Notes on the Mathematical Foundations of Classical and Quantum Field Theory

Marcelo Mendes Disconzi

These notes are inspired in two summer courses taken at the Brazilian "National Institute of Pure and Applied Mathematics", in 2005 and 2006. However, they reflect my own understanding and perspectives on the subject, hence the word "inspired".

My purpose in writing down these notes is to clarify my own ideas about the subject. If you find yourself with these notes and have any comment, critic or suggestion, feel free to write me: disconzi@math.sunysb.edu.

Eventually, I would like to stress that my intention is to modify and complete the text in so far as my understanding on the subject grows and my free time allows it. Therefore there will be no "final version" of it.

Contents

Chapter 1

Fiber bundles and connections

1.1 Fiber bundles

Principal and fiber bundles are in the heart of gauge theory and hence are indispensable for constructing field theories.

Definition 1. *Let* $\mathcal{M}$ *be a* C^∞ *manifold and* G *a Lie group. A* **principal bundle** $\pi : P \to \mathcal{M}$ *over* $\mathcal{M}$ *is a* C^∞ *manifold* P *such that:*

i. G *freely acts on the right*[1] *on* P*:* $P \times G \to P, (u, g) \mapsto R_g u = ug$.

ii. $\mathcal{M} \approx P/G$ *and* $\pi : P \to \mathcal{M}$ *is* C^∞.

iii. P *is locally trivial. This means that for every* $x \in \mathcal{M}$ *there exists a neighborhood* $\mathcal{U} \ni x$ *such that* $\pi^{-1}(\mathcal{U})$ *is diffeomorphic to* $\mathcal{U} \times G$*. This diffeomorphism* $\phi : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times G$ *satisfies* $u \mapsto (\pi(u), \psi(u))$ *where* $\psi : \pi^{-1}(\mathcal{U}) \to G$ *satisfies* $\psi(ug) = \psi(u)g$.

M is called **base space**, G **structural group**, π **projection** and P **total space**. The set $\pi^{-1}(x)$, $x \in \mathcal{M}$ is called the **fiber over** x and the maps $\pi^{-1}(x) \to \mathcal{U} \times G$ are called **local trivializations**. Notice that the fiber over x is diffeomorphic to G and that the action on the fibers is transitive ([1] p. 166). We shall denote the **identity** of G by e. A principal bundle is represented diagrammatically as:

$$\begin{array}{ccc} G & \hookrightarrow & P \\ & & \downarrow \\ & & \mathcal{M} \end{array}$$

[1]Notice that this action is given by an anti-homomorphism

Definition 2. *A C^∞ map $\sigma : \mathcal{M} \to P$ such that $\pi \circ \sigma = id_{\mathcal{M}}$ is called a* ***cross section*** *or simply a section. A C^∞ map $\sigma : \mathcal{U} \subset \mathcal{M} \to P, \mathcal{U}$ open, such that $\pi \circ \sigma = id_{\mathcal{M}}$ is called a* ***local cross section*** *or simply a local section.*

Let $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ be an open cover of $\mathcal{M}$ such that for each $\mathcal{U}_\alpha$ there exists a trivialization: $\phi_\alpha : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times G, u \mapsto (\pi(u), \psi_\alpha(u))$ For $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ define $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ as follows. Consider $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and $u \in \pi^{-1}(x)$. Put $g_{\alpha\beta}(x) = \psi_\alpha(u)(\psi_\beta(u))^{-1}$. This definition is independent of of u: if $u' \in \pi^{-1}(x)$ then $u' = ug$ for some $g \in G$ (transitivity on the fibers) and $g_{\alpha\beta}(x) = \psi_\alpha(u')(\psi_\beta(u'))^{-1} = \psi_\alpha(u)g(\psi_\beta(u)g)^{-1} = \psi_\alpha(u)(\psi_\beta(u))^{-1}$. $g_{\alpha\beta}$ is called **transition function**. It is easy to check: $g_{\alpha\alpha}(x) = e$, $g_{\alpha\beta}(x) = (g_{\beta\alpha}(x))^{-1}$ and $g_{\gamma\beta}(x) = g_{\gamma\beta}(x)g_{\beta\alpha}(x)$ if $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$ — which is called the **cocycle condition**. The justification for the name transition function is the following:

Let $\mathcal{U}_\alpha$, $\mathcal{U}_\beta$ with $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and ϕ_α, ϕ_β the respective trivializations. Take $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and consider $\phi_\alpha|\pi^{-1}(x) : \pi^{-1}(x) \to G$ and $(\phi_\beta|\pi^{-1}(x))^{-1} : G \to \pi^{-1}(x)$. Then $\phi_\alpha|\pi^{-1}(x) \circ (\phi_\beta|\pi^{-1}(x))^{-1}$ defines a map from G to G. For $g \in G$ we have that $\phi_\alpha|\pi^{-1}(x) \circ (\phi_\beta|\pi^{-1}(x))^{-1}(g) \in G$. Notice that $\phi_\alpha|\pi^{-1}(x) = \psi_\alpha$ and $\phi_\beta|\pi^{-1}(x) = \psi_\beta$, therefore $\phi_\alpha|\pi^{-1}(x) \circ (\phi_\beta|\pi^{-1}(x))^{-1}(g) = \psi_\alpha \circ (\psi_\beta)^{-1}(g)$. Since $g_{\alpha\beta}$ is independent of the choice of $u \in P$, we can take $u = \psi_\beta^{-1}(g)$. Then: $g_{\alpha\beta}(x) = \psi_\alpha(u)(\psi_\beta(u))^{-1} = \psi_\alpha(\psi_\beta^{-1}(g))(\psi_\beta(\psi_\beta^{-1}(g)))^{-1} = \psi_\alpha \circ \psi_\beta^{-1}(g)g^{-1}$, hence $g_{\alpha\beta}(x)g = \psi_\alpha \circ \psi_\beta^{-1}(g) = \phi_\alpha|\pi^{-1}(x) \circ (\phi_\beta|\pi^{-1}(x))^{-1}(g)$.

If we have $\phi_\alpha : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times G$ and define for each $x \in \mathcal{U}_\alpha$ $\sigma_\alpha(x) = \phi_\alpha^{-1}(x, e)$, we obtain a local section $\sigma_\alpha : \mathcal{U}_\alpha \to P$ such that $\phi(\sigma_\alpha(x))) = (x, e)$. In this case if $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ we have that σ_α and σ_β are related by $\sigma_\beta(x) = \sigma_\alpha(x)g_{\alpha\beta}(x)$. Indeed, for each $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ holds that $\psi_\alpha \circ \psi_\beta^{-1}(g) = g_{\alpha\beta}(x)g$, so $\psi_\beta^{-1}(g) = \psi_\alpha^{-1}(g_{\alpha\beta}(x)g) = \psi_\alpha^{-1}(g_{\alpha\beta}(x))g \Rightarrow (\phi_\beta|(\pi^{-1}(x)))^{-1}(g) = \phi_\beta^{-1}(x, g) = (\phi_\alpha|(\pi^{-1}(x)))^{-1}(g)(g_{\alpha\beta}(x))g = \phi_\alpha^{-1}(x, g_{\alpha\beta}(x))g = \phi_\alpha^{-1}(x, e)g_{\alpha\beta}(x)g$. In particular, for $g = e$ we have: $\phi_\beta^{-1}(x, e) = \sigma_\beta(x) = \phi_\alpha^{-1}(x, e)g_{\alpha\beta}(x) = \sigma_\alpha(x)g_{\alpha\beta}(x)$.

Reciprocally, given a local section $\sigma_\alpha : \mathcal{U}_\alpha \to P$ we define a local trivialization in the following way: for each $u \in \pi^{-1}(x)$, $x \in \mathcal{U}_\alpha$ there exists an unique $g_u \in G$ such that $u = \sigma_\alpha(x)g_u$. Define $\phi_\alpha(u) = (x, g_u)$. For this trivialization we have $\sigma_\alpha(x) = \phi^{-1}(x, e)$. On the intersection: $\sigma_\alpha(x) = \phi_\alpha^{-1}(x, e) = \phi_\beta^{-1}(x, e)g_{\beta\alpha}(x) = \sigma_\beta(x)g_{\beta\alpha}(x)$. Such trivializations are called **canonical local trivialization**.

Definition 3. *Let F, $\mathcal{M}$, E be C^∞ manifolds. A* ***fiber bundle*** *of fiber F, base space $\mathcal{M}$ and total space E is a submersion $\pi : E \to \mathcal{M}$ with the follow-*

ing property: there exist an open cover $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ *of* M *and diffeomorphisms* $\phi_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times F$ *such that* $\pi_1 \circ \phi = \pi$ *where* $\pi_1 : (x, y) \mapsto x$*.*

Fiber bundles are represented diagrammatically as:

$$\begin{array}{ccc} F & \hookrightarrow & E \\ & & \downarrow \\ & & \mathcal{M} \end{array}$$

It follows that for each $x \in M$, the fiber over x, $E_x = \pi^{-1}(x)$ is diffeomorphic to F. We also have that there are maps $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to Diff(F)$ such that $\phi_\beta \circ \phi_\alpha^{-1} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times F \to \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times F$ is written as:

$$\phi_\beta \circ \phi_\alpha^{-1}(x, y) = (x, g_{\alpha\beta}(x)(y))$$

Such $g_{\alpha\beta}$ have a cocycle condition: $\phi_{\alpha\alpha} = e$, $g_{\alpha\beta}(x) = (g_{\beta\alpha}(x))^{-1}$ and $g_{\gamma\beta}(x) = g_{\gamma\beta}(x)g_{\beta\alpha}(x)$ if $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$.

Remark 1. *Roughly speaking, we may say that a principal bundle is a fiber bundle whose fiber is a Lie group* G*.*

In general, if $\mathcal{M}$ is a manifold with an open cover $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ and G is a Lie group, a family of functions $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ satisfying the three properties above is called a **cocycle** for $\mathcal{M}$. For example, the transition functions define a cocycle. As before:

Definition 4. *A* C^∞ *map* $\sigma : \mathcal{M} \to E$ *such that* $\pi \circ \sigma = id_\mathcal{M}$ *is called a* ***cross section*** *or simply a section. A* C^∞ *map* $\sigma : \mathcal{U} \subset \mathcal{M} \to E$*,* $\mathcal{U}$ *open, such that* $\pi \circ \sigma = id_\mathcal{M}$ *is called a* ***local cross section*** *or simply a local section. We denote the space of sections by* $\Gamma(E)$*.*

A section $\sigma \in \Gamma(E)$ defines a family of functions $\sigma_\alpha : U_\alpha \to F$ such that $\sigma(x) = \phi_\alpha(x, \sigma_\alpha(x))$. If follows that this family satisfies

$$\sigma_\beta(x) = g_{\beta\alpha}(x)(\sigma_\alpha(x))$$

Reciprocally, every family of functions satisfying the above equation defines a section.

We should make a comment about a notational confusion that might arise. For principal bundles the relation among local sections whose domains intersect non-trivially is given by right multiplication by transition functions: $\sigma_\beta(x) = \sigma_\alpha(x)g_{\alpha\beta}(x)$, whereas in the general fiber bundle setting just introduced the cocycles appear on the left $\sigma_\beta(x) = g_{\beta\alpha}(x)(\sigma_\alpha(x))$. Since principal bundles are particular cases of fiber bundles (i.e., fiber bundles whose fiber is G) this might sound inconsistent. However, notice that in $\sigma_\beta(x) = \sigma_\alpha(x)g_{\alpha\beta}(x)$ we have the group multiplication between the elements $\sigma_\alpha(x) \in G$ and $g_{\alpha\beta}(x) \in G$, while in $\sigma_\beta(x) = g_{\beta\alpha}(x)(\sigma_\alpha(x))$ we are evaluating the diffeomorphism $g_{\beta\alpha}(x) \in Diff(F)$ at the element $\sigma_\alpha(x) \in F$ (of couse, right multiplication corresponds to the case there the diffeomorphism is given by $g_{\beta\alpha}(x) = R_{g_{\beta\alpha}(x)}$).

Remark 2. *If the fiber F has some structure (vector space, algebra, group etc) which is preserved by the maps $g_{\alpha\beta}$ then each fiber E_x also has that structure. It follows that the space of sections $\Gamma(E)$ also has the structure.*

Definition 5. *If F is a vector space and $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(F)$ we say that $\pi : E \to \mathcal{M}$ is a* ***vector bundle****. In this case $\Gamma(E)$ is an infinite dimensional vector space; indeed, a module over the ring of C^∞ functions $f : \mathcal{M} \to \mathbb{R}$ or $\mathbb{C}$.*

Definition 6. *A* ***representation*** *of G in $Diff(F)$ is a homomorphism $\rho : G \to Diff(F)$ such that the map $G \times F \to F$, $(g, y) \mapsto \rho(g)(y)$ is differentiable. This map is called the* ***action on the left*** *of G on F. Some usual notations are $\rho(g)(y) = \rho(g)y = g \cdot y = gy$.*

Proposition 1. *Let $\rho : G \to Diff(F)$ a representation, $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ an open cover of $\mathcal{M}$ and $\gamma_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ a cocycle. Then there exists a fiber bundle $\pi : E \to \mathcal{M}$ with transition functions $\rho_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to Diff(F)$ given by $\rho_{\alpha\beta} = \rho \circ \gamma_{\alpha\beta}$.*

Proof: Let $\tilde{E} = \bigsqcup \mathcal{U}_\alpha \times F$ be a disjoint union. For $(x, y) \in \mathcal{U}_\alpha \times F$ and $(x', y') \in \mathcal{U}_\beta \times F$ define $(x, y) \sim (x', y') \Leftrightarrow x' = x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and $y' = \rho_{\beta\alpha}(y)$. From the definition of cocycle we have that $\sim$ is an equivalence relation. Let $E = \tilde{E}/\sim$ be the quotient space and $q : \tilde{E} \to E$ the quotient map. If $\tilde{\pi} : \tilde{E} \to E$ is the projection $\tilde{\pi}(x, y) = x$ we have that $q(\tilde{p}) = q(\tilde{q}) \Rightarrow \tilde{\pi}(\tilde{p}) = \tilde{\pi}(\tilde{q})$. Then $\tilde{\pi}$ induces a map $\pi : E \to \mathcal{M}$. The composition of q with the inclusion $\mathcal{U}_\alpha \times F \to \tilde{E}$ defines a homomorphism $U_\alpha \times F \to \pi^{-1}(\mathcal{U}_\alpha) \subset E$ and a fiber bundle structure on E. $\square$

The group G is also called **structural group** of the fiber bundle. Summarizing, the above proposition states that a representation and a cocycle characterize the fiber bundle. In particular we have:

Proposition 2. *Let G a Lie group, $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ an open cover of $\mathcal{M}$ and $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ a cocycle. Then we can reconstruct the principal bundle $\pi : P \to \mathcal{M}$ whose transition functions corresponding to $\{\mathcal{U}_\alpha\}_{\alpha\in I}$. are $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$.*

Proof: Similar to the last proposition. □

Definition 7. *A **morphism** between fiber bundles $\pi_E : E \to \mathcal{M}$ and $\pi_{E'} : E' \to \mathcal{M}'$ is a pair of maps $\tilde{f} : E \to E'$, $f : \mathcal{M} \to \mathcal{M}'$ such that $f \circ \pi_E = \pi_{E'} \circ \tilde{f}$. We require that $\tilde{f}$ be compatible with the structure of the fibers (e.g. in the case of vector bundles the restriction to each fiber needs to be linear). If $\mathcal{M} = \mathcal{M}'$, $\tilde{f}$ is a diffeomorphism and $f = id$, i.e., $\pi_E = \pi_{E'} \circ \tilde{f}$ then we say that the bundles E and E' are **equivalent**[2].*

1.1.1 Examples

Example 1. ***Trivial bundle.***

A product manifold $\mathcal{M} \times G$ is turned into a principal bundle when it is provided with the right action of G on itself in such a way that $(x, u) \mapsto (x, ug)$, $x \in \mathcal{M}$, $u, g \in G$. It is called a **trivial principal bundle**. A principal bundle is trivial if and only if it admits a section defined on all $\mathcal{M}$ ([2] p. 36). Therefore non-trivial principal bundles can have only local sections. It is obvious that given manifolds $\mathcal{M}$ and F it is possible to endow $\mathcal{M} \times F$ with the structure of a fiber bundle which is called **trivial bundle**. Any bundle equivalent (in the sense of definition 7) to the trivial bundle is also called trivial. So in statements such as "if (...) then P is trivial" we mean equivalent to the trivial bundle.

Example 2. ***Tangent bundle.***

If $\mathcal{M}$ is a manifold and $\{\phi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n\}$ is an atlas on $\mathcal{M}$ we define $\gamma_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(\mathbb{R}^n)$ by:

$$\gamma_{\alpha\beta}(x) = d(\phi_\alpha \circ \phi_\beta^{-1})(\phi_\beta(x))$$

[2] The classification of bundles according to their equivalence classes is a field of study in its own; see [2]

which is a cocycle. We have a representation:

$$\rho = \text{identity} : GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$$

The tangent bundle $T\mathcal{M}$ is defined by $\gamma_{\alpha\beta}$ and ρ. The sections of $T\mathcal{M}$ are the vector fields, i.e., $\Gamma(T\mathcal{M}) = \mathfrak{X}(\mathcal{M})$.

The tangent bundle is a fiber bundle over $\mathcal{M}$ whose fiber is $\mathbb{R}^n$ and whose structural group is $GL(\mathbb{R}^n)$ — hence a vector bundle.

Example 3. ***Cotangent bundle.***

Using $(\mathbb{R}^n)^* \approx \mathbb{R}^n$ and proceeding analogously to the above example we get a fiber bundle whose fiber is $(\mathbb{R}^n)^*$ instead of $\mathbb{R}^n$. This is the cotangent bundle $T^*\mathcal{M}$.

The cotangent bundle is a fiber bundle over $\mathcal{M}$ whose fiber is $(\mathbb{R}^n)^*$ and whose structural group is $GL(\mathbb{R}^n)$ — hence a vector bundle.

Example 4. ***Tensor bundle.***

Denote by $\mathcal{T}^{r,s}$ the space of r-covariant and s-contravariant tensors:

$$\mathcal{T}^{r,s} = \underbrace{\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n}_{r} \times \underbrace{(\mathbb{R}^n)^* \otimes \ldots (\mathbb{R}^n)^*}_{s}$$

This space is isomorphic to the space of multi-linear maps:

$$L(\underbrace{(\mathbb{R}^n)^*, \ldots, (\mathbb{R}^n)^*}_{r}, \underbrace{\mathbb{R}^n, \ldots, \mathbb{R}^n}_{s}; \mathbb{R})$$

Given $T \in \mathcal{T}^{r,s}$ and $\phi \in GL(\mathbb{R}^n)$ the pull-back is defined by:

$$(\phi^* T)(\lambda_1, \ldots, \lambda_r, v_1, \ldots, v_s) = T(\lambda_1 \circ \phi, \ldots, \lambda_s \circ \phi, \phi(v_1), \ldots, \phi(v_r))$$

We have that $\phi^* : \mathcal{T}^{r,s} \to \mathcal{T}^{r,s}$ is linear and:

$$(\phi \circ \psi)^* = \psi^* \circ \phi^*$$
$$(\phi^*)^{-1} = (\phi^{-1})^*$$

If $\rho : G \to GL(\mathbb{R}^n)$ is a representation then:

$$\rho_* : G \to \mathcal{T}^{r,s}, \;\; \rho_* g = \rho(g^{-1})^* \tag{1.1}$$

is also a representation. The bundle of r-covariant and s-contravariant tensors, or tensor bundle for short, is defined by the cocycle $\{\gamma_{\alpha\beta}\}$ of example 2 and the representation ρ_* of $GL(\mathbb{R}^n)$. The sections of this bundle are called **tensor fields** or simply **tensor**.

Notice that $\mathcal{T}^{1,0}(\mathcal{M})$ coincides with $T\mathcal{M}$ and $\mathcal{T}^{0,1}(\mathcal{M})$ with $T^*\mathcal{M}$. We have that $T \in \Gamma(\mathcal{T}^{r,s}(\mathcal{M}))$ if and only if:

$$T : \underbrace{\mathfrak{X}(\mathcal{M}) \otimes \cdots \otimes \mathfrak{X}(\mathcal{M})}_{r} \times \underbrace{T^*\mathcal{M} \otimes \cdots \otimes T^*\mathcal{M}}_{s} \to \mathbb{R}$$

is multi-linear and

$$T(X_1, \ldots, fX_i, \ldots, X_r, \lambda_1, \ldots, \lambda_s) = fT(X_1, \ldots, X_i, \ldots, X_r, \lambda_1, \ldots, \lambda_s)$$
$$T(X_1, \ldots, X_r, \lambda_1, \ldots, f\lambda_i, \ldots, \lambda_s) = fT(X_1, \ldots, X_r, \lambda_1, \ldots, \lambda_i, \ldots, \lambda_s)$$

for every $f \in C^\infty(\mathcal{M})$

Since the subspace $\Lambda^s(\mathbb{R}^n) \subset \mathcal{T}^{0,s}(\mathbb{R}^n)$ of skew linear forms is invariant under the representation ρ_* we have a sub-bundle $\bigwedge^s(T^*\mathcal{M})$ of $\mathcal{T}^{0,s}(\mathcal{M})$. The sections of $\bigwedge^s(T^*\mathcal{M})$ are the s-differential forms. We warn the reader that physicists sometimes write $\bigwedge^s(T^*\mathcal{M})$ for the space of sections $\Gamma(\bigwedge^s(T^*\mathcal{M}))$. The space of section of $\bigwedge^s(T^*\mathcal{M})$, i.e, $\Gamma(\bigwedge^s(T^*\mathcal{M}))$, is also written $\Omega(\mathcal{M})$.

The tensor bundle is a fiber bundle over $\mathcal{M}$ whose fiber is $\mathcal{T}^{r,s}$ and whose structural group is $GL(\mathbb{R}^n)$ — hence a vector bundle.

Of particular interest is the case of $(1,1)$-tensors: the fiber is $\mathbb{R}^n \otimes (\mathbb{R}^n)^*$ which is isomorphic to $\mathrm{End}(\mathbb{R}^n)$. If we denote by E the fiber bundle with fiber $\mathbb{R}^n$ then we denote by $\mathrm{End}(E)$ that with fiber $\mathrm{End}(\mathbb{R}^n)$ — and we call it the bundle of endomorphisms. It is important in the study of the curvature of some bundles and will appear in the Yang-Mills equations.

Example 5. ***Frame bundle.***

The cocycle is the same of example 2 and the representation given by left composition. A point of the principal bundle $L(\mathcal{M})$ corresponds to a point $x \in \mathcal{M}$ and a basis $\{v_1(x), \ldots, v_n(x)\}$ of $T\mathcal{M}_x$. The right action is given by:

$$R_L(x, \{v_1(x), \ldots, v_n(x)\}) = (x, \{L^{-1}(v_1(x)), \ldots, L^{-1}(v_n(x))\}), \;\; L \in GL(\mathbb{R}^n)$$

The frame bundle is a principal bundle over $\mathcal{M}$ whose structural group is $GL(\mathbb{R}^n)$.

Example 6. ***Orthonormal frame bundle.***

The construction is similar to the last example, but we take as group $SO(n)$ instead of $GL(\mathbb{R}^n)$ and $\mathcal{M}$ is an oriented Riemannian manifold.

Construction of the cocycle: take $\{\phi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}^n\}$ an positive atlas. Given $\{v_1(x), \ldots, v_n(x)\}$ basis of $T\mathcal{M}_x$. $d\phi_\alpha(x)v^\alpha_\beta(x) = \frac{\partial}{\partial x_\beta}$. Using Gram-Schmidt we get an orthonormal basis $\{e^\alpha_1(x), \ldots, e^\alpha_n(x)\}$ of $T\mathcal{M}_x$. Then $\gamma_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to SO(n)$ is given by $\gamma_{\alpha\beta}(x)$=matrix of coordinate changing from $\{e^\alpha_1(x), \ldots, e^\alpha_n(x)\}$ to $\{e^\beta_1(x), \ldots, e^\beta_n(x)\}$.

The orthonormal frame bundle is a principal bundle over $\mathcal{M}$ whose structural group is $SO(n)$.

Example 7. ***Hopf bundles.***

Define the right action:

$$S^3 \times S^1 \to S^3, \qquad ((z_1, z_2), \lambda) \mapsto (\lambda z_1, \lambda z_2)$$

and the projection:

$$\pi : S^3 \to \mathbb{CP}^1, \qquad (z_1, z_2) \mapsto \{(\lambda z_1, \lambda z_2), \lambda \in \mathbb{C}\}$$

Then we get a principal bundle whose structural group is $S^1 = \{z \in \mathbb{C}; |z| = 1\}$, the base space is $S^2 \approx \mathbb{CP}^1$ and the total space is $S^3 = \{(z_1, z_2) \in \mathbb{C} \times \mathbb{C}; |z_1|^2 + |z_2|^2 = 1\}$.

Analogously, consider the group of quaternions:

$$q = x_0 + x_1 I + x_2 J + x_3 K, \ x_0, x_1, x_2, x_3 \in \mathbb{R}$$
$$I^2 = J^2 = K^2 = -1$$
$$IJ = K, \ JK = I, \ KI = J$$
$$IJ = -JI, \ JK = -KJ, \ KI = -IK$$
$$|q|^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2$$

Take $S^3 = \{q \in \mathbb{H} : |q| = 1\}$; S^7=unit sphere in $\mathbb{H} \times \mathbb{H}$ and $S^4 = \mathbb{HP}^2$ the one-dimensional quaternionic subspace in $\mathbb{H} \times \mathbb{H}$.

We have the right action:

$$S^7 \times S^3 \to S^7$$
$$((z_1, z_2), \lambda) \mapsto (z_1 \lambda, z_2 \lambda)$$

and the projection:

$$\pi : S^7 \to \mathbb{HP}^2 = S^4$$
$$(z_1, z_2) \mapsto \{\lambda z_1, \lambda z_2; \lambda \in \mathbb{H}\}$$

We obtain a principal bundle whose structural group is S^3, the base space is S^4 and the total space S^7.

Example 8. ***Grassmanian bundle.***

The base space is $\mathcal{M} = G(n,k) = \{L \subset \mathbb{R}^n : L$ is a k-dimensional subspace $\}$; the total space is $E = \tilde{G}(n,k) = \{(L, y) : L \in G(n,k),\ y \in L\}$ and the projection $\pi : \tilde{G}(n,k) \to G(n,k)$, $(L, y) \mapsto L$. $G(n,k)$ is a compact $k(n-k)$-dimensional manifold

$\pi : \tilde{G}(n,k) \to G(n,k)$ is a vector bundle with structural group $GL(k)$.

1.1.2 Bundle constructions

It is natural to know how to construct new bundles from old ones.

Proposition 3. *Let $\pi : P \to \mathcal{M}$ a principal bundle and F a manifold in which G acts on the left. Define an action on the right on $P \times F$ by $(u, f) \mapsto (u, f)g := (ug, \rho(g)^{-1}(f))$, where $\rho : G \to Diff(F)$ is a representation. Then $E := (P \times F)/G$ is a fiber bundle with fiber F and base space $\mathcal{M}$.*

Remark 3. *It is usual to write gf or $g \cdot f$ instead of $\rho(g)(f)$.*

Proof: We are identifying $(u, f) \sim (ug, g^{-1}f)$. The projection is given by $\pi_E((u, f)G)) = \pi(u)$. $\pi^{-1}(\mathcal{U}) \approx \mathcal{U} \times G$ induces a diffeomorphism $\pi_E^{-1}(\mathcal{U}) \approx \mathcal{U} \times F$. Explicitly, given local sections $\sigma_\alpha : \mathcal{U}_\alpha \to P$ and $\sigma_\beta : \mathcal{U}_\beta \to P$ define $\phi_\alpha : \mathcal{U}_\alpha \times F \to \pi_E^{-1}(\mathcal{U}_\alpha)$ as $\phi_\alpha(x, f) = [(\sigma_\alpha(x), f)]$, where $[(\sigma_\alpha(x), f)]$ is the class of $(\sigma_\alpha(x), f)$. Put $\phi_{\alpha,x} := \phi_\alpha|\{x\} : F \to \pi_E^{-1}(x)$. For each orbit on $\pi^{-1}(x)$ there exists an unique $f \in F$ such that the orbit passes by $(\sigma_\alpha(x), f)$, hence $\phi_{\alpha,x}$ is bijective and so it is ϕ_α. Since we have $\sigma_\beta(x) = \sigma_\alpha(x) g_{\alpha\beta}(x)$, where $g_{\alpha\beta}(x) : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ is a transition function and $[(hg, f)] = [(h, gf)]$ we obtain that $\phi_\alpha^{-1} \circ \phi_\beta(x, f) = (x, g_{\alpha\beta}(x)f)$. □

Definition 8. *The fiber bundle $\pi : E \to \mathcal{M}$ in proposition is called* ***fiber bundle associated with*** *$\pi : P \to \mathcal{M}$* ***and with the representation*** *ρ or simply* ***associated bundle****.*

From the proof, we see that each $u \in P$ gives an isomorphism (which we also denote by u), as $u : F \to E_{\pi(u)},\ f \mapsto [u, f]$. This isomorphism satisfies $ug(f) = u(gf),\ g \in G$.

We remark that instead of starting with the principal bundle and then construct the associated bundle, we may start with a fiber bundle $\pi : E \to \mathcal{M}$ then construct a principal bundle $\pi : P \to \mathcal{M}$ such that $\pi : E \to \mathcal{M}$ is associated with $\pi : P \to \mathcal{M}$. For example, we may start with the tangent bundle and then construct the principal frame bundle (see the examples above). For details see [3] p. 41.

Theorem 1. *Let $\pi : E \to \mathcal{M}$ and $\pi' : E' \to \mathcal{M}$ be vector bundles with fibers F and F' and transition functions $\{g_{\alpha\beta}\}$, $\{h_{\alpha\beta}\}$ respectively. Then there are bundles over $\mathcal{M}$ with fibers $F \oplus F'$, $F \otimes F'$, F^* and $\bigwedge^p F^*$*

Proof: From proposition 1, it suffices to provide the cocycles. We have:

$$k_{\alpha\beta} = \begin{pmatrix} g_{\alpha\beta} & 0 \\ 0 & h_{\alpha\beta} \end{pmatrix} \in GL(F \oplus F')$$

is a cocycle which gives rise to a bundle with fiber $F \oplus F'$. This bundle is called **direct sum of bundles** and it is denoted by $\pi : E \oplus E' \to \mathcal{M}$.

$$k_{\alpha\beta} = g_{\alpha\beta} \otimes h_{\alpha\beta} \in GL(F \otimes F')$$

is a cocycle which gives rise to a bundle with fiber $F \otimes F'$. This bundle is called **tensor product of bundles** and it is denoted by $\pi : E \otimes E' \to \mathcal{M}$.

$$k_{\alpha\beta} = (\bar{g}_{\alpha\beta})^t \in GL(F)$$

where $^-$ denotes complex conjugation (in case of complex vector space) and ^{t} is the transpose. $k_{\alpha\beta}$ is a cocycle which gives rise to a bundle with fiber F^*. This bundle is called **dual bundle** and it is denoted by $\pi : E^* \to \mathcal{M}$.

$$k_{\alpha\beta} = \bigwedge^p (g_{\alpha\beta}) \in GL(\bigwedge^p F^*)$$

is a cocycle which gives rise to a bundle with fiber $\bigwedge^p(F^*)$. This bundle is called **bundle of forms, exterior bundle etc** and it is denoted by $\pi : \bigwedge^p(E^*) \to \mathcal{M}$. □

Definition 9. *A set $e_1, \ldots, e_n$ of sections of a vector bundle $\pi : E \to \mathcal{M}$ is a **base of sections** if any $\sigma \in \Gamma(E)$ can be written as $\sigma = f^\alpha e_\alpha,\ f^\alpha \in C^\infty(\mathcal{M})$.*

Locally there always exists a base of sections, but it exists globally if and only if E is trivial (use $(x, v) \mapsto v^\alpha e_\alpha(x)$). Do not confuse base of sections with sections: global sections always exist for vector bundles because at least the null sections exists, although a basis of sections may exist only locally. On the other hand, as we mentioned in the examples, for a principal bundle, there exists a global section if and only if the principal bundle is trivial.

Let G be Lie group. Define a representation $ad : G \to Diff(G)$ by $ad(g) : G \to G,\ h \mapsto ad(g)(h) := ad_g(h) := ghg^{-1}$. Since $ad_g(e) = e$, differentiating at the identity we get an isomorphism: $Ad_g := Dad_g(e) \equiv ad_{g_*} : \mathfrak{g} \xrightarrow{\approx} \mathfrak{g}$, where $\mathfrak{g} \approx TG_e$ is the Lie algebra of G. Notice that $Ad_{gh} = Ad_g \circ Ad_h$ (use the chain rule). Therefore we have a representation $Ad : G \to \mathrm{Aut}(\mathfrak{g}),\ g \mapsto Ad_g$.

Definition 10. *The representations ad and Ad are (both) called **adjoint representations**. The context will differ between ad and Ad when "adjoint representation" is referred to.*

Given a principal bundle $\pi : P \to \mathcal{M}$, Ad defines a left action of G on $\mathfrak{g}$. Therefore, by proposition 3 we have a vector bundle $AdP := (P \times \mathfrak{g})/G$ with fiber $\mathfrak{g}$.

Definition 11. *AdP is called the **adjoint bundle** of P.*

The following will be useful in chapter 3:

Definition 12. *Given two fiber bundles over $\mathcal{M}$, $\pi_1 : E_1 \to \mathcal{M}$ and $\pi_2 : E_2 \to \mathcal{M}$ their **fibered product** or **Whitney sum** is the fiber bundle $q : E_1 \tilde{\times} E_2$, where $E_1 \tilde{\times} E_2$ is the subspace of all pairs $(x_1, x_2) \in E_1 \times E_2$ such that $\pi_1(x_1) = \pi_2(x_2)$, and $q(x_1, x_2) = \pi_1(x_1) = \pi_2(x_2)$. It follows that the fiber over p is $\pi_1^{-1}(p) \times \pi_2^{-1}(p)$. Because this generalizes the direct sum of vector bundles sometimes the fiber product is written as $E_1 \oplus E_2$.*

Definition 13. *Suppose that $\pi : P \to \mathcal{M}$ is a principal bundle and that $f : \mathcal{N} \to \mathcal{M}$ is a continuous map. Then we can form the* ***pullback bundle*** *$\pi : f^*P \to \mathcal{N}$ — which is a principal bundle with same structural group G — as the fibered product of the following diagram:*

$$\begin{array}{ccc} & & P \\ & & \downarrow \pi \\ \mathcal{N} & \xrightarrow{f} & \mathcal{M} \end{array}$$

i.e., $f^*P \subset \mathcal{N} \times P$ is the set of pairs $\{(a,p) \in \mathcal{N} \times P : f(a) = \pi(p)\}$. The action of G on the total space is induced fro the action of G on P. The pullback of an associated bundle is defined analogously.

1.2 Connections on fiber bundles

Connections will give us a way of differentiating sections. It will also allow us to define the curvature, which is a measure of the non-triviality of the bundle.

Definition 14. *Let $\pi : P \to \mathcal{M}$ be a principal bundle. Denote by G its structural group. If G_x is the fiber over $x \in \mathcal{M}$ we define the* ***vertical space at*** *$u \in G_x$, denoted by V_uP as the subspace of TP_u tangent to G_x.*

V_uP can be constructed as follows: let $A \in \mathfrak{g}$ be any vector. $R_{\exp(tA)}u = u\exp(tA),\ t \in (-\epsilon, \epsilon)$ defines a path in P passing through u in $t = 0$. Since $x = \pi(u) = \pi(R_{\exp(tA)}u)$, this path is contained in G_x. Define $A^{\#}(u) \in T(G_x)_u \subset TP_u$ by $\frac{d}{dt}(R_{\exp(tA)}u)|_{t=0}$.

