The Classical Gauge Theoretic Structure of the Standard Model

Elijah Sheridan

May 20, 2022

Contents

1 Introduction

This text is an undergraduate thesis completed by the author to graduate with honors in mathematics at Vanderbilt University. In particular, this is an expository paper encapsulating independent learning—overseen, critically, by Professor Marcelo Disconzi—conducted by the author from January 2021 to May 2022.

This project was motivated by the author's interest in theoretical high energy physics: while the modern paradigms—quantum field theory, in particular the Standard Model—remain mathematically non-rigorous, rigor can in fact be achieved by stepping away from the quantum domain back to a classical setting. That is, a consideration of classical fields rather than quantum fields enables a confident employment of mathematics: thus, the author set out to understand this mathematics, or the area of mathematical gauge theory, which deals with connections on principal bundles. Of course, as physics constituted the original motivation, this text extends a little bit beyond bundles and connections to consider other objects relevant to the Standard Model in particular: namely, Lie group representation theory and spin geometry. However, the emphasis on predominantly on mathematical gauge theory: the latter sections are much shorter and contain fewer proofs.

This paper is organized as follows.

First, the theory of bundles is developed: fibre, vector, principal, and associated bundles are considered, with some attention allocated to the important example of vertical bundles (vector bundles which arise canonically from principal bundles).

Second, the theory of connections on bundles is introduced: in particular, the author chooses a less canonical approach by centering the discussion on the so-called "Atiyah sequence" which clarifies the relationship between Ehresmann connections, connection one-forms, and horizontal lifts. This section includes discussion of important ideas including (but not limited to) local connection forms, curvature, and exterior covariant differentiation, while briefly touching upon how these objects carry over from principal bundles to their associated bundles.

Third, Lie representation theory is briefly reviewed: some general results are stated, and some attention is given to some of the Lie groups most important for the Standard Model: namely, $SU(3), SU(2)$, and $U(1)$.

Fourth, spin geometry is swiftly elaborated upon. Algebraically, Clifford algebras are constructed and the spin group is defined, from which the geometric objects of spin bundles and Dirac operators may be introduced.

Fifth, and finally, qualitative connections between the developed mathematics and the (classical) Standard Model are drawn. First, we define the principal and associated bundles whose sections and connections constitute the domain of the Standard Model Lagrangian. Second, we state the Standard Model Lagrangian as it typically appears in the physical literature and informally comment on how the terms are constructed mathematically from the aforementioned domain.

1.1 Notation and Conventions

In no particular order, we adopt the following notation and conventions.

- All functions, atlases, manifolds, and actions are smooth unless otherwise specified.
- Representations are always assumed to be linear

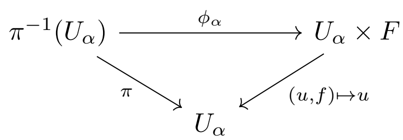
- By proj_i we mean the projection onto the ith factor of a direct sum
- Ex. 2.20 introduces some notation concerning Lie theory
- We say a map between spaces with G-actions is G-equivariant if $f(gx) = g(f(x))$ or $g^{-1}(f(x))$, depending upon whether the respective actions are the same (both right or both left) or different (one right, one left)
- We employ the symbol $\cong$ to denote equivalence in the category of smooth manifolds (that is, via diffeomorphism) unless otherwise specified
- By function, unless otherwise specified, we mean a real-valued map on whatever manifold is presently being discussed (discernible from context)
- By $\mathbb{1}_A$ we always mean the identity function on A: sometimes we will omit the subscript and trust that context reveals the (co)domain
- For wedge products and antisymmetrizations of tensors, we never imply a normalization factor unless we explicitly denote one
- We use [ ] notation to denote cosets in quotient spaces and matrix representations of linear transformations: we trust context will indicate when we are referring to each
- As a moderate abuse of notation, we allow $f^{-1}(m)$ to stand in for the preimage $f^{-1}(\{m\})$
- If we say a symmetric nondegenerate bilinear form has signature (k, ℓ), this denotes k negative eigenvalues and ℓ positive eigenvalues

2 Bundles

2.1 Fibre Bundles

In the same way that manifolds are locally identifiable with Euclidean space, we consider an object which is locally identifiable with a product of manifolds.

Definition 2.1 (Fibre Bundle)**.** A *fibre bundle* is a four-tuple (E, M, F, π) such that E, M, F are manifolds and $\pi : E \to M$ is a surjection admitting an open cover $\{U_\alpha\}_{\alpha \in A}$ of M for which $\pi^{-1}(U_\alpha) \cong U_\alpha \times F$ by a map ϕ_α and the following diagram commutes.



If we wish to refer to the projection of a fibre bundle and the fibre bundle is ambiguous, we will use the notation π_E to denote the projection associated to the fibre bundle E.

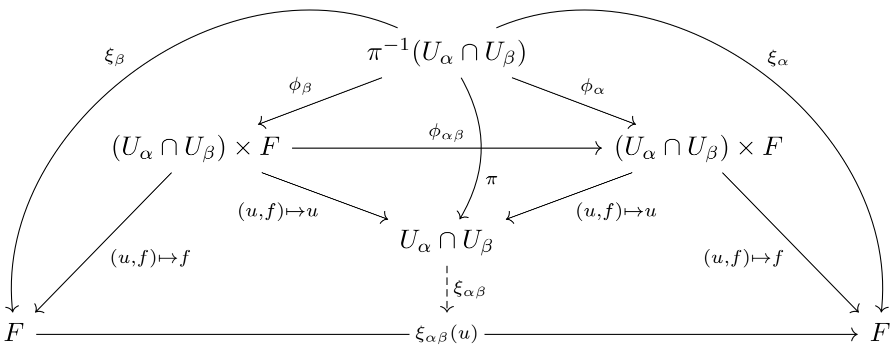
Let (E, M, F, π) be a fibre bundle. In general, we may denote such fibre bundles by $F \hookrightarrow E \to M$, any subset of this diagram, or just by E. If $\{U_\alpha\}$ and $\{\phi_\alpha\}$ (now omitting the indexing set) are the open cover and associated diffeomorphisms, the pairs (U_α, ϕ_α) are each referred to as a *local trivialization*. We refer to E as the *entire space*, M as the *base space*, and F as the *fibre*. Given $p \in E$, note that $\phi_\alpha(p) = (\pi(p), \xi_\alpha(p))$ for some $\xi_\alpha : \pi^{-1}(U_\alpha) \to F$; the name "fibre" comes from the observation that, given $m \in M$, $\xi_\alpha|_{\pi^{-1}(m)}$ yields a diffeomorphism $\pi^{-1}(m) \cong F$: hence, a F is the fibre over m by π. Fibre bundles thus assign a copy of F to each point in M, but in general these copies cannot all be simultaneously canonically identified with each other: instead, this need only be possible locally, through the ϕ_α. We do, however, have a relevant result revealing a case when these fibres can be globally canonically identified.

Theorem 2.2. *If M is a contractible manifold, the fibre bundle $F \hookrightarrow E \to M$ satisfies $E \cong M \times F$.*

Proof. See [11, Cor. 11.6]. □

In physics, frequently the case of a flat spacetime, or $M = \mathbb{R}^{1,3}$ (Minkowski space), is emphasized. As this space is contractible, many of bundles in practice will be globally trivial, or diffeomorphic to a Cartesian product. Nevertheless, fibre bundles prove an important framework ultimately for non-triviality in a geometric sense, rather than a topological sense: we will clarify this notion in time.

Observe that on E we have maps $\phi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1} : (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F$ which necessarily act by $(m, f) \mapsto (m, \xi_{\alpha\beta}(m)(f))$, where $\xi_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{Diff}(F)$ is a map we denote a *transition function*. In particular, $\xi_{\alpha\beta}(m)(f) = (\xi_\alpha \circ \xi_\beta^{-1})(f)$. We summarize these constructions in the following diagram, using a dashed line to denote $\xi_{\alpha\beta}$ to emphasize the fact that it should be understood as being "function-valued" (having codomain $\mathrm{Diff}(F)$) whereas all other functions are "manifold-valued."



Now, notice that transition functions obey $\xi_{\alpha\alpha} = \mathbb{1}$, $\xi_{\alpha\beta} = \xi_{\beta\alpha}^{-1}$, and $\xi_{\alpha\beta} \circ \xi_{\beta\delta} = \xi_{\alpha\delta}$ on triple intersections $U_\alpha \cap U_\beta \cap U_\delta$; this last property is the *cocycle condition*, which is deserving of a name due to the following.

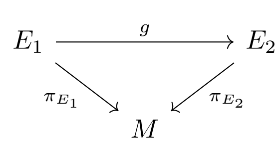
Definition 2.3 (Fibre Bundle, Alternative Definition)**.** A *fibre bundle* is a pair (M, F) of manifolds, an open cover $\{U_\alpha\}$ for M, and a set of maps $\{\xi_{\alpha\beta}\}$ where $\xi_{\alpha\beta} : U_\alpha \cap U_\beta \to \text{Diff}(F)$ and the cocycle condition $\xi_{\alpha\beta} \circ \xi_{\beta\delta} = \xi_{\alpha\delta}$ holds on triple intersections.

Proposition 2.4. Def. 2.1 and Def. 2.3 coincide.

Proof. We have already seen Def. 2.1 implies Def. 2.3, so it suffices to consider the converse. Let (M, F), $\{U_\alpha\}$, and $\{\xi_{\alpha\beta}\}$ constitute a fibre bundle in the sense of Def. 2.3. Let $E = \sqcup_\alpha U_\alpha \times F$ modulo the equivalence relation $(\beta, u, f) \sim (\alpha, u, \xi_{\alpha\beta}(u)(f))$ for $u \in U_\alpha \cap U_\beta$. Define π by the map $[(\alpha, u, f)] \mapsto u$ and form local trivializations by letting $\phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F$ act by $[(\alpha, u, f)] \mapsto (u, f)$, which is well-defined as $[(\alpha, u, f)]$ has only one element with first component α. Then (E, M, F, π) is a fibre bundle in the sense of Def. 2.1. □

It follows that a collection of transition functions satisfying the cocycle condition determine a fibre bundle. Having clarified this structure, we now turn our attention to structure-preserving maps between fibre bundles.

Definition 2.5 (Similar-Base Fibre Bundle Morphism)**.** A *morphism* between fibre bundles E_1, E_2 with the same base M is a map $g : E_1 \to E_2$ such that the following diagram commutes.



The map g is called an *isomorphism* if it is a diffeomorphism.

At times, we will restrict the codomain of the $\xi_{\alpha\beta}$, giving rise to the following definition.

Definition 2.6 (G-atlas)**.** A G-atlas is a covering set of local trivializations $\{(U_\alpha, \phi_\alpha)\}$ on a fibre bundle E such that the maps $\phi_{\alpha\beta}$ induce transition functions $\xi_{\alpha\beta}$ whose codomains are a Lie group G that acts on the fibre F on the left by diffeomorphisms.

Similar to how the discussion of atlases on a manifold canonically proceeds to the definition of a maximal atlas, here we define a G*-bundle* to be a fibre bundle with a maximal G-atlas (that is, a G-atlas that is not properly contained within any other G-atlas). Let our fibre bundle E be equipped with a G-atlas: we then say G is the *structure group* of E.

From here, we move to the notion of "attaching" or "selecting" a particular element of F to each point in M, which will prove fundamental.

Definition 2.7 (Section)**.** A *section* of a fibre bundle E is a map $X : M \to E$ such that $\pi \circ X = \mathbb{1}$.

It follows that a section X maps each $m \in M$ to $X(m) \in \pi^{-1}(M) \cong F$, thereby picking out an element of F (recalling again that these different fibres cannot all simultaneously be identified with F). We let the space of sections on E be denoted $\Gamma(E)$. Def. 2.3 inspires a useful alternative understanding of sections.

Proposition 2.8. Sections $X \in \Gamma(E)$ are equivalent to families of functions $X_\alpha : U_\alpha \to F$ such that $X_\alpha(u) = \xi_{\alpha\beta}(u)(X_\beta(u))$.

