$, respectively. Then, for all $X, Y \in \mathcal{X}(Q)$ we have:*

$$\tilde{\nabla}_X Y = \nabla_X Y + d\rho(X)Y + d\rho(Y)X - \langle X, Y\rangle grad\rho \tag{5.12}$$

Proof By the definition of $\tilde{\nabla}$ and making $\ll , \gg = e^{2\rho}\langle , \rangle$, formula (4.19) gives

$$2 \ll \tilde{\nabla}_X Y, Z \gg = Y \ll X, Z \gg + X \ll Z, Y \gg - Z \ll X, Y \gg$$
$$- \ll [Y, Z], X \gg - \ll [X, Z], Y \gg - \ll [Y, X], Z \gg .$$

On the other hand we have

$$\begin{aligned} Y \ll X, Z \gg &= Y(e^{2\rho}\langle X, Z\rangle) = e^{2\rho}Y\langle X, Z\rangle + \langle X, Z\rangle Y(e^{2\rho}) = \\ &= e^{2\rho}[Y\langle X, Z\rangle + \langle X, Z\rangle Y(2\rho)], \end{aligned}$$

so,

$$
\begin{aligned}
2 \ll \tilde{\nabla}_X Y, Z \gg &= e^{2\rho}\{Y\langle X, Z\rangle + \langle X, Z\rangle Y(2\rho) + X\langle Z, Y\rangle + \\
&+ \langle Z, Y\rangle X(2\rho) - Z\langle X, Y\rangle - \langle X, Y\rangle Z(2\rho) \\
&- \langle [Y, Z], X\rangle - \langle [X, Z], Y\rangle - \langle [Y, X], Z\rangle\}.
\end{aligned}
$$

From (4.19) one obtains

$$
\begin{aligned}
2 \ll \tilde{\nabla}_X Y, Z \gg &= 2e^{2\rho} < \nabla_X Y, Z > + e^{2\rho}\{\langle X, Z\rangle Y(2\rho) \\
&+ \langle Z, Y\rangle X(2\rho) - \langle X, Y\rangle Z(2\rho)\} \\
&= 2 \ll \nabla_X Y, Z \gg + \ll X, Z \gg Y(2\rho) \\
&+ \ll Z, Y \gg X(2\rho) - \ll X, Y \gg Z(2\rho).
\end{aligned}
$$

Since $Y(2\rho) = 2Y(\rho) = 2d\rho(Y)$ we have

$$
\begin{aligned}
\ll \tilde{\nabla}_X Y, Z \gg &= \ll \nabla_X Y, Z \gg + \ll X, Z \gg d\rho(Y) \\
&+ \ll Z, Y \gg d\rho(X) - \ll X, Y \gg d\rho(Z).
\end{aligned}
$$

The definition of $grad\ \rho$ gives

$$d\rho(Z) = \langle grad\ \rho, Z\rangle$$

for all Z, thus

$$
\begin{aligned}
\langle \tilde{\nabla}_X Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle X, Z\rangle d\rho(Y) + \langle Z, Y\rangle d\rho(X) \\
&- \langle X, Y\rangle\langle grad\ \rho, Z\rangle \quad \text{for all} \quad Z.
\end{aligned}
$$

So, one obtains (5.12).

Proof of Proposition 2.5

One defines $\rho : Q_h \to \mathbf{R}$ by the equality $e^{2\rho} = 2(h - U)$ so $e^{2\rho} d\rho = -dU$ and then

$$e^{2\rho} grad\ \rho = -grad\ U \quad \text{with respect to} \ \langle , \rangle, \tag{5.13}$$

that is

$$2(h - U)d\rho = -dU. \tag{5.14}$$

Let $\gamma = \gamma(t)$ be a motion of $(Q, \langle , \rangle, -dU)$ with mechanical energy h and contained in Q_h. By (5.7) we have

$$\nabla_{\dot{\gamma}} \dot{\gamma} = -(grad\ U)(\gamma(t)) \tag{5.15}$$

because

$$2K(\dot{\gamma}) = \langle \dot{\gamma}, \dot{\gamma}\rangle = 2(h - U(\gamma(t)) = e^{2\rho(\gamma(t))},$$

that implies $\dot{\gamma}(t) \neq 0$ for all t in the maximal interval of γ.

Using (5.12), (5.15), (5.13) and (5.14) one can write

$$\begin{aligned} \tilde{\nabla}_{\dot{\gamma}}\dot{\gamma} &= \nabla_{\dot{\gamma}}\dot{\gamma} + 2d\rho(\dot{\gamma})\dot{\gamma} - \langle\dot{\gamma},\dot{\gamma}\rangle grad\ \rho \\ &= -(grad\ U)(\gamma(t)) + 2d\rho(\dot{\gamma})\dot{\gamma} - e^{2\rho(\gamma(t))} grad\ \rho, \quad \text{so} \end{aligned}$$

$$\tilde{\nabla}_{\dot{\gamma}}\dot{\gamma} = 2d\rho(\dot{\gamma})\dot{\gamma}. \tag{5.16}$$

Let s and $\tilde{s}$ be the arc lengths in $\langle,\rangle$ and $\ll,\gg$ respectively. Call $\mu(s) = \gamma(t(s))$ and $c(\tilde{s}) = \mu(s(\tilde{s}))$. So $c(\tilde{s}) = \gamma(t(s(\tilde{s})))$ and $c'(\tilde{s}) = \frac{dc(\tilde{s})}{d\tilde{s}} = \dot{\gamma}(t(s(\tilde{s})))\frac{dt}{d\tilde{s}}(s(\tilde{s})) = \dot{\gamma}(t(s(\tilde{s}))).\frac{dt(s)}{ds}.\frac{ds(\tilde{s})}{d\tilde{s}}$.

But $(\frac{dt(s)}{ds})^2 = (\frac{ds(t)}{dt})^{-2} = \langle\dot{\gamma},\dot{\gamma}\rangle^{-1} = e^{-2\rho(\gamma(t(s)))}$ and then

$$\frac{dt(s)}{ds} = e^{-\rho(\gamma(t(s)))}. \tag{5.17}$$

Analogously

$$\begin{aligned} \left(\frac{ds(\tilde{s})}{d\tilde{s}}\right)^2 &= \left(\frac{d\tilde{s}(s)}{ds}\right)^{-2} = \ll \mu'(s), \mu'(s) \gg^{-1} \\ &= \ll \dot{\gamma}(t(s))\frac{dt(s)}{ds}, \dot{\gamma}(t(s))\frac{dt(s)}{ds} \gg^{-1} \\ &= \left(\frac{dt(s)}{ds}\right)^{-2} \ll \dot{\gamma}(t(s)), \dot{\gamma}(t(s)) \gg^{-1} \end{aligned}$$

that gives

$$\begin{aligned} \left(\frac{ds(\tilde{s})}{d\tilde{s}}\right).\left(\frac{dt(s)}{ds}\right) &= \ll \dot{\gamma}(t(s)), \dot{\gamma}(t(s)) \gg^{-1/2} \\ &= e^{-\rho(\gamma(t(s)))}\langle\dot{\gamma}(t(s)), \dot{\gamma}(t(s))\rangle^{-1/2} \end{aligned}$$

then $(\frac{ds(\tilde{s})}{d\tilde{s}}).(\frac{dt(s)}{ds}) = e^{-2\rho(\gamma(t(s)))}$ and

$$c'(\tilde{s}) = \dot{\gamma}(t(s(\tilde{s}))).e^{-2\rho(\gamma(t(s)))}. \tag{5.18}$$

Now compute $\tilde{\nabla}_{c'(\tilde{s})}c'(\tilde{s})$ using (5.18) and obtain

$$\begin{aligned} \tilde{\nabla}_{c'}c' &= \tilde{\nabla}_{e^{-2\rho}\dot{\gamma}}e^{-2\rho}\dot{\gamma} = e^{-2\rho}\tilde{\nabla}_{\dot{\gamma}}e^{-2\rho}\dot{\gamma} = e^{-2\rho}[e^{-2\rho}\tilde{\nabla}_{\dot{\gamma}}\dot{\gamma} + [d(e^{-2\rho})\dot{\gamma}]\dot{\gamma}] \\ &= e^{-4\rho}[\tilde{\nabla}_{\dot{\gamma}}\dot{\gamma} - 2d\rho(\dot{\gamma})\dot{\gamma}]; \end{aligned}$$

using (5.16) we get $\tilde{\nabla}_{c'}c' = 0$, so $c(\tilde{s}) = \gamma(t(s(\tilde{s})))$ is a geodesic in the Jacobi metric. ∎

5.3 Mechanical systems as second order vector fields

Let $(Q, \langle,\rangle, \mathcal{F})$ be a mechanical system on the riemannian manifold $(Q, \langle,\rangle)$ and $q(t)$ a motion, that is, a solution of the generalized Newton law $(\frac{D\dot{q}}{dt}) = \mu^{-1}(\mathcal{F}(\dot{q}))$.

In local coordinates we have (see (5.3)):

$$\sum_{k=1}^{n}\left[\ddot{q}_k + \sum_{ij}\Gamma_{ij}^{k}\dot{q}_i\dot{q}_j\right]\frac{\partial}{\partial q_k} = \sum_{k=1}^{n} f_k(q,\dot{q})\frac{\partial}{\partial q_k}$$

where the $f_k(q,\dot{q})$ are the components of $\mu^{-1}(\mathcal{F}(\dot{q}))$, that is, the Newton law is locally equivalent to the 2nd order system of ordinary differential equations:

$$\ddot{q}_k = -\sum_{i,j}\Gamma_{ij}^{k}\dot{q}_i\dot{q}_j + f_k(q,\dot{q}), \quad k = 1,\ldots,n,$$

or, to the first order system of ordinary differential equations:

$$\begin{cases} \dot{q}_k = v_k \\ \dot{v}_k = -\sum_{i,j}\Gamma_{ij}^{k}(q_1,\ldots,q_n)v_iv_j + f_k(q,\dot{q}), \end{cases} \tag{5.19}$$

$k = 1,\ldots,n$.

Using (5.6) we also have

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_j} - \frac{\partial K}{\partial q_j} = \sum_{k=1}^{n} g_{jk}f_k(q,\dot{q}), \quad j = 1,\ldots,n \tag{5.20}$$

that are called the **Lagrange equations** for the system.

This way, in natural coordinates $(q,\dot{q}) = (q,v)$ of TQ we have, well defined, the vector-field

$$E : (q,v) \longmapsto ((q,v),(\dot{q},\dot{v}))$$

where the $(\dot{q},\dot{v})$ are given by (5.19). The map above is a vector field E on TQ,

$$E : v_p \in TQ \longmapsto E(v_p) \in T(TQ).$$

The tangent space TQ is called the **phase space** and the vector field E defined on TQ is said to be a **second order vector field** because the first equation (see (5.19)) is $\dot{q} = v$. This is equivalent to say that any trajectory of $E = E(v_p)$ is the derivative of its projection on Q. In the special case where $\mathcal{F} = 0$, the vector field E reduces to the geodesic flow S of $\langle,\rangle$, (see (4.21)), given locally by

$$S : (q,v) \longmapsto ((q,v),(v,\gamma))$$

where $\gamma = (\gamma_k)$ is given by $\gamma_k = -\sum_{ij} \Gamma^k_{ij} v_i v_j$.

In order to write an explicit expression for $E = E(v_p)$, let us introduce the concept of **vertical lifting operator** . Is an operator denoted by C_{v_p} associated to an element $v_p \in T_pQ$. C_{v_p} is a map

$$C_{v_p} : T_pQ \longrightarrow T_{v_p}(TQ)$$

defined by

$$C_{v_p}(w_p) = \frac{d}{ds}(v_p + sw_p) \mid_{s=0} . \tag{5.21}$$

C_{v_p} takes $w_p \in T_pQ$ into a tangent vector of $T(TQ)$ at the point $v_p \in TQ$. This tangent vector $C_{v_p}(w_p)$ is **vertical**, that is, is tangent at the point v_p to the fiber T_pQ since the curve $s \mapsto v_p + sw_p$ passes through v_p at $s = 0$ and has values on T_pQ for all s. In local coordinates, if $v_p = (q_i, v_i)$ and $w_p = (q_i, w_i)$, we have

$$C_{v_p} : (q_i, w_i) \longmapsto ((q_i, v_i), (0, w_i))$$

because the curve $v_p + sw_p$ is given, in local coordinates by $v_p + sw_p = (q_i, v_i + sw_i)$ and its tangent vector at $s = 0$ is written as $((q_i, v_i), (0, w_i))$.

The map C_{v_p} is linear and injective so is an isomorphism of T_pQ onto its image

$$C_{v_p}(T_pQ) = T_{v_p}(\tau^{-1}(p)).$$

So, the vector field $E = E(v_p)$ is given, in local coordinates, by the expression

$$E(v_p) = E((q_i, v_i)) = ((q_i, v_i), (v_i, \gamma_i + f_i))$$

where $\gamma_i = -\sum_{r,s} \Gamma^i_{rs} v_r v_s$ and the (f_i) defined by

$$\mu^{-1}(\mathcal{F}(v_p)) = \sum_{i=1}^{n} f_i \frac{\partial}{\partial q_i}(p).$$

Then

$$E(v_p) = ((q_i, v_i), (v_i, \gamma_i)) + ((q_i, v_i), (0, f_i)), \quad \text{so}$$
$$E(v_p) = S(v_p) + C_{v_p}(\mu^{-1}(\mathcal{F}(v_p))). \tag{5.22}$$

Proposition 4.5 *The second order vector field $E = E(v_p)$ defined on TQ and associated to the generalized Newton law of the mechanical system $(Q, \langle , \rangle, \mathcal{F})$, is given by the expression (5.22) where $S = S(v_p)$ is the geodesic flow of $\langle , \rangle$. The trajectories of E are the derivatives of the motions satisfying $\mu(\frac{D\dot{q}}{dt}) = \mathcal{F}(\dot{q})$. When $\mathcal{F}(v_p) = -dU(p)$, and h is a regular value of E_m, the manifold $E_m^{-1}(h)$ is invariant under the flow of the vector field $E = E(v_p)$.*

5.4 Mechanical systems with holonomic constraints

Let $\mathcal{F} : TQ \to T^*Q$ be a C^1-field of external forces acting on a riemannian manifold $(Q, \langle,\rangle)$.

A **holonomic constraint** is a submanifold $N \subset Q$ such that $dim\ N < dim\ Q$. A C^2-curve $q : I \subset \mathbf{R} \to Q$ is said to be **compatible** with N if $q(t) \in N$ for all $t \in I$. In order to obtain motions compatible with N we have to introduce a **field of reactive external forces** $\mathcal{R} : TN \longrightarrow T^*Q$ depending on $Q, \langle,\rangle, N$ and $\mathcal{F}$, only, and to consider the **generalized Newton law**

$$\mu(\frac{D\dot{q}}{dt}) = (\mathcal{F} + \mathcal{R})(\dot{q}). \tag{5.23}$$

The constraint N is said to be **perfect** (or to satisfy **d'Alembert principle**) if, for a given $\mathcal{F}$, the field of reactive external forces $\mathcal{R}$ has to satisfy "$\mu^{-1}\mathcal{R}(v_q)$ is orthogonal to T_qN for all $v_q \in TN$". Here orthogonality is understood with respect to $\langle,\rangle$, μ is the mass operator and ∇ is the Levi-Civita connection associated to the riemannian structure $(Q, \langle,\rangle)$. Using the decomposition ${v_q}^T + {v_q}^\perp = v_q$ for all $q \in N$ and $v_q \in T_qQ$, that is

$$T_qN \oplus (T_qN)^\perp = T_qQ, \quad q \in N,$$

one obtains from (5.23), assuming $\dot{q} \neq 0$, the following relations:

$$(\nabla_{\dot{q}}\dot{q})^T - [\mu^{-1}(\mathcal{F}(\dot{q}))]^T = 0 \tag{5.24}$$

$$\mu^{-1}\mathcal{R}(\dot{q}) = (\nabla_{\dot{q}}\dot{q})^\perp - [\mu^{-1}(\mathcal{F}(\dot{q}))]^\perp. \tag{5.25}$$

If we denote by D the Levi-Civita connection associated to the riemannian metric $\ll,\gg$ induced by $\langle,\rangle$ on N, the next Exercise 2.5 shows that if N is perfect, the C^2 solution curves compatible with N are precisely the motions of the mechanical system (without constraints) $(N, \ll,\gg, \mathcal{F}_N)$ where $\mathcal{F}_N(v_q) = \mu_N[(\mu^{-1}\mathcal{F}(v_q))^T]$, $v_q \in T_qN$, the μ_N being the mass operator of $(N, \ll,\gg)$.

In fact, since $D_{\dot{q}}\dot{q} = (\nabla_{\dot{q}}\dot{q})^T$ (by Exercise 2.5) one obtains from (5.24) that

$$\mu_N(D_{\dot{q}}\dot{q}) = \mathcal{F}_N(\dot{q}) = \mu_N([\mu^{-1}\mathcal{F}(\dot{q})]^T) \tag{5.26}$$

which is the generalized Newton law corresponding to $(N, \ll,\gg, \mathcal{F}_N)$.

Also, from (5.25) we see that

$$\begin{aligned} \mu^{-1}\mathcal{R}(\dot{q}) &= \nabla_{\dot{q}}\dot{q} - (\nabla_{\dot{q}}\dot{q})^T - [\mu^{-1}\mathcal{F}(\dot{q})]^\perp \\ &= \nabla_{\dot{q}}\dot{q} - D_{\dot{q}}\dot{q} - [\mu^{-1}\mathcal{F}(\dot{q})]^\perp. \end{aligned} \tag{5.27}$$

If X, Y are local vector fields on N and $\bar{X}, \bar{Y}$ be local extensions to Q, we have

$$B(X,Y) = \nabla_{\bar{X}}\bar{Y} - D_XY \tag{5.28}$$

where B is bilinear and symmetric with $B(X,Y)(q)$ depending only on $X(q)$ and $Y(q)$; B is called the **second fundamental form of the imbedding** $i : N \to Q$ (see "Geometria Riemanniana" by M.P. do Carmo, 2ª edição pag. 128). So, from (5.27) and (5.28) we can write $\mu^{-1}\mathcal{R}(\dot{q}) = B(\dot{q},\dot{q}) - [\mu^{-1}\mathcal{F}(\dot{q})]^{\perp}$ that suggests

$$\mathcal{R}(v_q) = \mu[B(v_q, v_q) - [\mu^{-1}\mathcal{F}(v_q)]^{\perp}] \in T_q^*Q \tag{5.29}$$

for all $q \in N$ and $v_q \in T_qN$. The last expression gives the way to compute the **reactive external force**introduced in (5.23) when the constraint is perfect.

Using (5.6) for $\mu_N(D_{\dot{q}}\dot{q})$ with $\dot{q} \neq 0$, in local coordinates of N, and also (5.26), we obtain the so-called **Lagrange equations** for obtaining the motions compatible with the perfect constraints without computing the reaction of the constraints.

Exercise 2.5 Let N be a submanifold of a pseudo-riemannian manifold $(Q, \langle,\rangle)$ with Levi-Civita connection ∇. For any pair of vector fields X, Y on N we define D_XY as the vector field on N that at the point $p \in N$ is equal to $(D_XY)(p) = [(\nabla_{\bar{X}}\bar{Y})(p)]^T$ where $\bar{X}, \bar{Y}$ are local vector fields that extend X and Y in a neighborhood of $p \in Q$, respectively, $[(\nabla_{\bar{X}}\bar{Y})(p)]^T$ being the orthogonal projection of $(\nabla_{\bar{X}}\bar{Y})(p)$ onto T_pN, under $\langle,\rangle$.

Show that $(D_XY)p$ does not depend on the chosen extensions and that

$$D : \mathcal{X}(N) \times \mathcal{X}(N) \to \mathcal{X}(N)$$

has the properties of an affine connection. Verify also that D is symmetric and compatible with the pseudo-riemannian metric $\ll,\gg$ induced by $\langle,\rangle$ on N. So, D is the Levi-Civita connection associated to the pseudo-riemannian manifold $(N, \ll,\gg)$.

5.5 Some classical examples

The study of a system of particles with or without constraints starts, in classical analytical mechanics, with the consideration of a manifold of configurations Q endowed, in general, with two metrics: $(,)$ and $\langle,\rangle$; the first one is called the **spatial metric** and the second is the one corresponding to the kinetic energy that defines the mass operator $\mu : TQ \to T^*Q$. With the two metrics one introduces the **tensor** of inertia $I : \mathcal{X}(Q) \to \mathcal{X}(Q)$ characterized by the relation

$$(I(X), Z) = \langle X, Z\rangle \tag{5.30}$$

for all $X, Z \in \mathcal{X}(Q)$. It is clear that:

i) I is non degenerate with respect to $(,)$ so I^{-1} exists.

ii) I is symmetric with respect to $(,)$, since:

$$(I(X), Z) = \langle X, Z\rangle = \langle Z, X\rangle = (I(Z), X) = (X, I(Z)).$$

iii) I is symmetric with respect to $\langle,\rangle$. In fact,

$$(I(I(X)), Z) = \langle I(X), Z\rangle \quad \text{and}$$
$$(I(I(X)), Z) = (I(X), I(Z)) = (I(I(Z)), X) = \langle I(Z), X\rangle$$

iv) I^{-1} is symmetric with respect to $\langle,\rangle$ and $(,)$:

$$\langle I^{-1}(X), Z\rangle = (X, Z) = (X, I(I^{-1}(Z))) = (I(I^{-1}(Z)), X) = \langle I^{-1}(Z), X\rangle$$

and

$$\begin{aligned}(I^{-1}(X), Z) &= (I^{-1}(X), I(I^{-1}(Z))) = \langle I^{-1}(X), I^{-1}(Z)\rangle \\ &= (I(I^{-1}(X)), I^{-1}(Z)) = (X, I^{-1}(Z)).\end{aligned}$$

v) Assume $(,)$ and $\langle,\rangle$ are positive definite. Then I and I^{-1} are positive definite with respect to the metrics:

$$\begin{aligned}(I(X), X) &= \langle X, X\rangle; \\ \langle I(X), X\rangle &= \langle I(X), I^{-1}(I(X))\rangle = (I(X), I(X)); \\ \langle I^{-1}(X), X\rangle &= (X, X); \\ (I^{-1}(X), X) &= (I^{-1}(X), I(I^{-1}(X))) = \langle I^{-1}(X), I^{-1}(X)\rangle.\end{aligned}$$

In the applications, the usual forces are given by a map $F : TQ \to TQ$ which is fiber preserving that is $F(T_pQ) \subset T_pQ$ for all $p \in Q$; the notion of **work** is introduced using the spatial metric. So, the **work of $F(v_p)$ along w_p** is defined as $(F(v_p), w_p)$. To obtain the external field of forces $\mathcal{F} : TQ \to T^*Q$ from F we write

$$\mathcal{F} \stackrel{def}{=} \mu I^{-1} F \tag{5.31}$$

and, then, the generalized Newton law can be written under one of the two equivalent forms:

$$(\frac{D\dot{q}}{dt}) = I^{-1}F(\dot{q}) \quad \text{or} \quad I(\frac{D\dot{q}}{dt}) = F(\dot{q})$$

(In (5.31), so in the last formulae, I is considered as a fiber preserving map $I : TQ \longrightarrow TQ$.)

Example 1.5 - The system of n mass points

Let k be a three dimensional oriented euclidean vector space also considered as affine space associated to itself. A pair (q_i, m_i) such that $q_i \in k$ and $m_i > 0$ is said to be a **mass point** and m_i is the **mass** of point $q_i, i = 1, \ldots, n$. To give n mass points is to consider $q = (q_1, \ldots, q_n) \in k^n$ and $(m_1, \ldots, m_n) \in \mathbb{R}_+{}^n$.

Assume that at each point $q_i \in k$ acts an **external force** $f_i{}^{ext} = f_i{}^{ext}(q, \dot{q}) \in k$ and $(n-1)$ **internal forces** $f_{ij} \in k, j \in \{1, \ldots, n\} \backslash \{i\}$, due to the action of q_j on q_i. The laws, in classical mechanics, relative to the motions $q_i(t)$ of the mass points (q_i, m_i) are the following:

I - **Newton laws:**

$$m_i \ddot{q}_i = f_i \stackrel{def}{=} (f_i{}^{ext} + \sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ij}), \qquad i = 1, \ldots, n.$$

II - **Principle of action and reaction:**

f_{ij} and $(q_i - q_j)$ are linearly dependent and $f_{ij} = -f_{ji}$.

Remark: The two laws above imply the following:

(a) $\sum_{i=1}^{n} m_i \ddot{q}_i = \sum_{i=1}^{n} f_i{}^{ext}$

(b) $\sum_{i=1}^{n} m_i \ddot{q}_i \times (q_i - c) = \sum_{i=1}^{n} f_i{}^{ext} \times (q_i - c)$ for any $c \in k$.
(here $\times$ means the usual vector product in k).

In fact, case (a) is trivial. Using Newton's law one proves case (b) under the hypothesis $c = 0$, provided that $\sum_{i,j} f_{ij} \times q_i = 0$; but since $f_{ij} \times (q_i - q_j) = 0$, we have

$$\sum_{i,j} f_{ij} \times q_i = \sum_{i,j} f_{ij} \times q_j = -\sum_{i,j} f_{ji} \times q_j = -\sum_{i,j} f_{ij} \times q_i = 0.$$

The case (b) for arbitrary $c \in k$ follows from case (a) and from case (b) with $c = 0$.

The **kinetic energy of a motion** is $K = \frac{1}{2} \sum_{i=1}^{n} m_i (\dot{q}_i, \dot{q}_i)$ where $(,)$ is the inner product of k. The manifold $Q = k^n$ is the configuration space that can be endowed with two riemannian metrics:
$(u, v) = (u_1, v_1) + \ldots + (u_n, v_n)$, the **spatial metric**, and $\langle u, v \rangle = m_1(u_1, v_1) + \ldots + m_n(u_n, v_n)$, the metric corresponding to the kinetic energy, where the masses appear.

The Levi-Civita connection ∇ associated to $\langle , \rangle$ has the g_{ij} as constant functions, so the Christoffel symbols are all zero (see Exercise 5.4) and then

$$(\frac{D\dot{q}}{dt}) = \ddot{q} = (\ddot{q}_1, \ldots, \ddot{q}_n).$$

The mass operator $\mu : Tk^n \to T^*k^n$ is defined by $\mu(w_x)(.) = \langle w_x, . \rangle$ for all $w_x \in T_x k^n \cong k^n$. If the usual forces are given by $F : Tk^n \to Tk^n$ with $F = (f_1, \dots, f_n)$, one defines $\mathcal{F} : Tk^n \to T^*k^n$, the field of external forces, using the formula $\mathcal{F} = \mu I^{-1} F$ where I is given by (5.30). Then one can write:

$$\begin{aligned} \mathcal{F}(v_x)u_x &= (\mu I^{-1}F)(v_x)u_x = \langle I^{-1}F(v_x), u_x \rangle \\ &= (I \circ I^{-1}F(v_x), u_x) = (F(v_x), u_x), \end{aligned}$$

so,

$$\mathcal{F}(v_x)u_x = \sum_{i=1}^{n} (f_i(v_x), u_x{}^i), \quad \text{where} \quad u_x = (u_x{}^1, \dots, u_x{}^n). \tag{5.32}$$

Then $\mathcal{F}(v_x)u_x$ is the total work of the external forces $f_i(v_x)$ along $u_x{}^i$.

From the generalized Newton law (5.7) we have

$$\mathcal{F}(\dot{q})u_x = (\mu(\frac{D\dot{q}}{dt}))u_x = \mu(\ddot{q})u_x = \langle \ddot{q}, u_x \rangle = \sum_{i=1}^{n} (m_i \ddot{q}_i, u_x{}^i)$$

and (5.32) implies $\mathcal{F}(\dot{q})u_x = \sum_{i=1}^{n} (f_i(q, \dot{q}), u_x{}^i)$; so, since u_x is arbitrary in k^n one obtains the classical Newton's law:

$$m_i \ddot{q}_i = f_i(q, \dot{q}), \quad i = 1, \dots, n,$$

and conversely.

Example 2.5 - The planar double pendulum

One may think about two mass points (q_1, m_1) and (q_2, m_2), $q_i \in \mathbf{R}^2, i = 1, 2$, in the configuration space $Q = \mathbf{R}^2 \times \mathbf{R}^2 = \mathbf{R}^4$ and about a holonomic constraint N, the submanifold defined by the conditions:

$$\begin{aligned} | q_1 - 0 |^2 &= \ell_1{}^2 && (5.33) \\ | q_2 - q_1 |^2 &= \ell_2{}^2, && (5.34) \end{aligned}$$

where $0 \in \mathbf{R}^2$ is the origin.

If $a, b \in \mathbf{R}^2$, $a.b$ denotes the usual inner product of $\mathbf{R}^2$. Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ vectors in $\mathbf{R}^4$, that is, $u_i, v_i, \in \mathbf{R}^2, i = 1, 2$.

The spatial metric in $\mathbf{R}^4$ is given by

$$(u, v) = u_1.v_1 + u_2.v_2,$$

and

$$\langle u, v\rangle = m_1 u_1.v_1 + m_2 u_2.v_2$$

is the metric corresponding to the kinetic energy

$$K(\dot q) = \frac{1}{2}[m_1\dot q_1.\dot q_1 + m_2\dot q_2.\dot q_2], \quad \dot q = (\dot q_1, \dot q_2) \in \mathbf{R}^4.$$

The Levi-Civita connection ∇ associated to the metric $\langle,\rangle$ gives the acceleration of $q(t) = (q_1(t), q_2(t)) \in \mathbf{R}^4$ with Christoffel symbols equal to zero:

$$(\frac{D\dot q}{dt}) = \ddot q = (\ddot q_1, \ddot q_2). \tag{5.35}$$

The usual external forces acting on q_1 and q_2 are

$$F_1 = (0, m_1 g) \quad \text{and} \quad F_2 = (0, m_2 g),$$

respectively. As in the previous Example 1.5, one defines the field of external forces

$$\mathcal{F} : T(\mathbf{R}^2 \times \mathbf{R}^2) \to T^*(\mathbf{R}^2 \times \mathbf{R}^2)$$

using the total work of the physical external forces:

$$\mathcal{F}(\dot q)(u_1, u_2) = (F_1(\dot q), u_1) + (F_2(\dot q), u_2) \tag{5.36}$$

where $F_i(\dot q) = F_i = (0, m_i g), \quad i = 1, 2.$

Assuming that the submanifold N defined by (5.33) and (5.34) is a perfect constraint, that is, satisfies the d'Alembert principle, we have by (5.23) that for any C^2 curve compatible with N,

$$\mathcal{R}(\dot q) = \mu(\frac{D\dot q}{dt}) - \mathcal{F}(\dot q), \quad \mathcal{R}(\dot q) \in T^*_{q(t)}Q,$$

is such that the vector $\mu^{-1}(\mathcal{R}(\dot q))$ is, at the point $q(t) \in N$, orthogonal to $T_{q(t)}N$ with respect to the metric $\langle,\rangle$, for all t; that is,

$$\langle \mu^{-1}\mathcal{R}(\dot q), (v_1, v_2)\rangle = 0 \tag{5.37}$$

for all $(v_1, v_2) \in T_{q(t)}N$. But $(v_1, v_2) \in T_{q(t)}N$ means that v_1 and v_2 in $\mathbf{R}^2$ have to satisfy:

$$v_1.(q_1 - 0) = 0 \tag{5.38}$$
$$(v_2 - v_1).(q_2 - q_1) = 0 \tag{5.39}$$

where (5.38) and (5.39) were obtained by derivative, with respect to time, of (5.33) and (5.34), respectively. If one denotes

$$I\mu^{-1}\mathcal{R}(\dot q) \stackrel{def}{=} (R_1(\dot q), R_2(\dot q)), \tag{5.40}$$

condition (5.37) and the definitions (5.30) and (5.40) give

$$\begin{aligned} 0 &= \langle \mu^{-1}\mathcal{R}(\dot q), (v_1, v_2)\rangle = (I\mu^{-1}\mathcal{R}(\dot q), (v_1, v_2)) \\ &= ((R_1(\dot q), R_2(\dot q)), (v_1, v_2)) = (R_1(\dot q)).v_1 + (R_2(\dot q)).v_2 \end{aligned}$$

so, $R_1(\dot q)$ and $R_2(\dot q)$ defined in (5.40) satisfy

$$(R_1(\dot q)).v_1 + (R_2(\dot q)).v_2 = 0 \tag{5.41}$$

for all v_1, v_2 in $\mathbb{R}^2$ that verify (5.38) and (5.39).

