Preparation for Gauge Theory

George Svetlichny[1]

February 19, 1999

[1]Departamento de Matemática, Pontifícia Universidade Católica, Rua Marquês de São Vicente 225, 22453-900 Gávea, Rio de Janeiro, RJ, Brazil
e-mail: svetlich@mat.puc-rio.br

Contents

1 Introduction

These are the class notes for a course preparatory to classical gauge theory given at the Mathematics Department of the Pontifícia Universidade Católica of Rio de Janeiro during the (southern-hemisphere) spring term of 1997. They purport to provide the necessary mathematical background at a beginning graduate level for someone interested in gauge theory and who has an elementary understanding of differentiable manifolds, Lie groups, Lie algebras, and multilinear algebra. These notes are intended to take the reader to the point at which he or she can understand what gauge theories *are*, but, unfortunately, stop short of doing anything with them. They do not treat the more advanced, and more exciting topics, such as moduli spaces, topological invariants, and quantum aspects.

Field theory was a scientific revolution initiated in the nineteenth century with the development of Maxwell's electrodynamics. It gained a strongly geometric character with Einstein's general theory of relativity. A century later, the revolution still continues with new insights and ideas appearing without abating. Much of the motivation and inspiration comes from the still largely incomplete theory of quantum fields and its generalizations, such as string and M-theory. These quantum theories exercise a strong influence on classical field theory and many aspects of research into classical field theory are incomprehensible without understanding the effort to gain insight into quantum theories. Since these notes are directed toward classical gauge theory, certain construct may seem arbitrary or unmotivated as their full appreciation can only come through quantum theory. These would have to be taken mostly on faith that they are appropriate and interesting, as we have no means to explore the quantum aspects in these notes.

A second revolution in field theory occurred with the advent of gauge theory to describe fundamental particle interactions. There were many fundamental shifts in viewpoint, among them the realization that physical fields, originally represented by functions on space-time, had to be treated by topologically more sophisticated objects, technically know as sections of fiber bundles. This change was already implicitly presaged by general relativity, which for a long time led the process of introducing geometric reasoning into physics. The full changeover however had to await the flowering of gauge theory.

This "gauge revolution" had remarkable and totally unexpected mathematical consequences. The very same equations that were instrumental in constructing the most successful theory of matter know to date, provided subtle tools for exploring the structure of low dimensional manifolds and objects therein, by means of entirely new invariants. A mathematical revolution followed at the heels of the physical one. These notes are dedicated to shorten the path of the interested reader to the point where he or she can begin to appreciate the fruits of these remarkable developments.

I am indebted to the participants of the course for many helpful remarks and suggestions.

2 Preliminaries, Notation, Conventions

We assume the reader is familiar with the basic notions of general topology, group theory, multilinear algebra, differentiable manifolds, Lie groups and Lie algebras, and a smattering of categorical ideas. Although fiber-bundle theory is treated in these notes, we assume the reader has some elementary understanding of the standard bundles found in manifold theory, such as the tangent bundle, the cotangent bundle, the bundle of exterior forms, and tensor bundles.

We shall have occasion in these notes to deal with similar constructs in several categories, the main ones being sets, topological spaces, and $\mathcal{C}^\infty$ differentiable manifolds. By words such as "map", "morphism", "isomorphism", etc. we shall mean the notion appropriate to the category. Thus when talking about manifolds, by "map" we shall mean a $\mathcal{C}^\infty$ map, when talking about topological spaces, a continuous map, and when talking about sets, just a map. Similarly for other notions. Some expositions will be done in one category and the translation to other categories, when it is straightforward, will be left to the reader.

We shall use various notational conventions which are summarized in the appendices. Mention to the appendices are made upon first use of the conventions in the text. Appendix A resumes the main conventions of these notes as a whole.

Most of the material in these notes is elementary and the proofs are easy. We leave out some proofs that are either very technical or do not contribute to understanding the essentials of the subject.

3 Groups

3.1 Group Action

In this section we are in the category of sets. With the understanding of Section 2, most of the definitions and results can be carried over without change when the objects are topological spaces or manifolds.

Let G be a group and X a set. By a *left action* of G on X we mean a map $\alpha : G \times X \to X$ which satisfies

1. $\alpha(e, x) = x$
2. $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$.

For convenience we shall normally write $g \cdot x$ instead of $\alpha(g, x)$. In this notation the two axioms become $e \cdot x = x$ and $g \cdot (h \cdot x) = (gh) \cdot x$. One easily sees that each map $\alpha(g, \cdot)$ is invertible with inverse $\alpha(g^{-1}, \cdot)$, and that by axiom (2) the map $g \mapsto \alpha(g, \cdot)$ is a group homomorphism $G \to \mathrm{Aut}(X)$. Reciprocally, any group homomorphism of G into $\mathrm{Aut}(X)$ defines a left action. Similarly, by a *right action* of G on X we mean a map $\beta : X \times G \to X$ which satisfies

1. $\beta(x, e) = x$
2. $\beta(\beta(x, g), h) = \beta(x, gh)$.

For convenience we shall normally write $x \cdot g$ instead of $\beta(x, g)$. Properties analogous to those of left action hold also for right action, however now $g \mapsto \beta(\cdot, g)$ is a group *anti-homomorphism* $G \to \mathrm{Aut}(X)$.

We have a natural left and a natural right action of a group G on itself, given by multiplication: $\alpha(g, h) = gh$ and $\beta(h, g) = hg$.

If we have a left action on X by a group G and simultaneously a right action by a group H, we say the two actions *commute* if $(g \cdot x) \cdot h = g \cdot (x \cdot h)$ for all $x \in X$, $g \in G$, and $h \in H$.

Let G be a group and H a subgroup. We have a left action of G on the right coset spaces kH of G defined by $g \cdot (kH) = gkH$, and we likewise have a right action of G on the left coset spaces by $(Hk) \cdot g = Hkg$. When the subgroup H is normal, then the left and right coset spaces coincide and so one has both a left and a right action on them and the two actions commute. The actions of G on its (left and right) coset spaces defined above are called *canonical* actions.

Let (G, X, α) and (G, Y, γ) be two left actions of the same group G. A morphism from the first action to the second is a map $f : X \to Y$ such that

$$f(g \cdot x) = g \cdot f(x) \tag{1}$$

This is equivalent to the commutativity of the following diagram

$$\begin{tikzcd} G \times X \arrow[r, "\alpha"] \arrow[d, "\mathrm{Id} \times f"'] & X \arrow[d, "f"] \\ G \times Y \arrow[r, "\gamma"] & Y \end{tikzcd}$$

We say such a morphism is an isomorphism if f is invertible. Isomorphic group actions will also be called *equivalent*. Analogous definitions hold for right actions.

If f is invertible then, given $y \in Y$, one can write (1) as

$$g \cdot y = f(g \cdot f^{-1}(y)) \tag{2}$$

This equation can be used, whenever one has an action in X but not in Y, to *define* one in Y that is equivalent to the one in X. It is an easy exercise to show that (2) does indeed define an action.

Note that the definition of left and right action of a group does not use the existence of the inverse in a group, and an identical definition can be given for a semigroup S. We thus have the more general notion of *semigroup action*, which at times is useful to consider.

If (G, X, α) is a left action, then $G_0 = \{g \in G \,|\, \forall x \in X, g \cdot x = x\}$ is a normal subgroup of G, being the kernel of the homomorphism $G \to \mathrm{Aut}(X)$. One then has the action of the quotient group G/G_0 on X given by $gG_0 \cdot x = g \cdot x$. It is easy to verify that this is well defined and that the axioms are satisfied. An action for which $G_0 = \{e\}$ is called *effective*. We shall normally deal only with effective actions though at times non-effective actions will arise.

Given a left action of G on X and a point $x \in X$ we call the set $\mathcal{O}_x = \{g \cdot x \,|g \in G\}$ the *orbit* of x. We have the following

Theorem 1 *The set of orbits $\{\mathcal{O}_x \,|\, x \in X\}$ partitions the set X.*

Proof: As $x \in \mathcal{O}_x$ the set of orbits cover X. Suppose $z \in \mathcal{O}_x \cap \mathcal{O}_y$ and $w \in \mathcal{O}_x$. We have $z = g \cdot x$, $z = k \cdot y$ and $w = h \cdot x$ for some elements $g, h, k \in G$. From this we find $w = hg^{-1}k \cdot y \in \mathcal{O}_y$ and so $\mathcal{O}_x \subset \mathcal{O}_y$ but by symmetry we also have $\mathcal{O}_y \subset \mathcal{O}_x$ and so the two orbits coincide. Thus two orbits either are disjoint or coincide and we have a partition. Q.E.D

We say an action is *transitive* if there is only one orbit, that is, given any $x_1, x_2 \in X$ there is a $g \in G$ such that $x_2 = g \cdot x_1$.

If $Y \subset X$ is a union of orbits and $y \in Y$, then $g \cdot y \in Y$ for all $g \in G$. We can thus restrict the action of G to Y. Clearly, the minimal subsets for which we can do this are single orbits.

For $x \in X$ define $K_x = \{g \in G \,|\, g \cdot x = x\}$. Obviously K_x is a subgroup of G. It is called the *stability subgroup* of x.

We say an action is *free* if $K_x = \{e\}$ for all x. In particular this means that given two points on an orbit, there is a unique element of the group that connects the two.

Theorem 2 *Let $y \in \mathcal{O}_x$ then $K_y = gK_xg^{-1}$ where g is any element of G such that $y = g \cdot x$.*

Proof: We have $h \cdot y = y$ if and only if $hg \cdot x = g \cdot x$ if and only if $g^{-1}hg \cdot x = x$. Q.E.D

Theorem 3 *The action restricted to any orbit $\mathcal{O}_x$ is equivalent to the canonical action of G on the right coset spaces of the stability subgroup K_x of x.*

Proof: Let $Z = \{gK_x \,|\, g \in G\}$ and define $f : \mathcal{O}_x \to Z$ by associating to $y = a \cdot x \in \mathcal{O}_x$ the coset aK_x. One has $a \cdot x = b \cdot x$ if and only if $a^{-1}b \cdot x = x$ that is, if and only if $a^{-1}b \in K_x$, and this if and only if $aK_x = bK_x$. Thus the correspondence is well defined and injective. It is obviously surjective, and so bijective. One has $f(g \cdot y) = f((ga) \cdot x) = gaK_x = g \cdot f(y)$ and so f defines an equivalence of actions. Q.E.D

> **Example 1** *The special orthogonal group $SO(3)$ acts naturally on $\mathbb{R}^3$. The orbit of any non-zero vector is the sphere centered at the origin that contains it. The stability subgroup of $(0, 0, 1)$ is $SO(2)$, the group of rotations of the x-y plane. One thus has for the coset space*
>
> $$SO(3)/SO(2) \simeq S^2$$
>
> *One should view this as an isomorphism of manifolds.*

Theorem 4 *Two canonical left actions on right coset spaces G/H and G/K are equivalent if and only if the subgroups H and K are conjugate.*

Proof: Suppose $f : G/H \to G/K$ establishes the equivalence. Let $f(H) = wK$. For $g \in H$ one has $g \cdot f(H) = f(gH) = f(H)$ and so $gwK = wK$. Thus $w^{-1}gw \in K$ and we conclude that $w^{-1}Hw \subset K$. On the other hand, given $k \in K$ one has $f(wkw^{-1}H) = wkw^{-1}f(H) = wkw^{-1}wK = wK = f(H)$. As f is an isomorphism one has $wkw^{-1}H = H$ meaning that $wKw^{-1} \subset H$ which with the previous inclusion implies that $H = wKw^{-1}$. Q.E.D

We see thus that there are universal models for the action within any single orbit, the canonical action on coset sets, and that the equivalence classes of actions on single orbits is in bijective correspondence with the set of conjugacy classes of subgroups of G. The actions within an orbit can be studied entirely within the structure of the group G. The study of group actions thus divides neatly into two problems: the *internal* problem of understanding the action within single orbits (equivalent to studying the canonical action in coset spaces) and the *external* problem of understanding how the orbits are put together to form the set X.

Since the orbits partition X, they define an equivalence relation and the quotient space by this relation is called the *space of orbits*. Understanding the space of orbits is often of utmost importance.

3.2 Lie Groups and Lie Group Actions

In this section we review, without proofs, some basic facts about Lie groups. Readers desiring greater detail should consult appropriate textbooks on the subject.

Let G be a Lie group. When convenient, we shall in this section denote the right action of G on G by ρ. Thus $\rho_g h = hg$. Similarly for the left action, which we shall denote by λ, and write $\lambda_g h = gh$. These maps induce isomorphisms $d\rho_g : T_hG \to T_{hg}G$ and $d\lambda_g : T_hG \to T_{gh}G$. To facilitate notation we shall write $d\rho_g(v) = v \cdot g$ and $d\lambda_g(v) = g \cdot v$.

Of particular interest is the tangent space at the identity T_eG which is known as the *Lie algebra* $\mathfrak{g}$ of G. The tangent space T_gG at any other point can be canonically identified with $\mathfrak{g}$ in two ways, either by the map $d\rho_{g^{-1}}$ or by $d\lambda_{g^{-1}}$. In these notes we shall conventionally use the identification by $d\rho$. Under such an identification, a vector $v \in T_gG$ will have the form $v = L \cdot g$

for some $L \in \mathfrak{g}$. Given a map $\phi : G \to M$ where M is any manifold, one has $d\phi_g : T_g G \to T_{\phi(g)} M$. With the identification of $T_g G$ with $\mathfrak{g}$ by right action one has $d\phi_g v = \tilde{d}\phi_g(v \cdot g^{-1})$ for a map $\tilde{d}\phi_g = d\phi_g \circ d\rho_{g^{-1}} : \mathfrak{g} \to T_{\phi(g)} M$. Thus

$$v = L \cdot g \quad \Rightarrow \quad d\phi_g v = \tilde{d}\phi_g L \tag{3}$$

We shall make frequent use of this construct.

The adjoint map

$$\mathrm{Ad}_g(L) = g \cdot L \cdot g^{-1}$$

maps $\mathfrak{g}$ into itself.

Given a vector $L \in \mathfrak{g}$ one can extend it to a vector field $\mathcal{X}_L^\lambda$ on G by $\mathcal{X}_L^\lambda(g) = g \cdot L$ using left action, and likewise to a vector field $\mathcal{X}_L^\rho(g) = L \cdot g$ by right action. These vector fields are *invariant* in the sense that $g \cdot \mathcal{X}_L^\lambda(h) = \mathcal{X}_L^\lambda(gh)$ and $\mathcal{X}_L^\rho(h) \cdot g = \mathcal{X}_L^\rho(hg)$. Given these vector fields one can introduce a Lie bracket in $\mathfrak{g}$ by

$$[L, K] = [\mathcal{X}_L^\lambda, \mathcal{X}_K^\lambda](e) = [\mathcal{X}_L^\rho, \mathcal{X}_K^\rho](e)$$

which happens to be the same using either left-invariant or right invariant extension. We recall the properties of the Lie bracket

$$[L, K] = -[K, L] \tag{4}$$

$$[L, [K, M]] + [K, [M, L]] + [M, [L, K]] = 0 \tag{5}$$

Property (5) is called the *Jacobi Identity*.

Furthermore, each vector field $\mathcal{X}_L^\lambda$ and $\mathcal{X}_L^\rho$ defines a flow, $\exp(t\mathcal{X}_L^\lambda)$ and $\exp(t\mathcal{X}_L^\rho)$ respectively. The integral curve passing through the identity e is the same for the two vector fields and we define

$$\exp(tL) = \exp(t\mathcal{X}_L^\lambda)e = \exp(t\mathcal{X}_L^\rho)e$$

Thus $\exp(tL) \in G$, and $L \mapsto \exp(L)$ defines a map $\exp : \mathfrak{g} \to G$ called the *exponential map*. We also write e^L for $\exp(L)$. One has $\exp(t\mathcal{X}_L^\lambda)g = e^{tL}g$ and $\exp(t\mathcal{X}_L^\rho)g = ge^{tL}$ so that both flows are easily expressed through the exponential map. One has the useful formula

$$g \exp(L) g^{-1} = \exp(g \cdot L \cdot g^{-1}) = \exp(\mathrm{Ad}_g L)$$

Let now α be a left action of G on a manifold M. The differential $d\alpha_{(g,x)} : T_{(g,x)}(G \times M) \to T_{g \cdot x} M$ can be written as $(v, \xi) \mapsto d_1\alpha_{(g,x)} v + d_2\alpha_{(g,x)} \xi$ where

$d_1\alpha_{(g,x)}$ is the partial differential with respect to the first variable, that is, the differential at g of the map $g \mapsto g \cdot x$ for x fixed, and $d_2\alpha_{(g,x)}$ is the differential at x of the map $x \mapsto g \cdot x$ for g fixed. One has $T_{(g,x)}(G \times M) \simeq T_g G \times T_x M \simeq \mathfrak{g} \times T_x M$ where we used the identification of $T_g G$ with $\mathfrak{g}$ by right action. Thus we can write

$$d_1\alpha_{(g,x)} v = \tilde{d}_1\alpha_{(g,x)}(v \cdot g^{-1}) = \tilde{d}_1\alpha_{(g,x)} L \tag{6}$$

where $L \in \mathfrak{g}$ corresponds to $v \in T_g G$ under the mentioned identification.

When $G \subset \mathrm{GL}(X)$ for a finite-dimensional vector space X with the base field $\mathbb{F}$ being $\mathbb{R}$ or $\mathbb{C}$, the Lie algebra $\mathfrak{g}$ is the linear space of endomorphisms $L \in \mathrm{End}(X)$ such that $\exp(tL) \in G$ for all t, where now exp is the usual exponential defined either by functional calculus of by the exponential series

$$\exp(tL) = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n$$

The Lie bracket in this case is the commutator $[L, K] = LK - KL$. Each tangent space $T_g G$ is a linear space consisting of endomorphisms of the form gL, or equivalently of the form Lg, for $L \in \mathfrak{g}$, and where the product is ordinary composition. The differentials of the right and left action of G on itself are likewise given by composition. Thus if $v \in T_g G$, then $g \cdot v = gv$ and $v \cdot g = vg$. In particular $\mathrm{Ad}_g L = gLg^{-1}$. When $G = \mathrm{GL}(X)$, then $\mathfrak{g} = \mathrm{End}(X)$.

When $X = \mathbb{F}^n$, then G is a matrix group, and all the spaces mentioned above are spaces of matrices. Composition is ordinary matrix multiplication. In particular, when $G = \mathrm{GL}(n, \mathbb{F})$, then $\mathfrak{g} = \mathcal{M}(n, \mathbb{F})$, the space of all $n \times n$ matrices over $\mathbb{F}$.

3.3 Group and Lie Algebra Representations

A particular type of group action is given by *group representations*. This is a left action of a group G on a *vector space* X over a field $\mathbb{F}$ in which each map $x \mapsto g \cdot x$ is $\mathbb{F}$-linear. One usually writes $g \cdot x = R(g)x$ where $R(g) \in \mathrm{GL}(X)$. One easily sees that the set of linear transformations $R(g)$ satisfy $R(e) = I$ and $R(gh) = R(g)R(h)$, that is, one has a group homomorphism $G \to \mathrm{GL}(X)$. Reciprocally, given such a set of linear maps one has a group representation. Whenever X is finite dimensional of dimension n, one can by a choice of a fixed basis, represent each $R(g)$ as an $n \times n$

matrix $M(g) \in \mathrm{GL}(n, \mathbb{F})$. These matrices obviously satisfy $M(e) = I$ and $M(gh) = M(g)M(h)$. We say in this case that we have a *matrix representation* of G. A matrix representation is thus nothing more than a group homomorphism $G \to \mathrm{GL}(n, \mathbb{F})$.

Given two representations R_1 and R_2 of G on vector spaces X_1 and X_2 over the same field, a linear map $T : X_1 \to X_2$ is called an *intertwiner* for the two representations if

$$R_2(g)T = TR_1(g) \tag{7}$$

A similar notion holds for two matrix representation, in which case, T is an $n_2 \times n_1$ matrix and $M_i(g) \in \mathrm{GL}(n_i, \mathbb{F})$.

The intertwiner expresses the notion of a morphism of actions (1), in the context of linear spaces.

We say two representations are *equivalent* if they have an invertible intertwiner.

We say a representation R is *reducible* if there is a non-trivial subspace $Y \subset X$ invariant under R, that is, $R(g)Y \subset Y$ for all $g \in G$. For a matrix representation we can take $X = \mathbb{F}^n$ and the same definition applies. A representation that is not reducible is said to be *irreducible*.

Given a representation R of a group G in a vector space X one has a natural representation R^*, called the *dual representation* in the dual space X' defined, for a linear functional ϕ, by $(R^*(g)\phi)(x) = \phi(R(g^{-1})x)$, that is, $R^*(g) = R(g^{-1})'$.

When $G \subset \mathrm{GL}(X)$, there is a natural representation of G on X given by $(g, x) \mapsto gx$. When G is a matrix group, $G \subset \mathrm{GL}(n, \mathbb{F})$ then there are *two* natural matrix representations on $\mathbb{F}^n$ defined by the actions

$$(m, x) \mapsto mx \tag{8}$$
$$(m, x) \mapsto (m^t)^{-1}x \tag{9}$$

Note that representation (9) is equivalent to the dual of (8).

Let $\mathcal{L}$ be a Lie algebra over a base field $\mathbb{F}$. A *representation* of $\mathcal{L}$ in an associative algebra $\mathcal{A}$ over $\mathbb{F}$ is a map $r : \mathcal{L} \to \mathcal{A}$ such that $r([K, L]) = r(K)r(L) - r(L)r(K)$. Note that in the associative algebra $\mathcal{A}$, one can define a Lie bracket $[\cdot, \cdot]$ as the commutator $[a, b] = ab - ba$. With this definition one has $r([K, L]) = [r(K), r(L)]$. Particularly important cases are when $\mathcal{A}$ is $\mathrm{End}(X)$ for a vector space X over $\mathbb{F}$ and when $\mathcal{A}$ is $\mathcal{M}(n, \mathbb{F})$ the algebra of $n \times n$-matrices over $\mathbb{F}$.

Let now G be a Lie group and suppose X finite dimensional with the base field being either $\mathbb{R}$ or $\mathbb{C}$. Let $L \in \mathfrak{g}$. One can now define a representation $\mathfrak{r}$ of $\mathfrak{g}$ in $\mathrm{End}(X)$ by

$$\mathfrak{r}(L) = \frac{d}{dt} R(e^{tL})\Big|_{t=0} \tag{10}$$

One easily finds that $\mathfrak{r}$ is indeed a representation and that

$$\begin{aligned} R(g)\mathfrak{r}(L)R(g^{-1}) &= \mathfrak{r}(\mathrm{Ad}_g L) \\ R(\exp(L)) &= \exp(\mathfrak{r}(L)) \end{aligned}$$

3.4 Affine Actions

Let X be a vector space over a field $\mathbb{F}$. An *affine map* $A : X \to X$ is one of the form $Ax = Bx + a$ where B is linear and $a \in X$. Denote an affine map by the pair (B, a). The set of affine maps is an algebra over $\mathbb{F}$ with the linear structure given by $\alpha(B, a) + \beta(D, e) = (\alpha B + \beta D, \alpha a + \beta e)$ for $\alpha, \beta \in \mathbb{F}$, and the product by composition: $(B, a)(D, e) = (BD, a + Be)$. An affine map (B, a) is invertible if and only if B is invertible, in which case the inverse is $(B^{-1}, -B^{-1}a)$. The set of all invertible affine maps is called the *general affine group* of X and will be denoted by $\mathrm{GA}(X)$. For $X = \mathbb{F}^n$ we write $\mathrm{GA}(n, \mathbb{F})$ and for $X = \mathbb{R}^n$ we write $\mathrm{GA}(n)$. An *affine representation* of a group G is a homomorphism of G into $\mathrm{GA}(X)$ for some vector space X. The pairs $(B(g), a(g))$ then satisfy $B(gh) = B(g)B(h)$ and $a(gh) = a(g) + B(g)a(h)$. Note that B is then a representation of G as defined in Section 3.3.

4 Bundles

4.1 Fiber Bundles

A *fiber bundle* is a mathematical object with the following ingredients:

1. Three topological spaces: the *total space* E, the *base space* X, and the *fiber* F.

2. A map $\pi : E \to X$, the *projection*.

3. A covering $\mathcal{U}$ of X by a family $(U_\alpha)_{\alpha \in A}$, of open sets.

4. For each U_α in $\mathcal{U}$ a homeomorphism, the *local trivialization*,

$$h_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F$$

such that for U_β in $\mathcal{U}$ with $U_\alpha \cap U_\beta \neq \emptyset$ the map

$$h_\alpha \circ h_\beta^{-1} : (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F$$

is given by

$$h_\alpha \circ h_\beta^{-1}(x, f) = (x, h_{\alpha\beta}(x)(f)) \tag{11}$$

where

$$h_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Aut}(F)$$

and are called the *transition maps*.

Note that $h_{\alpha\beta}$ depends on the *ordered* pair (α, β). It's clear that these maps satisfy the following relations for all U_α in $\mathcal{U}$, for all pairs U_α, U_β in $\mathcal{U}$ such that $U_\alpha \cap U_\beta \neq \emptyset$, and for all triples $U_\alpha, U_\beta, U_\gamma$ in $\mathcal{U}$ such that $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$:

$$\begin{aligned} h_{\alpha\alpha}(x) &= \mathrm{Id}_F & (12)\\ h_{\beta\alpha}(x) &= h_{\alpha\beta}(x)^{-1} & (13)\\ h_{\alpha\beta}(x) \circ h_{\beta\gamma}(x) \circ h_{\gamma\alpha}(x) &= \mathrm{Id}_F & (14) \end{aligned}$$

These three relations are not independent as either (12) or (13) follows from the other two, but it is conceptually useful to write down all three.

For simplicity we shall sometimes use abbreviated expressions such as "fiber bundle E" or "fiber bundle E over X with fiber F" without specifying all of the data required by the definition. It is to be understood however that such data is always present.

The transition maps can be construed as *gluing instructions* by which the total space E is constructed by gluing together the cartesian products $U_\alpha \times F$.

In fact, let

$$\tilde{E} = \coprod_{\alpha \in A} U_\alpha \times F = \bigcup_{\alpha \in A} U_\alpha \times F \times \{\alpha\}$$

be the disjoint union and let $\sim$ be the equivalence relation generated by the equivalences

$$(x, f, \alpha) \sim (x, h_{\alpha\beta}(x)(f), \beta) \tag{15}$$

for all $x \in U_\alpha \cap U_\beta$ and all U_α, U_β in $\mathcal{U}$ with non-empty intersection. Let $E = \tilde{E}/\sim$ be the quotient space. We now have:

Theorem 5 *Let X and F be topological spaces, $\mathcal{U}$ a family $(U_\alpha)_{\alpha\in A}$ of open sets covering X, and $h_{\alpha\beta}$ maps satisfying (12–14) above, then the quotient space E constructed in the preceding paragraph is a fiber bundle satisfying the given data.*

The proof is utterly straightforward and is left to the reader.

Two very familiar examples of such gluing leads to the Moebius strip and to the tangent bundle of a manifold.

> **Example 2** *Consider the circle S^1 as the set of unimodular complex numbers, and let $\phi(x) = e^{i\pi x}$. Let $U_1 = \phi(-\frac{3}{4}, \frac{3}{4})$ and $U_2 = \phi(\frac{1}{4}, \frac{7}{4})$ be an open cover. The intersection $W = U_1 \cap U_2$ consists of two components $W_1 = \phi(-\frac{3}{4}, -\frac{1}{4})$, and $W_2 = \phi(\frac{1}{4}, \frac{3}{4})$. Let the fiber F be the interval $[-1, 1]$ and define the transition map $h_{21} : W \to End(F)$ to be $h_{21}(x)f = f$ for $x \in W_1$ and $h_{21}(x)f = -f$ for $x \in W_2$.*

It is easy to see that the resulting space E is the Moebius strip. This should be considered as the prototypical example of gluing together a fiber bundle, the general one is seen to be the result of gluing together many cartesian products "twisted" by the transition maps. If all the transition maps $h_{\alpha\beta}(x)$ are the identity then the construction of the previous theorem obviously leads to $E = X \times F$.

> **Example 3** *Let M be a manifold of dimension n, and consider an atlas $\mathcal{U}$ given by a family of open sets $(U_\alpha)_{\alpha\in A}$, along with coordinate functions $x^1_\alpha, \ldots, x^n_\alpha$ in each U_α. For $x \in U_\alpha$ a tangent vector is given by*
>
> $$v = \sum_j v^j_\alpha \frac{\partial}{\partial x^j_\alpha}$$
>
> *If $x \in U_\alpha \cap U_\beta$ then also*
>
> $$v = \sum_j v^j_\beta \frac{\partial}{\partial x^j_\beta}$$
>
> *where one has*
>
> $$v^j_\beta = \sum_k \frac{\partial x^j_\beta}{\partial x^k_\alpha} v^k_\alpha$$

We can use this last formula to define the transition functions for a bundle with fiber $\mathbb{R}^n$ and base space M. For $x \in U_\alpha \cap U_\beta$, and $f = (f^1, \dots, f^n) \in \mathbb{R}^n$ define

$$(h_{\beta\alpha}(x)f)^j = \sum_k \frac{\partial x_\beta^j}{\partial x_\alpha^k}(x) f^k$$

Property (13) follows from the inverse function theorem and (14) from the chain rule. The resulting bundle is the *tangent bundle* TM of the manifold, which, as a set, consists of all the tangent vectors at all points of M. The local trivialization map $h_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n$ is, for a vector v at a point $x \in U_\alpha$, given by $h_\alpha(v) = (x, (v_\alpha^1, \dots, v_\alpha^n))$. The tangent bundle is a manifold of dimension $2n$ with local coordinates in $\pi^{-1}(U_\alpha)$ being $(x_\alpha^1, \dots, x_\alpha^n, v_\alpha^1, \dots, v_\alpha^n)$.

We now extend the notion of local trivializations beyond the meaning that occurs in the definition. By a *local trivialization* in the extended sense we now mean a pair (W, h_W) consisting of an open subset $W \subset X$ and a homeomorphism $h_W : \pi^{-1}(W) \to W \times F$ such that for any U_α in $\mathcal{U}$ such that $U_\alpha \cap W \neq \emptyset$, the map $h_W \circ h_\alpha^{-1} : (U_\alpha \cap W) \times F \to (U_\alpha \cap W) \times F$ has the form

$$h_W \circ h_\alpha^{-1}(x, f) = (x, h_{W\alpha}(x)(f)) \tag{16}$$

where $h_{W\alpha} : U_\alpha \cap W \to \mathrm{Aut}(F)$. Each of the local trivializations h_α for U_α in $\mathcal{U}$ specified in the definition of course continues being a local trivialization in this extended sense. Note also that if (W_1, h_1) and (W_2, h_2) are two local trivializations in the extended sense, with $W_1 \cap W_2 \neq \emptyset$, then the transition map $h_2 \circ h_1^{-1} : (W_1 \cap W_2) \times F \to (W_1 \cap W_2) \times F$ also has the same form, to wit, $h_2 \circ h_1^{-1}(x, f) = (x, h_{21}(x)(f))$ where $h_{21} : W_1 \cap W_2 \to \mathrm{Aut}(F)$. This means that if we add to the family of defining local trivializations $((U_\alpha, h_\alpha))_{\alpha \in A}$ any family of local trivializations in the extended sense, we again have data defining a fiber bundle with a larger set of local trivializations. By definitions that follow shortly, this new bundle will be equivalent to the original one. In what follows we shall conventionally drop the adjective "local" from the expression "local trivialization", and speak simply of a "trivialization".