Proof. First, we see how sections of E induce the described families of functions. We can exploit local trivializations and understand the $X|_{U_\alpha}$ as a map $U_\alpha \to U_\alpha \times G$, except the first component of this function is necessarily the identity, so all structure is preserved if we merely consider $X_\alpha = \mathrm{proj}_2 \circ \phi_\alpha \circ X$, or $u \mapsto (\xi_\alpha \circ X)(u)$. On an overlap $U_\alpha \cap U_\beta$ we are already assured $\xi_\alpha(p) = \xi_{\alpha\beta}(\pi(p))(\xi_\beta(p))$, and substituting $\pi(p) = u$ and $p = X(u)$ gives the desired result.

To see the other direction, note that the aforementioned families of functions define a section X by defining $X(u) = \phi_\alpha^{-1}(u, X_\alpha(u))$, because the criterion relating X_α, X_β is precisely the one that assures us that X is well-defined, independent of the chart. $\square$

Note that Prop. 2.8 certainly entails that sections are also equivalent to families of maps $X'_\alpha : U_\alpha \to U_\alpha \times F$ such that the first component is the identity and the above property holds in the second component. Both vantages are useful. We refer to the X_α associated with $X \in \Gamma(E)$ as *local sections*.

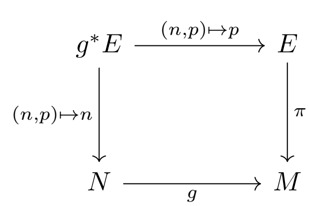
We conclude by considering a particular mechanism for creating new fibre bundles from existing ones.

Definition 2.9 (Pullback Bundle)**.** Given a manifold N and a map $g : N \to M$, the *pullback bundle of* E *by* g, or just the *pullback bundle*, is the set

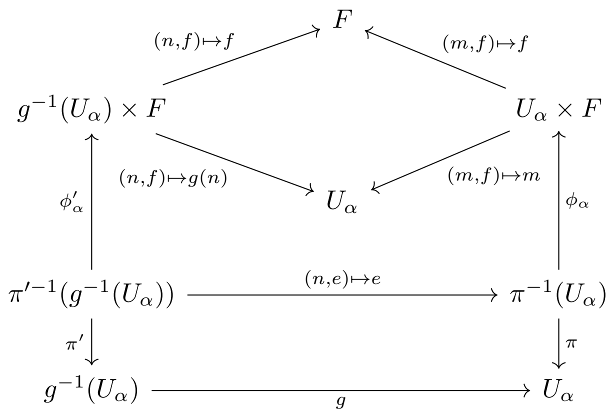
$$g^*E = \{(n, p) \in N \times E \mid g(n) = \pi(p)\} \tag{1}$$

understood as a fibre bundle over N with projection π' given by $(n, p) \mapsto n$ and local trivializations $\{(g^{-1}(U_\alpha), \phi'_\alpha)\}$ for $\phi'_\alpha : \pi'^{-1}(g^{-1}(U_\alpha)) \mapsto g^{-1}(U_\alpha) \times F$ given by $(n, p) \mapsto (n, f)$ where $\phi_\alpha(p) = (m, f) \in U_\alpha \times F$.

Said differently, the pullback bundle g^*E is the subset of $N \times E$ such that the following diagram commutes.



The local trivializations can also be understood as follows, recalling the projection π' on g^*E is the map $(n,p) \mapsto n$.



In essence, the pullback bundle just assigns the fibres $\pi^{-1}(m)$ above m in E to $g^{-1}(m)$ (i.e., the base point goes from m to $g^{-1}(m)$).

We conclude by noting that from any two fibre bundles we are free to form the direct sum of fibre bundles by taking products of the projection map and of local trivializations in the obvious fashion. Similarly straightforward constructions give rise to tensor and exterior products of fibre bundles.

2.2 Vector Bundles

We now consider a class of fibre bundles with a very convenient property: namely, that fibres are endowed with vector space structure. In particular, these bundles will ultimately provide a natural framework for the matter fields arising in gauge theory (although these vector bundles will associated bundles in particular).

Definition 2.10 (Vector Bundle)**.** A *vector bundle of rank k*, or just a *vector bundle* is a fibre bundle $V \hookrightarrow W \to M$ for V a k-dimensional vector space with projection π such that $\pi^{-1}(m)$ has the structure of a k-dimensional vector space for all $m \in M$ and for each local trivialization $\phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times V$ the map $v \mapsto \phi_\alpha^{-1}(m,v)$ from $V \to \pi^{-1}(m)$ is linear.

Let $W \to M$ be a vector bundle. Note that this definition requires that the transition functions associated to a rank-n vector bundle (evaluated at a point $m \in M$) are linear as well (being the composition of linear maps). In particular, vector bundles with fibre (vector space) V could be equivalently characterized as being $GL(V)$-bundles.

Example 2.11 (Tangent Bundle)**.** Given a manifold M, for each $m \in M$ we have the tangent space T_mM. If we take the union $TM = \cup_{m\in M} T_mM$, we refer to TM as the *tangent bundle* to M, TM is a vector bundle of rank $\dim(M)$. In particular, if $\{(U_\alpha, \phi_\alpha)\}$ is a atlas on M and $\pi : TM \to M$ is the projection mapping $v \in T_mM$ to m, then $\{(\pi^{-1}(U_\alpha), \varphi_\alpha)\}$ is a local trivialization where $\varphi_\alpha(v) = (\pi(v), (\phi_\alpha)_*(v)) \in U_\alpha \times \mathbb{R}^n$. Evidently $\pi^{-1}(m) = T_mM$, so we have a vector space structure, and $v \mapsto \varphi_\alpha^{-1}(m,v) = (\varphi_\alpha)_*^{-1}(v)$ is linear, assuring us that TM is indeed a vector bundle.

Definition 2.12 (Vector Subbundle)**.** A *vector subbundle of rank ℓ*, or just a *subbundle* of W is $W' = \cup_{m\in M} W'_m$ where each W'_m is a linear subspace of $\pi^{-1}(m) \subset W$, and W' is a vector bundle with projection $\pi|_{W'}$.

Definition 2.13 (Similar-Base Vector Bundle Morphism)**.** An *(iso)morphism* between vector bundles $W_1 \to M, W_2 \to M$ is an (iso)morphism of fibres bundles such that each $f|_{\pi_{W_1}^{-1}(m)} : \pi_{W_1}^{-1}(m) \mapsto \pi_{W_2}^{-1}(f(m))$ is a linear map.

Said differently, vector bundle morphisms preserve fibres (points in $\pi_{W_1}^{-1}(m)$ are mapped to $\pi_{W_2}^{-1}(m)$) and act linearly on those fibres.

We note that the linearity of vector bundle transition functions assures us that there is a canonical 0 in each fibre $\pi^{-1}(m)$ ($m \in M$), hence we are free to consider the 0 section of a vector bundle, or equivalently, the trivial subbundle $M \times \{0\}$.

Direct sums and tensor products of vector bundles remain vector bundles, as one would expect and hope. For the remainder of the text, we adopt the shorthand $\Gamma(M) = \Gamma(TM)$, $\mathcal{T}_\ell^k(M) = \Gamma([\otimes_{i=1}^k TM] \otimes [\otimes_{i=1}^k T^*M])$, and $\Omega^k(M)$ denoting $\Gamma(\Lambda^k T^*M)$ or, equivalently, the antisymmetrization of $\mathcal{T}_k^0(M)$ (that is, if $\omega \in \Omega^k(M)$, then $\omega(m)$ is a k-form on T_mM).

2.3 Principal Bundles

We now consider a distinct class of bundles which extra structure that equips them to ultimately describe the gauge bosons arising in gauge theory.

Definition 2.14 (Principal G-Bundle)**.** A *principal G-bundle*, or just a principal bundle, is a G-bundle P whose fibre is the Lie group G and whose transition functions are left multiplication by elements of G (i.e., G is the structure group).

Let (P, M, G, π) be a principal G-bundle. Let $p \in P$ satisfy $\pi(p) = u \in U_\alpha$ and $\phi_\alpha(p) = (u, h)$. There is a right action of $G \ni g$ on P given by $pg = \phi_\alpha^{-1}(u, hg)$. We assure ourselves that this action is well-defined by noting that if $p \in U_\beta$ as well, then

$$pg = \phi_\beta^{-1}(u, \xi_{\beta\alpha}(u)(h)g) = \phi_\beta^{-1}(u, \xi_{\beta\alpha}(u)(hg)) = \phi_\alpha^{-1}(u, hg), \tag{2}$$

where the second equality exploits the fact that $\xi_{\beta\alpha}(u)$ is actually an element of G acting by left multiplication, and left and right multiplication in G commute. Note that we will write $\xi_{\beta\alpha}(u)g$ sometimes instead of $\xi_{\beta\alpha}(u)(g)$ because $\xi_{\beta\alpha}(u)$ is a geniune element of G acting by left multiplication on g. Finally, we note the following important fact.

Corollary 2.15. Let (U_α, ϕ_α) be a local trivialization on P: then ξ_α is G-equivariant.

Proof. Let $\phi_\alpha(p) = (m, h)$. Then, evidently, $\xi_\alpha(pg) = \xi_\alpha(\phi_\alpha^{-1}(u, hg)) = hg = \xi_\alpha(p)g$. ☐

We now turn our attention to principal bundle morphisms.

Definition 2.16 (Similar-Base Morphism of Principal Bundles)**.** An *(iso)morphism* between principal bundles $P \to M, Q \to M$ is an (iso)morphism of fibres bundles such that each $f|_{\pi_P^{-1}(m)} : \pi_Q^{-1}(m) \mapsto \pi_W^{-1}(f(m))$ is a group (isomorphism) homomorphism.

Definition 2.17. An *automorphism* of P is a G-equivariant isomorphism $P \to P$.

Example 2.18 (Frame Bundle)**.** Recall that a frame on T_mM is a choice of basis

$$\{v_1, \ldots, v_n\} \subset T_mM \tag{3}$$

Let F_m denote the space of all such bases of T_mM and let $F(TM) = \cup_{m \in M} F_m$. We refer to $F(TM)$ as the *frame bundle of* TM, and we now exhibit its principal $GL(\mathbb{R}^n)$-bundle structure. Certainly we have the natural projection $\pi : F(TM) \to M$ by $v \mapsto m$ if v is a frame for T_mM. Now, let $\{(U_\alpha, \phi_\alpha)\}$ be an atlas for M. Fix $v = \{v_1, \ldots, v_n\} \in F_m \subset F(TM)$ for $m \in U_\alpha$ and observe that to each v_j we can associate $v_j^\alpha \in \mathbb{R}^n$ by $v_j^\alpha = (\phi_\alpha)_*(v_j)$ (identifying tangent spaces to points in $\mathbb{R}^n$ with $\mathbb{R}^n$ via translation). Moreover, because v is a frame, the matrix $v^\alpha \in \mathbb{M}_n$ given by $(v^\alpha)_{ij} = (v_j^\alpha)_i$ is invertible, hence $v^\alpha \in GL(\mathbb{R}^n)$. We can thus define local trivializations $\{(U_\alpha, \psi_\alpha)\}$ on $F(TM)$ by $v \mapsto (\pi(v), v^\alpha) \in U_\alpha \times GL(\mathbb{R}^n)$. If, additionally, $v \in U_\beta$ then observe that $v_j^\alpha = (\phi_\alpha \circ \phi_\beta^{-1})_*(v_j^\beta)$, hence the transition function $\xi_{\alpha\beta}(m) : GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ associated with the overlap $\pi^{-1}(U_\alpha \cap U_\beta)$ is the map $v^\beta \mapsto [\phi_\alpha \circ \phi_\beta^{-1}]v^\alpha$ where $[\phi_\alpha \circ \phi_\beta^{-1}] \in GL(\mathbb{R}^n)$ is the matrix given by expressing the map $(\phi_\alpha \circ \phi_\beta^{-1})_*$ in the bases of $T_{\phi_\alpha(m)}\mathbb{R}^n$, $T_{\phi_\beta(m)}\mathbb{R}^n$ induced by the charts ϕ_α, ϕ_β. This is left multiplication by $GL(\mathbb{R}^n)$, hence $F(TM)$ is a principal $GL(\mathbb{R}^n)$-bundle.

For brevity, we refer to $F(TM)$ as $F(M)$. Note that the above construction can be generalized to arbitrary vector bundles $E \to M$ with fibre V: then the bundle of frames $F(E) \to M$ is a principal $GL(V)$ bundle.