From (5.35), and the definition of μ we obtain

$$\begin{aligned} \mu(\frac{D\dot q}{dt})(u_1, u_2) &= \langle(\frac{D\dot q}{dt}), (u_1, u_2)\rangle = \langle(\ddot q_1, \ddot q_2), (u_1, u_2)\rangle \\ &= m_1\ddot q_1.u_1 + m_2\ddot q_2.u_2. \end{aligned} \tag{5.42}$$

From (5.23), (5.36), (5.40) and (5.42) we have

$$\begin{aligned} m_1\ddot q_1.u_1 + m_2\ddot q_2.u_2 &= (F_1(\dot q)).u_1 + (F_2(\dot q)).u_2 + \mathcal{R}(\dot q)(u_1, u_2) \\ &= (F_1(\dot q)).u_1 + (F_2(\dot q)).u_2 + (R_1(\dot q)).u_1 + (R_2(\dot q)).u_2; \end{aligned}$$

in fact,

$$\begin{aligned} \mathcal{R}(\dot q)(u_1, u_2) &= \mu I^{-1}(R_1(\dot q), R_2(\dot q))(u_1, u_2) \\ &= \langle I^{-1}(R_1(\dot q), R_2(\dot q)), (u_1, u_2)\rangle \\ &= ((R_1(\dot q), R_2(\dot q)), (u_1, u_2)) \\ &= (R_1(\dot q)).u_1 + (R_2(\dot q)).u_2 \quad , \end{aligned}$$

then

$$m_1\ddot q_1.u_1 + m_2\ddot q_i u_2 = (F_1(\dot q) + R_1(\dot q)).u_1 + (F_2(\dot q) + R_2(\dot q)).u_2 \quad ;$$

and since $(u_1, u_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ is arbitrary (see (5.24)) we have

$$\begin{aligned} m_1\ddot q_1 &= F_1(\dot q) + R_1(\dot q) \\ m_2\ddot q_2 &= F_2(\dot q) + R_2(\dot q). \end{aligned} \tag{5.43}$$

Equations (5.43) are the classical Newton law for two mass points; $R_1(\dot q)$, $R_2(\dot q)$ are the constraint's reactions that have to satisfy (5.41) for all (v_1, v_2) such that (5.38) and (5.39) hold, that is, "the virtual work of the reactive forces is equal to zero (classical d'Alembert principle)".

Exercise 3.5 Show that (5.41) for all (v_1, v_2) such that (5.38) and (5.39) hold is equivalent to

$$\begin{aligned} R_2(\dot q) &= \rho(q_2 - q_1) \\ R_1(\dot q) + R_2(\dot q) &= \alpha(q_1 - 0), (\rho, \alpha \in \mathbb{R}). \end{aligned}$$

Let us derive now the **Lagrange equations** (5.20) corresponding to the generalized Newton law (5.26) for the planar double pendulum. From (5.36) the field of external forces is given by

$$\mathcal{F}(\dot{q})(u_1, u_2) = (F_1(\dot{q}), u_1) + (F_2(\dot{q}), u_2) = (m_1 u_1^y + m_2 u_2^y)g$$

provided that $u_1 = (u_1^x, u_1^y)$ and $u_2 = (u_2^x, u_2^y)$.

The function $U : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ defined by

$$U(q_1, q_2) = -m_1 g y_1 - m_2 g y_2, \quad \text{where} \quad q_1 = (x_1, y_1)$$

and $q_2 = (x_2, y_2)$, is such that $\mathcal{F}(v_p) = -dU(p), v_p \in T_p\mathbf{R}^4$.

So, $\mathcal{F}$ is a conservative field of forces. The manifold N is a torus with coordinates (φ, θ), so, the potential energy U and the kinetic energy K restricted to N are $\bar{U}$ and $\bar{K}$ respectively:

$$\begin{aligned}
\bar{U} &= -m_1 g \ell_1 \cos\theta - m_2 g(\ell_1 \cos\theta + \ell_2 \cos\varphi) \\
\bar{K} &= \frac{1}{2}[m_1(\dot{q}_1, \dot{q}_1) + m_2(\dot{q}_2, \dot{q}_2)] = \frac{1}{2}\sum_{i=1}^{2} m_i(\dot{x}_i^2 + \dot{y}_i^2)
\end{aligned}$$

where $\dot{q}_1 = (\dot{x}_1, \dot{y}_1)$ and $\dot{q}_2 = (\dot{x}_2, \dot{y}_2)$ for $x_1 = \ell_1 \sin\theta$, $y_1 = \ell_1 \cos\theta$, $x_2 = \ell_1 \sin\theta + \ell_2 \sin\varphi$, $y_2 = \ell_1 \cos\theta + \ell_2 \cos\varphi$. Then $\dot{x}_1 = \ell_1 \dot{\theta} \cos\theta$, $\dot{y}_1 = -\ell_1 \dot{\theta} \sin\theta$, $\dot{x}_2 = \ell_1 \dot{\theta} \cos\theta + \ell_2 \dot{\varphi} \cos\varphi$, $\dot{y}_2 = -\ell_1 \dot{\theta} \sin\theta - \ell_2 \dot{\varphi} \sin\varphi$ and consequently:

$$\frac{\partial \bar{U}}{\partial \theta} = (m_1 + m_2) g \ell_1 \sin\theta, \qquad \frac{\partial \bar{U}}{\partial \varphi} = m_2 g \ell_2 \sin\varphi,$$

$$\begin{aligned}
\frac{\partial \bar{K}}{\partial \theta} &= m_1 \dot{x}_1 \frac{\partial \dot{x}_1}{\partial \theta} + m_1 \dot{y}_1 \frac{\partial \dot{y}_1}{\partial \theta} + m_2 \dot{x}_2 \frac{\partial \dot{x}_2}{\partial \theta} + m_2 \dot{y}_2 \frac{\partial \dot{y}_2}{\partial \theta} \\
&= m_1 \ell_1 \dot{\theta} \cos\theta(-\ell_1 \dot{\theta} \sin\theta) + m_1 \ell_1 \dot{\theta} \sin\theta(\ell_1 \dot{\theta} \cos\theta) \\
&+ m_2(\ell_1 \dot{\theta} \cos\theta + \ell_2 \dot{\varphi} \cos\varphi)(-\ell_1 \dot{\theta} \sin\theta) \\
&+ m_2(\ell_1 \dot{\theta} \sin\theta + \ell_2 \dot{\varphi} \sin\varphi) \ell_1 \dot{\theta} \cos\theta,
\end{aligned}$$

so,

$$\frac{\partial \bar{K}}{\partial \theta} = m_2 \ell_1 \ell_2 \dot{\varphi} \dot{\theta} \sin(\varphi - \theta).$$

$$\begin{aligned}
\frac{\partial \bar{K}}{\partial \varphi} &= m_2 \dot{x}_2 \frac{\partial \dot{x}_2}{\partial \varphi} + m_2 \dot{y}_2 \frac{\partial \dot{y}_2}{\partial \varphi} \\
&= m_2(\ell_1 \dot{\theta} \cos\theta + \ell_2 \dot{\varphi} \cos\varphi)(-\ell_2 \dot{\varphi} \sin\varphi) \\
&+ m_2(\ell_1 \dot{\theta} \sin\theta + \ell_2 \dot{\varphi} \sin\varphi) \ell_2 \dot{\varphi} \cos\varphi,
\end{aligned}$$

so,

$$\frac{\partial \bar{K}}{\partial \varphi} = m_2 \ell_1 \ell_2 \dot{\varphi} \dot{\theta} \sin(\theta - \varphi).$$

$$\begin{aligned}\frac{\partial \bar{K}}{\partial \dot{\theta}} &= m_1 {\ell_1}^2 \dot{\theta} \cos^2\theta + m_1 {\ell_1}^2 \dot{\theta} \sin^2\theta + m_2(\ell_1 \dot{\theta}\cos\theta + \ell_2 \dot{\varphi}\cos\varphi)\ell_1 \cos\theta \\ &+ m_2(\ell_1 \dot{\theta}\sin\theta + \ell\dot{\varphi}\sin\varphi)\ell_1 \sin\theta,\end{aligned}$$

so,

$$\frac{\partial \bar{K}}{\partial \dot{\theta}} = m_1 {\ell_1}^2 \dot{\theta} + m_2 \ell_1^2 \dot{\theta} + m_2 \ell_1 \ell_2 \dot{\varphi}\cos(\theta - \varphi).$$

$$\begin{aligned}\frac{\partial \bar{K}}{\partial \dot{\varphi}} &= m_2(\ell_1 \dot{\theta}\cos\theta + \ell_2 \dot{\varphi}\cos\varphi)\ell_2 \cos\varphi \\ &+ m_2(\ell_1 \dot{\theta}\sin\theta + \ell_2 \dot{\varphi}\sin\varphi)\ell_2 \sin\varphi,\end{aligned}$$

so,

$$\frac{\partial \bar{K}}{\partial \dot{\varphi}} = m_2 {\ell_2}^2 \dot{\varphi} + m_2 \ell_1 \ell_2 \dot{\theta}\cos(\theta - \varphi).$$

The two Lagrange's equations are

$$\frac{d}{dt}\frac{\partial \bar{K}}{\partial \dot{\theta}} - \frac{\partial \bar{K}}{\partial \theta} = -\frac{\partial \bar{U}}{\partial \theta}, \qquad \frac{d}{dt}\frac{\partial \bar{K}}{\partial \dot{\varphi}} - \frac{\partial \bar{K}}{\partial \varphi} = -\frac{\partial \bar{U}}{\partial \varphi}$$

and then

$$\begin{aligned}\frac{d}{dt}[m_1 {\ell_1}^2 \dot{\theta} + m_2 {\ell_1}^2 \dot{\theta} + m_2 \ell_1 \ell_2 \dot{\varphi}\cos(\theta - \varphi)] &- m_2 \ell_1 \ell_2 \dot{\varphi}\dot{\theta}\sin(\theta - \varphi) \\ &= -(m_1 + m_2) g \ell_1 \sin\theta \\ \frac{d}{dt}[m_2 {\ell_2}^2 \dot{\varphi} + m_2 \ell_1 \ell_2 \dot{\theta}\cos(\theta - \varphi)] &- m_2 \ell_1 \ell_2 \dot{\varphi}\dot{\theta}\sin(\theta - \varphi) \\ &= -m_2 g \ell_2 \sin\varphi.\end{aligned}$$

These two equations determine a second order system of ordinary differential equations on the torus of coordinates (θ, φ):

$$\begin{aligned}(m_1 + m_2)\ell_1^2 \ddot{\theta} &+ m_2 \ell_1 \ell_2 [\ddot{\varphi}\cos(\theta - \varphi) - \dot{\varphi}(\dot{\theta} - \dot{\varphi})\sin(\theta - \varphi)] - \\ &- m_2 \ell_1 \ell_2 \dot{\varphi}\dot{\theta}\sin(\theta - \varphi) + (m_1 + m_2) g \ell_1 \sin\theta = 0, \quad (5.44)\end{aligned}$$

$$\begin{aligned}m_2 \ell_2^2 \ddot{\varphi} &+ m_2 \ell_1 \ell_2 [\ddot{\theta}\cos(\theta - \varphi) - \dot{\theta}(\dot{\theta} - \dot{\varphi})\sin(\theta - \varphi)] - \\ &- m_2 \ell_1 \ell_2 \dot{\varphi}\dot{\theta}\sin(\theta - \varphi) + m_2 g \ell_2 \sin\varphi = 0. \qquad (5.45)\end{aligned}$$

One can compute $\ddot{\theta}$ and $\ddot{\varphi}$ in (5.44) and (5.45) and get a system of two ordinary differential equations in the normal form; in fact the matrix

$$\begin{bmatrix} (m_1+m_2)\ell_1^2 & m_2\ell_1\ell_2\cos(\theta-\varphi) \\ & \\ m_2\ell_1\ell_2\cos(\theta-\varphi) & m_2\ell_2^2 \end{bmatrix}$$

is positive definite, with determinant equal to

$$m_1 m_2 \ell_1^2 \ell_2^2 + m_2^2 \ell_1^2 \ell_2^2 \sin^2(\theta-\varphi) > 0.$$

The mechanical energy $E_m = \bar{K} + \bar{U}$ is a first integral of system (5.44), (5.45) (see Proposition 1.5) expressed as:

$$\begin{aligned} E_m &= \frac{1}{2}(m_1+m_2)\ell_1^2\dot{\theta}^2 + \frac{1}{2}m_2\ell_2^2\dot{\varphi}^2 + m_2\ell_1\ell_2\dot{\theta}\dot{\varphi}\cos(\theta-\varphi) - \\ &- (m_1+m_2)g\ell_1\cos\theta - m_2 g\ell_2\cos\varphi. \end{aligned}$$

The critical points are the zero vectors $0_p \in T_pN$ such that $d\bar{U}(p) = 0$, that is, $\frac{\partial \bar{U}}{\partial \theta}(p) = \frac{\partial \bar{U}}{\partial \varphi}(p) = 0$, or, equivalently, $p = (\theta, \varphi)$ such that $\sin\theta = \sin\varphi = 0$; so, one has 4 critical configurations on the torus N:

$$p_1 = (0,0), \quad p_2 = (0,\pi), \quad p_3 = (-\pi, 0) \ \text{ and } \ p_4 = (\pi,\pi).$$

5.6 The dynamics of rigid bodies

Let K and k be two oriented euclidean vector spaces also considered as affine spaces associated to K and k, respectively. Assume that both spaces have dimension 3 so, each one has well defined the vector product operation (denoted by $\times$) corresponding to the inner product $(\, ,)$.

An **isometry** $M : K \to k$ is a distance preserving map, that is, $\|X - Y\| = \|MX - MY\|$ for all $X, Y \in K$. The induced map $M^* : K \to k$ is defined by: ($0 \in K$ is the zero vector)

$$M^*X = M(X) - M(0), \qquad \text{for all} \quad X \in K \tag{5.46}$$

Proposition 5.5 Let M^* be the induced map of an isometry M. Then one has the following:

1. M^* is modulus preserving.

2. M^* preserves inner product and is linear.

3. M^* is a bijection, so M is an affine (bijective) transformation.

4. The inverse of M is an isometry.

5. If M^* is orientation preserving then M^* preserves vector product.

Proof

1. $\|M^*X\| = \|M(X) - M(0)\| = \|X - 0\| = \|X\|$.

2.

$$\begin{aligned}(M^*X, M^*Y) &= \frac{1}{2}(\|M^*X\|^2 + \|M^*Y\|^2 - \|M^*X - M^*Y\|^2) \\ &= \frac{1}{2}(\|X\|^2 + \|Y\|^2 - \|X - Y\|^2) = (X, Y)\end{aligned}$$

so, M^* preserves inner product. Moreover M^* is linear because: for any $\alpha \in \mathbb{R}$ and $X \in K$ we have

$$\begin{aligned}\|M^*(\alpha X) - \alpha M^*X\|^2 &= \|M^*(\alpha X)\|^2 + \alpha^2\|M^*X\|^2 - 2(M^*(\alpha x), \alpha M^*X) \\ &= \|\alpha X\|^2 + \alpha^2\|X\|^2 - 2\alpha(M^*(\alpha X), M^*X) \\ &= 2\alpha^2\|X\|^2 - 2\alpha(\alpha X, X) = 0;\end{aligned}$$

and also:

$$\begin{aligned}&\|M^*(X - Y) - (M^*X - M^*Y)\|^2 = \\ = \; &\|M^*(X - Y)\|^2 + \|M^*X - M^*Y\|^2 - 2(M^*(X - Y), M^*X - M^*Y) = \\ = \; &\|X - Y\|^2 + \|X - Y\|^2 - 2(X - Y, X) + 2(X - Y, Y) = 0,\end{aligned}$$

that shows that M^* is linear.

3. Since M^* is linear, it is enough to prove that M^* is an injection; but if $M^*X = o$ ($o \in k$ is the zero vector) one has $\|M^*X\| = \|X\| = 0$, so $X = 0$ and M^* has an inverse $(M^*)^{-1}$.

4. The map $N : k \to K$ defined by

$$N(x) = (M^*)^{-1}(x - M(0)) \quad \text{for all} \quad x \in k, \tag{5.47}$$

is the inverse of M since by (5.46) and (5.47) we have:

$$M(N(x)) = M(0) + M^*(N(x)) = M(0) + (x - M(0)) = x$$

But (5.47) gives $N(o) = -(M^*)^{-1}(M(0))$, so,

$$N(x) = (M^*)^{-1}x - (M^*)^{-1}(M(0)) = N(o) + (M^*)^{-1}x \tag{5.48}$$

and N is an isometry with $N^* = (M^*)^{-1}$ as induced map. In fact (5.48) shows that $N^* = (M^*)^{-1}$ and (5.47) implies:

$$\|N(x)-N(y)\| = \|(M^*)x-(M^*)^{-1}y\| = \|M^*(M^*)^{-1}x-M^*(M^*)^{-1}y\| = \|x-y\|,$$

so N preserves distances. ∎

Exercise 4.5 Prove property 5. in Proposition 5.5.

An isometry $M : K \to k$ is said to be a **proper isometry** if its induced map $M^* : K \to k$ is orientation preserving.

A rigid motion of K relative to k is a C^2 curve

$$M : t \longmapsto M_t$$

where M_t is a proper isometry. If, moreover, $M_t(0) = o$ for all t, then M is said to be a **rotation**.

Proposition 6.5 *Any rigid motion M of K relative to k is such that M_t has a unique decomposition $M_t = T_t \circ R_t$ where $R_t = M_t{}^* : K \to k$ defines a rotation and $T_t : k \to k$ is given by $T_t x = x + r(t)$, that is, T_t is a translation in k, for each t.*

Proof From (5.46) we have:

$$\begin{aligned} M_t(X) &= M_t{}^*(X-0) + M_t(0) = M_t{}^*X + M_t(0) \\ &= R_tX + M_t(0) = T_t(R_tX) = (T_t \circ R_t)X \end{aligned}$$

where $T_t(x) \stackrel{def}{=} x + r(t)$ for all $x \in k$, $\quad r(t) \stackrel{def}{=} M_t(0)$. If $M_t = \bar{T}_t \circ \bar{R}_t$ is another decomposition such that $\bar{T}_t(x) = x + \bar{r}(t)$ for all $x \in k$ and $\bar{R}_t 0 = o$ then $\bar{T}_t(\bar{R}_tX) = T_t(M_t{}^*X)$ or $\bar{R}_tX + \bar{r}(t) = M_t^*X + r(t)$ for all $X \in K$; in particular for $X = 0$ one gets $r(t) = \bar{r}(t)$ and consequently $\bar{R}_t = M_t{}^*$. ∎

A rigid motion M is said to be **translational** if in the (unique) decomposition $M_t = T_t \circ M_t{}^*$, the linear isometry $M_t{}^*$ does not depend on t, that is, $M_t{}^* = M_{t_o}{}^*$ for some t_o. In that case we have $M_t(X) = M_{t_o}{}^*X + r(t)$.

We will derive now, the expression that describes the kinematics of a rigid motion M of a **(moving)** system K with respect to a **(stationary)** system k, that is, for t in some interval I of the real line, $M_t : K \to k$ is the corresponding proper isometry. Let us denote by $Q(t) \in K$ a moving C^2 radius vector also defined in I and let $q(t) = M_t(Q(t))$ be the radius vector, in k, corresponding to the action of M_t on the moving point $Q(t)$. Let us denote by $r(t) \in k$ the vector $r(t) = M_t(0)$.

Taking into account that $M_t(X) = M_t{}^* X + M_t(0)$ for all $X \in K$ one obtains:

$$q(t) = M_t(Q(t)) = M_t^* Q(t) + r(t). \tag{5.49}$$

By derivative of (5.49) with respect to time one has

$$\dot q(t) = \dot M_t^* Q(t) + M_t^* \dot Q(t) + \dot r(t). \tag{5.50}$$

Special cases:

a) If the rigid motion M is **translational**, that is, $M_t{}^* = M_{t_o}{}^*$ for all t, one obtains from (5.50) that

$$\dot q(t) = M_{t_o}{}^* \dot Q(t) + \dot r(t) \tag{5.51}$$

and so, the **absolute velocity** $\dot q(t)$ is equal to the sum of the **relative velocity** $M_{t_o}{}^* \dot Q(t)$ with the **velocity** $\dot r(t)$ (of the origin 0) **of the moving system** K.

b) If the rigid motion M is a **rotation** of the moving system K with respect to the stationary system k, that is, if $r(t) = 0$ for all t, one obtains from (5.49):

$$q(t) = M_t{}^* Q(t) \quad \text{and} \quad \dot q(t) = \dot M_t{}^* Q(t) + M_t{}^* \dot Q(t). \tag{5.52}$$

If, moreover, $Q(t) = \xi =$ constant, (5.52) shows that

$$q(t) = M_t{}^* \xi \qquad \text{for all} \quad t. \tag{5.53}$$

and the motion of $q(t)$ is called a **transferred rotation of** ξ.

Exercise 5.5 Assume it is given a skew-symmetric linear operator $A : V \to V$ acting on an oriented 3-dimensional euclidean vector space V. Then there exists a unique vector $\omega \in V$ such that $Ay = \omega \times y$ for all $y \in V$. Also $\omega = 0$ if and only if $A = 0$. We use to denote simply $A = \omega \times$.

Let us consider the induced linear map $M_t{}^*$ associated to a rigid motion $M : t \to M_t$ of K with respect to k. One can construct two linear operators (depending C^1 on time):

$$\dot M_t^* (M_t{}^*)^{-1} : k \to k \quad \text{and} \quad (M_t{}^*)^{-1} \dot M_t^* : K \to K.$$

From Proposition 5.5 (2. and 3.) $M_t{}^*$ is a linear isometry:

$$(M_t^* X, M_t^* Y) = (X, Y), \quad \text{for all} \quad X, Y \in K. \tag{5.54}$$

By derivative of (5.54) with respect to time we obtain

$$(\dot{M}_t^* X, M_t^* Y) + (M_t^* X, \dot{M}_t^* Y) = 0, \quad \text{for all} \quad X, Y \in K. \tag{5.55}$$

Since $(M_t^*)^{-1}$ is also a linear isometry one gets from (5.55) that

$$((M_t^*)^{-1}\dot{M}_t^* X, Y) + (X, (M_t^*)^{-1}\dot{M}_t^* Y)) = 0, \quad \text{for all} \quad X, Y \in K \tag{5.56}$$

and also

$$(\dot{M}_t^*(M_t^*)^{-1}x, y) + (x, \dot{M}_t^*(M_t^*)^{-1}y) = 0 \quad \text{for all} \quad x, y \in k, \tag{5.57}$$

where $x = M_t^* X$ and $y = M_t^* Y$ are arbitrary in k. Then (5.56) and (5.57) show that $(M_t^*)^{-1}\dot{M}_t^*$ and $\dot{M}_t^*(M_t^*)^{-1}$ are skew-symmetric linear operators acting on K and k, respectively. Using the result of Exercise 5.5 above one can state the following:

Proposition 7.5 *Let $M : t \to M_t$ be a rigid motion of K with respect to k and M_t^* its induced linear isometry. Then there exist unique vectors $\Omega(t) \in K$ and $\omega(t) \in k$ such that $(M_t^*)^{-1}\dot{M}_t^* = \Omega(t)\times$ and $\dot{M}_t^*(M_t^*)^{-1} = \omega(t)\times$. Moreover $\omega(t) = M_t^*\Omega(t)$.*

Proof We only need to prove that $\omega(t) = M_t^*\Omega(t)$. But from the definition of $\Omega(t)$ we know that

$$(M_t^*)^{-1}\dot{M}_t^* Y = \Omega(t) \times Y \quad \text{for all} \quad Y \in K;$$

so, making $Y = (M_t^*)^{-1}y$, one obtains

$$(M_t^*)^{-1}\dot{M}_t^*(M_t^*)^{-1}y = \Omega(t) \times (M_t^*)^{-1}y,$$

and then

$$\dot{M}_t^*(M_t)^{-1}y = M_t^*[\Omega(t) \times (M_t^*)^{-1}y].$$

The last expression and Proposition 5.5 (5.) show that

$$\dot{M}_t^* M_t^{-1} y = [M_t^*\Omega(t)] \times y \quad \text{for all} \quad y \in k,$$

thus the definition and the uniqueness of $\omega(t)$ enable us to conclude the result. ∎

We will give now the interpretation of $\omega(t)$ and $\Omega(t)$ when we are dealing with the special cases considered above. Start with a rotation M ($r(t) = 0$ for all t) such that $Q(t) = \xi =$ constant, that is, the motion of $q(t)$ is a transferred rotation of $\xi \in K$. We have the following result:

Proposition 8.5 *If $q(t)$ is a transferred rotation of ξ, to each time t for which $\dot{M}_t^* \neq o$ there corresponds an axis of rotation, that is, a line in k through the* *origin whose points have zero velocity at that time, and at the same time, each point out of the axis of rotation has velocity orthogonal to the axis with the modulus proportional to the distance from the point to the mentioned axis; if, otherwise, we have $\dot{M}_t^* = 0$, then all the points in k have zero velocity at this time t.*

Proof By (5.53) we have

$$\dot{q}(t) = \dot{M}_t^* \xi. \tag{5.58}$$

If $\dot{M}_t^* = 0$, (5.58) shows that $\dot{q}(t) = 0$. Assume otherwise $\dot{M}_t^* \neq 0$; in this last case (5.53) and (5.58) imply that

$$\dot{q}(t) = \dot{M}_t^*(M_t{}^*)^{-1} q(t). \tag{5.59}$$

One sees that the linear operator $\dot{M}_t^*(M_t{}^*)^{-1} : k \to k$ is non zero and skew-symmetric; in fact $\dot{M}_t^*(M_t{}^*)^{-1} = 0$ implies $\dot{M}_t^* = 0$ (contradiction). From Proposition 7.5 $\dot{M}_t^*(M_t{}^*)^{-1}$ is skew-symmetric and moreover, there exists a unique non zero vector $\omega(t) \in k$ such that

$$\dot{M}_t^*(M_t{}^*)^{-1} = \omega(t)\times; \tag{5.60}$$

then equations (5.59) and (5.60) imply that

$$\dot{q}(t) = \omega(t) \times q(t). \tag{5.61}$$

The **instantaneous axis of rotation** at the time t is the line in k through the origin and direction $\rho\omega(t)$, $\rho \in \mathbf{R}$, and (5.61) shows that $|\dot{q}(t)| = |\omega(t)| \ |q(t)| \sin\theta$ where $|q(t)| \ \sin\theta$ is the distance from $q(t)$ to the axis of rotation . ∎

Another case to be considered is a general rotation ($r(t) = 0$ for all t); so equations (5.52) imply

$$\dot{q}(t) = \dot{M}_t^*(M_t{}^*)^{-1} q(t) + M_t{}^* \dot{Q}(t) \tag{5.62}$$

and using Proposition 7.5 there exists a unique $\omega(t) \in k$ so that equation (5.62) can be written

$$\dot{q}(t) = \omega(t) \times q(t) + M_t^* \dot{Q}(t). \tag{5.63}$$

So, for a rotation M, the **absolute velocity** $\dot{q}(t)$ is equal to the sum of the **relative velocity** $M_t{}^*\dot{Q}(t)$ and the **transferred velocity of rotation** $\omega(t) \times q(t)$.

The **dynamics of mass points in a non-inertial frame** can be studied by assuming that k is an inertial and that K is a non-inertial coordinate system subjected to a rigid motion $M : t \to M_t$. From (5.50) we know that $\dot{q}(t) = \dot{M}_t^* Q(t) + M_t{}^* \dot{Q}(t) + \dot{r}(t)$. Let us suppose also that the motion of the point $q \in k$ with mass $m > 0$ satisfies the Newton's equation

$$m\ddot{q} = f(q, \dot{q}); \tag{5.64}$$

so we have:

$$f(q,\dot{q}) = m\ddot{q} = m[\ddot{M}_t^* Q(t) + 2\dot{M}_t^* \dot{Q}(t) + M_t{}^* \ddot{Q}(t) + \ddot{r}(t)]. \tag{5.65}$$

The special case in which M is translational ($M_t^* = M_{t_o}^* =$ constant) implies that

$$mM_{t_o}^* \ddot{Q}(t) = m(\ddot{q} - \ddot{r}) = f(q,\dot{q}) - m\ddot{r}(t)$$

or

$$m\ddot{Q}(t) = (M_{t_o}{}^*)^{-1} f(q,\dot{q}) - (M_{t_o}{}^*)^{-1} m\ddot{r}(t).$$

The case in which M is a rotation ($r(t) = 0$ for all t) gives from (5.65):

$$m\ddot{Q}(t) = (M_t{}^*)^{-1}[f(q,\dot{q}) - m\ddot{M}_t^* Q(t) - 2m\dot{M}_t^* \dot{Q}(t)],$$

so

$$m\ddot{Q}(t) = (M_t{}^*)^{-1} f(q,\dot{q}) - 2m\Omega(t) \times \dot{Q}(t) - m(M_t{}^*)^{-1}\ddot{M}_t^* Q(t). \tag{5.66}$$

From the definition of $\Omega(t)$ we have

$$(M_t{}^*)^{-1}\dot{M}_t^* Y = \Omega(t) \times Y \qquad \text{or}$$

$$\dot{M}_t^* Y = M_t{}^*(\Omega(t) \times Y) \quad \text{for all} \quad Y \in K; \tag{5.67}$$

The derivative of (5.67) gives

$$\ddot{M}_t^* Y = \dot{M}_t^*(\Omega(t) \times Y) + M_t{}^*(\dot{\Omega}(t) \times Y)$$

and so,

$$(M_t{}^*)^{-1}\ddot{M}_t^* Y = \Omega(t) \times (\Omega(t) \times Y) + \dot{\Omega}(t) \times Y$$

for all $Y \in K$ and, in particular, for $Y = Q(t)$, that is,

$$(M_t{}^*)^{-1}\ddot{M}_t^* Q(t) = \Omega(t) \times (\Omega(t) \times Q(t)) + \dot{\Omega}(t) \times Q(t)$$

and this last equality can be introduced in (5.66) giving, after setting $(M_t{}^*)^{-1} f(q,\dot{q}) = F(t,q,\dot{q})$:

$$m\ddot{Q}(t) = -m\Omega(t) \times (\Omega(t) \times Q(t)) - 2m\Omega(t) \times \dot{Q}(t) - m\dot{\Omega}(t) \times Q(t) + F(t,q,\dot{q})$$

where one calls
$F_1 = -m\dot{\Omega}(t) \times Q(t)$: **the inertial force of rotation,**
$F_2 = -2m\Omega(t) \times \dot{Q}(t)$: **the Coriolis force,**
$F_3 = -m\Omega(t) \times (\Omega(t) \times Q(t))$: **the centrifugal force.**

Thus one can state the following:

Proposition 9.5 *The motion in a (non inertial) rotating coordinate system takes place as if three additional inertial forces (the inertial force of rotation*

F_1, the Coriolis force F_2 and the centrifugal force F_3) together with the external force $F(t,q,\dot q) = (M_t{}^*)^{-1} f(q,\dot q)$ acted on every moving point $Q(t)$ of mass m.

