Proyecto Fin de Máster en Investigación Matemática
Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid

Hamiltonian formulation of the Yang-Mills equations

Alumna: Marta Farré Puiggalí
Director: Marco Castrillón López

Curso 2011-2012

Autorización de difusión

La abajo firmante, matriculada en el Máster en Investigación Matemática de la Facultad de Ciencias Matemáticas, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: "Hamiltonian formulation of the Yang-Mills equations", realizado durante el curso académico 2011-2012 bajo la dirección de Marco Castrillón López en el Departamento de Geometría y Topología, y a la biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

Marta Farré Puiggalí

Abstract: Given a principal fiber bundle P over a semi-Riemannian manifold (M, g), Yang-Mills equations are defined over the space of principal connections on P. They are the Euler-Lagrange equations corresponding to a certain Lagrangian defined on J^1C, the first jet bundle of C, where C is the bundle of connections on P. We write the Hamiltonian counterpart of the variational problem defined by the Yang-Mills Lagrangian on the polysymplectic bundle Π and recover Yang-Mills equations from Hamilton-Cartan equations on Π.

Key words: Connections, Euler-Lagrange equations, Hamilton equations, jet bundle, polysymplectic bundle, Variational Calculus, Yang-Mills.

Resumen: Dado un fibrado principal P sobre una variedad semi-Riemanniana (M, g), las ecuaciones de Yang-Mills se definen en el espacio de las conexiones principales en P. Son las ecuaciones de Euler-Lagrange correspondientes a una cierta Lagrangiana definida en J^1C, el fibrado de jets de primer orden de C, donde C es el fibrado de las conexiones en P. Escribimos la versión Hamiltoniana del problema variacional definido por la Lagrangiana de Yang-Mills en el fibrado polisimpléctico Π y recuperamos las ecuaciones de Yang-Mills a partir de las ecuaciones de Hamilton-Cartan en Π.

Palabras clave: Cálculo Variacional, conexiones, ecuaciones de Euler-Lagrange, ecuaciones de Hamilton, fibrado de jets, fibrado polisimpléctico, Yang-Mills.

Código MSC2010: Primary: 70S05 Secondary: 53C05 , 58A20, 70S15

Contents

1 Introduction

The Calculus of Variations has a long and rich history since its formal beginning with the works of Euler, Lagrange and Hamilton. The goal is to find extreme values of certain functionals. The main difficulty that one encounters when dealing with the functionals involved in this branch of Mathematics is the complex nature of the spaces where they are defined. More precisely, these functionals are generally integro-differential operators on infinite dimensional spaces as, for example, spaces of functions or spaces of sections of bundles.

Many of the interesting problems in the Calculus of Variations come from Physics and describe many of the fundamental equations in Mechanics, Electromagnetism, Relativity, etc. The objects of all these equations are sections of bundles. More precisely, if M is a model of space-time, the electromagnetic potential is a section of the bundle connections. This is because electromagnetic potential can be identified with connections of an appropriate principal bundle on M. This instance can be further generalized to the so-called Yang-Mills equations. Since their introduction in the 50's, these equations have proved to be the convenient framework to model interactions between quantum particles (interactions including the aforementioned electromagnetism as well as weak or strong nuclear interactions). In fact, it is interesting to note that the understanding of these forces as described through connections was a late result form the 70's. In any case, the Yang-Mills equations can be obtained as variational equations on the set $\mathcal{A}$ of connections of certain principal bundles. Moreover, the integro-differential operator is given as

$$\begin{array}{ccc} \mathcal{A} & \longrightarrow & \mathbb{R} \\ A & \mapsto & \displaystyle\int_M L(j^1 A)\mathbf{v} \end{array}$$

where $\mathbf{v}$ is a volume form in M and L is a function depending on the first-order Taylor expansion $j^1 A$ of A. One can give a more geometrical definition to L through the language of jets. In this case, the function L (called the Lagrangian) is thus defined as

$$L : J^1 C \longrightarrow \mathbb{R}$$

where $C \to M$ is the bundle of connections and $J^1 C$ stands for the jet space.

The equations for extremal values of the operators in Variational Calculus defined by functions of the type above (that is, operators given by integration of functions depending on the first jet) have a well known structure. They are the so-called Euler-Lagrange equations and are ubiquitous in the literature. These equations, in particular, give a standard way to formulate the Mechanics when the bundle is $\mathbb{R} \times Q \to \mathbb{R}$, the sections of which are just curves in the configuration space Q. The equations have a similar structure in arbitrary bundles. As we learn from Mechanics, the Hamiltonian picture of these equations is essential in many instances. This framework deals with positions and momenta and the evolution equations of these variables are the Hamilton equations. Again, in fiber bundles, the analogue idea works. In this case, the "position" variables are encoded through a bundle $Y \to M$ and the momenta are in a composite bundle $\Pi \to Y \to M$. The new Hamilton equations are defined on sections of this composite bundle. The importance of these equations is connected with notions as symplectic (or multisymplectic) forms, quantization, etc.

The main goal of this work consists of giving a precise formulation of the Hamilton equations for the Yang-Mills Lagrangian by giving the correct exposition of the geometric objects involved in the construction.

The structure of the work is as follows:

In Section 2 we introduce the basic tools about bundles and connections, including jet bundles and bundles of connections.

In Section 3 we first introduce briefly the Lagrangian formalism for an arbitrary configuration bundle and then define the polysymplectic bundle Π and Hamiltonian systems on it. Finally we consider the case when the configuration bundle is a principal bundle P and take the quotient of Π by the action of the structure group on TP. We introduce a bracket on the quotient which coincides with the Poisson bracket when evaluated on G-invariant Poisson forms and finally include the reduced Hamilton-Cartan equations on the quotient.

In Section 4 we start introducing Yang-Mills equations on the space of connections of a principal bundle P, consider the action of the gauge group of P on C and prolong it to J^1C, drop the Yang-Mills Lagrangian to the quotient space of J^1C by this action and recover Yang-Mills equations from the dropped variational problem. Then we go on to specify the Hamiltonian approach defined on Section 3 taking the Yang-Mills Lagrangian, that is, we write the linear Legendre transformation, the Yang-Mills Hamiltonian and Hamilton-Cartan equations working with local coordinates on Π, more precisely on the image of the linear Legendre transformation $\mathcal{P}$. We check that Hamilton-Cartan equations on $\mathcal{P}$ coincide with Yang-Mills equations, that is, solutions of the Hamilton-Cartan equations come from solutions of the variational problem defined by the Yang-Mills Lagrangian, taking the prolongation to the first jet bundle and composing with the Legendre transformation. Finally we compute the expression of the Poisson $(n-1)$-forms and Poisson bracket on the constraint manifold $\mathcal{P}$ and give the characterization of solutions of a Hamiltonian system on $\mathcal{P}$ in terms of Poisson $(n-1)$-forms on $\mathcal{P}$ analogous to the result in Section 3. In general Section 4 is an attempt to generalize the Electromagnetism example in [1] to the Yang-Mills case (where we change S^1 for an arbitrary Lie group G). The last point, reduction of the equations by the action of the gauge group on Π cannot be done in an analogous way to [1] since in contrast to the Electromagnetism Hamiltonian, the Yang-Mills Hamiltonian is not invariant under this action. This is left for future work.

2 Bundles

2.1 Introduction

Some of the main ingredients we will be dealing with are fiber bundles, particularly vector, affine and principal bundles, and also connections on the principal fiber bundles. Here we give a quick introduction to the concepts. For further details see [6].

Although it is not always required, we will always assume that the spaces are differentiable manifolds and all functions are smooth (we might for instance reformulate the same concepts considering topological spaces and continuous functions).

The idea of a fiber bundle is that of a manifold that can be viewed as the disjoint union of copies of a manifold F indexed by another manifold M. These copies will be the fibers of a smooth function. We will require that the dimensions of both manifolds are greater than 0, for otherwise this definition would be trivial. Locally we will be able to express this disjoint union as a product manifold of the form $U \times F$, where U is an open subset of M. If we do not require that the fibers are diffeomorphic, then we have the weaker notion of fibration. Let us give some formal definitions:

Definition 1 (Fibration)**.** *Let E and M be differentiable manifolds. A smooth map $\pi : E \longrightarrow M$ is called a fibration if it is a surjective submersion. E, M, π and the triple (E, π, M) will be called total space, base space, projection and fibred manifold respectively.*

We will sometimes abbreviate (E, π, M) by E or π depending on whether we want to distinguish between different fibred manifolds with the same total space. Usually we will write $\pi : E \longrightarrow M$ or just $E \longrightarrow M$.

Remark 1. Using the local structure of submersions we have that for each $p \in E$ there is an open subset $U_p \subset E$, $p \in U_p$, and a diffeomorphism $\phi : \pi(U_p) \times F_p \longrightarrow U_p$ such that $\pi|_{U_p} \circ \phi = pr_1$, where F_p is a differentiable manifold of dimension $dim(E) - dim(M)$ and pr_1 denotes the projection onto the first factor of the product manifold. Note that we do not require that π be injective, so one can see the fiber $\pi^{-1}(x)$ as a disjoint union of manifolds of the same dimension. Imagine for instance a spiral made with a ribbon projecting over a circle.

Note that with the above definition the fibers need not be diffeomorphic. Consider for instance the fibred manifold $(\mathbb{R} \times \mathbb{R} \backslash \{0\}, pr_1, \mathbb{R})$, where the fiber $\pi^{-1}(0)$ is different from the rest.

In order to avoid this inconvenience we will deal with a particular case of fibred manifolds called fiber bundles. Dealing with fiber bundles will be easier than dealing with fibred manifolds and they will provide a suitable model for a variety of physical situations. One can for instance think of a metal sheet and all possible temperatures at each point as a fiber bundle since all possible values at each point are the same. As an example in which the fiber bundle would not be trivial (we will see what this means) one can think of the Earth as a sphere and consider all possible velocities that the wind can have at each point, that is, the tangent space at each point. Note that despite the fact that the sets of velocities are different at each point, they are all diffeomorphic, which is what will matter. Note also that in this example the fibers are more than just a smooth manifold and more than just diffeomorphic to each other. This example is known as the tangent bundle and we will come back to it later.

So we give the following definitions:

Definition 2 (Local trivialisation)**.** *Let (E, π, M) be a fibred manifold. Given $p \in M$, an open neighborhood $U_p \subset M$ of p, and a differentiable manifold F_p, a diffeomorphism*

$$\varphi_{U_p} : \pi^{-1}(U_p) \xrightarrow{\cong} U_p \times F_p$$

such that $\pi|_{\pi^{-1}(U_p)} = pr_1 \circ \varphi_{U_p}$ is called a local trivialisation around p, or just a local trivialisation if we do not specify any particular point and just take an open subset $U \subset M$. Note that this condition implies that $\pi^{-1}(y) \cong F_p$ for all $y \in U_p$.

Definition 3 (Fiber bundle)**.** *If it is possible to define local trivialisations around any point $p \in M$ then (E, π, M) will be called a fiber budle.*

Remark 2. In a fiber bundle we will necessarily have $F_p \cong F_q =: F$ for all $p, q \in M$ in case M is connected. Otherwise we will have the same conclusion in each of the connected components of M. In order to justify this, take a local trivialisation $\pi^{-1}(U_p) \cong U_p \times F_p$ around $p \in M$. Then consider all local trivialisations with fibers diffeomorphic to F_p and take the union of all corresponding open subsets of M, obtaining an open subset of M with fibers diffeomorphic to F_p. Since the union of all open subsets of M corresponding to local trivialisations with fibers not diffeomorphic to F_p should also be an open subset, and since we are assuming that M is connected, we get that the first union must be equal to the whole base space M, for it is not empty.

Then we will refer to F as its typical fiber and denote a fiber bundle by (E, π, M, F), just by (E, π, M) if we do not need to mention the fiber or by some of the abbreviations given for a fibred manifold.

If we have two local trivialisations φ_U and φ_V with overlapping domains, then the change of trivialisation will be of the form

$$\begin{array}{rccc} \varphi_V \circ \varphi_U^{-1}: & U \cap V \times F & \longrightarrow & U \cap V \times F \\ & (x, y) & \longmapsto & (x, \tilde{\varphi}(x, y)), \end{array}$$

where for each $x \in U \cap V$, $\tilde{\varphi}(x, \cdot) : F \longrightarrow F$ is a diffeomorphism. We will call $\tilde{\varphi}(x, \cdot)$ a transition function and denote it by g_{UV}.

An atlas on the total space E can be constructed from an atlas on the base space M and an atlas on the typical fiber F using the local trivialisations, for they give diffeomorphisms $\varphi_U : \pi^{-1}(U) \longrightarrow U \times F$ satisfying $pr_1 \circ \varphi_U = \pi|_{\pi^{-1}(U)}$, where U can be taken to be a chart on M, and then $U \times F$ can be given a product atlas. For these charts the first $\dim(M)$ coordinates of points in the same fiber will be equal. We will write coordinates as (x^i, y^a), where x^i denote the coordinates in the base space and y^a the coordinates in the fiber, and we will call them adapted coordinates or fiber coordinates.

Definition 4. *In case there is a global trivialisation, that is, a diffeomorphism*

$$\phi : E \xrightarrow{\cong} M \times F$$

with $\pi = pr_1 \circ \phi$, then E is called a trivial bundle.

Remark 3. One can define a fiber bundle starting with a tuple (E, π, M, F) satisfying the property of having local trivialisations around any $p \in M$ and then obtain that π must be a submersion, for it is locally a projection.

We will usually assume an additional structure on F and also on the transition functions. For instance one might require that F be a vector space and g_{UV} be linear functions.

As in the following example, fiber bundles are often defined as projections from a set to a manifold. To ensure that the projection gives a fiber bundle we need the following theorem:

Theorem 1. *Let M and F be differentiable manifolds of dimensions n and m respectively, E a set and $\pi : E \longrightarrow M$ a map such that the fibers $\pi^{-1}(x)$, $x \in M$, have the structure of a differentiable manifold of dimension m. Assume also that for each $x \in M$ there is an open neighborhood $x \in W_x$ and a bijection*

$$\Phi_x : \pi^{-1}(W_x) \longrightarrow W_x \times F$$

such that $pr_1 \circ \Phi_x = \pi|_{\pi^{-1}(W_x)}$ and $pr_2 \circ \Phi_x|_{\pi^{-1}(y)} : \pi^{-1}(y) \longrightarrow F$ is a diffeomorphism for all $y \in W_x$.

Then E admits a unique differentiable structure such that $\pi : E \longrightarrow M$ becomes a fiber bundle and the maps Φ_x are local trivialisations.

A proof of this theorem can be found in [9].

Note that a product manifold $M \times F$ gets its differentiable structure from the corresponding differentiable structures on M and F. Here we are getting a differentiable structure on each $\pi^{-1}(W_x)$ by writing it as a product manifold.

Example 1 (Tangent and cotangent bundles)**.** Let M be a smooth manifold with $\dim(M) = n$. Consider the disjoint union

$$TM = \bigcup_{x \in M} T_x M$$

and let π be the projection

$$\begin{array}{rccc} \pi: & TM & \longrightarrow & M \\ & v \in T_x M & \longmapsto & x. \end{array}$$

We can define a fiber atlas on TM from an atlas on M. Let (U, x^i) be a chart on M. There is a bijection between $\pi^{-1}(U)$ and $U \times \mathbb{R}^n$ given by $v \longmapsto (x^i, v_i)$, where $v = v_i \left.\frac{\partial}{\partial x^i}\right|_x \in \pi^{-1}(U)$. Note that the other conditions of theorem 1 are satisfied, for $\pi^{-1}(x)$ is a vector space diffeomorphic to $\mathbb{R}^n$. (TM, τ_M, M) will be called the tangent bundle.

We can also define a fiber bundle structure on the disjoint union

$$T^*M = \bigcup_{x \in M} T_x^* M$$

in a similar fashion. The resulting bundle (T^*M, τ_M^*, M) will be called the cotangent bundle.

Note that $\dim(TM) = \dim(T^*M) = 2n$.

Definition 5 (Sections)**.** *Let (E, π, M) be a fiber bundle.*

- *A smooth function $s : M \longrightarrow E$ is called a global section of π if $\pi \circ s = id_M$.*
- *A smooth function $s : U \longrightarrow E$, where $U \subset M$ is an open subset, is called a local section of π if $\pi \circ s = id_U$.*

The notation used for the set of all sections of a fiber bundle $E \longrightarrow M$ will be $\Gamma(E)$. We will denote the set of local sections defined in some neighborhood of a point $x \in M$ by $\Gamma_x(E)$ and use this set in the definition of jet bundles later. When dealing with jet bundles we will also use the notation $\Gamma_W(E)$, where $W \subset M$ is an open subset, for all local sections defined on W.

Example 2. Sections of the tangent and cotangent bundle are vector fields and differential forms respectively.

Definition 6 (Bundle morphism)**.** *Let (E, π, M) and (E', π', M') be two fiber bundles. A bundle morphism is a pair of smooth maps (F, f), $F : E \longrightarrow E'$, $f : M \longrightarrow M'$, such that the following diagram commutes:*

$$\begin{tikzcd} E \arrow[r, "F"] \arrow[d, "\pi"'] & E' \arrow[d, "\pi'"] \\ M \arrow[r, "f"] & M', \end{tikzcd}$$

that is, it is a morphism that sends fibers of E to fibers of E'.

Note that F determines f (once we know its existence), namely $f(x) = (\pi' \circ F)|_{\pi^{-1}(x)}$.

There are several ways to construct new fiber bundles from given fiber bundles. We define one of these which will be used later:

Definition 7 (Pull-back bundle)**.** *Let (E, π, M) be a fiber bundle and $h : N \longrightarrow M$ a smooth map. The pull-back bundle $(h^*E, h^*\pi, N)$ is the fiber bundle with total space*

$$h^*E = \{(x, p) \in N \times E : h(x) = \pi(p)\}$$

and projection

$$\begin{array}{rccc} h^*\pi : & h^*E & \longrightarrow & N \\ & (x,p) & \longmapsto & x. \end{array}$$

Remark 4. The pull-back bundle has the same typical fiber as the original bundle. In fact the fiber of $h^*\pi$ over $x \in N$ is diffeomorphic to the fiber of π over $h(x)$.

2.2 Vector bundles and affine bundles

Vector bundles will appear constantly. We will also work with first order jet bundles and bundles of connections, which are affine bundles, so we give the definitions here. First let us introduce the notation E_x for the fiber of π over $x \in M$, that is $E_x := \pi^{-1}(x)$.

Definition 8 (Vector bundle)**.** *A vector bundle is a fiber bundle (E, π, M, F) where the typical fiber F is a vector space and the transition functions are linear isomorphisms.*

Definition 9 (Affine bundle)**.** *An affine bundle is a fiber bundle (E, π, M, F) where the typical fiber F is an affine space and the transition functions are affine isomorphisms.*

We will say that an affine bundle (E, π, M, F) is modelled on a vector bundle (E', π', M, V) if the affine space F is modelled on the vector space V.

Remark 5. Note that if (E, π, M, F) is a vector (affine) bundle then E_x is a vector (affine) space. Consider a local trivialisation around $x \in M$

$$\varphi_{U_x} : \pi^{-1}(U_x) \xrightarrow{\cong} U_x \times F$$

and define a sum and a scalar multiplication on E_x as

$$y + y' := \varphi_{U_x}^{-1}(x, pr_2(\varphi_{U_x}(y)) + pr_2(\varphi_{U_x}(y'))), \text{ for all } y, y' \in E_x,$$

$$\lambda y := \varphi_{U_x}^{-1}(x, \lambda pr_2(\varphi_{U_x}(y))), \text{ for all } \lambda \in \mathbb{R}, y \in E_x.$$

Note that because of the conditions imposed on the transition functions, that is, linearity, this definition does not depend on the local trivialisation. Let $\varphi_{V_x} : \pi^{-1}(V_x) \xrightarrow{\cong} V_x \times F$ be another local trivialisation around x. Then

$$\varphi_{V_x} \circ \varphi_{U_x}^{-1}(x, pr_2\varphi_{U_x}(y) + pr_2\varphi_{U_x}(y')) = (x, g_{U_xV_x}(pr_2\varphi_{U_x}(y) + pr_2\varphi_{U_x}(y')))$$
$$= (x, g_{U_xV_x}(pr_2\varphi_{U_x}(y))) + (x, g_{U_xV_x}(pr_2\varphi_{U_x}(y'))) = \varphi_{V_x}(y) + \varphi_{V_x}(y')$$
$$= (x, pr_2\varphi_{V_x}y + pr_2\varphi_{V_x}y'),$$

so $\varphi_{V_x}^{-1}(\varphi_{V_x}(y) + \varphi_{V_x}(y')) = y + y'$, using bijectivity. Similarly one deals with scalar multiplication and with the affine case.

Definition 10 (Dual bundle). *Given a vector bundle (E, π, M, F), the dual vector bundle $(E^*, \hat{\pi}, M, F^*)$ is defined to be the vector bundle with total space*

$$E^* = \bigcup_{x \in M} E_x^*$$

and projection

$$\begin{array}{rccc} \hat{\pi}: & E^* & \longrightarrow & M \\ & f_x & \longmapsto & x, \end{array}$$

where $f_x \in E_x^*$. If we have a local trivialisation on E given by (x^i, y^a) ((x^i) are coordinates on an open subset $U \subset M$ and (x^i, y^a) are coordinates on $\pi^{-1}(U)$), then we can define a bijection $f_x \longmapsto (x^i, y_a)$ between $\hat{\pi}^{-1}(U)$ and $U \times F^*$ as $x^i(f_x) = x^i(\hat{\pi}(x))$ and $y_a(f_x) = f_x(B_a)$, where $\{B_1, \ldots, B_m\}$ is the basis of E_x corresponding to the coordinates $(y^1, \ldots, y^m)$, that is, $B_a = (0, \ldots, 1, \ldots, 0)$, where the 1 is in position a and the rest are zeroes.

Definition 11 (Tensor bundle). *Given two vector bundles (E, π, M, F), (E', π', M, F') over the same base space M, the tensor bundle $(E \otimes E', \tilde{\pi}, M, F \otimes F')$ is defined to be the vector bundle with total space*

$$E \otimes E' = \bigcup_{x \in M} E_x \otimes E'_x$$

and projection

$$\begin{array}{rccc} \tilde{\pi}: & E \otimes E' & \longrightarrow & M \\ & v_x \otimes v'_x & \longmapsto & x, \end{array}$$

where $v_x \in E_x$ and $v'_x \in E'_x$. Given local trivialisations of E and E', $\pi^{-1}(U) \cong U \times F$ and $(\pi')^{-1}(U) \cong U \times F'$ with coordinates (x^i, y^a) and (x^i, z^b) respectively, we define a local trivialisation $\tilde{\pi}^{-1}(U) \cong U \times (F \otimes F')$ of $E \otimes E'$ with coordinates (x^i, t^{ab}), where $t^{ab}(v_x \otimes v'_x) = y^a(v_x) z^b(v'_x)$.

Remark 6. Analogously one can define the tensor bundle of an arbitrary finite number of vector bundles. An interesting example arises when considering tensor products of a vector bundle and its dual: $E \otimes \overset{r}{\ldots} \otimes E \otimes E^* \otimes \overset{s}{\ldots} \otimes E^*$. In particular, if we take $E = TM$ then we get a fiber bundle whose sections are s-covariant and r-contravariant tensor fields. For example we can consider the vector bundle $T^*M \otimes TM$, which has fibers isomorphic to $End(T_xM)$ and whose sections are 1-forms taking values in TM. We will later consider 1-forms taking values in a Lie algebra $\mathfrak{g}$ which will be sections of the fiber bundle $T^*M \otimes \mathfrak{g}$.