For our convenience in future sections, let R_g, L_g, A_g denote right multiplication, left multiplication, and conjugation on G by $g \in G$ (either on G or upon a set on which G acts, trusting that context will clarify this ambiguity). We conclude by noting that because G acts upon itself freely and transitively, the G action on P must also be free and transitive.

2.4 Associated Bundles

A left action of G on a space W induces a new bundle from P in the following way; as we have alluded to, these bundles will ultimately facilitate the description of matter particles.

Definition 2.19 (Associated Bundle)**.** The *associated bundle* to P by a left action of G on S is the fibre bundle with entire space $E = (P \times V)/G$, where the G action here is $(p, x)g = (pg, g^{-1}x)$; a base space M; a base space open cover $\{U_\alpha\}$ inherited from P; and transition functions $\xi'_{\alpha\beta}$ given by $u \mapsto (x \mapsto \xi_{\alpha\beta}(u)x)$.

Let E be the fibre bundle associated to P by a representation ρ. We sometimes denote this fibre bundle by $P \times_\rho V$, where ρ is the G-action on V. We observe that the transition functions defined above apply the fact that $\xi_{\alpha\beta}(u)$ is an element of G and G acts on V on the left. Additionally, we note that we have appealed to Def. 2.3 (rather than our original definition) in constructing the associated bundle. Finally, it's worth noticing that E is, by definition, a G-bundle, albeit not a principal one.

Example 2.20 (Adjoint Bundle)**.** Given a principal G-bundle, any representation of G upon a vector space induces an associated bundle. To exhibit this, we consider a particular associated bundle arising in this fashion from any principal bundle P.

Letting $\mathrm{Aut}(G)$ denote the automorphism group of the Lie group G, recall the adjoint action of G upon itself: namely, the map $\mathrm{Ad} : G \to \mathrm{Aut}(G)$ given by $g \mapsto \mathrm{Ad}(g)$ such that $\mathrm{Ad}(g)(h) = g^{-1}hg$.

Now, recall that the *Lie algebra* $\mathfrak{g}$ to a *Lie group* G is the space of *left-invariant vector fields* endowed with the Lie bracket $[\,,\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ given by the Lie bracket of vector fields. By left-invariant vector fields we mean vector fields v on G which obey the property $(L_g)_*v(h) = v(gh)$ for $g, h \in G$. It follows that Ad induces a map $\mathrm{ad} : G \to \mathrm{Aut}(\mathfrak{g})$ by $g \mapsto \mathrm{ad}(g) = \mathrm{Ad}(g)_*$. Hence, by Def. 2.19 we are free to consider the bundle associated the principal G-bundle P by the adjoint representation on G on $\mathfrak{g}$.

The *adjoint bundle associated to* P, or just the *adjoint bundle*, is then the vector bundle $\mathfrak{g} \hookrightarrow \operatorname{ad} P \to M$ associated to P by the adjoint representation of G on $\mathfrak{g}$.

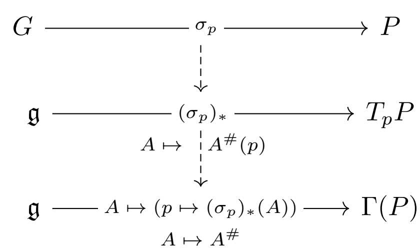
Many familiar examples of vector bundles are associated bundles in disguise: for example, the tangent bundle TM is associated to $F(M)$ by the fundamental representation of $GL(\mathbb{R}^n)$ on $\mathbb{R}^n$, and the tensor bundles $\mathcal{T}^k_\ell(M)$ (tensor are associated to $F(M)$ by tensor products of this representation.

2.5 Vertical Bundles

In anticipation of our discussion of connections, before concluding our consideration of bundles it would be prudent to briefly build some theory for an important particular type of vector bundle: in particular, the vertical bundle, which is a canonical subbundle to the tangent bundle of a principal bundle.

Note that the G action on our principal bundle P entails that $p \in P$ induces a map $\sigma_p : G \to P$ by $g \mapsto pg$. The pushforward of these maps then associate a canonical vector field on P to each element of $\mathfrak{g}$, which we here interpret as T_eG (e being the identity element of G).

Definition 2.21 (Fundamental Vector Field)**.** The *fundamental vector field* associated to $A \in \mathfrak{g}$ on P is the element $A^\# \in \Gamma(P)$ given by $A^\#(p) = (\sigma_p)_*(A)$ for $p \in P$.



We denote elements in the image of fundamental vector fields as *fundamental vectors*. The set of fundamental vectors in a given T_pP coincides with an intrinsically defined canonical linear subspace of T_pP, and the map $(\sigma_p)_*$ identifies $\mathfrak{g}$ with this space isomorphically. Specifically, recall that, given a manifold M and a submanifold $N \subset M$, for $m \in N$ the tangent space T_mN is a linear subspace of T_mM. We have canonical submanifolds $\pi^{-1}(m) \subset P$ for $m \in M$—the fibres of P—and it turns out that the fundamental vector fields form exactly the tangent spaces to this submanifold.

To prove this, though, it is actually useful to introduce an alternative characterization for fundamental vector fields. In particular, observe that if $A \in \mathfrak{g}$ and we interpret A as a tangent vector in T_eG, then if A were the tangent vector $c'_A(0)$[1] to a curve c_A on G, then $A^\#(p)$ would necessarily be the tangent vector $(\sigma \circ c_A)'(0)$. To define $A^\#$ in terms of a curve, then, suffices to find a curve c_A for any A. There is, of course, no unique candidate, but there is a canonical one: namely, $c_A(t) = \exp(At)$. In fact, $\exp : \mathfrak{g} \to G$ is often defined such that this is its fundamental property. Hence, we have the following definition.

Definition 2.22 (Fundamental Vector Field, Alternative Definition)**.** The *fundamental vector field* associated to $A \in \mathfrak{g}$ on P is the element $A^\# \in \Gamma(P)$ such that $A^\#(p)$ is the tangent vector $c'_{p,A}(0)$ where the curve $c_{p,A}$ is given by $t \mapsto p\exp(At)$.

[1]We comment once, as a reminder, that, given a curve g, by $g'(\tau) = \frac{d}{dt} g|_{t=\tau}$ we mean $g_*(1)$ for $1 \in T_\tau\mathbb{R}$.

With this definition in hand, we are now prepared to demonstrate the aforementioned desired result.

Proposition 2.23. The linear map $A \mapsto A^{\#}(p)$ is an isomorphism $\mathfrak{g} \to T_p\,\pi^{-1}(m)$ for $m = \pi(p)$

Proof. Certainly, this is possible because $\pi^{-1}(b) \cong G$, hence

$$\dim T_p\,\pi^{-1}(b) = \dim T_e G = \dim \mathfrak{g}. \tag{4}$$

Then, because the G action on P preserves fibres, the curve $c_{p,A}(t)$ is contained in $\pi^{-1}(m)$ ($\pi(pg) = \pi(p) = m$), hence elements of the image of the pushforward $(c_{p,A})_*$ are necessarily contained in the linear subspace $T_{pg}\pi^{-1}(m) \subset T_{pg}P$.

It suffices now to show that the map is, in fact, surjective on this subspace. Let $v \in T_p\,\pi^{-1}(m)$: it follows that we have a curve $c : (-\varepsilon, \varepsilon) \to \pi^{-1}(m)$ such that $c'(0) = v$. Because the G action on this fibre is transitive (being merely left multiplication), we are assured that for each $t \in (-\varepsilon, \varepsilon)$ there is a $g(t)$ satisfying $c(t) = c(0)g(t) = pg(t)$; in particular, $g(0) = e$. Because this action is smooth, $g(t) : (-\varepsilon, \varepsilon) \to G$ is a smooth curve. Moreover, recalling earlier definitions, we have $c = \sigma_p \circ g$, so

$$v = c'(0) = (\sigma_p \circ g)'(0) = (\sigma_p)_*\,(g'(0)) = A^{\#}(p) \tag{5}$$

as long as we define $A = g'(0) \in \mathfrak{g}$ (identifying the Lie algebra with T_eG). The surjectivity of $A \mapsto A^{\#}(p)$ follows, concluding our proof that this map is an isomorphism. □

This proposition motivates the following definition.

Definition 2.24 (Vertical subspace)**.** The *vertical subspace* of T_pP is $V_p = T_p\,\pi^{-1}(m)$ for $\pi(p) = m$.

To reiterate, Prop. 2.23 illustrates that $V_p \cong \mathfrak{g}$ as vector spaces. Additionally, from the proof of Prop. 2.23 we note that tangent vectors in V_p are necessarily mapped to 0 by π_* (as they are tangent to curves that are trivial when composed with π), giving us the alternative definition $V_p = \ker(\pi_*|_{T_pP})$.

Returning to our vertical subspace, we comment that the name "vertical subspace" itself arises because the fibre, in some sense, extends out "vertically" from the base manifold. Hence, we can understand T_pP as consisting of "vertical vectors" directly tangent to the fibre. We will shortly have a (non-canonical) notion of "horizontal vectors" that are intuitively tangent to the base manifold (enabling a decomposition of the T_pM into vertical and horizontal subspaces, similar to how the tangent space to a product is the product of tangent spaces). This notion will be intimately related to our ultimate definition of a connection.

These vertical subspaces come together to form a bundle in their own right, a subbundle of TP.

Definition 2.25 (Vertical Bundle)**.** The *vertical bundle* of P, VP, is the vector bundle

$$VP = \bigcup_{p \in P} V_p \tag{6}$$

equipped with the local trivializations of TP restricted to V_p in each T_pP.

We refer to vectors in $V_p \subset T_pP$ and sections of VP as vertical. Note that all vertical vector fields are fundamental by Prop. 2.23. We now exhibit an important property of fundamental vector fields (sections of VP). First, however, we require some preliminary ideas.

Definition 2.26. A *flow* on M is a map $\phi(t,m) : (-\varepsilon,\varepsilon) \times M \to M$ such that $\phi(0,m) = m$ and $\phi(t,\phi(s,m)) = \phi(t+s,m)$.

Recall that a vector field v on M induces a flow ϕ on M such that, defining ϕ_m to be the curve $t \mapsto \phi(t,m)$, we have $\phi'_m(t) = v(m(t))$. Indeed, by the existence and uniqueness of ordinary differential equations, such curves exist and form a well-defined $\phi(t,m)$, although we acknowledge that the permitted values of t will, in general, vary with m. We let $\phi_t : M \to M$ denote the diffeomorphism $m \mapsto \phi(t,m)$, and we will often let these maps ϕ_t denote the flow.

Lemma 2.27. Given a vector field $X \in \Gamma(M)$, its associated flow $\phi_t : M \to M$, and a diffeomorphism $g : M \to M$, the vector field $g_*(X)$ has associated flow $g \circ \phi_t \circ g^{-1}$.

Proof. First, recall the following differential geometric fact.

$$X(f) = \lim_{t\to 0} \frac{1}{t}\Big[(f \circ \phi_t)(m) - f(m)\Big]. \tag{7}$$

Now, recall that, because tangent vectors map functions to scalars (as a directional derivative), vector fields map functions to functions. Hence, we can consider $g_*(X)(f)(m)$ for $f : M \to \mathbb{R}$ and $m \in M$: the evaluation of the function $g_*(X)(f)$ at $m \in M$.

$$\begin{aligned} g_*(X)(f)(m) &= X(f \circ g)(g^{-1}(m)) \\ &= \lim_{t\to 0} \frac{1}{t}\Big[([f \circ g] \circ \phi_t)(g^{-1}(m)) - (f \circ g)(g^{-1}(m))\Big] \\ &= \lim_{t\to 0} \frac{1}{t}\Big[(f \circ (g \circ \phi_t \circ g^{-1}))(m) - f(m)\Big]. \end{aligned}$$

It follows that $g \circ \phi_t \circ g^{-1}$ is the flow associated with $g_*(X)$[2]. ∎

This next result is a foundational Lie theoretic result.

Proposition 2.28. Given Lie groups G, H and a Lie group homomorphism $\phi : G \to H$, the following diagram commutes.

$$\begin{array}{ccc} G & \xrightarrow{\ \phi\ } & H \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[\ \phi_*\ ]{} & \mathfrak{h} \end{array}$$

Proof. See Th. 3.32 in [16] ∎

We are now prepared to prove our main result.