For the purposes of giving a mathematical definition of a rigid body, we start saying that a **body** is a bounded borelian $S \subset K$. Since the physical meaning of a body corresponds to the consideration of a system of mass points, the **rigid body** is a body in which the distance between points is constant; so, the characterization of a rigid body $S \subset K$ is done saying that S is a bounded connected Borel set in K such that during the action of any rigid motion $M : t \mapsto M_t$ of K relative to k, the points $\xi \in S$ do not move, that is

$$Q(t,\xi) = \xi \quad \text{for any } t \quad \text{and any } \xi \in S. \tag{5.68}$$

The distribution of the masses on S will be considered in the sequel. Without loss of generality one assumes, from now on, that the origin O of K belongs to S.

A rigid motion M of K relative to k induces, by restriction, a motion of S relative to k, and, when S is a rigid body, we have from (5.49) and (5.68):

$$q(t,\xi) = M_t(Q(t,\xi)) = M_t(\xi) = M_t{}^*\xi + r(t) \tag{5.69}$$

for any t and any $\xi \in S$.

If a rigid motion is a rotation ($r(t) \equiv 0$), its action on the rigid body S is given, from (5.69), by the equation

$$q(t,\xi) = M_t{}^*\xi, \quad \text{for all} \quad \xi \in S, \tag{5.70}$$

that is, by a transferred rotation of each $\xi \in S$; so, a rotation acting on a rigid body S is said to be a **motion of S with a fixed point**, the origin $0 \in K$, since $r(t) = M_t(0) = o$. At each instant t, either the image $M_t(S)$ of S has an instantaneous axis of rotation passing through $o \in k$, the points $q(t,\xi) \in M_t(S)$ with velocities $\omega(t) \times q(t,\xi)$, or all the points of $M_t(S)$ have zero velocity, according what states Proposition 8.5 above.

If M is translational ($M_t{}^* = M_{t_o}{}^*$ for all t), its action on a rigid body S is given, from (5.69) by the equation

$$q(t,\xi) = M_{t_o}{}^*\xi + r(t) = M_{t_o}{}^*\xi + M_t(0)$$

so, $\dot q(t,\xi) = \dot r(t)$, that is, the velocity of any point of $M_t(S)$ is equal to the velocity $\dot r(t)$ of $M_t(0)$.

We go to introduce now the notions of **mass, center of mass, kinetic energy** and **kinetic momentum** of a rigid body S.

A **distribution of mass** on a rigid body S is defined through a **positive scalar measure** m on K; the following hypothesis is often used:

$$\text{"}m(U) > 0 \ \ \textit{for all non empty open set } U \textit{ of } S\text{"}. \tag{5.71}$$

The **center of mass of** S corresponding to a distribution of mass m is the point $G \in K$ given by

$$G = \frac{1}{m(S)} \int_S \xi dm(\xi) \tag{5.72}$$

where $m(S)$ is the **total mass** of the rigid body S which is a positive number (see the fundamental hypothesis).

Under the action of a rigid motion $t \to M_t$, the center of mass describes a curve in k given by:

$$g(t) \stackrel{def}{=} M_t(G) = \frac{1}{m(S)} \int_S M_t \xi dm(\xi) = \frac{1}{m(S)} \int_S q(t,\xi) dm(\xi) \tag{5.73}$$

Proposition 10.5 *The velocity $\dot{q}(t,\xi)$ of a point ξ of a given rigid body S under the action of a rigid motion $t \to M_t$ is given by*

$$\dot{q}(t,\xi) = \dot{g}(t) + \omega(t) \times [q(t,\xi) - g(t)]$$

where $\omega(t)\times = \dot{M}_t^*(M_t{}^*)^{-1}$.

Proof By (5.68) and (5.69) we have for all $\xi \in K$:

$$q(t,\xi) = M_t{}^*\xi + r(t) \quad \text{and} \quad \xi = (M_t{}^*)^{-1}[q(t,\xi) - r(t)];$$

so, by derivative one obtains:

$$\begin{aligned} \dot{q}(t,\xi) &= \dot{M}_t^*\xi + \dot{r}(t) = \dot{M}_t^*(M_t{}^*)^{-1}[q(t,\xi) - r(t)] + \dot{r}(t) \quad \text{or} \\ \dot{q}(t,\xi) &= \omega(t) \times [q(t,\xi) - r(t)] + \dot{r}(t), \quad \text{for all } \xi \in K. \end{aligned} \tag{5.74}$$

Choosing $\xi = G$ we get

$$\dot{g}(t) = \omega(t) \times [g(t) - r(t)] + \dot{r}(t); \tag{5.75}$$

then (5.74) and (5.75) prove the result. ∎

The **kinetic energy**of the motion of a rigid body S at a certain time t is, by definition,

$$K^c(t) = \frac{1}{2} \int_S |\, \dot{q}(t,\xi) \,|^2 dm(\xi) \tag{5.76}$$

The vectors $\omega(t)$ and $\Omega(t) = (M_t^*)^{-1}\omega(t)$ characterized by the equalities $\dot{M}_t^*(M_t^*)^{-1} = \omega(t)\times$ and $(M_t^*)^{-1}\dot{M}_t^* = \Omega(t)\times$ are called the **instantaneous angular velocities relative to** k and K, respectively.

The **kinetic momentum relative to k** of the motion of S at a certain time t is the vector

$$p(t) = \int_S [q(t,\xi) \times \dot{q}(t,\xi)]dm(\xi) \tag{5.77}$$

and the **kinetic momentum relative to the body** is

$$P(t) = (M_t{}^*)^{-1}p(t) \tag{5.78}$$

Special case: **rigid body with a fixed point.**

In that case $r(t) = 0$ for all t and then:

$$\begin{aligned}
q(t,\xi) &= M_t^*\xi, \qquad \dot{q}(t,\xi) = \omega(t) \times q(t,\xi); \\
K^c(t) &= \frac{1}{2}\int_S |\,\omega(t) \times q(t,\xi)|^2 dm(\xi) = \frac{1}{2}\int_S |\Omega(t) \times \xi|^2 dm(\xi) \\
p(t) &= \int_S [M_t^*\xi \times (\omega(t) \times M_t^*\xi)]dm(\xi); \\
P(t) &= \int_S [\xi \times (\Omega(t) \times \xi)]dm(\xi).
\end{aligned} \tag{5.79}$$

The last expression in (5.79) suggests how to give a definition for the **inertia operator of a rigid body** S:

$$A : X \in K \longmapsto [\int_S \xi \times (X \times \xi)dm(\xi)] \in K. \tag{5.80}$$

Proposition 11.5 *The inertia operator A of a rigid body $S \subset K$ is symmetric and positive with respect to the inner product of K. If, moreover, S has at least two points whose radii vectors are linearly independent and the distribution of mass satisfies (5.71), then A is positive definite.*

Proof

$$(AX, Y) = (Y, \int_S \xi \times (X \times \xi)dm(\xi)) = \int_S (Y, \xi \times (X \times \xi))dm(\xi)$$

then

$$(AX, Y) = \int_S (X \times \xi, Y \times \xi)dm(\xi) = (X, AY), \tag{5.81}$$

so A is symmetric. Assume now that $(AY, Y) = \int_S |Y \times \xi|^2 dm(\xi) = 0$. This implies that the set $E = \{\xi \in S \,|\, |Y \times \xi| \neq 0\}$ has measure $m(E) = 0$. On the other hand there exist $a, b \in S$ and neighborhoods U_a, U_b in K of a and b, such that v_1, v_2 are linearly independent for all $v_1 \in U_a$ and $v_2 \in U_b$. From the hypothesis on the measure m we have $m(U_a \cap S) > 0$ and $m(U_b \cap S) > 0$; so, there exist $u \in U_a \cap S$ and $v \in U_b \cap S$ such that $u, v \notin E$, that is $|Y \times u| = |Y \times v| = 0$, that imply $Y = 0$, that is, A is positive definite. ∎

If we come back to the special case of the motion of a rigid body S with a fixed point $O \in K$, we have from (5.79):

$$\begin{aligned} P(t) &= A\Omega(t) \\ K^c(t) &= \frac{1}{2}(A\Omega(t), \Omega(t)). \end{aligned} \tag{5.82}$$

In fact,

$$\begin{aligned} K^c(t) &= \frac{1}{2}\int_S |\Omega(t) \times \xi|^2 dm(\xi) = \frac{1}{2}\int_S (\Omega(t), \xi \times (\Omega(t) \times \xi)) dm(\xi) \\ &= \frac{1}{2}(\Omega(t), \int_S \xi \times (\Omega(t) \times \xi) dm(\xi)) \\ &= \frac{1}{2}(\Omega(t), A\Omega(t)). \end{aligned}$$

Another remark on the inertia operator A is the following: since A is linear and symmetric, there exists an orthonormal basis (E_1, E_2, E_3) in K where E_i is an eigenvector of a (real) eigenvalue I_i of A; since A is positive, $I_i \geq 0, i = 1, 2, 3$. If $\Omega(t) = \Omega_1(t)E_1 + \Omega_2(t)E_2 + \Omega_3(t)E_3$ we have

$$K^c(t) = \frac{1}{2}(I_1 {\Omega_1}^2(t) + I_2 {\Omega_2}^2(t) + I_3 {\Omega_3}^2(t)).$$

Since $AE_i = I_i E_i, i = 1, 2, 3$, and because we had assumed, without loss of generality, that the fixed point 0 belongs to S, the three lines: $0 + \lambda E_i, \lambda \in \mathbf{R}, i = 1, 2, 3$, are mutually orthogonal, and are called the **principal axis** of S at the point 0.

The set $\{\Omega \in K | (A\Omega, \Omega) = 1\}$ is called the **inertia ellipsoid of the rigid body** S at the point 0. The equation of such ellipsoid, with respect to the reference frame $(0, E_1, E_2, E_3)$, is

$$I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2 = 1$$

where $\Omega = \Omega_1 E_1 + \Omega_2 E_2 + \Omega_3 E_3$.

Special case: **motion of a rigid body with a fixed axis.**

If $S \subset K$ is a rigid body with a fixed point ($r(t) = M_t(0) = o$ for all t) and if $\omega(t) = \omega \neq 0$ is constant, we say that S rotates around the **axis** $e = \frac{\omega}{|\omega|} \in k$ with constant **angular velocity** ω. In this case, the motions $q(t, \xi)$ of S satisfy:

$$\begin{aligned} \dot{q}(t,\xi) &= \omega \times q(t,\xi) \\ q(0,\xi) &= M_o^* \xi. \end{aligned}$$

The solution of that ordinary differential equation, with the initial condition above, can be easily found. In fact let $\tilde{\omega} = \omega \times$ be the skew symmetric operator corresponding to the vector $\omega \neq 0$; so the solution is

$$q(t,\xi) = exp(t\tilde{\omega}) M_o^* \xi$$

Since in the present case $q(t,\xi) = M_t^* \xi$ one has:

$$M_t{}^* = exp(t\tilde{\omega}).M_o{}^*$$

Exercise 6.5 Assume that S rotates around the axis $e = \frac{\omega}{|\omega|}$ with constant angular velocity; then:

1) The distance $\rho(\xi)$ between $q(t,\xi)$ and the axis $\{\lambda e | \lambda \in \mathbf{R}\}$ does not depend on t.

2) The kinetic energy is given by

$$K^c(t) = \frac{1}{2} I_e |\omega|^2, \quad \text{where} \quad I_e = \int_S \rho^2(\xi) dm(\xi)$$

is called the moment of inertia of the rigid body with respect to the axis $\{\lambda e | \lambda \in \mathbf{R}\}$.

3) $\Omega(t) = (M_t{}^*)^{-1}\omega = \Omega$ is constant and

$$K^c(t) = \frac{1}{2} I_\Omega |\Omega|^2, \qquad \text{where}$$

$$I_\Omega = \int_S |E \times \xi|^2 dm(\xi)$$

is the moment of inertia of the rigid body with respect to the axis $\{\lambda E | \lambda \in \mathbf{R}\}$, $E = \frac{\Omega}{|\Omega|}$.

4) The eigenvalues I_1, I_2 and I_3 of the inertia operator A are the momenta of inertia of the rigid body with respect to the principal axis of S.

Exercise 7.5 (Steiner) The moment of inertia of the rigid body with respect to an axis is equal to the sum of the moment of inertia with respect to another axis through the center of mass and parallel to the first one plus $m(S)d^2$ where d is the distance between the two axis.

The dynamics of a rigid body S is introduced for the bodies S that have at least three points not in a straight line. Let us fix, from now on, a proper linear isometry $B : K \to k$. The Lie group $SO(k;3)$ of all proper (linear) orthogonal operators of k is a compact manifold with dimension three. The **configuration space** of a rigid body is a six-dimensional manifold, namely $k \times SO(k;3)$.

Proposition 12.5 *The set of all proper isometries M of K onto k is diffeomorphic to the six-dimensional manifold $k \times SO(k;3)$.*

Proof Let us consider the map

$$\Phi_B : M \longmapsto (M(0), M^*B^{-1}) \tag{5.83}$$

where B is the linear isometry fixed above and M^* is the linear map associated to M, that is,

$$M^*(X) = M(X) - M(0) \quad \text{for all} \quad X \in K.$$

It is easy to see that Φ_B is differentiable, injective and has a differentiable inverse Ψ_B given by

$$\Psi_B : (r, h) \in k \times SO(k;3) \longmapsto N$$

where N is the proper isometry defined by $N(X) = r + hB(X)$.

By (5.69) the motion of S is given by

$$q(t,\xi) = M_t{}^*(\xi) + r(t), \qquad r(t) = M_t(0);$$

taking into account the map Φ_B (see (5.69)), to the proper isometry M_t there corresponds a pair $(r(t), h(t)) \in k \times SO(k;3)$ that is:

$$\Phi_B(M_t) = (r(t), h(t) = M_t{}^*B^{-1}). \tag{5.84}$$

So, we can write:

$$q(t,\xi) = r(t) + M_t{}^*(\xi) = r(t) + h(t)B\xi. \tag{5.85}$$

Let us denote by β the **σ-algebra of all Borel sets of** K, by λ **a real-valued measure on** (K, β) and let $f : K \to \mathbf{R}$ be **a (real-valued) λ-measurable function.** The correspondence

$$\nu : E \in \beta \longmapsto \int_E f(\xi) d\lambda(\xi) \tag{5.86}$$

is a real-valued measure on (K, β). Moreover, for any λ-measurable function $g : K \to \mathbf{R}$, one has

$$\int_E g(\xi) d\nu(\xi) \stackrel{def}{=} \int_E g(\xi) f(\xi) d\lambda(\xi). \tag{5.87}$$

Given a **vector-valued λ-measurable function** $G : K \to k$, one obtains (taking in k a positive orthonormal basis) its components $g_i, i = 1, 2, 3$, that are (real-values) λ-measurable functions. So, the vector $\nu(E) = \int_E G(\xi) d\lambda(\xi)$ has three components:

$$\nu_i(E) = \int_E g_i(\xi) d\lambda(\xi), \qquad i = 1, 2, 3. \tag{5.88}$$

It can be also introduced the notion of **vector-valued measure on**(K, β) or **measure on** (K, β) **with values on** k, through the utilization of its three components. In fact if Φ is a measure on (K, β) with values on k and Φ_1, Φ_2, Φ_3 its components in a positive orthonormal basis of k, and given a Φ-measurable (real-valued) function $f : K \to \mathbf{R}$, one denotes by $\int_E f(\xi) d\Phi(\xi)$ the vector in k with components $\int_E f(\xi) d\Phi_i(\xi), i = 1, 2, 3$. Given a Φ-measurable vector-valued function $v : K \to k$, we have that $\int_E v(\xi).d\Phi(\xi)$ is the number given by $\sum_i (\int_E v_i(\xi) d\Phi_i(\xi))$ and $\int_E v(\xi) \times d\Phi(\xi)$ is the vector in k with components:

$$\int_E v_2(\xi) d\Phi_3(\xi) \quad - \quad \int_E v_3(\xi) d\Phi_2(\xi); \int_E v_3(\xi) d\Phi_1(\xi) - \int_E v_1(\xi) d\Phi_3(\xi);$$
$$\int_E v_1(\xi) d\Phi_2(\xi) - \int_E v_2(\xi) d\Phi_1(\xi).$$

If ν is the vector-valued measure introduced by (5.88) depending on a λ-measurable function $G : K \to k$ with components $g_i : K \to \mathbf{R}$, we have

$$\int_E (v(\xi), d\nu(\xi)) = \int_E (v(\xi), G(\xi)) d\lambda(\xi) \quad \text{and} \quad \int_E v(\xi) \times d\nu(\xi) = \int_E [v(\xi) \times G(\xi)] d\lambda(\xi).$$

We want to consider now the notion of **(physical) fields of forces** acting on a rigid body S. If S is under the action of the gravitational acceleration $\vec{g} \in k, |\vec{g}| = g$, one understands that each m-measurable subset $E \subset S$ with mass $m(E)$, is subjected to an external force $m(E)\vec{g}$. So, one can define the weight field of forces as a vector-valued measure on S:

$$E \subset S \longmapsto m(E)\vec{g} = \int_E \vec{g} dm(\xi). \tag{5.89}$$

In general, a **field of forces acting on** $S \subset K$ is a law

$$w \in T(k \times SO(k;3)) \longrightarrow f_w$$

where f_w is a vector-valued measure on S with values on k.

Since $q(t,\xi) = r(t) + h(t)B\xi$ (see 5.85) and so:

$$\dot{q}(t,\xi) = \dot{r}(t) + \dot{h}B\xi, \tag{5.90}$$

we see that to each $w = (u,s) \in T_{(r,h)}(k \times SO(k;3))$ there correspond the maps $q, v : K \to k$ defined by

$$q(\xi) = r + hB\xi, \qquad v(\xi) = u + sB\xi. \tag{5.91}$$

It is usual, in Physics, to consider surface forces, volume forces, etc, in the following way: one defines on S a (real-valued) measure σ and a bounded function $\alpha : k \times k \to k$ such that it is well defined the vector-valued measure on S, with values on k, by:

$$f_w(E) = \int_E \alpha(q(\xi), v(\xi))d\sigma(\xi) \tag{5.92}$$

for any Borel subset $E \subset S$.

Like in the case of a finite system of mass points, it is usual to consider the **field of external forces** $f_w{}^{ext}$ and the **field of internal forces** $f_w{}^{int}$. Given a rigid motion $M : t \mapsto M_t$ of K with respect to k, from (5.85) and (5.90) each proper isometry M_t is represented by the pair $(r(t), h(t)) \in k \times SO(k,3)$ and, at this point, the tangent vector $w(t) = (\dot{r}(t), \dot{h}(t))$ determines the measures

$$f_t{}^{ext} = f_{w(t)}{}^{ext} \quad \text{and} \quad f_t{}^{int} = f_{w(t)}{}^{int}, \quad \text{for each } t.$$

We say that two fields of forces f_w and g_w, acting on a rigid body $S \subset K$, are said to be **equivalent with respect to** M_t if

$$\int_S df_t(\xi) = \int_S dg_t(\xi) \quad \text{and}$$

$$\int_S M_t\xi \times df_t(\xi) = \int_S M_t\xi \times dg_t(\xi) \tag{5.93}$$

As in the case of a finite number of mass points, the fundamental laws, in classical mechanics, relative to the motions of a rigid body S, are:

I - **Newton law**

"The sum of the internal and external fields of forces is, at each time t, equal to the kinematical distribution D_t (assumed to be well defined)", that is:

$$D_t(E) \stackrel{def}{=} \int_E \ddot{q}(t,\xi)dm(\xi) = \int_E df_t^{ext}(\xi) + \int_E df_t^{int}(\xi),$$

for all Borel sets E of S.

II - **Action and reaction principle:**

" The field of internal forces $f_w{}^{int}$ is equivalent to zero, with respect to any proper isometry M_t of an arbitrary rigid motion M of K with respect to k.

The general equations for the motion of a rigid body S are the equations EG_1) and EG_2) below that follow from I and II:

$$EG_1) \qquad \int_S \ddot{q}(t,\xi)dm(\xi) = \int_S df_t^{ext}(\xi) \stackrel{def}{=} F_t^{ext} \tag{5.94}$$

$$EG_2) \qquad \int_S [(q(t,\xi)-c)\times \ddot{q}(t,\xi)]dm(\xi) = \int_S (q(t,\xi)-c)\times df_t^{ext}(\xi) \stackrel{def}{=} P_{t,c}{}^{ext}$$

$$\text{for all } \; c \in k. \tag{5.95}$$

Exercise 8.5 Prove the following formula that gives the variation (derivative) of the kinetic energy $K^c(t)$ (see (5.76)):

$$\frac{dK^c(t)}{dt} = \int_S (\dot{q}(t,\xi), df_t^{ext}(\xi)) = (\dot{g}(t), F_t^{ext}) + (\omega(t), P_{t,g(t)}{}^{ext}),$$

where F_t^{ext} and $P_{t,c}{}^{ext}$ (for $c = g(t)$) appear in EG_1 and EG_2.

A rigid body S is said to be **free** under the action of a rigid motion $M : t \mapsto M_t$ of K relatively to k, if $f_t{}^{ext}$ is equivalent to zero with respect to M_t for all t. In particular, if $f_w{}^{ext} = 0$ that is, in the absence of external forces, the rigid body is said to be **isolated**; for an (approximate) example we can think about the rolling of a spaceship.

If G is the center of mass of S, that is, $G = \frac{1}{m(S)}\int_S \xi dm(\xi)$, then $g(t) = M_t G = \frac{1}{m(S)}\int_S M_t\xi dm(\xi) = \frac{1}{m(S)}\int_S q(t,\xi)dm(\xi)$.

After two derivatives with respect to time one has: $m(S)\ddot{g}(t) = \int_S \ddot{q}(t,\xi)dm(\xi)$; by $EG_1)$ and assuming that S is free, one obtains $\ddot{g}(t) = 0$ for all t:

Proposition 13.5 *If a rigid body S is free under the action of $M : t \mapsto M_t$, its center of mass moves uniformly and linearly. Moreover, the kinetic momentum and the kinetic energy are constants of motion.*

Proof From (5.77) one obtains

$$\dot{p}(t) = \int_S [q(t,\xi) \times \ddot{q}(t,\xi)]dm(\xi)$$

and $EG_{2)}$ (with $c = 0$) implies:

$$\dot{p}(t) = \int_S q(t,\xi) \times df_t{}^{ext}(\xi) = \int_S M_t(\xi) \times df_t{}^{ext}(\xi);$$

but, the fact that S is free under the action of $M : t \mapsto M_t$ gives, together with (5.93), that $\dot{p}(t) = 0$. With an analogous argument with the expression of $\frac{dK^c(t)}{dt}$ given by the result of Exercise 8.5 we see that $\frac{dK^c(t)}{dt} = 0$; so, $p(t)$ and $K^c(t)$ are constants of motion. More precisely, since $p(t)$ is a vector-valued constant of motion, one obtains four (scalar valued) constants of motion for any rigid body S free under the action of M. ∎

Assume we are looking at an inertial coordinate system where the center of mass is stationary. Then

Proposition 14.5 *A free rigid body rotates around its center of mass as if the center of mass was fixed.*

Let us consider the motion of a rigid body around a stationary point, in the absence of external forces. In this case, there exist four real valued constants of motion given by Proposition 13.5. One can also consider the induced functions

$$K^c : T(SO(k;3)) \longrightarrow R \qquad p : T(SO(k;3)) \longrightarrow k, \tag{5.96}$$

defined by

$$\begin{aligned} s_h &\in T(SO(k;3)) \longmapsto K^c(s_h) = \frac{1}{2}\int_S |sB\xi|^2 dm(\xi), \\ s_h &\in T(SO(k;3)) \longmapsto p(s_h) = \int_S (hB\xi \times sB\xi)dm(\xi), \end{aligned} \tag{5.97}$$

respectively. In general (if the rigid body does not have any particular symmetry) the four scalar-valued maps (K^c and the components p_i of p in a basis of k) defined on the six-dimensional manifold $T(SO(k,3))$ are independent in the sense that they do not have critical points, that is, the inverse image of any value (K_o, p_o) (if non empty) is a two dimensional orientable compact invariant manifold, provided that the value K_o of $K^c(s_h)$ be positive. Moreover, $K_o > 0$ implies that the vector field induced on the inverse image of (K_o, p_o) by (K^c, p) has no singular points, that is, each connected component $(K^c, p)^{-1}\,(K_o, p_o)$ is a bi-dimensional torus.

Proposition 15.5 *The kinetic momentum $P(t)$ relative to a rigid body S that is free under the action of $M : t \mapsto M_t$, satisfies the* **Euler equation**: $\dot{P}(t) = P(t) \times \Omega(t)$.*Moreover, $\Omega(t)$ is given by the relation, $A\dot{\Omega}(t) = [A\Omega(t)] \times \Omega(t)$, A being the inertia operator.*

Proof In fact, $p(t) = M_t{}^* P(t)$, so by Proposition 13.5 we have

$$\dot{p}(t) = \dot{M}_t^* P(t) + M_t^* \dot{P}(t) = 0, \qquad \text{and so}$$

$$\dot{P}(t) = -(M_t{}^*)^{-1} \dot{M}_t^* P(t) = -\Omega(t) \times P(t) = P(t) \times \Omega(t).$$

But, since $P(t) = A\Omega(t)$, we also have $A\dot{\Omega}(t) = [A\Omega(t)] \times \Omega(t)$. ∎

Proposition 16.5 *In the motion of a rigid body S with a fixed point, subjected to a field of external forces, the kinetic momenta $p(t)$ and $P(t)$ satisfy the equations*

$$\dot{p}(t) = \int_S (M_t{}^*\xi) \times df_t{}^{ext}(\xi),$$

$$\dot{P}(t) = P(t) \times \Omega(t) + \int_S \xi \times [(M_t{}^*)^{-1} df_t{}^{ext}(\xi)].$$

Proof From (5.77) one obtains $\dot{p}(t) = \int_S [q(t,\xi) \times \ddot{q}(t,\xi)] dm(\xi)$ and since there is a fixed point we can write $q(t,\xi) = M_t{}^*\xi$; using EG_2) with $c = 0$ we have the equation for $\dot{p}(t)$. Since $P(t) = (M_t{}^*)^{-1} p(t)$ and using again (5.77) one can write by derivative:

$$\dot{P}(t) = (\dot{M}_t^*)^{-1} \int_S [q(t,\xi) \times \dot{q}(t,\xi)] dm(\xi) + (M_t{}^*)^{-1} \dot{p}(t);$$

but $M_t{}^*(M_t{}^*)^{-1} = Id$ implies, by derivative, that $(\dot{M}_t^*)^{-1} = -(M_t{}^*)^{-1} \dot{M}_t^* (M_t{}^*)^{-1}$ so,

$$\dot{P}(t) = \int_S [\xi \times (M_t{}^*)^{-1} df_t{}^{ext}(\xi)] - \Omega(t) \times [(M_t{}^*)^{-1} \int_S [q(t,\xi) \times \dot{q}(t,\xi)] dm(\xi)]$$

and finally,

$$\dot{P}(t) = P(t) \times \Omega(t) + \int_S \xi \times [(M_t{}^*)^{-1} df_t{}^{ext}(\xi)]. \quad ∎$$

In order to relate the properties EG_1) and EG_2) with the abstract Newton law, we start by defining the metric $\langle , \rangle$ on $k \times SO(k;3)$. This metric is induced by the kinetic energy. Since (see (5.90))

$$q(t,\xi) = r(t) + h(t)B\xi \qquad \text{and}$$
$$\dot{q}(t,\xi) = \dot{r}(t) + \dot{h}(t)B\xi,$$

we have

$$K^c(t) = \frac{1}{2}\int_S |\dot{r}(t) + \dot{h}B\xi|^2 dm(\xi); \tag{5.98}$$

We will assume that the origin $0 \in K$ coincides with the center of mass $G = \frac{1}{m(S)}\int_S \xi dm(\xi)$; so, we have $\int_S \xi dm(\xi) = 0$, that implies

$$K^c(t) = \frac{1}{2}m(S)|\dot{r}(t)|^2 + \frac{1}{2}\int_S |\dot{h}(t)B\xi|^2 dm(\xi).$$

The last expression suggests the introduction of a metric on $k \times SO(k;3)$; in fact, given two tangent vectors $(u,s), (\bar{u},\bar{s})$ at the point $(r,h) \in k \times SO(k;3)$ one defines

$$\langle (u,s),(\bar{u},\bar{s})\rangle_{(r,h)} \stackrel{\text{def}}{=} m(S)(u,\bar{u}) + \int_S (sB\xi, \bar{s}B\xi)dm(\xi)$$

in which the right hand side defines two inner products,

$$\langle u,\bar{u}\rangle_r = m(S)(u,\bar{u}) \qquad \text{and} \qquad \langle s,\bar{s}\rangle_h = \int_S (sB\xi, \bar{s}B\xi)dm(\xi), \tag{5.99}$$

on k and $SO(k;3)$, respectively. Recall that s and $\bar{s}$ are tangent vectors at $h \in SO(k;3)$. So, it is defined on $SO(k;3)$ a riemannian metric which is left invariant, that is, the left translations are isometries. In fact, given $g \in SO(k;3)$, the left translation L_g is defined by the expression $L_g(x) = gx$, for all $x \in SO(k;3)$ and, since g is a linear transformation acting on k, its derivative satisfies $dL_g(x) = L_g$; so one obtains

$$\begin{aligned}
\langle \quad dL_g(h)s, dL_g(h)\bar{s}\rangle_{gh} &= \langle gs, g\bar{s}\rangle_{gh} \\
= &\int_S (gsB\xi, g\bar{s}B\xi)dm(\xi) = \int_S (sB\xi, \bar{s}B\xi)dm(\xi) \\
= &\langle s,\bar{s}\rangle_h.
\end{aligned}$$

The acceleration, in the product metric, corresponding to a vector $\dot{q} = (\dot{r},\dot{h})$ tangent to $k \times SO(k;3)$ at the point (r,h), is equal to

$$\frac{D\dot{q}}{dt} = \frac{D}{dt}(\dot{r},\dot{h}) = (\ddot{r}, \frac{D\dot{h}}{dt}).$$

The mass operator in the product metric acts on $\frac{D\dot{q}}{dt}$ as

$$\mu(\frac{D\dot{q}}{dt})(u,s) = \langle \ddot{r}, u\rangle_r + \langle \frac{D\dot{h}}{dt}, s\rangle_h.$$

Let us introduce now an abstract field of forces $\mathcal{F} : T(k \times SO(k;3)) \longrightarrow T^*(k \times SO(k;3))$ in a suitable way such that the generalized Newton law

$$\mu(\frac{D\dot{q}}{dt}) = \mathcal{F}(\dot{q})$$

becomes equivalent to the general equations EG$_1$) and EG$_2$), for the motion of a rigid body. The way we define $\mathcal{F}$ is the following: for $(\bar{u}, \bar{s})$ and $w = (u, s)$ in $T_{r,h}(k \times SO(k;3))$ we state:

$$(\mathcal{F}(u,s))(\bar{u},\bar{s}) = \int_S (\bar{u}, df_w^{ext}(\xi)) + \int_S (\bar{s}B\xi, df_w^{ext}(\xi)). \tag{5.100}$$

Recall (see (5.94), (5.95)) the general equations:

EG$_1$) $$\int_S \ddot{q}(t,\xi)dm(\xi) = \int_S df_t^{ext}(\xi) = F_t^{ext}$$

EG$_2$) $$\int_S (q(t,\xi) - c) \times \ddot{q}(t,\xi)dm(\xi) = \int_S (q(t,\xi) - c) \times df_t^{ext}(\xi) = P_{t,c}^{ext},$$

for all $c \in k$.