Definition 15. *Defining $A^{\#}$ in every $u \in P$ we have a vector field called* ***fundamental vector field*** *(of A).*

It follows that $\# : \mathfrak{g} \to V_uP,\ A \mapsto A^{\#}(u)$ gives an isomorphism of vector spaces $\mathfrak{g} \approx V_uP$. It is easy to verify that $\#$ preserves the Lie bracket ($[A,B]^{\#} = [A^{\#}, B^{\#}]$) and hence it is an isomorphism of algebras. Notice that V_uP is also given by $V_uP = \ker(D\pi(u))$. It can be shown that $DR_g(u)(A^{\#}(u))$ is the fundamental vector field corresponding to $Ad_{g^{-1}}(A) \in \mathfrak{g}$ ([4] p. 81).

Definition 16. *If we have a decomposition $V_uP \oplus H_uP = TP_u$, $u \in P$, then H_uP is called* ***horizontal space*** *at u.*

The horizontal and vertical spaces give us sub-bundles of TP denoted by HP and VP respectively. We shall say that a vector field is horizontal or vertical according it belongs to HP or VP.

Definition 17. *A* ***connection*** *on a principal bundle $\pi : P \to \mathcal{M}$ is a family of subspaces $\{H_uP\}_{u\in P}$, $H_uP \subset TP_u$ satisfying: (i) $TP_u = V_uP \oplus H_uP$; (ii) $DR_g(u)(H_uP) = H_{ug}P$ and (iii) H_uP depends differentiably on u — together with (i) this means that a vector field X on P can be written (in an obvious notation) as $X = X_H + X_V$, where X_H, X_V are differentiable vector fields.*

Because $\pi \circ R_g = R_g \circ \pi$ it follows directly that $\pi_H \circ DR_g = DR_g \circ \pi_H$ and that $\pi_V \circ DR_g = DR_g \circ \pi_V$ ([5] p. 51). An application which associates to each point a subspace of the tangent space is called a **distribution**.

Proposition 4. *$D\pi(u)|H_uP$ gives an isomorphism $H_uP \approx T\mathcal{M}_{\pi(u)}$.*

Proof: π is a submersion[3] and H_uP is complementary to $\ker(D\pi) \approx V_uP$. □

The following definitions aims to fix some notation and terminology.

Definition 18. *Recall that we denote by $\Gamma(\cdot)$ the space of sections of some bundle. If F is a vector space and E is a vector bundle over $\mathcal{M}$ we shall write $F \otimes E$ to denote the tensor product of E with the trivial bundle $\mathcal{M} \times F$, hence $\Gamma(F \otimes E)$ denotes sections of this tensor bundle[4]. Sections of the bundle $E \otimes \bigwedge^p(T^*\mathcal{M})$ are called E-**valued (differential) p-forms**. If E is trivial, $E = \mathcal{M} \times F$, then call elements of $\Gamma(E \otimes \bigwedge^p(T^*\mathcal{M})) \equiv \Gamma(F \otimes \bigwedge^p(T^*\mathcal{M}))$ F-**valued (differential) p-forms**. In other words, when $E = \mathcal{M} \times F$, we slightly abuse the terminology and write $F \otimes \bigwedge^p(T^*\mathcal{M})$ instead of $E \otimes \bigwedge^p(T^*\mathcal{M})$ and call $\Gamma(E \otimes \bigwedge^p(T^*\mathcal{M}))$ F-valued forms instead of E-valued forms.*

Proposition 5. *Given a connection on P there exists an unique $\mathfrak{g}$ valued 1-form w, i.e., an element of $\Gamma(\mathfrak{g} \otimes T^*P)$, such that $w(A^\#) = A$ and $w(X_H) = 0$, for every $A \in \mathfrak{g}$ and every $X_H \in HP$.*

---

[3] there exists only one differentiable structure on $\mathcal{M}$ which makes π a submersion, see [5] p. 50.

[4] Here some authors use the notation $F \otimes \Gamma(E)$, but this is also an abuse of notation, actually meaning $C^\infty(\mathcal{M}, F) \otimes \Gamma(E)$ (see remark 5). Different authors use different notations, we follow to some extent the notation of [6].

Proof: Simply put $w : TP_u \xrightarrow{\pi_V} V_uP \xrightarrow{\#^{-1}} \mathfrak{g}$. □

Definition 19. *w above is called the* ***connection 1-form*** *or sometimes* ***Ehresmann connection***

Notice that the use of "the" in the sentence "the connection 1-form" might be a bit misleading: it is true that for a *given* connection in P there exists a unique w satisfying the conditions of proposition 5. We can, however, have different connections in P what will give rise to different connection one forms. Notice, also, that a connection one form is more than just an element of $\Gamma(\mathfrak{g} \otimes T^*P)$: it is a section of $\mathfrak{g} \otimes T^*P$ which satisfy the properties of proposition 5. Some authors introduce another notation for the space of connection one forms, we shall keep using the notation $w \in \Gamma(\mathfrak{g} \otimes T^*P)$ always bearing in mind that w satisfies the extra conditions of 5.

Now it is not difficult to see that given an element $w \in \Gamma(\mathfrak{g} \otimes T^*P)$ which satisfies the conditions of proposition 5 we can obtain a connection in P whose corresponding connection one form is exactly w. This is done by defiding the horizontal spaces $H_uP \subset TP_u$ as the kernel of $w_u : TP_u \to \mathfrak{g}$ (see [7] for more details). Therefore the converse of proposition 5 is true and the study of connections in P can be accomplished by the study of connection one forms.

Proposition 6. $R_g^*(w) = Ad_{g^{-1}} \circ w$

Proof: Since $X = X_H + X_V$ and w is linear, it suffices to check separately. X is horizontal: then $R_g^*(w)(u)(X(u)) = w(ug)(DR_g(u)(X(u))) = w(ug)(X(ug)) = 0$. On the other hand $w(u)(X(u)) = 0$ and therefore $Ad_{g^{-1}}(w(u)(X(u))) = Ad_{g^{-1}}(0) = 0$. Suppose now that X is vertical. Then it is the fundamental vector field corresponding to some $A \in \mathfrak{g}$ and we get $R_g^*(w)(u)(X(u)) = R_g^*(w)(u)(A^\#(u)) = w(ug)(DR_g(u)(A^\#(u)))$; this last term equals $w(ug)((Ad_{g^{-1}}(A))^\#)$ because $DR_g(u)(A^\#(u))$ is the fundamental vector field corresponding to $Ad_{g^{-1}}(A)$. Then $R_g^*(w)(u)(X(u)) = w(ug)((Ad_{g^{-1}}(A))^\#) = Ad_{g^{-1}}(A) = Ad_{g^{-1}}(w(u)(A^\#(u)))$, where we used that $w(V^\#) = V$ □

Theorem 2. *Let $\mathcal{U}_\alpha$,$\mathcal{U}_\beta$ open sets on $\mathcal{M}$ and $\sigma_\alpha, \sigma_\beta$ local sections. Define $\mathcal{A}_\alpha = \sigma_\alpha^* w$ and $\mathcal{A}_\beta = \sigma_\beta^* w$ (which are $\mathfrak{g}$-valued one forms on open sets of $\mathcal{M}$).*

On $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ *we have* $\mathcal{A}_\beta = Ad_{g^{-1}_{\alpha\beta}} \circ \mathcal{A}_\alpha + g^{-1}_{\alpha\beta} Dg_{\alpha\beta}$*, where* $g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ *is the transition functions.*

Remark 4. *Notice that there is no sense in multiplying* $g^{-1}_{\alpha\beta} Dg_{\alpha\beta}$*. Here* $g^{-1}_{\alpha\beta} Dg_{\alpha\beta}$ *is a notation for the following:* $Dg_{\alpha\beta}(x) : T(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)_x \to TG_{g_{\alpha\beta}(x)}$*. Then by* $g^{-1}_{\alpha\beta}(x) Dg_{\alpha\beta}(x)(v)$*,* $v \in T(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)_x$*, we mean the vector in* $TG_e \approx \mathfrak{g}$ *obtained by left translation of* $Dg_{\alpha\beta}(x)(v)$*, i.e,* $g^{-1}_{\alpha\beta}(x) Dg_{\alpha\beta}(x) = D(L_{g^{-1}_{\alpha\beta}} \circ g_{\alpha\beta})(x) : T(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)_x \to TG_e \approx \mathfrak{g}$*, where* $L_{g^{-1}_{\alpha\beta}(x)} : h \mapsto g_{\alpha\beta}(x)^{-1} h$ *is the left translation by* $g^{-1}_{\alpha\beta}$*. The notation is justified by the fact that for* matrix groups $g^{-1}_{\alpha\beta} Dg_{\alpha\beta}$ *corresponds to ordinary multiplication of matrices, but most physics books simply write the formula in its full generality without mentioning this abuse of notation (see [5] p. 52).*

In order to prove the above theorem we must remember some elementary facts about differentiating:

Definition 20. *Let* M, N *and* W *be differentiable manifolds and* $p : M \times N \to W$*,* $(x, y) \mapsto p(x, y)$ *a differentiable map. The restriction to* TM_x *of the differential* $Dp(x, y)$ *is denoted by* $\frac{\partial p}{\partial x}(x, y) : TM_x \to TW_{p(x,y)}$ *and it is called* ***partial derivative of*** p ***with respect to*** x*; we define* $\frac{\partial p}{\partial y}$ *analogously. Notice that* $Dp(x, y)(X, Y) = \frac{\partial p}{\partial x}(x, y)(X) + \frac{\partial p}{\partial y}(x, y)(Y)$*. The generalization to* $M_1 \times \cdots \times M_k$ *is obvious.*

Lemma 1. *Let* M *be a differentiable manifold and let* G *be Lie group; let* $g, h : M \to G$ *and* $j : M \to G \times G$ *be such that* $j(x) = (g(x), h(x))$*; put* $f = p \circ j$*, where* $p : G \times G \to G$ *is the multiplication of* G*. Then* $Df(x)(X) = \frac{\partial p}{\partial u}(g(x), h(x))(Dg(x)(X)) + \frac{\partial p}{\partial v}(g(x), h(x))(Dh(x)(X))$*, where* $(u, v) \in G \times G$ *and* $X \in TM_x$*.*

Proof: First observe that $Dj(x)(X) = (Dg(x)(X), Dh(x)(X))$; indeed, if $\gamma : (-\epsilon, \epsilon) \to M$ is a differentiable path such that $\gamma(0) = x$ and $\gamma'(0) = X$ we have $\frac{d}{dt}(j \circ \gamma(t))|_{t=0} = \frac{d}{dt}(g \circ \gamma(t), h \circ \gamma(t))|_{t=0} = ((g \circ \gamma)'(0), (g \circ \gamma)'(0)) = (Dg(x)(X), Dh(x)(X))$. Then:

$$\begin{aligned} Df(x)(X) &= Dp(j(x)) \circ Dj(x)(X) = Dp(g(x), h(x))(Dg(x)(X), Dh(x)(X)) \\ &= \frac{\partial p}{\partial u}(g(x), h(x))(Dg(x)(X)) + \frac{\partial p}{\partial v}(g(x), h(x))(Dh(x)(X)) \end{aligned}$$

□

Proof of theorem 2: For $x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$ we have $\sigma_\beta(x) = \sigma_\alpha(x) g_{\alpha\beta}(x)$. As $\pi^{-1}(x) \approx G$ we identify these two spaces without mentioning it again. From the above lemma:

$$D\sigma_\beta(x)(X) = \frac{\partial p}{\partial u}(\sigma_\alpha(x), g_{\alpha\beta}(x))(D\sigma_\alpha(x)(X)) + \frac{\partial p}{\partial v}(\sigma_\alpha(x), g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X))$$

Since on $\frac{\partial p}{\partial u}(\sigma_\alpha(x), g_{\alpha\beta}(x))$ the term $g_{\alpha\beta}(x)$ is fixed we can write:

$$\frac{\partial p}{\partial u}(\sigma_\alpha(x), g_{\alpha\beta}(x)) \circ D\sigma_\alpha(x) = DR_{g_{\alpha\beta}(x)}(\sigma_\alpha(x)) \circ D\sigma_\alpha(x)$$

Analogously:

$$\frac{\partial p}{\partial v}(\sigma_\alpha(x), g_{\alpha\beta}(x)) \circ Dg_{\alpha\beta}(x) = DL_{\sigma_\alpha(x)}(g_{\alpha\beta}(x)) \circ Dg_{\alpha\beta}(x)$$

(notice that we have right or left translation according to we multiply the fixed element on the right or on the left). Then:

$$\begin{aligned} D\sigma_\beta(x)(X) &= DR_{g_{\alpha\beta}(x)}(\sigma_\alpha(x))(D\sigma_\alpha(x)(X)) + DL_{\sigma_\alpha(x)}(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X)) \\ &= DR_{g_{\alpha\beta}(x)}(\sigma_\alpha(x))(D\sigma_\alpha(x)(X)) + DL_{\sigma_\beta(x)g_{\alpha\beta}(x)^{-1}}(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X)) \quad (*) \end{aligned}$$

Applying w to $(*)$ and using that

$$w(\sigma_\beta(x)) = w(\sigma_\alpha(x) g_{\alpha\beta}(x)) = w(R_{g_{\alpha\beta}(x)}(\sigma_\alpha(x))$$

we have:

$$\begin{aligned} w(\sigma_\beta(x))(D\sigma_\beta(x)(X)) = w(R_{g_{\alpha\beta}(x)}(\sigma_\alpha(x)))\Big(DR_{g_{\alpha\beta}(x)}(\sigma_\alpha(x))(D\sigma_\alpha(x)(X))\Big) + \\ w(\sigma_\beta(x))\Big(DL_{\sigma_\beta(x)g_{\alpha\beta}(x)^{-1}}(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X))\Big) \quad (**) \end{aligned}$$

We analyze the first and the second terms on the right side of the above expression. The first term equals to $R^*_{g_{\alpha\beta}(x)}(w)(\sigma_\alpha(x)(D\sigma_\alpha(x)(X))$ which, by proposition 6 equals to $Ad_{g_{\alpha\beta}(x)^{-1}}\big(w(\sigma_\alpha(x))(D\sigma_\alpha(x)(X))\big)$. This last expressions is means $Ad_{g_{\alpha\beta}(x)^{-1}} \circ \sigma^*_\alpha(w)(x)(X)$ which equals $Ad_{g_{\alpha\beta}(x)^{-1}} \circ \mathcal{A}_\alpha(x)(X)$. Since $DL_{\sigma_\beta(x)g_{\alpha\beta}(x)^{-1}} = D(L_{\sigma_\beta(x)} \circ L_{g_{\alpha\beta}(x)^{-1}})$, the second term equals to $w(\sigma_\beta(x))\Big(D(L_{\sigma_\beta(x)} \circ L_{g_{\alpha\beta}(x)^{-1}})(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X))\Big)$. This can be worked out to equal:

$$w(\sigma_\beta(x))\Big(DL_{\sigma_\beta(x)}(\overbrace{L_{g_{\alpha\beta}(x)^{-1}}(g_{\alpha\beta}(x))}^{=e}) \circ \underbrace{DL_{g_{\alpha\beta}(x)^{-1}}(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X))}_{=A\in\mathfrak{g}}\Big)$$

Here A is an element of $\mathfrak{g}$ because $L_{g_{\alpha\beta}^{-1}} \circ g_{\alpha\beta} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ implies $D(L_{g_{\alpha\beta}^{-1}} \circ g_{\alpha\beta})(x) : T(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)_x \to TG_{L_{g_{\alpha\beta}^{-1}(x)}(g_{\alpha\beta}(x))} = TG_e$. Therefore the preceding expression equals to $w(\sigma_\beta(x))(DL_{\sigma_\beta(x)}(e)(A))$. But

$$DL_{\sigma_\beta(x)}(e)(A) = \frac{d}{dt}(L_{\sigma_\beta(x)}(e^{tA}))|_{t=0} = \frac{d}{dt}(\sigma_\beta(x)(e^{tA}))|_{t=0} = \frac{d}{dt}(R_{e^{tA}}(\sigma_\beta(x)))|_{t=0}$$

$$= A^\# \equiv \text{fundamental vector field corresponding to } A \text{ at } \sigma_\beta(x)$$

hence $w(\sigma_\beta(x))(DL_{\sigma_\beta(x)}(e)(A)) = w(\sigma_\beta(x))(A^\#)$. By the definition of w: $w(\sigma_\beta(x))(A^\#) = A$. Returning to $(**)$ we get: $w(\sigma_\beta(x)) = Ad_{g_{\alpha\beta}(x)^{-1}} \circ \mathcal{A}_\alpha(x)(X) + A$. As the left side of this expression is $\mathcal{A}_\beta(x)(X)$ and $A = DL_{g_{\alpha\beta}(x)^{-1}}(g_{\alpha\beta}(x))(Dg_{\alpha\beta}(x)(X)) = D(L_{g_{\alpha\beta}^{-1}} \circ g_{\alpha\beta})(x)(X)$ we get $\mathcal{A}_\beta = Ad_{g_{\alpha\beta}^{-1}} \circ \mathcal{A}_\alpha + g_{\alpha\beta}^{-1} Dg_{\alpha\beta}$. □

Definition 21. *The condition $\mathcal{A}_\beta = Ad_{g_{\alpha\beta}^{-1}} \circ \mathcal{A}_\alpha + g_{\alpha\beta}^{-1} Dg_{\alpha\beta}$ is called **compatibility condition**.*

Now, it should be clear that given *local* connection one forms $\{A_\alpha\}_{\alpha\in I}$ satisfying the compatibility conditions we can reconstruct the connection one form w. Thus we shall sometimes refer to w and A_α simply as "the connection". This is a standard use in the literature, but in order to avoid confusion, we should point out the following: w is an "honest" connection, in the sense that it is a section of a bundle over P (an element of $\Gamma(\mathfrak{g} \otimes T^*P)$). When written in local coordinates in P, w transforms according to the transition functions of the tensor bundle $\mathfrak{g} \otimes T^*P$. The compatibility conditions described above, however, *are not* a transformation law for a section of a bundle over $\mathcal{M}$; or, put differently, the family $\{A_\alpha\}_{\alpha\in I}$ *does not* define a section of a bundle over $\mathcal{M}$. In other words, while $w \in \Gamma(\mathfrak{g} \otimes T^*P)$, patching together the A_α's does not give rise an element of $\Gamma(\mathfrak{g} \otimes T^*\mathcal{M})$ (even though the original w which is deinfed in P can be reconstructed from the A_α's). The reason, of course, is that the maps σ_α — through which we pullback the connection one form w — are defined only locally , differently from what happens when we pullback a differential form via a (globally defined) map $f : \mathcal{M} \to P$.

The following remark will be used throughtout the text.

Remark 5. *Given vector bundles E_1 and E_2 over $\mathcal{M}$ we have $\Gamma(E_1 \otimes E_2) \cong \Gamma(E_1) \otimes \Gamma(E_2)$ (the tensor product on the right hand side is taken over smooth* *function on $\mathcal{M}$, i.e., $\Gamma(E_1) \otimes_{C^\infty(\mathcal{M})} \Gamma(E_2)$ and hence this is an isomorphism of $C^\infty(\mathcal{M})$-modules). Therefore any element of $\Gamma(E_1 \otimes E_2)$ can be written as a linear combination of elements of the form $\sigma \otimes \tau$, $\sigma \in \Gamma(E_1)$, $\tau \in \Gamma(E_2)$. In particular, if one of the bundles, say E_1 is a trivial bundle $E_1 = \mathcal{M} \times F$, we can choose a global basis of sections for $\Gamma(E_1)$ by picking a basis $\{e_\ell\}$ of F and then defining the constant sections $\sigma_\ell(x) = e_\ell$. In this case any section of $\Gamma(E_1 \otimes E_2)$ can be written as a linear combination of elements of the form $\sigma_\ell \otimes \tau_\ell$ (notice that we are not saying that the sections τ_ℓ are a basis for the sections of the bundle E_2. All we are using in writing this is the aforementioned isomorphism and the triviality of E_1).*

Let F be a vector space and consider F-valued p-forms on P, i.e, sections of $\Gamma(F \otimes \bigwedge^p(T^*P))$. Pick a basis r_ℓ of F. Then any $\nu \in \Gamma(F \otimes \bigwedge^p(T^*P))$ can be written as $\nu = r_\ell \otimes \nu^\ell$ (by remark 5 above both "components" of the tensor product can run over the same index ℓ even though F and $\bigwedge^p(T^*P_x)$ in general have different dimensions.)

Definition 22. ***Exterior derivative:*** *In the previous notation and assumptions: $d\nu = r_\ell \otimes d\nu^\ell$.*

Notice that in defining the exterior of F-valued forms we are explicitely using the triviality of the bundle $\mathcal{M} \times F$. We notice that there is not a canonical way of defining the exterior derivative of forms with values in arbitrary bundles.

Definition 23. *Given a connection w on a principal fiber bundle $\pi : P \to \mathcal{M}$ we define the* ***exterior covariant derivative*** *$D^w : \Gamma(\mathfrak{g} \otimes \bigwedge^p(T^*P)) \to \Gamma(\mathfrak{g} \otimes \bigwedge^{p+1}(T^*P))$ as $D^w\nu(X_1, \ldots, X_{p+1}) := d\nu(\pi_H(X_1), \ldots, \pi_H(X_{p+1}))$, where (according to the last definition) $d\nu := e_\ell \otimes d\nu^\ell$ and $\pi_H : TP \to HP$ is the horizontal projection on the sub-bundle given by the connection.*

Notice D^w depends on the connection since there is a projection on the horizontal space given by it.

Definition 24. *The* ***curvature of a connection*** *on a principal bundle $\pi : P \to \mathcal{M}$ is an element of $\Gamma(\mathfrak{g} \otimes \bigwedge^2(T^*P))$ defined as $\Omega := D^w w$. We write $\Omega(w)$, Ω_w etc if we want to stress the dependence on w.*

Proposition 7. *$R_g^*\Omega = Ad_{g^{-1}} \circ \Omega$.*

Proof: [1] p.312. □

We do not have a wedge product for forms with values in arbitrary bundles. We can define it only when we have a natural "product" between the bundles. There are the following cases of interest. First consider an usual (i.e., $\mathbb{R}$-valued) p-form μ and a E-valued q-form ν. Then:

Definition 25. *The* ***wedge product*** *is defined as* $\nu \wedge \mu = e_\alpha \otimes \nu^\alpha \wedge \mu$, *where* $\nu = e_\alpha \otimes \nu^\alpha$

Consider now that we have E-valued forms and F (i.e., the fiber) is an algebra.

Definition 26. *The* ***wedge product*** *of E-valued p and q- forms is defined by* $\mu \wedge \nu = T_\ell \cdot U_m \otimes (\mu^\ell \wedge \nu^m)$, *where $\cdot$ is the product of the algebra. In particular, if F is a Lie algebra we have $\cdot = [,]$ and if it is an algebra of endomorphisms we have $\cdot = \circ =$ usual composition of endomorphisms.*

Remark 6. *This product is C^∞-linear in each factor ([8] p. 258).*

Another way of writing this is the following:

Proposition 8. *For E-valued p and q forms μ and ν we have: that* $\mu \wedge \nu(X_1, \dots, X_{p+q})$ *equals to*[5]*:*

$$\frac{1}{p!q!} \sum_{\sigma \in S(p+q)} \operatorname{sign}(\sigma) \mu(X_{\sigma(1)}, \dots, X_{\sigma(p)}) \cdot \mu(X_{\sigma(p+1)}, \dots, X_{\sigma(p+q)})$$

Proof: This is a simple computation. Write $\mu = T_\alpha \otimes \mu^\alpha$ and $\nu = U_\beta \otimes \nu^\beta$.

$$\mu \wedge \nu(X_1, \dots, X_{p+q}) = \Big((T_\alpha \otimes \mu^\alpha) \wedge (U_\beta \otimes \nu^\beta)\Big)(X_1, \dots, X_{p+q}) =$$
$$(T_\alpha \cdot U_\beta)\Big(\mu^\alpha \wedge \nu^\beta(X_1, \dots, X_{p+q})\Big) =$$
$$\Big(T_\alpha \cdot U_\beta\Big) \frac{1}{p!q!} \sum_{\sigma \in S(p+q)} \operatorname{sign}(\sigma) \mu^\alpha(X_{\sigma(1)}, \dots, X_{\sigma(p)}) \nu^\beta(X_{\sigma(p+1)}, \dots, X_{\sigma(p+q)})$$

[5] The factorial terms, of course, come from the definitions $\omega \wedge \eta = \frac{(p+q)!}{p!q!} \operatorname{Alt}(\omega \otimes \eta)$ and $\operatorname{Alt}(\omega) = \frac{1}{p!} \sum_{\sigma \in S_p} \varepsilon(\sigma) \omega \circ \sigma$

Changing the order of sums:

$$= \frac{1}{p!q!} \sum_{\sigma \in S(p+q)} \mathrm{sign}(\sigma)(T_\alpha \cdot U_\beta)\mu^\alpha(X_{\sigma(1)}, \ldots, X_{\sigma(p)})\nu^\beta(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)})$$

$$= \frac{1}{p!q!} \sum_{\sigma \in S(p+q)} \mathrm{sign}(\sigma)\Big((T_\alpha \otimes \mu^\alpha)(X_{\sigma(1)}, \ldots, X_{\sigma(p)})\Big) \cdot \Big((U_\beta \otimes \nu^\beta)(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)})\Big)$$

$$= \frac{1}{p!q!} \sum_{\sigma \in S(p+q)} \mathrm{sign}(\sigma)\mu(X_{\sigma(1)}, \ldots, X_{\sigma(p)}) \cdot \nu(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)})$$

□

Corollary 1. *If μ is a $\mathfrak{g}$-valued 1-forms we have $\mu \wedge \mu = 2[\mu, \mu]$, where $[\mu, \mu](X, Y) = [\mu(X), \mu(Y)]$.*

Proof: For any two $\mathfrak{g}$-valued 1-forms:

$$\Big((T_\alpha \otimes \mu^\alpha) \wedge (U_\beta \otimes \nu^\beta)\Big)(X, Y) = [T_\alpha, U_\beta]\Big(\mu^\alpha \wedge \nu^\beta(X, Y)\Big) =$$
$$[T_\alpha, U_\beta](\mu^\alpha(X)\nu^\beta(Y) - \mu^\alpha(Y)\nu^\beta(X)) =$$
$$[\mu^\alpha(X)T_\alpha, \nu^\beta(Y)U_\beta] - [\mu^\alpha(Y)T_\alpha, \nu^\beta(X)U_\beta] =$$
$$[\mu(X), \nu(Y)] - [\mu(Y), \nu(X)]) = [\mu(X), \nu(Y)] + [\nu(X), \mu(Y)]$$

For $\mu = \nu$:

$$[\mu(X), \mu(Y)] + [\mu(X), \mu(Y)] = 2[\mu, \mu](X, Y)$$

□

For the unattentive reader, we should point out that for a $\mathfrak{g}$-valued one form μ we do *not* have $\mu \wedge \mu = 0$ as it happens for ordinary differential forms. The reason, of course, is due to the presence of the commutator term.

Finally, consider the product between an E-valued and an $\mathrm{End}(E)$-valued form (i.e., the fiber bundle whose fiber is $(\mathrm{End}(E))_x \approx \mathrm{End}(F)$).

Definition 27. *Let $T \otimes \mu$ be a p-form with values in $\mathrm{End}(E)$. Let $\sigma \otimes \nu$ be a q-form with values in E. We define their **wedge product**, which is a $p + q$-form with values in E, as $(T \otimes \mu) \wedge (\sigma \otimes \nu) = T(\sigma) \otimes (\mu \wedge \nu)$. For arbitrary $\mathrm{End}(E)$ and E valued forms we simply extend this product linearly.*

In all these definitions we expand the E-valued forms as tensor products of section and ordinary forms. This way of writing is not unique, but it is easy to check that everything is well defined.

Definition 28. *The **graded commutator** of F-valued p and q- forms is defined by* $\{\mu, \nu\} = \mu \wedge \nu - (-1)^{pq} \nu \wedge \mu$.

A straightforward computation shows that for a 1-form μ the graded commutator satisfies: $\mu \wedge \mu = \frac{1}{2}\{\mu, \mu\}$

Proposition 9. *(Cartan's structure equation)* $\Omega = dw + [w, w]$, *where* w *is the connection one-form (notice that corollary 1 allow us to write* $\Omega = dw + \frac{1}{2} w \wedge w$ [6]*).*

Proof: It suffices to prove for the following cases:
(i) X and Y are horizontal. In this case $w(X) = w(Y) = 0$ and the equation follows from the definition of Ω.
(ii) X and Y are fundamental vector fields. In this case $X = A^{\#}$ and $Y = B^{\#}$, $A, B \in \mathfrak{g}$. From the definition of exterior derivative:

$$dw(X, Y) = Xw(Y) - Yw(X) - w([X, Y]) = -w([X, Y])$$

where the last equality holds because $w(X) = w(A^{\#}) = A$ and $w(Y) = w(B^{\#}) = B$ are constant functions on P. Therefore

$$dw(X, Y) = -w([A^{\#}, B^{\#}]) = -[A, B]^{\#} = -[A, B] = [w(X), w(Y)]$$
$$\text{while } \Omega(A^{\#}, B^{\#}) = D^w w(A^{\#}, B^{\#}) = dw(\pi_H(A^{\#}), \pi_H(B^{\#})) = 0$$

(iii) X is horizontal and $Y = A^{\#}$ is vertical. Then $Xw(Y) = 0$ because $w(Y)$ is constant and $Yw(X) = 0$ because $w(X) = 0$. $[X, A^{\#}]$ is also horizontal so

$$dw(X, Y) = Xw(Y) - Yw(X) - w([X, Y]) = 0$$

The result follows. □

[6] Different authors use different conventions so this equation can also be found as $\Omega = dw + w \wedge w$, $\Omega = dw + \frac{1}{2}[w, w]$ etc. For exemple, [7] *defines* $w \wedge w$ as $w \wedge w(X, Y) := [w(X), w(Y)]$. Also, some authors use $[w, w]$ to denote the graded commutaros instead.

Remark 7. *Notice that we are using the definition of exterior derivative without a normalization factor.*

$$d\mu(X_1, \dots, X_{p+1}) := \sum_{j=1}^{p+1} (-1)^{j+1} X_j \mu(X_1, \dots, \hat{X}_j \dots, X_{p+1}) +$$
$$\sum_{1 \le j < k \le p+1} (-1)^{j+k} \mu([X_j, X_k], X_1, \dots, \hat{X}_j, \dots, \hat{X}_k, \dots, X_{p+1})$$

Some authors, e.g., [3], use the definition like this, whereas others, e.g., [4], put a factor $\frac{1}{p+1}$ before the summands. Because of such conventions, we may found different expressions for Cartan's equation, such as $dw + \frac{1}{2}[w, w]$.

Definition 29. *A connection such that $\Omega = 0$ is said to be **flat**.*

We can extract relevant information about the geometry and topology of P by knowing if it admits a flat connection. First, notice that the vertical distribution $VP \subset TP$ is integrable, the fibers being exactly the integrable manifolds. If the curvature vanishes, the horizontal distribution is integrable as well. To see this, suppose that $\Omega = 0$. Then in particular $D^w w(X, Y) = 0 = dw(X, Y) + [w(X), w(Y)]$ for any pair of horizontal vectors X, Y. Since w vanishes on horizontal vectors this gives

$$0 = dw(X, Y) + w([X, Y]) =$$
$$dw(X, Y) = Xw(Y) - Yw(X) - w([X, Y]) = -w([X, Y])$$

Hence $w([X, Y]) = 0$ for any pair of horizontal vectors. Since w vanishes only on horizontal vectors, this means that $[X, Y]$ is also horizontal, so the horizontal distribution is involutive and hence integrable by Frobenius' theorem.

Being integrable the horizontal distribuition defines a foliation on P which can be used to parallel transport along the leaves of this foliation (see below for the definition of parallel transport). Combining this facts, we can actually show that P is a trivial bundle. Conversely, it is not difficult to show that a trivial bundle admits a flat connection. Hence (see [6] p 48ff, [7] p. 73ff, [9] p. 37ff)

Theorem 3. *A principal bundle admits a flat connection if and only if it is trivial.*

Proposition 10. *Given a connection on* $\pi : P \to \mathcal{M}$ *and a vector field* $X \in \mathfrak{X}^\infty(\mathcal{M})$ *there exists an unique vector field* $\tilde{X} \in \mathfrak{X}^\infty(P)$ *such that* $D\pi(\tilde{X}) = X$*. Moreover* $DR_g(\tilde{X}) = \tilde{X} \circ R_g$*.*

Proof: The existence and uniqueness follow from $T\mathcal{M}_{\pi(u)} \approx H_uP$. The second claim is true by the definition of connection (see property (ii) of the definition). □

Definition 30. *We call the vector field* $\tilde{X}$ *given by the above proposition the* ***horizontal lift*** *of* X*.*

Definition 31. *Given a differentiable path* $\gamma : [0,1] \to \mathcal{M}$ *we say that a path* $\tilde{\gamma} : [0,1] \to P$ *is the* ***horizontal lift*** *of* γ *if* $\pi \circ \tilde{\gamma} = \gamma$ *and* $\tilde{\gamma}'(t) \in H_{\tilde{\gamma}(t)}P$ *for every* $t \in [0,1]$*.*

Proposition 11. *given a path* $\gamma : [0,1] \to \mathcal{M}$ *and a point* $u_0 \in \pi^{\gamma(0)}$ *there exists a unique horizontal lift* $\tilde{\gamma}$ *of* γ *such that* $\tilde{\gamma}(0) = u_0$*.*

Proof: It follows directly from the last proposition and the existence and uniqueness of ordinary differential equations. □

Definition 32. *The* $\tilde{\gamma}$ *given in the last proposition is called the* ***horizontal lift of*** γ ***through*** u_0*. Notice that given such an horizontal lift, it is uniquely defined the point* $u_1 = \tilde{\gamma}(1) \in \pi^{-1}(\gamma(1))$*.* u_1 *is called the* ***parallel transport of*** u_0 ***through*** $\tilde{\gamma}$

Notice that if $\tilde{\gamma}$ lifts γ through u_0 then $R_g \circ \tilde{\gamma}$ lifts γ through u_0g. We have therefore a map between fibers: $\Upsilon(\tilde{\gamma}) : \pi^{-1}(\gamma(0)) \to \pi^{-1}(\gamma(1))$ such that $\Upsilon(\tilde{\gamma}) \circ R_g = R_g \circ \Upsilon(\tilde{\gamma})$.

Definition 33. *Let* $\pi : P \to \mathcal{M}$ *be a principal bundle,* F *a manifold in which* G *acts on the left and* $E = (P \times F)/G$ *the associated bundle. A* ***connection*** *on* E *is a distribution which associates to each* $p \in E$ *a* n*-dimensional vector space* Q_p *such that:*

(i) $TE_p = TF_p \oplus Q_p$*, where* TF_p *is the tangent space to the fiber* $E_p \approx F$ *at* p*.*

(ii) given a differentiable path $\gamma : [0,1] \to \mathcal{M}$ *and a point* $f_0 \in pi^{-1}(\gamma(0)$ *there exists* $\tilde{\gamma} : [0,1] \to E$ *such that* $\tilde{\gamma}(0) = f_0$ *and* $\pi_E \circ \tilde{\gamma} = \gamma$*;* $\tilde{\gamma}$ *defines a*

isomorphism between $E_{\gamma(0)}$ *and* $E_{\gamma(1)}$ *and this isomorphism depends differentiably on* t.

(iii) $p \mapsto Q_p$ *is differentiable.*

Theorem 4. *Given a connection on* $\pi : P \to \mathcal{M}$ *there exists a connection on* $\pi : E \to \mathcal{M}$.

