Notes on Mathematical Physics for Mathematicians

Daniel V. Tausk

Contents

CHAPTER 1

Classical Mechanics

There are three basic ingredients for the formulation of a physical theory: spacetime, ontology and dynamics. Spacetime is represented, within the theory, by a set endowed with some extra structure; the points of such set are called *events*. For Classical Mechanics, the adequate type of spacetime is *Galilean spacetime*, which consists of a four dimensional real affine space endowed with global absolute time and Euclidean metric structure over the simultaneity hyperplanes defined by the absolute time function. We will give the details of the definition in Section 1.1. To describe the *ontology* of the theory means to say what are the objects whose existence is posited by the theory, i.e., the thing the theory is all about. Such objects should be in some way connected to spacetime. For Classical Mechanics, the ontology consists of *particles*, having *worldlines* which are one-dimensional submanifolds of spacetime. Finally, the *dynamics* says how the ontology is allowed to behave. In the case of Classical Mechanics, this means specifying when a given set of worldlines for the particles is admissible by the theory. This is done by means of a second order ordinary differential equation. We present the ontology and the dynamics of Classical Mechanics in Section 1.2.

1.1. Galilean spacetime

Recall that an *affine space* consists of a non empty set E, a vector space V and a transitive free action $V \times E \to E$ of the additive group $(V, +)$ on the set E. The elements of E are called *points* and the elements of V are called *vectors*; the action of a vector $v \in V$ upon a point $e \in E$ is denoted by $v + e \in E$ or, alternatively, by $e + v \in E$. The fact that the action is free and transitive means that given points $e_1, e_2 \in E$, there exists a unique vector $v \in V$ such that $e_2 = v + e_1$; such vector v is denoted by $e_2 - e_1$. We normally refer to the set E as the affine space and to V as the *underlying vector space* of E. The *dimension* of E and the *scalar field* of E are, by definition, the dimension of V and the scalar field of V, respectively.

1.1.1. Definition. A *Galilean spacetime* consists of:

- a four-dimensional real affine space E;
- a non zero linear functional $\mathfrak{t} : V \to \mathbb{R}$ on the underlying vector space V of E;
- a (positive definite) inner product $\langle \cdot, \cdot \rangle$ on the kernel $\mathrm{Ker}(\mathfrak{t})$ of $\mathfrak{t}$.

We call $\mathfrak{t}$ the *time functional*.

Given events $e_1, e_2 \in E$, the *elapsed time* from e_1 to e_2 is defined as $\mathfrak{t}(e_2 - e_1)$. Two events $e_1, e_2 \in E$ are called *simultaneous* if the elapsed time from e_1 to e_2 is zero, i.e., if the vector $e_2 - e_1$ is in the kernel of $\mathfrak{t}$. When $e_1, e_2 \in E$ are simultaneous, we can define the *distance* between e_1 and e_2 to be $\|e_1 - e_2\|$, where $\|\cdot\|$ is the norm on $\mathrm{Ker}(\mathfrak{t})$ associated to the given inner product. The simultaneity relation is an equivalence relation, which defines a partition of E into equivalence classes. Such equivalence classes are precisely the orbits of the action of $\mathrm{Ker}(\mathfrak{t})$ on E. Those are three-dimensional real affine spaces with underlying vector space $\mathrm{Ker}(\mathfrak{t})$ (see Exercise 1.3); we call them the *simultaneity hyperplanes.*

Notice that, unlike vector spaces in which the origin is a "special" point, affine spaces have no "special" points. That is one of the reasons why spacetime is taken to be an affine space, rather than a vector space. Also, notice that, given a Galilean spacetime, one cannot define an object which represents space. Rather, one has a family of simultaneity hyperplanes (one space for each instant of time). This happens because in Galilean spacetime it is meaningless to say that two events e_1, e_2 on distinct simultaneity hyperplanes are "at the same place in space". To say that two events happened at the same time is meaningful, but to say that they happened at the same place is not!

1.1.1. Units of measurement. The definition of Galilean spacetime given above is not quite what it should be. Notice that, according to our definitions, the elapsed time between two events $e_1, e_2 \in E$ is a real number and this is not quite right. The elapsed time between events should not be a number, but rather something that is expressed using a unit of measurement of time, such as *seconds.* In order to obtain a more appropriate definition of Galilean spacetime, one should fix a real one-dimensional vector space S and demand that the time functional $\mathfrak{t}$ be a non zero linear map $\mathfrak{t} : V \to S$ taking values in S. Thus, elapsed times would be elements of S, not real numbers! The choice of a non zero element of S amounts to the choice of both a *time orientation* and of a *unit of measurement of time* and it allows one to identify S with $\mathbb{R}$. Notice that, by taking $\mathbb{R}$ as the counter-domain of $\mathfrak{t}$ in our definition of Galilean spacetime, we have automatically taken a time orientation to be part of the structure of spacetime (i.e., we can say that an event e_2 is in the *future* of the event e_1 when $\mathfrak{t}(e_2 - e_1) > 0$). One might argue that this is not appropriate[1].

Just as the counter-domain of $\mathfrak{t}$ shouldn't be $\mathbb{R}$, also the counter-domain of the inner product on the kernel of $\mathfrak{t}$ shouldn't be $\mathbb{R}$. The outcome of an inner product should be expressed in units of *squared length.* So, just as in the case of time, one should fix a real one-dimensional vector space M

[1] Orientation of time is not needed for specifying the dynamics of Classical Mechanics. One might argue that it would be more appropriate to leave the orientation of time out of the structure of spacetime, and define it in terms of entropy. We shall not discuss this issue here.

whose elements are to be interpreted as lengths; in this case, a orientation should be given in M (as there should be a canonical notion of positive length). Square lengths should be elements of the tensor product $M \otimes M$, so instead of an inner product on $\mathrm{Ker}(\mathfrak{t})$, one should have a positive-definite symmetric bilinear form on $\mathrm{Ker}(\mathfrak{t})$ taking values in $M \otimes M$. A quotient of a length by an elapsed time (i.e., a *velocity*), would then be an element of the one-dimensional vector space $M \otimes S^*$, where S^* denotes the dual space of S (Exercise 1.6 clarifies the role of tensor products and dual spaces here).

We shall not pursue this programme further in this text. We just would like the reader to be aware that a better way to formulate things in Physics is to have lots of given one-dimensional vector spaces, each one appropriate for the values of a given type of physical dimension (such as elapsed time, length, mass, charge, etc) and lots of other one-dimensional vector spaces constructed from the given ones using, for instance, tensor products and duals, which are spaces for the values of things such as velocities, accelerations, forces, etc. Since working like this systematically would be too annoying and distractive, we shall simply assume that units of measurement have been chosen and we will identify all such would-be one-dimensional vector spaces with $\mathbb{R}$ once and for all.

1.1.2. Inertial coordinate systems and the Galileo group. We start by defining a notion of isomorphism for Galilean spacetimes. Recall that, given affine spaces E, E', with underlying vector spaces V, V', respectively, then a map $A : E \to E'$ is called *affine* if there exists a linear map $L : V \to V'$ such that:

$$A(v + e) = L(v) + A(e),$$

for all $v \in V$, $e \in E$. The linear map L is unique when it exists and it is called the *underlying linear map* of A. An affine map A is called an *affine isomorphism* if it is bijective; this happens if and only if the underlying linear map L is bijective. When A is an affine isomorphism, then its inverse A^{-1} is also affine, with underlying linear map L^{-1}. The affine automorphisms of E (i.e., the affine isomorphisms from E to E) form a group $\mathrm{Aff}(E)$ under composition and the map $A \mapsto L$ that associates to each affine isomorphism A its underlying linear map L is a group homomorphism from $\mathrm{Aff}(E)$ onto the general linear group $\mathrm{GL}(V)$ (i.e., the group of linear isomorphisms of V). The kernel of such homomorphism is a normal subgroup of the group of affine automorphisms of E; its elements are of the form $E \ni e \mapsto v + e \in E$, with $v \in V$, and they are called *translations* of E. The group of all translations of E is obviously isomorphic to the additive group of V (see Exercise 1.5).

1.1.2. Definition. Let $(E, V, \mathfrak{t}, \langle\cdot,\cdot\rangle)$, $(E', V', \mathfrak{t}', \langle\cdot,\cdot\rangle')$ be Galilean spacetimes. An *isomorphism* $A : (E, V, t, \langle\cdot,\cdot\rangle) \to (E', V', t', \langle\cdot,\cdot\rangle')$ is an affine isomorphism $A : E \to E'$ with underlying linear isomorphism $L : V \to V'$ such that:

(a) *L preserves time*, i.e., $\mathfrak{t}' \circ L = \mathfrak{t}$;

(b) L carries the inner product $\langle\cdot,\cdot\rangle$ on $\mathrm{Ker}(\mathfrak{t})$ to the inner product $\langle\cdot,\cdot\rangle'$ on $\mathrm{Ker}(\mathfrak{t}')$, i.e., $\langle L(v_1), L(v_2)\rangle' = \langle v_1, v_2\rangle$, for all $v_1, v_2 \in \mathrm{Ker}(\mathfrak{t})$.

Notice that condition (a) above implies that L sends $\mathrm{Ker}(\mathfrak{t})$ onto $\mathrm{Ker}(\mathfrak{t}')$. Galilean spacetimes and its isomorphisms form a category. Two Galilean spacetimes are always isomorphic, i.e., given two Galilean spacetimes there always exist an isomorphism from one to the other (see Exercise 1.7).

One might wonder why we care about defining isomorphisms of Galilean spacetimes or why should we ever talk about "Galilean spacetimes" in the plural when presenting a physical theory; after all, in Physics we should have only one spacetime. We have two purposes in mind for isomorphisms: one, is to consider the set of all automorphisms of a given Galilean spacetime. Such set forms a group under composition and it is called the *Galileo group*. The other purpose is to give an elegant definition of a privileged class of coordinate systems over Galilean spacetime, which is what we are going to do next.

Consider the affine space $\mathbb{R}^4$ canonically obtained from the vector space $\mathbb{R}^4$ (see Exercise 1.1), the time functional on the vector space $\mathbb{R}^4$ given by the projection onto the first coordinate and the canonical Euclidean inner product on the kernel $\{0\} \times \mathbb{R}^3 \cong \mathbb{R}^3$ of such time functional. We have just defined a Galilean spacetime, which we call the *Galilean spacetime of coordinates*. In what follows, we denote by $(E, V, \mathfrak{t}, \langle\cdot,\cdot\rangle)$ the Galilean spacetime over which Classical Mechanics is going to be formulated.

1.1.3. Definition. An isomorphism $\phi : E \to \mathbb{R}^4$ from the Galilean space time E to the Galilean spacetime of coordinates $\mathbb{R}^4$ is called an *inertial coordinate system*.

Notice that we have avoided the (very common) terminology "inertial observer". This is no accident. There is no need to talk about "observers" here. We will discuss this point in more detail later.

1.1.4. Definition. The group of automorphisms of the Galilean spacetime E is called the *active Galileo group* and the group of automorphisms of the Galilean spacetime of coordinates $\mathbb{R}^4$ is called the *passive Galileo group*.

Obviously, the active and the passive Galileo groups are isomorphic, as any choice of inertial coordinate system induces an isomorphism between them (see Exercise 1.8); but such isomorphism is not canonical, in the sense that it depends on the choice of the inertial coordinate system. This two Galileo groups have distinct physical interpretations: elements of the active Galileo group are used to transform stuff inside spacetime (say, move particle worldlines around), while elements of the passive Galileo group are used to relate two inertial coordinate systems. More explicitly, if $\phi_1 : E \to \mathbb{R}^4$, $\phi_2 : E \to \mathbb{R}^4$ are inertial coordinate systems then the only map $A : \mathbb{R}^4 \to \mathbb{R}^4$

such that $\phi_2 = A \circ \phi_1$, i.e., such that the diagram:

$$\begin{tikzcd} & E \arrow[dl, "\phi_1"'] \arrow[dr, "\phi_2"] & \\ \mathbb{R}^4 \arrow[rr, "A"'] & & \mathbb{R}^4 \end{tikzcd} \tag{1.1.1}$$

commutes is an element of the passive Galileo group.

We finish the section by taking a closer look at the Galileo group. Let $A : \mathbb{R}^4 \to \mathbb{R}^4$ be an element of the passive Galileo group. Then A can be written as the composition of a linear isomorphism $L : \mathbb{R}^4 \to \mathbb{R}^4$ with a translation of $\mathbb{R}^4$:

$$A(t, x) = L(t, x) + (t_0, x_0), \quad (t, x) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4, \tag{1.1.2}$$

where $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$ is fixed. The linear isomorphism L is required to satisfy conditions (a) and (b) in Definition 1.1.2. Condition (a) says that L is of the form:

$$L(t, x) = \big(t, L_0(x) - vt\big), \quad (t, x) \in \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4, \tag{1.1.3}$$

for some fixed $v \in \mathbb{R}^3$ and some linear isomorphism $L_0 : \mathbb{R}^3 \to \mathbb{R}^3$. The expression $L_0(x) - vt$ just expresses an arbitrary $\mathbb{R}^3$-valued linear function of the pair (t, x); the reason for us choosing to write a minus sign in front of v will be apparent in a moment. Condition (b) now requires that the linear isomorphism L_0 be an isometry of $\mathbb{R}^3$, i.e., an element of the orthogonal group $\mathrm{O}(3)$. Let us consider the following subgroups of the passive Galileo group:

- the group of translations of $\mathbb{R}^4$, which is isomorphic to the additive group $\mathbb{R}^4$;
- the group of isometries (i.e., rotations and reflections) of $\mathbb{R}^3$ (or, more precisely, of $\{0\} \times \mathbb{R}^3$), which is the orthogonal group $\mathrm{O}(3)$;
- the group of linear isomorphisms of $\mathbb{R}^4$ of the form:

$$\mathbb{R} \times \mathbb{R}^3 \ni (t, x) \longmapsto (t, x - vt) \in \mathbb{R} \times \mathbb{R}^3, \tag{1.1.4}$$

  with $v \in \mathbb{R}^3$. This group is isomorphic to the additive group $\mathbb{R}^3$ and its elements are called *Galilean boosts* (see Exercise 1.9 for the physical interpretation of Galilean boosts and for an explanation of why we prefer to use a minus sign in front of v).

One immediately sees that every element of the passive Galileo group can be written in a unique way as a composition of a translation of $\mathbb{R}^4$, an isometry of $\mathbb{R}^3$ and a Galilean boost. We can identify the passive Galileo group with the cartesian product:

$$\mathbb{R}^4 \times \mathrm{O}(3) \times \mathbb{R}^3, \tag{1.1.5}$$

by identifying the element A defined by (1.1.2) and (1.1.3) with the triple $\big((t_0, x_0), L_0, v\big)$. It is easy to check that the passive Galileo group is a closed

Lie subgroup of the Lie group[2] $\mathrm{Aff}(\mathbb{R}^4)$ of affine isomorphisms of $\mathbb{R}^4$ and that the map that identifies it with the cartesian product (1.1.5) is a smooth diffeomorphism (in particular, the Galileo group is a *ten dimensional* Lie group). We emphasize, however, that the group structure on the cartesian product (1.1.5) that turns such map into a group isomorphism is *not* the one of the direct product of the groups $\mathbb{R}^4$, $\mathrm{O}(3)$ and $\mathbb{R}^3$. We left to the reader as an exercise to write down explicitly the appropriate group structure on (1.1.5) and to check that it coincides with the one of a *semi-direct product* of the form:

$$\mathbb{R}^4 \rtimes \big(\mathrm{O}(3) \ltimes \mathbb{R}^3\big),$$

where $\mathrm{O}(3)$ acts on $\mathbb{R}^3$ and $\mathrm{O}(3) \ltimes \mathbb{R}^3$ acts on $\mathbb{R}^4$.

As for the active Galileo group, we have already observed that it is isomorphic to the passive Galileo group (by means of a choice of an inertial coordinate system) and any isomorphism between the two Galileo groups carries the three subgroups of the passive Galileo group defined above to subgroups of the active Galileo group. However, except for the group of translations, the corresponding subgroups of the active Galileo group *depend* on the choice of inertial coordinate system. Namely, it makes sense to say that an affine isomorphism of the affine space E is a (pure) translation, but it doesn't make sense to say that it is purely linear (it only makes sense to say that an affine map is linear when its domain and counter-domain are affine spaces which have been canonically obtained from vector spaces).

1.2. Ontology and dynamics

Let us now present the ontology and the dynamics of Classical Mechanics. We will do this using an inertial coordinate system, which will be fixed throughout the section. One must then check that the dynamics defined (i.e., the particle worldlines which are admissible by the theory) do not depend on the choice of inertial coordinate system. This task is very simple and will be left to the reader (Exercises 1.10, 1.11 and 1.12). It is also possible to give an *intrinsic* formulation of Classical Mechanics, i.e., to formulate it directly over Galilean spacetime. To do so isn't terribly difficult, but we do not want to distract the reader with the technical complications that arise during such task, so we relegate such formulation to an optional section (Section 1.3).

Classical Mechanics is a theory about *particles*. Particles have *trajectories* that are represented within the theory by smooth curves $q : \mathbb{R} \to \mathbb{R}^3$ (what we mean when we use the word "particle" is precisely that such trajectories exist). The *graph* of the map q, i.e., the set:

$$\mathrm{gr}(q) = \big\{\big(t, q(t)\big) : t \in \mathbb{R}\big\} \subset \mathbb{R}^4$$

[2]The group $\mathrm{Aff}(\mathbb{R}^4)$ has a differential structure because it is an open subset of the real finite-dimensional vector space of all affine maps of $\mathbb{R}^4$. It is diffeomorphic to the cartesian product $\mathrm{GL}(\mathbb{R}^4) \times \mathbb{R}^4$, where the first coordinate represents the linear part and the second the translation part.

is called the *worldline* of the particle. More precisely, the worldline of the particle is the subset of the Galilean spacetime E that is mapped onto $\mathrm{gr}(q)$ by our fixed inertial coordinate system; but we will keep referring to $\mathrm{gr}(q)$ as the worldline of the particle, anyway. Also, it should be mentioned that the map $q : \mathbb{R} \to \mathbb{R}^3$ becomes well-defined only after a choice of an inertial coordinate system (as have been mentioned before, in the absence of a choice of an inertial coordinate system, we do not even have an object to use as the counter-domain of the map q, i.e., we do not have an object representing space).

A universe described by Classical Mechanics contains a certain number n of particles, with trajectories[3]:

$$q_j : \mathbb{R} \longrightarrow \mathbb{R}^3, \quad j = 1, 2, \ldots, n.$$

To each particle it is associated a positive real number m_j called the *mass* of the particle. The natural number n and the positive real numbers m_j are parameters of the theory. This completes the description of the ontology. Now, let us describe the dynamics. For each $j = 1, \ldots, n$, the trajectory q_j of the j-th particle must satisfy the differential equation:

$$\frac{\mathrm{d}^2 q_j}{\mathrm{d}t^2}(t) = F_j\big(t, q_1(t), \ldots, q_n(t), \tfrac{\mathrm{d}q_1}{\mathrm{d}t}(t), \ldots, \tfrac{\mathrm{d}q_n}{\mathrm{d}t}(t)\big), \quad t \in \mathbb{R}, \tag{1.2.1}$$

where $F_j : \mathrm{dom}(F_j) \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n \to \mathbb{R}^3$ is a smooth map defined over some open subset $\mathrm{dom}(F_j)$ of $\mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n$ and it is called the *total force acting upon the j-th particle* (actually, it is more appropriate to call the map F_j the *force law* and the righthand side of (1.2.1) the *force*).

What we have just presented is merely a prototype of a theory. In order to complete the formulation of the theory, we have to present also a table of force laws, i.e., we have to say what the maps F_j are. We will now present just a very small table of force laws. One should keep in mind that by expanding such table one can obtain new theories compatible with the prototype described above.

The force laws described below will only depend upon the positions $q_j(t)$ and thus we will omit the other variables.

- The *gravitational force*. For $i \neq j$, we set:

  $$F_{ij}^{\mathrm{gr}}(q_1, \ldots, q_n) = \frac{G m_i m_j}{\|q_i - q_j\|^3}(q_j - q_i),$$

  and we call it the *gravitational force of particle j acting upon particle i* (or, alternatively, the gravitational force that particle j *exerts* upon particle i). Here G denotes a constant called *Newton's gravitational constant.*

[3] A particular correspondence between the particles and the numbers 1, 2, ..., n is obviously not meant to have physical meaning. It would be more appropriate to use a general n-element set for labeling the particles, rather than the set $\{1, 2, \ldots, n\}$.

- The *electrical force*. For $i \neq j$, we set:
$$F^{\mathrm{el}}_{ij}(q_1, \ldots, q_n) = -\frac{Ce_ie_j}{\|q_i - q_j\|^3}(q_j - q_i),$$
and we call it the *electrical force of particle j acting upon particle i* (or, alternatively, the electrical force that particle j *exerts* upon particle i). Here C denotes a constant called *Coulomb's constant* and $e_j \in \mathbb{R}$, $j = 1, \ldots, n$, is a new parameter of the theory called the *charge* of the j-th particle.

We define the *total gravitational force acting upon particle j* by:
$$F^{\mathrm{gr}}_j = \sum_{i \neq j} F^{\mathrm{gr}}_{ji},$$
and, similarly, the *total electrical force acting upon particle j* by:
$$F^{\mathrm{el}}_j = \sum_{i \neq j} F^{\mathrm{el}}_{ji}.$$
Now we state a rule that says that *the forces in our table should be added in order that the total force F_j be obtained.* Thus, the total force F_j acting upon particle j is given by:
$$F_j = F^{\mathrm{gr}}_j + F^{\mathrm{el}}_j.$$

Notice that both for the gravitational and for the electrical force we have:
$$F_{ij} = -F_{ji}. \tag{1.2.2}$$
This is sometimes called *Newton's law of reciprocal actions.* It implies that:
$$\sum_{j=1}^{n} F_j = \sum_{j=1}^{n} \sum_{i \neq j} F_{ji} = 0. \tag{1.2.3}$$

The gravitational and the electrical forces are the *fundamental* forces of Classical Mechanics. When dealing with practical physics problems, other forces do arise, such as friction, viscosity, contact forces, etc. Those other forces are supposed to be *emergent* forces; one would not need them if all the microscopic details of the interactions were to be taken into account[4].

Also, we should mention *external forces.* A physical theory, in principle, is supposed to be about the entire universe. So, in the universe described by Classical Mechanics as formulated above, there is nothing but those n particles. Evidently, one would like also to apply the theory to subsystems of the universe. For Classical Mechanics, a subsystem of the universe is obtained by choosing a subset of the set of all particles, i.e., a subset I of the set of labels $\{1, 2, \ldots, n\}$. Set $I^{\mathrm{c}} = \{1, 2, \ldots, n\} \setminus I$. One could then

[4]Of course, this cannot be taken too seriously, as we know that Classical Mechanics is not a fundamental theory. Classical Mechanics does not work properly at the microscopic scale.

compare the dynamics of the particles with labels in I when the presence of the particles with labels in I^{c} is taken into account with the dynamics of the particles with labels in I when the presence of the particles with labels in I^{c} is *not* taken into account. One should not expect the two dynamics to be identical, but in many cases it might happen that those two dynamics are very similar (observe, for instance, that F_{ij} becomes very small when $\|q_i - q_j\|$ becomes very large). In those cases, we say that the subsystem defined by I is *almost isolated*. There is another possibility: maybe we don't get a good approximation of the dynamics of the particles with labels in I by ignoring the existence of the particles with labels in I^{c}, but we do get a good approximation of the dynamics of the particles with labels in I by taking into account only the forces exerted upon particles with labels in I by particles with labels in I^{c} and by ignoring the forces exerted upon particles with labels in I^{c} by particles with labels in I (for example, we can study the dynamics of a tennis ball near the Earth by taking into account the gravitational force exerted by the Earth upon the tennis ball, and by ignoring the gravitational force exerted by the tennis ball upon the Earth). Thus, we could write down the differential equations for the trajectories of the particles with labels in I using the *internal forces* of the subsystem defined by I (i.e., the forces between particles with labels in I) and also the *external forces* (i.e., the forces exerted upon the particles with labels in I by particles with labels in I^{c}). There is also another type of situation in which we would like to talk about external forces. Sometimes, in Physics, we need to make "Frankenstein theories", mixing up pieces of (not necessarily fully compatible) distinct theories. For instance, one might like to consider the particles of Classical Mechanics moving inside a magnetic field. Magnetic fields are not part of Classical Mechanics: they are part of Maxwell's electromagnetism, a theory which is not even formulated within Galilean spacetime (it is formulated within *Minkowski spacetime*). Nevertheless, for practical applications, one might well be willing to consider the particles of Classical Mechanics inside a magnetic field and thus one would have to consider the force exerted by the magnetic field upon the particles (the *Lorentz force*).

1.2.1. Forces with a potential. Taking together all the force maps F_j we obtain a map:

$$F = (F_1, \ldots, F_n) : \mathrm{dom}(F) \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n \longrightarrow (\mathbb{R}^3)^n.$$

If the righthand side of (1.2.1) does not depend on the velocities $\frac{\mathrm{d}q_j}{\mathrm{d}t}(t)$ then we can think of F as a map of the form:

$$F : \mathrm{dom}(F) \subset \mathbb{R} \times (\mathbb{R}^3)^n \longrightarrow (\mathbb{R}^3)^n.$$

That is a *time-dependent vector field* over $(\mathbb{R}^3)^n$, i.e., for each $t \in \mathbb{R}$, we have a vector field $F(t, \cdot)$ over (an open subset of) $(\mathbb{R}^3)^n$. When the forces F_j do not depend on the velocities and there exists a smooth map:

$$V : \mathrm{dom}(V) = \mathrm{dom}(F) \subset \mathbb{R} \times (\mathbb{R}^3)^n \longrightarrow \mathbb{R}$$

such that[5]:

$$F(t,q) = -\nabla_q V(t,q), \quad (t,q) \in \mathrm{dom}(F) \subset \mathbb{R} \times (\mathbb{R}^3)^n, \tag{1.2.4}$$

then we call V a *potential* for the force F. In (1.2.4) we have denoted by $\nabla_q V(t,q)$ the gradient of the map $V(t,\cdot)$ evaluated at the point q. Forces that admit a potential will be essential for the variational formulation of Classical Mechanics (Subsection 1.4.1). When $F(t,q)$ does not depend on t, then, as is well-known from elementary calculus, the existence of the potential V is equivalent to the condition that the line integral $\int_q F$ of the vector field F over a piecewise smooth curve $q : [a,b] \to (\mathbb{R}^3)^n$ depend only on the endpoints $q(a)$, $q(b)$ of the curve[6]. In that case the force F is called *conservative*. We leave as a simple exercise to the reader to check that the following:

$$V^{\mathrm{gr}}(q_1,\ldots,q_n) = -\sum_{i<j} \frac{Gm_im_j}{\|q_i - q_j\|}, \quad V^{\mathrm{el}}(q_1,\ldots,q_n) = \sum_{i<j} \frac{Ce_ie_j}{\|q_i - q_j\|},$$

are potentials for the gravitational and for the electrical forces, respectively. Obviously, $V = V^{\mathrm{gr}} + V^{\mathrm{el}}$ is a potential for $F = F^{\mathrm{gr}} + F^{\mathrm{el}}$.

1.3. Optional section: intrinsic formulation

In this section we present an intrinsic (i.e., coordinate system free) formulation of the ontology and dynamics of Classical Mechanics. Material from this section will not be used elsewhere, so uninterested readers may safely skip it. While some might complain that this section contains too much "abstract nonsense", we think that something is learned from this exercise. In what follows, a Galilean spacetime $(E, V, \mathfrak{t}, \langle\cdot,\cdot\rangle)$ is fixed.

Consider the quotient:

$$\mathfrak{T} = E/\operatorname{Ker}(\mathfrak{t}),$$

which is a one-dimensional real affine space with underlying vector space $V/\operatorname{Ker}(\mathfrak{t})$ (see Exercise 1.4). The linear functional $\mathfrak{t}$ induces an isomorphism between the quotient $V/\operatorname{Ker}(\mathfrak{t})$ and the real line $\mathbb{R}$; by means of such isomorphism, we may regard $\mathbb{R}$ as the underlying vector space of the affine space $\mathfrak{T}$. Denote by:

$$\bar{\mathfrak{t}} : E \longrightarrow \mathfrak{T}$$

the quotient map, which is an affine map whose underlying linear map is $\mathfrak{t}$. A point t of the affine space $\mathfrak{T}$ is called an *instant of time*; the inverse image of t by $\bar{\mathfrak{t}}$ is a simultaneity hyperplane, which is to be interpreted as *space at the instant t*.

---

[5]Why do we use a minus sign in the righthand side of (1.2.4)? It is a good convention. For instance, it allows us to build some intuition about the behavior of the particles by thinking about the graph of V as some sort of roller coaster.

[6]If the curve is $q = (q_1,\ldots,q_n)$, with q_j denoting the trajectory of the j-th particle, then such line integral is called the *work* done by the force F.

Recall that a real finite-dimensional affine space is, in a natural way, a differentiable manifold whose tangent space at an arbitrary point is naturally identified with the underlying vector space of the affine space (see Exercise 1.13). The affine map $\bar{\mathfrak{t}}$ is smooth[7] (see Exercise 1.14). By a *section* of $\bar{\mathfrak{t}}$ we mean a map $q : \mathfrak{T} \to E$ such that $\bar{\mathfrak{t}} \circ q$ is the identity map of $\mathfrak{T}$. We can define a *particle worldline* to be the image of a smooth section q of $\bar{\mathfrak{t}}$; clearly, the section q is uniquely determined by the corresponding worldline. Equivalently, one can define a particle worldline to be a smooth submanifold of E that is mapped diffeomorphically onto $\mathfrak{T}$ by the map $\bar{\mathfrak{t}}$.

Let $q : \mathfrak{T} \to E$ be a smooth section of $\bar{\mathfrak{t}}$. Given $t \in \mathfrak{T}$, then the tangent space $T_t\mathfrak{T}$ is canonically identified with $\mathbb{R}$ and the tangent space $T_{q(t)}E$ is canonically identified with V; thus, the differential $\mathrm{d}q(t)$ is a linear map from $\mathbb{R}$ to V. We set:

$$\dot{q}(t) = \mathrm{d}q(t)1 \in V,$$

and we call it the *velocity at the instant* $t \in \mathfrak{T}$ of a particle whose worldline is the image of q. Since $\bar{\mathfrak{t}} \circ q$ is the identity of $\mathfrak{T}$, it follows by differentiation that $\mathfrak{t} \circ \mathrm{d}q(t)$ is the identity of $\mathbb{R}$, so that, in particular:

$$\mathfrak{t}\big(\dot{q}(t)\big) = 1.$$

The velocity $\dot{q}(t)$ is a generator of the (one-dimensional) tangent space at the point $q(t)$ of the worldline $q(\mathfrak{T})$; actually, it is the only vector in that tangent space that is mapped by $\mathfrak{t}$ to the number 1. The map $\dot{q}$ (which is essentially the differential of q) is smooth and it takes values in the affine subspace $\mathfrak{t}^{-1}(1)$ of V:

$$\dot{q} : \mathfrak{T} \longrightarrow \mathfrak{t}^{-1}(1) \subset V.$$

The underlying vector space of the affine space $\mathfrak{t}^{-1}(1)$ is $\mathrm{Ker}(\mathfrak{t})$. We can differentiate the map $\dot{q}$ at some $t \in \mathfrak{T}$ to obtain a linear map:

$$\mathrm{d}\dot{q}(t) : \mathbb{R} \longrightarrow \mathrm{Ker}(\mathfrak{t}).$$

We set:

$$\ddot{q}(t) = \mathrm{d}\dot{q}(t)1 \in \mathrm{Ker}(\mathfrak{t}).$$

We call $\ddot{q}(t)$ the *acceleration at the instant* $t \in \mathfrak{T}$ of a particle whose worldline is the image of q.

We have learned some interesting things: velocities $\dot{q}(t)$ are elements of the affine space $\mathfrak{t}^{-1}(1)$. So we cannot add two velocities and get a new velocity! On the other hand, we can subtract two velocities and obtain an element of the three-dimensional vector space $\mathrm{Ker}(\mathfrak{t})$. A difference of velocities is a *relative velocity*; since $\mathrm{Ker}(\mathfrak{t})$ is endowed with an inner product, we can talk about the norm of a relative velocity. That's a *relative speed.* Notice that, unlike velocities, accelerations $\ddot{q}(t)$ are elements of the vector space $\mathrm{Ker}(\mathfrak{t})$, so we can add them and multiply them by real numbers obtaining new elements of $\mathrm{Ker}(\mathfrak{t})$. It is also meaningful to take the norm of an acceleration.

---

[7] Actually, it is a smooth fibration.

The intrinsic formulation of Classical Mechanics is almost done. The trajectories $q_j : \mathbb{R} \to \mathbb{R}^3$ used in the formulation of Section 1.2 must be replaced by sections $q_j : \mathfrak{T} \to E$ of $\bar{\mathfrak{t}}$. The lefthand side of equality (1.2.1) is replaced with $m_j\ddot{q}_j(t)$. We just have to explain what type of objects should the maps F_j be replaced with.

Consider the cartesian product E^n of n copies of E. This is an affine space with underlying vector space V^n. Let Q denote the subset of E^n defined by:

$$Q = \big\{(q_1, \ldots, q_n) \in E^n : \mathfrak{t}(q_i - q_j) = 0,\ i, j = 1, \ldots, n\big\}.$$

The set Q is an affine subspace of E^n with underlying vector space:

$$\big\{(v_1, \ldots, v_n) \in V^n : \mathfrak{t}(v_i - v_j) = 0,\ i, j = 1, \ldots, n\big\}.$$

The affine space Q is called *configuration spacetime*. Notice that, for each t in $\mathfrak{T}$, the n-tuple $\big(q_1(t), \ldots, q_n(t)\big)$ belongs to Q. In the intrinsic formulation of Classical Mechanics, the force laws F_j are maps of the form:

$$F_j : \mathrm{dom}(F_j) \subset Q \times \mathfrak{t}^{-1}(1)^n \longrightarrow \mathrm{Ker}(\mathfrak{t}). \tag{1.3.1}$$

It is readily checked that both the gravitational and the electrical forces are well-defined maps of the form (1.3.1).

1.4. An introduction to variational calculus

A *variational problem* is a particular case of the problem of finding a critical point of a map whose domain is typically infinite-dimensional. We proceed to describe an important class of variational problems for curves.

Let $[a, b]$ be an interval and consider the vector space $C^\infty\big([a, b], \mathbb{R}^n\big)$ of smooth maps $q : [a, b] \to \mathbb{R}^n$. Given points $q_a, q_b \in \mathbb{R}^n$, then the set:

$$C^\infty_{q_a q_b}\big([a, b], \mathbb{R}^n\big) = \big\{q \in C^\infty\big([a, b], \mathbb{R}^n\big) : q(a) = q_a,\ q(b) = q_b\big\}$$

is an affine subspace of $C^\infty\big([a, b], \mathbb{R}^n\big)$ whose underlying vector space is the space $C^\infty_{00}\big([a, b], \mathbb{R}^n\big)$ of smooth maps from $[a, b]$ to $\mathbb{R}^n$ vanishing at the endpoints of the interval $[a, b]$. Consider a smooth map:

$$L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

and define:

$$S_L : C^\infty_{q_a q_b}\big([a, b], \mathbb{R}^n\big) \longrightarrow \mathbb{R}$$

by setting:

$$S_L(q) = \int_a^b L\big(t, q(t), \dot{q}(t)\big)\,\mathrm{d}t,$$

where $\dot{q}(t) = \frac{\mathrm{d}q}{\mathrm{d}t}(t)$. The map L is called a *Lagrangian* and S_L is called the corresponding *action functional*. More precisely, we have a *family* of action functionals associated to the Lagrangian L (one action functional for each interval $[a, b]$ and for each choice of points $q_a, q_b \in \mathbb{R}^n$); but we will be a little sloppy and refer to any of them as "the action functional". The variational problem that we are going to consider is the problem of finding the critical points of S_L. But we do not intend to discuss infinite-dimensional calculus seriously, because we don't have to. We will just present a definition of critical point for this specific context. If we were to discuss infinite-dimensional calculus seriously, then it would be preferable to replace $C^\infty\big([a,b],\mathbb{R}^n\big)$ with the space $C^k\big([a,b],\mathbb{R}^n\big)$ of maps of class C^k, for some fixed finite k. The space $C^k\big([a,b],\mathbb{R}^n\big)$, endowed with the appropriate norm (for instance, the sum of the supremum norm of the function with the supremum norms of its first k derivatives) is a Banach space[8], while $C^\infty\big([a,b],\mathbb{R}^n\big)$ can only handle the structure of a *Fréchet space*. Calculus on Banach spaces is simpler to handle (it is more similar to finite-dimensional calculus) than calculus on Fréchet spaces. But we are not going to need any theorems from infinite-dimensional calculus, so we don't have to worry about any of that.

1.4.1. Definition. Let $q : [a,b] \to \mathbb{R}^n$ be a smooth curve. By a *variation* of q we mean a family $(q_s)_{s\in I}$ of smooth curves $q_s : [a,b] \to \mathbb{R}^n$, where $I \subset \mathbb{R}$ is an open interval with $0 \in I$, q_0 equals q and the map:

$$I \times [a,b] \ni (s,t) \longmapsto q_s(t) \in \mathbb{R}^n$$

is smooth. We say that the variation $(q_s)_{s\in I}$ has *fixed endpoints* if the maps:

$$I \ni s \longmapsto q_s(a) \in \mathbb{R}^n, \quad I \ni s \longmapsto q_s(b) \in \mathbb{R}^n$$

are constant. The *variational vector field* of a variation $(q_s)_{s\in I}$ is the smooth map $v : [a,b] \to \mathbb{R}^n$ defined by:

$$v(t) = \frac{\mathrm{d}}{\mathrm{d}s}\, q_s(t)\Big|_{s=0}, \quad t \in [a,b]. \tag{1.4.1}$$

Clearly, if v is the variational vector field of a variation with fixed endpoints then $v(a) = v(b) = 0$.

Notice that a smooth curve $q : [a,b] \to \mathbb{R}^n$ is a point of the space $C^\infty_{q_a q_b}\big([a,b],\mathbb{R}^n\big)$, where $q_a = q(a)$, $q_b = q(b)$; a variation with fixed endpoints of q is a curve $s \mapsto q_s$ in $C^\infty_{q_a q_b}\big([a,b],\mathbb{R}^n\big)$ passing through the point q at $s = 0$. The corresponding variational vector field v is like the vector tangent to that curve at $s = 0$.

1.4.2. Definition. We say that a smooth curve $q : [a,b] \to \mathbb{R}^n$ is a *critical point* of the action functional S_L if

$$\frac{\mathrm{d}}{\mathrm{d}s}\, S_L(q_s)\Big|_{s=0} = 0, \tag{1.4.2}$$

for every variation with fixed endpoints $(q_s)_{s\in I}$ of q.