[2]Proof inspired by [10, Th. 3.2, Sec. 1.2.3]

Proposition 2.29. The element $\mathrm{ad}(g^{-1})(A) \in \mathfrak{g}$ has fundamental vector field $(R_g)_* A^\#$ for $g \in G$.

Proof. Fix $A \in \mathfrak{g}$ and let $a : \mathbb{R} \to G$ denote the map $t \mapsto \exp(At)$. By Def. 2.22, $A^\#(p)$ is tangent to the curve $\sigma_p \circ a$ at $t = 0$, hence the flow associated to $A^\#$ is $\phi(t, p) = (\sigma_p \circ a)(t) = pa(t)$. It follows that $\phi_t : P \to P$ is then $R_{a(t)}$. Thus, by Lem. 2.27, we are assured that $(R_g)_* A^\#$ has associated flow $R_{ga(t)g^{-1}}$. From here, it follows from Prop. 2.28 that

$$ga(t)g^{-1} = A_{g^{-1}} \exp(At) = \exp(A(g^{-1})_*(At)) = \exp(t\, \mathrm{ad}(g^{-1})(A)). \tag{8}$$

Hence, the flow associated to $(R_a)_* A^\#$ is $R_{\mathrm{ad}(g^{-1})(A)}$, which in turn is the flow associated to the fundamental vector field $\mathrm{ad}(g^{-1})(A)$, delivering the desired result[3]. ☐

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{A \mapsto A^\#} & \Gamma(VP) \\ \big\downarrow{\scriptstyle \mathrm{ad}(g^{-1})} & & \big\downarrow{\scriptstyle (R_g)_*} \\ \mathfrak{g} & \xrightarrow[A \mapsto A^\#]{} & \Gamma(VP) \end{array}$$

We close by commenting on the fact that the adjoint bundle $\mathrm{ad}\, P$ featured in Ex. 2.20 and the vertical bundle VP are actually closely related. First, we can construct a natural vector bundle isomorphism between VP and the trivial vector bundle $P \times \mathfrak{g}$. In particular, any element of VP is given by the evaluation of a fundamental vector field at a point $p \in P$: i.e., every element of VP is of the form $A^\#(p)$, which can merely be mapped to $(p, A) \in P \times \mathfrak{g}$. Now, let G act on $P \times \mathfrak{g}$ in the natural way: $(p, A) \mapsto (pg, \mathrm{ad}(g^{-1})(A))$, the so-called diagonal action. We can then quotient $P \times \mathfrak{g}$ by this action, identifying $(p, A) \sim (p, A)g$. Comparison with Def. 2.19 reveals that this space $(P \times \mathfrak{g})/G$ is exactly $\mathrm{ad}\, P$: we summarize this result as follows.

Proposition 2.30. The quotient of VP by the natural diagonal G action can be identified with $\mathrm{ad}\, P$.

This entails a correspondence on the level of sections.

Corollary 2.31. The space of G-equivariant sections of VP can be identified with the space of sections of $\mathrm{ad}\, P$.

Proof. Given $X \in \Gamma(VP)$, define $\overline{X} : M \to VP/G$ by $\overline{X}(m) = [X(p)]$ for $p \in \pi^{-1}(m)$ given $m \in M$. This is well-defined because for $p, q \in \pi_P^{-1}(m)$, we are assured that $p = qg$ for some $g \in G$, so

$$[X(p)] = [X(qg)] = [((R_g)_* \circ X)(q)] = [X(q)] \tag{9}$$

as $(R_g)_*$ preserves cosets on VP/G. Because VP/G is identifiable with $\mathrm{ad}\, P$, the map $X \mapsto \overline{X}$ sends G-equivariant sections of VP to sections of $\mathrm{ad}\, P$.

Conversely, given $\overline{X} \in \Gamma(\mathrm{ad}\, P)$, we can define a G-equivariant $X \in \Gamma(VP)$ by letting $X(p)$ be such that $[X(p)] = \overline{X}(\pi_P(p))$. This is well-defined because each coset in VP/G contains a unique element in the fibre $\pi_{TP}^{-1}(p)$. In particular, we are merely letting $X(p)$ be that unique element for the coset $[\overline{X}(\pi_P(p))]$. ☐

[3]Proof inspired by [10, Prop. 1.2, Sec. 2.2.1]

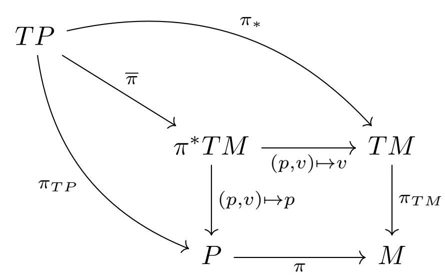
3 Connections

3.1 Defining Connections on Principal Bundles

3.1.1 The Atiyah Sequence

We chose to introduce our first notion of connections—connections on principal bundles—through a slightly longer route than is conventional, building on the bundle theory we have now developed. This has the advantage of elegantly elucidating precisely how and why three equivalent formulations of connections coincide. To achieve this goal, we must construct the so-called Atiyah sequence, a process we now commence. This approach is inspired by [9, Appendix A]

Recalling our principal bundle $P \to M$ with projection π, consider the pullback bundle π^*TM, a vector bundle over P. There is a natural map $\overline{\pi} : TP \to \pi^*TM$ given by $v \mapsto (\pi_{TP}(v), \pi_*(v)) \in P \times TM$. The codomain of $\overline{\pi}$ is indeed π^*TM: if we let $w \in T_pP$ and $\pi(p) = m \in M$, then $\pi_*(w) \in T_mM$, hence $\pi_{TM}(\overline{\pi}(v))$ is indeed $m = \pi(p)$ as we require by Def. 2.9. Said differently, $\overline{\pi}$ is the map that ensures the commutativity of the following diagram.



In particular, by this commutativity and the linearity of pushforwards, the map $\overline{\pi}$ is a morphism between vector bundles of the same base. Moreover, we actually have the following short exact sequence of vector bundles over P.

$$0 \longrightarrow VP \xrightarrow{\iota} TP \xrightarrow{\overline{\pi}} \pi^*TM \longrightarrow 0$$

Here, ι is the natural inclusion of $VP \subset TP$. Now, consider the G action on TP given by $v \mapsto (R_g)_*(v)$ and the G action on π^*TM given by $(p, v) \mapsto (pg, v)$. Because we already established a G action on VP in the discussion leading up to Prop. 2.30 (the adjoint action, achieved through identification with $P \times \mathfrak{g}$, not the restriction of the action on TP), we now have a G action on each of the non-trivial terms in the above sequence. In particular, the two non-trivial maps in our sequence—ι and $\overline{\pi}$—are equivariant with respect to these G actions. That is, the following diagram commutes (identifying VP with $P \times \mathfrak{g}$ as we have shown).

$$\begin{array}{ccccc} VP & \xrightarrow{\iota} & TP & \xrightarrow{\overline{\pi}} & \pi^*TM \\ {\scriptstyle (p,A)\mapsto(pg,\,\mathrm{ad}(g^{-1})(A))}\big\downarrow & & \big\downarrow{\scriptstyle v\,\mapsto\,(R_g)_*(v)} & & \big\downarrow{\scriptstyle (p,v)\mapsto(pg,\,v)} \\ VP & \xrightarrow{\iota} & TP & \xrightarrow{\overline{\pi}} & \pi^*TM \end{array}$$

Note that the left square here follows from Prop. 2.29. The commutativity of the right square follows from a lemma we show now.

Lemma 3.1. The map π_* is invariant under $(R_g)_*$.

Proof. To show this, we exploit local trivializations. Let $p \in P$, $m = \pi(p)$, and (U, ϕ) be a local trivialization of P such that $p \in U$. For the duration of this proof, we identify $\pi^{-1}(U_\alpha)$ with $U_\alpha \times G$, which entails an identification of their respective tangent bundles. Let $v \in T_pP$, hence $v' = (R_g)_*(v) \in T_{pg}P$. It suffices to show π_* acts the same on v, v'. Exploiting our identification, $v = x + y$ and $v' = x' + y'$ for $x, x' \in T_mU$ and $y, y' \in T_hG$ (recalling $T_{(m,h)}U_\alpha \times G \cong T_mU_\alpha \oplus T_hG$ in the sense of vector spaces). Note now that R_g acts on this local trivialization by $(m, h) \mapsto (m, hg)$, hence if $v = c'(t)$ for $c(t) = (m(t), h(t))$ (shrinking the domain of c to keep its image in the local trivialization), then v' is tangent to the curve $c(t)g = (m(t), h(t)g)$ at t. In other words, the first component of the curves corresponding to v, v' coincide, hence x and x' coincide, and in particular

$$\pi_*(v) = \pi_*(x) = \pi(x') = \pi_*(v') \tag{10}$$

because $T_hG \subset T_{(m,h)}U_\alpha \times G$ lies in the kernel of $(\pi_U)_*$. $\square$

The short exact sequence we are considering here is referred to as the *Atiyah sequence*. Actually, it is typically the quotient of this sequence by the G actions which bears that name, but it is equivalent to consider this sequence equipped with equivariant maps, so we follow this simpler route instead[4].

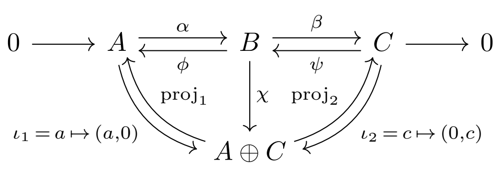
We recall that in the category of, for example, Abelian groups, if we have a short exact sequence

$$0 \longrightarrow A \xrightarrow{\alpha} B \xrightarrow{\beta} C \longrightarrow 0$$

then we say that the sequence splits if we have an isomorphism $B \cong A \oplus C$; moreover, we actually have that the following three are equivalent.

Direct Sum Split:	an isomorphism $\chi : B \to A \oplus C$ such that $\chi \circ \alpha = \iota_1$ and $\beta \circ \chi^{-1} = \mathrm{proj}_2$
Left Split:	a morphism $\phi : B \to A$ such that $\phi \circ \alpha = \mathbb{1}$
Right Split:	a morphism $\psi : C \to B$ such that $\beta \circ \psi = \mathbb{1}$

Note that the domain of $\mathbb{1}$ varies in the above definitions. In particular, χ induces $\phi = \mathrm{proj}_1 \circ \chi$ and $\psi = \chi^{-1} \circ \iota_2$ while ϕ, ψ induce $\chi = \mathrm{im}(\alpha) \oplus \ker(\phi)$ and $\chi = \mathrm{im}(\psi) \oplus \ker(\beta)$, respectively, thereby completing the equivalence. This is all exhibited through the following diagram.

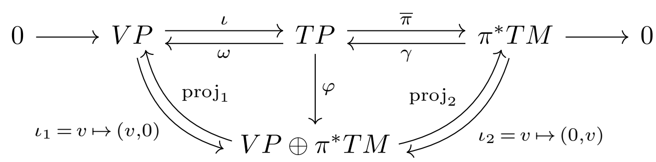


Category theory generalizes this notion with the splitting lemma, which holds for all Abelian categories: while the category of vector bundles over a fixed base manifold isn't Abelian, it ends up being sufficiently nice such that the splitting lemma still holds.

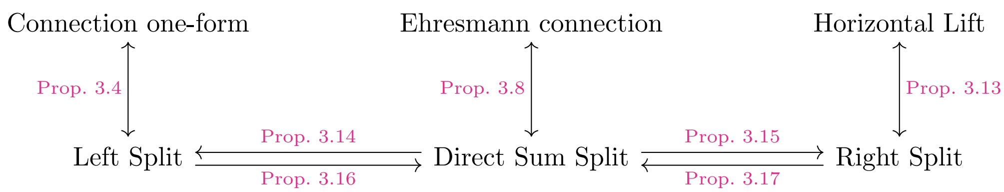
[4]Otherwise, we would have to jump through some unpleasant hoops, like proving that TP/G is a vector bundle.