Proof: ([4] p. 109) Let u and e be points in P and E respectively. There exists a point f in F such that $uf = e$. For f fixed, we have a differentiable mapping of P into E given by $u \in P \mapsto ef = e \in E$, and it induces the map of the fiber V_uP of P into the fiber F_x over $x = \pi(u) \in \mathcal{M}$. We define the tangent subspace Q_e to be the image of V_uP by this map. Since $(R_g)_*(H_uP) = H_{ug}P$ and $uaa^{-1}f = e$, this map does not depend on the choice of u in $\pi^{-1}(x)$. It is easy to see that $TE_e = F_e \oplus E_e$ and that the distribution Q_e is differentiable. Let $\gamma(t), t \in [0,1]$ be a curve in $\mathcal{M}$ starting at x_0 and ending at x_1. There exists a unique horizontal curve $\tilde{\gamma}(t)$ in P which starts at u_0 and covers $\gamma(t)$. Therefore $\tilde{\gamma}(t)f = e(t)$ is an integral curve of the distribution $e \mapsto Q_e$, which covers $\gamma(t)$. Since $e(t)$ is the map $u_0 f \in F_{x_0} \mapsto u_1 f \in F_{x_1}$, where $u_1 = \widetilde{\gamma(1)}$, it gives an isomorphisms of F_{x_0} onto F_{x_1}. Hence $e \mapsto Q_e$ defines a connection in E.

□

1.2.1 Connection on vector bundles

From now up to the end of this chapter $\pi : E \to \mathcal{M}$ will be a *vector bundle.*

Consider a principal bundle $\pi : P \to \mathcal{M}$ with a connection w and an associated vector bundle $\pi : E \to \mathcal{M}$. We want to define a differential operator on the sections of E. Given a local section $\sigma_\alpha : \mathcal{U}_\alpha \to \pi^{-1}(\mathcal{U}_\alpha)$ of P define $\mathcal{A}_\alpha = \sigma_\alpha^* w$. Fix a vector field $X \in \mathfrak{X}^\infty(\mathcal{M})$. Notice that $\mathcal{A}_\alpha(X)$ is a function $\mathcal{A}_\alpha(X) : \mathcal{U}_\alpha \to \mathfrak{g}$. The associated bundle is constructed using a representation $\rho : G \to \mathrm{Aut}(F)$. Differentiating at the identity we have a representation of the Lie algebra $\rho_* : \mathfrak{g} \to \mathrm{End}(F)$.

Define $\mathcal{A}_\alpha^X : \mathcal{U}_\alpha \to \mathrm{End}(F)$, $\mathcal{A}_\alpha^X := \rho_* \circ \mathcal{A}_\alpha(X)$. Given any section $\sigma \in \Gamma(E)$ the operator $\nabla_X : \Gamma(E) \to \Gamma(E)$ given by

$$(\nabla_X \sigma)(p) = D\sigma(p)(X(p)) + \mathcal{A}_\alpha^X(p)(\sigma(p))$$

is globally defined because of the compatibility conditions (see theorem 2; see also theorem 16). If we want to stress the dependence of ∇_X on w or on the corresponding family $\{\mathcal{A}_\alpha\}$ we write ∇^w_X, $\nabla^{\mathcal{A}}_X$.

Definition 34. *The operator $\nabla_X : \Gamma(E) \to \Gamma(E)$ is called the **covariant derivative with respect to** X.*

It is easy to check that for every $a, b \in \mathbb{R}$ (or $\mathbb{C}$), every $\sigma, \tau \in \Gamma(E)$, every $X, Y \in \mathfrak{X}^\infty(\mathcal{M})$ and every $f \in C^\infty(\mathcal{M})$ hold:
(i) $\nabla_X(a\sigma + b\tau) = a\nabla_X\sigma + b\nabla_X\tau$.
(ii) $\nabla_{aX+bY}\sigma = a\nabla_X\sigma + b\nabla_Y\sigma$.
(iii) $\nabla_X(f\sigma) = X(f)\sigma + f\nabla_X\sigma$ (Leibniz's rule).
(iv) $\nabla_{fX}\sigma = f\nabla_X\sigma$.

Definition 35. *Let E_1 and E_2 be vector bundles over $\mathcal{M}$ and $L : \Gamma(E_1) \to \Gamma(E_2)$ a linear map. We say that L is a **differential operator of rank** r if for every $f \in C^\infty(\mathcal{M})$ such that $f(p) = 0,\ p \in \mathcal{M}$ we have $L(g^{r+1}\sigma)(p) = 0$ for every section $\sigma \in \Gamma(E_1)$. The space of all differential operators of rank r is a vector space and it is denoted by $Diff^r(E_1, E_2)$.*

From the two definitions above it follows easily that $\nabla_X \in Diff^1(E, E)$. Notice that we have a linear map: $\nabla : \Gamma(E) \to \Gamma(E \otimes T^*\mathcal{M})$ given by $\sigma \mapsto (X \mapsto \nabla_X\sigma)$. Given such ∇ we can define parallel transport and hence reconstruct the connection on E. Therefore we also call ∇ **connection** on E — exactly as in the principal bundle case where we use the word connection for both the distribution of horizontal spaces and the 1-form w. We emphasize this with a definition:

Definition 36. *The linear map $\nabla : \Gamma(E) \to \Gamma(E \otimes T^*\mathcal{M})$ given by $\sigma \mapsto (X \mapsto \nabla_X\sigma)$ is (also) called **connection** on E.*

It easily follows:

Proposition 12. *Properties:*
(i) $\nabla \in Diff^1(E, E \otimes T^*\mathcal{M})$.
(ii) $\nabla(f\sigma) = f\nabla\sigma + \sigma \otimes Df$, where f is a function on $\mathcal{M}$ (Leibniz's rule).

In general, when we have a map $f : \Gamma(E) \to \Gamma(E)$ which depends on one or more vector fields, i.e, $f = f_{X_1,\dots,X_p}$, we can consider (as above) the map $\Gamma(E) \to \Gamma(E \otimes \bigwedge^p(T^*\mathcal{M}))$ given by $\sigma \mapsto \Big((X_1, \dots, X_p) \mapsto f_{X_1,\dots,X_p}(\sigma)\Big)$. We refer to this simply as "letting $X_1, \dots, X_p$ vary".

If we have two vector bundles E_1 and E_2 with connections ∇_1 and ∇_2 we can form a connection on the sum bundle $\nabla_1 \oplus \nabla_2 : \Gamma(E_1 \oplus E_2) \to \Gamma((E_1 \oplus E_2) \otimes T^*\mathcal{M})$. We also have a connection on the tensor bundle: $\nabla_1 \otimes \nabla_2 := \nabla_1 \otimes \mathbf{1}_{E_2} + \mathbf{1}_{E_1} \otimes \nabla_2 : \Gamma(E_1 \otimes E_2) \to \Gamma((E_1 \otimes E_2) \otimes T^*\mathcal{M})$.

The following theorem give us a natural way to "transport" structures from $\pi : P \to \mathcal{M}$ to $\pi : E \to \mathcal{M}$.

Theorem 5. *Let $\pi : P \to \mathcal{M}$ be a principal bundle and $\pi : E \to \mathcal{M}$ an associated bundle with fiber F. Let $C^\infty(P, F)^G$ denote the space of maps $f : P \to F$ which are equivariant to G, i.e., $f(ug) = g^{-1}f(u),\ u \in P,\ g \in G$. Then there exists a one-to-one correspondence between $C^\infty(P, F)^G$ and $\Gamma(E)$.*

Proof: Let $f \in C^\infty(P, F)^G$. Define $\hat{f} : P \to P \times F$ by $\hat{f}(u) = (u, f(u))$. $\hat{f}$ is equivariant: $\hat{f}ug = (ug, f(ug)) = (ug, g^{-1}f(u)) = g^{-1}(u, f(u))$, where in the last equality we used the definition of right action of G on $P \times F$. Take a local section $\sigma_\alpha : \mathcal{U}_\alpha \to P$ and define a section of E by $s_\alpha(x) := [\hat{f}(\sigma_\alpha(x))] = [(\sigma_\alpha(x), f(\sigma_\alpha(x)))]$. This is well defined because if $\sigma_\beta : \mathcal{U}_\beta \to P$ is such that $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ then $\sigma_\beta(x) = \sigma_\alpha(x)g_{\alpha\beta}(x)$. Hence

$$[\hat{f}(\sigma_\beta(x))] = [(\sigma_\beta(x), f(\sigma_\beta(x)))] = [(\sigma_\alpha(x)g_{\alpha\beta}(x), f(\sigma_\alpha(x))g_{\alpha\beta}(x))] = \\ [(\sigma_\alpha(x), f(\sigma_\alpha(x)))g_{\alpha\beta}(x)] = [(\sigma_\alpha(x), f(\sigma_\alpha(x)))]$$

Notice that s is indeed a section since $\pi_E(s(x)) = \pi(\sigma_\alpha(x)) = x$ and all maps involved are C^∞.

For the reciprocal, we remember that the associated bundle involves a representation ρ. We have that (recall proposition 1):

$$\sigma_\alpha \in \Gamma(E) \Leftrightarrow \begin{cases} \sigma_\alpha : \mathcal{U}_\alpha \to F \\ \sigma_\beta(x) = \rho(\gamma_{\beta\alpha}(x))\sigma_\alpha(x) \end{cases}$$

Now we want that σ_α defines a function of $C^\infty(P, F)^G$. But then it must be a function $\tilde{\sigma}_\alpha : \mathcal{U}_\alpha \times G \to F$ satisfying (identifying $\pi^{-1}(\mathcal{U}_\alpha)$ with $G \times F$):

$$\tilde{\sigma}_\alpha(x, gh) = \rho(h)^{-1}\sigma_\alpha(x) \text{ and } \tilde{\sigma}_\beta(x, g) = \tilde{\sigma}_\alpha(x, \gamma_{\alpha\beta}(x)g)$$

Then we have $\sigma \in \Gamma(E) \mapsto \{\tilde{\sigma}_\alpha\}$ which satisfies $\tilde{\sigma}_\alpha(x, g) = \rho(g)^{-1}\sigma_\alpha(x)$. Therefore

$$\tilde{\sigma}_\alpha(x, gh) = \rho(gh)^{-1}\sigma_\alpha(x) = \rho(h)^{-1}\rho(g)^{-1}\sigma_\alpha(x) = \rho(h)^{-1}\tilde{\sigma}_\alpha(x, g)$$

so:

$$\tilde{\sigma}_\beta(x,g) = \rho(g)^{-1}\sigma_\beta(x) = \rho(g)^{-1}\rho(\gamma_{\beta\alpha}(x))\sigma_\alpha(x) = \\ = \rho(g^{-1}\gamma_{\beta\alpha}(x))\sigma_\alpha(x) = \rho(\gamma_{\alpha\beta}(x)g)^{-1}\sigma_\alpha(x) = \tilde{\sigma}_\alpha(x, \gamma_{\alpha\beta}(x)g)$$

hence $\{\tilde{\sigma}_\alpha\} \in C^\infty(P,F)^G$. □

Summarizing, the above theorem states that an equivariant function f defines a section σ_f by $\sigma_f(x) = [u, f(u)]$, $u \in \pi^{-1}(x)$, and a section σ defines an equivariant function f_σ by $f_\sigma = u^{-1}(\sigma(\pi(u)))$ where u is seen as an isomorphism $u : F \to E_{\pi(u)}$ ([3] p.51).

We have an action of G on $\mathrm{End}(F)$ defined by

$$\Sigma : G \to \mathrm{Aut}(\mathrm{End}(F)),\ \Sigma(g)(L) := \rho(g)L\rho(g)^{-1},\ L \in \mathrm{End}(F),\ g \in G$$

Proposition 13. *Let $\pi : P \to \mathcal{M}$ be a principal bundle and $\pi : E \to \mathcal{M}$ and associated vector bundle with fiber F. Let w, w' be connections on $\pi : P \to \mathcal{M}$. Then $w' - w$ defines a section of* $\mathrm{End}(E) \otimes T^*\mathcal{M}$.

Proof: By the last theorem is suffices to show that $w' - w$ defines an equivariant map $P \to \mathrm{End}(F) \otimes (\mathbb{R}^N)^*(\approx$ fiber of $\mathrm{End}(E) \otimes T^*\mathcal{M})$. Let $X \in \mathfrak{X}^\infty(\mathcal{M})$. Given $x \in \mathcal{M}$ consider $X(x)$ and $p \in \pi^{-1}(x)$. At p consider $Y \in TP$ such that $D\pi(p)(Y(p)) = X(p)$. Define $\alpha_X : P \to \mathfrak{g}$ by $\alpha_X(p) = w'(p)(Y(p)) - w(p)(Y(p)) = w'(p)(Y_H(p) + Y_V(p)) + w(p)(Y_H(p) + Y_V(p)) = w'(p)(Y_H(p)) \in \mathfrak{g}$, where ${}_{H,V}$ denote the splitting with respect to w (and not w') so that $w(p)(Y_H(p)) = 0$ but not necessarily $w'(p)(Y_H(p)) = 0$; notice also for every connection the action on vertical components is the same and hence $w'(p)(Y_V(p)) - w(p)(Y_V(p)) = 0$. Now we claim that α_X defines a map $P \to \mathfrak{g}$, i.e, it does not depends on Y. Indeed, if W is another vector field such that $D\pi(p)(W(p)) = X(x)$ then $W_H(p) = Y_H(p)$ (they coincide on the vertical). Now, since G acts on F through a representation ρ, differentiating at the identity we have a map $\rho_* : \mathfrak{g} \to \mathrm{End}(F)$ and hence $\rho_* \circ \alpha_X : P \to \mathrm{End}(F)$. If we let X to vary, we obtain a map $\rho_* \circ \alpha : P \to \mathrm{End}(F) \otimes (\mathbb{R}^N)^*$. It is easy to check that this map is equivariant. □

Analogously:

Proposition 14. *The curvature Ω on P defines a section of* $\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{M})$.

Proof: By hypothesis there is a representation $\rho : G \to \mathrm{Aut}(F)$. Define $\mathcal{F} : P \to \mathrm{End}(F) \otimes \bigwedge^2(\mathbb{R}^N)(\approx$ fiber of $\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{M}))$ by $\mathcal{F}(p)(X,Y) := \rho_*(\Omega(p)(\tilde{X},\tilde{Y}))$, where $\tilde{}$ denotes the (horizontal) lifting. Now

$$\begin{aligned}\mathcal{F}(pg)(X(pg),Y(pg)) &= \rho_*(\Omega(pg)(\tilde{X}(pg),\tilde{Y}(pg))) = \rho_*(R_g^*\Omega(p)(\tilde{X}(p),\tilde{Y}(p))) \\ &= \rho_*(Ad_{g^{-1}}(\Omega(p)(\tilde{X}(p),\tilde{Y}(p))) = \Sigma(g^{-1})(\mathcal{F}(p)(X(p),Y(p)))\end{aligned}$$

since the diagram:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{Ad_{g^{-1}}} & \mathfrak{g} \\ \rho_* \downarrow & & \downarrow \rho_* \\ \mathrm{End}(F) & \xrightarrow{\Sigma(g^{-1})} & \mathrm{End}(F) \end{array}$$

commutes; in the second step we used that by the definition of horizontal space $\rho_*(\Omega(pg)(\tilde{X}(pg),\tilde{Y}(pg))) = \rho_*(\Omega(pg)(DR_g(X(p)),DR_g(Y(p))))$. □

We have seen that $w'-w$ defines a section of $\mathrm{End}(E)\otimes T^*\mathcal{M}$. We also have seen that for each connection w', w on $\pi : P \to \mathcal{M}$ we have a connection on $\pi : E \to \mathcal{M}$ (see theorem 4). So it would be natural to expect that $\nabla^{w'} - \nabla^{w}$ be a section of $\mathrm{End}(E) \otimes T^*\mathcal{M}$. This is indeed the case:

Proposition 15. $\nabla^{w'} - \nabla^{w} \in \Gamma(\mathrm{End}(E) \otimes T^*\mathcal{M})$.

Proof: Locally we have:

$$\begin{aligned}(\nabla_X^{w'}\sigma)(p) &= D\sigma(p)(X(p)) + \mathcal{A}_\alpha^{X'}(p)(\sigma(p)) \\ (\nabla_X^{w}\sigma)(p) &= D\sigma(p)(X(p)) + \mathcal{A}_\alpha^{X}(p)(\sigma(p)) \text{ then:} \\ (\nabla_X^{w'} - \nabla_X^{w})\sigma(p) &= \underbrace{(\mathcal{A}_\alpha^{X'}(p) - \mathcal{A}_\alpha^{X}(p))}_{=\mathcal{B}(p)}(\sigma(p))\end{aligned}$$

where $\mathcal{B}$ is an endomorphism of the fiber (since it is the difference of two endomorphisms). □

Another way of expressing this results is to say that $\nabla_X^{w'} - \nabla_X^{w}$ is an operator of rank zero, i.e., it depends pointwise only on the value of the section and not on its derivatives. The content of this proposition is already present in the definition of the covariant derivative, since the ordinary differential is locally a covariant derivative, so we have that the difference $\nabla_X - D$ is an endomorphism. We can go one step further:

Let $\mathfrak{C}(E)$ the space of connections on E. Given $\nabla^0 \in \mathfrak{C}(E)$ and given a End(E)-valued 1-form $\mathcal{A}$ on $\mathcal{M}$ we define $\nabla^{\mathcal{A}} : \Gamma(E) \to \Gamma(E \otimes T^*\mathcal{M})$ by $\nabla^{\mathcal{A}}\sigma = \nabla^0\sigma + \mathcal{A} \otimes \sigma$, where $(\mathcal{A} \otimes \sigma)(p)(X) = \mathcal{A}(p)(X)(\sigma(p))$. Then:

Corollary 2. $\nabla^{\mathcal{A}} \in \mathfrak{C}(E)$ *and* $\Gamma(\mathrm{End}(E) \otimes T^*\mathcal{M}) \approx \mathfrak{C}(E)$

Proof: It is a direct consequence of the last proposition. □

This means that $\mathcal{A} \in \Gamma(\mathrm{End}(E) \otimes T^*\mathcal{M})$ defines a connection only with respect to another one; in other words $\mathfrak{C}(E)$ is an *affine space*. Bearing in mind this subtlety, physicists usually make an abuse of language and call $\mathcal{A}$ itself "the connection".

We conclude that in order to obtain a non-empty space $\mathfrak{C}(E)$ all that we have to prove is the existence of *one* connection; all the others being obtained by adding an endomorphism.

Proposition 16. *([7]) Any principal bundle $\pi : P \to \mathcal{M}$ has a connection.*

Proof: As we have seen the differential is locally a connection, so that its sum with an endomorphism defines another connection. In order to define the connection globally we take an open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ over which P is trivial. For each α we have a connection w_α. Let $\{\lambda_\alpha\}_{\alpha \in I}$ be a partition of unity subordinate to the covering $\{\mathcal{U}_\alpha\}_{\alpha \in I}$. Define $w = \sum_{\alpha \in I} \lambda_\alpha w_\alpha$. This is a one-form on P with values in $\mathfrak{g}$. Near the pre-image of any point $x \in \mathcal{M}$ this one form is an affine combination of connections one-forms and hence is a connection one-form. But if that is true in the neighborhood of the pre-image of each point of $\mathcal{M}$ then it means that w is a connection one-form on all of P. □

Corollary 3. $\dim \mathfrak{C}(E) = \infty$

Proof: $\dim \Gamma(E) = \infty$ □

Now let us express the curvature in terms of the connection, exactly in the same way we did for the curvature on P. First we introduce an algebraic expression to the curvature. Then we shall express it as an *exterior* *covariant derivative* (analogously to $\Omega = D^w w$). Finally, we shall show that this definition coincides with that induced by Ω (the curvature on the principal bundle P) — both define the same section of $\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{M})$ (see proposition 14).

Given vector fields $X, Y \in \mathfrak{X}^\infty(\mathcal{M})$ put:

$$\mathcal{F}(X,Y)\sigma = \nabla_X\nabla_Y\sigma - \nabla_Y\nabla_X\sigma - \nabla_{[X,Y]}\sigma, \ \sigma \in \Gamma(E)$$

(to be honest we should use other symbol instead of $\mathcal{F}$, since we are defining it and only afterwards we shall show that this definition agrees with our previous concept of curvature, i.e., from proposition 14; but we hope that with this warning there will be no misunderstanding). We write $\mathcal{F}(\nabla)$, $\mathcal{F}_\nabla$ etc when we want to stress the dependence on ∇.

We notice that $\mathcal{F}$ is $C^\infty(\mathcal{M})$-linear:

$$\begin{aligned}\mathcal{F}(X,fY)\sigma &= \nabla_X\nabla_{fY}\sigma - \nabla_{fY}\nabla_X\sigma - \nabla_{[X,fY]}\sigma, f \in C^\infty(\mathcal{M}) \\ &= \nabla_X(f\nabla_Y\sigma) - f\nabla_Y\nabla_X\sigma - \nabla_{f[X,Y]+X(f)Y}\sigma \\ &= f\nabla_X\nabla_Y\sigma + X(f)\nabla_Y\sigma - f\nabla_Y\nabla_X\sigma - f\nabla_{[X,Y]}\sigma - X(f)\nabla_y\sigma \\ &= f\mathcal{F}(X,Y)\sigma\end{aligned}$$

In order to prove $C^\infty(\mathcal{M})$-linearity in the entry $\mathcal{F}(\cdot, Y)$ we simply use the anti-symmetry $\mathcal{F}(X,Y) = -\mathcal{F}(Y,X)$ (which is obvious from the definition).

Now we show that it is $C^\infty(\mathcal{M})$-linear with respect to sections:

$$\begin{gathered}\mathcal{F}(X,Y)(f\sigma) = \nabla_X\nabla_Y(f\sigma) - \nabla_Y\nabla_X(f\sigma) - \nabla_{[X,Y]}(f\sigma) \\ = \nabla_X(f\nabla_Y\sigma + Y(f)\sigma) - \nabla_Y(f\nabla_X\sigma + X(f)\sigma) - (f\nabla_{[X,Y]}\sigma + [X,Y]\sigma) \\ = f\nabla_X\nabla_Y\sigma + X(f)\nabla_Y\sigma + Y(f)\nabla_X\sigma + X(Y(f))\sigma - f\nabla_Y\nabla_X\sigma - \\ Y(f)\nabla_X\sigma - X(f)\nabla_Y\sigma - Y(X(f))\sigma - f\nabla_{[X,Y]}\sigma - ([X,Y]f)\sigma \\ = f\mathcal{F}(X,Y)\sigma + X(Y(f))\sigma - Y(X(f))\sigma - ([X,Y]f)\sigma = f\mathcal{F}(X,Y)\sigma\end{gathered}$$

What these calculations actually show is that $\mathcal{F}(X,Y)$ is a section of $\mathrm{End}(E)$ because:

Proposition 17. *Let $g : \Gamma(E) \to \Gamma(E)$ be a $C^\infty(\mathcal{M})$-linear map. Then g defines a section of* $\mathrm{End}(E)$.

Remark 8. *Notice that there is a straightforward reciprocal: given a section $T \in \mathrm{End}(E)$, T defines a map from $\Gamma(E)$ to $\Gamma(E)$ by $\sigma \in \Gamma(E)$ by $(T\sigma)(x) = T(x)\sigma(x)$.*

Sketch proof: Put $g_p = g|p$, $p \in \mathcal{M}$. We define an endomorphism of the fiber E_p, $T_p : E_p \to E_p$ by $v \in E_p \mapsto g_p(\sigma(p)) = (g\sigma)(p)$, where σ is a section such that $\sigma(p) = v$. The definition of T_p does not depend on σ, for if σ' is other section such that $\sigma'(p) = v$ then $0 = g_p(\sigma(p) - \sigma'(p)) = g_p(\sigma(p)) - g_p(\sigma'(p))$. To define T_p globally we use a partition of unity (more details see [8], p.220). □

Now, if we let X, Y to vary we obtain that $\mathcal{F} \in \Gamma(\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{M}))$. Now our aim is to express $\mathcal{F}$ as some "derivative" of a a End(E)-valued 1-form. For this we need:

Definition 37. *Given a connection ∇ on a vector bundle $\pi : E \to \mathcal{M}$ we define the **exterior covariant derivative** $D^\nabla : \Gamma(E \otimes \bigwedge^p(T^*\mathcal{M})) \to \Gamma(E \otimes \bigwedge^{p+1}(T^*\mathcal{M}))$ by*

$$D^\nabla \mu(X_1, \ldots, X_{p+1}) := \sum_{j=1}^{p+1} (-1)^{j+1} \nabla_{X_j} \mu(X_1, \ldots, \hat{X}_j \ldots, X_{p+1}) + \sum_{1 \le j < k \le p+1} (-1)^{j+k} \mu([X_j, X_k], X_1, \ldots, \hat{X}_j, \ldots, \hat{X}_k, \ldots, X_{p+1})$$

Notice that D^∇ naturally extends ∇, i.e., $D^\nabla = \nabla$ for $p = 0$. Obviously this operator is linear. It also satisfies a Leibniz's rule; to see this, first we apply the definition to a E-valued 1-form; it suffices to compute it for $e \otimes \mu$:

$$\begin{aligned} D^\nabla(e \otimes w)(X, Y) = \nabla_X(ew(Y)) - \nabla_Y(ew(X)) - ew([X, Y]) = \\ (Xw(Y))e + w(Y)\nabla_X e - (Yw(X))e - w(X)\nabla_Y e - ew([X, Y]) = \\ w(Y)\nabla_X e - w(X)\nabla_Y e + e\big(Xw(Y) - Yw(X) - w([X, Y])\big) = \\ (\nabla e \wedge w + e \otimes dw)(X, Y) = (D^\nabla e \wedge w + e \otimes dw)(X, Y) \end{aligned}$$

where in the last step we used that $D^\nabla = \nabla$ for $p = 0$. Proceeding inductively, for a E-valued p-form $\sigma = e \otimes w$ we get $D^\nabla(\sigma \otimes w) = D^\nabla e \wedge w + e \otimes dw$.

Now, if f is a differential form on $\mathcal{M}$ and w a form on $\mathcal{M}$ with values on

E, writing $w = T_\alpha \otimes w^\alpha$ we obatin:

$$D^\nabla(f \wedge w) = D^\nabla(f \wedge T_\alpha \otimes w^\alpha) = D^\nabla(T_\alpha \otimes (f \wedge w^\alpha)) =$$
$$D^\nabla T_\alpha \wedge f \wedge w^\alpha + T_\alpha \otimes D^\nabla(f \wedge w^\alpha) =$$
$$D^\nabla T_\alpha \wedge f \wedge w^\alpha + T_\alpha \otimes (df \wedge w^\alpha + (-1)^{\deg(f)} f \wedge w^\alpha) =$$
$$D^\nabla T_\alpha \wedge f \wedge w^\alpha + T_\alpha \otimes (df \wedge w^\alpha) + (-1)^{\deg(f)} T_\alpha \otimes (f \wedge w^\alpha) =$$
$$df \wedge (T_\alpha \otimes w^\alpha) + D^\nabla T_\alpha \wedge f \wedge w^\alpha + (-1)^{\deg(f)} T_\alpha \otimes (f \wedge w^\alpha)$$

Since $D^\nabla T_\alpha = \nabla T_\alpha$ is a 1-form with values in E we have $D^\nabla T_\alpha \wedge f = (-1)^{\deg(f)\cdot 1} f \wedge D^\nabla T_\alpha$, so we get

$$df \wedge (T_\alpha \otimes w^\alpha) + (-1)^{\deg(f)} f \wedge \left(D^\nabla T_\alpha \wedge w^\alpha + T_\alpha \otimes w^\alpha\right) =$$
$$df \wedge w + (-1)^{\deg(f)} f \wedge D^\nabla w$$

which is the Leibniz rule (see [10] p. 110ff for more properties of D^∇).

We summarize these results:

Proposition 18. *(Leibniz rule) (i)* $D^\nabla(\sigma \otimes w) = D^\nabla \sigma \wedge w + \sigma \otimes dw$*; (ii)* $D^\nabla(f \wedge w) = df \wedge w + (-1)^{\deg(f)} f \wedge D^\nabla w$*.*

As we said D^∇ extends ∇; because of that some writers also denote D^∇ by ∇. It is important to notice, however, that D^∇ is *not* a connection (and that is why we avoid to use the notation ∇). By saying that it is not a connection, we mean the following. Consider the bundle $E' = E \otimes T^*\mathcal{M}$; it has a natural connection $\nabla' : \Gamma(E') \to \Gamma(E' \otimes T^*\mathcal{M}) = \Gamma(E \otimes T^*\mathcal{M} \otimes T^*\mathcal{M})$ which is the tensor connection $\nabla' = \nabla \otimes \nabla_L$, where $\nabla : \Gamma(E) \to \Gamma(E \otimes T^*\mathcal{M})$ is the connection on E and $\nabla_L : \Gamma(T^*\mathcal{M}) \to \Gamma(T^*\mathcal{M} \otimes T^*\mathcal{M})$ is the connection on $T^*\mathcal{M}$ induced by the Levi-Civita connection. So we can also say that ∇' extends ∇, as does D^∇:

$$\Gamma(E) \xrightarrow{\nabla} \Gamma(E \otimes T^*\mathcal{M}) \xrightarrow{\nabla'} \Gamma(E \otimes T^*\mathcal{M} \otimes T^*\mathcal{M})$$
$$\Gamma(E) \xrightarrow{\nabla} \Gamma(E \otimes T^*\mathcal{M}) \xrightarrow{D^\nabla} \Gamma(E \otimes \bigwedge{}^2(T^*\mathcal{M}))$$

But ∇' is, by construction, a connection on E' whereas D^∇ is not since it does not satisfy the Leibniz rule for a connection: if f is a function on $\mathcal{M}$ and $\sigma' = \sigma \otimes \omega$ is a section of E' then

$$D^\nabla(f\sigma') = df \wedge \sigma' + f D^\nabla \sigma'$$

Which, in order to be a connection should equal $df \otimes \sigma' + f D^\nabla \sigma'$ (see property(ii) of proposition 12)[7]. The reason why we are bringing attention to such trivial statements is that the notation used by different authors can be very confusing: both ∇' and D^∇ are sometimes denoted by ∇ since both can be considered extensions of the original ∇ [8]. Some authors try to distinguish ∇ from their extensions by denoting the extension by ∇^2, but again this is not very helpful since both ∇' and D^∇ are sometimes denoted by ∇^2 (e.g., [7] p. 85).

Theorem 6. *Given a connection ∇ on $\pi : E \to \mathcal{M}$, for any section $\sigma \in \Gamma(E)$ and any pair of vector fields $X, Y \in \mathfrak{X}^\infty(\mathcal{M})$ we have*

$$(D^\nabla \circ \nabla)(X,Y)\sigma = \mathcal{F}(X,Y)\sigma$$

i.e., the curvature $\mathcal{F}$ is given by the composition:

$$\mathcal{F} : \Gamma(E) \xrightarrow{\nabla} \Gamma(E \otimes T^*\mathcal{M}) \xrightarrow{D^\nabla} \Gamma(E \otimes {\textstyle\bigwedge}^2(T^*\mathcal{M}))$$

Remark 9. *Now we have the following setting: the connection ∇ is identified with an* End(E)*-valued* 1*-form $\mathcal{A}$ (corollary 2) and the curvature $\mathcal{F}$ is written as an exterior covariant derivative of such connection (actually, as the composition $D^\nabla \circ \nabla$). So, this is the exact analogous of $\Omega = D^w w$.*

Proof: Recall from definition 36 that $(\nabla\sigma)(X) = \nabla_X \sigma$. Then applying the definition of exterior covariant derivative to the E-valued 1-form $\nabla\sigma$ we have:

$$\begin{aligned}(D^\nabla \nabla\sigma)(X_1, X_2) &= \nabla_{X_1}\Big((\nabla\sigma)(X_2)\Big) - \nabla_{X_2}\Big((\nabla\sigma)(X_1)\Big) - (\nabla\sigma)([X_1, X_2]) \\ &= \nabla_{X_1}\nabla_{X_2}\sigma - \nabla_{X_2}\nabla_{X_1}\sigma - \nabla_{[X_1,X_2]}\sigma\end{aligned}$$

□

Now we eventually prove that the two ways of constructing the curvature expressed above in fact coincide. Recall from proposition 14 that the curvature Ω on $\pi : P \to \mathcal{M}$ defines a section $\mathcal{F}$ of End$(E) \otimes \bigwedge^2(T^*\mathcal{M})$ by $\mathcal{F}(X,Y) = \rho_*(\Omega(\tilde{X}, \tilde{Y}))$, where X, Y are vector fields on $\mathcal{M}$, $\tilde{}$ denotes their horizontal lift to P and $\rho : G \to \mathrm{Aut}(F)$ is a representation.

[7] D^∇ does satisfy, of course, the Leibniz rule for an exterior covariant derivative.

[8] Of course, this situation is not different from the standard situation in Riemannian geometry, where the same symbol ∇ is used to denote the Levi-Citiva connection and its extension to the tensor bundle.

Theorem 7. *For any section $\tau \in \Gamma(E)$ we have $\mathcal{F}\tau = D^{\nabla}\nabla\tau$.*

Proof: We work locally. Consider a trivialization of $\sigma_\alpha : \mathcal{U} \to P$. It induces a trivialization of E. We shall show in proposition 20 that locally $D^{\nabla}\nabla = d\mathcal{A}_\alpha + \mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$, where $\mathcal{A}_\alpha$ is the End(E) valued form on $\mathcal{U}$ given by $\mathcal{A}_\alpha = \rho_* \circ \sigma_\alpha^* w$ (recall also definition of covariant derivative; the reason why $d\mathcal{A}_\alpha + \mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$ does not have a factor $\frac{1}{2}$ as the Cartan's structure equation becomes clear below in the proof).

We will show that the diagrams

$$\begin{array}{ccc} \Gamma(\mathfrak{g} \otimes T^*\mathcal{U}) & \stackrel{d}{\longrightarrow} & \Gamma(\mathfrak{g} \otimes \bigwedge^2(T^*\mathcal{U})) \\ \rho_* \downarrow & & \downarrow \rho_* \\ \Gamma(\mathrm{End}(E) \otimes T^*\mathcal{U}) & \stackrel{d}{\longrightarrow} & \Gamma(\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{U})) \end{array}$$

and

$$\begin{array}{ccc} \Gamma(\mathfrak{g} \otimes T^*\mathcal{U}) \times \Gamma(\mathfrak{g} \otimes T^*\mathcal{U}) & \stackrel{\frac{1}{2}\wedge}{\longrightarrow} & \Gamma(\mathfrak{g} \otimes \bigwedge^2(T^*\mathcal{U})) \\ (\rho_*, \rho_*) \quad \downarrow & & \downarrow \rho_* \\ \Gamma(\mathrm{End}(E) \otimes T^*\mathcal{U}) \times \Gamma(\mathrm{End}(E) \otimes T^*\mathcal{U}) & \stackrel{\wedge}{\longrightarrow} & \Gamma(\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{U})) \end{array}$$

commute (again, the reason for the factor $\frac{1}{2}$ becomes clear in the proof). This is a direct consequence of definitions sine d and $\wedge$ act only on the "form part" (recall definitions 22 and 26). Explicitly, writing $\mu = T_i \otimes \mu^i$ and $\nu = U_j \otimes \nu^j$ for $\mathfrak{g}$-valued forms we have that for any vector X: $\mu(X) = T_i \mu^i(X)$; since ρ_* is a homomorphism $\rho_*(\mu(X)) = \rho_*(T_i\mu^i(X)) = \mu_i(X)\rho_*(T_i) = (\rho_*(T_i) \otimes \mu^i)(X)$, i.e., $\rho_* \circ (T_i \otimes \mu^i) = \rho_*(T_i) \otimes \mu^i$. Applying this for $d\mu$ instead of μ and using definition 22: $\rho_* \circ d\mu = \rho_* \circ (T_i \otimes d\mu^i) = \rho_*(T_i) \otimes d\mu^i = d(\rho_*(T_i) \otimes \mu^i) = d\rho_* \circ \mu$. Analogously: $\rho_* \circ (\mu \wedge \nu) = \rho_* \circ ([T_i, U_j] \otimes \mu^i \wedge \nu^j) = \rho_*([T_i, U_j]) \otimes \mu^i \wedge \nu^j$. This equals $[\rho_*(T_i), \rho_*(U_j)] \otimes \mu^i \wedge \nu^j$ since ρ_* is a (Lie algebra) representation.