Given a bundle $\pi : E \to X$ we write $F_x = \pi^{-1}(\{x\})$ and call F_x the *fiber over* x. Each F_x is homeomorphic to the fiber F though not necessarily in any canonical way.

We shall in general deal with bundles defined over some fixed base set X. These form a category in which the morphisms, also called *bundle maps*, are maps $\phi : E_1 \to E_2$ such that the following diagram commutes:

$$\begin{tikzcd} E_1 \arrow[rr, "\phi"] \arrow[dr, "\pi_1"'] & & E_2 \arrow[dl, "\pi_2"] \\ & X & \end{tikzcd}$$

Note that this definition implies that ϕ maps fiber to fiber, $\phi(F_{1x}) \subset F_{2x}$. An isomorphism in this category is a map ϕ as above which is an isomorphism between the total spaces E_i. Isomorphic bundles are said to be *equivalent*.

We shall generally consider fiber-bundle theory as concerning bundles only up to isomorphism. With this notion we first establish isomorphism of bundles that merely change the system of trivializations. Suppose we have the complete data $(E, F, X, \pi, \mathcal{U}, (h_\alpha)_{\alpha\in A})$ of a fiber bundle as given in the definition. Let now $\mathcal{V}$ denote another family $(V_\lambda)_{\lambda\in\Lambda}$ of open sets covering X and $(k_\lambda)_{\lambda\in\Lambda}$ a family of corresponding trivialization homeomorphisms. It is easily seen that $(E, F, X, \pi, \mathcal{V}, (k_\lambda)_{\lambda\in\Lambda})$ provides a complete set of data for a fiber bundle as given by the definition. It is also easily seen that the identity map $\mathrm{Id}_E : E \to E$ is a bundle isomorphism. By this observation we can now free ourselves of the original defining set of trivializations and pass on to any other set defining an equivalent bundle. A particular case of this is to pass on to a *refinement* of $\mathcal{U}$, that is a cover $\mathcal{V}$ such that each U_α is a union of a subfamily of V_λ in $\mathcal{V}$. Consider now the cover by the family $V_{(\alpha,\lambda)} = V_\lambda$ indexed by the subset of pairs (α, λ) in $A \times \Lambda$ such that $V_\lambda \subset U_\alpha$. Define $h_{(\alpha,\lambda)}$ to be h_α restricted to V_λ. The new resulting covering family of open sets and trivializations defines an isomorphic bundle. With this in mind, if we now have a finite set of bundles over the same base space, we can, by choosing a common refinement of all of the covering families of the individual bundles, consider that all bundles have the same covering family of open sets over which the trivializations are defined.

Returning now to the notion of a bundle morphism, let $\mathcal{U}$ be a covering with respect to which both bundles trivialize. Since ϕ maps fiber to fiber one has the following commutative diagram for U_α in $\mathcal{U}$

$$
\begin{tikzcd}
\pi_1^{-1}(U_\alpha) \arrow[r, "h_\alpha^{(1)}"] \arrow[d, "\phi"'] & U_\alpha \times F_1 \arrow[d, "\mathrm{Id} \times \phi_\alpha"] \\
\pi_2^{-1}(U_\alpha) \arrow[r, "h_\alpha^{(2)}"] & U_\alpha \times F_2
\end{tikzcd}
$$

Here the right vertical arrow should be read as

$$\mathrm{Id} \times \phi_\alpha : (x, f) \mapsto (x, \phi_\alpha(x)(f))$$

where the maps $\phi_\alpha : U_\alpha \to \mathrm{Hom}(F_1, F_2)$ *represent* ϕ in the trivializations.

We now have $h_\alpha^{(2)} \circ \phi = (\mathrm{Id} \times \phi_\alpha) \circ h_\alpha^{(1)}$ from which

$$\phi = (h_\alpha^{(2)})^{-1} \circ (\mathrm{Id} \times \phi_\alpha) \circ h_\alpha^{(1)}$$

valid in $\pi_1^{-1}(U_\alpha)$. One has a similar expression for ϕ in $\pi_1^{-1}(U_\beta)$ for another open set $U_\beta \in \mathcal{U}$. If now $U_\alpha \cap U_\beta \neq \emptyset$ then one has

$$(h_\alpha^{(2)})^{-1} \circ (\mathrm{Id} \times \phi_\alpha) \circ h_\alpha^{(1)} = (h_\beta^{(2)})^{-1} \circ (\mathrm{Id} \times \phi_\beta) \circ h_\beta^{(1)}$$

valid in $\pi_1^{-1}(U_\alpha \cap U_\beta)$. From this we deduce

$$(\mathrm{Id} \times \phi_\alpha) \circ h_\alpha^{(1)} \circ (h_\beta^{(1)})^{-1} = h_\alpha^{(2)} \circ (h_\beta^{(2)})^{-1} \circ (\mathrm{Id} \times \phi_\beta)$$

By (11) this can now be easily expressed in terms of the transition maps

$$h_{\alpha\beta}^{(2)}(x)(\phi_\beta(x)(f)) = \phi_\alpha(x)(h_{\alpha\beta}^{(1)}(x)(f)) \tag{17}$$

for all $x \in U_\alpha \cap U_\beta$ and $f \in F_1$. One should note the resemblance of formula (17) to that of the morphism of group actions (1) and the intertwiner of group representations (7).

Reciprocally, we have the following theorem whose proof is utterly straightforward:

Theorem 6 *Let E_1 and E_2 be two fiber bundles with fibers F_1 and F_2 over the same base space X. Given a family $(U_\alpha)_{\alpha \in A}$ of open sets covering X over which both bundles trivialize, and given maps $\phi_\alpha : U_\alpha \to Hom(F_1, F_2)$ satisfying (17) then there is a unique bundle morphism $\phi : E_1 \to E_2$ for which the ϕ_α are the local representatives.*

We shall now adopt a convention by which we abbreviate (17) to

$$h^{(2)}_{\alpha\beta} \circ \phi_\beta = \phi_\alpha \circ h^{(1)}_{\alpha\beta} \tag{18}$$

In this convention, maps such as the ones appearing in (18) are considered as *parameterized maps*, that is a family of maps indexed by elements of some set ($U_\alpha \cap U_\beta$ in this particular case) and equation such as (18) is supposed to hold at each point of the indexing set. This translates to (17). Details of this convention are to found in Appendix B. This convention will be generally in force from now on.

In the particular case that ϕ is an isomorphism then F_1 and F_2 are isomorphic and we can identify the two and denote each by F. Each ϕ_α is invertible, belonging to $\mathrm{Aut}(F)$. In this case we can now write

$$h^{(2)}_{\alpha\beta} = \phi_\alpha \circ h^{(1)}_{\alpha\beta} \circ \phi_\beta^{-1} \tag{19}$$

We say a fiber bundle is *trivial* if it is equivalent to the *product bundle* $X \times F$ with one single defining trivialization $h_X = \mathrm{Id}_{X\times F}$. We see from (19) that a bundle is trivial if and only if its transition maps can be written as

$$h_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1}$$

for a family of maps $\phi_\alpha : U_\alpha \to \mathrm{Aut}(F)$.

> **Example 4** *Consider the trivial bundle* $E = [0,1] \times \mathbb{R}$ *with base space* $[0,1]$*, fiber* $\mathbb{R}$*, and the single identity map* $Id : E \to E$ *as the defining family of trivializations.*

We shall find it instructive to consider the maps h_i, $i = 1, 2, 3$ from E to E given by

$$\begin{aligned} h_1(x,f) &= (x, (1+x^2)f) \\ h_2(x,f) &= (x, 2x + (1+x^2)f) \\ h_3(x,f) &= (x, (1+x)f + f^3) \end{aligned}$$

each of these is a bundle isomorphism. This example call attention to the fact that though a bundle may be isomorphic to a product bundle, one should not think that certain relations that are natural to the cartesian product carry over to a trivial bundle. Thus while it is natural to think of the subset $[0,1] \times \{1\} \subset [0,1] \times \mathbb{R}$ as "horizontal" being of the form $\pi_{\mathbb{R}}^{-1}(\{1\})$ where

$\pi_{\mathbb{R}} : [0,1] \times \mathbb{R} \to \mathbb{R}$ is the canonical projection, in a trivial bundle there is in general no canonical projection on the fiber. Thus the image of the same set by the h_i are in no way "horizontal". Metaphorically speaking, a trivial bundle is a cartesian product which lost one of its canonical projections. Though a trivial bundle has a trivialization over the full base space, which could be called a "global" trivialization, there is in general no one canonical such.

> **Example 5** *Let G be a Lie group. The tangent bundle TG is trivial, a global trivialization $TG \to G \times \mathfrak{g}$ is given by the map that associates to $v \in T_gG$ the pair $(g, v \cdot g^{-1})$. Associating the pair $(g, g^{-1} \cdot v)$ defines a second trivialization.*

By a *local section* of a bundle we mean an open set $U \subset X$ and a map $\sigma : U \to E$ such that $\pi \circ \sigma(x) = x$ for all $x \in U$. We shall denote by $\Gamma(U)$ the set of local sections over U. A *global section* is a section in which $U = X$. A bundle may not have any global sections. A simple obvious example of this is the double cover $S^1 \to S^1$ which can be realized as the map $z \mapsto z^2$ of unimodular complex numbers. This is a bundle with fiber being a two point set.

Let $\sigma : U \to E$ be a local section of a bundle and consider a trivialization over an open set $W \subset U$. The map $h_W \circ \sigma : W \to W \times F$ then has the form

$$h_W \circ \sigma(x) = (x, s_W(x)) \tag{20}$$

The map $s_W : W \to F$ is called the *representative of* σ in the given trivialization. Local sections are therefore generalizations of (partial) maps of X to F. Local sections are also often know as *fields*. The value of a field at $x \in U$ being $\sigma(x) \in F_x$. Familiar examples are, for instance, local sections of TM known as *vector fields* or of T^*M, known as *covector fields* or 1-*forms*. Note that the value of a field in general lies in a space, the fiber F_x over the point x, that varies with the point. In a trivialization though, the representative of the section takes its value in a fixed space, the fiber F. Introducing local coordinates $y^1, \ldots, y^m$ in F, the representative can then be described by a set of *components* $s_W(x) = (s^1_W(x), \ldots, s^m_W(x))$. In the sequel we shall usually drop the adjective "local" and speak simply of a section. Global sections will be referred to as such.

We shall in the sequel often have to deal with how mathematical structures related to fiber bundles appear, or as is usually said, are *represented* in trivializations. For a trivialization h_W in an open set W, such local representatives should in principle carry some indication that identifies the trivialization (such as the index W on the map s_W in (20) that represents a local section σ). The systematic use of such indices would generally overburden the notation and we shall often drop them whenever one is dealing only with one trivialization. We shall of course be forced to introduce them any time we must compare local representatives in two trivializations over intersecting open sets. Thus for local representatives for σ one has

$$s_V(x) = h_{VW}(x)(s_W(x)) \tag{21}$$

Such *transition formulas* will be common in what follows. It should also be pointed out that although a trivialization consists of a *pair*, an open set W and a homeomorphism $h_W : \pi^{-1}(W) \to U \times F$, it is often customary to use only the open set W as a label. This is somewhat awkward as it is absolutely legitimate to consider two different trivializations over the same open set W, that is, to consider two different homeomorphisms. This should be kept in mind in understanding transition formulas such as (21), and interpret it also as expressing a relation for two trivializing homeomorphism over the same set.

4.2 G-bundles

Let E be a fiber bundle with base space X and fiber F. Let G be a group and suppose we have a (usually taken to be effective) left action of G on F and let $\rho : G \to \mathrm{Aut}(F)$ the the corresponding group homomorphism. Suppose $h_{\alpha\beta}(x) \in \rho(G)$ for all $x \in U_\alpha \cap U_\beta$ and all (α, β) with $U_\alpha \cap U_\beta \neq \emptyset$, then we say that G is the *structure group* of the bundle. This expression is a bit abusive as the structure group is not uniquely defined, and a bundle has in general many structure groups. One can of course always consider $\mathrm{Aut}(F)$ as the structure group. Just to what extent a structure group of a bundle is to be considered as part of its structure depends mostly on the application one has in mind. It is often useful to consider bundles with a given structure group G in mind, in which case one speaks of a *G-bundle*. This is particularly true in gauge theory.

To be more precise, a G-bundle with a given left action of G on F is one where the transition maps have the form $h_{\alpha\beta}(x)(f) = g_{\alpha\beta}(x) \cdot f$ and where the maps $g_{\alpha\beta} : U \cap V \to G$ are required to satisfy

$$
\begin{aligned}
g_{\alpha\alpha}(x) &= e & (22)\\
g_{\beta\alpha}(x) &= g_{\alpha\beta}(x)^{-1} & (23)\\
g_{\alpha\beta}(x)g_{\beta\gamma}(x)g_{\gamma\alpha}(x) &= e. & (24)
\end{aligned}
$$

These conditions automatically follow from properties (12-14) of transition maps if the action is effective. Condition (24) above is know as the (Čech) 1-*cocycle* condition. By abuse of language, we shall also refer to the maps $g_{\alpha\beta}$ as *transition maps*, or when convenient to specify the group, as G*-transition maps*.

When dealing with a G-bundle, we can relativize some of the notions defined above to reflect the fixed structure group. Thus by a trivialization of a G-bundle in an open set W we shall now mean one in which the maps $h_{W\alpha}$ in equation (16) are of the form $h_{W\alpha}(x)(f) = g_{W\alpha}(x) \cdot f$ for a map $g_{W\alpha} : U_\alpha \cap W \to G$. Likewise an isomorphism of two G-bundles with the same base-space and same fiber will be given by maps ϕ_α appearing in (19) of the form $\phi_\alpha(x)(f) = g_\alpha(x) \cdot f$ where $g_\alpha : U_\alpha \to G$. Equation (19) is now to be read in terms of group multiplication as

$$
g^{(2)}_{\alpha\beta} = g_\alpha g^{(1)}_{\alpha\beta} g_\beta^{-1} \tag{25}
$$

In Čech cohomology, a set of maps $g_\alpha : U_\alpha \to G$ is called a 0-*cochain with values in* G. Given such a cochain, the maps $g_\alpha g_\beta^{-1}$ are easily seen to satisfy conditions for being the G-transition maps of a G-bundle, in particular they define a 1-cocycle. This cocycle is called the *coboundary* of the cochain. Thus a G-bundle is trivial if and only if its cocycle is a coboundary. Similarly we can say that two bundles are equivalent if the cocycles are intertwined by a cochain.

Under the viewpoint of considering the structure group as part and parcel of fiber-bundle structure, isomorphism of bundles may now depend on the structure group chosen. The largest isomorphism classes of course correspond to $G = \mathrm{Aut}(F)$. In Example 4 if we take the structure group G to be $\mathrm{GL}(1)$ then h_1 is a G-bundle isomorphism while h_2 and h_3 are not. Expanding G to $\mathrm{GA}(1)$, makes h_2 a G-bundle isomorphism, while h_3 continues not being, and expanding G to $\mathrm{Diff}(\mathbb{R})$, the diffeomorphism group of $\mathbb{R}$, now includes h_3.

4.3 Structured Fibers

Very often the fiber F of a bundle has additional structure that is transferred to each individual fiber F_x. This happens when each of the transition maps $h_{\alpha\beta}$ is an isomorphism of the structure in F. In this section we shall be concerned only with certain algebraic structures, such as vector space and algebra. Other structures will appear in subsequent sections. Whenever F is a vector space and each transition map $h_{\alpha\beta}$ is a linear isomorphism then each fiber F_x is also canonically a vector space. This is because the equivalence classes defined by the gluing instructions (see Theorem 5 and (15)) are compatible with the linear operations in F, that is, given that $(x, f_1, \alpha) \sim (x, h_{\alpha\beta}(x)f_1, \beta)$ and $(x, f_2, \alpha) \sim (x, h_{\alpha\beta}(x)f_2, \beta)$ and that a_1 and a_2 are elements of the base field, then

$$(x, a_1 f_1 + a_2 f_2, \alpha) \sim (x, a_1 h_{\alpha\beta}(x) f_1 + a_2 h_{\alpha\beta}(x) f_2, \beta)$$

so that in F_x one can define

$$a_1[(x, f_1, \alpha)] + a_2[(x, f_2, \alpha)] = [(x, a_1 f_1 + a_2 f_2, \alpha)]$$

which makes F_x canonically into a vector space. A bundle with this structure is called a *vector bundle*. Note that the *zero section* which takes each $x \in X$ to the zero element of F_x is a well-defined global section. This global section, by abuse of notation, is usually denoted simply by 0. A vector G-bundle arises whenever the transition maps are of the form $h_{\alpha\beta} = R(g_{\alpha\beta})$ for a representation R of G and G-transition maps $g_{\alpha\beta}$.

A slightly weaker notion is that of an *affine bundle*, which arises whenever F is a vector space but the transition maps $h_{\alpha\beta}$ are affine isomorphisms. One now has $h_{\alpha\beta} f = B_{\alpha\beta} f + c_{\alpha\beta}$ where the $B_{\alpha\beta}$ are linear isomorphism of F and $c_{\alpha\beta} \in F$. In this case the $B_{\alpha\beta}$ are linear transition maps and the $c_{\alpha\beta}$ satisfy $B_{\alpha\beta} c_{\beta\gamma} + c_{\alpha\beta} = c_{\alpha\gamma}$, whenever the open sets corresponding to the indices have a non-empty intersection. Note that in this process, part of the structure of F is lost when going to the fibers F_x. Due to the presence of the terms $c_{\alpha\beta}$, the linear structure is weakened as there is no way to canonically identify the zero element of F_x, but linear relations between differences of elements of F_x continue to be well defined. Vector bundles are obviously special case of affine bundles, those for which $c_{\alpha\beta} = 0$ for all (α, β). An affine G-bundle would be one in which $h_{\alpha\beta} = H(g_{\alpha\beta})$ for an affine representation H of G and G-transition maps $g_{\alpha\beta}$.

Finally one has an *algebra bundle*, or, a *bundle of algebras*, whenever the fiber is an algebra $\mathcal{A}$ and each $h_{\alpha\beta}$ is an $\mathcal{A}$-isomorphism. In particular an algebra bundle is a vector bundle. An algebra G-bundle arises whenever there is a group homomorphism $\rho : G \to \mathrm{Aut}(\mathcal{A})$ and one has $h_{\alpha\beta} = \rho(g_{\alpha\beta})$ for G-transition maps $g_{\alpha\beta}$.

Obviously the above type of constructions can be extended to a wide variety of algebraic structure.

Let E be a vector bundle over X and $U \in X$ an open set. Let $\mathcal{F}(U)$ be the set of maps $U \to \mathbb{F}$ of U to the base field $\mathbb{F}$ of F. $\mathcal{F}(U)$ is naturally an $\mathbb{F}$-algebra under pointwise operations.

The set of local sections $\Gamma(U)$ is obviously a vector space under pointwise linear operations. One also defines a pointwise product of elements $f \in \mathcal{F}(U)$ and $\sigma \in \Gamma(U)$ by $(f\sigma)(x) = f(x)\sigma(x)$. If E is an algebra bundle, then $\Gamma(U)$ is in addition an algebra under pointwise multiplication.

4.4 Principal Bundles

Let G be a topological group. A very special kind of bundle with structured fibers is a *principal G-bundle* in which each fiber is a homogeneous space with free transitive action of G. Let there be given a topological space X, a family $\mathcal{U}$ of open sets that cover X, and a family of G-transition maps subordinate to $\mathcal{U}$. The group G acts on itself by left multiplication and so we can now use this action to construct, by gluing, a bundle with fiber G and structure group G using the given transition maps. This bundle is called a *principal G-bundle* normally denoted by PG. For a given base space the principal G-bundles are not necessarily unique as they depend on the chosen cocycle and there are in general many inequivalent ones.

> **Example 6** *There are two inequivalent principal $\mathbb{Z}_2$-bundles over S^1. One has $E = S^1 \coprod S^1$ with π being the identity on each copy. The other has $E = S^1$ as the double cover over S^1. Identifying S^1 with the unimodular complex numbers, π is the map $z \mapsto z^2$.*

A group acts upon itself also by multiplication on the right and this commutes with the left action. This right action thus is compatible with the gluing procedure dictated by the transition maps, that is the relation $(x, g, \beta) \sim (x, g_{\beta\alpha}(x)g, \alpha)$ is invariant under right multiplication of g by $h \in G$. Using this, there is a natural *right* action of G on PG. Under this action each fiber is a single orbit and the stability subgroup of any point is the trivial subgroup $\{e\}$. Thus G acts transitively and freely on each fiber.

Let now F be any topological space and ρ a left action of G on F. Consider the space $PG \times F$. This is easily seen to be a bundle with projection $\hat{\pi} : PG \times F \to X$ given by $\hat{\pi}(p, f) = \pi(p)$, with fiber $G \times F$, and the same cocycle as that of PG using the left action $(g, (h, f)) \mapsto (gh, f)$ of G on $G \times F$. Define now on $PG \times F$ another left action $\gamma(g, (p, f)) = g \cdot (p, f) = (p \cdot g^{-1}, g \cdot f)$. Denote by $PG \times_\rho F$ the space of orbits by this action and by $\pi_\rho : PG \times F \to PG \times_\rho F$ the canonical map to the quotient. It is easy to see that $\hat{\pi}$ factors trough π_ρ and we have a map $\tilde{\pi} : PG \times_\rho F \to X$.

Theorem 7 *$PG \times_\rho F$ is a G-bundle isomorphic to the bundle with fiber F, projection $\tilde{\pi}$, group action ρ, and the cocycle of PG.*

Proof: Consider the action map $\rho : G \times F \to F$ taking (g, f) into $g \cdot f$. Consider also on $G \times F$ the action $\eta(h, (g, f)) = h \cdot (g, f) = (gh^{-1}, h \cdot f)$ and let Q be the space of orbits under this action endowed with the quotient topology. Now $\rho(gh^{-1}, h \cdot f) = gh^{-1} \cdot (h \cdot f) = g \cdot f$ so ρ is constant on the orbits and thus factors through Q:

$$\begin{tikzcd} G \times F \arrow[r, "\rho"] \arrow[d, "\pi_\eta"] & F \\ Q \arrow[ur, "\sigma"'] & \end{tikzcd}$$

where π_η is the canonical map to the quotient.

Since ρ maps (e, f) to f, it is surjective, and thus so is σ. Suppose now that $g_1 \cdot f_1 = g_2 \cdot f_2$. One has $(g_2^{-1} g_1) \cdot (g_1, f_1) = (g_1(g_2^{-1} g_1)^{-1}, (g_2^{-1} g_1) \cdot f_1) = (g_1 g_1^{-1} g_2, g_2^{-1} \cdot (g_2 \cdot f)) = (g_2, f_2)$. So (g_1, f_1) and (g_2, f_2) are on the same orbit. Since one has $\sigma(\mathcal{O}_{(g,f)}) = g \cdot f$ this result means that σ is also injective and hence bijective. The continuity of σ follows from the universal property of the quotient topology. We now prove it is open. A subset $A \subset Q$ is open if and only if the union of orbits $W = \pi_\eta^{-1}(A)$ is an open subset of $G \times F$. If now $(g, f) \in W$ then $(e, g^{-1} \cdot f) \in W \cap (\{e\} \times F) = \{e\} \times W_e$ where W_e is an open set in F. It is now clear that

$$W = \bigcup_{f \in W_e} \mathcal{O}_{(e,f)}$$

and that $\sigma(A) = W_e$ so that indeed σ is open and thus a homeomorphism. Using this fact, we transfer, by (2), the action ρ on F to an equivalent action ρ_Q on Q.

We shall now show that $PG \times_\rho F$ is a G-bundle with fiber Q, with the group action ρ_Q, and the cocycle of PG. Since Q and F are homeomorphic and ρ and ρ_Q equivalent, this will prove the theorem.

Let h_α be a defining local trivializing homeomorphism of PG, then one has the map

$$h_\alpha \times \mathrm{Id} : \hat{\pi}^{-1}(U_\alpha) \to U_\alpha \times G \times F$$

which gives a trivialization of $PG \times F$. This map is an isomorphism of actions γ and $\mathrm{Id} \times \eta$ so one can pass to quotients and get a homeomorphism

$$\tilde{h}_\alpha : \pi_\rho^{-1}(U_\alpha) \to U_\alpha \times Q$$

which we take as a defining local trivialization. One now has the following diagram:

$$\begin{tikzcd}
\hat{\pi}^{-1}(U_\alpha \cap U_\beta) \arrow[r, "h_\alpha \times \mathrm{Id}"] \arrow[d, "\mathrm{Id}"] & U_\alpha \cap U_\beta \times G \times F \arrow[r, "\mathrm{Id} \times \pi_\eta"] \arrow[d, "\mathrm{Id} \times g_{\beta\alpha} \times \mathrm{Id}"] & U_\alpha \cap U_\beta \times Q \\
\hat{\pi}^{-1}(U_\alpha \cap U_\beta) \arrow[r, "h_\beta \times \mathrm{Id}"] & U_\alpha \cap U_\beta \times G \times F \arrow[r, "\mathrm{Id} \times \pi_\eta"] & U_\alpha \cap U_\beta \times Q
\end{tikzcd}$$

Note that $\rho(g_{\beta\alpha}(x)g, f) = (g_{\beta\alpha}(x)g) \cdot f = g_{\beta\alpha}(x) \cdot \rho(g, f)$ which, composing with σ^{-1}, implies $\pi_\eta(g_{\beta\alpha}(x)g, f) = g_{\beta\alpha}(x) \cdot \pi_\eta(g, f)$. Thus the above diagram can be completed with a rightmost downward vertical arrow $(x, q) \mapsto (x, g_{\beta\alpha}(x) \cdot q)$, establishing the proper transition map. This completes the proof. Q.E.D

With this construction we can now form all the bundles based on a given cocycle in a global fashion starting with the principal bundle PG. Such bundles are said to be *associated* to the principal bundle. We can also in this way extend certain constructions on the principal bundle to all its associated bundles in a systematic fashion.

Example 7 *Let M be an n-dimensional manifold and $\mathcal{U}$ and atlas. Let $U, V \in \mathcal{U}$ with $U \cap V \neq \emptyset$ and let $x^1, \dots, x^n$ and $y^1, \dots, y^n$*

be local coordinates in U and V respectively. One then has for $x \in U \cap V$ the Jacobian matrix

$$J^i{}_j(x) = \frac{\partial y^i}{\partial x^j}(x) \in GL(n) \tag{26}$$

As was seen in Example 3, these matrices satisfy the conditions for being the transition maps of a bundle. The principal $GL(n)$-bundle $\mathcal{F}(M)$ formed from this cocycle is called the frame bundle *of M.*

The reason for the name *frame bundle* is that one can identify the elements of each fiber F_x of the bundle with ordered bases $e_1, \ldots, e_n$ of T_xM. If we use the coefficients $\epsilon^i{}_j$ in the local expansion

$$e_i = \sum_k \epsilon^k{}_i \frac{\partial}{\partial x^k}$$

as local coordinates of the ordered base at x, then one has

$$e_i = \sum_k \sum_\ell J^k{}_\ell \epsilon^\ell{}_i \frac{\partial}{\partial y^k}$$

Now $\epsilon^i{}_j \in \mathrm{GL}(n)$ and $\epsilon^i{}_j \mapsto \sum_\ell J^k{}_\ell \epsilon^\ell{}_i$ is a left action of $J \in \mathrm{GL}(n)$ on $\epsilon \in \mathrm{GL}(n)$ by ordinary matrix multiplication. From this one sees that indeed the set of frames is a principal $\mathrm{GL}(n)$-bundle with cocycle J. One has the right action of $\mathrm{GL}(n)$ on the bundle of frames by having $A^i{}_j \in \mathrm{GL}(n)$ transform the frame $e_1, \ldots, e_n$ to the frame $f_1, \ldots, f_n$ by $f_j = \sum_i e_i A^i{}_j$. Note that we were able to define the right action without recourse to local coordinates.

Consider now the actions by the two natural representations $\rho(A, z) = Az$ and $\rho^*(A, z) = (A^t)^{-1} z$ of $\mathrm{GL}(n)$ on $\mathbb{R}^n$. One has the identifications:

$$\begin{aligned} \mathcal{F}(M) \times_\rho \mathbb{R}^n &\simeq TM \\ \mathcal{F}(M) \times_{\rho^*} \mathbb{R}^n &\simeq T^*M \end{aligned}$$

For a bundle to be associated to a principal bundle is not very special. All that is necessary, by Theorem 7, is to interpret the transition maps as due to the action of a topological group. We shall only need this for vector bundles E with a finite dimensional fiber F. One has $h_{\alpha\beta} \in \mathrm{GL}(F)$ and one can construct the corresponding principal bundle $P\mathrm{GL}(F)$. The action $\rho : \mathrm{GL}(F) \times F \to F$ is the natural one $(T, f) \to Tf$. As a corollary of Theorem 7 one has

Theorem 8

$$PGL(F) \times_\rho F \simeq E$$

4.5 Bundle Operations

Consider a family of fiber bundles $(E_\lambda)_{\lambda\in\Lambda}$ over the same base space X and with fibers F_λ. It is often possible to extend an operation that produces a new space F from the F_λ to act fiber-wise on the spaces E_λ to produce a new fiber bundle E with fiber F. This is the case whenever one can trivialize all the E_λ over the open sets of the same covering family and then combine the transition maps $h^\lambda_{\alpha\beta}$ of each bundle E_λ to transition maps $h_{\alpha\beta}$ for the space F. The first step is always possible if the family is finite, and we shall only deal with this case in these notes. So assume there is a fixed family of open sets $(U_\alpha)_{\alpha\in A}$ covering X and let h^λ_α and $h^\lambda_{\alpha\beta}$ be respectively the trivialization and the transition maps of the bundles E_λ.

A simple example of such a construction is the cartesian product $F = \prod_\lambda F_\lambda$. Define a new bundle E with transition maps $h_{\alpha\beta} = \prod_\lambda h^\lambda_{\alpha\beta}$. This bundle is called the *product bundle* and one writes $E = \prod_\lambda E_\lambda$. Note that this notation is misleading as the total space E is *not* the cartesian product of the total spaces E_λ. The product bundle is the product object in the category of bundles over a fixed base space. The canonical categorical projections $p_\mu : E \to E_\mu$ are bundle maps whose local representatives $U_\alpha \times \prod_\lambda F_\lambda \to U_\alpha \times F_\mu$ are $(x, f) \mapsto (x, f_\mu)$. Theorem 6 then provides a bundle map. A special case of this construction is when each E_λ is a principal G_λ-bundle, $E_\lambda = PG_\lambda$. It is easy to see that $E = \prod_\lambda E_\lambda$ is a principal $\prod_\lambda G_\lambda$ bundle.

For vector bundles with the fibers being vector spaces over the same base field one can form the direct sum, also known as the Whitney sum, and the tensor product of bundles. The direct sum has fiber $F = \bigoplus_\lambda F_\lambda$ with transition maps $h_{\alpha\beta} = \bigoplus_\lambda h^\lambda_{\alpha\beta}$ and the tensor product has fiber $F = \bigotimes_\lambda F_\lambda$ with transition maps $h_{\alpha\beta} = \bigotimes_\lambda h^\lambda_{\alpha\beta}$. The corresponding bundles are denoted by $\bigoplus_\lambda E_\lambda$ and $\bigotimes_\lambda E_\lambda$ respectively. Notations such as $E_1 \oplus E_2$, or $E_1 \otimes E_2 \otimes E_3$ are also used and are self-explanatory. As a topological fiber bundle, $\bigoplus_\lambda E_\lambda$ coincides with $\prod_\lambda E_\lambda$. It is the algebraic structure that justifies a different notation. For a fixed vector bundle E one can also form the exterior powers $\bigwedge^p(E)$ defined by the transition maps $\bigwedge^p(h_{\alpha\beta})$.