2.3 Principal fiber bundles and connections

A principal fiber bundle is a fiber bundle in which the fibers are diffeomorphic to a Lie group G and the transition functions are given by a product on the Lie group. More precisely:

Definition 12 (Principal fiber bundle). *Let (P, π, M, G) be a fiber bundle with typical fiber a Lie group G. It will be called a principal fiber bundle if we can find local trivialisations*

$$\phi_\alpha : \pi^{-1}(U_\alpha) \longrightarrow U_\alpha \times G$$

satisfying

$$\begin{array}{rccc} \phi_\alpha \circ \phi_\beta^{-1}: & U_{\alpha\beta} \times G & \longrightarrow & U_{\alpha\beta} \times G \\ & (x,h) & \longmapsto & (x, g_{\alpha\beta}(x) \cdot h) \end{array}$$

for some smooth map $g_{\alpha\beta} : U_{\alpha\beta} \longrightarrow G$*, where* $U_\alpha, U_\beta \subset M$ *are open subsets and* $U_{\alpha\beta} = U_\alpha \cap U_\beta$.

Alternatively,

Definition 13 (Principal fiber bundle)**.** *Consider a tuple* (P, π, M, G)*, where* P *and* M *are differentiable manifolds,* G *is a Lie group and* $\pi : P \longrightarrow M$ *is a smooth map. We will say that* (P, π, M, G) *is a principal fiber bundle if the following properties are satified:*

- G *acts freely on* P *on the right, with the action denoted by*

$$\begin{array}{ccc} P \times G & \longrightarrow & P \\ (p,g) & \longmapsto & p \cdot g, \end{array}$$

- M *is the quotient manifold* P/G,
- π *is the canonical projection,*
- P *is locally trivial, meaning in this case that for each* $x \in M$ *there exists an open neighborhood* $x \in U_x$ *and a smooth map* $\varphi : \pi^{-1}(U_x) \longrightarrow G$ *such that*

$$\begin{array}{ccc} \pi^{-1}(U_x) & \stackrel{\cong}{\longrightarrow} & U_x \times G \\ p & \longmapsto & (\pi(p), \varphi(p)) \end{array}$$

  is a diffeomorphism and $\varphi(p \cdot g) = \varphi(p) \cdot g$ *for all* $g \in G$.

G *will be referred to as the structure group.*

Remark 7. Note that, since the action is free, the fibers will be diffeomorphic to G.

Proposition 1. *A principal fiber bundle* (P, π, M, G) *has a global section if and only if it is trivial.*

Proof. If the fiber bundle P is trivial, that is, we have a diffeomorphism $\Phi : P \stackrel{\cong}{\longrightarrow} M \times G$ satisfying $\pi = pr_1 \circ \Phi$, then we can define the global section

$$\begin{array}{rccccc} s: & M & \longrightarrow & M \times G & \longrightarrow & P \\ & x & \longmapsto & (x,e) & \longmapsto & \Phi^{-1}(x,e). \end{array}$$

Conversely, if there is a global section $s : M \longrightarrow P$, then a global trivialisation is defined by

$$\begin{array}{ccc} M \times G & \longrightarrow & P \\ (x,g) & \longmapsto & s(x) \cdot g. \end{array}$$

Let $p \in P$. Since $s(\pi(p))$ and p are in the same fiber, and the orbits of G coincide with the fibers in P, there exists necessarily an element $g \in G$ such that $p = s(\pi(p)) \cdot g$. Hence the map is surjective.

Now we see that it is injective: suppose that $s(x) \cdot g = s(x') \cdot g'$. Then $s(x)$ and $s(x')$ are in the same fiber and necessarily $x = x'$. Since the action is free, also $g = g'$.

$\square$

Remark 8. We can formulate the local version of the previous proposition. If (P, π, M, G) is a principal fiber bundle, then we have that for an open subset $U \subset M$ there is a local trivialisation $\pi^{-1}(U) \cong U \times G$ if and only if there is a local section defined on U.

Definition 14 (Principal bundle morphism)**.** *Given two principal fiber bundles (P, π, M, G) and (P', π', M', G'), we define a principal bundle morphism as a pair of smooth functions (F, f), $F : P \longrightarrow P'$, $f : M \longrightarrow M'$, such that $\pi' \circ F = f \circ \pi$, that is, the diagram*

$$\begin{tikzcd} P \arrow[r, "F"] \arrow[d, "\pi"'] & P' \arrow[d, "\pi'"] \\ M \arrow[r, "f"] & M' \end{tikzcd}$$

commutes, together with a homomorphism of groups $\Phi : G \longrightarrow G'$ such that $F(p \cdot g) = F(p) \cdot \Phi(g)$ for all $p \in P$, $g \in G$. Note again that f is determined by F.

Definition 15 (Principal bundle automorphism)**.** *A principal bundle morphism between a principal bundle P and itself given by diffeomorphisms is called an automorphism of P. We denote the group of all such morphisms by $AutP$.*

Definition 16 (Gauge transformation)**.** *A principal bundle automorphism (F, f) such that the induced function f is the identity map on M will be called a gauge transformation or vertical morphism. The group of all gauge transformations is denoted by $GauP$.*

A fiber bundle that will appear quite often is the adjoint bundle, which is a particular case of associated bundle. The definition of associated bundle, as the one of principal bundle, also involves a group action on the fibers but does not require the fibers to be diffeomorphic to the group.

Definition 17 (Associated bundle)**.** *Let (P, π, M, G) be a principal fiber bundle and F a differentiable manifold on which G acts on the left; we will construct another fiber bundle with these ingredients (the associated bundle to the given principal bundle and to the action on F). Let us consider the product manifold $P \times F$ and define a right action of G on it by*

$$\begin{array}{ccc} (P \times F) \times G & \longrightarrow & P \times F \\ ((p, \xi), g) & \longmapsto & (p \cdot g, g^{-1} \cdot \xi). \end{array}$$

This action defines an equivalence relation on $P \times F$ and the quotient $(P \times F)/G$ is denoted by $P \times_G F$. The associated bundle has $E := P \times_G F$ as its total space, M as its base space and the projection is given by

$$\begin{array}{cccc} \pi_E : & P \times_G F & \longrightarrow & M \\ & [(p, \xi)] & \mapsto & \pi(p). \end{array}$$

The idea of the associated bundle is to replace the fiber G of a principal bundle by some other differentiable manifold F. Let us check that this is what we are doing with the above definition:

The fiber of π_E over $x \in M$ is, by definition,

$$\pi_E^{-1}(x) = \{[(p, \xi)] : p \in P, \xi \in F \text{ with } \pi(p) = x\}.$$

If we fix a point $p_0 \in \pi^{-1}(x)$ then

$$\pi_E^{-1}(x) = \{[(p_0 \cdot g, \xi)] : g \in G, \xi \in F\} = \{[(p_0, g \cdot \xi)] : g \in G, \xi \in F\}$$

$$= \{[(p_0, \xi)] : \xi \in F\},$$

so the fiber of the associated bundle is diffeomorphic to F.

Now we construct local trivialisations of $P \times_G F$ from local trivialisations of P. Let $\varphi_U : \pi^{-1}(U) \longrightarrow U \times G$ be a local trivialisation of P. Then

$$\pi_E^{-1}(U) \cong (\pi^{-1}(U) \times F)/G \cong (U \times G \times F)/G \cong U \times F,$$

where the last bijection is given by $[(x, g, \xi)] \longmapsto (x, g \cdot \xi)$, which is well-defined since $[(x, g, \xi)] = [(x, gh, h^{-1} \cdot \xi)]$ is mapped to $[(x, g \cdot \xi)] = [(x, (gh) \cdot (h^{-1} \cdot \xi))]$. Then by theorem 1 we get that $P \times_G F$ is the total space of a fiber bundle.

If P is a trivial bundle we can reason as above with a global trivialisation and get that $P \times_G F \cong M \times F$ is trivial.

Example 3. We now introduce the frame fiber bundle. Let M be a differentiable manifold. We define a frame as a basis of T_xM for some $x \in M$ and denote by F_xM the set of all frames at $x \in M$. The total space of the frame bundle will be the set of all frames, that is

$$FM = \bigcup_{x \in M} F_xM,$$

and the projection will be

$$\begin{array}{rccc} \pi : & FM & \longrightarrow & M \\ & p & \longmapsto & x, \end{array}$$

where $p \in F_xM$.

Using theorem 1 we can give (FM, π, M) the structure of a fiber bundle. Note that $\pi^{-1}(x) \cong Gl(n, \mathbb{R})$. Let (U, x^i) be a chart on M. Then each frame $p \in \pi^{-1}(U)$ can be given coordinates (x^i, X^i_k), where $p = \left(X^i_1 \frac{\partial}{\partial x^i}\big|_x, \ldots, X^i_n \frac{\partial}{\partial x^i}\big|_x\right)$. This gives us the bijection required to apply theorem 1.

FM is in fact a principal fiber bundle with structure group $Gl(n, \mathbb{R})$. Let $p = (v_1, \ldots, v_n) \in FM$ be a frame, then the action (on the right) of $g = (a^i_j) \in Gl(n, \mathbb{R})$ on p is defined as $p \cdot g = (y_1, \ldots, y_n)$, where $y_k = v_j a^j_k$. Note that the action is free, the orbits coincide with the fibers and FM is locally trivial in the sense of the second definition of principal bundle (definition 13).

Now we consider the action of $Gl(n, \mathbb{R})$ on $\mathbb{R}^n$ on the left given by the usual product, that is, if $\xi = (\xi^1, \ldots, \xi^n) \in \mathbb{R}^n, g = (a^i_j) \in Gl(n, \mathbb{R})$ then $g \cdot \xi = (a^1_j \xi^j, \ldots, a^n_j \xi^j)$. The associated bundle to the frame bundle and this action of $Gl(n, \mathbb{R})$ on $\mathbb{R}^n$ is the tangent bundle. The identification is given by the map $[(p, \xi)] \longmapsto v_j \xi^j$ which is clearly well-defined and gives and isomorphism between $\pi_E^{-1}(\pi(p))$ and $T_{\pi(p)}M$.

Proposition 2. *If the action of G on F is trivial, that is, $g \cdot \xi = \xi$ for all $g \in G$ and $\xi \in F$, then the associated bundle $P \times_G F$ is trivial.*

Proof. Since $[(p, \xi)] = [(p \cdot g, g^{-1} \cdot \xi)] = [(p \cdot g, \xi)]$ for all $p \in P$, $g \in G$, $\xi \in F$, we get

$$P \times_G F \cong P/G \times F = M \times F.$$

□

A particular example of associated bundle is the adjoint bundle, which replaces the fiber G by its Lie algebra $\mathfrak{g}$.

Definition 18 (Adjoint bundle). *Let (P, π, M, G) be a principal fiber bundle and consider the action of the structure group G on its Lie algebra $\mathfrak{g}$ given by the adjoint representation*

$$\begin{array}{rccc} Ad: & G & \longrightarrow & Aut(\mathfrak{g}) \\ & g & \longmapsto & Ad_g. \end{array}$$

Recall that Ad_g was defined to be the differential of the conjugation map at the identity, that is $Ad_g = (dC_g)_e$, where

$$\begin{array}{rccc} C_g: & G & \longrightarrow & G \\ & h & \longmapsto & ghg^{-1}. \end{array}$$

The adjoint bundle is the associated bundle to π and to this action of G on $\mathfrak{g}$; its fibers are therefore isomorphic to $\mathfrak{g}$. We will denote it by $\tilde{\mathfrak{g}}$. Another frequent notation is $ad(P)$.

For a rewiev of the adjoint representation or other related concepts see [10].

Remark 9. If G is abelian then $\tilde{\mathfrak{g}}$ is trivial, since the conjugation map is the identity. Then the adjoint representation is trivial and proposition 2 gives that $\tilde{\mathfrak{g}}$ is trivial.

There are several ways to define a connection on a principal fiber bundle P, for instance as a horizontal distribution on P, as a 1-form on P taking values in the Lie algebra $\mathfrak{g}$ of the structure group G (which corresponds to the vertical projection with respect to the horizontal distribution) or as a covariant derivative in the linear case.

The tangent space at $p \in P$, T_pP, has a canonical direction, the vertical direction, determined by the action of the structure group. The subspace of vertical vectors at p, that is, vectors tangent to the fiber, is called the vertical subspace and denoted by V_pP or just V_p. The subbundle of TP with fibers V_p will be denoted by VP. If we consider the differential of the projection π, $(d\pi)_p : T_pP \longrightarrow T_{\pi(p)}M$, then this subspace is defined as

$$V_pP = \{v \in T_pP : (d\pi)_p(v) = 0\},$$

since the vectors tangent to the fibers are the vectors tangent to curves on which π is constant. In order to express any vector $v \in T_pP$ as a sum of a vertical component and some other component we need to choose some other direction, that is, define what will be horizontal. This will allow us to derive sections of arbitrary vector bundles, but for now let us give the first definition of connection, which is just the choice of a horizontal space.

Definition 19 (Connection 1). *A connection on a principal fiber bundle (P, π, M, G) is a differentiable distribution H on P of rank $dim(M)$ such that the following properties are satisfied:*

- *$T_pP = V_p \oplus H_p$, for all $p \in P$, so any $v \in T_pP$ can be uniquely written as $v = x + y$ with $x \in V_p$ and $y \in H_p$,*
- *$H_{p\cdot g} = (dR_g)_p(H_p)$, for all $g \in G$, $p \in P$, where*

$$\begin{array}{rccc} R_g: & P & \longrightarrow & P \\ & p & \longmapsto & p \cdot g. \end{array}$$

For the splitting of $v \in T_pP$ when a connection is given, we will use the notation $v = v^v + v^h$, with $v^v \in V_pP$ and $v^h \in H_pP$.

A particular type of vector field on P which will be important is the following:

Definition 20 (Horizontal vector field)**.** *A vector field $X \in \mathfrak{X}(P)$ is called horizontal if $X_p \in H_p$ for all $p \in P$.*

We will now introduce another type of vector field on P which will be needed in the second definition of connection:

Definition 21 (Fundamental vector field)**.** *For each $A \in \mathfrak{g}$, the fundamental vector field A^* associated to it is the vector field on P with flow given by*

$$\Phi(p,t) = p \cdot exp(tA), \text{ for all } p \in P, t \in \mathbb{R},$$

that is, $A^*_p = \frac{d}{dt}\big|_{t=0} p \cdot exp(tA)$.

Remark 10. Since the integral curves remain on the fiber, fundamental vector fields are vertical.

The interesting fact about these vector fields is that for each $p \in P$ they give a linear isomorphism with the vertical subspaces:

$$\begin{array}{rccc} \sigma_p : & \mathfrak{g} & \longrightarrow & V_pP \\ & A & \longmapsto & A^*_p. \end{array}$$

Note that $(dR_g)_p A^*_p = (Ad_{g^{-1}}A)^*_{p\cdot g}$. Indeed,

$$\begin{aligned} (dR_g)_p A^*_p &= \frac{d}{dt}\bigg|_{t=0} R_g(p \cdot exp(tA)) = \frac{d}{dt}\bigg|_{t=0} p \cdot exp(tA)g \\ &= \frac{d}{dt}\bigg|_{t=0} (p \cdot g) \cdot g^{-1} exp(tA) g = (Ad_{g^{-1}}A)^*_{p\cdot g}. \end{aligned}$$

As we mentioned above, the second definition of a connection follows from the first one by taking the projection of tangent vectors to its vertical component with respect to the chosen distribution. Then we assign the corresponding element in $\mathfrak{g}$ via the isomorphism σ_p, so we obtain a 1-form w on P with values in $\mathfrak{g}$ called the associated connection form. This 1-form satisfies the following properties:

- $w(A^*) = A$ for all $A \in \mathfrak{g}$,
- $(R_g^* w)(X) = Ad_{g^{-1}} w(X)$, for all $g \in G$, $X \in \mathfrak{X}(P)$.

On the other hand, if we consider a 1-form w on P with values in $\mathfrak{g}$ satisfying the above properties, we can take

$$H_p = \{v \in T_pP : w(v) = 0\},$$

which defines a distribution on P satisfying the properties of the first definition of connection. Therefore, instead of referring to w as the 1-form associated to the connection, we will also call w a connection.

Definition 22 (Connection 2)**.** *A connection w on P is a* 1*-form on P taking values in* $\mathfrak{g}$ *which satisfies*

- $w(A^*) = A$ *for all* $A \in \mathfrak{g}$,
- $(R_g^* w)(X) = Ad_{g^{-1}} w(X)$, *for all* $g \in G$, $X \in \mathfrak{X}(P)$.

Let us see the equivalence between the two defintions with a little more detail:

Proposition 3. *The two previous definitions of connection are equivalent.*

Proof. Let H be a distribution satisfying the properties of the first definition. We define a 1-form on TP with values in $\mathfrak{g}$ as

$$\begin{array}{rccc} w: & TP & \longrightarrow & \mathfrak{g} \\ & v & \longmapsto & A, \end{array}$$

where $v \in T_pP$ and $A_p^* = v^v$. Since A_p^* is vertical, the first property is satisfied (recall that w is just the vertical projection composed with the isomorphism σ_p^{-1}). Now we check the second property. Let $v \in T_pP$, then

$$\begin{aligned} R_g^* w(v) &= w(dR_g(v)) = w(dR_g(v^v + v^h)) = w(dR_g(v^v)) + w(dR_g(v^h)) \\ &= w(dR_g A_p^*) = w((Ad_{g^{-1}} A)^*_{p \cdot g}) = Ad_{g^{-1}} A = Ad_{g^{-1}}(w(A_p^*)), \end{aligned}$$

where we are using the second property of the first definition ($(dR_g)_p(v^h)$ is horizontal), the property of fundamental vector fields we mentioned above and the first property of the second definition.

Now, if w is a 1-form on P with values in $\mathfrak{g}$ satisfying the properties of the second definition, define $H_p = \{v \in T_pP : w(v) = 0\}$. Then $T_pP = V_pP \oplus H_p$, for $w : T_pP \longrightarrow \mathfrak{g} \cong V_pP$ is surjective, because of the first property. To see the second property we use the following equalities:

$$w((dR_g)_p v) = R_g^* w(v) = Ad_{g^{-1}} w(v),$$

where $v \in T_pP$. $Ad_{g^{-1}}$ is an isomorphism and therefore $w(v) = 0 \Leftrightarrow Ad_{g^{-1}} w(v) = 0 \Leftrightarrow w((dR_g)_p v) = 0$, that is, we have $H_{p \cdot g} = (dR_g)_p(H_p)$.

Note that if we start with the distribution, define the associated 1-form and then take the null spaces then we recover the original distribution. □

Now assume that a connection H on P is given, according to the first definition. Then $(d\pi)_p : T_pP \longrightarrow T_{\pi(p)}M$ gives an isomorphism between H_p and $T_{\pi(p)}M$, so each vector $v \in T_xM$ can be assigned a unique vector on T_pP for each $p \in \pi^{-1}(x)$ via these isomorphisms. Given a vector field $X \in \mathfrak{X}(M)$, let us consider the vector field on P defined in this manner:

Definition 23 (Horizontal lift)**.** *Given a principal fiber bundle (P, π, M, G), a connection on it and a vector field $X \in \mathfrak{X}(M)$, the horizontal lift X^* of X is defined as the only horizontal vector field on P such that*

$$(d\pi)_p X_p^* = X_{\pi(p)}, \text{ for all } p \in P.$$

Note that the horizontal lift X^* is G-invariant, that is,

$$(dR_g)_p X_p^* = X_{p\cdot g}^* \text{ for all } g \in G, p \in P,$$

because of the G-invariance of the connection.

Remark 11. Any horizontal vector field Y on P which is G-invariant is the horizontal lift of a vector field X on M, namely $X = d\pi(Y)$, which means $X_q = (d\pi)_p Y_p$ for an arbitrary $p \in \pi^{-1}(q)$ and which is well-defined because of the G-invariance.

Now we give some properties of horizontal lifts and fundamental vector fields:

Proposition 4. *Let $X, Y \in \mathfrak{X}(M)$, $f \in \mathcal{F}(M)$ and $A, B \in \mathfrak{g}$. We have*

- $(X+Y)^* = X^* + Y^*$,
- $(fX)^* = f^* X^*$, *where* $f^* = f \circ \pi$,
- $[X,Y]^* = [X^*, Y^*]^h$ *(recall that* $\cdot^h$ *denotes the horizontal projection), so in general* $[X^*, Y^*]$ *is not horizontal,*
- *if* X *is horizontal, then* $[X, A^*]$ *is also horizontal,*
- $[A,B]^* = [A^*, B^*]$.

Proof. For the first three properties apply $d\pi$ to both sides. For example, the third property follows from

$$d\pi[X^*, Y^*]^h = d\pi[X^*, Y^*] = [d\pi X^*, d\pi Y^*] = [X, Y].$$

For the fourth property just write the Lie bracket as the Lie derivative

$$[A^*, X]_p = \lim_{t\to 0} \frac{(d\phi_{-t})_{\phi_t(p)}(X_{\phi_t(p)}) - X_p}{t},$$

where $\phi(t,p)$ is the flow of the vector field A^*, that is, $\phi(p,t) = p \cdot exp(tA)$, and therefore $\phi_t = R_{exp(tA)}$. Then $d\phi_{-t} = dR_{-exp(tA)}$, so $(d\phi_{-t})_{\phi_t(p)}(X_{\phi_t(p)})$ is horizontal and also $[A^*, X]_p$.

For the fifth property we use $[\sigma_p A, \sigma_p B] = \sigma_p([A,B])$ which follows again from the expression of the Lie bracket as the Lie derivative and the fact that σ_p is an isomorphism:

$$[\sigma_p A, \sigma_p B] = \lim_{t\to 0} \frac{(d\phi_{-t})_{\phi_t(p)} \sigma_{\phi_t(p)} B - \sigma_p B}{t} = \lim_{t\to 0} \frac{(dR_{-exp(tA)})_{\phi_t(p)} B^*_{\phi_t(p)} - \sigma_p B}{t}$$

$$= \lim_{t\to 0} \frac{(Ad_{exp(tA)} B)^*_{\phi_t(p)\cdot(-exp(tA))} - \sigma_p B}{t} = \lim_{t\to 0} \frac{\sigma_{p\cdot exp(tA)\cdot(-exp(tA))} Ad_{exp(tA)} B - \sigma_p B}{t}$$

$$= \sigma_p \lim_{t\to 0} \frac{Ad_{exp(tA)} B - B}{t} = \sigma_p [A, B],$$

for $Ad_{exp(tA)} B = dR_{-exp(tA)} dL_{exp(tA)} B = dR_{-exp(tA)} B$ (here R_g and L_g denote right and left multiplication on the Lie group). Then

$$[A_p^*, B_p^*] = [\sigma_p A, \sigma_p B] = \sigma_p([A,B]) = [A,B]_p^*.$$

□

Remark 12. Note that in general if $f, g \in \mathcal{F}(P)$ then $fX^* + gY^*$ will not be a horizontal lift, although it is a horizontal vector field, since it will not necessarily be G-invariant. For example, if the two vector fields are independent then $fX^* + gY^*$ will be G-invariant if and only if $f = \tilde{f} \circ \pi$ and $g = \tilde{g} \circ \pi$ for some $\tilde{f}, \tilde{g} \in \mathcal{F}(M)$ (that is, $f = \tilde{f}^*$ and $g = \tilde{g}^*$ with the notation in the second property).