Now, since d and $\wedge$ commute with pull-backs, using Cartan's structure equation and the previuos calculation we obtain $\rho_* \circ (\sigma_\alpha^*(\Omega)) = \rho_* \circ (\sigma_\alpha^*(dw + \frac{1}{2}w \wedge w)) = \rho_* \circ (d\sigma_\alpha^*(w)) + \frac{1}{2}\rho_* \circ (\sigma_\alpha^*(w) \wedge \sigma_\alpha^*(w)) = d(\rho_* \circ \sigma_\alpha^*(w)) + \frac{1}{2}\rho_* \circ (\sigma_\alpha^*(w) \wedge \sigma_\alpha^*(w))$. The first term gives: $d(\rho_* \circ \sigma_\alpha^*(w)) = d\mathcal{A}_\alpha$. For the second term, write $\sigma_\alpha^* w = \sigma_\alpha^*(T_i \otimes w^i) = \sigma_\alpha^* T_i \otimes \sigma_\alpha^* w_i$. Again using the previous calculation:

$$\rho_* \circ (\sigma_\alpha^*(w) \wedge \sigma_\alpha^*(w)) = \rho_*(\sigma_\alpha^* T_i \otimes \sigma_\alpha^* w_i \wedge \sigma_\alpha^* T_j \otimes \sigma_\alpha^* w_j) =$$
$$\rho_* \circ ([\sigma_\alpha^* T_i, \sigma_\alpha^* T_j] \otimes \sigma_\alpha^* w^i \wedge \sigma_\alpha^* w^j) = [\rho_* \circ \sigma_\alpha^* T_i, \rho_* \circ \sigma_\alpha^* T_j] \otimes \sigma_\alpha^* w^i \wedge \sigma_\alpha^* w^j$$

Notice that $\rho_* \circ \sigma_\alpha^* T_i$ are endomorphism of the fibers of the bundle E and w^i are ordinary differential forms on $\mathcal{U}$. Hence we can write $\sigma_\alpha^* w^i = f_\mu^i dx^\mu$, with f_μ^i real valued functions on $\mathcal{U}$, and denote $f_\mu^i \rho_* \circ \sigma^* T_i$ by B_μ. Then

$$[\rho_* \circ \sigma_\alpha^* T_i, \rho_* \circ \sigma_\alpha^* T_j] \otimes \sigma_\alpha^* w^i \wedge \sigma_\alpha^* w^j = [\rho_* \circ \sigma_\alpha^* T_i, \rho_* \circ \sigma_\alpha^* T_j] \otimes f_\mu^i dx^\mu \wedge f_\nu^j dx^\nu = [f_\mu^i \rho_* \circ \sigma_\alpha^* T_i, f_\nu^j \rho_* \circ \sigma_\alpha^* T_j] \otimes dx^\mu \wedge dx^\nu = [B_\mu, B_\nu] \otimes dx^\mu \wedge dx^\nu$$

Since the sum is over all μ, ν, using the anti-symmetry of $[,]$ and $dx^\mu \wedge dx^\nu$ the last expression equals to $2B_\mu B_\nu \otimes dx^\mu \wedge dx^\nu$, where $B_\mu B_\nu = B_\mu \circ B_\nu$ =usual composition of endomorphisms. Hence

$$2B_\mu B_\nu \otimes dx^\mu \wedge dx^\nu = 2(f_\mu^i \rho_* \circ \sigma_\alpha^* T_i)(f_\nu^j \rho_* \circ \sigma_\alpha^* T_j) \otimes dx^\mu \wedge dx^\nu = 2(\rho_* \circ \sigma_\alpha^* T_i)(\rho_* \circ \sigma_\alpha^* T_j) \otimes f_\mu^i dx^\mu \wedge f_\nu^j dx^\nu = 2\big((\rho_* \circ \sigma_\alpha^* T_i)(\rho_* \circ \sigma_\alpha^* T_j)\big) \otimes \sigma_\alpha^* w^i \wedge \sigma_\alpha^* w^j =$$

By definition of wedge product of End(E)-valued forms this last expression equals to

$$2(\rho_* \circ \sigma_\alpha^* T_i \otimes \sigma_\alpha^* w^i) \wedge (\rho_* \circ \sigma_\alpha^* T_j \otimes \sigma_\alpha^* w^j)$$

Using the equality $\rho_* \circ (T_i \otimes \mu^i) = \rho_*(T_i) \otimes \mu^i$ we have then $\rho_* \circ \sigma_\alpha^* T_i \otimes \sigma_\alpha^* w^i = \rho_* \circ (\sigma_\alpha^* T_i \otimes \sigma_\alpha^* w^i) = \rho_* \circ \sigma_\alpha^*(T_\otimes w^i) = \rho_* \circ \sigma_\alpha^* w$. Hence

$$2(\rho_* \circ \sigma_\alpha^* T_i \otimes \sigma_\alpha^* w^i) \wedge (\rho_* \circ \sigma_\alpha^* T_j \otimes \sigma_\alpha^* w^j) = 2\rho_* \circ \sigma_\alpha^* w \wedge \rho_* \circ \sigma_\alpha^* w = 2\mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$$

Therefore the term $\frac{1}{2}\rho_* \circ (\sigma_\alpha^*(w) \wedge \sigma_\alpha^*(w))$ equals to $\mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$, what finishes the proof. □

We now have the following nice picture. A connection on P induces a connection on E, and the connection 1-form w on P (which is thought of as the connection itself), induces a covariant derivative ∇ on E. On P exterior covariant derivative of w gives rise to the curvature Ω. On E exterior covariant derivative of ∇ gives rise to the curvature $\mathcal{F}$. And this two curvatures are connected by the representation ρ as shown in the last theorem. Moreover, we have a Cartan's structure equation $\Omega = dw + \frac{1}{2} w \wedge w$ on P and, taking into account proposition 20 below, a structure-like equation $d\mathcal{A}_\alpha + \mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$. Moreover, the reason why we have a factor $\frac{1}{2}$ is completely consistent with the definition of the wedge product for $\mathfrak{g}$-valued and for End(E)-valued forms:

in the first case we combine the anti-symmetric wedge product of one forms with the anti-symmetric bracket of the Lie algebra what yields a symmetric term and hence a double counting, while in the second case we just have the wedge product combined with usual multiplication of two endomorphisms (see [6] p 36).

Summarizing:

$$D^w w = \Omega = dw + \frac{1}{2} w \wedge w \text{ on } P$$
$$D^\nabla \nabla = \mathcal{F} = d\mathcal{A}_\alpha + \mathcal{A}_\alpha \wedge \mathcal{A}_\alpha \text{ on } E$$

and these quantities are related by the representation ρ.

Recall that If we have two vector bundles E_1 and E_2 with connections ∇_1 and ∇_2 we can form a connection on tensor bundle. $\nabla_1 \otimes \nabla_2 : \Gamma(E_1 \otimes E_2) \to \Gamma((E_1 \otimes E_2) \otimes T^*\mathcal{M})$. In particular, since $\mathrm{End}(E) \approx E \otimes E^*$, we have an exterior covariant derivative induced on the bundle of endomorphisms.

1.2.2 The local form of the connection and curvature

Here we investigate properties of the curvature $\mathcal{F}$ in local coordinates and derive some useful formulas.

This is First, we fix some notation: we shall denote (as usual) $\{\partial_\mu\}$, $\mu = 1, \ldots, n = \dim(\mathcal{M})$, a local basis of $T\mathcal{M}_x$; we should write $\{\partial_\mu(x)\}$ for stressing that it is a basis of $T\mathcal{M}_x$, but we shall omit x when no confusion can arise. Remember that locally it is always possible to write a basis of sections for a vector bundle. We denote $\{e_i\}$, $i = 1, \ldots, m = \dim(F)$ a local basis of sections. Again, we should write $e_i(x)$ for E_x, but we omit the point x when it is possible. We shall indicate indices which run from 1 to m by Latin letters and indices which run from 1 to n by Greek letters. Having $\{e_i\}$, we can arrange a local basis of sections of E^* as $\{(e_i)^* =: e^i\}$ and a local basis of sections of $\mathrm{End}(E)$ as $\{e_i \otimes e^j\}$. *Over all this section we shall assume that we are working in local coordinates, so that a local trivialization has always been chosen.* Therefore, for simplifying notation we shall omit the indices of local trivialization. So we shall write, for example, $\mathcal{A}$ instead of $\mathcal{A}_\alpha$ for $\rho_* \circ \sigma_\alpha w$ appearing in the expression of the connection on E.

Recall that the connection ∇ induces a connection and an exterior covariant derivative on the bundle $\mathrm{End}(E)$ — which we also denote by ∇ and D^∇

We start writing the connection in local coordinates. Following physicists notation, we shall write $\nabla_\mu := \nabla_{\partial_\mu}$. Since ∇ associates to each pair section-tangent vector a new section we may write:

$$\nabla_\mu e_j = \mathcal{A}^i_{\mu j} e_i$$

and for a general section $\sigma = \sigma^i e_i$ and vector $X = X^\mu \partial_\mu$:

$$\nabla_X \sigma = \nabla_{X^\mu \partial_\mu}(\sigma^i e_i) = X^\mu \nabla_\mu(\sigma^i e_i) = \\ X^\mu((\partial_\mu \sigma^i)e_i + \mathcal{A}^j_{\mu i}\sigma^i e_j)$$

where in the last step we use the Leibniz rule. Renaming the dummy indices in the last term we get:

$$\nabla_X \sigma = X^\mu(\partial_\mu \sigma^i + \mathcal{A}^i_{\mu j}\sigma^j)e_i$$

If we we define the component functions $\nabla_\mu \sigma = (\nabla_\mu \sigma)^i e_i$ the above expression gives:

$$(\nabla_\mu \sigma)^i = \partial_\mu \sigma^i + \mathcal{A}^i_{\mu j}\sigma^j$$

which is very familiar to physicists[9].

Remark 10. *Physicists would write $\nabla_\mu \sigma^i$ instead of $(\nabla_\mu \sigma)^i$. Definitely this is an abuse of notation since* there is no sense *in taking the covariant derivative of σ^i.*

Comparing the above expression for $(\nabla_\mu \sigma)^i$ with the expression for the covariant derivative given in chapter 1.2:

$$(\nabla_X \sigma)(p) = D\sigma(p)(X(p)) + \mathcal{A}^X_\alpha(p)(\sigma(p))$$

we may suspect that $\mathcal{A}^i_{\mu j}$ are the components of the End(E)-valued form $\mathcal{A}_\alpha$ in this particular coordinate systems. This is indeed the case. First notice that

[9] If σ is a section of the tangent bundle, i.e., a vector field, then $\mathcal{A}^i_{\mu j}$ are exactly the Christoffel symbols.

$$X^\mu \mathcal{A}^i_{\mu j} \sigma^j e_i = \mathcal{A}^i_{\mu j}(e_i \otimes e^j \otimes dx^\mu)(\sigma \otimes X)$$

Now observe that $(e_i \otimes e^j \otimes dx^\mu) \in \Gamma(\mathrm{End}(E|\mathcal{U}) \otimes T^*\mathcal{U})$, i.e., it is a local section of the (trivial) bundle $\mathrm{End}(E|\mathcal{U}) \otimes T^*\mathcal{U}$ over $\mathcal{U}$, where $\mathcal{U} \subset \mathcal{M}$ is an open set on which we perform the trivialization. We can also write

$$\mathcal{A}(X)\sigma = X^\mu \mathcal{A}^i_{\mu j} \sigma^j e_i = X^\mu \mathcal{A}^i_{\mu j}(e_i \otimes e^j)(\sigma)$$

Putting all pieces together we have the following expression for the covariant derivative expressed in terms of a given local basis of sections and local basis of the tangent space:

$$(\nabla_X \sigma)^i = X\sigma^i + (\mathcal{A}(X)\sigma)^i = X^\mu \partial_\mu \sigma^i + (\mathcal{A}(X)\sigma)^i$$

Remark 11. *Physicist like to suppress the internal indices i, j which are associated to the sections e_i and write $\mathcal{A}$ in terms of components:*

$$\mathcal{A}_\mu = \mathcal{A}^i_{\mu j} e_i \otimes e^j$$

In other words, by "supressing" the internal indices, we can think of $\mathcal{A}_\mu$ as matrices representing the corresponding endomorphism in the given trivialization; the matrix $\mathcal{A}_\mu$ has entries $\mathcal{A}^i_{\mu j}$.

Now let us investigate the local form of the curvature. Define:

$$\mathcal{F}_{\mu\nu} = \mathcal{F}(\partial_\mu, \partial_\nu)$$

$\mathcal{F}_{\mu\nu}$ is a section of $\mathrm{End}(E)$. Notice that since $[\partial_\mu, \partial_\nu] = 0$ we also have $\mathcal{F}_{\mu\nu} = [\nabla_\mu, \nabla_\nu]$. The linearity properties of the curvature allow us to write for any pair of tangent vectors $X = X^\mu \partial_\mu$ and $Y = Y^\nu \partial_\nu$:

$$\mathcal{F}(X, Y) = X^\mu Y^\nu \mathcal{F}_{\mu\nu}$$

Now we compute:

$$\mathcal{F}_{\mu\nu}e_i = \nabla_\mu\nabla_\nu e_i - \nabla_\nu\nabla_\nu e_{-}\nabla_{[\partial_\mu,\partial_\nu]}e_i$$
$$\nabla_\mu(\mathcal{A}^j_{\nu i}e_j) - \nabla_\nu(\mathcal{A}^j_{\mu i}e_j) =$$
$$(\partial_\mu\mathcal{A}^j_{\nu i})e_j + \mathcal{A}^k_{\mu j}\mathcal{A}^j_{\nu i}e_k - (\partial_\nu\mathcal{A}^j_{\mu i})e_j - \mathcal{A}^k_{\nu j}\mathcal{A}^j_{\nu i}e_k =$$
$$\big((\partial_\mu\mathcal{A}^j_{\nu i}) - (\partial_\nu\mathcal{A}^j_{\mu i}) + \mathcal{A}^j_{\mu k}\mathcal{A}^k_{\nu i} - \mathcal{A}^j_{\nu k}\mathcal{A}^k_{\nu i}\big)e_j \quad (*)$$

where we used $[\partial_\mu, \partial_\nu] = 0$ and in the last step we simply relabeled the indices. Analogously to what we did for the connection we may expand the curvature in a local basis $\{e_i \otimes e^j\}$ of End(E):

$$\mathcal{F}_{\mu\nu} = \mathcal{F}^j_{\mu\nu i}e_j \otimes e^i$$

the functions $\mathcal{F}^j_{\mu\nu i}$ are called components of the curvature (in that local basis). Notice:

$$\mathcal{F}_{\mu\nu}e_i = \mathcal{F}^j_{\mu\nu i}e_j$$

So, comparing with $(*)$ we have:

$$\mathcal{F}^j_{\mu\nu i} = (\partial_\mu\mathcal{A}^j_{\nu i}) - (\partial_\nu\mathcal{A}^j_{\mu i}) + \mathcal{A}^j_{\mu k}\mathcal{A}^k_{\nu i} - \mathcal{A}^j_{\nu k}\mathcal{A}^k_{\nu i}$$

Again, physicists like to suppress the internal indices i, j, k associated to the basis of sections of E over $\mathcal{U}$, meaning that $\mathcal{A}_\mu$ is thought of as a matrix and then we may write the above expression as:

$$\mathcal{F}_{\mu\nu} = \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [A_\mu, A_\nu]$$

Now that we expressed the functions $\mathcal{F}_{\mu\nu}$ in a local basis $\{e_i \otimes e^j\}$ of End(E) we can write the curvature itself in terms of the "components" $\mathcal{F}_{\mu\nu}$:

$$\mathcal{F} = \frac{1}{2}\mathcal{F}_{\mu\nu}dx^\mu \wedge dx^\nu$$

the factor $\frac{1}{2}$ appears because we are summing over all indices and not only over $\mu < \nu$.

Now we show that $D^\nabla \circ \nabla$ is proportional to the curvature.

Lemma 2. *For any section α of E we have*

$$\nabla\alpha = \nabla_\mu\alpha \otimes dx^\mu$$

Proof: For any vector field X:

$$(\nabla\alpha)(X) = \nabla_X\alpha = X^\mu\nabla_\mu\alpha = (\nabla_\mu\alpha \otimes dx^\mu)(X)$$

□

Lemma 3. *For any E-valued p-form α, written locally as $\alpha = \alpha^i e_i \otimes w_I dx^I = \alpha^i e_i w_I \otimes dx^I := \alpha_I \otimes dx^I$ with $dx^I = dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p}$, we have:*

$$D^\nabla\alpha = \nabla_\nu\alpha_I \otimes dx^\nu \wedge dx^I$$

Proof: Compute:

$$\begin{aligned} D^\nabla\alpha = D^\nabla(\alpha_I \otimes dx^I) = (\nabla\alpha_I) \wedge dx^I + \alpha_I \otimes ddx^I = (\nabla\alpha_I) \wedge dx^I \\ = \nabla_\nu\alpha_I \otimes dx^\nu \wedge dx^I \end{aligned}$$

□

Proposition 19. *For any E-valued p-form α we have:*

$$D^\nabla \circ D^\nabla\alpha = \mathcal{F} \wedge \alpha$$

Proof: As before, write $\alpha = \alpha_I \otimes dx^I$, where $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_p}$.

$$\begin{aligned} D^\nabla D^\nabla\alpha &= D^\nabla(\nabla_\nu\alpha_I \otimes dx^\nu \wedge dx^I) \\ &= \nabla_\tau\nabla_\nu\alpha_I \otimes dx^\tau \wedge dx^\nu \wedge dx^I \\ &= \frac{1}{2}[\nabla_\tau, \nabla_\nu](\alpha_I) \otimes dx^\tau \wedge dx^\nu \wedge dx^I \\ &= \frac{1}{2}\mathcal{F}_{\tau\nu}(\alpha_I) \otimes dx^\tau \wedge dx^\nu \wedge dx^I = \mathcal{F} \wedge \alpha \end{aligned}$$

□

It is a easy calculation to show that if μ is and End(E)-valued p-form and σ is a E-valued form we have $D^\nabla(\mu \wedge \sigma) = (D^\nabla\mu) \wedge \sigma + (-1)^p\mu \wedge D^\nabla\sigma$.

Theorem 8. *(Bianchi Identity)* $D^\nabla \mathcal{F} = 0$.

Proof: For any E-valued form we α have

$$(D^\nabla)^3 = D^\nabla(D^\nabla D^\nabla \alpha) = D^\nabla(\mathcal{F} \wedge \alpha) = (D^\nabla \mathcal{F}) \wedge \alpha + \mathcal{F} \wedge (D^\nabla \alpha)$$

On the other hand:

$$(D^\nabla)^3 \alpha = (D^\nabla)^2 (D^\nabla \alpha) = \mathcal{F} \wedge D^\nabla \alpha$$

So $D^\nabla \mathcal{F} = 0$. □

Remark 12. *Although the Bianchi has been deduced used some computations carried out on local coordinates — namely, proposition 19 — it is easy to see that it holds globally. Indeed, it states that* $D^\nabla \mathcal{F} = 0$ *at leat locally. But the null section is defined globally and and equals* $D^\nabla \mathcal{F}$ *on every local trivialization. So we should have* $D^\nabla \mathcal{F} = 0$ *globally. We claim — without proving, however — that proposition 19 itself holds globally. These remarks will be useful on section 4.3.2.*

Lemma 4. *For any E-valued form α we have* $D^\nabla \alpha = d\alpha + \mathcal{A} \wedge \alpha$

Proof: Using lemma 3 and recalling $\mathcal{A} = \mathcal{A}^i_{\mu j} e_i \otimes e^j \otimes dx^\mu = \mathcal{A}_\mu \otimes dx^\mu$

$$\begin{gathered} D^\nabla \alpha = \nabla_\mu \alpha_I \otimes dx^\mu \wedge dx^I = (\partial_\mu + \mathcal{A}_\mu)\alpha_I \otimes dx^\mu \wedge dx^I = \\ \partial_\mu \alpha_I \otimes dx^\mu \wedge dx^I + \mathcal{A}_\mu(\alpha_I) \otimes dx^\mu \wedge dx^I = \\ d\alpha + \mathcal{A} \wedge \alpha \end{gathered}$$

Notice that we are using that locally the usual derivative is a covariant derivative. □

Proposition 20. $\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$.

Proof: For any E-valued form α:

$$\begin{gathered} D^\nabla D^\nabla \alpha = D^\nabla(d\alpha + \mathcal{A} \wedge \alpha) = d(d\alpha + \mathcal{A} \wedge \alpha) + \mathcal{A} \wedge (d\alpha + \mathcal{A} \wedge \alpha) = \\ d(\mathcal{A} \wedge \alpha) + \mathcal{A} \wedge d\alpha + \mathcal{A} \wedge \mathcal{A} \wedge \alpha = \\ d\mathcal{A} \wedge \alpha - \mathcal{A} \wedge d\alpha + \mathcal{A} \wedge d\alpha + \mathcal{A} \wedge \mathcal{A} \wedge \alpha = \\ = (d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}) \wedge \alpha \end{gathered}$$

On the other hand $D^\nabla D^\nabla \alpha = \mathcal{F} \wedge \alpha$. The result follows. □

Remark 13. *This is only another way of writing:*

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + [A_\mu, A_\nu]$$

Chapter 2

Clifford Algebras

Let V be a finite dimensional vector space over $\mathbb{R}$ with a quadratic form $<,>$. Denote by $\| x \|= \sqrt{< x, x >}$ We are using the notation of inner product because in most applications we have so, but $<,>$ needs to be only a quadratic form. Consider the **tensorial algebra**

$$T(V) = \bigoplus_{n\geq 0} \underbrace{V \otimes V \otimes \cdots \otimes V}_{n \text{ times}} = \\ \mathbb{R} \oplus V \oplus V \otimes V \oplus \cdots$$

Denote by $I(V)$ the two-sided ideal generated by $v \otimes w + w \otimes v + 2 < v, w > \cdot 1$.

Definition 38. *In the above notation the* ***Clifford algebra*** *of* V *is* $Cl(V) := T(V)/I(V)$.

In other words, $Cl(V)$ is an associative $\mathbb{R}$-algebra with unity generated by V subjected to the relations $v \otimes w + w \otimes v = -2 < v, w > \cdot 1$. We shall usually write vw instead of $v \otimes w$ and $< v, w >$ instead of $< v, w > \cdot 1$. Notice that if $\{e_1, \ldots, e_n\}$ is an orthonormal basis of V we have $e_j e_k = -e_k ej$ for $j \neq k$. It also follows from the definition that $v^2 = - \| v \|^2$.

Lemma 5. *(Fundamental lemma of Clifford algebras). Let* $\phi : V \to A$ *a linear map, where* A *is an associative algebra with unity. If it happens* $\phi(v) \cdot \phi(v) = - \| v \|^2 f$ *for every* $v \in V$ *then* ϕ *has a unique extension for an homomorphism of algebras (also denoted by* ϕ*)* $\phi : Cl(V) \to A$.

Proof: The fundamental theorem for tensorial algebras guarantees that ϕ uniquely extends to $T(V)$ as an homomorphism of algebras. The condition

$\phi(v) \cdot \phi(v) = (\phi(v))^2 = - \parallel v \parallel^2$ imply that $\ker(\phi : T(V) \to A)$ contains $I(V)$ and hence ϕ descends to the quotient. □

Remark 14. *If* $\dim(V) = n$ *we usually write* $Cl(V) = Cl(n)$.

Notation: From now up to the end of this chapter $\{e_1, \ldots, e_n\}$ will denote an orthonormal basis of V.

Notice that any $\xi \in Cl(V)$ is uniquely written as

$$\xi = a_0 + \sum_{k=1}^{n} \sum_{j_1 < \cdots < j_k} a_{j_1,\ldots,j_k} e_{j_1} \cdots e_{j_k}$$

where $a_{j_1,\ldots,j_k} \in \mathbb{R}$. If follows that $Cl(V)$ is a vector space of dimension 2^n (notice that $Cl(V)$ is isomorphic, *as vector space*, to $\bigwedge^* V$). If we denote by $Cl_0(V)$ ($Cl_1(V)$) the subspaces of $Cl(V)$ containing only an even (odd) number of elements on the base we have a $\mathbb{Z}_2$-graded splitting $Cl(V) = Cl_0(V) \oplus Cl_1(V)$. Notice that $Cl_0(V)$ is a sub-algebra and $Cl_1(V)$ is a module over $Cl_0(V)$. The elements of $Cl_0(V)$ ($Cl_1(V)$) commute (anti-commute) among themselves[1].

Proposition 21. *If* $(V_1, <,>_1)$ *and* $(V_2, <,>_2)$ *are vector spaces endowed with quadratic forms which are preserved by a linear map* $f : V_1 \to V_2$ *then* f *uniquely extends to* $f : Cl(V_1) \to Cl(V_2)$.

Proof:

$$f(x)f(x) = - < f(x), f(x) >_2 = < x, x >_1 = - \parallel x \parallel^2$$

The results follows from the fundamental lemma. □

Remark 15. *If* $(V_3, <,>_3)$ *is other vector space with a quadratic form preserved by* $g : V_2 \to V_3$*, then* $g \circ f : V_1 \to V_3$ *uniquely extends to a map between the algebras. By uniqueness this must be the composition of the extensions of* f *and* g.

Denote by $\widetilde{\mathrm{Aut}}(Cl(V))$ the automorphisms of $Cl(V)$ which preserve V.

[1] $\mathbb{Z}_2$-graded algebras such that the splitting if given in terms of commuting and anti-commuting parts are called in physicists' literature *superalgebras*

Proposition 22. $\widetilde{\mathrm{Aut}}(Cl(V)) \approx O(V)$.

Proof: $f \in O(V)$ uniquely extends to an automorphism of $Cl(V)$ (use the fundamental lemma and the above remark applied to $g = f^{-1}$). Reciprocally, if $f \in \widetilde{\mathrm{Aut}}(Cl(V))$ then

$$< f(v), f(v) >= -f(v)f(v) = f(-v^2) = f(\| v \|^2) =\| v \|^2 f(1) =< v, v >$$

□

In the next section we shall give an explicit description of the action of $O(V)$ on $Cl(V)$. Such description will be useful when we consider spinor bundles.

2.1 The isomorphism $\widetilde{\mathrm{Aut}}(Cl(V)) \approx O(V)$

Since $Cl(V)$ is an algebra with unity we can consider the *group of units* of $Cl(V)$ (i.e., the group of elements of $Cl(V)$ which are invertible), denoted by $Cl^{\times}(V)$. Notice that if $< v, v > \neq 0$ then $v \in Cl^{\times}(V)$; indeed $v^{-1} = -\frac{v}{\|v\|^2}$.

$Cl^{\times}(V)$ is a Lie group of dimension 2^n whose Lie algebra $\mathfrak{Cl}(V)$ is $Cl(V)$ itself. We have the **adjoint representation** $ad : Cl^{\times}(V) \to \mathrm{Aut}(Cl(V))$, $ad_u(x) = uxu^{-1}$ (a simple calculation shows that ad is indeed a representation). Derivating at the identity:

$$\frac{d}{dt}(ad_{\exp(tY)}(X))|_{t=0} = YX - XY = [Y, X]$$

so we have a representation of the Lie algebra: $Ad : \mathfrak{Cl}(V) \to \mathrm{End}(Cl(V))$, $Y \mapsto [Y, \cdot]$. Since

$$[Y, X]Z + X[Y, Z] = [Y, XZ]$$

we have that $[Y, \cdot]$ is actually a *derivation* of the Lie algebra, so we have $Ad : \mathfrak{Cl}(V) \to \mathrm{Der}(Cl(V))$.

Proposition 23. *Let $v, w \in V$, $< v, v > \neq 0$ and $v^{\perp}$ the space orthogonal to v. Then $-ad_v(w) = -vwv^{-1} = w - 2\frac{<v,w>}{<v,v>}v = R_{v^{\perp}}(w)$, where $R_{v^{\perp}}$ denotes the reflexion through $v^{\perp}$.*

Proof: Since $v^{-1} = -v/ < v, v >$ we have (using $vw + wv = -2 < v, w >$):

$$\begin{aligned} - < v, v > ad_v(w) = - < v, v > vwv^{-1} = vwv \\ = -v^2 w - 2 < v, w > v =< v, v > w - 2 < v, m > v \end{aligned}$$

□

Corollary 4. *$ad \in \widetilde{\mathrm{Aut}}(Cl(V))$.*

Proof: From the above proportion we have $ad_v(V) = V$. □

Up to a minus signal, we have that ad_v is a reflection through $v^{\perp}$. In order to get a reflection we consider:

Definition 39. *The **twisted adjoint representation** $\tilde{ad} : Cl^{\times}(V) \to \mathrm{Aut}(Cl(V))$ is given by $\tilde{ad}_u(x) = \alpha(u)xu^{-1}$, where $\alpha : Cl(V) \to Cl(V)$ is the automorphism which extends $\alpha : V \to V$, $\alpha(v) = -v$ (the existence of such extension is guaranteed by the fundamental lemma).*

A straightforward calculation gives:

Proposition 24. $\tilde{ad}_v(w) = w - 2\frac{<v,w>}{<v,v>}v = R_{v^{\perp}}(w)$

Denote by $\tilde{P}(V)$ the subgroup of $Cl^{\times}(V)$ whose elements are such that $\tilde{ad}(V) = V$. Restricting $\tilde{ad}$ we get a representation $\tilde{ad} : \tilde{P}(V) \to \widetilde{Aut}(CL(V)$.

Proposition 25. *The kernel of $\tilde{ad} : tildeP(V) \to \widetilde{Aut}(CL(V)$ is the set of non-null multiples of 1 (i.e. $\mathbb{R}$ or occasionally $\mathbb{C}$).*

Proof: [11] p. 14 □

Proposition 26. *$\tilde{ad}_u : V \to V$ preserves $<,>$ for every $u \in \tilde{P}(V)$. So we have a homomorphism $\tilde{ad} : \tilde{P}(V) \to O(V)$*

Proof: [11] p. 16 □

Consider now the subgroup $P(V)$ of $\tilde{P}(V)$ generated by elements $v \in V$ such that $< v, v > \neq 0$, i.e.,

$$P(V) = \{v_1 \cdots v_r \in Cl(V) :< v_i, v_i > \neq 0, v_1, \ldots, v_r \text{ is a finite sequence of } V\}$$

Restricting $\tilde{ad}$ to $P(V) \subset \tilde{P}(V)$ we get a homomorphism $\tilde{ad} : P(V) \to O(V)$. Since for each $v \in V$ $\tilde{ad}_v = R_{v^\perp}$ and every element of $P(V)$ is a product of vectors $v \in V$, we get:

$$\tilde{ad}_{v_1 \cdots v_r} = R_{v_1^\perp} \circ \cdots \circ R_{v_r^\perp}$$

By the Cartan-Dieudonné theorem $O(V)$ is generated by reflections, so $\tilde{ad} : P(V) \to O(V)$ is onto. On the other hand, restricting $\tilde{ad} : \tilde{P}(V) \to \widetilde{\mathrm{Aut}}(Cl(V))$ to $P(V)$ we get another homomorphism $\tilde{ad} : P(V) \to \widetilde{\mathrm{Aut}}(Cl(V))$. We have therefore two homomorphisms:

$$\tilde{ad} : P(V) \begin{array}{l} \nearrow \ \widetilde{\mathrm{Aut}}(CL(V)) \\ \\ \searrow \ O(V) \end{array}$$

But by 22 we have:

$$\tilde{ad} : P(V) \begin{array}{l} \nearrow \ \widetilde{\mathrm{Aut}}(CL(V)) \\ \quad\ \ \downarrow\approx \\ \searrow \ O(V) \end{array}$$

As the non-null multiples of 1 belong to $P(V)$, proposition 25 gives, together with the surjectivity of $\tilde{ad} : P(V) \to O(V)$ that:

Theorem 9. $\tilde{ad} : P(V)/\ker\{\tilde{ad} : P(V) \to O(V)\} \approx O(V) \approx \widetilde{\mathrm{Aut}}(CL(V)$

This is the main result of this section. It means that the action of $O(V)$ on $CL(V)$ is adjoint-like, i.e., any automorphism of $\widetilde{\mathrm{Aut}}(CL(V))$ is written as

$$\tilde{ad}_{v_1 \cdots v_r} = \tilde{ad}_{v_1} \circ \cdots \circ \tilde{ad}_{v_r}$$

where $v_1, \ldots, v_r \in P(V)$.

The Cartan-Dieudonné theorem also tells us that each element of $SO(V)$ is written as an even number of reflections. Considering $SP(V) := P(V) \cap Cl_0(V)$ and proceeding in analogous fashion we get

Theorem 10. $\tilde{ad} : SP(V)/\ker\{\tilde{ad} : SP(V) \to SO(V)\} \approx SO(V) \subset \widetilde{\mathrm{Aut}}(CL(V)$

This means that $SO(V)$ acts on $Cl(V)$ by

$$\tilde{ad}_{v_1 \cdots v_{2r}} = \tilde{ad}_{v_1} \circ \cdots \circ \tilde{ad}_{v_{2r}}$$

2.2 The groups *Pin* and *Spin*

Definition 40. *$Pin(V)$ is the subgroup of $Cl^{\times}(V)$ generated by elements $v \in V$ such that $\| v \|^2 = 1$. In other words:*

$$Pin(V) := \{\xi \in Cl^{\times}(V) : \xi = u_1 \cdots u_k,\ u_j \in V,\ \ \| u_j \| = 1\}$$

We also define: $Spin(V) := Pin(V) \cap Cl_0(V)$, in other words:

$$Spin(V) := \{\xi \in Cl^{\times}(V) : \xi = u_1 \cdots u_{2k},\ u_j \in V,\ \ \| u_j \| = 1\}$$

Some important remarks: (i) There exists an obvious homomorphism $Cl(V) \to \mathbb{Z}_2$, $Cl_0(V) \mapsto 1$, $Cl_1(V) \mapsto -1$. If we restrict to $Pin(V)$ then $Spin(V)$ is the kernel of $Pin(V) \to \mathbb{Z}_2$. (ii) If $\xi = u_1 \cdots u_k \in Pin(V)$ then $\xi^{-1} = (-u_k) \cdots (-u_1) \in Pin(V)$ (since $v^{-1} = -\frac{v}{\|v\|^2}$ for $< v, v > \neq 0$).

The group $Spin(V)$ acts on $Cl(V)$ by $Spin(V) \times Cl(V) \to Cl(V)$, $(\sigma, \xi) \mapsto ad_{\sigma} = \sigma \xi \sigma^{-1}$. Notice that here is not necessary to consider the twisted adjoint representation because the elements of $Spin(V)$ are products of an *even* number of vectors $v \in V$ — so that the cumbersome minus sign disappears. By the results of the last section ad_{σ} preserves V for every $\sigma \in Spin(V)$, hence we have a *representation* of $Spin(V)$ on $Cl(V)$, i.e., $ad : Spin(V) \to \widetilde{\mathrm{Aut}}(Cl(V))$, $\sigma \mapsto ad_{\sigma}$.

Some of the next results could be derived directly from theorem 10, but we shall provide explicit calculations.