Another construct for vector bundles E_1 and E_2 is to consider $F = \mathrm{Hom}(F_1, F_2)$, the space of linear maps $F_1 \to F_2$. Define for $\phi \in \mathrm{Hom}(F_1, F_2)$

the transition formula $h_{\alpha\beta}\phi = h^2_{\alpha\beta} \circ \phi \circ {h^1_{\alpha\beta}}^{-1}$. These are the transition maps for the vector bundle $\mathrm{Hom}(E_1, E_2)$ of linear homomorphisms.

Consider now F', the dual of F. For $\phi \in F'$ define

$$h'_{\alpha\beta}\phi = \phi \circ {h_{\alpha\beta}}^{-1} \tag{27}$$

These are the transition maps of the bundle E', the dual bundle. This is a particular case of $\mathrm{Hom}(E_1, E_2)$ where for E_2 we take the trivial bundle $X \times \mathbb{F}$ where $\mathbb{F}$ is the base field of the fiber.

For the case of G-bundles a simplification occurs whenever all the bundles use the same G-transition maps and only differ in the action of G defined on each fiber F_λ. Provided these actions can be combined to an action on F, the resulting bundle continues to be a G-bundle for the same group. Thus for vector bundles, one can form direct sums and tensor products of group representations and so the Whitney sum and the tensor product of vector G-bundles can again be considered as being G-bundles. The same is true of the bundle of linear homomorphisms, the dual bundle, and the exterior powers.

Starting with the tangent bundle TM, the dual bundle is the cotangent bundle T^*M. Tensor products of these two bundles leads to the usual tensor bundles. The exterior powers $\bigwedge^p(T^*M)$ are the familiar bundles of exterior p-covectors whose local sections are p-forms. All these bundles can be construed as $\mathrm{GL}(n)$-bundles. In Section 5.1 we shall see how this group may be reduced.

If E_i for $i = 1, \ldots, n$ are vector bundles with fibers F_i, and V is a vector bundle with fiber W, all defined over the same base field, then a bundle map $\alpha : E_1 \times \cdots \times E_n \to V$ is said to be *multilinear* or *n-linear* if each fiberwise restriction $F_{1x} \times \cdots F_{nx} \to W_x$ of α is an n-linear map. Such bundle maps obviously form a vector space.

If all the E_i are the same bundle E with fiber F, then the natural S_n (permutation group) action on the n-fold cartesian product $F \times \cdots \times F$ gives rise, by Theorem 6, to an action on $E \times \cdots \times E$. We say an n-linear map $\alpha : E \times \cdots \times E \to V$ is *antisymmetric* if $\alpha(\pi \cdot p) = \sigma_\pi \alpha(p)$ for $\pi \in S_n$ where σ_π is ± 1 depending on whether π is even or odd.

4.6 Connections

Consider a fiber bundle $\pi : E \to M$ with fiber F and in which all spaces are manifolds. At each point $p \in E$ there is a canonical subspace $V_pE \subset T_pE$

of the tangent space at p to E called the *vertical subspace* consisting of all vectors v such that $d\pi\, v = 0$. These are vectors tangent to the fiber $F_{\pi(p)}$, which is a submanifold. On the other hand, there is in general no canonical way of choosing a complementary "horizontal" subspace (see the discussion following Example 4). By a *connection* on E we mean a smooth choice (to be shortly explained) at each $p \in E$ of a subspace $H_pE \subset T_pE$ complementary to V_pE. Given such a connection then at each point $p \in E$ we have projections π_p^v and π_p^h onto the vertical and horizontal subspaces respectively. We shall assume that the distribution H_pE varies in a $\mathcal{C}^\infty$ way with p, which means that for any $\mathcal{C}^\infty$ vector field $\mathcal{X}$ on E, the horizontally projected field $\pi^h\mathcal{X}$ is likewise $\mathcal{C}^\infty$.

Let us examine this in a trivialization. One has $T_{(x,f)}(U \times F) \simeq T_xU \times T_fF$. Let $(\xi, y) \in T_xU \times T_fF$. Such a vector is vertical if and only if $\xi = 0$. Thus $\pi^v_{(x,f)}$ must have the form

$$\pi^v_{(x,f)}(\xi, y) = (0, y + \Gamma(x, f)\xi)$$

where $\Gamma(x, f)$ is a *linear* map $T_xU \to T_fF$. From this it follows that

$$\pi^h_{(x,f)}(\xi, y) = (\xi, -\Gamma(x, f)\xi)$$

It is useful to calculate the change in the connection map $\Gamma(x, f)$ in a different trivialization. In this case define the map

$$\psi(x, f) = (x, h_{VU}(x)(f))$$

The differential of this map must intertwine the local representatives of π^v, the projection maps on the vertical tangent subspaces, specifically, $d\psi_{(x,f)} \circ \pi^v_{(x,f)} = \pi^v_{\psi(x,f)} \circ d\psi_{(x,f)}$. One has

$$dh_{VU}(x)(f)(\xi, y) = (\xi, d_1h_{VU}(x)(f)\xi + d_2h_{VU}(x)(f)y)$$

The intertwining relation with π^v gives us

$$d_1h_{VU}(x)(f) + \Gamma_V(x, h_{VU}(x)(f)) = d_2h_{VU}(x)(f)\Gamma_U(x, f)$$

which can be solved for

$$\begin{aligned} \Gamma_V(x, f) \quad &= \quad d_2h_{VU}(x)(h^{-1}_{VU}(x)(f))\Gamma_U(x, h^{-1}_{VU}(x)(f)) \\ &\quad\quad -d_1h_{VU}(x)(h^{-1}_{VU}(x)(f)) \end{aligned} \tag{28}$$

Reciprocally, one can use (28) to define a connection.

Theorem 9 *Let $\pi : E \to X$ be a bundle with fiber F. Suppose we have a set of trivializations $h_U : \pi^{-1}(U) \to U \times F$ such that the open sets U cover X. Suppose that for each trivialization and each $(x, f) \in U \times F$ we have linear maps $\Gamma_U(x, f) : T_xU \to T_fF$ varying smoothly with respect to (x, f) and such that relation (28) is satisfied for each pair of trivializations such that $U \cap V \neq \emptyset$. Then there is a unique connection of the bundle for which the representative with respect to the given trivializations are the given $\Gamma_U(x, f)$.*

The proof is entirely straightforward.

In the case of a G-bundle, one has $h_{UV}(x)(f) = g_{UV}(x) \cdot f = \alpha(g_{UV}(x), f)$ where for convenience we have explicitly introduced the action α of G on F. Equation (28) now becomes

$$\begin{aligned} \Gamma_V(x, f) &= d_2\alpha(g_{VU}(x), g_{VU}^{-1}(x) \cdot f)\Gamma_U(x, g_{VU}^{-1}(x) \cdot f) \\ &\quad -d_1\alpha(g_{VU}(x), g_{VU}^{-1}(x) \cdot f)dg_{VU}(x) \end{aligned} \tag{29}$$

where dg_{VU} is the differential of the map $g_{VU} : U \cap V \to G$.

We now specialize to vector bundles. In this case one has a canonical identification $V_pE \simeq F_{\pi(p)}$ as follows: If $q \in F_{\pi(p)}$ then it makes sense to consider the curve $\gamma(t) = p + tq \in F_{\pi(p)}$ as this space is a vector space. Identify $\gamma'(0) \in V_pE$ with $q \in F_{\pi(p)}$. We shall say that a connection is *linear* if π_p^h varies affinely with p in a fixed fiber F_x, and the canonical zero section of E is horizontal.

In a locally trivialized vector bundle using the mentioned identification $T_{(x,f)}(U \times F) \simeq T_xU \times T_fF \simeq T_xU \times F$ for a connection to be linear one must have $\Gamma(x, f)\xi = \Gamma(x)(\xi)f$ where $\Gamma(x)$ can now be interpreted as a linear map $T_xU \to \mathrm{End}(F)$, in other words an $\mathrm{End}(F)$-valued 1-form on U. Such *vector-valued* forms are explained in Appendix B.

We now have

$$\pi^h_{(x,f)}(\xi, w) = (\xi, -\Gamma(x)(\xi)f) \tag{30}$$

$$\pi^v_{(x,f)}(\xi, w) = (0, w + \Gamma(x)(\xi)f) \tag{31}$$

Using the parameterized map convention of Appendix B, (28) now reads

$$\Gamma_V = h_{VU}\Gamma_U h_{VU}^{-1} - dh_{VU}h_{VU}^{-1} \tag{32}$$

This has a form analogous to a *gauge transformation*, which shall be defined in section 4.7.

Another useful expression is that of (29) for a linear connection on a vector G-bundle. The action is now $g \cdot f = R(g)f$ where R is a group representation. Using the conventions introduced in section 3.2, equation (6), we have $d_1\alpha(g,f)L \cdot g^{-1} = \tilde{d}_1\alpha(g,f)L = \mathfrak{r}(L)f$ where $L \in \mathfrak{g}$. Also $d_2\alpha(g,f)y = R(g)y$. From these we find, again using the parameterized map convention of Appendix B, that

$$\Gamma_V = R(g_{VU})\Gamma_U R(g_{VU})^{-1} - \mathfrak{r}(dg_{VU} \cdot g_{VU}^{-1}) \tag{33}$$

Just as (32) this has the form of a *gauge transformation*, defined in section 4.7.

Let now G be a Lie group and PG a principal G-bundle over an n-dimensional manifold M. In this case one has a canonical identification $V_pPG \simeq \mathfrak{g}$ as follows. Let $p \in PG$ and $L \in \mathfrak{g}$. Consider the curve $\gamma(t) = p \cdot \exp(tL)$ in PG. Identify $L_p = \gamma'(0) \in V_pPG$ with $L \in \mathfrak{g}$. With this identification, the vertical projection π_p^v is represented for every $p \in PG$ by a map $\omega_p : T_pPG \to \mathfrak{g}$ which is thus a $\mathfrak{g}$-*valued* 1-*form* on PG, which we shall call the *connection* 1-*form*. Such a 1-form obviously satisfies

$$\omega_p(L_p) = L \tag{34}$$

A connection is said to be *invariant* if the horizontal subspaces H_p satisfy $H_{p\cdot g} = dR_g(H_p) = H_p \cdot g$, where R_g is the right action map $p \mapsto p \cdot g$ on PG. This means that if at p one has $v = v_V \oplus v_H$, a decomposition of a tangent vector into the vertical and horizontal components, then $v \cdot g = v_V \cdot g \oplus v_H \cdot g$ is the decomposition at $p \cdot g$.

One has

$$L_p \cdot g = \frac{d}{dt} p \cdot \exp(tL)g \bigg|_{t=0} = \frac{d}{dt} p \cdot g(g^{-1}\exp(tL)g) \bigg|_{t=0} = (\mathrm{Ad}_{g^{-1}}L)_{p\cdot g}$$

so the action of dR_g on the spaces V_p, when transferred to $\mathfrak{g}$, translates to $\mathrm{Ad}_{g^{-1}}$. Now for an invariant connection, one must have $\pi_{p\cdot g}^v(v \cdot g) = \pi_p^v(v) \cdot g$, which by what was said above means

$$\omega_{p\cdot g}(v \cdot g) = \mathrm{Ad}_{g^{-1}}\omega_p(v) \tag{35}$$

Let now ω be an invariant connection 1-form on PG. Let h_U be a trivialization map on an open set U. We shall analyze the structure of $\omega^U = (h_U^{-1})^*(\omega)$ on $U \times G$. Since the trivialization map commutes with right action, the image form also corresponds to an invariant connection. One has the identification

$$T_{(x,h)}(U \times G) \simeq T_xU \times T_hG \simeq T_xU \times \mathfrak{g}$$

where of course the identification of T_hG with $\mathfrak{g}$ is via $dR_{h^{-1}}$. Assuming this identification for the time being, consider $v = (\xi, L) \in T_xU \times \mathfrak{g}$. One has $v \cdot g = (\xi, \mathrm{Ad}_{g^{-1}}L)$, and by (34) $\omega^U_{(x,h)}(\xi, L) = \omega^U_{(x,h)}(\xi, 0) + L$. By (35) one has

$$\omega^U_{(x,h)}(\xi, h^{-1} \cdot L \cdot h) = h^{-1} \cdot \omega^U_{(x,e)}(\xi, L) \cdot h = h^{-1} \cdot (\omega^U_{(x,e)}(\xi, 0) + L) \cdot h$$

so that $\omega^U_{(x,h)}(\xi, 0) = h^{-1} \cdot A_U(x)(\xi) \cdot h$ where A_U is a $\mathfrak{g}$-valued 1-form on U. One finally has has

$$\omega^U_{(x,h)}(\xi, L) = h^{-1} \cdot (A_U(\xi) + L) \cdot h \tag{36}$$

In the trivialized bundle therefore an invariant connection is represented by the $\mathfrak{g}$-valued 1-form A_U on U, which we shall call the *local principal gauge potential*. Note that in doing so we have passed from an object defined on the total space PG to one defined on the base space M. As in physics M is generally space-time, this is in keeping with the notion that physical field theory deals with objects defined directly on space-time. In the physical literature, for a matrix group G, what is know as the *gauge potential* arises from A_U after a choice of a local chart, as will be explained in Section 6.1.

We will now find the relation between A_U and A_V in $U \cap V$, corresponding to two trivializations. Let $\psi_{VU} = h_V \circ h_U^{-1}$. One has

$$\omega^V_{(x,g_{VU}h)}(\tilde{d}\psi_{VU}(\xi, L)) = \omega^U_{(x,h)}(\xi, L) \tag{37}$$

where by $\tilde{d}\psi_{VU}$ we mean the translation of $d\psi_{VU}$ under the identification of T_hG with $\mathfrak{g}$. Now $\psi_{VU}(x, h) = (x, g_{VU}(x)h)$. We can then write $d\psi_{VU}(\xi, v) = (\xi, g_{VU} \cdot v + dg_{VU}\xi \cdot h)$. One has $v = L \cdot h$ for some $L \in \mathfrak{g}$, which gives $d\psi_{VU}(\xi, L \cdot h) = (\xi, g_{VU} \cdot L \cdot h + dg_{VU}\xi \cdot h)$. The vertical component is at the point $(x, g_{VU}(x)h)$ and we must bring it to $\mathfrak{g}$ via right action. One thus has: $\tilde{d}\psi_{VU}(\xi, L) = (\xi, g_{VU} \cdot L \cdot g_{VU}^{-1} + dg_{VU}\xi \cdot g_{VU}^{-1})$ Using this and equations (36) and (37), one deduces the transformation law of the local principal gauge potentials:

$$A_V = g_{VU} \cdot A_U \cdot g_{VU}^{-1} - dg_{VU} \cdot g_{VU}^{-1} = \mathrm{Ad}_{g_{VU}} A_U - dg_{VU} \cdot g_{VU}^{-1} \tag{38}$$

We shall see in section 4.7 that this has the form of a *gauge transformation.*

Just as before, we can use (38) to define an invariant connection on a principal G-bundle.

Theorem 10 *Let $\pi : PG \to X$ be a principal G-bundle. Suppose we have a set of trivializations $h_U : \pi^{-1}(U) \to U \times G$ such that the open sets U cover X. Suppose that for each trivialization we have a $\mathfrak{g}$-valued 1-form A_U such that relation (38) is satisfied for each pair of trivializations such that $U \cap V \neq \emptyset$. Then there is a unique invariant connection in PG for which the representatives with respect to the given trivializations are the given A_U.*

Again, the proof is entirely straightforward. It is often through such local representatives that an invariant connection on a principal bundle is defined.

Given a Lie group G and a principal G-bundle over a manifold M, let F be a manifold that carries a left action ρ of G. Let $PG \times_\rho F$ be the associated bundle with map $\pi_\rho : PG \times F \to PG \times_\rho F$. If now ω is an invariant connection on PG, there is a canonical induced connection on $PG \times F$ given by $H_{(p,w)}(PG \times F) = H_p PG \times \{0\}$. This distributions of tangent subspaces is obviously invariant by the action $(p, f) \mapsto (p \cdot g^{-1}, g \cdot f)$ used in the construction of the associated bundle and so descends to $PG \times_\rho F$ by the differential of π_ρ, providing us with the *induced connection* in $PG \times_\rho F$.

Consider now the case in which F is a vector space and ρ is given by a representation $g \cdot f = R(g) f$ of G. Let $\mathfrak{r}$ be the corresponding representation of $\mathfrak{g}$ (see Section 3.2). In a trivialization, π_ρ is represented by $\mathrm{Id} \times \rho : U \times G \times F \to U \times F$. The horizontal subspace at $(x, f) \in U \times F$ is the image by the differential of $\mathrm{Id} \times \rho$ of the horizontal subspace at (x, e, f). Now $(\xi, L, y) \in T_x U \times \mathfrak{g} \times F$ is horizontal at (x, e, f) if and only if $y = 0$ and $A(x)(\xi) + L = 0$, furthermore, the differential of $\mathrm{Id} \times \rho$ at this point acts as $(\xi, L, y) \mapsto (\xi, \mathfrak{r}(L) f + y)$. One now sees that the horizontal subspace at (x, f) consists of vectors of the form $(\xi, -\mathfrak{r}(A(x)(\xi)) f)$. As the second component is linear in f we see that the induced connection is linear. One also has the following expression for the projection onto the vertical subspace

$$\pi^v_{(x,f)}(\xi, y) = y + \mathfrak{r}(A(x)(\xi)) f \tag{39}$$

There is something like a reciprocal to this construction. Let E be a vector bundle and γ a linear connection. By Theorem 8, E is naturally an $\mathrm{End}(F)$-bundle. Equations (33) and (38) are now identical and by Theorem

10 one has an invariant connection on $P\mathrm{End}(F)$ which by (39) induces the original connection γ on E. This situation will always be assumed for a vector bundle if the structure group in not explicitly indicated.

Consider now bundle operations as discussed in Section 4.5. If each bundle E_λ has a connection, it is often possible to combine them in a canonical fashion to define a connection in E. We shall only need this for the case of the product bundle $\prod_\lambda E_\lambda$. One defines the *product connection* by stipulating that $v \in H_pE$ if and only if $dp_\lambda v \in H_{p_\lambda(p)}E_\lambda$. It is easy to see that the product of invariant connections on principal G_λ-bundles is an invariant connection on the principal $\prod G_\lambda$-bundle, which is the product.

4.7 Gauge Transformations

Let PG be a principal G-bundle over a base space X. A *gauge transformation* is a bundle isomorphism $\phi : PG \to PG$ that commutes with right action. Because of the free right action of G on PG we can write $\phi(p) = p \cdot \gamma(p)$ for a unique function $\gamma : PG \to G$. For this to commute with right action it is necessary and sufficient that $(p \cdot g) \cdot \gamma(p \cdot g) = (p \cdot \gamma(p)) \cdot g$ and so

$$\gamma(p \cdot g) = g^{-1}\gamma(p)g \tag{40}$$

Thus the set of all gauge transformations is the set of all maps $\gamma : PG \to G$ satisfying (40). The set of all gauge transformations of PG is called the *gauge group* of PG and we denote it by $\mathcal{G}(PG)$.

When X is a manifold and G a Lie group, one can consider $\mathcal{G}(PG)$ as (generally) an infinite dimensional Lie group. One can discover what one should take to be the Lie algebra of the gauge group by considering a one-parameter family $\gamma_t(p)$ of gauge transformations of the form $\gamma_t(p) = \exp(t\theta(p))$ where $\theta(p) \in \mathfrak{g}$. Property (40) for all t translates to

$$\theta(p \cdot g) = \mathrm{Ad}_{g^{-1}}\theta(p) \tag{41}$$

Maps $\theta : PG \to \mathfrak{g}$ satisfying (41) are taken to constitute the Lie algebra of $\mathcal{G}(PG)$ and are called *infinitesimal gauge transformations*. The Lie bracket of two infinitesimal gauge transformations $[\theta_1, \theta_2]$ is calculated pointwise $[\theta_1, \theta_2](p) = [\theta_1(p), \theta_2(p)]$. Since the adjoint action of G in $\mathfrak{g}$ commutes with the Lie bracket one sees that $[\theta_1, \theta_2]$ satisfies (41) and so indeed is likewise an infinitesimal gauge transformation.

In a trivialization the map ϕ is represented by

$$\mathrm{Id} \times \phi_U : U \times G \to U \times G$$

acting as $(x, g) \mapsto (x, \phi_U(x)(g))$. Because of commutativity with right action by G one has $\phi_U(x)(g) = \phi_U(x)(eg) = \phi_U(x)(e)g$ so ϕ_U is expressible through a local map $\phi_U^0 : U \to G$ as $\phi_U(x)(g) = \phi_U^0(x)g$. The map γ in a trivialization is expressed by a map $\mathrm{Id} \times \gamma_U : U \times G \to G$ satisfying $\gamma_U(x)(gh) = h^{-1}\gamma_U(x)(g)h$ hence $\gamma_U(x)(h) = h^{-1}\gamma_U(x)(e)h = h^{-1}\gamma_U^0(x)h$ and γ also is expressible through a function $\gamma_U^0 : U \to G$. From $\phi_U(x)(g) = g\gamma_U(x)(g)$ one in fact concludes that $\gamma_U^0 = \phi_U^0$ and so the map γ can be locally suppressed. We shall from now on drop the superscript 0 and write simply ϕ_U.

From the above discussion and (25) it is easy to see that the gauge transformation changes the transition maps g_{UV} to the equivalent maps $\phi_U g_{UV} \phi_V^{-1}$. Note that there are no restrictions on the maps ϕ_U which can be an arbitrary 0-cocycle with values in G. In fact, from the above discussion we can state

Theorem 11 *Let $\pi : PG \to X$ be a principle G-bundle. Suppose we have a set of trivializations $h_U : \pi^{-1}(U) \to U \times G$ such that the open sets U cover X. Suppose that for each such open set we have a maps $\phi_U : U \to G$, then there is a unique gauge transformation $\phi : PG \to PG$ for which the representative with respect to the given trivializations are the given ϕ_U.*

Let now ω be an invariant connection on PG. A gauge transformation ϕ acts on ω on the left by push-forward. The new connection $\phi \cdot \omega$ being defined by

$$(\phi \cdot \omega)_{\phi(p)}(d\phi\, v) = \omega_p(v) \tag{42}$$

In section 4.6 we have seen that in a trivialization, an invariant connection is represented by the local principal gauge potential A which is a $\mathfrak{g}$-valued 1-form on U. We shall now calculate how a gauge transformation acts on these local potentials. Identifying once again $T_{(x,g)}U \times G$ with $T_xU \times \mathfrak{g}$, one has

$$(\phi \cdot A)_U(x)(\xi) = (\phi \cdot \omega)^U_{(x,e)}(\xi, 0) = \omega^U_{(x,\phi(x)^{-1})}(\tilde{d}\phi_U^{-1}(\xi, 0))$$

where by $\tilde{d}\phi_U^{-1}$ we mean the differential of the map $(x, g) \mapsto (x, \phi_U^{-1}(x)g)$ with the tangent spaces of G identified with $\mathfrak{g}$ through right action. Taking into account these identifications we see that $\tilde{d}\phi_U^{-1}(\xi, 0) = (\xi, d\phi_U^{-1}(x)\xi \cdot \phi_U(x))$ where in this formula $d\phi_U^{-1}$ means the differential of the map $\phi_U^{-1} : U \to G$. From this and (36) one has

$$(\phi \cdot A)(x)(\xi) = \mathrm{Ad}_{\phi_U(x)}(A(x)(\xi) + d\phi_U^{-1}(x)(\xi) \cdot \phi_U)$$

Now $\phi_U(x)\phi_U^{-1}(x) = e$, so $d\phi_U(x)(\xi) \cdot \phi_U(x)^{-1} + \phi_U(x) \cdot d\phi_U^{-1}(x)(\xi) = 0$ and we finally have the formula for a gauge transformation of the local principal gauge potential.

$$(\phi \cdot A) = \phi_U \cdot A \cdot \phi_U^{-1} - d\phi_U \cdot \phi_U^{-1} \tag{43}$$

It is useful to call attention to formulas (32), (33) and (38) which show that the transition formulas for the representatives of a connection in two different trivializations have the same abstract form as a gauge transformation.

It is also useful to calculate the effect of an infinitesimal gauge transformation on the local potential. Let the gauge transformation ϕ_t be defined through $\phi_{t,U}(x) = \exp(t\theta_U(x))$ and let δA_U be the coefficient of t in the Taylor expansion of $\phi_t \cdot A$ in (43). One easily calculates that

$$\delta A = [\theta, A] - d\theta$$

We shall not need this formula in these notes, but it is very often used in the physical literature.

4.8 Parallel Transport

We are in the category of manifolds. Let E be a fiber bundle with base space M, and fiber F. Consider a connection on E with vertical tangent space projectors π_p^v. A smooth curve in E is said to be *horizontal* if at each point its tangent lies in the horizontal tangent space of that point. To be horizontal is to be an integral curve of a differential equation. In fact consider a parameterized smooth curve in E and consider its image $e(t) = (x(t), f(t))$ in a trivialization $U \times F$. One has $\pi^v(e'(t)) = (0, f'(t) + \Gamma(x(t), f(t))x'(t))$ and so the condition for being horizontal is

$$f'(t) + \Gamma(x(t), f(t))x'(t) = 0 \tag{44}$$

This in local coordinates is an ordinary differential equation for $e(t)$. Note that only the components $f(t)$ are required to obey a differential equation and that $x(t)$ can be freely given with arbitrary parameterization. This allows us to determine $f(t)$ from $x(t)$. Given a smooth curve C in M a *horizontal lifting* of C is a horizontal curve $\tilde{C}$ in E such that $\pi(\tilde{C}) = C$. In a trivialization this means that once C is parameterized, then $\tilde{C}$ satisfies (44) inheriting a parameterization from that of C. By the existence, uniqueness, and regularity theorems for solutions of ordinary differential equations, any smooth curve in M has at least a local unique horizontal lifting passing through any point $f \in F_x$ for any $x \in C$.

Let C be a smooth curve in M with initial point x_0 and end point x_1. Let $f_0 \in F_{x_0}$, and assume that there is a global horizontal lifting of C with initial point f_0. The endpoint $f_1 \in F_{x_1}$ of $\tilde{C}$ is called the *parallel transport of* f_0 *along* C. It is obviously unique if it exists. If the parallel transport exists for all $f_0 \in F_{x_0}$ then the map $f_0 \mapsto f_1$ defines a diffeomorphism $F_{x_0} \to F_{x_1}$.

For a vector bundle with a linear connection, equation (44) has the form $f'(t) + \Gamma(x(t))(x'(t))f(t) = 0$ which is a *linear* equation. Thus parallel transport is always globally defined and the parallel transport map $F_{x_0} \to F_{x_1}$ is a linear isomorphism.

It is useful to have explicit forms for the parallel transport map. This is given by a construction known as time-ordered exponential integrals. Let F be a finite dimensional vector space and consider the following non-autonomous differential equation in F:

$$\frac{df}{dt} = A(t)f \tag{45}$$

where $A(t)$ is a linear operator which is a $\mathcal{C}^\infty$ function of t. By the existence, uniqueness, and regularity theorem for the solution of ordinary differential equations, for any $f \in F$, there is a unique solution $f(t)$ with $f(a)$ given. In differential equation theory, one generally introduces what is known as the *fundamental solution* of (45), that is, an $\mathrm{End}(F)$-valued function $W(t,a)$ which satisfies

$$\frac{\partial}{\partial t}W(t,a) = A(t)W(t,a) \tag{46}$$

$$W(a,a) = I \tag{47}$$

One now has

$$f(t) = W(t,a)f(a)$$

In physical literature one however often sees the solution as given by the *time-ordered exponential integral*, to be explained below:

$$f(t) = T\exp\left(\int_a^t A(s)\,ds\right)f(a) \tag{48}$$

The operator in front of $f(a)$ is not literally the exponential of an integral but a symbolic way of expressing a limit of a Riemann product, analogous to a Riemann sum. Let $[a, b]$ be an interval and partition it as $a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b$, let $\Delta_i t = t_i - t_{i-1}$, chose $t'_i \in [t_{i-1}, t_i]$, and consider the product

$$\begin{aligned} & T \prod_{i=1}^{N} \exp(A(t'_i)\Delta_i t) \\ = \ & \exp(A(t'_N)\Delta_N t) \exp(A(t'_{N-1})\Delta_{N-1} t) \cdots \exp(A(t'_1)\Delta_1 t) \end{aligned}$$

Note that the order of the factors in this product is important as the various linear operators $A(t)$ do not necessarily commute with each other. The chosen order is that of decreasing values of the t_i as one goes through the product from left to right. This is called *time order* and the symbol T symbolizes this choice. One has by definition

$$T \exp\left(\int_a^b A(t)\, dt\right) = \lim_{N\to\infty} T \prod_{i=1}^{N} \exp\left(A(t'_i)\Delta_i t\right) \tag{49}$$

where the limit is taken in the same sense as the one that defines a Riemann integral. The limit of course is nothing more than the fundamental solution $W(b, a)$, thus one has the obvious composition property for time-ordered exponential integrals: if $a < b < c$, then

$$T \exp\left(\int_a^c A(t)\, dt\right) = T \exp\left(\int_b^c A(t)\, dt\right) T \exp\left(\int_a^b A(t)\, dt\right) \tag{50}$$

The time-ordered exponential integral is the non-commutative analog of the continuous product of a function. Let $f(t)$ be a *positive* function on an interval $[a, b]$ and define the *continuous product*

$$\prod_a^b f(t)^{dt} = \lim_{N\to\infty} \prod_{i=1}^{N} f(t'_i)^{\Delta_i t}$$

with the limit understood as in the previous paragraph. One easily shows that if $\ln f$ is Riemann integrable in $[a, b]$, then

$$\prod_a^b f(t)^{dt} = \exp\left(\int_a^b \ln f(t)\, dt\right)$$

and so the continuous product reduces to an ordinary integral, a continuous sum, and for this reason there is no separately developed theory for continuous products. In the non-commutative case however one cannot reduce continuous products to integrals and a separate treatment is necessary. To complete the identification note that in particular if $f(t) = \exp(a(t))$ then

$$\prod_a^b f(t)^{dt} = \exp\left(\int_a^b a(t)\,dt\right)$$

Thus $T\exp\left(\int_a^b A(t)\,dt\right)$ is rightly thought of as the continuous product of $\exp(A(t))$ ordered with time decreasing left to right.

Let now A be an $\mathrm{End}(F)$-valued 1-form on a manifold M and let C be a smooth oriented curve in M. One can also define the *path-ordered exponential integral*

$$P\exp\left(\int_C A\right) = \lim_{N\to\infty} P\prod_{i=1}^N \exp\left(\int_{C_i} A\right) \tag{51}$$

where the curve C has been partitioned into *successive* arcs $C_1, \ldots C_N$ each one inheriting its orientation from C. The limit is to be understood in relation to a fixed parameterization of C with the maximum parameter length of the C_i tending to zero. Such a path-ordered exponential integral is an element of $\mathrm{End}(F)$.

It is useful to note that in products (49) and (51) one can replace the factor $\exp\left(A(t_i')\Delta_i t\right)$ by $I + A(t_i')\Delta_i t$ and respectively $\exp\left(\int_{C_i} A\right)$ by $I + \int_{C_i} A$, and obtain the same limits. This has the advantage of not relying on the existence of the exponential.