In the context of infinite-dimensional calculus, the map S_L is smooth and the lefthand side of (1.4.2) is precisely the differential of S_L at the point q in the direction v, where v is the variational vector field of $(q_s)_{s\in I}$. With our approach, however, it is not clear in principle that the lefthand side of

[8]The space $C^k_{q_a q_b}\big([a,b],\mathbb{R}^n\big)$ is then a closed affine subspace of a Banach space and therefore it is a Banach manifold.

(1.4.2) depends only on v (i.e., that two variations with the same variational vector field would yield the same value for the lefthand side of (1.4.2)) and that it is linear in v. Nevertheless, as we will see now, a very simple direct computation of the lefthand side of (1.4.2) shows that both things are true. Using any standard result about differentiation under the integral sign and the chain rule we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}s} S_L(q_s)\Big|_{s=0} = \int_a^b \frac{\partial L}{\partial q}\big(t, q(t), \dot{q}(t)\big)v(t) + \frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)\dot{v}(t)\,\mathrm{d}t, \tag{1.4.3}$$

where $\dot{v}(t) = \frac{\mathrm{d}v}{\mathrm{d}t}(t)$. Above, we have denoted by $\frac{\partial L}{\partial q}$ and by $\frac{\partial L}{\partial \dot{q}}$ the differential of L with respect to its second and third variables; thus, in $\frac{\partial L}{\partial \dot{q}}$ the symbol $\dot{q}$ is merely a label, not "the derivative of q". Notice that the expressions:

$$\frac{\partial L}{\partial q}\big(t, q(t), \dot{q}(t)\big), \quad \frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)$$

are differentials evaluated at a point of real valued functions over $\mathbb{R}^n$ and therefore they are elements of the dual space $\mathbb{R}^{n*}$ (thus they can be applied to vectors $v(t)$, $\dot{v}(t)$ of $\mathbb{R}^n$, as we did in (1.4.3)). We won't systematically take too seriously the difference between $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ and we will sometimes identify the two spaces in the usual way.

Now we want to use integration by parts in (1.4.3) to get rid of $\dot{v}(t)$. In other words, we observe that:

$$\frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)\dot{v}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)v(t)\Big) - \Big(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)\Big)v(t). \tag{1.4.4}$$

Using the fundamental theorem of calculus and the fact that:

$$v(a) = v(b) = 0$$

for variations with fixed endpoints, we see that the first term on the righthand side of (1.4.4) vanishes once it goes inside the integral sign. Thus:

$$\frac{\mathrm{d}}{\mathrm{d}s} S_L(q_s)\Big|_{s=0} = \int_a^b \Big(-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big) + \frac{\partial L}{\partial q}\big(t, q(t), \dot{q}(t)\big)\Big)v(t)\,\mathrm{d}t. \tag{1.4.5}$$

We want to infer from (1.4.5) that q is a critical point of S_L if and only if the big expression inside the parenthesis in (1.4.5) vanishes, i.e., if and only if:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big) = \frac{\partial L}{\partial q}\big(t, q(t), \dot{q}(t)\big), \quad t \in [a, b]. \tag{1.4.6}$$

The differential equation (1.4.6) is called the *Euler–Lagrange equation*. Obviously, if the Euler–Lagrange equation holds then, by (1.4.5), q is a critical point of S_L. In order to prove the converse, we need two ingredients.

The first, is the observation that any smooth map $v : [a, b] \to \mathbb{R}^n$ with $v(a) = v(b) = 0$ is the variational vector field of some variation of q with fixed endpoints. For instance, we can consider the variation[9]:

$$q_s(t) = q(t) + sv(t).$$

By this observation, the assumption that q be a critical point of S_L implies that the righthand side of (1.4.5) vanishes, for any smooth map v such that $v(a) = v(b) = 0$. The second ingredient is this.

1.4.3. LEMMA (fundamental lemma of the calculus of variations). *Let $\alpha : [a, b] \to \mathbb{R}^{n*}$ be a continuous map and assume that:*

$$\int_a^b \alpha(t)v(t)\,\mathrm{d}t = 0, \tag{1.4.7}$$

for any smooth map $v : [a, b] \to \mathbb{R}^n$ having support contained in the open interval $]a, b[$. Then $\alpha = 0$.

PROOF. Assuming by contradiction that α is not zero, then the j-th coordinate $\alpha_j : [a, b] \to \mathbb{R}$ of α is not zero for some $j = 1, \ldots, n$. By continuity, α_j never vanishes (and has a fixed sign) over some interval $[c, d]$ contained in $]a, b[$. One can construct a smooth v whose only non vanishing coordinate is the j-th coordinate v_j and such that v_j is non negative (but not identically zero) over $[c, d]$ and vanishes outside $[c, d]$. Then the integral in (1.4.7) will not be zero. □

It is possible to prove a more general version of the fundamental lemma of the calculus of variations, by assuming α to be merely Lebesgue integrable; in that case, the thesis says that α vanishes almost everywhere. This version is a little harder to prove and we are not going to need it.

As we have seen, if q is a critical point of S_L then the integral in (1.4.5) vanishes for every smooth $v : [a, b] \to \mathbb{R}^n$ such that $v(a) = v(b) = 0$ and thus, by the fundamental lemma of the calculus of variations, it follows that the Euler–Lagrange equation is satisfied. We have just proven:

1.4.4. THEOREM. *A smooth curve $q : [a, b] \to \mathbb{R}^n$ is a critical point of S_L if and only if it satisfies the Euler–Lagrange equation* (1.4.6). □

The lefthand side of the Euler–Lagrange equation (1.4.6) should *not* be confused with $\frac{\partial^2 L}{\partial t \partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)$ (for instance, such expression is automatically zero if L does not depend on t). The lefthand side of the Euler–Lagrange equation is the derivative of the map $t \mapsto \frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big)$.

1.4.5. REMARK. We have worked so far with a Lagrangian whose domain is $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$. We could have used a Lagrangian whose domain is an open subset of $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ instead. Some obvious adaptations would have to be

[9]Some authors define critical points using exclusively variations of this type. This is not a good idea, since this breaks the manifest invariance under general diffeomorphisms of the notion of critical point!

done in our definitions. For instance, the domain of the action functional S_L would consist only of curves q such that $\big(t, q(t), \dot{q}(t)\big)$ is in the domain of L, for all $t \in [a,b]$. Also, in the definition of critical point, one should only consider curves $q : [a,b] \to \mathbb{R}^n$ that are in the domain of S_L and only variations $(q_s)_{s\in I}$ of q such that q_s is in the domain of S_L, for all $s \in I$. Observe that if $(q_s)_{s\in I}$ is an arbitrary variation of a curve q that is in the domain of S_L then there exists an open interval $I' \subset I$ containing the origin such that q_s is in the domain of S_L, for all $s \in I'$. Namely, the set of pairs $(s,t) \in I \times [a,b]$ such that $\big(t, q_s(t), \frac{\mathrm{d}}{\mathrm{d}t}\, q_s(t)\big)$ is in the domain of L is open in $I \times [a,b]$ and contains $\{0\} \times [a,b]$; since $[a,b]$ is compact, it follows that such set contains $I' \times [a,b]$ for some open interval $I' \subset I$ containing the origin. Obviously, Theorem 1.4.4 also holds for a Lagrangian defined in an open subset of $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$.

1.4.6. Remark. We have actually proved a little more than what is stated in Theorem 1.4.4. A variation $(q_s)_{s\in I}$ is said to have *compact support* if the map:

$$I \ni s \longmapsto q_s(t) \in \mathbb{R}^n$$

is constant for t in a neighborhood of a in $[a,b]$ and for t in a neighborhood of b in $[a,b]$. Our proof of Theorem 1.4.4 has actually shown that if (1.4.2) holds for all variations of q having compact support then q satisfies the Euler–Lagrange equation.

1.4.1. Variational formulation of Classical Mechanics. Now it is easy to see that when the force $F = (F_1, \ldots, F_n)$ admits a potential V (see Subsection 1.2.1) then the differential equation (1.2.1) defining the dynamics of Classical Mechanics is precisely the Euler–Lagrange equation of the Lagrangian $L : \mathrm{dom}(V) \times (\mathbb{R}^3)^n \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n \to \mathbb{R}$ defined by:

$$L(t, q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) = \sum_{j=1}^{n} \frac{1}{2}\, m_j \|\dot{q}_j\|^2 - V(t, q_1, \ldots, q_n). \tag{1.4.8}$$

Namely, the Euler–Lagrange equation for L can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_j}\big(t, q(t), \dot{q}(t)\big) = \frac{\partial L}{\partial q_j}\big(t, q(t), \dot{q}(t)\big), \quad j = 1, \ldots, n,$$

and (identifying $\mathbb{R}^3$ with its dual space):

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_j}\big(t, q(t), \dot{q}(t)\big) = \frac{\mathrm{d}}{\mathrm{d}t}\big(m_j \dot{q}_j(t)\big) = m_j \frac{\mathrm{d}^2 q_j}{\mathrm{d}t^2}(t),$$

$$\frac{\partial L}{\partial q_j}\big(t, q(t), \dot{q}(t)\big) = -\nabla_{q_j} V\big(t, q(t)\big) = F_j\big(t, q(t)\big).$$

The Lagrangian (1.4.8) motivates the following definition.

1.4.7. Definition. We call $\frac{1}{2}\, m_j \|\dot{q}_j(t)\|^2$ the *kinetic energy* of the j-th particle at the instant t and the sum $\sum_{j=1}^{n} \frac{1}{2}\, m_j \|\dot{q}_j(t)\|^2$ appearing in (1.4.8) the *total kinetic energy* at the instant t. We also call $V\big(t, q(t)\big)$ the *potential energy* at the instant t.

It is common to denote the total kinetic energy by T, so that the Lagrangian becomes $L = T - V$. Notice that we can define a kinetic energy for just one particle, but not a potential energy for just one particle.

What is so good about rewriting the dynamics of Classical Mechanics in terms of an Euler–Lagrange equation? One motivation is that the Euler–Lagrange equation is good for doing arbitrary transformations of coordinates (even time-dependent transformations of coordinates[10]), i.e., after a transformation of coordinates the Euler–Lagrange equation is transformed into the Euler–Lagrange equation of a transformed Lagrangian (what we mean exactly by this will become clearer when we do variational calculus on manifolds). On the other hand, equation (1.2.1) isn't good for arbitrary transformations of coordinates (for instance, under a general transformation of coordinates the lefthand side of (1.2.1) gets a term with a first order derivative of q). Also, essentially all equations in Physics (the field equations of Electromagnetism, of General Relativity and of the Gauge theories of the Standard Model, for instance) can be obtained from variational problems. So, one usually talks about "the Lagrangian" of the theory. We will also see that Lagrangians (and *Hamiltonians*) are essential for doing Quantum Theory.

1.4.8. Remark. The Lorentz force (exerted by a magnetic field upon a charged particle) is an important example of a force which depends on the velocity of the particle and therefore it is not a force with a potential (in the sense defined in Subsection 1.2.1). Nevertheless, as we will see later, it is possible to handle it using a Lagrangian.

1.5. Lagrangians on manifolds

The variational problem discussed in Section 1.4 can be straightforwardly reformulated for curves on manifolds. One can then prove that the critical points of the action functional are again the curves which satisfy the Euler–Lagrange equation. But, in a manifold, in order to give meaning to the statement that a curve satisfies the Euler–Lagrange equation, one has to use a coordinate chart; it turns out that such statement does not depend on the choice of the coordinate chart. There are some minor technical complications that arise in the context of manifolds and for the reader's convenience we give all details. We won't try to give a manifestly coordinate independent formulation of the Euler–Lagrange equation; attempts at doing so tend to get nasty and it happens that for an important class of Lagrangians (*hyper-regular* Lagrangians), the Euler–Lagrange equation

[10] One could be willing to use, say, spherical coordinates based on an orthonormal basis that rotates!

is equivalent to *Hamilton's equations* (which we will present later on) and those are easily formulated in a manifestly coordinate independent manner.

Let Q be a differentiable manifold. Given an interval $[a, b]$, we denote by $C^\infty\big([a, b], Q\big)$ the set of all smooth maps $q : [a, b] \to Q$ and, given points $q_a, q_b \in Q$, we denote by $C^\infty_{q_a q_b}\big([a, b], Q\big)$ the subset of $C^\infty\big([a, b], Q\big)$ consisting of maps q with $q(a) = q_a$, $q(b) = q_b$. In order to define an action functional $S_L : C^\infty_{q_a q_b}\big([a, b], Q\big) \to \mathbb{R}$ of the form:

$$S_L(q) = \int_a^b L\big(t, q(t), \dot{q}(t)\big)\,\mathrm{d}t, \tag{1.5.1}$$

the appropriate domain for the map L is the product $\mathbb{R} \times TQ$, where TQ denotes the tangent bundle of Q. So, we define a *Lagrangian* on the manifold Q to be a smooth map:

$$L : \mathbb{R} \times TQ \longrightarrow \mathbb{R}$$

and the map S_L defined above is called the corresponding *action functional*. As before, we should point out that in fact there is a family of action functionals for a given Lagrangian L (one for each interval $[a, b]$ and for each choice of points $q_a, q_b \in Q$), but as before we will be a little sloppy and use the same name and notation for all of them. Also, as before (see Remark 1.4.5), it is possible to work with a map L whose domain is some open subset of $\mathbb{R} \times TQ$ and in order to do that there are obvious adaptations in the definitions and proofs that follow; we don't want to distract the reader with things like that. In order to avoid awkward moments in the future, let us make a warning about notation: if $q : [a, b] \to Q$ is a differentiable curve in a manifold then for each $t \in [a, b]$ the derivative $\dot{q}(t)$ of q at t is an element of the tangent space $T_{q(t)}Q$ which is a subset of the tangent bundle TQ. So, strictly speaking, the notation in (1.5.1) is wrong; it is $\big(t, \dot{q}(t)\big)$, and not $\big(t, q(t), \dot{q}(t)\big)$, which is a point of the domain $\mathbb{R} \times TQ$ of the Lagrangian L. Nevertheless, we find it convenient in many cases to write elements of the tangent bundle TQ as ordered pairs consisting of a point of Q and a tangent vector at that point. It happens that typical constructions of the tangent space of a manifold at a point have the property that tangent spaces at distinct points are disjoint, so that the set TQ can be taken to be literally the union of all tangent spaces. But when Q is a submanifold of $\mathbb{R}^n$ one identifies the tangent space T_qQ at a point $q \in Q$ with a subspace of $\mathbb{R}^n$, and such subspaces of $\mathbb{R}^n$ are not disjoint and thus one is forced to take TQ to be the *disjoint* union $TQ = \bigcup_{q\in Q} \big(\{q\} \times T_qQ\big)$; in that case, one *has to* write elements of the tangent bundle as ordered pairs (a point and a vector). So, for reasons of uniformity, we find it convenient to do so also when Q is a general manifold.

Definitions 1.4.1 and 1.4.2 can be readily adapted to the present context. Given a smooth curve $q : [a, b] \to Q$, we define a *variation* of q to be a family $(q_s)_{s\in I}$ of smooth curves $q_s : [a, b] \to Q$, where $I \subset \mathbb{R}$ is an open interval with $0 \in I$, q_0 equals q and the map:

$$I \times [a,b] \ni (s,t) \longmapsto q_s(t) \in Q$$

is smooth. If the maps $s \mapsto q_s(a)$, $s \mapsto q_s(b)$ are constant, we say that the variation $(q_s)_{s\in I}$ has *fixed endpoints*. The *variational vector field* of a variation $(q_s)_{s\in I}$ is defined again by formula (1.4.1), but now v is a smooth map from $[a,b]$ to the tangent bundle TQ such that $v(t) \in T_{q(t)}Q$ for all $t \in [a,b]$, i.e., v is a vector field along the curve q. Again, for a variation with fixed endpoints, the variational vector field v satisfies $v(a) = v(b) = 0$. As before, we say that q is a *critical point* of S_L if (1.4.1) holds for every variation with fixed endpoints $(q_s)_{s\in I}$ of q.

If we were to do serious infinite-dimensional calculus here, we would have to show how to turn $C^\infty\big([a,b],Q\big)$ into an infinite-dimensional manifold (it is a Fréchet manifold and, by replacing C^∞ with C^k, for some fixed finite k, it would be a Banach manifold), how to identify the space of smooth vector fields along $q \in C^\infty\big([a,b],Q\big)$ with the tangent space $T_qC^\infty\big([a,b],Q\big)$, we would have to show that $C^\infty_{q_aq_b}\big([a,b],Q\big)$ is a submanifold of $C^\infty\big([a,b],Q\big)$ whose tangent space at a point q consists of the smooth vector fields v along q with $v(a) = v(b) = 0$ and we would have to show that the action functional S_L is smooth. All of that would require a considerable amount of work. The construction of the manifold structure of $C^\infty\big([a,b],Q\big)$ is completely standard, but not completely straightforward. Luckily, we don't have to worry about any of that here, as we won't be using any theorems from infinite-dimensional calculus.

If $L : \mathbb{R}\times TQ \to \mathbb{R}$ is a Lagrangian, $\widetilde{Q}$ is another differentiable manifold and $\varphi : Q \to \widetilde{Q}$ is a smooth diffeomorphism, then there is an obvious way to push L to the manifold $\widetilde{Q}$ using φ. Namely, we define a Lagrangian $L_\varphi : \mathbb{R}\times T\widetilde{Q} \to \mathbb{R}$ by requiring that the diagram:

$$\begin{tikzcd}
\mathbb{R}\times TQ \arrow[rr, "\mathrm{Id}\times \mathrm{d}\varphi"] \arrow[dr, "L"'] & & \mathbb{R}\times T\widetilde{Q} \arrow[dl, "L_\varphi"] \\
& \mathbb{R} &
\end{tikzcd}$$

be commutative, where Id denotes the identity map of $\mathbb{R}$ and $\mathrm{d}\varphi : TQ \to T\widetilde{Q}$ denotes the differential of the map φ (the map whose restriction to the tangent space T_qQ is the differential[11] $\mathrm{d}\varphi_q : T_qQ \to T_{\varphi(q)}\widetilde{Q}$, for all $q \in Q$). Obviously, if $q : [a,b] \to Q$ is a smooth map and if $\tilde{q} = \varphi\circ q : [a,b] \to \widetilde{Q}$ is

[11]Some authors prefer to use $T\varphi$ instead of $\mathrm{d}\varphi$; that is indeed the natural notation if one thinks of T as a functor that carries Q to TQ. Also, if φ is a map from $\mathbb{R}^m$ to $\mathbb{R}^n$ then there is some conflict of notation: in that context, $\mathrm{d}\varphi$ usually refers to the map that associates to each $x \in \mathbb{R}^m$ a linear map $\mathrm{d}\varphi_x$ from $\mathbb{R}^m$ to $\mathbb{R}^n$, while the map $d\varphi$ from the tangent bundle $T\mathbb{R}^m \cong \mathbb{R}^m \times \mathbb{R}^m$ to the tangent bundle $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ is not exactly that. Having said that, we will continue to use $\mathrm{d}\varphi$ instead of $T\varphi$.

the curve obtained by pushing q using φ, then:

$$L_\varphi\big(t,\tilde q(t),\dot{\tilde q}(t)\big)=L\big(t,q(t),\dot q(t)\big),$$

for all $t\in[a,b]$, so that:

$$S_L(q)=S_{L_\varphi}(\tilde q).$$

The equality $\tilde q_s=\varphi\circ q_s$ defines a bijection between variations with fixed endpoints $(q_s)_{s\in I}$ of q and variations with fixed endpoints $(\tilde q_s)_{s\in I}$ of $\tilde q$ and then, since $S_L(q_s)=S_{L_\varphi}(\tilde q_s)$ for all $s\in I$, it follows that q is a critical point of S_L if and only if $\tilde q$ is a critical point of S_{L_φ}. What we have just observed is pretty obvious: any definition that makes sense for manifolds must give rise to a concept that is invariant under diffeomorphisms (just like, say, any definition that makes sense for groups must give rise to a concept that is invariant under group isomorphisms, and so on).

An important particular case of the construction of pushing a Lagrangian using a smooth diffeomorphism is this: consider a local chart $\varphi:U\to\widetilde U$, where U is an open subset of Q and $\widetilde U$ is an open subset of $\mathbb{R}^n$. We can push the Lagrangian L (more precisely, the restriction of L to the open subset $\mathbb{R}\times TU$ of $\mathbb{R}\times TQ$) using φ obtaining a Lagrangian:

$$L_\varphi:\mathbb{R}\times T\widetilde U=\mathbb{R}\times\widetilde U\times\mathbb{R}^n\longrightarrow\mathbb{R}$$

on the manifold $\widetilde U$. We call L_φ the *representation of L with respect to the chart φ*. A smooth curve $q:[a,b]\to Q$ with $q\big([a,b]\big)\subset U$ is a critical point of S_L if and only if the curve $\tilde q=\varphi\circ q$ is a critical point of S_{L_φ}. Actually, in order to get to that conclusion, there is a minor detail one should pay attention to: φ only induces a bijection between variations of q *that stay inside of* U and variations of $\tilde q$ that stay inside of $\widetilde U$. But if the image of q is contained in U then, for any variation $(q_s)_{s\in I}$ of q, we have that the image of q_s is contained in U for s in some open interval $I'\subset I$ containing the origin (see the argument that appears in Remark 1.4.5). Thus, for the definition of critical point, it doesn't make any difference to consider only variations of the curve that stay inside some given open set containing the image of the curve. By Theorem 1.4.4, the curve $\tilde q$ is a critical point of S_{L_φ} if and only if[12]:

$$\frac{\mathrm d}{\mathrm dt}\frac{\partial L_\varphi}{\partial\dot q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big)=\frac{\partial L_\varphi}{\partial q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big),\tag{1.5.2}$$

for all $t\in[a,b]$. When (1.5.2) holds for all $t\in[a,b]$ (where $\tilde q=\varphi\circ q$), we say that the curve q *satisfies the Euler–Lagrange equation with respect to the chart* φ. We have shown that a smooth curve with image contained in the domain of a chart is a critical point of S_L if and only if it satisfies the Euler–Lagrange equation with respect to that chart. It follows that, if φ_1, φ_2 are local charts whose domains contain the image of a smooth curve

[12]We denote by $\frac{\partial L_\varphi}{\partial q}$ and $\frac{\partial L_\varphi}{\partial\dot q}$ the derivatives of L_φ with respect to its second and third variables, respectively, even though $(t,q,\dot q)$ is the typical name of a point of $\mathbb{R}\times TQ$, not of a point of the domain of L_φ.

q then q satisfies the Euler–Lagrange equation with respect to φ_1 if and only if q satisfies the Euler–Lagrange equation with respect to φ_2. That could also be checked by a direct computation using the transition function $\varphi_2 \circ \varphi_1^{-1}$ (see Exercises 1.18 and 1.19). But we have just proven that fact and no such computation was necessary[13]; that is because being a solution of the Euler–Lagrange equation is equivalent to being a critical point of the action functional and the latter notion is manifestly invariant under diffeomorphisms!

What we have done above is to prove a version of Theorem 1.4.4 for curves on manifolds, but only for curves whose image is contained in the domain of a local chart! Now we have to handle arbitrary curves. Consider a smooth curve $q : [a, b] \to Q$ and a local chart $\varphi : U \to \widetilde{U}$ (whose domain does not necessarily contain the image of q). We will say that *q satisfies the Euler–Lagrange equation with respect to φ* if (1.5.2) holds for all $t \in q^{-1}(U)$, where:

$$\tilde{q} = \varphi \circ q|_{q^{-1}(U)}. \tag{1.5.3}$$

Notice that $q^{-1}(U)$ is an open subset of $[a, b]$ (open with respect to $[a, b]$, of course). When the image of q does not intercept U then $q^{-1}(U)$ is the empty set and the condition that q satisfy the Euler–Lagrange equation with respect to φ is vacuously satisfied. The right generalization of Theorem 1.4.4 is:

1.5.1. THEOREM. *Let $L : \mathbb{R}\times TQ \to \mathbb{R}$ be a Lagrangian and $q : [a, b] \to Q$ be a smooth curve. Then the following statements are equivalent:*

(a) *q is a critical point of S_L;*
(b) *for every local chart φ on Q, q satisfies the Euler–Lagrange equation with respect to φ;*
(c) *there exists a family of local charts on Q, whose domains cover the image of q, such that q satisfies the Euler–Lagrange equation with respect to any chart belonging to that family.*

PROOF. The implication (b)⇒(c) is obvious. The implications (a)⇒(b) and (c)⇒(a) will be proven in Lemmas 1.5.2 and 1.5.6 below. □

The fact that (a)⇒(b) is useful when one knows that q is a critical point of S_L; then you can choose whatever local chart you like and you know that q satisfies the Euler–Lagrange equation with respect to that chart. The fact that (c)⇒(a), on the other hand, is useful when you want to check that q is a critical point of S_L; in that case, you can choose your favorite set of charts (as long as their domains are able to cover the image of q) for checking that q satisfies the Euler–Lagrange equation with respect to them.

[13]Suppose that a curve q satisfies the Euler–Lagrange equation with respect to a chart φ_1 just at *one given instant* $t \in [a, b]$. Is it true that it satisfies the Euler–Lagrange equation with respect to a different chart φ_2 at that same instant? The argument that we have just presented does not allow us to conclude that. But that is indeed true and it follows from the result of Exercise 1.19.

Let's get to the proof of (a)$\Rightarrow$(b).

1.5.2. LEMMA. *Condition (a) in the statement of Theorem 1.5.1 implies condition (b).*

PROOF. Let $\varphi : U \to \widetilde{U} \subset \mathbb{R}^n$ be a local chart on Q. Defining $\tilde{q}$ as in (1.5.3), it suffices to check that (1.5.2) holds for all $t \in [c, d]$, where $[c, d]$ is an arbitrary interval contained in $q^{-1}(U)$. Because of Remark 1.4.6, it suffices to prove that:

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_c^d L_\varphi\big(t, \tilde{q}_s(t), \dot{\tilde{q}}_s(t)\big)\,\mathrm{d}t \Big|_{s=0} = 0, \tag{1.5.4}$$

for any variation with *compact support*:

$$I \times [c, d] \ni (s, t) \longmapsto \tilde{q}_s(t) \in \mathbb{R}^n \tag{1.5.5}$$

of the curve $\tilde{q}|_{[c,d]}$. By replacing I with a smaller interval, we can assume that $\tilde{q}_s(t)$ belongs to $\widetilde{U}$, for all $s \in I$, $t \in [c, d]$. Set:

$$q_s(t) = \varphi^{-1}\big(\tilde{q}_s(t)\big),$$

for $s \in I$, $t \in [c, d]$ and $q_s(t) = q(t)$, for $s \in I$, $t \in [a, b] \setminus [c, d]$. The fact that the variation (1.5.5) has compact support implies easily that the map:

$$I \times [a, b] \ni (s, t) \longmapsto q_s(t) \in Q$$

is smooth and therefore it is a variation with fixed endpoints of q. Since q is a critical point of S_L, we have:

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_a^b L\big(t, q_s(t), \dot{q}_s(t)\big)\,\mathrm{d}t \Big|_{s=0} = 0. \tag{1.5.6}$$

The difference:

$$\int_a^b L\big(t, q_s(t), \dot{q}_s(t)\big)\,\mathrm{d}t - \int_c^d L\big(t, q_s(t), \dot{q}_s(t)\big)\,\mathrm{d}t$$

is independent of $s \in I$ and since:

$$\int_c^d L\big(t, q_s(t), \dot{q}_s(t)\big)\,\mathrm{d}t = \int_c^d L_\varphi\big(t, \tilde{q}_s(t), \dot{\tilde{q}}_s(t)\big)\,\mathrm{d}t$$

for all $s \in I$, it follows from (1.5.6) that (1.5.4) holds. □

The implication (c)$\Rightarrow$(a) is a bit trickier. The plan is to show that the derivative of $s \mapsto S_L(q_s)$ at $s = 0$ is zero for variations $(q_s)_{s\in I}$ whose variational vector field v has small support. Then we have to show that this suffices for establishing that q is a critical point of S_L. That happens because the derivative of $s \mapsto S_L(q_s)$ defines a linear function of the variational vector field v (our next lemma) and because variational vector fields with small support span all smooth vector fields v along q with $v(a) = v(b) = 0$. Curiously, we won't need to show that every smooth vector field v along q with $v(a) = v(b) = 0$ is the variational vector field of some variation with fixed endpoints. That wouldn't be hard to do (but it isn't as straightforward as when the manifold is $\mathbb{R}^n$); one could, for instance, use the exponential map of some arbitrary Riemannian metric of Q to define a variation of q by $q_s(t) = \exp_{q(t)}\big(sv(t)\big)$ (or one could embed Q in $\mathbb{R}^N$, using Whitney's theorem, then construct the desired variation in $\mathbb{R}^N$, and then retract it back to Q using a tubular neighborhood). But we simply don't need to prove that.

1.5.3. LEMMA. *Let $q : [a, b] \to Q$ be a smooth curve. There exists a real valued linear map D defined in the space of all smooth vector fields along q such that:*

$$D(v) = \frac{\mathrm{d}}{\mathrm{d}s} S_L(q_s)\Big|_{s=0},$$

for any variation $(q_s)_{s\in I}$ of q, where v denotes the variational vector field.

PROOF. If the image of q is contained in the domain of a local chart $\varphi : U \to \widetilde{U} \subset \mathbb{R}^n$ then one can simply define D by:

$$D(v) = \int_a^b \frac{\partial L_\varphi}{\partial q}\big(t, \tilde{q}(t), \dot{\tilde{q}}(t)\big)\tilde{v}(t) + \frac{\partial L_\varphi}{\partial \dot{q}}\big(t, \tilde{q}(t), \dot{\tilde{q}}(t)\big)\dot{\tilde{v}}(t)\,\mathrm{d}t,$$

where $\tilde{q} = \varphi \circ q$, $\tilde{v}(t) = \mathrm{d}\varphi_{q(t)}\big(v(t)\big)$ and $t \in [a, b]$. For the general case, choose a partition $a = t_0 < t_1 < \cdots < t_k = b$ of $[a, b]$ such that $q\big([t_i, t_{i+1}]\big)$ is contained in the domain of some chart, for all i. Then (using the same symbol S_L to denote the action functional for curves defined on the smaller intervals):

$$S_L(q_s) = \sum_{i=0}^{k-1} S_L(q_s|_{[t_i,t_{i+1}]}), \quad s \in I,$$

and if D_i is a linear map that satisfies the thesis of the lemma for the restricted curve $q|_{[t_i,t_{i+1}]}$, we simply define D by setting:

$$D(v) = \sum_{i=0}^{k-1} D_i(v|_{[t_i,t_{i+1}]}). \qquad \square$$

Now we show that smooth vector fields with small support span all smooth vector fields that vanish at the endpoints.

1.5.4. LEMMA. *Let $q : [a, b] \to Q$ be a smooth curve and let $\mathcal{U}$ be an open cover of the interval $[a, b]$. Every smooth vector field v along q with $v(a) = v(b) = 0$ can be written as a finite sum $\sum_{i=1}^k v_i$ of smooth vector fields v_i along q with $v_i(a) = v_i(b) = 0$ and such that the support of v_i is contained in some element of $\mathcal{U}$.*

PROOF. Replace $\mathcal{U}$ with a finite subcover $\{U_1, \ldots, U_k\}$ and consider a smooth partition of unit subordinated to it, i.e., smooth maps $\xi_i : [a, b] \to \mathbb{R}$, $i = 1, \ldots, k$, with $\sum_{i=1}^k \xi_i = 1$, such that the support of ξ_i is contained in U_i, for all i. Set $v_i = v\xi_i$, $i = 1, \ldots, k$. $\square$

Now we check that, under condition (c) in the statement of Theorem 1.5.1, the derivative of $s \mapsto S_L(q_s)$ at $s = 0$ vanishes when the variational vector field has small support.

1.5.5. LEMMA. *Let $q : [a, b] \to Q$ be a smooth curve and let $\varphi : U \to \widetilde{U}$ be a local chart such that q satisfies the Euler–Lagrange equation with respect to φ. If v is a smooth vector field along q such that $v(a) = v(b) = 0$ and such that the support of v is contained in an interval $[c, d]$ contained in $q^{-1}(U)$ then $D(v) = 0$, where D is a linear map satisfying the condition in the statement of Lemma 1.5.3.*

PROOF. We define a variation $(q_s)_{s\in I}$ of q as follows: if t is in $[a, b]\setminus[c, d]$, we set $q_s(t) = q(t)$, for all $s \in I$. For $t \in [c, d]$, we set:

$$q_s(t) = \varphi^{-1}\big(\tilde{q}(t) + s\tilde{v}(t)\big), \tag{1.5.7}$$

for all $s \in I$, where $\tilde{q}(t) = \varphi\big(q(t)\big)$, $\tilde{v}(t) = \mathrm{d}\varphi_{q(t)}\big(v(t)\big)$ and I is chosen small enough so that $\tilde{q}(t) + s\tilde{v}(t)$ is in $\widetilde{U}$, for all $s \in I$, $t \in [c, d]$. The map $(s, t) \mapsto q_s(t)$ is smooth in $I \times [a, b]$ because equality (1.5.7) actually holds for t in the neighborhood $q^{-1}(U)$ of $[c, d]$. Thus $(q_s)_{s\in I}$ is a variation of q with variational vector field v, so that:

$$D(v) = \frac{\mathrm{d}}{\mathrm{d}s} S_L(q_s)\Big|_{s=0}.$$

Arguing as in the proof of Lemma 1.5.2, we see that:

$$\frac{\mathrm{d}}{\mathrm{d}s} S_L(q_s)\Big|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}s}\int_c^d L_\varphi\big(t, \tilde{q}_s(t), \dot{\tilde{q}}_s(t)\big)\,\mathrm{d}t\,\Big|_{s=0}, \tag{1.5.8}$$

where $\tilde{q}_s(t) = \varphi\big(q_s(t)\big) = \tilde{q}(t) + s\tilde{v}(t)$, for $t \in [c, d]$, $s \in I$. Since:

$$I \times [c, d] \ni (s, t) \longmapsto \tilde{q}_s(t)$$

is a variation with fixed endpoints of $\tilde{q}|_{[c,d]}$ and since $\tilde{q}|_{[c,d]}$ is a solution of the Euler–Lagrange equation (1.5.2), it follows that the righthand side of (1.5.8) vanishes. This concludes the proof. □

1.5.6. LEMMA. *Condition (c) in the statement of Theorem 1.5.1 implies condition (a).*

PROOF. Let $\mathcal{U}$ be the set of all intervals J that are open in $[a, b]$ and such that $q(J)$ is contained in the domain of one of the charts belonging to the family whose existence is assumed in condition (c). By Lemma 1.5.5, $D(v) = 0$ if $v(a) = v(b) = 0$ and the support of v is contained in some $J \in \mathcal{U}$ (obviously, if the support of v is contained in J then it is contained in some closed interval $[c, d]$ contained in J). By Lemma 1.5.4 and by the linearity of D, it follows that $D(v) = 0$ for any v with $v(a) = v(b) = 0$, proving that q is a critical point of S_L. □

Assume now that the manifold Q is a submanifold of $\mathbb{R}^n$ (we have been using n to denote the dimension of Q, but, obviously, now we are not) and assume that we have a Lagrangian $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ on $\mathbb{R}^n$. Since Q is a submanifold of $\mathbb{R}^n$, TQ is naturally identified with a submanifold of $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ and thus we can restrict L to $\mathbb{R} \times TQ$. What are the critical points of the action functional of the restriction of L? The following proposition answers that.

1.5.7. PROPOSITION. *Let L be a Lagrangian on $\mathbb{R}^n$, Q be a submanifold of $\mathbb{R}^n$ and L^{restr} denote the Lagrangian on Q obtained by restricting L to $\mathbb{R} \times TQ$. A smooth curve $q : [a, b] \to Q$ is a critical point of the action functional $S_{L^{\mathrm{restr}}}$ if and only if:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q}\big(t, q(t), \dot q(t)\big) - \frac{\partial L}{\partial q}\big(t, q(t), \dot q(t)\big) \in (T_{q(t)}Q)^{\mathrm{o}},$$

for all $t \in [a, b]$, where $(T_{q(t)}Q)^{\mathrm{o}} \subset \mathbb{R}^{n*}$ denotes the annihilator of the tangent space $T_{q(t)}Q \subset \mathbb{R}^n$ (if one identifies the dual space $\mathbb{R}^{n*}$ with $\mathbb{R}^n$ in the usual way then then annihilator $(T_{q(t)}Q)^{\mathrm{o}}$ is identified with the orthogonal complement $(T_{q(t)}Q)^{\perp}$).

PROOF. We leave details to the reader as an exercise. The hint is *don't* (really, don't!) use Theorem 1.5.1 and the representation of L^{restr} with respect to a local chart on Q. It is much easier to use directly the definition of critical point, observing that $S_{L^{\mathrm{restr}}}(q_s) = S_L(q_s)$, for any variation $(q_s)_{s\in I}$ of q in Q. The derivative $\frac{\mathrm{d}}{\mathrm{d}s}S_L(q_s)|_{s=0}$ can be computed exactly as in Section 1.4. The only difference from what was done there is that now we are going obtain that q is a critical point of $S_{L^{\mathrm{restr}}}$ if and only if the righthand side of (1.4.5) vanishes for smooth maps $v : [a, b] \to \mathbb{R}^n$ such that $v(a) = v(b) = 0$ *and* $v(t) \in T_{q(t)}Q$, for all $t \in [a, b]$. You will then need a generalization of the Fundamental Lemma of the Calculus of Variations (which is stated as Exercise 1.17) to conclude the proof. If you prefer not to use the fact that any v is the variational vector field of a variation of q in Q, notice that in order to conclude the proof it is enough to consider vector fields v along q having small support (contained in $q^{-1}(U)$, where $U \subset Q$ is the domain of a local chart) and for those one straightforwardly constructs a variation of q using a local chart on Q. $\square$

1.5.1. Back to Classical Mechanics. Constraints. Are the variational problems on manifolds discussed in this section of any use to Classical Mechanics? They are. This stuff is useful when we work on a problem in which there is something which constrains the motion of the particles: a track from which a particle cannot get away, a string or a rigid bar connecting two particles, etc. Of course, at the fundamental level, there are no such constraints; tracks, strings and bars are themselves made out of particles. The problem of n constrained particles which we are going to discuss in this subsection would, at the fundamental level, be a problem of $N > n$ free particles, interacting through the fundamental forces of Classical Mechanics (the electrical and the gravitational forces[14]).

Let us consider a system of n particles subject to forces $F = (F_1, \dots, F_n)$ that admit a potential V. We then know that their trajectories satisfy the Euler–Lagrange equation of the Lagrangian L defined in (1.4.8). Now, assume that apart from the forces encoded in the potential V, there are also additional forces acting on the particles, in such a way that the configuration $\big(q_1(t), \dots, q_n(t)\big) \in (\mathbb{R}^3)^n$ must stay inside a subset Q of $(\mathbb{R}^3)^n$, for all t. Such additional forces will be called the *forces from the constraint*. We will assume that the subset Q is a smooth submanifold of $(\mathbb{R}^3)^n$.