Using the connection we can define an alternative way to derive r-forms to the exterior derivative, namely the exterior covariant derivative. First we introduce some particular types of r-forms:

Definition 24. *Let φ be an r-form on P taking values in a vector space V of finite dimension on which G acts on the left. We say that φ is*

- *horizontal, if*

$$\varphi(X_1, \ldots, X_r) = 0 \textit{ whenever } X_i \textit{ is vertical for some } i,$$

- *pseudotensorial, if*

$$R_g^* \varphi(p) = g^{-1} \cdot \varphi(p),$$

  where we are using the notation $\cdot$ for the action of G on V, and

- *tensorial, if it is pseudotensorial and horizontal.*

If ρ denotes the action on V then we say that the pseudotensorial or tensorial form is of type (ρ, V).

Remark 13. The connection form w is pseudotensorial of type $(Ad, \mathfrak{g})$.

Definition 25 (Exterior covariant derivative)**.** *Let φ be an r-form on P. The exterior covariant derivative $D\varphi$ of φ is the $(r+1)$-form defined as*

$$D\varphi(X_1, \ldots, X_{r+1}) = d\varphi(X_1^h, \ldots, X_{r+1}^h),$$

where $X_1, \ldots, X_{r+1} \in \mathfrak{X}(P)$.

If we want to put emphasis on the connection w we will write $D = d^w$.

Proposition 5. *Let φ be a pseudotensorial r-form on P of type (ρ, V). Then*

- *$d\varphi$ is a pseudotensorial $(r+1)$-form of type (ρ, V),*
- *$\varphi \circ \pi_h$ is a tensorial r-form of type (ρ, V), where π_h is the projection to the horizontal space given by the connection,*
- *$D\varphi$ is a tensorial $(r+1)$-form of type (ρ, V).*

Definition 26 (Curvature form)**.** *The curvature form Ω is the exterior covariant derivative of the connection form ω, that is,*

$$\Omega = D\omega.$$

The curvature form is therefore a tensorial 2-form of type $(Ad, \mathfrak{g})$.

Proposition 6. *A connection (as a distribution) is integrable in the sense of Frobenius if and only if its curvature vanishes.*

Proof. Since

$$\Omega(X,Y) = dw(X^h,Y^h) = X^h(w(Y^h)) - Y^h(w(X^h)) - w([X^h,Y^h]) = -w([X^h,Y^h]),$$

we have

$$\Omega = 0 \Leftrightarrow [X^h,Y^h] \text{ is horizontal},$$

that is,

$$\Omega = 0 \Leftrightarrow \text{ the distribution } H \text{ is involutive }.$$

Furthermore H can be locally generated by $\dim(M)$ vector fields, that is, H is differentiable. Hence, by Frobenius theorem, we get that H is integrable if and only if $\Omega \equiv 0$. □

Proposition 7 (Structure equation)**.** *If w is a connection on a principal fiber bundle and Ω is its curvature form, then*

$$\Omega = d\omega + [w,w],$$

where by definition, $[w,w](X,Y) = [w(X),w(Y)]$.

Remark 14. In coordinates, if $a = a^\alpha_j dx^j \otimes B_\alpha$ is a 1-form with values in $\mathfrak{g}$ ($\{B_\alpha\}_\alpha$ is a basis of $\mathfrak{g}$), then

$$[a,a] = a^\alpha_j a^\beta_k dx^j \wedge dx^k \otimes [B_\alpha, B_\beta].$$

In order to prove proposition 7, just do the calcuations separating the cases when both vector fields are horizontal, when both are vertical and when one is horizontal and the other one is vertical. Some of the properties of proposition 4 are used.

Proposition 8. *If φ is a tensorial 1-form of type $(Ad,\mathfrak{g})$ then*

$$d^\omega\varphi = d\varphi + \omega\wedge\varphi,$$

where $\omega\wedge\varphi(X,Y) = [\omega(X),\varphi(Y)] - [\omega(Y),\varphi(X)]$.

A proof is given in [6].

Remark 15. The connection form is not tensorial, so there is no contradiction with the structure equation.

Remark 16. In fact the formula $d^\omega\alpha = d\alpha + \omega\wedge\alpha$ is also valid for a tensorial k-form α of type $(Ad,\mathfrak{g})$, taking into account all permutations in the definition of $\omega\wedge\alpha$.

Proposition 9 (Bianchi identity)**.** *Let Ω be the curvature form of a connection. Then*

$$D\Omega = 0.$$

For the proof take the exterior derivative on the structure equation and see that the resulting 3-form vanishes when applied to three horizontal vectors.

Remark 17. In general it is not true that $D^2 = 0$, in fact $D^2(\cdot) = \Omega\wedge\cdot$.

We will now study another side of the connection, namely the parallel transport and covariant derivative, which provides a way to derive sections of a vector bundle using the connection.

Definition 27. *Let P be a principal fiber bundle with a connection. A smooth curve on P is called horizontal if its tangent vectors are horizontal with respect to the given connection.*

Proposition 10. *Let (P, π, M, G) be a principal fiber bundle with a connection and let $\alpha : [0,1] \longrightarrow M$ be a smooth curve in M. Then for each $p \in \pi^{-1}(\alpha(0))$ there exists a unique horizontal curve $\alpha_p^* : [0,1] \longrightarrow P$ in P such that $\alpha_p^*(0) = p$ and $\pi(\alpha^*(t)) = \alpha(t)$ for all $t \in [0,1]$. We call α^* the horizontal lift of α.*

See [6] for a proof.

The horizontal lift α^* induces a diffeomorphism between $\pi^{-1}(\alpha(0))$ and $\pi^{-1}(\alpha(1))$, called parallel transport:

$$\begin{array}{rrcl} c_{0,1}: & \pi^{-1}(\alpha(0)) & \longrightarrow & \pi^{-1}(\alpha(1)) \\ & p = \alpha_p^*(0) & \longmapsto & \alpha_p^*(1). \end{array}$$

In fact, the horizontal lift of the curve induces a diffeomorphism between $\pi^{-1}(\alpha(s))$ and $\pi^{-1}(\alpha(t))$ for all $s, t \in [0,1]$, denoted by $c_{s,t}$. Note that the parallel transport commutes with the group action.

For each associated bundle to P we can define a notion of horizontality from the connection on the original bundle. Since the associated bundle is not necessarily principal we need to adapt the definition of connection:

Definition 28 (Ehresmann connection)**.** *An Ehresmann connection A on a fiber bundle $\pi : E \longrightarrow M$ is a 1-form on E taking values in the vertical subbundle, that is,*

$$A_p : T_pE \longrightarrow V_pE, \text{ for all } p \in E$$

and satisfying also $A_p(v_p) = v_p$ for all $v_p \in V_pE$.

Remark 18. Note that in the particular case of a principal bundle, an Ehresmann connection and a connection take values in different spaces.

An Ehresmann connection can also be defined as a horizontal distribution as in the case of connections but without the condition concerning the action of the structure group.

Given a connection H on the principal fiber bundle P, we can define an Ehresmann connection on $E = P \times_G F$ by $H_{[(p,\xi)]} = d\xi(H_p)$, where

$$\begin{array}{rrcl} \xi: & P & \longrightarrow & E \\ & p & \longmapsto & [(p,\xi)]. \end{array}$$

In this case we have an analogous result to proposition 10 and can define the parallel transport.

If the associated bundle E is a vector bundle, the parallel transport is an isomorphism and provides a way to define the derivative of a section, a covariant derivative. Let

$s \in \Gamma(E)$, and denote by $\dot{\alpha}(t)$ the tangent vector of α at t. Then the covariant derivative of s at $\alpha(t)$ in the direction of $\dot{\alpha}(t)$ is given by the formula

$$\nabla_{\dot{\alpha}(t)} s = \lim_{h \to 0} \frac{1}{h} \left(c_{t,t+h}^{-1}(s(\alpha(t+h))) - s(\alpha(t)) \right).$$

If $X \in \mathfrak{X}(M)$, we can define the covariant derivative $\nabla_X s$ taking the integral curves of X in the above definition. $\nabla_X s$ is a section of E defined by

$$\nabla_X s(x) = \nabla_{\dot{\alpha}(t)} s$$

where α is an integral curve of X, with $\alpha(t) = x$ and $\dot{\alpha}(t) = X_x$.

In general a covariant derivative in a vector bundle (E, π, M) is a map

$$\begin{array}{rccc} \nabla: & \mathfrak{X}(M) \times \Gamma(E) & \longrightarrow & \Gamma(E) \\ & (X, s) & \longmapsto & \nabla_X s \end{array}$$

satisfying the following properties:

- $\nabla_{X+Y} s = \nabla_X s + \nabla_Y s$,
- $\nabla_X (s + t) = \nabla_X s + \nabla_X t$,
- $\nabla_{fX} s = f \nabla_X s$,
- $\nabla_X (fs) = X(f)s + f\nabla_X s$,

for all $X, Y \in \mathfrak{X}(M), f \in \mathcal{F}(M), s, s' \in \Gamma(E)$.

If the associated bundle is a vector bundle, then we can relate the covariant derivative defined by the parallel transport in the associated bundle with the exterior covariant derivative in the original bundle. This relation is given by the following result:

Proposition 11. *Let (P, π, M, G) be a principal fiber bundle with a connection and let (E, ρ, M, V) be an associated vector bundle with an induced connection. Let $Y \in T_pP$, $X \in T_{\pi(p)}M$ with $d\pi(Y) = X$. Let s be a section of E and $f : P \longrightarrow V$ be the smooth map defined by $f(p) = p^{-1}s(\pi(p))$, where*

$$\begin{array}{rccc} p: & V & \longrightarrow & E_{\pi(p)} \\ & \xi & \longmapsto & [(p, \xi)] \end{array}$$

is an isomorphism. Then we get

$$\nabla_X s = p(Df(Y)).$$

Now let (x^i, y^a) be adapted coordinates on a fiber bundle $E \longrightarrow M$. Locally we can define an Ehresmann connection by giving the horizontal lift:

$$\frac{\partial}{\partial x^i} \longmapsto \frac{\partial}{\partial x^i} - \Gamma_i^a(x, y) \frac{\partial}{\partial y^a},$$

where Γ_a^i depend both on x^i and y^a.

In the case of a vector bundle, if there exist smooth maps $\Gamma_{ij}^a \in \mathcal{F}(M)$ such that $\Gamma_i^a(x, y) = \Gamma_{ij}^a(x) y^j$ then we say that the connection is linear. In this case a covariant derivative is another alternative definition of connection, for we have

Theorem 2. *There is a bijection between linear connections and covariant derivatives in a vector bundle* (E, π, M).

See chapter 3 of [6].

Example 4. For the tangent bundle $TM \longrightarrow M$, sections are vector fields on M, so a covariant derivative (linear connection) is a map

$$\begin{array}{rccc} \nabla : & \mathfrak{X}(M) \times \mathfrak{X}(M) & \longrightarrow & \mathfrak{X}(M) \\ & (X, Y) & \longmapsto & \nabla_X Y \end{array}$$

satisfying the corresponding properties which coincide with the usual definition of affine connection on M from Riemannian geometry. In this case the coefficients of the connection coincide with the Christoffel symbols.

In the case of a principal connection there is locally a basis of the vertical space such that we can write the horizontal lift with coefficients depending only on the coordinates (x^i) of the base space. Let $\{B_1, \ldots, B_m\}$ be a basis of $\mathfrak{g}$ and let $\pi^{-1}(U) \cong U \times G$ be a local trivialisation of P. We define the following vector fields on $\pi^{-1}(U)$:

$$(\tilde{B}_a)_p = \left.\frac{d}{dt}\right|_{t=0} (x, exp(tB_a) \cdot g),$$

where $p = (x, g) \in U \times G \cong \pi^{-1}(U)$. These vector fields are clearly vertical and give a basis for each vertical subspace. In contrast with the fundamental vector fields they depend on the trivialisation, but have the advantage of being G-invariant, that is,

$$(dR_g)_p(\tilde{B}_a)_p = (\tilde{B}_a)_{p \cdot g}.$$

Recall that this was not true for fundamental vector fields, for they satisfied $(dR_g)B^* = (Ad_{g^{-1}}B)^*$.

We write the local expression of the horizontal lift given by the G-invariant vertical vector fields $\tilde{B}_a$ as

$$\left(\frac{\partial}{\partial x^i}\right)_x \longmapsto \left(\frac{\partial}{\partial x^i}\right)_{(x,g)} - \Gamma_i^a(x, g)(\tilde{B}_a)_{(x,g)}.$$

If w denotes the connection form (with values in VP) then the G-invariance of $\tilde{B}_a$ gives

$$w\left(\left(\frac{\partial}{\partial x^i}\right)_{(x,gh)}\right) = w\left(\Gamma_i^a(x, gh)(\tilde{B}_a)_{(x,gh)}\right) = dR_h\left(\Gamma_i^a(x, gh)(\tilde{B}_a)_{(x,g)}\right).$$

Since

$$w\left(\frac{\partial}{\partial x^i}\right)_{(x,gh)} = w\left(dR_h\left(\frac{\partial}{\partial x^i}\right)_{(x,g)}\right) = dR_h\left(w\left(\frac{\partial}{\partial x^i}\right)_{(x,g)}\right),$$

we also have

$$dR_h\left(\Gamma_i^a(x, gh)(\tilde{B}_a)_{(x,g)}\right) = dR_h\left(\Gamma_i^a(x, g)(\tilde{B}_a)_{(x,g)}\right),$$

and hence

$$\Gamma_i^a(x, g) = \Gamma_i^a(x, gh), \text{ for all } h \in G.$$

Then we can write the horizontal lift as

$$\frac{\partial}{\partial x^i} \longmapsto \frac{\partial}{\partial x^i} - \Gamma_i^a(x)\tilde{B}_a,$$

with $\Gamma_i^a \in \mathcal{F}(U)$.

2.4 Jet bundles

The idea of jet bundles is to put together all local sections of a given fiber bundle having the same Taylor expansion of first order at a point, that is,

Definition 29 (1-jet of a section)**.** *Given a fiber bundle $\pi : E \longrightarrow M$ and a point $x \in M$, we say that two local sections $s, s' \in \Gamma_x(E)$ belong to the same equivalence class $j^1_x s$ if $s(x) = s'(x)$ and $(ds)_x = (ds')_x$. We call $j^1_x s$ the 1-jet of s at x.*

1-jets can also be viewed as a generalisation of the concept of tangent vector, which is what we get if we consider the trivial bundle $\pi : \mathbb{R} \times Q \longrightarrow \mathbb{R}$ for some differentiable manifold Q. The sections of this bundle can be identified with smooth curves on Q, and 1-jets of sections will be equivalence classes of smooth curves on Q with the same differential at a point, that is, tangent vectors at a point.

It can be shown that this set of equivalence classes is a manifold and that respective projections to the total space and the base space give rise to fiber bundles. A thorough study on jet bundles is given in [9].

Definition 30 (First jet manifold)**.** *Let $\pi : E \longrightarrow M$ be a fiber bundle. The set of all 1-jets*

$$\left\{ j^1_x s : x \in M, s \in \Gamma_x(E) \right\}$$

is called the first jet manifold of E and is denoted by J^1E. The projection of J^1E onto the total space

$$\begin{array}{rccc} \pi_{1,0}: & J^1E & \longrightarrow & E \\ & j^1_x s & \longmapsto & s(x) \end{array}$$

will be called the target projection. The notation indicates that the projection goes from the first jet manifold to E, which would be the 0-jet manifold. Higher order jet manifolds can be defined and corresponding projections can be written as $\pi_{k,k-1}$. The projection onto the base space

$$\begin{array}{rccc} \pi_{1}: & J^1E & \longrightarrow & M \\ & j^1_x s & \longmapsto & x \end{array}$$

is called the source projection (corresponding π_k source projections for higher orders are also defined).

An atlas on J^1E can be defined from an atlas on E in a way analogous to how an atlas on M induced an atlas on TM.

Let us give coordinates on J^1E. If (x^i, y^α) are local adapted coordinates on E, then $(x^i, y^\alpha, y^\alpha_i)$ with $x^i(j^1_x s) := x^i(x) = x^i(s(x))$, $y^\alpha(j^1_x s) := y^\alpha(s(x))$ and

$$y^\alpha_i(j^1_x s) := \left. \frac{\partial(y^\alpha \circ s)}{\partial x^i} \right|_x$$

are local adapted coordinates on $J^1E \longrightarrow E$. Note that if (x^i, y^α) are coordinates on an open subset $U \subset E$ then the coordinates defined on J^1E are valid for $j^1_x s$ such that $s(x) \in U$.

Now we want to see that the first jet manifold is indeed a manifold. We will not prove this result in detail, but it can be found in [9]. An outline of the proof would be as follows:

Lemma 1. *If two local coordinate systems $u = (x^i, y^\alpha, y^\alpha_i)$ and $v = (z^j, t^\beta, t^\beta_j)$ on J^1E have overlappig domains then the composition $v \circ u^{-1}$ is smooth.*

Lemma 2. *The source and target projections are smooth submersions.*

Lemma 3. $\pi_{1,0} : J^1E \longrightarrow E$ *admits local affine trivialisations around any point.*

Finally applying theorem 1 we get that the first jet manifold J^1E is a manifold, namely the total space of the now fiber bundle $\pi_{1,0} : J^1E \longrightarrow E$.

In fact we have

Proposition 12. *The jet bundle* $\pi_{1,0} : J^1E \longrightarrow E$ *is an affine bundle modelled on the vector bundle* $T^*M \otimes VE$ *(formally we should write* $\pi^*(T^*M) \otimes VE$*).*

Consider $(J^1E)_y$ the fiber of $\pi_{1,0} : J^1E \longrightarrow E$ over $y \in E$. Take $j^1_x s,\ j^1_x s' \in (J^1E)_y$, where necessarily $x = \pi(y)$ and $s(x) = s'(x) = y$. Then $v \in T_xM$ can be lifted to T_yE in two different ways according to each of the 1-jets and the difference is a vertical vector. We can define the difference $j^1_x s - j^1_x s'$ to be the co-vector that sends v to this vertical vector and so obtain that the fiber $(J^1E)_y$ is modelled on $T^*_xM \otimes V_yE$ and J^1E is modelled on $T^*M \otimes VE$, that is, on the fiber bundle of 1-forms on M with values in VE.

Note that the same is not valid for $\pi_1 : J^1E \longrightarrow M$ since vectors on M would not necessarily be lifted to the same tangent space on E when considering the horizontal spaces determined by jets in the same fiber. Then we cannot take the difference. Nevertheless, we have the following proposition:

Proposition 13. $\pi_1 : J^1E \longrightarrow M$ *is a fiber bundle.*

This result follows from the fact that $\pi_1 = \pi \circ \pi_{1,0}$, that is, π is a composite fiber bundle. See [4] or [5].

Remark 19. Note that a 1-jet $j^1_x s$ is a choice of horizontal space at $s(x)$. Therefore, given a section of $\pi_{1,0}$ we obtain an Ehresmann connection on E, taking $H_{s(x)} = Im(ds)_x$.

Let us see a couple of useful examples:

Example 5. Let $pr_1 : M \times \mathbb{R} \longrightarrow M$ be a trivial fiber bundle. The first jet manifold $J^1(M \times \mathbb{R})$ is canonically diffeomorphic to $T^*M \times \mathbb{R}$ via the map

$$\begin{array}{ccc} J^1(M \times \mathbb{R}) & \longrightarrow & T^*M \times \mathbb{R} \\ j^1_x s & \longmapsto & ((d\bar{s})_x, \bar{s}(x))\,, \end{array}$$

where if s is a local section of $M \times \mathbb{R}$, then $\bar{s} = pr_2 \circ s$ is a smooth map on some open subset of M. Note that in this case the condition of first jet equivalence can be rewritten as $j^1_x s = j^1_x s'$ if $\bar{s}(x) = \bar{s'}(x)$ and $(d\bar{s})_x = (d\bar{s'})_x$, where $s, s' \in \Gamma_x(M \times \mathbb{R})$, for $pr_1 \circ s$ is the identity map on its domain. Therefore the map is well-defined and injective. Note that it is also surjective and that its local expression is just a swap of coordinates:

$$(x^i, y, y_i) \longmapsto (x^i, y_i, y),$$

so we get that it is a diffeomorphism.

Example 6. Let $pr_1 : \mathbb{R} \times M \longrightarrow \mathbb{R}$ be a trivial fiber bundle. The first jet manifold $J^1(\mathbb{R} \times M)$ is canonically diffeomorphic to $\mathbb{R} \times TM$ via the map

$$\begin{array}{ccc} J^1(\mathbb{R} \times M) & \longrightarrow & \mathbb{R} \times TM \\ j^1_x s & \longmapsto & (x,\ \frac{d}{dt}\big|_{t=x} pr_2 \circ s(t)). \end{array}$$

Remark 20. Note that in both examples the first jet manifold is a vector bundle. In fact one can strengthen proposition 12 and get that if the fiber bundle $E \longrightarrow M$ is trivial, via a diffeomorphism Φ, then $\pi_{1,0} : J^1E \longrightarrow E$ is a vector bundle. In order to get this, a global section of $\pi_{1,0}$ is defined, determining a zero for each affine fiber. This global section is given by $z(a) = j^1_{\pi(a)} s_a$, where $a \in E$ and s_a is the section of E defined by $s_a(x) = \Phi^{-1}(x, pr_2(\Phi(a)))$.

We now introduce the concepts of prolongation of a section and prolongation of a morphism to the first jet manifold.

Definition 31 (Prolongation of a section)**.** *Let $\pi : E \longrightarrow M$ be a fiber bundle and let $s \in \Gamma_W(\pi)$ be a local section defined on $W \subset M$. We define the prolongation of s to the first jet manifold as*

$$\begin{array}{rccc} j^1 s : & W & \longrightarrow & J^1E \\ & x & \longmapsto & j^1 s(x) := j^1_x s. \end{array}$$

Remark 21. Note that the following diagram is commutative:

$$\begin{tikzcd} J^1E \arrow[d, "\pi_{10}"'] \arrow[dd, bend right=60, "\pi_1"'] \\ E \arrow[d, shift right, "\pi"'] \\ W. \arrow[u, shift right, bend right, "s"'] \arrow[uu, bend right=60, "j^1 s"'] \end{tikzcd}$$

In particular we have that $j^1 s$ is a section of π_1, for $\pi_1 \circ j^1 s(x) = \pi_1(j^1_x s) = x$, and that $\pi_{1,0} \circ j^1 s(x) = \pi_{1,0}(j^1_x s) = s(x)$.

Remark 22. There are sections of $J^1E \longrightarrow M$ which are not the prolongation of a section of $E \longrightarrow M$. If we have local coordinates $(x^i, y^\alpha, y^\alpha_i)$ on J^1E, a local section s of $J^1E \longrightarrow M$ can be written as

$$(x^i) \longmapsto (x^i, s^\alpha, s^\alpha_i),$$

where no relation between s^α and s^α_i is assumed. If s happens to be the prolongation of a section of $E \longrightarrow M$ then the expression of s in coordinates is

$$(x^i) \longmapsto \left(x^i, s^\alpha, \frac{\partial s^\alpha}{\partial x^i} \right).$$

Remark 23. If ψ is a local section of $J^1E \longrightarrow M$, then ψ satisfies $j^1(\pi_{1,0} \circ \psi) = \psi$ if and only if ψ is the prolongation of some section of $E \longrightarrow M$.