Lemma 6. *The image of the representation of $Spin(V)$ on $Cl(V)$ consists of automorphisms which preserve V and orientation. Therefore we have a map from $Spin(V)$ onto $SO(V)$.*

Proof: We have already seen that ad_{σ} preserves V. Notice that for each $w \in V$ with $\| v \| = 1$, we have that $ad_w(v) = -R_{x^{\perp}}(v)$. Then, for $\xi = u_1 \cdots u_{2k} \in Spin(V)$, ad_{ξ} is an even composition of reflections and hence it preserves orientation. Since each element of $SO(V)$ is written as an even number of reflections, we have the result. □

Lemma 7. *The kernel of the representation* $ad : Spin(V) \to \widetilde{\mathrm{Aut}}(Cl(V))$ *is* $Spin(V) \cap \mathrm{cen}(Cl(V))$, *where* $\mathrm{cen}(Cl(V))$ *is the center of the Clifford algebra.*

Proof: obvious. □

From the surjectivity of $Spin(V)$ onto $SO(V)$ we get an isomorphism $Spin(V)/(Spin(V) \cap \mathrm{cen}(Cl(V)) \approx SO(V)$. So it is important to understand $\mathrm{cen}(Cl(V))$.

Lemma 8. *If* $\dim(V)$ *is even then* $\mathrm{cen}(Cl(V)) = \mathbb{R}$*; if* $\dim(V)$ *is odd than* $\mathrm{cen}(Cl(V)) = span\{1, e_1 \cdots e_n\} \approx \mathbb{R} \oplus \mathbb{R}$.

Proof: Fix an orthonormal basis $\{e_1, \ldots, e_n\}$ and notice that $e_j \cdot (e_{j_1} \cdots e_{j_t}) = \pm(e_{j_1} \cdots e_{j_t}) \cdot e_j$, where the sign is $(-1)^t$ if $j \neq j_r$ for every $1 \leq j \leq t$ and $(-1)^{t-1}$ if $j = j_r$ for some $r \in \{1, \ldots, t\}$. The result follows by inspection. □

Lemma 9. $Spin(V) \cap \mathrm{cen}(Cl(V)) = \{+1, -1\}$.

Proof: From the definition of $Spin(V)$ it follows that $Spin(V) \cap \mathrm{cen}(Cl(V)) = Spin(V) \cap \mathrm{cen}(Cl_0(V)) = Spin(V) \cap \mathbb{R}$. $Spin(V)$ is generated by unitary vectors v and the multiplication of two such vectors can result only in either $+1$ or -1 ($v^2 = -1$, $v(-v) = 1$). Then $Spin(V) \cap \mathbb{R} = \{+1, -1\}$. □

Putting these lemmas together:

Theorem 11. $Spin(V)/\{+1, -1\} \approx SO(V)$.

The important fact here is that $Spin(V)$ is not a double copy of $SO(V)$:

Theorem 12. *If* $\dim(V) \geq 2$ *then* $Spin(V) \to SO(V)$ *is a non-trivial double cover.*

Sketch proof: Let W be a 2-dimensional subspace of V. Notice that the pre-image of $SO(W) \subset SO(V)$ is $Spin(W) \subset Spin(V)$. It suffices to show that $Spin(W) \to SO(W)$ is a non-trivial double cover (since $\pi_1(SO(W)) \to \pi_1(SO(V))$ is surjective). Now notice that $Cl(W) \approx \mathbb{H}, Spin(W) \approx S^1 \subset \mathbb{C}$ and $W \approx Span\{\mathbf{j}, \mathbf{k}\}$. Therefore the conjugation of $Spin(W)$ on W is simply the square of the usual action of S^1 on W. □

It follows from the above theorem that if $\dim(V) = n$ then $Spin(V)$ is a compact Lie group of dimension $n(n-1)/2$, which is connected if $n > 1$ and simply connected if $n > 2$ — so for $n > 2$ the group $Spin(V)$ is a universal cover of $SO(V)$ and then the Lie algebras $\mathfrak{spin}(V)$ of $Spin(V)$ and $\mathfrak{so}(V)$ of $SO(V)$ are isomorphic.

2.3 Classification of Clifford algebras

Let $\mathbb{C}l(V)$ be the complexification of $Cl(V)$, i.e., $\mathbb{C}l(V) = Cl(V) \otimes \mathbb{C}$. This is equivalent of considering complex coefficients on the linear combinations. Given an orthonormal basis $\{e_1, \ldots, e_n\}$ define:

$$w_{\mathbb{C}} := i^{[\frac{n+1}{2}]} e_1 \cdots e_n$$

where i is the imaginary number and $[\cdot]$ is the integer part. A simply calculation shows that $w_{\mathbb{C}}$ is independent of the choice of basis and that $w_{\mathbb{C}}^2 = 1$. $w_{\mathbb{C}}$ acts on $\mathbb{C}l(V)$ by left multiplication. From $w_{\mathbb{C}}^2 = 1$ it follows that we can decompose $\mathbb{C}l(V)$ in $\mathbb{C}l^+(V)$ and $\mathbb{C}l^-(V)$, eigenspaces associated to ± 1 eigenvalues ($v \in \mathbb{C}l^{\pm}(V)$ if and only if $w_{\mathbb{C}} v = \pm v$).

Notice that if $\dim(V)$ is even, then $w_{\mathbb{C}} \in \text{cen}(\mathbb{C}l_0(V))$ and $w_{\mathbb{C}}$ anti-commutes with $\mathbb{C}l_1(V)$. If $\dim(V)$ is odd, then $w_{\mathbb{C}} \in \text{cen}(\mathbb{C}l(V))$.

Proposition 27. *If $\dim(V)$ is odd then the splitting $\mathbb{C}l^+(V) \oplus \mathbb{C}l^-(V)$ is orthogonal (i.e.: $\mathbb{C}l^{\pm}(V)$ are subalgebras which annihilates each other) and $\mathbb{C}l^{\pm}(V)$ are both isomorphic to $\mathbb{C}l_0(V)$ (isomorphism of algebras).*

Proof: Take $u \in \mathbb{C}l^+(V)$ and $v \in \mathbb{C}l^-(V)$. Then $w_{\mathbb{C}} u = u$ and $w_{\mathbb{C}} v = -v$. Since $w_{\mathbb{C}}$ is on the center of $\mathbb{C}l(V)$ (because the dimension of V is odd) we have $w_{\mathbb{C}} u w_{\mathbb{C}} v = u(-v) = w_{\mathbb{C}}^2 uv = uv$, so $uv = -uv$ and therefore $uv = 0$.

For the isomorphism, notice that $w_{\mathbb{C}}$ alternates $\mathbb{C}l_0(V)$ and $\mathbb{C}l_1(V)$, hence $\mathbb{C}l_0(V) \cap \mathbb{C}l^{\pm}(V) = \{0\}$. It follows that the compositions:

$$\mathbb{C}l_0(V) \hookrightarrow \mathbb{C}l(V) \xrightarrow{\pi^{\pm}} \mathbb{C}l^{\pm}(V)$$

are algebra isomorphisms. □

Remark 16. *notice that this embedding of* $\mathbb{C}l_0(V)$ *is the graph o f an isomorphism* $\mathbb{C}l^+(V) \approx \mathbb{C}l^-(V)$*.*

We have an analogous result for the case of even dimension:

Proposition 28. *If dim*(V) *is even then there is an orthogonal splitting* $\mathbb{C}l_0(V) = \mathbb{C}l_0^+(V) \oplus \mathbb{C}l_0^-(V)$ *and there is an algebra isomorphism* $\mathbb{C}l_0^+(V) \approx \mathbb{C}l_0^-(V)$*. Moreover,* $\mathbb{C}l_0^+(V)$ *is isomorphic (as algebra) to* $\mathbb{C}l(W)$*, where* $W \subset V$ *is a codimension 2 subspace.*

Proof: The orthogonal splitting follows, analogously to the above proposition, from the fact that for each n even we have $w_{\mathbb{C}} \in \mathrm{cen}(\mathbb{C}l_0(V))$. Notice that there is an isomorphism $\mathbb{C}l_0(V) \approx \mathbb{C}l(W')$, where W' is a subspace of codimension 1. Indeed $\dim(\mathbb{C}l_0(V)) = 2^n/2 = 2^{n-1} = \dim(\mathbb{C}l(W'))$, given a vector space isomorphism. Since $\dim(W')$=odd, by the above proposition we have $\mathbb{C}l(W') = \mathbb{C}l^+(W') \oplus \mathbb{C}l^-(W')$, hence we have the algebra isomorphism $\mathbb{C}l_0^+(V) \approx \mathbb{C}l^+(W')$. Using again the above proposition we have $\mathbb{C}l^+(W') \approx \mathbb{C}l_0(W')$, so $\mathbb{C}l_0^+(V) \approx \mathbb{C}l_0(W')$ (algebra isomorphism). Take now $W \subset W'$ a subspace of codimension 1. We have an algebra isomorphism $\mathbb{C}l_0(W') \approx \mathbb{C}l(W)$. The result follows. □

Proposition 29. *Similar result hold for* $\mathbb{C}l_1(V)$*, i.e.,* $\mathbb{C}l_1(V) = \mathbb{C}l_1^+(V) \oplus \mathbb{C}l_1^-(V)$*.*

Proof: Analogous. □

Summarizing, if the dimension of V is:

even:

$w_{\mathbb{C}} \in \mathrm{cen}(\mathbb{C}l_0(V)),\ \mathbb{C}l_0(V) \approx \mathbb{C}l_0^+(V) \oplus \mathbb{C}l_0^-(V),\ \mathbb{C}l_1(V) \approx \mathbb{C}l_1^+(V) \oplus \mathbb{C}l_1^-(V),$
$\mathbb{C}l_0^+(V) \approx \mathbb{C}l_0(W) \approx \mathbb{C}l_0^-(V)$

odd:

$w_{\mathbb{C}} \in \mathrm{cen}(\mathbb{C}l(V)),\ \mathbb{C}l(V) \approx \mathbb{C}l^+(V) \oplus \mathbb{C}l^-(V),$
$\mathbb{C}l_0(V) \approx \mathbb{C}l^+(W) \approx \mathbb{C}l^-(V)$

The lemma 11 below is the key lemma to classify Clifford algebras. In order to prove this lemma we must understand the tensor product of Clifford algebras.

We start recalling that if $\mathfrak{A}$ and $\mathfrak{B}$ are algebras with unit over K then the tensor product $\mathfrak{A} \otimes \mathfrak{B}$ is the algebra whose the underlying vector space is the tensor product of $\mathfrak{A}$ and $\mathfrak{B}$ and whose multiplication on simple elements is given by

$$(a \otimes b) \cdot (a' \otimes b') = (aa') \otimes (bb')$$

If, however $\mathfrak{A} = \mathfrak{A}^0 \oplus \mathfrak{A}^1$ and $\mathfrak{B} = \mathfrak{B}^0 \oplus \mathfrak{B}^1$ are $\mathbb{Z}_2$-graded algebras then we can introduce a second $\mathbb{Z}_2$-graded multiplication by the rule

$$(a \otimes b) \cdot (a' \otimes b') = (-1)^{\deg(b)\deg(a')}(aa') \otimes (bb')$$

for pure degree elements. The resulting algebra is called $\mathbb{Z}_2$-**graded tensor product** and is denoted by $\tilde{\otimes}$. It is also $\mathbb{Z}_2$-graded:

$$(\mathfrak{A}\tilde{\otimes}\mathfrak{B})^0 = \mathfrak{A}^0 \otimes \mathfrak{B}^0 + \mathfrak{A}^1 \otimes \mathfrak{B}^1$$
$$(\mathfrak{A}\tilde{\otimes}\mathfrak{B})^1 = \mathfrak{A}^1 \otimes \mathfrak{B}^0 + \mathfrak{A}^0 \otimes \mathfrak{B}^1$$

Lemma 10. *Let $V = V_1 \oplus V_2$ be a $<,>$-orthogonal decomposition of V (i.e., $< v_1 + v_2, v_1 + v_2 >=< v_1, v_1 > + < v_2, v_2 >$ for $v_1 \in V_1$ and $v_2 \in V_2$). Then there exists an isomorphism of Clifford algebras:*

$$\mathbb{C}l(V) \approx \mathbb{C}l(V_1)\tilde{\otimes}\mathbb{C}l(V_2)$$

where the quadratic form on V_i is $<,>$ restricted to V_i.

Proof: Consider the map $f : V \to \mathbb{C}l(V_1)\tilde{\otimes}\mathbb{C}l(V_2)$ given by $f(v) = v_1 \otimes 1 + 1 \otimes v_2$ where $v = v_1 + v_2$ is the decomposition of v with respect to the splitting $V = v_1 \oplus V_2$. From the $<,>$-orthogonality and the definition of $\mathbb{Z}_2$-graded tensor product we have that

$$f(v) \cdot f(v) = (v_1 \otimes 1 + 1 \otimes v_2)^2 = v_1^2 \otimes 1 + 1 \otimes v_2^2 =$$
$$- < v_1 + v_2, v_1 + v_2 > 1 \otimes 1 = - < v, v > 1 \otimes 1$$

Hence by the fundamental lemma f extends to an algebra homomorphism $f : \mathbb{C}l(V) \approx \mathbb{C}l(V_1)\tilde{\otimes}\mathbb{C}l(V_2)$. The image of f is a subalgebra which contains $\mathbb{C}l(V_1) \otimes 1$ and $1 \otimes \mathbb{C}l(V_2)$. Therefore f is onto. The injectivity follows easily by considering a basis for $\mathbb{C}l(V)$ generated by a basis of V which is compatible with the splitting. □

Lemma 11. $\mathbb{C}l(V \oplus \mathbb{R}^2) \approx \mathbb{C}l(V) \otimes \mathbb{C}l(\mathbb{R}^2)$

Proof: Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of V and $\{e_1, e_2\}$ an orthonormal basis of $\mathbb{R}^2$. Define: $\phi : V \oplus \mathbb{R}^2 \to \mathbb{C}l(V) \otimes \mathbb{C}l(\mathbb{R}^2)$ by $v_j \mapsto iv_j \oplus e_1 e_2$, $1 \leq j \leq n$ and $e_j \mapsto 1 \otimes e_j$, $j = 1, 2$. We shall show that for every u_i, u_j belonging to the basis of $V \oplus \mathbb{R}^2$ we have $(\phi(u_i))^2 = - \parallel u_i \parallel^2$ and $\phi(u_i)\phi(u_j) + \phi(u_j)\phi(u_i) = 0$ if $i \neq j$. A straightforward calculation shows that this is equivalent to the fundamental lemma and so it gives an extension $\phi : \mathbb{C}l(V \oplus \mathbb{R}^2) \to \mathbb{C}l(V) \otimes \mathbb{C}l(\mathbb{R}^2)$.

Compute:

$$\begin{aligned} \phi(v_j)\phi(v_j) = (iv_j \otimes e_1e_2)(iv_j \otimes e_1e_2) = -v_jv_j \otimes e_1e_2e_1e_2 = \\ = v_jv_j \otimes e_2e_1e_1e_2 = v_j^2 \otimes 1 = - \parallel v_j \parallel^2 \end{aligned}$$

The remaining calculations are analogous. So we have the desired extension. Notice that since ϕ takes the generators of $\mathbb{C}l(V \oplus \mathbb{R}^2)$ onto the generators of the $\mathbb{C}l(V) \otimes \mathbb{C}l(\mathbb{R}^2)$ we have an isomorphic of algebras since both have the same dimension. $\square$

Now we notice that $Cl(\mathbb{R}) \approx \mathbb{C}$, what implies $\mathbb{C}l(\mathbb{R}) \approx \mathbb{C} \oplus \mathbb{C}$; and $\mathbb{C}l(\mathbb{R}^2) \approx \mathbb{H}$, what implies $\mathbb{C}l(\mathbb{R}^2) \approx \mathbb{H} \oplus \mathbb{C}$. Define a map $\mathbb{H} \to M_{\mathbb{C}}(2)$ (= complex 2×2 matrices) by

$$\alpha + j\beta \mapsto \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix}$$

$\alpha, \beta \in \mathbb{C}$, we have an algebra isomorphism $\mathbb{H} \otimes \mathbb{C} \approx M_{\mathbb{C}}(2)$. We obtain therefore the following

Inductive rule: using the lemma 11 and $\mathbb{H} \otimes \mathbb{C} \approx M_{\mathbb{C}}(2)$ we can get inductively all the complex Clifford algebras.

Theorem 13. *If* $\dim(V) = 2m$ *than* $\mathbb{C}l(V) \approx M_{\mathbb{C}}(2^m)$*; if* $\dim(V) = 2m + 1$ *than* $\mathbb{C}l(V) \approx M_{\mathbb{C}}(2^m) \oplus M_{\mathbb{C}}(2^m)$.

Proof: Use the inductive rule and the isomorphic $M_{\mathbb{C}}(m) \otimes M_{\mathbb{C}}(k) \approx M_{\mathbb{C}}(mk)$. $\square$

Corollary 5. *If* $\dim(V) = 2m$ *then* $\mathbb{C}l_0^+(V) \approx M_{\mathbb{C}}(2^{m-1})$.

Proof: Use theorem 13 and proposition 28. □

We recall some definitions:

Definition 41. *If A is an algebra with unity and V is a finite dimensional vector space than a linear map $\rho : A \to \mathrm{End}(V)$ such that $\rho(ab) = \rho(a)\rho(b)$ and $1 \mapsto id$ is called a* ***representation*** *of A on V. The* ***dimension of the representation*** *is the dimension of V. Two representations $\rho_1 : A \to \mathrm{End}(V_1)$ and $\rho_2 : A \to \mathrm{End}(V_2)$ are said to be* ***equivalent*** *if there exists a linear and invertible map $f : V_1 \to V_2$ such that $\rho_2(a) = f \circ \rho_1(a) \circ f^{-1}$ for every $a \in A$.*

If $\rho_1 : A \to \mathrm{End}(V_1)$ and $\rho_2 : A \to \mathrm{End}(V_2)$ are two representations of A it follows that $\rho = \rho_1 \oplus \rho_2 : A \to \mathrm{End}(V_1 \oplus V_2)$, given by $\rho(a) = \rho_1(a) \oplus \rho_2(a)$ is also a representation.

Definition 42. *A representation ρ which can be written as $\rho = \rho_1 \oplus \rho_2$ is called* ***reducible****; otherwise it is called* ***irreducible****.*

Theorem 14. *If $\dim(V) = 2m$ then $\mathbb{C}l(V)$ has an unique (up to isomorphism) irreducible representation of dimension 2^m.*

Proof: By theorem 13 we have an isomorphism $\mathbb{C}l(V) \approx M_{\mathbb{C}}(2^m)$. By the Werderburn theorem $\mathrm{End}(\mathbb{C}^{2^m})$ has an unique (up to isomorphism) representation of dimension 2^m and in fact $\mathrm{End}(\mathbb{C}^{2^m}) \approx M_{\mathbb{C}}(2^m)$. □

Notice that the action of $\mathbb{C}l(V)$ on $\mathbb{C}^{2^m}$ induces a isomorphism $\mathbb{C}l(V) \approx \mathrm{End}(\mathbb{C}^{2^m}) \approx \mathbb{C}^{2^m} \otimes (\mathbb{C}^{2^m})^*$.

Theorem 15. *If $\dim(V) = 2m+1$ then $\mathbb{C}l(V)$ has two irreducible representation (up to isomorphism) of dimension 2^m. Moreover: $\mathbb{C}l_0(V) \approx \mathrm{End}(\mathbb{C}^{2^m})$.*

Proof: We have the isomorphisms:

$$\mathbb{C}l(V) \approx \mathbb{C}l^+(V) \oplus \mathbb{C}l^-(V) \overset{\text{prop.27}}{\approx} \mathbb{C}l_0(V) \oplus \mathbb{C}l_0(V)$$

Using theorem 13 on the other hand

$$\mathbb{C}l(V) \approx M_{\mathbb{C}}(2^m) \oplus M_{\mathbb{C}}(2^m) \overset{\text{Wederburn}}{\approx} \mathrm{End}(\mathbb{C}^{2^m}) \oplus \mathrm{End}(\mathbb{C}^{2^m})$$

□

Definition 43. *The action of* $\mathbb{C}l(V)$ *on* $\mathbb{C}^{2^m}$ *and* $\mathbb{C}^{2^m} \oplus \mathbb{C}^{2^m}$ *through the above isomorphisms is called* ***Clifford multiplication****. It is sometimes denoted by a dot: "·".*

Let $\dim(V) = 2m$ and consider the irreducible representation of $\mathbb{C}l(V)$ on $\mathbb{C}^{2^m}$. Through this representation $w_{\mathbb{C}}$ becomes a map $\hat{w}_{\mathbb{C}} : \mathbb{C}^{2^m} \to \mathbb{C}^{2^m}$ and therefore we have $\mathbb{C}^{2^m} = (\mathbb{C}^{2^m})^+ \oplus (\mathbb{C}^{2^m})^-$. We use this to prove:

Theorem 16. *Let* $\dim(V) = 2m$ *and consider the irreducible representation of* $\mathbb{C}l(V)$ *on* $\mathbb{C}^{2^m}$*. Clifford multiplication induces the following isomorphismm:*

$$\begin{aligned}
\mathbb{C}l_0^+(V) &\approx \mathrm{End}((\mathbb{C}^{2m})^+) \\
\mathbb{C}l_0^-(V) &\approx \mathrm{End}((\mathbb{C}^{2m})^-) \\
\mathbb{C}l_1^+(V) &\approx \mathrm{Hom}((\mathbb{C}^{2m})^-, (\mathbb{C}^{2m})^+) \\
\mathbb{C}l_1^-(V) &\approx \mathrm{Hom}((\mathbb{C}^{2m})^+, (\mathbb{C}^{2m})^-)
\end{aligned}$$

Proof: We recall that if $\dim(V)$=even then $w_{\mathbb{C}}$ commutes with $\mathbb{C}l_0(V)$ and anti-commutes with $\mathbb{C}l_1(V)$. Denote by ˆ the images of elements in $\mathbb{C}l(V)$ by the isomorphism $\mathrm{End}(\mathbb{C}^{2^m})$.

If $u \in \mathbb{C}l_0^+(V)$ then $w_{\mathbb{C}}u = uw_{\mathbb{C}}$ (because it belongs to $Cl_0(V)$) and $w_{\mathbb{C}}u = u$ (because it belongs to $Cl^+(V)$). Take $h \in (\mathbb{C}^{2^m})^+$. We have:

$$\hat{u} \circ \hat{w}_{\mathbb{C}}(h) = \hat{u}(h) = \hat{w}_{\mathbb{C}}(\hat{u}(h))$$

what means that $\hat{u}(h)$ is an eigenvector of $\hat{w}_{\mathbb{C}}$ with eigenvalue $+1$ and therefore $\hat{u}(h) \in (\mathbb{C}^{2^m})^+$, i.e., $\hat{u}$ leaves $(\mathbb{C}^{2^m})^+$ invariant. On the other hand, if $h \in (\mathbb{C}^{2^m})^-$ then:

$$\hat{u} \circ \hat{w}_{\mathbb{C}}(h) = -\hat{u}(h) = \hat{u}(h)$$

since $w_{\mathbb{C}}u = u$. Therefore $\hat{u}(h) = 0$ and hence $\hat{u} \in \mathrm{End}((\mathbb{C}^{2^m})^+)$. The isomorphism $\mathbb{C}l_0^-(V) \approx \mathrm{End}((\mathbb{C}^{2^m})^-)$ is proven in analogous fashion.

Take now $v \in \mathbb{C}l_1^+(V)$. Then $w_{\mathbb{C}}v = -vw_{\mathbb{C}}$ and $w_{\mathbb{C}}v = v$. If $h \in (\mathbb{C}^{2^m})^-$ we have

$$\hat{w}_{\mathbb{C}}(\hat{v}(h)) = -\hat{v}(-h) = \hat{v}(h)$$

what implies $\hat{v}(h) \in (\mathbb{C}^{2^m})^+$. Now for $h \in (\mathbb{C}^{2^m})^-$ we have

$$\hat{w}_{\mathbb{C}} \circ \hat{v}(h) = \hat{v}(h)$$

and

$$\hat{w}_{\mathbb{C}} \circ \hat{v}(h) = -\hat{v} \circ \hat{w}_{\mathbb{C}}(h) = -\hat{v}(h)$$

what implies $\hat{v}(h) = 0$ and therefore $\hat{h} \in \mathrm{Hom}((\mathbb{C}^{2^m})^-, (\mathbb{C}^{2^m})^+)$. The proof is analogous for $\mathbb{C}l_1^-(V)$. □

Remark 17. *For* $\dim(V) = 2$, *considering the isomorphism* $\mathbb{C}l(V) \approx \mathrm{End}(\mathbb{C}^{2^m})$ *we can give a nice interpretation to the last theorem. The splitting* $\mathbb{C}l(V) = \mathbb{C}l_0^+(V) \oplus \mathbb{C}l_0^- \oplus \mathbb{C}l_1^+(V) \oplus \mathbb{C}l_1^-(V)$ *implies that if we write a vector* $v \in \mathbb{C}^{2^m} = (\mathbb{C}^{2^m})^+ \oplus (\mathbb{C}^{2^m})^-$ *as* $v = v^+ + v^-$, *then a matrix* $M \in M_{\mathbb{C}}(2^m)$ *can be written as:*

$$M = \underbrace{\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}}_{(\mathbb{C}^{2m})^+ \to (\mathbb{C}^{2m})^+} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix}}_{(\mathbb{C}^{2m})^- \to (\mathbb{C}^{2m})^-} + \underbrace{\begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}}_{(\mathbb{C}^{2m})^- \to (\mathbb{C}^{2m})^+} + \underbrace{\begin{pmatrix} 0 & 0 \\ D & 0 \end{pmatrix}}_{(\mathbb{C}^{2m})^+ \to (\mathbb{C}^{2m})^-}$$

where $A, B, C, D \in M_{\mathbb{C}}(2^{m-1})$.

Since $Spin(V) \subset \mathbb{C}l(V)$ we can restrict the representations of $\mathbb{C}l(V)$ and such a way we obtain:

Proposition 30. *There exists an unique (complex) representation of* $Spin(V)$, $\Delta_{\mathbb{C}} : Spin(V) \to \mathrm{Aut}(\mathbb{C}^{2^m})$, *which is induced by irreducible representations of* $\mathbb{C}l(V)$.

Proof: It follows from $Spin(V) \subset \mathbb{C}l_0(V)$ and theorem 14. □

Proposition 31. *If* $\dim(V) = 2m$ *then* $\Delta_{\mathbb{C}}$ *splits in two inequivalent representations,* $\Delta_{\mathbb{C}} = \Delta_{\mathbb{C}}^+ \oplus \Delta_{\mathbb{C}}^-$, *each one of dimension* 2^{m-1}. *If* $\dim(V) = 2m+1$ *then* $\Delta_{\mathbb{C}}$ *is a* 2^m*-dimensional irreducible representation of* $Spin(V)$.

Sketch proof: For $\dim(V) = 2m$ we have $\mathbb{C}l_0(V) = \mathbb{C}l_0^+(V) \oplus \mathbb{C}l_0^-(V)$, $\mathbb{C}l_0^+(V) \approx \mathrm{End}((\mathbb{C}^{2^m})^+)$ and $\mathbb{C}l_0^-(V) \approx \mathrm{End}((\mathbb{C}^{2^m})^-)$. So, restricting to $Spin(V) \subset \mathbb{C}l_0(V)$ these isomorphisms give $\Delta_{\mathbb{C}}^+ : Spin(V) \to \mathrm{Aut}((\mathbb{C}^{2^m})^+)$ and $\Delta_{\mathbb{C}}^+ : Spin(V) \to \mathrm{Aut}((\mathbb{C}^{2^m})^+)$. It follows from theorem 16 that these are the only irreducible inequivalent representations (up to isomorphism) of $Spin(V)$. We also have $\dim((\mathbb{C}^{2^m})^+) = \dim((\mathbb{C}^{2^m})^-) = 2^m/2 = 2^{m-1}$.

For $\dim(V) = 2m+1$ theorem 15 give us two 2^m-dimensional irreducible representations of $\mathbb{C}l(V)$ and $\mathbb{C}l_0(V) \approx \mathrm{End}(\mathbb{C}^{2^m})$. The result follows from $Spin(V) \subset \mathbb{C}l_0(V)$.

A detailed proof can be found in [12] p. 19. □

Remark 18. *We are not saying — and it is* not *true — that $\Delta_{\mathbb{C}}$ is the unique complex irreducible representation of $Spin(V)$ for* $\dim(V) = 2m+1$ *neither that $\Delta^{\pm}_{\mathbb{C}}$ are the only two irreducible representations of $Spin(V)$ for* $\dim(V) = 2m$. *What we claim is that these are the* only ones which extend to (automatically irreducible) representations of $\mathbb{C}l_0(V)$.

2.4 The group $Spin^c$

Definition 44. *The subgroup of $Cl^{\times}(V)$ generated by $Spin(V)$ and S^1 is called the $Spin^c(V)$ group, i.e.,*

$$Spin^c(V) := \{a \in Cl^{\times}(V) : a = a_0 u_1 \cdots u_2 k, a_0 \in S^1, u_i \in V, \| u_i \| = 1\}$$

Proposition 32. $Spin^c(V) \approx Spin(V) \times S^1/\{(1,1),(-1,-1)\}$.

Proof: Define a map $Spin(V) \times S^1 \to Spin^c(V)$ by $(\sigma, a) \mapsto a\sigma$. Since elements of S^1 commute with elements of $Spin(V)$ such map is onto. Its kernel are elements of the form (σ, σ^{-1}) because $a\sigma = 1$ gives $a = \sigma^{-1}$. It follows from lemmas 8 and 9 that$Spin(V) \cap S^1 = \{+1, -1\}$. The result follows. □

Using the theorem 11 we get:

Corollary 6. $Spin(V)/\pm 1 \approx SO(V) \times S^1$.

It follows that $Spin^c(V)$ is a non-trivial double cover of $SO(V) \times S^1$.

The action of $Spin(V)$ on $\mathbb{C}l(V)$ extends to an action of $Spin^c(V)$ on $\mathbb{C}l(V)$. Notice that $Spin(V)$ is identified with $(Spin(V) \times \pm 1)/\{\pm 1\}$. Moreover: since $\Delta_{\mathbb{C}}$ is obtained by restricting a representation of $\mathbb{C}l(V)$, it follows that $\Delta_{\mathbb{C}}$ extends to a representation $\tilde{\Delta}_{\mathbb{C}} : Spin^c(V) \to \mathrm{Aut}(\mathbb{C}^{2^m})$ and that this representation is "unique" in the sense of remark 18. It also follows that if $\dim(V) = 2m$ then $\tilde{\Delta}_{\mathbb{C}} = \tilde{\Delta}^{+}_{\mathbb{C}} \oplus \tilde{\Delta}^{-}_{\mathbb{C}}$, where $\tilde{\Delta}^{\pm}_{\mathbb{C}}$ are extensions of $\Delta^{\pm}_{\mathbb{C}}$.

Chapter 3

Spinor bundles

3.1 Spin structures and Dirac operator

We shall assume that the vector space V for constructing Clifford algebras and spin-like groups has dimension ≥ 2. Instead of writing $SO(V), Spin(V)$ etc we shall write $SO(n), Spin(n)$ etc where $n = \dim(V)$.

Throughout this chapter E denotes a vector bundle over a Riemannian manifold $\mathcal{M}$ and $P_{SO(n)}(E)$ its bundle of positively oriented orthonormal frames, i.e., a $SO(n)$ principal bundle whose fiber on $x \in \mathcal{M}$ is the set of all positively oriented orthonormal basis of E_x ([11] p 78). Recall that for $n \geq 3$ we have a universal double cover $\lambda : Spin(n) \to SO(n)$.

Definition 45. *Let $n \geq 3$. A **spin structure** on E is a $Spin(n)$-principal bundle $P_{Spin(n)}(E)$ together with a 2-sheeting cover $\Lambda : P_{Spin(n)}(E) \to P_{SO(n)}(E)$ such that $\Lambda(pg) = \Lambda(p)\lambda(g)$ for every $p \in P_{Spin(n)}(E)$ and every $g \in Spin(n)$. When $E = T\mathcal{M}$ we call $\mathcal{M}$ a **spin manifold**.*

It is also possible to define spin structures for $n = 1, 2$ (see [11] p.80). Notice that Λ restricted to fibers equals λ. We have the following commutative diagram of bundles:

$$
\begin{array}{ccccc}
 & & Spin(n) & \xrightarrow{\lambda} & SO(n) \\
 & \nearrow & & & \\
\mathbb{Z}_2 & & \downarrow & & \downarrow \\
 & \searrow & & & \\
 & & P_{Spin(n)}(E) & \xrightarrow{\Lambda} & P_{SO(n)}(E) \\
 & & \searrow & & \swarrow \\
 & & & \mathcal{M} &
\end{array}
$$

Denoting by $\rho : SO(n) \to Aut(V)$ the trivial representation and by $P_{SO(n)}(\mathcal{M})$ the bundle of positively oriented orthonormal frames of a oriented Riemannian manifold $\mathcal{M}$, we have that $T\mathcal{M} = P_{SO(n)}(\mathcal{M}) \times_\rho \mathbb{R}^n$. Analogously if E is a Riemannian oriented vector bundle we have $E = P_{SO(n)}(E) \times_\rho \mathbb{R}^n$. As we have seen the action of $SO(n)$ on $\mathbb{R}^n$ extends to an action on $Cl(n)$, so there is a map $cl : SO(n) \to \mathrm{Aut}(Cl(n))$, i.e., a representation. Then, analogously to the cases $T\mathcal{M}$ and E we can define:

Definition 46. *The **Clifford bundle** of a Riemannian oriented vector bundle E is defined as $Cl(E) := P_{SO(n)}(E) \times_{cl} Cl(n)$, where $cl : SO(n) \to \mathrm{Aut}(Cl(n))$ is the representation which extends the natural action of $SO(n)$ on $\mathbb{R}^n$.*

Therefore, the Clifford bundle is a fiber bundle whose fibers are Clifford algebras.

Definition 47. *Let $\Lambda : P_{Spin(n)}(E) \to P_{SO(n)}(E)$ be a spin structure on E. Let $\mathfrak{m}$ be a left module over $Cl(n)$ and $\mu : Spin(n) \to SO(\mathfrak{m}) \subset \mathrm{Aut}(\mathfrak{m})$ a representation of $Spin(n)$ given by left multiplication of elements of $Spin(n) \subset Cl(n)$. A **spinor bundle** for E is a fiber bundle $S(E)$ given by $S(E) := P_{Spin(n)}(E) \times_\mu \mathfrak{m}$. If $\mathfrak{m}_{\mathbb{C}}$ is a left module over $\mathbb{C}l(n)$ then $S_{\mathbb{C}}(E) := P_{Spin(n)}(E) \times_\mu \mathfrak{m}_{\mathbb{C}}$ is called **complex spinor bundle***

Definition 48. *A section of $S(E)$ is called a **spinor field**.*

Proposition 33. *$S(E)$ is a bundle of modules over the bundle of algebras $Cl(E)$.*

Proof: Consider the right action of $Spin(n)$ on $P_{Spin(n)}(E) \times (Cl(n) \times \mathfrak{m})$ and on $P_{Spin(n)}(E) \times \mathfrak{m}$ given by

$$(p, \phi, m) \mapsto (R_g(p), ad_{g^{-1}}\phi, g^{-1}m) = (pg, g^{-1}\phi g, g^{-1}m)$$
$$(p, m) \mapsto (R_g(p), g^{-1}m) = (pg, g^{-1}m)$$

respectively. So the diagram:

$$\begin{array}{ccc} P_{Spin(n)}(E)\times(Cl(n)\times\mathfrak{m}) & \longrightarrow & P_{Spin(n)}(E)\times\mathfrak{m} \\ {\scriptstyle (R_g, ad_{g^{-1}}, g^{-1})}\Big\downarrow & & \Big\downarrow{\scriptstyle (R_g, g^{-1})} \\ P_{Spin(n)}(E)\times(Cl(n)\times\mathfrak{m}) & \longrightarrow & P_{Spin(n)}(E)\times\mathfrak{m} \end{array}$$

commutes; in the diagram the horizontal arrows mean multiplication of $\mathfrak{m}$ by $Cl(n)$: $(p,\phi,m)\mapsto(p,\phi m)$. In other words, the multiplication $Cl(n)\times\mathfrak{m}\to\mathfrak{m}$ is equivariant with respect to the action of $Spin(n)$, hence it descends to the quotient: $Cl(E)\oplus S(E)\to S(E)$. □

It is instructive to see directly the action on the quotient. The multiplication $Cl(E)\oplus S(E)\to S(E)$ is given by $[(p,\phi,x)]\mapsto[(p,\phi x)]$. This is independent of the choice (p,ϕ,x) because if $(p',\phi',x')\in[(p,\phi x)]$ then

$$(p',\phi',x')=(pg,g^{-1}\phi g,g^{-1})\mapsto(pg,g^{-1}\phi gg^{-1}x)=(pg,g^{-1})\sim(p,\phi x)$$

In particular, $\Gamma(S(E))$ is a module over $\Gamma(Cl(E))$:

Corollary 7. *For every $\phi\in\Gamma(Cl(E))$, every $\sigma\in\Gamma(S(E))$ and every $g\in Spin(n)$ we have: $\mu(g)(\phi\sigma)=cl(g)(\phi)\mu(g)(\sigma)$, where $cl: Spin(n)\to\mathrm{Aut}(Cl(n))$ is given by $cl(g))(\phi)=g\phi g^{-1}=ad_g(\phi)$.*

Proof: $\mu(g)$ acting on $\phi\sigma$ is, by definition, left multiplication by g, hence

$$\mu(g)(\phi\sigma)=g\phi\sigma=g\phi g^{-1}g\sigma=cl(g)(\phi)\mu(g)(\sigma)$$

□

Given a connection w on $P_{SO(n)}(E)$ we have a covariant derivative ∇^w on $Cl(E)$. If E has a spin structure then the connection on $P_{SO(n)}(E)$ lifts to a connection $\tilde{w}$ on $P_{Spin(n)}(E)$, inducing a covariant derivative $\nabla^{\tilde{w}}$ on $S(E)$.