The type of constraint that we are considering here are constraints on the *positions* of the particles (known as *holonomic* constraints). One could also consider constraints formulated in terms of the *velocities* of the particles that cannot be reduced to constraints on the positions alone (*non holonomic* constraints); we are not going to consider those in this course. The constraints under consideration in this subsection are also being assumed to be independent of time (otherwise, we should consider a submanifold Q_t of $(\mathbb{R}^3)^n$, for each t).

Let us try to take the Lagrangian L defined in (1.4.8) and consider its restriction to a Lagrangian L^{cons} on the submanifold Q. What are the critical points of the action functional? The answer is obtained from Proposition 1.5.7. A curve $q = (q_1, \dots, q_n) : [a, b] \to Q \subset (\mathbb{R}^3)^n$ is a critical point of the action functional $S_{L^{\mathrm{cons}}}$ if and only if the vector:

$$(1.5.9) \quad \Big(m_j \frac{\mathrm{d}^2 q_j}{\mathrm{d}t^2}(t) + \nabla_{q_j} V\big(t, q(t)\big)\Big)_{j=1,\dots,n} = \Big(m_j \frac{\mathrm{d}^2 q_j}{\mathrm{d}t^2}(t) - F_j\big(t, q(t)\big)\Big)_{j=1,\dots,n} \in (\mathbb{R}^3)^n$$

is orthogonal[15] to the tangent space $T_{q(t)}Q \subset (\mathbb{R}^3)^n$, for all $t \in [a, b]$. We can reformulate this by saying that:

$$(1.5.10) \quad m_j \frac{\mathrm{d}^2 q_j}{\mathrm{d}t^2}(t) = -\nabla_{q_j} V\big(t, q(t)\big) + R_j(t) = F_j\big(t, q(t)\big) + R_j(t), \quad j = 1, \dots, n,$$

for all $t \in [a, b]$, for some map $R = (R_1, \dots, R_n) : [a, b] \to (\mathbb{R}^3)^n$ such that $R(t) \in (\mathbb{R}^3)^n$ is orthogonal to the tangent space $T_{q(t)}Q$, for all $t \in [a, b]$. If $q = (q_1, \dots, q_n)$ are the actual trajectories of the particles, then (1.5.10) holds if and only if the vector $R_j(t) \in \mathbb{R}^3$ is the force exerted upon the j-th particle by the constraint at the instant t.

[14] As noted earlier, this cannot be taken too seriously, as we know that Classical Mechanics is not a fundamental theory.

[15] Orthogonality here means orthogonality with respect to the standard inner product of $(\mathbb{R}^3)^n$. We observe that sometimes it is useful to consider an inner product on $(\mathbb{R}^3)^n$ that is scaled by the masses of the particles. See Exercise 1.22 for details.

Of course, there is no *logical* reason why the critical points of the action functional $S_{L^{\rm cons}}$ should correspond to the trajectories of n particles subject to the forces encoded in the potential V and to the forces from the constraint. That isn't a logical consequence of the dynamics of Classical Mechanics, as presented in Section 1.2. What we have shown is that it is true that the critical points of the action functional $S_{L^{\rm cons}}$ correspond to the trajectories of the n constrained particles if and only if the forces $R_j(t)$ exerted by the constraint upon the particles constitute a vector $R(t) = \big(R_1(t), \ldots, R_n(t)\big)$ that is orthogonal to the tangent space $T_{q(t)}Q$, for all t. Is the orthogonality between $R(t)$ and $T_{q(t)}Q$ a reasonable assumption? That can only be judged within the context of a specific example. When $n = 1$, so that the manifold Q corresponds to a curve or a surface inside physical space from which the particle cannot get away, then the assumption of orthogonality between the vector $R(t)$ (in $\mathbb{R}^3$) and the tangent space $T_{q(t)}Q$ is easy to be physically interpreted. It means that, whatever keeps the particle inside Q, does that by exerting upon the particle a force that is orthogonal to Q; such condition can usually be understood as the absence of *friction* between the particle and the surface. For the general case, the orthogonality between $R(t)$ and the vectors in $T_{q(t)}Q$ is an orthogonality between vectors in *configuration space* $(\mathbb{R}^3)^n$, not an orthogonality between vectors in physical space, so it is hard to understand the physical meaning of such orthogonality condition. In Exercises 1.20 and 1.21 we ask the reader to work with concrete examples (the *double pendulum* and the *double spherical pendulum*) and to discover what the orthogonality condition means in those cases.

1.6. Hamiltonian formalism

The ordinary differential equation (1.2.1) defining the dynamics of Classical Mechanics is of second order. Everyone knows that a second order ordinary differential equation over a certain space can be reformulated in terms of a first order ordinary differential equation over a new space whose dimension is the double of the dimension of the original space. Such reformulation has some advantages. For an ordinary differential equation of first order the initial condition that determines the solution is just the value of the solution at one given initial instant and thus one can see the equation as defining a family of maps (the flow) from a space to itself. One can then ask questions such as "is there any interesting structure that is invariant by that flow?". One obvious way of reformulating the differential equation that defines the dynamics of Classical Mechanics as a first order equation is to introduce a new independent variable for $\dot q(t)$. It turns out that this is not the best idea; a small modification of that idea yields much better results.

1.6.1. Definition. The *momentum* of the j-th particle at time $t \in \mathbb{R}$ is defined by:

$$p_j(t) = m_j \dot q_j(t). \tag{1.6.1}$$

Notice that (by (1.2.1)) the derivative of p_j at an instant t is equal to the total force F_j acting upon the j-th particle at that instant. Assuming (1.2.3) it follows immediately that the *total momentum*:

$$\sum_{j=1}^{n} p_j(t)$$

is independent of t, i.e., it is *conserved*. Condition (1.2.3) holds if we don't have external forces (either we are dealing with the entire universe or with a system which is almost isolated, so that external forces are neglected) and if the force laws satisfy Newton's law of reciprocal actions (1.2.2) (which holds for the fundamental forces of Classical Mechanics, i.e., the gravitational and the electrical forces). In the presence of external forces (and assuming Newton's law of reciprocal actions) the derivative of the total momentum is equal to the sum of the external forces; therefore, even when the total momentum of a system is not conserved due to external forces, it is not hard to keep track of its evolution, since we can ignore the internal forces which are often the ones which are complicated to handle (imagine keeping track of the internal forces among all particles of some macroscopic system!). Later on we will take a closer look at the subject of conservation laws and we will understand the relationship between those and *symmetry*; for now, that is all that must be said about the conservation of the total momentum. We are just trying to give some motivation for Definition 1.6.1, so that it doesn't look like we have given a name for some arbitrary formula involving the particle trajectories q_j.

The maps $q_j : \mathbb{R} \to \mathbb{R}^3$, $p_j : \mathbb{R} \to \mathbb{R}^3$, $j = 1, \ldots, n$, define a curve:

$$(q, p) = (q_1, \ldots, q_n, p_1, \ldots, p_n) : \mathbb{R} \longrightarrow (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n.$$

If the maps q_j satisfy (1.2.1) and the maps p_j are defined by (1.6.1) then the curve (q, p) satisfies the first order differential equation:

$$\frac{\mathrm{d}q_j}{\mathrm{d}t}(t) = \frac{p_j(t)}{m_j}, \quad \frac{\mathrm{d}p_j}{\mathrm{d}t}(t) = F_j, \qquad j = 1, \ldots, n, \tag{1.6.2}$$

where, in the second equation, F_j is understood to be evaluated at the point:

$$\big(t, q_1(t), \ldots, q_n(t), \tfrac{1}{m_1}\, p_1(t), \ldots, \tfrac{1}{m_n}\, p_n(t)\big).$$

Conversely, if the curve (q, p) satisfies (1.6.2) then the maps q_j satisfy (1.2.1) and the maps p_j are given by (1.6.1). Let us now assume that the force F admits a potential $V : \mathrm{dom}(V) \subset \mathbb{R}\times(\mathbb{R}^3)^n \to \mathbb{R}$. Equation (1.6.2) becomes:

$$\frac{\mathrm{d}q_j}{\mathrm{d}t}(t) = \frac{p_j(t)}{m_j}, \quad \frac{\mathrm{d}p_j}{\mathrm{d}t}(t) = -\nabla_{q_j} V\big(t, q(t)\big), \qquad j = 1, \ldots, n. \tag{1.6.3}$$

Recall from Definition 1.4.7 the notions of kinetic and potential energy. The *total energy* (also known as *mechanical energy* or just *energy*) at the instant

t is defined to be the sum of the total kinetic energy with the potential energy:

$$\sum_{j=1}^{n} \frac{1}{2} m_j \|\dot{q}_j(t)\|^2 + V\big(t, q(t)\big). \tag{1.6.4}$$

The kinetic energy of the j-th particle can be written in terms of its momentum p_j:

$$\frac{1}{2} m_j \|\dot{q}_j(t)\|^2 = \frac{\|p_j(t)\|^2}{2m_j}$$

and so the total energy is equal to:

$$\sum_{j=1}^{n} \frac{\|p_j(t)\|^2}{2m_j} + V\big(t, q(t)\big).$$

Define a map $H : \mathrm{dom}(V) \times (\mathbb{R}^3)^n \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n \to \mathbb{R}$ by setting[16]:

$$H(t, q, p) = H(t, q_1, \ldots, q_n, p_1, \ldots, p_n) = \sum_{j=1}^{n} \frac{\|p_j\|^2}{2m_j} + V(t, q), \tag{1.6.5}$$

for all $(t, q) \in \mathrm{dom}(V) \subset \mathbb{R} \times (\mathbb{R}^3)^n$ and all $p \in (\mathbb{R}^3)^n$. The total energy at the instant t is then equal to $H\big(t, q(t), p(t)\big)$.

1.6.2. Definition. The map H defined in (1.6.5) is called the *Hamiltonian* of Classical Mechanics.

Notice that:

$$\frac{\partial H}{\partial q_j}(t, q, p) = \nabla_{q_j} V(t, q), \quad \frac{\partial H}{\partial p_j}(t, q, p) = \frac{p_j}{m_j}, \quad j = 1, \ldots, n, \tag{1.6.6}$$

where we have used the standard identification between $\mathbb{R}^3$ and the dual space $\mathbb{R}^{3^*}$ (notice that the lefthand sides of the equalities in (1.6.6) are elements of $\mathbb{R}^{3^*}$, while the righthand sides are elements of $\mathbb{R}^3$). It follows from (1.6.6) that equation (1.6.3) is equivalent to:

$$\frac{\mathrm{d}q}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p}\big(t, q(t), p(t)\big), \quad \frac{\mathrm{d}p}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q}\big(t, q(t), p(t)\big), \tag{1.6.7}$$

where now we have used the standard identification between $(\mathbb{R}^3)^n$ and the dual space $(\mathbb{R}^3)^{n*}$ (the lefthand sides of the equalities in (1.6.7) are elements of $(\mathbb{R}^3)^n$, while the righthand sides are elements of the dual space $(\mathbb{R}^3)^{n*}$). In this section we will continue to use such identification. Later, when we work with manifolds, we will discover that the more appropriate domain for H is an open subset of $\mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^{n*}$ and that $p(t)$ should be regarded as an element of $(\mathbb{R}^3)^{n*}$. Under such conditions, both sides of the first equation in (1.6.7) become elements of $(\mathbb{R}^3)^n$ (actually, the righthand

[16] We are using the symbols q, p for names of curves $t \mapsto q(t)$, $t \mapsto p(t)$ and also for names of points $q, p \in (\mathbb{R}^3)^n$. This type of notation abuse is very convenient and it hardly generates any misunderstandings.

side is an element of the *bidual* of $(\mathbb{R}^3)^n$, which is naturally identified with $(\mathbb{R}^3)^n$) and both sides of the second equation in (1.6.7) become elements of $(\mathbb{R}^3)^{n*}$. For now, insisting on not identifying $(\mathbb{R}^3)^n$ with $(\mathbb{R}^3)^{n*}$ would just be annoying.

1.6.3. DEFINITION. Equations (1.6.7) are called *Hamilton's equations.* The space $(\mathbb{R}^3)^n \times (\mathbb{R}^3)^n$ on which the curve $t \mapsto \big(q(t), p(t)\big)$ takes values is called the *phase space.*

Assume that $H(t, q, p)$ does not depend on t, so that we write $H(q, p)$ instead of $H(t, q, p)$ (if H is of the form (1.6.5), this happens if and only if the potential $V(t, q)$ does not depend on t). If $t \mapsto \big(q(t), p(t)\big)$ is a solution of Hamilton's equations (1.6.7) then:

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(q(t), p(t)\big) = \frac{\partial H}{\partial q}\big(q(t), p(t)\big)\frac{\mathrm{d}q}{\mathrm{d}t}(t) + \frac{\partial H}{\partial p}\big(q(t), p(t)\big)\frac{\mathrm{d}p}{\mathrm{d}t}(t)$$

$$= \frac{\partial H}{\partial q}\big(q(t), p(t)\big)\frac{\partial H}{\partial p}\big(q(t), p(t)\big) - \frac{\partial H}{\partial p}\big(q(t), p(t)\big)\frac{\partial H}{\partial q}\big(q(t), p(t)\big). \tag{1.6.8}$$

In the formula above, the expressions:

$$\frac{\partial H}{\partial q}\big(q(t), p(t)\big)\frac{\partial H}{\partial p}\big(q(t), p(t)\big), \quad \frac{\partial H}{\partial p}\big(q(t), p(t)\big)\frac{\partial H}{\partial q}\big(q(t), p(t)\big)$$

denote the evaluation of an element of $(\mathbb{R}^3)^{n*}$ at an element of $(\mathbb{R}^3)^n$. Under the identification of $(\mathbb{R}^3)^{n*}$ with $(\mathbb{R}^3)^n$, both expressions become the (standard) inner product between the vectors $\frac{\partial H}{\partial q}\big(q(t), p(t)\big)$, $\frac{\partial H}{\partial p}\big(q(t), p(t)\big)$ of $(\mathbb{R}^3)^n$. Hence the difference (1.6.8) vanishes and:

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(q(t), p(t)\big) = 0.$$

We have proven:

1.6.4. PROPOSITION. *If $H(t, q, p)$ does not depend on t then H is constant along the solutions of Hamilton's equations* (1.6.7). □

In other words, the Hamiltonian H (when it does not depend on time) is a *first integral* for Hamilton's equations. It follows that, when the potential energy does not depend on time, the *total energy is conserved* in Classical Mechanics. Of course, we could have discovered the conservation of total energy simply by computing the derivative of (1.6.4), but it is interesting to see that such conservation law is a particular case of the more general fact that H is a first integral of Hamilton's equations.

Hamilton's equations state that the curve $t \mapsto \big(q(t), p(t)\big)$ is an integral curve of the time-dependent vector field over $(\mathbb{R}^3)^n \times (\mathbb{R}^3)^n$ defined by:

$$(t, q, p) \longmapsto \Big(\frac{\partial H}{\partial p}(t, q, p), -\frac{\partial H}{\partial q}(t, q, p)\Big) \in (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n. \tag{1.6.9}$$

This vector field is *almost* the gradient of H, which is given by:

$$\nabla_{(q,p)}H(t,q,p)=\Big(\frac{\partial H}{\partial q}(t,q,p),\frac{\partial H}{\partial p}(t,q,p)\Big)\in(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n.$$

However, (1.6.9) differs from the gradient by an order switch and by a minus sign. We now introduce a modified notion of gradient that yields exactly (1.6.9). This is done as follows: recall that the gradient of a map is the vector that, when contracted with a given inner product, yields the differential of that map; in other words, the gradient is the vector that corresponds to the differential via the isomorphism between the space and the dual space induced by an inner product. For example, if $\langle\cdot,\cdot\rangle$ denotes the standard inner product of $(\mathbb{R}^3)^n$, then:

$$\langle\nabla_{(q,p)}H(t,q,p),(x,y)\rangle=\partial_{(q,p)}H(t,q,p)(x,y)=\frac{\partial H}{\partial q}(t,q,p)x+\frac{\partial H}{\partial p}(t,q,p)y,$$

for all $(x,y)\in(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$. We obtain a new notion of gradient simply by replacing the inner product with something else. We define an *anti-symmetric* bilinear map ω over the vector space $(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$ by setting:

$$\omega\big((x,y),(\bar{x},\bar{y})\big)=\langle x,\bar{y}\rangle-\langle\bar{x},y\rangle,\qquad (x,y),(\bar{x},\bar{y})\in(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n. \tag{1.6.10}$$

The bilinear form ω induces an isomorphism between $(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$ and its dual space $\big((\mathbb{R}^3)^n\times(\mathbb{R}^3)^n\big)^*$ and we can therefore consider the vector:

$$\vec{H}(t,q,p)\in(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$$

that corresponds to the differential $\partial_{(q,p)}H(t,q,p)\in\big((\mathbb{R}^3)^n\times(\mathbb{R}^3)^n\big)^*$ via such isomorphism. More explicitly, given (t,q,p) in the domain of H, it is easily checked that there exists a unique vector $\vec{H}(t,q,p)$ in $(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$ that satisfies the equality:

$$\omega\big(\vec{H}(t,q,p),(x,y)\big)=\partial_{(q,p)}H(t,q,p)(x,y)=\frac{\partial H}{\partial q}(t,q,p)x+\frac{\partial H}{\partial p}(t,q,p)y,$$

for all $(x,y)\in(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$ and that such vector is given by:

$$\vec{H}(t,q,p)=\Big(\frac{\partial H}{\partial p}(t,q,p),-\frac{\partial H}{\partial q}(t,q,p)\Big).$$

Thus $\vec{H}$ is precisely the time-dependent vector field (1.6.9) whose integral curves are the solutions $t\mapsto\big(q(t),p(t)\big)$ of Hamilton's equations!

1.6.5. Definition. The bilinear form ω defined in (1.6.10) is called the *canonical symplectic form* of the phase space $(\mathbb{R}^3)^n\times(\mathbb{R}^3)^n$ and the vector field $\vec{H}$ is called the *symplectic gradient* of the map H.

In this section we have taken the first steps towards presenting what is called the *Hamiltonian formalism.* Here is an overview of what are going to be our next steps. Hamilton's equations can be formulated in the context of manifolds and for that purpose the manifold is required to be endowed with what will be called a *symplectic form*. We will study Hamiltonians and symplectic forms on manifolds in Section 1.8, but first we have to know what is meant by a symplectic form over a vector space, which is the subject of Section 1.7. After Section 1.8, we will see that given a Lagrangian L over a manifold Q (so that the domain of L is an open subset of $\mathbb{R}\times TQ$) then, under certain conditions (known as *hyper-regularity*), we can construct a map H (the *Legendre transform* of L), called the *Hamiltonian* corresponding to L, defined over an open subset of $\mathbb{R} \times TQ^*$, where TQ^* denotes the cotangent bundle of Q. When the Lagrangian L is the difference between total kinetic energy and potential energy (so that — assuming that the force exerted by the constraint is normal to Q — the Euler–Lagrange equation is the dynamical equation of Classical Mechanics), the corresponding Hamiltonian H will be exactly the sum of the total kinetic energy with the potential energy (with the total kinetic energy rewritten in terms of new variables p). We will see that a cotangent bundle TQ^* carries a canonical symplectic form, so that it makes sense to talk about Hamilton's equations associated to a Hamiltonian H over a cotangent bundle. Moreover, we will see that when H is constructed from L, the solutions of Hamilton's equations associated to H are the same as the solutions of the Euler–Lagrange equation associated to L; more precisely, q is a solution of the Euler–Lagrange equation associated to L if and only if (q, p) is a solution of Hamilton's equations associated to H, for some curve p. There will be a explicit description of such curve p. When $Q = (\mathbb{R}^3)^n$ and L is the difference between total kinetic energy and potential energy, p agrees with the momentum (1.6.1).

1.7. Symplectic forms over vector spaces

In Section 1.6 we have defined the canonical symplectic form of the phase space $(\mathbb{R}^3)^n \times (\mathbb{R}^3)^n$. Let us now explain what we mean by a symplectic form over a vector space and let us prove some elementary results about those. This section is just a bunch of results from elementary linear algebra (which are usually not taught in standard elementary linear algebra courses). In Subsection 1.7.1 we show how to construct a volume form from a symplectic form using the exterior product and that requires a bit of multilinear algebra. For the reader's convenience, we have included a short review of multilinear algebra in the appendix (Section A.1).

1.7.1. Definition. Let V be a real finite-dimensional vector space. A *symplectic form* over V is an anti-symmetric bilinear form $\omega : V \times V \to \mathbb{R}$ that is also *non degenerate*, i.e., given $v \in V$, if $\omega(v, w) = 0$ for all $w \in V$ then $v = 0$. The pair (V, ω) is called a *symplectic vector space* (or just a *symplectic space*).

We have restricted our definition to the context of finite-dimensional vector spaces over the field of real numbers. Of course, the same definition makes sense for arbitrary vector spaces over any scalar field, but real finite-dimensional vector spaces are sufficient for our purposes. Notice that the condition that ω be non degenerate is equivalent to the condition that the linear map canonically associated to ω:

$$V \ni v \longmapsto \omega(v, \cdot) \in V^* \tag{1.7.1}$$

be injective. Since V is finite-dimensional, this is the same as requiring that (1.7.1) be an isomorphism. In the infinite-dimensional case, one should talk about *weak non degeneracy* (when (1.7.1) is injective) and *strong non degeneracy* (when (1.7.1) is an isomorphism[17]). But let us focus on the finite-dimensional case. Just like inner products, symplectic forms induce an isomorphism between the space V and its dual V^*, so one can talk about the vector v that *represents* a linear functional $\alpha \in V^*$ with respect to ω, i.e., $v \in V$ is the only vector such that $\omega(v, \cdot) = \alpha$. This is the crucial property that allowed us to define the symplectic gradient $\vec{H}$ in Section 1.6. Notice that in the case of an inner product $\langle \cdot, \cdot \rangle$, the linear functionals $\langle v, \cdot \rangle$ and $\langle \cdot, v \rangle$ are equal, while if ω is a symplectic form then $\omega(v, \cdot) = -\omega(\cdot, v)$, so one must be more careful and pay attention to our convention of putting the v in the *first* variable of ω in the definition (1.7.1) of the isomorphism between V and its dual space V^*.

1.7.2. Example. For any natural number n, we define the *canonical symplectic form* of $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ by:

$$\omega_0\big((x, y), (\bar{x}, \bar{y})\big) = \langle x, \bar{y} \rangle - \langle \bar{x}, y \rangle, \quad x, y, \bar{x}, \bar{y} \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product of $\mathbb{R}^n$. It is a simple exercise to check that ω_0 is indeed anti-symmetric and non degenerate. We also define a *canonical symplectic form* for the space $\mathbb{R}^n \times \mathbb{R}^{n*}$ (again denoted by ω_0), by setting:

$$\omega_0\big((x, \alpha), (\bar{x}, \bar{\alpha})\big) = \bar{\alpha}(x) - \alpha(\bar{x}), \quad x, \bar{x} \in \mathbb{R}^n, \ \alpha, \bar{\alpha} \in \mathbb{R}^{n*}.$$

Notice that when using $\mathbb{R}^n \times \mathbb{R}^{n*}$ instead of $\mathbb{R}^n \times \mathbb{R}^n$ one does not need an inner product to define the symplectic form!

Not surprisingly, there is a notion of isomorphism for symplectic spaces.

1.7.3. Definition. Given symplectic spaces (V, ω), $(\widetilde{V}, \tilde{\omega})$, then a *symplectomorphism* from (V, ω) to $(\widetilde{V}, \tilde{\omega})$ is a linear isomorphism:

$$T : V \longrightarrow \widetilde{V}$$

such that:

$$\tilde{\omega}\big(T(v), T(w)\big) = \omega(v, w), \tag{1.7.2}$$

for all $v, w \in V$. One could rephrase (1.7.2) by saying that the *pull-back* $T^*\tilde{\omega}$ (which is defined by the lefthand side of (1.7.2)) is equal to ω.

[17] Of course, in the infinite-dimensional case, V would normally be assumed to be a *topological* vector space and one would consider only its topological dual space, consisting of *continuous* linear functionals over V.

Clearly, the composition of symplectomorphisms is a symplectomorphism and the inverse of a symplectomorphism is a symplectomorphism. Symplectomorphisms in the theory of symplectic spaces play the same role that orthogonal transformations (i.e., linear isometries) play in the theory of vector spaces with inner product. We now define the analogue for the theory of symplectic spaces of the notion of orthonormal basis.

1.7.4. Definition. If (V, ω) is a symplectic space, then a *symplectic basis* for (V, ω) is a basis $(e_1, \ldots, e_n, e_1', \ldots, e_n')$ for V such that:

$$\omega(e_i, e_j') = \delta_{ij}, \quad \omega(e_i, e_j) = 0, \quad \omega(e_i', e_j') = 0,$$

for all $i, j = 1, \ldots, n$, where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$.

Clearly, the canonical basis of $\mathbb{R}^{2n}$ (and the canonical basis of $\mathbb{R}^n \times \mathbb{R}^{n*}$) is symplectic with respect to the canonical symplectic form. Just like in the theory of spaces with inner product, we have:

1.7.5. Proposition. *Any symplectic space admits a symplectic basis.*

Proof. This is a simple linear algebra exercise whose details are left to the reader. The idea is to use induction in the dimension of the symplectic space (V, ω). If V is not the null space then we can find vectors $e_1, e_1' \in V$ such that $\omega(e_1, e_1') \neq 0$ and obviously such vectors can be chosen with $\omega(e_1, e_1') = 1$. The restriction of ω to the space $\mathrm{span}\{e_1, e_1'\}$ spanned by e_1, e_1' is easily seen to be non degenerate and thus it follows from the result of Exercise 1.28 that V is the direct sum of $\mathrm{span}\{e_1, e_1'\}$ with the orthogonal complement (with respect to ω) of $\mathrm{span}\{e_1, e_1'\}$. Now apply the induction hypothesis to the restriction of ω to the orthogonal complement of $\mathrm{span}\{e_1, e_1'\}$ and conclude the proof. □

In Exercise 1.30 we ask the reader to prove a generalization of Proposition 1.7.5 that yields a special basis for anti-symmetric bilinear forms ω that are not necessarily non degenerate.

1.7.6. Corollary. *Any symplectic space is even dimensional.*

Proof. A symplectic basis has an even number of elements. □

Another proof of Corollary 1.7.6 is given in the statement of Exercise 1.27.

We have the following result, which is analogous to the result that two vector spaces having the same dimension, endowed with inner products, are linearly isometric.

1.7.7. Corollary. *If (V, ω) and $(\widetilde{V}, \tilde{\omega})$ are symplectic spaces having the same dimension, then there exists a symplectomorphism from (V, ω) to $(\widetilde{V}, \tilde{\omega})$.*

Proof. Choose a symplectic basis of (V,ω) and a symplectic basis of $(\widetilde{V},\tilde{\omega})$. Define a linear isomorphism $T: V \to \widetilde{V}$ sending one basis to the other. Apply the result of Exercise 1.29. $\square$

Because of Corollary 1.7.7, in order to prove a theorem about arbitrary symplectic spaces, it suffices to prove it for $\mathbb{R}^{2n}$, endowed with the canonical symplectic form.

1.7.1. The volume form induced by a symplectic form. Let (V,ω) be a symplectic space with $\dim(V) = 2n$. The symplectic form ω is an element of $\bigwedge_2 V^*$ and the wedge product:

$$\omega^n = \omega \wedge \cdots \wedge \omega$$

of n copies of ω is an element of the one-dimensional space $\bigwedge_{2n} V^*$. We will show that ω^n is not zero and therefore it is a volume form over V. Because of Corollary 1.7.7 it suffices to prove that ω^n is not zero when $V = \mathbb{R}^{2n}$ and $\omega = \omega_0$ is the canonical symplectic form of $\mathbb{R}^{2n}$. Let us denote[18] by $(\mathrm{d}q^1, \ldots, \mathrm{d}q^n, \mathrm{d}p_1, \ldots, \mathrm{d}p_n)$ the dual basis of the canonical basis of $\mathbb{R}^{2n}$. The canonical symplectic form ω_0 of $\mathbb{R}^{2n}$ is given by:

$$\omega_0 = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i. \tag{1.7.3}$$

Namely:

$$\begin{aligned}\sum_{i=1}^{n}(\mathrm{d}q^i \wedge \mathrm{d}p_i)\big((x,y),(\bar{x},\bar{y})\big) &= \sum_{i=1}^{n} \mathrm{d}q^i(x,y)\mathrm{d}p_i(\bar{x},\bar{y}) - \mathrm{d}q^i(\bar{x},\bar{y})\mathrm{d}p_i(x,y)\\ &= \sum_{i=1}^{n} x^i\bar{y}_i - \bar{x}^i y_i = \langle x,\bar{y}\rangle - \langle \bar{x},y\rangle,\end{aligned}$$

for all $x = (x^1,\ldots,x^n)$, $y = (y_1,\ldots,y_n)$, $\bar{x} = (\bar{x}^1,\ldots,\bar{x}^n)$, $\bar{y} = (\bar{y}_1,\ldots,\bar{y}_n)$ in $\mathbb{R}^n$. The *canonical volume form* of $\mathbb{R}^{2n}$ is given by:

$$\mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n.$$

Let us compute the wedge product ω_0^n of n copies of ω_0. We have:

$$\omega_0^n = \sum_{i_1,\ldots,i_n=1}^{n} \mathrm{d}q^{i_1} \wedge \mathrm{d}p_{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_n} \wedge \mathrm{d}p_{i_n}.$$

Clearly:

$$\mathrm{d}q^{i_1} \wedge \mathrm{d}p_{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_n} \wedge \mathrm{d}p_{i_n} = 0$$

unless $i_1,\ldots,i_n \in \{1,\ldots,n\}$ are pairwise distinct. Therefore:

$$\omega_0^n = \sum_{\sigma \in S^n} \mathrm{d}q^{\sigma(1)} \wedge \mathrm{d}p_{\sigma(1)} \wedge \cdots \wedge \mathrm{d}q^{\sigma(n)} \wedge \mathrm{d}p_{\sigma(n)},$$

[18] This is actually more than just a notation. One can see $\mathrm{d}q^i$, $\mathrm{d}p_j$ as the (constant) one-forms over $\mathbb{R}^{2n}$ that are the differentials of the scalar functions $(q,p) \mapsto q^i$, $(q,p) \mapsto p_j$.

where S^n denotes the group of all bijections of the set $\{1, \dots, n\}$. By counting order switches the reader can easily check that:

$$\begin{aligned}&\mathrm{d}q^{\sigma(1)} \wedge \mathrm{d}p_{\sigma(1)} \wedge \cdots \wedge \mathrm{d}q^{\sigma(n)} \wedge \mathrm{d}p_{\sigma(n)} \\ &\qquad = (-1)^{\frac{n(n-1)}{2}} \mathrm{d}q^{\sigma(1)} \wedge \cdots \wedge \mathrm{d}q^{\sigma(n)} \wedge \mathrm{d}p_{\sigma(1)} \wedge \cdots \wedge \mathrm{d}p_{\sigma(n)}.\end{aligned}$$

We have:

$$\begin{aligned}&\mathrm{d}q^{\sigma(1)} \wedge \cdots \wedge \mathrm{d}q^{\sigma(n)} \wedge \mathrm{d}p_{\sigma(1)} \wedge \cdots \wedge \mathrm{d}p_{\sigma(n)} \\ &\qquad = \operatorname{sgn}(\sigma)^2 \mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n \\ &\qquad = \mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n,\end{aligned}$$

where $\operatorname{sgn}(\sigma)$ is the *sign* of the permutation σ, i.e., $\operatorname{sgn}(\sigma) = 1$ if σ is even and $\operatorname{sgn}(\sigma) = -1$ if σ is odd. Hence:

$$\omega_0^n = (-1)^{\frac{n(n-1)}{2}} n!\, \mathrm{d}q^1 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n.$$

This motivates the following:

1.7.8. Definition. If (V, ω) is a symplectic space with $\dim(V) = 2n$ then the *volume form induced by* ω is:

$$(-1)^{\frac{n(n-1)}{2}} \frac{1}{n!} \omega^n, \tag{1.7.4}$$

where ω^n denotes the wedge product of n copies of ω.

Our computations have shown that the volume form induced by the canonical symplectic form of $\mathbb{R}^{2n}$ is just the canonical volume form of $\mathbb{R}^{2n}$ (obviously, the same holds if $\mathbb{R}^{2n}$ is replaced with $\mathbb{R}^n \times \mathbb{R}^{n*}$). Moreover, as remarked earlier, it follows from Corollary 1.7.7 that (1.7.4) is indeed a volume form over V, i.e., it is not zero. Namely, the pull-back of (1.7.4) by a symplectomorphism from $(\mathbb{R}^{2n}, \omega_0)$ to (V, ω) is equal to the canonical volume form of $\mathbb{R}^{2n}$, so that (1.7.4) cannot be zero.

1.8. Symplectic manifolds and Hamiltonians

In Section 1.6 we have seen that the solutions to Hamilton's equations are precisely the integral curves of the symplectic gradient of the Hamiltonian. Let us now formulate Hamilton's equations in the context of manifolds. The required ingredients are a Hamiltonian and a symplectic form. In this section we are going to use several facts that are taught during courses on calculus on manifolds. The reader is assumed to be familiar with such facts, but there is a short summary of those in the appendix (Section A.2).

1.8.1. Definition. Let M be a differentiable manifold. By a *symplectic form* over M we mean a smooth two-form ω over M such that:

(a) ω_x is non degenerate (i.e., ω_x is a symplectic form over the tangent space T_xM), for all $x \in M$;
(b) ω is closed, i.e., the exterior derivative $\mathrm{d}\omega$ vanishes.

The pair (M, ω) is called a *symplectic manifold*.

Observe that if (V, ω) is a symplectic space (in the sense of Section 1.7) and if we regard ω as a *constant* two-form over V (i.e., the two-form that associates ω to every point $x \in V$) then ω is automatically closed and therefore (V, ω) is a symplectic manifold. Actually, we have a bit of a terminological conflict here: if V is a real finite-dimensional vector space then a symplectic form over the vector space V (in the sense of Section 1.7) is not the same thing as a symplectic form over V when V is regarded as a manifold. When V is regarded as a manifold, then a symplectic form ω over V gives us a symplectic form ω_x over the vector space V *for each* $x \in V$. Terminological conflicts of this type are common during courses on differentiable manifolds and they cause no disastrous misunderstandings.

It is an obvious consequence of Corollary 1.7.6 that every symplectic manifold is even dimensional.

It is not at all obvious why we have chosen to require a symplectic form on a manifold to be closed. It happens that, because of such assumption, all symplectic manifolds are locally alike (Darboux's theorem below). The reader will see that the assumption that the symplectic form be closed is crucial for the most basic theorems on the subject. A (not necessarily closed) smooth two-form that is non degenerate at every point is usually called *almost symplectic*.

Let us now formulate Hamilton's equations in a symplectic manifold.

1.8.2. Definition. Let (M, ω) be a symplectic manifold and $H : M \to \mathbb{R}$ be a smooth map. We call it a *Hamiltonian* over M (we will use that terminology sometimes also when H is only defined in an open subset of M). The *symplectic gradient* of H is the unique smooth vector field $\vec{H}$ over M such that:

$$\omega\big(\vec{H}(x), \cdot\big) = \mathrm{d}H(x) \in T_xM^*,$$

for all $x \in M$. A smooth map $H : \mathbb{R} \times M \to \mathbb{R}$ (perhaps defined only over some open subset of $\mathbb{R} \times M$) is called a *time-dependent Hamiltonian* over M. It's *symplectic gradient* is defined to be the unique time-dependent vector field $\vec{H}$ over M such that:

$$\omega\big(\vec{H}(t, x), \cdot\big) = \partial_x H(t, x) \in T_xM^*,$$

for all $(t, x) \in \mathbb{R} \times M$, where $\partial_x H(t, x)$ denotes the differential at the point x of the map $H(t, \cdot)$. An integral curve of the symplectic gradient $\vec{H}$ of a (possibly time-dependent) Hamiltonian H is called a *solution to Hamilton's equations*.

If H is a time-dependent Hamiltonian then, for each $t \in \mathbb{R}$, $H(t, \cdot)$ is a (time *in*dependent) Hamiltonian and the symplectic gradient of $H(t, \cdot)$ is the vector field $\vec{H}(t, \cdot)$, where $\vec{H}$ denotes the symplectic gradient of H. In other words, the symplectic gradient of a time-dependent Hamiltonian at

$(t,x) \in \mathbb{R} \times M$ can be obtained by first "freezing time", obtaining $H(t,\cdot)$, and then computing the symplectic gradient of $H(t,\cdot)$ at x.

Here is the generalization of the notion of symplectomorphism to the context of manifolds.

1.8.3. Definition. Given symplectic manifolds (M,ω), $(\widetilde{M},\tilde{\omega})$ then a *symplectomorphism* from (M,ω) to $(\widetilde{M},\tilde{\omega})$ is a smooth diffeomorphism:

$$\Phi : M \longrightarrow \widetilde{M}$$

such that:

$$\Phi^*\tilde{\omega} = \omega,$$

i.e., such that $\mathrm{d}\Phi_x : T_xM \to T_{\Phi(x)}\widetilde{M}$ is a symplectomorphism from the symplectic space (T_xM, ω_x) to the symplectic space $(T_{\Phi(x)}\widetilde{M}, \tilde{\omega}_{\Phi(x)})$. A *symplectic chart* over a symplectic manifold (M,ω) is a local chart:

$$\Phi : U \subset M \longrightarrow \widetilde{U} \subset \mathbb{R}^{2n}$$

that is a symplectomorphism from U endowed with (the restriction of) ω to $\widetilde{U}$ endowed with (the restriction of) the canonical symplectic form of $\mathbb{R}^{2n}$ (we also allow the counter-domain of a symplectic chart to be an open subset of $\mathbb{R}^n \times \mathbb{R}^{n*}$).

1.8.4. Theorem (Darboux). *If (M,ω) is a symplectic manifold then for every point of M there exists a symplectic chart whose domain contains that point.*

Proof. See Exercise 1.32. □

Darboux's theorem won't be terribly important for us, since for what is going to be our central example of symplectic manifold (cotangent bundles), the symplectic charts can be easily constructed. It is easy to see that the integral curves of a symplectic gradient $\vec{H}$ are represented, with respect to a symplectic chart Φ, by solutions of the (standard) Hamilton's equations (1.6.7) corresponding to the Hamiltonian that represents H with respect to Φ (see Exercise 1.31 for details).

We have proven in Section 1.6 (Proposition 1.6.4) that a Hamiltonian that does not depend on time is a first integral of Hamilton's equations. Such fact generalizes to Hamiltonians on symplectic manifolds.