Definition 32 (Prolongation of a bundle morphism)**.** *Let $\pi : E \longrightarrow M$ and $\pi' : E' \longrightarrow M'$ be fiber bundles and let (F, f) be a bundle morphism between E and E' such that f is a diffeomorphism. The 1-jet prolongation of (F, f) is defined as*

$$\begin{array}{rccc} j^1 F : & J^1E & \longrightarrow & J^1E' \\ & j^1_x s & \longmapsto & j^1_{f(x)}(F \circ s \circ f^{-1}). \end{array}$$

See the following diagram:

$$\begin{tikzcd}
J^1E \arrow[r, "j^1F"] \arrow[d, "\pi_{10}"'] & J^1E' \arrow[d, "\pi'_{10}"] \\
E \arrow[r, "F"] \arrow[d, "\pi"] & E' \arrow[d, "\pi'"'] \\
M \arrow[r, "f"] \arrow[u, bend left, "s"] & M'. \arrow[u, bend right, "F\circ s\circ f^{-1}"']
\end{tikzcd}$$

Remark 24. Given $j^1_x s = j^1_x s'$, we have $j^1F(j^1_x s) = j^1F(j^1_x s')$, that is, the prolongation of a morphism is well-defined, since the conditions $F \circ s \circ f^{-1}(f(x)) = F \circ s' \circ f^{-1}(f(x))$ and $\left.\frac{\partial(F\circ s\circ f^{-1})}{\partial x^i}\right|_{f(x)} = \left.\frac{\partial(F\circ s'\circ f^{-1})}{\partial x^i}\right|_{f(x)}$ depend on the values and first derivatives of s and s', which coincide.

Remark 25. The definition of prolongation of a bundle morphism is a generalisation of the definition of prolongation of a section, which is the prolongation of the bundle morphism (s, id_M) between (M, id_M, M) and (E, π, M):

$$\begin{tikzcd}
M \arrow[r, "j^1s"] \arrow[d, "id_M"'] & J^1E \arrow[d, "\pi_{1,0}"] \\
M \arrow[r, "s"] \arrow[d, "id_M"'] & E \arrow[d, "\pi"] \\
M \arrow[r, "id_M"] & M.
\end{tikzcd}$$

Proposition 14. *The pairs (j^1F, F) and (j^1F, f) are bundle morphisms between π_{10} and π'_{10}, and between π_1 and π'_1 respectively.*

Proof. We have to check that the diagram

$$\begin{tikzcd}
J^1E \arrow[r, "j^1F"] \arrow[d, "\pi_{10}"] \arrow[dd, bend right=60, "\pi_1"'] & J^1E' \arrow[d, "\pi'_{10}"'] \arrow[dd, bend left=60, "\pi'_1"] \\
E \arrow[r, "F"] \arrow[d, "\pi"] & E' \arrow[d, "\pi'"'] \\
M \arrow[r, "f"] & M'
\end{tikzcd}$$

commutes, and to that end it is enough to check that

$$\begin{tikzcd}
J^1E \arrow[r, "j^1F"] \arrow[d, "\pi_{10}"'] & J^1E' \arrow[d, "\pi'_{10}"] \\
E \arrow[r, "F"] & E'
\end{tikzcd}$$

commutes, which follows easily:

$$\pi'_{10} \circ j^1F(j^1_x s) = \pi'_{10}(j^1_{f(x)}(F \circ s \circ f^{-1})) = F \circ s \circ f^{-1}(f(x)) = F \circ s(x) = F \circ \pi_{10}(j^1_x s).$$

$\square$

Proposition 15. *Let $\pi'' : E'' \longrightarrow M''$ be another fiber bundle and (G, g) a bundle morphism between π' and π'' with g a diffeomorphism. Then we have $j^1(G \circ F) = j^1 G \circ j^1 F$. Moreover $j^1 id_E = id_{J^1E}$.*

Proof. Just write the definitions. $\square$

If $(z^j, t^\beta, t^\beta_j)$ are coordinates in J^1E', then using the chain rule we get that the local expression of j^1F is written as

$$(x^i, y^\alpha, y^\alpha_i) \longmapsto \left(F^j, F^\beta, \left(\left.\frac{\partial F^\beta}{\partial x^i}\right|_x + y^\alpha_i \left.\frac{\partial F^\beta}{\partial y^\alpha}\right|_x \right) \left.\frac{\partial (f^{-1})^i}{\partial x^j}\right|_{f(x)} \right).$$

Therefore we have that the morphism j^1F restricted to each fiber is affine for its local expression is

$$y^\alpha_i \longmapsto \left.\frac{\partial F^\beta}{\partial x^i}\right|_x \left.\frac{\partial (f^{-1})^i}{\partial x^j}\right|_{f(x)} + \left.\frac{\partial F^\beta}{\partial y^\alpha}\right|_x \left.\frac{\partial (f^{-1})^i}{\partial x^j}\right|_{f(x)} y^\alpha_i.$$

Note that conditions on the last set of coordinates are imposed, so we get again that not every bundle morphism between jet bundles will be a prolongation.

Let X be a projectable vector field on E, that is, there exists a vector field $Y \in \mathfrak{X}(M)$ such that $d\pi(X) = Y$, that is, such that the diagram

$$\begin{tikzcd} E \arrow[r, "X"] \arrow[d, "\pi"'] & TE \arrow[d, "d\pi"] \\ M \arrow[r, "Y"] & TM \end{tikzcd}$$

commutes. Then the flow of X, $\{\Phi_t\}_{t\in\mathbb{R}}$, are bundle morphisms and its prolongations to the first jet manifold $\{j^1\Phi_t\}_{t\in\mathbb{R}}$ are the flow for a vector field on J^1E which will we called the prolongation of X and denoted by j^1X. A projectable vector field is written in coordinates as

$$X = f^i \frac{\partial}{\partial x^i} + g^\alpha \frac{\partial}{\partial y^\alpha},$$

where $f^i = f^i(x^j)$ and $g^\alpha = g^\alpha(x^i, y^\beta)$.

If we locally write $\Phi_t = (\Phi^i, \Phi^\alpha)$ for the flow of X, and then $\Psi_t = (\Phi^i)$ for the flow of Y, the prolongation of Φ_t to J^1E is given by

$$(x^j, y^\beta, y^\beta_j) \longmapsto \left(\Phi^i, \Phi^\alpha, \left(\left.\frac{\partial \Phi^\alpha}{\partial x^j}\right|_x + y^\beta_j \left.\frac{\partial \Phi^\alpha}{\partial y^\beta}\right|_x \right) \left.\frac{\partial (\Psi_t^{-1})^j}{\partial x^i}\right|_{\Psi_t(x)} \right).$$

Then taking derivatives with respect to t we get that the expression of j^1X in coordinates is

$$j^1X = f^i \frac{\partial}{\partial x^i} + g^\alpha \frac{\partial}{\partial y^\alpha} + \left(\frac{\partial g^\alpha}{\partial x^i} + \frac{\partial g^\alpha}{\partial y^\beta} y^\beta_i - \frac{\partial f^j}{\partial x^i} y^\alpha_j \right) \frac{\partial}{\partial y^\alpha_i}.$$

Now let us define a 1-form on J^1E with values in VE which is useful to distinguish which sections and vector fields on J^1E come from sections and vector fields on E and which do not. It is called the contact structure and it is defined as

$$\Theta(v) = d\pi_{1,0}(v) - ds(d\pi_1(v)) \in V_{s(x)}E,$$

where $v \in T_{j^1_x s}(J^1E)$. The expression in coordinates is given by

$$\Theta = (dy^\alpha - y^\alpha_i dx^i) \otimes \frac{\partial}{\partial y^\alpha}.$$

Proposition 16. *Let $\pi : E \longrightarrow M$ be a fiber bundle and $\bar{s}$ a section of $J^1E \longrightarrow M$. We have that $\bar{s} = j^1 s$ for some section s of $E \longrightarrow M$ if and only if $\bar{s}^*\Theta = 0$.*

Proof. We write the section $\bar{s}$ in coordinates as $\bar{s}(x^i) = (x^i, \bar{s}^\alpha, \bar{s}^\alpha_i)$, where $\bar{s}^\alpha = y^\alpha \circ \bar{s}$ and $\bar{s}^\alpha_i = y^\alpha_i \circ \bar{s}$ as usual. Then

$$\bar{s}^*\Theta = (d\bar{s}^\alpha - \bar{s}^\alpha_i dx^i) \otimes \frac{\partial}{\partial y^\alpha} = \left(\frac{\partial \bar{s}^\alpha}{\partial x^i} dx^i - \bar{s}^\alpha_i dx^i \right) \otimes \frac{\partial}{\partial y^\alpha}$$

and

$$\bar{s}^*\Theta = 0 \Leftrightarrow \frac{\partial \bar{s}^\alpha}{\partial x^i} = \bar{s}^\alpha_i \Leftrightarrow \bar{s} = j^1 s \text{ with } s = \pi_{1,0} \circ \bar{s}.$$

□

The contact structure also characterizes which bundle morphisms defined on the jet bundle are prolongations:

Theorem 3. *Let $\pi : E \longrightarrow M$ and $\pi' : E' \longrightarrow M'$ be fiber bundles and let (F, f) be a bundle morphism between $\pi_1 : J^1E \longrightarrow M$ and $\pi'_1 : J^1E' \longrightarrow M'$ such that f is a diffeomorphism. Then $dF(ker(\Theta_\pi)) \subset ker(\Theta_{\pi'})$ if and only if F is the prolongation of a bundle morphism (f_0, f) between E and E', where Θ_π and $\Theta_{\pi'}$ denote the contact structures on J^1E and J^1E' respectively.*

See [9] for a proof.

Remark 26. In [9] it is shown that the pull-back bundle $\pi^*_{1,0}(TE)$ has a canonical decomposition

$$\pi^*_{1,0}(TE) = \pi^*_{1,0}(VE) \oplus H,$$

and then the contact structure can be defined as

$$pr_1 \circ (d\pi_{1,0}, \tau_{J^1E}) : TJ^1E \longrightarrow \pi^*_{1,0}(VE),$$

where

$$(d\pi_{1,0}, \tau_{J^1E}) : v \in T_{j^1_x s}J^1E \longmapsto (d\pi_{1,0}(v), j^1_x s) \in \pi^*_{1,0}(TE)$$

and pr_1 is the projection onto the first factor with respect to the above decomposition. Recall that $\pi^*_{1,0}(TE) \subset TE \times J^1E$, so it is intuitively clear that the element $j^1_x s$ in $(v, j^1_x s) \in \pi^*_{1,0}(TE)$ provides a way to decompose $v \in T_{s(x)}E$ into a vertical and a horizontal component. If $pr_1 \circ (d\pi_{1,0}, \tau_{J^1E}) : v \in T_{j^1_x s}J^1E \longmapsto (w, j^1_x s)$, with $w \in V_{s(x)}E$, then $w = \Theta(v)$.

Finally we state how the contact structure characterizes the prolongued vector fields (see [9] again). First we need the following definition:

Definition 33 (Infinitesimal symetry)**.** *We say that a vector field X on J^1E is an infinitesimal symmetry if it satisfies that for every vector field $Y \in Ker(\Theta)$, also $[X, Y] \in Ker(\Theta)$.*

Proposition 17. *Let X be a projectable vector field on J^1E with respect to $\pi_{1,0}$. X is an infinitesimal symmetry if and only if X is the prolongation of a vector field on E.*

2.5 Bundles of connections

Let $\pi : P \longrightarrow M$ be a principal bundle. We can define a bundle (the bundle of connections) whose sections will be the connections on P. See [2] or [4].

Let $p \in P$ with $\pi(p) = x$. The short sequence

$$0 \longrightarrow VP \overset{i}{\hookrightarrow} TP \overset{d\pi}{\longrightarrow} TM \longrightarrow 0$$

is exact (recall that π is a submersion).

A splitting h of this sequence, that is, a map $h : TM \longrightarrow TP$ satisfying $d\pi \circ h = Id_{TM}$, gives a connection on P; we take $h(T_{\pi(p)}M)$ as the horizontal subspace in T_pP and then apply dR_g for each $g \in G$ to define the horizontal distribution at each point of the fiber of x. Alternatively we can consider the sequence

$$0 \longrightarrow VP/G \overset{i}{\hookrightarrow} TP/G \overset{d\pi}{\longrightarrow} TM \longrightarrow 0,$$

where the action of G is given by dR_g, and then a splitting gives us directly a connection on P. This sequence is called the Atiyah sequence.

Remark 27. Note that $VP/G \cong \tilde{\mathfrak{g}}$. We already said that, for $A \in \mathfrak{g}$ and $p \in P$, $A \longmapsto A_p^*$ gives an isomorphism and, since $(Ad_{g^{-1}}A)^*_{p\cdot g} = dR_g(A_p^*)$, the map

$$\begin{array}{ccc} \tilde{\mathfrak{g}} & \longrightarrow & VP/G \\ [(p,A)] & \longmapsto & \left[A_p^*\right] \end{array}$$

is well-defined (recall that $[(p,A)] = [(p \cdot g, Ad_{g^{-1}}A)]$ and that A_p^* and $dR_g(A_p^*)$ represent the same class in VP/G).

The following definition is the one given in [2]:

Definition 34. *The bundle of connections of $\pi : P \longrightarrow M$ is the fiber bundle with total space*

$$C(P) = \{\lambda_x : T_xM \longrightarrow (TP/G)_x : \lambda_x \textit{ is linear and } d\pi \circ \lambda_x = Id_{T_xM}, x \in M\}$$

and projection

$$\begin{array}{cccc} \rho : & C(P) & \longrightarrow & M \\ & \lambda_x & \longmapsto & x. \end{array}$$

A section of the bundle of connections gives a splitting of the above sequence and therefore a connection on P.

For each $x \in M$,

$$C(P)_x = \{\lambda \in T_x^*M \otimes (TP/G)_x : d\pi \circ \lambda = Id_{T_xM}\}$$

is an affine subspace of $T_x^*M \otimes (TP/G)_x$ modelled on the vector space

$$\{\lambda \in T_x^*M \otimes (TP/G)_x : d\pi \circ \lambda = 0\} \cong \{\lambda \in T_x^*M \otimes \tilde{\mathfrak{g}}_x\},$$

where we are using the above isomorphism between (VP/G) and $\tilde{\mathfrak{g}}$. Therefore, the bundle of connections $C(P)$ is an affine bundle, an affine subbundle of $T^*M \otimes (TP/G)$, modelled on the vector bundle $T^*M \otimes \tilde{\mathfrak{g}} \longrightarrow M$.

Recall that the first jet bundle J^1P was an affine bundle modelled on the vector bundle $T^*M \otimes VE$. If we consider the action of G on J^1P given by $j^1_x s = j^1_x(R_g \circ s)$, that is, the action given by the first jet prolongation $j^1 R_g$ of (R_g, Id_M), then we get that $J^1P/G \longrightarrow P/G$ is a fiber bundle diffeomorphic to $C(P) \longrightarrow M$.

We said in the previous subsection that a section of $J^1P \longrightarrow P$ gave an Ehresmann connection on P. Then a section of $J^1P/G \longrightarrow P/G = M$ gives a G-invariant Ehresmann connection, that is, a principal connection. We could as well have taken $J^1P/G \longrightarrow M$ as the definition of the bundle of connections and a section of it as the definition of a principal connection, as in [4].

Example 7. If $P = M \times G$ is a trivial bundle, then $C(P) = T^*M \otimes \tilde{\mathfrak{g}}$, for we can take a gobal section of P and define a corresponding zero connection. In fact, we already saw in remark 20 that J^1P can be identified with $T^*M \otimes VP$ in this case, and then $C(P) = J^1P/G$ with $T^*M \otimes \tilde{\mathfrak{g}}$. We will use this later, in the fourth section. Note that since P is trivial, the adjoint bundle must be trivial too and then $C(P)$ can be thought of as $T^*M \otimes \mathfrak{g}$.

Now we introduce coordinates in the bundle of connections. Let $\{B_1, \ldots, B_n\}$ be a basis of $\mathfrak{g}$ and let $\pi^{-1}(U) \cong U \times G$ be a local trivialisation of P, so that we can construct the vector fields $\tilde{B}_1, \ldots, \tilde{B}_m$ introduced in section 2.3. The fibers of the bundle of connections over U can be thought of as $T_xM \otimes \mathfrak{g}$ and for each $A \in C(P)_x$ with $x \in U$ we can write $A = A^\alpha_i dx^i \otimes B_\alpha$. Coordinates (x^i, A^α_i) give a local trivialisation of $C(P)$.

Remark 28. If $A(x) \in \Gamma(C(P))$ is a connection on P with horizontal lift given locally by

$$\frac{\partial}{\partial x^i} \longmapsto \frac{\partial}{\partial x^i} - \Gamma^\alpha_i(x)\tilde{B}_\alpha,$$

then $A^\alpha_i(A(x)) = \Gamma^\alpha_i(x)$.

3 Geometric variational calculus

3.1 Geometric Mechanics

Here we give a brief introduction to the Lagrangian and Hamiltonian formalisms in Mechanics and refer the reader to [7] for a detailed exposition. We start with the Lagrangian formalism.

Let Q be a differentiable manifold which represents the space where particles move; it is called the configuration space. In order to model how particles move we consider the tangent bundle TQ, that is, positions and velocities. Now to describe the evolution of a system we want to define a vector field on TQ whose integral curves will be the trajectories of particles on TQ. We denote by $(q^i, \dot{q}^i)$ the local coordinates on TQ.

In this context a Lagrangian is a smooth map $L : TQ \longrightarrow \mathbb{R}$. Generally L is given by the difference between kinetic and potential energy.

Consider the space of all $\mathcal{C}^2$ curves on Q with fixed endpoints:

$$\mathcal{C}^2(q_0, q_1, [a,b]) = \left\{ c : [a,b] \longrightarrow Q : c \in \mathcal{C}^2([a,b]), c(a) = q_0 \text{ and } c(b) = q_1 \right\}$$

and the functional

$$\begin{array}{rccc} S : & \mathcal{C}^2(q_0, q_1, [a,b]) & \longrightarrow & \mathbb{R} \\ & c & \longmapsto & \int_a^b L(c(t), \dot{c}(t))dt. \end{array}$$

Hamilton's variational principle states that a curve $c \in \mathcal{C}^2(q_0, q_1, [a,b])$ describes the evolution of the system defined by the Lagrangian L if and only if c is a critical point of the functional S, meaning that

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_a^b L(c_\epsilon(t), \dot{c}_\epsilon(t))dt = 0,$$

for all variations $c_\epsilon \in \mathcal{C}^2(q_0, q_1, [a,b])$ of c.

This condition is equivalent to

$$\int_a^b \left(\frac{\partial L}{\partial q^i} \frac{d}{d\epsilon}\bigg|_{\epsilon=0} c_\epsilon(t) + \frac{\partial L}{\partial \dot{q}^i} \frac{d}{dt} \frac{d}{d\epsilon}\bigg|_{\epsilon=0} c_\epsilon(t) \right) dt = 0,$$

using the chain rule, and to

$$\int_a^b \left(\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right) \frac{d}{d\epsilon}\bigg|_{\epsilon=0} c_\epsilon(t) dt = 0,$$

using integration by parts in the second addend and the fact that

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} c_\epsilon(a) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} c_\epsilon(b) = 0.$$

Since we are considering all variations of c, we get the Euler-Lagrange equations:

$$\left\{ \begin{array}{rcl} \frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) & = & 0, \\ \dot{q}^i - \frac{dq^i}{dt} & = & 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rcl} \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i} \frac{d\dot{q}^j}{dt} & = & -\frac{\partial^2 L}{\partial q^j \partial \dot{q}^i} \frac{dq^j}{dt} + \frac{\partial L}{\partial q^i}, \\ \dot{q}^i & = & \frac{dq^i}{dt}. \end{array} \right.$$

If the Lagrangian is regular, that is, $det\left(\frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\right) \neq 0$, then we can write

$$\left\{ \begin{array}{rcl} \frac{d\dot{q}^j}{dt} & = & F^j(q^i, \dot{q}^i), \\ \dot{q}^i & = & \frac{dq^i}{dt} \end{array} \right.$$

and the Euler-Lagrange equations define a second order differential equation on Q:

$$\frac{d^2 c(t)}{dt} = F\left(c(t), \frac{dc(t)}{dt}\right).$$

For the Hamiltonian formalism we work on T^*Q instead of TQ. We define a Hamiltonian to be a smooth map $H : T^*Q \longrightarrow \mathbb{R}$ and denote by (q^i, p_i) the coordinates on T^*Q.

In T^*Q we can define a vector field from the Hamiltonian using the canonical symplectic form Ω on T^*Q (in coordinates $\Omega = dq^i \wedge dp_i$), namely the vector field $X_H \in \mathfrak{X}(T^*Q)$ such that

$$i_{X_H}\Omega = dH.$$

The equations associated to this vector field, that is, $\dot{z} = X_H(z)$, are called Hamilton's equations. In coordinates they are

$$\left\{ \begin{array}{rcl} \frac{dq^i}{dt} & = & \frac{\partial H}{\partial p_i} \\ \frac{dp^i}{dt} & = & -\frac{\partial H}{\partial q^i}. \end{array} \right.$$

Given a Lagrangian L, we can relate both viewpoints through the Legendre transformation $\mathbb{F}L : TQ \longrightarrow T^*Q$ defined by

$$\mathbb{F}L(v)(w) = \left.\frac{d}{ds}\right|_{s=0} L(v + sw),$$

where $v, w \in T_pQ$. The expression in coordinates is $(q^i, \dot{q}^i) \longmapsto (q^i, \frac{\partial L}{\partial \dot{q}^i})$.

The condition we used to define a regular Lagrangian is equivalent to $\mathbb{F}L$ being locally a diffeomorphism. If $\mathbb{F}L$ is a global diffeomorphism then we say that the Lagrangian L is hyperregular.

The following result states that, under some circumstances, Euler-Lagrange equations and Hamilton's equations are equivalent:

Proposition 18. *Let L be a hyperregular Lagrangian and define an associated Hamiltonian $H = E \circ \mathbb{F}L^{-1} \in \mathcal{F}(T^*Q)$, where*

$$\begin{array}{rccc} E: & TQ & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & \mathbb{F}L(v)(v) - L(v) \end{array}$$

is called the energy of L. If Z is the vector field on TQ associated to the Lagrangian L and X_H is the vector field on T^*Q associated to the Hamiltonian H, then

$$X_H \circ \mathbb{F}L = Z.$$

If $c(t)$ and $d(t)$ are integral curves of Z and X_H respectively such that $\mathbb{F}L(c(0)) = d(0)$, then $\mathbb{F}L(c(t)) = d(t)$ and they project to the same curve on the base space Q, that is $\tau_Q(c(t)) = \tau_Q^*(d(t))$.

Remark 29. If we consider the pull-back form $\Omega_L = \mathbb{F}L^*\Omega$ on TQ, then the vector field $Z \in \mathfrak{X}(TQ)$ in the above proposition is the unique vector field on TQ such that

$$i_Z \Omega_L = dE$$

(writing Ω_L in coordinates shows that it is nondegenerate if and only if L is regular). It is denoted by X_E.