Theorem 17. *The covariant derivative ∇^w on $Cl(E)$ acts as a derivation, i.e.,*

$$\nabla^w(\psi\phi)=(\nabla^w\psi)\phi+\phi(\nabla^w\phi)$$

for every $\psi,\phi\in\Gamma(Cl(E)$.

Proof: Let $\phi, \psi \in \Gamma(Cl(E))$ and $\sigma \in \Gamma(S(E))$. Since the representation $cl : Spin(n) \to \mathrm{Aut}(Cl(n))$ is given by $cl(g) = Ad_g$, we have by a standard argument (similar to that of section 3.1) that $cl_* : \mathfrak{spin}(n)(\approx \mathfrak{so}(n)) \to \mathrm{Der}(Cl(n)) \subset \mathrm{End}(Cl(n))$, $Y \mapsto [Y, \cdot]$. It follows that

$$cl_*(\psi\phi) = (cl_*(Y)(\psi))\phi + \psi(cl_*(Y)(\phi))$$

Recall from the chapter 1 that ∇^w is written as:

$$\nabla^w_X(\psi\phi)(p) = D(\psi\phi)(p)(X(p)) + \mathcal{A}^X_\alpha(p)((\psi\phi)(p))$$

where X is a vector field on the base space $\mathcal{M}$. Remembering the definition of $\mathcal{A}_\alpha$ (chapter 1) we have $\mathcal{A}^X_\alpha(p) = cl_*(A_\alpha(p)(X(p))$. Therefore

$$\begin{aligned}\nabla^w_X(\psi\phi)(p) = &\big(D\psi(p)(X(p))\big)\phi(p) + \psi(p)\big(D\phi(p)(X(p))\big) + \\ &cl_*\big(\underbrace{A_\alpha(p)(X(p))}_{=Y\in\mathfrak{spin}(n)}\big)((\psi\phi)(p)) \\ = &\big(D\psi(p)(X(p))\big)\phi(p) + \psi(p)\big(D\phi(p)(X(p))\big) + \\ &\big(cl_*(Y)(\psi)(p)\big)\phi(p) + \psi(p)\big(cl_*(Y)(\phi)(p)\big)\end{aligned}$$

Rearranging the terms we have the result. □

Theorem 18. *The covariant derivative $\nabla^{\tilde{w}}$ on $S(E)$ acts as a derivation over the module structure of $\mathfrak{m}$ over $Cl(E)$, i.e.,*

$$\nabla^{\tilde{w}}(\phi\sigma) = (\nabla^w\phi)\sigma + \phi(\nabla^{\tilde{w}}\sigma)$$

for every $\phi \in \Gamma(Cl(E)$ and every $\sigma \in \Gamma(S(E))$

Proof: From corollary 7 we have $\mu(g)(\phi\sigma) = cl(g)(\phi)\mu(g)(\sigma)$. Derivating at the identity we get:

$$\mu_*(\phi\sigma) = (cl_*(\phi))\sigma + \phi(\mu_*(\sigma))$$

Now the argument is very similar to the last proportion. □

When $E = T\mathcal{M}$ it is usual do denote $Cl(E)$ by $CL(\mathcal{M})$ and $S(E)$ by $S(\mathcal{M})$. We viewed that in $S(\mathcal{M})$ we have $\nabla^{\tilde{w}} : \Gamma(S(\mathcal{M})) \to \Gamma(S(\mathcal{M}) \otimes T^*\mathcal{M})$.

Definition 49. *The* ***Dirac operator*** $\not{D} : \Gamma(S(\mathcal{M})) \to \Gamma(S(\mathcal{M}))$ *is defined by*

$$(\not{D}\sigma)(x) = \sum_{j=1}^{n} e_j \cdot \nabla_{e_j}^{\tilde{w}} \sigma(x)$$

where $\{e_j\}$ *is an orthonormal basis of* $T\mathcal{M}_x$ *and* $\cdot$ *is the Clifford module multiplication.*

It can be shown that the Dirac operator does not depend on the particular choice of basis.

3.2 *Spin*^{c} structures

Up to this point we have proceed in full generality. From now on we shall give a more explicit — although particular — description of spinor bundles. For concreteness, we shall state the results even though some of them are a repetition, in a particular case, of the above results.

Consider the complexified Clifford algebra $\mathbb{C}l(n)$. We know from theorems 14 and 15 that if $n = 2m$ then $\mathbb{C}l(n)$ has an unique irreducible representation of dimension 2^m; and if $n = 2m + 1$ it has two inequivalent irreducible representations of dimension 2^m, i.e., $\mathbb{C}l(2m) \approx \text{End}(\mathbb{C}^{2^m})$ and $\mathbb{C}l(2m+1) \approx \text{End}(\mathbb{C}^{2^m}) \oplus \text{End}(\mathbb{C}^{2^m})$. Notice that the endomorphisms can be seen as a left module over the vector space in which they act, so the Clifford algebra is a left module over $\mathbb{C}^{2^m}$ or $\mathbb{C}^{2^m} \oplus \mathbb{C}^{2^m}$ (depending on if n is even or odd), giving the module $\mathfrak{m}$ of the above constructions. These constructions involve, also, representations $cl : SO(n) \to \text{Aut}(Cl(n))$ and $\mu : Spin(n) \to Aut(\mathfrak{m})$. The first is given by 10 and the second by proposition 31: we have two inequivalent irreducible representations of dimension 2^{m-1} if $n = 2m$ and a unique 2^m irreducible representations of dimension 2^m if $n = 2m+1$. Therefore we have the spin structure $\Lambda : P_{Spin(n)}(E) \to P_{SO(n)}(E)$ and respective associated spinor bundles. For example, for $n = 2m + 1$ we have:

$$
\begin{array}{ccccccccc}
Spin(n) & \hookrightarrow & P_{Spin(n)}(E) & \overset{\mu}{\Rightarrow} & \mathbb{C}^{2m} & \hookrightarrow & S(E) & \searrow & \\
\big\downarrow \lambda & & \big\downarrow \Lambda & & & & & & \mathcal{M} \\
SO(n) & \hookrightarrow & P_{SO(n)}(E) & \overset{cl}{\Rightarrow} & \mathbb{C}l(n) & \hookrightarrow & \mathbb{C}l(E) & \nearrow & \\
 & & \big\downarrow & & & & & & \\
 & & \mathcal{M} & & & & & &
\end{array}
$$

together with the Clifford multiplication $\Gamma(Cl(E))\times\Gamma(S(E)) \to \Gamma(S(E))$.

Recall that $Spin^c(n)$ is given by $(Spin(n) \times S^1)/\{\pm 1\} = Spin(n) \times_{\mathbb{Z}_2} S^1$ (see proposition 32), i.e., in $Spin(n) \times S^1$ we identify $(g, z) \sim (-g, -z)$. We have the following homomorphisms:

- $\lambda : Spin^c((n) \to SO(n), \lambda([g, z]) = \lambda([g])$.
- $i_1 : Spin(n) \to Spin^c(n), i_1(g) = [g, 1]$.
- $i_2 : S^1 \to Spin^c(n), i_2(z) = [1, z]$.
- $l : Spin^c(n) \to S^1, l([g, z]) = z^2$.
- $p : Spin^c(n) \to SO(n) \times S^1, p([g, z]) = (\lambda(g), z^2)$, i.e., $p = \lambda \times l$.

As usual, we shall denote by $\mathcal{M}$ a positively oriented Riemannian manifold and $P_{SO(n)}(E)$ its bundle of positively oriented frames. A known result states that there exists a unique torsion free connection on $\mathcal{M}$ compatible with the metric. Seen on $P_{SO(n)}(E)$ this connection is a 1-form $w : TP_{SO(n)}(E)(\mathcal{M}) \to \mathfrak{so}$.

Definition 50. *If $\pi : E \to \mathcal{M}$ is a oriented Riemannian vector bundle, a* ***spin^c structure*** *on E is a $Spin^c(n)$-bundle $P_{Spin^c(n)}(E)$ together with a map $\Lambda : P_{Spin^c(n)}(E) \to P_{SO(n)}(E)$ such that:*

$$
\begin{array}{ccc}
P_{Spin^c(n)}(E) \times Spin^c(n) & \longrightarrow & P_{Spin^c(n)}(E) \\
\big\downarrow \Lambda \times \lambda & & \big\downarrow \Lambda \\
P_{SO(n)}(E) \times SO(n) & \longrightarrow & P_{SO(n)}(E)
\end{array}
$$

commutes

Remark 19. *Every spin structure induces a spin^c structure through the inclusion $i_1 : Spin(n) \to Spin^c(n)$.*

As before, we shall denote by $P_{Spin^c(n)}(\mathcal{M})$ the bundle $P_{Spin^c(n)}(E)$ when $E = T\mathcal{M}$. Notice that $P_{Spin^c(n)}(\mathcal{M})/S^1$ is a $Spin(n)/\{\pm 1\} = SO(n)$-principal bundle isomorphic to $P_{SO(n)}(M)$ and $P_1 := P_{Spin^c(n)}(\mathcal{M})/Spin(n)$ is a $S^1/\{\pm\} = S^1$-principal bundle over $\mathcal{M}$. We have therefore a double cover $\pi : P_{Spin^c(n)}(\mathcal{M}) \to P_{SO(n)}(M)\tilde{\times}P_1$.

Take a connection $\alpha : TP_1 \to i\mathbb{R}$ on P_1; we are identifying the Lie algebra of S^1 — which is $\mathbb{R}$ — with the purely imaginary numbers. the connections w and α define a connection $w \times \alpha : T(PsM\tilde{\times}P_1) \to \mathfrak{so}(n) \oplus i\mathbb{R}$. Such connection lift through $\pi : P_{Spin^c(n)}(\mathcal{M}) \to P_{SO(n)}(M)\tilde{\times}P_1$ to a connection $\widetilde{w \times \alpha}$ on $P_{Spin^c(n)}(\mathcal{M})$ ([13] p. 57). The following diagram commutes:

$$\begin{array}{ccc} TP_{Spin^c(n)}(\mathcal{M}) & \xrightarrow{\widetilde{w\times\alpha}} & \mathfrak{spin}^c(n) \\ \Big\downarrow D\pi & & \Big\downarrow p_* = (\lambda_*, l_*) \\ T(P_{SO(n)}(M)\tilde{\times}P_1) & \xrightarrow{\widetilde{w\times\alpha}} & \mathfrak{so}(n) \oplus i\mathbb{R} \end{array}$$

where p_* is the differential of $p : Spin^c(n) \to SO(n) \times S^1$ and $\mathfrak{spin}^c(n)$ is the Lie algebra of $Spin^c(n) \approx \mathfrak{so}(n) \times i\mathbb{R}$ is the Lie algebra of $Spin^c(n)$ (recall that $\mathfrak{spin}(n) \approx \mathfrak{so}(n)$).

Since $Spin^c(n)$ is contained in $\mathbb{C}l(n)$, the representation $\mu : Spin(n) \to \mathrm{Aut}(\mathfrak{m})$ extends to a representation of $Spin^c(n)$, also denoted by μ. Recall that here $\mathfrak{m}$ is either $\mathbb{C}^{2^{m-1}} \oplus \mathbb{C}^{2^{m-1}}$ or $\mathbb{C}^{2^m}$ depending on $n = 2m$ or $n = 2m+1$. So, this representation of $Spin^c(n)$ gives rise to a spinor bundle with structural group $Spin^c(n)$ — which we still denote by $S(\mathcal{M})$.

Now the connection $\widetilde{w \times \alpha}$ on $P_{Spin^c(n)}(\mathcal{M})$ determines a connection / covariant derivative $\nabla^\alpha : \Gamma(S(\mathcal{M})) \to \Gamma(T^*\mathcal{M} \otimes S(\mathcal{M}))$. We denote it by ∇^α instead of $\nabla^{w\times\alpha}$ because w is thought as "fixed", i.e., it will always be the only torsion free connection compatible with the metric, whereas α is "chosen" on P_1.

Given a vector field X on $\mathcal{M}$ and a spinor field ψ, i.e., a section of $S(\mathcal{M})$, the Clifford multiplication allow us to multiply $X \cdot \psi$, $(X\psi)(p) = X(p) \cdot \psi(p)$. Indeed, recall that $\mathbb{R} \subset Cl(n) \subset \mathbb{C}l(n) \approx \mathrm{End}(\mathfrak{m})$, so $X(p) \in \mathbb{R}^n$ can be identified with an endomorphism of the fiber over p and hence it acts on $\psi(p) \in S(\mathcal{M})_p \approx \mathfrak{m}$. Now we can state:

Theorem 19. *Let X, Y be vector fields on $\mathcal{M}$ and $\psi \in \Gamma(S(\mathcal{M}))$. Then, for every connection α on P_1 we have:*

$$\nabla^\alpha_Y(X \cdot \psi) = X \cdot (\nabla^\alpha_Y \psi) + (\nabla_Y X) \cdot \psi$$

and

$$X < \psi_1, \psi_2 >=< \nabla^\alpha_X \psi_1, \psi_2 > + < \psi_1, \nabla^\alpha_X \psi_2 >$$

for every $\psi_1, \psi_2 \in \Gamma(S(\mathcal{M}))$.

The proof of the first statement is "essentially" the same of theorems 17 and 18. However, we put it here.

Proof: ∇^α is written as:

$$\nabla^w(X \cdot \psi)(p)(Y(p)) = D(X\psi)(p)(Y(p)) + \mathcal{A}^Y_\sigma(p)((X\psi)(p))$$

where $\mathcal{A}_\sigma$ is, as usual, given by

$$\mathcal{A}^Y_\sigma(p) = \mu_* \circ \mathcal{A}_\sigma(p)(Y(p)) = \mu_* \circ \sigma^*(\widetilde{w \times \alpha})(p)(Y(p))$$

Since $\sigma^*(\widetilde{w \times \alpha})(p)(Y(p)) \in \mathfrak{spin}^c(n)$ we can write $\sigma^*(\widetilde{w \times \alpha})(p)(Y(p)) = y + it$, where $y \in \mathfrak{so}(n)$ and $t \in \mathbb{R}$. Since $\mu_* : \mathfrak{spin}^c(n) \to \mathrm{End}(\mathfrak{m}) \approx \mathbb{C}l(n)$, we have that $\mu_*(y + it)$ is simply Clifford multiplication by $y + it$. Therefore

$$\mathcal{A}^Y_\sigma(p)((X\psi)(p)) = (y + it) \cdot X(p)\psi(p) = y \cdot X(p) \times \psi(p) + X(p) \cdot (it\psi(p))$$

The covering map $\lambda : Spin^c((n) \to SO(n)$ satisfies:

$$\lambda_*(z)x = zx - xz$$

for every $z \in \mathfrak{spin}(n)$ and every $x \in \mathbb{R}^n$. The proof of this equality is simple but takes long time. The reader may see [13] p. 16ff. By the commutativeness of 3.2 we have $\lambda_*(y) = wD\pi(v)$, where $v \in TP_{Spin^c(n)}(\mathcal{M})$; indeed:

$$y + it = \sigma^*(\widetilde{w \times \alpha})(p)(Y(p)) = (\widetilde{w \times \alpha})(\sigma(p))(D\sigma(p)Y(p))$$

what implies

$$y = \tilde{w}(\underbrace{D\sigma(p)(Y(p))}_{=v})$$

Then:

$$\begin{gathered}\mathcal{A}^Y_\sigma(p)((X\pi)(p)) = X(p) \cdot y \cdot \psi(p) + (\lambda_*(y)(X(p))) \cdot \psi(p) + X(p) \cdot (it\psi(p)) = \\ X(p) \cdot y \cdot \psi(p) + (w(D\pi(v))(X(p))) \cdot \psi(p) + X(p) \cdot (it\psi(p)) = \\ X(p) \cdot (y + it) \cdot \psi(p) + (w(D\pi(v))(X(p))) \cdot \psi(p) = \\ X(p) \cdot \Big(\mu_* \circ \sigma^*(\widetilde{w \times \alpha})(p)(Y(p))(\psi(p))\Big) + \Big(w(D\pi(v))(X(p))\Big) \cdot \psi(p)\end{gathered}$$

Notice that

$$w(D\pi(v))(X(p)) = \lambda_*\Big(\underbrace{\tilde{w}(D\sigma(p)(Y(p)))}_{=\sigma^*(\tilde{w})(p)(Y(p))}\Big)(X(p))$$

Since $\lambda : Spin(n) \to SO(n) \subset \mathrm{Aut}(\mathbb{R}^n)$ gives a representation, the above equality means that $w(D\pi(v))(X(p))$ corresponds to the endomorphism term $\mathcal{A}_\sigma^Y$ in the expression of $\nabla_Y X$. Working out the derivative $D(X\psi)$ and rearranging the terms we get the result.

In order to prove $X < \psi_1, \psi_2 >=< \nabla_X^\alpha \psi_1, \psi_2 > + < \psi_1, \nabla_X^\alpha \psi_2 >$ we simply notice that the representation $\mu : Spin^c(n) \to \mathrm{Aut}(\mathfrak{m})$ is unitary. □

Now, analogously to what we did before, it is possible to define the Dirac operator for this spinor bundle. We shall use the Dirac operator when we deal with interacting fields.

Chapter 4

Classical field theory

We shall denote by $\mathcal{M}$ a pseudo-Riemannian oriented n-dimensional manifold. We shall denote the pseudo-metric by g and usually refers to it simply as the "metric".

4.1 Maxwell's electromagnetism

4.1.1 The Hodge $\star$ operator

We start recalling some operations on forms.

On each tangent space $T\mathcal{M}_x$ the map $v \mapsto g(v, \cdot)$ is an isomorphism (this is direct consequence of g being non-degenerated). We may use this isomorphism to define an inner product on $T^*\mathcal{M}_x$. Denote by $H : T\mathcal{M}_x \to T^*\mathcal{M}_x$ the isomorphism and put $< \mu, \nu >_x := < H^{-1}(\mu), H^{-1}(\nu) >_x$. We shall write simply $<,>$ instead of $<,>_x$.

Fist let us obtain an explicit expression for this inner product in local coordinates. Let $\{e_1, \ldots, e_n\}$ be a basis of local sections around x (for example $e_i = \frac{\partial}{\partial x^i}$). Let $\{e^1, \ldots, e^n\} = \{e_1^*, \ldots, e_n^*\}$ be the dual basis. Then

$$H(v)(u) = g(v, \cdot)(u) = g(v, u) = g_{\alpha\beta} v^\alpha u^\beta = g_{\alpha\beta} v^\alpha e^\beta(u)$$

So $g(v, \cdot) = g_{\alpha\beta} v^\alpha e^\beta$. Writing in matrix form:

$$v = (v_1, \ldots, v_n) \mapsto (v_1 \cdots v_n)[g] = [g] \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

where $[g]$ denotes the matrix of g on the chosen basis and in the last equality we use the symmetry of the metric. Notice that in this matrix notation it is clear that the the non-degenerescence implies an isomorphism.

If $\mu = \mu_\alpha e^\alpha, \nu = \nu_\beta e^\beta$ we have $H^{-1}(\mu) = g^{\alpha\beta}\mu_\beta e_\alpha$ and $H^{-1}(\nu) = g^{\alpha\beta}\nu_\beta e_\alpha$, where $g^{\alpha\beta}$ denotes the elements of $[g]^{-1}$. Then:

$$< \mu, \nu >=< H^{-1}(\mu), H^{-1}(\nu) >=< g^{\alpha\beta}\mu_\beta e_\alpha, g^{\gamma\delta}\nu_\delta e_\gamma >= g^{\alpha\beta}\mu_\beta g^{\gamma\delta}\nu_\delta \overbrace{< e_\alpha, e_\gamma >}^{=g_{\alpha\gamma}}$$
$$= g^{\alpha\beta} \underbrace{g^{\gamma\delta} g_{\alpha\gamma}}_{=g^\delta_\alpha} \mu_\beta\nu_\delta = \underbrace{g^{\alpha\beta} g^\delta_\alpha}_{=g^{\delta\beta}=g^{\beta\delta}} \mu_\beta\nu_\delta = g^{\beta\delta}\mu_\beta\nu_\delta$$

All in all, we have proven:

Proposition 34. *There is an isomorphism $T\mathcal{M}_x \approx T^*\mathcal{M}_x$ given by $v \mapsto g(v, \cdot)$. This isomorphism induces an inner product on $T^*\mathcal{M}_x$ given in local coordinates by $< \mu, \nu >= g^{\alpha\beta}\mu_\alpha\nu_\beta$.*

Now we extend the inner product to $\bigwedge^p(T^*\mathcal{M}_x)$.

Definition 51. *Let μ, ν be p-forms. Take now $\{e_1, \ldots, e_n\}$ to be an* orthonormal *basis of local sections, so that the corresponding basis of* 1*-forms is also orthonormal (notice that in this case generally we can not take e_i to be $\frac{\partial}{\partial x^i}$). Define $< \mu, \nu >_x$ as the unique inner product such that $\{e^{i_1} \wedge \cdots \wedge e^{i_p}\}_{i_1 < \cdots < i_p}$ is an orthonormal basis of $\bigwedge^p(T^*\mathcal{M}_x)$.*

It is easy to verify that this inner product is given on this basis explicitly by:

$$< e^{i_1} \wedge \cdots \wedge e^{i_p}, e^{j_1} \wedge \cdots \wedge e^{j_p} >= \det \begin{pmatrix} < e^{i_1}, e^{j_1} > & \cdots & < e^{i_1}, e^{j_p} > \\ \vdots & & \vdots \\ < e^{i_p}, e^{j_1} > & \cdots & < e^{i_p}, e^{j_p} > \end{pmatrix}$$

Indeed if $i_\ell = j_\ell, \ell = 1, \ldots, p$ then $\det(\cdot) = \pm 1$. If there is at least on ℓ such that $i_\ell \neq j_k, k = 1, \ldots, p$ then $< e^{i_\ell}, e^{j_k} >= 0$ and then $\det(\cdot) = 0$.

It should be noticed that if we express the the p-forms in terms of 1-forms which are not necessarily orthonormal then we still have a similar formula for their inner product, although the "cross terms" will not vanish. So, in the basis $\{dx^{i_1} \wedge \cdots \wedge dx^{i_p}\}_{i_1 < \cdots < i_p}$ we have

$$< dx^{i_1} \wedge \cdots \wedge dx^{i_p}, dx^{j_1} \wedge \cdots \wedge dx^{j_p} >= \det \begin{pmatrix} < dx^{i_1}, dx^{j_1} > & \cdots & < dx^{i_1}, dx^{j_p} > \\ \vdots & & \vdots \\ < dx^{i_p}, dx^{j_1} > & \cdots & < dx^{i_p}, dx^{j_p} > \end{pmatrix}$$

Now that we define an inner product on each $\bigwedge^p(T^*\mathcal{M}_x)$ we can define a global inner product, i.e., an inner product on $\bigwedge^p(T^*\mathcal{M})$:

Definition 52. *Give p-forms $\mu, \nu \in \bigwedge^p(T^*\mathcal{M})$ we define their inner product as*

$$< \mu, \nu >:= \int_{\mathcal{M}} < \mu, \nu >_x vol(x)$$

provided this integral converges. Here vol is the natural volume form induced by the metric.

Remark 20. *The space of forms such that $< \mu, \nu > < \infty$ is not a Hilbert space because such space is not complete on the metric given by $<,>$; it is necessary to take its completion.*

Now we can define:

Definition 53. *The **Hodge $\star$ operator** is defined as the the (unique) linear operator $\star : \bigwedge^p(T^*\mathcal{M}) \to \bigwedge^{n-p}(T^*\mathcal{M})$ such that $\mu \wedge \star\nu =< \mu, \nu >_x vol$.*

It follows that

Proposition 35. *$\star$ gives an isomorphism $\bigwedge^p(T^*\mathcal{M}) \approx \bigwedge^{n-p}(T^*\mathcal{M})$.*

Some important properties which steam from $\star$ are the following. Notice that we can define a second order differential operator $\bigwedge^p(T^*\mathcal{M}) \to \bigwedge^p(T^*\mathcal{M})$ given by

$$\begin{aligned} w \in \bigwedge{}^p(T^*\mathcal{M}) \mapsto dw \in \bigwedge{}^{p+1}(T^*\mathcal{M}) \mapsto \star dw \in \bigwedge{}^{n-p-1}(T^*\mathcal{M}) \\ \mapsto d \star dw \in \bigwedge{}^{n-p}(T^*\mathcal{M})(\mathcal{M}) \mapsto \star d \star dw \in \bigwedge{}^p(T^*\mathcal{M}) \end{aligned}$$

Definition 54. *The operator $\Delta : \bigwedge^p(T^*\mathcal{M}) \to \bigwedge^p(T^*\mathcal{M})$, $\Delta := \star d \star d$ is called **Hodge Laplacian**. A form such that $\Delta w = 0$ is called **harmonic**.*

It is possible to show:

Theorem 20. *There exists an unique harmonic form in each cohomology class, i.e,* $H^k(\mathcal{M}) \approx \ker(\Delta_k)$.

Notice that Δ gives rise to a partial differential equation. The theorem then states that the space of solutions of this partial differential equation is a topological invariant.

4.1.2 Maxwell's equations

Let us take now $\mathcal{M}$ as the Minkovskian space-time $\mathbb{R}^{1,3}$. Let us consider the electric and magnetic fields $(t, \vec{x}) \mapsto E(t, \vec{x}) \in \mathbb{R}^3$ and $(t, \vec{x}) \mapsto B(t, \vec{x}) \in \mathbb{R}^3$. It is well known that actually B is not a vector because it is unchanged through $\vec{x} \mapsto -\vec{x}$. But it can be written as a 2-form. So we have:

$$E = (E_x, E_y, E_z) = E_x dx + E_y dy + E_z dz$$
$$B = (B_x, B_y, B_z) = B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy$$

We define $F := B + E \wedge dt$ — which is known in physics literature as the "energy-momentum tensor". Differentiating:

$$\begin{gathered} dF = dB + dE \wedge dt = d(B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy) + \\ d(E_x dx + E_y dy + E_z dz) \wedge dt = (\partial_\mu B_x dx^\mu) \wedge dy \wedge dz + (\partial_\mu B_y dx^\mu) \wedge dz \wedge dx \\ +(\partial_\mu B_z dx^\mu) \wedge dx \wedge dy + [(\partial_\mu E_x dx^\mu) \wedge dx + (\partial_\mu E_y dx^\mu) \wedge dy \\ +(\partial_\mu E_z dx^\mu) \wedge dz] \wedge dt = \partial_\nu B_\nu dx \wedge dy \wedge dz + (\partial_\nu E_\alpha dx^\nu \wedge dx^\alpha) \wedge dt \\ +[\partial_t B_x dy \wedge dz + \partial_t B_y dz \wedge dx + \partial_t B_z dx \wedge dy] \wedge dt \end{gathered}$$

where the indices ν, α run only through spacial coordinates, i.e, $\nu, \alpha = 1, 2, 3$. If we write d_S for the exterior derivative acting only on spatial coordinates we have $dF = d_S B + (\partial_t B + d_S E) \wedge dt$. Since the first term does not involve dt, $dF = 0$ if and only if $d_S B = 0$ and $\partial_t B + d_S E = 0$ separately — what is equivalent to:

$$\nabla \cdot B = 0 \text{ and } \nabla \times E + \partial_t B = 0$$

which is the first pair of Maxwell's equations. Therefore these two equations can be written as $dF = 0$.

Now define $J = \rho dt + j_x dx + j_y dy + j_z dz$, where ρ is the charge distribution and $j = (j_x, j_y, j_z)$ the current density. A calculation similar to that above shows that the second pair of Maxwell's equations

$$\nabla \cdot E = \rho \text{ and } \nabla \times B - \partial_t E = j$$

can be written as $\star_s d_S \star_s E = \rho$ and $\star_s d_S \star_s B - \partial_t E = j$, where $\star_s$ denotes the Hodge star operator acting only on spatial coordinates. This can be written briefly as $\star d \star F = J$. Summarizing, the **Maxwell's equations** are:

$$dF = 0 \text{ and } \star d \star F = J$$

Now we want to study the symmetries of Maxwell's equations. In other words, we want to find out what maps $\phi : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$ preserve the equations.

Theorem 21. *A diffeomorphism $\phi : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$ preserves the Maxwell's equations if it is an isometry of the Lorentz metric $g = \mathrm{diag}[1, -1, -1, -1]$.*

Proof: For the first equation $dF = 0$ this is trivial because if $F' = \phi^* F$ then $dF' = d\phi^* F = \phi^*(dF) = \phi^*(0) = 0$. For the second pair of equations we shall need that ϕ^* commutes with $\star$. Since $\star$ depends on the metric we have $\phi^* \star = \star \phi^*$ for isometries of the Lorentz metric. We shall show this explicitly.

Recall that ([8] p. 81) $\star(dx^{i_1} \wedge \cdots \wedge dx^{i_p}) = \pm dx^{i_{p+1}} \wedge \cdots dx^{i_n}$, where $\{i_{p+1}, \ldots, i_n\}$ denotes the set of indices complementary to $\{i_1, \ldots, i_p\}$, i.e., $\{i_{p+1}, \ldots, i_n\} = \{1, 2, \ldots, n\} - \{i_1, \ldots, i_p\}$. The signal ± 1 is given by $\mathrm{sign}(i_1, \ldots, i_n)\epsilon(i_1) \cdots \epsilon(i_n)$, where $\mathrm{sign}(i_1, \ldots, i_n)$ is the signal of the permutation $(1, \ldots, n) \mapsto (i_1, \ldots, i_n)$ and $\epsilon(i) =< dx^i, dx^i >$. So we compute:

$$\star dx^0 \wedge dx^1 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix} \epsilon(0)\epsilon(1) dx^2 \wedge dx^3 = -dx^2 \wedge dx^3$$

$$\star dx^0 \wedge dx^2 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 1 & 3 \end{pmatrix} \epsilon(0)\epsilon(2) dx^1 \wedge dx^3 = -dx^1 \wedge dx^3$$

$$\star dx^0 \wedge dx^3 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 3 & 1 & 2 \end{pmatrix} \epsilon(0)\epsilon(3) dx^1 \wedge dx^2 = -dx^1 \wedge dx^2$$

$$\star dx^1 \wedge dx^2 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 0 & 3 \end{pmatrix} \epsilon(1)\epsilon(2) dx^0 \wedge dx^3 = dx^0 \wedge dx^3$$

$$\star dx^1 \wedge dx^3 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 0 & 2 \end{pmatrix} \epsilon(1)\epsilon(3) dx^0 \wedge dx^2 = dx^0 \wedge dx^2$$

$$\star dx^2 \wedge dx^3 = \mathrm{sign}\begin{pmatrix} 0 & 1 & 2 & 3 \\ 2 & 3 & 0 & 1 \end{pmatrix} \epsilon(2)\epsilon(3) dx^0 \wedge dx^1 = dx^0 \wedge dx^1$$

(notice that for the last three equalities we could have used that for any metric of signature $(s, n-s)$ holds $\star^2 = (-1)^{p(n-p)+s}$, [8] p.91). Therefore we have $\star dx^\mu \wedge dx^\nu = \pm dx^\sigma \wedge dx^\tau$ where σ, τ are complementary indices to μ, ν and $\mu < \nu, \sigma < \tau$. Given $\phi : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$ we have:

$$\phi^*(\star dx^\mu \wedge dx^\nu) = \phi^*(dx^\sigma \wedge dx^\tau) = \pm\phi^*(dx^\sigma) \wedge \phi^*(dx^\tau) = \\ \pm d\phi^*(x^\sigma) \wedge d\phi^*(x^\tau) = \pm d\phi^\sigma \wedge d\phi^\tau$$

If ϕ is an isometry then $< e_i, e_j >_x = < d\phi(x)(e_i), d\phi(x)(e_j) >_{\phi(x)}$ which implies $< dx^\mu, dx^\nu > = < d\phi^\mu, d\phi^\nu >$. Since the signal of the permutations which exchanges μ, ν and σ, τ is the same for dx^i and $d\phi^i$ we have:

$$\pm d\phi^\sigma \wedge d\phi^\tau = \star d\phi^\mu \wedge d\phi^\nu = \star\phi^*(dx^\mu \wedge dx^\nu)$$

It follows that if $F' = \phi^* F, J' = \phi^* J$ then

$$\star d \star F' = \star d \star \phi^* F = \phi^*(\star d \star F) = \phi^* J = J'$$

□

We have more than the above result. First, it is trivial to show invariance under translations (the exterior derivative "kills" the translated elements), so we actually have invariance of Maxwell's equations under the Poincaré group (which is the semi-direct product of the Lorentz group with translations; see chapter 6 for more details)[1]. It also can be shown that diffeomorphisms that change the metric by a positive scalar function also leave the Maxwell's equations invariant. These form the **conformal group** — this was the first step towards the discovery of gauge theories.

4.2 Fields and Lagrangians

The precise definition of what a field is depends on the particular problem considered. Generally a field is a C^∞ section of some fiber bundle $\pi : E \to \mathcal{M}$.

[1]"The fact that Maxwell's equations are not invariant under the Galilean group connecting inertial frames was one the major aspects of the crises that erupted in fundamental classical physics towards the end of the nineteenth century. Despite many contributions from Lorentz, Poincaré, and others, the situation remained murky till Einstein clarify the situation completely" [14] p. 8

For instance, the field may be a map $\phi : \mathcal{M} \to X$ where X is some manifold. In this case fields are sections of the trivial bundle $\mathcal{M} \times X$ (this is the case of the so-called σ-models [15] p. 155). In a gauge theory fields are connections $\mathcal{A}$ with gauge group a given Lie group G. If $\pi : P \to \mathcal{M}$ is a G-principal bundle then $\mathcal{A}$ is a section of some associated vector bundle. We shall consider this situation later when we study the Yang-Mills theory. If $\mathcal{M}$ is a spin manifold (i.e., a manifold which admits a spin structure) then we may consider spinor fields: sections of a spinor bundle over $\mathcal{M}$ (usually tensored with some vector bundle). In the case of gravitational field, fields are metrics over $\mathcal{M}$ — which are a subset of sections of $(0, 2)$ symmetric tensors. A theory is supposed to posses many fields and in this case we consider the fiber product $E = \tilde{\times} E_i$. Roughly, each E_i is decomposed in an extrinsic and an intrinsic part. This last one is associated to the frame bundle of $\mathcal{M}$ through a representation of $GL(n)$ or $Spin(n)$ and the representation determines the *kind of field*. For example, the scalar field is associated to the trivial representation ([15] p. 155, [16] p. 205).