It is a simple matter to see that EG$_1$) and EG$_2$) are equivalent to EG$_1$) and EG'$_2$), where

EG$'_2$) : $$\int_S (q(t,\xi) - g(t)) \times \ddot{q}(t,\xi)dm(\xi) = \int_S (q(t,\xi) - g(t)) \times df_t^{ext}(\xi) = P_{t,g(t)}^{ext}$$

with

$$g(t) = M_tG = M_t\left[\frac{1}{m(S)}\int_S \xi dm(\xi)\right] = \frac{1}{m(S)}\int_S q(t,\xi)dm(\xi),$$

G being the center of mass of S, that we already made equal to the origin 0 of K. Thus we can write:

$$\int_S \xi dm(\xi) = 0. \tag{5.101}$$

The expression of $q(t,\xi) = M_t(\xi)$ is, in this case, $q(t,\xi) = M_t(0) + M_t^*\xi = g(t) + h(t)B\xi$, with $M_t^* = h(t)B$. So we have

$$\dot{q}(t,\xi) = \dot{g}(t) + \dot{M}_t^*\xi \qquad \text{and} \qquad \ddot{q}(t,\xi) = \ddot{g}(t) + \ddot{M}_t^*\xi,$$

then EG$_1$) becomes equivalent to

$$\int_S \ddot{g}(t)dm(\xi) + \ddot{M}_t^* \int_S \xi dm(\xi) = F_t^{ext} = m(S)\ddot{g}(t),$$

and, by (5.101), we have EG$_1$) equivalent to

$$m(S)(\ddot{g}(t), \bar{u}) = (F_t^{ext}, \bar{u}), \qquad \text{for all} \qquad \bar{u} \in k. \tag{5.102}$$

On the other hand $\mathrm{EG}_2)'$ is equivalent to

$$\begin{aligned}
P^{ext}_{t,g(t)} &= \int_S M_t^*\xi \times (\ddot{g}(t) + \ddot{M}_t^*\xi)dm(\xi) \\
&= \left(\int_S M_t^*\xi dm(\xi)\right) \times \ddot{g}(t) + \int_S \frac{d}{dt}(M_t^*\xi \times \dot{M}_t^*\xi)dm(\xi) = \\
&= M_t^*\left(\int_S \xi dm(\xi)\right) \times \ddot{g}(t) + \frac{d}{dt}\int_S (M_t^*\xi \times \dot{M}_t^*\xi)dm(\xi),
\end{aligned}$$

and by (5.101) $\mathrm{EG}_2)'$ is equivalent to

$$\left(\frac{d}{dt}\int_S (M_t^*\xi \times \dot{M}_t^*\xi)dm(\xi), \quad \bar{u}\right) = (P^{ext}_{t,g(t)}, \quad \bar{u}) \qquad \text{for all} \qquad \bar{u} \in k. \quad (5.103)$$

From what is said in Exercise 5.5 there is a linear isomorphism Φ between k and the space $s(k)$ of all linear skew-symmetric operators of k. In fact, for any $A \in s(k)$, $\Phi(A)$ is the unique vector in k such that $Av = \Phi(A) \times v$ for all $v \in k$. With that notation, $\mathrm{EG}_2)'$ be equivalent to (5.103) means be equivalent to

$$\begin{aligned}
(P^{ext}_{t,g(t)}, \Phi(A)) &= \frac{d}{dt}\int_S (M_t^*\xi \times \dot{M}_t^*\xi, \Phi(A))dm(\xi) \\
&= \frac{d}{dt}\int_S (\Phi(A) \times M_t^*\xi, \dot{M}_t^*\xi)dm(\xi);
\end{aligned}$$

then $\mathrm{EG}_2)'$ is equivalent to

$$(P^{ext}_{t,g(t)}, \Phi(A)) = \frac{d}{dt}\int_S (AM_t^*\xi, \dot{M}_t^*\xi)dm(\xi), \qquad \text{for all} \qquad A \in s(k). \tag{5.104}$$

There is also a linear isomorphism between the tangent space $T_hSO(k;3)$ and $s(k)$ (see Exercise 9.5 below) using the map

$$\dot{\bar{h}} \in T_hSO(k,3) \longmapsto \dot{\bar{h}}\ h^{-1} \in s(k) \tag{5.105}$$

which is the derivative of the right translation $R_{h^{-1}}$ defined as $R_{h^{-1}}(x) = xh^{-1}$, for all $x \in SO(k;3)$.

Exercise 9.5 Prove that $\dot{\bar{h}}\ h^{-1} \in s(k)$ in (5.104) and that the map is a linear isomorphism.

We recall that $M_t^* = h(t)B$, so (5.104) and (5.105) imply that $\mathrm{EG}_2)'$ is equivalent to

$$\begin{aligned}
(P^{ext}_{t,g(t)}, \Phi(\dot{\bar{h}}\ h^{-1}(t))) &= \frac{d}{dt}\int_S (\dot{\bar{h}}\ h^{-1}(t)h(t)B\xi, \dot{h}(t)B\xi)dm(\xi) \\
&= \frac{d}{dt}\int_S (\dot{\bar{h}}B\xi, \dot{h}(t)B\xi)dm(\xi) \qquad \text{for all} \qquad \dot{\bar{h}} \in T_{h(t)}SO(k;3).
\end{aligned}$$

From (5.99), (5.102) and the last expression, one can say that EG$_1$) and EG$_2$)$'$ are equivalent to

$$
\begin{aligned}
&(F_t^{ext}, \bar{u}) + (P_{t,g(t)}^{ext}, \Phi(\dot{\bar{h}}\ \ h^{-1}(t)) = \\
= \ & \frac{d}{dt}\left[m(S)(\dot{g}(t), \bar{u}) + \int_S (\dot{\bar{h}} B\xi, \dot{h}(t) B\xi) dm(\xi)\right] \\
= \ & \frac{d}{dt}\langle (\dot{g}(t), \dot{h}(t)), (\bar{u}, \dot{\bar{h}})\rangle \\
& \text{for all} \quad (\bar{u}, \dot{\bar{h}}) \in T_{(g(t),h(t))}\ k \times SO(k;3).
\end{aligned}
\tag{5.106}
$$

Remark that if we extend, by parallel transport, the vector $(\bar{u}, \dot{\bar{h}})$ along the motion $q(t) = (g(t), h(t))$, one obtains a vector field along $q(t)$ still denoted by $(\bar{u}, \dot{\bar{h}})$ so that $\frac{D}{dt}(\bar{u}, \dot{\bar{h}}) = 0$ and then the second member of (5.106) can be written as

$$
\frac{d}{dt}\langle \dot{q}, (\bar{u}, \dot{\bar{h}})\rangle = \langle \frac{D\dot{q}}{dt}, (\bar{u}, \dot{\bar{h}})\rangle + \langle \dot{q}, \frac{D}{dt}(\bar{u}, \dot{\bar{h}})\rangle = \langle \frac{D\dot{q}}{dt}, (\bar{u}, \dot{\bar{h}})\rangle. \tag{5.107}
$$

Let us recall the field of forces

$$
\mathcal{F} : T(k \times SO(k;3)) \longrightarrow T^*(k \times SO(k;3))
$$

given in the following way: if $(u, s) \in T_{(r,h)}(k \times SO(k;3))$ then we have $\mathcal{F}(u, s) \in T^*_{(r,h)}(k \times SO(k,3))$ if, and only if (5.100) holds, that is, for $(u, s) = \dot{q}$:

$$
(\mathcal{F}(\dot{q}))(\bar{u}, \dot{\bar{h}}) = (F_t^{ext}, \bar{u}) + (P_{t,g(t)}^{ext}, \Phi(\dot{\bar{h}} h^{-1}(t))). \tag{5.108}
$$

The constructions of $h^{-1}(t)$, F_t^{ext} and $P_{t,g(t)}^{ext}$ are possible because given $(r, h) \in k \times SO(k;3)$ and $(u, s) \in T_{(r,h)}(k \times SO(k;3))$ we are able to find $q(t, \xi)$ and so $\dot{q}(t, \xi)$ that determine $h^{-1}(t), F_t^{ext}$ and $P_{t,g(t)}^{ext}$. The conclusion is then the following result:

Proposition 17.5 *The general equations* EG$_1$) *and* EG$_2$) *that govern the motions of a rigid body S (see (5.94) and (5.95)) are equivalent to the generalized Newton law $\mu(\frac{D\dot{q}}{dt}) = \mathcal{F}(\dot{q})$ on the manifold $k \times SO(3)$ with the riemannian metric given by equations (5.99) and the field of forces $\mathcal{F}$ characterized by (5.100).*

Proof As we saw, the equations EG$_1$) and EG$_2$) are equivalent to (5.106); using (5.106) and (5.107) we see that

$$
\langle \frac{D\dot{q}}{dt}, v\rangle = [\mathcal{F}(\dot{q})]v \quad \text{for all} \quad v \in T_{q(t)}[k \times SO(k;3)],
$$

so,

$$
\mu(\frac{D\dot{q}}{dt}) = \mathcal{F}(\dot{q}). \qquad \blacksquare
$$

We intend, now, to derive the **Lagrange equations for the motion of a rigid body** S; then the take a positive orthonormal basis $\{e_1, e_2, e_3\}$ for the vector space k and denote by (r_1, r_2, r_3) the coordinates of a vector $r \in k$. Let (h_1, h_2, h_3) be a local system of coordinates for $SO(k;3)$. So if $(\bar{u}, \bar{s}) \in T_{(r,h)}(k \times SO(k;3))$ we have $\bar{u} = \Sigma_{i=1}^3 \bar{u}_i e_i$ and $\bar{s} = \Sigma_{i=1}^3 \bar{s}_i \frac{\partial}{\partial h_i}(h)$. The force $\mathcal{F}$ defined above in (5.100) has the following expression in those local coordinates

$$\mathcal{F}(r,h))(\bar{u},\bar{s}) = \int_S (\bar{u}, df_w^{ext}(\xi)) + \int_S (\bar{s}B\xi, df_w^{ext}(\xi)) =$$

$$= \sum_{i=1}^3 \bar{u}_i (e_i, \int_S df_w^{ext}(\xi)) + \sum_{i=1}^3 \bar{s}_i \int_S (\frac{\partial}{\partial h_i}(h)B\xi, df_w^{ext}(\xi)) =$$

$$= \sum_{i=1}^3 (\int_S df_w^{ext}(\xi))_i, dr_i(\bar{u}) + \sum_{i=1}^3 (\int_S (\frac{\partial}{\partial h_i}(h)B\xi, df_w^{ext}(\xi)))dh_i(\bar{s}) \quad (5.109)$$

Then if $t \to (r(t), h(t)) \in k \times SO(k;3)$ is a motion of S under the external forces f^{ext} and being $K^c(t)$ the kinetic energy along this motion, the Newton law gives

$$\frac{d}{dt}\frac{\partial K^c}{\partial \dot{r}_i} - \frac{\partial K^c}{\partial r_i} = (\int_S df_t^{ext}(\xi))_i, \quad i = 1,2,3, \quad (5.110)$$

$$\frac{d}{dt}\frac{\partial K^c}{\partial \dot{h}_i} - \frac{\partial K^c}{\partial h_i} = \int_S (\frac{\partial}{\partial h_i}(h)B\xi, df_t^{ext}(\xi)), \quad i = 1,2,3. \quad (5.111)$$

We will relate the right hand sides of equations (5.110) and (5.111) above, with the physical notions of total force and momentum of external forces with respect to a point.

Since $\frac{\partial}{\partial h_i}(h)h^{-1}(t) \in T_e(SO(k,3))$, it follows that, for each t, there exist vectors $\omega_i(t) \in k$ such that

$$\omega_i(t) \times = \frac{\partial}{\partial h_i}(h)h^{-1}(t), \quad i = 1,2,3. \quad (5.112)$$

This implies

$$\begin{aligned} \frac{d}{dt}\frac{\partial K^c}{\partial \dot{h}_i} - \frac{\partial K^c}{\partial h_i} &= \int_S (\omega_i(t) \times h(t)B\xi, df_t^{ext}(\xi)) = \\ &= (\omega_i(t), \int_S hB\xi \times df_t^{ext}(\xi)), \quad i = 1,2,3. \quad (5.113) \end{aligned}$$

Introducing the usual notation $F_t^{ext} = \int_S df_t^{ext}(\xi)$ (total force at t) and $P_t^{ext} = P_{t,r(t)}^{ext} = \int_S (q(t,\xi) - r(t)) \times df_t^{ext}(\xi) = \int_S hB\xi \times df_t^{ext}(\xi)$ (the momentum of external forces with respect to $r(t)$ at the time t) we obtain the **Lagrange equations for the motions of a rigid body** S:

$$\frac{d}{dt}\frac{\partial K^c}{\partial \dot{r}_i} - \frac{\partial K^c}{\partial r_i} = (F_t^{ext})_i, \qquad i = 1, 2, 3 \tag{5.114}$$

$$\frac{d}{dt}\frac{\partial K^c}{\partial \dot{h}_i} - \frac{\partial K^c}{\partial h_i} = (\omega_i(t), P_t^{ext}), \quad i = 1, 2, 3. \tag{5.115}$$

Since $K^c(t) = \frac{1}{2}m(S)|\dot{r}|^2 + \frac{1}{2}\int_S |\dot{h}B\xi|^2 dm(\xi)$ the first Lagrange equation gives us

$$m(S)\ddot{r}(t) = F_t^{ext},$$

and the hypothesis $G = 0$ implies $r(t) = g(t)$ so we obtain the classical Newton law for the motion of G. If the rigid body moves with a fixed point, the second of the Lagrange equations are the only ones to be considered.

Exercise 10.5 Let $S \subset K$ be a rigid body with fixed point $O \in S$. Assume $K = k$, $B = id$, (O, e_x, e_y, e_z) and (O, e_1, e_2, e_3) orthogonal positively oriented frames fixed in k and in S, respectively. If $e_z \times e_3 \neq 0$, let $e_N = \frac{e_z \times e_3}{|e_z \times e_3|}$. The **nodal line** passes through O and has direction e_N. The **Euler angles** (φ, θ, ψ) are defined as follows: φ is the angle of rotation along the axis $(0, e_z)$ and sends e_x to e_N; θ is the angle of rotation along $(0, e_N)$ and sends e_z to e_3; ψ is the rotation along $(0, e_3)$ and sends e_N to e_1. Show that to each (φ, θ, ψ) satisfying $0 < \varphi < 2\pi$, $0 < \psi < 2\pi$, $0 < \theta < \pi$, corresponds a rotation $R(\varphi, \theta, \psi)$ defining local coordinates for $SO(k; 3)$.Denote by I_1, I_2, I_3 the moments of inertia of S relative to (e_1, e_2, e_3) and prove that $\Omega = Ae_1 + Be_2 + Ce_3$, $\omega = \bar{A}e_x + \bar{B}e_y + \bar{C}e_z$, $K^c = \frac{1}{2}(I_1A^2 + I_2B^2 + I_3C^2)$ where $A = \dot{\varphi}\sin(\psi)\sin(\theta) + \dot{\varphi}\cos(\psi)$, $B = \dot{\varphi}\cos(\psi)\sin(\theta) - \dot{\theta}\sin(\psi)$ and $C = \dot{\varphi}\cos(\theta) + \dot{\varphi}$. Compute $\bar{A}$, $\bar{B}$ and $\bar{C}$.

5.7 Dynamics of pseudo-rigid bodies

The present section corresponds to Dirichlet-Riemann formulation of ellipsoidal motions for fluid masses also called pseudo-rigid bodies.

As in the previous section, k and K are two 3-dimensional euclidean vector spaces considered as affine spaces; they represent the fixed (inertial) space and the moving space respectively.

A motion $t \mapsto M_t$ is a smooth map where each $M_t : K \to k$ is an orientation preserving affine transformation (bijection) such that takes the zero vector $O \in K$ into the zero $0 \in k$.

If we fix a ball $B_r \subset K$ of radius r and centered in O, a motion of a pseudo-rigid body is the motion

$$t \mapsto M_t(B_r) \subset k$$

of a solid ellipsoid.

Given M_t, we call $B = M_{t=0}$ and set $Q_t = M_t \circ B^{-1} : k \to k$, so $Q_t \in GL^+(k,3)$. The derivative $\dot{Q}_t = \dot{M}_t \circ B^{-1}$ represents the tangent vector at the point $Q_t \in GL^+(k,3)$ to the curve $t \mapsto Q_t$. Take a point $X \in \mathcal{B}_r$; then $q(t,X) = M_t X$ is a curve in k with velocity $\dot{q}(t,X) = \dot{M}_t X$.

The **kinetic energy** of the motion of the solid ellipsoid is

$$K^c(t) = \frac{1}{2}\int_{\mathcal{B}_r} |\dot{q}(t,X)|^2 \, dm(X)$$

where the positive measure m is the distribution of mass. So

$$K^c(t) = \frac{1}{2}\int_{\mathcal{B}_r} |\dot{Q}_t \circ BX|^2 \, dm(X) = \frac{1}{2}\int_{\mathcal{B}_r} |\dot{Q}_t \circ BX|^2 \, \rho dV(X)$$

where ρ is the density and V is the Lebesgue volume. When $\rho =$ constant,

$$\frac{\rho}{2}\int_{\mathcal{B}_r} |\dot{Q}_t \circ BX|^2 \, dV(X).$$

In order to work with matrices, we fix two positive orthonormal bases (e_1, e_2, e_3) and (E_1, E_2, E_3) in k and K, respectively, so that the matrix B is Id, the identity matrix. We shall denote by Q_t and X the corresponding matrices of Q_t and X with respect to the fixed bases. Then

$$K^c(t) = \frac{\rho}{2}\int_{\mathcal{B}_r} |\dot{Q}_t X|^2 \, dV(X). \tag{5.116}$$

Proposition 18.5 Any real $n \times n$ matrix G has a (not unique) bipolar decomposition $G = LDR$, that is L, R are orthogonal matrices and $D = \operatorname{diag}(\sqrt{\sigma_1}, \ldots, \sqrt{\sigma_n})$. Moreover $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ are the non negative eigenvalues of $G^T G$ (G^T is the transpose of G).

Proposition 19.5 The matrix $\mathcal{E}_0 = \rho \int_{\mathcal{B}_r} XX^T dV(X)$ is given by $\mathcal{E}_0 = \frac{4\rho\pi r^5}{15}$.

Proposition 20.5 The kinetic energy (5.116) is given by $K^c(t) = \frac{1}{2}\operatorname{tr}(\dot{Q}_t \, \mathcal{E}_0 \, \dot{Q}_t^T)$. (Here $\operatorname{tr} A$ denotes the trace of the matrix A).

From the propositions above it follows that

$$K^c(t) = \frac{1}{2}\bar{m}\operatorname{tr}(\dot{Q}_t \dot{Q}_t^T). \tag{5.117}$$

Exercise 11.5 Prove the three last propositions.

Let us assume, from now on, that $\bar{m} = 1$.

Remark The expression (5.117) suggests the following Riemannian metric for the group $GL^+(3)$ of all 3×3 matrices of positive determinant:

$$\langle A\,,\,B\,\rangle_Q := \mathrm{tr}\,(AB^T), \tag{5.118}$$

for all $Q \in GL^+(3)$ and all $A, B \in T_Q GL^+(3)$.

Assume that a smooth motion has a (not necessarily unique) smooth bipolar decomposition $Q_t = T_t^T\, A_t\, S_t$ (i.e. three smooth paths: A_t diagonal, and T_t, S_t orthogonal paths).

In the case Q_t is analytic, this is always possible; also, if the eigenvalues of $Q_t Q_t^T$ are two by two distinct and Q_t is not analytic, the smooth decomposition is still possible. However, there are examples of C^∞ paths Q_t for which there is no continuous bipolar decomposition (see Montaldi [Mon], Kato [Ka] and Roberts- S. Dias [RS]). We have:

Proposition 21.5 From the equation of continuity in hydrodynamics and $\rho =$ constant, it follows that a smooth path $Q_t = M_t \circ B^{-1}$ corresponding to an ellipsoidal motion satisfies $\det Q_t = 1$, that is, Q_t is a curve in the Lie group $SL(3)$.

Proof: Assume $Q_t = T_t^T\, A_t S_t$ and call

$$x = T_t q(t, X) = T_t M_t X = T_t Q_t B X$$

where $T_t = (T_{ki})$ means a rotation that takes (e_1, e_2, e_3) to the orthonormal basis $(\bar{e}_1(t), \bar{e}_2(t), \bar{e}_3(t))$, that is $\bar{e}_i(t) = \sum_{k=1}^{3} T_{ki} e_k$, $i = 1, 2, 3$. One may visualize the isomorphisms:

$$A_t : K \to k, \quad S_t : K \to K, \quad T_t : k \to k, \quad \text{and} \quad Q_t = M_t \circ B^{-1}, \quad k \overset{B^{-1}}{\to} K \overset{M_t}{\to} K.$$

Then $u := \dot{x} = \left(\dot{T}_t Q_t + T_t \dot{Q}_t\right) BX$ and $BX = Q_t^{-1} T_t^T x$ so,

$$u = \left(\dot{T}_t T_t^T + T_t \dot{Q}_t Q_t^{-1} T_t^T\right) x$$

and $div\, u = \sum_k \frac{\partial u_k}{\partial x_k} = \mathrm{tr}\,(\dot{Q}_t Q_t^{-1}) = \frac{1}{\det Q_t} \frac{d}{dt}\,(\det Q_t)$. Finally $div\, u = 0$ iff $\frac{d}{dt}\,(\det Q_t) = 0$ iff $\det Q_t = cte$. Thus $\det Q_t = 1$ because for $t = 0$ we have $\det Q_0 = \det\,(BB^{-1}) = 1$.

From Dirichlet-Riemann formulation (see Chandrasekhar [Ch] and Montaldi[Mon]) the motions of pseudo-rigid bodies are given by a generalized Newton law describing a mechanical system on the configuration space $GL^+(3)$ with a holonomic constraint defined by the submanifold $SL(3)$ of $GL^+(3)$, that is:

$$\mu\frac{D\dot{Q}}{dt} = -dV + \lambda df, \qquad Q \in SL(3). \tag{5.119}$$

Here $f : GL(3) \to \mathbb{R}$ is the determinant function and $\lambda : TSL(3) \to \mathbb{R}$ is the Lagrange multiplier; also, $SL(3) = f^{-1}(1) \subset GL^+(3)$ is an analytic 8-dimensional orientable submanifold of $GL^+(3)$,

$$\mu : TGL^+(3) \to T^*GL^+(3)$$

is the mass operator (Legendre transformation) relative to the trace metric, $\mu(v)(\cdot) := \langle v\,,\cdot\rangle$, see (5.118), and $\frac{D\dot{Q}}{dt}$ is the covariant derivative of $\dot{Q}(t)$ (acceleration) along $Q(t)$ is that metric of trace. The map $df : TGL^+(3) \to T^*GL^+(3)$ is given by

$$v \mapsto df(\pi v)$$

where $\pi : TGL^+(3) \to GL^+(3)$ is the canonical bundle projection. We still denote by df its restriction to $TSL(3)$. We will show that $\mu^{-1}df : TSL(3) \to TGL^+(3)$ satisfies **d'Alembert principle.** In fact for any $A \in TSL(3)$ we have

$$(\mu^{-1}df)\,A = w \in T_{\pi(A)}GL^+(3)$$

where w is such that $\langle w\,,\cdot\rangle = \left[df_{\pi(A)}\right](\cdot)$, so w is orthogonal to $T_{\pi(A)}SL(3)$.

Then, as we already saw in 5.4, there exists a unique function $\lambda : TSL(3) \to \mathbb{R}$, the so called Lagrange multiplier. The function

$$V : GL^+(3) \to \mathbb{R}$$

is the **potential energy** and corresponds to the gravitational potential (see examples below).

Proposition 22.5 The generalized Newton law (5.119) is equivalent to the system

$$\ddot{Q} = -\frac{\partial V}{\partial Q} + \lambda\frac{\partial f}{\partial Q}, \qquad \det Q = 1. \tag{5.120}$$

Proof: Here $Q, \ddot{Q}, \frac{\partial V}{\partial Q}, \frac{\partial f}{\partial Q}$ are 3×3 matrices: $Q = (q_{ij}), \ddot{Q} = (\ddot{q}_{ij}), \frac{\partial V}{\partial Q} = (\frac{\partial V}{\partial q_{ij}})$ and $\frac{\partial f}{\partial Q} = (\frac{\partial f}{\partial q_{ij}})$, respectively. We also have that (see Exercise 1.5):

$$\mu\left(\frac{D\dot{Q}}{dt}\right) = \sum_{i,j}\left[\frac{d}{dt}\frac{dK^c}{d\dot{q}_{ij}} - \frac{dK^c}{dq_{ij}}\right]dq_{ij},$$

where

$$K^c = \frac{1}{2}\langle \dot{Q}, \dot{Q}\rangle = \frac{1}{2}\left[\dot{q}_{11}^2 + \dot{q}_{12}^2 + \cdots + \dot{q}_{33}^2\right].$$

Then

$$\mu\left(\frac{D\dot{Q}}{dt}\right) = -dV + \lambda df = \sum_{ij} \ddot{q}_{ij}\, dq_{ij} = \sum_{ij}\left(-\frac{\partial V}{\partial q_{ij}} + \lambda \frac{\partial f}{\partial q_{ij}}\right) dq_{ij}$$

and the proof is complete.

For the Dirichlet-Riemann formulation (see[Ch]) one considers, from the smooth bi-polar decomposition $Q_t = T_t^T\, A_t\, S_t$, the new variables

$$\Omega^* := \dot{T}T^T \qquad\qquad \Lambda^* := \dot{S}S^T$$

which are skew symmetric paths because the derivative of $TT^T = SS^T = I$ gives

$$\dot{T}T^T + T\dot{T}^T = 0 = \dot{S}S^T + S\dot{S}^T.$$

We thus obtain:

$$\dot{Q} = T^T\left(\Omega^{*T}A + \dot{A} + A\Lambda^*\right) S$$

and also

$$\begin{aligned}
\ddot{Q} &= \dot{T}^T\left(\Omega^{*T}A + \dot{A} + A\Lambda^*\right) S + T^T\left(\Omega^{*T}A + \dot{A} + A\Lambda^*\right)\dot{S} + \\
&\quad + T^T\ddot{A}S + T^T\left[\tfrac{d}{dt}\left(A\Lambda^* - \Omega^* A\right)\right] S = \\
&= \left[-\tfrac{\partial V}{\partial Q} + \lambda\tfrac{\partial(\det Q)}{\partial Q}\right]_{Q = T^T AS}.
\end{aligned}$$

Then one obtains the **equation of motion**:

$$\ddot{A} + \Omega^*\left(\Omega^* A - \dot{A} - A\Lambda^*\right) + \left(-\Omega^* A + \dot{A} + A\Lambda^*\right)\Lambda^* + \tfrac{d}{dt}\left(A\Lambda^* - \Omega^* A\right)$$

$$= \left[-T\left(\tfrac{\partial V}{\partial Q}\right)_{Q=T^TAS} S^T + \lambda T\left(\tfrac{\partial(\det Q)}{\partial Q}\right)_{Q=T^TAS} S^T\right]. \tag{5.121}$$

Exercise 12.5

I. If $f = \det Q$, $Q \in GL^+(3)$, then $df_Q(B) = (\det Q)\,\mathrm{tr}(Q^{-1}B)$ for any 3×3 real matrix B.

II. For any function $\phi : GL^+(3) \to \mathbb{R}$ then $\frac{\partial\phi}{\partial Q} = [d\phi_Q(B_{ij})]_{ij}$ where B_{ij} is the matrix with 1 at the (ij)–entry and zero otherwise.

III. $T\,\frac{\partial(\det Q)}{\partial Q} S^T = A^{-1}(\det A)$ for any $Q \in GL(3)$.

Examples of potentials

Assume that $V : L(3) \to \mathbb{R}$ is of the form:

$$V(Q) = \bar{V}\left(\mathrm{I}(C), \mathrm{II}(C), \mathrm{III}(C)\right)$$

where $C = QQ^T$ and $\mathrm{I}(C) = \operatorname{tr} C$, $\mathrm{II}(C) = \frac{1}{2}\left[(\operatorname{tr} C)^2 - \operatorname{tr}(C^2))\right]$, $\mathrm{III}(C) = \det C$.