1.8.5. Theorem. *If H is a (time* independent*) Hamiltonian over a symplectic manifold (M,ω) then H is constant over the integral curves of the symplectic gradient $\vec{H}$, i.e., H is a first integral of $\vec{H}$.*

Proof. If $t \mapsto x(t)$ is an integral curve of $\vec{H}$ then:

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(x(t)\big) = \mathrm{d}H_{x(t)}\big(\dot{x}(t)\big) = \mathrm{d}H_{x(t)}(\vec{H}_{x(t)}) = \omega\big(\vec{H}_{x(t)}, \vec{H}_{x(t)}\big) = 0,$$

where $\dot{x}(t) = \frac{\mathrm{d}x}{\mathrm{d}t}(t)$. □

Theorem 1.8.5 does not hold for time-dependent Hamiltonians. Actually, the argument used in its proof shows that if $t \mapsto x(t)$ is an integral curve of $\vec{H}$ then:

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(t, x(t)\big) = \frac{\partial H}{\partial t}\big(t, x(t)\big).$$

Now we establish a deeper relationship between the flow of $\vec{H}$ and the symplectic structure.

1.8.6. THEOREM. *If H is a time-dependent Hamiltonian over a symplectic manifold (M, ω) then the symplectic form ω is invariant under the flow of the symplectic gradient $\vec{H}$.*

PROOF. We just have to check that the Lie derivative $\mathbb{L}_{\vec{H}_t}\omega$ is zero, for all $t \in \mathbb{R}$, where $\vec{H}_t = \vec{H}(t, \cdot)$ (see Proposition A.2.7 if you are not familiar with this). We use the standard formula for the Lie derivative of a differential form (see (A.2.9)):

$$\mathbb{L}_{\vec{H}_t}\omega = \mathrm{d}i_{\vec{H}_t}\omega + i_{\vec{H}_t}\mathrm{d}\omega.$$

Since ω is closed, the second term on the righthand side of the equality above vanishes. As for the first term, observe that, by the definition of $\vec{H}_t$, $i_{\vec{H}_t}\omega$ is equal to $\mathrm{d}H_t$, where $H_t = H(t, \cdot)$. Since $\mathrm{d}(\mathrm{d}H_t) = 0$, the proof is concluded. □

We have seen in Subsection 1.7.1 (recall Definition 1.7.8) that a symplectic form over a vector space induces a volume form over that vector space. Obviously, the same construction can be used (pointwise) for manifolds.

1.8.7. DEFINITION. The *volume form* induced by a symplectic form ω on a manifold M is defined by:

$$(-1)^{\frac{n(n-1)}{2}} \frac{1}{n!}\,\omega^n, \tag{1.8.1}$$

where ω^n denotes the wedge product of n copies of ω and n denotes half the dimension of M.

We have already shown that ω^n never vanishes, so that (1.8.1) is indeed a volume form over M.

1.8.8. COROLLARY (Liouville's theorem). *The volume form induced by the symplectic form is invariant under the flow of the symplectic gradient of a (possibly time-dependent) Hamiltonian.*

PROOF. A diffeomorphism that preserves the symplectic form preserves the volume form. □

Liouville's theorem is very important for Statistical Mechanics. Also, there are many theorems about the dynamics of a flow that preserves a measure (which, by Liouville's theorem, is the case of a Hamiltonian flow).

1.9. Canonical forms in a cotangent bundle

In this section we will show that the cotangent bundle of a differentiable manifold carries a canonical one-form and a canonical symplectic form; such symplectic form is, up to a sign, equal to the exterior differential of the canonical one-form. The symplectic form will allow us to talk about Hamilton's equations in a cotangent bundle.

Let Q be a differentiable manifold and denote by TQ^* its cotangent bundle. A point of TQ^* will be written as an ordered pair (q,p), where q is a point of Q and $p \in T_qQ^*$ is a linear functional over the tangent space T_qQ. Denote by $\pi : TQ^* \to Q$ the *canonical projection*, i.e., $\pi(q,p) = q$, for all $(q,p) \in TQ^*$. The cotangent bundle TQ^* is a differentiable manifold and the projection π is a smooth map (it is also a smooth submersion). We are going to define a one-form θ over the differentiable manifold TQ^*, i.e., for each point $(q,p) \in TQ^*$ we are going to associate a linear functional $\theta_{(q,p)}$ over the tangent space $T_{(q,p)}TQ^*$. The construction of θ is straightforward: given a point $(q,p) \in TQ^*$, we consider the differential $\mathrm{d}\pi_{(q,p)}$ of the projection π at the point (q,p), which is a linear map:

$$\mathrm{d}\pi_{(q,p)} : T_{(q,p)}TQ^* \longrightarrow T_qQ.$$

Now, p is a linear functional over T_qQ and therefore we can compose it with $\mathrm{d}\pi_{(q,p)}$ to obtain a linear functional over $T_{(q,p)}TQ^*$. We set:

$$\theta_{(q,p)} = p \circ \mathrm{d}\pi_{(q,p)} : T_{(q,p)}TQ^* \longrightarrow \mathbb{R}, \tag{1.9.1}$$

for all $(q,p) \in TQ^*$, so that:

$$\theta_{(q,p)}(\zeta) = p\big(\mathrm{d}\pi_{(q,p)}(\zeta)\big),$$

for all $(q,p) \in TQ^*$ and all $\zeta \in T_{(q,p)}TQ^*$.

1.9.1. Definition. The one-form θ defined in (1.9.1) is called the *canonical one-form* of the cotangent bundle TQ^*.

1.9.2. Example. Let us compute explicitly the canonical one-form θ of the cotangent bundle TQ^* of an open subset Q of $\mathbb{R}^n$. Such cotangent bundle is identified with the product $Q \times \mathbb{R}^{n*}$ and the canonical projection $\pi : TQ^* \to Q$ is just the projection onto the first coordinate of the product $Q \times \mathbb{R}^{n*}$. Given a point $(q,p) \in TQ^*$, i.e., q is in Q and p is in $\mathbb{R}^{n*}$, then the tangent space $T_{(q,p)}TQ^*$ is identified with $\mathbb{R}^n \times \mathbb{R}^{n*}$ and the differential $\mathrm{d}\pi_{(q,p)}$ is the projection onto the first coordinate of the product $\mathbb{R}^n \times \mathbb{R}^{n*}$. The canonical one-form θ is given by:

$$\theta_{(q,p)}(\zeta_1,\zeta_2) = p(\zeta_1), \quad q \in Q,\ p \in \mathbb{R}^{n*},\ \zeta_1 \in \mathbb{R}^n,\ \zeta_2 \in \mathbb{R}^{n*}. \tag{1.9.2}$$

Let us write the one-form θ using the standard way of writing down differential forms. Denote by[19]:

$$q : Q \times \mathbb{R}^{n*} \longrightarrow Q, \quad p : Q \times \mathbb{R}^{n*} \longrightarrow \mathbb{R}^{n*}$$

[19] Yes, we are using q and p both for the names of the projection maps of the product $Q \times \mathbb{R}^{n*}$ and for the names of points of Q and of $\mathbb{R}^{n*}$, respectively.

the projection maps of the product $Q \times \mathbb{R}^{n*}$ and write:

$$q = (q^1, \dots, q^n), \quad p = (p_1, \dots, p_n),$$

so that q^i and p_i, $i = 1, \dots, n$, are real valued maps over $Q \times \mathbb{R}^{n*}$ (when we write $p = (p_1, \dots, p_n)$, we are using the standard identification between $\mathbb{R}^{n*}$ and $\mathbb{R}^n$). Denote by $\mathrm{d}q^i$, $\mathrm{d}p_i$ the (ordinary or exterior) differential of such maps, which are (constant) one-forms over $Q \times \mathbb{R}^{n*}$. The one-forms $\mathrm{d}q^i$, $\mathrm{d}p_i$ are simply the dual basis of the canonical basis of $\mathbb{R}^n \times \mathbb{R}^{n*}$, i.e., they are the $2n$ projections of $\mathbb{R}^n \times \mathbb{R}^{n*} \cong \mathbb{R}^{2n}$. Given $\zeta = (\zeta_1, \zeta_2) \in \mathbb{R}^n \times \mathbb{R}^{n*}$ then the n coordinates of $\zeta_1 \in \mathbb{R}^n$ are $\mathrm{d}q^i(\zeta)$, $i = 1, \dots, n$ and therefore, by (1.9.2), the one-form θ is equal to:

$$\theta = \sum_{i=1}^{n} p_i \mathrm{d}q^i. \tag{1.9.3}$$

As we will see in one moment, by computing the canonical one-form of the cotangent bundle of an open subset of $\mathbb{R}^n$, we have actually computed the canonical one-form of any cotangent bundle.

Given differentiable manifolds Q, $\widetilde{Q}$, if $\varphi : Q \to \widetilde{Q}$ is a smooth diffeomorphism then it induces a smooth diffeomorphism:

$$\mathrm{d}\varphi : TQ \longrightarrow T\widetilde{Q}$$

between the tangent bundles TQ, $T\widetilde{Q}$ (which sends a point $(q, \dot{q})$ in TQ to the point $\big(\varphi(q), \mathrm{d}\varphi_q(\dot{q})\big)$ in $T\widetilde{Q}$) and it also induces a smooth diffeomorphism:

$$\mathrm{d}^*\varphi : TQ^* \longrightarrow T\widetilde{Q}^*$$

between the cotangent bundles TQ^*, $T\widetilde{Q}^*$, defined by:

$$\mathrm{d}^*\varphi(q, p) = \big(\varphi(q), p \circ \mathrm{d}\varphi_q^{-1}\big) = \big(\varphi(q), (\mathrm{d}\varphi_q^{-1})^*(p)\big), \quad (q, p) \in TQ^*. \tag{1.9.4}$$

Be aware that our notation might be a little misleading: the restriction of $\mathrm{d}^*\varphi$ to a fiber T_qQ^* of TQ^* is not the transpose $\mathrm{d}\varphi_q^*$ of the differential $\mathrm{d}\varphi_q$, but actually its *inverse* $(\mathrm{d}\varphi_q^*)^{-1} = (\mathrm{d}\varphi_q^{-1})^*$ (the map $\mathrm{d}\varphi_q^*$ goes in the wrong direction!). That is why we write $\mathrm{d}^*\varphi$, instead of $\mathrm{d}\varphi^*$; using $\mathrm{d}\varphi^*$ would be too misleading.

We have a commutative diagram:

$$\begin{tikzcd} TQ^* \arrow[r, "\mathrm{d}^*\varphi"] \arrow[d] & T\widetilde{Q}^* \arrow[d] \\ Q \arrow[r, "\varphi"'] & \widetilde{Q} \end{tikzcd} \tag{1.9.5}$$

in which the vertical arrows are the canonical projections.

If θ denotes the canonical one-form of TQ^* and $\tilde{\theta}$ denotes the canonical one-form of $T\widetilde{Q}^*$ then it is a simple exercise[20] to check that the diffeomorphism $\mathrm{d}^*\varphi$ carries θ to $\tilde{\theta}$, i.e, the pull-back of $\tilde{\theta}$ by $\mathrm{d}^*\varphi$ is equal to θ:

$$(\mathrm{d}^*\varphi)^*\tilde{\theta} = \theta. \tag{1.9.6}$$

The reader should have quickly guessed that (1.9.6) must hold! This is just a particular case of the fact that any concept whose definition makes sense for an arbitrary differentiable manifold must be preserved by smooth diffeomorphisms (just like any concept whose definition makes sense for rings must be preserved by ring isomorphisms, and so on).

If $\varphi : U \subset Q \to \widetilde{U} \subset \mathbb{R}^n$ is a local chart on Q then the smooth diffeomorphism:

$$\mathrm{d}^*\varphi : TU^* \subset TQ^* \longrightarrow T\widetilde{U}^* = \widetilde{U} \times \mathbb{R}^{n*}$$

is a local chart on the cotangent bundle TQ^*; it is the local chart *canonically associated* to φ on the cotangent bundle. The local chart $\mathrm{d}^*\varphi$ on TQ^* carries the restriction[21] to the open set TU^* of the canonical one-form θ of TQ^* to the canonical one-form of the cotangent bundle $T\widetilde{U}^*$ of the open subset $\widetilde{U}$ of $\mathbb{R}^n$. Thus, the canonical one-form of the cotangent bundle $T\widetilde{U}^*$ (which we have computed in Example 1.9.2) is the representation of the canonical one-form θ of TQ^* with respect to the local chart $\mathrm{d}^*\varphi$. Notice that our considerations have proven that *the canonical one-form of a cotangent bundle is smooth.*

There are two possibilities here for notation and we will let the reader pick her favorite one. We can, as in Example 1.9.2, denote by:

$$q^1, \ldots, q^n, p_1, \ldots, p_n, \tag{1.9.7}$$

the $2n$ projections of $\widetilde{U} \times \mathbb{R}^{n*} \subset \mathbb{R}^n \times \mathbb{R}^{n*}$, so that the canonical one-form of the cotangent bundle $T\widetilde{U}^*$ is given by the righthand side of (1.9.3) and the (restriction to the open set TU^* of the) canonical one-form θ of TQ^* is given by the pull-back:

$$\theta = (\mathrm{d}^*\varphi)^*\Big(\sum_{i=1}^n p_i \mathrm{d}q^i\Big).$$

The other possibility is to use (1.9.7) to denote the coordinate functions of the map $\mathrm{d}^*\varphi$, so that $\mathrm{d}^*\varphi = (q^1, \ldots, q^n, p_1, \ldots, p_n)$ and (1.9.7) denote real valued maps over the open subset TU^* of TQ^*. Under such notation, the (restriction to TU^* of) the canonical one-form θ of TQ^* is given exactly by the righthand side of equality (1.9.3).

[20] Just differentiate the arrows of diagram (1.9.5) and follow the definitions!

[21] Obviously, the canonical one-form of the cotangent bundle TU^* of the open subset U of Q is the restriction to TU^* of the canonical one-form of the cotangent bundle TQ^*.

1.9.3. Definition. Given a differentiable manifold Q, then the *canonical symplectic form* of the cotangent bundle TQ^* is defined by:

$$\omega = -\mathrm{d}\theta, \tag{1.9.8}$$

where θ denotes the canonical one-form of TQ^*.

We use a minus sign in (1.9.8) because we want ω to agree with the canonical symplectic form of $\mathbb{R}^n \times \mathbb{R}^{n*}$ when $Q = \mathbb{R}^n$. We will see in a moment that ω is indeed a symplectic form over TQ^*. It is already clear that it is smooth (because θ is smooth) and that it is closed (because it is exact). If Q, $\widetilde{Q}$ are differentiable manifolds and $\varphi : Q \to \widetilde{Q}$ is a smooth diffeomorphism then, by (1.9.6), since pull-backs commute with exterior differentiation, it follows that:

$$(\mathrm{d}^*\varphi)^*\widetilde{\omega} = \omega,$$

where ω denotes the canonical symplectic form of TQ^* and $\widetilde{\omega}$ the canonical symplectic form of $T\widetilde{Q}^*$.

1.9.4. Example. If Q is an open subset of $\mathbb{R}^n$ then, using the same notation used in Example 1.9.2, we see (by taking the exterior derivative on both sides of (1.9.3)) that the canonical symplectic form ω of TQ^* is given by:

$$\omega = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i, \tag{1.9.9}$$

and therefore it agrees with the canonical symplectic form (1.7.3) of the space $\mathbb{R}^n \times \mathbb{R}^{n*}$. If $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* = \mathbb{R} \times Q \times \mathbb{R}^{n*} \to \mathbb{R}$ is a time-dependent Hamiltonian over Q and if TQ^* is endowed with its canonical symplectic form ω then the symplectic gradient of H is given by:

$$\mathrm{dom}(H) \ni (t, q, p) \longmapsto \Big(\frac{\partial H}{\partial p}(t, q, p), -\frac{\partial H}{\partial q}(t, q, p)\Big) \in \mathbb{R}^n \times \mathbb{R}^{n*}.$$

Given a smooth curve $(q, p) : I \to TQ^*$ (defined over some interval $I \subset \mathbb{R}$) then, for $t \in I$ with $\big(t, q(t), p(t)\big) \in \mathrm{dom}(H)$, the condition:

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(q(t), p(t)\big) = \vec{H}\big(t, q(t), p(t)\big)$$

is equivalent to (the satisfaction at t of) Hamilton's equations:

$$\frac{\mathrm{d}q}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p}\big(t, q(t), p(t)\big), \quad \frac{\mathrm{d}p}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q}\big(t, q(t), p(t)\big). \tag{1.9.10}$$

If $\varphi : U \subset Q \to \widetilde{U} \subset \mathbb{R}^n$ is a local chart on Q then the (restriction to TU^* of the) symplectic form ω of TQ^* is the pull-back by the local chart $\mathrm{d}^*\varphi$ of the (restriction to $\widetilde{U} \times \mathbb{R}^{n*}$ of the) canonical symplectic form of $\mathbb{R}^n \times \mathbb{R}^{n*}$. It follows that the bilinear form $\omega_{(q,p)}$ on $T_{(q,p)}TQ^*$ is non degenerate, for all $(q, p) \in TQ^*$.

Again, we have two possibilities for notation: use (1.9.7) to denote the projections of $\widetilde{U} \times \mathbb{R}^{n*} \subset \mathbb{R}^n \times \mathbb{R}^{n*}$, so that the (restriction to TU^* of the) canonical symplectic form ω of TQ^* is the pull-back by $\mathrm{d}^*\varphi$ of the righthand side of (1.9.9) or use (1.9.7) to denote the coordinate functions of the map $\mathrm{d}^*\varphi$, so that the righthand side of (1.9.9) is precisely the (restriction to TU^* of the) canonical symplectic form ω of TQ^*.

Our considerations so far have proven the following:

1.9.5. Proposition. *Given a differentiable manifold Q then:*

(a) *the canonical symplectic form of TQ^* is indeed a symplectic form;*

(b) *if $\widetilde{Q}$ is another differentiable manifold and if $\varphi : Q \to \widetilde{Q}$ is a smooth diffeomorphism then $\mathrm{d}^*\varphi : TQ^* \to T\widetilde{Q}^*$ is a symplectomorphism if both tangent bundles are endowed with their canonical symplectic forms;*

(c) *if $\varphi : U \subset Q \to \widetilde{U} \subset \mathbb{R}^n$ is a local chart on Q then the local chart $\mathrm{d}^*\varphi$ on TQ^* is symplectic if TQ^* is endowed with its canonical symplectic form.* □

Let $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ be a time-dependent Hamiltonian over the cotangent bundle TQ^* of a differentiable manifold Q. Given a local chart $\varphi : U \subset Q \to \widetilde{U} \subset \mathbb{R}^n$ on Q then, since the local chart $\mathrm{d}^*\varphi$ on TQ^* is symplectic, we have that a curve on TQ^* is an integral curve of the symplectic gradient $\vec{H}$ if and only if the representation of such curve with respect to the chart $\mathrm{d}^*\varphi$ is a solution of Hamilton's equations (1.9.10) corresponding to the Hamiltonian that represents H with respect to the chart $\mathrm{d}^*\varphi$ (see Exercise 1.31 for details).

The first of Hamilton's equations can be formulated without the aid of a coordinate chart. More explicitly, given $t \in \mathbb{R}$, $q \in Q$, we can freeze the first two variables of H, obtaining a map $H(t, q, \cdot)$ that sends each p in the open subset:

$$\mathrm{dom}\big(H(t, q, \cdot)\big) = \big\{p \in T_qQ^* : (t, q, p) \in \mathrm{dom}(H)\big\}$$

of the cotangent space T_qQ^* to the real number $H(t, q, p)$. The differential of the map $H(t, q, \cdot)$ at a point p of its domain will be denoted by $\frac{\partial H}{\partial p}(t, q, p)$ and it is a linear functional over T_qQ^*, i.e., it is an element of the bidual T_qQ^{**}, which we identify with an element of the tangent space T_qQ. Unfortunately, the partial derivative $\frac{\partial H}{\partial q}(t, q, p)$ does not make sense for a time-dependent Hamiltonian H on a cotangent bundle, as one can fix q and move p, but one cannot fix p and move q (one would need a connection[22] to make sense out of $\frac{\partial H}{\partial q}(t, q, p)$ without the aid of a local chart).

[22] A connection on Q induces a direct sum decomposition of the tangent space $T_{(q,p)}TQ^*$ into a *vertical subspace* (the space tangent to the fiber) and a *horizontal subspace* (which is defined by the connection). The derivative $\frac{\partial H}{\partial p}(t, q, p)$ is the differential

The following proposition simply says that if a curve $t \mapsto \big(q(t), p(t)\big)$ on a cotangent bundle satisfies Hamilton's equations (formulated intrinsically, in terms of the symplectic gradient with respect to the canonical symplectic form of the cotangent bundle) then it also satisfies the first of Hamilton's equations (which can be formulated without the aid of a coordinate chart).

1.9.6. PROPOSITION. *If a smooth curve* $(q, p) : I \to TQ^*$ *(defined on some interval* $I \subset \mathbb{R}$*) satisfies* $\big(t, q(t), p(t)\big) \in \mathrm{dom}(H)$ *and:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(q(t), p(t)\big) = \vec{H}\big(t, q(t), p(t)\big) \tag{1.9.11}$$

for a certain $t \in I$ *then:*

$$\frac{\mathrm{d}q}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p}\big(t, q(t), p(t)\big) \in T_{q(t)}Q. \tag{1.9.12}$$

PROOF. If two manifolds Q, $\widetilde{Q}$ are diffeomorphic and if the thesis of the proposition holds for $\widetilde{Q}$ then it also holds for Q (if this is not obvious to you, see Exercise 1.34). Moreover, in order to prove the thesis for a certain manifold Q, we can replace Q with an open neighborhood of the point $q(t)$ in Q. Since a sufficiently small neighborhood of a point of Q is diffeomorphic to an open subset of $\mathbb{R}^n$, it suffices to prove the proposition in the case when Q is an open subset of $\mathbb{R}^n$. In that case, condition (1.9.11) means that the curve (q, p) satisfies Hamilton's equations (1.9.10) (at the given instant t) and condition (1.9.12) means that the curve (q, p) satisfies the first of Hamilton's equations (at the given instant t), so the conclusion is obvious. □

The method that we used to prove Proposition 1.9.6 is a very nice strategy for proving theorems about differentiable manifolds. It works as follows: suppose that we want to prove a certain statement about differentiable manifolds[23]. For each differentiable manifold Q, the statement may or may not hold for Q, i.e., the statement defines a subclass $\mathfrak{C}$ of the class of all differentiable manifolds. Proving the statement means to prove that $\mathfrak{C}$ is the class of all differentiable manifolds. For a statement that *really is* a statement about differentiable manifolds (i.e., it concerns the manifold structure alone), it should be the case that if a manifold diffeomorphic to Q is in $\mathfrak{C}$ than also Q is in $\mathfrak{C}$. Let us say in this case that the class $\mathfrak{C}$ is *closed under diffeomorphisms*. Proving that the class $\mathfrak{C}$ is closed under diffeomorphisms normally is a completely follow-your-nose type of exercise, i.e., one just needs to use a given diffeomorphism to keep carrying things over from one manifold to the other. We say that the statement under consideration is

of H along the vertical subspace and the derivative $\frac{\partial H}{\partial q}(t, q, p)$ can be defined as the differential of H along the horizontal subspace (which, obviously, depends on the choice of connection).

[23]More generally, the statement might be about a family of manifolds. The strategy can be easily adapted for that situation as well.

local (or that the class $\mathfrak{C}$ is local) if the class $\mathfrak{C}$ has, in addition, the following property: given a differentiable manifold Q, if every point of Q has an open neighborhood in Q that (regarded as a manifold in its own right) belongs to $\mathfrak{C}$, then Q belongs to $\mathfrak{C}$. Normally, when a statement is local, it is very easy to check that it really is local. Obviously, if $\mathfrak{C}$ is closed under diffeomorphisms and local then, if every open subset of $\mathbb{R}^n$ is in $\mathfrak{C}$, it follows that every differentiable manifold is in $\mathfrak{C}$. We have therefore the following strategy for proving a local statement about differentiable manifolds: (i) observe that the class $\mathfrak{C}$ defined by the statement is closed under diffeomorphisms; (ii) check that the class $\mathfrak{C}$ is local; (iii) prove the statement for open subsets of $\mathbb{R}^n$. There is nothing spectacular about this strategy, but it allows one to focuss on what is really important, i.e., proving the statement for open subsets of $\mathbb{R}^n$. Many authors tend to present proofs of theorems about differentiable manifolds in which the proof of the statement for open subsets of $\mathbb{R}^n$ gets mixed up with a lot of procedures of carrying things around using coordinate charts. Our strategy relieves the mind from such distractions.

1.10. The Legendre transform

The Legendre transform is the procedure that is used to transform a Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$ into a Hamiltonian $H : \mathbb{R} \times TQ^* \to \mathbb{R}$. This procedure is performed *fiberwise*: for each $t \in \mathbb{R}$ and each $q \in Q$, we freeze the first two variables of L, obtaining a map $L(t, q, \cdot)$ over the tangent space T_qQ; on such map, we perform a certain procedure (which we also call "Legendre transform") that yields the map $H(t, q, \cdot)$ over the cotangent space T_qQ^*. We start by studying the procedure which is applied to the map $L(t, q, \cdot)$; that is a general procedure that can be applied to real valued maps over a real finite-dimensional vector space.

In what follows, E denotes a fixed real finite-dimensional vector space. Let $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ be a map of class C^2 defined over an open subset $\mathrm{dom}(f)$ of E. The differential $\mathrm{d}f(x)$ of f at a point $x \in \mathrm{dom}(f)$ is a linear functional over E, i.e., an element of the dual space E^*. The differential $\mathrm{d}f$ is therefore a map (of class C^1):

$$\mathrm{d}f : \mathrm{dom}(f) \subset E \longrightarrow E^* \tag{1.10.1}$$

from the open subset $\mathrm{dom}(f)$ of E to the dual space E^*.

1.10.1. Definition. The map $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ is called *regular* if its differential (1.10.1) is a local diffeomorphism and it is called *hyper-regular* if its differential is a diffeomorphism onto an open subset of E^* (the image $\mathrm{Im}(\mathrm{d}f)$ of the map $\mathrm{d}f$).

It follows from the inverse function theorem that f is regular if and only if its second differential:

$$\mathrm{d}(\mathrm{d}f)(x) : E \longrightarrow E^*$$

is a linear isomorphism for all $x \in \mathrm{dom}(f) \subset E$. Moreover, f is hyper-regular if and only if f is regular and the map $\mathrm{d}f$ is injective.

Given a map $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ of class C^2 over an open subset $\mathrm{dom}(f)$ of E, we consider the map $\phi : \mathrm{dom}(f) \subset E \to \mathbb{R}$ (of class C^1) defined by:

$$\phi(x) = \mathrm{d}f(x)x - f(x), \quad x \in \mathrm{dom}(f). \tag{1.10.2}$$

The expression $\mathrm{d}f(x)x$ denotes the evaluation of the linear functional $\mathrm{d}f(x)$ at the vector x.

1.10.2. Definition. If $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ is a hyper-regular map of class C^2 then its *Legendre transform* is the map $f^* : \mathrm{dom}(f^*) \subset E^* \to \mathbb{R}$ defined by:

$$f^* = \phi \circ (\mathrm{d}f)^{-1},$$

with ϕ given by (1.10.2). The domain $\mathrm{dom}(f^*)$ of f^* is the image $\mathrm{Im}(\mathrm{d}f)$ of the map $\mathrm{d}f$.

The map f^* looks like just a map of class C^1 (since both ϕ and the inverse of $\mathrm{d}f$ are maps of class C^1), but, curiously, we will prove below that the map f^* is of class C^2. In any case, we are really interested only in the case when f is smooth and clearly in that case f^* is smooth as well. The reader might be a little puzzled by Definition 1.10.2. It seems as if the formula for f^* just fell from the sky. Many authors define the Legendre transform in terms of the solution to a maximization problem[24], but we do not. In a moment, the reader should be convinced that the Legendre transform is a very clever construction.

1.10.3. Example. Consider the Lagrangian:

$$L : \mathrm{dom}(V) \times (\mathbb{R}^3)^n \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n \to \mathbb{R}$$

defined in Subsection 1.4.1. Its Euler–Lagrange equation is the dynamical equation for a system of n particles subject to (the forces with) a potential $V : \mathrm{dom}(V) \subset \mathbb{R}\times(\mathbb{R}^3)^n \to \mathbb{R}$. Given $(t, q) \in \mathrm{dom}(V) \subset \mathbb{R}\times(\mathbb{R}^3)^n$, consider the map $f : (\mathbb{R}^3)^n \to \mathbb{R}$ obtained by freezing the first two variables of L:

$$f(\dot{q}) = L(t, q, \dot{q}), \quad \dot{q} \in (\mathbb{R}^3)^n.$$

[24]One defines the value of the Legendre transform f^* at a linear functional $\alpha \in E^*$ to be the maximum of $\alpha(x) - f(x)$, with x running through the domain of f. When f is convex, such maximum is attained precisely at the point x such that $\mathrm{d}f(x) = \alpha$ (if such point exists). Therefore, in that case, the definition using the maximization problem agrees with ours. The maximization problem might create a sense of motivation on some readers, but it is nevertheless a bad idea: it creates an illusion of difficulty in the case when f is not convex. It is an "illusion of difficulty", because the difficulty is being created by the bad choice of definition alone! With our definition, there are no difficulties when f is not convex and all the nice properties of the Legendre transform can be (easily) proven as well.

The differential of f at a point $\dot q$ is the partial differential $\frac{\partial L}{\partial \dot q}(t,q,\dot q)$ of L with respect to its third variable and it is an element of the dual space $(\mathbb{R}^3)^{n*}$; identifying $(\mathbb{R}^3)^{n*}$ with $(\mathbb{R}^3)^n$ in the standard way, we obtain:

$$\mathrm{d}f(\dot q) = \big(\tfrac{\partial L}{\partial \dot q_1}(t,q,\dot q),\ldots,\tfrac{\partial L}{\partial \dot q_n}(t,q,\dot q)\big) = (m_1\dot q_1,\ldots,m_n\dot q_n) \in (\mathbb{R}^3)^n,$$

i.e., $\mathrm{d}f(\dot q)$ is the vector $p=(p_1,\ldots,p_n)$ containing the momenta of all the particles (more precisely, p becomes the vector containing the momenta of all the particles when $\dot q$ is replaced with the derivative $\dot q(t)$ of the curve $t\mapsto q(t)$ describing the trajectories of all the particles). We have:

$$\mathrm{d}f(\dot q)\dot q = \frac{\partial L}{\partial \dot q}(t,q,\dot q)\dot q = \sum_{j=1}^n m_j\|\dot q_j\|^2$$

and therefore the map ϕ corresponding to f as in (1.10.2) is given by:

$$\phi(\dot q) = \frac{\partial L}{\partial \dot q}(t,q,\dot q)\dot q - L(t,q,\dot q) = \sum_{j=1}^n \frac{1}{2}m_j\|\dot q_j\|^2 + V(t,q).$$

Thus, ϕ yields the total energy (after q is replaced with $q(t)$ and $\dot q$ is replaced with $\dot q(t)$). The Legendre transform $f^*=\phi\circ(\mathrm{d}f)^{-1}$ is just the map ϕ written in terms of p, instead of $\dot q$:

$$f^*(p) = \sum_{j=1}^n \frac{\|p_j\|^2}{2m_j} + V(t,q).$$

Setting $H(t,q,p)=f^*(p)$ (with f defined by $f(\dot q)=L(t,q,\dot q)$), we see that H is precisely the Hamiltonian (1.6.5) of Classical Mechanics.

This is the first good news: the Legendre transform turns the Lagrangian of Classical Mechanics into the Hamiltonian of Classical Mechanics. In a moment we are going to see what happens in the case of Classical Mechanics with constraints (Subsection 1.10.1). But, first, let us prove a few more properties of the Legendre transform.

If $f:\mathrm{dom}(f)\subset E\to\mathbb{R}$ is a hyper-regular map of class C^2 then the differential of its Legendre transform $f^*:\mathrm{dom}(f^*)\subset E^*\to\mathbb{R}$ is a continuous map:

$$\mathrm{d}f^*:\mathrm{dom}(f^*)\subset E^*\to E^{**} \tag{1.10.3}$$

taking values in the bidual space E^{**}. We have the following:

1.10.4. Lemma. *Under the standard identification of E^{**} with E, the differential* (1.10.3) *of the Legendre transform f^* of a hyper-regular map $f:\mathrm{dom}(f)\subset E\to\mathbb{R}$ of class C^2 is equal to the inverse* $(\mathrm{d}f)^{-1}$ *of the differential of the map f.*

Proof. Let $\alpha\in\mathrm{dom}(f^*)=\mathrm{Im}(\mathrm{d}f)$ be given and let $x\in\mathrm{dom}(f)$ be such that $\mathrm{d}f(x)=\alpha$. We have to show that $\mathrm{d}f^*(\alpha)$ is the element of E^{**} that is identified with x, i.e., that:

$$\mathrm{d}f^*(\alpha)\dot\alpha = \dot\alpha(x), \tag{1.10.4}$$

for all $\dot{\alpha} \in E^*$. Since $f^* = \phi \circ (\mathrm{d}f)^{-1}$ (with ϕ defined in (1.10.2)), we have:

$$\mathrm{d}f^*(\alpha)\dot{\alpha} = \mathrm{d}\phi(x)\big[\mathrm{d}\big((\mathrm{d}f)^{-1}\big)(\alpha)\dot{\alpha}\big]. \tag{1.10.5}$$

Given $\dot{x} \in E$, we differentiate both sides of (1.10.2) with respect to x and evaluate at $\dot{x}$ (for the differentiation of the term $\mathrm{d}f(x)x$ we use the product rule[25]). The result is:

$$\mathrm{d}\phi(x)\dot{x} = \big(\mathrm{d}f(x)\dot{x}\big)x + \mathrm{d}f(x)\dot{x} - \mathrm{d}f(x)\dot{x} = \big(\mathrm{d}f(x)\dot{x}\big)x. \tag{1.10.6}$$

Now set $\dot{x} = \mathrm{d}\big((\mathrm{d}f)^{-1}\big)(\alpha)\dot{\alpha}$ in (1.10.6) in order to compute the righthand side of (1.10.5). Notice that, since the maps $\mathrm{d}f$ and $(\mathrm{d}f)^{-1}$ are mutually inverse, the differentials $\mathrm{d}f(x)$ and $\mathrm{d}\big((\mathrm{d}f)^{-1}\big)(\alpha)$ are also mutually inverse, so that:

$$\mathrm{d}f(x)\dot{x} = \mathrm{d}f(x)\big[\mathrm{d}\big((\mathrm{d}f)^{-1}\big)(\alpha)\dot{\alpha}\big] = \dot{\alpha}.$$

It follows that (1.10.4) holds. □

1.10.5. COROLLARY. *If f is a hyper-regular map of class C^2 then so is its Legendre transform.*

PROOF. Since $\mathrm{d}f$ is a diffeomorphism of class C^1, so is its inverse $(\mathrm{d}f)^{-1}$, which, by the lemma, is identified with $\mathrm{d}f^*$. It follows that f^* is a map of class C^2 and that f^* is hyper-regular. □

Since the Legendre transform of a hyper-regular map of class C^2 is again a hyper-regular map of class C^2, it makes sense to take the Legendre transform f^{**} of the Legendre transform f^*. It turns out that f^{**} is just f.

1.10.6. PROPOSITION. *Under the standard identification of E^{**} with E, the Legendre transform f^{**} of the Legendre transform f^* of a hyper-regular map $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ of class C^2 is equal to f (i.e., the Legendre transform is* involutive*).*

PROOF. Let $\psi : \mathrm{dom}(f^*) \subset E^* \to \mathbb{R}$ be the analogue of the map (1.10.2) for f^*, i.e., ψ is defined by:

$$\psi(\alpha) = \mathrm{d}f^*(\alpha)\alpha - f^*(\alpha), \quad \alpha \in \mathrm{dom}(f^*) = \mathrm{Im}(\mathrm{d}f).$$

The double Legendre transform f^{**} is equal to $\psi \circ (\mathrm{d}f^*)^{-1}$. Let $x \in \mathrm{dom}(f)$ be fixed and set $\alpha = \mathrm{d}f(x)$. By Lemma 1.10.4, $(\mathrm{d}f^*)^{-1} = \mathrm{d}f$ (up to the identification between E^{**} and E) and therefore:

$$f^{**}(x) = \psi(\alpha).$$

---

[25]Here is how you do this: if $(v, w) \mapsto v \cdot w$ denotes any bilinear map, then the differential of a map $p(x) = v(x) \cdot w(x)$ is given by:

$$\mathrm{d}p(x)\dot{x} = \big(\mathrm{d}v(x)\dot{x}\big) \cdot w(x) + v(x) \cdot \big(\mathrm{d}w(x)\dot{x}\big).$$

In order to prove that, think about p as the composite of the map $x \mapsto \big(v(x), w(x)\big)$ with the bilinear map that is denoted by the dot; use the chain rule and the standard formula for the differential of a bilinear map. In the case under consideration, v is the map $\mathrm{d}f$, w is the identity map $x \mapsto x$ and the bilinear map $E^* \times E \to \mathbb{R}$ is the evaluation map.

Since $\mathrm{d}f^* = (\mathrm{d}f)^{-1}$, we obtain that $\mathrm{d}f^*(\alpha)$ is equal to (the element of E^{**} that is identified with) x and therefore:

$$\psi(\alpha) = \mathrm{d}f^*(\alpha)\alpha - f^*(\alpha) = \alpha(x) - f^*(\alpha) = \alpha(x) - \phi(x),$$

with ϕ defined in (1.10.2). But $\phi(x) = \alpha(x) - f(x)$ and the conclusion follows. $\square$

Now we are ready to study the Legendre transform of a Lagrangian. Let Q be a differentiable manifold and L be a Lagrangian on Q, i.e., L is a smooth real valued map defined over an open subset $\mathrm{dom}(L)$ of $\mathbb{R} \times TQ$. For each $(t, q) \in \mathbb{R} \times Q$, we obtain a map $L(t, q, \cdot)$ by freezing the first two variables of L, i.e., $L(t, q, \cdot)$ is the map $\dot{q} \mapsto L(t, q, \dot{q})$ defined over the open subset:

$$\mathrm{dom}\big(L(t, q, \cdot)\big) = \big\{\dot{q} \in T_qQ : (t, q, \dot{q}) \in \mathrm{dom}(L)\big\}$$

of the tangent space T_qQ. The differential of the map $L(t, q, \cdot)$ at a point $\dot{q} \in \mathrm{dom}\big(L(t, q, \cdot)\big)$ will be denoted by $\frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})$ and it is an element of the cotangent space T_qQ^*.

1.10.7. Definition. The Lagrangian L is said to be *regular* (resp., *hyper-regular*) if the map $L(t, q, \cdot)$ is regular (resp., hyper-regular) for all (t, q) in $\mathbb{R} \times Q$.