3.2 Field theories

In this section we introduce the Lagrangian formalism for an arbitrary bundle $E \longrightarrow M$ (in contrast with $Q \times \mathbb{R} \longrightarrow \mathbb{R}$). The role of the tangent bundle will be played by the first jet bundle:

Definition 35 (Lagrangian density)**.** *Let (E, π, M) be a fiber bundle with M orientable and compact. A Lagrangian density is a bundle morphism*

$$\mathcal{L} : J^1E \longrightarrow \bigwedge^n T^*M.$$

If a volume form v is given on M then we can write $\mathcal{L} = Lv$, where $L \in \mathcal{F}(J^1E)$.

Now consider the functional

$$\begin{array}{rccl} S: & \Gamma(E) & \longrightarrow & \mathbb{R} \\ & s & \longmapsto & \int_M \mathcal{L} \circ j^1 s = \int_M \left(L \circ j^1 s\right) v, \end{array}$$

where s is a section of $E \longrightarrow M$. We want to minimize this functional in the following sense:

Definition 36. *A section s of $E \longrightarrow M$ is said to be an extremal of L if*

$$\left.\frac{d}{dt}\right|_{t=0} \int_M \left(L \circ j^1(s_t)\right) v = 0,$$

for all variations of s, where a variation of s is a smooth map

$$\begin{array}{ccc} M \times \mathbb{R} & \longrightarrow & E \\ (x,t) & \longmapsto & s(x,t) = s_t(x) \end{array}$$

such that $s_0(x) = s(x)$ for all $x \in M$.

Proposition 19. *A section s of $E \longrightarrow M$ is extremal if and only if the Euler-Lagrange equations*

$$\frac{\partial L}{\partial y^\alpha} \circ j^1 s - \frac{\partial}{\partial x^i}\left(\frac{\partial L}{\partial y_i^\alpha} \circ j^1 s\right) = 0$$

are satisfied for all α, for adapted coordinates such that $v = d^n x$.

This is proved by doing analogous calculations to the case of Mechanics in 3.1 and applying Stokes theorem.

Definition 37 (Poincaré-Cartan form)**.** *Given a Lagrangian L, the associated Poincaré-Cartan form is the n-form on J^1E with expression in coordinates given by*

$$\Theta_L = \frac{\partial L}{\partial y_i^\alpha} dy^\alpha \wedge d^{n-1}x_i + \left(L - \frac{\partial L}{\partial y_i^\alpha} y_i^\alpha\right) d^n x.$$

It can be shown that this expression does not depend on the choice of coordinates; an intrinsic definition is given in the following subsection.

Note that the n-form $\mathcal{L} \circ j^1 s$ on M can be rewritten in terms of the Poincaré-Cartan form Θ_L as

$$\mathcal{L} \circ j^1 s = (j^1 s)^* \Theta_L.$$

Indeed,

$$(j^1 s)^* \Theta_L = (j^1 s)^* \left(\frac{\partial L}{\partial y_i^\alpha} dy^\alpha \wedge d^{n-1}x_i + \left(L - \frac{\partial L}{\partial y_i^\alpha} y_i^\alpha\right) d^n x\right)$$

$$= \frac{\partial L}{\partial y_i^\alpha} \circ j^1 s \; ds^\alpha \wedge d^{n-1}x_i + \left(L \circ j^1 s - s_i^\alpha \frac{\partial L}{\partial y_i^\alpha} \circ j^1 s\right) d^n x = L \circ j^1 s \; d^n x.$$

The $(n+1)$-form $\Omega_L = -d\Theta_L$ gives another characterization for critical sections:

Proposition 20. *A section s of $E \longrightarrow M$ is extremal if and only if*

$$(j^1 s)^*(i_X \Omega_L) = 0,$$

for all vertical vector fields X on J^1E (vertical with respect to π_1).

This result can be shown by direct computation and using proposition 19.

Remark 30. The last proposition implies that the Euler-Lagrange equations are independent of the choice of coordinates.

Example 8. If we consider the trivial bundle $\mathbb{R} \times Q \longrightarrow \mathbb{R}$ for some smooth manifold Q then $J^1(\mathbb{R} \times Q) \cong \mathbb{R} \times TQ$ and we recover the situation described in 3.1. Note that we were looking for curves on Q that would describe the evolution of a system, that is, sections of $\mathbb{R} \times Q \longrightarrow \mathbb{R}$.

If we denote by $(t, q^i, \dot{q}^i)$ coordinates on $\mathbb{R} \times TQ$ then the Poincaré-Cartan form is written as

$$\Theta_L = \frac{\partial L}{\partial \dot{q}^i} dq^i + \left(L - \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i \right) dt = \frac{\partial L}{\partial \dot{q}^i} dq^i - E(q^i, \dot{q}^i) dt$$

and Euler-Lagrange equations as

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) = 0.$$

3.3 Hamiltonian formulation, the polysymplectic bundle and brackets

Now we introduce the analogues of the cotangent bundle for arbitrary bundles.

A first attempt to develop the Hamiltonian formalism could be to work on the dual jet bundle:

Definition 38 (Dual jet bundle)**.** *Let $\pi : E \longrightarrow M$ be an arbitrary fiber bundle. The dual jet bundle is a fiber bundle over E, $(J^1E)^* \longrightarrow E$, with fibers given by the affine duals of the fibers of the jet bundle $J^1E \longrightarrow E$ (which are affine spaces). We will take as affine dual the set of affine maps taking values in $\bigwedge^n T^*M$.*

Given adapted coordinates (x^i, y^a) on the base space E, we define adapted coordinates on $(J^1E)^*$ by (x^i, y^a, p^i_a, p), where if (x^i, y^a, y^a_i) are coordinates on the fiber of J^1E over (x^i, y^a) then p^i_a and p define all possible affine maps on this fiber by

$$y^a_i \longmapsto (p^i_a y^a_i + p) dx^1 \wedge \cdots \wedge dx^n.$$

One advantage that the dual jet bundle $(J^1E)^*$ presents is that it can be endowed with a canonical multisymplectic form. This is accomplished by identifying it with a subbundle Z of $\bigwedge^n T^*E$, namely the subbundle consisting of all n-covectors that vanish after contraction by two vertical vectors. The fiber of Z over $y \in E$ is given by

$$Z_y = \left\{ z \in \left(\bigwedge^n T^*E \right)_y : i_u i_v z = 0, \text{ for all } u, v \in V_y E \right\}.$$

Note that the n-covectors that vanish after contraction by two vertical vectors are the ones that have just one factor dy^a or none, so for each $z \in Z$ we can write

$$z = p dx^1 \wedge \cdots \wedge dx^n + p^i_a dy^a \wedge d^{n-1} x_i.$$

Now we relate Z and $(J^1E)^*$ through the fiber map

$$\begin{array}{rccc} \Phi : & Z & \longrightarrow & (J^1E)^* \\ & z & \longmapsto & \Phi(z), \end{array}$$

where $\Phi(z)(j^1_x s) = s^* z \in \bigwedge^n T^*M$ and, if $z \in Z_y$, then $\Phi(z) \in \left((J^1E)^* \right)_y$, that is, $s(x) = y$. Φ is a vector bundle isomorphism. Note that if we write $s(x^i) = (x^i, s^a)$, then $s^*z = pdx^n + p^i_a ds^a \wedge d^{n-1}x_i = pdx^n + p^i_a \frac{\partial s^a}{\partial x^j} dx^j \wedge d^{n-1}x_i = \left(p + p^i_a \frac{\partial s^a}{\partial x^i}\right) d^n x = \left(p + p^i_a y^a_i\right) d^n x$, that is, $\Phi(z)$ is the map $y^a_i \mapsto (p + p^i_a y^a_i) d^n x$, so the expression of Φ in coordinates is

$$(x^i, y^a, p^i_a, p) \longmapsto (x^i, y^a, p^i_a, p).$$

In $\bigwedge^n T^*E$ we can define a canonical n-form Θ, which can be pull-backed to Z and then to $(J^1E)^*$. It is given by

$$\Theta(W)(X_1, \ldots, X_n) = W(d\pi(X_1), \ldots, d\pi(X_n)),$$

where $W \in \bigwedge^n T^*E$, $X_1, \ldots, X_n \in T_W\left(\bigwedge^n T^*E\right)$, $\pi : \bigwedge^n T^*E \longrightarrow E$ is the projection, and so $d\pi : T(\bigwedge^n T^*E) \longrightarrow TE$.

Now if $i : Z \hookrightarrow \bigwedge^n T^*E$ denotes the inclusion then $(\Phi^{-1})^* i^* \Theta$ gives us a canonical n-form on $(J^1E)^*$ which we will also denote by Θ. The canonical multisymplectic $(n+1)$-form is defined to be $\Omega = -d\Theta$.

In local coordinates one can write

$$\Theta = p^i_a dy^a \wedge d^{n-1}x_i + pd^n x,$$

$$\Omega = dy^a \wedge dp^i_a \wedge d^{n-1}x_i - dp \wedge d^n x.$$

Now we want a bridge between J^1E and $(J^1E)^*$. Given a Lagrangian density $\mathcal{L} : J^1E \longrightarrow \bigwedge^n T^*M$ we can define a fiber bundle morphism as follows:

Definition 39 (Legendre transformation)**.** *The Legendre transformation is the bundle morphism over E defined as*

$$\begin{array}{rccc} \mathbb{F}\mathcal{L}: & J^1E & \longrightarrow & (J^1E)^* \\ & j^1_x s & \longmapsto & \mathbb{F}\mathcal{L}(j^1_x s), \end{array}$$

where

$$\mathbb{F}\mathcal{L}(j^1_x s)(j^1_x s') = \mathcal{L}(j^1_x s) + \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathcal{L}(j^1_x s + \epsilon(j^1_x s' - j^1_x s)).$$

Remark 31. We define $\mathbb{F}\mathcal{L}$ to be a gauge bundle morphism over E. This means that the first $n+m$ coordinates of $j^1_x s$ and $j^1_x s'$ are equal, so we are taking derivatives only in the vertical directions and $\mathbb{F}\mathcal{L}$ is just the first-order vertical Taylor approximation to $\mathcal{L}$.

In coordinates the Legendre transform gives

$$p^i_a = \frac{\partial L}{\partial y^a_i},$$

$$p = L - \frac{\partial L}{\partial y^a_i} y^a_i.$$

Remark 32. The dimensions of J^1E and $(J^1E)^*$ are different, so we cannot expect the Lengendre transformation to be a diffeomorphism.

Remark 33. The n-form $\mathbb{F}\mathcal{L}^*\Theta$ and the $(n+1)$-form $\mathbb{F}\mathcal{L}^*\Omega$ on J^1E coincide respectively with the forms Θ_L and Ω_L defined in the previous subsection.

We will now introduce an alternative fiber bundle to develop the Hamiltonian formalism which will have the same dimension as J^1E:

Definition 40 (Polysymplectic bundle)**.** *Let $\pi : E \longrightarrow M$ be a fiber bundle. The polysymplectic bundle $\Pi \longrightarrow E$ is defined as*

$$\Pi = \pi^*(TM) \otimes_E V^*E \otimes_E \pi^* \left(\bigwedge^n T^*M \right).$$

We will usually write

$$\Pi = TM \otimes V^*E \otimes \bigwedge^n T^*M,$$

but keep in mind that it is a bundle over E. Note that indeed the polysymplectic bundle has the same dimension as the jet bundle.

Remark 34. The polysymplectic bundle is the linear version of the dual jet bundle. Note that $TM \otimes V^*E$ is the dual bundle of $T^*M \otimes VE$, precisely the vector bundle on which J^1E is modelled. So, instead of taking the bundle with fibers the set of affine maps on the fibers of J^1E, we take the bundle with fibers the set of linear maps on the vector spaces on which the fibers of J^1E are modelled.

Coordinates on Π will be written as (x^i, y^a, π_a^i), so that elements in Π are written as

$$\pi_a^i \frac{\partial}{\partial x^i} \otimes dy^a \otimes d^n x.$$

Definition 41 (Linear Legendre transformation)**.** *The linear Legendre transformation is defined as*

$$\begin{array}{rccc} \widehat{\mathbb{F}\mathcal{L}} : & J^1E & \longrightarrow & \Pi \\ & j_x^1 s & \longmapsto & \widehat{\mathbb{F}\mathcal{L}}(j_x^1 s), \end{array}$$

where

$$\widehat{\mathbb{F}\mathcal{L}}(j_x^1 s)(\omega) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \mathcal{L}(j_x^1 s + \epsilon\omega),$$

for any $\omega \in T^*M \otimes VE$.

In coordinates it gives

$$\pi_a^i = \frac{\partial L}{\partial y_i^a}.$$

Let $k : (J^1E)^* \longrightarrow \Pi$ denote the bundle morphism which assigns to each affine map the corresponding linearization. Since we had a canonical $(n+1)$-form on $(J^1E)^*$, it is enough to have a section of $k : (J^1E)^* \longrightarrow \Pi$ in order to write Hamilton equations on Π, so let us define a Hamiltonian system as a pair (Π, δ) where δ is a section of k. Then we can take the pull-backs of Ω and Θ to Π, which will be denoted by Ω_δ and Θ_δ respectively, and write the equation

$$\pi^* \left(i_X \Omega_\delta \right) = 0,$$

where X is any vertical vector field on Π and π is a section of $\Pi \longrightarrow M$. If the equation is satisfied then the section π is said to be a solution of the Hamiltonian system.

Furthermore, if we have a connection A on E, then one can equivalently define a Hamiltonian system as a pair $(\Pi, \mathcal{H})$, where $\mathcal{H}$ is a smooth map

$$\mathcal{H} : \Pi \longrightarrow \bigwedge^n T^*M,$$

called a Hamiltonian density. We will write $\mathcal{H} = -H d^n x$. To relate both definitions note that given two sections δ_1, δ_2 of $k : (J^1E)^* \longrightarrow \Pi$, the difference $\Theta_{\delta_1} - \Theta_{\delta_2}$ is a Hamiltonian density. Indeed, if we write

$$\delta_j(x^i, y^a, \pi^i_a) = (x^i, y^a, \pi^i_a, -H_{\delta_j}),\ j = 1, 2,$$

then

$$\Theta_{\delta_1} - \Theta_{\delta_2} = \pi^i_a dy^a \wedge d^{n-1}x_i - H_{\delta_1} d^n x - \pi^i_a dy^a \wedge d^{n-1}x_i + H_{\delta_2} d^n x = (H_{\delta_2} - H_{\delta_1}) d^n x.$$

Now a section of $(J^1E)^* \longrightarrow \Pi$ can be defined from an Ehresmann connection $A : TE \longrightarrow VE$ on $E \longrightarrow M$ as

$$\begin{array}{rccc} \delta_A|_{\Pi_y} : & T_xM \otimes V^*_y E \otimes \bigwedge^n T^*_x M & \longrightarrow & Z_y \cong (J^1E)^*_y \\ & w \otimes \xi \otimes d^n x & \longmapsto & (\xi \circ A) \wedge i_w d^n x. \end{array}$$

In coordinates we write $w = w^i \frac{\partial}{\partial x^i}$, $\xi = \xi_\alpha dy^\alpha$ and $w \otimes \xi \otimes d^n x = \pi^i_\alpha \frac{\partial}{\partial x^i} \otimes dy^\alpha \otimes d^n x$. Then $i_w d^n x = w^i d^{n-1} x_i$ and, if the connection is locally given by

$$\frac{\partial}{\partial x^i} \longmapsto \frac{\partial}{\partial x^i} - \Gamma^\alpha_i \frac{\partial}{\partial y^\alpha},$$

then $A = \Gamma^\alpha_i dx^i \otimes \frac{\partial}{\partial y^\alpha} + dy^\alpha \otimes \frac{\partial}{\partial y^\alpha}$, for $A\left(\frac{\partial}{\partial y^\alpha}\right) = \frac{\partial}{\partial y^\alpha}$ and $A\left(\frac{\partial}{\partial x^i} - \Gamma^\alpha_i \frac{\partial}{\partial y^\alpha}\right) = 0$ implies $A\left(\frac{\partial}{\partial x^i}\right) = \Gamma^\alpha_i \frac{\partial}{\partial y^\alpha}$.

Hence the local expression of $\xi \circ A$ is

$$\begin{aligned} \xi_a dy^a \circ \left(\Gamma^\alpha_i dx^i \otimes \frac{\partial}{\partial y^\alpha} + dy^\alpha \otimes \frac{\partial}{\partial y^\alpha}\right) & \\ = \xi_\alpha \Gamma^\alpha_i dx^i + \xi_\alpha dy^\alpha & \end{aligned}$$

and finally $(\xi \circ A) \wedge i_w v$ is locally written as

$$w^i \xi_\alpha \Gamma^\alpha_i d^n x + w^i \xi_\alpha dy^\alpha \wedge d^{n-1} x_i = \pi^i_\alpha \Gamma^\alpha_i d^n x + \pi^i_\alpha dy^\alpha \wedge d^{n-1} x_i.$$

So if we have a Hamiltonian system (Π, δ) and we have a connection A on E, then define $(\Pi, \mathcal{H})$ with $-H d^n x = \mathcal{H}$ and $H = H_\delta - H_{\delta_A}$ and conversely, given $\mathcal{H}$ we can recover $\Theta_\delta = \Theta_{\delta_A} + \mathcal{H}$ and Θ_δ defines δ.

Having a connection A on E and taking the definition of a Hamiltonian system as $(\Pi, \mathcal{H})$, we can write Hamilton's equations as

$$\pi^*(i_X d(\Theta_{\delta_A} + \mathcal{H})) = 0,$$

where X is a vertical vector field and π is a section of $\Pi \longrightarrow M$.

Remark 35. We can also get a Hamiltonian density $\mathcal{H}^A_{\mathcal{L}}$ from a hyperregular Lagrangian density $\mathcal{L}$, where hyperregular means that $\widehat{\mathbb{F}\mathcal{L}}$ is a diffeomorphism. We just need to take $\delta = \mathbb{F}\mathcal{L} \circ \widehat{\mathbb{F}\mathcal{L}}^{-1}$ and define $\mathcal{H}^A_{\mathcal{L}} = \Theta_\delta - \Theta_{\delta_A}$. The pair $(\Pi, \mathcal{H}^A_{\mathcal{L}})$ is called the Hamiltonian system associated to $\mathcal{L}$ and A.

Theorem 4. *Let $\mathcal{L} : J^1E \longrightarrow \bigwedge^n T^*M$ be a hyperregular Lagrangian. A section $s : M \longrightarrow E$ is a solution of the variational problem defined by $\mathcal{L}$ if and only if the section $\pi = \widehat{\mathbb{F}\mathcal{L}} \circ j^1 s$ of $\Pi \longrightarrow M$ is a solution of the Hamiltonian system $(\Pi, \mathcal{H}^A_{\mathcal{L}})$.*

$$\begin{tikzcd}
J^1E \arrow[r, "\widehat{\mathbb{F}\mathcal{L}}"] \arrow[d] & \Pi \arrow[dd] \\
E \arrow[d] & \\
M \arrow[u, bend right, "s"'] \arrow[uu, bend left, "j^1 s"] & M \arrow[uu, bend right, "\pi"']
\end{tikzcd}$$

See [8] for a proof.

Proposition 21. *The expression of the equation* $\pi^*(i_X d(\Theta_{\delta_A} + \mathcal{H})) = 0$ *in coordinates is*

$$\left(\frac{\partial H}{\partial \pi^i_\alpha}\right)_\pi = \left(\frac{\partial y^\alpha}{\partial x^i} + \Gamma^\alpha_i\right)_\pi, \quad \left(\frac{\partial H}{\partial y^\alpha}\right)_\pi = -\left(\frac{\partial \pi^j_\alpha}{\partial x^j} - \frac{\partial \Gamma^\beta_i}{\partial y^\alpha}\pi^i_\beta\right)_\pi.$$

They are called Hamilton-Cartan equations.

Proof. Recall from the calculation above that

$$\Theta_{\delta_A} = \pi^i_\alpha \Gamma^\alpha_i d^n x + \pi^i_\alpha dy^\alpha \wedge d^{n-1}x_i.$$

Then

$$\Theta_{\delta_A} + \mathcal{H}^A = (\pi^i_\alpha \Gamma^\alpha_i - H)d^n x + \pi^i_\alpha dy^\alpha \wedge d^{n-1}x_i$$

and

$$d(\Theta_{\delta_A} + \mathcal{H}^A) = \left(\pi^i_\alpha \frac{\partial \Gamma^\alpha_i}{\partial y^a} - \frac{\partial H}{\partial y^a}\right) dy^a \wedge d^n x$$

$$+ \left(\frac{\partial \pi^i_\alpha}{\partial \pi^j_a}\Gamma^\alpha_i - \frac{\partial H}{\partial \pi^j_a}\right) d\pi^j_a \wedge d^n x + \frac{\partial \pi^i_\alpha}{\partial \pi^j_a} d\pi^j_a \wedge dy^\alpha \wedge d^{n-1}x_i$$

$$= \left(\pi^i_\alpha \frac{\partial \Gamma^\alpha_i}{\partial y^a} - \frac{\partial H}{\partial y^a}\right) dy^a \wedge d^n x + \left(\Gamma^a_j - \frac{\partial H}{\partial \pi^j_a}\right) d\pi^j_a \wedge d^n x + d\pi^j_a \wedge dy^a \wedge d^{n-1}x_j.$$

Taking a vertical vector field $X = X^\gamma \frac{\partial}{\partial y^\gamma} + X^k_\beta \frac{\partial}{\pi^k_\beta}$ then

$$i_X d(\Theta_{\delta_A} + \mathcal{H}^A) = X^a \left(\pi^i_\alpha \frac{\partial \Gamma^\alpha_i}{\partial y^a} - \frac{\partial H}{\partial y^a}\right) d^n x$$

$$+ X^j_a \left(\Gamma^a_j - \frac{\partial H}{\partial \pi^j_a}\right) d^n x + X^j_a dy^a \wedge d^{n-1}x_j - X^a d\pi^j_a \wedge d^{n-1}x_j,$$

and writing a section s of $\Pi \longrightarrow M$ as (x^i, s^a, s^i_a) we get

$$s^*(i_X d(\Theta_{\delta_A} + \mathcal{H}^A)) = X^a \left(\pi^i_\alpha \frac{\partial \Gamma^\alpha_i}{\partial y^a} - \frac{\partial H}{\partial y^a}\right)_s d^n x$$

$$+ X^j_a \left(\Gamma^a_j - \frac{\partial H}{\partial \pi^j_a}\right)_s d^n x + X^j_a ds^a \wedge d^{n-1}x_j - X^a ds^j_a \wedge d^{n-1}x_j$$

$$= X^a \left(\pi^i_\alpha \frac{\partial \Gamma^\alpha_i}{\partial y^a} - \frac{\partial H}{\partial y^a}\right)_s d^n x + X^j_a \left(\Gamma^a_j - \frac{\partial H}{\partial \pi^j_a}\right)_s d^n x + X^j_a \frac{\partial s^a}{\partial x^j} d^n x - X^a \frac{\partial s^j_a}{\partial x^j} d^n x.$$

Therefore

$$\left(\frac{\partial H}{\partial y^a}\right)_s = \left(-\frac{\partial \pi_a^j}{\partial x^j} + \pi_\alpha^i \frac{\partial \Gamma_i^\alpha}{\partial y^a}\right)_s ,$$

$$\left(\frac{\partial H}{\partial \pi_a^j}\right)_s = \left(\frac{\partial y^a}{\partial x^j} + \Gamma_j^a\right)_s .$$

□

Remark 36. Note that taking $\mathbb{R} \times Q \longrightarrow \mathbb{R}$ and the trivial connection $\frac{\partial}{\partial t} \longmapsto \frac{\partial}{\partial t}$, we recover Hamilton's equations from section 3.1.