1. **Gravitational potential**

$$\bar{V} = -2\pi G\rho \int_0^\infty \frac{ds}{\left[(s^3 + \mathrm{I}(C)\, s^2 + \mathrm{II}(C)\, s + \mathrm{III}(C)\right]^{1/2}}.$$

2. **Ciarlet-Geymonat material** (see[LS])

$$\bar{V} = \frac{1}{2}\lambda\left(\mathrm{III}(C) - 1 - \ln \mathrm{II}(C)\right) + \frac{1}{2}\mu\left(\mathrm{I}(C) - 3 - \ln \mathrm{III}(C)\right).$$

3. **Saint Venant-Kirchhoff material** (see[LS])

$$\bar{V} = \frac{1}{2}\lambda\left(\operatorname{tr}(C - Id)\right)^2 + \mu\left(\operatorname{tr}(C - Id)^2\right).$$

Remark In equations (5.121) we need

$$\frac{\partial V}{\partial Q} = \frac{\partial \bar{V}}{\partial \mathrm{I}}\frac{\partial \mathrm{I}(C)}{\partial Q} + \frac{\partial \bar{V}}{\partial \mathrm{II}}\frac{\partial \mathrm{II}(C)}{\partial Q} + \frac{\partial \bar{V}}{\partial \mathrm{III}}\frac{\partial \mathrm{III}(C)}{\partial Q}.$$

Proposition 23.5 (see [SD])

$$\frac{\partial \mathrm{I}(C)}{\partial Q} = 2Q$$

$$\frac{\partial \mathrm{II}(C)}{\partial Q} = 2\left[Id \operatorname{tr}(QQ^T) - QQ^T\right] Q$$

$$\frac{\partial \mathrm{III}(C)}{\partial Q} = 2\det(QQ^T)\,(Q^{-1})^T.$$

Final Remark Using the expression of the gravitational potential and the result of exercise 12.5, III, we see that the equations (5.121) are precisely the so-called **Dirichlet-Riemann** equations (see [Ch] p. 71).

5.8 Dissipative mechanical systems

The results we will present in this section have their proofs in the article "Dissipative Mechanical Systems", by I. Kupka and W.M. Oliva, appeared in *Resenhas* IME-USP 1993, vol. 1, no. 1, 69-115 (see [KO]).

A mechanical system $(Q, \langle,\rangle, \mathcal{F})$, is said to be **dissipative** if the field of external forces $\mathcal{F} : TQ \to T^*Q$ is given by

$$\mathcal{F}(v) = -dV(p) + \tilde{D}(v) \quad \text{for all} \quad v \in T_pQ;$$

where $V : Q \to \mathbb{R}$ is a $C^{r+1}(r \geq 1)$ potential energy and $\tilde{D} \in C^1$ verifies $(\tilde{D}(v))v < 0$ for all $0 \neq v \in TQ$. $\tilde{D}$ is called a **dissipative** external field of forces (or simply a **dissipative** force) and $(-dV)$ is said to be the **conservative** force.

Remarks 1. $\tilde{D}(0_p) = 0 \ \forall p \in Q$ (0_p is the zero vector of T_pQ). In fact, continuity of $\tilde{D}$ shows that $(\tilde{D}(0_p))v = \lim_{\lambda \to 0} \frac{1}{\lambda}(\tilde{D}(\lambda v))\lambda v \leq 0$ for $\lambda > 0$ and $0 \neq v \in T_pQ$ implies $(\tilde{D}(0_p))v = 0$ (otherwise $(\tilde{D}(\epsilon v))v < 0$ for small $\epsilon < 0$ and then $(\tilde{D}(\epsilon\ v))\epsilon\ v\rangle 0$ which is a contradiction).

2. The mass operator $\mu : TQ \to T^*Q$ defines $D = \mu^{-1}\tilde{D} : TQ \to TQ$ and $(\tilde{D}(v))v < 0$ is equivalent to $(\langle D(v), v\rangle) < 0$ for all $0 \neq v \in TQ$.

It is usual to say that D is a dissipative force when $\tilde{D} = \mu D$ is a dissipative force.

Let us denote by DMS the set of all vector fields $X \in C^r(TQ, TTQ)$ such that X is defined by a dissipative mechanical system, that is, by a pair (V, D) as above. If z is a trajectory of (V, D) and q its projection on Q, then $z = \frac{dq}{dt} = \dot{q}$ and the motion $q = q(t)$ satisfies the generalized Newton law

$$\frac{D\dot{q}}{dt} = -(grad\ \ V)(q) + D(\dot{q}). \tag{5.122}$$

It is useful to remark that the mechanical energy E_m decreases along non trivial integral curves of any mechanical system (V, D). In fact, we have:

$$\dot{E}_m = \frac{d}{dt}(\frac{1}{2}\langle \dot{q}, \dot{q}\rangle + V(q(t))) = \langle D\dot{q}, \dot{q}\rangle$$

which shows that E_m decreases on all integral curves not reduced to a singular point. The singular points of X lie on the zero section $O(Q)$; moreover $0 \in O(Q)$ is a singular point if and only if p is critical for V.

We say that a function $V \in C^{r+1}(Q, \mathbb{R})$ is said to be a **Morse function** if the Hessian of V at each critical point is a non-degenerate quadratic form. It is well known that the set of all Morse functions is an open dense subset of $C^{r+1}(Q, \mathbb{R})$ with the standard C^{r+1} topology.

A dissipative mechanical system (V, D) is said to be **strongly dissipative** if V is a Morse function and D comes from a **strongly dissipative force** that is, satisfies the following additional condition: for all $p \in Q$ and all $\omega \neq 0, \omega \in T_pQ$, one has $\langle\langle d_v D(0_p)\omega, \omega\rangle\rangle < 0$ where $d_v D$ denotes the vertical differential of D.

From now on let us denote by SDMS the set of all $X \in$ DMS such that $X = (V, D)$ is strongly dissipative and by $\mathcal{D}$ the set of all strongly dissipative forces D.

Proposition 24.5 *Let (V, D) be a strongly dissipative mechanical system. Then the following properties hold:*

i) *The singular points of (V, D) are hyperbolic.*

ii) *The stable and unstable manifolds $W^s(0)$ and $W^u(0)$ of a singular point 0 are properly embedded.*

iii) $\dim W^u(0)$ *is the Morse index of V at $\tau(0) \in Q$.*

iv) $\dim W^u(0) \leq \dim Q \leq \dim W^s(0)$.

Exercise 11.5 Prove property (ii) in the last proposition.

Two submanifolds S_1 and S_2 of a manifold M are said to be in **general position** or **transversal** indexsubmanifold ! transversalif either $S_1 \cap S_2$ is empty or at each point $x \in S_1 \cap S_2$ the tangent spaces T_xS_1 and T_xS_2 span the tangent space T_xM.

Let us denote by $SDMS(D)$ the set of all C^r strongly dissipative mechanical systems $X = (V, D)$ with a fixed D. Analogously we introduce the set $SDMS(V)$.

All the subsets of DMS are endowed with the topology induced by the C^r-Whitney topology of $C^r(TQ, TTQ)$.

This topology possesses the Baire property.

Proposition 25.5 *The set of all systems X in $SDMS$ such that their stable and unstable manifolds are pairwise transversal is open in $SDMS$.*

Proposition 26.5 *Assume* $\dim Q > 1$, $r > 3(1+\dim Q)$ *and let $\mathcal{G}$ be the subset of $SDMS(D)$ (resp. $SDMS(V)$) of all systems X such that their invariant manifolds are pairwise transversal. Then $\mathcal{G}$ is open dense in $SDMS(D)$ (resp. $SDMS(V)$).*

As usual, we say that $X \in SDMS$ is **structurally stable** if there exists a neighborhood W of X (in the Whitney C^r-topology) and a continuous map

h from W into the set of all homeomorphisms of TQ (with the compact open topology), such that:

1) $h(X)$ is the identity map;

2) $h(Y)$ takes orbits of X into orbits of Y, for all $Y \in W$, that is, $h(Y)$ is a **topological equivalence** between X and Y.

   If the topological equivalence $h(Y)$ preserves the time, that is, if X_t (resp. Y_t) is the flow map of X (resp Y) and $h(Y) \circ X_t = Y_t \circ h(Y)$ for all $t \in \mathbb{R}$, then we say that $h(Y)$ is a **conjugacy** between X and Y.

   Recall that the subset of all complete C^r vector fields X on a manifold M (the flow map X_t of X is defined for all $t \in \mathbb{R}$) is open in the set of all C^r-vector fields with the Whitney C^r-topology.

Proposition 27.5 *Any complete strongly dissipative mechanical system such that all the stable and unstable manifolds of singular points are in general position is structurally stable and the topological equivalence is a conjugacy.*

If in the last proposition we do not assume the mechanical system to be complete, the same arguments used in the proof also show that the corresponding time-one map flow is a Morse-Smale map in the sense presented in [HMO], then stable with respect to the attractor $\mathcal{A}$ (V, D), which in this case is the union of the unstable manifolds of all singular points of (V, D).

Let us consider an example of a strongly dissipative mechanical system which does not satisfy the conclusions of Proposition 26.5 in the sense that it does not belong to $\mathcal{G}$; it is the system which describes the motions of a particle (unit mass) constrained on the surface Q of a symmetric vertical solid torus of $\mathbb{R}^3$ obtained by the rotation around the x-axis, of a circle defined by the equations $y = 0$ and $x^2 + (z - 3)^2 = 1$. The potential is proportional to the height function of Q and the dissipative force D is given by $D(v) = -cv$, $c > 0$, for all $v \in TQ$. These data define a strongly dissipative mechanical system with Q as the configuration space. The metric of Q is the one induced by the usual inner product of $\mathbb{R}^3$ and the potential is a well known Morse function with four critical points. The symmetry of the problem shows that the unstable manifold of dimension one of a saddle is contained in the stable manifold of dimension 3 of the other saddle hence they are not in general position since $\dim TQ = 4$.

A dissipative force D is said to be **complete** if, for any Morse function V, the vector field associated to (V, D) is complete, that is, all of its integral curves are defined for all time.

Let us consider a linear dissipative field of forces, that is, a function D defined by

$$D(v) = -c(\tau(v))v, \quad \text{for all} \quad v \in TQ$$

where $c : Q \to \mathbf{R}$ is a strictly positive C^r function and Q is compact. It is a simple matter to show that D is a strongly dissipative force. We will show that D is complete. If it were not the case, there would exist a smooth function $V : Q \to \mathbf{R}$ and a motion $t \to q(t)$ of (V, D) whose maximal interval of a existence is $]\alpha, +\infty[$ with $-\infty < \alpha < 0$. We know that $\frac{d}{dt}(E_m(\dot{q})) = \langle D(\dot{q}), \dot{q}\rangle$ is negative and also that

$$0\langle|\langle D(\dot{q}), \dot{q}\rangle| \leq \mu|\dot{q}|^2 \leq 2\mu(E_m(\dot{q}) + k)$$

where $\mu > 0$ is the maximum of c on Q and $k = |\nu|$, ν being the minimum of V on Q. For all t, $\alpha < t < 0$, we may write

$$-2\mu(E_m(\dot{q}) + k) \leq \dot{E}_m(\dot{q}) \leq \frac{d}{dt}(E_m(\dot{q}) + k) < 0$$

or

$$\frac{d(E_m(\dot{q}) + k)}{E_m(\dot{q}) + k} \geq -2\mu dt$$

which implies

$$E_m(\dot{q}) + k \leq (E_m(\dot{q}(0)) + k)e^{-2\mu t}$$

and then $E_m(\dot{q}(t))$ is bounded and strictly decreasing, so there exists $\lim_{t\to\alpha_-} E_m(\dot{q}(t)) = L < +\infty$. This shows that $|\dot{q}|^2 = 2(E_m(\dot{q}) - V(q(t))$ is also bounded, because V is bounded; now it is immediate that we have a contradiction.

Chapter 6

Mechanical systems with non-holonomic constraints

6.1 D'Alembert principle

Let $(Q, \langle , \rangle)$ be a C^∞ riemannian manifold where Q is still called the **configuration space**. A **constraint** Σ is a distribution of subspaces on Q, that is, a map

$$\Sigma : q \in Q \longrightarrow \Sigma_q$$

where Σ_q is a (linear) subspace of T_qQ with $\dim \Sigma_q = m < n = \dim Q$, for all $q \in Q$. Assume also that Σ is C^∞, that is, there exist a neighborhood of each point $q \in Q$ and m C^∞ local vector fields $Y^1, \ldots, Y^m$ that generate Σ_x in all the points x of the neighborhood above. The riemannian metric $\langle , \rangle$ enable us to construct $\Sigma_q^\perp$, the orthogonal subspace to Σ_q, for any $q \in Q$. We have, then, two complementary vector subbundles

$$\Sigma Q = \bigcup_{q \in Q} \Sigma_q \qquad \text{and} \qquad \Sigma^\perp Q = \bigcup_{q \in Q} \Sigma_q^\perp$$

with dimensions $(n + m)$ and $n + (n - m)$, respectively.

There are, well defined, two C^∞ projections denoted by

$$P : TQ \longrightarrow \Sigma Q \qquad \text{and} \qquad P^\perp : TQ \longrightarrow \Sigma^\perp Q$$

that project each $v_q \in T_qQ$ into the orthogonal components $P(v_q) \in \Sigma_q$ and $P^\perp(v_q) \in \Sigma_q^\perp$, respectively.

Let F^k, $k \geq 1$, to be the set of all C^k fields of external forces and F_Σ^k be the subset of all $\mathcal{G} \in F^k$ such that $\mathcal{G}(v) = \mathcal{G}(Pv)$ for all $v \in TQ$. Recall that any

$\mathcal{G} \in F^k$ sends the fiber T_qQ into the fiber T_q^*Q, for all $q \in Q$, and that to define $\mathcal{G} \in F_\Sigma^k$ it is enough to know the values of $\mathcal{G}$ on the vectors $v \in \Sigma Q$.

A **mechanical system with constraints** on a riemannian manifold (Q, $\langle,\rangle$) is a set $(Q, \langle,\rangle, \Sigma, \mathcal{F})$ of data where $\mathcal{F} \in F^k$ is an external field of forces and Σ is a C^∞ constraint.

For our purposes it is convenient to recall now the classical Frobenius theorem. A C^∞ distribution Σ of dimension m on the manifold Q admits, at each point $q \in Q$, the local C^∞ generators vector-fields $Y^1, \ldots, Y^m$, defined in a neighborhood U_q of q. We use to say that a vector-field Y belongs to $\Sigma(Y \in \Sigma)$ if $Y_p \in \Sigma_p$ for all p where Y is defined. It is clear, by this definition, that the $Y^i \in \Sigma, i = 1, \ldots, m$. A distribution Σ is said to be **integrable** if through each point $q \in Q$ passes an integral submanifold $\mathcal{M}$ of Q, that is, $T_x\mathcal{M} = \Sigma_x$ for all $x \in \mathcal{M}$. A **leaf** of an integrable distribution Σ is a maximal integral submanifold $\mathcal{M}$ (so any leaf is connected).

Frobenius theorem - *A C^∞ distribution Σ on Q is integrable if, and only if, $Y, \bar{Y} \in \Sigma$ implies that $[Y, \bar{Y}] \in \Sigma$.*

To check the integrability of Σ it is enough that to each point $q \in Q$ and any corresponding local generators $Y^1, \ldots, Y^m$ we have that $[Y^i, Y^j] \in \Sigma$ for all $i, j = 1, \ldots, m$.

The distribution Σ is often given locally (in an open set U) by the zeros $(n-m)$ linearly independent 1-differential forms $\omega_1, \ldots, \omega_{n-m}$, that is, a vector $v_p \in TQ$, $p \in U$ is also in ΣQ if, and only if, $\omega_\nu(p)(v_p) = 0$ for all $\nu = 1, \ldots, (n-m)$. A dual statement for the Frobenius theorem is the following: *The C^∞ distribution is integrable if, and only if, to each point $q \in Q$ and $(n-m)$ local forms ω_ν, defined in a neighborhood of q, whose zeros span Σ, we have that $d\omega_\nu \wedge \omega_1 \wedge \ldots \wedge \omega_{n-m} = 0$, for all $\nu = 1, \ldots, n-m$.*

When Σ is a non integrable distribution the mechanical system is said to be **non-holonomic**. If, otherwise, Σ is integrable, the mechanical system is said to be **semi-holonomic** . A **true non-holonomic** mechanical system is a non-holonomic one such that the restriction of Σ to any neighborhood of any point of Q is a (local) non-integrable distribution. If Q is a connected analytic manifold with an analytic distribution Σ, the concepts "non-holonomic" and "true non holonomic" coincide.

A C^2-curve $t \mapsto q(t)$ on Q is said to be **compatible** with a distribution Σ if $\dot{q}(t) \in \Sigma_{q(t)}$ for all t.

Given a mechanical system with constraints: $(Q, \langle,\rangle, \Sigma, \mathcal{F})$, and, in order to obtain motions on Q compatible with Σ, we have to introduce **a field of reactive forces** $\mathcal{R} \in F_\Sigma^k$ depending on $Q, \langle,\rangle, \Sigma$ and $\mathcal{F}$ and to consider the

generalized Newton law:

$$\mu(\frac{D\dot{q}}{dt}) = (\mathcal{F} + \mathcal{R})(\dot{q}).$$

A constraint Σ is said to be **perfect** (or to satisfy **d'Alembert principle for constraints**) if, for any $\mathcal{F} \in \mathcal{F}^k$, the field of reactive external forces $\mathcal{R}$ has to satisfy

$$\mu^{-1}\mathcal{R}(v_q) \in \Sigma_q^{\perp} \qquad \text{for any} \qquad v_q \in \Sigma Q.$$

Example 1.6 A planar disc of radius r rolls without slipping along another disc of radius R on the same plane. The equality $rd\theta_1 = Rd\theta_2$ is the physical condition corresponding to the motion without slipping, θ_1 and θ_2 being angles that measure the two rotations. One considers $Q = S^1 \times S^1$ and Σ spanned by the vector fields $v = A\frac{\partial}{\partial\theta_1} + B\frac{\partial}{\partial\theta_2}$ such that $\omega(v) = 0$, $\omega = rd\theta_1 - Rd\theta_2$. Since $n = 2$ and $\dim\Sigma = 1$, Σ is an integrable distribution on the manifold of configurations Q.

Example 2.6 Consider the motion of a vertical knife that is free to slip along itself on a horizontal plane and also free to make pivotations around the vertical line passing through a point P of the knife. Let (x, y) to be cartesian coordinates of P in the horizontal plane and φ the angle between the knife and the x-axis. The manifold of configurations Q is $\mathbb{R}^2 \times S^1$ with (local) coordinates (x, y, φ) and there are physical conditions $dx = ds\cos\varphi$ and $dy = ds\sin\varphi$, ds being the elementary slipping; that implies $(\sin\varphi)dx = (\cos\varphi)dy$ and so Σ on the manifold Q is spanned by the vectors that makes equal to zero the 1-differential form

$$\omega = (\sin\varphi)dx - (\cos\varphi)dy.$$

It can be seen that Σ is a non integrable distribution.

Example 3.6 Motions of a vertical planar disc that one allows to roll without slipping on a horizontal plane and also can make pivotations around the vertical line passing through the center. The manifold of configurations Q is $\mathbb{R}^2 \times S^1 \times S^1$ with local coordinates (x, y, φ, ψ) where (x, y) are coordinates on the horizontal plane for the point of contact between the disc and the plane, φ is the angle between the x-axis and the trace of the plane of the disc with the horizontal plane, and ψ measures the rotation of the disc when rolling. If r is the radius of the given disc, the physical conditions imply that

$$dx = ds\cos\varphi, \qquad dy = ds\sin\varphi \qquad \text{and} \qquad ds = rd\psi.$$

We define two 1-differential forms ω_1 and ω_2 by

$$\omega_1 = dx - r(\cos\varphi)d\psi \qquad \text{and} \qquad \omega_2 = dy - r(\sin\varphi)d\psi$$

and the distribution Σ is spanned by the vectors $v \in TQ$ such that $\omega_1(v) = \omega_2(v) = 0$. So the distribution Σ has dimension $m = 2$ and the dimension of Q is $n = 4$. This analytic distribution is non-integrable.

Exercise 1.6 Prove, using the Frobenius theorem and also through physical arguments, that the constraints in Examples 2.6 and 3.6 are non-integrable.

Given a mechanical system with constraints $(Q, \langle,\rangle, \Sigma, \mathcal{F})$, then a C^2-motion $t \mapsto$ q(t) on Q is compatible with Σ if, and only if,

$$\langle \dot{q}, Z^i \rangle = 0, \qquad i = 1, \dots, (n-m), \tag{6.1}$$

for each t in the interval of definition of the curve, where the local C^∞-vector-fields $(Z^1, \dots, Z^{n-m})$ form an orthonormal set at each point, are defined in a neighborhood of $q(t)$ and span the distribution $\Sigma^\perp$ in all the points of that neighborhood.

To prove the existence of the field of reactive forces, we start by introducing the **total second fundamental form of** Σ:

$$B : TQ \times_Q \Sigma Q \longrightarrow \Sigma^\perp Q, \tag{6.2}$$

defined as follows: if $\xi \in T_qQ, \eta \in \Sigma_q$ and $z \in \Sigma_q^\perp$, let X, Y, Z, three germs of vector-fields at $q \in Q$, $Y \in \Sigma$ and $Z \in \Sigma^\perp$, such that $X(q) = \xi$, $Y(q) = \eta$ and $Z(q) = z$; one defines the bilinear form $B(\xi, \eta)$ by

$$\langle B(\xi, \eta), z \rangle = \langle \nabla_X Y, Z \rangle (q). \tag{6.3}$$

Remark that

$$\langle B(\xi, \eta), z \rangle = -\langle \nabla_X Z, Y \rangle (q), \tag{6.4}$$

and that the number $\langle \nabla_X Y, Z \rangle(q) = -\langle \nabla_X Z, Y \rangle(q)$ depends on the values $X(q), Y(q), Z(q)$, only.

Proposition 1.6 *Given a mechanical system with perfect constraints* $(Q, \langle,\rangle, \Sigma, \mathcal{F})$, $\mathcal{F} \in F^k, k \geq 1$, *then there exists a unique field of reactive forces* $\mathcal{R} \in F_\Sigma^k$ *such that:*

(i) $\mu^{-1}\mathcal{R}(v_q) \in \Sigma_q^\perp$ *for all* $v_q \in \Sigma Q$;

(ii) *for each* $v_q \in \Sigma Q$, *the maximal solution* $t \mapsto q(t)$ *that satisfies*

$$\mu\left(\frac{D\dot{q}}{dt}\right) = (\mathcal{F} + \mathcal{R})(\dot{q}) \tag{6.5}$$

and initial condition $\dot{q}(0) = v_q$, *is compatible with* Σ. *Moreover,*

(iii) *the motion in* (ii) *is of class* C^{k+2} *and is uniquely determined by* $v_q \in \Sigma Q$;

(iv) *The reactive field of forces $\mathcal{R}$ is given by*

$$\mathcal{R}(v_q) = \mu B(v_q, v_q) - \mu([\mu^{-1}\mathcal{F}(v_q)]^{\perp}), \qquad \forall v_q \in \Sigma Q \tag{6.6}$$

$$\mathcal{R}(w_q) = \mathcal{R}(Pw_q), \qquad \text{otherwise.} \tag{6.7}$$

Proof Let $\bar{R} \in \mathcal{F}^k_{\Sigma}$ be another field of forces in the conditions (i), (ii), (iii) of the proposition; then by (ii) we obtain

$$P^{\perp}(\frac{D\dot{q}}{dt}) = P^{\perp}\mu^{-1}\mathcal{F}(\dot{q}) + P^{\perp}\mu^{-1}\bar{R}(\dot{q}),$$

and so

$$\mu^{-1}\bar{R}(\dot{q}) = P^{\perp}\mu^{-1}\bar{R}(\dot{q}) = P^{\perp}(\frac{D\dot{q}}{dt}) - P^{\perp}\mu^{-1}\mathcal{F}(\dot{q}). \tag{6.8}$$

From (6.1) one obtains by covariant derivative:

$$\langle\frac{D\dot{q}}{dt}, Z^i\rangle + \langle\dot{q}, \nabla_{\dot{q}}Z^i\rangle = 0 \tag{6.9}$$

and, since $(Z^1, \ldots, Z^{n-m})$ is orthonormal we have

$$P^{\perp}(\frac{D\dot{q}}{dt}) = \sum_{i=1}^{n-m}\langle(\frac{D\dot{q}}{dt}), Z^i\rangle Z^i. \tag{6.10}$$

From (6.9) and (6.10) it follows

$$P^{\perp}(\frac{D\dot{q}}{dt}) = -\sum_{i=1}^{n-m}\langle\dot{q}, \nabla_{\dot{q}}Z^i\rangle Z^i \tag{6.11}$$

and for $t = 0$ (6.11) implies

$$P^{\perp}(\frac{D\dot{q}}{dt})|_{t=0} = -\sum_{i=1}^{n-m}\langle v_q, \nabla_{v_q}Z^i\rangle Z^i = B(v_q, v_q). \tag{6.12}$$

Then (6.8), for $t = 0$, gives

$$\mu^{-1}\bar{R}(v_q) = B(v_q, v_q) - P^{\perp}\mu^{-1}\mathcal{F}(v_q), \qquad \forall v_q \in \Sigma_q \tag{6.13}$$

and the uniqueness of the field of forces $\mathcal{R} \in \mathcal{F}^k_{\Sigma}$ follows. The conditions (i) and (iii) with $\mathcal{R}$ given by (iv) are trivial ones. It remains only to prove condition (ii). Using the expression (6.6) of $\mathcal{R}(v_q), v_q \in \Sigma Q$, we can look for a C^2 curve $t \longrightarrow q(t)$ on Q, compatible with Σ and satisfying (6.5), or, in other words, verifying

$$\begin{aligned}\frac{D\dot{q}}{dt} &= \mu^{-1}\mathcal{F}(\dot{q}) + B(\dot{q}, \dot{q}) - [\mu^{-1}\mathcal{F}(\dot{q})]^{\perp} = \\ &= P[\mu^{-1}\mathcal{F}(\dot{q})] + B(\dot{q}, \dot{q}).\end{aligned} \tag{6.14}$$

For that, let $(Y^1, \ldots, Y^m)$ be an orthonormal local basis for Σ, and so we need to find functions $v_r(t)$, $r = 1, \ldots, m$, such that one has, locally,

$$\dot{q}(t) = \sum_{k=1}^{m} v_k(t) Y^k. \tag{6.15}$$

Equation (6.14) is equivalent to the following two equations:

$$\sum_{r=1}^{m} \langle \frac{D\dot{q}}{dt}, Y^r \rangle Y^r = P[\mu^{-1}\mathcal{F}(\dot{q})], \tag{6.16}$$

$$\sum_{j=1}^{n-m} \langle \frac{D\dot{q}}{dt}, Z^j \rangle Z^j = B(\dot{q}, \dot{q}). \tag{6.17}$$

But (6.15) and (6.16) give, for any $r = 1, \ldots, m$:

$$\dot{v}_r(t) + \sum_{k=1}^{m} v_k(t) \langle \frac{DY^k}{dt}, Y^r \rangle = \langle P[\mu^{-1}\mathcal{F}(\dot{q})], Y^r \rangle. \tag{6.18}$$

Equations (6.18) define a system of ordinary differential equations that has a unique solution $(v_r(t)), r = 1, \ldots, m$, provided that the values $v_r(0), r = 1, \ldots, m$, are fixed as the components of the vector $v_q \in \Sigma_q$ with respect to the basis $(Y^1(q), \ldots, Y^m(q))$ of Σ_q. On the other hand condition (6.17) is automatically satisfied because it is precisely (6.10). It is clear that (6.15) can be integrated giving us $t \to q(t)$, compatible with Σ, uniquely, since we can fix $q(0) = q \in Q$. The proof of Proposition 1.6 is then complete. ∎

Let us consider again, the vertical lifting operator

$$C_{v_q} : T_qQ \longrightarrow T_{v_q}(TQ)$$

given by the formula (5.21):

$$C_{v_q}(w_q) \stackrel{def}{=} \frac{d}{ds}(v_q + sw_q)|_{s=0},$$

that, in natural coordinates of TQ, if

$$v_q = (q, v) = (q_1, \ldots, q_n, v_1, \ldots, v_n),$$

then C_{v_q} has the expression

$$w_q = (q, w) = (q_1, \ldots, q_n, w_1, \ldots, w_n) \longmapsto ((q, v), (0, w)). \tag{6.19}$$

The following formula holds:

$$C_{v_q}(Pw_q) = TP(C_{v_q} w_q) \quad \text{for all} \quad v_q \in \Sigma Q \quad \text{and} \quad w_q \in TQ \tag{6.20}$$

where TP denotes the derivative of the projection $P : TQ \longmapsto \Sigma Q$. In fact,

$$\begin{aligned} C_{v_q}(Pw_q) &= \frac{d}{ds}(v_q + sPw_q)|_{s=0} = \frac{d}{ds}P(v_q + sw_q)|_{s=0} \\ &= TP\frac{d}{ds}(v_q + sw_q)|_{s=0} = TPC_{v_q}(w_q). \end{aligned}$$

In these local coordinates, since for any C^2 curve $t \longrightarrow q(t)$ one has $\frac{D\dot{q}}{dt} = \sum_{k=1}^{n}(\ddot{q}_k + \sum_{i,j}\Gamma_{ij}^k\dot{q}_i\dot{q}_j)\frac{\partial}{\partial q_k}$, the expression (6.19) for C_{v_q} implies that

$$\begin{aligned} C_{\dot{q}}(\frac{D\dot{q}}{dt}) &= ((q,\dot{q}),(0,(\ddot{q}_k + \sum_{i,j}\Gamma_{ij}^k\dot{q}_i\dot{q}_j)_k)) \quad \text{or} \\ C_{\dot{q}}(\frac{D\dot{q}}{dt}) &= \ddot{q} - S(\dot{q}). \end{aligned} \tag{6.21}$$

where we recall the expression of the geodesic flow of $\langle\ ,\ \rangle$:

$$S(\dot{q}) = ((q,\dot{q}),(\dot{q},(-\sum_{i,j}\Gamma_{ij}^k\dot{q}_i\dot{q}_j)_k) \tag{6.22}$$

and we also have

$$\ddot{q} = ((q,\dot{q}),(\dot{q},(\ddot{q}_k)_k). \tag{6.23}$$

Since $C_{\dot{q}}$ is injective, (6.5) is locally equivalent to the second order ordinary differential equation

$$\ddot{q} = E(\dot{q}) \stackrel{def}{=} S(\dot{q}) + C_{\dot{q}}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}) \tag{6.24}$$

obtained using (6.5) and (6.21). From (6.20) and (6.24) one obtains

$$\begin{aligned} E(\dot{q}) &= S(\dot{q}) + C_{\dot{q}}P([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}) + C_{\dot{q}}P^{\perp}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}) \\ &= TP(S(\dot{q}) + C_{\dot{q}}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q})) + \\ &+ S(\dot{q}) - TP(S(\dot{q})) + C_{\dot{q}}P^{\perp}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}). \end{aligned} \tag{6.25}$$

Since, by the last proposition , the solution $t \rightarrow q(t)$ is compatible with Σ, that is, $\dot{q} = P\dot{q}$, with (6.11) and (6.21) on can write

$$\begin{aligned} C_{\dot{q}}P^{\perp}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}) = C_{\dot{q}}P^{\perp}(\frac{D\dot{q}}{dt}) &= \\ C_{\dot{q}}(\frac{D\dot{q}}{dt}) - C_{\dot{q}}(P\frac{D\dot{q}}{dt}) &= \\ \ddot{q} - S(\dot{q}) - TP(\ddot{q} - S(\dot{q})) = TP(S(\dot{q})) - S(\dot{q}) & \end{aligned} \tag{6.26}$$

because $\dot{q} = P\dot{q}$ implies $\ddot{q} = P\ddot{q}$. Then (6.25) and (6.26) gives us,

$$E(\dot{q}) = S(\dot{q}) + C_{\dot{q}}([\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R}]\dot{q}) = TP(E(\dot{q})). \tag{6.27}$$

The last condition shows, in particular, that given a mechanical system with constraints $(Q, \langle , \rangle, \Sigma, \mathcal{F})$ there is well defined a vector field $v_q \mapsto E(v_q)$ on the vector bundle $\Sigma Q \subset TQ$. In fact we have explicitly:

$$E(v_q) = S(v_q) + C_{v_q}[(\mu^{-1}\mathcal{F} + \mu^{-1}\mathcal{R})v_q] = TP(E(v_q)) \tag{6.28}$$

for all $v_q \in \Sigma_q$.