There are three equivalent notions of a connection on a principal bundle P, which are referred to as an *Ehresmann connection*, a *connection one-form*, and a *horizontal lift*. These three objects correspond to the direct sum split, left split, and right split, respectively, of the Atiyah sequence associated with P that we just developed in the previous section. Moreover, these respective objects induce each other in the same way outlined in the previous paragraph.

Before commencing, for utility, we present the previous diagram in the context of the Atiyah sequence.



In the next three sections, we will show the correspondence between the three equivalent notions of principal bundle connections and the three splits of the Atiyah sequence. In the following section, we prove the splitting lemma in the context of the Atiyah sequence, showing that these splits are all equivalent. A roadmap for our endeavor is as follows.



Given the abstraction of this formalism and approach, we conclude here by providing a concrete example of how this splitting works.

Example 3.2 (Split Short Exact Sequence in Category of Vector Spaces)**.** This example operated in the category of vector spaces: hence, within this example, all morphisms are in that category. Let

$$0 \longrightarrow U \xrightarrow{\alpha} V \xrightarrow{\beta} W \longrightarrow 0$$

be a short exact sequences of vector spaces. Necessarily, then, if k, ℓ, m are the dimensions of U, V, W, respectively, then $\ell = k+m$. We now briefly show how short exact sequences induce each other. We follow the same path specified in the above roadmap: namely, showing the direct sum split implies and is implied by each of the other two splits. Intuitively, α can be thought of as an inclusion and β as a projection with kernel $\alpha(U)$.

We begin by showing how the direct sum split entails left and right splits. Let $\varphi : V \to U \oplus W$ be a direct sum split: this is an isomorphism such that $\varphi \circ \alpha$ is the same thing as the map $U \ni u \mapsto (u, 0) \in U \oplus W$ and $\beta \circ \varphi^{-1}$ is the same thing as the map $U \oplus W \ni (u, w) \mapsto w \in W$. In particular, we can define a distinguished linear subspace in V by $Y = \varphi^{-1}(\{0\} \oplus W)$. Observe that this subspace is necessarily complementary to $X = \alpha(U) = \varphi^{-1}(U \oplus \{0\})$. It is from this linear subspace X that we can construct left and right splits.

Concretely, let $a_1, \ldots, a_k, b_1, v_m$ be a basis for V such that $a_i \in X$ and $b_j \in Y$. From here, we can define a left split ω by $a_i \mapsto \alpha^{-1}(a_i)$ (noting that α is an isomorphism onto its image $\alpha(U)$) and $b_j \mapsto 0$. That is, by distinguishing a complementary subspace Y to $\alpha(U)$ in V, the direct sum split enabled the definition of a projection back onto $\alpha(U) \cong U$[5].

A right split then arises by defining γ by $w \mapsto \varphi^{-1}(0, w)$. Similar to the left split, the distinguishment of the complementary subspace Y permitted the definition of an isomorphism between W and a subspace of V[6].

Now we consider how the left and right splits enable direct sum splits. Certainly our direct sum split φ must be the product map $\alpha^{-1} \times 0$ on $\alpha(U)$ to satisfy the direct sum split definition: thus, it suffices to determine the behavior of φ on a complementary subspace. But the left and right splits each determine a canonical complementary subspace Y: for a left split ω, this is the kernel of ω, and for a right split γ, this is the image of γ. In either case, φ is determined by identifying Y with $\{0\} \oplus W \subset U \oplus W$ (the inverse of our approach when we began with the direct sum split). The precise details of this identification are fixed by the condition $\beta \circ \varphi^{-1} = \mathrm{proj}_2$: namely, on Y we must have that φ must be the product map $0 \times \beta$.

3.1.2 Ehresmann Connection

Ehresmann connections on principal bundles are typically initially introduced as follows.

Definition 3.3 (Ehresmann Connection)**.** A *principal G-bundle Ehresmann connection*, or just an Ehresmann connection, is the selection of a subbundle $HP \subset TP$ of rank n such that $VP + HP = TP$, $(R_g)_*$ leaves HP invariant for $g \in G$, and $H_p = HP \cap T_pP$ depends smoothly on $p \in P$.

We refer to the H_p as horizontal subspaces and their elements as horizontal vectors. Note that it follows from this definition and some linear algebra that $VP \cap HP$ is the trivial subbundle of TP, and that $(R_g)_*(H_p) = H_{pg}$ (that is, $(R_g)_*$ is a surjection between horizontal subspaces). Indeed, it is this trivial intersection that justifies our use of $+$ (as compared to $\oplus$). Additionally, our smoothness condition may seem initially ill-posed, but we really mean that for any $p \in P$ there should be a neighborhood U for which we can select a local set of (smooth) vector fields $v^1, \dots, v^{\dim(M)} \in \Gamma(U)$ such that at each $q \in U$ we have that $v^1(q), \dots, v^{\dim(M)}(q)$ span H_q. Finally, we comment that it follows from the rank-nullity theorem that $\pi_*|_{HP}$ is a vector bundle isomorphism $HP \to TM$: indeed, this holds more generally for any subbundle complementary to VP in TP.

We argue that this notion of a connection coincides with that of an equivariant direct sum split of the Atiyah sequence in the sense defined earlier. That is, Ehresmann connections HP are in correspondence with equivariant maps $\varphi : TP \to VP \oplus \pi^*TM$ (where the G action on the direct sum is the just the canonical action in each space, separately).

Proposition 3.4. There is a bijection between the set of Ehresmann connections on P and the set of equivariant direct sum splits of the Atiyah sequence.

Proof. First, let HP be an Ehresmann connection. Let $u \in TP$: thus, $u = v + h$ for $v \in VP$, $h \in HP$. Define a map $\varphi : TP \to VP \oplus \pi^*TM$ by $v + h \mapsto (v, \overline{\pi}(h)) = (v, (p, \pi_*(h)))$. This is a direct sum split as it is a vector bundle isomorphism that commutes appropriately with

[5] Otherwise, there was no such canonical projection without additional structure, such as an inner product, which would've given us an orthogonal subspace.

[6] The previous footnote is just as relevant here

our full Atiyah sequence diagram. We now exhibit equivariance: observe the following.

$$\varphi((R_g)_*(u)) = \Big(\varphi((R_g)_*(v)), \varphi((R_g)_*(h))\Big) = \mathrm{ad}(g^{-1})(v) + \varphi(h') \tag{11}$$

Note here that $h' \in HP$ because of the $(R_g)_*$ invariance of HP; thus, because $\pi_*|_{HP}$ is a vector bundle isomorphism, there is a unique $w \in TM$ such that $\pi_*(h) = w = \pi_*(h')$ (recalling from Lem. 3.1 that π_* is invariant under $(R_g)_*$). Hence, $\varphi(h') = (0, (pg, \pi_*(h')) = (0, (pg, \pi_*(h))$ and, summarizing,

$$\varphi((R_g)_*(v + h)) = \Big(\mathrm{ad}(g^{-1})(v), (pg, \pi_*(h)\Big) \tag{12}$$

as we desire. Moreover, this induced equivariant direct sum is unique because of the aforementioned uniqueness of $w \in TM$.

Now let φ be a direct sum split of the Atiyah sequence: that is, an equivariant vector bundle isomorphism $TP \to VP \oplus \pi^*TM$. Define an Ehresmann connection HP by $\varphi^{-1}(0 \oplus \pi^*TM)$. Evidently this is rank n and $VP + HP = TP$. Consider invariance under $(R_g)_*$: if we let $h \in HP$, then $h = \varphi^{-1}(0, (p, w))$ for $w = \pi_*(h) \in TM$, then

$$\begin{aligned}(R_g)_*(h) &= (R_g)_*(\varphi^{-1}(0, (p, w))) = (R_g)_*(\varphi^{-1}(0, (p, w))) \\ &= \varphi^{-1}(0, (p, w)g) = \varphi^{-1}(0, (pg, w)) \in HP\end{aligned} \tag{13}$$

Hence, we have $(R_g)_*$ invariance. Finally, to show smooth dependence on p, pick any local frame $\{v^1(q), \ldots, v^n(q)\}$ of TM around $\pi(p)$ (e.g., the one induced by a chart around p on M), pull this frame back to π^*TM to form the local sections $\{(p, v^1(\pi(p))), \ldots, (p, v^n(\pi(p)))\}$, then pull these vector fields back to TP to make $\{\varphi^{-1}(0, (p, v^1(\pi(p)))), \varphi^{-1}(0, (p, v^n(\pi(p))))\}$ and we have local (smooth) vector fields spanning H_q at each q in the neighborhood around p. It follows that HP is an Ehresmann connection. □

3.1.3 Connection One-form

We begin by constructing the notion of a vector-valued form.

Definition 3.5 (E-Valued Form)**.** An *E-valued k-form on M* for a vector bundle $E \mapsto M$ is a section of the vector bundle $E \otimes \mathcal{T}^0_k(M)$.

The name comes from the idea that the evaluation of an E-valued k-form on k vector fields on M leaves us with a section of E. We denote the space of E-valued k-forms on M, by $\Omega^k(M, E)$ We now consider a special, simple case.

Definition 3.6 (Vector-Valued Form)**.** A *vector-valued k-form on M* is an $(M \times V)$-valued k-form on M, where V is some vector space.

We denote the space of vector-valued k-forms on M by $\Omega^k(M, V)$ if V is the vector space. Often times, we merely want our k-forms to return elements of a vector space, without the full structure of a non-trivial vector bundle. Indeed, we immediately apply this new notion; canonically, discussions of connections on principal bundles transition swiftly to the definition of a special vector-valued on P: the connection one-form.

Definition 3.7 (Connection One Form)**.** A *connection one-form on P* is a $\mathfrak{g}$-valued one-form ω on P such that $\omega(A^{\#})(p) = (p, A)$ and ω is equivariant as a map $TP \to \mathfrak{g} \times P$, recalling $\mathfrak{g} \times P \cong VP$ in the sense of vector bundles isomorphisms and the natural action of G on this space discussed adjacent to Prop. 2.30.

Recall that $P \times \mathfrak{g}$ is isomorphic to VP. Additionally, we comment at this point that it follows from Cor. 2.31 that, equivalently, connection one-forms are $\operatorname{ad} P$-valued one forms on M.

A connection HP induces a connection one-form ω (and vice versa) by the definition $\ker(\omega) = HP$, but we will return to this correspondence later. Right now, we are interested in showing the equivalence between connection one-forms and equivariant left splits of the Atiyah sequence, or equivariant vector bundle morphisms $\omega : TP \to VP$. This is a rather straightforward task.

Proposition 3.8. There is a bijection between the set of connection one-forms on P and the set of equivariant left splits of the Atiyah sequence.

Proof. First, let ω be a connection one-form on P. This can already be understood as a vector bundle morphism $TP \to VP$, the necessary composition law is automatically satisfied by the requirement $\omega(A^{\#})(p) = (p, A)$, and equivariance follows trivially from the very definition of a connection one-form.

Now, let ω be a left split of the Atiyah sequence: that is, an equivariant vector bundle morphisms $TP \to VP$. As a linear map $TP \to P \times \mathfrak{g}$, we immediately have that this is a $\mathfrak{g}$-valued one-form on P. Moreover, because $\iota \circ \omega$ must be the identity, we have the desired equality $\omega(A^{\#})(p) = (p, A)$, and the equivariance is again trivially assured. □

We conclude by noting a particular important way in which the horizontal lift can act.

Definition 3.9 (Horizontal Lift of Vector Field)**.** Given a horizontal lift γ, the *horizontal lift of a vector field* $V \in \Gamma(M)$, denoted by $\tilde{V}$, is given by $p \mapsto \gamma((p, V(\pi_P(p))))$

3.1.4 Horizontal Lift

We introduce the general notion of a lift.

Definition 3.10 (Lift of Curves)**.** Given a curve $c : (-\varepsilon, \varepsilon) \to M$, a *lift* of c to $p \in \pi^{-1}(c(0))$ in P is a curve $\bar{c}_p : (-\varepsilon, \varepsilon) \to P$ such that $\pi_*(\bar{c}'(t)) = c'(t)$ and $\bar{c}_{pg}(t) = \bar{c}_p(t)g$.

In the presence of a connection, we say a lift is horizontal if the tangent vectors to any lifted curve are horizontal. In fact, connections give rise to unique horizontal lifts, and horizontal lifts give rise to connections by taking the union of all horizontal lifts of all curves on M. Among lifts, horizontal lifts tend to be the ones worth emphasizing.