Though the time-ordered exponential integral is a purely formal expression, it suggests the following expansion

$$W(b,a) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_a^b \int_a^b \cdots \int_a^b T(A(t_1)A(t_2)\cdots A(t_n))\,dt_1\,dt_2\cdots dt_n \tag{52}$$

where, by convention, the $n = 0$ term is I, and $T(A(t_1)A(t_2)\cdots A(t_n))$ means the product of the $A(t_i)$ in order of decreasing times. Thus

$$T(A(t_1)A(t_2)) = \begin{cases} A(t_2)A(t_1) & \text{if} \quad t_1 < t_2 \\ A(t_1)A(t_2) & \text{if} \quad t_1 > t_2 \end{cases}$$

The right-hand side of (52) indeed does converges to $W(b,a)$ and is known as the *Dyson series*. Because of the time-ordering instruction T, each term in fact is a sum of $n!$ equal contributions, each one an integral over a simplex

$$\{(t_1, t_2, \ldots, t_n) \,|\, t_{\pi(1)} \geq t_{\pi(2)} \geq \cdots \geq t_{\pi(n)}\}$$

where π is a permutation of $\{1, 2, \ldots, n\}$. We finally get

$$W(b,a) = \sum_{n=0}^{\infty} \int_a^b \int_a^{t_1} \int_a^{t_2} \cdots \int_a^{t_{n-1}} A(t_1)A(t_2)\cdots A(t_n)\, dt_n \cdots dt_2\, dt_1 \qquad (53)$$

where, again by convention, the $n = 0$ term is I. This expansion can be obtained directly by applying Picard's method to the fundamental solution equations (46-47). Note that the existence of the exponential is not needed to define each term of (53) which makes it useful as a formal series in contexts in which A(t) belongs to an algebra for which the exponential is not defined.

We can now use path-ordered exponential integrals to express the effect of parallel transport. In a trivialization of a vector bundle the parallel transport equation reads

$$\frac{df}{dt} = -\Gamma(x(t))(x'(t))f$$

where $x(t)$ is a parameterized path C in U with initial point $x(0)$ and final point $x(1)$. Thus we have

$$f(1) = T\exp\left(-\int_0^1 \Gamma(x(t))(x'(t))\, dt\right) f(0) = P\exp\left(-\int_C \Gamma\right) f(0)$$

and so $P\exp\left(-\int_C \Gamma\right)$ is the parallel transport operator for the oriented curve C.

Finally, let G be a Lie group with Lie algebra $\mathfrak{g}$ and A a $\mathfrak{g}$-valued 1-form on an open set U in a manifold M. One can, for a curve contained in U, likewise define

$$P\exp\left(\int_C A\right) = \lim_{N\to\infty} P \prod_{i=1}^{N} \exp\left(\int_{C_i} A\right) \qquad (54)$$

which results in an element of G.

Such a path exponential integral solves the parallel transport equation for a principal G-bundle with an invariant connection. In fact, consider a trivialization $U \times G$ of such a bundle and a parameterized curve $p(t) = (x(t), g(t))$ in it. The tangent vector at $p(t)$ is $(x'(t), g'(t)) \in T_{x(t)}U \times T_{g(t)}G$. With $T_{g(t)}G$ identified with $\mathfrak{g}$ by right action we represent this tangent vector now by $(x'(t), g'(t)\cdot g(t)^{-1}) \in T_xU \times \mathfrak{g}$. By (36) such a vector is horizontal if and only if

$$g(t)^{-1} \cdot (A(x'(t)) + g'(t)\cdot g(t)^{-1}) \cdot g(t) = 0$$

that is, if and only if $A(x'(t)) + g'(t) \cdot g(t)^{-1} = 0$. Thus the parallel transport equation in a principal G-bundle with an invariant connection is

$$\frac{dg}{dt} = -A(x(t))(x'(t)) \cdot g \tag{55}$$

It is now easy to see that

$$g(t) = T \exp\left(-\int_0^t A(x(s))(x'(s))\, dt\right) g(0)$$

In fact, one has by (50) that

$$g(t+r) = T \exp\left(-\int_t^{t+r} A(x(t))(x'(t))\, ds\right) g(t)$$

and differentiating this with respect to r at $r = 0$ gives (55). Thus the effect of parallel transporting a group element g over an oriented curve C in U is to multiply it on the left by

$$P \exp\left(-\int_C A\right)$$

This result also shows that parallel transport is globally defined as the path-ordered exponential integral exists for any compact oriented curve with end points. The proof of this is analogous to the proof of the existence of Riemann integrals of continuous functions on compact intervals.

We now investigate how the local representatives of parallel transport depend on the trivialization. Let C be an oriented curve in M with initial point x_0 and final point x_1. Suppose parallel transport along C is defined for all $f_0 \in F_{x_0}$. Let $T : F_{x_0} \to F_{x_1}$ be the parallel transport diffeomorphism $f_0 \mapsto f_1$. If C lies in an open set U in which the bundle is trivialized, then T is represented by a diffeomorphism $T_U : F \to F$. If now V is another such open set and if $\tilde{C}_U$ is the horizontal lifting of C in the bundle trivialized over U, then $\tilde{C}_V = (\mathrm{Id} \times h_{VU})(\tilde{C}_U)$ is the lifting in the bundle trivialized over V. We thus have $T_V(h_{VU}(x_0)(f_0)) = h_{VU}(x_1)(f_1) = h_{VU}(x_1)(T_U(f_0))$ from which

$$T_V = h_{VU}(x_1) \cdot T_U \cdot h_{VU}(x_0)^{-1} \tag{56}$$

Note that this transformation is point-wise, that is, it does not involve the differential of h_{VU}.

4.9 Curvature

We are in the category of manifolds. Consider a bundle $\pi : E \to M$ with fiber F and a connection Γ. Let $x \in M$, and $p \in F_x$. Remember that parallel transport of p is always uniquely defined locally along curves passing through x. We can choose local charts $U \subset M$, $W \subset V \subset F$ with local coordinates $x^1, \ldots, x^n$ in U and $f^1, \ldots, f^m$ in V such that the bundle trivializes over U, and given a smooth curve C in U with end points x_0 and x_1 and any $f \in W$, then parallel transport of f exists over C and lies in V. Now parallel transport in general depends on the path C joining the two points, and in particular if $x_1 = x_0$ it may not be the identity. The *curvature* of the connection is a measure of by how much the transport depends on C, or equivalently by how much it differs from the identity along closed curves. A more precise statement of this, which we shall not prove, is the theorem that states that the curvature is zero if and only if parallel transport is the same along homotopy equivalent paths with the same end points. Parallel transport may still depend on the path even with curvature zero if there is more than one homotopy class of paths. Connections with zero curvature are called *flat*.

In local coordinates the equation of parallel transport is:

$$\frac{df^a}{dt} = -\sum_i \Gamma^a{}_i(x(t), f(t)) \frac{dx^i}{dt}$$

Let us choose numbers $\xi^1, \ldots, \xi^n$ and $\eta^1, \ldots, \eta^n$ such that the coordinates $x^i + t\xi^i + s\eta^i$ define points in U for $t, s \in [0, 1]$. Consider the piece-wise smooth curve C defined by the border of this rectangle passing successively through the vertices $x_0^i = x^i$, $x_1^i = x^i + \xi^i$, $x_2^i = x^i + \xi^i + \eta^i$, $x_3^i = x^i + \eta^i$, and back to x_0. We shall now calculate the effect of parallel transport along C to second order in ξ and η.

Now to second order in ξ along the first leg, from x_0 to x_1, parallel transport can be represented by the map

$$T_\xi^{(0)} : (x^i, f^a) \mapsto (x^i + \xi^i, f^a - \sum_i \Gamma^a{}_i(x, f)\xi^i + \sum_{ijb} \frac{\partial \Gamma^a{}_i}{\partial f^b} \Gamma^b{}_j \xi^i \xi^j)$$

We must now compute, retaining only term up to second order in ξ and η, the quantity $T_{-\eta}^{(3)} \circ T_{-\xi}^{(2)} \circ T_\eta^{(1)} \circ T_\xi^{(0)}$ where the superscript in $T^{(k)}$ means that all functions must be evaluates at the vertex point x_k. This is a straightforward though tedious calculation and we find that the result of this parallel transport is

$$(x^i, f^a) \mapsto (x^i, f^a + \sum_{ij} R^a{}_{ij}(x,f)\eta^i\xi^j)$$

where

$$R^a{}_{ij} = \frac{\partial \Gamma^a{}_j}{\partial x^i} - \frac{\partial \Gamma^a{}_i}{\partial x^j} + \sum_b \frac{\partial \Gamma^a{}_i}{\partial f^b}\Gamma^b{}_j - \sum_b \frac{\partial \Gamma^a{}_j}{\partial f^b}\Gamma^b{}_i \tag{57}$$

In terms of the non-trivialized bundle, the quantity R at a point $p \in E$ is seen to be a bilinear, anti-symmetric map $T_{\pi(p)}M \times T_{\pi(p)}M \to V_pE$. For this reason, R is referred to as the *curvature* 2*-form*.

One can use equation (57) to calculate the curvature of a linear connection on a vector bundle or of an invariant connection on a principal bundle. In these cases however it is instructive to use the path-ordered exponential integrals. For this we must recall the Baker-Campbell-Hausdorff formula. Let a and b be elements of a an associative algebra. One has, as formal power series in a and b:

$$\exp(a)\exp(b) = \exp\left(a + b + \frac{1}{2}[a,b] + []^2 + []^3 + \cdots\right) \tag{58}$$

where by $[]^k$ we mean a sum of terms each one of which is a k-fold nested bracket of a and b, for example $[]^2 = \frac{1}{12}([a,[a,b]] + [b,[b,a]])$. Although exact expressions for the $[]^k$ are known, we shall not need them explicitly beyond the terms already shown. The same formula holds for a and b being elements of the Lie algebra $\mathfrak{g}$ of some Lie group G where the exponentials are now actual group elements.

For a vector bundle we can take $V = W = F$, and denote now by $\Gamma = \sum_i \Gamma_i \, dx^i$ the $\mathrm{End}(F)$-valued connection 1-form. Consider once again parallel transport along the rectangle C used above. For the first leg of this we have the following parallel transport operator

$$P\exp\left(-\int_{C_1} \Gamma\right) \tag{59}$$

where C_1 is the first segment of the path, from x_0 to x_1. This to second order in ξ is

$$T^{(0)}_\xi = \exp\left(-\Gamma(x_0)(\xi) + (D\Gamma(x_0)\cdot\xi)(\xi)\right)$$

where by $D\Gamma \cdot \xi$ we mean the 1-form

$$\sum_{ij} \frac{\partial \Gamma_i}{\partial x^j}\xi^j \, dx^i$$

As before, up to second order in ξ and η, we must calculate $T^{(3)}_{-\eta}T^{(2)}_{-\xi}T^{(1)}_{\eta}T^{(0)}_{\xi}$, where again the superscript in $T^{(k)}$ means that all functions must be evaluates at the vertex point x_k. This can be done in a straightforward though tedious manner using the Baker-Campbell-Hausdorff formula and the result, to second order is $\exp(R(\eta,\xi))$ where

$$R(\eta,\xi) = d\Gamma(\eta,\xi) + [\Gamma(\eta),\Gamma(\xi)] \tag{60}$$

We see this defines an $\mathrm{End}(F)$-valued 2-form called the *curvature 2-form*.

From (56) one has $\exp(R_V(\eta,\xi) = h_{VU}(x_0)\exp(R_U(\eta,\xi))h_{VU}(x_0)^{-1} = \exp(h_{VU}(x_0)R_U(\eta,\xi)h_{VU}(x_0)^{-1})$ and so the curvature 2-form has the following transition formula

$$R_V = h_{VU}R_U h_{VU}^{-1} \tag{61}$$

We see on the right hand side the transition map of the bundle $\mathrm{End}(E) = \mathrm{Hom}(E,E)$ (see Section 4.5). This means that R can be viewed as an antisymmetric bilinear bundle map $TM \times TM \to \mathrm{End}(E)$, that is an $\mathrm{End}(E)$-valued 2-form.

For the case of a principal G-bundle one has the group element

$$P\exp\left(-\int_{C_1} A\right)$$

instead of (59). Proceeding in an entirely analogous manner, using again the Baker-Campbell-Hausdorff formula, we conclude that up to second order in ξ and η, parallel transport around C corresponds to multiplication by the group element $\exp(F(\eta,\xi))$ where

$$F(\eta,\xi) = dA(\eta,\xi) + [A(\eta),A(\xi)] \tag{62}$$

defining thus a $\mathfrak{g}$-valued 2-form in U called the *curvature 2-form*. It is customary to write (62) as

$$F = dA + [A,A]$$

Similar to (61), the transition formula for F is

$$F_V = g_{VU}\cdot F_U \cdot g_{VU}^{-1} \tag{63}$$

Given that G acts on $\mathfrak{g}$ by adjoint action $L \mapsto Ad_g L$ there is an associated bundle $P\mathfrak{g} = PG \times_{Ad} \mathfrak{g}$. We see on the right-hand side of (63) the transition map of $P\mathfrak{g}$ and so F can be considered as an antisymmetric bilinear bundle map $TM \times TM \to P\mathfrak{g}$, that is a $P\mathfrak{g}$-valued 2-form.

In the literature one often sees formulas slightly different from (60) and (62). Sometimes there is a factor of $\frac{1}{2}$ in front of the second term. This is due to adopting a different convention concerning alternating forms and exterior products as explained in Section A. Sometimes there is a difference in relative sign. This is due to conventionally using, in certain definitions, the negative of what is our connection 1-form. The reader beware.

If now $\phi : PG \to PG$ is a gauge transformation, then ϕ takes horizontal curves with respect to an invariant connection ω to horizontal curves with respect to the connection $\phi \cdot \omega$. Now in a trivialization, ϕ is represented by left multiplication by the map $\phi_U(x)$ (see Section 4.7), exactly as in a change of trivialization. This observation, along with (63) means that we have the following gauge transformation law for the curvature:

$$F \mapsto \phi_U \cdot F \cdot \phi_U^{-1} = \mathrm{Ad}_{\phi_U} F \tag{64}$$

Similarly we easily seen that the effect of an infinitesimal gauge transformation (see last paragraph of Section 4.7) is

$$\delta F = [\theta, F]$$

This last equation is very frequently used in the physical literature, but we shall not need it in these notes.

4.10 Covariant Derivatives

Consider a fiber bundle $\pi : E \to M$ with fiber F and structure group G in which all spaces are manifolds. Suppose we are given a connection on E with projections π_p^v on the vertical tangent subspaces. Recall that a local section of E on an open set $U \subset M$ is a map $\sigma : U \to E$ such that $\pi \circ \sigma(x) = x$. Let ξ be a tangent vector at $x \in U$. By the *covariant derivative of* σ *at* x *in direction* ξ we mean

$$\nabla_\xi \sigma(x) = \pi_{\sigma(x)}^v d\sigma_x \, \xi$$

One sees that $\nabla_\xi \sigma(x) \in V_{\sigma(x)} E$. Recall that for a vector bundle one has a canonical identification $V_p E \simeq F_{\pi(p)}$. We shall assume this identification and consider exclusively vector bundles with F finite dimensional. One should consider $\nabla_\xi \sigma(x)$ as the analog of the notion of a directional derivative of a bundle section.

One may wonder why one could not simply calculate the ordinary directional derivative of a section σ. This however is not well defined. Let $x(t)$ be a curve passing through x at $t = 0$ with $x'(0) = \xi$. The directional derivative should intuitively correspond to

$$\lim_{t\to 0} \frac{\sigma(x(t)) - \sigma(x(0))}{t}$$

The trouble with this is that $\sigma(x(t))$ and $\sigma(x(0))$ lie in different fibers and there is no well defined way of calculating their difference. If we have a connection though, we can parallel transport $\sigma(x(t))$ back along the given curve to the fiber over $x(0)$, and then compute the difference and the limit. This procedure defines the covariant derivative as an easy exercise shows.

To get further insight into this situation, trivialize the bundle in an open set U with local coordinates $x^1, \ldots, x^n$ and introduce a basis $f_1, \ldots, f_n$ of F. A section in U now is represented by a function $s(x) = s^i(x)f_i$ of U to F. Consider the partial derivative $\frac{\partial s}{\partial x^j} = \sum_i \frac{\partial s^i}{\partial x^j} f_i$, which would be a naive "directional derivative". In another trivialization over the same open set one has another representative $\tilde{s}(x) = \sum_i \tilde{s}^i(x) f_i$ where $\tilde{s}^i(x) = \sum h^i{}_k(x)s^k(x)$, and where the $h^i{}_k$ represent the transition map. Write this in matrix form as $\tilde{s} = hs$. One has

$$\frac{\partial \tilde{s}}{\partial x^j} = \frac{\partial h}{\partial x^j} s + h \frac{\partial s}{\partial x^j}$$

If the first term were absent, one would have that $\frac{\partial s}{\partial x^j}$ would likewise be a section of the same bundle and so partial derivatives (and thus directional derivatives) would be well defined as operators on sections. This is not so, but from the fact that the first term is linear in s one can try to construct a new section by an expression of the form

$$\frac{\partial s}{\partial x^j} + L_j s$$

where $L_j(x) \in \mathrm{End}(F)$. Imposing a transformation law on L so that

$$\frac{\partial \tilde{s}}{\partial x^j} + \tilde{L}_j \tilde{s} = h\left(\frac{\partial s}{\partial x^j} + L_j s\right)$$

one finds

$$\tilde{L}_j = hL_j h^{-1} - \frac{\partial h}{\partial x^j} h^{-1}$$

Comparing this to (32) one sees that L must define a connection in the bundle. Connections thus arise naturally once one tries to introduce differential calculus for sections of bundles.

If now $\mathcal{X}$ is a vector field in U one can calculate $\nabla_{\mathcal{X}(x)}\sigma(x)$ for each $x \in U$, this gives us a new local section $\nabla_{\mathcal{X}}\sigma$ of E. Trivializing over U, a section σ is represented by a map $\mathrm{Id} \times s : U \to U \times F$ and a linear connection by an $\mathrm{End}(F)$-valued 1-form Γ. At point $x \in U$, the differential of $\mathrm{Id} \times s$ acts as $\xi \mapsto (\xi, ds_x\,\xi)$ so $\nabla_\xi\sigma(x)$ is represented by:

$$ds_x\,\xi + \Gamma(x)(\xi)s(x) \tag{65}$$

where ds is the differential of s. If $f_1, \ldots, f_m$ is a basis for F, then $s(x) = \sum_{j=1}^m s^j(x) f_i$ and $ds_x\,\xi = \sum_{j=1}^m \xi(s^j)(x) f_i$ so for convenience we shall write $ds\,\xi$ as $\xi(s)$. Note that $\xi(s) \in F$. For a vector field $\mathcal{X}$ one has:

$$\nabla_{\mathcal{X}} s = \mathcal{X}(s) + \Gamma(\mathcal{X})s \tag{66}$$

Let W be a vector space and consider maps from open sets $U \subset M$ to W. One can view these as local sections of the cartesian product $M \times W$ considered as a trivial bundle with the identity map as the one defining trivialization. One now has the canonical identification $T_{(x,w)}(M \times W) \simeq T_x M \times W$ and can define the horizontal subspace as $H_{(x,w)} = \{0\} \times W$. In this trivialization, $\Gamma(x) = 0$ and so

$$\nabla_{\mathcal{X}} s = \mathcal{X}(s)$$

In particular for real or complex valued partial functions f on M we shall always take $\nabla_{\mathcal{X}} f = \mathcal{X}(f)$.

The map $\mathcal{X} \mapsto \nabla_{\mathcal{X}}$ fails to be a Lie algebra homomorphism and this fact is related to the existence of curvature. It is instructive to calculate

$$\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}} - \nabla_{\mathcal{Y}}\nabla_{\mathcal{X}} - \nabla_{[\mathcal{X},\mathcal{Y}]}$$

One has

$$\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}} s = \mathcal{X}(\mathcal{Y}(s)) + \mathcal{X}(\Gamma(\mathcal{Y}))s + \Gamma(\mathcal{Y})\mathcal{X}(s) + \Gamma(\mathcal{X})\mathcal{Y}(s) + \Gamma(\mathcal{X})\Gamma(\mathcal{Y})s$$

Taking into account that $\mathcal{X}(\mathcal{Y}(s)) - \mathcal{Y}(\mathcal{X}(s)) = [\mathcal{X},\mathcal{Y}](s)$ and, by (122), that $\mathcal{X}(\Gamma(\mathcal{Y})) - \mathcal{Y}(\Gamma(\mathcal{X}) - \Gamma([\mathcal{X},\mathcal{Y}]) = d\Gamma(\mathcal{X},\mathcal{Y})$ one concludes that

$$(\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}} - \nabla_{\mathcal{Y}}\nabla_{\mathcal{X}} - \nabla_{[\mathcal{X},\mathcal{Y}]})s = R(\mathcal{X},\mathcal{Y})s \tag{67}$$

where R is the curvature 2-form. Note that $R(\mathcal{X},\mathcal{Y})$ at point x depends only on the values $\mathcal{X}(x)$ and $\mathcal{Y}(x)$ and not on how these are extended to vector fields $\mathcal{X}$ and $\mathcal{Y}$.

Recall that an invariant connection on a principal bundle induces a linear connection on all the associated vector bundles, and consequently a covariant derivative on each such bundle. Recall also by the discussion preceding Theorem 8 that any vector bundle can be considered an associated bundle of $P\mathrm{End}(F)$ and any linear connection on the bundle as one induced from an invariant connection on the principal bundle. Furthermore given any finite family E_i, $i = 1, \dots, n$ of vector bundles with fibers F_i one can consider them all as being associated to a single principal bundle PG where $G = \mathrm{End}_1(F_1) \times \cdots \times \mathrm{End}(F_n)$ and whose action on F_i is projection to the i-th factor followed by the natural action of $\mathrm{End}(F_i)$. With this in mind one can interpret the discussion that follows as also pertaining to vector bundles and linear connections even when no explicit structure group is indicated.

The covariant derivatives induced from a fixed principal bundle onto its associated vector bundles have natural properties with respect to various bundle constructions. We shall consider in particular tensor products and bundles of linear homomorphisms, and deduce what may be called the Leibniz rule for the associated connections. Let therefore PG be a fixed principal G-bundle and E an associated vector bundle with corresponding representation R of G. By (39) and (66) one sees that in a trivialization

$$\nabla_{\mathcal{X}} s = \mathcal{X}(s) + \mathfrak{r}(A(\mathcal{X}))s \tag{68}$$

If now E_1 and E_2 are two associated vector bundles defined by representations R_1 and R_2 of G, one easily sees that $E_1 \otimes E_2$ is the associated bundle defined by the tensor product representation $R = R_1 \otimes R_2$. The associated Lie algebra representation, as can be easily deduced from (10), is given by $\mathfrak{r}(L) = \mathfrak{r}_1(L) \otimes I + I \otimes \mathfrak{r}_2(L)$. Let $\nabla^{(1)}$ and $\nabla^{(2)}$ be the covariant derivatives in E_1 and E_2 respectively and ∇ the covariant derivative in E. One deduces

$$\nabla_{\mathcal{X}} s_1 \otimes s_2 = \mathcal{X}(s_1 \otimes s_2) + \mathfrak{r}_1(A(\mathcal{X}))s_1 \otimes s_2 + s_1 \otimes \mathfrak{r}_2(A(\mathcal{X}))s_2$$

which, as $\mathcal{X}(s_1 \otimes s_2) = \mathcal{X}(s_1) \otimes s_2 + s_1 \otimes \mathcal{X}(s_2)$, is to say

$$\nabla = \nabla^{(1)} \otimes I + I \otimes \nabla^{(2)}$$

This is obviously an analog of the Leibniz rule. Whenever dealing with bundles that are all associated to a given principal bundle with a fixed invariant connection, we shall in general neglect to label the symbol ∇ to indicate in which bundle the covariant derivative is acting, as this should be clear from the context. Thus we shall simply write $\nabla_{\mathcal{X}}(s_1 \otimes s_2) = (\nabla_{\mathcal{X}} s_1) \otimes s_2 + s_1 \otimes (\nabla_{\mathcal{X}} s_2)$ which makes this result resemble even more the Leibniz rule. This result obviously extends to the tensor product of any finite number of associated vector bundles.

Let E be a vector bundle with fiber F and E' the dual bundle with fiber F', the dual of F. One sees from (39), (27) and (66) that in a trivialization, if t represents a local section of E' that $\nabla_{\mathcal{X}} t = \mathcal{X}(t) - \mathfrak{r}(A(\mathcal{X}))' t$.

Consider now the bundle $\mathrm{Hom}(E_1, E_2)$. This is associated to PG via a representation that acts on $\phi \in \mathrm{Hom}(F_1, F_2)$ via $R(g)\phi = R_2(g)\phi R_1(g^{-1})$. The corresponding Lie algebra representation is given by $\mathfrak{r}(L)\phi = \mathfrak{r}_2(L)\phi - \phi\mathfrak{r}_1(L)$. If now t represents a local section of $\mathrm{Hom}(E_1, E_2)$ and s represents a local section of E_1 over the same open set, then ts is a local section of E_2 depending linearly on s. One has in a trivialization,

$$\nabla_{\mathcal{X}} t = \mathcal{X}(t) + \mathfrak{r}_2(A(\mathcal{X}))t - t\mathfrak{r}_1(A(\mathcal{X})) \tag{69}$$

Now ts is linear in s so $\mathcal{X}(ts) = \mathcal{X}(t)s + t\mathcal{X}(s)$. Using this and (68), one concludes that

$$\nabla_{\mathcal{X}} ts = (\nabla_{\mathcal{X}} t)s + t\nabla_{\mathcal{X}} s \tag{70}$$

In particular if $\nabla_{\mathcal{X}} t = 0$ then $\nabla_{\mathcal{X}} ts = t\nabla_{\mathcal{X}} s$. As an example of this consider the map $E' \otimes E \to M \times \mathbb{F}$ which pointwise is given by the natural duality between F'_x and F_x. This map corresponds to a section of $\mathrm{Hom}(E' \otimes E, M \times \mathbb{F})$. In a trivialization this is represented by the constant map t with value being the natural duality $F' \otimes F \to \mathbb{F}$. Hence $\mathcal{X}(t) = 0$. Furthermore, specializing in (69) to $\mathfrak{r}_2 = 0$ as the representation in the trivial bundle is trivial, and to $\mathfrak{r}_1(L) = -\mathfrak{r}(L)' \otimes I + I \otimes \mathfrak{r}(L)$, one sees that the last two term are also zero and so $\nabla_{\mathcal{X}} t = 0$. From this we get $\mathcal{X} < f, s >= \nabla_{\mathcal{X}}(t(f \otimes s)) = t(\nabla_{\mathcal{X}}(f \otimes s))$, and so

$$\mathcal{X} < f, s >=< \nabla_{\mathcal{X}} f, s > + < f, \nabla_{\mathcal{X}} s >$$

It is often useful to consider the covariant derivative $\nabla_{\mathcal{X}}\sigma$ as being a function of both $\mathcal{X}$ and σ. Seeing that it is linear in $\mathcal{X}$ we can define it as a map $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$. To see this, suppose that in some open set U one has n vector fields $e_1, \dots, e_n$ which at each point provide a basis for the tangent space. Let $e^1, \dots, e^n$ be the corresponding dual set of 1-forms. We now define

$$\nabla\sigma = \sum_i e^i \otimes \nabla_{e_i}\sigma \tag{71}$$

To see that this is independent of the choice of the e_i, let $\tilde{e}_i$ be another choice, then one has $\tilde{e}_i = \sum_j M_i{}^j e_j$ for some field of invertible matrices $M_i{}^j$. One then has $\tilde{e}^i = \sum_j M^i{}_j E^j$, where $M^i{}_j$ is the inverse of the transpose of $M_i{}^j$. Is is now easy to see that the right-hand side of (71) is the same using $\tilde{e}_i$ instead of e_i. The characteristic property of ∇ is a form of the Leibniz rule

$$\nabla(f\sigma) = df \otimes \sigma + f\nabla\sigma$$

Let E be a vector bundle. An antisymmetric p-linear bundle map $\alpha : TM \times \cdots \times TM \to E$ is called a *E-valued p-forms*, or generically a *bundle-valued p-forms.* Let $\mathcal{X}_1, \ldots, \mathcal{X}_p$ be vector fields, then $\alpha(\mathcal{X}_1, \ldots, \mathcal{X}_p)$ is a section of E. If E has a linear connection γ, we can define, analogously to (123) the *covariant exterior derivative* $d_\gamma\alpha$ of α by

$$\begin{aligned} d_\gamma\alpha(\mathcal{X}_1, \ldots, \mathcal{X}_{p+1}) = \sum_i (-1)^{i+1} \nabla_{\mathcal{X}_i} \alpha(\mathcal{X}_1, \ldots, \hat{\mathcal{X}}_i, \ldots \mathcal{X}_{p+1}) + \\ \sum_{1 \le i < j \le p+1} (-1)^{i+j+1} \alpha([\mathcal{X}_i, \mathcal{X}_j], X_1, \ldots, \hat{\mathcal{X}}_i, \ldots, \hat{\mathcal{X}}_j, \ldots, \mathcal{X}_{p+1}) \end{aligned} \tag{72}$$

which as we shall see shortly is an E-valued $(p+1)$-form. The ordinary exterior derivative is not well defined in this context for the same reasons that one cannot define the ordinary directional derivative for bundle sections. Note however that, just as for the ordinary exterior derivative, no connection is needed on TM. We now determine the representative of $d_\gamma\alpha$ in a trivialization of E. Let Γ be the connection 1-form in the trivialization, and a the local representative of α, then $\nabla_{\mathcal{Y}}(a(\mathcal{X}_1, \ldots, \mathcal{X}_p)) = \mathcal{Y}(a(\mathcal{X}_1, \ldots, \mathcal{X}_p)) + \Gamma(\mathcal{Y})a(\mathcal{X}_1, \ldots, \mathcal{X}_p)$. Using this and (123) one arrives at the local representative

$$\begin{aligned} d_\gamma a(\mathcal{X}_1, \ldots, \mathcal{X}_{p+1}) = da(\mathcal{X}_1, \ldots, \mathcal{X}_{p+1}) + \\ \sum_i (-1)^{i+1} \Gamma(\mathcal{X}_i) a(\mathcal{X}_1, \ldots, \hat{\mathcal{X}}_i, \ldots \mathcal{X}_{p+1}) \end{aligned}$$

From this we see that locally, $d_\gamma\alpha$ is a vector-valued (F-valued) $(p+1)$-form. To show that globally it defines a bundle-valued (E-valued) $(p+1)$-form, one needs to argue that it has the right transition formula, but this is automatic given it's intrinsic definition (72).

We have seen that the curvature of a linear connection on a vector bundle E can be viewed as a $\mathrm{End}(E)$-valued 2-form. We now have the *Bianchi Identities*

Theorem 12 (Bianchi Identities) *Let E be a vector bundle with a linear connection γ and let R be the curvature 2-form of γ, then*

$$d_\gamma R = 0 \tag{73}$$

Proof: Apply the Jacobi Identity for covariant derivatives

$$[\nabla_{\mathcal{X}}, [\nabla_{\mathcal{Y}}, \nabla_{\mathcal{Z}}]] + [\nabla_{\mathcal{Y}}, [\nabla_{\mathcal{Z}}, \nabla_{\mathcal{X}}]] + [\nabla_{\mathcal{Z}}, [\nabla_{\mathcal{X}}, \nabla_{\mathcal{Y}}]] = 0$$

to any section of E. Using (70, 67) and the definition of d_γ, one arrives at the conclusion. Q.E.D

In contrast to the ordinary exterior derivative, we do not have $d_\gamma^2 = 0$. In fact, a simple calculation shows

$$d_\gamma^2 \alpha(\mathcal{X}_1, \ldots, \mathcal{X}_{p+2}) = \sum_{1 \le i < j \le p+2} (-1)^{i+j+1} R(\mathcal{X}_i, \mathcal{X}_j) \alpha(X_1, \ldots, \hat{\mathcal{X}}_i, \ldots, \hat{\mathcal{X}}_j, \ldots, \mathcal{X}_{p+2})$$

It is worth noting that d_γ^2 is a 0-th order differential operator, that is, it involves no differentiation.