The vector-field (6.28) defined on the manifold ΣQ is a second order vector-field and then any trajectory is the derivative of its projection on the configuration space Q.

Use the proof above to show, from equation (6.24) and further considerations, that one can solve the following:

Exercise 2.6 Given a mechanical system with perfect constraints $(Q, \langle , \rangle, \Sigma, \mathcal{F})$, $\mathcal{F} \in \mathcal{F}^k (k \geq 1)$, and denoting by $X_{\mathcal{F}}$ the vector-field on TQ corresponding to the mechanical system (without constraints) $(Q, \langle , \rangle, \mathcal{F})$, then the vector field $E = E(v_q)$ associated to $(Q, \langle , \rangle, \Sigma, \mathcal{F})$ is given by $E = TP(X_{\mathcal{F}})$.

The geometrical meaning of the last statement is that at each point v_q of ΣQ we have two elements of $T_{v_q}(TQ)$: the first one is $X_{\mathcal{F}}(v_q)$ and the other is its projection $E(v_q) = TP(X_{\mathcal{F}}\ (v_q))$ that belongs to $T_{v_q}(\Sigma Q)$, that is, we have on ΣQ the equality $E = TP(X_{\mathcal{F}})$.

Example 4.6 A rigid body $S \subset K$ which beside having a fixed point is constrained to move in a such a way that the angular velocity is always orthogonal to a straight line ℓ fixed in S passing through the fixed point. We assume $K = k$ and $B = id$. In this case $Q = SO(k; 3)$ and $m = 2$. Let (e_1, e_2, e_3) a positively oriented basis with e_3 in the direction of ℓ then, as local coordinates in a neighborhood of any given position of S, one can take the Euler's angles (φ, θ, ψ) (see Exercise 10.5) and the distribution Σ is characterized by $C = \dot{\varphi}\cos\theta + \dot{\psi} = 0$ that is, it has a basis given by the vector fields

$$Y^1 = \frac{\partial}{\partial\theta} \quad , \quad Y^2 = \frac{\partial}{\partial\varphi} - \cos\theta\frac{\partial}{\partial\psi},$$

or by the zeros of the one form $w = d\psi + \cos\theta d\varphi$, so, by the Frobenius theorem, Σ is true non holonomic. In fact $[X_1, X_2] = \sin\theta\frac{\partial}{\partial\psi}$. Assume I_1, I_2, I_3 are the moments of inertia with respect to e_1, e_2, e_3, respectively, and that $I_1 = I_2 \neq I_3 > 0$. The kinetic energy is given by

$$K^c = \frac{1}{2}(I_1(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2) + I_3(\dot{\varphi}\cos\theta + \dot{\psi})^2).$$

In the metric of $SO(k;3)$ defined by K^c, the vector field $Z = I_3^{-\frac{1}{2}} \frac{\partial}{\partial \psi}$ is a unit vector orthogonal to Σ. Let us show that, in the present case, $B(\dot{q}, \dot{q}) = 0$. To compute $B(\dot{q}, \dot{q})$, we recall that $\alpha = \sum_{j=1}^{3} (\frac{d}{dt} \frac{\partial K^c}{\partial \dot{q}_j} - \frac{\partial K^c}{\partial q_j}) dq_j$ (here $(q_1, q_2, q_3) = (\varphi, \theta, \psi)$) satisfies $\alpha(v) = \langle \frac{D\dot{q}}{dt}, v \rangle, v \in TQ$. Therefore we have

$$\begin{aligned} -\langle \dot{q}, \nabla_{\dot{q}} Z \rangle &= \langle P^{\perp} \frac{D\dot{q}}{dt}, Z \rangle = \langle \frac{D\dot{q}}{dt}, Z \rangle \\ &= (\frac{d}{dt} \frac{\partial K^c}{\partial \dot{\psi}} - \frac{\partial K^c}{\partial \psi}) = I_3^{\frac{1}{2}} (\dot{\varphi} \cos \theta + \dot{\psi})^{\cdot}. \end{aligned} \tag{6.29}$$

In the present case equation (6.1) becomes $\dot{\varphi} \cos \theta + \dot{\psi} = 0$ that together with (6.23) implies $B(\dot{q}, \dot{q}) = 0$.

6.2 Orientability of a distribution and conservation of volume

Given a mechanical system with constraints say, with data $(Q, \langle , \rangle, \Sigma, \mathcal{F})$, we will come back to the flow defined by the vector field on ΣQ of equation (6.28); such a vector field is also called GMA which stands for Gibbs, Maggi and Appell, who first derived the equations for mechanical systems with non holonomic constraints. The statement of Proposition 1.6 describes the way of finding the C^2-motions $t \to q(t)$ on Q, compatible with the distribution Σ. In fact we have to look for a C^2-curve on Q such that $q(0) = p \in Q$ and $\dot{q}(0) = v_p \in \Sigma Q$ and satisfying the equation (6.14), that is

$$\frac{D\dot{q}}{dt} = P[\mu^{-1} \mathcal{F}(\dot{q})] + B(\dot{q}, \dot{q}).$$

Using the E. Cartan structural equations (see section 4.5) it is also possible to derive the second order ordinary differential equation above. In fact, take in an open neighborhood N_p of $p \in Q$, an orthonormal basis $(X_1, \ldots, X_m)$ for Σ and also an orthonormal basis $(X_{m+1}, \ldots, X_n)$ for $\Sigma^{\perp}$. Then they define the orthonormal basis $(X_1, \ldots, X_m, X_{m+1}, \ldots, X_n)$ of TN_p.

Now we are able to introduce the 1-forms ω^i on N_p by the relations $\omega^i(X_j) = \delta^i_j$, $i, j = 1, \ldots, n$, and we obtain the dual basis $(\omega^1, \omega^2, \ldots, \omega^m, \omega^{m+1}, \ldots, \omega^n)$ of $(X_1, X_2, \ldots, X_n)$. The corresponding structural equations (4.59) and (4.61) are:

$$d\omega^i + \sum_{p=1}^{n} \omega^i_p \wedge \omega^p = 0, \qquad i = 1, \ldots, n,$$

$$\omega^i_p + \omega^p_i = 0, \qquad i, p = 1, \ldots, n,$$

and the distribution Σ is given in terms of these local forms as

$$\Sigma_q = \cap_{\alpha=m+1}^{n} ker\, \omega^{\alpha}(q), \qquad \text{for any} \quad q \in N_p. \tag{6.30}$$

Assume that Σ is perfect, that is, for any given field of external forces $\mathcal{F}$ d'Alembert motions $t \in I \to q(t) \in Q$ imply, for all $t \in I$:

$$\frac{D\dot{q}}{dt} - \mu^{-1}\mathcal{F}(\dot{q}) \in \Sigma^{\perp}_{q(t)},$$

that is, $\dot{q} = \dot{q}(t)$ satisfies, for all $t \in I$:

$$\omega^{\alpha}(\dot{q}) = 0, \qquad \alpha = m+1, \ldots, n, \tag{6.31}$$

$$\omega^{k}\left(\frac{D\dot{q}}{dt} - \mu^{-1}\mathcal{F}(\dot{q})\right) = 0, \qquad k = 1, \ldots, m. \tag{6.32}$$

Let us suppose that $\dot{q}(t) \neq 0$ and also that a local vector field W extends $\dot{q}(t)$, that is, $W(q(t)) = \dot{q}(t)$. Then by (4.26) we have

$$(\nabla_W \omega^{\alpha})(W) = W(\omega^{\alpha}(W)) - \omega^{\alpha}(\nabla_W W),$$

that, computed at $q(t)$ gives

$$(\nabla_{\dot{q}}\omega^{\alpha})(\dot{q}) = \dot{q}(\omega^{\alpha}(\dot{q})) - \omega^{\alpha}(\frac{D\dot{q}}{dt});$$

from (6.31) and (4.58) we obtain

$$\omega^{\alpha}(\frac{D\dot{q}}{dt}) + \sum_{i=1}^{m}\omega_i^{\alpha}(\dot{q})\omega^{i}(\dot{q}) = 0, \qquad \alpha = m+1, \ldots, n. \tag{6.33}$$

The total second fundamental form introduced in (6.2) gives to us:

$$\begin{aligned}
B(\dot{q},\dot{q}) &= \sum_{\alpha=m+1}^{n} \langle B(\dot{q},\dot{q}), X_{\alpha}(q(t))\rangle X_{\alpha}(q(t)) \\
&= \sum_{\alpha=m+1}^{n} \langle (\nabla_{\dot{q}}W)(q(t)), X_{\alpha}(q(t))\rangle X_{\alpha}(q(t)) \\
&= \sum_{\alpha=m+1}^{n} \langle \frac{D\dot{q}}{dt}, X_{\alpha}(q(t))\rangle X_{\alpha}(q(t)) \\
&= \sum_{\alpha=m+1}^{n} \left(\omega^{\alpha}\left(\frac{D\dot{q}}{dt}\right)\right) X_{\alpha}(q(t)),
\end{aligned}$$

then,

$$B(\dot{q},\dot{q}) = -\sum_{\alpha=m+1}^{n}[\sum_{i=1}^{m}\omega_i^{\alpha}(\dot{q})\omega^{i}(\dot{q})]X_{\alpha}(q(t)). \tag{6.34}$$

So, (6.33) and (6.34) imply

$$\omega^{\alpha}(\frac{D\dot{q}}{dt} - B(\dot{q},\dot{q})) = 0, \quad \alpha = m+1, \ldots, n. \tag{6.35}$$

Equations (6.32) also gives:

$$P^{\perp}(\frac{D\dot{q}}{dt}-\mu^{-1}\mathcal{F}(\dot{q}))=\frac{D\dot{q}}{dt}-\mu^{-1}(\mathcal{F}(\dot{q})),\qquad \text{so, from (6.35)}$$

we have

$$P(\frac{D\dot{q}}{dt}-B(\dot{q},\dot{q}))=\frac{D\dot{q}}{dt}-B(\dot{q},\dot{q}).$$

Adding the two last equalities we obtain (6.14). If, otherwise, $\dot{q}(t)=0$ for some $t\in I$, the reactive field of forces $\mathcal{R}$ can be introduced, anyway, by continuity. In fact we obtain (6.6) since (6.34) makes sense for any $v_q\in\Sigma Q$:

$$B(v_q,v_q)=-\sum_{\alpha=m+1}^{n}[\sum_{i=1}^{m}\omega_i^{\alpha}(v_q)\omega^{i}(v_q)]X_{\alpha}(q);\tag{6.36}$$

then, $\mathcal{R}$ is defined by the next two equalities:

$$\begin{aligned}\mathcal{R}(v_q)&\stackrel{def}{=}\mu B(v_q,v_q)-\mu P^{\perp}\mu^{-1}\mathcal{F}(v_q),\qquad \forall v_q\in\Sigma Q,\\ \mathcal{R}(w_q)&\stackrel{def}{=}\mathcal{R}(Pw_q),\qquad \forall w_q\in TQ.\end{aligned}$$

Thus, the generalized Newton law $\mu\frac{D\dot{q}}{dt}=\mathcal{F}(\dot{q})+\mathcal{R}(\dot{q})$ has a meaning on TQ and its flow on TQ leaves ΣQ invariant.

The conservative field of forces $\mathcal{F}(v_q)=-dV(q)$ defined by a C^2 potential energy $V:Q\to\mathbf{R}$ allow us to rewrite (6.14) as

$$\frac{D\dot{q}}{dt}=-(P\ grad\ V)q(t)+B(\dot{q},\dot{q})\tag{6.37}$$

and there is the **conservation of energy** along trajectories on ΣQ. In fact, if $q(t)$ is such that $\dot{q}(0)\in\Sigma_{q(0)}$ and satisfies (6.37) we know by Proposition 1.6 that $\dot{q}(t)\in\Sigma_{q(t)}$ for all t and we have

$$\begin{aligned}\frac{d}{dt}[E_m(\dot{q}(t))]&=\frac{d}{dt}(\frac{1}{2}\langle\dot{q},\dot{q}\rangle+V(q(t)))=\langle\frac{D\dot{q}}{dt},\dot{q}\rangle+[dV(q(t))]\dot{q}(t)\\&=\langle-(P\ grad\ V)q(t),\dot{q}\rangle+\langle(grad\ V)(q(t)),\dot{q}\rangle=0.\end{aligned}$$

The orientability of a distribution Σ, that is, the orientability of the vector subbundle Σ, can be explained in the following way (see Definition 2.1 of [KO1]): "A distribution Σ on the riemannian manifold $(Q,\langle,\rangle)$ is **orientable** if there exists a differentiable exterior $(n-m)$-form Ψ on Q such that, for any $q\in Q$, and any sequence $(z_1,\ldots,z_{n-m})$ of elements in $\Sigma_q^{\perp}$, $\Psi_q(z_1,\ldots,z_{n-m})\neq 0$ if, and only if, $(z_1,\ldots,z_{n-m})$ is a basis of $\sum_q^{\perp}$". In fact this is equivalent to say that ΣQ is orientable. In the codimension one case $(m=n-1)$, $\mathcal{D}$ orientable is equivalent to the existence of a globally defined unitary vector field N, orthogonal to $\sum_q$, $\forall q\in Q$.

In([KO1] Proposition 2.2) it appears a necessary and sufficient condition for the conservation of a volume form in ΣQ:

Proposition 2.6 (Kupka and Oliva) *If Σ is orientable there is a volume form on ΣQ invariant under the flow defined by the mechanical system $(Q, \langle,\rangle, \Sigma, \mathcal{F} = -dV)$ if, and only if, the trace of the restriction of $B^{\perp}$ (total second fundamental form of $\Sigma^{\perp}$) to $\Sigma^{\perp}Q \times_Q \Sigma^{\perp}Q$, vanishes.*

The conservation of a volume form means that there is a (global) non zero exterior $(n+m)$-form ω on ΣQ such that the Lie derivative $L_X\omega = 0$, X being the GMA vector field associated to the data $(Q, \langle,\rangle, \Sigma, \mathcal{F} = -dV)$.

Remark finally that Proposition 2.6 remains true for the flow defined on ΣQ by the equation

$$\frac{D\dot{q}}{dt} = P[\mu^{-1}\mathcal{F}(\dot{q})] + B(\dot{q}, \dot{q})$$

when $\mathcal{F}$ is a positional field of external forces that is, $\mu^{-1}\mathcal{F}$ is a vector field on Q (not necessarily a gradient vector field).

6.3 Semi-holonomic constraints

Let $N \subset Q$, $0 < n = \dim N < \dim Q = m$, a C^∞ submanifold, that is, a C^∞ holonomic constraint of a mechanical system $(Q, \langle,\rangle, \mathcal{F})$. Take a tubular neighborhood of N in Q (see Proposition 9.4) and $p : \Omega \to N$ the projection from the tube Ω onto N (recall that Ω is an open set of Q that contains N).

Fix $x \in N$ and consider the fiber $p^{-1}(x) \subset \Omega \subset Q$. Take $y \in p^{-1}(x)$ and use the Levi-Civita connection to construct Σ_y as the subspace of $T_y\Omega$ whose vectors are obtained from the elements of T_xN by parallel transport along the unique geodesic $\gamma = \gamma(s)$ passing through x at $t = 0$ with velocity $\dot{\gamma}(0) = \exp_x^{-1}(y)$; we also have $\gamma(1) = y$. If we make x vary in N one obtains on Ω a C^∞ distribution. The sequence of data $(\Omega, \langle,\rangle, \Sigma, \mathcal{F})$ defines on Ω a mechanical system with an integrable constraint Σ. This way the holonomic constraint has been considered as a constraint of a **semi-holonomic** system.

Exercise 3.6 Consider Proposition 1.6 applied to the mechanical system with constraints $(\Omega, \langle,\rangle, \Sigma, \mathcal{F})$; assume the submanifold $N \subset \Omega \subset Q$ thought as a holonomic constraint for the holonomic mechanical system $(Q, \langle,\rangle, \mathcal{F})$; compare the field of reactive forces given by (6.6) and (6.7) with the reaction of the constraint defined in 5.29. Show that the motions compatible with N are the same in both cases.

6.4 The attractor of a dissipative system

The next notions and results that will be state in this chapter, appear in the paper [FO] "*Dissipative systems with constraints*" by G. Fusco and W.M. Oliva, Journal of Differential Equations, vol. 63, n° 3, July 1986, p. 362-388. We will describe the discussion that was made there on the qualitative behavior of the flow defined by the vector field on ΣQ given by equation (6.28) called the GMA vector field. We shall focus our attention on the set $\mathcal{A}$ given by the initial conditions in ΣQ of all global bounded solutions of (6.28). As we shall see, strictly dissipativeness implies that $\mathcal{A}$ is a global attractor.

For the study and the statements we will present from now on, the GMA or $(Q, \langle , \rangle, \Sigma, \mathcal{F})$ has Q compact and Σ perfect. Assume the field of forces $\mathcal{F} : TQ \to T^*Q$ is a C^k function, $k \geq 1$, given by $\mathcal{F} = d(V \circ \tau) + \tilde{D}$, such that $V : Q \to \mathbf{R}$ is a C^{k+1} function and $D = \mu^{-1}\tilde{D}$ is **dissipative with respect to** Σ, that is, $\langle PD(v), v\rangle \leq 0$ for each $v \in \Sigma Q$, **strictly dissipative** if $\langle PD(v), v\rangle = 0$ implies $v = 0$, **strongly dissipative** if there is a continuous function $c : Q \to \mathbf{R}^+ \setminus \{0\}$ such that $\langle PDv, v\rangle \leq -c|v|^2$. The GMA is said to be **dissipative (strictly dissipative)** if the function $V : Q \to \mathbf{R}$ is C^{k+1} and D is dissipative (strictly dissipative) with respect to Σ. A strictly dissipative GMA is said to be **strongly dissipative** if V is a Morse function and D is strongly dissipative. Denote by $O : Q \to \Sigma Q \subset TQ$ the zero section and by X_V the vector field on Q defined as the orthogonal projection on Σ_q of $(\text{grad} V)(q)$, for any $q \in Q$, that is, $X_V = P\text{grad}V$.

Exercise 4.6 Compare these notions of dissipativeness with the ones presented in chapter 5.

Proposition 3.6 (i) *The set G^{k+1} of potential functions $V \in C^{k+1}(Q, \mathbf{R})$ ($k \geq 1$) such that $X_V(Q)$ and $O(Q)$ are transversal is open and dense in $C^{k+1}(Q, .\mathbf{R})$;*

(ii) *If $V \in G^{k+1}$, then the set C_V of the critical points of* GMA, *or equivalently, the set ϵ_V of the equilibria of the underlying dissipative system is a C^k compact manifold of dimension $r = \dim Q - \dim \Sigma$;*

iii) *C_V, ϵ_V depend C^k continuously on $V \in G^{k+1}$.*

From this theorem it follows that, generically, for a holonomic mechanical system, the set of equilibria is made of a finite number of points; when $r = 1$ as in the case of the rigid body in Example 4.6, the set of equilibria is generically the union of a finite number of circles.

Proposition 4.6 *The trajectories $t \to v(t)$ of the* GMA *vector field associated with a dissipative system are globally defined in the future and bounded. If the system is strictly dissipative all trajectories approach the set C_V of the critical points as $t \to \infty$. Moreover if $t \to v(t)$ is defined also for negative time and* *bounded, then $v(t)$ approaches C_V as $t \to -\infty$.*

Strictly dissipativeness implies that all trajectories of GMA approach the set of critical points but it is not a sufficient condition in order that the ω-limit set of any orbit contains just one point. For instance, when Q is a circle $C \subset \mathbf{R}^3$, s the curvilinear abscissa along C, T the unit vector tangent to C at s, $V = 0$, $D(vT)(wT) = -v^2 w$ for all $v, w \in \mathbf{R}$, the equations of motion take the form $\dot{s} = v$, $\dot{v} = -v^3$. From that, $v \to 0$ as $t \to \infty$ while s grows unboundedly if the initial value is not zero. Therefore the ω-limit of any orbit through any point in $TC \setminus O(C)$ is all the $O(C)$.

The main point in this example is the nongenericity of V; in fact we know that for $r = 0$ and $V \in G^{k+1}$ the critical points of GMA are isolated and then the ω-limit set of any orbit must be a single point if the system is strictly dissipative. In the case $r = 1$, even for $V \in G^{k+1}$, the critical points are not isolated. Using a general theorem in transversality theory have the following result:

Proposition 5.6 *Let $r = 1$. Then there is an open and dense set in G^{k+1}, $k \geq 2$, such that if a function V is in this set, V is a Morse function and there are at most a finite number of points in C_V for which $V|C_V$ is not strictly monotonic. Moreover, if the system is strictly dissipative, then the ω-limit set of any orbit of the GMA contains just one point. The same is true for the α-limit set of any negatively bounded orbit.*

The next proposition concerns the case of a generic value of r and gives conditions in order that the ω-limit of any orbit contains just one point. We state the theorem without specific reference to the GMA because the result can be applied to any evolutionary equation that satisfies the property that the ω-limit set of any bounded orbit contains only critical points.

Proposition 6.6 *Suppose that the ω-limit set $\omega(\gamma)$ of a bounded orbit γ of a vector field $X \in C^1(\Omega, \mathbf{R}^n)$ contains only critical points. Then a sufficient condition in order that $\omega(\gamma)$ contains just one point is that the local center manifold at each critical point coincides locally with the set of critical points. A similar result holds true for the α-limit set of a negatively bounded orbit.*

We now begin the study of $\mathcal{A}$ by giving a characterization of the attractor and some of its properties.

Proposition 7.6 *If $\Phi : \mathcal{D} \subset (\Sigma Q) \times \mathbf{R} \to \Sigma Q$ is the dynamical system associated with a Σ-strictly dissipative mechanical system and*

$$\mathcal{A} = \{x \in \Sigma Q | \Phi(x,t) \text{ is defined for } t \in (-\infty, +\infty) \text{ and bounded}\},$$

then:

(i) $\mathcal{A}$ it is compact, connected, invariant and maximal.

(ii) *$\mathcal{A}$ is uniformly asymptotically stable for the flow Φ.*

(iii) *$\mathcal{A}$ is an upper semicontinuous function of the potential V and of the dissipative field of force D.*

(iv) *If Φ_1 is the time one map associated with Φ and $\mathcal{B} = \{x \in \Sigma Q | E_m(x) < a\}$ with a sufficiently large $a > 0$, then $\mathcal{A} = \bigcap_{n \geq 0} \Phi_1^n \ \mathcal{B}$.*

It is interesting to remark that, if the α-limit set of any negatively bounded orbit contains just one point, as for instance in the cases described in Propositions 3.6, 4.6 and 5.6, then $\mathcal{A} = \bigcup_{x \in C_V} \ W_x^u$, W_x^u being the unstable manifold corresponding to the critical point x.

One of the basic questions in the description of the structure of $\mathcal{A}$, which is a subset of ΣQ, is to see how is its relation with the configuration space Q. The following theorem says that $\mathcal{A}$ is at least as large as Q.

Proposition 8.6 *Let $\mathcal{A}$ be the attractor of a strictly dissipative system. Then the image of $\mathcal{A}$ under the canonical projection $\tau : TQ \to Q$ is all the (compact) configuration space.*

This result implies that given any point $q \in Q$ there is a $v_q \in \Sigma_q$ such that the orbit of GMA through v_q is globally defined and bounded.

The next theorem gives conditions in order that the attractor and the configuration space have the same dimension. **Proposition 9.6** *If the GMA is strongly dissipative (so V is a Morse function) and $\mathcal{A}$ is a differentiable manifold then* $\dim \mathcal{A} = \dim Q$.

In the remaining part of this section we shall discuss some aspects of the dependence of the attractor on the potential function V and on the dissipative field of forces D.

Proposition 10.6 *Given a strongly dissipative field of forces $D \in C^k$ (with the Whitney topology) there is a neighborhood $\mathcal{N}$ of $0 \in C^{k+1}(Q, \mathbb{R})$ such that, if $\mathcal{A}^V$ is the attractor corresponding to $V \in \mathcal{N}$ and the given D, then*

(i) *$\mathcal{A}^V$ is a C^k differentiable manifold and $\tau | \mathcal{A}^V$ is a C^k diffeomorphism of $\mathcal{A}^V$ onto Q.*

(ii) *$\mathcal{A}^V$ depends C^k continuously on $V \in \mathcal{N}$ and $\mathcal{A}^0 = O(Q)$.*

Since all the orbits of GMA approach the attractor as $t \to \infty$, once $\mathcal{A}$ is known, an important step towards understanding the flow is to know the flow of GMA on the attractor. When, as in the situation of the last theorem, $\mathcal{A}$ is diffeomorphic to Q, to study the flow on $\mathcal{A}$ is the same as to study a first order equation on Q. We consider a potential of type ϵV, $V \in C^2(Q, \mathbf{R}), \epsilon \geq 0$, and a strongly dissipative field of forces $D \in C^1$; then Theorem 9.6 implies that the attractor $\mathcal{A}^\epsilon$ is a C^1 manifold diffeomorphic to $O(Q)$ and approaches $O(Q)$ in the C^1 sense as $\epsilon \to 0$. This implies that $\tau|\mathcal{A}^\epsilon$ is a diffeomorphism of $\mathcal{A}^\epsilon$ onto Q if ϵ is sufficiently small. It follows that given $q \in Q$, there is a unique point $(\tau|\mathcal{A}^\epsilon)^{-1}(q)$ in $\Sigma_q \cap \mathcal{A}^\epsilon$. Therefore $t \longrightarrow \dot{q}_\epsilon(t)$ is an orbit of GMA in $\mathcal{A}^\epsilon$ if and only if

$$\dot{q}_\epsilon(t) = (\tau|\mathcal{A}^\epsilon)^{-1}(q_\epsilon(t)),$$

that is, if and only if the corresponding motion $t \to q_\epsilon(t)$ is a solution of the first order equation

$$\dot{q}_\epsilon = X^\epsilon(q) \stackrel{def}{=} (\tau|\mathcal{A}^\epsilon)^{-1}(q).$$

The vector field X^ϵ depends on $\mathcal{A}^\epsilon$ and cannot be computed explicitly unless one knows $\mathcal{A}^\epsilon$ which is not, in general, the case. Since $\mathcal{A}^\epsilon$ approaches $O(Q)$ as $\epsilon \to 0$, we have $X^\epsilon(q)$ approaches zero as $\epsilon \to 0$, thus we consider the vector field $Y^\epsilon \stackrel{def}{=} \epsilon^{-1} X^\epsilon$ which has the same orbits as X^ϵ and study the limit Y^0 of Y^ϵ as $\epsilon \to 0$. If Y^0 exists and is structurally stable then for ϵ sufficiently small the flow of X^ϵ is topologically equivalent to Y^0. If $(q, \dot{q}) = (q, v)$ are natural local coordinates on TQ then the function $\sigma^\epsilon := O(Q) \to \Sigma Q$ describing $\mathcal{A}^\epsilon$ has a local representation $q \to (\bar{q}(\epsilon, q), \bar{v}(\epsilon, q))$, $\bar{q}(\epsilon, .)$, $\bar{v}(\epsilon, .)$ are C^1 functions such that $\bar{q}(\epsilon, .) \to \mathrm{id}$, $\bar{v}(\epsilon, .) \to 0$, in the C^1 topology, as $\epsilon \to 0$.

Moreover $\bar{q}(\epsilon, .)$ has a C^1 inverse because $(\tau|\mathcal{A}^\epsilon)$ is a C^1 diffeomorphism and the same is true for σ^ϵ.

Proposition 11.6 *If $\mathcal{A}^\epsilon$ is a smooth function of ϵ in the sense that $\bar{q}, \bar{v}$ and their derivatives with respect to q are continuously differentiable with respect to ϵ then, as $\epsilon \to 0$, Y^ϵ converges in the C^1 sense to the C^1 vector field given by $Y^0 = -(P \circ (FD))^{-1} P\ \mathrm{grad} V$, FD being the fiber (vertical) derivative of D.*

Remark that $P \circ (FD) : \Sigma Q \to \Sigma Q$ is a diffeomorphism because D is a strongly dissipative field of forces.

Chapter 7

Hyperbolicity and Anosov systems. Vaconomic mechanics

7.1 Hyperbolic and partial hyperbolic structures

In Chapter 5, section 5.7, we saw that, generically, holonomic dissipative mechanical systems have a very simple dynamics with a Morse-Smale flow and, moreover, they are structurally stable and the topological equivalence is a conjugacy (see Propositions 18.5, 19.5, 20.5, 21.5, and [KO]).

During many years the mathematical community believed that the structural stability of flows was generically related with simple structures; in fact, that is true in two dimensions. But, in 1967, D.V. Anosov studied, extensively, special flows, nowadays called **Anosov flows**, which are structurally stable and constitute a class of non trivial and complex dynamical systems. Moreover, an Anosov flow which is Hölder C^1 and has an invariant measure (generated by a volume's form) is ergodic.

The structural stability for Hölder C^1 Anosov flows, as well as the **ergodicity** when there is an invariant measure, were proved by Anosov in his book [A] where one can also see a proof of the fact that the geodesic flow on the unitary tangent bundle of a compact riemannian manifold having strictly negative all their sectional curvatures satisfies definition 7.1 below and, moreover, is Anosov; other geometrical proofs of this last fact are also available in Arnold and Avez [AA] as well as an analytical proof in Moser [M]. As a matter of fact, the last result goes back to Hadamard's work that, essentially, gave a proof for it; in [H] 1934, Hedlund proved the ergodicity of the geodesic flow on the (3-dimensional) unitary tangent bundle of a closed surface, with constant and strictly negative curvature, and Hopf in [Ho], 1940, extended the result for the general case of surfaces with strictly negative curvature.