We will now give some definitions and state some results for which we would like to write analogues in the Yang-Mills case, where we will not be working in the whole Π.

Definition 42 (Horizontal forms)**.** *An r-form on $(J^1E)^*$ is said to be horizontal if it vanishes after contraction with any vertical vector field with respect to $(J^1E)^* \longrightarrow M$.*

Definition 43 (Poisson forms)**.** *A horizontal r-form F on $(J^1E)^*$ is said to be Poisson if there exists an $(n-r)$-multivector field $X_F \in \Gamma\left(\bigwedge^{n-r} T(J^1E)^*\right)$ such that $i_{X_F}\Omega = dF$.*

Proposition 22. *Every (horizontal) Poisson r-form F on $(J^1E)^*$ with $r > 0$ is projectable to Π, that is, the expression in coordinates does not depend on p.*

See [1] for a proof of this result (and also for the following in this section). In the proof it is seen that a multivector field X such that $i_X\Omega = dF$ must be vertical, meaning that it contains no elements of the form $X^{i_1\cdots i_s}(\partial/\partial x^{i_1}) \wedge \cdots \wedge (\partial/\partial x^{i_s})$.

Proposition 23. *Every function $F : (J^1E)^* \longrightarrow \mathbb{R}$ which does not depend on the affine coordinate, alternatively every function $F : \Pi \longrightarrow \mathbb{R}$, is a Poisson 0-form.*

Proof. Take

$$X_F = -\frac{\partial F}{\partial \pi_\alpha^i}\frac{\partial}{\partial y^\alpha} \wedge v_i^* + \frac{\partial F}{\partial y^\alpha}\frac{\partial}{\partial \pi_\alpha^i} \wedge v_i^* - \frac{\partial F}{\partial x^i}\frac{\partial}{\partial p} \wedge v_i^*,$$

where $v^* = (\partial/\partial x^1) \wedge \cdots \wedge (\partial/\partial x^n)$ and $v_i^* = i_{dx^i}v^*$. □

Poisson $(n-1)$-forms will be of interest for us since they characterize the solutions of a Hamiltonian system (see proposition 25 below). The local expression of a Poisson $(n-1)$-form F is given by

$$F = (\pi_\alpha^i X^\alpha + g^i)d^{n-1}x_i + z,$$

where $X^\alpha, g^i \in \mathcal{F}(E)$ and z is a horizontal closed $(n-1)$-form on Π.

We will see more details later in the Yang-Mills case.

Poisson $(n-1)$-forms can be written without using coordinates as

$$F = \theta_X + \pi^* w + Z,$$

where X is a vertical vector field on $E \longrightarrow M$, θ_X is the map

$$\theta_X : \Pi = \pi^*(TM) \otimes V^*E \otimes \pi^*\left(\bigwedge^n T^*M\right) \longrightarrow \pi^*\left(\bigwedge^{n-1} T^*M\right)$$

which contracts the vertical form component part with the vertical vector field and then contracts vectors with n-forms giving $(n-1)$-forms, w is a horizontal $(n-1)$-form on E and Z is a closed horizontal $(n-1)$-form on Π.

Definition 44 (Poisson bracket)**.** *Let F and H be (horizontal) Poisson forms on $(J^1E)^*$ of degrees r and s respectively and X_F and X_H denote associated multivector fields. We define the Poisson bracket to be the $(r+s+1-n)$-form*

$$\{F, H\} = -i_{X_F} i_{X_H} \Omega.$$

If F is a (horizontal) Poisson $(n-1)$-form and $H \in \mathcal{F}(\Pi)$ (so that it is a Poisson 0-form) then the local expression for the Poisson bracket is

$$\{F, H\} = \frac{\partial F^i}{\partial y^a} \frac{\partial H}{\partial \pi^i_a} - \frac{\partial F^i}{\partial \pi^i_a} \frac{\partial H}{\partial y^a}.$$

Proposition 24. *Given an Ehresmann connection A on $E \longrightarrow M$ and a Riemannian connection on M, there is a canonical connection on $\Pi \longrightarrow M$ with horizontal lift given by*

$$\frac{\partial}{\partial x^i} \longmapsto \frac{\partial}{\partial x^i} - \Gamma^\alpha_i \frac{\partial}{\partial y^\alpha} - \left(-\frac{\partial \Gamma^\beta_i}{\partial y^\alpha} \pi^j_\beta + \Gamma^j_{ik} \pi^k_\alpha - \Gamma^k_{ik} \pi^j_\alpha \right) \frac{\partial}{\partial \pi^j_\alpha},$$

where Γ^α_i are the coefficients of the connection on E and Γ^i_{jk} are the Christoffel symbols of the connection on M.

Proposition 25. *A section π of the bundle $\Pi \longrightarrow M$ is a solution of the Hamiltonian system (Π, A, H) if and only if for every horizontal Poisson $(n-1)$-form F we have*

$$\{F, H\}\, d^n x \circ \pi = d(\pi^* F) - (d^h F) \circ \pi,$$

where $d^h F$ is the horizontal differential of F with respect to the connection on Π given in the above proposition.

If we have a principal bundle $P \longrightarrow M$ with structure group G, then

$$\frac{\Pi}{G} \cong TM \otimes \tilde{\mathfrak{g}}^* \otimes \bigwedge^n T^* M$$

(recall from the subsection on bundles of connections that the action of G on VP and TM is given by dR_g, and hence $VP/G \cong \tilde{\mathfrak{g}}$ and $TM/G \cong TM$).

If we ask a (horizontal) Poisson $(n-1)$-form F on Π to be G-invariant then

$$F = \theta_X + \pi^* w + Z,$$

where X is a G-invariant vertical vector field, w is an $(n-1)$-form on M and Z is a closed horizontal G-invariant $(n-1)$-form on Π.

Remark 37. A vector field $X \in \mathfrak{X}(P)$ is G-invariant if and only if its flow Φ_t is an automorphism of P for all $t \in \mathbb{R}$ and it is G-invariant and vertical if and only if its flow Φ_t is a gauge transformation for all $t \in \mathbb{R}$. These last vector fields are called gauge vector fields and its set is denoted by $gauP$. Sections of $TP/G \longrightarrow M$ can be identified with G-invariant vector fields and sections of $\tilde{\mathfrak{g}} \cong VP/G \longrightarrow M$ with gauge vector fields.

Let f be an $(n-1)$-form on Π/G which is the projection of a G-invariant (horizontal) Poisson $(n-1)$-form on Π. Using the identification $gauP \cong \Gamma(\tilde{\mathfrak{g}})$ we can write

$$f = \theta_\xi + \pi^* w + Z$$

for some $\xi \in \Gamma(\tilde{\mathfrak{g}})$, where θ_ξ is defined in an analogous way to θ_X, w is an $(n-1)$-form on M and Z is a closed horizontal $(n-1)$-form on Π/G.

Definition 45 (Vertical derivative). *The vertical derivative of $h \in \mathcal{F}(\Pi/G)$ is the fiber bundle morphism*

$$\begin{array}{rccc} \frac{\delta h}{\delta \mu}: & TM \otimes \tilde{\mathfrak{g}}^* \otimes \bigwedge^n T^*M & \longrightarrow & T^*M \otimes \tilde{\mathfrak{g}} \otimes \bigwedge^n TM \\ & \mu & \longmapsto & \frac{\delta h}{\delta \mu}(\mu) \end{array}$$

defined by

$$\frac{\delta h}{\delta \mu}(\mu)(\tau) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} h(\mu + \epsilon\tau),$$

with $\mu, \tau \in T_xM \otimes \tilde{\mathfrak{g}}_x^* \otimes \bigwedge^n T_x^*M$.

Definition 46 (Lie-Poisson brackets). *Let f be an $(n-1)$-form on Π/G which is the projection of a G-invariant horizontal Poisson $(n-1)$-form on Π and let h be a function on Π/G. We define the Lie-Poisson brackets on Π/G as*

$$\{f, h\}_{\pm}(\mu) = \pm \left\langle \mu, \left[\xi, \frac{\delta h}{\delta \mu}(\mu)\right]\right\rangle,$$

where ξ is the section of $\tilde{\mathfrak{g}} \longrightarrow M$ such that $f = \theta_\xi + \pi^*\omega$, $\mu \in T_xM \otimes \tilde{\mathfrak{g}}_x^* \otimes \bigwedge^n T_x^*M$, $[,]$ takes the Lie bracket on the corresponding component giving an element in $(\Pi/G)_x^*$ and $\langle , \rangle$ is the pairing between Π/G and $(\Pi/G)^*$.

Theorem 5. *Let (P, π, M, G) be a principal fiber bundle, f an $(n-1)$-form on Π/G which is the projection of a G-invariant horizontal Poisson $(n-1)$-form on Π and h a function on Π/G. Then*

$$\{p^* f, p^* h\} = p^* \{f, h\}_+$$

where $p : \Pi \longrightarrow \Pi/G$ is the projection.

Definition 47 (Divergence). *Let $P \longrightarrow M$ be a principal fiber bundle with a connection A and let V be an associated vector bundle. The divergence with respect to A is defined to be the only $\mathbb{R}$-linear operator*

$$div^A : \Gamma(TM \otimes V) \longrightarrow \Gamma(V)$$

such that

$$div \langle X, \eta \rangle = \langle div^A X, \eta \rangle + \langle X, \nabla^A \eta \rangle,$$

for all $X \in \Gamma(TM \otimes V)$ and $\eta \in \Gamma(V^*)$, where div is the usual divergence defined on vector fields and ∇^A denotes the covariant derivative defined by the connection induced by A on V^*. Here we see ∇^A as $\nabla^A : \Gamma(V^*) \longrightarrow \Gamma(T^*M \otimes V^*)$ rather than $\nabla^A : \mathfrak{X}(M) \times \Gamma(V^*) \longrightarrow \Gamma(V^*)$.

Remark 38. The usual divergence of a vector field $X \in \mathfrak{X}(M)$ with respect to an n-form $\omega \in \Omega^n(M)$ is defined to be the function $f = divX$ such that $\mathcal{L}_X \omega = f\omega$; it is a particular case of the above definition taking $E = M \times \mathbb{R}$.

Theorem 6. *Let (P, π, M, G) be a principal fiber bundle with $n =$dim(M) and v a volume form on M. Let A be a connection on $P \longrightarrow M$ and $\mathcal{H}$ a G-invariant Hamiltonian density on Π. We write h for the dropped density to Π/G and for each section π of $\Pi \longrightarrow M$ we write μ for the reduced section $p \circ \pi$.*

Then the following assertions are equivalent:

1. *for every horizontal Poisson $(n-1)$-form F on Π*
$$\pi^* \{F, H\} v = d(\pi^* F) - d^h F \circ \pi,$$
2. *the section π satisfies the Hamilton-Cartan equations,*
3. *for every dropped horizontal Poisson $(n-1)$-form f on Π/G*
$$\mu^* \{f, h\}_+ v = d(\mu^* f) - d^h f \circ \mu,$$
4. *the section μ satisfies*
$$div^A \mu = ad^*_{\frac{\delta h}{\delta \mu}(\mu)} \mu.$$

$$\begin{tikzcd}
\Pi \arrow[dr, "p"] \arrow[dd] & \\
& \Pi/G \cong TM \otimes \tilde{\mathfrak{g}}^* \arrow[dl] \\
M \arrow[uu, bend left, "\pi"] \arrow[ur, bend right, "\mu"'] &
\end{tikzcd}$$

Remark 39. The operator

$$\begin{array}{rccc} ad^* : & \mathfrak{g} \times \mathfrak{g}^* & \longrightarrow & \mathfrak{g}^* \\ & (v, \alpha) & \longmapsto & ad^*_v \alpha \end{array}$$

is defined as $ad^*_v \alpha(u) = \alpha([u, v]) = \langle \alpha, [u, v] \rangle$, for all $u \in \mathfrak{g}$. Note that for each $\xi \in \Gamma(\tilde{\mathfrak{g}})$,

$$ad^*_{\frac{\delta h}{\delta \mu}(\mu)} \mu(\xi) = \left\langle \mu, \left[\xi, \frac{\delta h}{\delta \mu}(\mu) \right] \right\rangle,$$

which is the definition of the Lie-Poisson bracket.

4 Yang-Mills equations

4.1 Introduction

In order to define the Yang-Mills equations we first need to recall the definition of the Hodge star operator:

Definition 48 (Hodge star operator)**.** *Let (M, g) be a semi-Riemannian manifold of dimension n. The Hodge star operator is defined to be the only linear operator*

$$\begin{array}{rccc} * : & \Omega^k(M) & \longrightarrow & \Omega^{n-k}(M) \\ & \alpha & \longmapsto & *\alpha \end{array}$$

satisfying that $\beta \wedge (*\alpha) = g(\alpha, \beta) v_g$ for all $\beta \in \Omega^k(M)$.

Proposition 26. *Let $(x_1, \ldots, x_n)$ be a local coordinate system, and write $\alpha = \alpha_{i_1 \cdots i_k} dx^1 \wedge \cdots \wedge dx^{i_k}$, then*

$$*\alpha = \frac{1}{(n-k)!} \epsilon_{i_1 \cdots i_n} \sqrt{det(g)} \alpha_{j_1 \cdots j_k} g^{i_1 j_1} \cdots g^{i_k j_k} dx^{i_{k+1}} \wedge \cdots \wedge dx^{i_n},$$

where $\epsilon_{i_1 \cdots i_n}$ is the sign of the permutation $(1, \ldots, n) \longmapsto (i_1, \ldots, i_n)$.

We will constantly write $\sqrt{det(g)}$ for $\sqrt{|det(g)|}$.

The following properties of the Hodge star operator are satisfied:

- $*v_g = 1,\ *1 = v_g$,
- $*(*\alpha) = (-1)^{k(n-k)}\alpha$.

Let $\pi : P \longrightarrow M$ be a principal fiber bundle with structure group G over a compact semi-Riemannian manifold (M, g). Yang-Mills equations are

$$*d^A * F_A = 0,$$

where A is a connection on P and is the unknown of the equation, and F_A is the curvature. $*d^A*$ is often denoted by δ^A.

Remark 40. Note that the Hodge star operator is defined on $\Omega^k(M)$, so here we are seeing F_A as an element of $\Omega^2(M, \tilde{\mathfrak{g}})$ rather than $\Omega^2(P, \mathfrak{g})$. Then if $w \otimes \eta \in \Omega^2(M, \tilde{\mathfrak{g}})$, we define

$$*(w \otimes \eta) = (*w) \otimes \eta.$$

To see F_A in $\Omega^2(M, \tilde{\mathfrak{g}})$ we define, for each $u, v \in T_xM$,

$$F_A(u, v) = [(p, F_A(u_p^h, v_p^h))],$$

where $p \in \pi^{-1}(x)$ and u_p^h and v_p^h are the horizontal lifts of u and v to H_p. It is well-defined since $R_g^* F_A = Ad_{g^{-1}} \circ F_A$, so that

$$[(p \cdot g, F_A(u_{p\cdot g}^h, v_{p\cdot g}^h))] = [(p \cdot g, F_A(dR_g(u_p^h), dR_g(v_p^h)))]$$
$$= [(p \cdot g, Ad_{g^{-1}} \circ F_A(u_p^h, v_p^h))] = [(p, F_A(u_p^h, v_p^h))].$$

Yang-Mills equations are the Euler-Lagrange equations corresponding to the Lagrangian in the next subsection. Now we just give a lemma and a couple of calculations which will be useful in the next subsection:

Lemma 4. *If for* $\alpha = \omega_1 \otimes B_1, \beta = \omega_2 \otimes B_2 \in \bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}}$ *we write*

$$\alpha \dot{\wedge} \beta = h(B_1, B_2)\omega_1 \wedge \omega_2,$$

then

$$d(\alpha \dot{\wedge} \beta) = (d^A\alpha)\dot{\wedge}\beta + \alpha\dot{\wedge}(d^A\beta).$$

We will need the following:

- If A is a connection, that is, a section of the bundle of connections, ω a section of the underlying vector bundle and ϵ a real number, then

$$F_{A+\epsilon\omega} = d(A + \epsilon\omega) + [A + \epsilon\omega, A + \epsilon\omega] = dA + \epsilon d\omega + [A, A] + \epsilon A \wedge \omega + \epsilon^2[\omega, \omega]$$
$$= dA + [A, A] + \epsilon(d\omega + A \wedge \omega) + \epsilon^2[\omega, \omega] = F_A + \epsilon d^A\omega + \epsilon^2[\omega, \omega].$$

- If M is a compact manifold, then

$$\int_M (d^A\omega)\dot{\wedge} * F_A = \int_M d(\omega\dot{\wedge} * F_A) - \int_M \omega\dot{\wedge}(d^A * F_A) = -\int_M \omega\dot{\wedge}(d^A * F_A),$$

using the previous lemma and Stokes theorem.

4.2 Covariant Lagrangian reduction for Yang-Mills

Let $\pi : P \longrightarrow M$ be a principal fiber bundle with structure group G over a semi-Riemmanian compact manifold (M, g). Consider the corresponding bundle of connections $C \longrightarrow M$ over M, and define the Lagrangian density

$$\begin{array}{rccc} \mathcal{L}: & J^1(C) & \longrightarrow & \bigwedge^n T^*M \\ & j^1_x A & \longmapsto & \frac{1}{2}((F_A)_x, (F_A)_x)_{g,h} v_g \end{array},$$

where, if $\omega_1 \otimes \alpha_1, \omega_2 \otimes \alpha_2 \in \bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}}$ then $(\omega_1 \otimes \alpha_1, \omega_2 \otimes \alpha_2)_{g,h} = g(\omega_1, \omega_2)h(\alpha_1, \alpha_2)$, being h the Killing form (or any other Ad-invariant bilinear form on $\mathfrak{g}$).

Note that we just applied the Killing form to elements in $\tilde{\mathfrak{g}}$: if $\alpha_1 = [(p, B_1)]$ and $\alpha_2 = [(p, B_2)]$, we put

$$h(\alpha_1, \alpha_2) = h(B_1, B_2),$$

which is well-defined because of the Ad-invariance of h.

We will see the expression of the Lagrangian in coordinates in the following subsection.

Let $\Phi : P \longrightarrow P$ be a gauge transformation and write locally $\Phi(x, g) = (x, \gamma(x) \cdot g)$ for some smooth function $\gamma : U \subset M \longrightarrow G$, which can be seen as a section of $U \times G \longrightarrow U$.

Remark 41. Indeed, we can write Φ on $U \times G$ as $\Phi(x, g) = (x, \gamma(x, g))$. Since Φ satisfies $\Phi(x, g \cdot h) = \Phi(x, g) \cdot h$, we get $\gamma(x, g \cdot h) = \gamma(x, g) \cdot h$ and therefore we can write $\gamma(x, g) = \gamma(x, e) \cdot g = \gamma(x) \cdot g$.

For each connection A on P, Φ induces the transformed connection given locally by

$$A' = Ad_\gamma A + d\gamma\gamma^{-1}$$

(if ω is the connection form associated to A, then this expression corresponds to $(\Phi^{-1})^*\omega$).

Then we have an action of $J^1(U \times G)$ on $C|_U$ defined by

$$\begin{array}{ccc} J^1(U \times G) \times C|_U & \longrightarrow & C|_U \\ (j^1_x\gamma, A_x) & \longmapsto & Ad_{\gamma(x)}A_x + (d\gamma)_x\gamma^{-1}(x) \end{array},$$

and by 1-jet prolongation we get an action of $J^2(U \times G)$ on $J^1(C|_U)$:

$$\begin{array}{ccc} J^2(U \times G) \times J^1(C|_U) & \longrightarrow & J^1(C|_U) \\ (j^2_x\gamma, j^1_x A) & \longmapsto & j^1_x(Ad_\gamma A + d\gamma\gamma^{-1}) \end{array}.$$

Proposition 27. *The quotient space $J^1(C)/J^2(AdP)$ can be identified with $\bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}}$ and the projection with the curvature bundle map*

$$\begin{array}{rccc} \Omega: & J^1C & \longrightarrow & \bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}} \\ & j^1_x A & \longmapsto & (F_A)_x \end{array},$$

which is a surjective submersion with connected fibers, and where $AdP = P \times_G G$ is the associated bundle with action of G on G given by $\tilde{g} \cdot g = \tilde{g}g\tilde{g}^{-1}$.

See [3] for a proof.

Remark 42. Alternatively one can think of $\bigwedge^2 T^*M\otimes\tilde{\mathfrak{g}}$ as $J^1C/GauP_x$, where the group

$$GauP_x = \left\{\Phi \in GauP : \Phi|_{P_x} = Id_{P_x}\right\}$$

acts on J^1C with orbits that coincide with the fibers of the curvature map, and where $GauP$ denotes the group of all gauge transformations on P. Note that we are using $GauP_x \cong GauP_y$.

Since $F_{A'} = Ad_\gamma F_A$ and the Killing form is Ad-invariant, the Lagrangian density $\mathcal{L}$ remains invariant under gauge transformations, that is $\mathcal{L}(j^1_x A) = \mathcal{L}(j^1_x(Ad_\gamma A + \gamma^{-1}d\gamma))$, and hence it drops to the quotient space as

$$l : \bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}} \longrightarrow \bigwedge^n T^*M.$$

Remark 43. If $\Phi_C : C \longrightarrow C$ denotes the induced gauge transformation on the bundle of connections and $j^1\Phi_C$ its lifting to J^1C, that is, $j^1\Phi_C(j^1_xA) = j^1_x(\Phi_C(A))$ then the previous statement is a particular case of the following theorem.

Theorem 7 (Utiyama). *A smooth function $L : J^1C \longrightarrow \mathbb{R}$ is gauge invariant, that is $L \circ j^1\Phi_C = L$ for all $\Phi \in GauP$, if and only if $L = \hat{L} \circ \Omega$, where $\hat{L} : \bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}} \longrightarrow \mathbb{R}$ is an Ad-invariant smooth function.*

See [3] again for a proof of this theorem.

Now consider variations of a section A of $C \longrightarrow M$ of the form $A_\epsilon = A + \epsilon\omega$, where ω is a section of $T^*M \otimes \tilde{\mathfrak{g}} \longrightarrow M$. These variations will drop to the quotient space as $F_{A+\epsilon\omega}$. Note that

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} F_{A+\epsilon\omega} = d^A\omega,$$

for using the formula $F_{A+\epsilon\omega} = d(A+\epsilon\omega) + [A+\epsilon\omega, A+\epsilon\omega]$ we get $F_{A+\epsilon\omega} = F_A + \epsilon d^A w + \epsilon^2[\omega,\omega]$.

Hence, for a section F of $\bigwedge^2 T^*M \otimes \tilde{\mathfrak{g}} \longrightarrow M$ such that $F = F_A$ for some $A \in \Gamma(C)$, the variations along F we will take in the formulation of the variational problem in the quotient space will be sections of the form $F_A + \epsilon d^A w$.

The variational principle then yields

$$\begin{aligned} 0 = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_M l(F_\epsilon)v_g &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_M \frac{1}{2}(F + \epsilon d^A w, F + \epsilon d^A w)_{gh} v_g \\ &= \int_M (F, d^A\omega)_{gh} v_g = \int_M (*d^A * F, \omega)_{gh} v_g, \end{aligned}$$

for arbitrary w (here we are using the last calculation in the previous subsection and the fact that $(\alpha,\beta)_{gh}v_g = \alpha\dot{\wedge} * \beta$). Hence we get the Yang-Mills equation

$$\delta^A F_A := *d^A * F_A = 0.$$

Remark 44. It can be shown that it is only necessary to consider variations of a specific type, namely variations of the form

$$\Psi_\epsilon = \Phi^X_\epsilon \circ s,$$

where s is the section and X is any vertical vector field, with flow denoted by Φ^X_ϵ.