Horizontal lifting of curves actually reduces to the following notion of horizontal lifting of individual vectors. We formalize this as follows.

Definition 3.11 (Lift of Vectors)**.** A *horizontal lift* is a bundle morphism $\gamma : \pi^*TM \to TP$ such that $\overline{\pi} \circ \gamma = \mathbb{1}$ and $\gamma(pg, w) = (R_g)_*\gamma(p, w)$.

Given a vector $v \in T_mM$, the lift allows us to "lift" v into any tangent space T_pP for any p above m. The pullback bundle π^*TM is precisely the natural object to describe the pairs (p, v) which are both "above" m in their respective bundles. This lift is termed horizontal because its inverse is $\overline{\pi}$, whose kernel consists precisely of vertical vectors, hence the image of the lift is complementary to these vertical vectors, or is "horizontal."

We then have the following result.

Proposition 3.12. Horizontal lifts (of vectors) give rise to a unique horizontal lift for any curve on M: that is, a lift such that the tangent vectors to the lifted curves all lie within the image of the lift (of vectors).

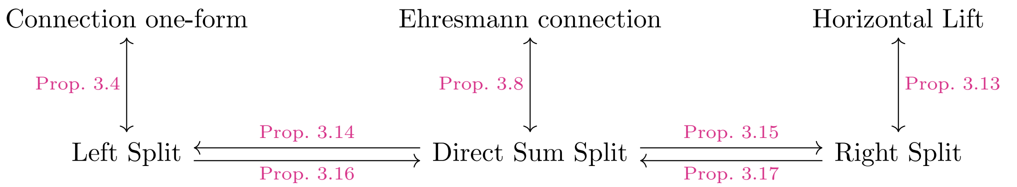
Proof. See Prop. 2.1 in Chapter 2.2 of [10]. ☐

There really isn't much to say regarding the equivalence between horizontal lifts and left splits of the Atiyah sequence: it follows directly.

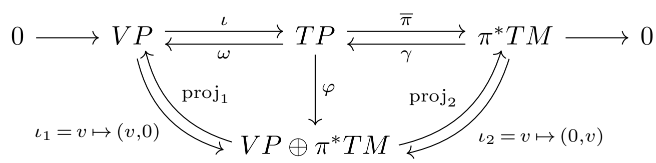
Proposition 3.13. There is a bijection between the set of horizontal lift operations on P and the set of equivariant right splits of the Atiyah sequence.

3.1.5 Equivalence of Connection Definitions

We now prove a series of propositions assuring us that our three notions of connection—Ehresmann connection, connection one-form, and horizontal lift—coincide. That is, we complete the bottom row of our diagram of correspondences, which for convenience we recall here.



Additionally, also for convenience, we recall the Atiyah sequence and its splits. In particular, this diagram exhibits how right and left splits are formed from direct sum splits in obvious fashions.



Proposition 3.14. Direct sum splits φ induce right splits ω in the Atiyah sequence.

Proof. Given a direct sum split φ, take $\omega = \text{proj}_1 \circ \varphi$. Recalling our first map in the sequence, $\iota : VP \to TP$, note that $\omega \circ \iota = \text{proj}_1 \circ \varphi \circ \iota = \text{proj}_1 \circ \iota_1 = \mathbb{1}$, where $\iota_1 : VP \to VP \oplus \pi^*TM$ is the canonical inclusion. Moreover, ω is necessarily an equivariant morphism, being the composition of equivariant morphisms. ☐

Proposition 3.15. Direct sum splits φ induce left splits γ in the Atiyah sequence.

Proof. Given a direct sum split φ, take $\gamma = \varphi^{-1} \circ \iota_2$, recalling $\iota_2 : \pi^*TM \to VP \oplus \pi^*TM$ is the canonical inclusion. Recalling our first map in the sequence, $\iota : VP \to TP$, note that $\overline{\pi} \circ \gamma = \overline{\pi} \circ \varphi^{-1} \circ \iota_2 = \text{proj}_2 \circ \iota_2 = \mathbb{1}$. Moreover, γ is necessarily an equivariant morphism, being the composition of equivariant morphisms. ☐

Proposition 3.16. Left splits ω induce direct sum splits φ in the Atiyah sequence.

Proof. By the rank-nullity theorem, a left split ω induces the decomposition of TP into $\text{Im}(\iota) \oplus \ker(\omega) \cong VP \oplus \ker(\omega)$ (isomorphic as vector bundles). We can define φ by mapping $(v, h) \in VP \oplus \ker(\omega)$ to $(v, \overline{\pi}(h))$. Also due to rank-nullity (as previously mentioned), $\overline{\pi}$ is an (equivariant) vector bundle isomorphism when restricted to a subbundle complementary to VP, hence φ is an equivariant vector bundle isomorphism (being equivariant in each component). Finally, we certainly have that $\iota \circ \varphi$ is the canonical injection ι_1 and $\overline{\pi} \circ \varphi^{-1} = \text{proj}_2$. ☐

Proposition 3.17. Right splits γ induce direct sum splits HP in the Atiyah sequence.

Proof. This argument effectively coincides with that of Prop. 3.16, with the decomposition $TP \cong \mathrm{Im}(\iota) \oplus \ker(\omega) \cong VP \oplus \ker(\omega)$ replaced by $TP \cong \ker(\overline{\pi}) \oplus \mathrm{Im}(\gamma) \cong VP \oplus \mathrm{Im}(\gamma)$. Equivariance and diagram commutativity follow readily. □

Having exhibited this equivalence, we will now refer to principal bundle connections in general, by which we mean an Ehresmann connection, connection one-form, and horizontal lift, simultaneously.

3.2 Properties of Connections

3.2.1 Lie Bracket on Principal Bundles

We will find it useful to understand how fundamental and horizontal vectors behave with respect to the Lie bracket of vector fields on P. We recall that $[X,Y] = \mathcal{L}_X(Y) = \frac{d}{dt}(\phi_t)_* X|_{t=0}$ where $X, Y \in \Gamma(P)$, $\mathcal{L}$ is the Lie derivative, and ϕ_t is the flow associated with Y. First, we consider the case where both vector fields are fundamental: this behavior is characterized as follows.

Proposition 3.18. The map $A \mapsto A^\#$ is a Lie algebra homomorphism from the Lie algebra $\mathfrak{g}$ to the Lie algebra of vector fields on P: that is, given $A, B \in \mathfrak{g}$, $[A^\#, B^\#] = [A,B]^\#$.

Proof. Recall that the flow associated to $B^\#$ is $R_{b(t)}$ for $b(t) = \exp(Bt)$, as shown in the proof of Prop. 2.29. Thus, as we have just recalled,

$$\begin{aligned}
[A^\#, B^\#] &= \frac{d}{dt}(R_{b(t)})_*(A^\#)|_{t=0} \\
&= \frac{d}{dt}(\mathrm{ad}(b(t)^{-1})(A))^\#|_{t=0} \\
&= \frac{d}{dt}(\mathrm{ad}(-b(t))(A))^\#|_{t=0} \\
&= \left(\frac{d}{dt}(\mathrm{ad} \circ -b)(-t)|_{t=0}(A)\right)^\#
\end{aligned}$$

Here, we have used Prop. 2.29. Embedded in this expression we have

$$\frac{d}{dt}(\mathrm{ad} \circ -b)(t)|_{t=0} = ((\mathrm{ad})_* \circ (-b)_*)(1) = (\mathrm{ad})_*(-B) \tag{14}$$

where $1 \in T_0\mathbb{R}$ and we observe that $(-b)_*(1) = (-b)'(0) = -B$. Moreover, it is a Lie theoretic fact that $(\mathrm{ad})_*(B)(C) = [B,C]$ for $B, C \in \mathfrak{g}$, hence

$$\begin{aligned}
[A^\#, B^\#] &= ((\mathrm{ad})_*(-B)(A))^\# \\
&= (\mathrm{ad}_*(-B)(A))^\# \\
&= [-B, A]^\# \\
&= [A, B]^\#
\end{aligned}$$

□

Notice that, as a corollary to this result, given two vertical vector fields $X, Y \in \Gamma(VP)$ such that for $p \in P$ we have $X(p) = A^\#(p), Y(p) = B^\#(p)$ for $A, B \in \mathfrak{g}$, it follows that $[X, Y](p) = [A, B]^\#(p)$. Additionally, we also conclude form this result that $\Gamma(VP)$ is closed under the Lie bracket.

Now we consider the case where one vector field is fundamental and the other is horizontal.

Proposition 3.19. The Lie bracket $[X, A^\#]$ for $X \in \Gamma(HP)$ and $A \in \mathfrak{g}$ is horizontal.

Proof. $[X, A^\#] = \frac{d}{dt}(R_{\exp(At)})_* X|_{t=0}$ and HP is invariant under $(R_g)_*$ for $g \in G$, thus $(R_{\exp(At)})_* X$ is horizontal for any t, making the difference quotient for this derivative also a horizontal vector. □

Finally, we consider the case of two horizontal vector fields

Proposition 3.20. The Lie bracket of two horizontal vector fields $X, Y \in \Gamma(HP)$ is horizontal if and only if $\Omega(X, Y) = 0$, where Ω is the curvature of the connection.

This is an initial motivation for curvature on principal bundles. We will prove this result upon our introduction of the curvature in a few sections.

We conclude by considering a useful, related result.

Proposition 3.21. The Lie bracket $[X, A^\#]$ vanishes for $A \in \mathfrak{g}$ and $X \in \Gamma(P)$ such that $(R_g)_*(X) = X$ for each $g \in G$.

Proof.

$$[X, A^\#] = \frac{d}{dt}(R_{\exp(At)})_* X|_{t=0} = \frac{d}{dt} X|_{t=0} = 0 \tag{15}$$

□

We say that such an X is *G-invariant*.

Example 3.22. A horizontal lift γ (that is, a G-equivariant vector bundle morphism $\pi^{-1}TM \to TP$ which behaves the right way with the Atiyah sequence) induces a map $\Gamma(M) \to \Gamma(P)$. In particular, letting π be the projection $P \to M$, given $X \in \Gamma(M)$ we can form a $X' \in \Gamma(\pi^* TM)$ by $X'(p) = (p, (X \circ \pi)(p))$. We can then define $X''(p) = (\gamma \circ X')(p) \in \Gamma(P)$. Moreover, letting $g \in G$,

$$(R_g)_* X''(p) = (R_g)_* \gamma(p, (X \circ \pi)(p)) = \gamma(pg, (X \circ \pi)(p)) = \gamma(pg, (X \circ \pi)(pg)) = X''(pg) \tag{16}$$

where we have exploited that $\pi(p) = \pi(pg)$. Hence, X'' is G-invariant. More broadly, horizontal lifts of vector fields on M are a nice class of examples of G-invariant vector fields on P.

3.2.2 Basic Forms

Of the tensor fields on P, those with a certain specific set of properties is distinguished, both for their convenient behavior with the structure we have and will introduce (namely, connections and covariant derivatives) and for their physical relevance. We take the time here to quickly develop their theory. The definitions are, at present, not well-motivated, but their utility will be rapidly revealed in the ensuing sections.

Let ρ be a G-action on a vector space V and let $E = P \times_\rho V$ be an associated bundle to P, a principal bundle with connection.

Definition 3.23 (Horizontal Form)**.** A V-valued differential k-form on P is *horizontal* if it vanishes whenever any of its arguments are vertical.

If we understand V-valued differential k-forms as maps sending k vector fields to a section of $P \times V$, we have G-actions in both the domain (applying $(R_g)_*$ to the arguments of the form, or applying $(R_g)^*$ to the form) and the codomain (ρ, which we often suppress and just write gv for $v \in V$). Thus, we have a well-defined notion of equivariance.

Definition 3.24 (Basic Form)**.** G-equivariant (with respect to ρ) horizontal V-valued k-forms on P are called *basic.* We use $\Omega^k_\rho(P, V)$ to the space of such basic forms.