A linear connection on the tangent bundle TM allows the following useful construction for two vector fields $\mathcal{X}$ and $\mathcal{Y}$

$$T(\mathcal{X}, \mathcal{Y}) = \nabla_{\mathcal{X}} \mathcal{Y} - \nabla_{\mathcal{Y}} \mathcal{X} - [\mathcal{X}, \mathcal{Y}]$$

Computing this in a trivialization one has, calling the representative of T again by the same letter, that $T(\mathcal{X}, \mathcal{Y}) = \Gamma(\mathcal{X})\mathcal{Y} - \Gamma(\mathcal{Y})\mathcal{X}$. From this it is clear that at a point $x \in M$, the vector $T(\mathcal{X}, \mathcal{Y})$ depends only on the vectors $\mathcal{X}(x)$ and $\mathcal{Y}(x)$ at x and not on how these are extended to the actual vector fields $\mathcal{X}$ and $\mathcal{Y}$. Thus T can be identified with a tensor in $TM \otimes T^*M \otimes T^*M$ and $T(\mathcal{X}, \mathcal{Y})$ is anti-symmetric in its two arguments. This tensor is known as the *torsion tensor* or simply the *torsion* of the connection. A connection for which $T = 0$ is known as *torsion-free*, or *torsionless*.

5 Manifolds

5.1 Pseudo-Riemannian Manifolds

A *pseudo-Riemannian* manifold is a manifold M with a non-degenerate symmetric bilinear form $g(x)$, called the *pseudo-metric* defined in each tangent space T_xM. We shall often write $< v, w >_x$ in place of $g(x)(v, w)$ and often suppress mention of the point x. We shall assume that $g(x)$ is a $\mathcal{C}^\infty$ function of x by which we mean that for two $\mathcal{C}^\infty$ vector fields $\mathcal{X}$ and $\mathcal{Y}$, $g(\mathcal{X}, \mathcal{Y})$ is a $\mathcal{C}^\infty$ function on M.

For any symmetric bilinear non-degenerate form β on a real n-dimensional vector space W, there is a basis $e_1, \ldots, e_n$ such that $\beta(e_i, e_j) = \eta_{ij}$ where the matrix $\eta = \mathrm{diag}\,(1, \ldots, 1, -1, \ldots, -1)$ with r entries of 1 and s entries of -1, where $r + s = n$. The pair (r, s) is called the *signature* of β. Such a basis is called *orthonormal*. The group of linear transformations $L : W \to W$ such that $\beta(Lx, Ly) = \beta(x, y)$ is denoted by $\mathrm{O}(\beta)$ and is know as the *orthogonal group* of β. The subgroup of $\mathrm{GL}(n)$ of matrices Λ such that $\Lambda\eta\Lambda^t = \eta$ is denoted by $\mathrm{O}(r, s)$ and is know as the *pseudo-orthogonal group of signature* (r, s). Obviously $\mathrm{O}(\beta) = \mathrm{O}(-\beta)$, $\mathrm{O}(r, s) \simeq \mathrm{O}(s, r)$, and $\mathrm{O}(\beta) \simeq \mathrm{O}(r, s)$. When $s = 0$ we write simply $\mathrm{O}(n)$. From $\Lambda\eta\Lambda^t = \eta$ one concludes $\det(\Lambda)^2 = 1$, that is $\det(\Lambda) = \pm 1$ for any element of the orthogonal groups. The elements of determinant 1 form a subgroup called the *special (pseudo)-orthogonal* group and we denote these correspondingly by $\mathrm{SO}(\beta)$, $\mathrm{SO}(r, s)$ and $\mathrm{SO}(n)$.

In a pseudo-Riemannian manifold each $g(x)$ has some signature. This signature is constant on the connected components of M since to pass continuously to a different signature, the form would have to become degenerate at some point. We shall assume the signature is constant on M. A manifold of signature $(n, 0)$ is called a *Riemannian* manifold. A *space-time* is a manifold of signature $(1, n-1)$, or, according to some authors, one of signature $(n-1, 1)$. Since replacing g by $-g$ does not change most geometric facts, for many effects manifolds of signature (r, s) are equivalent to those of (s, r).

If M is pseudo-Riemannian, then in a neighborhood of any point x_0 one can, by appropriate linear combinations of any set of local coordinates, introduce coordinates $x^1, \ldots, x^n$ such that

$$< \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} >_{x_0} = \eta_{ij} \tag{74}$$

It should be emphasized that in general it is impossible to introduce local coordinates so that (74) hold at all points in a neighborhood of x_0. The possibility of doing so depends on the vanishing of the curvature of an appropriate connection. In any case, consider the vectors fields $v_i = \frac{\partial}{\partial x^i}$. In a small enough neighborhood of x_0 one can apply the usual Gram-Schmidt orthonormalization procedure to $v_1, \ldots, v_n$ to obtain vector fields $e_1, \ldots, e_n$

such that $< e_i, e_j >= \eta_{ij}$. Such a set of vector fields is called a (local) *n-bein*.

Let now $\mathcal{F}_{\mathrm{O}}(M)$ be the *orthonormal frame-bundle* of M, that is, the fiber at any point x is the set of all ordered orthonormal bases in T_xM with respect to the bilinear form $g(x)$. This bundle is trivialized over open sets U in M in which one has an n-bein $e_1^U, \ldots, e_n^U$. Given an orthonormal basis $b_1, \ldots, b_n$ at $x \in U$ one has $b_i = \sum_j e_j R^j{}_i$ for some matrix $R \in \mathrm{O}(r,s)$. The trivializing homeomorphism assigns to the basis b the point $(x, R) \in U \times \mathrm{O}(r,s)$. If V with n-bein $e_1^V, \ldots, e_n^V$ define now another trivialization, then one has $e_i^U = \sum_j e_j^V G^j{}_i$ for some matrix $G \in \mathrm{O}(r,s)$ and one can easily see that the transition map g_{VU} is given by $R \mapsto GR$, so G defines a transition map and the structure group of the bundle is $\mathrm{O}(r,s)$. In fact it is a *principal* $\mathrm{O}(r,s)$-bundle as the global right action is easily defined by $v_i \mapsto \sum_j v_j L^j{}_i$ for $L \in \mathrm{O}(r,s)$.

Using the pseudo-Riemannian structure we can reduce the structure group of the tangent bundle from $\mathrm{GL}(n)$ to $\mathrm{O}(r,s)$. In fact, given an open set U with n-bein $e_1^U, \ldots, e_n^U$ and a tangent vector $v \in T_xM$ for $x \in U$ one has $v = \sum_j v_U^j e_j$. The tangent bundle can now be trivialized by assigning to v the point $(x, (v^1, \ldots, v^n)) \in U \times \mathbb{R}^n$. As $v_V^j = L^j{}_i v_U^i$ for some matrix $L \in \mathrm{O}(r,s)$ the transition maps now belong to $\mathrm{O}(r,s)$ and we have expressed TM as a bundle with structure group $\mathrm{O}(r,s)$. This is an explicit example of the important process of considering a given fiber bundle as having several structure groups. We shall see in subsequent sections how the reduction to $\mathrm{O}(r,s)$ allows for constructions, spin bundles in our case, that would be impossible with $\mathrm{GL}(n)$ as the structure group.

As before, introducing the two natural representations $\rho(L,z) = Lz$ and $\rho^*(L,z) = (L^t)^{-1}z$ of of $\mathrm{O}(r,s)$ on $\mathbb{R}^n$ one has the identifications:

$$\mathcal{F}_{\mathrm{O}}(M) \times_\rho \mathbb{R}^n \simeq TM \tag{75}$$

$$\mathcal{F}_{\mathrm{O}}(M) \times_{\rho^*} \mathbb{R}^n \simeq T^*M \tag{76}$$

Note that in the Riemannian case, the two representations are identical which shows that the tangent and cotangent bundles are then isomorphic $\mathrm{O}(n)$-bundles. Since one can introduce a Riemannian metric in any paracompact Hausdorff manifold, the tangent and cotangent bundles in these are always isomorphic vector bundles.

By a deliberate abuse of notation we shall also denote by $< \cdot, \cdot >$ the duality between T_x^*M and T_xM, that is if $\alpha \in T_x^*M$ and $v \in T_xM$ we shall write $< \alpha, v >$ for $\alpha(v)$. We shall also write $< v, \alpha >$ for the same thing.

Given a pseudo-Riemannian structure g one can define a map $\ell : T_xM \to T^*_xM$ by the relation $\ell(v)(w) =< v, w >$ for all $w \in T_xM$. By our abuse of notation this can be written as $< \ell(v), w >=< v, w >$. One also has the inverse map $r : T^*_xM \to T_xM$ defined by $< r(\alpha), v >= \alpha(v)$ which by our abuse of notation can also be expressed as $< r(\alpha), v >=< \alpha, v >$. Obviously r and ℓ are inverses of each other. These maps can be used to introduce a non-degenerate symmetric bilinear form in each T^*_xM by the formula $< \alpha, \beta >=< r(\alpha), r(\beta) >$.

It is useful to determine local expressions for these maps and bilinear forms in any trivialization of TM and T^*M. Such a trivialization is given by a set of vector fields $e_1, \dots, e_n$ in an open set such that at any point x they form a basis for T_xM. Let $e^1, \dots, e^n$ be the corresponding dual set of 1-forms.

Define

$$g_{ij} =< e_i, e_j >$$

If now $v = \sum_i v^i e_i$ and $w = \sum_i w^i e_i$ then one has

$$< v, w >= \sum_{ij} g_{ij} v^i w^j$$

One will also have

$$r(e^i) = \sum_j g^{ij} e_j$$

for some matrix g^{ij}. From $< r(e^i), e_j >=< e^i, e_j >= \delta^i{}_j$ one deduces $\sum_k g^{ik} g_{kj} = \delta^i{}_j$ so g^{ij} is the matrix inverse of g_{ij}. One has for $\alpha = \sum_i \alpha_i e^i$ and $\beta = \sum_i \beta_i e^i$ that

$$< \alpha, \beta >= \sum_{ij} g^{ij} \alpha_i \beta_j$$

Of course, by definition,

$$< \alpha, v >= \sum_i \alpha_i v^i$$

Interpreting the metric g as a section of $T^*M \otimes T^*M$, one has $g = \sum_{ij} g_{ij}\, e^i \otimes e^j$. This is known as the *(pseudo)-metric tensor*. For the particular case of a coordinate basis $e_i = \frac{\partial}{\partial x^i}$ of a set of local coordinates $x^1, \dots, x^n$, one has $e^i = dx^i$ and the g_{ij} are then the well-known usual components of the metric tensor $g = \sum_{ij} g_{ij}\, dx^i \otimes dx^j$. For an n-bein $g^{ij} = g_{ij} = \eta_{ij}$ and so $g = \sum_{ij} \eta_{ij}\, e^i \otimes e^j$.

In the physical literature, r is known as *raising indices*. This is because if $\alpha = \sum_i \alpha_i \, dx^i$, then it is customary to write $r(\alpha) = \sum_i \alpha^i \frac{\partial}{\partial x^i}$ and so the index of α_i was "raised". Similarly ℓ is known as *lowering indices*.

Let $S \in \mathrm{GL}(n)$, then for any integer n and any real number s the maps $\nu_n : S \mapsto \det(S)^{-n}$ and $\mu_s : S \mapsto |\det(S)|^{-s}$ define one-dimensional representations of $\mathrm{GL}(n)$. Remember (26), the transition map given by the Jacobian matrix of change of local coordinates. A line bundle with fiber $\mathbb{R}$ using the transition map J of local coordinate changes and representation ν_n is called the *bundle of signed densities of weight* n, and representation μ_s, the *bundle of (absolute) densities of weight* s.

For a bundle trivialized in local chart U, a density D is represented by a function $D_U(x)$. One has in the intersection of charts U and V, for a signed density $D_V = \det(J)^{-n} D_U$, and for an absolute density $D_V = |\det(J)|^{-s} D_U$, where n and s are the respective weights.

Consider now an absolute density D of weight 1 and its local representatives D_U and D_V where U carries local coordinates $x^1, \ldots, x^n$ and V coordinates $y^1, \ldots, y^n$. For an open set $W \subset U \cap V$ one has

$$\int \cdots \int_{y(W)} D_V(y) \, dy^1 \cdots dy^n =$$
$$\int \cdots \int_{x(W)} |\det(J)|^{-1} D_U(x) |\det(J)| \, dx^1 \cdots dx^n =$$
$$\int \cdots \int_{x(W)} D_U(x) \, dx^1 \cdots dx^n$$

In the second term the factor $|\det(J)|^{-1}$ arises due to change of trivialization and the factor $|\det(J)|$ due to the formula for change of variables in integration; the two factors cancel. This means that absolute densities of weight 1 can be integrated over the manifold using local coordinates and a partition of unity:

$$\int_M D = \sum_{U \in \mathcal{U}} \int_{x_U(U)} \xi_U D_U(x_U) \, dx_U$$

where x_U are local coordinates in U and the ξ_U form a partition of unity subordinate to an atlas $\mathcal{U}$. We see thus that such a density defines a signed measure on M and given any function f on M its integral with respect to this measure is given by $\int_M fD$.

An n-form $\Omega \in \bigwedge^n(T^*M)$ has a representative in local coordinates as $D_U(x) \, dx^1 \wedge \cdots \wedge dx^n$ and one sees that the set of local coefficient D_U define a *signed* density of weight 1. This can be used to define an absolute density of weight 1 if the atlas of local charts can be so chosen as to have $\det(J) > 0$ for all pairs of intersecting charts. When this is possible we say the manifold is *orientable*, and a choice of an atlas in which $\det(J) > 0$ is called a choice of *orientation*. Once such an atlas is chosen we say the manifold is *oriented*. In this case we can integrate n-forms on M.

On an orientable manifold the frame-bundle $\mathcal{F}(M)$ is the disjoint union of two bundles that correspond to frames of a fixed orientation. Likewise the orthogonal frame bundle $\mathcal{F}_{\mathrm{O}}(M)$ separates into a disjoint union corresponding to orthonormal frames of a fixed orientation. We shall denote any one of these sub-bundles by $\mathcal{F}_{\mathrm{SO}}(M)$ which is obviously a principal $\mathrm{SO}(r, s)$-bundle. The tangent and cotangent bundles can now be considered as associated to $\mathcal{F}_{\mathrm{SO}}(M)$

$$\begin{aligned} \mathcal{F}_{\mathrm{SO}}(M) \times_\rho \mathbb{R}^n &\simeq TM \\ \mathcal{F}_{\mathrm{SO}}(M) \times_{\rho^*} \mathbb{R}^n &\simeq T^*M \end{aligned}$$

by restricting the the previously used representations in (75, 76) to $\mathrm{SO}(r, s)$.

Let now M be a pseudo-Riemannian manifold. In a local chart U one has the local representative $\sum_{ij} g^U_{ij}\, dx^i \wedge dx^j$ of the metric tensor as defined above. From its definition, one easily deduces $\sum_{k\ell} g^V_{k\ell} J^k i J^\ell{}_j = g^U_{ij}$ which means that $\sqrt{|\det(g^U)|} = \sqrt{|\det(g^U_{ij})|}$ is an absolute density of weight 1 called the *volume element*. If M is oriented, then

$$\Omega = \sqrt{|\det(g^U)|}\, dx^1 \wedge \cdots \wedge dx^n$$

defines an n-form called the *volume n-form*.

The pseudo-metric g defines a symmetric bilinear non-degenerate form in the space of exterior p-covectors $\bigwedge^p(T^*_x M)$ at a point $x \in M$ by the formula

$$< \alpha_1 \wedge \cdots \wedge \alpha_p, \beta_1 \wedge \cdots \wedge \beta_p >= \det_{ij} < \alpha_i, \beta_j >$$

where the right-hand side is the determinant of the matrix $< \alpha_i, \beta_j >$. This in turn defines the *Hodge star* operator

$$* : \textstyle\bigwedge^p(T^*_x M) \to \bigwedge^{n-p}(T^*_x M)$$

by the relation

$$\phi \wedge *\psi =< \phi, \psi > \Omega \tag{77}$$

for any two exterior p-forms ϕ and ψ. One obviously has $*1 = \Omega$. One can get a concrete idea of the Hodge operator by considering an orthonormal basis $e_1, \dots, e_n$ of the tangent space at a point. Let $e^a = r(e_a)$, be the corresponding covectors. One has $< e^a, e^b >= \eta^{ab}$ where η^{ab} is the inverse matrix of η_{ab}, obviously equal to it, introduced for notational convenience. One has $dx^i = \sum_a h^i{}_a e^a$ for some matrix of coefficients $h^i{}_a$. Thus $\Omega = \sqrt{|\det(g)|}\det(h) e^1 \wedge \cdots \wedge e^n$. On the other hand, $g^{ij} = \sum_{ab} h^i{}_a h^j{}_b \eta^{ab}$ from which one deduces $\det(g)^{-1} = \det(h)^2 \det(\eta)$. If we denote by σ_h the sign of $\det(h)$ one concludes $\Omega = \sigma_h e^1 \wedge \cdots \wedge e^n$. One can now easily find that

$$* (e^{i_1} \wedge \cdots \wedge e^{i_p}) = \sigma_{IJ} \sigma_h (\textstyle\prod_{a=1}^p \eta^{i_a i_a})\, e^{j_1} \wedge \cdots \wedge e^{j_{n-p}} \tag{78}$$

where $\{j_1, \dots, j_{n-p}\}$ is the complementary subset, in ascending order, to $\{i_1, \dots, i_p\}$ in $\{1, \dots, n\}$ and σ_{IJ} is the parity of the permutation

$$(1, \dots, n) \mapsto (i_1, \dots, i_p, j_1, \dots, j_{n-p})$$

A double application of this formula gives

$$* * \psi = (-1)^{s+p(n-p)} \psi \tag{79}$$

If n is even, $n = 2m$, and ψ is an m-form, then $*\psi$ is again an m-form. An m-form ψ such that $*\psi = \psi$ is said to be *self-dual*, and one such that $*\psi = -\psi$ is said to be *anti-self-dual*. If $(-1)^{s+m} = 1$, then from (79) one has $* * \psi = \psi$ for any m-form. In this case we can decompose any m-form into a sum of a self-dual and an anti-self-dual form. Define

$$\psi^{\pm} = \frac{\psi \pm *\psi}{2}$$

then one sees that $\psi = \psi^+ + \psi^-$ and $*\psi^{\pm} = \pm\psi^{\pm}$.

Using the Hodge star one can define the differential operator $\delta = *d*$ where d is the exterior derivative. One sees that δ transforms a p-form into a $p-1$ form. From (79) one sees that $\delta^2 = 0$. The operator $d\delta + \delta d$ is a second order operator on the space of p-forms and in the Riemannian case is knows as the *Hodge Laplacian*.

5.2 The Levi-Civita Connection

Theorem 13 *On any pseudo-Riemannian manifold M there is a unique invariant connection on the orthogonal frame bundle $\mathcal{F}_O(M)$ called the Levi-Civita connection which is characterized by the following properties of the associated connection ∇ on the tangent bundle $T(M)$:*

$$\begin{aligned} \nabla_{\mathcal{X}}\mathcal{Y} - \nabla_{\mathcal{Y}}\mathcal{X} &= [\mathcal{X}, \mathcal{Y}] && (80) \\ \mathcal{X}(g(\mathcal{Y}, \mathcal{Z})) &= g(\nabla_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}) + g(\mathcal{Y}, \nabla_{\mathcal{X}}\mathcal{Z}) && (81) \end{aligned}$$

Of these, condition (80) says that the Levi-Civita connection is torsion-free, and condition (81) says that the covariant derivative of the pseudo-metric tensor vanishes. Viewing g as a section of $T^*(M) \otimes T^*(M)$, condition (81) is equivalent to $\nabla g = 0$ (see Section 4.10).

Proof: Let $e_1, \ldots, e_n$ be vector fields in an open set such that at any point x they form a basis for T_xM. Use this basis to locally trivialize TM. One has $\nabla_{e_j} e_k = \sum_\ell \Gamma^\ell{}_{kj} e_\ell$ and $[e_j, e_k] = \sum_\ell C^\ell{}_{jk} e_\ell$ for some functions $C^\ell{}_{jk}$. Condition (80) now means that

$$\Gamma^\ell{}_{kj} - \Gamma^\ell{}_{jk} = C^\ell{}_{jk}$$

Let now $\Gamma_{ijk} = \sum_\ell g_{i\ell}\Gamma^\ell{}_{jk}$ and $C_{ijk} = \sum_\ell g_{i\ell} C^\ell{}_{jk}$ then condition (81) becomes

$$e_k(g_{ij}) = \Gamma_{jki} + \Gamma_{ikj} = \Gamma_{jki} + \Gamma_{ijk} + C_{ijk} \tag{82}$$

Refer to equation (82) as $E(ijk)$ and now add $E(ijk)$ to $E(jki)$ and subtract $E(kij)$. One arrives at

$$2\Gamma_{jki} = e_k(g_{ij}) + e_i(g_{jk}) - e_j(g_{ik}) + C_{kij} - C_{ijk} - C_{jki}$$

and finally

$$\Gamma^i{}_{jk} = \frac{1}{2}\sum_\ell g^{i\ell}(e_j(g_{k\ell}) + e_k(g_{j\ell}) - e_\ell(g_{jk}) + C_{jk\ell} - C_{k\ell j} - C_{\ell jk}) \tag{83}$$

In a coordinate basis, $e_i = \frac{\partial}{\partial x^i}$, one has $C^\ell{}_{ij} = 0$ and we arrive at the classic expression:

$$\Gamma^i{}_{jk} = \frac{1}{2}\sum_\ell g^{i\ell}\left(\frac{\partial g_{k\ell}}{\partial x^j} + \frac{\partial g_{j\ell}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^\ell}\right) \tag{84}$$

For an n-bein e_i one has $g_{ij} = \eta_{ij}$ thus $e_k(g_{ij}) = 0$ and so

$$\Gamma^i{}_{jk} = \frac{1}{2}\sum_\ell \eta^{i\ell}(C_{jk\ell} - C_{k\ell j} - C_{\ell jk}) \tag{85}$$

Thus, the connection, if it exists, is uniquely determined. To show existence it is necessary to show that (83), or alternatively the specialized form (84) or (85), does indeed define a connection. One way to do this is to show that Γ has the proper transformation law (33) under change of trivialization, which is straightforward if not tedious. One could also argue that as conditions (80)and (81) are trivialization-independent, and as they can be met in a unique way in any trivialization, Γ cannot help but transform in the right way under a change of trivialization. Q.E.D

The curvature 2-form associated to the Levi-Civita connection is thus an anti-symmetric map $T^*M \otimes T^*M \to \mathrm{End}(TM)$ and so can be interpreted as a tensor in $TM \otimes T^*M \otimes T^*M \otimes T^*M$. As such it is known as the *Riemann curvature tensor*.

6 Lagrangian Theories

6.1 Lagrangians

Physical field theory is preponderately Lagrangian theory. In its simplest terms one deals with a set of fields $\psi^1, \ldots, \psi^m$ which for now we simply consider as functions on a Euclidean space $\mathbb{R}^n$. For $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_i \geq 0$ being integers, denote by $|\alpha|$ the sum $\alpha_1 + \cdots + \alpha_n$ and by ∂_α the partial derivative

$$\frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \cdots \partial {x_n}^{\alpha_n}}$$

A *k-order Lagrangian density* (or simply *Lagrangian*) in this context is just a function $\mathcal{L}(x, \partial_\alpha \psi^a(x))$ of the point $x \in \mathbb{R}^n$ and of the values of the fields and their partial derivatives at x up to order k ($\alpha = 0$ corresponds to just the field values). Associated to a Lagrangian density and the fields ψ^a is the *action*

$$S(\psi) = \int_{\mathbb{R}^n} \mathcal{L}(x, \partial_\alpha \psi^a(x))\, dx \tag{86}$$

which defines a functional on the set of fields. The word "action" here is borrowed from physics where it means a quantity which has the physical dimension of energy times time. It should not be confused with the use we make elsewhere in these notes to mean group action or its analogs. The Lagrangian defines a set of differential equations that constitute the physical laws obeyed by the fields. These equations express the property of the fields being a singular point of the action. Let γ^a be a set of functions of compact support and consider the action $S(\psi + r\gamma)$. The requirement that ψ be a singular point of S then translates to the requirement that for all such γ

$$\frac{d}{dr} S(\psi + r\gamma)\bigg|_{r=0} = 0$$

One has

$$\frac{d}{dr} S(\psi + r\gamma)\bigg|_{r=0} = \int_{\mathbb{R}^n} \sum_a \sum_\alpha (-1)^{|\alpha|} \partial_\alpha \frac{\partial \mathcal{L}}{\partial y^a_\alpha} \partial_\alpha \gamma^a \, dx$$

where by $\frac{\partial \mathcal{L}}{\partial y^a_\alpha}$ we mean the partial derivative of $\mathcal{L}$ in relation to the variable for which one substitutes $\partial_\alpha \psi^a$ to form the integrand of the action. After integration by parts, the right-hand side becomes

$$\int_{\mathbb{R}^n} \sum_a \sum_\alpha (-1)^{|\alpha|} \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial y^a_\alpha} \right) \gamma^a \, dx$$

As this must be true for all γ with compact support one deduces that ψ must satisfy the *Euler-Lagrange equations* (one for each a):

$$\sum_\alpha (-1)^{|\alpha|} \partial_\alpha \left(\frac{\partial \mathcal{L}}{\partial y^a_\alpha} (x, \partial_\beta \psi^b(x)) \right) = 0$$

This is often abbreviated to the strange-looking expression

$$\sum_\alpha (-1)^{|\alpha|} \partial_\alpha \frac{\partial \mathcal{L}}{\partial \left(\partial_\alpha \psi^a \right)} = 0$$

When one of the variables x^i corresponds to time, these are the dynamical equation of a physical theory described by $\mathcal{L}$. That all fundamental physical theories are described by Lagrangians is a remarkable fact, which is not truly understood. Although quantum theory sheds some light on this, as the action itself and not just its critical points has physical meaning, we shall simply accept this fact in these notes.

Once it is realized that physical fields should really be thought of as sections of fiber bundles, (86) has to be modified. In first place, $\mathbb{R}^n$ should be replaced by a manifold M (such as physical space-time), and in second place, the ψ^a ought to be considered local representatives of fiber-bundle sections in appropriate trivializations. One should then replace (86) by

$$S(\psi) = \sum_{U\in\mathcal{U}} \int_U \xi_U \mathcal{L}_U(x_U, {\psi_U}^a(x_U), {\psi_U}^a_i(x_U))\, dx_U \tag{87}$$

where $\mathcal{U}$ is an atlas for M of coordinate charts over which the bundles in question trivialize, x_U are local coordinates in U, ψ_U are representatives of the sections, and ξ_U is a partition of unity subordinate to $\mathcal{U}$. The $\mathcal{L}_U$ are local representatives of what would be the Lagrangian understood in global form. For this to make sense, the resulting action S must not depend on the details of the construction in (87), that is, on the choice of local coordinates and trivializations. While it is easy to restate the definition of a Lagrangian theory so that this is automatic, it is by no means entirely straightforward to provide examples, and even less so of describing all possibilities.

Most physical theories are defined by first order Lagrangians, though higher order ones do occur.

Let $\pi : E \to M$ be any bundle with fiber F (we are in the category of manifolds) and consider at any point $x \in M$ the following relation on sections of E. Let $f : E \to \mathbb{R}$ be any $\mathcal{C}^\infty$ function and let $\mathcal{X}_1, \ldots, \mathcal{X}_k$ be any k vector fields on M. We say two sections σ_1 and σ_2 have a *contact of order* k at x if for all such f and all such $\mathcal{X}_i$ one has for all $0 \le \ell \le k$

$$(\mathcal{X}_1\mathcal{X}_2\cdots\mathcal{X}_\ell f \circ \sigma_1)(x) = (\mathcal{X}_1\mathcal{X}_2\cdots\mathcal{X}_\ell f \circ \sigma_2)(x)$$

Having contact of order k is obviously an equivalence relation. It is easy to see that in a trivialization $U \times F$ with local coordinates $x^1, \ldots, x^n$ in U and $y^1, \ldots, y^m$ in a local chart W in F that if a section σ in a neighborhood of x is given by $y^a = s^a(x)$, then σ_1 and σ_2 have contact of order k at x if and only if s^a_1 and s^a_2 coincide at x along with all partial derivatives up to and including order k. By the *k-jet* of σ at x we mean the equivalence class of σ under the relation of having contact of order k. The *k-jet bundle* $J^k(E)$ is the bundle of k-jets of sections of E at all point $x \in M$. This is a manifold for which local coordinates are defined in terms of the coordinates x^i and y^a introduced above, where to the equivalence class $[\sigma]_x$ we associate the coordinates

$$\left(x^1, \ldots, x^n, s^1(x), \ldots, s^m(x), \frac{\partial s^1(x)}{\partial x^1}, \ldots, \frac{\partial^k s^m(x)}{\partial (x^n)^k}\right)$$

where one has to include all the partial derivatives up to order k of all the components s^a. The rather complex exact expression for the transition maps of this bundle will not be important for us, it can be calculated from the chain rule. Given a section σ of E we define its *k-th jet extension* by $j^k\sigma(x) = [\sigma]_x$ which is a section of $J^k(E)$.

Lagrangians as we have defined them are functions of the fields and their derivatives, so they are naturally defined on $J^k(E)$ which we see incorporates in a global manner sections and their derivatives up to order k. One must, to be able to calculate the action, integrate the Lagrangian density. This will be possible if we define the Lagrangian density as a bundle map

$$\mathcal{L} : J^k(E) \to \mathcal{V}(M)$$

where $\mathcal{V}(M)$ is the bundle of volume densities, that is, absolute densities of weight 1. If M is oriented, this bundle can be replaced by $\bigwedge^n(T^*M)$, the bundle of n-forms on M. The action of a section σ is now defined as

$$S(\sigma) = \int_M \mathcal{L} \circ j^k\sigma \tag{88}$$

The action is a functional on $\Gamma(M)$, the set of smooth sections of E. In most cases of interest, this set is formally an infinite dimensional differential manifold. Singular point of S on this manifold correspond to solutions of the Euler-Lagrange differential equations whose local form can be calculated as before.