In other words, the Lagrangian L is regular if the map:

$$T_qQ \supset \mathrm{dom}\big(L(t, q, \cdot)\big) \ni \dot{q} \longmapsto \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q}) \in T_qQ^* \tag{1.10.7}$$

is a local diffeomorphism for all $(t, q) \in \mathbb{R} \times Q$ and the Lagrangian L is hyper-regular if the map (1.10.7) is a diffeomorphism onto an open subset of T_qQ^*, for all $(t, q) \in \mathbb{R} \times Q$. We can join all the maps (1.10.7) with (t, q) running over $\mathbb{R} \times Q$, obtaining the so called *fiber derivative* of L, which is the map:

$$\mathbb{F}L : \mathrm{dom}(L) \subset \mathbb{R} \times TQ \longrightarrow \mathbb{R} \times TQ^*$$

defined by:

$$\mathbb{F}L(t, q, \dot{q}) = \Big(t, q, \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})\Big), \quad (t, q, \dot{q}) \in \mathrm{dom}(L).$$

The conditions of regularity and hyper-regularity of L can be stated in terms of the fiber derivative $\mathbb{F}L$.

1.10.8. Lemma. *The Lagrangian L is regular if and only if its fiber derivative $\mathbb{F}L$ is a local diffeomorphism and the Lagrangian L is hyper-regular if and only if its fiber derivative $\mathbb{F}L$ is a diffeomorphism onto an open subset of $\mathbb{R} \times TQ^*$ (such open subset is obviously the image of $\mathbb{F}L$).*

Proof. It is obvious that $\mathbb{F}L$ is injective if and only if the map (1.10.7) is injective for all $(t, q) \in \mathbb{R} \times Q$. Since a smooth map is a diffeomorphism onto an open subset of its counter-domain if and only if the map is an injective local diffeomorphism, it follows that in order to prove the lemma it suffices to prove the first part of its statement, i.e., to prove that L is regular if and only if $\mathbb{F}L$ is a local diffeomorphism. We employ the strategy discussed at the end of Section 1.9. Observe that:

(a) if two manifolds Q, $\widetilde{Q}$ are diffeomorphic and if the thesis (that a Lagrangian L on the manifold is regular if and only if the map $\mathbb{F}L$ is a local diffeomorphism) holds for the manifold $\widetilde{Q}$ then it also holds for the manifold Q (if this is not obvious to you, see Exercise 1.36);
(b) if every point of Q has an open neighborhood U in Q such that the thesis holds for the manifold U then the thesis holds for Q.

Because of (a) and (b), it suffices to prove the thesis if Q is an open subset of $\mathbb{R}^n$. In that case, the thesis follows easily from the inverse function theorem (see Exercise 1.37). □

1.10.9. Definition. If $L : \mathrm{dom}(L) \subset \mathbb{R} \times TQ \to \mathbb{R}$ is a hyper-regular Lagrangian, then its *Legendre transform* is the map:

$$L^* : \mathrm{dom}(L^*) = \mathrm{Im}(\mathbb{F}L) \subset \mathbb{R} \times TQ^* \longrightarrow \mathbb{R}$$

defined by:

$$L^*(t,q,p) = \big(L(t,q,\cdot)\big)^*(p), \quad (t,q,p) \in \mathrm{dom}(L^*) = \mathrm{Im}(\mathbb{F}L) \subset \mathbb{R} \times TQ^*,$$

where $\big(L(t,q,\cdot)\big)^*(p)$ denotes the value at the point $p \in T_qQ^*$ of the Legendre transform $\big(L(t,q,\cdot)\big)^*$ of the map $L(t,q,\cdot)$. The map L^* is also denoted by H and it is also called the *Hamiltonian* associated to the Lagrangian L.

It follows straightforwardly from the definition above that the Hamiltonian $H : \mathrm{dom}(H) = \mathrm{dom}(L^*) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ associated to a hyper-regular Lagrangian L satisfies:

$$H\big(\mathbb{F}L(t,q,\dot q)\big) = H\Big(t,q,\frac{\partial L}{\partial \dot q}(t,q,\dot q)\Big) = \frac{\partial L}{\partial \dot q}(t,q,\dot q)\dot q - L(t,q,\dot q),$$

for all $(t,q,\dot q) \in \mathrm{dom}(L)$, i.e., the Hamiltonian H is the composition of the map:

$$\mathbb{R} \times TQ \supset \mathrm{dom}(L) \ni (t,q,\dot q) \longmapsto \frac{\partial L}{\partial \dot q}(t,q,\dot q)\dot q - L(t,q,\dot q) \in \mathbb{R} \tag{1.10.8}$$

with the inverse of the map $\mathbb{F}L$. Lemma 1.10.8 implies that the domain of H is open and that H is smooth (since the map (1.10.8) is clearly smooth). We can also define a *fiber derivetive*:

$$\mathbb{F}H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \longrightarrow \mathbb{R} \times TQ$$

for H by setting:

$$\mathbb{F}H(t,q,p) = \Big(t,q,\frac{\partial H}{\partial p}(t,q,p)\Big) \in \mathbb{R} \times TQ, \quad (t,q,p) \in \mathrm{dom}(H).$$

The following proposition is an immediate consequence of the results that we have proven about the Legendre transform of maps on real finite-dimensional vector spaces.

1.10.10. PROPOSITION. *Let* $L : \mathrm{dom}(L) \subset \mathbb{R} \times TQ \to \mathbb{R}$ *be a hyper-regular Lagrangian and* $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ *be its Legendre transform. Then:*

(a) *for each* $(t, q) \in \mathbb{R} \times Q$*, the map* $H(t, q, \cdot)$ *is hyper-regular and its differential:*

$$T_qQ^* \supset \mathrm{dom}\big(H(t, q, \cdot)\big) \ni p \longmapsto \frac{\partial H}{\partial p}(t, q, p) \in \mathrm{dom}\big(L(t, q, \cdot)\big) \subset T_qQ$$

is the inverse of the map (1.10.7)*;*

(b) *the fiber derivatives* $\mathbb{F}L$*,* $\mathbb{F}H$ *are mutually inverse smooth diffeomorphisms;*

(c) *for each* $(t, q) \in \mathbb{R} \times Q$*, the Legendre transform of the map* $H(t, q, \cdot)$ *is equal to the map* $L(t, q, \cdot)$*.*

PROOF. Item (a) follows from Lemma 1.10.4, item (b) follows from item (a) and from Lemma 1.10.8 and item (c) follows from Proposition 1.10.6. □

For a Hamiltonian that is the Legendre transform of a hyper-regular Lagrangian, Proposition 1.9.6 yields the following:

1.10.11. PROPOSITION. *Under the conditions of Proposition 1.10.10, if a smooth curve* $(q, p) : I \to TQ^*$ *(defined on some interval* $I \subset \mathbb{R}$*) satisfies* $\big(t, q(t), p(t)\big) \in \mathrm{dom}(H)$ *and:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(q(t), p(t)\big) = \vec{H}\big(t, q(t), p(t)\big) \tag{1.10.9}$$

for a certain $t \in I$ *then* $\big(t, q(t), \dot{q}(t)\big) \in \mathrm{dom}(L)$ *and:*

$$p(t) = \frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big). \tag{1.10.10}$$

PROOF. Follows immediately from Proposition 1.9.6 and item (a) of Proposition 1.10.10. □

Proposition 1.10.11 says that if H is the Legendre transform of a hyper-regular Lagrangian L then, for a solution (q, p) to Hamilton's equations, the p is determined from the q by equality (1.10.10).

Finally, we prove that when H is the Legendre transform of L, the Euler–Lagrange equation and Hamilton's equations have the same solutions; more precisely, a map q is a solution to the Euler–Lagrange equation if and only if it is a solution to Hamilton's equations, with p defined by (1.10.10). The statement that a curve q on a manifold Q be a solution to the Euler–Lagrange equation is to be understood in terms of coordinate charts. Recall from Section 1.5 that if $\varphi : U \subset Q \to \widetilde{U} \subset \mathbb{R}^n$ is a local chart on Q and $L : \mathrm{dom}(L) \subset \mathbb{R} \times TQ \to \mathbb{R}$ is a Lagrangian on Q then the representation of L with respect to φ is the Lagrangian L_φ with domain (Id denotes the identity map of $\mathbb{R}$):

$$\mathrm{dom}(L_\varphi) = (\mathrm{Id}\times \mathrm{d}\varphi)\big(\mathrm{dom}(L)\cap(\mathbb{R}\times TU)\big)\subset \mathbb{R}\times T\widetilde{U} = \mathbb{R}\times\widetilde{U}\times\mathbb{R}^n$$

defined by:

$$L_\varphi\big(t,\varphi(q),\mathrm{d}\varphi_q(\dot q)\big) = L(t,q,\dot q),$$

for all $(t,q,\dot q)\in\mathrm{dom}(L)\cap(\mathbb{R}\times TU)$.

1.10.12. Theorem. *Under the conditions of Proposition 1.10.10, assume that we are given a smooth curve $q : I\to Q$ (defined over some interval $I\subset\mathbb{R}$) and a local chart $\varphi : U\subset Q\to\widetilde{U}\subset\mathbb{R}^n$. For those $t\in I$ with $\big(t,q(t),\dot q(t)\big)\in\mathrm{dom}(L)$, define $p(t)\in T_{q(t)}Q^*$ by equality* (1.10.10)*. If L_φ denotes the representation of L with respect to the chart φ then, for all $t\in I$ with $q(t)\in U$ and $\big(t,q(t),\dot q(t)\big)\in\mathrm{dom}(L)$, the following conditions are equivalent:*

(a) *the Euler–Lagrange equation:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_\varphi}{\partial\dot q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big) = \frac{\partial L_\varphi}{\partial q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big) \tag{1.10.11}$$

holds, where $\tilde q = \varphi\circ q|_{q^{-1}(U)}$;

(b) *equality* (1.10.9) *(Hamilton's equations) holds.*

Let us first prove the equivalence between the Euler–Lagrange equation and Hamilton's equations when $Q=\mathbb{R}^n$. Theorem 1.10.12 will then follow easily.

1.10.13. Lemma. *Let $L : \mathrm{dom}(L)\subset\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be a hyper-regular Lagrangian and $H : \mathrm{dom}(H)\subset\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^{n*}\to\mathbb{R}$ be its Legendre transform. Let $q : I\to\mathbb{R}^n$ be a smooth curve (defined over some interval $I\subset\mathbb{R}$) and, for all $t\in I$ with $\big(t,q(t),\dot q(t)\big)\in\mathrm{dom}(L)$, define $p(t)$ by equality* (1.10.10)*. Then, for all $t\in I$ with $\big(t,q(t),\dot q(t)\big)\in\mathrm{dom}(L)$, we have:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot q}\big(t,q(t),\dot q(t)\big) = \frac{\partial L}{\partial q}\big(t,q(t),\dot q(t)\big) \tag{1.10.12}$$

if and only if:

$$\frac{\mathrm{d}q}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p}\big(t,q(t),p(t)\big),\quad \frac{\mathrm{d}p}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q}\big(t,q(t),p(t)\big). \tag{1.10.13}$$

Proof. Given $t\in I$, since the map $H\big(t,q(t),\cdot\big)$ is the Legendre transform of the map $L\big(t,q(t),\cdot\big)$, it follows from Lemma 1.10.4 that their differentials are mutually inverse. Thus (1.10.10) implies that the first equation in (1.10.13) is satisfied. Moreover, (1.10.10) implies that the lefthand side of the second equation in (1.10.13) equals the lefthand side of the Euler–Lagrange equation (1.10.12). Thus, the proof will be completed once we check that:

$$\frac{\partial H}{\partial q}\big(t,q(t),p(t)\big) = -\frac{\partial L}{\partial q}\big(t,q(t),\dot q(t)\big).$$

By the definition of the Legendre transform (we have used above q, $\dot q$ for names of curves and now we are going to use them also for names of points):

$$H\Big(t, q, \frac{\partial L}{\partial \dot q}(t, q, \dot q)\Big) = \frac{\partial L}{\partial \dot q}(t, q, \dot q)\dot q - L(t, q, \dot q), \tag{1.10.14}$$

for all $(t, q, \dot q) \in \mathrm{dom}(L)$. Differentiating both sides of (1.10.14) with respect to q and evaluating at a vector $v \in \mathbb{R}^n$, we obtain:

$$\begin{aligned}\frac{\partial H}{\partial q}(t, q, p)v + \frac{\partial H}{\partial p}(t, q, p)\Big(\frac{\partial^2 L}{\partial q \partial \dot q}(t, q, \dot q)v\Big)&\\ = \Big(\frac{\partial^2 L}{\partial q \partial \dot q}(t, q, \dot q)v\Big)\dot q - \frac{\partial L}{\partial q}(t, q, \dot q)v&,\end{aligned}$$

where $p = \frac{\partial L}{\partial \dot q}(t, q, \dot q)$. Since the differentials of the maps $L(t, q, \cdot)$ and $H(t, q, \cdot)$ are mutually inverse, we obtain that $\frac{\partial H}{\partial p}(t, q, p)$ is (the element of the bidual of $\mathbb{R}^n$ that is identified with the vector) $\dot q \in \mathbb{R}^n$ and therefore:

$$\frac{\partial H}{\partial p}(t, q, p)\Big(\frac{\partial^2 L}{\partial q \partial \dot q}(t, q, \dot q)v\Big) = \Big(\frac{\partial^2 L}{\partial q \partial \dot q}(t, q, \dot q)v\Big)\dot q.$$

The conclusion follows. □

Proof of Theorem 1.10.12. Denote by H_Φ the representation of the Hamiltonian H with respect to the local chart $\Phi = \mathrm{d}^*\varphi$ of TQ^*, i.e., the domain of H_Φ is the image of $\mathrm{dom}(H) \cap (\mathbb{R} \times TU^*)$ under $\mathrm{Id} \times \Phi$ (with Id the identity map of $\mathbb{R}$) and:

$$H_\Phi\big(t, \Phi(q, p)\big) = H_\Phi\big(t, \varphi(q), (\mathrm{d}\varphi_q^{-1})^*(p)\big) = H(t, q, p),$$

for all $(t, q, p) \in \mathrm{dom}(H) \cap (\mathbb{R} \times TU^*)$. Set:

$$\big(\tilde q(t), \tilde p(t)\big) = \Phi\big(q(t), p(t)\big) \in \widetilde U \times \mathbb{R}^{n*} \subset \mathbb{R}^n \times \mathbb{R}^{n*},$$

for all $t \in q^{-1}(U) \subset I$. By the result of item (c) of Exercise 1.31, equality (1.10.9) is equivalent to:

$$\frac{\mathrm{d}\tilde q}{\mathrm{d}t}(t) = \frac{\partial H_\Phi}{\partial p}\big(t, \tilde q(t), \tilde p(t)\big), \quad \frac{\mathrm{d}\tilde p}{\mathrm{d}t}(t) = -\frac{\partial H_\Phi}{\partial q}\big(t, \tilde q(t), \tilde p(t)\big). \tag{1.10.15}$$

We have to show that (1.10.15) is equivalent to (1.10.11). This will follow from Lemma 1.10.13 once we check the validity of the following facts:

(i) L_φ is hyper-regular and H_Φ is the Legendre transform of L_φ;
(ii) for all $t \in q^{-1}(U) \subset I$ with $\big(t, \tilde q(t), \dot{\tilde q}(t)\big) \in \mathrm{dom}(L_\varphi)$, we have:

$$\tilde p(t) = \frac{\partial L_\varphi}{\partial \dot q}\big(t, \tilde q(t), \dot{\tilde q}(t)\big).$$

Let us check those facts. For $t \in \mathbb{R}$, $q \in U$, setting $\tilde q = \varphi(q)$, we have:

$$L_\varphi(t, \tilde q, \cdot) \circ \mathrm{d}\varphi_q = L(t, q, \cdot), \tag{1.10.16}$$

$$H_\Phi(t, \tilde q, \cdot) \circ (\mathrm{d}\varphi_q^*)^{-1} = H(t, q, \cdot). \tag{1.10.17}$$

Since $H(t, q, \cdot)$ is the Legendre transform of $L(t, q, \cdot)$, equalities (1.10.16), (1.10.17) and the result of Exercise 1.35 imply that $H_\Phi(t, \tilde{q}, \cdot)$ is the Legendre transform of $L_\varphi(t, \tilde{q}, \cdot)$. We have proven fact (i). Let us prove fact (ii).

Given $t \in q^{-1}(U) \subset I$ with $\big(t, \tilde{q}(t), \dot{\tilde{q}}(t)\big) \in \mathrm{dom}(L_\varphi)$ (or, equivalently, with $\big(t, q(t), \dot{q}(t)\big) \in \mathrm{dom}(L)$), we replace q with $q(t)$ and $\tilde{q}$ with $\tilde{q}(t)$ in (1.10.16); after that, we differentiate both sides at the point $\dot{q}(t)$, obtaining:

$$\frac{\partial L_\varphi}{\partial \dot{q}}\big(t, \tilde{q}(t), \dot{\tilde{q}}(t)\big) \circ \mathrm{d}\varphi_{q(t)} = \frac{\partial L}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big). \tag{1.10.18}$$

By (1.10.10), the righthand side of (1.10.18) equals $p(t)$ and therefore:

$$\frac{\partial L_\varphi}{\partial \dot{q}}\big(t, \tilde{q}(t), \dot{\tilde{q}}(t)\big) = p(t) \circ \mathrm{d}\varphi_{q(t)}^{-1} = (\mathrm{d}\varphi_{q(t)}^{-1})^*\big(p(t)\big) = \tilde{p}(t),$$

proving fact (ii). □

1.10.1. Mechanics with constraints again. In Example 1.10.3 we have seen that the Legendre transform of the Lagrangian of Classical Mechanics is precisely the Hamiltonian of Classical Mechanics. Let us now see what happens in the presence of constraints. As in Subsection 1.5.1 we consider n particles subject to a constraint defined by a submanifold Q of $(\mathbb{R}^3)^n$ and to (the forces with) a potential $V : \mathrm{dom}(V) \subset \mathbb{R} \times (\mathbb{R}^3)^n \to \mathbb{R}$. We consider the restriction L^{cons} to $\mathbb{R} \times TQ$ of the Lagrangian L of Classical Mechanics without constraints (defined in Subsection 1.4.1). We have seen in Subsection 1.5.1 that if the force exerted by the constraint is normal to Q then the curves $q = (q_1, \ldots, q_n)$ in Q which are possible trajectories for the particles are precisely the critical points of the action functional corresponding to L^{cons} (which are, in a given chart, solutions to the Euler–Lagrange equation of the Lagrangian that represents L^{cons} with respect to that chart). Let us now compute the Legendre transform H^{cons} of the Lagrangian L^{cons} (see also Exercise 1.38 for a slightly different approach). Since L^{cons} is the restriction of L, given $(t, q, \dot{q}) \in \mathbb{R} \times TQ$ with (t, q) in the domain of V, the partial derivative:

$$\frac{\partial L^{\mathrm{cons}}}{\partial \dot{q}}(t, q, \dot{q}) \in T_qQ^* \tag{1.10.19}$$

is simply the restriction to T_qQ of the linear functional $\frac{\partial L}{\partial \dot{q}}(t, q, \dot{q}) \in (\mathbb{R}^3)^{n*}$ (which was computed in Example 1.10.3). Thus, (1.10.19) is the restriction to T_qQ of the linear functional over $(\mathbb{R}^3)^n$ that is identified (via the standard identification of $(\mathbb{R}^3)^n$ with its dual space) with the vector:

$$(m_1\dot{q}_1, \ldots, m_n\dot{q}_n) \in (\mathbb{R}^3)^n.$$

Let us check that L^{cons} is hyper-regular. The map:

$$T_qQ \ni \dot{q} \longmapsto \frac{\partial L^{\mathrm{cons}}}{\partial \dot{q}}(t, q, \dot{q}) = \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})\Big|_{T_qQ} \in T_qQ^* \tag{1.10.20}$$

is linear (because the map $\dot{q} \mapsto \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})$ is linear). Given $\dot{q} \in T_qQ$, if (1.10.19) vanishes then:

$$\frac{\partial L^{\text{cons}}}{\partial \dot{q}}(t, q, \dot{q})\dot{q} = 0$$

and therefore:

$$\frac{\partial L^{\text{cons}}}{\partial \dot{q}}(t, q, \dot{q})\dot{q} = \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})\dot{q} = \sum_{j=1}^{n} m_j \|\dot{q}_j\|^2 = 0$$

which implies $\dot{q} = 0$. This proves that the linear map (1.10.20) is injective and it is therefore an isomorphism (and a diffeomorphism). Hence L^{cons} is hyper-regular. Let us compute its Legendre transform H^{cons}. Clearly:

$$\frac{\partial L^{\text{cons}}}{\partial \dot{q}}(t, q, \dot{q})\dot{q} - L^{\text{cons}}(t, q, \dot{q}) = \frac{\partial L}{\partial \dot{q}}(t, q, \dot{q})\dot{q} - L(t, q, \dot{q}), \tag{1.10.21}$$

for every $(t, q, \dot{q}) \in \mathbb{R} \times TQ$ with $(t, q) \in \text{dom}(V)$. The righthand side of (1.10.21) was computed in Example 1.10.3. The Hamiltonian H^{cons} is therefore given by:

$$H^{\text{cons}}(t, q, p) = \sum_{j=1}^{n} \frac{1}{2} m_j \|\dot{q}_j\|^2 + V(t, q),$$

for all $(t, q) \in \text{dom}(V)$, $p \in T_qQ^*$, where $\dot{q} \in T_qQ$ is the unique vector such that:

$$p = \frac{\partial L^{\text{cons}}}{\partial \dot{q}}(t, q, \dot{q}) \in T_qQ^* \tag{1.10.22}$$

holds (the existence and uniqueness of $\dot{q}$ follows from the fact that (1.10.20) is a linear isomorphism). A solution of Hamilton's equations corresponding to H^{cons} (i.e., an integral curve of the symplectic gradient of H^{cons} with respect to the canonical symplectic form of the cotangent bundle TQ^*) is, by Proposition 1.10.11 and Theorem 1.10.12, a curve $I \ni t \mapsto \big(q(t), p(t)\big)$ in TQ^* such that:

$$p(t) = \frac{\partial L^{\text{cons}}}{\partial \dot{q}}\big(t, q(t), \dot{q}(t)\big) \tag{1.10.23}$$

for all $t \in I$ and such that the representation of q with respect to any local chart of Q is a solution of the Euler–Lagrange equation corresponding to the Lagrangian that represents L^{cons} with respect to the given chart (by Theorem 1.5.1, such condition is equivalent to the condition that the restriction of q to any closed interval contained in I be a critical point of the action functional corresponding to L^{cons}). In other words, a solution of Hamilton's equations corresponding to H^{cons} consists of a curve (q, p) in TQ^*, with q a curve in Q describing the trajectories of n particles subject to the given constraints and the given potential (under the assumption that the force exerted by the constraint is normal to Q) and with p defined by (1.10.23). Notice that:

$$H^{\mathrm{cons}}\big(t, q(t), p(t)\big) = \sum_{j=1}^{n} \frac{1}{2} m_j \|\dot{q}_j(t)\|^2 + V\big(t, q(t)\big),$$

for all $t \in I$, i.e., $H^{\mathrm{cons}}\big(t, q(t), p(t)\big)$ equals the total energy at the instant t. By Theorem 1.8.5, if the potential V does not depend on time, we conclude that the total energy is conserved.

While it is standard to denote a solution to Hamilton's equations by (q, p), one should be aware that this notation might be misleading: in the absence of constraints (Example 1.10.3), $p(t)$ was precisely the (element of $(\mathbb{R}^3)^{n*}$ that is identified with the) vector containing the momenta of all the particles. However, in the present context, $p(t)$ is the *restriction* to the tangent space $T_{q(t)}Q$ of the linear functional over $(\mathbb{R}^3)^n$ that is identified with the vector containing the momenta of all the particles. Thus, in general, $p(t)$ is *not the momentum of anything*! It is customary, nevertheless, to call $p(t)$ the *canonical momentum* or the *canonically conjugate momentum*. If one considers a local chart φ on Q, then the curve (q, p) is represented (with respect to the local chart $\mathrm{d}^*\varphi$ induced on TQ^*) by a curve:

$$(\tilde{q}, \tilde{p}) = (\tilde{q}^{\,1}, \ldots, \tilde{q}^{\,m}, \tilde{p}_1, \ldots, \tilde{p}_m)$$

in $\mathbb{R}^m \times \mathbb{R}^{m*}$, with $m = \dim(Q)$. Elementary Physics texts would normally call the curves $\tilde{q}^{\,j}$, $j = 1, \ldots, m$, the description of the evolution of the mechanical system in terms of *generalized coordinates*[26]. One should pay attention to the following facts: the index j in $\tilde{q}^{\,j}$ *does not* refer to the number of a particle; the number $m = \dim(Q)$ is not the number of particles and it is normally called the number of *degrees of freedom* of the system. The $\tilde{p}_j$ is not the "momentum of the j-th particle", since p (and $\tilde{p}$) is not momentum and j does not refer to the number of a particle. The sum $\sum_{j=1}^{m} \tilde{p}_j$ is not the total momentum and in fact it does not have any physically relevant meaning (it is highly dependent on the choice of the coordinate chart) and it is not conserved. The actual total momentum $\sum_{j=1}^{n} m_j \dot{q}_j(t)$ is not in general conserved either because there are external forces (the forces exerted by the constraint) acting upon the system of n constrained particles. In Section 1.11 we will understand this violation of the conservation of the total momentum[27] as a break down of the spatial translations symmetry.

[26]In the standard notation used by elementary Physics texts for the description of the evolution of a mechanical system in terms of generalized coordinates, there are no tildes over the q's and the p's. We need the tildes because we are using the (q, p) without the tildes for the curve in TQ^*.

[27]Of course, there is no violation of the conservation of the total momentum if one considers the full (closed) system, containing both the n particles and the objects (also made out of particles) that are constraining the n particles.

1.11. Symmetry and conservation laws

In this section we will make a brief exposition of the relationship between symmetries and conservation laws. The result we present is a particular case of the celebrated *Nöther's theorem* (whose most general formulation concerns not only Mechanics, but also field theories[28]). The idea is that to a continuous symmetry of a mechanical system (described in terms of an element of the Lie algebra of a Lie group that acts on the configuration space and preserves the Hamiltonian) we associate a conservation law (described in terms of a first integral of the equations of motion defined over phase space). The result that we are going to present is not the best that can be done. For instance, time translation symmetry implies conservation of energy (Theorem 1.8.5), but that important relationship between symmetry and conservation law does not follow from the result we are going to present in this section. Nevertheless, such result is very simple and sufficient for our limited purposes. During the presentation, we are going to use some very basic facts about Lie groups and actions of Lie groups on manifolds which, for the reader's convenience, are recalled in the appendix (Section A.3).

Let Q be a differentiable manifold and $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ be a time-dependent Hamiltonian over TQ^*. Given a Lie group G and a smooth action:

$$\rho : G \times Q \ni (g, q) \longmapsto g \cdot q \in Q$$

of G on Q then, for each $g \in G$, the smooth diffeomorphism:

$$\rho_g : Q \ni q \longmapsto g \cdot q \in Q$$

of Q induces a smooth diffeomorphism $\mathrm{d}^*(\rho_g)$ of the cotangent bundle TQ^* (see (1.9.4)). For $g, h \in G$, we have $\rho_g \circ \rho_h = \rho_{gh}$ and therefore (see Exercise 1.33):

$$\mathrm{d}^*(\rho_g) \circ \mathrm{d}^*(\rho_h) = \mathrm{d}^*(\rho_{gh}).$$

We thus obtain a smooth action of G on TQ^* defined by:

$$g \cdot (q, p) = \mathrm{d}^*(\rho_g)(q, p) = \big(g \cdot q, \big(\mathrm{d}(\rho_g)_q^{-1}\big)^*(p)\big), \quad g \in G, \ (q, p) \in TQ^*.$$

We say that the action of G on Q is a *symmetry* of the Hamiltonian H if, for all $(t, q, p) \in \mathrm{dom}(H)$ and all $g \in G$, we have $\big(t, g \cdot (q, p)\big) \in \mathrm{dom}(H)$ and:

$$H\big(t, g \cdot (q, p)\big) = H(t, q, p).$$

For each X in the Lie algebra $\mathfrak{g}$ of G, the action of G on Q induces a vector field X^Q on Q and the action of G on TQ^* induces a vector field X^{TQ^*} on TQ^* (see Subsection A.3.1). Recall from Section 1.9 that θ denotes the canonical one-form of the cotangent bundle TQ^*. Here is the main result of the section.

[28]We observe also that most presentations of Nöther's theorem use the Lagrangian formalism, while we are going to use the Hamiltonian formalism.

1.11.1. THEOREM. *Given a differentiable manifold Q, a time-dependent Hamiltonian $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ and a smooth action of a Lie group G on Q that is a symmetry of H then, for all $X \in \mathfrak{g}$, the real valued map $\theta(X^{TQ^*})$ over TQ^* is a first integral of $\vec{H}$, i.e., it is constant along the integral curves of $\vec{H}$.*

PROOF. It is sufficient to prove that for all $t \in \mathbb{R}$ the directional derivative:

$$\mathrm{d}\big(\theta(X^{TQ^*})\big)(\vec{H}_t) = \vec{H}_t\big(\theta(X^{TQ^*})\big)$$

of the map $\theta(X^{TQ^*})$ along the vector field $\vec{H}_t = \vec{H}(t, \cdot)$ vanishes. We start by observing that the canonical one-form θ of TQ^* is invariant under the flow of the vector field X^{TQ^*}. Namely, the flow at time t of X^{TQ^*} is the diffeomorphism $\mathrm{d}^*(\rho_g)$ with $g = \exp(tX)$ and it follows from (1.9.6) that the pull-back of θ by $\mathrm{d}^*(\rho_g)$ equals θ. Thus, the Lie derivative of θ along X^{TQ^*} vanishes:

$$\mathbb{L}_{X^{TQ^*}}\theta = 0.$$

Since the action of G on Q is a symmetry of H, we see that for all $t \in \mathbb{R}$ the map $H_t = H(t, \cdot)$ is constant along the integral curves of X^{TQ^*} (because such integral curves are of the form $\mathbb{R} \ni s \mapsto \exp(sX) \cdot (q, p) \in TQ^*$) and therefore:

$$X^{TQ^*}(H_t) = \mathrm{d}H_t(X^{TQ^*}) = 0. \tag{1.11.1}$$

We compute the Lie derivative of θ along X^{TQ^*} in terms of exterior differentiation and interior products (formula (A.2.9)):

$$\mathbb{L}_{X^{TQ^*}}\theta = \mathrm{d}i_{X^{TQ^*}}\theta + i_{X^{TQ^*}}\mathrm{d}\theta = 0,$$

and we evaluate the result at $\vec{H}_t$:

$$\mathrm{d}\big(\theta(X^{TQ^*})\big)(\vec{H}_t) + \mathrm{d}\theta(X^{TQ^*}, \vec{H}_t) = 0. \tag{1.11.2}$$

In order to conclude the proof we have to check that the first term on the lefthand side of (1.11.2) vanishes, which will follow if we show that the second term vanishes. Since $\mathrm{d}\theta = -\omega$, keeping in mind the definition of the symplectic gradient $\vec{H}_t$, we obtain:

$$\mathrm{d}\theta(X^{TQ^*}, \vec{H}_t) = \omega(\vec{H}_t, X^{TQ^*}) = \mathrm{d}H_t(X^{TQ^*})$$

and the conclusion follows from (1.11.1). □

Theorem 1.11.1 can be easily generalized to the following result about Hamiltonians on symplectic manifolds: if (M, ω) is a symplectic manifold with exact symplectic form $\omega = -\mathrm{d}\theta$ and if a Lie group G acts on M in such a way that the action preserves the one-form θ and a time-dependent Hamiltonian H over M then, for all $X \in \mathfrak{g}$, the real valued map $\theta(X^M)$ over M is a first integral of $\vec{H}$. The proof of such result is identical to the proof of Theorem 1.11.1. A generalization of Theorem 1.11.1 to symplectic manifolds whose symplectic form is not exact (or to the case in which the symplectic form is exact but the action of the Lie group preserves the symplectic form and not the corresponding one-form) is given in Exercise 1.39.

The first integral $\theta(X^{TQ^*})$ of $\vec{H}$ given by Theorem 1.11.1 can be made more explicit as shown by the following:

1.11.2. PROPOSITION. *If a Lie group G acts smoothly on Q then for all $X \in \mathfrak{g}$, we have:*

$$\theta(X^{TQ^*})(q,p) = p\big(X^Q(q)\big),$$

for all $(q,p) \in TQ^*$.

PROOF. By the definition of θ, we have:

$$\theta(X^{TQ^*})(q,p) = p\big[\mathrm{d}\pi_{(q,p)}\big(X^{TQ^*}(q,p)\big)\big],$$

for all $(q,p) \in TQ^*$, where $\pi : TQ^* \to Q$ denotes the canonical projection. Since $X^{TQ^*}(q,p)$ is the image of X by the differential at $1 \in G$ of the map $g \mapsto g\cdot(q,p)$, we have (by the chain rule) that $\mathrm{d}\pi_{(q,p)}\big(X^{TQ^*}(q,p)\big)$ equals the image of X by the differential at $1 \in G$ of the map $g \mapsto \pi\big(g\cdot(q,p)\big) = g\cdot q$. Thus:

$$\mathrm{d}\pi_{(q,p)}\big(X^{TQ^*}(q,p)\big) = X^Q(q)$$

and the conclusion follows. $\square$

1.11.3. EXAMPLE. Let $V : \mathrm{dom}(V) \subset \mathbb{R}\times(\mathbb{R}^3)^n \to \mathbb{R}$ be a smooth potential and consider the Hamiltonian:

$$H(t,q,p) = \sum_{j=1}^{n} \frac{\|p_j\|^2}{2m_j} + V(t,q), \quad (t,q) \in \mathrm{dom}(V),\ p \in (\mathbb{R}^3)^{n*},$$

of Classical Mechanics. Let G be the abelian Lie group $(\mathbb{R}^3,+)$, whose Lie algebra $\mathfrak{g}$ is $\mathbb{R}^3$ endowed with the identically vanishing Lie bracket. We consider the action of G on $Q = (\mathbb{R}^3)^n$ by *spatial translations*, i.e., we set:

$$g\cdot q = (g+q_1,\ldots,g+q_n), \quad g \in \mathbb{R}^3,\ q = (q_1,\ldots,q_n) \in (\mathbb{R}^3)^n.$$

The differential of the map $q \mapsto g\cdot q$ at a point $q \in Q$ is the identity map (and its inverse transpose is also the identity map) and therefore the induced action of G on the cotangent bundle TQ^* is given by:

$$g\cdot(q,p) = (g\cdot q, p), \quad q \in (\mathbb{R}^3)^n,\ p \in (\mathbb{R}^3)^{n*}.$$

Since the action of g on p is trivial, the kinetic term of H is preserved by the action. Now, assume that the potential V depends on q only through the differences $q_i - q_j$, $i,j = 1,\ldots,n$; this is the case of the electrical and gravitational potentials. Under such assumption the action of G is a symmetry of the potential and thus also of the Hamiltonian. Let us determine the conservation law associated to an element $X \in \mathfrak{g} = \mathbb{R}^3$. The value at $q \in (\mathbb{R}^3)^n$ of the vector field X^Q is obtained by differentiating $g \mapsto g\cdot q$ at $g = 0$ and evaluating at the vector X; thus:

$$X^Q(q) = (X,\ldots,X) \in (\mathbb{R}^3)^n.$$

By Proposition 1.11.2, the conserved quantity $\theta(X^{TQ^*})$ is given by:

$$TQ^* \ni (q,p) \longmapsto p\big(X^Q(q)\big) = \sum_{j=1}^{n} p_j(X) \in \mathbb{R}.$$

Replacing X with the k-th vector of the canonical basis of $\mathbb{R}^3$ ($k = 1,2,3$), we obtain that *the k-th component of the total momentum is conserved if spatial translations are a symmetry of the Hamiltonian H.*

1.11.4. Example. Consider again the Hamiltonian of Classical Mechanics with a potential V (as in Example 1.11.3) and let now $G = \mathrm{SO}(3)$ be the Lie group of all orientation preserving linear isometries of $\mathbb{R}^3$ (or, equivalently, the group of 3×3 orthogonal matrices whose determinant is equal to 1). We consider the action of G on $Q = (\mathbb{R}^3)^n$ by *spatial rotations*, i.e., we set:

$$g\cdot q = \big(g(q_1),\ldots,g(q_n)\big), \quad g \in \mathrm{SO}(3),\ q = (q_1,\ldots,q_n) \in (\mathbb{R}^3)^n.$$

The map $q \mapsto g\cdot q$ is linear and thus its differential at any point equals itself; moreover, since $q \mapsto g\cdot q$ is a linear isometry of $(\mathbb{R}^3)^n$ (endowed with its canonical inner product) then the transpose inverse of $q \mapsto g\cdot q$ equals itself (upon identification of $(\mathbb{R}^3)^n$ with its dual space using the canonical inner product). Thus, the action of G induced on the cotangent bundle TQ^* is:

$$g\cdot(q,p) = (g\cdot q, g\cdot p), \quad g \in \mathrm{SO}(3),\ q \in (\mathbb{R}^3)^n,\ p \in (\mathbb{R}^3)^n \cong (\mathbb{R}^3)^{n*}.$$

Since each $g \in G$ is a linear isometry of $\mathbb{R}^3$, the action of G preserves the kinetic term of the Hamiltonian. Now, assuming that the potential V depends on q only through the *norms* of the differences $q_i - q_j$ (which is the case of the electric and the gravitational potentials) then the action of G is also a symmetry of the potential and hence of the Hamiltonian. The Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ of G is the Lie algebra of anti-symmetric linear endomorphisms of $\mathbb{R}^3$ (or, equivalently, of 3×3 anti-symmetric matrices) endowed with the standard commutator. Let us compute the conserved quantity corresponding to an element $X \in \mathfrak{so}(3)$. Given $q \in (\mathbb{R}^3)^n$ we differentiate the map $g \mapsto g\cdot q$ at the identity and we evaluate it at X, obtaining:

$$X^Q(q) = \big(X(q_1),\ldots,X(q_n)\big), \quad q \in (\mathbb{R}^3)^n.$$

Thus, by Proposition 1.11.2, the conserved quantity corresponding to X is:

$$TQ^* \ni (q,p) \longmapsto p\big(X^Q(q)\big) = \sum_{j=1}^{n} p_j\big(X(q_j)\big) \in \mathbb{R}.$$

Let us rewrite such conserved quantity in a nicer way. The Lie algebra of $\mathrm{SO}(3)$ can be identified with $\mathbb{R}^3$ endowed with the vector product (see Exercise 1.40); more explicitly, given $X \in \mathfrak{so}(3)$, there exists a unique $v \in \mathbb{R}^3$ such that:

$$X(w) = v \wedge w, \quad w \in \mathbb{R}^3,$$

where $\wedge$ denotes the vector product. Thus (identifying $p_j \in \mathbb{R}^{3^*}$ with a vector of $\mathbb{R}^3$):

$$p_j\big(X(q_j)\big) = \langle p_j, v \wedge q_j\rangle = \langle v, q_j \wedge p_j\rangle,$$

so that the conserved quantity associated to v is:

$$TQ^* \ni (q,p) \longmapsto \sum_{j=1}^{n} \langle v, q_j \wedge p_j\rangle \in \mathbb{R}. \tag{1.11.3}$$

This motivates the following:

1.11.5. Definition. The *angular momentum* of the j-th particle at time $t \in \mathbb{R}$ is defined by:

$$L_j(t) = q_j(t) \wedge p_j(t) = m_j\big(q_j(t) \wedge \dot{q}_j(t)\big).$$

Replacing v with the k-th vector of the canonical basis of $\mathbb{R}^3$ ($k = 1, 2, 3$), we obtain that *the k-th component of the total angular momentum*:

$$\sum_{j=1}^{n} L_j(t) = \sum_{j=1}^{n} q_j(t) \wedge p_j(t)$$

is conserved if spatial rotations are a symmetry of the Hamiltonian H.