It is straightforward to see that $\Omega^k_\rho(P, V)$ forms a submodule of $\Omega^k(P, V)$. Significantly, these basic forms on P that we are considering actually have two special properties: first, they descend in a well-defined fashion to forms on M, and second, these associated forms on M can be understood as taking values in E rather than merely in V. Thus, we can identify $\Omega^k_\rho(P, V)$ with $\Omega^k(M, E)$. We achieve this with the following construction. By definition, for each $p \in P$ (letting $\pi(p) = m$) we have a map $f_p : V \to \pi_E^{-1}(x)$ given by $v \mapsto [p, v]$. Now, let $\alpha \in \Omega^k_\rho(P, V)$: to this element we can associate $\alpha^\spadesuit \in \Omega^k(M, E)$ given by

$$\alpha^\spadesuit(m)(v_1, \ldots, v_k) = (f_p \circ \alpha(p))(\gamma(p, v_1), \ldots, \gamma(p, v_k)) \tag{17}$$

for $v_1, \ldots, v_k \in T_m M$, $p \in \pi^{-1}(m)$, and γ the horizontal lift associated to the given connection. Conversely, let $\beta \in \Omega^k(M, E)$, to this element we can associate $\beta^\clubsuit$ defined by

$$\beta^\clubsuit(p)(u_1, \ldots, u_k) = (f_p^{-1} \circ \beta(m))(\pi_*(u_1), \ldots, \pi_*(u_k)) \tag{18}$$

for $u_1, \ldots, u_k \in T_p P$ and π the projection on P. From here, we prove the following

Proposition 3.25. The maps $\alpha \mapsto \alpha^\spadesuit$ and $\beta \mapsto \beta^\clubsuit$ are isomorphisms $\Omega^k_\rho(P, V) \to \Omega^k(M, E)$ and $\Omega^k(M, E) \to \Omega^k_\rho(P, V)$ and are mutual inverses: in particular, $\Omega^k_\rho(P, V)$ and $\Omega^k(M, E)$ are isomorphic as modules.

Proof. That these constructions are inverses follows from the requirement $\overline{\pi} \circ \gamma = \mathbb{1}$ we've imposed upon horizontal lifts. For the following result to hold, then, it thus suffices to show that these are well-defined maps on the specified domains and codomains. We retain the definitions provided in the paragraphs leading up to this proposition.

We begin by exhibiting two important properties of f_p. Let $g \in G$ and $v \in V$: then, because $(pg, v), (p, gv)$ are both elements of the coset $[p, v] \in E$, it follows that

$$f_{pg}(v) = [pg, v] = [p, gv] = f_p(gv) \tag{19}$$

Moreover, if $w = f_p(v) = f_{pg}(g^{-1}v)$ we have

$$(f_{pg}^{-1} \circ f_{pg})(g^{-1}v) = (f_{pg}^{-1} \circ f_p)(v) \Longrightarrow g^{-1}v = f_{pg}^{-1}(w) \Longrightarrow g^{-1}f_p^{-1}(w) = f_{pg}^{-1}(w) \tag{20}$$

We now begin with the map $\alpha \mapsto \alpha^\spadesuit$. It suffices to show that $\alpha^\spadesuit(m)$ is well-defined independent of the choice of $p \in \pi^{-1}(m)$. Due to the transitivity of the G-action on fibres of P, it suffices to show equivalence for $p, pg \in \pi^{-1}(m)$. We can see this as follows, exploiting the (pseudo) equivariance we noted for f_p and the (genuine) equivariance of γ, α.

$$\begin{aligned}(f_{pg} \circ \alpha(p))(\gamma(pg, v_1), \ldots, \gamma(pg, v_k)) &= (f_p \circ L_g \circ \alpha(p))((R_g)_*\gamma(p, v_1), \ldots, (R_g)_*\gamma(p, v_k)) \\ &= (f_p \circ L_g \circ L_{g^{-1}} \circ \alpha(p))(\gamma(p, v_1), \ldots, \gamma(p, v_k)) \\ &= (f_p \circ \alpha(p))(\gamma(p, v_1), \ldots, \gamma(p, v_k))\end{aligned}$$

Now we turn our attention to $\beta \mapsto \beta^{\clubsuit}$. It suffices here to show that $\beta^{\clubsuit}$ is basic. It is evident that it is horizontal (we are applying π_* to the arguments, whose kernel is vertical vectors), so G-equivariance is the sole remaining property to be shown. We see this as follows, recalling that π_* is invariant under $(R_g)_*$.

$$\begin{aligned}
\beta^{\clubsuit}(pg)((R_g)_*(u_1), \ldots, (R_g)_*(u_k)) &= (f_{pg}^{-1} \circ \beta(m))(\pi_*(u_1), \ldots, \pi_*(u_k)) \\
&= g^{-1}(f_p^{-1} \circ \beta(m))(\pi_*(u_1), \ldots, \pi_*(u_k)) \\
&= g^{-1}(f_p^{-1} \circ \beta(m))(\pi_*(u_1), \ldots, \pi_*(u_k)) \\
&= g^{-1}\beta^{\clubsuit}(p)(u_1, \ldots, u_k)
\end{aligned}$$[7]

□

We comment that ♠ need not have been constructed with exactly the horizontal lift compatible with the connection in hand: it could have been any choice of horizontal lift, but given the fact that a connection was already necessarily present for $\Omega^k_\rho(P, V)$ to have been well-defined, it was convenient to take the canonical choice of horizontal lift.

To conclude, we present a particular definition we will exploit later. Specifically, we have a natural way for $\Omega^k_\rho(P, \mathfrak{g})$ to act upon $\Omega^k(P, V)$; namely, we exploit the fact that $\rho : G \to GL(V)$ induces $\rho_* : \mathfrak{g} \to \mathfrak{gl}(V) \cong \mathrm{End}(V) \cong V \otimes V^*$ (isomorphic as vector spaces). Additionally, at this point we (sparingly) begin our use Penrose abstract index notation [15, Sec. 2.2.4].

Definition 3.26. Given $\alpha \in \Omega^k_\rho(P, \mathfrak{g})$ and $\beta \in \Omega^\ell(P, V)$, we define $\alpha \wedge_\rho \beta$ by

$$(\alpha \wedge_\rho \beta)_{a_1 \ldots a_{k+\ell}} = (\rho_* \circ \alpha)_{[a_1 \ldots a_k}(\beta_{a_{k+1} \ldots a_{k+\ell}]}) \tag{21}$$

For clarity, here $\rho_* \circ \alpha$ is interpreted as as map $T_pM \otimes \cdots \otimes T_pM \to V^* \otimes V$ at each $p \in P$ (with vector field inputs being associated to Penrose abstract indices above) and likewise $\beta : T_pM \otimes \cdots \otimes T_pM \to V$, hence the evaluation of $\rho_* \circ \alpha$ can, in turn, be applied to the evaluation of β. We recall that square brackets are used to denote antisymmetrization. Finally, we use the $\wedge_\rho$ notation because of the resemblence that the formula has to the regular wedge product.

For our utility, we take a moment to provide a more verbose formula for the above construction that lends itself toward computation.

$$(\alpha \wedge_\rho \beta)(p)(v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \mathrm{sgn}(\sigma) \rho_*\Big(\alpha(p)(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\Big)\Big(\beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)})\Big) \tag{22}$$

Here, $S_{k+\ell}$ is the symmetric group on $k + \ell$ objects with elements $\sigma : \{1, \ldots, k + \ell\} \to \{1, \ldots, k + \ell\}$, understood as permutations (i.e., bijections).

Before departing from this section, we insert a final comment emphasizing a particular case of this construction: namely, the case $\alpha \in \Omega^k_{\mathrm{ad}}(P, \mathfrak{g})$ and $\beta \in \Omega^\ell(P, \mathfrak{g})$. Because $\mathrm{ad}_* : \mathfrak{g} \to \mathrm{End}(G)$ acts by $A \mapsto [A, \cdot]$, in this case Eq. 22 becomes

$$(\alpha \wedge_{\mathrm{ad}} \beta)(p)(v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \mathrm{sgn}(\sigma)[\alpha(p)(v_{\sigma(1)}, \ldots, v_{\sigma(k)}), \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)})] \tag{23}$$

Consequently, in this special case, we use the notation $(\alpha \wedge_{\text{ad}} \beta) = [\alpha, \beta]$, but be warned that this notation obscures the antisymmetrization that is still happening here. For instance, if $k = \ell = 1$ for α, β, we have that $[\alpha, \beta](X, Y)$ does not equal $[\alpha(X), \beta(Y)]$, as one might intuitively think from the notation, but rather $[\alpha(X), \beta(Y)] - [\alpha(X), \beta(Y)]$. Indeed, we even have that $[\alpha, \alpha](X, Y) = 2[\alpha(X), \alpha(Y)]$, as will be manifest in the proof of Th. 3.39.

3.2.3 Space of Connections

We have established a correspondence between Ehresmann connections, connection one-forms, and horizontal lifts. We now wish to consider the set of such objects for a given principal bundle P, and to do this we adopt the perspective of connection one-forms, given their natural algebraic structure. In particular, letting the space of connection one-forms on P be $\mathcal{A}(P)$, because elements of $\mathcal{A}(P)$ are sections of the tensor bundle $\operatorname{ad} P \otimes \Omega^1(P)$, $\mathcal{A}(P)$ is a subset of the $C^\infty(P)$-module $\Gamma(\operatorname{ad} P \otimes \Omega^1(P)) = \Omega^1(P, \operatorname{ad} P)$. But there is an immediate a subtlety here. We are free to apple the module operations to elements of $\mathcal{A}(P)$, but $\mathcal{A}(P)$ is not a submodule itself: indeed, given a fundamental vector field $A^\#$ on P associated to $A \in \mathfrak{g}$, we require that $\omega \in \mathcal{A}(P)$ satisfy $\omega(A^\#) = A$, but given $\omega, \omega' \in \mathcal{A}(P)$, we have both that $f\omega(A^\#) = fA \neq A$ for $f \in C^\infty(P)$ (where $(fA)(p) = (p, f(p)A)$, identifying $\operatorname{ad} P$ with $P \times \mathfrak{g}$) and $(\omega + \omega')(A^\#) = 2A \neq A$. That is, $\mathcal{A}(P)$ isn't closed under either module operation. Evidently, the structure of $\mathcal{A}(P)$ is more complicated.

To move toward understanding $\mathcal{A}(P)$, consider the evaluation of $\omega - \omega'$ $(\omega, \omega' \in \mathcal{A}(P))$ on $V \in \Gamma(P)$ such that $V(p) \in VP$ for all $p \in P$: that is, V is strictly vertical. By definition ω, ω' are required to map fundamental vectors to their associated lie algebra elements. Hence ω, ω' coincide on V, so their difference vanishes, and more broadly the difference of two connection one-forms vanishes when their argument is vertical. It follows that a difference of connection one-forms is horizontal. Moreover, we are already aware that connections are G-equivariant, hence their differences are as well. Hence, in tandem, differences of connection one-forms are basic, or $\omega - \omega' \in \Omega^1_{\text{ad}}(P, \mathfrak{g})$.

As the reader may recognize, a set where the elements themselves do not appear to have any kind of linear structure but their differences do often admits a characterization as an affine space. Because any space of basic one-forms forms a module, we consider the following definition in particular.

Definition 3.27. An *affine module* is a set A acted upon regularly (that is, freely and transitively) by the additive group associated with a module B (that is, the group given by B and module addition). We say *A is an affine module modelled on B*

We use the terminology "modelled on" because the affine module A is really in bijection with B: fix any element of $a \in A$, identify it with $0 \in B$, then identify $a' \in A$ with the $b \in B$ such that $a' = b + a$ (again, interpreting a as 0).

By definition, given an affine module A with group B and $a, a' \in A$, there is a unique $b \in B$ such that $a + b = a'$, hence it is natural to define $a - a' = b$: now we see how this lines up with what we have seen regarding $\mathcal{A}(P)$. Indeed, basic one-forms are the natural objects to add to connection one-forms: they don't affect the behavior of the original connection one-form on vertical vectors and the preserve the equivariance. This next result then naturally follows.

Proposition 3.28. $\mathcal{A}(P)$ is an affine $C^\infty(P)$-module modelled on $\Omega^1_{\text{ad}}(P, \mathfrak{g})$.

3.2.4 Local Connection One-Form

The local trivializations of P enable us to locally pullback a connection one-form to M. We achieve this as follows.