Gauge theory is a Lagrangian theory of connections on principal fiber bundles along with sections of vector bundles, invariant under action of the gauge group. To be able to talk about a Lagrangian defined for a connection on a principal bundle we must see how a connection can be seen as a section of a bundle. To this end remember formula (38) giving the relation between the local principal gauge potentials in two trivializations. We interpret this now as expressing the transition map of a fiber bundle. If we trivialize in a local chart one has $A = \sum_i A_i^U\, dx^i$ where $A_i^U \in \mathfrak{g}$. If G is a matrix group, then the A_i are matrices, known in the physical literature as *gauge potentials*. If we take the A_i^U as local representatives of ω we see that our bundle must have fiber $\mathfrak{g}^n$. One has from (38) that $A_i^V = \sum_j J_i^{*j}(g_{VU} \cdot A_j^U \cdot g_{VU}^{-1} + \gamma_j)$ where $\gamma_i \in \mathfrak{g}$ are defined by $-dg_{VU} \cdot g_{VU}^{-1} = \sum_i \gamma_i\, dx^i$, and J_i^{*j} is the matrix of the adjoint action of $\mathrm{GL}(n)$ corresponding to the cocycle of change of local coordinates (26), that is, the inverse of the transpose of J. This is a transition map of an affine bundle with the following data: The group $\tilde{G}$ of the bundle as a set is $\mathrm{GL}(n) \times G \times \mathfrak{g}^n$. The group multiplication law is

$$(J, g, L_1, \ldots, L_n)(J', g', L'_1, \ldots, L'_n) = (JJ', gg', \mathrm{Ad}_g L'_1 + L_1, \ldots, \mathrm{Ad}_g L'_n + L_n)$$

The action of $\tilde{G}$ on the fiber is given by

$$(J, g, L_1, \ldots, L_n) \cdot (K_1, \ldots, K_n) = (K'_1, \ldots, K'_n)$$

where

$$K'_i = \sum_j J^{*j}_i (\mathrm{Ad}_g K_j + L_j)$$

and the transition map of the bundle is $\tilde{g}_{VU} = (J_{VU}, g_{VU}, \gamma_1, \ldots, \gamma_n)$. The bundle so constructed is called the *connection bundle* of PG. With this construction, invariant connections on PG correspond to sections of the connection bundle. One can now introduce Lagrangian densities of invariant connections.

A gauge transformation $\phi : PG \to PG$ induces a transformation on an invariant connection $\omega \to \phi \cdot \omega$ and consequently a transformation of a section A of the connection bundle given essentially by (43). We say a Lagrangian theory for a gauge potential is *gauge invariant* if the action satisfies $S(\phi \cdot A) = S(A)$ for all gauge transformations ϕ. Often the Lagrangian density itself is gauge invariant which guarantees the invariance of the action. In this case, as the transition map relating two trivializations (38) is also in the form of a gauge transformation, the Lagrangian density automatically takes care of the requirement that the action not depend on the choice of trivialization in the construct (87) above.

As a first example, let M be any oriented pseudo-Riemannian manifold with a principal G-bundle. Let A be a section of the connection bundle and F the local representative of the corresponding curvature two-form. The *Yang-Mills* Lagrangian density for a principal G-bundle of a matrix group G is given by

$$\mathcal{L}_{YM} = -k \mathrm{Tr}(F \wedge *F) \tag{89}$$

where k is a conventional constant, and $*$ is the Hodge star (see (77)). This equation has to be well interpret as F has values in $\mathfrak{g}$ which is a linear space of matrices. Given a $\mathfrak{g}$-valued p-form α and q-form β, each one is a matrix $\alpha^a{}_b$ and $\beta^a{}_b$ of ordinary p- and q-forms respectively. What is meant by $\alpha \wedge \beta$ is the matrix $\gamma^a{}_b$ of $p+q$-forms where $\gamma^a{}_b = \sum_c \alpha^a{}_c \wedge \beta^c{}_b$. Note that because of the trace and formulas (38) and (64), the resulting n-form is well defined over the manifold and is gauge invariant. We thus have a gauge-invariant theory, called the Yang-Mills theory. Recall that the curvature 2-form can be interpreted as a $P\mathfrak{g}$-valued 2-form. The Euler-Lagrange equations for the Yang-Mills theory are $d_A * F = 0$ where d_A is the covariant exterior derivative. Along with the Bianchi Identities (73), the Yang-Mills 2-form F therefore satisfies

$$\begin{aligned} d_A F &= 0 \quad (90) \\ d_A * F &= 0 \quad (91) \end{aligned}$$

Note that if A leads to a self-dual or anti-self-dual curvature, $F = \pm * F$, then by the Bianchi Identities, which always hold, the Yang-Mills equations are automatically satisfied.

Another much studied example is the Chern-Simons theory. Let M be an oriented 3-manifold and consider a principal bundle of a matrix group G. The Chern-Simons Lagrangian is

$$\mathcal{L}_{CS} = k \operatorname{Tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A) \quad (92)$$

where k is a constant. The Euler-Lagrange equations are $F = 0$, that is, the connection must be flat. The transformation properties of this Lagrangian under gauge transformations of A are a lot more sophisticated than for the Yang-Mills theory since A transforms in a more complicated fashion than F. In fact it is not even immediately clear that the action is well defined, as (92) is not invariant under gauge transformations, so one cannot join overlapping trivializations as is done in the Yang-Mills case using the construction of (87). However for some groups, such as $\mathrm{SU}(n)$, the principal bundles are trivial and one can use one global trivialization. The gauge invariance continues to be subtle. Under a gauge transformation defined by the function $\phi : M \to G$, (92) by

$$\mathcal{L}_{WZ}(\phi) = \frac{k}{3} \operatorname{Tr}(d\phi\, \phi^{-1} \wedge d\phi\, \phi^{-1} \wedge d\phi\, \phi^{-1})$$

which is known as the *Wess-Zumino* term. If ϕ_s is a smooth curve in the gauge group (in the sense that $(x, s) \mapsto \phi_s(x)$ is $\mathcal{C}^\infty$), then an easy calculation shows that the derivative of $\mathcal{L}_{WZ}(\phi_s)$ with respect to s is an exact differential. Thus on a manifold without boundary, the Chern-Simons action is invariant under those transformations that can be joined to the identity by a smooth curve in the gauge group (in fact for those in the component of the identity), but is not generally invariant. This subtle property is part of the reason for the great interest in this theory.

6.2 Minimal Coupling

For simplicity consider initially a system of first-order partial differential equations on $\mathbb{R}^n$ for a set of m real functions ψ^a, $a = 1, \ldots, m$ given by

$$\sum_{i=1}^{n} L_i \frac{\partial \psi}{\partial x^i} + M\psi = 0 \tag{93}$$

where the L_i and M are $m \times m$ constant matrices. There is no loss of generality in supposing real functions, for if the ψ were complex, one could consider the $2m$ real functions comprised of their real and imaginary parts, and these would satisfy a system of equations of the same form as (93). Suppose furthermore that there is a matrix group G of $m \times m$ matrices that commute with the L_i and M. If now ψ is a solution of (93) and $\Lambda \in G$ then $\Lambda\psi$ is also a solution and we say that G is a *global* symmetry group of the system. By "global" one means in this case that the same matrix Λ is applied to $\psi(x)$ for all points x. If however now we apply a G-valued function $\Lambda(x)$ to $\psi(x)$ and consider the new function $\tilde{\psi}(x) = \Lambda(x)\psi(x)$, then $\tilde{\psi}$ is not a solution of the new system. One however easily shows that $\tilde{\psi}$ satisfies the equation

$$\sum_{i=1}^{n} L_i (\frac{\partial}{\partial x^i} - \frac{\partial \Lambda}{\partial x^i}\Lambda^{-1})\tilde{\psi} + M\tilde{\psi} = 0$$

We see therefore that the group of *local* transformations $\psi(x) \mapsto \Lambda(x)\psi(x)$ is not a symmetry group of the system as the system itself changes under the transformation. One could however enlarge the system to make the local transformations symmetries. One interprets the term $-\frac{\partial \Lambda}{\partial x^j}\Lambda^{-1}$ as representing the change under the action of the local group of an additional set of functions represented by an n-tuple of $m \times m$ matrices $A_i(x)$. Rewrite now the original equation as

$$\sum_{i=1}^{n} L_i (\frac{\partial}{\partial x^i} + A_i)\psi + M\psi = 0 \tag{94}$$

and extend the local group to now act on both ψ and A as

$$\psi(x) \mapsto \Lambda(x)\psi(x) \tag{95}$$

$$A_i(x) \mapsto \Lambda(x)A_i(x)\Lambda(x)^{-1} - \frac{\partial\Lambda(x)}{\partial x^i}\Lambda(x)^{-1} \tag{96}$$

Comparing (96) to (43) one sees that it is of the same form as the change in a gauge potential due to the action of a gauge group. Gauge transformations thus arise naturally if one tries to make a global symmetry group local. Furthermore, comparing the operator $\frac{\partial}{\partial x^i} + A_i$ to (66) one sees that it is of the form of a covariant derivative with respect to the connection whose gauge potential is A. The change in going from (93) to (94) is precisely that of replacing ordinary partial derivatives with the corresponding covariant derivatives. The new set of functions comprised of ψ and A is however not satisfactory from various aspects. Whereas the ψ are governed by a system of differential equations, the A are not. From a physical point of view this not natural as all physical fields are to be thought of as fundamentally dynamical objects. What is lacking is thus a system of differential equations that would govern A. Since on A the local group acts in the same way as a gauge transformation, it is natural to posit that A is in fact a local gauge potential and that that it is governed by a gauge-invariant system of differential equations. One saw in Section 6.1 that Lagrangian gauge theories provide such equations. Particular cases of equation (93) can likewise be obtained from a Lagrangian. Suppose there is an invertible $m \times m$ matrix R such that $RL_i = -L_i^t R^t$ and $RM = M^t R^t$, and define $\bar{\psi} = \psi^t R$. The Lagrangian density

$$\mathcal{L}_\psi = \bar{\psi}(\sum_i L_i \frac{\partial}{\partial x^i} + M)\psi \tag{97}$$

is easily shown to have (93) as the Euler-Lagrange equations. Many important physical equations of free non-interacting particles exemplify this particular case. Let now $\mathcal{L}_A$ be a gauge-invariant Lagrangian density for a gauge-invariant theory for the local potential A of a principal G-bundle for the $m \times m$ matrix group G. The sum of the two, $\mathcal{L} = \mathcal{L}_\psi + \mathcal{L}_A$ gives rise to the system of Euler-Lagrange equations consisting of the independent systems (93) and the Euler-Lagrange equations of $\mathcal{L}_A$. Let now $\mathcal{L}'_\psi$ be the result of replacing in $\mathcal{L}_\psi$ the partial derivative $\frac{\partial}{\partial x^i}$ by the covariant derivative $\nabla_i = \frac{\partial}{\partial x^i} + A_i$, then $\mathcal{L}_{(\psi,A)} = \mathcal{L}'_\psi + \mathcal{L}_A$ has as its Euler-Lagrange system one that now couples the fields ψ and A. For instance, if $\mathcal{L}_A$ is the Yang-Mills

Lagrangian, then (91) changes to

$$* d_A * F = j \tag{98}$$

where j is a 1-form called the *current 1-form* constructed from the ψ field (j is proportional to $\sum_i \bar{\psi} L_i \psi \, dx^i$ for our example (97)). This process is called *minimal coupling* and is the method used in physical theories to pass from a set of equations such as (93) representing non-interacting free fields to coupled interacting fields whose interaction is mediated by local gauge potentials as dynamical objects. In this sense physical theories such as the standard model of elementary particles are constructed from practically nothing, at least as far as their general form is concerned. Confining ourselves to second order partial differential equations, the choices for the Lagrangian density of noninteracting systems, such as $\mathcal{L}_\psi$, and the choices for gauge invariant $\mathcal{L}_A$ are severely limited. If interacting theories are to be defined by minimal coupling, then there is only an extremely reduced number of possibilities. It is a most remarkable fact that such theories are so successful in describing the vast majority of physical interactions.

7 Electromagnetism

7.1 Maxwell's Equations

The very first physical gauge theory is classical Maxwell electrodynamics, though it was not originally presented as such. In appropriate physical units, Maxwell's equations for the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are:

$$\nabla \cdot \mathbf{B} = 0 \tag{99}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{100}$$

$$\nabla \cdot \mathbf{E} = \rho \tag{101}$$

$$\nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J} \tag{102}$$

where ρ is the charge density and $\mathbf{J}$ the electric current density. Equations (99) and (100) are know as the *homogeneous* Maxwell's equations and the other two as the *non-homogeneous* ones. One of the immediate consequences of Maxwell's equations is the conservation law for electric charge:

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0$$

From the homogeneous equations we deduce that there is a function V, called the *electric, or scalar, potential*, and a vector field $\mathbf{A}$, called the *magnetic, or vector, potential* such that

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \end{aligned}$$

Our next step is to interpret these equations in four-dimensional space-time, that is, $\mathbb{R}^4$ with a constant pseudo-Riemannian metric with signature $(1,3)$. We introduce linear coordinates $(x^0, x^1, x^2, x^3) = (t, x, y, z)$ such that the corresponding coordinate vector fields $\frac{\partial}{\partial x^\mu}$ are orthonormal. We denote by (D_x, D_y, D_z) the components of a vector $\mathbf{D}$ in $\mathbb{R}^3$.

Consider now the 1-form

$$A = \sum_\mu A_\mu \, dx^\mu = -V \, dt + A_x \, dx + A_y \, dy + A_z \, dz$$

A direct calculation of $F = dA$ now provides

$$\begin{aligned} &F = \frac{1}{2} \sum_{\mu\nu} F_{\mu\nu} \, dx^\mu \wedge dx^\nu = \\ &-E_x \, dt \wedge dx - E_y \, dt \wedge dy - E_z \, dt \wedge dz + \\ &\quad B_x \, dy \wedge dz + B_y \, dz \wedge dx + B_z \, dx \wedge dy \end{aligned} \tag{103}$$

The tensor $\sum_{\mu\nu} F_{\mu\nu} \, dx^\mu \wedge dx^\nu$ is called the *electromagnetic tensor* which combines the electric and the magnetic fields into a single object.

The homogeneous Maxwell's equations are now seen to just be $dF = 0$ which follows immediately from $dF = d^2 A = 0$. To formulate the non-homogeneous equation introduce the 1-form

$$j = \sum_\mu j_\mu \, dx^\mu = -\rho \, dt + J_x \, dx + J_y \, dy + J_z \, dz \tag{104}$$

A simple calculation, using the Hodge star operator, now shows that $\delta F = *d * F = j$ so that Maxwell's equation are now reduced to

$$dF = 0 \tag{105}$$

$$\delta F = j \tag{106}$$

Note that the conservation of charge is now expressed as $\delta j = 0$ which follows immediately from (106) and $\delta^2 = 0$.

The last step in interpreting electromagnetism as a gauge theory is to consider the 1-form A as the local principal gauge potential of an invariant connection on a principal U(1)-bundle and the electromagnetic 2-form F as its curvature. The Lie algebra of U(1) is just the real line $\mathbb{R}$, so the local principal gauge potential and the curvature are just normal real-valued differential forms. Note that U(1) is abelian and so equation (62) for the curvature gives just $F = dA$. The form of a gauge transformation is also simplified. A function $\phi : U \to \mathrm{U}(1)$ can be locally expressed as $\phi = e^{i\Lambda}$. An easy calculation now shows that the corresponding gauge transformation (43) is $A \mapsto A - d\Lambda$. Source-free ($j = 0$) electromagnetism is a Lagrangian theory with the conventional Lagrangian density being

$$\mathcal{L}_{EM} = -\frac{1}{2} F \wedge *F$$

Comparing this with (89) we see that electromagnetism is just a Yang-Mills theory for the group U(1). To get a theory with sources, it is customary to use minimal coupling to couple the electromagnetic field to other fields. Compare (105,106) to (90,98).

7.2 Dirac's Magnetic Monopole

Interpreting electromagnetism as a gauge theory on $P\mathrm{U}(1)$ does not offer any special advantages except when one considers quantum theory or extensions to situations that transcend classical Maxwellian theory. One such advantage is seen in trying to define magnetic monopoles. The homogeneous Maxwell's equations state that there are no magnetic sources, that is, classical electromagnetism has no magnetic monopoles. In analogy with electrostatics, a magnetostatic magnetic monopole situated at the origin in $\mathbb{R}^3$ would be a source of a magnetic field away from the origin given by

$$\mathbf{B} = g \frac{\mathbf{r}}{r^3} \tag{107}$$

Here $\mathbf{r} = (x, y, z)$ is the position vector, r it's norm, and g is a physical constant called the *magnetic charge*. There is no magnetic potential $\mathbf{A}$ defined away from the origin such that $\mathbf{B} = \nabla \times \mathbf{A}$ since then by Stokes's theorem the flux of $\mathbf{B}$ through a spherical surface centered at the origin would be zero, contradicting it's explicit form (107). Dirac's solution to this problem was to introduce a magnetic potential that is singular along a curve (the "Dirac string") joining the origin to infinity. Such a singularity is non-physical since the curve can be chosen arbitrarily. Reformulating the monopole in terms of principal bundles does away with this singularity.

Let M be $\mathbb{R}^3$ minus the origin. Cover M with two open sets which are M minus one of the half z-axes:

$$\begin{aligned} U_s &= M \setminus \{(0,0,z) \mid z > 0\} \\ U_n &= M \setminus \{(0,0,z) \mid z < 0\} \end{aligned}$$

In U_s introduce the 1-form

$$A_s = g\frac{y}{r(r-z)}\,dx - g\frac{x}{r(r-z)}\,dy$$

Likewise in U_n introduce the 1-form

$$A_n = -g\frac{y}{r(r+z)}\,dx + g\frac{x}{r(r+z)}\,dy$$

It is easy to show that in each open set the exterior derivative of each 1-form is precisely $B_x\,dy \wedge dz + B_y\,dz \wedge dx + B_z\,dx \wedge dy$ where $\mathbf{B}$ is the monopole field (107). One has

$$A_n - A_s = 2g\frac{y}{x^2+y^2}\,dx - 2g\frac{x}{x^2+y^2}\,dy \tag{108}$$

Suppose now that $2g$ is an integer, then one easily calculates that (108) is equal to $-d\phi \cdot \phi^{-1}$ where $\phi : U_s \cap U_n \to \mathrm{U}(1)$ is given by

$$\phi(x,y,z) = \frac{(x+iy)^{2g}}{|x+iy|^{2g}}$$

that is, $\phi = e^{2gi\theta}$, where θ is the azimuthal spherical coordinate. But this, by Theorem 10, is just the condition that the two 1-forms define an invariant connection on a principal U(1)-bundle over M. The curvature of this connection is precisely the monopole field (107) interpreted as a 2-form. Note that no singular Dirac string is needed to define the monopole now. The individual 1-forms when interpreted as vector potentials have Dirac stings; for A_s it is the positive z-axis, and for A_n the negative one. The Dirac monopole is not a realistic candidate for a possible physical particle, but more complex gauge theories allow for the existence of realistic magnetic monopoles. Experimental searches for magnetic monopoles have up to now failed to find any. If they do exist, they are very rare in the universe. Note that in the above description the magnetic charge g is *quantized*, that is, it assumes discrete values since $2g$ must be an integer. In a more detailed quantum mechanical analysis, when both electric and magnetic charges are allowed one finds that it is the product of the two charges that must be quantized. Physicists found this result quite intriguing for, as the argument goes, if there is just one magnetic monopole in the universe, quantization of electric charge, observed empirically, follows.

8 Spin

8.1 Clifford Algebras

Let V be a vector space over a base-field $\mathbb{F}$ of characteristic different from 2, β a symmetric bilinear form on V, and $q(v) = \beta(v, v)$ the corresponding quadratic form. Let $T(V) = \sum_{p=0}^{\infty} T^p(V)$ be the full tensor algebra of V where $T^0(V) = \mathbb{F}$ and for $p > 0$, $T^p(V) = V^{p\otimes} = V \otimes \cdots \otimes V$ is the p-fold tensor product of V with itself. Let $I \subset T(V)$ be the two-sided ideal generated by elements of the form $v \otimes v + q(v)$ for all $v \in V$. The *Clifford Algebra* $\mathcal{C}\ell(V, q)$ is defined as the quotient $T(V)/I$. Since $V \cap I = \{0\}$, the inclusion $V \subset T(V)$ descends to an inclusion of V into $\mathcal{C}\ell(V, q)$ and we shall thus identify V with a subspace of $\mathcal{C}\ell(V, q)$. We have for elements of V the fundamental Clifford relations

$$v^2 + q(v) = 0 \tag{109}$$

Substituting $v + u$ for u in (109), one easily deduces the equivalent polarized form of (109):

$$uv + vu + 2\beta(u, v) = 0 \tag{110}$$

Note that when $q = 0$ then $\mathcal{C}\ell(V, q) = \bigwedge(V)$, the exterior algebra of V.

As an arbitrary element of $T(V)$ is a sum of tensor products of elements of V one sees that V generates $\mathcal{C}\ell(V, q)$ as an algebra (by convention an empty product of elements of V is equal to 1)

Let now $v_1, v_2, \ldots, v_m \in V$ and consider the element $v_1 v_2 \cdots v_m \in \mathcal{C}\ell(V, q)$. Let π be any permutation of $\{1, \ldots, m\}$. By repeated use of (110) one sees that

$$v_1 v_2 \cdots v_m = \sigma_\pi v_{\pi(1)} v_{\pi(2)} \cdots v_{\pi(m)} + w \tag{111}$$

where w consists of terms which are products of *at most* $m-2$ factors v_i and σ_π is ± 1 according to whether the permutation is even or odd. Among the vectors $v_1, \ldots, v_m$ there might be some that are equal. One can now use (111) to bring any two such next to each other, and then use (109) to eliminate them in favor of a numerical coefficients. Proceeding in this manner we see that any product $v_1 v_2 \cdots v_m$ can now be rewritten as a linear combination of products of at most m factors chosen from the same set of vector and in which all factors in each product are different.

To save on notation we shall often suppress some or all of the data (V, q) from $\mathcal{C}\ell(V, q)$ when the context makes clear what is meant.

Theorem 14 *$\mathcal{C}\ell(V, q)$ satisfies the following* universal property*. Let $\mathcal{A}$ be any associative algebra with unit e, and $\phi : V \to \mathcal{A}$ a map that satisfies $\phi(v)^2 + q(v)e = 0$, then ϕ extends to a unique algebra homomorphism $\phi^\sharp : \mathcal{C}\ell(V, q) \to \mathcal{A}$.*

Proof: By the universal property of the tensor algebra, ϕ extends to a unique algebra homomorphism $\hat{\phi} : T(V) \to \mathcal{A}$. Now $\hat{\phi}$ vanishes on the ideal I, and so descends to an extension $\phi^\sharp$ to $\mathcal{C}\ell(V, q)$ showing existence. Uniqueness follows immediately from the fact that V is a set of generators for $\mathcal{C}\ell$. Q.E.D

Denote by $\mathrm{O}(\beta)$ the set of linear transformations $\gamma : V \to V$ such that $\beta(\gamma v, \gamma w) = \beta(v, w)$, or equivalently $q(\gamma v) = q(v)$. If q is non-degenerate, then then $\mathrm{O}(\beta)$ is the indicated orthogonal group, otherwise $\mathrm{O}(\beta)$ contains non-invertible elements. In any case it is a semigroup. Let now $\gamma \in \mathrm{O}(\beta)$ and consider the map $V \to \mathcal{C}\ell(V, q)$ given by $v \mapsto \gamma \cdot v \in \mathcal{C}\ell$. Since $(\gamma \cdot v)^2 = q(\gamma \cdot v) = q(v)$, we have by Theorem 14 an extension of this map to all of $\mathcal{C}\ell$ thus associating to each $\gamma \in \mathrm{O}(\beta)$ an element of $\mathrm{End}(\mathcal{C}\ell(V, q))$ defining thus an action of $\mathrm{O}(\beta)$ on $\mathcal{C}\ell$ by algebra endomorphisms (isomorphisms in case q is non-degenerate). We call this the *canonical action* of $\mathrm{O}(\beta)$ on $\mathcal{C}\ell$.

An *ordered semigroup* S is a semigroup with a partial order such that if $a, b \in S$ with $a < b$ then $ra < rb$ and $ar < br$ for all $r \in S$.

Let S be an ordered semigroup. An associative algebra $\mathcal{A}$ is an *S-filtered* algebra if there are *subspaces* $\mathcal{A}^{(s)}$ for $s \in S$ such that $\mathcal{A} = \bigcup_{s \in S} \mathcal{A}^{(s)}$, $\mathcal{A}^{(s)} \mathcal{A}^{(r)} \subset \mathcal{A}^{(sr)}$, and if $r \leq s$ then $\mathcal{A}^{(r)} \subset \mathcal{A}^{(s)}$.

A Clifford algebra $\mathcal{C}\ell$ is canonically an $\mathbb{N}$-filtered algebra where the set of natural numbers forms an ordered semigroup under addition and the usual order. Define $\mathcal{C}\ell^{(r)}$ as the image of $\oplus_{p=0}^{r} T^p(V)$ under the quotient map.

Let S be a semigroup. An associative algebra $\mathcal{A}$ is an *S-graded* algebra if there are vector *subspaces* $\mathcal{A}_s$ for $s \in S$ such that $\mathcal{A} = \oplus_{s \in S} \mathcal{A}_s$ and

$\mathcal{A}_s\mathcal{A}_r \subset \mathcal{A}_{sr}$. Note that if $e \in S$ is the identity then $\mathcal{A}_e$ is a subalgebra. An element $a \in \mathcal{A}$ is said to be *homogeneous* if $a \in \mathcal{A}_s$ for some s, in which case we shall call s the *degree* of a and denote it by $|a|$.

Associated to any S-filtered algebra there is a canonical S-graded algebra $Gr(\mathcal{A}) = \oplus_{s\in S} Gr_s(\mathcal{A})$ where $Gr_s(\mathcal{A}) = \mathcal{A}^{(s)}/(\bigcup_{r<s} \mathcal{A}^{(r)})$.

Theorem 15 $Gr(\mathcal{C}\ell(V,q)) \simeq \bigwedge(V)$

Proof: The map $V^p \to Gr_p(\mathcal{C}\ell)$ given by $(v_1, v_2, \ldots, v_p) \mapsto [v_1 v_2 \cdots v_p]$ is p-linear and, by (111), is anti-symmetric. Thus it descends to a map $\gamma_p : \bigwedge^p(V) \to Gr_p(\mathcal{C}\ell)$. The direct sum of these maps is an algebra homomorphism $\gamma : \bigwedge(V) \to Gr(\mathcal{C}\ell(V,q))$, which is obviously surjective as V generates $\mathcal{C}\ell$. An element ϕ is in the kernel of γ if and only if each p-homogeneous part is in the kernel of γ_p, so suppose ϕ is p-homogeneous. It is thus a sum of terms of the form $\alpha_1 \wedge \cdots \wedge \alpha_p$. Consider the corresponding sum of Clifford products $\alpha_1 \cdots \alpha_p$ in $\mathcal{C}\ell^{(p)}$. If this sum is zero in Gr_p this means that upon a finite number of applications of (111) it can be reduced to an element of $\mathcal{C}\ell^{(p-1)}$. This fact is independent of the actual quadratic form q that is used and is equally valid for $q = 0$ but this means that ϕ was already zero in $\bigwedge^p(V)$ and so γ is injective. Q.E.D

As a corollary we have

Theorem 16 *$\mathcal{C}\ell(V,q)$ and $\bigwedge(V)$ have the same dimension, in particular if V is finite dimensional of dimension n, both algebras have dimension 2^n.*

It is in fact easy to establish a bijection between linear bases of $\mathcal{C}\ell(V,q)$ and $\bigwedge(V)$ which is done in the proof of the following theorem

Theorem 17 *Let $(v_\alpha)_{\alpha\in A}$ be a Hammel basis for V. Assume A is totally ordered. In $\mathcal{C}\ell(V,q)$, the products $v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p}$ with $\alpha_1 < \alpha_2 < \cdots < \alpha_p$ and any $p \geq 0$ form a basis for $\mathcal{C}\ell(V,q)$ as a vector space. By convention the product for $p = 0$ is 1.*

Proof: Since finite tensor products of the v_α form a basis for $T(V)$, the finite Clifford products of the v_α generate $\mathcal{C}\ell(V,q)$. By what was said above, any finite product of the v_α is a linear combination of terms of the form $v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p}$ with $\alpha_1 < \alpha_2 < \cdots < \alpha_p$. The image of $v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p}$ in $\bigwedge(V)$ is $v_{\alpha_1} \wedge v_{\alpha_2} \wedge \cdots \wedge v_{\alpha_p}$. As these images are linearly independent in $\bigwedge(V)$, the original terms must be linearly independent in $\mathcal{C}\ell(V,q)$ and so form a basis. Q.E.D

The most commonly seen gradings are $\mathbb{Z}$ and $\mathbb{N}$ gradings. A Clifford algebra has however a canonical $\mathbb{Z}_2$-grading. Let $\mathcal{C}\ell_0$ be the image under the quotient map of the *even* tensor powers $\oplus_{p=0}^{\infty} T^{2p}(V)$ and $\mathcal{C}\ell_1$ the image of the *odd* tensor powers $\oplus_{p=0}^{\infty} T^{2p+1}(V)$. A $\mathbb{Z}_2$-graded algebra $\mathcal{A}$ is also called a *superalgebra* where $\mathcal{A}_0$ is called the *bosonic subalgebra* and $\mathcal{A}_1$ the *fermionic subspace* (which is not a subalgebra).

Given two $\mathbb{Z}_2$-graded algebras $\mathcal{A}$ and $\mathcal{B}$, the *twisted, or graded, tensor product* $\mathcal{A}\hat{\otimes}\mathcal{B}$ of the two is a $\mathbb{Z}_2$-graded algebra which, as a vector space, coincides with the ordinary tensor product $\mathcal{A} \otimes \mathcal{B}$ but for which the multiplication, given homogeneous elements $a_1, a_2 \in \mathcal{A}$ and $b_1, b_2 \in \mathcal{B}$, is defined by $(a_1 \otimes b_1)(a_2 \otimes b_2) = (-1)^{|a_2||b_1|}(a_1 a_2) \otimes (b_1 b_2)$. The sign on the right hand side is negative exactly when in going from the left to the right-hand side, two fermionic elements exchange position. The $\mathbb{Z}_2$-grading of $\mathcal{C} = \mathcal{A}\hat{\otimes}\mathcal{B}$ is given by $\mathcal{C}_0 = (A_0 \otimes B_0) \oplus (A_1 \otimes B_1)$ and $\mathcal{C}_1 = (A_0 \otimes B_1) \oplus (A_1 \otimes B_0)$.

Suppose now that $V = V_1 \oplus V_2$ is a direct sum decomposition such that if $v = v_1 \oplus v_2$ then $q(v) = q(v_1) + q(v_2)$. We call this an *orthogonal* decomposition. Let q_i for $i = 1, 2$ be the restriction of q to V_i. One has

Theorem 18 *$\mathcal{C}\ell(V, q) \simeq \mathcal{C}\ell(V_1, q_1)\hat{\otimes}\mathcal{C}\ell(V_2, q_2)$.*

Proof: Consider the map $f : V \to \mathcal{C}\ell(V_1, q_1)\hat{\otimes}\mathcal{C}\ell(V_2, q_2)$ given by $f(v) = v_1 \otimes 1 + 1 \otimes v_2$. Since the v_i are fermionic, one calculates that $f(v)^2 = v_1^2 \otimes 1 + 1 \otimes v_2^2 = q(v_1) + q(v_2) = q(v)$. By Theorem 14 this map extends to an algebra homomorphism $f^\sharp : \mathcal{C}\ell(V, q) \to \mathcal{C}\ell(V_1, q_1)\hat{\otimes}\mathcal{C}\ell(V_2, q_2)$. Let $(v_\alpha)_{\alpha \in A}$ be a Hammel basis for V_1 and $(w_\beta)_{\beta \in B}$ a Hammel basis for V_2. Assume A and B disjoint and consider a total order on $A \cup B$ in which any element of A is less than any element of B. By Theorem 17 terms of the form $v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p} w_{\beta_1} w_{\beta_2} \cdots w_{\beta_q}$ with $\alpha_1 < \alpha_2 < \cdots < \alpha_p < \beta_1 < \beta_2 < \cdots < \beta_q$ form a basis for $\mathcal{C}\ell(V, q)$ and the corresponding terms $v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p} \otimes w_{\beta_1} w_{\beta_2} \cdots w_{\beta_q}$ form a basis for $\mathcal{C}\ell(V_1, q_1)\hat{\otimes}\mathcal{C}\ell(V_2, q_2)$. But one has $f^\sharp(v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p} w_{\beta_1} w_{\beta_2} \cdots w_{\beta_q}) = v_{\alpha_1} v_{\alpha_2} \cdots v_{\alpha_p} \otimes w_{\beta_1} w_{\beta_2} \cdots w_{\beta_q}$ and so $f^\sharp$ is an algebra isomorphism. Q.E.D

If V is real and finite dimensional of dimension n, there is a basis $e_1, \ldots, e_n$ such that $\beta(e_i, e_j) = \mathrm{diag}\,(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$ where there are r entries of 1, s entries of -1 and $t = n - r - s$ entries of 0. If V is complex, then one can choose a basis for which $\beta(e_i, e_j) = \mathrm{diag}\,(1, \ldots, 1, 0, \ldots, 0)$ with r entries of 1, and $t = n - r$ entries of 0. The bilinear form β is non-degenerate if and only if $t = 0$. Non-degenerate forms are thus in the real case classified by their signature (r, s) and in the complex case they are all equivalent. Bases such as the ones here introduced are called *orthonormal*. Choose one such. By our previous discussion the products $e_{i_1}e_{i_2}\cdots e_{i_p}$ with strictly increasing indices form a basis for $\mathcal{C}\ell$.