1.12. The Poisson bracket

In this section we define the Poisson bracket, which is a binary operation on the space of smooth real valued maps over a symplectic manifold. We will need the Poisson bracket for our forthcoming discussion of *quantization*. The concept of Poisson bracket allows one to establish some nice algebraic analogies between Classical and Quantum Mechanics (in fact, such analogies are somewhat misleading, but interesting nevertheless). We will prove some simple properties of the Poisson bracket, which should allow the reader to have an idea of the relevance of the concept for Classical Mechanics.

In what follows, (M, ω) denotes a fixed symplectic manifold and $C^\infty(M)$ denotes the vector space of all smooth real valued maps over M.

1.12.1. Definition. Given $f, g \in C^\infty(M)$, then the *Poisson bracket* $\{f, g\}$ is the element of $C^\infty(M)$ defined by:

$$\{f, g\} = \omega(\vec{f}, \vec{g}),$$

where $\vec{f}$, $\vec{g}$ denote the symplectic gradients of f and g, respectively (recall Definition 1.8.2).

Here is an alternative definition of the Poisson bracket: for each x in M, the symplectic form ω_x over the tangent space T_xM induces a linear isomorphism:

$$T_xM \ni v \longmapsto \omega_x(v, \cdot) \in T_xM^*,$$

which is precisely the isomorphism that carries the symplectic gradient $\vec{f}(x)$ to the differential $\mathrm{d}f(x)$. Such isomorphism can be used to carry the symplectic form ω_x over T_xM to a symplectic form π_x over the dual space T_xM^*, i.e., π_x satisfies:

$$\pi_x\big(\omega_x(v,\cdot),\omega_x(w,\cdot)\big) = \omega_x(v,w), \quad v,w \in T_xM.$$

We have that π is a smooth anti-symmetric $(0,2)$-tensor field (or twice contravariant tensor field) over M (see Subsection A.2.2). Obviously, the Poisson bracket of two maps $f,g \in C^\infty(M)$ is given by:

$$\{f,g\} = \pi(\mathrm{d}f,\mathrm{d}g).$$

The direct relationship between the Poisson bracket and Classical Mechanics is that the Poisson bracket can be used to describe the time evolution of the value of a smooth function $f : M \to \mathbb{R}$ along the flow of a Hamiltonian.

1.12.2. Proposition. *Let $H : \mathrm{dom}(H) \subset \mathbb{R} \times M \to \mathbb{R}$ be a time-dependent Hamiltonian and let $t \mapsto x(t)$ be an integral curve of $\vec{H}$. Then, given a smooth map $f : M \to \mathbb{R}$, we have:*

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(x(t)\big) = \{f,H_t\}\big(x(t)\big), \tag{1.12.1}$$

where $H_t = H(t,\cdot)$.

Notice that the map H_t in the statement of Proposition 1.12.2 is defined only over some open subset of M, so that, to be completely precise, the Poisson bracket in (1.12.1) is the Poisson bracket between *the restriction* of f to the domain of H_t and the map H_t. Of course, the Poisson bracket is also defined for maps whose domain is an open subset of the symplectic manifold M (in fact, open subsets of a symplectic manifold are themselves symplectic manifolds, endowed with the restriction of the symplectic form).

Proof of Proposition 1.12.2. It is a straightforward computation:

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(x(t)\big) = \mathrm{d}f_{x(t)}\big[\vec{H}\big(t,x(t)\big)\big] = \omega_{x(t)}\big[\vec{f}\big(x(t)\big),\vec{H}_t\big(x(t)\big)\big]. \qquad \square$$

1.12.3. Corollary. *A smooth map $f : M \to \mathbb{R}$ is a first integral of the symplectic gradient $\vec{H}$ of a time-dependent Hamiltonian H if and only if the Poisson bracket $\{f,H_t\}$ vanishes, for all $t \in \mathbb{R}$.* $\square$

Poisson brackets are also useful for writing down in a nice way the condition that a local chart be symplectic.

1.12.4. Proposition. *Let $\Phi : U \subset M \to \widetilde{U} \subset \mathbb{R}^{2n}$ be a local chart on M; write $\Phi = (q^1,\ldots,q^n,p^1,\ldots,p^n)$. The local chart Φ is symplectic if and only if:*

$$\{q^i,q^j\} = 0, \quad \{p^i,p^j\} = 0, \quad \{q^i,p^j\} = \delta^{ij}, \quad i,j = 1,\ldots,n,$$

where $\delta^{ij} = 1$ for $i = j$ and $\delta^{ij} = 0$ for $i \neq j$.

PROOF. For $x \in U$, let:

$$\frac{\partial}{\partial q^1}(x), \ldots, \frac{\partial}{\partial q^n}(x), \frac{\partial}{\partial p^1}(x), \ldots, \frac{\partial}{\partial p^n}(x), \tag{1.12.2}$$

denote the basis of T_xM that is carried by $\mathrm{d}\Phi(x)$ to the canonical basis of $\mathbb{R}^{2n}$. The dual basis of (1.12.2) is:

$$\mathrm{d}q^1(x), \ldots, \mathrm{d}q^n(x), \mathrm{d}p^1(x), \ldots, \mathrm{d}p^n(x). \tag{1.12.3}$$

The chart Φ is symplectic if and only if (1.12.2) is a symplectic basis of (T_xM, ω_x), for all $x \in U$. By the result of Exercise 1.42, the basis (1.12.2) is symplectic for ω_x if and only if the dual basis (1.12.3) is symplectic for π_x, which happens (for all $x \in U$) if and only if:

$$\pi(\mathrm{d}q^i, \mathrm{d}q^j) = \{q^i, q^j\} = 0, \quad \pi(\mathrm{d}p^i, \mathrm{d}p^j) = \{p^i, p^j\} = 0,$$
$$\pi(\mathrm{d}q^i, \mathrm{d}p^j) = \{q^i, p^j\} = \delta^{ij},$$

for all $i, j = 1, \ldots, n$. □

Let us now investigate some algebraic properties of the Poisson bracket and its relationship with other operations defined on manifolds. We start by noticing that, for $f, g \in C^\infty(M)$, we have:

$$\{f, g\} = \omega(\vec{f}, \vec{g}) = -\omega(\vec{g}, \vec{f}) = -\mathrm{d}g(\vec{f}) = -\vec{f}(g).$$

In other words, if we identify the vector field $\vec{f}$ with the linear endomorphism of $C^\infty(M)$ that sends g to $\vec{f}(g) = \mathrm{d}g(\vec{f})$ then:

$$\{f, \cdot\} = -\vec{f}, \tag{1.12.4}$$

where $\{f, \cdot\}$ denotes the linear endomorphism of $C^\infty(M)$ that sends g to $\{f, g\}$. It follows that $\{f, \cdot\}$ is a *derivation* of the algebra $C^\infty(M)$, i.e.:

$$\{f, g_1 g_2\} = \{f, g_1\}g_2 + g_1\{f, g_2\}, \quad f, g_1, g_2 \in C^\infty(M). \tag{1.12.5}$$

Another interesting algebraic property of the Poisson bracket is given by the following theorem.

1.12.5. THEOREM. *The real vector space $C^\infty(M)$ endowed with the Poisson bracket is a Lie algebra, i.e., the Poisson bracket is bilinear, anti-symmetric and satisfies the* Jacobi identity*:*

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad f, g, h \in C^\infty(M). \tag{1.12.6}$$

PROOF. The only non trivial part of the statement is the Jacobi identity. Since the symplectic form ω is closed, we have:

$$\mathrm{d}\omega(\vec{f}, \vec{g}, \vec{h}) = 0. \tag{1.12.7}$$

We use formula (A.2.8) for computing the lefthand side of (1.12.7):

$$\begin{aligned} &\vec{f}\big(\omega(\vec{g}, \vec{h})\big) - \vec{g}\big(\omega(\vec{f}, \vec{h})\big) + \vec{h}\big(\omega(\vec{f}, \vec{g})\big) \\ &\qquad - \omega\big([\vec{f}, \vec{g}], \vec{h}\big) + \omega\big([\vec{f}, \vec{h}], \vec{g}\big) - \omega\big([\vec{g}, \vec{h}], \vec{f}\big) = 0. \end{aligned} \tag{1.12.8}$$

We rewrite the six terms in the lefthand side of (1.12.8) in terms of iterated Poisson brackets. Notice that:

$$\vec{f}\big(\omega(\vec{g},\vec{h})\big) = \vec{f}\big(\{g,h\}\big),$$

and using (1.12.4) we obtain:

$$\vec{f}\big(\omega(\vec{g},\vec{h})\big) = -\{f,\{g,h\}\}.$$

The second and third terms in the lefthand side of (1.12.8) can in an analogous way be rewritten in terms of iterated Poisson brackets. Let us work with the remaining terms. We have:

$$-\omega\big([\vec{f},\vec{g}],\vec{h}\big) = \omega\big(\vec{h},[\vec{f},\vec{g}]\big) = \mathrm{d}h\big([\vec{f},\vec{g}]\big) = [\vec{f},\vec{g}](h) = \vec{f}\big(\vec{g}(h)\big) - \vec{g}\big(\vec{f}(h)\big).$$

Using (1.12.4) we obtain:

$$-\omega\big([\vec{f},\vec{g}],\vec{h}\big) = \{f,\{g,h\}\} - \{g,\{f,h\}\}.$$

The fifth and sixth terms in the lefthand side of (1.12.8) can in an analogous way be rewritten in terms of iterated Poisson brackets. Once all terms in the lefthand side of (1.12.8) are rewritten in terms of iterated Poisson brackets and the appropriate cancelations are performed (and taking into account the anti-symmetry of the Poisson bracket), one obtains the Jacobi identity (1.12.6). $\square$

We have then that the vector space $C^\infty(M)$ is an associative algebra, endowed with the pointwise product of real valued functions, and also a Lie algebra, endowed with the Poisson bracket. Moreover, the two structures are related by the fact that, for all $f \in C^\infty(M)$, the linear endomorphism $\{f,\cdot\}$ of $C^\infty(M)$ is a derivation with respect to the associative product, i.e., (1.12.5) holds.

1.12.6. Definition. A *Poisson algebra* is a vector space V endowed with both an associative algebra structure $V \times V \ni (x,y) \mapsto xy \in V$ and a Lie algebra structure $V \times V \ni (x,y) \mapsto [x,y] \in V$, in such a way that, for all $x \in V$, the linear endomorphism $[x,\cdot]$ of V is a derivation of the associative product, i.e.:

$$[x, y_1y_2] = [x,y_1]y_2 + y_1[x,y_2], \quad x, y_1, y_2 \in V.$$

We have shown that if (M,ω) is a symplectic manifold, then $C^\infty(M)$ is a Poisson algebra, endowed with the pointwise product and the Poisson bracket. Notice that the associative product of the Poisson algebra $C^\infty(M)$ is also commutative, but such commutativity is not a requirement of the definition of Poisson algebra. In Exercise 1.43 we give an example of a family of Poisson algebras whose associative product might not be commutative (such example is related to Quantum Theory, as we will learn later).

The Jacobi identity (1.12.6) (or, more generally, the Jacobi identity in any Lie algebra) can be interpreted in two ways: first, it says that for all

$f \in C^\infty(M)$, the linear endomorphism $\{f, \cdot\}$ is a derivation of the Poisson bracket:

$$\{f, \{g, h\}\} = \{\{f, g\}, h\} + \{g, \{f, h\}\}, \quad f, g, h \in C^\infty(M).$$

The second interpretation is that the map $f \mapsto \{f, \cdot\}$ (the *adjoint representation* of the Lie algebra) is a Lie algebra homomorphism from the Lie algebra $C^\infty(M)$ (endowed with the Poisson bracket) to the Lie algebra of linear endomorphisms of $C^\infty(M)$ (endowed with the standard commutator of linear operators):

$$[\{f, \cdot\}, \{g, \cdot\}] = \{\{f, g\}, \cdot\}, \tag{1.12.9}$$

or, more explicitly:

$$\{f, \{g, h\}\} - \{g, \{f, h\}\} = \{\{f, g\}, h\}, \quad f, g, h \in C^\infty(M).$$

This second interpretation, coupled with (1.12.4), yields the following:

1.12.7. PROPOSITION. *For all $f, g \in C^\infty(M)$, the Lie bracket of the symplectic gradients $\vec{f}$, $\vec{g}$ is given by minus the symplectic gradient of the Poisson bracket $\{f, g\}$:*

$$[\vec{f}, \vec{g}] = -\overrightarrow{\{f, g\}}.$$

PROOF. Use (1.12.9) and (1.12.4). □

1.13. Measurements and Observables

Exercises

Affine spaces and Galilean spacetimes.

EXERCISE 1.1. Let V be a vector space and consider the action of the additive group $(V, +)$ on the set $E = V$ given by:

$$V \times E \ni (v, e) \longmapsto v + e \in E.$$

Show that $E = V$ is an affine space with underlying vector space V. This is called the affine space *canonically obtained* from the vector space V.

EXERCISE 1.2. Let E be an affine space with underlying vector space V. Show that for every point $O \in E$, the map:

$$V \ni v \longmapsto v + O \in E$$

is an affine isomorphism from the affine space canonically obtained from V onto the affine space E.

EXERCISE 1.3. Let E be an affine space with underlying vector space V and let V_0 be a vector subspace of V. Let E_0 be an orbit of the action of V_0 on E. Show that E_0 can be made into an affine space with underlying vector space V_0 in such a way that the inclusion map of E_0 into E is affine, with the inclusion map of V_0 into V its underlying linear map. We call E_0 an *affine subspace* of E.

EXERCISE 1.4. Let E be an affine space with underlying vector space V and let V_0 be a vector subspace of V. Denote by E/V_0 the quotient of E by the action of V_0 (i.e., E/V_0 is the set of orbits of the action of V_0 on E). Show that the action:

$$(V/V_0) \times (E/V_0) \ni (v + V_0, e + V_0) \longmapsto (e + v) + V_0 \in E/V_0$$

is well-defined and turns E/V_0 into an affine space with underlying vector space V/V_0. Show that the quotient map $E \to E/V_0$ is an affine map whose underlying linear map is the quotient map $V \to V/V_0$.

EXERCISE 1.5. Given an affine space E with underlying vector space V, show that there exists a short exact sequence of groups:

$$1 \longrightarrow V \longrightarrow \mathrm{Aff}(E) \longrightarrow \mathrm{GL}(V) \longrightarrow 1,$$

where 1 denotes the trivial group.

Units of measurement.

EXERCISE 1.6. Let M be a real one-dimensional vector space. Every non zero vector $m \in M$ defines a basis of M, a basis $m \otimes m$ of $M \otimes M$ and a basis m^{-1} of the dual space M^*, where the linear functional m^{-1} satisfies $m^{-1}(m) = 1$ (m^{-1} is simply the dual basis of m). Given non zero vectors $m_1, m_2 \in M$ with $m_2 = cm_1$, check that:

$$m_2 \otimes m_2 = c^2 m_1 \otimes m_1, \quad m_2^{-1} = \tfrac{1}{c}\, m_1^{-1}.$$

Conclude that if the elements of M are to be interpreted as lengths then the elements of $M \otimes M$ are to be interpreted as square lengths and the elements of M^* are to be interpreted as inverse lengths.

Inertial coordinate systems and the Galileo group.

EXERCISE 1.7. Show that any two Galilean spacetimes are isomorphic.

EXERCISE 1.8. Let X, Y be objects of an arbitrary category (for instance, Galilean spacetimes) and assume that we are given an isomorphism $f : X \to Y$. Show that f induces a group isomorphism:

$$f_* : \mathrm{Aut}(X) \longrightarrow \mathrm{Aut}(Y)$$

from the group of automorphisms of X to the group of automorphisms of Y defined by:

$$f_*(g) = f \circ g \circ f^{-1}, \quad g \in \mathrm{Aut}(X).$$

Show that a subgroup G of $\mathrm{Aut}(X)$ has the property that the subgroup $f_*(G)$ of $\mathrm{Aut}(Y)$ is independent of the isomorphism $f : X \to Y$ if and only if G is normal in $\mathrm{Aut}(X)$.

EXERCISE 1.9. Given an inertial coordinate system $\phi : E \to \mathbb{R}^4$, then the inverse image under ϕ of $\mathbb{R} \times \{0\} \subset \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$ is an affine one-dimensional subspace of E which we call the *moving origin* of the inertial coordinate system ϕ. Let ϕ_1, ϕ_2 be inertial coordinate systems related by an element A of the passive Galileo group as in diagram (1.1.1). The moving origin of the inertial coordinate system ϕ_2 is mapped by ϕ_1 onto the following one-dimensional affine subspace of $\mathbb{R}^4$:

$$\phi_1\big[\phi_2^{-1}\big(\mathbb{R}\times\{0\}\big)\big] = A^{-1}\big(\mathbb{R}\times\{0\}\big). \tag{1.13.1}$$

Assume that A is the Galilean boost (1.1.4). Show that (1.13.1) is equal to:

$$\big\{(t, vt) : t \in \mathbb{R}\big\}.$$

This means that *the origin of ϕ_2 moves with uniform velocity v with respect to the coordinate system ϕ_1*.

Ontology and dynamics.

EXERCISE 1.10. Let ϕ_1, ϕ_2 be inertial coordinate systems related by an element A of the passive Galileo group as in diagram (1.1.1). Let A be given as in (1.1.2) and (1.1.3). Given maps $q, \tilde{q} : \mathbb{R} \to \mathbb{R}^3$, show that:

$$\phi_1^{-1}\big(\mathrm{gr}(q)\big) = \phi_2^{-1}\big(\mathrm{gr}(\tilde{q})\big) \tag{1.13.2}$$

if and only if q and $\tilde{q}$ are related by:

$$\tilde{q}(t+t_0) = L_0\big(q(t)\big) - vt + x_0, \quad t \in \mathbb{R}. \tag{1.13.3}$$

Equality (1.13.2) means that q and $\tilde{q}$ are representations with respect to the inertial coordinate systems ϕ_1, ϕ_2, respectively, of the *same* particle worldline.

EXERCISE 1.11. Suppose that the force maps F_j, $j = 1, \ldots, n$, satisfy the condition:

$$\begin{aligned} L_0\big(F_j(t, q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)\big) = F_j\big(&t+t_0, L_0(q_1) - vt + x_0, \ldots, \\ &L_0(q_n) - vt + x_0, L_0(\dot{q}_1) - v, \ldots, L_0(\dot{q}_n) - v\big), \end{aligned}$$

for all $(t, q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) \in \mathrm{dom}(F_j) \subset \mathbb{R} \times (\mathbb{R}^3)^n \times (\mathbb{R}^3)^n$, where $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^3$, $v \in \mathbb{R}^3$ and a linear isometry $L_0 : \mathbb{R}^3 \to \mathbb{R}^3$ are fixed. Show that if:

$$q_j : \mathbb{R} \longrightarrow \mathbb{R}^3, \quad \tilde{q}_j : \mathbb{R} \longrightarrow \mathbb{R}^3, \quad j = 1, \ldots, n,$$

are smooth maps related as in (1.13.3) then the maps q_j satisfy the differential equation (1.2.1) if and only if the maps $\tilde{q}_j$ satisfy the differential equation (1.2.1).

EXERCISE 1.12. Show that the gravitational and the electrical forces satisfy the condition given in the statement of Exercise 1.11.

Intrinsic formulation.

EXERCISE 1.13. Let E be a finite-dimensional real affine space with underlying vector space V. Show that there is a unique way to turn E into a differentiable manifold in such a way that, for every point $O \in E$, the map defined in Exercise 1.2 is a smooth diffeomorphism. Show that, given a point $e \in E$, then the linear isomorphism from V to the tangent space T_eE given by the differential of the map defined in Exercise 1.2 at the point $e-O$ is independent of the choice of the point O. We use such linear isomorphism to identify once and for all the tangent space T_eE with the vector space V.

EXERCISE 1.14. Let E, E' be real finite-dimensional affine spaces with underlying vector spaces V, V', respectively. Show that every affine map $A : E \to E'$ is smooth and that for every $e \in E$, the differential:

$$\mathrm{d}A(e) : T_eE \cong V \longrightarrow V' \cong T_{A(e)}E'$$

is the underlying linear map of A.

EXERCISE 1.15. Let $q : \mathfrak{T} \to E$ be a smooth section of $\bar{\mathfrak{t}}$. Given a velocity $v \in \mathfrak{t}^{-1}(1)$, show that the following statements are equivalent:

(a) $\dot{q}(t) = v$, for all $t \in \mathfrak{T}$;
(b) q is an affine map whose underlying linear map $\mathbb{R} \to V$ is given by multiplication by the vector v;
(c) the image of q is an affine subspace of E with underlying vector space spanned by v.

A particle whose worldline is the image of a section $q : \mathfrak{T} \to E$ satisfying one of the equivalent conditions above is said to have *rectilinear uniform motion with velocity* v.

EXERCISE 1.16. Let $\phi : E \to \mathbb{R}^4$ be an inertial coordinate system. The affine isomorphism ϕ passes to the quotient and defines an affine isomorphism τ from $\mathfrak{T} = E/\operatorname{Ker}(\mathfrak{t})$ to $\mathbb{R} \cong \mathbb{R}^4/\big(\{0\} \times \mathbb{R}^3\big)$. Let $q_0 : \mathfrak{T} \to E$ be a smooth section of $\bar{\mathfrak{t}}$ and let $q : \mathbb{R} \to \mathbb{R}^3$ be the smooth map such that:

$$\phi\big(q_0(\mathfrak{T})\big) = \mathrm{gr}(q).$$

We have a commutative diagram:

$$\begin{tikzcd} \mathfrak{T} \arrow[r, "q_0"] \arrow[d, "\tau"'] & E \arrow[d, "\phi"] \\ \mathbb{R} \arrow[r, "(\mathrm{Id},q)"'] & \mathbb{R}^4 \end{tikzcd}$$

where Id denotes the identity map of $\mathbb{R}$. Consider the moving origin of the inertial coordinate system ϕ (see Exercise 1.9); it is like the worldline of a particle having rectilinear uniform motion with some velocity $v \in \mathfrak{t}^{-1}(1)$. Given $t_0 \in \mathfrak{T}$, show that the relative velocity $\dot{q}_0(t_0) - v \in \operatorname{Ker}(\mathfrak{t})$ is mapped by the underlying linear map of ϕ to the vector $\big(0, \frac{\mathrm{d}q}{\mathrm{d}t}(t)\big)$, where $t = \tau(t_0)$.

Lagrangians on manifolds.

EXERCISE 1.17. Let Q be a differentiable manifold, $q:[a,b]\to Q$ be a smooth curve and $\alpha:[a,b]\to TQ^*$ be a continuous map, where $\alpha(t)$ belongs to the dual space $T_{q(t)}Q^*$, for all $t\in[a,b]$. Assume that:

$$\int_a^b \alpha(t)v(t)=0,$$

for any smooth vector field $v:[a,b]\to TQ$ along q whose support is contained in the open interval $]a,b[$. Show that $\alpha=0$.

EXERCISE 1.18. Let U be an open subset of $\mathbb{R}^n$, $L:\mathbb{R}\times U\times\mathbb{R}^n\to\mathbb{R}$ be a Lagrangian, $\sigma:U\to\sigma(U)$ be a smooth diffeomorphism onto an open subset $\sigma(U)$ of $\mathbb{R}^n$ and $L_\sigma:\mathbb{R}\times\sigma(U)\times\mathbb{R}^n\to\mathbb{R}$ be the Lagrangian obtained by pushing L using σ, i.e.:

$$L_\sigma\big(t,\sigma(q),\mathrm{d}\sigma_q(\dot q)\big)=L(t,q,\dot q),$$

for all $t\in\mathbb{R}$, $q\in U$, $\dot q\in\mathbb{R}^n$. Let $q:[a,b]\to U$ be a smooth curve and set $\tilde q=\sigma\circ q$. Show that:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_\sigma}{\partial\dot q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big)-\frac{\partial L_\sigma}{\partial q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big)=&\\ [(\mathrm{d}\sigma_{q(t)})^*]^{-1}\Big(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot q}\big(t,q(t),\dot q(t)\big)-\frac{\partial L}{\partial q}\big(t,q(t),\dot q(t)\big)\Big),&\end{aligned}$$

for all $t\in[a,b]$, where $(\mathrm{d}\sigma_{q(t)})^*:\mathbb{R}^{n*}\to\mathbb{R}^{n*}$ denotes the transpose of the linear map $\mathrm{d}\sigma_{q(t)}$.

EXERCISE 1.19. Let Q be a differentiable manifold, $L:\mathbb{R}\times TQ\to\mathbb{R}$ be a Lagrangian, $q:[a,b]\to Q$ be a smooth map, $\varphi:U\to\widetilde U\subset\mathbb{R}^n$ be a local chart on Q and $\tilde q=\varphi\circ q|_{q^{-1}(U)}$. Use the result of Exercise 1.18 (with σ the transition function between two coordinate charts) to show that, for all $t\in q^{-1}(U)$, the linear functional:

$$(\mathrm{d}\varphi_{q(t)})^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_\varphi}{\partial\dot q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big)-\frac{\partial L_\varphi}{\partial q}\big(t,\tilde q(t),\dot{\tilde q}(t)\big)\Big)\in T_{q(t)}Q^*$$

does not depend on the choice of the chart φ.

EXERCISE 1.20 (the double pendulum). For this exercise, let us pretend that physical space is two-dimensional and let us identify it with the complex plane $\mathbb{C}$, which is more convenient for writing down the formulas (three-dimensional physical space will be considered in Exercise 1.21 below). The *double pendulum* is the system consisting of two particles in $\mathbb{C}$ constrained in the following way: the first particle remains over a circle (say, centered at the origin of $\mathbb{C}$) of radius $r_1>0$ and the second particle remains over a circle of radius $r_2>0$ centered at the position of the first particle. The set $Q\subset\mathbb{C}^2$ of allowed pairs of positions for the two particles is then the image of the map:

$$\mathbb{R}\times\mathbb{R}\ni(\theta_1,\theta_2)\longmapsto(r_1e^{i\theta_1},r_1e^{i\theta_1}+r_2e^{i\theta_2})\in\mathbb{C}\times\mathbb{C}. \tag{1.13.4}$$

(a) Compute the differential of the map (1.13.4) and show that such map is a smooth immersion.
(b) Show that the map (1.13.4) passes to the quotient and induces an embedding from the torus $(\mathbb{R}/2\pi\mathbb{Z})^2$ to $\mathbb{C}^2$. Conclude that the image Q of (1.13.4) is a smooth submanifold of $\mathbb{C}^2$ diffeomorphic (through the map (1.13.4)) to the torus.
(c) Compute the tangent space of Q at a point $q = (q_1, q_2) \in \mathbb{C}^2$. Show that a force $R = (R_1, R_2) \in \mathbb{C}^2$ is orthogonal to $T_q Q$ if and only if R_2 is parallel to the line connecting the points $q_1, q_2 \in \mathbb{C}$ and R_1 is the sum of $-R_2$ with a vector parallel to the line connecting the point q_1 and the origin.
(d) Consider the potential $V : \mathbb{C}^2 \to \mathbb{R}$ defined by:

$$V(q_1, q_2) = -m_1 g \Re(q_1) - m_2 g \Re(q_2), \quad q_1, q_2 \in \mathbb{C},$$

where g is a positive constant, $m_1, m_2 > 0$ denote the masses of the particles and $\Re(z)$ denotes the real part of $z \in \mathbb{C}$ (this corresponds to the potential of a force of magnitude $m_j g$ pointing to the direction of the positive real axis, acting on the j-th particle. For instance, this could arise from what is called a *homogeneous gravitational field of magnitude* g pointing to the direction of the positive real axis). Write down the representation of the Lagrangian L^{cons} (corresponding to the potential V and the constraint given by Q) with respect to a local chart on Q given by the inverse of the restriction of the map (1.13.4) to some open set in which it is injective (for instance, an open square of side 2π). Write down also the corresponding Euler–Lagrange equation.

The condition on the forces R_1, R_2 that you discovered when solving item (c) is precisely the condition that one would expect under the assumption that particle number 1 is attached to the origin and particle number 2 is attached to particle number 1 by means of inextensible strings of negligible mass. The force R_2 is the *tension* exerted upon particle number 2 by the string connecting the two particles and it should be parallel to that string (i.e., parallel to the line connecting the two particles) — that is the standard assumption about string tension. The force R_1 is the sum of two string tensions, one exerted upon particle number 1 by the string connecting particle number 1 to the origin (which should be parallel to that string) and the other exerted upon particle number 1 by the string connecting both particles (which should[29] be equal to $-R_2$).

EXERCISE 1.21 (the double spherical pendulum). The *double spherical pendulum* is the system consisting of two particles in $\mathbb{R}^3$ constrained in the

[29] Here's the argument: assuming Newton's law of reciprocal actions, then the force exerted by particle number 2 upon the string is equal to $-R_2$. If R_2' denotes the force exerted by the string upon particle number 1 then the force exerted by particle number 1 upon the string is $-R_2'$. The total force on the string is then $-R_2 - R_2'$ and since the mass of the string is being neglected, we take such total force to be zero, i.e., $R_2' = -R_2$.

following way: the first particle remains over a sphere (say, centered at the origin of $\mathbb{R}^3$) of radius $r_1 > 0$ and the second particle remains over a sphere of radius $r_2 > 0$ centered at the position of the first particle. Check that the submanifold Q of $\mathbb{R}^3 \times \mathbb{R}^3$ corresponding to such constraint is the image under the linear isomorphism:

$$\mathbb{R}^3 \times \mathbb{R}^3 \ni (u_1, u_2) \longmapsto (r_1 u_1, r_1 u_1 + r_2 u_2) \in \mathbb{R}^3 \times \mathbb{R}^3 \tag{1.13.5}$$

of the product $S^2 \times S^2$, where $S^2 \subset \mathbb{R}^3$ denotes the unit sphere centered at the origin. Conclude that Q is a smooth submanifold of $\mathbb{R}^3 \times \mathbb{R}^3$. Compute the tangent space $T_q Q$ at a point $q = (q_1, q_2) \in Q$ and show that a force $R = (R_1, R_2) \in \mathbb{R}^3 \times \mathbb{R}^3$ is orthogonal to $T_q Q$ if and only if it satisfies the condition that appears in item (c) of Exercise 1.20 above. If you want to have some fun, you can choose a potential V and write down the representation of the Lagrangian L^{cons} and the Euler–Lagrange equation with respect to your favorite chart on Q (for instance, you can obtain one by means of the map (1.13.5) and of spherical coordinates on the sphere S^2).

EXERCISE 1.22. Let $m_1, \ldots, m_n > 0$ denote the masses of the particles and consider the linear isomorphism $M : (\mathbb{R}^3)^n \to (\mathbb{R}^3)^n$ defined by:

$$M(q_1, \ldots, q_n) = (m_1 q_1, \ldots, m_n q_n), \quad q_1, \ldots, q_n \in \mathbb{R}^3.$$

Denote by $\langle \cdot, \cdot \rangle$ the standard inner product of $(\mathbb{R}^3)^n$ and by $\langle \cdot, \cdot \rangle_M$ the inner product of $(\mathbb{R}^3)^n$ defined by:

$$\langle q, q' \rangle_M = \langle M(q), q' \rangle, \quad q, q' \in (\mathbb{R}^3)^n.$$

If $V : \mathrm{dom}(V) \subset \mathbb{R} \times (\mathbb{R}^3)^n \to \mathbb{R}$ is a smooth map, we denote by $\nabla_q^M V(t, q)$ the *gradient of V relative to the inner product $\langle \cdot, \cdot \rangle_M$* (with respect to the second variable), i.e.:

$$\langle \nabla_q^M V(t, q), \cdot \rangle_M = \frac{\partial V}{\partial q}(t, q) \in (\mathbb{R}^3)^{n*}, \quad (t, q) \in \mathrm{dom}(V).$$

Show that:

(a) $\nabla_q^M V(t, q) = M^{-1}\big(\nabla_q V(t, q)\big)$, for all $(t, q) \in \mathrm{dom}(V)$;

(b) given a subspace S of $(\mathbb{R}^3)^n$, then M maps the orthogonal complement of S with respect to the inner product $\langle \cdot, \cdot \rangle_M$ onto the orthogonal complement of S with respect to the standard inner product of $(\mathbb{R}^3)^n$;

(c) the vector (1.5.9) is orthogonal to $T_{q(t)} Q$ with respect to the standard inner product of $(\mathbb{R}^3)^n$ if and only if the vector:

$$\frac{\mathrm{d}^2 q}{\mathrm{d}t^2}(t) + \nabla_q^M V(t, q)$$

is orthogonal to $T_{q(t)} Q$ with respect to the inner product $\langle \cdot, \cdot \rangle_M$.

EXERCISE 1.23. Let Q be a smooth submanifold of $\mathbb{R}^n$. Given a point $q_0 \in Q$, show that there exists a unique map:

$$\alpha : T_{q_0} Q \times T_{q_0} Q \longrightarrow \mathbb{R}^n / T_{q_0} Q$$

having the following property: if $q : I \to Q$, $v : I \to \mathbb{R}^n$ are smooth curves defined in an interval I such that $v(t) \in T_{q(t)}Q$ for all $t \in I$ and if $q(t_0) = q_0$ for some $t_0 \in I$ then:

$$\alpha\big(\dot{q}(t_0), v(t_0)\big) = \dot{v}(t_0) + T_{q_0}Q \in \mathbb{R}^n / T_{q_0}Q.$$

Show that such map α is bilinear (*hint*: let f be an $\mathbb{R}^m$-valued smooth map defined in an open neighborhood U of q_0 in $\mathbb{R}^n$ such that $0 \in \mathbb{R}^m$ is a regular value of f and $Q \cap U = f^{-1}(0)$. Differentiate the equality $\mathrm{d}f_{q(t)}\big(v(t)\big) = 0$ at $t = t_0$). The map α is called the *second fundamental form* of the submanifold Q at the point q_0 and it is denoted by $\alpha^Q_{q_0}$. If one chooses an inner product on $\mathbb{R}^n$ (not necessarily the standard one) then one can identify the quotient $\mathbb{R}^n/T_{q_0}Q$ with the orthogonal complement of $T_{q_0}Q$ with respect to that inner product, obtaining from $\alpha^Q_{q_0}$ a bilinear form taking values in that orthogonal complement. That's the *second fundamental form relative to the chosen inner product.*

Exercise 1.24. Consider the map M and the inner product $\langle\cdot,\cdot\rangle_M$ defined in Exercise 1.22. Assuming that the trajectories $q = (q_1, \ldots, q_n)$ of the particles are obtained as critical points of the action functional $S_{L^{\mathrm{cons}}}$, show that the forces $R(t)$ exerted by the constraint are given by:

$$R(t) = M\Big(\alpha^Q_{q(t)}\big(\dot{q}(t), \dot{q}(t)\big) + P_{q(t)}\big[\nabla^M_q V\big(t, q(t)\big)\big]\Big),$$

where the second fundamental form of Q is taken to be relative to the inner product $\langle\cdot,\cdot\rangle_M$ (see Exercise 1.23) and $P_{q(t)}$ denotes the orthogonal projection with respect to $\langle\cdot,\cdot\rangle_M$ onto the orthogonal complement of $T_{q(t)}Q$ with respect to $\langle\cdot,\cdot\rangle_M$.

Exercise 1.25. Consider the inner product $\langle\cdot,\cdot\rangle_M$ defined in Exercise 1.22 and the Riemannian metric on the submanifold Q of $(\mathbb{R}^3)^n$ induced by such inner product. Show that if the potential V is zero then a curve $q : [a, b] \to Q$ is a critical point of the action functional $S_{L^{\mathrm{cons}}}$ if and only if q is a geodesic of Q (*hint*: what is the variational problem whose solutions are the geodesics of a Riemannian manifold?).

Symplectic forms over vector spaces.

Exercise 1.26. Let V, W be finite-dimensional real vector spaces and $B : V \times W \to \mathbb{R}$ be a bilinear map. Consider the linear map:

$$V \ni v \longmapsto B(v, \cdot) \in W^* \tag{1.13.6}$$

canonically associated to B. Given bases $\mathcal{E} = (e_i)_{i=1}^n$, $\mathcal{F} = (f_j)_{j=1}^m$ of V and W, respectively, we can associate an $n \times m$ matrix $[B]_{\mathcal{E}\mathcal{F}}$ to B whose entry at row i and column j is $B(e_i, f_j)$. Show that:

(a) given $v \in V$, $w \in W$, then:

$$B(v,w) = \sum_{i=1}^{n}\sum_{j=1}^{m} B(e_i, f_j)v_i w_j = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} [B]_{\mathcal{E}\mathcal{F}} \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix},$$

where $(v_1, \ldots, v_n)$ denotes the coordinates of v with respect to $\mathcal{E}$ and $(w_1, \ldots, w_m)$ denotes the coordinates of w with respect to $\mathcal{F}$;

(b) if V is endowed with the basis $\mathcal{E}$ and W^* with the dual basis of $\mathcal{F}$, show that the matrix that represents the linear map (1.13.6) is the transpose of $[B]_{\mathcal{E}\mathcal{F}}$;

(c) if V and W have the same dimension, show that B is *non degenerate* (i.e., given $v \in V$, if $B(v,w) = 0$ for all $w \in W$ then $v = 0$) if and only if the matrix $[B]_{\mathcal{E}\mathcal{F}}$ is invertible (*hint*: B is non degenerate if and only if the linear map (1.13.6) is injective).

Exercise 1.27. If (V,ω) is a symplectic space then, by the result of item (c) in Exercise 1.26, the matrix $H = [\omega]_{\mathcal{E}\mathcal{E}}$ is invertible, where $\mathcal{E}$ is any basis of V. Conclude from the fact that H is both invertible and anti-symmetric that V is even-dimensional (*hint*: take the determinant on both sides of $H^{\mathrm{t}} = -H$).