Definition 7.1 Let M be a C^∞ compact riemannian manifold. A non singular flow $T^t : M \to M$ is partially hyperbolic if the (derivative) variational flow $DT^t : TM \to TM$ satisfies:

(i) for any $p \in M$, $T_pM = \mathcal{X}_p \oplus \mathcal{Y}_p \oplus \mathcal{Z}_p$, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are invariant sub-bundles of TM, $dim\mathcal{X}_p = \ell \geq 1$, $dim\mathcal{Y}_p = k \geq 1$, $\mathcal{Z}_p \supset [(T^tp)'_{t=0}]$;

(ii) there exist $a, c > 0$ such that

$$\begin{aligned} |DT^t\xi| &\leq a|\xi|e^{-ct}, \quad \forall t \geq 0, \quad \forall \xi \in \mathcal{X}_p, \quad p \in M, \\ \text{that is} \quad |DT^t\xi| &\geq a^{-1}|\xi|e^{-ct}, \quad \forall t \leq 0, \quad \forall \xi \in \mathcal{X}_p, \quad p \in M; \\ |DT^t\mu| &\leq a|\mu|e^{ct}, \quad \forall t \leq 0, \quad \forall \mu \in \mathcal{Y}_p, \quad p \in M, \\ \text{that is} \quad |DT^t\mu| &\geq a^{-1}|\mu|e^{ct}, \quad \forall t \geq 0, \quad \forall \mu \in \mathcal{Y}_p, \quad p \in M; \end{aligned}$$

$\mathcal{X}$ and $\mathcal{Y}$ are said to be uniformly contracting and expanding, respectively;

(iii) $\mathcal{Z}$ is neutral in the sense that it is neither uniformly contracting nor uniformly expanding;

If, in particular, $\mathcal{Z}_p = [(T^tp)'_{t=0}]$, (iii) is satisfied and the flow is said to be hyperbolic or Anosov.

Under that definition one uses to say that the manifold M has a **partial hyperbolic structure under T^t** (**hyperbolic structure** in the Anosov case).

The flows with hyperbolic behavior on the trajectories , and the structure of manifolds of non positive curvature were considered in two surveys, respectively, by Pesin in [Pe], 1981, and by Eberlein in [E], 1984; both papers present an extensive and fundamental list of references on the subjects under consideration.

Castro and Oliva in [CO], constructed an Anosov flow obtained as the quotient by a suitable vector field of a partially hyperbolic flow over a codimension one true non-holonomic orientable distribution of a compact riemannian manifold. The distribution is constant umbilical (see [CO]) and conserves volume (see the previous Chapter 6, section 6.2.). The manifold is supposed to have sufficiently negative sectional curvatures on the 2-planes contained in the distribution and only on them. An explicit example is also presented there.

In [CKO] the authors presented examples of partially hyperbolic flows motivated by the study of the Σ-geodesic flows i.e., dynamic free systems (see (6.37) with $V = 0$). Suitable conditions properly decouple its variational equation and imply the hyperbolic properties of the trajectories; the cases of a general Lie group and of a semi-simple Lie group are also analyzed.

In [G], 1997, Gouda regarded a magnetic field as a closed 2-form-$\tilde{B}$ on a riemannian manifold Q and defined a magnetic flow which is, in fact, a perturbation of a geodesic flow. A sufficient condition is presented there for a magnetic flow to become an Anosov flow (see [G], Theorem 7.2). The second order differential system considered in [G] is a holonomic mechanical system; the closed 2-form $\tilde{B}$ on Q defines the Lorentz field of forces:

$$\Omega : TQ \to T^*Q,$$

by $\Omega(v_p)(w_p) = \langle w_p, \tilde{B}_p(w_p, v_p)\rangle$ for all $v_p, w_p \in T_pQ$. The generalized Newton law (see Chapter 5, section 5.1) defines, for that field of forces, the second order mechanical system introduced by Gouda:

$$\mu\left(\frac{D\dot{q}}{dt}\right) = \Omega(\dot{q}),$$

where $\mu : TQ \to T^*Q$ is the mass operator.

It is our understanding that many interesting questions, especially in the non-holonomic context, can be analyzed trying to obtain more examples giving rise to other kinds of complex and hyperbolic dynamics.

7.2 Vaconomic mechanics

The non-holonomic mechanics has two fundamental approaches for its development. One is based in the d'Alembert principle for which we gave the foundations in Chapter 6. But, since many years, it is known that the D'alembert approach, for the (true) non-holonomic mechanics, does not have a parallel within the so called variational principles.

In [AK], Arnold and Kozlov introduced the non-holonomic Mechanics under the Lagrange variational point of view for constrained systems; then it appeared the so-called **Vaconomic Mechanics**. Vershic and Gershkovich also developed that approach including in the survey [VG] most of the recent contributions that appeared on that field of geometric mechanics.

Kupka and Oliva wrote a monograph (see [KO1]) on the non-holonomic mechanics trying to put in evidence the main differences between the d'Alembertian and the Vaconomic approaches. The last one also works with motions compatible with a (non-integrable) distribution by they are determined by a variational method; in fact, each motion $q(t)$ with values on the configuration space Q corresponding to a potential V and a distribution Σ, is a stationary point $q = q(t)$ of a functional $\mathcal{L}$, given by $\mathcal{L}(q) = \int_o^T [\frac{1}{2} \parallel \dot{q} \parallel^2 - V(q)]dt$; $\mathcal{L}$ is defined in a suitable Hilbert manifold. The stationary points of $\mathcal{L}$ are the **non-holonomic variational trajectories**; there are **singular** and **regular** (non singular) trajectories and they are characterized in [KO1] . The (global) second order ordinary differential equation for the regular vaconomic trajectories is also derived there; it defines a flow of a Hamiltonian vector field on the tangent bundle considered as the Whitney sum $\Sigma Q \oplus \Sigma^{\perp} Q$, on the configuration space. The solutions of that vector-field are, then, of type $(\dot{q}(t), P(t))$ where $\dot{q}(t) \in \Sigma_{q(t)}$ and $P(t) \in \Sigma^{\perp}_{q(t)}$, $q = q(t)$ being a regular vaconomic trajectory. The component $q(t)$ is compatible with Σ and the kinetic energy is conserved, but the bundle ΣQ is **not invariant** under the flow; the component $P(t)$ gives, locally, the classical Lagrange multipliers.

It is particularly interesting, to analyze the hyperbolic and all the ergodic aspects of the vaconomic flows; some of them, already appear in [VG] and this investigation still remains as a very nice field of research.

Chapter 8

Special relativity

As we already said in Chapter 1, the first difficulty that arises in Newtonian mechanics is the fact that no material object has been observed traveling faster than the speed $\mathbf{c}$ of the light in a vacuum. The way to eliminate that is to consider the tangent vector $(1, \dot{\alpha}(t))$ to the particle's world line $(t, \alpha(t))$ in $(\mathbf{R} \times \mathbf{R}^3)$ (relatively to an inertial coordinate system), and compare $|\dot{\alpha}(t)|$ with $\mathbf{c}$. Since one needs to obtain $|\dot{\alpha}(t)| < \mathbf{c}$, it is enough to observe that the tangent directions of pulses of light, always at constant speed $\mathbf{c}$, define a circular cone at each point of $\mathbf{R} \times \mathbf{R}^3$, with vertex in that point, semi-angle φ equal to $\arctan \mathbf{c}$ and axis parallel to the time axis $\mathbf{R}$; then we require the motion $\alpha(t)$ be such that, for each t, the vector $(1, \dot{\alpha}(t))$ is inside the corresponding cone at the point $(t, \alpha(t))$ of the world line.

In the context of pseudo-riemannian geometry, the idea is to change the sign in the time coordinate of the metric tensor on $\mathbf{R} \times \mathbf{R}^3$; with this idea one constructs some special quadratic cones to argüe with, as above. As we will see, this is a starting point to introduce **special relativity**.

From now on we will assume that units were chosen so that the fundamental constant, the speed of light, is unity, that is, we shall assume $\mathbf{c} = 1$, so $\varphi = \pi/4$.

This Chapter 8 has its presentation based in part on chapter 5. and 6. of the book [ON] "Semi-Riemannian Geometry - with applications to Relativity" by B. O'Neill, Academic Press, 1983.

8.1 Lorentz manifolds

Let $(Q, \langle , \rangle)$ be a pseudo-riemannian manifold. The **index** of $\langle , \rangle$ at $p \in Q$ is the largest integer which is the dimension of a subspace $W \subset T_pQ$ such that the restriction of the quadratic form $\langle , \rangle_p$ to W is negative definite. Since Q is supposed to be connected and the bilinear form $\langle u_p, v_p\rangle$ is symmetric and non degenerate, the index is constant with $p \in Q$. So, one can talk about the index ν of $(Q, \langle,\rangle)$. We have $0 \le \nu \le n = \dim Q$ and it is clear that $\nu = 0$, if, and only if, $\langle,\rangle$ is a riemannian metric. If we fix an orthonormal basis $(e_1, \ldots, e_n)$ for T_pQ (with respect to $\langle,\rangle$), for each vector $v_p = \sum_{i=1}^n v_i e_i$ one can write $v_p = \sum_{i=1}^n \varepsilon_i \langle v_p, e_i\rangle e_i$ where $\varepsilon_i = \langle e_i, e_i\rangle = +1$ or -1. The number of ε_i equal to -1 is the index ν.

Examples

8.1. $Q = \mathbf{R}^2$, $\langle v, w\rangle = v_1w_1 - v_2w_2$ for $v = (v_1, v_2), w = (w_1, w_2)$. In this case $\nu = 1$.

8.2. $Q = \mathbf{R}^{n+1}$ and $\langle,\rangle$ is such that $v = (v_1, \ldots, v_{n+1})$ implies

$$\langle v, v\rangle = -(v_1^2 + \ldots + v_\nu^2) + v_{\nu+1}^2 + \ldots + v_{n+1}^2. \tag{8.1}$$

In this case the index of $(\mathbf{R}^{n+1}, \langle,\rangle)$ is equal to ν and the pseudo-riemannian manifold $(\mathbf{R}^{n+1}, \langle,\rangle)$ is simply denoted by $\mathbf{R}_\nu^{n+1}$.

Definition 8.1 A Lorentz manifold is a pseudo-riemannian manifold with index $\nu = 1$. The Lorentz manifold $\mathbf{R}_1^n$ is called the **Minkowski n-space**.

Let $(Q, \langle,\rangle)$ be a Lorentz manifold. There are three categories of tangent vectors:

Definition 8.2 A vector $v \in T_pQ$ is said to be

(i) **space like** if $\langle v, v\rangle > 0$ or $v = 0$;

(ii) **light like** or **null** if $\langle v, v\rangle = 0$ and $v \neq 0$;

(iii) **time like** if $\langle v, v\rangle < 0$.

The set of null vectors in T_pQ is called the **null cone at** $p \in Q$.

Proposition 8.1 *Let ∇ be the Levi-Civita connection associated to $\langle,\rangle$ in the Lorentz manifold $(Q, \langle,\rangle)$. Then, the tangent vectors to a geodesic of ∇ belong always to one and the same category.*

Proof If $t \longrightarrow q(t)$ is a geodesic and $\frac{D}{dt}$ is the covariant derivative associated to ∇ one obtains:

$$\frac{d}{dt}\langle \dot q, \dot q\rangle = 2\langle \dot q, \frac{D\dot q}{dt}\rangle = 0, \quad \text{because} \quad \frac{D\dot q}{dt} = 0;$$

so, $\langle \dot{q}(t), \dot{q}(t)\rangle$ does not depend on t. ∎

Exercise 8.1 The last proposition is not true for a general smooth curve on Q; show this with a counter-example.

8.2 The quadratic map of $\mathbf{R}_1^{n+1}$

The quadratic form q on $\mathbf{R}_1^{n+1}$ defined with $\langle , \rangle$ and $\nu = 1$ in (8.1) is given by

$$q(u) = \langle u, u\rangle = -u_1^2 + u_2^2 + \ldots + u_{n+1}^2, \quad u = (u_1 \ldots u_{n+1}) \in \mathbf{R}^{n+1} \tag{8.2}$$

and is called the **quadratic map** of the Minkowski space $\mathbf{R}_1^{n+1}$. If $u = (u_1, \ldots, u_{n+1})$, one can write: $u = \Sigma_{i=1}^{n+1} u_i e_i$, $(e_1, \ldots, e_{n+1})$ being the canonical basis of $\mathbf{R}^{n+1}$. So,

$$q(u) = \sum_{i,j=1}^{n+1} g_{ij} u_i u_j \tag{8.3}$$

where $g_{ij} = \langle e_i, e_j\rangle = g_{ji}$.

Proposition 8.2 *The symmetric matrix (g_{ij}) of the bilinear form $\langle u, v\rangle$, associated with the quadratic map (8.2) of $\mathbf{R}_1^{n+1}$, is diagonal with $g_{11} = -1$ and $g_{ii} = +1, i = 2, \ldots, n+1$.*

Proof In fact, as usually, the formula

$$\langle u + v, u + v\rangle = \langle u, v\rangle + \langle v, v\rangle + 2\langle u, v\rangle \tag{8.4}$$

gives

$$\langle u, v\rangle = \frac{1}{2}\{\langle u + v, u + v\rangle - \langle u, u\rangle - \langle v, v\rangle\}. \tag{8.5}$$

Since by (8.2) we have

$$\langle u, u\rangle = \langle (u_1, \ldots, u_{n+1}), (u_1, \ldots, u_{n+1})\rangle = -u_1^2 + u_2^2 + \ldots + u_{n+1}^2, \tag{8.6}$$

and because $g_{ij} = \langle e_i, e_j\rangle$, we use (8.5) and (8.6) and one completes the proof. ∎

The only critical point of the map q in (8.2) (or (8.6)) is the origin in $\mathbf{R}^{n+1}$; so, any real number (except zero) is a regular value of q. The vector gradient of q at $w \in \mathbf{R}_1^{n+1}$, is, by definition, given by

$$\langle (grad\ q)(w), v\rangle = [dq(w)](v), \forall v \in T_w \mathbf{R}_1^{n+1} = \mathbf{R}^{n+1}. \tag{8.7}$$

Using (8.3) and Proposition 8.2 we see that $dq(w) = -2w_1du_1 + 2\sum_{i=2}^{n+1} w_idu_i$, so

$$[dq(w)](v) = -2v_1w_1 + 2v_2w_2 + \ldots + 2v_{n+1}w_{n+1}. \tag{8.8}$$

Since (8.7) and (8.8) imply

$$\langle (grad\ \ q)(w), e_1\rangle = -2w_1,\ \ \langle (grad\ \ q)(w), e_i\rangle = 2w_i,\quad i = 2,\ldots,n+1,$$

then

$$(grad\ \ q)(w) = 2w_1e_1 + 2\sum_{i=2}^{n+1} w_ie_i = 2w \tag{8.9}$$

and

$$\langle (grad\ \ q)(w), (grad\ \ q)(w)\rangle = -4w_1^2 + 4w_2^2 + \ldots + 4w_{n+1}^2 = 4q(w). \tag{8.10}$$

Given $r > 0$ and $\varepsilon = \pm 1$, the number εr^2 is a regular value of q; so, $Q^n \overset{def}{=} q^{-1}(\varepsilon r^2)$ is an imbedded n-dimensional submanifold of $\mathbf{R}^{n+1}$ called a **central hyperquadric**. Take $w \in Q^n$; by (8.7) and (8.10) we have

$$q(w) = -w_1^2 + w_2^2 + \ldots + w_{n+1}^2 = \varepsilon r^2$$

and

$$dq(w)[(grad\ \ q)(w)] = 4q(w) = 4\varepsilon r^2.$$

But the tangent space T_wQ^n is the set $T_wQ^n = \{v \in \mathbf{R}^{n+1} | [dq(w)]v = 0\}$, that is, by (8.7), $T_wQ^n = \{v \in \mathbf{R}^{n+1} | \langle (grad\ \ q)(w), v\rangle = 0\}$.

If one considers an orthogonal basis $(v_1,\ldots,v_n)$ of T_wQ^n (with respect to the metric induced on Q^n by $\langle,\rangle$), then $(\frac{1}{2r}(grad\ \ q)(w), v_1,\ldots,v_n)$ is an orthonormal basis of $T_w\mathbf{R}^{n+1} = \mathbf{R}^{n+1}$. Since (8.8) and $w \in Q^n$ imply

$$\langle \frac{(grad\ \ q)(w)}{2r}, \frac{(grad\ \ q)(w)}{2r}\rangle = \frac{1}{4r^2}4q(w) = \varepsilon,$$

we have the following result:

Proposition 8.3 *Let r be a positive number. Then if $\varepsilon = +1$, the central hyperquadric $S_1^n = q^{-1}(\varepsilon r^2) = q^{-1}(r^2)$ is a Lorentz manifold (the Lorentz sphere). If $\varepsilon = -1$, $q^{-1}(\varepsilon r^2) = q^{-1}(-r^2)$ is a riemannian manifold.*

Proposition 8.4 *Let α be a nonconstant geodesic of the Lorentz sphere S_1^n. Then:*

(i) *If α is time like, α is a parametrization of one branch of a hyperbola in $\mathbf{R}_1^{n+1}$.*

(ii) *If α is light like, α is a straight line, that is, a geodesic of* $\mathbf{R}_1^{n+1}$.

(iii) *If α is space like, α is a periodic parametrization of an ellipse in* $\mathbf{R}_1^{n+1}$.

Proof Let $p \in S_1^n$, that is, $q(p) = r^2$, so p is space like. Consider a 2-plane $\pi \subset \mathbf{R}_1^{n+1}$ through the origin of $\mathbf{R}^{n+1}$ and p. Now one considers the restriction of g, the metric of $\mathbf{R}_1^{n+1}$, to the plane π. We have three possibilities: (a) $g|\pi$ is nondegenerate with index 1. Let (e_1, e_2) be an orthonormal basis of π with respect to $g|\pi$ such that $e_2 = \frac{p}{r}$, so e_1 is necessarily time like. A generic point $(ae_1 + be_2) \in \pi \cap S_1^n$ satisfies $r^2 = -a^2 + b^2$. This implies that $\pi \cap S_1^n$ is a hyperbola in π and the branch through p can be parametrized by

$$\begin{aligned} \alpha(t) &= rSh(t)e_1 + rCh(t)e_2, \quad t \in \mathbf{R}; \\ so, \quad \dot{\alpha}(t) &= rCh(t)e_1 + rSh(t)e_2, \end{aligned}$$

and then $\langle \dot{\alpha}, \dot{\alpha} \rangle = -r^2Ch^2(t) + r^2Sh^2(t) = -r^2$, that means, α is time like.

On the other hand $\ddot{\alpha}(t) = \alpha(t)$ and from (8.9) we get

$$\ddot{\alpha}(t) = \frac{1}{2}(grad \ \ q)(\alpha(t)),$$

so, $\ddot{\alpha}(t)$ is orthogonal to S_1^n at the point $\alpha(t)$. Then $\alpha(t)$ is a time like geodesic (see Exercise 2.5) that proves (i). The second possibility is: (b) $g|\pi$ is positive definite. In this case we take an orthonormal basis (e_2, e_3) for π then a point $ae_2 + be_3$ on π belongs to $S_1^n = \{v \in \mathbf{R}_1^{n+1} | \langle v, v \rangle = r^2\}$ if, and only if, $a^2 + b^2 = r^2$. Thus, the parametrization $\alpha(t) = r(\cos t)e_2 + r(\sin t)e_3$ satisfies $\langle \alpha(t), \alpha(t) \rangle = r^2$ and α is space like. But $\ddot{\alpha}(t) = -\alpha(t) = -\frac{1}{2}grad \ \ \alpha(t)$, so $\alpha(t)$ is a space like geodesic of S_1^n, that proves (iii). The third and last possibility is: (c) $g|\pi$ is degenerate with a null space of dimension 1. If $v \neq 0$ is a null vector, the pair (p, v) is an orthogonal basis for π and $ap + bv \in \pi \cap S_1^n$ if, and only if, $q(ap + bv) = r^2$ or $\langle ap + bv, ap + bv \rangle = a^2r^2 = r^2$ that gives $a = \pm 1$. The set $\pi \cap S_1^n$ is the union of two parallel straight lines, one of them containing p, parametrized by

$$\alpha(t) = p + tv$$

and such that $\dot{\alpha}(t) = v$; so α is light like and since $\ddot{\alpha}(t) = 0$, α is a light like geodesic of S_1^n that proves (ii). Finally, any other geodesic of S_1^n passing through $p \in S_1^n$ is in one of three classes considered above. In fact, if $\beta = \beta(t)$ is such that $\beta(0) = p$ and $\dot{\beta}(0)$ is its tangent vector at p, we construct the 2-plane π passing through the origin $0 \in \mathbf{R}^{n+1}$ and p, and also containing the vector $\dot{\beta}(0)$; by uniqueness, $\beta(t)$ is in one of the classes above. ∎

The set $q^{-1}(0)$ is the union of the null cone $\mathcal{N} = q^{-1}(0) - \{0\}$ with the origin $\{0\}$. In coordinates we have that

$$q^{-1}(0) = \{u \in \mathbf{R}^{n+1} | u_1^2 = u_2^2 + \ldots + u_{n+1}^2\}.$$

Remark The null cone $\mathcal{N}$ has two connected components and is a submanifold of codimension one of $\mathbf{R}^{n+1}$, because $0 \in \mathbf{R}$ is a regular value of q restricted to $\mathbf{R}^{n+1} - \{0\}$. $\mathcal{N}$ is invariant under multiplication by a real number $\lambda \neq 0$; moreover, it is diffeomorphic to $(\mathbf{R} - \{0\}) \times S^{n-1}$ (where S^{n-1} is a $(n-1)$-dimensional sphere) and is not a pseudo-riemannian manifold (in fact, any $u \in \mathcal{N}$ is, at the same time, tangent and orthogonal to $\mathcal{N}$, so the restriction of $\langle , \rangle$ to $\mathcal{N}$ is degenerate).

8.3 Time-cones and time-orientability of a Lorentz manifold

We will introduce, in the sequel, the notion of **time-orientability** of a Lorentz manifold $(Q, \langle , \rangle)$ of dimension $n \geq 2$. Fix a point $p \in Q$ and consider a subspace $W \subset T_pQ$. As in the case of vectors, there are three categories of subspaces:

Definition 8.3

(i) W **is space like** if $\langle , \rangle|_W$ is positive definite;

(ii) W **is time like** if $\langle , \rangle|_W$ is non degenerate of index 1;

(iii) W **is light like** if $\langle , \rangle|_W$ is degenerate.

Observe that the category of a vector $v \in T_pQ$ is the category of the subspace $\mathbf{R}v$, spanned by v.

Let $W^\perp$ denote the linear subspace of all vectors v in T_pQ such that $\langle v, u \rangle_p = 0$ for all $u \in W$. It is easy to show that $\dim W^\perp = n - \dim W$ and that $W = (W^\perp)^\perp$. The standard identity

$$\dim W + \dim W^\perp = \dim(W \cap W^\perp) + \dim(W + W^\perp)$$

implies that $W \cap W^\perp = \{0\}$ if, and only if, $W + W^\perp = T_pQ$ and that W is non degenerate if, and only if, $W \cap W^\perp = \{0\}$. As a counter example, take in $\mathbf{R}^2_1$ the subspace W spanned by the vector $v = (1, 1)$. Since $\langle v, v \rangle = 0$ we have $W \cap W^\perp \neq \{0\}$ and then $W + W^\perp \neq T_pQ$.

Proposition 8.5 *If $z \in T_pQ$ is a time like vector ($\langle z, z \rangle < 0$), then $z^\perp = \{u \in T_pQ | \langle u, z \rangle = 0\}$ is a $(n-1)$-dimensional space like subspace such that $T_pQ = \mathbf{R}z \oplus z^\perp$.*

Proof Since z is a time like vector, the subspace $\mathbf{R}z$ is a nondegenerate. Then we only need to check that $z^\perp$ is space like. But this follows because the index of $(Q, \langle , \rangle)$ is equal to 1. ∎

Corollary 8.1 *A subspace $W \subset T_pQ$ is time like if, and only if, $W^\perp$ is space like. Since $W = (W^\perp)^\perp$ then W is space like if, and only if, $W^\perp$ is time like.*

Let us denote by τ the set of all time like vectors of T_pQ, that is $u \in \tau$ means that $\langle u, u\rangle < 0$. For a given $u \in \tau$, the set

$$C(u) = \{v \in \tau | \langle u, v\rangle < 0\} \tag{8.11}$$

is called the **time cone** of T_pQ containing u. It is clear that $v \in C(u)$ implies $\lambda v \in C(u)$ for all $\lambda > 0$; also $C(-u) = -C(u)$ is the **opposite cone** to $C(u)$.

Proposition 8.6 *τ is the (disjoint) union of $C(u)$ and $C(-u)$.*

Proof In fact $v \in \tau$ implies either $\langle u, v\rangle < 0$ $(v \in C(u))$ or $\langle u, v\rangle > 0$ $(v \in C(-u))$, because $\langle u, v\rangle = 0$ means $v \in u^\perp$ and $u^\perp$ is space like by Proposition 8.5, that is, $\langle v, v\rangle > 0$ (contradiction). Then $\tau \subset C(u) \cup C(-u)$. Conversely, $v \in C(u) \cup C(-u)$ means $v \in \tau$ that follows from (8.11). ∎

Proposition 8.7 *Two time like vectors v, w belong to the same time cone if, and only if, $\langle v, w\rangle < 0$.*

Proof Use Proposition 8.5 and write for $u \in \tau$:

$$\begin{aligned} v &= au + \bar{v}, \quad \bar{v} \in u^\perp \\ w &= bu + \bar{w}, \quad \bar{w} \in u^\perp. \end{aligned}$$

The time cone being $C(u) = C(\frac{u}{|u|})$, one assumes, for simplicity, that $|u| = 1$. But v and w are time like, and so $\langle v, v\rangle = (a^2\langle u, u\rangle + \langle \bar{v}, \bar{v}\rangle) < 0$, or $|v|^2 = -a^2 + |\bar{v}|^2$, because $\bar{v}$ is space like and $u \in \tau$. Then $|a| > |\bar{v}|$; analogously, $|b| > |\bar{w}|$. Since $\langle v, w\rangle = -ab + \langle \bar{v}, \bar{w}\rangle$ and $\bar{v}^\perp, \bar{w}^\perp$ are space like (so, for then one can apply Schwarz inequality), we have $|\langle \bar{v}, \bar{w}\rangle| \le |\bar{v}||\bar{w}| < |ab|$. Assume now, by hypothesis, that v and w are in $C(u)$; then $\langle v, u\rangle$ and $\langle w, u\rangle$ are strictly negative numbers and that implies $a > 0$ and $b > 0$ and by consequence $\langle v, w\rangle < 0$. Conversely, if $\langle v, w\rangle < 0$, the condition $(-ab + \langle \bar{v}, \bar{w}\rangle) < 0$ implies $ab > 0$. So if $a > 0$ (then $b > 0$) we have:

$$\begin{aligned} \langle v, u\rangle &= -a < 0 \quad \text{so} \quad v \in C(u), \\ \langle w, u\rangle &= -b < 0 \quad \text{so} \quad w \in C(u); \end{aligned}$$

the case $a < 0$ (and then $b < 0$) gives, analogously: $\langle v, u\rangle = |a| > 0$ and

$\langle w, u\rangle = |b| > 0$ that means w and v belong to $C(-u)$. ∎

Corollary 8.2 *If u, v are time like vectors then*

$$u \in C(v) \Longleftrightarrow v \in C(u) \Longleftrightarrow C(u) = C(v).$$

Moreover, time cones are convex sets.

Proof We only prove convexity; if v, w are in $C(u)$ and $a \geq 0$, $b \geq 0$ with $a^2 + b^2 > 0$, then $\langle (av + bw), u\rangle < 0$ or $av + bw \in C(u)$. ∎

Proposition 8.8 *Let $v, w \in \tau$. Then:*
(i) $|\langle v, w\rangle| \geq |v|.|w|$, *with equality if, and only if, v and w are linearly dependent* **(backwards Schwarz inequality)**.

(ii) *If v, w belong to the same cone of T_pQ, there is a unique $\varphi \geq 0$ (the hyperbolic angle between v and w) such that $\langle v, w\rangle = -|v||w|Ch\varphi$.*

Proof (i) By Proposition 8.5 we have $w = av + \bar{w}, \bar{w} \in v^{\perp}$; and, since w is time like we have $\langle w, w\rangle = (a^2\langle v, v\rangle + \langle \bar{w}, \bar{w}\rangle) < 0$. Then

$$\begin{aligned} \langle v, w\rangle^2 &= a^2\langle v, v\rangle^2 = (\langle w, w\rangle - \langle \bar{w}, \bar{w}\rangle).\langle v, v\rangle \\ &\geq \langle w, w\rangle.\langle v, v\rangle = |w|^2.|v|^2 \end{aligned}$$

(because $\langle \bar{w}, \bar{w}\rangle > 0$ and $\langle v, v\rangle < 0$). The equality holds if, and only if $\langle \bar{w}, \bar{w}\rangle = 0$ (or $\bar{w} = 0$), that means $w = av$.

(ii) By Proposition 8.7 we have $\langle v, w\rangle < 0$, hence $-\langle v, w\rangle/|v|.|w| \geq 1$ and so, by the definition and elementary properties of the hyperbolic cosine one has the result. ∎

Corollary 8.3 (backwards triangle inequality)
If v and $w \in \tau$ and are in the same time cone, then $|v| + |w| \leq |v + w|$, with equality if, and only if, v and w are linearly dependent.

Proof Since $\langle v, w\rangle < 0$ (Proposition 8.7), backwards Schwarz inequality gives $|v||w| \leq -\langle v, w\rangle$ then

$$(|v| + |w|)^2 = |v|^2 + |w|^2 + 2|v|.|w| \leq -\langle v + w, v + w\rangle = |v + w|^2.$$

The equality comes if, and only if, $|v|.|w| = -\langle v, w\rangle = |\langle v, w\rangle|$; then, Proposition 8.7 gives the result. ∎

Remarks 1. It is against our euclidean intuition that a straight line segment is no longer the shortest route between two points. As we will see, this result (see Corollary 8.3) is fundamental in some applications to relativity theory.

2. In each tangent space T_pQ of a Lorentz manifold $(Q, \langle , \rangle)$ there are two time cones (see Corollary 8.2) and there is no intrinsic way to distinguish them. When we choose one we are time orienting T_pQ.

The **time orientability** of a Lorentz manifold is related with the choice of a time cone in each tangent space T_pQ, in a continuous way. So, let C be a function on Q that, to each $p \in Q$ assigns a time cone C_p in T_pQ; we say that C is **smooth** if for each $p \in Q$ there corresponds a smooth (local) vector field V defined in a neighborhood $\mathcal{U}$ of p such that $V_q \in C_q$ for each $q \in \mathcal{U}$. Such a smooth function C is said to be a **time orientation** of Q. If $(Q, \langle , \rangle)$ admits a time orientation we say that $(Q, \langle , \rangle)$ is **time orientable** and if we choose a specific time orientation we use to say that $(Q, \langle , \rangle)$ is **time oriented**. The Minkowski space $\mathbf{R}_1^{n+1}$ is time orientable; the usual time orientation is the one of a cone containing $\frac{\partial}{\partial u_1}$ corresponding to the natural coordinates $(u_1, u_2, \ldots, u_{n+1})$.