In our case, since the fibers are affine spaces, any such variation can be written as $A_\epsilon = A + \epsilon w$, with the notations above.

4.3 Hamilton-Cartan equations for Yang-Mills

For a fiber bundle $E \longrightarrow M$, the polysymplectic bundle was defined as $\Pi = TM \otimes V^*E \otimes \bigwedge^n T^*M$. Note that in our case we have $\Pi = TM \otimes (TM \otimes \tilde{\mathfrak{g}}^*) \otimes \bigwedge^n T^*M$ since $C \longrightarrow M$ is an affine bundle modelled on $T^*M \otimes \tilde{\mathfrak{g}}$. We will usually write $\Pi = TM \otimes (TM \otimes \tilde{\mathfrak{g}}^*)$ and assume that a fixed volume form $d^n x$ is given. $\alpha \in \Pi$ will be written in coordinates as $\alpha = \pi_\alpha^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \otimes B^\alpha$, where $\{B_1, \ldots, B_m\}$ is a basis for $\mathfrak{g}$ and $\left\{B^1, \ldots, B^m\right\}$ denotes its dual basis. If (x^i, A_i^α) are coordinates on C then we write $(x^i, A_i^\alpha, A_{ij}^\alpha)$ for the coordinates on J^1C. Using the formula $F_A = dA + [A, A]$ we can write the Yang-Mills Lagrangian in coordinates as

$$L(x^i, A_i^\alpha, A_{ij}^\alpha) = \frac{1}{2}\left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) g^{js} g^{ir} \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\eta A_s^\tau c_{\eta\tau}^\beta\right) h_{\alpha\beta} \sqrt{det(g)},$$

where $h_{\alpha\beta}$ is an abbreviation for $h(B_\alpha, B_\beta)$ and, in the right side of the equality, $j < i$ and $s < r$.

Indeed, if $A(x) = A_i^\alpha(x) dx^i \otimes B_\alpha$ denotes a section of C, that is, a connection on P, then

$$(dA)_x = (A_{ij}^\alpha - A_{ji}^\alpha) dx^j \wedge dx^i \otimes B_\alpha,\ j < i$$

and

$$[A_x, A_x] = \left[A_j^\nu dx^j \otimes B_\nu, A_i^\mu dx^i \otimes B_\mu\right] = A_j^\nu A_i^\mu \left[dx^j \otimes B_\nu, dx^i \otimes B_\mu\right]$$
$$= A_j^\nu A_i^\mu c_{\nu\mu}^\alpha dx^j \wedge dx^i \otimes B_\alpha = -A_j^\nu A_i^\mu c_{\mu\nu}^\alpha dx^j \wedge dx^i \otimes B_\alpha,\ j < i,$$

so

$$(F_A)_x = \left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) dx^j \wedge dx^i \otimes B_\alpha,\ j < i.$$

Therefore we get

$$L(j_x^1 A) = \frac{1}{2}((F_A)_x, (F_A)_x)_{gh} \sqrt{det(g)}$$

$$= \frac{1}{2}\left(\left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) dx^j \wedge dx^i \otimes B_\alpha, (A_{rs}^\beta - A_{sr}^\beta - A_r^\eta A_s^\tau c_{\eta\tau}^\beta) dx^s \wedge dx^r \otimes B_\beta\right)_{gh}$$

$$= \frac{1}{2}(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha) g^{js} g^{ir} (A_{rs}^\beta - A_{sr}^\beta - A_r^\eta A_s^\tau c_{\eta\tau}^\beta) h_{\alpha\beta} \sqrt{det(g)}.$$

Then the expression of $\widehat{\mathbb{F}\mathcal{L}} : J^1C \longrightarrow \Pi$ in coordinates will be

$$\widehat{\mathbb{F}\mathcal{L}}(j_x^1 A) = \widehat{\mathbb{F}\mathcal{L}}(x^i, A_i^\alpha, A_{ij}^\alpha) = \left(x^i, A_i^\alpha, \frac{\partial L}{\partial A_{ij}^\alpha}\right)$$

$$= \left(x^i, A_i^\alpha, \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\eta A_s^\tau c_{\eta\tau}^\beta\right) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)}\right),$$

with $s < r$ and no restrictions on i and j.

So we get that the image are the curvature forms. Since we are working on a given point and locally any 2-form is the curvature form of a 1-form, really we can reach any 2-form. The image is therefore $\mathrm{Im}\widehat{\mathbb{F}\mathcal{L}} = TM \wedge TM \otimes \tilde{\mathfrak{g}}^*$, or alternatively $\mathrm{Im}\widehat{\mathbb{F}\mathcal{L}} = \left\{\pi_\alpha^{ij} + \pi_\alpha^{ji} = 0\right\}$.

Note that $\widehat{\mathbb{F}\mathcal{L}}$ is not a diffeomorphism, that is, $\mathcal{L}$ is not hyperregular, since the image is strictly contained in Π, and also not a diffeomorphism with the image since $\dim(\mathrm{Im}\widehat{\mathbb{F}\mathcal{L}}) \neq \dim J^1C$. Therefore we cannot apply theorem 4. Anyway, we will be able to develop the Hamiltonian approach working on the image $\mathcal{P} := \mathrm{Im}\widehat{\mathbb{F}\mathcal{L}} = TM \wedge TM \otimes \tilde{\mathfrak{g}}^*$.

Let $\omega \in TM \wedge TM \otimes \tilde{\mathfrak{g}}^*$. The elements of its fiber by $\widehat{\mathbb{F}\mathcal{L}}^{-1}$ will be of the form $j^1_x(A+df)$, where f is a smooth function taking values in $\mathfrak{g}$ with $\frac{\partial f^\alpha}{\partial x^i}(x) = 0$ for all α, i and $\widehat{\mathbb{F}\mathcal{L}}(j^1_x A) = \omega$. In coordinates the elements of the fiber can be written as

$$\left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right).$$

We see that $\mathbb{F}\mathcal{L}$ is constant along the fiber of ω. In coordinates we have

$$\mathbb{F}\mathcal{L}\left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right) = \left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, \frac{\partial L}{\partial A^\alpha_{ij}}, L - \frac{\partial L}{\partial A^\alpha_{ij}} A^\alpha_{ij} \right)$$

$$= \left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, \left(A^\beta_{kl} + \frac{\partial^2 f^\beta}{\partial x^l \partial x^k} - \left(A^\beta_{lk} + \frac{\partial^2 f^\beta}{\partial x^k \partial x^l} \right) \right.\right.$$

$$\left. - \left(A^\mu_k + \frac{\partial f^\mu}{\partial x^k} \right) \left(A^\nu_l + \frac{\partial f^\nu}{\partial x^l} \right) c^\beta_{\mu\nu} \right) g^{jl} g^{ik} h_{\alpha\beta} \sqrt{det(g)},$$

$$L - \left(A^\beta_{kl} + \frac{\partial^2 f^\beta}{\partial x^l \partial x^k} - \left(A^\beta_{lk} + \frac{\partial^2 f^\beta}{\partial x^k \partial x^l} \right) - \left(A^\mu_k + \frac{\partial f^\mu}{\partial x^k} \right) \left(A^\nu_l + \frac{\partial f^\nu}{\partial x^l} \right) c^\beta_{\mu\nu} \right)$$

$$\left. g^{jl} g^{ik} h_{\alpha\beta} \sqrt{det(g)} \left(A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right) \right)$$

$$= \left(x^i, A^\alpha_i, \left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) g^{jl} g^{ik} h_{\alpha\beta} \sqrt{det(g)}, \right.$$

$$\left. L - \left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) g^{jl} g^{ik} h_{\alpha\beta} \sqrt{det(g)} \left(A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right) \right).$$

Note that

$$L\left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right) = L(x^i, A^\alpha_i, A^\alpha_{ij}),$$

and also that

$$\left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) g^{jl} g^{ik} \left(A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right)$$

$$= g\left(\left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) dx^l \otimes dx^k, \left(A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} \right) dx^j \otimes dx^i \right).$$

Now, since $T^*M \otimes T^*M = (T^*M \wedge T^*M) \oplus^\perp (T^*M \vee T^*M)$ and $\frac{\partial^2 f^\alpha}{\partial x^j \partial x^i} dx^i \otimes dx^j \in T^*M \vee T^*M$, we get

$$g\left(\left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) dx^l \otimes dx^k, \left(A_{ij} + \frac{\partial^2 f^\alpha}{\partial x^i \partial x^j} \right) dx^j \otimes dx^i \right)$$

$$= g\left(\left(A^\beta_{kl} - A^\beta_{lk} - A^\mu_k A^\nu_l c^\beta_{\mu\nu} \right) dx^l \otimes dx^k, A_{ij} dx^j \otimes dx^i \right),$$

for $g\left((A^\beta_{kl} - A^\beta_{lk} + A^\mu_k A^\nu_l c^\beta_{\mu\nu})dx^l \otimes dx^k, \frac{\partial^2 f^\alpha}{\partial x^i \partial x^j} dx^j \otimes dx^i\right)$ vanishes.

Consequently, $\mathbb{F}\mathcal{L}$ does not depend on the element of the fiber, that is

$$\mathbb{F}\mathcal{L}\left(x^i, A^\alpha_i + \frac{\partial f^\alpha}{\partial x^i}, A^\alpha_{ij} + \frac{\partial^2 f^\alpha}{\partial x^j \partial x^i}\right) = \mathbb{F}\mathcal{L}(x^i, A^\alpha_i, A^\alpha_{ij}).$$

Therefore we have a well-defined section $\delta = \mathbb{F}\mathcal{L} \circ \widehat{\mathbb{F}\mathcal{L}}^{-1}$ of $k : (J^1C)^* \longrightarrow \Pi$ along $\mathcal{P}$ and hence the multisymplectic structure on $(J^1C)^*$ can be transferred to $\mathcal{P}$.

$$\begin{tikzcd} J^1C \arrow[r, "\mathbb{F}\mathcal{L}"] \arrow[dr, "\widehat{\mathbb{F}\mathcal{L}}"'] & (J^1C)^* \arrow[d, shift right, "k"'] \\ & \mathcal{P} \subset \Pi \arrow[u, shift right, "\delta"'] \end{tikzcd}$$

Let Ω be the canonical multisymplectic $(n+1)$-form on $(J^1C)^*$ we introduced earlier. In coordinates,

$$\Omega = dA^\alpha_i \wedge d\pi^{ij}_\alpha \wedge d^{n-1}x_j - dp \wedge d^n x,$$

where we are identifying the coordinates p^{ij}_α with π^{ij}_α.

We denote the pull-back of Ω to $k^{-1}(\mathcal{P})$ by $\Omega^{\mathcal{P}} = i^*\Omega$ and the pull-back to $\mathcal{P}$ by $\Omega^{\mathcal{P}}_\delta = \delta^* i^* \Omega$, where $i : k^{-1}(\mathcal{P}) \longrightarrow (J^1C)^*$ is the inclusion.

We will call $(\mathcal{P}, \delta)$ a Hamiltonian system with constraints and say that a section π of $\mathcal{P} \longrightarrow M$ is a solution of the Hamiltonian system if the equation

$$\pi^*\left(i_X \Omega^{\mathcal{P}}_\delta\right) = 0$$

is satisfied for any vertical vector field X on $\mathcal{P}$.

We define the following change of coordinates on Π (and on $(J^1C)^*$):

$$R^{ij}_\alpha = \frac{1}{2}(\pi^{ij}_\alpha - \pi^{ji}_\alpha) \text{ , if } i < j$$

$$S^{ij}_\alpha = \frac{1}{2}(\pi^{ij}_\alpha + \pi^{ji}_\alpha) \text{ , if } i \leq j$$

Note that this change of coordinates corresponds to the expression of π^{ij}_α as the direct sum of its symmetric and antisymmetric parts:

$$\pi^{ij}_\alpha = R^{ij}_\alpha + S^{ij}_\alpha,\ R^{ij}_\alpha \in TM \wedge TM \otimes \tilde{\mathfrak{g}}^*,\ S^{ij}_\alpha \in TM \vee TM \otimes \tilde{\mathfrak{g}}^*.$$

When $i > j$, we have $R^{ij}_\alpha = -R^{ji}_\alpha$ and $S^{ij}_\alpha = S^{ji}_\alpha$.

The constraint manifold $\mathcal{P}$ is defined by $S^{ij}_\alpha = 0$ for all $i \leq j$ and the pull-back of the multisymplectic form Ω on $k^{-1}(\mathcal{P})$ will have the following expression in local coordinates:

$$\Omega^{\mathcal{P}} = -\sum_{i<j} dR^{ij}_\alpha \wedge (dA^\alpha_i \wedge d^{n-1}x_j - dA^\alpha_j \wedge d^{n-1}x_i) - dp \wedge d^n x.$$

Indeed, the local expression of Ω on $(J^1C)^*$ was

$$\Omega = dA^\alpha_i \wedge d\pi^{ij}_\alpha \wedge d^{n-1}x_j - dp \wedge d^n x.$$

Applying the change of coordinates we get

$$\Omega = dA_i^\alpha \wedge d(R_\alpha^{ij} + S_\alpha^{ij}) \wedge d^{n-1}x_j - dp \wedge d^n x.$$

If $\Omega^{\mathcal{P}}$ is the restriction of Ω to $k^{-1}(\mathcal{P})$, since on $\mathcal{P}$ (and therefore on $k^{-1}(\mathcal{P})$) we have $S_\alpha^{ij} = 0$ and $R_\alpha^{ii} = 0$ for all i, j, α, we obtain

$$\Omega^{\mathcal{P}} = -dR_\alpha^{ij} \wedge dA_i^\alpha \wedge d^{n-1}x_j - dp \wedge d^n x$$

$$= \sum_{i<j} -dR_\alpha^{ij} \wedge dA_i^\alpha \wedge d^{n-1}x_j + \sum_{i>j} dR_\alpha^{ji} \wedge dA_i^\alpha \wedge d^{n-1}x_j - dp \wedge d^n x$$

$$= -\sum_{i<j} dR_\alpha^{ij} \wedge (dA_i^\alpha \wedge d^{n-1}x_j - dA_j^\alpha \wedge d^{n-1}x_i) - dp \wedge d^n x,$$

where we have used $R_\alpha^{ij} = -R_\alpha^{ji}$.

Now we will write Hamilton's equations.

Recall that

$$\Omega = -d\Theta,$$

where the expression of Θ in local coordinates is given by

$$\Theta = \pi_\alpha^{ij} dA_i^\alpha \wedge d^{n-1}x_j + p d^n x.$$

When we had an Ehresmann connection A on $C \longrightarrow M$, we were able to define a section of $(J^1C)^* \longrightarrow \Pi$ associated to it in the following way:

$$\begin{array}{rccc} \delta_A: & \Pi = TM \otimes (TM \otimes \tilde{\mathfrak{g}}^*) \otimes \bigwedge^n T^*M & \longrightarrow & Z \cong (J^1C)^* \\ & w \otimes \xi \otimes v & \longmapsto & (\xi \circ A) \wedge i_w v, \end{array}$$

where $w \in TM$, $\xi \in TM \otimes \tilde{\mathfrak{g}}^*$ and $v \in \bigwedge^n T^*M$.

The expression in coordinates was

$$\Theta_{\delta_A} = \pi_\alpha^{ij} dA_i^\alpha \wedge d^{n-1}x_j + \pi_\alpha^{ij} \Gamma_{ij}^\alpha d^n x,$$

where Γ_{ij}^α denote the coefficients of the horizontal lift

$$\frac{\partial}{\partial x^j} \longmapsto \frac{\partial}{\partial x^j} - \Gamma_{ij}^\alpha dx^i \otimes B_\alpha$$

defined by the connection A. Hence $H_{\delta_A} = -\pi_\alpha^{ij} \Gamma_{ij}^\alpha$. Note that we are using the vector bundle structure to identify $\frac{\partial}{\partial A_i^\alpha}$ with $dx^i \otimes B_\alpha$.

In the case of a linear connection we can write

$$\frac{\partial}{\partial x^j} \longmapsto \frac{\partial}{\partial x^j} - \Gamma_{ij\sigma}^{\alpha k} A_k^\sigma dx^i \otimes B_\alpha.$$

We will assume that the symbols of the connection are symmetric, that is, $\Gamma_{ij}^\alpha = \Gamma_{ji}^\alpha$.

In the coordinates introduced, Hamilton-Cartan equations have the following expression:

$$\left(\frac{\partial H}{\partial \pi_\alpha^{ij}}\right)_\pi = \left(\frac{\partial A_i^\alpha}{\partial x^j} + \Gamma_{ij}^\alpha\right)_\pi, \quad \left(\frac{\partial H}{\partial A_i^\alpha}\right)_\pi = -\left(\frac{\partial \pi_\alpha^{ij}}{\partial x^j} - \frac{\partial \Gamma_{kj}^\beta}{\partial A_i^\alpha} \pi_\beta^{kj}\right)_\pi.$$

Let us rewrite them on $\mathcal{P}$ after the change of coordinates. Using the chain rule we get, from the first set of equations,

$$\left(\frac{\partial H}{\partial R^{\mu\nu}_{\beta}}\right) = \left(\frac{\partial H}{\partial \pi^{ij}_{\alpha}}\right)\frac{\partial \pi^{ij}_{\alpha}}{\partial R^{\mu\nu}_{\beta}} = \frac{\partial H}{\partial \pi^{\mu\nu}_{\beta}} - \frac{\partial H}{\partial \pi^{\nu\mu}_{\beta}}$$

$$= \left(\frac{\partial A^{\beta}_{\mu}}{\partial x^{\nu}} + \Gamma^{\beta}_{\mu\nu}\right) - \left(\frac{\partial A^{\beta}_{\nu}}{\partial x^{\mu}} + \Gamma^{\beta}_{\nu\mu}\right) = \frac{\partial A^{\beta}_{\mu}}{\partial x^{\nu}} - \frac{\partial A^{\beta}_{\nu}}{\partial x^{\mu}},$$

where we are using the symmetry of the connection. Note that in these equations $\mu < \nu$.

The other set of equations gives us

$$\frac{\partial H}{\partial A^{\alpha}_{i}} = -\left(\frac{\partial \pi^{ij}_{\alpha}}{\partial x^{j}} - \frac{\partial \Gamma^{\beta}_{kj}}{\partial A^{\alpha}_{i}}\pi^{kj}_{\beta}\right) = -\left(\frac{\partial\left(R^{ij}_{\alpha} + S^{ij}_{\alpha}\right)}{\partial x^{j}} - \frac{\partial \Gamma^{\beta}_{kj}}{\partial A^{\alpha}_{i}}\left(R^{kj}_{\beta} + S^{kj}_{\beta}\right)\right)$$

$$= -\left(\frac{\partial R^{ij}_{\alpha}}{\partial x^{j}} - \frac{\partial \Gamma^{\beta}_{kj}}{\partial A^{\alpha}_{i}}R^{kj}_{\beta}\right) = -\frac{\partial R^{ij}_{\alpha}}{\partial x^{j}},$$

where in the last equality we are using that $\frac{\partial \Gamma^{\beta}_{kj}}{\partial A^{\alpha}_{i}}R^{kj}_{\beta} = 0$ since $\Gamma^{\beta}_{kj} = \Gamma^{\beta}_{jk}$ and $R^{kj}_{\beta} = -R^{jk}_{\beta}$.

Now let us write the expression of H in coordinates, which is simply the difference between the last coordinate of δ_A and of δ. This comes from taking the difference $\Theta_{\delta} - \Theta_{\delta_A} = \mathcal{H} = -H d^n x$ and ignoring the volume form $d^n x$, so $H = H_{\delta} - H_{\delta_A}$.

The expression of δ in coordinates was given by

$$\left(x^{i}, A^{\alpha}_{i}, \pi^{ij}_{\alpha}\right) \longmapsto \left(x^{i}, A^{\alpha}_{i}, \pi^{ij}_{\alpha}, L \circ (\widehat{\mathbb{F}\mathcal{L}})^{-1} - \pi^{ij}_{\alpha}\frac{\partial A^{\alpha}_{i}}{\partial x^{j}}\right).$$

Then the expression of H on Π is

$$H = -\left(L \circ (\widehat{\mathbb{F}\mathcal{L}})^{-1} - \pi^{ij}_{\alpha}\frac{\partial A^{\alpha}_{i}}{\partial x^{j}}\right) + \pi^{ij}_{\alpha}\Gamma^{\alpha}_{ij},$$

and composing with $\widehat{\mathbb{F}\mathcal{L}}$ yields

$$H \circ \widehat{\mathbb{F}\mathcal{L}} = -L + \left(A^{\beta}_{rs} - A^{\beta}_{sr} - A^{\mu}_{r}A^{\nu}_{s}c^{\beta}_{\mu\nu}\right) g^{js}g^{ir}h_{\alpha\beta}\sqrt{det(g)}(A^{\alpha}_{ij} - A^{\alpha}_{ji})$$

$$+ \left(A^{\beta}_{rs} - A^{\beta}_{sr} - A^{\mu}_{r}A^{\nu}_{s}c^{\beta}_{\mu\nu}\right) g^{js}g^{ir}h_{\alpha\beta}\sqrt{det(g)}(\Gamma^{\alpha}_{ij} - \Gamma^{\alpha}_{ji}),$$

where the sum is taken over $j < i$ and $s < r$, and where we are using that $\pi^{ij}_{\alpha} = -\pi^{ji}_{\alpha}$ in $\mathcal{P}$. Since $\Gamma^{\alpha}_{ij} = \Gamma^{\alpha}_{ji}$, the term depending on the connection vanishes. Now we write the definition of L and compute:

$$H \circ \widehat{\mathbb{F}\mathcal{L}} = -L + \left(A^{\beta}_{rs} - A^{\beta}_{sr} - A^{\mu}_{r}A^{\nu}_{s}c^{\beta}_{\mu\nu}\right) g^{js}g^{ir}h_{\alpha\beta}\sqrt{det(g)}(A^{\alpha}_{ij} - A^{\alpha}_{ji})$$

$$= -\frac{1}{2}\left(A^{\alpha}_{ij} - A^{\alpha}_{ji} - A^{\mu}_{i}A^{\nu}_{j}c^{\alpha}_{\mu\nu}\right) g^{js}g^{ir}\left(A^{\beta}_{rs} - A^{\beta}_{sr} - A^{\eta}_{r}A^{\tau}_{s}c^{\beta}_{\eta\tau}\right) h_{\alpha\beta}\sqrt{det(g)}$$

$$+ \left(A^{\beta}_{rs} - A^{\beta}_{sr} - A^{\mu}_{r}A^{\nu}_{s}c^{\beta}_{\mu\nu}\right) g^{js}g^{ir}h_{\alpha\beta}\sqrt{det(g)}(A^{\alpha}_{ij} - A^{\alpha}_{ji})$$

$$= \frac{1}{2}(A^\beta_{rs} - A^\beta_{sr} - A^\eta_r A^\tau_s c^\beta_{\eta\tau}) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)} \left(A^\alpha_{ij} - A^\alpha_{ji} + A^\mu_i A^\nu_j c^\alpha_{\mu\nu}\right)$$

$$= \frac{1}{2}(A^\beta_{rs} - A^\beta_{sr} - A^\eta_r A^\tau_s c^\beta_{\eta\tau}) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)} \left(A^\alpha_{ij} - A^\alpha_{ji} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu} + 2A^\mu_i A^\nu_j c^\alpha_{\mu\nu}\right)$$

$$= \frac{1}{2}(A^\beta_{rs} - A^\beta_{sr} - A^\eta_r A^\tau_s c^\beta_{\eta\tau}) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)} \left(A^\alpha_{ij} - A^\alpha_{ji} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu}\right)$$

$$+(A^\beta_{rs} - A^\beta_{sr} - A^\eta_r A^\tau_s c^\beta_{\eta\tau}) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)} A^\mu_i A^\nu_j c^\alpha_{\mu\nu}.$$

Note that the first addend is equal to

$$\frac{1}{2}(F_A, F_A)_{gh} \sqrt{det(g)}.$$

Substituting $\pi^{ij}_\alpha = (A^\beta_{rs} - A^\beta_{sr} - A^\mu_r A^\nu_s c^\beta_{\mu\nu}) g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)}$ we get the expression of H on $\mathcal{P}$

$$H = \frac{1}{2} \pi^{ij}_\alpha \pi^{rs}_\beta g_{ri} g_{sj} h^{\alpha\beta} \left(\sqrt{det(g)}\right)^{-1} + \pi^{ij}_\alpha A^\mu_i A^\nu_j c^\alpha_{\mu\nu},$$

which in the new coordinates is

$$H = \frac{1}{2} R^{ji}_\alpha R^{sr}_\beta g_{ri} g_{sj} h^{\alpha\beta} \left(\sqrt{det(g)}\right)^{-1} - R^{ji}_\alpha A^\mu_i A^\nu_j c^\alpha_{\mu\nu},$$

with $j < i$ and $s < r$.