Exercise 1.28. Let V be a (not necessarily finite-dimensional) real vector space and S be a subspace of V. Given a bilinear map $B : V \times V \to \mathbb{R}$, then the *orthogonal complement* of S with respect to B is the subspace $S^\perp$ of V defined by:

$$S^\perp = \big\{ v \in V : B(v,w) = 0, \text{ for all } w \in S \big\}.$$

Show that, if S is finite-dimensional, then the following conditions are equivalent:

(a) $S \cap S^\perp = \{0\}$;

(b) $B|_{S\times S}$ is non degenerate;

(c) $V = S \oplus S^\perp$.

(*hint*: the only non trivial part is to prove that (b) implies $V = S + S^\perp$. For that, notice that, assuming (b), since S is finite-dimensional, the linear map $S \ni v \mapsto B(v,\cdot)|_S \in S^*$ is an isomorphism. Given $w \in V$, we can then find $v \in S$ such that the linear functional $B(v,\cdot)|_S$ is equal to $B(w,\cdot)|_S$. Notice that $w - v \in S^\perp$).

Exercise 1.29. Given symplectic spaces (V,ω), $(\widetilde{V},\tilde{\omega})$, show that the following conditions are equivalent for a linear map $T : V \to \widetilde{V}$:

(a) T is a symplectomorphism;

(b) the image under T of *any* symplectic basis of (V,ω) is a symplectic basis of $(\widetilde{V},\tilde{\omega})$;

(c) *there exists* a symplectic basis of (V,ω) whose image under T is a symplectic basis of $(\widetilde{V},\tilde{\omega})$.

EXERCISE 1.30. Let V be a real finite-dimensional vector space and $\omega : V \times V \to \mathbb{R}$ be a (not necessarily non degenerate) anti-symmetric bilinear form.

(a) Show that there exists a basis $(e_1, \ldots, e_n, e'_1, \ldots, e'_n, f_1, \ldots, f_m)$ of V such that:

$$\omega(e_i, e'_j) = \delta_{ij}, \quad \omega(e_i, e_j) = 0, \quad \omega(e'_i, e'_j) = 0,$$
$$\omega(e_i, f_k) = 0, \quad \omega(e'_i, f_k) = 0, \quad \omega(f_k, f_l) = 0,$$

for all $i, j = 1, \ldots, n$, $k, l = 1, \ldots, m$, where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$ (*hint*: the proof is almost identical to the proof of Proposition 1.7.5).

(b) Given a basis of V as above, show that $(f_1, \ldots, f_m)$ is a basis of the *kernel* of ω, i.e., the kernel of the linear map (1.7.1).

(c) Given a basis of V as above, show that ω is non degenerate if and only if $m = 0$.

Symplectic manifolds and Hamiltonians.

EXERCISE 1.31. Let (M, ω), $(\widetilde{M}, \tilde{\omega})$ be symplectic manifolds and let $\Phi : M \to \widetilde{M}$ be a symplectomorphism. Given a time-dependent Hamiltonian $H : \mathrm{dom}(H) \subset \mathbb{R} \times M \to \mathbb{R}$, we can push it to the manifold $\widetilde{M}$ using Φ, obtaining a time-dependent Hamiltonian $H_\Phi : \mathrm{dom}(H_\Phi) \subset \mathbb{R} \times \widetilde{M} \to \mathbb{R}$ such that $\mathrm{dom}(H_\Phi) = (\mathrm{Id} \times \Phi)\big(\mathrm{dom}(H)\big)$ (Id denotes the identity map of $\mathbb{R}$) and:

$$H_\Phi\big(t, \Phi(x)\big) = H(t, x),$$

for all $(t, x) \in \mathrm{dom}(H)$.

(a) Show that, for $(t, x) \in \mathrm{dom}(H)$, we have:

$$\mathrm{d}\Phi_x\big(\vec{H}(t, x)\big) = \overrightarrow{H_\Phi}\big(t, \Phi(x)\big).$$

(b) Let $x : I \to M$ be a smooth curve (defined over some interval $I \subset \mathbb{R}$) and let $t \in I$ be such that $\big(t, x(t)\big) \in \mathrm{dom}(H)$. Setting $\tilde{x} = \Phi \circ x$, show that:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = \vec{H}\big(t, x(t)\big), \tag{1.13.7}$$

if and only if:

$$\frac{\mathrm{d}\tilde{x}}{\mathrm{d}t}(t) = \overrightarrow{H_\Phi}\big(t, \tilde{x}(t)\big).$$

(c) Let now $\Phi : U \subset M \to \widetilde{U} \subset \mathbb{R}^{2n}$ be a symplectic chart and define H_Φ as above (of course, before pushing H using Φ, we have to restrict H to $\mathrm{dom}(H) \cap (\mathbb{R} \times U)$). Let $x : I \to M$ be a smooth curve and set $\tilde{x} = \Phi \circ x|_{x^{-1}(U)}$. Write $\tilde{x} = (\tilde{q}, \tilde{p})$, with:

$$\tilde{q} : x^{-1}(U) \subset I \longrightarrow \mathbb{R}^n, \quad \tilde{p} : x^{-1}(U) \subset I \longrightarrow \mathbb{R}^n.$$

Given $t \in I$ with $x(t) \in U$ and $\big(t, x(t)\big) \in \mathrm{dom}(H)$, show that (1.13.7) holds if and only if:

$$\frac{\mathrm{d}\tilde{q}}{\mathrm{d}t}(t) = \frac{\partial H_\Phi}{\partial p}\big(t, \tilde{q}(t), \tilde{p}(t)\big), \quad \frac{\mathrm{d}\tilde{p}}{\mathrm{d}t}(t) = -\frac{\partial H_\Phi}{\partial q}\big(t, \tilde{q}(t), \tilde{p}(t)\big).$$

Exercise 1.32. The goal of this exercise is to prove Darboux's theorem (Theorem 1.8.4).

(a) Show that, in order to prove Darboux's theorem, it is sufficient to prove the following result: if ω is a symplectic form over an open subset U of $\mathbb{R}^{2n}$ with $0 \in U$ and if $\omega(0)$ is the canonical symplectic form ω_0 of $\mathbb{R}^{2n}$ then there exists a smooth diffeomorphism Φ from an open neighborhood of 0 in $\mathbb{R}^{2n}$ onto an open neighborhood of 0 in U such that $\Phi^*\omega$ is (the restriction to the domain of Φ of) ω_0. (*hint*: use Corollary 1.7.7).

(b) Let ω and U be as in item (a). Show that there exists a smooth one-form λ over some open neighborhood V of 0 in U such that $\lambda(0) = 0$ and $\mathrm{d}\lambda = \omega_0 - \omega$.

(c) For $t \in \mathbb{R}$, consider the smooth two-form over U defined by:

$$\omega_t = (1 - t)\omega_0 + t\omega.$$

The set:

$$\big\{(t, x) \in \mathbb{R} \times V : \omega_t(x) \text{ is non degenerate}\big\} \tag{1.13.8}$$

is open in $\mathbb{R} \times V$ and contains $\mathbb{R} \times \{0\}$. We can define a smooth time-dependent vector field X with domain (1.13.8) such that:

$$i_{X_t}\omega_t = \lambda,$$

where $X_t = X(t, \cdot)$. Denote by F the flow of X with initial time $t_0 = 0$ (i.e., for each x, $t \mapsto F(t, x)$ is the maximal integral curve of X passing through x at $t = 0$). Show that for all $t \in \mathbb{R}$, the map $F_t = F(t, \cdot)$ is defined over an open neighborhood of the origin (*hint*: the domain of F is open in $\mathbb{R} \times V$ and contains $\mathbb{R} \times \{0\}$, since $X_t(0) = 0$, for all $t \in \mathbb{R}$).

(d) Use the result of Exercise A.2 and formulas (A.2.6) and (A.2.9) to prove that:

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^*\omega_t) = 0.$$

Conclude the proof of Darboux's theorem by observing that:

$$F_1^*\omega = \omega_0.$$

Canonical forms in a cotangent bundle.

Exercise 1.33. Let Q_1, Q_2, Q_3 be differentiable manifolds and:

$$\varphi : Q_1 \longrightarrow Q_2, \quad \psi : Q_2 \longrightarrow Q_3$$

be smooth diffeomorphisms. Show that:

$$\mathrm{d}^*(\psi \circ \varphi) = \mathrm{d}^*\psi \circ \mathrm{d}^*\varphi.$$

EXERCISE 1.34. Let Q, $\widetilde{Q}$ be differentiable manifolds and $\varphi : Q \to \widetilde{Q}$ be a smooth diffeomorphism. Let $H : \mathrm{dom}(H) \subset \mathbb{R} \times TQ^* \to \mathbb{R}$ be a time-dependent Hamiltonian and $H_\Phi : \mathrm{dom}(H_\Phi) \subset \mathbb{R} \times T\widetilde{Q}^* \to \mathbb{R}$ be the time-dependent Hamiltonian obtained by pushing H using $\Phi = \mathrm{d}\varphi^*$ (see Exercise 1.31). Given $(t, q, p) \in \mathrm{dom}(H)$, show that:

$$\mathrm{d}\varphi_q\Big(\frac{\partial H}{\partial p}(t, q, p)\Big) = \frac{\partial H_\Phi}{\partial p}\big(t, \tilde{q}, \tilde{p}\big),$$

where $(\tilde{q}, \tilde{p}) = \mathrm{d}^*\varphi(q, p)$. Conclude, using also the result of Exercise 1.31, that if the thesis of Proposition 1.9.6 holds for $\widetilde{Q}$ then it also holds for Q.

The Legendre transform.

EXERCISE 1.35. Let E, E' be real finite-dimensional vector spaces, let $f : \mathrm{dom}(f) \subset E \to \mathbb{R}$ be a map of class C^2 defined over some open subset $\mathrm{dom}(f)$ of E and let $T : E' \to E$ be a linear isomorphism. Consider the map $f \circ T : T^{-1}\big(\mathrm{dom}(f)\big) \to \mathbb{R}$. Show that:

(a) f is regular (resp., hyper-regular) if and only if $f \circ T$ is regular (resp., hyper-regular);

(b) if f is hyper-regular then the Legendre transform $(f \circ T)^*$ of $f \circ T$ is defined in the open subset $T^*\big(\mathrm{dom}(f^*)\big)$ of E'^* and it is equal to $f^* \circ T^{*-1}$, where $T^* : E^* \to E'^*$ denotes the transpose of the linear map T and f^* denotes the Legendre transform of f.

EXERCISE 1.36. Let Q, $\widetilde{Q}$ be differentiable manifolds and $\varphi : Q \to \widetilde{Q}$ be a smooth diffeomorphism. Let $L : \mathrm{dom}(L) \subset \mathbb{R} \times TQ \to \mathbb{R}$ be a Lagrangian on Q and let $L_\varphi : \mathrm{dom}(L_\varphi) \subset \mathbb{R} \times T\widetilde{Q} \to \mathbb{R}$ be the Lagrangian on $\widetilde{Q}$ obtained by pushing L using φ, i.e., $\mathrm{dom}(L_\varphi) = (\mathrm{Id} \times \mathrm{d}\varphi)\big(\mathrm{dom}(L)\big)$ (Id denotes the identity map of $\mathbb{R}$) and:

$$L_\varphi \circ (\mathrm{Id} \times \mathrm{d}\varphi)|_{\mathrm{dom}(L)} = L.$$

Show that the diagram:

$$\begin{tikzcd}
\mathrm{dom}(L) \arrow[r, "\mathbb{F}L"] \arrow[d, "(\mathrm{Id}\times \mathrm{d}\varphi)|_{\mathrm{dom}(L)}"'] & \mathbb{R} \times TQ^* \arrow[d, "\mathrm{Id}\times \mathrm{d}^*\varphi"] \\
\mathrm{dom}(L_\varphi) \arrow[r, "\mathbb{F}L_\varphi"'] & \mathbb{R} \times T\widetilde{Q}^*
\end{tikzcd}$$

commutes. Conclude that claim (a) in the proof of Lemma 1.10.8 is true.

EXERCISE 1.37. Let $f : U \subset \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map defined over an open subset U of $\mathbb{R}^m \times \mathbb{R}^n$. Show that the following conditions are equivalent:

(a) for all $x \in \mathbb{R}^m$, the map:

$$f(x, \cdot) : \big\{y \in \mathbb{R}^n : (x, y) \in U\big\} \ni y \longmapsto f(x, y) \in \mathbb{R}^n$$

is a local diffeomorphism;

(b) the map $U \ni (x,y) \mapsto \big(x, f(x,y)\big) \in \mathbb{R}^m \times \mathbb{R}^n$ is a local diffeomorphism.

(*hint*: compute the differential of the map $(x,y) \mapsto \big(x, f(x,y)\big)$ and use the inverse function theorem).

Exercise 1.38. Let (Q,g) be a Riemannian manifold and:

$$V : \mathrm{dom}(V) \subset \mathbb{R} \times Q \longrightarrow \mathbb{R}$$

be a smooth map defined over an open subset $\mathrm{dom}(V)$ of $\mathbb{R} \times Q$. Define a Lagrangian L on Q by setting:

$$L(t,q,\dot q) = \frac{1}{2}\, g_q(\dot q, \dot q) - V(t,q), \tag{1.13.9}$$

for all $(t,q) \in \mathrm{dom}(V)$ and all $\dot q \in T_qQ$. If the Riemannian metric of Q is the one defined in Exercise 1.25 (and if V is the restriction to $\mathbb{R} \times Q$ of the potential defined over an open subset of $\mathbb{R} \times (\mathbb{R}^3)^n$) then the Lagrangian (1.13.9) is precisely the Lagrangian L^{cons} of Subsection 1.5.1. Show that L is hyper-regular and that its Legendre transform H is given by:

$$H(t,q,p) = \frac{1}{2}\, g_q^{-1}(p,p) + V(t,q),$$

for all $(t,q) \in \mathrm{dom}(V)$ and all $p \in T_qQ^*$, where g_q^{-1} is the inner product on the dual space T_qQ^* that turns the linear isomorphism:

$$T_qQ \ni \dot q \longmapsto g_q(\dot q, \cdot) \in T_qQ^*$$

into a linear isometry[30].

Symmetry and conservation laws.

Exercise 1.39. Let (M,ω) be a symplectic manifold and:

$$H : \mathrm{dom}(H) \subset \mathbb{R} \times M \longrightarrow \mathbb{R}$$

be a time-dependent Hamiltonian over M. Let G be a Lie group and:

$$\rho : G \times M \ni (g,x) \longmapsto g \cdot x \in M$$

be a smooth action of G on M. We say that ρ is a *symmetry* of the triple (M,ω,H) if for all $g \in G$ the diffeomorphism $\rho_g = \rho(g,\cdot)$ is a symplectomorphism of (M,ω) and if for all $(t,x) \in \mathrm{dom}(H)$ and all $g \in G$ we have $(t, g\cdot x) \in \mathrm{dom}(H)$ and $H(t, g\cdot x) = H(t,x)$.

(a) Given $X \in \mathfrak{g}$, show that the one-form $i_{X^M}\omega = \omega(X^M, \cdot)$ is closed (*hint*: the Lie derivative $\mathbb{L}_{X^M}\omega$ vanishes) and therefore locally given as a differential of a real valued smooth map.

[30] Given a basis of T_qQ and considering T_qQ^* to be endowed with the dual basis, then the matrix that represents g_q^{-1} is precisely the inverse of the matrix that represents g_q (see Exercise 1.41).

(b) If f is a real valued smooth map over an open subset $\mathrm{dom}(f)$ of M whose differential $\mathrm{d}f$ equals (the restriction to $\mathrm{dom}(f)$ of) $i_{X^M}\omega$, show that f is a first integral of the vector field $\vec{H}$, i.e., f is constant along integral curves of $\vec{H}$ that stay inside $\mathrm{dom}(f)$.

When the symplectic form ω is exact, so that $\omega = -\mathrm{d}\theta$ for some smooth one-form θ, and the action of G preserves θ then, for all $X \in \mathfrak{g}$, the one-form $i_{X^M}\omega$ is exact, since the map $f = \theta(X^M)$ satisfies $\mathrm{d}f = i_{X^M}\omega$ (see item (a) of Exercise 1.46).

Exercise 1.40. Show that:

(a) for each $v \in \mathbb{R}^3$, the linear endomorphism $X_v : w \mapsto v \wedge w$ of $\mathbb{R}^3$ is anti-symmetric and it is therefore an element of the Lie algebra $\mathfrak{so}(3)$ of the Lie group $\mathrm{SO}(3)$;
(b) the map $\mathbb{R}^3 \ni v \mapsto X_v \in \mathfrak{so}(3)$ is a linear isomorphism;
(c) given $v, w \in \mathbb{R}^3$, then the commutator:

$$[X_v, X_w] = X_v \circ X_w - X_w \circ X_v$$

is equal to $X_{v\wedge w}$. Conclude that $\mathbb{R}^3$ endowed with the vector product is a Lie algebra and that the map $v \mapsto X_v$ is an isomorphism from the Lie algebra $(\mathbb{R}^3, \wedge)$ onto the Lie algebra $\mathfrak{so}(3)$.

The Poisson bracket.

Exercise 1.41. Let V be a real finite-dimensional vector space and $B : V \times V \to \mathbb{R}$ be a non degenerate bilinear form, so that the linear map:

$$V \ni v \longmapsto B(v, \cdot) \in V^* \tag{1.13.10}$$

canonically associated to B is an isomorphism. Define a bilinear form:

$$B' : V^* \times V^* \longrightarrow \mathbb{R}$$

by carrying B to V^* using the linear isomorphism (1.13.10), i.e.:

$$B'\big(B(v, \cdot), B(w, \cdot)\big) = B(v, w), \quad v, w \in V.$$

Show that the linear map:

$$V^* \ni \alpha \longmapsto B'(\alpha, \cdot) \in V^{**} \cong V$$

canonically associated to B' is the inverse of the transpose of the linear map (1.13.10) (*hint*: check first that the transpose of (1.13.10) is given by $V \ni v \mapsto B(\cdot, v) \in V^*$). Conclude that, if V is endowed with a certain basis and V^* is endowed with the corresponding dual basis, then the matrix that represents B' is the inverse of the transpose of the matrix that represents B (see Exercise 1.26).

Exercise 1.42. Let (V, ω) be a symplectic space and $\pi : V^* \times V^* \to \mathbb{R}$ be the symplectic form over V^* for which the linear isomorphism:

$$V \ni v \longmapsto \omega(v, \cdot) \in V^*$$

is a symplectomorphism, i.e.:

$$\pi\big(\omega(v,\cdot),\omega(w,\cdot)\big)=\omega(v,w),\quad v,w\in V.$$

Show that a basis of V is symplectic for ω if and only if its dual basis is symplectic for π (*hint*: a basis is symplectic if and only if the matrix that represents the symplectic form with respect to that basis is $A=\left(\begin{smallmatrix}0_n & \mathrm{I}_n\\ -\mathrm{I}_n & 0_n\end{smallmatrix}\right)$, where 0_n and I_n denote the $n\times n$ zero matrix and the $n\times n$ identity matrix, respectively. Use the result of Exercise 1.41 and the fact that $A^{-1}=-A$).

Exercise 1.43. Let A be an associative algebra (for instance, the algebra of linear endomorphisms of a vector space) and define:

$$[x,y]=xy-yx,\quad x,y\in A. \tag{1.13.11}$$

Show that A is a Poisson algebra endowed with its associative product and with the commutator (1.13.11) (it is also a Poisson algebra if we replace the commutator (1.13.11) with any scalar multiple of it).

Exercise 1.44. Let (M,ω), $(\widetilde{M},\tilde{\omega})$ be symplectic manifolds and let $\Phi:M\to\widetilde{M}$ be a symplectomorphism. Show that the map:

$$\Phi^*:C^\infty(\widetilde{M})\ni f\longmapsto f\circ\Phi\in C^\infty(M)$$

is an *isomorphism of Poisson algebras*, i.e., it is a linear isomorphism that preserves both the associative (pointwise) product:

$$\Phi^*(fg)=\Phi^*(f)\Phi^*(g),\quad f,g\in C^\infty(\widetilde{M}),$$

and the Poisson bracket:

$$\Phi^*\big(\{f,g\}\big)=\{\Phi^*(f),\Phi^*(g)\},\quad f,g\in C^\infty(\widetilde{M}).$$

Exercise 1.45. Let M be an open subset of $\mathbb{R}^{2n}$ endowed with the canonical symplectic form (Example 1.7.2). Given a map $f:M\to\mathbb{R}$, denote its $2n$ partial derivatives by:

$$\frac{\partial f}{\partial q^1},\dots,\frac{\partial f}{\partial q^n},\frac{\partial f}{\partial p^1},\dots,\frac{\partial f}{\partial p^n}.$$

Given $f,g\in C^\infty(M)$, show that their Poisson bracket is given by:

$$\{f,g\}=\sum_{j=1}^n\frac{\partial f}{\partial q^j}\frac{\partial g}{\partial p^j}-\frac{\partial f}{\partial p^j}\frac{\partial g}{\partial q^j}.$$

Exercise 1.46. Let (M,ω) be a symplectic manifold whose symplectic form is exact and let θ be a smooth one-form over M with $\omega=-\mathrm{d}\theta$. Assume that we are given a smooth action $\rho:G\times M\to M$ of a Lie group G on M that preserves the one-form θ, i.e., $\rho_g^*\theta=\theta$ for all $g\in G$, where $\rho_g=\rho(g,\cdot)$ (this is the case, for instance, if M is a cotangent bundle TQ^*, θ is the canonical one-form and the action of G on $M=TQ^*$ is obtained from an action of G on Q). Each X in the Lie algebra $\mathfrak{g}$ induces a vector field X^M on M (when $M=TQ^*$ and the action of G is a symmetry of a time-dependent

Hamiltonian then the map $\theta(X^M)$ is precisely the conserved quantity given by Theorem 1.11.1).

(a) For $X \in \mathfrak{g}$, show that the differential of the map $\theta(X^M)$ is:

$$i_{X^M}\omega = \omega(X^M, \cdot)$$

(*hint*: write the Lie derivative of θ along X^M using formula (A.2.9) and observe that such Lie derivative must vanish). Conclude that the symplectic gradient of $\theta(X^M)$ is X^M.

(b) For $X, Y \in \mathfrak{g}$, show that the Poisson bracket $\{\theta(X^M), \theta(Y^M)\}$ equals $\omega(X^M, Y^M)$.

(c) For $X, Y \in \mathfrak{g}$, show that $\omega(X^M, Y^M) = -\theta\big([X^M, Y^M]\big)$ (*hint*: compute $\mathrm{d}\theta(X^M, Y^M)$ using formula (A.2.8). Use the result of item (a) to conclude that $X^M\big(\theta(Y^M)\big) = \omega(Y^M, X^M)$).

(d) Show that the map $\mathfrak{g} \ni X \longmapsto \theta(X^M) \in C^\infty(M)$ is a Lie algebra homomorphism if $C^\infty(M)$ is endowed with the Poisson bracket (*hint*: use formula (A.3.4)).

APPENDIX A

A summary of certain prerequisites

This notes are intendend as a course for mathematicians and graduate students in Mathematics. Therefore, a lot of standard material from graduate mathematical courses are taken as prerequisites. Nevertheless, in this appendix we make a quick presentation of some of those prerequisites. If you don't have any familiarity with such prerequisites, you probably won't be able to learn them using this appendix, but if you have some familiarity with them, this appendix might be useful for a quick review or as a quick reference guide. Most results will be stated without proof.

A.1. Quick review of multilinear algebra

Throughout the section, V denotes a fixed real finite-dimensional vector space. For most of what is presented in the section, the field of real numbers can be replaced with an arbitrary field[1] and for *everything* that is presented in the section it can be replaced with an arbitrary field of characteristic zero. We choose to use the field of real numbers for the presentation only for psychological reasons.

Given natural numbers r, s, then an *(r,s)-tensor over V* (also called a tensor that is *r times covariant and s times contravariant*) is a multilinear map:

$$\tau : V \times \cdots \times V \times V^* \times \cdots \times V^* \longrightarrow \mathbb{R}$$

in which there are r copies of V and s copies of the dual space V^*. The set of (r,s)-tensors over V is, in a natural way, a real vector space. Such vector space is naturally isomorphic to the tensor product of r copies of the dual space V^* and s copies of the space V:

$$\Big(\bigotimes_r V^*\Big) \otimes \Big(\bigotimes_s V\Big). \tag{A.1.1}$$

We won't need such identification between spaces of multilinear maps and tensor products of vector spaces, but we will use (A.1.1) as a notation for the space of (r,s)-tensors over V. The space of $(0,0)$-tensors over V is simply the field of real numbers. The space of $(0,1)$-tensors over V is the

[1] In the definition of wedge product, the factorial of the degree of the forms appears in the denominator and that doesn't make sense in general if the characteristic of the field is not zero. Also, if the characteristic of the field is two, then anti-symmetry is the same as symmetry and it is not true that an anti-symmetric map vanishes when two of its entries are equal.

bidual V^{**} and it will be identified in the usual way with the space V itself, so that elements of V are regarded as $(0,1)$-tensors over V. The space of $(1,0)$-tensors over V is simply the dual space V^*. The space of linear endomorphisms of V can be naturally identified with the space of $(1,1)$-tensors over V: namely, we identify a linear endomorphism $T : V \to V$ with the bilinear map $V \times V^* \ni (v, \alpha) \mapsto \alpha\big(T(v)\big) \in \mathbb{R}$.

Given some other real finite-dimensional vector space W, a linear isomorphism $T : W \to V$ and an (r,s)-tensor τ over V then the *pull-back* $T^*\tau$ is the (r,s)-tensor over W defined by:

$$(T^*\tau)(w_1, \ldots, w_r, \beta_1, \ldots, \beta_s) = \tau\big(T(w_1), \ldots, T(w_r), \beta_1 \circ T^{-1}, \ldots, \beta_s \circ T^{-1}\big), \tag{A.1.2}$$

for all $w_1, \ldots, w_r \in W$, $\beta_1, \ldots, \beta_s \in W^*$. When the tensor τ is *purely covariant*, i.e., when $s = 0$, then the pull-back $T^*\tau$ is defined for *any* linear map $T : W \to V$ (because we don't need to use T^{-1} in (A.1.2) when $s = 0$). The operation $\tau \mapsto T^*\tau$ defines a linear map:

$$T^* : \Big(\bigotimes_r V^*\Big) \otimes \Big(\bigotimes_s V\Big) \longrightarrow \Big(\bigotimes_r W^*\Big) \otimes \Big(\bigotimes_s W\Big)$$

in which it is assumed that T be an isomorphism if $s \neq 0$. Given some other real finite-dimensional vector space P and a linear map $S : P \to W$ then:

$$(T \circ S)^*\tau = S^*T^*\tau,$$

for any tensor τ over V, in which it is necessary to assume that T and S be isomorphisms if τ is not purely covariant.

Given an (r,s)-tensor τ over V and an (r',s')-tensor τ' over V, then their *tensor product* is the $(r+r', s+s')$-tensor $\tau \otimes \tau'$ over V defined by:

$$(\tau \otimes \tau')(v_1, \ldots, v_{r+r'}, \alpha_1, \ldots, \alpha_{s+s'}) = \tau(v_1, \ldots, v_r, \alpha_1, \ldots, \alpha_s)\tau'(v_{r+1}, \ldots, v_{r+r'}, \alpha_{s+1}, \ldots, \alpha_{s+s'}),$$

for all $v_1, \ldots, v_{r+r'} \in V$, $\alpha_1, \ldots, \alpha_{s+s'} \in V^*$. The product $c\tau$ of a tensor τ by a real number c coincides with the tensor product $c \otimes \tau$ (and also with $\tau \otimes c$), where the real number c is seen as a $(0,0)$-tensor. Clearly, the tensor product operation is associative, i.e., if τ, τ', τ'' are tensors over V then:

$$(\tau \otimes \tau') \otimes \tau'' = \tau \otimes (\tau' \otimes \tau''),$$

so that we can write tensor products of several tensors without parenthesis. If $(e_1, \ldots, e_n)$ is a basis of V and $(e^1, \ldots, e^n)$ denotes its dual basis then:

$$e^{i_1} \otimes \cdots \otimes e^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}, \quad i_1, \ldots, i_r, j_1, \ldots, j_s = 1, \ldots, n, \tag{A.1.3}$$

is a basis of the space of (r,s)-tensors over V. The dimension of the space of (r,s)-tensors over V is therefore equal to n^{r+s}. The coordinates of an (r,s)-tensor τ over V with respect to the basis (A.1.3) are given by:

$$\tau_{i_1 \ldots i_r}^{j_1 \ldots j_s} = \tau(e_{i_1}, \ldots, e_{i_r}, e^{j_1}, \ldots, e^{j_s}). \tag{A.1.4}$$

We say that the family of real numbers (A.1.4) *represents* the tensor τ with respect to the basis $(e_1, \ldots, e_n)$. The r lower indexes in (A.1.4) are normally called the *covariant* indexes and the s upper indexes the *contravariant* indexes.

Pull-backs preserve tensor products: if $T : W \to V$ is a linear map and τ, τ' are tensors over V then:

$$T^*(\tau \otimes \tau') = (T^*\tau) \otimes (T^*\tau'), \tag{A.1.5}$$

where we have to assume that T be an isomorphism if either τ or τ' is not purely covariant.

In what follows, we will focus on purely covariant tensors and we will use the symbol $\bigotimes_k V^*$ to denote the space of all k-linear maps $\tau : V^k \to \mathbb{R}$ (i.e., the space of all $(k, 0)$-tensors over V). The subspace of $\bigotimes_k V^*$ consisting of *anti-symmetric* k-linear maps will be denoted by $\bigwedge_k V^*$. For $k = 0$, both $\bigotimes_k V^*$ and $\bigwedge_k V^*$ are just the scalar field $\mathbb{R}$. If k is larger than the dimension of V, then $\bigwedge_k V^*$ is the null space.

If $\kappa \in \bigwedge_k V^*$ is an anti-symmetric k-linear map and $\lambda \in \bigwedge_l V^*$ is an anti-symmetric l-linear map then the *wedge product* $\kappa \wedge \lambda \in \bigwedge_{k+l} V^*$ is the $(k + l)$-linear anti-symmetric map defined by[2]:

$$(\kappa \wedge \lambda)(v_1, \ldots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S^{k+l}} \mathrm{sgn}(\sigma)(\kappa \otimes \lambda)(v_{\sigma(1)}, \ldots, v_{\sigma(k+l)}),$$

for all $v_1, \ldots, v_{k+l} \in V$, where S^{k+l} denotes the group of all bijections of the set $\{1, \ldots, k + l\}$ and $\mathrm{sgn}(\sigma)$ denotes the *sign* of the permutation σ, i.e., $\mathrm{sgn}(\sigma) = 1$ if σ is even and $\mathrm{sgn}(\sigma) = -1$ if σ is odd. The product $c\kappa$ of an anti-symmetric k-linear map κ by a real number c coincides with the wedge product $c \wedge \kappa$ (and also with $\kappa \wedge c$), where the real number c is seen as an element of $\bigwedge_0 V^*$. The wedge product operation is associative, i.e., if κ, λ and μ are anti-symmetric purely covariant tensors over V then:

$$(\kappa \wedge \lambda) \wedge \mu = \kappa \wedge (\lambda \wedge \mu),$$

so that we can write wedge products of several anti-symmetric covariant tensors without parenthesis. Given $\kappa_i \in \bigwedge_{k_i} V^*$, $i = 1, \ldots, r$, then the following formula holds:

$$\begin{aligned}(\kappa_1 \wedge \cdots \wedge \kappa_r)&(v_1, \ldots, v_k)\\ &= \frac{1}{k_1! \cdots k_r!} \sum_{\sigma \in S^k} \mathrm{sgn}(\sigma)(\kappa_1 \otimes \cdots \otimes \kappa_r)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),\end{aligned}$$

[2]Some authors use $\frac{1}{(k+l)!}$ in front of the summation sign. This difference in the definition of the exterior product also influences the definition of exterior differentiation of differential forms (see Subsection A.2.2), since the properties that characterize the exterior differential depend on the definition of the wedge product.

for all $v_1, \ldots, v_k \in V$, where $k = k_1 + \cdots + k_r$. In particular, given linear functionals $\alpha_i \in V^* = \bigwedge_1 V^*$, $i = 1, \ldots, r$, then:

$$(\alpha_1 \wedge \cdots \wedge \alpha_r)(v_1, \ldots, v_r) = \sum_{\sigma \in S^r} \operatorname{sgn}(\sigma)\alpha_1(v_{\sigma(1)}) \cdots \alpha_r(v_{\sigma(r)}),$$

i.e.:

$$(\alpha_1 \wedge \cdots \wedge \alpha_r)(v_1, \ldots, v_r) = \det\big(\alpha_i(v_j)\big)_{r \times r}, \tag{A.1.6}$$

for all $v_1, \ldots, v_r \in V$. If $(e_1, \ldots, e_n)$ is a basis of V and $(e^1, \ldots, e^n)$ denotes its dual basis then:

$$e^{i_1} \wedge \cdots \wedge e^{i_k}, \quad 1 \le i_1 < i_2 < \cdots < i_k \le n, \tag{A.1.7}$$

is a basis of $\bigwedge_k V^*$. Thus, for $0 \le k \le n$, the dimension of $\bigwedge_k V^*$ is $\binom{n}{k}$ and the dimension of $\bigwedge_n V^*$ is equal to 1. The non zero elements of $\bigwedge_n V^*$ are called *volume forms* over V. The coordinates of $\kappa \in \bigwedge_k V^*$ with respect to the basis (A.1.7) are:

$$\kappa_{i_1 \ldots i_k} = \kappa(e_{i_1}, \ldots, e_{i_k}).$$

For $\kappa \in \bigwedge_k V^*$, $\lambda \in \bigwedge_l V^*$, we have:

$$\kappa \wedge \lambda = (-1)^{kl} \lambda \wedge \kappa,$$

so that $\kappa \wedge \kappa = 0$ if $\kappa \in \bigwedge_k V^*$ and k is *odd.*

Set:

$$\bigotimes V^* = \bigoplus_{k=0}^{\infty} \bigotimes_k V^*, \quad \bigwedge V^* = \bigoplus_{k=0}^{\infty} \bigwedge_k V^* = \bigoplus_{k=0}^{n} \bigwedge_k V^*,$$

where $n = \dim(V)$. The tensor product operation of purely covariant tensors extends in a unique way to a bilinear binary operation in the space $\bigotimes V^*$ and the wedge product operation of anti-symmetric purely covariant tensors extends in a unique way to a bilinear binary operation in the space $\bigwedge V^*$. Both $\bigotimes V^*$ and $\bigwedge V^*$ become associative (graded) real algebras with unit endowed with such binary operations. Observe that $\bigwedge V^*$ is a vector subspace but *not a subalgebra* of $\bigotimes V^*$.

Given a linear map $T : W \to V$, then the pull-back operation $\tau \mapsto T^*\tau$ (on purely covariant tensors over V) extends to an algebra homomorphism:

$$T^* : \bigotimes V^* \longrightarrow \bigotimes W^*,$$

so that (A.1.5) holds for all $\tau, \tau' \in \bigotimes V^*$. The restriction of T^* to $\bigwedge V^*$ gives an algebra homomorphism:

$$T^* : \bigwedge V^* \longrightarrow \bigwedge W^*,$$

so that:

$$T^*(\kappa \wedge \lambda) = (T^*\kappa) \wedge (T^*\lambda),$$

for all $\kappa, \lambda \in \bigwedge V^*$.

Given a vector $v \in V$ and a tensor $\tau \in \bigotimes_k V^*$, $k \geq 1$, we define the *interior product of* τ *by* v to be the tensor $i_v\tau \in \bigotimes_{k-1} V^*$ given by:

$$(i_v\tau)(v_1, \ldots, v_{k-1}) = \tau(v, v_1, \ldots, v_{k-1}),$$

for all $v_1, \ldots, v_{k-1} \in V$. For $\tau \in \bigotimes_0 V^*$ we set $i_v\tau = 0$. The map i_v extends in a unique way to a linear endomorphism (not an algebra homomorphism!) of $\bigotimes V^*$. Such endomorphism sends $\bigwedge V^*$ to $\bigwedge V^*$ and for $\kappa \in \bigwedge_k V^*$, $\lambda \in \bigwedge_l V^*$ we have:

$$i_v(\kappa \wedge \lambda) = (i_v\kappa) \wedge \lambda + (-1)^k \kappa \wedge (i_v\lambda). \tag{A.1.8}$$

A.2. Quick review of calculus on manifolds

We have selected for presentation some topics which are taught during courses on calculus on manifolds. We won't present either the definition of differentiable manifold or the construction of the tangent bundle. The word "smooth" always refers to "class C^∞". Differentiable manifolds are always assumed to be smooth (i.e., endowed with a smooth atlas) and to have a Hausdorff topology. For maps, the word "differentiable" is to be taken literally, i.e., a differentiable map is map that can be differentiated once.