Proposition 8.9 *A Lorentz manifold $(Q, \langle , \rangle)$ is time orientable if, and only if, there exists a time like vector field $X \in \mathcal{X}(Q)$.*

Proof If $X \in \mathcal{X}(Q)$ satisfies $X_p \in \tau \subset T_pQ$, one defines the map C by $C_q = C(X_q)$, for all $q \in Q$. Conversely, let C be a time orientation of $(Q, \langle , \rangle)$. Since C is smooth we have a covering of Q by neighborhoods $\mathcal{U}$ and in each one of which there exists a vector field $X_{\mathcal{U}}$ and $C_q = C(u)$ where u is the value of $X_{\mathcal{U}}$ at q, for all $q \in \mathcal{U}$. Now let $\{f_\alpha | \alpha \in \mathcal{A}\}$ be a differentiable partition of unity subordinate to the covering of Q by the neighborhoods $\mathcal{U}$ (see Proposition 4.2). Thus, the support of each f_α, is contained in some element $\mathcal{U}(\alpha)$ of that covering. The functions f_α are non negative and time cones are convex sets. Thus $X = \Sigma f_\alpha X_{\mathcal{U}(\alpha)}$ is time like. ∎.

Exercise 8.2 The Lorentz sphere $S_1^n = q^{-1}(r^2)$ introduced in Proposition 8.3 is time orientable. **Hint**: use the projection to S_1^n of $\frac{\partial}{\partial u_1}$.

8.4 Lorentz geometry notions in special relativity

Let $(Q, \langle , \rangle)$ be a Lorentz manifold and $p \in Q$.

Definition 8.4 An element $v \in T_pQ$ is said to be a **causal vector** if it is not space like (so, either null or time like). For a time like vector $(u \in \tau)$, the set $\bar{C}(u)$ of all causal vectors v such that $\langle u, v \rangle < 0$ is the **causal cone** in T_pQ containing u. A **causal curve** $t \mapsto \alpha(t)$ in Q is a smooth curve such that $\dot{\alpha}(t)$

is a causal vector, for all t.

Exercise 8.3 Show that for vectors in T_pQ of $(Q, \langle , \rangle)$:

(a) Causal vectors v, w are in the same causal cone if and only if either $\langle v, w\rangle < 0$ or v and w are null such that $w = av$, $a > 0$.

(b) If $u \in \tau$, $\bar{C}(u) =$ closure of $(C(u) - \{0\})$.

(c) Causal cones are convex.

(d) The components of the set of all causal vectors in T_pQ are the two causal cones in T_pQ.

Definition 8.5 A **space time** is a connected time-orientable four-dimensional Lorentz manifold $(Q, \langle , \rangle)$. A **Minkowski space time** Q is a space time that is isometric to the Minkowski 4-space $\mathbf{R}^4_1$.

If the space time $(Q, \langle , \rangle)$ is time oriented, the time orientation is called the **future** and its negative is the **past**. A tangent vector $v \in T_pQ$ in a future causal cone is said to be **future pointing**. A causal curve is future pointing if all its velocity vectors are future pointing.

Definition 8.6 Any isometry that takes a time oriented space time $(Q, \langle , \rangle)$ onto the Minkowski 4-space $\mathbf{R}^4_1$ and preserves time orientation is called an **inertial coordinate system** of $(Q, \langle , \rangle)$.

Proposition 8.10 *Given a basis (e_0, e_1, e_2, e_3) in a tangent space T_pQ of a time oriented space time $(Q, \langle , \rangle)$ such that e_0 is future pointing, then there is a unique inertial coordinate system ξ of $(Q, \langle , \rangle)$ such that $\frac{\partial}{\partial x_i}(p) = e_i, i = 0, 1, 2, 3$.*

Proof The existence of the isometry $\xi : Q \longrightarrow \mathbf{R}^4_1$ is obtained from a normal coordinate system (see Exercise 6.4). The uniqueness of such an isometry follows from the fact that two local isometries of a connected pseudo-riemannian manifold whose differentials coincide at a single point are necessarily equal. ∎

As we did in the case of Newtonian mechanics (see Chapter 1) we keep ourselves, here, calling **events** the points of the space time Q and **particles** will correspond to parametrized curves. We do not have a canonical time function as in the case of a galilean space-time structure but we follow assuming the existence of inertial coordinate systems. The particles are defined as follows:

Definition 8.7 A **light like particle** is a future null geodesic of a time oriented space time $(Q, \langle , \rangle)$. A **material particle** (also called an **observer**) is a time like future pointing smooth curve $\alpha : s \in I \longmapsto \alpha(s) \in Q$ such that $|\alpha'(s)| = 1$ for all $s \in I$; its image $\alpha(I)$ is the **world line** of α and the parameter s is called the **proper time** of the material particle. A material particle which is a geodesic is said to be **freely falling**.

Remarks 1. The world line of a material particle is a one-dimensional submanifold of Q.
2. We can think that each material particle has a "clock" in order to measure its proper time.
3. The fact that light moves geodesically is a fundamental hypothesis in Relativity; since in this case $\langle\dot\gamma,\dot\gamma\rangle = 0$ (see Definition 8.7 above), the parametrization by proper time is impossible. One says that "it cannot carry a clock".

8.5 Minkowski space time geometry

From Definition 8.5 there is an isometry ξ between a given Minkowski space time $(Q, \langle,\rangle)$ and $\mathbf{R}^4_1$; it is usual to denote a Minkowski space time by $(Q, \langle,\rangle, \xi)$. At this point it is clear that given two points p, q in Q, there is a unique geodesic α such that $\alpha(0) = p$ and $\alpha(1) = q$. Also there is a natural linear isometry identifying T_pQ and T_qQ called the **distant parallelism** and the exponential map $\exp_p : T_pQ \longrightarrow Q$ is an isometry. In fact between the points $\xi(p)$ and $\xi(q)$ of $\mathbf{R}^4_1$ there is a unique (straight line) geodesic going from $\xi(p)$ to $\xi(q)$ and a translation of the affine space $\mathbf{R}^4_1$ taking $\xi(p)$ into $\xi(q)$. The manifold $(Q, \langle,\rangle, \xi)$ is viewed from p in the same geometric way as T_pQ is viewed from zero. Also Q is a normal neighborhood of each of its points. The vector $\dot\alpha(0)$ is called the **displacement vector** , it satisfies $\exp_p \dot\alpha(0) = q$ and it is denoted by $\vec{pq}$. One can move the notion of causality from the tangent spaces of M to M itself. For an event $p \in Q$, the **future time cone of p** is the set

$$\{q \in Q \mid \vec{pq} \in T_pQ \ \text{ is time like and future pointing}\}.$$

The **future light cone of p** is the set

$$\{q \in Q \mid \vec{pq} \in T_pQ \ \text{ is null and future pointing}\}.$$

The union of these two sets is the **future causal cone** of p. Past analogues are defined similarly. Of course all these notions depend on the isometry ξ. From now on one assumes that ξ is an inertial coordinate system of $(Q, \langle,\rangle)$, that is, ξ preserves time orientation.

In order to give a clear understanding of the term "causal" used in definition 8.4 it is usual and natural to say that an event p can influence an event q if, and only if, there exists a particle from p to q (see definition 8.7). It can be proved the following:

Exercise 8.4 The only events that can be influenced by event p are those in its future causal cone. The only events that can influence an event p are those in its past causal cone.

Definition 8.8 Given two points p, q in a Minkowski space time $(Q, \langle,\rangle, \xi)$, the square root of the absolute value of $\langle \vec{pq}, \vec{pq}\rangle$ is called the **separation** between p and q, and is denoted by pq, that is

$$pq = |\langle \vec{pq}, \vec{pq}\rangle|^{1/2}.$$

Then if $\vec{pq}$ is time like future pointing, pq represents the *time from the event p to event q* computed as the proper time of the unique freely falling material particle from p to q. It is also clear that if $pq = 0$ the displacement vector $\vec{pq}$ is light like and there is a light like particle going from p to q.

If three events p, q, o belong to a Minkowski space time $(Q, \langle,\rangle, \xi)$ and p, q are in the same time cone of o, then the hyperbolic angle $\varphi = p\hat{o}q$ is, by definition, the hyperbolic angle between the time like tangent vectors $\vec{op}$ and $\vec{oq}$ (see Proposition 8.8 - (ii)).

Proposition 8.11 *Let $p, q \in Q$ in the same time cone of $o \in Q$. Then if $\vec{op}$ is orthogonal to $\vec{pq}$ we have:*

(i) $(oq)^2 = (op)^2 - (pq)^2$

(ii) $(op) = (oq)\ Ch\,\varphi$ and $(pq) = (oq)\ Sh\,\varphi$.

Proof From Proposition 8.5 we see that $\vec{pq}$ is space like. Now, moving $\vec{pq}$ by distant parallelism to o we can write $\vec{oq} = \vec{op} + \vec{pq}$ and then scalar products yield

$$\begin{aligned}\langle \vec{oq}, \vec{oq}\rangle &= \langle \vec{op}, \vec{op}\rangle + \langle \vec{pq}, \vec{pq}\rangle + 2\langle \vec{op}, \vec{pq}\rangle = \\ &= \langle \vec{op}, \vec{op}\rangle + \langle \vec{pq}, \vec{pq}\rangle\end{aligned}$$

that is $-(oq)^2 = -(op)^2 + (pq)^2$ that proves (i). Condition (ii) follows from the fact that $\vec{op}$ and $\vec{oq}$ are time like, that is, from Proposition 8.8 we have

$$\langle \vec{op}, \vec{oq}\rangle = -(op)(oq)\ Ch\,\varphi = \langle \vec{op}, \vec{op} + \vec{pq}\rangle = -(op)^2, \quad \text{with } \varphi \geq 0,$$

then $(op) = (oq)\ Ch\,\varphi$ and $(pq)^2 = -(oq)^2 + (op)^2 = (oq)^2[Ch^2\varphi - 1] = (oq)^2 Sh^2\varphi$; but

$$Sh\,\varphi = \frac{e^{\varphi} - e^{-\varphi}}{2} = \frac{e^{\varphi}}{2}(1 - e^{-2\varphi}) \geq 0, \text{ so, } (pq) = (oq)\ Sh\,\varphi \quad \blacksquare$$

In a Minkowski space time $(Q, \langle,\rangle, \xi)$ the (time) x^o axis of ξ through $p \in Q$ is the world line of a freely falling observer ω; the natural parametrization of ω

has $t = x^o(\omega(t))$ and t is the proper time of ω. We have to keep in mind that ω depends on ξ.

To $p \in Q$ there corresponds $\xi(p)$ given by $\xi(p) = (x^o(p), x^1(p), x^2(p), x^3(p)) \in \mathbf{R}^4_1$. The first component $x^o(p)$ is said to be the **ξ-time** of p and $\vec{p} = (x^1(p), x^2(p), x^3(p)) \in \mathbf{R}^3$ is the **ξ-position** of p.

Now if $\alpha : I \longrightarrow Q$ is a particle of a space time $(Q, \langle , \rangle)$ and $s \in I$, the ξ-time of $\alpha(s)$ is $t = x^o(\alpha(s))$ and its ξ-position is $(x^1(\alpha(s)), x^2(\alpha(s)), x^3(\alpha(s)))$. Since α is time like and future pointing (see Definition 8.7) then

$$\frac{dt}{ds} = \frac{d(x^o \circ \alpha)}{ds} = -\langle \alpha', \frac{\partial}{\partial x^o} \rangle \neq 0,$$

so, $(x^o \circ \alpha)$ is a diffeomorphism of I onto some interval $J \subset \mathbf{R}$ with inverse $u : J \to I$. At a ξ-time $t \in J$, the ξ-position of α is

$$\vec{\alpha}(t) = (x^1 \alpha u(t), x^2 \alpha u(t), x^3 \alpha u(t)).$$

The curve $\vec{\alpha}(t)$ is the ξ-associated Newtonian particle of α and one uses to say that $\vec{\alpha}$ is what the observer ω observes of α.

One main point in special relativity is to relate the Newtonian concepts applied to $\vec{\alpha}$ with the relativistic analogues for α.

If the particle $\alpha : I \to Q$ is light like in $(Q, \langle , \rangle)$ and ξ is an inertial coordinate system, the associated Newtonian particle $\vec{\alpha}$ of α is a straight line in $\mathbf{R}^3$ with speed **c** =1. In fact, α is a future null geodesic in $(Q, \langle , \rangle)$ so $\xi \circ \alpha$ is a geodesic in $\mathbf{R}^4_1$. Thus

$$x^i \alpha(s) = a_i s + b_i \qquad i = 0, 1, 2, 3.$$

Then $\vec{\alpha}(s) = (x^1\alpha(s), x^2\alpha(s), x^3\alpha(s))$ is a straight line in $\mathbf{R}^3$ and its reparametrization $\vec{\alpha}(t)$ follows this straight line and the vector $\frac{d\alpha}{ds}$ is null with $\frac{dt}{ds} > 0$. It follows that the speed v of $\vec{\alpha}$ is

$$v = \left|\frac{d\vec{\alpha}}{dt}\right| = \left(\frac{d\vec{\alpha}}{dt}\right) \cdot \left(\frac{dt}{ds}\right)^{-1} = 1.$$

Proposition 8.12 *Light has the same constant speed v =***c***= 1 relative to every inertial coordinate system ξ and then relative to every freely falling observer.*

Proposition 8.13 *If the particle $\alpha : I \longrightarrow Q$ is material, we have that*

(i) *the speed $\left|\frac{d\vec{\alpha}}{dt}\right|$ of the ξ-associated Newtonian particle $\vec{\alpha}$ is $v = \left|\frac{d\vec{\alpha}}{dt}\right| =$ Tanh φ where φ is the hyperbolic angle between $\alpha' = \frac{d\alpha}{ds}$ and the time coordinate vector $\frac{\partial}{\partial x^o}$ of ξ, which implies, in particular, that $0 \leq v < 1$.*

(ii) *The proper time s of α and its ξ-time t are related by*

$$\frac{dt}{ds} = \frac{d(x^o \circ \alpha)}{ds} = Ch\ \ \varphi = \frac{1}{\sqrt{1 - v^2}} \geq 1.$$

Proof In fact, $\alpha' = \frac{d\alpha}{ds}$ and the time coordinate $\frac{\partial}{\partial x^o}$ of ξ are time like and future pointing, so there is a unique hyperbolic angle $\varphi \geq 0$ determined by $-\langle \alpha', \frac{\partial}{\partial x^o} \rangle = Ch\ \ \varphi \geq 1$. Since $\alpha' = \sum_{i=0}^{3} \frac{d(x^i \circ \alpha)}{ds} \frac{\partial}{\partial x^i}$ we have

$$\frac{dt}{ds} = -\langle \alpha', \frac{\partial}{\partial x^o} \rangle = Ch\ \ \varphi$$

and $\langle \alpha', \alpha' \rangle = -1$ gives

$$-\left(\frac{dt}{ds}\right)^2 + \left|\frac{d\vec{\alpha}}{ds}\right|^2 = -1.$$

Since $\varphi \geq 0$ it follows

$$\left|\frac{d\vec{\alpha}}{ds}\right| = \sqrt{Ch^2\varphi - 1} = Sh\ \ \varphi \geq 0;$$

thus $\vec{\alpha}(t)$ has speed

$$v = \left|\frac{d\vec{\alpha}}{dt}\right| = \left(\frac{d\vec{\alpha}}{ds}\right)\left(\frac{dt}{ds}\right)^{-1} = \frac{Sh\ \ \varphi}{Ch\ \ \varphi} = Tanh\ \ \varphi.$$

Finally one obtains that $Ch\ \ \varphi = \frac{1}{\sqrt{1-v^2}}$. ∎

The so called **time dilation effect of Larmor and Lorentz** is interpreted through Proposition 8.13-(ii) for a particle with proper time (s); the faster the particle is moving relative to the observer, that is, the larger v is, the slower the particle's clock (s) runs relative to the observer clock (t).

We saw that to an inertial coordinate system ξ of a Minkowski space time $(Q, \langle , \rangle, \xi)$ there corresponds a freely falling observer ω. But, conversely, given a freely falling observer $\omega = \omega(t)$ of a time oriented space time $(Q, \langle , \rangle)$ such that $\omega(0) = p \in Q$, one can talk about the space like (euclidean) tridimensional subspace $E_o = (\dot{\omega}(0))^{\perp}$ and define an isometry ξ provided that we choose an orthonormal basis of E_o (see Proposition 8.10). The subspace E_o is the same for all choices of ξ and the image of E_o on Q under the exponential map $exp_p : T_pQ \longrightarrow Q$ is called the **rest space** of ω at p; it is the set of events in Q that the observer ω considers simultaneous with p. One can argue, analogously, with the space like subspace $E_t = (\dot{\omega}(t))^{\perp} \subset T_{\omega(t)}Q$ and talk about the rest space of ω at $\omega(t)$ which is formed by the events in Q that ω considers simultaneous with $\omega(t)$. The euclidean rest spaces E_o and E_t are canonically isometric.

The **relativistic addition of velocities** is another effect that holds in a Minkowski is space time $(Q, \langle , \rangle, \xi)$ when one considers two material particles on $Q : \alpha = \alpha(\tau)$ and $\beta = \beta(\sigma)$. We can define the hyperbolic angle φ between $\alpha'(\tau)$ and $\beta'(\sigma)$ if we make use of the distant parallelism and also define $v = Tanh\ \ \varphi$ as the corresponding **instantaneous relative speed**. Assume that a rocketship ρ leaves a space station α and also that both are freely falling particles. Let

$v_1 > 0$ be their instantaneous relative speed. A space-man μ is ejected from ρ in the plane of ρ and α with constant speed v_2 relative to ρ. Let us compute the speed v of the space-man relative to α. Let $v_1 = Tanh\ \varphi_1$ and $v_2 = Tanh\ \varphi_2$. One can argue on $\mathbf{R}_1^4$ using the isometry ξ. So, if $v_2 > 0$, by distant parallelism the tangent vector ρ is between the vectors α' and μ' and the angle φ defined by α' and μ' is given by $\varphi = \varphi_1 + \varphi_2$, so

$$v = Tanh\ \varphi = Tanh\ (\varphi_1 + \varphi_2) = \frac{Tanh\ \varphi_1 + Tanh\ \varphi_2}{1 + Tanh\ \varphi_1 . Tanh\ \varphi_2} = \frac{v_1 + v_2}{1 + v_1 v_2}.$$

Exercise 8.5 Prove that the same formula holds if $v_2 < 0$.

The next classical example is the so called **twin paradox**, described as follows: "On their 21st birthday Peter leaves his twin Paul behind on their freely falling spaceship and departs at the event q with constant relative speed $v = 24/25$ for a free fall of seven years of his proper time. Then he turns and comes back symmetrically in another seven years. Upon his arrival at the event o he is thus 35 years old, but Paul is 71". We have to drop a perpendicular px from the turn p to the world line of the spaceship. By Propositions 8.11 and 8.13 we have

$$ox = op\ Ch\ \varphi = \frac{7}{[1 - (24/25)^2]^{1/2}} = 25.$$

If the separations ox and xq are equal (symmetry) then $xq = 25$. Thus Paul's age at Peter's return is $21 + 2(25) = 71$ years.

Definition 8.9 The **energy-momentum vector field of a material particle** $\alpha : I \longrightarrow Q$ of mass m is the vector field $P = m\frac{d\alpha}{ds}$ on α (s is the proper time of α).

For an associated Lorentz coordinate ξ corresponding to a freely falling observer ω, the components of P are

$$P^i = m\frac{d(x^i \circ \alpha)}{ds}, \quad i = 0, 1, 2, 3.$$

If t is the proper time of the observer, we have

$$P^i = m\frac{d(x^i \circ \alpha)}{dt} . \frac{dt}{ds}$$

where $\frac{dt}{ds} = \frac{1}{\sqrt{1-v^2}}$ and v is the speed of the ξ-associated Newtonian particle.

The space components P^1, P^2, P^3 define a vector field

$$\vec{P} = \frac{m}{\sqrt{1 - v^2}} \frac{d\vec{\alpha}}{dt}$$

on the associated Newtonian particle $\vec{\alpha}$ in $E_0 = \mathbf{R}^3$.

The time component P^0 is given by

$$P^0 = m\frac{d(x^0 \circ \alpha)}{ds} = m\frac{dt}{ds} = \frac{m}{\sqrt{1-v^2}} = m + \frac{1}{2}mv^2 + O(v^4).$$

Einstein identified P^0 as the total energy E of the particle as measured by ω, concluding, in particular, that mass is merely one form of energy, the **rest energy** E_{rest}. Converting to conventional units we have the famous formula

$$E_{rest} = mc^2,$$

where c is the speed of light.

Definition 8.10 The **energy momentum vector field of a light like particle** $\gamma : I \longrightarrow Q$ is its 4-velocity $P = \gamma' = \frac{d\gamma}{ds}$.

Then any freely falling observer ω splits P into **energy** E and **momentum** P (both relative to ω) by setting $P = E\frac{\partial}{\partial x^0} + \vec{P}$ with $\vec{P}$ orthogonal to $\frac{\partial}{\partial x^0}$, just as in the case of material particles; in this case $E = |\vec{P}| = -\langle\gamma', \frac{\partial}{\partial x^0}\rangle$. But γ and ω are both geodesics then $E = |\vec{P}|$ is constant and $\vec{P}$ is parallel. For a material particle we have that $E^2 = m^2 + |\vec{P}|^2$, so one concludes that, by analogy, a light like particle does not have mass.

The wave character of light follows from the next observation; for instance, a photon of energy E, relative to some observer, has **frequency** $\nu = \frac{E}{h}$ where h is the constant of Planck. Usually one says that frequency times **wave length** λ is speed c. In geometric units $\lambda\nu = 1$. Since frequency and wave length derive from energy, they too depend on the observer. Thus, "visible light" for one observer is "radio waves" for another and "x rays" for a third observer.

8.6 Lorentz and Poincar groups

The set of all linear isometries of $\mathbf{R}^n_1$ is called the **Lorentz group**; it is a subgroup of the group of all isometries of $\mathbf{R}^n_1$. The translation $T_x : v \in \mathbf{R}^n_1 \to v + x \in \mathbf{R}^n_1$, defined by an element $x \in \mathbf{R}^n_1$, is also an isometry of $\mathbf{R}^n_1$; in fact the set of all translations of $\mathbf{R}^n_1$ is an abelian subgroup of the group of all isometries. The group of all isometries of $\mathbf{R}^n_1$ is called the **Poincar group**.

Proposition 8.14 *Each isometry Φ of $\mathbf{R}^n_1 (n \geq 2)$ has a unique decomposition $\Phi = T_x \circ \theta$ where T_x is the translation defined by an element $x \in \mathbf{R}^n_1$ and θ is a linear (homogeneous) isometry of $\mathbf{R}^n_1$. Furthermore $T_x\theta_1 T_y\theta_2 = T_{x+\theta_1 y}\theta_1\theta_2$. In particular the group of all isometries of $\mathbf{R}^n_1$ is a subgroup of the group of all affine transformations of $\mathbf{R}^4_1$.*

Proof Start with Φ such that $\Phi(0) = 0$. Let us show that, necessarily, Φ is linear (and homogeneous). In fact $d\Phi(0)$ is a linear isometry of $T_o(\mathbf{R}_1^n)$ and, so, of $\mathbf{R}_1^n$, because $T_o(\mathbf{R}_1^n)$ is canonically linearly isometric to $\mathbf{R}_1^n$. Let θ be the linear isometry of $\mathbf{R}_1^n$ corresponding to $d\Phi(0)$; since $d\bar{\theta}(0) = d\Phi(0)$ we have $\bar{\theta} = \Phi$ (see the proof of Proposition 8.10). If $\Phi(0) = x \neq 0$ we have $(T_{-x}\Phi)(0) = 0$ and by the same argument, $T_{-x}\Phi$ is equal to some linear isometry θ of $\mathbf{R}_1^n$ then $\Phi = T_x\theta$ and the decomposition follows. The uniqueness of the decomposition is trivial because if $T_x\theta = T_y\tilde{\theta}$ then $x = (T_x\theta)(0) = (T_y\tilde{\theta})(0) = y$ and also $\tilde{\theta} = \theta$. Finally, for all $v \in \mathbf{R}_1^n$ we have $(\theta T_y)(v) = \theta(y+v) = \theta(y) + \theta(v) = (T_{\theta y}\theta)(v)$. Hence $\theta T_y = T_{\theta y}\theta$ that makes true the multiplication rule. ∎

The last result shows that given a Minkowski space time $(Q, \langle,\rangle, \xi)$, all the possible inertial coordinate systems are obtained by making the composition of ξ with all the elements of the Poincar group of $\mathbf{R}_1^4$. As one can see, in Special Relativity we do not consider anymore the absolute time and no speed is larger than the speed of light. But we cannot avoid the inertial coordinate systems.

Bibliography

[A] Anosov, D.V. - Geodesic flows on closed riemannian manifolds with negative curvature *Proc. Inst. Steklov*, 90 (1967), 1-235.

[AM] Abraham, R., Marsden, J.E. - *Foundations of Mechanics*, Benjamin, (1978).

[Ar] Arnold, V.I. - *Mathematical Methods of Classical Mechanics*, Springer-Verlag, (1980).

[AA] Arnold, V.I., Avez, A. - *Problèmes Ergodiques de la Mécanique Classique*, Gauthier-Villars, (1967).

[ABS] Adler, R., Bazin, M., Schiffer, M. - *Introduction to General Relativity*, Mc. Graw-Hill (1975).

[AKN] Arnold, V.I., Kozlov, V.V., Neishtadt, A.I. - Mathematical Aspects of Classical and Celestial Mechanics, *Dynamical Systems III*, Encyclopaedia of Mathematical Sciences, Ed. Arnold V.I., chapter I (1988), p. 1-286.

[Ch] Chandrasekhar, S. - *Ellipsoidal figures of equilibrium*, Dover, New York (1987).

[CO] Castro, H.M.A., Oliva, W.M. - Anosov flows induced by partially hyperbolic Σ-geodesic flows, *Resenhas* IME-USP, 1999, Vol. 4, nº 2, 227–246 .

[CKO] Castro, HMA., Kobayashi, M.H., Oliva, W.M. - Partially hyperbolic Σ-geodesic flows, to appear in *J. Differential equations.*

[E] Eberlein, P. - Structure of manifolds of nonpositive curvature, *Lect. Notes in Math.* 1156, Springer-Verlag (1984), 86-153.

[F] Frankel, T. - *Gravitational Curvature*,W.H. Freeman, 1979.

[FO] Fusco, G., Oliva, W.M. - Dissipative systems with constraints - *J. Differential Equations*, vol. 63 (3) 1986.

[G] Gouda, N. - Magnetic flows of Anosov type, *Tohoku Math. J.* 49, (1997) 165-183.

[H] Hedlund, G .A. - On the metric transitivity of the geodesics on closed surface of constant negative curvature, *Ann. of Math.* 35(2) (1934), 787-808.

[Ho] Hopf, E. - Statistik der Geodätischen linen in mannigfaltigkeit-ein negativer krümmung, *Ber Verh. Sochs. Akad. Wiss*, Leipzig 91 (1939), 261-304.

[HMO] Hale, J.K., Magalhães, L., Oliva, W.M. - *An Introduction to Infinite Dimension Dynamical Systems: Geometric Theory*, Springer-Verlag, 1984.

[Ka] Kato, T. - *Perturbation theory of linear operators*, Springer-Verlag, 1980.

[KO] Kupka, I., Oliva, W.M. - Dissipative mechanical systems, *Resenhas* IME-USP (1993) vol. 1, n° 1, 69-115.

[KO1] Kupka, I., Oliva, W.M. - The nonholonomic mechanics, *preprint IST-UTL*, Lisbon (1998), to appear in J. of Differential equations.

[L] Lang, S. - *Differential manifolds*, Addison Wesley.

[LS] Lewis, D. and Simo, J.C. - Nonlinear stability of rotating pseudo-rigid bodies, *Proc. Roy. Soc. London Ser.* A, 427, n° 1873 (1990), 271-319.

[Ma] Marsden, J. E. - *Lectures on Mechanics*, Cambridge University Press, (1992).

[MR] Marsden, J. E., Ratiu T. - *Introduction to Mechanics and Symmetry. A basic exposition of classical mechanical systems*, 2nd ed., Texts in Appl. Math., 17, Springer, (1999).

[M] Moser, J. - On a theorem of Anosov, *J. Differential Equations* 5 (1969), 411-440.

[Mon] Montaldi, J. - *Introduction to Geometric Mechanics*, Lecture Notes for DEA, INLN, University of Nice-Sophia Antipolis, France (1995).

[ON] O'Neill, B. - *Semi-riemannian geometry with applications to relativity*, Academic Press, 1983.

[R] de Rham, G. - *Differentiable manifolds*, Springer-Verlag, 1984.

[RS] Roberts, R.M., Sousa-Dias, E. - Symmetries of Riemann ellipsoids, To appear in , *Resenhas* IME. USP, Vol. 4, n° 2, 1999.

[SD] Sousa Dias, E. - Relative equilibria in linear ellasticity, in *Symmetry and Perturbation Theory*, STP98, Degasperis A. and Gaeta, G. editors, World Scientific, (1999) 258-267.

[VG] Vershik, A.M., Gershkovich, V. Ya. - Nonholonomic Dynamical Systems, Geometry of Distributions and Variational Problems, *Dynamical Systems VII*, Encyclopaedia of Mathematical Sciences, Eds. Arnold, V.I., and Novikov S.P. (1990), pp. 4-79.

Index

RELATÓRIOS TÉCNICOS DO DEPARTAMENTO DE MATEMÁTICA APLICADA

2000

RT-MAP-0001 – Laércio C. Barros, Suzana A. O. Souza & Pedro Aladar Tonelli
"Two Cases of Asymptotic Smoothness for Fuzzy Dynamical Systems"
February 16, 2000 - São Paulo - IME-USP - 10 pg.

RT-MAP-0002 – Carlos Juiti Watanabe, Paulo Sérgio Pereira da Silva & Pedro Aladar Tonelli
"Álgebra Diferencial em Teoria de Controle"
Abril, 2000 - São Paulo - IME-USP - 14 pg.

RT-MAP-0003 – Eduardo Colli & Edson Vargas
"An Open Set of Initial Conditions Can Draw a Hyperbolic Horseshoe"
April, 3, 2000 - São Paulo - IME-USP - 32 pg.

RT-MAP-0004 – Eduardo Colli
"A Starting Condition Approach to Parameter Distortion in Generalized Renormalization"
Mayl, 4, 2000 - São Paulo - IME-USP - 60 pg.