Now we can work out the left-hand sides of Hamilton-Cartan equations:

$$\frac{\partial H}{\partial R^{ji}_\alpha} = R^{sr}_\beta g_{ri} g_{sj} h^{\alpha\beta} \left(\sqrt{det(g)}\right)^{-1} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu},$$

$$\frac{\partial H}{\partial A^\mu_i} = -R^{ji}_\alpha A^\nu_j c^\alpha_{\mu\nu}.$$

So the Hamilton-Cartan equations will be

$$R^{sr}_\beta g_{ri} g_{sj} h^{\alpha\beta} \left(\sqrt{det(g)}\right)^{-1} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu} = A^\alpha_{ji} - A^\alpha_{ij},$$

$$-R^{ji}_\alpha A^\nu_j c^\alpha_{\mu\nu} = -\frac{\partial R^{il}_\mu}{\partial x^l},$$

where $j < i$, $s < r$ and $i < l$.

The first set of equations is equivalent to

$$-R^{sr}_\beta = \left(A^\alpha_{ij} - A^\alpha_{ji} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu}\right) g^{ri} g^{sj} h_{\alpha\beta} \sqrt{det(g)}.$$

If we write

$$F^\alpha_{ji} = (A^\alpha_{ij} - A^\alpha_{ji} - A^\mu_i A^\nu_j c^\alpha_{\mu\nu}),\ j < i,$$

that is, $F_A = F^\alpha_{ji} dx^j \wedge dx^i \otimes B_\alpha$, then

$$\pi^{ij}_\alpha = F^\beta_{sr} g^{js} g^{ir} h_{\alpha\beta} \sqrt{det(g)},\ s < r$$

which we write as F_α^{ji}, so what we get is

$$R_\beta^{sr} = -\left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) g^{ri} g^{sj} h_{\alpha\beta} \sqrt{det(g)}$$
$$= -F_{ji}^\alpha g^{ri} g^{sj} h_{\alpha\beta} \sqrt{det(g)} = -F_\beta^{sr} = F_\beta^{rs},$$

which implies that solutions of Hamilton-Cartan equations for Yang-Mills come from sections of J^1C composed with the linear Legendre transformation, in fact prolongations of sections of C composed with the linear Legendre transformation.

Now we see that the second set of equations is equivalent to Yang-Mills equation

$$*d^A * F_A = 0.$$

In coordinates we write

$$F_A = \left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) dx^j \wedge dx^i \otimes B_\alpha.$$

Then using the expression in coordinates of the Hodge star operator we gave in the introduction of this section we obtain

$$*F_A = \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{js} g^{ir} \frac{\sqrt{det(g)}}{(n-2)!} \epsilon_{jii_3\cdots i_n} dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\beta,$$

where $j < i$ and $s < r$.

Using the other set of equations,

$$-R_\beta^{sr} = \left(A_{ij}^\alpha - A_{ji}^\alpha - A_i^\mu A_j^\nu c_{\mu\nu}^\alpha\right) g^{ri} g^{sj} h_{\alpha\beta} \sqrt{det(g)},$$

we get that

$$*F_A = \frac{-R_\alpha^{ji} h^{\alpha\beta}}{(n-2)!} \epsilon_{jii_3\cdots i_n} dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\beta,\ \ j < i.$$

We will use the formula

$$d^A(*F_A) = d(*F_A) + A \wedge (*F_A).$$

In the first addend we only need to derive with repect to the coordinates on M (recall that we are looking for sections of $\Pi \longrightarrow M$):

$$d(*F_A) = -\frac{\partial R_\alpha^{ji}}{\partial x^j} \frac{h^{\alpha\beta}}{(n-2)!} \epsilon_{jii_3\cdots i_n} dx^j \wedge dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\beta$$
$$-\frac{R_\alpha^{ji}}{\partial x^i} \frac{h^{\alpha\beta}}{(n-2)!} \epsilon_{jii_3\cdots i_n} dx^i \wedge dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\beta.$$

The second gives

$$A \wedge (*F_A) = \left(A_k^\alpha dx^k \otimes B_\alpha\right)$$
$$\wedge \left(\left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \frac{\sqrt{det(g)}}{(n-2)!} \epsilon_{jii_3\cdots i_n} dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\beta\right)$$

$$= A_i^\alpha \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \frac{\sqrt{det(g)}}{(n-2)!} c_{\alpha\beta}^\gamma \epsilon_{jii_3\cdots i_n} dx^i \wedge dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\gamma$$

$$+A_j^\alpha \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \frac{\sqrt{det(g)}}{(n-2)!} c_{\alpha\beta}^\gamma \epsilon_{jii_3\cdots i_n} dx^j \wedge dx^{i_3} \wedge \cdots \wedge dx^{i_n} \otimes B_\gamma.$$

So Yang-Mills equation

$$*(d * F_A) + *(A \wedge (*F_A)) = 0, \text{ equivalently, } d * F_A + A \wedge (*F_A) = 0,$$

yield

$$-\frac{\partial R_\alpha^{ji}}{\partial x^j} \frac{h^{\alpha\gamma}}{(n-2)!} \epsilon_{jii_3\cdots i_n} = -A_j^\alpha \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \frac{\sqrt{det(g)}}{(n-2)!} c_{\alpha\beta}^\gamma \epsilon_{jii_3\cdots i_n},$$

$$-\frac{\partial R_\alpha^{ji}}{\partial x^i} \frac{h^{\alpha\gamma}}{(n-2)!} \epsilon_{jii_3\cdots i_n} = -A_i^\alpha \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \frac{\sqrt{det(g)}}{(n-2)!} c_{\alpha\beta}^\gamma \epsilon_{jii_3\cdots i_n},$$

that is,

$$-\frac{\partial R_\mu^{ji}}{\partial x^j} h^{\mu\gamma} = -A_j^\alpha \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \sqrt{det(g)} c_{\alpha\beta}^\gamma.$$

Substituting

$$-R_l^{ji} h^{l\beta} = \left(A_{rs}^\beta - A_{sr}^\beta - A_r^\mu A_s^\nu c_{\mu\nu}^\beta\right) g^{ri} g^{sj} \sqrt{det(g)},$$

we get

$$-\frac{\partial R_\mu^{ji}}{\partial x^j} h^{\mu\gamma} = A_j^\alpha R_l^{ji} h^{\beta l} c_{\alpha\beta}^\gamma$$

and finally

$$-\frac{\partial R_\mu^{ji}}{\partial x^j} = A_j^\alpha R_\gamma^{ji} c_{\alpha\mu}^\gamma = -A_j^\alpha R_\gamma^{ji} c_{\mu\alpha}^\gamma, \ j < i,$$

or

$$-\frac{\partial R_\mu^{il}}{\partial x^l} = -A_j^\alpha R_\gamma^{ji} c_{\mu\alpha}^\gamma, \ i < l.$$

So we have that solutions of Hamilton-Cartan equations are the composition of the linear Legendre transformation with the prolongation of solutions of Yang-Mills equations.

4.4 Poisson forms on the constraint manifold $\mathcal{P}$

Recall that we can view Poisson r-forms on Π ($r > 0$) since they do not depend on the affine coordinate of $(J^1E)^*$. Using $\Omega^{\mathcal{P}}$ we can view them on $\mathcal{P}$:

Definition 49 (Poisson forms on the constraied manifold)**.** *A horizontal r-form E on $k^{-1}(\mathcal{P})$ is said to be Poisson if there esists an $(n-r)$-multivector field X_E on $k^{-1}(\mathcal{P})$ such that $i_{X_E}\Omega^{\mathcal{P}} = dE$.*

Recall that the expression of $\Omega^{\mathcal{P}}$ in coordinates was

$$-\sum_{i<j} dR_\alpha^{ij} \wedge (dA_i^\alpha \wedge d^{n-1}x_j - dA_j^\alpha \wedge d^{n-1}x_i) - dp \wedge d^n x.$$

Proposition 28. *Poisson $(n-1)$-forms on $\mathcal{P}$ can be locally written as*

$$E = (R_\alpha^{j\mu} X_j^\alpha + G^\mu) d^{n-1} x_\mu,$$

where $X_j^\alpha, G^\mu \in \mathcal{F}(T^*M \otimes \mathfrak{g})$ and $-\frac{\partial E^\mu}{\partial A_i^\alpha} = \frac{\partial E^i}{\partial A_\mu^\alpha}$.

Proof. Let $E = E^\mu d^{n-1} x_\mu$ be a Poisson $(n-1)$-form on $\mathcal{P}$. Then

$$dE = \frac{\partial E^\mu}{\partial x^\mu} d^n x + \frac{\partial E^\mu}{\partial A_i^\alpha} dA_i^\alpha \wedge d^{n-1} x_\mu + \frac{\partial E^\mu}{\partial R_\alpha^{ij}} dR_\alpha^{ij} \wedge d^{n-1} x_\mu,$$

where $i < j$. Now let

$$X = X \frac{\partial}{\partial p} + X_i^\alpha \frac{\partial}{\partial A_i^\alpha} + X_\alpha^{ij} \frac{\partial}{\partial R_\alpha^{ij}}$$

be a vertical vector field on $k^{-1}(\mathcal{P})$. Then

$$i_X \Omega^{\mathcal{P}} = -X d^n x + \sum_{i<j} (X_i^\alpha dR_\alpha^{ij} \wedge d^{n-1} x_j - X_j^\alpha dR_\alpha^{ij} \wedge d^{n-1} x_i)$$

$$- \sum_{i<j} X_\alpha^{ij} (dA_i^\alpha \wedge d^{n-1} x_j - dA_j^\alpha \wedge d^{n-1} x_i),$$

so

$$\frac{\partial E^\mu}{\partial x^\mu} = -X, \quad X_\alpha^{i\mu} = -\frac{\partial E^\mu}{\partial A_i^\alpha} = \frac{\partial E^i}{\partial A_\mu^\alpha} \text{ for } i < \mu, \quad \text{and } \frac{\partial E^\mu}{\partial R_\alpha^{ij}} = \delta_\mu^j X_i^\alpha - \delta_\mu^i X_j^\alpha.$$

Note that the first condition imposes no restrictions on E. The third condition gives $E = (R_\alpha^{j\mu} X_j^\alpha + G^\mu) d^{n-1} x_\mu$ with X_j^α and G^i functions on $T^*M \otimes \mathfrak{g}$, where we are using $R_\alpha^{\mu j} = -R_\alpha^{j\mu}$ and not necessarily $j < \mu$. Then the expression of E is as stated. □

Definition 50 (Poisson bracket)**.** *Let E and H be Poisson forms on $\mathcal{P}$ of degrees r and s respectively and X_E and X_H denote associated multivector fields. We define the Poisson bracket to be the $(r+s+1-n)$-form*

$$\{E, H\} = -i_{X_E} i_{X_H} \Omega^{\mathcal{P}}.$$

We calculate the local expression of the Poisson bracket for a Poisson $(n-1)$-form E and a function H not depending on p:

Since H is Poisson on $\mathcal{P}$ we have

$$i_{X_E} i_{X_H} \Omega^{\mathcal{P}} = i_{X_E} dH = i_{X_E} \left(\frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial A_i^\alpha} dA_i^\alpha + \frac{\partial H}{\partial R_\alpha^{ij}} dR_\alpha^{ij} \right).$$

On the other hand if

$$E = E^\mu d^{n-1} x_\mu, \quad X_E = X_i^\alpha \frac{\partial}{\partial A_i^\alpha} + X_\alpha^{ij} \frac{\partial}{\partial R_\alpha^{ij}} + X \frac{\partial}{\partial p},$$

then the condition $i_{X_E} \Omega^{\mathcal{P}} = dE$ in coordinates gives

$$i_{X_E} \left(-\sum_{i<j} dR_\alpha^{ij} \wedge (dA_i^\alpha \wedge d^{n-1} x_j - dA_j^\alpha \wedge d^{n-1} x_i) - dp \wedge d^n x \right)$$

$$= \sum_{i<j} \left(X_i^\alpha dR_\alpha^{ij} \wedge d^{n-1}x_j - X_j^\alpha dR_\alpha^{ij} \wedge d^{n-1}x_i \right)$$

$$- \sum_{i<j} X_\alpha^{ij} (dA_i^\alpha \wedge d^{n-1}x_j - dA_j^\alpha \wedge d^{n-1}x_i) - X d^n x$$

$$= \frac{\partial E^\mu}{\partial x^\mu} d^n x + \frac{\partial E^\mu}{\partial A_j^\alpha} dA_j^\alpha \wedge d^{n-1}x_\mu + \frac{\partial E^\mu}{\partial R_\alpha^{ij}} dR_\alpha^{ij} \wedge d^{n-1}x_\mu,$$

which implies that

$$X_E = \sum_{i<j} \frac{\partial E^j}{\partial R_\alpha^{ij}} \frac{\partial}{\partial A_i^\alpha} - \sum_{i<j} \frac{\partial E^j}{\partial A_i^\alpha} \frac{\partial}{\partial R_\alpha^{ij}} - \frac{\partial E^\mu}{\partial x^\mu} \frac{\partial}{\partial p}.$$

Then

$$\{E, H\} = - \sum_{i<j} \left(\frac{\partial E^j}{\partial A_i^\alpha} \frac{\partial H}{\partial R_\alpha^{ij}} - \frac{\partial E^j}{\partial R_\alpha^{ij}} \frac{\partial H}{\partial A_i^\alpha} \right).$$

Proposition 29. *A section π of the constrained bundle $\mathcal{P} \longrightarrow M$ is a solution of the Hamiltonian system $(\mathcal{P}, A, H)$ if and only if for every horizontal Poisson $(n-1)$-form E on $\mathcal{P}$ we have*

$$\{E, H\}\, d^n x = d(\pi^* E) - (d^h E) \circ \pi,$$

where d^h denotes the horizontal differential with respect to the connection on the bundle $\Pi \longrightarrow M$.

Proof. One the one hand, we have $E = (R_\alpha^{j\mu} X_j^\alpha + G^\mu) d^{n-1}x_\mu$, with $-\frac{\partial E^\mu}{\partial A_i^\alpha} = \frac{\partial E^i}{\partial A_\mu^\alpha}$ and $\frac{\partial E^\mu}{\partial R_\alpha^{ij}} = \delta_\mu^j X_i^\alpha - \delta_\mu^i X_j^\alpha$ and

$$\{E, H\} = \sum_{i<j} \left(\frac{\partial E^j}{\partial A_i^\alpha} \frac{\partial H}{\partial R_\alpha^{ij}} - \frac{\partial E^j}{\partial R_\alpha^{ij}} \frac{\partial H}{\partial A_i^\alpha} \right) = \sum_{i<j} \left(\frac{\partial E^j}{\partial A_i^\alpha} \frac{\partial H}{\partial R_\alpha^{ij}} - X_i^\alpha \frac{\partial H}{\partial A_i^\alpha} \right).$$

On the other hand, the expression for the horizontal lift with respect to the connection on Π given in proposition 24 is

$$d(\pi^* E) - (d^h E) \circ \pi = \frac{\partial E^\mu}{\partial x^\mu} + \frac{\partial E^\mu}{\partial A_i^\alpha} \frac{\partial A_i^\alpha}{\partial x^\mu} + \frac{\partial E^\mu}{\partial R_\alpha^{ij}} \frac{\partial R_\alpha^{ij}}{\partial x^\mu}$$

$$- \frac{\partial E^\mu}{\partial x^\mu} + \Gamma_{i\mu}^\alpha \frac{\partial E^\mu}{\partial A_i^\alpha} + \left(- \frac{\partial \Gamma_{k\mu}^\beta}{\partial A_i^\alpha} R_\beta^{kj} + \tilde{\Gamma}_{\mu l}^j R_\alpha^{il} - \tilde{\Gamma}_{\mu l}^l R_\alpha^{ij} \right) \frac{\partial E^\mu}{\partial R_\alpha^{ij}}.$$

Using that $\frac{\partial E^\mu}{\partial A_i^\alpha} = -\frac{\partial E^i}{\partial A_\mu^\alpha}$ and $\frac{\partial E^\mu}{\partial R_\alpha^{ij}} = \delta_\mu^j X_i^\alpha$ (for $E = (R_\alpha^{j\mu} X_j^\alpha + G^\mu) d^{n-1}x_\mu$) we get

$$d(\pi^* E) - (d^h E) \circ \pi = \sum_{\mu<i} \frac{\partial E^\mu}{\partial A_i^\alpha} \left(\frac{\partial A_i^\alpha}{\partial x^\mu} - \frac{\partial A_\mu^\alpha}{\partial x^i} \right) + \delta_\mu^j X_i^\alpha \frac{\partial R_\alpha^{ij}}{\partial x^\mu} - \sum_{\mu<i} \frac{\partial E^\mu}{\partial A_i^\alpha} \left(\Gamma_{\mu i}^\alpha - \Gamma_{i\mu}^\alpha \right)$$

$$- \delta_\mu^j X_i^\alpha \left(- \frac{\partial \Gamma_{k\mu}^\beta}{\partial A_i^\alpha} R_\beta^{kj} + \tilde{\Gamma}_{\mu l}^j R_\alpha^{il} - \tilde{\Gamma}_{\mu l}^l R_\alpha^{ij} \right)$$

$$= \sum_{\mu<i} \frac{\partial E^\mu}{\partial A_i^\alpha}\left(\frac{\partial A_i^\alpha}{\partial x^\mu} - \frac{\partial A_\mu^\alpha}{\partial x^i}\right) + X_i^\alpha \frac{\partial R_\alpha^{ij}}{\partial x^j} - X_i^\alpha \left(-\frac{\partial \Gamma_{kj}^\beta}{\partial A_i^\alpha} R_\beta^{kj} + \tilde{\Gamma}_{jl}^j R_\alpha^{il} - \tilde{\Gamma}_{jl}^l R_\alpha^{ij}\right)$$

$$= \sum_{\mu<i} \frac{\partial E^\mu}{\partial A_i^\alpha}\left(\frac{\partial A_i^\alpha}{\partial x^\mu} - \frac{\partial A_\mu^\alpha}{\partial x^i}\right) + X_i^\alpha \frac{\partial R_\alpha^{ij}}{\partial x^j},$$

$$= \sum_{i<\mu} \frac{\partial E^\mu}{\partial A_i^\alpha}\left(\frac{\partial A_i^\alpha}{\partial x^\mu} - \frac{\partial A_\mu^\alpha}{\partial x^i}\right) + X_i^\alpha \frac{\partial R_\alpha^{ij}}{\partial x^j},$$

where we are using $\tilde{\Gamma}_{jl}^j = \tilde{\Gamma}_{lj}^j$, $\Gamma_{kj}^\beta = \Gamma_{jk}^\beta$ and $R_\beta^{kj} = -R_\beta^{jk}$. Therefore we get that

$$\frac{\partial H}{\partial R_\alpha^{ij}} = \frac{\partial A_i^\alpha}{\partial x^j} - \frac{\partial A_j^\alpha}{\partial x^i} \quad \text{and} \quad \frac{\partial H}{\partial A_i^\alpha} = -\frac{\partial R_\alpha^{ij}}{\partial x^j}$$

are equivalent to

$$\{E, H\}\, d^n x = d(\pi^* E) - (d^h E) \circ \pi.$$

$\square$

4.5 Future work

If we denote by $\mathcal{G}$ the action of $J^1(AdP)$ on Π and $\mathcal{P}$ we get

$$\frac{\Pi}{\mathcal{G}} = TM \otimes TM \otimes \tilde{\mathfrak{g}}^* \otimes \bigwedge^n T^*M \longrightarrow M,$$

$$\frac{\mathcal{P}}{\mathcal{G}} = TM \wedge TM \otimes \tilde{\mathfrak{g}}^* \otimes \bigwedge^n T^*M \longrightarrow M.$$

We define coordinates (x^i, r_α^{ij}), $i < j$, on $\mathcal{P}/\mathcal{G}$ such that the projection $\mathcal{P} \longrightarrow \mathcal{P}/\mathcal{G}$ is given by $(x^i, A_i^\alpha, R_\alpha^{ij}) \longmapsto (x^i, r_\alpha^{ij})$, so $\mathcal{G}$-invariant functions and forms on $\mathcal{P}$ are the ones not depending on the coordinates (A_i^α).

In Electromagnetism the Hamiltonian is $\mathcal{G}$-invariant, so the Poisson bracket vanishes when applied also to $\mathcal{G}$-invariant Poisson forms. Then the Lie-Poisson bracket also vanishes and an analog of theorem 6 is easily written (see [1]). The Yang-Mills Hamiltonian is not invariant under this gauge action, so we cannot proceed as in [1] in order to reduce the equations. We intend to tackle the issue from a different perspective in a near future.

References

[1] M. Castrillón, J. E. Marsden, *Some remarks on Lagrangian and Poisson reduction for field theories*, Journal of Geometry and Physics (2003)

[2] M. Castrillón, J. M. Masqué, *The geometry of the bundle of connections*, Math. Z. 236 (2001)

[3] P. L. García, *Gauge algebras, curvature and symplectic structure*, Journal of Geometry and Physics (1977)

[4] G. Giachetta, L. Mangiarotti, G. Sardanashvily, *New Lagrangian and Hamiltonian Methods in Field Theories*, World Scientific (1997)

[5] D. Husemoller, *Fiber bundles*, Springer (1994)

[6] S. Kobayashi, K. Nomizu, *Foundations of Differentiable Geometry, Vol I*, John Wiley and Sons, Interscience (1963)

[7] J.E. Marsden, T.S. Ratiu *Introduction to Mechanics and Symmetry*, Springer (1999)

[8] G. Sardanashvily, *Generalized Hamiltonian Formulation for Field Theory. Constrained Systems*, World Scientific (1995)

[9] D.J. Saunders, *The Geometry of Jet Bundles*, Cambridge University Press (1989)

[10] F.W. Warner, *Foundations of differentiable manifolds and Lie Groups*, Springer (1983)