Let us consider the case $\mathcal{M} = \mathbb{R}^{1,3}$ and a scalar field $\phi : \mathbb{R}^{1,3} \to \mathbb{R}$ (or $\mathbb{C}$), which can be viewed as a section of the trivial bundle. Let us describe the main idea in the passage from particle mechanics to fields, i.e., the passage from $q_j(t)$ to $\phi(x) = \phi(t, \vec{x})$. A field ϕ such as the electromagnetic field is a quantity which can be measured at any spatial point and any instant of time; in other words it can be measured at any space-time point. Analogously, $q_j(t)$ is a quantity which can be measured at any time t and for any index j. In the passage particle $\to$ field, ϕ corresponds to q and $\vec{x}$ corresponds to j; the time t remains as before. Sums over j become integrals over $\vec{x}$. This is the reason why field theories are sometimes described as mechanics with an infinite number of degrees of freedom ([16],196ff). Of great importance are the free field theories.

Definition 55. *A field theory is said to be a* ***free field theory****, and consequently its fields are called* ***free fields*** *if the space of fields is linear. A field which is not free is said* ***interacting***[2].

For theories defined by Lagrangians this restricts the Lagrangians to be quadratic on the field; so the Euler-Lagrange equations will be linear. The whole theory is made for free fields. Interacting fields are treated perturbatively.

[2] Actually there are some more technical conditions on the definition of a free field, see [17] p.18

Since we are interested in relativistic theories we must construct them so that the Lorentz group acts on the solution space of the equations of motion of fields and particles[3]. There is a physical reasoning to consider this: inertial observers in relative movement compute different coordinates x^μ and $x^{\mu\prime}$ for the same event. The passage from a coordinate system to another is made by the Poincaré group. Since a "physical fact" (e.g: A causes B) must be true (or false) in both coordinate systems, we demand that invariance of motion under the Poincaré group and that solutions be taken on solution by this action. The simples manner of doing this is to construct actions which are invariant under the Poincaré group. As a consequence, by the Noether theorem, our theory will have conserved quantities; they are the energy, the momentum and the angular momentum. Notice that this action of the Poncaré group on the space of solutions allows us to interpret each orbit as an unique physical solution seen from the coordinate systems of all possible inertial observers ([16] p. 203).

The discussion of the last paragraph justify, at least partially, to consider the action:

$$S(\phi) = \frac{1}{2}\int dx(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2)$$

which defines a free scalar field. As always: $\partial_\mu\phi\partial^\mu\phi = g^{\mu\nu}\partial_\mu\partial_\nu\phi$ It is not difficult to see that this action is invariant under the Poincaré group.

In order to deduce the Euler-Lagrange equation first we shall remember some terminology and notation from particle mechanics.

If $q : [0,1] \to \mathbb{R}^{1,3}$ is a path with $q(0) = q_0$ and $q(1) = q_1$ and $f : [0,1] \to \mathbb{R}^{1,3}$ is a function such that $f(0) = f(1) = 0$, then we can deform $q(t)$ considering $q_s(t) := q(t) + sf(t)$, where s is a small real number. The functions are assumed to be differentiable enough. f is called the **variation of** q (since $f(t) = \frac{q_s(t)-q(t)}{s}$) and it is usually denoted by δq. Notice that

$$\frac{dq_s(t)}{ds}|_{s=0} = \lim_{s\to 0}\frac{q_s(t) - \overbrace{q_0(t)}^{=q(t)}}{s} = \lim_{s\to 0}\frac{q(t) + sf(t) - q(t)}{s} = f(t) \equiv \delta q(t)$$

More generally, if G is a function depending on paths jointing q_0 and q_1 we define:

$$\delta G := \frac{dG(q_s(t))}{ds}|_{s=0}$$

[3]We shall see that the group is not the Lorentz group but the universal covering of the restricted Poincaré group, but here we can think of the Lorentz group

With this we can compute the variation of the action $S = \int_{t_0}^{t_1} L(q,\dot{q})dt$. Recall that L is a real function on the tangent bundle, $L : T\mathcal{M} \to \mathbb{R}$, where $\mathcal{M}$ is the configuration space; in our case $\mathcal{M} = \mathbb{R}^{1,3}$ and $T\mathcal{M} = \mathbb{R}^{1,3} \times \mathbb{R}^{1,3}$. We notice that physicists usually write $L(q,\dot{q})$ for both the function $L : TM \to \mathbb{R}$ and the composition $L \circ \tilde{q}$, where $\tilde{q} : \mathbb{R} \to T\mathcal{M}$ is the map naturally associated with a path $q : \mathbb{R} \to \mathcal{M}, t \mapsto q(t)$, i.e., $\tilde{q} : t \mapsto (q(t), \frac{dq(t)}{dt})$. In the first case $(q,\dot{q})$ denotes coordinates on the tangent bundle and $\dot{q}$ *is not* a tangent vector to a path (because there is no path involved). In the second case $(q,\dot{q})$ denotes the $2n$-tuple associated with a curve $q(t)$ under $\tilde{q}$ ([18] p. 170).

From $\delta S = 0$ we obtain the Euler-Lagrange equations. The deduction is well known, but we want to stress a specific point; denote $(q,\dot{q}) = (q_1, \ldots, q_n, \dot{q_1}, \ldots, \dot{q_n})$,

$$\begin{aligned}\delta L &= \frac{d}{ds} L(q_s, \dot{q}_s)|_{s=0} = DL_{(q_s,\dot{q}_s)}(\frac{dq_s}{ds}, \frac{d\dot{q}_s}{ds})|_{s=0} \\ &= \frac{\partial L}{\partial q_1}|_{q_0=q(t)} \underbrace{\frac{dq_{s1}}{ds}|_{s=0}}_{=\delta q_1} + \cdots + \frac{\partial L}{\partial \dot{q}_n}|_{q_0=q(t)} \underbrace{\frac{d\dot{q}_{s1}}{ds}|_{s=0}}_{=\delta \dot{q}_n}\end{aligned}$$

So $\delta L = \frac{\partial L}{\partial q_1}\delta q_1 + \cdots + \frac{\partial L}{\partial \dot{q}_n}\delta \dot{q}_n$. The fundamental fact here is that L is a real function defined on a *finite-dimensional* manifold, what allow us to calculate the partial derivatives.

Let us return to the action

$$S(\phi) = \frac{1}{2}\int dx(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2) = \int dx^0 \int d\vec{x}\frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2)$$

Since the action is the time integral of the Lagrangian we get

$$L = L(\phi) = \int d\vec{x}\frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2) = \int d\vec{x}\mathcal{L}(\phi, \partial_\mu\phi)$$

where $\mathcal{L}(\phi, \partial_\mu\phi)$ is called (for obvious reasons) the **Lagrangian density**. In field theory Lagrangians are typically spatial integrals of such densities. Let us remark the striking difference to the particle mechanics case: for each $\phi : \mathbb{R}^{1,3} \to \mathbb{R}$ and each fixed t the Lagrangian evaluated at ϕ yields a number; such as before the Lagrangian evaluated at $(q,\dot{q}) \in T\mathcal{M}$ yielded a real value. What means: if before we had $L : T\mathcal{M} \to \mathbb{R}$ now L is a function(al) defined *on the space of fields and its derivatives* (i.e., on the space of maps $\phi : \mathbb{R}^{1,3} \to \mathbb{R}$). In other words, the configuration space of field theory is the space of fields — which clearly is an *infinite dimensional space* ([16] p. 199). Now in order to get the analogous of $\delta L = \frac{\partial L}{\partial q_1}\delta q_1 + \cdots + \frac{\partial L}{\partial \dot{q}_n}\delta \dot{q}_n$ we need to derivate in this infinite-dimensional space, so we use the so-called Fréchet derivative of a functional. Recall that this is defined in the usual way of derivative: the derivative of a functional F at ϕ is obtained linearizing F: $F(\phi+\delta\phi) = F(\phi) + F_\phi(\delta\phi) + O(\delta\phi^2)$ and the linear map F_ϕ is defined as the derivative.

Physicist, however, would rather write F_ϕ in terms of its integral kernel ([16]p.198):

$$F_\phi(\delta\phi) = \int dx \delta\phi(x) \frac{\delta F}{\delta\phi(x)}$$

The "function" $\frac{\delta F}{\delta\phi(x)}$ in general does not exists, being rather a distribution (see chapter 5). An alternative definition of $\frac{\delta F}{\delta\phi(x)}$ is ([16] p. 198, [19] p. 25):

$$\frac{\delta F}{\delta\phi(y)} := \lim_{\epsilon\to 0}\big(F[\phi(x) - \epsilon\delta(x-y)] - F[\phi(x)]\big)$$

So, applying such techniques to S we obtain the Euler-Lagrange equations for the field ϕ ([19] p.26 [16] 196):

$$\frac{\delta S}{\delta\phi(x)} = 0 \Rightarrow \frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi} = 0$$

Of course, the question of the differentiability of S and $\mathcal{L}$ is not a trivial question. The derivatives above are not expected to exist in general. So, this is a problem that must be investigated for each particular Lagrangian. Another point of view is to treat all the derivatives as formal expressions, the partial derivatives appearing in the above formula being formal derivatives of $\mathcal{L}$ thought of as a function of the independent variables ϕ and $\partial_\mu\phi$. Specifically, for the Lagrangian density $\frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2)$ the Euler-Lagrange equations yield the so-called **Klein-Gordon equation**:

$$(\Box + m^2)\phi = 0$$

where $\Box = \frac{\partial}{\partial t} - \triangle$ is the D'Alambertian.

4.3 Gauge theory and Yang-Mills equations

The basic data of a gauge theory is a principal bundle $\pi : P \to \mathcal{M}$; usually it is assumed that G is compact. In physical applications $\mathcal{M}$ plays the role of space-time and G is called the **gauge group**. The basic dynamical variable is a connection on P called the **gauge field**. Recalling the results of chapter 1, mainly corollary 2, we can consider the gauge field as a connection on an associated End(E)-bundle. A local trivialization of P or $End(E)$ is called a **gauge**.

We return to the setting of section 1.2.1, i.e., $\pi : E \to \mathcal{M}$ is a vector bundle which is thought of as an associated bundle of a principal bundle $\pi : P \to \mathcal{M}$. The connection on $\pi : P \to \mathcal{M}$ induces a connection on $\pi : E \to \mathcal{M}$ and this gives rise to the curvature $\mathcal{F} \in \Gamma(\mathrm{End}(E) \otimes \bigwedge^2(T^*\mathcal{M}))$. We also consider a (pseudo-)Riemannian metric on $\mathcal{M}$.

4.3.1 Gauge transformations and gauge equivalence

We still denote by $\pi : P \to \mathcal{M}$ a principal bundle with structure group G and $\pi : E \to \mathcal{M}$ an associated vector bundle given by some representation ρ of G.

Definition 56. *An* ***automorphism*** *of P is a diffeomorphism $\phi : P \to P$ which is G-equivariant and which such that $\pi \circ \phi = \pi$.*

Clearly the set of automorphisms of P forms a group under composition. This group acts on the left on P commuting with the right action of G.

Definition 57. *The group of automorphisms of P is denoted by* Aut(P) *and is called the* ***group of gauge transformations*** *of P. Therefore and element of* Aut(P) *is also called a* ***gauge transformation****.*

There is some confusion in the terminology and and Aut(P) is sometimes called the *gauge group*, but we should keep this name for the group G only.

Definition 58. *An* ***automorphism*** *of E is a fiberwise linear isomorphism $\phi : E \to E$ which is G-equivariant and which such that $\pi \circ \phi = \pi$.*

Again, the set of automorphisms of E forms a group under composition. This group acts on the left on E.

Definition 59. *The group of automorphisms of E is denoted by* Aut(E) *and is called the* ***group of gauge transformations*** *of E. Therefore and element of* Aut(E) *is also called a* ***gauge transformation****.*

Here, we shall deal more with gauge theory on the associated vector bundle $\pi : E \to \mathcal{M}$, but the notions on $\pi : P \to \mathcal{M}$ are similar.

Definition 60. *Two connections ∇ and ∇' on E are said to be* ***gauge equivalent*** *if $\nabla = g^{-1}\nabla' g$ for some gauge transformation $g \in$* Aut(E)*.*

Notice that if ∇ and ∇' are E gauge equivalent then for any pair of vector field X and Y on $\mathcal{M}$ and for any section σ of E we have:

$$\mathcal{F}(X,Y)'\sigma = \nabla'_X\nabla'_Y\sigma - \nabla'_Y\nabla'_X\sigma - \nabla'_[X,Y]\sigma$$
$$= g^{-1}\nabla_X gg^{-1}\nabla_Y g\sigma - g^{-1}\nabla_Y gg^{-1}\nabla_X g\sigma - g^{-1}\nabla_[X,Y]g\sigma =$$
$$= g^{-1}\mathcal{F}(X,Y)g\sigma$$

or $\mathcal{F}' = g^{-1}\mathcal{F}g$ for brief. This means that locally $\mathcal{F}'_{\mu\nu} = g^{-1}\mathcal{F}_{\mu\nu}g$.

4.3.2 Yang-Mills equations

We want to define the Yang-Mills action and the then deduce from it the Yang-Mills equations. Our basic dynamical variable will be a connection on E. Actions are usually given in term of integral of Lagrangians, so we need to produce an n form to integrate.

Given a vector bundle $\pi : E \to \mathcal{M}$ consider the bundle of endomorphisms of E: End(E). Suppose also that there is a metric. Let $\alpha = T^\mu \otimes \alpha_\mu$ be an End(E)-valued form. We can define the usual operations.

Definition 61. ***Hodge star operator:*** $\alpha = T^\mu \otimes \alpha_\mu$

Recall that the isomorphism End$(F) \approx F \otimes F^*$ allow us to define a **trace** Tr : End$(F) \to \mathbb{R}$ given by $v \otimes f \mapsto f(v)$. This implies that if we have a section T of End(E) we can define a *function* $Tr(T)$ on $\mathcal{M}$ whose value at $x \in \mathcal{M}$ is given by $Tr(T)(x) = Tr(T(x))$. Now if we have an End(E)-valued form $\alpha = T^\mu \otimes \alpha_\mu$ we define:

Definition 62. ***trace:*** $Tr(\alpha) = Tr(T^\mu)\alpha_\mu$.

Now we have the ingredients to define the **Yang-Mills Lagrangian**

$$\mathcal{L}_{YM}(\nabla) := \frac{1}{2}Tr(\mathcal{F}_\nabla \wedge \star\mathcal{F}_\nabla) = \frac{1}{4}Tr(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu})$$

where the last equality is the expression of $\mathcal{L}$ in local coordinates. Since $\mathcal{F}$ depends on ∇ we write $\mathcal{L}_{YM}(\nabla)$ and $\mathcal{F}_\nabla$ to stress this dependence. Of course, $\mathcal{L}_{YM}(\nabla)$ also depends on the metric, but we consider the metric fixed here. Now we define the **Yang-Mills action** simply as

$$S_{YM}(\nabla) := \int_{\mathcal{M}} \mathcal{L}_{YM}(\nabla) = \frac{1}{2}\int_{\mathcal{M}} Tr(\mathcal{F}_\nabla \wedge \star\mathcal{F}_\nabla) = \frac{1}{4}\int_{\mathcal{M}} Tr(\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu})$$

whenever this integral makes sense. One way of guaranteeing this is to take $\mathcal{M}$ compact or to use forms with compact support.

Now we want to minimize the Yang-Mills action with respect to variations on the connection. This will lead us to the Yang-Mills equations. But before that, the first thing to do is to ask about the symmetries of $\mathcal{L}_{YM}(\nabla)$. It is a direct computation to see that if ∇ and ∇' are gauge equivalent then $\mathcal{L}_{YM}(\nabla) = \mathcal{L}_{YM}(\nabla')$. Therefore it is interesting to think of the Yang-Mills equations *modulo gauge equivalence*. This also tells us that the Yang-Mills equations will be invariant under gauge transformations.

Recall that the space of connection is an affine space. So, if we fix a connection ∇^0 we can obtain any other connection just adding an End(E)-valued one form $\mathcal{A}$: $\nabla^{\mathcal{A}} = \nabla^0 + \mathcal{A}$. So we may vary the connections just varying the $\mathcal{A}$. We then have a one-parameter family $\mathcal{A}_s$ of End(E)-valued one-forms $\mathcal{A}_s = \mathcal{A} + s\delta\mathcal{A}$, where s is a real parameter and $\delta\mathcal{A}$ is an End(E)-valued one-form. For any function f of $\mathcal{A}$ we define its variation δf with respect to $\mathcal{A}$ as $\delta f = \frac{d}{ds}f(\mathcal{A}_s)|_{s=0}$. $\delta f = 0$ means that $\frac{d}{ds}f(\mathcal{A}_s)|_{s=0} = 0$ for all variations $\delta\mathcal{A}$.

Remark 21. *In physical literature $\mathcal{A}$ is known as* ***vector potential****, so a physicist would talk about "varying the (vector) potential"*

In order to investigate how $\mathcal{F}$ varies with $\mathcal{A}$ it is worthwhile to find a expression for $\mathcal{F}$ which allows us to make computations. We have seen that in a local trivialization σ_α we have $\mathcal{F} = d\mathcal{A}_\alpha + \mathcal{A}_\alpha \wedge \mathcal{A}_\alpha$. If our bundle is trivial, it follows that this expression holds globally and then we can use it to make our computations. For a non-trivial bundle we would like to have a similar expressions. In order to do it, we recall remark 12. Denoting by $\mathcal{F}_0$ the curvature of ∇^0 and by $\mathcal{F}_\mathcal{A}$ that of $\nabla^\mathcal{A}$ we have for any E-valued form α:

$$D^{\nabla^0} D^{\nabla^0}\alpha = \mathcal{F}_0 \wedge \alpha, \quad D^{\nabla^\mathcal{A}} D^{\nabla^\mathcal{A}}\alpha = \mathcal{F}_\mathcal{A} \wedge \alpha$$

On the other hand, since $D^{\nabla^\mathcal{A}}\alpha = D^{\nabla^0}\alpha + \mathcal{A} \wedge \alpha$ we get:

$$\begin{aligned} D^{\nabla^\mathcal{A}} D^{\nabla^\mathcal{A}}\alpha &= D^{\nabla^\mathcal{A}}(D^{\nabla^0}\alpha + \mathcal{A} \wedge \alpha) \\ D^{\nabla^0}(D^{\nabla^0}\alpha + \mathcal{A} &= \wedge\alpha) + \mathcal{A} \wedge (D^{\nabla^0}\alpha + \mathcal{A} \wedge \alpha) \\ &= \mathcal{F}_0 \wedge \alpha + D^{\nabla^0}\mathcal{A} \wedge \alpha + \mathcal{A} \wedge \mathcal{A} \wedge \alpha \end{aligned}$$

So $\mathcal{F} = \mathcal{F}_0 + D^{\nabla^0}\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$. Now we compute:

$$\begin{aligned} \delta\mathcal{F} &= \frac{d}{ds}(\mathcal{F}_0 + D^{\nabla^0}\mathcal{A}_s + \mathcal{A}_s \wedge \mathcal{A}_s)|_{s=0} \\ &= \big(D^{\nabla^0}(\frac{d}{ds}\mathcal{A}_s) + (\frac{d}{ds}\mathcal{A}_s) \wedge \mathcal{A} + \mathcal{A} \wedge (\frac{d}{ds}\mathcal{A}_s)\big)|_{s=0} \\ &= D^{\nabla^0}\delta\mathcal{A} + \delta\mathcal{A} \wedge \mathcal{A} + \mathcal{A} \wedge \delta\mathcal{A} \end{aligned}$$

A calculation similar to that performed in lemma 4 shows that for any End(E)-valued form α we have $D^{\nabla^\mathcal{A}}\alpha = D^{\nabla^0}\alpha + \{\mathcal{A}, \alpha\}$; moreover, this is defined globally. Therefore

$$\delta\mathcal{F} = D^{\nabla^0}\delta\mathcal{A} + \{\mathcal{A}, \delta\mathcal{A}\} = D^{\nabla^\mathcal{A}}\delta\mathcal{A}$$

Now we state without proving some properties of the trace.

Proposition 36. *For any* End(E)*-valued p-form α and any* End(E)*-valued q-form β: (i) $Tr(\alpha \wedge \beta) = (-1)^{pq} Tr(\beta \wedge \alpha)$; (ii) $Tr(\{\alpha, \beta\}) = 0$; (iii) $Tr(D^\nabla \alpha) = dTr(\alpha)$; (iv) if $p + q = \dim(\mathcal{M}) - 1$ then*

$$\int_\mathcal{M} Tr(D^\nabla \alpha \wedge \beta) = (-1)^{p+1} \int_\mathcal{M} Tr(\alpha \wedge D^\nabla \beta)$$

(v) if $p + q = \dim(\mathcal{M})$ and we have a pseudo-Riemannian metric on $\mathcal{M}$:

$$\int_\mathcal{M} Tr(\alpha \wedge \star\beta) = \int_\mathcal{M} Tr(\beta \wedge \star\alpha)$$

We use such identities to compute:

$$\delta S_{YM} = \delta \int_{\mathcal{M}} \mathcal{L}_{YM}(\nabla^{\mathcal{A}}) = \frac{1}{2}\delta \int_{\mathcal{M}} Tr(\mathcal{F}_{\mathcal{A}} \wedge \star\mathcal{F}_{\mathcal{A}})$$
$$= \int_{\mathcal{M}} Tr(\delta\mathcal{F}_{\mathcal{A}} \wedge \star\mathcal{F}_{\mathcal{A}} + \mathcal{F}_{\mathcal{A}} \wedge \star\delta\mathcal{F}_{\mathcal{A}})$$
$$= \int_{\mathcal{M}} Tr(\delta\mathcal{F}_{\mathcal{A}} \wedge \star\mathcal{F}_{\mathcal{A}} + \mathcal{F}_{\mathcal{A}} \wedge \star\delta\mathcal{F}_{\mathcal{A}})$$
$$= \int_{\mathcal{M}} Tr(\delta\mathcal{F}_{\mathcal{A}} \wedge \star\mathcal{F}_{\mathcal{A}}) = \int_{\mathcal{M}} Tr(D^{\nabla^{\mathcal{A}}}\delta\mathcal{A} \wedge \star\mathcal{F}_{\mathcal{A}})$$
$$= \int_{\mathcal{M}} Tr(\delta\mathcal{A} \wedge D^{\nabla^{\mathcal{A}}} \star \mathcal{F}_{\mathcal{A}})$$

Therefore $\delta S_{YM} = 0$, i.e., the above integral vanishes for arbitrary variations $\delta\mathcal{A}$ if and only if

$$D^{\nabla^{\mathcal{A}}} \star \mathcal{F}_{\mathcal{A}} = 0$$

This is the so-called **Yang-Mills equation**

Remark 22. *Usually we talk about the Yang-Mills equations — in the plural. Historically they are*

$$D^{\nabla}\mathcal{F} = 0 \text{ and } D^{\nabla} \star \mathcal{F} = 0$$

But we have seen that the first equation is actually an identity: the Bianchi identity. Therefore all the information is contained in the second equation.

The reader should now compare these equations with free-source Maxwell equations: they are "the same" replacing the curvature $\mathcal{F}$ by the energy-momentum tensor and the covariant exterior derivative D^{∇} by the usual exterior derivative d. This is not a coincidence: it can be show that Yang-Mills equations reduce to Maxwell's equations in case $\mathcal{M} = \mathbb{R}^{1,3}$, $G = U(1)$ and $P = \mathcal{M} \times G = \mathbb{R}^{1,3} \times U(1)$, with the associated bundle being the trivial line bundle $\mathbb{R}^{1,3} \times i\mathbb{R}$. In this case the connection which minimizes the action is the flat d (which is globally defined since our bundles are trivial). See, for example, [3, 8, 1].

Chapter 5

Distributions

5.1 Test functions and Schwartz functions

We shall denote $|\alpha| = \alpha + \cdots + \alpha_n$ and $D^\alpha = \frac{\partial}{\partial x_1^{\alpha_1}} \cdots \frac{\partial}{\partial x_n^{\alpha_n}}$.

Definition 63. *A* ***test function*** *is a real or complex valued C^∞ function with compact support on an $\mathbb{R}^N$. The space of test functions is denoted by $\mathcal{D}$.*

For example, the function defined by:

$$\phi(x) = \begin{cases} e^{\frac{1}{x^2-1}}, & |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

is a test function on $\mathbb{R}$. Clearly, $\mathcal{D}$ is a vector space.

Definition 64. *Let $\{\phi_n\}$ be a sequence in $\mathcal{D}$. We say that ϕ_n* ***converges*** *to $\phi \in \mathcal{D}$ in $\mathcal{D}$, and denote it by $\phi \xrightarrow{\mathcal{D}} \phi$ if: $\phi_1, \phi_2, \ldots$ and ϕ vanish out of a bounded set $S \subset \mathbb{R}^N$; (ii) $D^\alpha \phi_n \to D^\alpha \phi$ uniformly on $\mathbb{R}^N$ for every multi-index α ([20] p.286).*

Some properties of test functions are:

Proposition 37. *If $\phi, \psi \in \mathcal{D}$ then (i) $f\phi \in \mathcal{D}$ for ever C^∞ function f; (ii) $\phi(Ax) \in \mathcal{D}$ for every affine map $A : \mathbb{R}^N \to \mathbb{R}^N$ and (iii) $\phi * \psi \in \mathcal{D}$ (convolution).*

Proof: [20] p 286. □

Corollary 8. *If* $\phi_n \xrightarrow{\mathcal{D}} \phi$ *and* $\psi_n \xrightarrow{\mathcal{D}} \psi$ *then (i)* $a\phi_n + b\psi_n \xrightarrow{\mathcal{D}} a\phi + b\psi$*; (ii)* $f\phi_n \xrightarrow{\mathcal{D}} f\phi$ *for ever* C^∞ *function* f*; (iii)* $\phi_n \circ A \xrightarrow{\mathcal{D}} \phi \circ A$ *for every affine map* $A : \mathbb{R}^N \to \mathbb{R}^N$ *and (iv)* $D^\alpha \phi_n \xrightarrow{\mathcal{D}} D^\alpha \phi$ *for every multi-index* α*.*

Proof: [20] p 287. □

Definition 65. *Let us denote by* $\mathcal{S}$ *the space of real or complex valued* C^∞ *functions on* $\mathbb{R}^N$ *with the following property: the function and all its derivatives decrease fast than every power of* $\frac{1}{\|x\|}$ *when* $\| x \| \to \infty$*, i.e.*

$$\lim_{\|x\| \to \infty} |x^k D^\alpha \phi(x)| = 0$$

for every multi-index k *and* α*, where* x^k *means* $x_1^{k_1} \cdots x_n^{k_n}$*.* $\mathcal{S}$ *is called the* ***space of Schwartz*** *([21], p. 234). The functions belonging to* $\mathcal{S}$ *are called* C^∞ ***functions with fast decreasing****.*

It follows that $\mathcal{S}$ is a vector space.

Remark 23. *Notice that the above condition on* $\| x \| \to \infty$ *implies that* $\| x^k D^\alpha \phi(x) \| \leq C_{k\alpha} = C_{k_1 \cdots k_n \alpha_1 \cdots \alpha_n}$*,* $k_1, \ldots, k_n, \alpha_1, \ldots, \alpha_n = 0, 1, \ldots$ *([22] p. 16).*

Definition 66. *Let* $\{\phi_n\}$ *be a sequence in* $\mathcal{S}$*. We say that* ϕ_n ***converges*** *to* ϕ *in* $\mathcal{S}$*, and denote it by* $\phi_n \xrightarrow{\mathcal{S}} \phi$ *if: (i) all the derivatives of* ϕ_n *converge to the respective derivatives of* ϕ *and the convergence is uniform on each bounded domain of* $\mathbb{R}^N$ *and (ii) all the inequalities* $\| x^k D^\alpha \phi_n(x) \| \leq C_{k\alpha}$ *hold and the constants* $C_{k\alpha}$ *do not depend on* n *([22],p. 16).*

Notice that taking the limit $n \to \infty$ we see that the limit function ϕ satisfies $\| x^k D^\alpha \phi(x) \| \leq C_{k\alpha}$; then, making $\| x \| \to \infty$ we obtain that $D^\alpha \phi$ decrease fast than every power of x and therefore $\phi \in \mathcal{S}$ ([22],p. 16).

In other words, $\phi_n \xrightarrow{\mathcal{S}} 0$ means that

$$\lim_{n \to \infty} \sup_{x \in \mathbb{R}^N} |x^k D^\alpha \phi_n(x)| = 0 \tag{5.1}$$

for every multi-indices k and α.

Proposition 38. *$\mathcal{D}$ is dense on $\mathcal{S}$.*

Proof: Take a C^∞ function satisfying:

$$e_1(x) = \begin{cases} 1 & \text{if } |x_j| \leq 1, j = 1, \ldots, n \\ 0 & \text{if } |x_j| > 2, j = 1, \ldots, n \end{cases}$$

Then $e_1 \in \mathcal{D}$. Define $e_n(x) = e_1(\frac{x}{n}), n = 1, 2, \ldots$. If $\phi \in \mathcal{S}$ then $e_n\phi \in \mathcal{D}$ and $e_n\phi \overset{\mathcal{S}}{\to} \phi$. □

The existence of the function e_1 can be showed as follows: take a compact $K \subset \mathbb{R}^N$ and an open set $U \supset K$. Then there exists a test function $\phi \in \mathcal{D}$ with support on U such that $Im\phi \subset [0, 1]$ and $\phi(x) = 1$ for every $x \in K$. Indeed, define

$$\rho_\epsilon(x) = \begin{cases} \frac{k}{\epsilon^n} \exp(\frac{-\epsilon^2}{\epsilon^2 - \|x\|^2}) & \text{if } \| x \| < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

where k is such that $k \int_{\|x\| \leq 1} \exp(\frac{-\epsilon^2}{\epsilon^2 - \|x\|^2}) dx = 1$. Take a bounded set G such that $K \subset G \subset \bar{G} \subset U$. Then the function ϕ defined by $\phi(x) = \int \chi_G(y)\rho_\epsilon(y - x)dy$ satisfies the desired properties, where χ_G is the characteristic function of G. For more details see [23] p.29.

Remark 24. $\phi_n \overset{\mathcal{D}}{\to} \phi \Rightarrow \phi_n \overset{\mathcal{S}}{\to} \phi$ *([22] p.17).*

5.2 Distributions and tempered distributions

Definition 67. *A **distribution** T in $\mathbb{R}^N$ is a continuous linear functional on $\mathcal{D}$. It means: (i) $T(a\phi + b\psi) = aT(\phi) + bT(\psi)$, (ii) if $\phi_n \overset{\mathcal{D}}{\to} \phi$ then $T(\phi_n) \to T(\phi)$ in $\mathbb{C}$. The space of distributions is denoted by $\mathcal{D}'$ ([20],p.287).*

Definition 68. *A **tempered distribution** T in $\mathbb{R}^N$ is a continuous linear functional on $\mathcal{S}$. It means: (i) $T(a\phi + b\psi) = aT(\phi) + bT(\psi)$, (ii) if $\phi_n \overset{\mathcal{S}}{\to} \phi$ then $T(\phi_n) \to T(\phi)$ in $\mathbb{C}$. The space of tempered distributions is denoted by $\mathcal{S}'$ ([23],p.32).*

It follows that $\mathcal{S}' \subset \mathcal{D}'$.
Notation: Physicist usually use an "inner-product-like" notation $< T, \phi >$ for $T(\phi)$.

Notice that every locally integrable function f on $\mathbb{R}^n$ may be identified with a distribution T_f: $T_f(\phi) \equiv < T_f, \phi > = \int f\phi$. This motivates:

Definition 69. *We call a distribution T **regular** if there exists a locally integrable function f such that $T(\phi) = \int f\phi$. In this case sometimes it is written T_f for T. Otherwise the distribution is called **singular**. By abuse of notation one writes $f(\phi) \equiv < f, \phi >$ instead of $T(\phi)$.*

Example. If Ω is a measurable set on $\mathbb{R}^N$ then the functional T_Ω defined by $T_\Omega(\phi) = \int_\Omega \phi$ defines a distribution; moreover, it is regular since $T_\Omega(\phi) = \int_\Omega \phi = \int \chi_\Omega \phi$. In particular, if $\Omega = (0,\infty) \times \cdots \times (0,\infty)$ then $H := T_\Omega$ satisfies $H(\phi) = \int \chi_\Omega \phi = \int_0^\infty \cdots \int_0^\infty \phi(x)dx$. H is called **Heaviside function**. Usually one writes $\int H\phi$ meaning $\int \chi_\Omega \phi$.

Example. The **Dirac δ-"function"** is a distribution defined by $\delta(\phi) = \phi(0)$. We also define δ_a by $\delta_a(\phi) = \phi(a)$. Other notations are $\delta(0)\phi =< \delta(0), \phi >= \phi(0)$, $\delta(x-a)\phi =< \delta(x-a), \phi >= \phi(a)$.

Proposition 39. *δ is singular.*

Proof: Let us suppose that there exists a locally integrable function f such that for every test function ϕ it holds $\delta(\phi) = \phi(0) = \int f\phi$. The function:

$$\phi(x,a) = \begin{cases} \exp(-\frac{a^2}{a^2 - \|x\|^2}) & \text{if } \| x \| < a \\ 0 & \text{otherwise} \end{cases}$$

belongs to $\mathcal{D}$. Then we have $\int_{\mathbb{R}^n} f(x)\phi(x,a)dx = \phi(0,a) = e^{-1}$. Taking $a \to 0$ we have $\int_{\mathbb{R}^n} f(x)\phi(x,a)dx \to 0 \neq e^{-1} = \phi(0,a)$ ([23] p. 32). □

Remark 25. *Although δ is singular, we found in the literature expressions such as $\int \delta(x-a)\phi(x) = \phi(a)$. This is an abuse of notation: if $\delta_a \equiv \delta(x-a)$ were regular then there would be a locally integrable function, which we could also denoted by $\delta(x-a)$, such that $\int \delta(x-a)\phi(x) = \phi(a)$. Although such a function does not exist, physicists like to preserve the notation of regular distributions. They go even one step further and use the notation of inner product $< \delta, \phi >= \int \delta(x-a)\phi(x)$.*

Definition 70. *Given $T \in \mathcal{D}'$ we define its* ***derivative with respect to*** x_j *as the distribution* $\frac{\partial T}{\partial x_j}$ *such that* $\frac{\partial T}{\partial x_j}(\phi) = -T(\frac{\partial \phi}{\partial x_j})$*. Applying the derivative successively we get* $D^\alpha T(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi)$ *([20],288).*

The definition of derivative makes sense because:

Proposition 40. *If $T \in \mathcal{D}'$ then $D^\alpha T \in \mathcal{D}'$ for every multi-index α.*

Proof:

$$D^\alpha T(a\phi + b\psi) = T(a(-1)^{|\alpha|} D^\alpha \phi + b(-1)^{|\alpha|} D^\alpha \psi) =$$
$$aT((-1)^{|\alpha|} D^\alpha \phi) + bT((-1)^{|\alpha|} D^\alpha \psi) = aD^\alpha T(\phi) + bD^\alpha T(\psi)$$

Since $\phi_n \xrightarrow{\mathcal{D}} \phi$ implies $D^\alpha \phi_n \xrightarrow{\mathcal{D}} D^\alpha \phi$ we have that $D^\alpha T$ is continuous. □

Example. Take the Heaviside function H on $\mathbb{R}$. We have

$$\frac{dH}{dx}(\phi) = -H(\frac{d\phi}{dx}) = -\int_0^\infty \frac{d\phi}{dx} dx = \phi(0) = \delta(\phi)$$

i.e, the derivative of the Heaviside function is the Dirac delta function: $\frac{dH}{dx} = \delta$. It is easy to show $\frac{d^m \delta}{dx^m} = (-1)^{|\alpha|} D^\alpha \phi(0)$.