We can now use Theorem 18 to determine the Clifford algebras of real and complex finite-dimensional vector spaces. Assume the quadratic form is non-degenerate. In the real case denote by (r, s) the signature of the form and the corresponding Clifford algebra by $\mathcal{C}\ell_{\mathbb{R}}(r, s)$. In the complex case we denote the Clifford algebra by $\mathcal{C}\ell_{\mathbb{C}}(n)$.

Consider the real case for $n = 1$ and let e be an orthonormal basis. Then $q(e)$ is either 1 or -1. In the first case $e^2 = 1$ and using the basis $f_1 = \frac{1}{2}(1+e)$ and $f_2 = \frac{1}{2}(1-e)$ one finds that $(\alpha_1 f_1+\alpha_2 f_2)(\beta_1 f_1+\beta_2 f_2) = (\alpha_1\beta_1 f_1+\alpha_2\beta_2 f_2)$. Hence $\mathcal{C}\ell_{\mathbb{R}}(1, 0) = \mathbb{R}\oplus\mathbb{R}$. In the second case $e^2 = -1$ and identifying e with the imaginary unit i of the complex numbers, one easily sees that $\mathcal{C}\ell_{\mathbb{R}}(0, 1) = \mathbb{C}$. In the complex case for $n = 1$ the orthonormal basis e satisfies $e^2 = 1$ and proceeding just as in the real case we conclude $\mathcal{C}\ell_{\mathbb{C}}(1) = \mathbb{C}\oplus\mathbb{C}$. By Theorem 18 one concludes $\mathcal{C}\ell_{\mathbb{R}}(r, s) = (\mathbb{R}\oplus\mathbb{R})^{r\hat{\otimes}}\hat{\otimes}\mathbb{C}^{s\hat{\otimes}}$ and $\mathcal{C}\ell_{\mathbb{C}}(n) = (\mathbb{C}\oplus\mathbb{C})^{n\hat{\otimes}}$ which allows us to explicitly calculate each Clifford algebra. We shall not do this calculation here but present the rather intriguing result. Let $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ denote respectively the real, complex and quaternionic field. For any one of these fields $\mathbb{F}$, let $\mathbb{F}(n)$ denote the algebra of $n\times n$ matrices with elements in $\mathbb{F}$. For the real case consider now the following correspondence:

0	1	2	3	4	5	6	7
$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H}\oplus\mathbb{H}$	$\mathbb{H}$	$\mathbb{C}$	$\mathbb{R}$	$\mathbb{R}\oplus\mathbb{R}$

To calculate $\mathcal{C}\ell_{\mathbb{R}}(r, s)$ find $r - s$ modulo 8 on the first line of the table. The Clifford algebra will be either $\mathbb{F}(m)$ or $\mathbb{F}(m)\oplus\mathbb{F}(m)$ where the scheme of the expression is read from the second line of the table and m is chosen so that the resulting dimension of the algebra is exactly 2^{r+s}. For example $\mathcal{C}\ell_{\mathbb{R}}(1, 3)\simeq\mathbb{R}(4)$ and $\mathcal{C}\ell_{\mathbb{R}}(3, 1)\simeq\mathbb{H}(2)$. For the complex algebra $\mathcal{C}\ell_{\mathbb{C}}(n)$, the correspondence is much simpler:

0	1
$\mathbb{C}$	$\mathbb{C}\oplus\mathbb{C}$

where the first line is n modulo 2. Thus $\mathcal{C}\ell_{\mathbb{C}}(4) = \mathbb{C}(4)$.

Let $\mathcal{A}$ be an associative algebra with unit e. An $\mathcal{A}$-*module* is a vector space M, with the same base field $\mathbb{F}$ as $\mathcal{A}$, along with a map $\cdot : \mathcal{A}\times M\to M$ such that

1. $(a \cdot (b \cdot m)) = (ab) \cdot m$
2. $(a + b) \cdot m = a \cdot m + b \cdot m$
3. $(\alpha a) \cdot m = \alpha(a \cdot m)$
4. $e \cdot m = m$

where $a, b \in \mathcal{A}$, $m \in M$, and $\alpha \in \mathbb{F}$.

This definition bears a strong resemblance to the definition a group action and can be thought of as defining the action of an algebra on a vector space by means of linear maps. We shall refer to the map $\cdot$ as an *algebra action*. Much of the following material on $\mathcal{A}$-modules will also bear a strong resemblance to analogous notions concerning group actions. In particular if by $\mathcal{A}^*$ we denote the group of invertible elements of $\mathcal{A}$, then the restriction of the algebra action of an $\mathcal{A}$-module M to $\mathcal{A}^* \times M$ results in a representation of the group $\mathcal{A}^*$.

A vector spaces V is an $\mathrm{End}(V)$-module where for $B \in \mathrm{End}(V)$ and $v \in V$ one defines $B \cdot v = Bv$. In the same way, if $\mathbb{F}$ is any field, then $\mathbb{F}^n$ is a $\mathcal{M}(n, \mathbb{F})$-module. The set of local sections $\Gamma(U)$ of a vector bundle is a $\mathcal{F}(U)$-module of the algebra of maps $U \to \mathbb{F}$ under pointwise multiplication.

Given an $\mathcal{A}$-module M, the map $m \mapsto a \cdot m$ defines an element $L(a) \in \mathrm{End}(M)$ and the map $a \mapsto L(a)$ is an algebra homomorphism $\mathcal{A} \to \mathrm{End}(M)$. Reciprocally given any vector space V and an algebra homomorphism $L : \mathcal{A} \to \mathrm{End}(V)$ one turns V into an $\mathcal{A}$-module by defining $a \cdot v = L(a)v$. $\mathcal{A}$-modules are also know as *representations* of $\mathcal{A}$.

Any algebra $\mathcal{A}$ is itself automatically an $\mathcal{A}$-module where the map $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is algebra multiplication. This action is known as the *regular representation*.

Given an algebra $\mathcal{A}$ with base field $\mathbb{F}$ and an extension of the field $\hat{\mathbb{F}} \supset \mathbb{F}$, by an $\hat{\mathbb{F}}$-module of $\mathcal{A}$ we mean an action of $\mathcal{A}$ on a vector space W over the extended field $\hat{\mathbb{F}}$ by $\hat{\mathbb{F}}$-linear endomorphisms, that is, a homomorphism of $\mathbb{F}$-algebras $\mathcal{A} \to \mathrm{End}_{\hat{\mathbb{F}}}(W)$. The case of interest for us is that of a *complex* module of a *real* algebra.

We say an $\mathcal{A}$-module M is *irreducible* if there is no proper non-zero subspace $W \subset M$ such that $\mathcal{A} \cdot W \subset W$. Given two $\mathcal{A}$-modules M_1 and M_2, an *intertwiner* is a linear map $T : M_1 \to M_2$ such that $a \cdot Tm = T(a \cdot m)$. Intertwiners are morphisms in the category whose objects are $\mathcal{A}$-modules. Two

$\mathcal{A}$-modules are said to be *equivalent* if they are isomorphic in this category. This means there is an invertible intertwiner.

We state without proof the following theorem.

Theorem 19 *Let $\mathbb{F}$ be any of the fields $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ then, up to isomorphism:*

1. *$\mathbb{F}(n)$ has a unique irreducible representation, namely the natural action of $\mathbb{F}(n)$ on $\mathbb{F}^n$*
2. *$\mathbb{F}(n) \oplus \mathbb{F}(n)$ has two irreducible representations given by projection onto one of the summands followed by the natural representation of the summand.*

As a consequence of this theorem we can assert that a Clifford algebra has either one or two irreducible modules depending on whether it is of the form $\mathbb{F}(m)$ or $\mathbb{F}(m) \oplus \mathbb{F}(m)$.

Let q be a quadratic form on a real vector space V. On the complexified space $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ one can introduce the complexified quadratic form $q_{\mathbb{C}}$ defined uniquely by requiring $q_{\mathbb{C}}(v \otimes z) = z^2 q(v)$. Define $\mathcal{C}\ell_{\mathbb{C}}(V, q)$ as $\mathcal{C}\ell(V, q) \otimes_{\mathbb{R}} \mathbb{C}$. One has

Theorem 20 *$\mathcal{C}\ell_{\mathbb{C}}(V, q) \simeq \mathcal{C}\ell(V_{\mathbb{C}}, q_{\mathbb{C}})$.*

Proof: Consider the map $V_{\mathbb{C}} \to \mathcal{C}\ell(V, q) \otimes_{\mathbb{R}} \mathbb{C}$ defined by $v \otimes z \mapsto v \otimes z \in \mathcal{C}\ell(V, q) \otimes_{\mathbb{R}} \mathbb{C}$. One has in $\mathcal{C}\ell(V, q) \otimes_{\mathbb{R}} \mathbb{C}$ that $(v \otimes z)^2 = v^2 \otimes z^2 = -q(v)z^2 = -q_{\mathbb{C}}(v \otimes z)$ which by the universal property of Clifford algebras means that the map extends to an algebra homomorphism $\mathcal{C}\ell(V_{\mathbb{C}}, q_{\mathbb{C}}) \to \mathcal{C}\ell(V, q) \otimes_{\mathbb{R}} \mathbb{C}$. Theorem 17 shows now that this homomorphism establishes a bijection between complex linear bases of the two algebras. Q.E.D

Consider now a *complex* $\mathcal{C}\ell(V, q)$-module W. One sees that there is then a natural unique extension of Clifford multiplication to turn W into a $\mathcal{C}\ell_{\mathbb{C}}(V, q)$-module, just set $(a \otimes z) \cdot w = z(a \cdot w)$.

Let now $e_1, \ldots, e_n$ be an orthonormal basis for V and consider the element $\eta = e_1 e_2 \cdots e_n$. One easily finds $\eta^2 = (-1)^{r + \frac{n(n-1)}{2}}$ where (r, s) is the signature of q. This means that there is an integer m such that the $\mathcal{C}\ell_{\mathbb{C}}(V, q)$ element $\omega = i^m \eta$, called the *volume element*, satisfies $\omega^2 = 1$. If $r + \frac{n(n-1)}{2}$ is even, then one can take $\omega = \eta$ and the volume element belongs to $\mathcal{C}\ell(V, q)$ itself. The volume element is, up to an overall sign, independent of the choice of orthonormal basis. Indeed if $e_1', \ldots, e_n'$ is another basis, then $e_i' = S e_i$ for some $S \in \mathrm{O}(V, q)$. As the e_i anticommute with each other, one has

$\eta' = \det(S)\eta = \pm\eta$. This ambiguity results in a sign ambiguity in the definition of ω. Note that if we change to another basis preserving orientation, then this ambiguity disappears, and so if we deal with a fixed orientation, there is no ambiguity.

Let now

$$p_\pm = \frac{1 \pm \omega}{2} \tag{112}$$

It is easy to verify that $p_\pm$ are idempotents $p_\pm^2 = p_\pm$ that $p_+ p_- = p_- p_+ = 0$ and that $p_+ + p_- = 1$. The sign ambiguity in ω lead to an ambiguity in the distinction between p_+ and p_-. When n is odd, η (and consequently also ω and $p_\pm$) commutes with elements of V and therefore is in the center of the algebra, when n is even, η (and consequently also ω) anticommutes with elements of V. For even n, any element $v \in V$ intertwines p_+ and p_-, that is, $p_+ v = v p_-$ and $p_- v = v p_+$.

If W is any complex module, then one has a canonical decomposition $W = W^+ \oplus W^-$ where $W^\pm = p_\pm W$. For n odd one has $\mathcal{C}\ell(V,q)W^\pm \subset W^\pm$ and so for an irreducible module one of the submodules $W^\pm$ must be zero. For even n, one has $\mathcal{C}\ell(V,q)W^\pm \subset W^\mp$, and since for $v \in V$ with $q(v) \neq 0$ one has $v^2 = -q(v)$, Clifford multiplication by v establishes a linear isomorphism between W^+ and W^-. In this case an irreducible module decomposes into two *subspaces* of equal dimension which are invariant under the *even* part (which is a subalgebra) of the Clifford algebra, and which are mapped into each other under the odd part. Entirely analogous considerations apply whenever W is a real $\mathcal{C}\ell(V,q)$-module and the volume element belongs to $\mathcal{C}\ell(V,q)$ itself.

8.2 Spin Groups

Consider a Clifford algebra $\mathcal{C}\ell(V,q)$ for an n-dimensional vector space V with a non-degenerate quadratic form q. Let α be the algebra isomorphism $\alpha : \mathcal{C}\ell \to \mathcal{C}\ell$ which (see Theorem 14) is induced by the map $V \to \mathcal{C}\ell$ given by $v \mapsto -v$. By the same theorem, $\alpha^2 = I$ as this holds on V. Denote by $\mathcal{C}\ell^*$ the group of invertible elements of the algebra. This contains a subgroup $\tilde{\mathcal{P}}$ of elements ϕ such that $\alpha(\phi)V\phi^{-1} \subset V$. Denote by $\tilde{\mathrm{Ad}}_\phi$, the *twisted* adjoint action $\tilde{\mathrm{Ad}}_\phi \psi = \alpha(\phi)\psi\phi^{-1}$. One has for $v \in V$ with $q(v) \neq 0$ that $v^{-1} = -q(v)^{-1}v$ so $q(v)\tilde{\mathrm{Ad}}_v w = vwv = -2\beta(w,v)v - wv^2 = -2\beta(w,v)v + q(v)w$ and we have

$$\tilde{\mathrm{Ad}}_v w = w - 2\frac{\beta(w,v)}{q(v)}v \tag{113}$$

This shows in particular that $\{v \in V \mid q(v) \neq 0\} \subset \tilde{\mathcal{P}}$. The map $\phi \mapsto \tilde{\mathrm{Ad}}_\phi$ defines a group homomorphism $\tilde{\mathcal{P}} \to \mathrm{GL}(V)$.

Theorem 21 *$\tilde{Ad}(\tilde{\mathcal{P}}) = O(V, q)$*

Proof: Let $w = \tilde{\mathrm{Ad}}_\phi v$, then $q(w) = -w^2 = \alpha(w)w = \alpha(\alpha(\phi)v\phi^{-1})\alpha(\phi)v\phi^{-1}$ which, using the fact that α is a homomorphism and that $\alpha^2 = I$, reduces to $\phi\alpha(v)v\phi^{-1}$ which is equal to $q(v)$. That the image of $\tilde{\mathcal{P}}$ under $\tilde{\mathrm{Ad}}$ is all of $\mathrm{O}(V, q)$ can be deduced from (113) as this formula defines an orthogonal reflection in a hyperplane, and by the Cartan-Dieudonné theorem any element of $\mathrm{O}(V, q)$ is a product of a finite number of such reflections. Q.E.D

Now in a certain sense the group $\tilde{\mathcal{P}}$ is too big as any scalar multiple of an element in $\tilde{\mathcal{P}}$ defines the same element of $\mathrm{O}(V, q)$ under $\tilde{\mathrm{Ad}}$. Just how this redundancy is to be removed to define a more convenient subgroup of $\tilde{\mathcal{P}}$ depends in part on the base field and on the application in mind. We shall here consider essentially two cases, the real finite-dimensional case and the complexified real case.

Let thus V be finite dimensional and q non-degenerate. We define the group $\mathrm{Pin}(V, q) \subset \tilde{\mathcal{P}}$ as being generated by elements $v \in V$ with $q(v) = \pm 1$ and $\mathrm{Spin}(V, q) \subset \mathrm{Pin}(V, q)$ as the subgroup of even elements. We state, without proof:

Theorem 22 *In the real case $\tilde{Ad}$ defines a two-to-one covering of $O(V, q)$ by Pin(V, q) and a two-to-one covering of SO(V, q) by Spin(V, q). Elements ϕ and $-\phi$ map to the same element of the orthogonal groups.*

For the case of Clifford algebras over $\mathbb{C}$ the spin group is not necessarily the most useful object. One often has to deal with a *complexified* Clifford algebra and the underlying real structure can be used to construct a complex extension of the real spin groups.

The group $\mathrm{Pin}^c(V, q) \subset \mathcal{C}\ell_{\mathbb{C}}(V, q)$ is defined as the subgroup generated by $\mathrm{Pin}(V, q) \otimes 1$ and $1 \otimes \mathrm{U}(1)$, and the group $\mathrm{Spin}^c(V, q)$ as the subgroup generated by $\mathrm{Spin}(V, q) \otimes 1$ and $1 \otimes \mathrm{U}(1)$. These group should not be confused with $\mathrm{Pin}(V_{\mathbb{C}}, q_{\mathbb{C}})$ and $\mathrm{Spin}(V_{\mathbb{C}}, q_{\mathbb{C}})$. Note that $u \otimes z$ and $(-u) \otimes (-z)$ define the same element in $\mathrm{Pin}^c(V, q)$. One can show that this is the only ambiguity. As u and $-u$ define the same element of $\mathrm{O}(q)$, one still has canonical maps $\lambda^c : \mathrm{Pin}^c(V, q) \to \mathrm{O}(q)$ and its restriction $\mathrm{Spin}^c(V, q) \to \mathrm{SO}(q)$.

Given a $\mathcal{C}\ell(V, q)$-module W it is often useful to have a symmetric bilinear or hermitian sesquilinear form on W that is invariant under the action of given subgroups (such as Pin or Spin or their subgroups) of $\mathcal{C}\ell^*$. We shall address this question only partially. For any $\mathcal{C}\ell(V,q)$-module W and any $\phi \in \mathcal{C}\ell(V,q)$ we denote by $\hat{\phi}$ the endomorphism in $\mathrm{End}(W)$ that corresponds to the algebra action by ϕ.

Consider first a real $\mathcal{C}\ell(n,0)$-module W and introduce any inner product $(\cdot,\cdot)_0$ on it. Let $e_1,\dots e_n \in \mathbb{R}^n$ be an orthonormal basis. The elements ± 1, along with $\pm e_{i_1}e_{i_2}\cdots e_{i_p}$ with $i_1 < i_2 \cdots i_p$ and $1 \le p \le n$ form, under Clifford multiplication, a finite group of order 2^{n+1} which we shall denote by $\mathbb{G}$. We now define a new inner product

$$(\phi,\psi) = \frac{1}{2^{n+1}} \sum_{\gamma \in \mathbb{G}} (\hat{\gamma}\phi, \hat{\gamma}\psi)_0$$

It is clear now that one has $(\hat{e}_i\phi, \hat{e}_i\psi) = (\phi,\psi)$, in other words, Clifford action by e_i is orthogonal. Furthermore one has $(\hat{e}_i\phi,\psi) = (\hat{e}_i^2\phi, \hat{e}_i\psi) = -(\phi, \hat{e}_i\psi)$ and Clifford action by e_i is anti-symmetric. Let now $v = \sum_i \lambda_i e_i$, then $(\hat{v}\phi, \hat{v}\psi) = \sum_i \lambda_i^2 (\hat{e}_i\phi, \hat{e}_i\psi) + \sum_{i\neq j} \lambda_i\lambda_j (\hat{e}_i\phi, \hat{e}_i\psi)$ The first sum is $(\sum_i \lambda_i^2)(\phi,\psi) = q(v)(\phi,\psi)$ and the second sum vanishes since for $i \neq j$ one has $(\hat{e}_i\phi, \hat{e}_j\psi) = -(\phi, \hat{e}_i\hat{e}_j\psi) = (\phi, \hat{e}_j\hat{e}_i\psi) = -(\hat{e}_j\phi, \hat{e}_i\psi)$. For $q(v) = 1$ therefore one has $(\hat{v}\phi, \hat{v}\psi) = (\phi,\psi)$, and so $(\cdot,\cdot)$ is $\mathrm{Pin}(n,0)$-invariant.

A precisely analogous construction can be used in case of a complex $\mathcal{C}\ell(n,0)$-module W to construct a $\mathrm{Pin}^c(n,0)$-invariant hermitian inner product on W. In this case Clifford action by e_i are anti-hermitian.

We now consider space-time signature and let W be a complex module of $\mathcal{C}\ell(1,n-1)$. Choose an orthonormal basis $e_0, e_1, \dots, e_{n-1}$, starting the labeling for convenience with 0. In what follows Greek indices run from 0 to $n-1$ and roman indices from 1 to $n-1$. In the complexified algebra $\mathcal{C}\ell_{\mathbb{C}}(1,n-1)$, the elements $e_0, ie_1, \dots, ie_{n-1}$ are orthonormal for the complexified quadratic form. So there is a hermitian inner product in W for which all these elements act as anti-self-adjoint isometries. We thus have $\hat{e}_0^* = -\hat{e}_0$ and $\hat{e}_i^* = \hat{e}_i$. From this one has $\hat{e}_0\hat{e}_\mu\hat{e}_0 = \hat{e}_\mu^*$ and so for a *real* vector $v = \sum_\mu v^\mu e_\mu$ one has $\hat{e}_0\hat{v}\hat{e}_0 = \hat{v}^*$. For $a \in W$ define now $\bar{a} \in W'$ by $\bar{a}b = \bar{a}(b) = (a, \hat{e}_0 b)$. The map $a \mapsto \bar{a}$ is an *anti-linear* isomorphism $W \to W'$. The hermitian product $(a,b) \mapsto \bar{a}b$ is non-degenerate but in general not positive definite. Let $v_1, \dots, v_p$ be real vectors and $\phi = v_1 \cdots v_p \in \mathcal{C}\ell$. Let $\phi^t = v_p \cdots v_1$. One has $(\hat{\phi}a, \hat{e}_0\hat{\phi}b) = (a, \hat{\phi}^*\hat{e}_0\hat{\phi}b) = (a, \hat{e}_0\hat{\phi}^t\hat{\phi}b) = \prod_i q(v_i)(a,b)$. This means that the new hermitian product is invariant by action of ϕ if $\prod_i q(v_i) = 1$. This in particular is true, though we shall not prove it here, for the component of identity of $\mathrm{Spin}(1, n-1)$. Because of this, the hermitian product $\bar{a}b$ is extremely useful in physical theories. An entirely analogous treatment can be made for signature $(n-1, 1)$ as well.

8.3 Spin Bundles

Let M be a pseudo-Riemannian manifold with pseudo-metric g. Let q_x be the quadratic form defined in T_xM by the pseudo-metric. One now has the Clifford algebra $\mathcal{C}\ell_x = \mathcal{C}\ell(T_xM, q_x)$ defined at each point x of the manifold. Since the pseudo-metric has the same signature (r, s) at every point, each $\mathcal{C}\ell_x$ is isomorphic to $\mathcal{C}\ell(r, s)$. The Clifford algebras $\mathcal{C}\ell_x$ are fibers of an algebra bundle, which we denote by $\mathcal{C}\ell(TM)$, with fiber $\mathcal{C}\ell(r, s)$. Let U be an open set with coordinate functions $x^1, \ldots, x^n$ and let $v_i = \frac{\partial}{\partial x^i}$ be the coordinate vector fields. Local coordinates for $a \in \mathcal{C}\ell_x$ can be taken to be the x^i along with the 2^n coefficients $a_{i_1 \cdots i_p}$ in the expansion $a = \sum_{p=0}^{n} a_{i_1 \cdots i_p} v_{i_1} \cdots v_{i_p}$. This defines $\mathcal{C}\ell(TM)$ as a manifold. To define the bundle structure, pick a fixed orthonormal basis $f_1, \ldots, f_n$ of $\mathbb{R}^{r+s} \subset \mathcal{C}\ell(r, s)$. Let now U be an open set with n-bein $e_1, \ldots, e_n$ and define, using Theorem 14, the isomorphism $\mathcal{C}\ell_x \to \mathcal{C}\ell(r, s)$ by extension from $e_i \mapsto f_i$. These can now be used to define a map $h_U : \pi^{-1}(U) \to U \times \mathcal{C}\ell(r, s)$, establishing thus the defining trivializations.

By the above construction of $\mathcal{C}\ell(TM)$ it is easy to see that this bundle is associated to the principal $\mathrm{O}(r, s)$-bundle $\mathcal{F}_{\mathrm{O}}(M)$ of ordered orthonormal bases, where the action of $\mathrm{O}(r, s)$ on $\mathcal{C}\ell(r, s)$ is the canonical one. Thus $\mathcal{F}_{\mathrm{O}}(M) \times_\rho \mathcal{C}\ell(r, s) \simeq \mathcal{C}\ell(TM)$, where ρ is the canonical action. For an oriented manifold we can restrict ρ to $\mathrm{O}(r, s)$ and get $\mathcal{C}\ell(TM)$ as a bundle associated to $\mathcal{F}_{\mathrm{SO}}(M)$.

A *spin bundle* on an pseudo-Riemannian manifold M is a vector bundle S, with fiber W, whose fiber W_x at each point x is a $\mathcal{C}\ell_x$-module. This means that there has to be a bundle map $\mathcal{C}\ell(TM) \times S \to S$ which restricted to $\mathcal{C}\ell_x \times W_x$ makes W_x into a $\mathcal{C}\ell_x$-module. One generally assumes that all these pointwise modules are equivalent. We say a spin bundle S is *irreducible* if W is an irreducible module.

An example is $S = \mathcal{C}\ell(TM)$ where the algebra action is algebra multiplication. Now given any $\mathcal{C}\ell(r, s)$-module W it is not in general possible to construct a spin bundle with fiber W so that the pointwise modules are equivalent to W.

Abstracting from the algebra bundle $\mathcal{C}\ell(TM)$, one can state the following problem. Given an algebra bundle A with fiber $\mathcal{A}$, and an $\mathcal{A}$-module W, does there exist a vector bundle S with fiber W and a bundle map $A \times S \to S$ such that restricted to fibers this map makes W_x into a $\mathcal{A}_x$-module equivalent to W. In general the answer is no as the following example shows.

> **Example 8** *Let* $\mathcal{A} = \mathbb{R} \oplus \mathbb{R}$, *and* $X = S^1$. *Consider* S^1 *as* $[0,1]$ *with endpoint identified. Let the bundle* A *be obtained from* $[0,1] \times \mathcal{A}$ *by the identification* $(0,(a,b)) \sim (1,(b,a))$. *Let* $W = \mathbb{R}$ *with algebra action* $(a,b) \cdot r = ar$.

Note that $\mathcal{A}$ has two one-dimensional ideals given by $I_1 = \mathbb{R} \times \{0\}$ and $I_2 = \{0\} \times \mathbb{R}$. Now $I_2 \cdot \mathbb{R} = 0$. Let us call I_2 the *annihilating ideal*. If a bundle S existed which answered to the posed problem, there would be defined at each point an annihilating ideal of $\mathcal{A}_x$ varying smoothly with x. However the bundle of one-dimensional ideals of the $\mathcal{A}_x$ is the double cover $S^1 \to S^1$ and so it has no global section. Thus the bundle S does not exist.

There is no easy answer to the general problem. For the case of Clifford algebras, one knows of topological criteria for orientable manifolds M that guarantee that *any* $\mathcal{C}\ell(r,s)$-module can define a spin bundle. Such manifolds are know as *spin manifolds*. Of particular interest are the irreducible modules. Notice that some modules, such as $\mathcal{C}\ell(r,s)$ itself, always define a spin bundle. Spin manifolds are special in that any module defines one.

An *(r,s)-spin structure* for an oriented pseudo-Riemannian manifold M with signature (r,s) is a principal $\mathrm{Spin}(r,s)$-bundle $P\mathrm{Spin}(r,s)$ over M and a bundle map $\xi : P\mathrm{Spin}(r,s) \to \mathcal{F}_{\mathrm{SO}}(M)$ which is equivariant with respect to the right actions, that is, $\xi(p \cdot u) = \xi(p) \cdot \lambda(u)$, where $\lambda : \mathrm{Spin}(r,s) \to \mathrm{SO}(r,s)$ is the canonical double cover. The manifold M is said to be an *(r,s)-spin manifold* if it has an (r,s)-spin structure. When no signature is mentioned, the Riemannian case is understood.

> **Example 9** *There are two inequivalent spin structures on the circle* S^1. *Pick any Riemannian metric and an orientation. Identify* $\pi : \mathcal{F}_{SO}(S^1) \to S^1$ *with the identity map* $Id : S^1 \to S^1$. *One has* $Spin(1) \simeq \mathbb{Z}_2$ *so there are two inequivalent principal* $Spin(1)$ *bundles given by Example 6. The two inequivalent spin structures* $\xi : PSpin(1) \to \mathcal{F}_{SO}(S^1)$, *each one given by one of the principal bundles, are given by the unique obvious maps.*

In the physical literature the first structure is known as the *Ramond* structure and the second one as the *Neveu-Schwartz* structure.

Notice that given any action of $\mathrm{SO}(r,s)$, one can define a $\mathrm{Spin}(r,s)$ action by composing it with λ. From this it is clear that using the spin structure, any bundle associated to $\mathcal{F}_{\mathrm{SO}}(M)$ can be redefined as a bundle associated to $P\mathrm{Spin}(r,s)$. In particular, this is true of the Clifford bundle $\mathcal{C}\ell(M)$. The action of $\mathrm{Spin}(r,s)$ on $\mathcal{C}\ell(r,s)$ via composition with λ and the canonical action of $\mathrm{SO}(r,s)$ is equivalent to the adjoint action, that is, if $u \in \mathrm{Spin}(r,s)$ and $\phi \in \mathcal{C}\ell(r,s)$ then $\lambda(u)\cdot\phi = \mathrm{Ad}_u\phi = u\phi u^{-1}$. Using this fact we can now create spin bundles from any $\mathcal{C}\ell(r,s)$-module W. Since $\mathrm{Spin}(r,s) \subset \mathcal{C}\ell(r,s)$, W carries a representation, call it ρ of $\mathrm{Spin}(r,s)$ being simply the restriction of the algebra action. We can now form the associated bundle $S = P\mathrm{Spin}(r,s) \times_\rho W$. To define the action of $\mathcal{C}\ell_x$ on W_x making S into a spin bundle, consider the map:

$$\mathrm{Id} \times \cdot : P\mathrm{Spin}(r,s) \times \mathcal{C}\ell(r,s) \times W \to P\mathrm{Spin}(r,s) \times W \tag{114}$$

There is an action of $\mathrm{Spin}(r,s)$ on the left-hand space given by $(r,\phi,z) \mapsto (ru^{-1}, \mathrm{Ad}_u\phi, u\cdot z)$ under the quotient of which one has the associated bundle $\mathcal{C}\ell(M) \times S$. Likewise there is an action on the right-hand space given by $(r,z) \mapsto (ru^{-1}, u\cdot z)$ under the quotient of which one has the associated bundle S. These two actions are compatible with the horizontal map as one has $\mathrm{Ad}_u\phi\cdot(u\cdot z) = (u\phi u^{-1}u)\cdot z = u\cdot(\phi\cdot z)$, which thus defines a bundle map $\mathcal{C}\ell(M)\times S \to S$, and so a spin-bundle.