A.2.1. Vector fields and flows. Let M be a differentiable manifold and let X be a smooth vector field over M, i.e., X is a smooth map from M to the tangent bundle TM such that $X(x) \in T_xM$, for all $x \in M$. By an *integral curve* of X we mean a differentiable map $x : I \to M$, defined over some interval $I \subset \mathbb{R}$, such that:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = X\big(x(t)\big),$$

for all $t \in I$. Given $t_0 \in \mathbb{R}$ and $x_0 \in M$, there exists a unique integral curve $x : I \to M$ of X with $t_0 \in I$, $x(t_0) = x_0$ and that is *maximal*, i.e., it cannot be extended to an integral curve of X defined in a strictly larger interval. Every integral curve of X is a restriction of a maximal integral curve. The *flow* of X is the map $F : \mathrm{dom}(F) \subset \mathbb{R} \times M \to M$ such that, for all $x_0 \in M$:

$$\big\{t \in \mathbb{R} : (t, x_0) \in \mathrm{dom}(F)\big\} \ni t \longmapsto F(t, x_0) \in M$$

is the maximal integral curve of X passing through x_0 at $t = 0$. The domain of F is open in $\mathbb{R} \times M$ and the map F is smooth. For each $t \in \mathbb{R}$, the map:

$$F_t : \mathrm{dom}(F_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(F)\big\} \ni x \longmapsto F(t, x) \in M$$

is a smooth diffeomorphism between open subsets of M whose inverse is the map F_{-t} (the image of F_t is precisely the domain of F_{-t}). The map F_0 is the identity map of M. Given $t, s \in \mathbb{R}$, if $x \in M$ is in the domain of F_t and $F_t(x)$ is in the domain of F_s then x is in the domain of F_{t+s} and:

$$F_{t+s}(x) = F_s\big(F_t(x)\big).$$

Let X be a smooth *time-dependent vector field* over a differentiable manifold M, i.e., X is a smooth map from an open subset $\mathrm{dom}(X)$ of $\mathbb{R} \times M$ to the tangent bundle TM such that $X(t,x) \in T_xM$, for all $(t,x) \in \mathrm{dom}(X)$. For each $t \in \mathbb{R}$, we obtain from X a smooth vector field $X_t = X(t,\cdot)$ over the open set:

$$\mathrm{dom}(X_t) = \big\{x \in X : (t,x) \in \mathrm{dom}(X)\big\}.$$

By an *integral curve* of X we mean a differentiable map $x : I \to M$, defined over some interval $I \subset \mathbb{R}$, such that $\big(t, x(t)\big) \in \mathrm{dom}(X)$ and:

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = X\big(t, x(t)\big),$$

for all $t \in I$. Given $t_0 \in \mathbb{R}$ and $x_0 \in M$, if $(t_0, x_0) \in \mathrm{dom}(X)$, there exists a unique integral curve $x : I \to M$ of X with $t_0 \in I$, $x(t_0) = x_0$ and that is maximal (in the sense explained above). Again, every integral curve of X is a restriction of a maximal integral curve. For vector fields that *do not* depend on time, it is true that the time translation of an integral curve is an integral curve (i.e., if $t \mapsto x(t)$ is an integral curve and $t_0 \in \mathbb{R}$ is given then $t \mapsto x(t_0 + t)$ is an integral curve); for that reason, one only considers integral curves satisfying some initial condition at $t = 0$ when defining the flow. For time-dependent vector fields it is not true that the time translation of an integral curve is an integral curve, so it is relevant to consider a flow with an arbitrary initial time t_0: for a fixed $t_0 \in \mathbb{R}$, we define the *flow with initial time* t_0 of the time-dependent vector field X to be the map $F^{t_0} : \mathrm{dom}(F^{t_0}) \subset \mathbb{R} \times \mathrm{dom}(X_{t_0}) \subset \mathbb{R} \times M \to M$ such that, for all $x_0 \in \mathrm{dom}(X_{t_0})$:

$$\big\{t \in \mathbb{R} : (t, x_0) \in \mathrm{dom}(F^{t_0})\big\} \ni t \longmapsto F^{t_0}(t, x_0) \in M$$

is the maximal integral curve of X passing through x_0 at $t = t_0$. The domain of F^{t_0} is open in $\mathbb{R} \times M$ and the map F^{t_0} is smooth. In fact, we can say more; the set:

$$\big\{(t_0, t, x) \in \mathbb{R} \times \mathbb{R} \times M : (t, x) \in \mathrm{dom}(F^{t_0})\big\}$$

is open in $\mathbb{R} \times \mathbb{R} \times M$ and the map $(t_0, t, x) \mapsto F^{t_0}(t, x) \in M$ (defined on such set) is smooth[3]. For each $t \in \mathbb{R}$, the map:

$$F^{t_0}_t : \mathrm{dom}(F^{t_0}_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(F^{t_0})\big\} \ni x \longmapsto F^{t_0}(t, x) \in M$$

is a smooth diffeomorphism between open subsets of M whose inverse is the map $F^t_{t_0}$ (the image of $F^{t_0}_t$ is the domain of $F^t_{t_0}$). Given $t_0, t, s \in \mathbb{R}$, if x is in the domain of $F^{t_0}_t$ and $F^{t_0}_t(x)$ is in the domain of F^t_s then x is in the domain of $F^{t_0}_s$ and:

$$F^{t_0}_s(x) = F^t_s\big(F^{t_0}_t(x)\big).$$

[3]Such properties of the flow of a time-dependent vector field X are easily established as a corollary of the properties of the flow of the (time independent) vector field over the manifold $\mathrm{dom}(X) \subset \mathbb{R} \times M$ defined by $(t, x) \mapsto \big(1, X(t, x)\big)$.

For any $t_0 \in \mathbb{R}$, the map $F_{t_0}^{t_0}$ is the identity map of $\mathrm{dom}(X_{t_0})$. When X *do not* depend on time, we have that $F_t^{t_0}$ coincides with $F_{t-t_0}^0$.

A.2.2. Tensor fields and differential forms. Given a differentiable manifold M, then an (r,s)*-tensor field over* M (also called a tensor field that is r *times covariant and* s *times contravariant*) is a map τ that associates to each point $x \in M$ an (r,s)-tensor $\tau(x)$ (denoted also by τ_x) over the tangent space T_xM (see Section A.1). Scalar fields (i.e., real valued functions) are $(0,0)$-tensor fields and vector fields are (identified with) $(0,1)$-tensor fields. By a *differential form of degree* k (or simply a k*-form*) over M we mean an *anti-symmetric* $(k,0)$-tensor field κ over M (i.e., κ associates to each $x \in M$ an element $\kappa(x)$ of $\bigwedge_k T_xM^*$). A zero-form is the same as a scalar field.

The natural counter-domain for an (r,s)-tensor field over M is the disjoint union:

$$\text{(A.2.1)} \quad \Big(\bigotimes_r TM^*\Big) \otimes \Big(\bigotimes_s TM\Big) = \bigcup_{x\in M} \{x\} \times \Big[\Big(\bigotimes_r T_xM^*\Big) \otimes \Big(\bigotimes_s T_xM\Big)\Big].$$

The set (A.2.1) can, in a natural way, be turned into a differentiable manifold, so that it makes sense to talk about *smooth* tensor fields. An atlas for (A.2.1) is obtained as follows: given a local chart $\varphi : U \subset M \to \widetilde{U} \subset \mathbb{R}^n$ on M, we define a local chart:

$$\bigcup_{x\in U} \{x\} \times \Big[\Big(\bigotimes_r T_xM^*\Big) \otimes \Big(\bigotimes_s T_xM\Big)\Big] \longrightarrow \widetilde{U} \times \Big[\Big(\bigotimes_r \mathbb{R}^{n*}\Big) \otimes \Big(\bigotimes_s \mathbb{R}^n\Big)\Big]$$

over (A.2.1) by:

$$(x,\tau) \longmapsto \big(\varphi(x), (\mathrm{d}\varphi_x^{-1})^*\tau\big).$$

The operations of tensor product and wedge product can be defined (pointwise) for fields, i.e., if τ is an (r,s)-tensor field over M and τ' is an (r',s')-tensor field over M then $\tau \otimes \tau'$ is the $(r+r', s+s')$-tensor field over M defined by:

$$(\tau \otimes \tau')_x = \tau_x \otimes \tau'_x, \quad x \in M,$$

and, similarly, if κ is a k-form over M and λ is an l-form over M then $\kappa \wedge \lambda$ is the $(k+l)$-form over M defined by:

$$(\kappa \wedge \lambda)_x = \kappa_x \wedge \lambda_x, \quad x \in M.$$

The tensor product $f \otimes \tau$ (or $\tau \otimes f$) of a scalar field f by a tensor field τ is just the ordinary product $f\tau$ (i.e., the map that sends $x \in M$ to $f(x)\tau(x)$) and the wedge product $f \wedge \kappa$ (or $\kappa \wedge f$) of a scalar field f by a differential form κ is the same as the ordinary product $f\kappa$. The tensor product of smooth tensor fields is smooth and the wedge product of smooth differential forms is smooth.

The operation of interior product by vectors can also be defined (pointwise) for fields: if X is a vector field over M and τ is a $(k,0)$-tensor field over M then, for $k \geq 1$, $i_X\tau$ denotes the $(k-1,0)$-tensor field over M defined by:

$$(i_X\tau)(x) = i_{X(x)}\tau_x, \quad x \in M.$$

We set $i_X\tau = 0$ for $k = 0$. The interior product $i_X\tau$ of a smooth tensor field τ by a smooth vector field X is a smooth tensor field.

If τ is an (r,s)-tensor field over M, $X_1, \ldots, X_r$ are vector fields over M and $\alpha_1, \ldots, \alpha_s$ are one-forms over M then $\tau(X_1, \ldots, X_r, \alpha_1, \ldots, \alpha_s)$ denotes the scalar field over M defined by:

$$M \ni x \longmapsto \tau_x\big(X_1(x), \ldots, X_r(x), \alpha_1(x), \ldots, \alpha_s(x)\big) \in \mathbb{R}. \tag{A.2.2}$$

The scalar field (A.2.2) is smooth if τ, $X_1, \ldots, X_r$, $\alpha_1, \ldots, \alpha_s$ are smooth.

If $\varphi : N \to M$ is a smooth local diffeomorphism defined over a differentiable manifold N and if τ is an (r,s)-tensor field over M, we define the *pull-back* $\varphi^*\tau$ to be the (r,s)-tensor field over N given by:

$$(\varphi^*\tau)_y = (\mathrm{d}\varphi_y)^*\tau_{\varphi(y)}, \quad y \in N.$$

When the tensor field τ is purely covariant (i.e., when $s = 0$) then the pull-back $\varphi^*\tau$ is defined for *any* smooth map $\varphi : N \to M$. In particular, the pull-back $\varphi^*\kappa$ is well-defined for any smooth map φ if κ is a differential form. If the tensor field τ is smooth then the pull-back $\varphi^*\tau$ is also smooth. Given a smooth local diffeomorphism $\psi : P \to N$ defined over a differentiable manifold P then:

$$(\varphi \circ \psi)^*\tau = \psi^*\varphi^*\tau.$$

The assumption that φ, ψ be local diffeomorphisms is not necessary if τ is purely covariant.

A (smooth) *local frame* over an open subset U of a differentiable manifold M is a sequence $(e_1, \ldots, e_n)$ of (smooth) vector fields over U such that $\big(e_1(x), \ldots, e_n(x)\big)$ is a basis of T_xM, for all $x \in U$. An (r,s)-tensor field τ over M is represented with respect to such a frame by a family of maps:

$$\tau^{j_1 \ldots j_s}_{i_1 \ldots i_r} : U \longrightarrow \mathbb{R}, \quad i_1, \ldots, i_r, j_1, \ldots, j_s = 1, \ldots, n, \tag{A.2.3}$$

in which, for $x \in U$, the real numbers $\tau^{j_1 \ldots j_s}_{i_1 \ldots i_r}(x)$ represent $\tau(x)$ with respect to the basis $\big(e_1(x), \ldots, e_n(x)\big)$ (see (A.1.4)). If τ is smooth and the local frame $(e_1, \ldots, e_n)$ is smooth then the maps (A.2.3) are smooth. Conversely, if for some family of smooth local frames whose domains cover M the corresponding maps (A.2.3) representing τ are smooth then τ is smooth.

Given a smooth map $f : M \to \mathbb{R}$ and a vector field X over M, we set:

$$X(f) = \mathrm{d}f(X).$$

Given two smooth vector fields X, Y over M, then there exists a unique vector field Z over M such that:

$$Z(f) = X\big(Y(f)\big) - Y\big(X(f)\big), \tag{A.2.4}$$

for any smooth map $f : M \to \mathrm{I\!R}$. Such vector field Z is smooth.

A.2.1. Definition. The only vector field Z satisfying (A.2.4) is denoted by $[X, Y]$ and it is called the *Lie bracket* of the vector fields X, Y.

A.2.2. Definition. Let $\varphi : N \to M$ be a smooth map defined over a differentiable manifold N. If X is a vector field over M and X' is a vector field over N then we say that X' and X are *φ-related* (or *related by φ*) if:

$$X\big(\varphi(y)\big) = \mathrm{d}\varphi_y\big(X'(y)\big),$$

for all $y \in N$.

If φ is a local diffeomorphism then X and X' are φ-related if and only if X' equals the pull-back $\varphi^* X$. We have the following:

A.2.3. Proposition. *If $\varphi : N \to M$ is a smooth map, X', Y' are smooth vector fields over N that are φ-related, respectively, to smooth vector fields X, Y over M then the Lie bracket $[X', Y']$ is φ-related to the Lie bracket $[X, Y]$.* □

A.2.4. Definition. Let τ be a smooth (r, s)-tensor field over a differentiable manifold M and let X be a smooth vector field over M. The *Lie derivative* of τ with respect to X is the smooth (r, s)-tensor field $\mathbb{L}_X\tau$ over M defined by:

$$\mathbb{L}_X\tau = \frac{\mathrm{d}}{\mathrm{d}t} F_t^*\tau \Big|_{t=0}, \tag{A.2.5}$$

where F denotes the flow of X.

Since the map F_t is a smooth diffeomorphism between open subsets of M, the pull-back $F_t^*\tau$ is always well-defined. The righthand side of (A.2.5) is to be understood as follows: given any $x \in M$, the value of the righthand side of (A.2.5) at the point x is the derivative at $t = 0$ of the curve:

$$t \longmapsto (F_t^*\tau)(x) \in \Big(\bigotimes_r T_xM^*\Big) \otimes \Big(\bigotimes_s T_xM\Big).$$

Here are the main properties of the Lie derivative. In what follows, X denotes a smooth vector field over the differentiable manifold M.

(1) the Lie derivative commutes with restriction to open sets, i.e., if τ is a smooth tensor field over M and U is an open subset of M then the Lie derivative of $\tau|_U$ with respect to $X|_U$ is the restriction of $\mathbb{L}_X\tau$ to U:
$$\mathbb{L}_{(X|_U)}(\tau|_U) = (\mathbb{L}_X\tau)|_U.$$
(2) If $f : M \to \mathrm{I\!R}$ is a smooth map (regarded as a $(0, 0)$-tensor field over M) then:
$$\mathbb{L}_X f = X(f).$$

(3) If Y is a smooth vector field over M (regarded as a $(0,1)$-tensor field over M) then:

$$\mathbb{L}_X Y = [X, Y].$$

(4) If τ is a smooth (r,s)-tensor field over M, $X_1, \ldots, X_r$ are smooth vector fields over M and $\alpha_1, \ldots, \alpha_s$ are smooth one-forms over M then:

$$X\big(\tau(X_1,\ldots,X_r,\alpha_1,\ldots,\alpha_s)\big) = \sum_{i=1}^{r} \tau(X_1,\ldots,\mathbb{L}_X X_i,\ldots,X_r,\alpha_1,\ldots,\alpha_s) + \sum_{i=1}^{s} \tau(X_1,\ldots,X_r,\alpha_1,\ldots,\mathbb{L}_X \alpha_i,\ldots,\alpha_s).$$

(5) If τ, τ' are smooth tensor fields over M then:

$$\mathbb{L}_X(\tau\otimes\tau') = (\mathbb{L}_X\tau)\otimes\tau' + \tau\otimes(\mathbb{L}_X\tau').$$

(6) If κ, λ are smooth differential forms over M then:

$$\mathbb{L}_X(\kappa\wedge\lambda) = (\mathbb{L}_X\kappa)\wedge\lambda + \kappa\wedge(\mathbb{L}_X\lambda).$$

If τ is a smooth tensor field over M and F is the flow of a smooth vector field X over M then the derivative of $t \mapsto F_t^*\tau$ at an arbitrary instant can also be written in terms of the Lie derivative. In fact, this can be done even when X is a time-dependent vector field.

A.2.5. PROPOSITION. *Let τ be a smooth (r,s)-tensor field over a differentiable manifold M and let X be a smooth time-dependent vector field over M. Given $t_0 \in \mathbb{R}$, if F^{t_0} denotes the flow of X with initial time t_0 then:*

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^{t_0})^*\tau = (F_t^{t_0})^*\mathbb{L}_{X_t}\tau, \tag{A.2.6}$$

where $X_t = X(t,\cdot)$.

PROOF. Let $t_1 \in \mathbb{R}$ be fixed and let us show that (A.2.6) holds at $t = t_1$. Let G denote the flow of the vector field X_{t_1}. Set:

$$\widetilde{F}_t(x) = G_{t-t_1}\big(F_{t_1}^{t_0}(x)\big),$$

for all $(t,x) \in \mathbb{R}\times M$ for which the righthand side of the equality is well-defined. We have $\widetilde{F}_{t_1} = F_{t_1}^{t_0}$ and:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}_t(x)\Big|_{t=t_1} = \frac{\mathrm{d}}{\mathrm{d}t}F_t^{t_0}(x)\Big|_{t=t_1},$$

for all $x \in \mathrm{dom}(F_{t_1}^{t_0})$. It follows from the result of Exercise A.1 that:

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^{t_0})^*\tau\Big|_{t=t_1} = \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}_t^*\tau\Big|_{t=t_1}.$$

Moreover:

$$(\widetilde{F}_t^*\tau)(x) = \mathrm{d}F_{t_1}^{t_0}(x)^*\big[(G_{t-t_1}^*\tau)\big(F_{t_1}^{t_0}(x)\big)\big],$$

for all $t \in \mathbb{R}$ and all $x \in M$ in the domain of $\widetilde{F}_t$. Taking the derivative at $t = t_1$ on both sides and taking into account that the map $\mathrm{d}F_{t_1}^{t_0}(x)^*$ is linear, we obtain:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\widetilde{F}_t^*\tau)(x)\Big|_{t=t_1} &= \mathrm{d}F_{t_1}^{t_0}(x)^*\Big[\,\frac{\mathrm{d}}{\mathrm{d}t}(G_{t-t_1}^*\tau)\big(F_{t_1}^{t_0}(x)\big)\Big|_{t=t_1}\,\Big]\\ &= \mathrm{d}F_{t_1}^{t_0}(x)^*\big[(\mathbb{L}_{X_{t_1}}\tau)\big(F_{t_1}^{t_0}(x)\big)\big],\end{aligned}$$

for all x in the domain of $F_{t_1}^{t_0}$. The conclusion follows. □

A.2.6. Definition. We say that a smooth (r, s)-tensor field τ over a differentiable manifold M is *invariant* under the flow F of a smooth vector field X over M if $F_t^*\tau$ is equal to (the restriction to the domain of F_t of) τ, for all $t \in \mathbb{R}$. If X is a smooth time-dependent vector field over M, we say that τ is invariant under the flow of X if $(F_t^{t_0})^*\tau$ is equal to (the restriction to the domain of $F_t^{t_0}$ of) τ, for all $t_0, t \in \mathbb{R}$.

A.2.7. Proposition. *Let τ be a smooth (r, s)-tensor field over a differentiable manifold M and X be a smooth vector field over M. Then τ is invariant under the flow of X if and only if $\mathbb{L}_X\tau = 0$. If X is a smooth time-dependent vector field over M then τ is invariant under the flow of X if and only if $\mathbb{L}_{X_t}\tau = 0$, for all $t \in \mathbb{R}$, where $X_t = X(t, \cdot)$.*

Proof. It suffices to consider the time-dependent case. The tensor field τ is invariant under the flow of X if and only if:

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^{t_0})^*\tau = 0, \tag{A.2.7}$$

for all $t_0, t \in \mathbb{R}$. If $\mathbb{L}_{X_t}\tau = 0$ for all $t \in \mathbb{R}$ then (A.2.7) follows from (A.2.6). If (A.2.7) holds for all $t \in \mathbb{R}$, then using (A.2.6) with $t = t_0$, we obtain that $\mathbb{L}_{X_{t_0}}\tau = 0$ for all $t_0 \in \mathbb{R}$. □

Exterior differentiation is an operation that takes a smooth k-form κ over a differential manifold to a smooth $(k + 1)$-form $\mathrm{d}\kappa$ over that same manifold. Such operation is characterized by the following set of properties:

(1) exterior differentiation commutes with restriction to open sets, i.e., if U is an open subset of a differentiable manifold M and κ is a smooth differential form over M then the exterior differential of the restriction of κ to U is the restriction of $\mathrm{d}\kappa$ to U:
$$\mathrm{d}(\kappa|_U) = (\mathrm{d}\kappa)|_U.$$
(2) Given a differentiable manifold M, then the map $\kappa \mapsto \mathrm{d}\kappa$ that takes smooth k-forms over M to smooth $(k + 1)$-forms over M is linear (over the field of real numbers).
(3) Exterior differentiation agrees with ordinary differentiation over smooth zero-forms (i.e., over smooth real valued functions).

(4) The exterior differential of the exterior differential of a smooth differential form κ vanishes:

$$\mathrm{d}(\mathrm{d}\kappa) = 0.$$

(5) If κ is a smooth k-form over a differentiable manifold M and λ is a smooth l-form over M then:

$$\mathrm{d}(\kappa \wedge \lambda) = (\mathrm{d}\kappa) \wedge \lambda + (-1)^k \kappa \wedge \mathrm{d}\lambda.$$

Let us show how properties (1)–(5) of exterior differentiation can be used to compute the exterior differential of a differential form using a local chart. Let $\varphi : U \subset M \to \widetilde{U} \subset \mathbb{R}^n$ be a local chart on a differentiable manifold M. If one denotes by $x^i : U \to \mathbb{R}$, $i = 1, \ldots, n$, the coordinate functions of φ (so that $\varphi = (x^1, \ldots, x^n)$) then it is customary to denote by $\frac{\partial}{\partial x^i}$, $i = 1, \ldots, n$, the local frame over U such that $\frac{\partial}{\partial x^i}(p)$ is mapped by $\mathrm{d}\varphi(p)$ to the i-th vector of the canonical basis of $\mathbb{R}^n$, for all $p \in U$, $i = 1, \ldots, n$. If $\mathrm{d}x^i$ denotes the one-form which is the (exterior or ordinary) differential of the scalar field x^i then $\big(\mathrm{d}x^1(p), \ldots, \mathrm{d}x^n(p)\big)$ is the dual basis of $\big(\frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\big)$, for all $p \in U$. If κ is a k-form over M then:

$$\kappa|_U = \sum_I \kappa_I \mathrm{d}x^I,$$

where I runs over the k-tuples $(i_1, \ldots, i_k)$ with $1 \le i_1 < \cdots < i_k \le n$ and:

$$\kappa_I = \kappa\big(\tfrac{\partial}{\partial x^{i_1}}, \ldots, \tfrac{\partial}{\partial x^{i_k}}\big), \quad \mathrm{d}x^I = \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

Using properties (1)–(5) of exterior differentiation it follows that:

$$(\mathrm{d}\kappa)|_U = \sum_I \mathrm{d}\kappa_I \wedge \mathrm{d}x^I.$$

Exterior differentiation commutes with pull-backs: if M, N are differentiable manifolds, $\varphi : N \to M$ is a smooth map and κ is a smooth differential form over M then:

$$\mathrm{d}(\varphi^* \kappa) = \varphi^* \mathrm{d}\kappa.$$

We have the following explicit formula for the exterior differential of a smooth k-form κ over a differentiable manifold M:

$$\begin{aligned} \text{(A.2.8)} \quad \mathrm{d}\kappa(X_0, X_1, \ldots, X_k) &= \sum_{i=0}^{k} (-1)^i X_i\big(\kappa(X_0, \ldots, \widehat{X_i}, \ldots, X_k)\big) \\ &+ \sum_{i<j} (-1)^{i+j} \kappa\big([X_i, X_j], X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k\big), \end{aligned}$$

for any smooth vector fields X_0, X_1, …, X_k over M, where the hat indicates that the corresponding term was omitted from the sequence.

A smooth differential form is said to be *closed* if its exterior differential vanishes; it is said to be *exact* if it is equal to the exterior differential of another smooth differential form. By property 4 above, every smooth exact form is closed. *Poincaré Lemma* says that every smooth closed form is locally exact, i.e., if κ is a smooth closed form over M then every point of M has an open neighborhood U such that the restriction $\kappa|_U$ is exact (actually, $\kappa|_U$ is exact whenever U is diffeomorphic to a star-shaped open subset of $\mathbb{R}^n$ or, more generally, whenever U is contractible).

There is a very nice formula that expresses the Lie derivative of a differential form in terms of exterior derivatives and interior products. If κ is a smooth differential form over a differentiable manifold M and X is a smooth vector field over M then:

$$\mathbb{L}_X\kappa = \mathrm{d}i_X\kappa + i_X\mathrm{d}\kappa. \tag{A.2.9}$$

In Exercise A.5 (which uses Exercises A.3 and A.4) we ask the reader to prove a generalization of formula (A.2.9).

A.3. A little bit of Lie groups

A *Lie group* is a differentiable manifold G, endowed with a group structure, in such a way that both the multiplication map:

$$G \times G \ni (g, h) \longmapsto gh \in G$$

and the inversion map:

$$G \ni g \longmapsto g^{-1} \in G$$

are smooth (in fact, the smoothness of the multiplication map implies the smoothness of the inversion map, by the implicit function theorem). The neutral element of a group G will be denoted by 1. The tangent space T_1G at the neutral element will be denoted by $\mathfrak{g}$. The space $\mathfrak{g}$ can be endowed with a binary operation (the Lie bracket) and with such operation it becomes a Lie algebra and it is called the *Lie algebra* of the Lie group G (more details are given below). To each $g \in G$, we can associate smooth diffeomorphisms:

$$L_g : G \ni x \longmapsto gx \in G, \quad R_g : G \ni x \longmapsto xg \in G,$$

known respectively as the *left translation map* and the *right translation map*. A vector field X over G is said to be *left invariant* (resp., *right invariant*) if X is L_g-related (resp., R_g-related) to X, for all $g \in G$ (recall Definition A.2.2). A left invariant (resp., right invariant) vector field is automatically smooth and it is uniquely determined by its value at $1 \in G$ by the formula:

$$X(g) = \mathrm{d}L_g(1)X(1), \quad g \in G,$$

(resp., by the formula $X(g) = \mathrm{d}R_g(1)X(1)$, $g \in G$). We use the following notation: letters like X, Y denote elements of $\mathfrak{g} = T_1G$; given $X \in \mathfrak{g}$, we denote by X^L (resp., by X^R) the unique left invariant vector field (resp., the unique right invariant vector field) such that $X^L(1) = X$ (resp., such that $X^R(1) = X$). By Proposition A.2.3, the Lie bracket of left invariant (resp., of right invariant) vector fields is left invariant (resp., right invariant). We define the *Lie bracket* $[X, Y] \in \mathfrak{g}$ of elements $X, Y \in \mathfrak{g}$ by setting:

$$[X, Y] = [X^L, Y^L](1),$$

so that:

$$[X,Y]^L = [X^L, Y^L],$$

for all $X, Y \in \mathfrak{g}$. We have also:

$$[X,Y]^R = -[X^R, Y^R], \tag{A.3.1}$$

for all $X, Y \in \mathfrak{g}$ (this follows from the observation that X^L and $-X^R$ are related by the inversion map $g \mapsto g^{-1}$). Endowed with the Lie bracket, the vector space $\mathfrak{g}$ becomes a *Lie algebra*, i.e., a vector space endowed with an anti-symmetric bilinear binary operation $(X, Y) \mapsto [X, Y]$ satisfying the *Jacobi identity*:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0, \quad X, Y, Z \in \mathfrak{g}.$$

Given $X \in \mathfrak{g}$ then there exists a unique smooth group homomorphism:

$$\gamma_X : \mathbb{R} \longrightarrow G$$

such that $\gamma_X'(0) = X$ (notice that $\gamma_X(0) = 1$, so that $\gamma_X'(0) \in \mathfrak{g}$). The map γ_X is also the maximal integral curve of X^L (and also of X^R) passing through 1 at $t = 0$. The map:

$$\exp : \mathfrak{g} \longrightarrow G$$

defined by:

$$\exp(X) = \gamma_X(1), \quad X \in \mathfrak{g},$$

is smooth and it is called the *exponential map* of G. We have:

$$\gamma_X(t) = \exp(tX),$$

for all $t \in \mathbb{R}$, $X \in \mathfrak{g}$.

A.3.1. Actions of Lie groups on manifolds. Let G be a Lie group, M be a differentiable manifold and:

$$\rho : G \times M \longrightarrow M$$

be a smooth (left) *action* of G on M, i.e., the map ρ is smooth and the map:

$$G \ni g \longmapsto \rho_g \overset{\text{def}}{=} \rho(g, \cdot) \in \mathrm{Diff}(M)$$

is a homomorphism from G to the group $\mathrm{Diff}(M)$ of all smooth diffeomorphisms of M. We write:

$$\rho(g, x) = g \cdot x,$$

for all $g \in G$, $x \in M$. Given $x \in M$, we obtain from the action ρ a smooth map:

$$\beta_x : G \longrightarrow M$$

defined by $\beta_x(g) = g \cdot x$, for all $g \in G$. Given $X \in \mathfrak{g}$, we define a vector field X^M over the manifold M by setting:

$$X^M(x) = \mathrm{d}\beta_x(1)X \in T_xM, \quad x \in M. \tag{A.3.2}$$

The vector field X^M is smooth. We have:

$$X^M(x) = \frac{\mathrm{d}}{\mathrm{d}t}\big(\exp(tX)\cdot x\big)\Big|_{t=0}, \tag{A.3.3}$$

for all $x \in M$. Obviously, in (A.3.3) we can replace $\exp(tX)$ with $\gamma(t)$, where γ is any differentiable curve in G with $\gamma(0) = 1$ and $\gamma'(0) = X$. It is easily checked that X^M is the only vector field on M that is β_x-related to the right invariant vector field X^R, for all $x \in M$. It follows from (A.3.1) and from Proposition A.2.3 that:

$$[X^M, Y^M] = -[X, Y]^M, \tag{A.3.4}$$

for all $X, Y \in \mathfrak{g}$. In other words, the map $X \mapsto X^M$ is an *anti-homomorphism* from the Lie algebra $\mathfrak{g}$ to the Lie algebra of smooth vector fields over M, endowed with the Lie bracket. For any $x \in M$, the curve:

$$\mathbb{R} \ni t \longmapsto \exp(tX)\cdot x \in M$$

is the maximal integral curve of X^M passing through x at $t = 0$; thus, if F denotes the flow of X^M, then:

$$F_t = \rho_{\exp(tX)},$$

for all $t \in \mathbb{R}$.

A.3.1. Remark. The group $\mathrm{Diff}(M)$ of smooth diffeomorphisms of a differentiable manifold M isn't a (finite-dimensional) Lie group, but it can be endowed with the structure of an infinite-dimensional Lie group (it is a Fréchet Lie group if M is compact and, in general, it is a Lie group modeled on a topological vector space which is an inductive limit of Fréchet spaces). Unfortunately, given a smooth action $\rho : G \times M \to M$, then the group homomorphism:

$$G \ni g \longmapsto \rho_g \in \mathrm{Diff}(M) \tag{A.3.5}$$

is not smooth, unless M is compact. Let us then assume that M is compact. The differential at the neutral element of a smooth homomorphism between Lie groups is a homomorphism between their Lie algebras. The Lie algebra homomorphism obtained by differentiating (A.3.5) at $1 \in G$ is precisely the map $\mathfrak{g} \ni X \mapsto X^M$. But we have seen above that such map is not a Lie algebra homomorphism, but an *anti*-homomorphism; what is going on here? It happens that the Lie algebra of the infinite-dimensional Lie group $\mathrm{Diff}(M)$ is identified with the space of smooth vector fields over M (vector fields of compact support, if M is not compact) endowed with the *negative* of the standard Lie bracket of vector fields!

Exercises

Quick review of calculus on manifolds.

EXERCISE A.1. Let M, N be differentiable manifolds and

$$F : \mathrm{dom}(F) \subset \mathbb{R} \times M \to N, \quad \widetilde{F} : \mathrm{dom}(\widetilde{F}) \subset \mathbb{R} \times M \to N$$

be smooth maps defined over open subsets of $\mathbb{R} \times M$. For each $t \in \mathbb{R}$, we define maps:

$$F_t : \mathrm{dom}(F_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(F)\big\} \ni x \longmapsto F(t, x) \in N,$$
$$\widetilde{F}_t : \mathrm{dom}(\widetilde{F}_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(\widetilde{F})\big\} \ni x \longmapsto \widetilde{F}(t, x) \in N.$$

Assume that for a certain $t_0 \in \mathbb{R}$ the maps F_{t_0} and $\widetilde{F}_{t_0}$ are equal and that:

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t, x)\Big|_{t=t_0} = \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}(t, x)\Big|_{t=t_0},$$

for all $x \in \mathrm{dom}(F_{t_0}) = \mathrm{dom}(\widetilde{F}_{t_0})$. Let τ be a smooth (r, s)-tensor field over N; if $s \neq 0$, assume that F_t and $\widetilde{F}_t$ are local diffeomorphisms, for all $t \in \mathbb{R}$. Show that:

$$\frac{\mathrm{d}}{\mathrm{d}t}F_t^*\tau\Big|_{t=t_0} = \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}_t^*\tau\Big|_{t=t_0}.$$

(*hint*: there is no loss of generality in assuming that both M, N are open subsets of Euclidean space. Notice that we can write $(F_t^*\tau)(x)$ in the form $\alpha\big(F_t(x), \mathrm{d}F_t(x)\big)$, for some smooth map α and that the derivative of:

$$t \longmapsto \alpha\big(F_t(x), \mathrm{d}F_t(x)\big)$$

at $t = t_0$ depends only on the value and derivative at $t = t_0$ of the map $t \mapsto F_t(x)$).

EXERCISE A.2. Let M, N be differentiable manifolds and:

$$F : \mathrm{dom}(F) \subset \mathbb{R} \times M \longrightarrow N$$

be a smooth map defined over some open subset $\mathrm{dom}(F)$ of $\mathbb{R} \times M$. For $t \in \mathbb{R}$, denote by F_t the smooth map $F(t, \cdot)$ defined over the open set $\mathrm{dom}(F_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(F)\big\}$. Let κ be a smooth *time-dependent k-form* over M, i.e., κ is a smooth map defined over some open subset $\mathrm{dom}(\kappa)$ of $\mathbb{R} \times M$, associating an element $\kappa(t, x)$ of $\bigwedge_k T_xM^*$ to each (t, x) in $\mathrm{dom}(\kappa)$. For each $t \in \mathbb{R}$, we have a smooth k-form $\kappa_t = \kappa(t, \cdot)$ over the open subset $\mathrm{dom}(\kappa_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(\kappa)\big\}$ of M. Given $t_0 \in \mathbb{R}$, show that the following equality holds over the open set $F_{t_0}^{-1}\big(\mathrm{dom}(\kappa_{t_0})\big)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}(F_t^*\kappa_t)\Big|_{t=t_0} = \frac{\mathrm{d}}{\mathrm{d}t}(F_t^*\kappa_{t_0})\Big|_{t=t_0} + F_{t_0}^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t\Big|_{t=t_0}\Big).$$

(*hint*: for a fixed x, define $\phi(t, s) = (F_t^*\kappa_s)(x)$, and then write the derivative of $t \mapsto \phi(t, t)$ in terms of the partial derivatives of ϕ).

Exercise A.3. Let M be a differentiable manifold and κ be a smooth time-dependent k-form over M. For $t \in \mathbb{R}$, set $\kappa_t = \kappa(t, \cdot)$ and denote by $\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t$ the smooth k-form over $\mathrm{dom}(\kappa_t)$ defined by:

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t\Big)(x) = \frac{\mathrm{d}}{\mathrm{d}t}\big(\kappa_t(x)\big), \quad x \in \mathrm{dom}(\kappa_t).$$

Show that the operation $\frac{\mathrm{d}}{\mathrm{d}t}$ commutes with the exterior derivative, i.e.:

$$\mathrm{d}\Big(\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{d}\kappa_t,$$

for all $t \in \mathbb{R}$ (*hint*: write down the representation of κ with respect to a local chart and observe that the exterior derivative and the operation $\frac{\mathrm{d}}{\mathrm{d}t}$ involve partial derivatives with respect to *distinct* variables).

Exercise A.4. Let M, N be differentiable manifolds and:

$$F : \mathrm{dom}(F) \subset \mathbb{R} \times M \longrightarrow N$$

be a smooth map defined over some open subset $\mathrm{dom}(F)$ of $\mathbb{R} \times M$. For $t \in \mathbb{R}$, denote by F_t the smooth map $F(t, \cdot)$ defined over the open set $\mathrm{dom}(F_t) = \big\{x \in M : (t, x) \in \mathrm{dom}(F)\big\}$. Given a k-form κ over N then the *generalized interior product* $i(F, \kappa)$ is the time-dependent $(k-1)$-form over M:

$$i(F, \kappa) : \mathrm{dom}(F) \ni (t, x) \longmapsto i(F, \kappa)(t, x) \in \bigwedge_{k-1} T_xM^*$$

defined by:

$$i(F, \kappa)(t, x) = \mathrm{d}F_t(x)^*\big[i_{v(t,x)}\kappa\big(F_t(x)\big)\big],$$

for all $(t, x) \in \mathrm{dom}(F)$, where $v(t, x) = \frac{\mathrm{d}}{\mathrm{d}t}F_t(x) \in T_{F_t(x)}N$. For $t \in \mathbb{R}$, denote by $i(F, \kappa; t)$ the k-form over $\mathrm{dom}(F_t)$ defined by $i(F, \kappa; t)(x) = i(F, \kappa)(t, x)$. Show that if κ is a k-form over N and λ is an l-form over N then:

$$i(F, \kappa \wedge \lambda; t) = i(F, \kappa; t) \wedge (F_t^*\lambda) + (-1)^k(F_t^*\kappa) \wedge i(F, \lambda; t), \tag{A.3.6}$$

for all $(t, x) \in \mathrm{dom}(F)$ (*hint*: use (A.1.8)).

Exercise A.5. Let M, N and F be as in the statement of Exercise A.4. The goal of this exercise is to show that for a smooth k-form κ over N the following formula holds:

$$\frac{\mathrm{d}}{\mathrm{d}t}F_t^*\kappa = \mathrm{d}\big(i(F, \kappa; t)\big) + i(F, \mathrm{d}\kappa; t). \tag{A.3.7}$$

Notice that if $N = M$ and F is the flow of a smooth vector field X over M then formula (A.3.7) (with $t = 0$) reduces to (A.2.9). Let $t \in \mathbb{R}$ be fixed and for a smooth k-form κ over N consider the smooth k-forms over $\mathrm{dom}(F_t) \subset M$ defined by:

$$D_1(\kappa) = \frac{\mathrm{d}}{\mathrm{d}t}F_t^*\kappa, \quad D_2(\kappa) = \mathrm{d}\big(i(F, \kappa; t)\big) + i(F, \mathrm{d}\kappa; t).$$

We have to show that $D_1(\kappa) = D_2(\kappa)$, for any smooth differential form κ over N.

(a) Check that the map $\kappa \mapsto D_i(\kappa)$ is linear (over the field of real numbers), for $i = 1, 2$.

(b) Check that, if κ and λ are smooth differential forms over N then:
$$D_i(\kappa \wedge \lambda) = D_i(\kappa) \wedge (F_t^* \lambda) + (F_t^* \kappa) \wedge D_i(\lambda),$$
for $i = 1, 2$ (*hint*: for $i = 2$ use formula (A.3.6)).

(c) Check that both D_1 and D_2 commute with the exterior derivative, i.e., that $D_i(\mathrm{d}\kappa) = \mathrm{d}\big(D_i(\kappa)\big)$, $i = 1, 2$ (*hint*: for $i = 1$ use the result of Exercise A.3).

(d) Use the results of the items above to conclude that the set of smooth differential forms κ over N for which $D_1(\kappa) = D_2(\kappa)$ is a real vector subspace closed under exterior products and exterior derivatives.

(e) Check that $D_1(\kappa) = D_2(\kappa)$ when κ is a smooth zero-form (i.e., a smooth real valued function) over N.

(f) When N admits a global chart then any smooth differential form over N can be obtained from smooth zero-forms using exterior derivatives, exterior products and sums. Explain why it suffices to prove (A.3.7) in the case when N admits a global chart. Conclude the proof of (A.3.7).