We want now to put a topology on $\mathcal{D}'$:

Definition 71. *A sequence $\{T_n\}$ of distributions is said to be* ***weakly convergent*** *to a distribution $T \in \mathcal{D}'$ if $T_n(\phi) \to T(\phi)$ for every test function $\phi \in \mathcal{D}$.*

Example. Consider distributions on $\mathbb{R}$. Define $f_n(x) = \frac{n}{\pi(1+n^n x^2)}$. If ϕ is a test function then $T_f(\phi) = \int f\phi \to \phi(0)$ when $n \to \infty$ ([20] p. 290). If we identify T_f with f we have that $f_n \to \delta$. But notice that as a function, f_n converges pointwise to 0 if $x \neq 0$ and to ∞ if $x = 0$. This is the exact sense of saying that de δ distribution is zero for $x \neq 0$ and "infinity" for $x = 0$: δ is the weak limit of a sequence of regular distributions such that the locally integrable functions defining such regular distributions are converging to zero for $x \neq 0$ and to infinity for $x = 0$. In particular, this example shows that the weak limit of regular distributions need not to be regular.

Proposition 41. *If $T_n \to T$ weakly then $D^\alpha T_n \to D^\alpha T$ weakly.*

Proof:

$$D^{\alpha}T(\phi) = (-1)^{|\alpha|}T_n(D^{\alpha}\phi) \to (-1)^{|\alpha|}T(D^{\alpha}\phi) = D^{\alpha}T(\phi)$$

□

Remark 26. *The above propositions allows us to differentiate convergent sequences or series of distributions term by term.*

Definition 72. *A distribution W is called an **anti-derivative** of a distribution T if $\frac{dW}{dx} = T$.*

Proposition 42. *Every distribution has an anti-derivative.*

Proof: [20] 293. □

Up to now we have defined everything on $\mathbb{R}^N$. In order to carry all these to a manifold all we need to do is to use a partition of unity.

Chapter 6

Free Quantum Field Theory

We start recalling some terminology of special relativity.

As always, we denote by $\mathbb{R}^{1,3}$ the $\mathbb{R}^4$ endowed with the Lorentz inner product $< x, y >= x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4$. It is usual to denote the matrix of the Lorentz metric by η, i.e., $\eta = \mathrm{diag}(1, -1, -1, -1)$. The **Lorentz group** $O(1, 3)$ is the group of linear isometries of $<, >$. We shall identify $\mathbb{R}^{1,3}$ with $\mathbb{R}^4$ when operations which do not involve the (pseudo-) metric takes place.

The Lorentz group is not connected, having four connected components ([19] p.54). We sketch the proof of this claim. Take $x \in \mathbb{R}^{1,3}$ and a Lorentz transformation $\Lambda \in O(1, 3)$. Write $x' = \Lambda x$. The requirement of preserving $<, >$ implies $\Lambda^T \eta \Lambda = \eta$. Taking the determinant we have $\det \Lambda = \pm 1$. From the fact that the transformations preserve $<, >$ it also follows that we can not find a continuous curve jointing the transformations which preserve the $x^0 > 0$ to the transformations which preserve the $x^0 < 0$ direction. We have therefore the four components, characterized by the signs of det and by the signs of the 00 component of Λ.

Explicitely, in the standard basis of $\mathbb{R}^4$ this four components are those which contain the matrices:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

The component containing the identity matrix is called **restricted Lorentz group** and denoted by $L_+^\uparrow$. The second and third transformation above are called parity and time-reversal respectively and have important roles in applications of QFT [1].

Recall that the light cone (based on $(0,0,0,0)$) is the set

$$\overline{V} := \{x \in \mathbb{R}^{1,3} :< x, x >\geq 0\}$$

and the forward light cone[2] is the set

$$\overline{V}_+ := \{x \in \mathbb{R}^{1,3} :< x, x >\geq 0, x^0 \geq 0\}$$

We denote $\mathrm{int}(\overline{V})$ by V and $\mathrm{int}(\overline{V}_+)$ by V_+. It is important to notice that $L_+^\uparrow$ preserves the forward light cone.

We want now to introduce the Poincaré group $\mathcal{P}$.

Definition 73. *Let G and H be groups and suppose that G acts on H: $(g,h) \in G \times H \mapsto g \cdot h \in H$. We define the* ***semi-direct product*** *of G and H, denoted by $G \ltimes H$, as the group whose elements are $(g,h) \in G \times H$ and multiplication is given by $(g,h)(g',h') = (gg', h(g \cdot h'))$.*

Since $O(1,3)$ acts on $\mathbb{R}^4$ we can take the their semi-direct product. Therefore we have:

Definition 74. *The* ***Poincaré group*** *$\mathcal{P}$ is defined as $O(1,3) \ltimes \mathbb{R}^4$ and the* ***restricted Poincaré group*** *$\mathcal{P}_+^\uparrow$ as $L_+^\uparrow \ltimes \mathbb{R}^4$.*

Explicitely the multiplication on $\mathcal{P}$ is given by

$$(\Lambda, a)(\Lambda', a') = (\Lambda\Lambda', \Lambda a' + a)$$

The mathematical content of special relativity is that $\mathcal{P}_+^\uparrow$ acts on the space of solutions of the equations of motions of particles and fields.

---

[1]In fact it is known that these other transformations are not exact symmetries of the laws of particle physics; not exact in the sense that they are sometimes violated. The violation of parity and time reversal invariance were among the revolutionary experimental discoveries of 1950s and 1960s [16]. In particle physics there is another (not exact) symmetry called *charge conjugation*, which means replace a particle by its anti-particle. There is a very famous theorem which states that every physical interaction is invariant if the three operations (parity, charge conjugation and time-reversal) are applied at once. This is the so-called CPT theorem (or PCT theorem...). For a mathematical treatment of CPT theorem see [24] p. 69; for a physical discussion see [19]. For a non-technical and nice discussion see [25] p. 638.

[2]also called positive light cone

6.1 The axioms of a free scalar QFT

We start with the a axioms of a *scalar* QFT. Beyond the purely mathematical interest, axiomatic quantum field theory helps us to understand which difficulties of the theory are inherent in its structure and which ones reflected just unjustified approximations or calculations methods. It also helps to clarify the conceptual foundations of the theory in so far as we adopt the principle that mathematics is the suitable language where we model physical phenomena.

A scalar QFT is a quadruple $(\mathcal{H}, U, \phi, D)$ satisfying:

Axiom 1 (relativistic invariance of states) $\mathcal{H}$ is a separable Hilbert space and $U(\cdot,\cdot)$ a strongly continuous unitary representation of the restricted Poincaré group on $\mathcal{H}$.

Axiom 2 (spectral condition) The projection-valued measure E_Ω on $\mathbb{R}^{1,3}$ corresponding to $U(id, a) = e^{ia^\mu P_\mu}$ has support in the closed forward light cone.

Axiom 3 (existence and uniqueness of the vacuum) There exists a unique vector $\psi_0 \in \mathcal{H}$ such that $U(id, a)\psi_0 = \psi_0$ for all $a \in \mathbb{R}^{1,3}$.

Axiom 4 (invariant domains for fields) There is a dense subspace $D \subset \mathcal{H}$ and a map ϕ from the space of Schwartz $\mathcal{S}(\mathbb{R}^{1,3})$ to the unbounded operators on $\mathcal{H}$ such that: (i) for each $f \in \mathcal{S}(\mathbb{R}^{1,3})$ we have $D \subset D(\phi(f))$, $D \subset D(\phi(f)^*)$ and $\phi(f)^*|D = \phi(\bar{f})|D$; (ii) $\psi_0 \in D$ and $\phi(f)D \subset D$ for all $f \in \mathcal{S}(\mathbb{R}^{1,3})$; (iii) for fixed $\psi \in D$ the map $f \mapsto \phi(f)\psi$ is linear.

Axiom 5 (regularity of the field) For any ψ_1 and ψ_2 in D, the map $f \mapsto < \psi_1, \phi(f)\psi_2 >$ is a tempered distribution.

Axiom 6 (Poincaré invariance of the field) For each $(\Lambda, a) \in \mathcal{P}_+^\uparrow$ we have $U(\Lambda, a)D \subset D$ and for all $f \in \mathcal{S}(\mathbb{R}^{1,3})$, $\psi \in D$

$$U(\Lambda, a)\phi(f)U(\Lambda, a)^{-1}\psi = \phi((\Lambda, a)f)\psi$$

where $(\Lambda, a)f(x) = f\big(\Lambda^{-1}(x - a)\big)$.

Axiom 7 (microscopic causality) If f and g in $\mathcal{S}(\mathbb{R}^{1,3})$ have supports which are spacelike separated then $(\phi(f)\phi(g) - \phi(g)\phi(f))\psi = 0$ for every $\psi \in D$.

Axiom 8 (cyclicity of the vacuum) The set D_0 of finite linear combinations of vectors of the form $\phi(f_1)\cdots\phi(f_n)\psi_0$ is dense in $\mathcal{H}$.

These axioms are known as **Garding-Wightman axioms**. Let us comment them and extract some consequences.

Fix $\Lambda = id$. Then $U(id, a)$ is a strongly continuous unitary representation of $\mathbb{R}^4$ and therefore the spectral theorem implies the existence of four commuting self-adjoint operators P_μ on $\mathcal{H}$, $\mu = 0, 1, 2, 3$, and a projection valued measure E_Ω on $\mathbb{R}^4$ such that ([24] p. 63)

$$< \psi, U(id, a)\psi >=< \psi, e^{ia^\mu P_\mu}\psi >= \int_{\mathbb{R}^4} e^{ia^\mu \lambda_\mu} d < \psi, E_\lambda \psi >$$

P_0 is called the **Hamiltonian** and the P_j, $j = 1, 2, 3$ are the **momentum operators**.

Axiom 2 is equivalent to say that the operators P_0 and $P_0^2 - P_1^2 - P_2^2 - P_3^2$ are both positive.

Axiom 3 implies that the point $(0, 0, 0, 0)$ has non-zero E_Ω-measure. It follows that $U(\Lambda, a)\psi_0 = \psi_0$ for all $(\Lambda, a) \in \mathcal{P}_+^\uparrow$ ([24])

The purpose of axiom 4 should be well understood. Analogously to what is made in quantum mechanics, where we associate an operator to each generalized coordinate q and p and the quantization process "classical" $\longrightarrow$ "quantum" is realized by $q \mapsto Q$ and $p \mapsto P$, in QFT we want to associate an operator to each space-time point x: $x \mapsto \phi(x)$. So, if we have a classical scalar field ϕ_{cl} which is a map $x \in \mathbb{R}^{1,3} \mapsto \phi_{cl}(x) \in \mathbb{R}$, the process of quantizing the field should be realized by a map $x \in \mathbb{R}^{1,3} \mapsto \phi(x) \in$ operators on $\mathcal{H}$, $\phi(x)$ being the "quantum field" at x. But we would like to have, also, some kind of equivariance (as in axiom 6). However, Wightman has proven that this is impossible [26]: a theory which equivariantly associates an operator to each $x \in \mathbb{R}^{1,3}$ is trivial in the sense that every operator is a multiple of the identity. This was not a surprise: Bohr and Rosenfeld [27] pointed out that from a physical point of view it is impossible (as a consequence of the uncertainty principle) to measure the field strength *at* a given point. Moreover, it seemed to be clear to the founders of QFT that it is a highly singular theory.

The way of avoiding such difficulties is to consider that the field is not determined *at* x but on a small *neighborhood* of x. This means that we may take a *test function* f supported on a neighborhood of x and associate an operator to f rather than to x: $f \mapsto \phi(f)$. We may consider $\phi(f)$ as a *smeared field*: if there were a well defined field $x \mapsto \phi(x)$ then $\phi(f)$ would be $\phi(x)$ averaged against f:

$$\phi(f) = \int_{\mathbb{R}^4} \phi(x) f(x) dx$$

Therefore we have mathematical *and* physical reasons to consider the quantum field as an operator valued distribution $f \mapsto \phi(f)$ instead of an operator valued function $x \mapsto \phi(x)$. The choice of $\mathcal{S}(\mathbb{R}^{1,3})$ as a test function space for ϕ rather than $C_0^\infty(\mathbb{R}^{1,3})$ is not absolutely necessary ([24] p. 64).

Recall that two sets $A, B \subset \mathbb{R}^{1,3}$ are **spacelike separated** if $< x, y >< 0$ for every $x \in A$ and every $y \in B$. Axiom 7 says that measurements at spacelike separated points cannot interfere with each other (what is mathematically expressed saying that the corresponding operators commute), what reflects the fact that nothing can travel fast than light.

Axiom 8 ensures that the Hilbert space $\mathcal{H}$ is not too large.

6.2 Quantizing the free scalar field

First quantization is a mystery,
but second quantization is a functor.
E. Nelson

Now we would like to construct models which satisfy the axioms of QFT. In doing so we will demonstrate that the axioms are consistent.

We will quantizing the free scalar field of mass m. As the name suggests, this field describe non-interacting particles of mass m. Although free field theories are not interesting from the physical point of view, it is important to know that these ones can be mathematically rigorously formulated in that the the most natural way to construct interacting field theories is to perturb a free theory.

6.2.1 Segal quantization

Definition 75. *Let $\mathcal{H}$ be a separable Hilbert space. Denote $\mathcal{H}^{(n)} = \otimes_{k=1}^{n} \mathcal{H}$. We define the* ***Fock space*** *(over $\mathcal{H}$) as $\mathfrak{F}(\mathcal{H}) = \oplus_{n=0}^{\infty} \mathcal{H}^{(n)}$.*

Let σ be a n-element permutation, i.e., $\sigma \in S(n)$. We define an operator on elements of $\mathcal{H}^{(n)}$, also denoted by σ, as $\sigma(\psi_{k_1} \otimes \cdots \otimes \psi_{k_n}) = \psi_{\sigma(k_1)} \otimes \cdots \otimes \psi_{\sigma(k_n)}$ (actually we are defining on basis elements, but then we extend it linearly to a bounded operator of norm 1).

Definition 76. *Define the **symmetrization** and **anti-symmetrization** operators:*

$$S_n = \frac{1}{n!}\sum_{\sigma\in S(n)}\sigma \text{ and } A_n = \frac{1}{n!}\sum_{\sigma\in S(n)}\text{sign}(\sigma)\sigma$$

Definition 77. *Then we define the **symmetric Fock space** or **Bosonic Fock Space** (over $\mathcal{H}$) as:*

$$\mathfrak{F}_s(\mathcal{H}) = \bigoplus_{n=0}^{\infty} S_n\mathcal{H}^{(n)}$$

*Analogously we define the **antisymmetric Fock space** or **Fermionic Fock space** (over $\mathcal{H}$)*

$$\mathfrak{F}_a(\mathcal{H}) = \bigoplus_{n=0}^{\infty} A_n\mathcal{H}^{(n)}$$

Notice that $S_n^2 = S_n$ and $S_n^* = S_n$, i.e., S_n is an orthogonal projection. Since we want to quantize a free *scalar* field, we shall work mainly with the symmetric Fock space.

Fix $f \in \mathcal{H}$. We shall define a map $b^-(f) : \mathfrak{F}(\mathcal{H}) \to \mathfrak{F}(\mathcal{H})$. First we define it on elements of $\mathcal{H}^{(n)}$ of the form $\psi = \psi_1 \otimes \cdots \otimes \psi_n$ as:

$$b^-(f)\psi =< f, \psi_1 > (\psi_2 \otimes \cdots \otimes \psi_n)$$

Now we extend $b^-(f)$ linearly to get a (bounded map of norm $\| f \|$) from $\mathcal{H}^{(n)}$ of $\mathcal{H}^{(n-1)}$. Putting $b^-(f) : \mathcal{H}^{(0)} \to 0$ we have a bounded map of norm $\| f \|$: $b^-(f) : \mathfrak{F}(\mathcal{H}) \to \mathfrak{F}(\mathcal{H})$. Write $b^+(f)$ for $(b^-(f))^*$.

Proposition 43. *$b^+(f)$ takes $\mathcal{H}^{(n)}$ into $\mathcal{H}^{(n+1)}$. Its action on product elements $\psi = \psi_1 \otimes \cdots \otimes \psi_n$ is*

$$b^+(f)\psi = f \otimes \psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n$$

Notice that $f \mapsto b^+(f)$ is linear and $f \mapsto b^-(f)$ is anti-linear.

Definition 78. *Denote $\mathcal{H}_s^{(n)} = S_n\mathcal{H}^{(n)}$. Then $\mathfrak{F}_s(\mathcal{H}) = \bigoplus_{n=0}^{\infty}\mathcal{H}_s^{(n)}$. We call $\mathcal{H}_s^{(n)}$ the n**-particle subspace of** $\mathfrak{F}_s(\mathcal{H})$. A vector $\psi = \{\psi^{(n)}\}_{n=0}^{\infty} \in \mathfrak{F}_s(\mathcal{H})$ for which $\psi^{(n)} = 0$ for all but finitely many n is called **finite particle vector**. Denote by F_0 the set of finite particle vectors. We call $\Omega = (1, 0, 0 \dots)$ the **vacuum***

It is easy to check that $b^-(f)$ takes $\mathfrak{F}_s(\mathcal{H})$ into itself. However $b^+(f)$ does not.

Definition 79. *Let A be any self-adjoint operator on $\mathcal{H}$, denote by D its domain of essential self-adjointness. Denote by*

$$D_A := \{\psi \in F_0 : \psi^{(n)} \in \bigotimes_{k=1}^{n} D \text{ for each } n\}$$

Definition 80. *Define $d\Gamma(A)$ on $D_A \cap \mathcal{H}_s^{(n)}$ as*

$$A \otimes id \otimes \cdots \otimes id + id \otimes A \otimes id \cdots \otimes id + \cdots + id \otimes \cdots \otimes A$$

We call $d\Gamma(A)$ the ***second quantization*** *of A.*

Proposition 44. *$d\Gamma(A)$ si essentially self-adjoint on D_A.*

Proof. [24] p. 208. □

In particular for $A = id$ we have $d\Gamma(id)\psi = n\psi$, $\psi^{(n)} \in \mathcal{H}_s^{(n)}$.

Definition 81. *We call $N = d\Gamma(id)$ the* ***number operator****.*

Notice that N is essentially-self adjoint on F_0. If U is a unitary operator on $\mathcal{H}$ we define $\Gamma(U)$ to be the unitary operator on $\mathfrak{F}_s(\mathcal{H})$ which equals $\otimes_{k=1}^{n}$ when restricted to $\mathcal{H}^{(n)}$ for $n > 0$ and equals the identity on $\mathcal{H}_s^{(0)}$. If e^{itA} is a continuous unitary group on $\mathcal{H}$ then $\Gamma(e^{itA})$ is the group generated by $d\Gamma(A)$, i.e., $\Gamma(e^{itA}) = e^{itd\Gamma(A)}$.

We can now introduce two fundamental operators of QFT.

Definition 82. *Given $f \in \mathcal{H}$ define $a^-(f)$ on $\mathfrak{F}_s(\mathcal{H})$ with domain F_0 by*

$$a^-(f) = \sqrt{N+1}b^-(f)$$

$a^-(f)$ is called the ***annihilation operator*** *because it takes each $(n+1)$-particle subspace into the n-particle subspace.*

Proposition 45. $(a^-(f))^*|F_0 = Sb^+(f)\sqrt{N+1}$

Proof. Just use that S is an orthogonal projection. Then

$$< \sqrt{N+1}b^-(f)\psi, \eta >=< \eta, Sb^+(f)\sqrt{N+1}\eta >$$

for every $\psi, \eta \in F_0$. □

Definition 83. *We call* $(a^-(f))^*|F_0$ *the* ***creation operator***.

Both $a^-(f)$ and $(a^-(f))^*|F_0$ are closable. We also denote their closures by $a^-(f)$ and $(a^-(f))^*|F_0$.

Definition 84. *Define the* ***Segal field operator*** $\Phi_S(f)$ *on* F_0 *by*

$$\Phi_S(f) = \frac{1}{\sqrt{2}}\big(a^-(f) + a^-(f)^*\big)$$

The map from $\mathcal{H}$ *to the self-adjoint operators on* $\mathfrak{F}_s(\mathcal{H})$ *(see theorem below) given by* $f \mapsto \Phi_S(f)$ *is called the* ***Segal quantization over*** $\mathcal{H}$.

Notice that the map $f \mapsto \Phi_S(f)$ is real linear but not complex linear since $f \mapsto a^-(f^*)$ is linear but $f \mapsto a^-(f)$ is anti-linear

Theorem 22. $\Phi_S(f)$ *is essentially self-adjoint on* F_0.

Proof. [24] p. 210. □

6.2.2 Lorentz invariant measures

We shall introduce a measure which is Lorentz invariant in the sense of the definition below. It plays a fundamental role in the quantization of free fields.

Definition 85. *For each* $m \geq 0$ *let* $H_m := \{x \in \mathbb{R}^{1,3} :< x, x >= m^2, x_0 > 0\}$. *These sets are called* ***mass hyperboloids*** *(of mass* m*).*

It is a straightforward consequence of the definition that H_m is invariant under $L_+^\uparrow$. Let j_m be the homeomorphism of H_m onto $\mathbb{R}^3$ (onto $\mathbb{R}^3 - \{0\}$ in case $m = 0$) given by $(x_0, x_1, x_2, x_3) \mapsto (x_1, x_2, x_2) \equiv \vec{x}$. We define a measure Ω_m on H_m by setting:

$$\Omega_m(E) = \int_{j_m(E)} \frac{d^3x}{\sqrt{m^2 + \parallel \vec{x} \parallel^2}}$$

for any measurable set $E \subset H_m$.

Proposition 46. Ω_m *on* H_m *is invariant under* $L_+^\uparrow$.

Proof: Notice that $\overline{V}_+ = \{0\} \cup (\cup_{m=0}^{\infty} H_m)$. As H_m, the set $\{0\}$ is invariant under $L_+^\uparrow$, so is $\overline{V}_+$.

d^4x is $L_+^\uparrow$-invariant since $\Lambda \in L_+^\uparrow$ implies that $\det \Lambda = 1$. Let $f \in C_c^\infty(0,\infty)$ (where $_c$ means "with compact support"). From the invariance of $\overline{V}_+$ follows that $f(< x, x >)\chi_{\overline{V}_+} d^4x$ is invariant under $L_+^\uparrow$ (χ=characteristic function). Now map V_+ homeomorphically onto $\mathbb{R}^3 \times \mathbb{R}^+$ by $h : (x_0, \vec{x}) \mapsto (\vec{x}, y)$ where $y =< x, x >$. Then $\frac{\partial y}{\partial x_0} = 2x_0$ so

$$d^4x = \frac{d^3xdy}{2\sqrt{m^2 + < x, x >^2}}$$

and hence the measure

$$\Omega^f(E) = \int_{h(E)} \frac{d^3xdy}{2\sqrt{m^2 + < x, x >^2}}$$

is $L_+^\uparrow$ invariant. Take a sequence in $C_c^\infty(0,\infty)$ converging to $\delta(y - m^2)$, $m > 0$. Then Ω^{f_n} converges to Ω_m in $\mathcal{S}'(\mathbb{R}^{1,3})$. Therefore $\Omega_m(g(x)) = \Omega_m(g(\Lambda x))$ for every $g \in \mathcal{S}(\mathbb{R}^{1,3})$ and $\Lambda \in L_+^\uparrow$. This generalizes to $\Omega_m(E) = \Omega_m(\Lambda E)$. Since $\Omega_m \to \Omega_0$ in $\mathcal{S}'(\mathbb{R}^{1,3})$ as $m \downarrow 0$ we also have proven the case when $m = 0$. □

Remark 27. *Physicists like to write $\delta(x^2 - m^2)d^4x$ or even $\delta(x^2 - m^2)$ for $d\Omega_m$.*

6.2.3 Quantizing the free scalar field

We can now quantize the free scalar field of mass m. Recall that the **Fourier transform**

$$f \in \mathcal{S}(\mathbb{R}^{1,3}) \mapsto \hat{f} \in \mathcal{S}(\mathbb{R}^{1,3})$$
$$\hat{f}(p) = \frac{1}{(2\pi)^2} \int e^{i<p,x>} f(x)dx$$

gives a linear bi-continuous bijection from $\mathcal{S}(\mathbb{R}^{1,3})$ onto $\mathcal{S}(\mathbb{R}^{1,3})$. In physical language, it allows us to pass from position space (x variable) to the momentum space (p variable) (obviously using the Riesz representation theorem to identify the space of linear functionals on $\mathbb{R}^{1,3}$ with $\mathbb{R}^{1,3}$). Notice that here we are defining the Fourier transform in terms of the (invariant) Lorentz inner product.

We shall take as Hilbert space $L^2(H_m, d\Omega_m)$. Let Φ_S be the Segal quantization over $L^2(H_m, d\Omega_m)$. We define for each $f \in \mathcal{S}(\mathbb{R}^{1,3})$ we define Ef in $L^2(H_m, d\Omega_m)$ by $Ef = \sqrt{2\pi}\hat{f}|H_m$; and for each real-valued $f \in \mathcal{S}(\mathbb{R}^{1,3})$:

$$\Omega_m(f) = \Omega_S(Ef)$$

For arbitrary $f \in \mathcal{S}(\mathbb{R}^{1,3})$ we put

$$\Omega_m(f) = \Omega_m(\Re f) + i\Phi_m(\Im f)$$

Definition 86. *The mapping $f \mapsto \Phi_m(f)$ is called* ***free Hermitian scalar field of mass*** *m.*

On $L^2(H_m, d\Omega_m)$ we define the following unitary representation of $\mathcal{P}_+^\uparrow$:

$$\big(U_m(\Lambda, a)\psi\big)(p) = e^{i<p,a>}\psi(\Lambda^{-1}p)$$

Then we have our main result:

Theorem 23. *(existence of free quantum fields) The quadruple:*

$$\Big(\mathfrak{F}_s\big(L^2(H_m, d\Omega_m)\big), \Gamma\big(U_m(\cdot,\cdot)\big), \Phi_m, F_0\Big)$$

satisfies the Wightman axioms.

Proof: [24] p. 213. □

Chapter 7

Interacting QFT

The mathematical treatment of interacting QFT is one of the most demanding areas of mathematical physics. Physicists have developed a lot of techniques to extract results from realistic models in QFT. Although most of such techniques are purely formal or even mathematically unjustifiable, the ability of physicist to obtain results with them encourages mathematicians to pursue a rigorous approach to them. It must be mentioned that by "results" we mean not only the predictions with extraordinary agreement with experiments but also some amazing mathematical conjectures and ideas which were finally proven by mathematicians (e.g.: the outcomes involving Seiberg-Witten equations and Donaldson's theory)[1].

7.1 Axioms for quantization of spinor fields

We want now to extend our quantization procedures to spinor fields. In the final we shall use a different approach to quantization, namely, the path-integral method. However, we think it is instructive to say what changes in axioms we should introduce in order to accommodate spinor fields.

In order to introduce spin in the theory we make the following modifications in the axioms: There is a two-to-one map Λ of $SL(2,\mathbb{C})$ onto $L_+^\uparrow$— the

[1]"Physicists have formulated a number of striking conjectures (...) and concepts (...). In many cases mathematicians have been able to verify the conjectures of physicists, but the proofs have dealt with each individual case and ignore the bigger picture which governs the physicists intuition. The basis of the physicists intuition is their belief that underlying quantum field theory and string theory is a (as yet undiscovered) self-consistent mathematical framework", in the introduction to [17].

role of $SL(2\mathbb{C})$ will be clarified below. We take $SL(2,\mathbb{C}) \times \mathbb{R}^4$ with product $(A,a)(B,b) = (AB, a + \Lambda(A)b)$. Let S be a finite-dimensional (not necessarily unitary) irreducible representation of $SL(2\mathbb{C})$ on a d-dimensional space A **spinor field of type** S is an object obeying the axioms for a Hermitian scalar field with the following changes ([24], p. 117):

Change 1 The single field ϕ is replaced by a d-tuple of fields $(\phi_1, \ldots, \phi_d)$.

Change 2 The field $\phi(f)$, f real, is not required to be symmetric.

Change 3 The transformation law (axiom 6) is replaced by:

$$U(\Lambda,a)\phi_i(f)U(\Lambda,a)^{-1}\psi = \sum_{j=1}^{d} S(A^{-1})_{ij}\phi((\Lambda,a)f)\psi$$

where $(\Lambda,a)f(x) = f\big(\Lambda^{-1}(x-a)\big)$ and $\Lambda = \Lambda(A)$.

Change 4 The vacuum is only required to be cyclic for

$$\{\phi_1(f), \ldots, \phi_d(f), \phi_1^*(f), \ldots, \phi_d^*(f) : f \in \mathcal{S}(\mathbb{R}^{1,3})\}$$

Change 5 If f and g in $\mathcal{S}(\mathbb{R}^{1,3})$ have supports which are spacelike separated then: (a) (Bose statistics) $\phi_i(f)\phi_j(g) - \phi_j(g)\phi_i(f) = 0$ and $\phi_i^*(f)\phi_j(g) - \phi_j(g)\phi_i^*(f) = 0$ if $S(-1) = 1$; (b) (Fermi statistics) $\phi_i(f)\phi_j(g) + \phi_j(g)\phi_i(f) = 0$ and $\phi_i^*(f)\phi_j(g) + \phi_j(g)\phi_i^*(f) = 0$ if $S(-1) = -1$.

7.2 The Dirac bundle

We want to start our investigation of interacting quantum fields. The first thing to do is to introduce interaction classically. After this, we shall quantize the fields in the path-integral formalism.

7.2.1 The Lorentz group and $SL(2,\mathbb{C})$

Our objective in this section is to show that $SL(2,\mathbb{C})$ is a double-cover of $L_+^\uparrow$. The arguments here a very standard and therefore we just sketch them. More details may be found in [28] p. 504ff and [19] chapter 7.

Recall that the Lorentz group $O(1,3)$ is the groups of linear isometries of the Minkowski metric, i.e.,linear transformations Λ in $\mathbb{R}^{1,3}$ such that $< \Lambda x, \Lambda y >=< x, y >$, and that we are interested in the connected component containing the identity $L_+^\uparrow$.

Define

$$H(2,\mathbb{C}) := \{2\times 2 \text{ matrices } A : A^* = A\}$$

It is straightforward to check that the Pauli matrices $\sigma_j,\ j = 1,2,3$ and the identity (denoted by σ_0) form a basis for $H(2,\mathbb{C})$. Define a map $_* : \mathbb{R}^{1,3} \to H(2,\mathbb{C})$ by $x \in \mathbb{R}^{1,3} \mapsto x_*$ where

$$x_* = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}$$

Solving for x we get that $x^j = \frac{1}{2}Tr(x_*\sigma_j)$. It follows $\det x_* = - < x, x >$. So we have an identification of $\mathbb{R}^{1,3}$ with $H(2,\mathbb{C})$.

The basic theorem is:

Theorem 24. *The assigment to $A \in SL(2,\mathbb{C})$ of the linear map Λ of Minkovski space*

$$\Lambda(A) : \mathbb{R}^{1,3} \approx H(2,\mathbb{C}) \to H(2,\mathbb{C})$$
$$\Lambda(A)(x)_* := Ax_*A^*$$

yields a $2:1$ homomorphism of $SL(2,\mathbb{C})$ onto $L_+^\uparrow$.

See [28] for a proof.

7.2.2 The Dirac bundle

Let $\mathcal{M}$ be a pseudo-Riemannian four dimensional spin manifold which both time and space orientable. Assume that the structural group of the tangent bundle is $L_+^\uparrow$. From these hypothesis, the previous section and chapter 3 we obtain that the spin structure on $T\mathcal{M}$ has structural group $SL(2,\mathbb{C})$. Consider the representation of $SL(2,\mathbb{C})$ on $\mathbb{C}^4$ given by

$$\rho : SL(2,\mathbb{C}) \to \text{Aut}(\mathbb{C}^4)$$
$$\rho(A) = \begin{pmatrix} A & 0 \\ 0 & (A^*)^{-1} \end{pmatrix}$$

Then we obtain a vector E bundle over $\mathcal{M}$ whose fiber is $\mathbb{C}^4$. We also have the Dirac operator on E.

Bibliography

[1] G. L. NABER. *Topology, geometry and gauge fields*. Springer, 2000.

[2] STEENROD. *The topology of fiber bundles*. Princeton University, 1951.

[3] MARATHE and MARTUCCI. *The mathematical foundations of gauge theories*. North-Holland, 1992.

[4] T. OKUBO. *Differential geometry*. M. Dekker, 1987.

[5] R. PALAIS. The geometrization of physics. 1981. www.math.neu.edu/ terng/GeometrizationOfPhysics.pdf.

[6] S. K. DONALDSON and P. B. KRONHEIMER. *The geometry of four manifolds*. Oxford Science Publications, Oxford, 1990.

[7] J.W. MORGAN. *An introducrtion to gauge theory, in: Gauge theory and the topology of four-manifolds*. AMS, 1998.

[8] J. BAEZ. *Gauge fields, knots and gravity*. World Scientific, 1994.

[9] P. W. MICHOR. *Gauge Theory for Fiber Bundles*. Monographs and Textbooks in Physical Sciences, Lecture Notes 19, Bibliopolis, Napoli, 1991.

[10] J. JOST. *Riemannian Geometry and Geometric Analysis (fourth edition)*. Springer, 2005.

[11] H. B. LAWSON and M. MICHELSHON. *Spin geometry*. Princeton University, 1989.

[12] J. W. MORAGAN. *The Seiberg-Witten equations and applications to the topology of smooth four-manifolds*. Princeton University, 1996.

[13] T. FRIEDRICH. *Dirac operators in Riemannian geometry*. AMS, 2000.

[14] VARADARAJAN. *Supersymmetry for mathematicians: an introduction*. Courant Institute of Mathematical Sciences, 2004.

[15] D. S. DELIGNE, P.; FREED. Classical field theory. *in: Quantum fields and strings: a course for mathematicians*, 1999.

[16] J. RABIN. Introduction to quantum field theory for mathematicians. *in: Geometry and quantum fields theory*, 1995.

[17] P. DELIGNE (Ed). Glossary. *in: Quantum fields and strings: a course for mathematicians*, 1999.

[18] Y. CHOQUET-BRUAT. *Analysis, Manifolds and Physics*. North-Holland, 1982.

[19] TICCIATI. *Quantum field theory for mathematicians*. Cambridge University Press, 1999.

[20] P DEBNATH, L.; MINUSINSKI. *Introduction to Hilbert spaces with applications*. Academic press, 1990.

[21] L. SCHWARTZ. *Théorie des distribuitions*. Hermann, 1998.

[22] G.E. GUELFAND, I.M.; CHILOV. *Les distributions*. Dumond, 1962.

[23] F CONSTANTINESCU. *Distributions and their applications in Physics*. Pergamon press, 1980.

[24] REED and SIMON. *Methods of Modern Mathematical Physics vol. 2*. Academic Press Inc, 1975.

[25] R. PENROSE. *The road to reality*. Alfred A. Knopf, 2004.

[26] A.S. WIGHTMAN. La théorie quantique locale et la théorie quantique des champs. *Ann. Inst. Henri Poincaré*, A1:403–420, 1964.

[27] N. BOHR and L. ROSENFELD. Zur frage der messbarkeit der elektromagnetischen feldgrössen. *Kgl. Danske Vid. Selsk. Mat. Fys. Medd.*, 12(8), 1933.

[28] T. FRANKEL. *The geometry of physics: an introduction*. Cambridge University Press, Cambridge, 2004.