Example 10 *Using the two spin structures of Example 9 one constructs two inequivalent spin bundles on S^1. As $\mathcal{C}\ell(1,0) \simeq \mathbb{C}$, the two bundles are the quotients of $PSpin(1)\times\mathbb{C}$ which results in the trivial bundle $S^1\times\mathbb{C}$ for the first $P\mathbb{Z}_2$ bundle of Example 6 and for the second in a Moebius-band type construction of Example 2 using $\mathbb{C}$ as the fiber instead of $[-1,1]$.*

The two bundles are inequivalent as Spin(1)-bundles since the cocycle of Example 6 is not a $\mathbb{Z}_2$ coboundary, as is easily seen.

Whereas a spin structure solves the problem of defining a spin bundle starting from any $\mathcal{C}\ell(r,s)$-module, a weaker condition, that of a spin^c structure solves the problem for *complex* modules. Since as was seen before, the algebra action on a complex module W can be extended to an algebra action of $\mathcal{C}\ell_{\mathbb{C}}(r,s)$, one naturally has an action of $\mathrm{Spin}^c(r,s)$ on W extending that of $\mathrm{Spin}(r,s)$.

An (r,s)-*spin^c structure* for an oriented pseudo-Riemannian manifold M with signature (r,s) is a principal $\mathrm{Spin}^c(r,s)$-bundle $P\mathrm{Spin}^c(r,s)$, a principal U(1)-bundle $P\mathrm{U}(1)$ over M, and a bundle map $\xi : P\mathrm{Spin}^c(r,s) \to \mathcal{F}_{\mathrm{SO}}(M) \times P\mathrm{U}(1)$ which is equivariant with respect to the right actions, that is, $\xi(p \cdot (u,z)) = \xi(p) \cdot (\lambda(u), z^2 d)$, where $\lambda : \mathrm{Spin}(r,s) \to \mathrm{SO}(r,s)$ is the canonical double cover. The manifold M is said to be an (r,s)-*spin^c manifold* if it has an (r,s)-spin^c structure.

Suppose now that M is an oriented (r,s)-spin^c manifold and let ξ be a (r,s)-spin^c structure. Because of the canonical map $\lambda^c : \mathrm{Spin}^c(r,s) \to \mathrm{SO}(r,s)$ one sees that any bundle associated to $\mathcal{F}_{\mathrm{SO}}(M)$ can be redefined, using the spin^c structure, as a bundle associated to $P\mathrm{Spin}^c(r,s)$. This of course is again true of the Clifford bundle $\mathcal{C}\ell(M)$. One has also the complexified Clifford bundle $\mathcal{C}\ell(M) \otimes \mathbb{C}$, viewed either as the tensor product of $\mathcal{C}\ell(M)$ with the trivial bundle with fiber $\mathbb{C}$, or as a bundle associated to $P\mathrm{Spin}^c(r,s)$ through the canonical action of $\mathrm{Spin}^c(r,s)$ on $\mathcal{C}\ell_{\mathbb{C}}(r,s)$.

Let now W be a complex $\mathcal{C}\ell(r,s)$-module. We can now carry out a construct entirely analogous to the one that constructed a spin bundle from a spin structure to now create a complex spin bundle from the spin^c structure. To do so, note that there is an action, call it ρ, of $\mathrm{Spin}^c(r,s)$ on W and so one can form the associated bundle $S = P\mathrm{Spin}^c(r,s) \times_\rho W$. We can proceed to define the action of $\mathcal{C}\ell_x \otimes_{\mathbb{R}} \mathbb{C}$ on W_x making S into a complex spin bundle. Similar to what we did before, consider now the map

$$\mathrm{Id} \times \cdot : P\mathrm{Spin}^c(r,s) \times \mathcal{C}\ell_{\mathbb{C}}(r,s) \times W \to P\mathrm{Spin}^c(r,s) \times W$$

We can now repeat almost word by word the paragraph following (114) to define a bundle map $\mathcal{C}\ell_{\mathbb{C}}(M) \times S \to S$, and so a spin-bundle.

If either $r + \frac{n(n-1)}{2}$ is even or we are dealing with a complex module, then either $\mathcal{C}\ell(V,q)$, or respectively $\mathcal{C}\ell(V,q) \otimes \mathbb{C}$, contains a volume element. In the corresponding situations for spin-bundles on oriented pseudo-Riemannian manifolds, one can choose a global section ω of the Clifford bundle which at each point $x \in M$ is the volume element of $\mathcal{C}\ell_x$. This is because there is no sign ambiguity in choosing the volume element if we only consider orthonormal basis with the same orientation. We now have global idempotents $p_\pm$ defined by (112) pointwise. Now $W = W^+ \oplus W^-$ and the corresponding bundle S splits into two sub-bundles $S^\pm = p_\pm S$. If n is even, these have the same fiber dimension, and are then sometimes called *half-spin* bundles.

Connections on spin bundles generally arise through their being associated to principal bundles. Let us first consider a spin structure

$\xi : P\mathrm{Spin}(r,s) \to \mathcal{F}_{\mathrm{SO}}$ on a pseudo-Riemannian manifold (M,g). Let S be a spin-bundle associated to $P\mathrm{Spin}(r,s)$ and ω the Levi-Civita connection on $\mathcal{F}_{\mathrm{SO}}$. Because ξ is two-to-one on fibers and equivariant, parallel transport in $\mathcal{F}_{\mathrm{SO}}$ lifts in a unique way to transport in $P\mathrm{Spin}(r,s)$ and this defines a unique lifting of ω to an invariant connection $\tilde{\omega}$ on $P\mathrm{Spin}(r,s)$ which then is transferred to S in the way explained in Section 4.6. As $\mathcal{C}\ell(M)$ is associated to $\mathcal{F}_{\mathrm{SO}}$, and consequently also to $P\mathrm{Spin}(r,s)$ as explained above in this section, this bundle also gains a connection associated to $\tilde{\omega}$. The connections in $\mathcal{C}\ell(M)$ and S are related through the Leibniz rule:

$$\nabla_{\mathcal{X}} a\psi = (\nabla_{\mathcal{X}} a)\psi + a\nabla_{\mathcal{X}}\psi \tag{115}$$

for $a \in \Gamma(\mathcal{C}\ell(M))$ and $\psi \in \Gamma(S))$ and where the product is Clifford action. To see this note that the Leibniz rule holds for sections of $\mathcal{C}\ell(M) \otimes S$ by virtue of Section 4.10. The Clifford multiplication map $\mathcal{C}\ell(M) \otimes S \to S$ is represented in a trivialization by a constant map $\mu : \mathcal{C}\ell(r,s) \otimes W \to W$. Following the argument of Section 4.10, one sees that $\nabla\mu = 0$ and so the Leibniz rule holds.

Now whereas a spin structure allows us to use the Levi-Civita connection to uniquely define an associated connection on spin bundles, the same is not true for a spin^c structure $\xi : P\mathrm{Spin}^c(r,s) \to \mathcal{F}_{\mathrm{SO}}(M) \times P\mathrm{U}(1)$ because of the presence of of the factor $P\mathrm{U}(1)$. However, given an invariant connection α on $P\mathrm{U}(1)$, one can combine it with the Levi-Civita connection ω on $\mathcal{F}_{\mathrm{SO}}$ to get the product connection (see end of Section 4.6) $\omega \times \alpha$ on $\mathcal{F}_{\mathrm{SO}}(M) \times P\mathrm{U}(1)$, and just as before lift it uniquely by ξ to an invariant connection on $P\mathrm{Spin}^c(r,s)$ which can now be used to define the associated connection on its associated bundles. This reproduces the connection associated to ω on all bundles that are also associated to $P\mathrm{Spin}(r,s)$, in particular those associated to $\mathcal{F}_{\mathrm{SO}}(M)$, but on S the connection is defined by both ω and α. The resulting covariant derivative continues to obey the Leibniz rule (115) for the same reason as in the previous paragraph.

One can use the idea behind the construction of the spin bundles to introduce further structure in them. For instance suppose that on the $\mathcal{C}\ell(r,s)$-module W one has a symmetric bilinear form $(\cdot,\cdot)$ that is invariant with respect to Clifford algebra action by elements of $\mathrm{Spin}(r,s)$. Consider now the map

$$\mathrm{Id} \times (\cdot,\cdot) : P\mathrm{Spin}(r,s) \times W \times W \to P\mathrm{Spin}(r,s) \times \mathbb{R}$$

There is an action of $\mathrm{Spin}(r,s)$ on the left-hand space given by $(r,\phi,\psi) \mapsto (ru^{-1}, u\phi, u\psi)$ under the quotient of which one has the associated bundle $S \times S$. Likewise there is an action on the right-hand space given by $(x,r,s) \mapsto (x, ru^{-1}, s)$ under the quotient of which one has the trivial bundle $M \times \mathbb{R}$. These two actions are compatible with the horizontal map as one has $(u\phi, u\psi) = (\phi, \psi)$ by hypothesis. One now has the bundle map $S \times S \to M \times \mathbb{R}$, defining a smoothly varying bilinear form, which we continue to denote by $(\cdot,\cdot)$, in each fiber W_x. Suppose we have done such a construction, then the pointwise bilinear form can be identified with a section t of $S' \otimes S$ where S' is the dual bundle to S. Now $S' \otimes S = P\mathrm{Spin}(r,s) \times_{\rho^* \otimes \rho} (W' \otimes W)$, where ρ^* is the dual representation. The bundle $S' \otimes S$ inherits a connection ω from the connection on the bundle $P\mathrm{Spin}(r,s)$ which in turn is inherited from the Levi-Civita connection on $\mathcal{F}_{\mathrm{SO}}(M)$. A trivialization of $S' \otimes S$ is obtained as a quotient of $U \times \mathrm{Spin}(r,s) \times (W' \otimes W)$. Let $\mathcal{X}$ be any vector field. One now has in this trivialization $\nabla_{\mathcal{X}} t = \mathcal{X}(t) + (-\rho(\omega(\mathcal{X}))' \otimes I + I \otimes \rho(\omega(\mathcal{X})))t$. The first term vanishes since t is represented by a *constant* section. The vanishing of the second term is just the infinitesimal expression of the invariance of t under the action of $\rho^* \otimes \rho$. Thus $\nabla t = 0$ and so for two sections ϕ and ψ of S and any vector field $\mathcal{X}$ we have

$$\mathcal{X}(\phi,\psi) = (\nabla_{\mathcal{X}}\phi, \psi) + (\phi, \nabla_{\mathcal{X}}\psi) \tag{116}$$

In entirely a similar fashion, starting with a $\mathrm{Spin}^c(r,s)$-invariant hermitian sesquilinear form on a complex $\mathrm{Spin}(r,s)$-module W, one can construct a pointwise hermitian sesquilinear form on the associated spin bundle S, satisfying property (116).

8.4 The Dirac Operator

In the standard model of elementary particles, matter is represented by spinor fields and interactions by connections on principal bundles. The Dirac operator is the basic differential operator acting on spinor fields, and the forces between the particles described by these fields is achieved through the minimal coupling ideas of Section 6.2. This imparts particular importance and usefulness to the Dirac operator. It is remarkable that the Dirac operator has also shown to have a fundamental mathematical importance in manifold theory.

Let S be a spin bundle on a pseudo-Riemannian manifold M. Assume S has a linear connection and let ∇ be the corresponding covariant derivative. Recall (see end of Section 4.10) that one can consider ∇ as a map $\nabla : \Gamma(S) \to \Gamma(T^*M \otimes S)$. The pseudo-Riemannian metric provides an isomorphism $r : T^*M \to TM$ (see Section 5.1) which extends to an isomorphism $r^\sharp : \Gamma(T^*M \otimes S) \to \Gamma(TM \otimes S)$. Clifford algebra action finally gives a map $\Gamma(T^*M \otimes S) \to \Gamma(S)$. We define the Dirac operator $D : \Gamma(S) \to \Gamma(S)$ as the composition of these three maps:

$$\Gamma(S) \xrightarrow{\nabla} \Gamma(T^*M \otimes S) \xrightarrow{r^\sharp} \Gamma(TM \otimes S) \xrightarrow{\cdot} \Gamma(S)$$

It is instructive to calculate the local form of this operator in a trivialization. Let $e_1, \ldots, e_n$ be a set of vector fields in an open set U which at each point form a basis for the tangent space, and let $e^1, \ldots, e^n$ be the corresponding dual 1-forms. Let $\gamma_i(x) \in \mathrm{End}(S_x)$ be the endomorphism that corresponds to Clifford action by e_i, and let ∇_i denote ∇_{e_i}. One has $\nabla\psi = \sum_i e^i \otimes \nabla_i\psi$. Under $r^\sharp$ this becomes $\sum_i r(e^i) \otimes \nabla_i\psi$. One has $r(e^i) = \sum_j g^{ij} e_j$. Under Clifford action one finally has $D\psi = \sum_{ij} g^{ij}\gamma_i\nabla_j\psi$. If we set $\gamma^i = \sum_j g^{ij}\gamma_j$ then one can write $D\psi = \sum_j \gamma^j\nabla_j\psi$. Two particular useful cases is to take for e_i an n-bein or the coordinate vector fields $\frac{\partial}{\partial x^i}$ for a set of local coordinates $x^1, \ldots, x^n$.

Example 11 *A complex representation of $\mathcal{C}\ell(0,3)$ is provided by associating to the canonical orthonormal basis (e_1, e_2, e_3) of $\mathbb{R}^3 \subset \mathcal{C}\ell(0,3)$ the following corresponding 2×2 matrices known as the* Pauli spin matrices*.*

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The corresponding spin bundle on $\mathbb{R}^3$ is just $\mathbb{R}^3 \times \mathbb{C}^2$ and the Dirac operator acting on $\psi : \mathbb{R}^3 \to \mathbb{C}^2$ is then

$$\begin{pmatrix} \frac{\partial}{\partial z} & \frac{\partial}{\partial x} - i\frac{\partial}{\partial y} \\ \frac{\partial}{\partial x} + i\frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix}$$

One easily calculates that $D^2 = \Delta I$ where Δ is the three-dimensional Laplacian. It is this property of being a "square root" of the Laplacian that is at the root of the great usefulness of the this operator.

We briefly describe now the historical case of four-dimensional space-time $\mathbb{R}^4$ with the notations and conventions of Section 7.1, and consider an irreducible complex module of $\mathcal{C}\ell(3,1)$. The reason for choosing signature $(3,1)$ and not that of the space-time, $(1,3)$, is that in the physical literature the Clifford algebra that is normally associated to the quadratic form q is what in the mathematical literature is normally associated to the form $-q$. The reader must also beware that some authors consider the signature of space-time to be $(3,1)$ and some even use imaginary coordinates.

Since $\mathcal{C}\ell_{\mathbb{C}}(4) \simeq \mathbb{C}(4)$ the unique irreducible complex module of $\mathcal{C}\ell(3,1)$ can be taken to be $\mathbb{C}^4$. The spin bundle S is trivial and can be taken to be $\mathbb{R}^4 \times \mathbb{C}^4$. Section of this bundle are called *Dirac spinors*.

The famous *Dirac equation* is

$$(iD - mI)\psi = 0$$

where m is a constant that in a physical particle theory corresponds to the mass of the particle.

Following the discussion at the end of Section 8.2 there is a hermitian product $\bar{\psi}\phi$ on spinors which is invariant under the action of the component of the identity of $\mathrm{Spin}(3,1)$. In the physical literature this is known as *Lorentz invariance*.

The Dirac equation is the Euler-Lagrange equation of a Lagrangian theory with

$$\mathcal{L}_D = \bar{\psi}(iD - mI)\psi\Omega$$

where $\Omega = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$ is the volume element of space-time. By the discussion in Section 8.3 the bundle S splits into a direct sum of two sub-bundles $S = S^+ \oplus S^-$. Sections of these sub-bundles are called *Weyl spinors*. As the Dirac operator maps $S^\pm$ to $S^\mp$, the Dirac equation only makes sense for Weyl spinors if $m = 0$, that is Weyl spinors correspond to massless particles. The group $\mathrm{Spin}(3,1)$ maps each sub-bundle into itself, but the two representation are not equivalent. In the physical literature it is said that they differ by *helicity*, which has to do with the intrinsic angular momentum carried by the particles.

The Dirac equation has a global $U(1)$ symmetry whose action is to multiply ψ by a unimodular complex number. Applying the minimal coupling idea of Section 6.2 to make the U(1) symmetry local, modifies the Dirac equation to

$$(iD + qA - mI)\psi = 0 \tag{117}$$

Where A is a vector field viewed as a section of the Clifford bundle (and so has Clifford algebra action on spinors) and physically identified with the electromagnetic potential of Section 7.1 through the correspondence, provided by the metric, of vector fields and 1-forms. We have introduced a physical constant q identified with the *electric charge*.

There is an antilinear map $\psi \to \psi_c$, which we shall not detail here, that takes a solution of (117) to one of $(iD - qA - mI)\psi = 0$ with the opposite sign of the charge. This map is called *charge conjugation*, though this is a bit of a misnomer, as in particle physics it relates particles and anti-particles and should be more properly called *matter-anti-matter conjugation*, which is defined also for neutral particles. Charge conjugation maps a Weyl spinor of one helicity to one of the other.

As $\mathcal{C}\ell(1,3) \simeq \mathbb{R}(4)$, there is a real irreducible spin bundle based on the module $\mathbb{R}^4$. Sections of this spin bundle are called *Majorana spinors*. In the physical literature, Majorana spinors are generally taken as sections of the complex Dirac spin bundle which satisfy a certain real-linear *reality condition*. As Majorana spinors do not have a global $U(1)$-symmetry, one cannot use the minimal coupling scheme to couple them to the electromagnetic field, and so they correspond to neutral particles, more specifically to particles that are identical to their anti-particles.

Using again the minimal coupling ideas of Section 6.2, the Lagrangian density for a Dirac spinor coupled to electromagnetism is taken to be

$$\mathcal{L} = \bar{\psi}(iD + qA - mI)\psi\Omega - \frac{1}{2}F \wedge *F$$

where F is the electromagnetic 2-form (103). The Euler-Lagrange equations for this Lagrangian are

$$\begin{gathered}(iD + qA - mI)\psi = 0 \\ \delta F = q\bar{\psi} \cdot \psi\end{gathered}$$

where $\bar{\psi} \cdot \psi$ is the 1-form that takes a vector v to $\bar{\psi}v\psi$. This gives an example of the current 1-form (104). This provides a rather accurate theory of electrons and positrons (the anti-particle of the electron) interacting with the electromagnetic field at low energies and treated quasi-classically, that is, the particles are treated quantum-mechanically and the electromagnetic field classically. A full quantum mechanical treatment starts with the same Lagrangian but follows a procedure of *quantization* of both the ψ and A fields and results in a *quantum field theory*, known as *quantum electrodynamics, or QED*, one of the most precise physical theories ever constructed.

Consider now the product $L = S^+ \otimes \mathbb{C}^N$ of S^+ (similar considerations apply to the other half-spin bundle) with the trivial $\mathbb{C}^N$-bundle, and the equation

$$D \otimes I\psi = 0 \tag{118}$$

for a section ψ of L. Introducing the canonical basis $e_1, \ldots, e_N$ of $\mathbb{C}^N$, one can write $\psi = \sum_{i=1}^N \psi_i \otimes e_i$ and (118) corresponds to N independent identical Dirac equations $D\psi_i = 0$. Equation (118) has a global $\mathrm{U}(N)$ symmetry whose action, given $T \in \mathrm{U}(N)$, is $T \cdot \psi = I \otimes T\psi$. We are now in the context of Section 6.2 and can make the symmetry global through minimal coupling introducing an invariant connection on a principal $\mathrm{U}(N)$-bundle. Extend the hermitian product on S^+ to L by setting $(\psi \otimes u, \phi \otimes v) = \bar{\psi}\phi(u, v)$ where $(u, v) = \sum_{i=1}^N u_i^* v_i$ is the usual hermitian inner product on $\mathbb{C}^N$. Equation (118) is the Euler-Lagrange equation of a Lagrangian density given by $\mathcal{L}_\psi = (\psi, (D \otimes I)\psi)\Omega$. The minimally coupled Dirac-Yang-Mills theory is governed by the Lagrangian density

$$\mathcal{L}_{DYM} = (\psi, (D \otimes I + g\hat{A})\psi)\Omega - \frac{1}{2}\operatorname{Tr}(F \wedge *F)$$

where g is a physical constant called the *coupling constant*, $A = \sum_i A_i dx^i$ is the gauge potential of the connection, F the curvature 2-form of the connection, and $\hat{A}$ is defined as the composition

$$\Gamma(L) \xrightarrow{I \otimes A} \Gamma(T^*M \otimes L) \xrightarrow{r^\sharp} \Gamma(TM \otimes L) \xrightarrow{\cdot \otimes I} \Gamma(L)$$

where, for concreteness' sake, $(I \otimes A)(\psi \otimes u) = \sum_i dx^i \otimes \psi \otimes A_i u$, $(\cdot \otimes I)(v \otimes \psi \otimes u) = v\psi \otimes u$, and $r^\sharp$ is the obvious extension of the isomorphism $r : T^*M \to TM$.

The Euler-Lagrange equations of this theory are

$$(D \otimes I + g\hat{A})\psi = 0$$
$$*d_A * F = g(\psi, (\cdot \otimes \cdot)\psi)$$

where the right-hand side of the second equation has to be interpreted as a 1-form whose value at a vector v is $g(\psi, (v \otimes \cdot)\psi)$ which in turn must be interpreted as a hermitian $N \times N$ matrix, an element of $\mathfrak{u}(N)$. Concretely the matrix elements are $g\bar{\psi}_i v \psi_j$. Hermiticity is assured by properties of the hermitian product on S^+.

It is essentially this construct that gave rise to the original Yang- Mills theory. A quantized elaboration of it is behind the Standard Model of elementary particle interactions, the spinor fields representing fundamental matter, and the gauge potentials the forces acting on it. An extra field called the *Higgs field* must be added to provide appropriate masses to the particles, but we won't elaborate on this here.

8.5 The Seiberg-Witten Equations

We are now in condition to introduce the Seiberg-Witten equations. Let M be a 4-dimensional oriented Riemann manifold. Such a manifold always has a spin^c structure, though we shall not prove this here. Chose one such structure and introduce an irreducible complex spin bundle S. Introduce now a connection in $P\mathrm{Spin}^c$ which is a lifting of the product connection in $\mathcal{F}_{\mathcal{O}}(M) \times P\mathrm{U}(1)$ consisting of the Levi-Civita connection on $\mathcal{F}_{\mathcal{O}}(M)$ and an invariant connection α on $P\mathrm{U}(1)$. This connection then induces one in the associated bundle S. By considerations introduced in section 8.1 one can introduce a fiberwise hermitian inner product $(\cdot, \cdot)$ in this bundle satisfying (116) with respect to the covariant derivative. The spin bundle S splits into the direct sum of two half-spin bundles $S^{\pm}$. Denote by D^+_α the Dirac operator restricted to S^+. Let F_α be the curvature 2-from of the connection, and F^+_α its self-dual part. Let now $e_1, \ldots, e_4$ be an orthonormal set of tangent vectors at some point, $e^1, \ldots, e^4$ the dual basis of covectors, and $\psi \in S^+$. Define

$$\sigma(\psi) = \sum_{ij} (e_i\psi, e_j\psi) e^i \wedge e^j$$

It is not hard to verify that $\sigma(\psi)$ is independent of the choice of orthonormal basis and that it is a purely imaginary, self-dual 2-form.

The famous Seiberg-Witten equations for a half-spinor field $\psi \in \Gamma(S^+)$ and the connection α are now:

$$F^+_\alpha = \frac{i}{4}\sigma(\psi) \tag{119}$$

$$D^+_\alpha \psi = 0 \tag{120}$$

Unfortunately we shall not explore the remarkable properties of these equations in these notes.

Appendix

A Basic Conventions

If X is an object of any category, by $\mathrm{End}(X)$ we mean the set of morphism of X to itself, that is $\mathrm{Hom}(X, X)$. By $\mathrm{Aut}(X)$ we mean the set of elements of $\mathrm{End}(X)$ that are invertible. Depending on the category, the sets $\mathrm{End}(X)$ and $\mathrm{Aut}(X)$ may have additional structures which are always assumed. Thus for linear spaces, $\mathrm{End}(X)$ is an algebra. Of course, $\mathrm{Aut}(X)$ is always a group.

If W is a vector space over a field $\mathbb{F}$, then we denote by W' its dual, that is, the set of linear maps $\phi : W \to \mathbb{F}$. We shall sometimes denote $\phi(w)$ by $< \phi, w >$. Suppose W finite dimensional and let $e_1, \ldots, e_n$ be a basis. We denote by $e^1, \ldots, e^n$ the corresponding dual basis, that is, one defined by

$$< e^i, e_j >= \delta^i{}_j$$

where the *Kroneker symbol* $\delta^i{}_j$ is defined as

$$\delta^i{}_j = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

We use the physicist's habit of indicating components by indices, sometimes they are subscripts and sometimes superscripts, as for e^i, e_j, and $\delta^i{}_j$ above. The placement is not capricious, but the reasons will not be explained here. The reader should not confuse a superscript index with a power.

Given $\phi_1, \phi_2, \ldots, \phi_p$ elements of W', our convention as to what $\phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_p$ means as an anti-symmetric p-linear form on W is

$$(\phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_p)(w_1, w_2, \ldots, w_p) = \det_{ij} < \phi_i, w_j > \tag{121}$$

where by $\det_{ij} a_{ij}$ we mean the determinant of the matrix a_{ij}. There is a contending convention in which the right-hand side of (121) is divided by $p!$. There are very good reasons for adopting either one of these, so the reader must beware.

If $A : W \to V$ is a linear map between two vector spaces over $\mathbb{F}$, we denote by $A' : V' \to W'$ the *dual map* defined by $(A'\phi)(x) = \phi(Ax)$. For an $n \times m$ matrix M over $\mathbb{F}$, we denote by M^t the matrix transpose.

The group $\mathrm{Aut}(W)$ of invertible linear maps $W \to W$ is called the *general linear group of* W and shall be denoted by $\mathrm{GL}(W)$. For $W = \mathbb{F}^n$ we shall write $\mathrm{GL}(n,\mathbb{F})$ and when $\mathbb{F}=\mathbb{R}$ we shall simply write $\mathrm{GL}(n)$. The identity of $\mathrm{GL}(W)$ we shall usually denote by I.

All manifolds are assumed to be Hausdorff and paracompact. If M is a manifold we denote by T_xM the space of tangent vectors at $x \in M$ and by TM the *tangent bundle*. Likewise we denote by T_x^*M the set of covectors at $x \in M$ and by T^*M the *cotangent bundle*. We assume the reader is familiar with tensor fields and the corresponding tensor bundles such as $TM \otimes T^*M \otimes T^*M$, etc. Likewise for differential forms.

Adopting convention (121) one has the following formula for a 1-form α and vector fields $\mathcal{X}$ and $\mathcal{Y}$:

$$d\alpha(\mathcal{X},\mathcal{Y}) = \mathcal{X}(\alpha(\mathcal{Y})) - \mathcal{Y}(\alpha(\mathcal{X})) - \alpha([\mathcal{X},\mathcal{Y}]) \tag{122}$$

With the contending convention there would be an overall factor of $\frac{1}{2}$ on the right-hand side.

The analog of (122) for p-forms is

$$\begin{aligned} d\alpha(\mathcal{X}_1,\ldots,\mathcal{X}_p) &= \sum_i (-1)^{i+1}\mathcal{X}_i(\alpha(\mathcal{X}_1,\ldots,\hat{\mathcal{X}}_i,\ldots\mathcal{X}_{p+1}))+ \\ &\sum_{1\le i<j\le p+1} (-1)^{i+j+1}\alpha([\mathcal{X}_i,\mathcal{X}_j],X_1,\ldots,\hat{\mathcal{X}}_i,\ldots,\hat{\mathcal{X}}_j,\ldots,\mathcal{X}_{p+1}) \end{aligned} \tag{123}$$

where a hat over an argument means that it is missing.

B Parameterized Maps

A map $f : X \times Y \to Z$ can be viewed alternatively as a family of maps from Y to Z, parameterized by $x \in X$. Formally, there is an associated map $f^\flat : X \to \mathrm{Hom}(Y,Z)$ given by $f^\flat(x)(y) = f(x,y)$. When Y reduces to a one-point set, f is essentially a map from X to Z, and $f^\flat$ is then thought of as a parameterized family of elements of Z. Under this circumstance one usually drops mention of the set Y. By abuse of notation we shall drop the indicator $\flat$ and use the same symbol for both maps, letting context clarify. Again, depending on the context, the most convenient view of such maps could be either as maps from cartesian products, or as parameterized maps. In fiber-bundle theory, the parameterized version is the most useful for many of the maps encountered there, and we shall use special notational conventions for these. When all such maps are parameterized by the same space X we shall often, for sake of brevity, not indicate the parameter x which is tacitly understood. Thus if $g : X \times W \to Y$ we shall write $f \circ g$ for the parameterized map corresponding to the map $X \times W \to Z$ given by $(x, w) \mapsto f(x, g(x, w))$. Alternatively we can write $f \circ g(x)(w) = (f(x) \circ g(x))(w)$. Similar construct hold when the maps, for each $x \in X$ take values in algebraic objects. Thus if $f(x) \in V$ where V is a vector space, and $L(x) \in \mathrm{End}(V)$ is a parameterized family of endomorphism of V, then by Lf we mean the map $x \mapsto L(x)f(x)$.

C Vector-valued Differential Forms

Let M be a manifold. Recall that a covector at $x \in M$ is a linear functional on T_xM, that is, a linear map $T_xM \to \mathbb{R}$. Let now W be a real vector space. A linear map $\omega : T_xM \to W$ is called a *W-valued covector at x*, or generically a *vector-valued* covector. If one has a W-valued covector $\omega(x)$ for each $x \in U \subset M$ then one speaks of a *W-valued* differential forms in U. Sometimes we shall write ω_x instead of $\omega(x)$.

Let ω be a W-valued form and $\phi \in W'$ be an element of the dual space, then $\phi \circ \omega$ is a usual differential form, that is, for each x, a map $T_xM \to \mathbb{R}$ given by $v \mapsto \phi(\omega_x(v))$. We say that ω is $\mathcal{C}^\infty$ if $\phi \circ \omega$ is $\mathcal{C}^\infty$ for all $\phi \in W^*$.

Higher order *vector-valued p-forms* are defined analogously. A W-valued p-form in an open set U is, for each, $x \in U$ a totally anti-symmetric p-linear map $T_xM \times \cdots \times T_xM \to W$. For any $\phi \in W'$ and any W-valued p-form α, $\phi \circ \alpha$ is an ordinary p-form, and as before we say α is is $\mathcal{C}^\infty$ if $\phi \circ \alpha$ is $\mathcal{C}^\infty$ for all $\phi \in W'$.

The exterior differential d can now be extended to W-valued forms. Given any $\mathcal{C}^\infty$ W-valued p-form α, we define the W-valued $p + 1$-form $d\alpha$ as that form for which for all $\phi \in W'$ one has $\phi \circ (d\alpha) = d(\phi \circ \alpha)$. The reader can easily verify that this defines $d\alpha$ uniquely. As before, $d^2 = 0$.

Index