Definition 3.29 (Canonical Local Section)**.** The *canonical local section* associated to a given local trivialization (U_α, ϕ_α) is the section $\psi_\alpha : U_\alpha \to \pi^{-1}(U_\alpha) \subset P$ given by $x \mapsto \phi_\alpha^{-1}(x, e)$ where $e \in G$ is the identity element.

Definition 3.30 (Local Connection 1-form)**.** Given a connection one-form ω on P, the *local connection one-form* on U_α, ω_α is the $\mathfrak{g}$-valued one-form on U_α defined by $(\psi_\alpha)^*\omega$.

We now turn our attention to the transformation behavior of the local connection one-forms between local trivializations on an overlap $U_\alpha \cap U_\beta$. We briefly recall a relevant Lie theoretic object, then present the main result.

Definition 3.31. The Maurer-Cartan one-form θ is the $\mathfrak{g}$-valued one-form on G acting canonically by $w(g)(v) = (L_{g^{-1}})_*(v) \in T_eG \cong \mathfrak{g}$ (isomorphic as vector spaces) for $v \in T_gG$

Theorem 3.32. Given local connection one-forms ω_α, ω_β defined on U_α and U_β, respectively, the following transformation law holds at $u \in U_\alpha \cap U_\beta$.

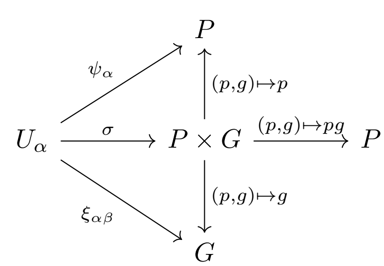
$$(\omega_\beta(u)) = \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha(u) + (L_{\xi_{\alpha\beta}^{-1}(u)} \circ \xi_{\alpha\beta})_* \tag{24}$$

$$= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha(u) + \xi_{\alpha\beta}^*\theta(u) \tag{25}$$

Proof. Let $p \in P$ and $u = \pi(p)$. We begin by observing that

$$\begin{aligned}\psi_\alpha(u) &= \phi_\alpha^{-1}(u, e)\\ &= \phi_\alpha^{-1}(u, \xi_{\alpha\beta}(u)\xi_{\beta\alpha}(u))\\ &= \phi_\alpha^{-1}(u, \xi_{\alpha\beta}(u)e)\xi_{\beta\alpha}(u)\\ &= \phi_\beta^{-1}(u, e)\xi_{\beta\alpha}(u)\\ &= \psi_\beta(u)\xi_{\beta\alpha}(u).\end{aligned}$$

Let $\sigma : U_\alpha \to P \times G$ be $u \mapsto (\psi_\alpha(u), \xi_{\alpha\beta}(u))$ and $\rho : P \times G \to P$ be the right action of G on P. This means $\psi_\beta = (\rho \circ \sigma)(u)$.



Consider the pushforward of both sides of $\psi_\beta = \psi_\beta\xi_{\beta\alpha}$ at $u \in U_\alpha$.

$$\begin{aligned}(\psi_\beta)_* &= (\rho \circ \sigma)_*\\ &= \rho_* \circ \sigma_*\\ &= \rho_*|_{T_{\psi_\alpha(u)}P \times \{0\}} \circ (\psi_\alpha)_* + \rho_*|_{\{0\} \times T_{\xi_{\alpha\beta}(u)}G} \circ (\xi_{\alpha\beta})_*\end{aligned} \tag{26}$$

Here, we have exploited the chain rule and the decomposition of a tangent space to a product manifold as a product of tangent spaces. At this point, observe that $T_{\psi_\alpha(u)}P \times \{0\}$ is the tangent space to the submanifold $P \times \{\xi_{\alpha\beta}(u)\} \subset P \times G$ at $(\psi_\alpha(u), \xi_{\alpha\beta}(u))$, and the restriction of ρ to this submanifold is $g \mapsto g\, \xi_{\alpha\beta}(u)$, or $R_{\xi_{\alpha\beta}(u)}$, hence

$$\rho_*|_{T_{\psi_\alpha(u)}P\times\{0\}} = (R_{\xi_{\alpha\beta}(u)})_*, \tag{27}$$

By an analogous argument, if we recall from Section 2.5 that $\sigma_{\psi_\alpha(u)} : G \to P$ acts by $g \mapsto \psi_\alpha(u)g$, we have that

$$\rho_*|_{\{0\}\times T_{\xi_{\alpha\beta}(u)}G} = (\sigma_{\psi_\alpha(u)})_*. \tag{28}$$

At this point, we consider it useful to explicitly clarify the domains of our pushforwards, so we briefly adopt the d notation for the pushforward in doing the next few computations. Consider the second term on the right hand side of Eq. 26.

$$\begin{aligned} d(\sigma_{\psi_\alpha(u)})_{\xi_{\alpha\beta}(u)} \circ d(\xi_{\alpha\beta})_u &= d(\sigma_{\psi_\beta(u)(\xi_{\alpha\beta}(u))^{-1}})_{\xi_{\alpha\beta}(u)} \circ d(\xi_{\alpha\beta})_u \\ &= d(\sigma_{\psi_\beta(u)} \circ L_{(\xi_{\alpha\beta}(u))^{-1}})_{\xi_{\alpha\beta}(u)} \circ d(\xi_{\alpha\beta})_u \\ &= d(\sigma_{\psi_\beta(u)})_e \circ d(L_{(\xi_{\alpha\beta}(u))^{-1}})_{\xi_{\alpha\beta}(u)} \circ d(\xi_{\alpha\beta})_u \end{aligned}$$

Hence,

$$d(\psi_\beta)_u = d(R_{\xi_{\alpha\beta}(u)})_{\psi_\alpha(u)} \circ d(\psi_\alpha)_u + d(\sigma_{\psi_\beta(u)})_e \circ d(L_{(\xi_{\alpha\beta}(u))^{-1}})_{\xi_{\alpha\beta}(u)} \circ d(\xi_{\alpha\beta})_u \tag{29}$$

Or, returning back to our subscript $*$ notation for pushforwards,

$$(\psi_\beta)_* = (R_{\xi_{\alpha\beta}(u)})_* \circ (\psi_\alpha)_* + (\sigma_{\psi_\beta(u)})_* \circ (L_{(\xi_{\alpha\beta}(u))^{-1}})_* \circ (\xi_{\alpha\beta})_* \tag{30}$$

We proceed by applying the connection one-form ω to both sides.

$$\omega \circ (\psi_\beta)_* = \omega \circ (R_{\xi_{\alpha\beta}(u)})_* \circ (\psi_\alpha)_* + \omega \circ (\sigma_{\psi_\beta(u)})_* \circ (L_{(\xi_{\alpha\beta}(u))^{-1}})_* \circ (\xi_{\alpha\beta})_* \tag{31}$$

The left hand side is $\omega \circ (\psi_\beta)_* = \psi_\beta^*\omega = \omega_\beta$. The first term on the right hand side is

$$\begin{aligned} \omega \circ (R_{\xi_{\alpha\beta}(u)})_* \circ (\psi_\alpha)_* &= (R^*_{\xi_{\alpha\beta}(u)}\omega) \circ (\psi_\alpha)_* \\ &= (\mathrm{ad}(\xi_{\alpha\beta}(u)^{-1}) \circ \omega) \circ (\psi_\alpha)_* \\ &= \mathrm{ad}(\xi_{\alpha\beta}(u)^{-1}) \circ (\omega \circ (\psi_\alpha)_*) \\ &= \mathrm{ad}(\xi_{\alpha\beta}(u)^{-1}) \circ \psi_\alpha^*\omega \\ &= \mathrm{ad}(\xi_{\alpha\beta}(u)^{-1}) \circ \omega_\alpha \end{aligned}$$

Now we consider the second term on the left hand side. The pushforward $(\sigma_{\psi_\beta(u)})_*$ maps to fundamental vectors, but the connection one-form then maps these directly back to the same Lie algebra element: i.e., $\omega \circ (\sigma_{\psi_\beta(u)})_* = \mathbb{1}$. Putting these results together, it follows that

$$\omega_\beta(u) = \mathrm{ad}(\xi^{-1}_{\alpha\beta}(u)) \circ \omega_\alpha(u) + (L_{\xi^{-1}_{\alpha\beta}(u)})_* \circ (\xi_{\alpha\beta})_* \tag{32}$$

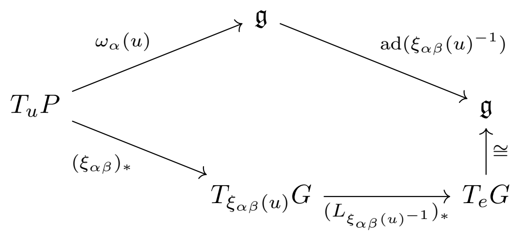
Finally, we note that the map $(L_{\xi^{-1}_{\alpha\beta}(u)})_*$ on G is precisely the Maurer-Cartan one-form, hence the second term in Eq. 32 is $\xi^*_{\alpha\beta}\theta(u)$, finishing the proof[8] □

[8] Proof inspired by [10, Th. 1.6, Sec. 2.2.1]

For clarity, because we have

$$U_\alpha \cap U_\beta \xrightarrow{\xi_{\alpha\beta}} G \xrightarrow{L_{\xi_{\alpha\beta}(u)^{-1}}} \{e\}$$

where this is merely a diagram and not a short exact sequence, we can understand Eq. 24 as follows.



Here, $\cong$ is of course an isomorphism of vector spaces.

Additionally, we comment that Eq. 24 doesn't take the form that conventionally appears in both the mathematical and physical literature. In particular, the following typically features instead.

$$\omega_\beta = \mathrm{ad}(g^{-1})\omega_\alpha + g^{-1}dg \tag{33}$$

Here, g is the transition function with values in G. In this notation, dg is the pushforward. The confusing part here is the object g^{-1}, which has somehow replaced the pushforward of left multiplication by g. This is substantiated by the idea that for matrix Lie groups, the push forward of left multiplication coincides with left multiplication. This follows directly from the differential geometric fact that pushforwards of linear transformations (e.g., multiplication by elements of matrix Lie groups) coincide with the linear transformation themselves in an extrinsic, embedded-in-$\mathbb{R}^n$ setting. However, we seize this opportunity to explore this idea a different way. Namely, we specialize to the case $G = SU(2)$, the special unitary group of degree 2, and demonstrate that the image of $T_g\, SU(2)$ under $g^{-1} \in SU(2)$ is, in fact, $T_e\, SU(2) \cong \mathfrak{su}(2)$ (vector space isomorphism) as we require. This at least shows that the codomain of term two of Eq. 33 makes sense.

Example 3.33. Recall the definition of $SU(2)$.

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \;\middle|\; \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\} \subset \mathbb{M}_2(\mathbb{C}) \tag{34}$$

We are interested in exhibiting that the pushforward $(L_g)_* : T_hSU(2) \mapsto T_{gh}SU(2)$ (omitting the restriction notation) coincides with the left action of the matrix g on the elements of $T_hSU(2)$ when this tangent space is viewed as a subset of $\mathbb{M}_2(\mathbb{C})$, the complex vector space of 2×2 complex matrices (i.e., explicitly using the embedding $SU(2) \subset \mathbb{M}_2(\mathbb{C})$). In particular, we demonstrate that $g^{-1}(T_gSU(2)) \subset T_eSU(2)$ (that they are isomorphic then follows from the invertibility of the matrix g^{-1}).

We begin by developing a convenient interpretation of $SU(2)$: namely, exhibiting that it is identifiable with S^3. We are free to identify $\mathbb{M}_2(\mathbb{C}) \cong \mathbb{C}^4 \cong \mathbb{R}^8$ (equivalence in the sense of vector space isomorphisms) in the canonical way: in particular,

$$\begin{pmatrix} a+bi & c+di \\ e+fi & g+hi \end{pmatrix} \mapsto \begin{pmatrix} a & b & c & d & e & f & g & h \end{pmatrix}^\top \tag{35}$$

As a subset of $\mathbb{R}^8$, $SU(2)$ actually belongs to the 4-dimensional linear subspace $V \subset \mathbb{R}^8$ spanned by

$$\left\{ \begin{pmatrix}1\\0\\0\\0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\\-1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\\-1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\1\\0\\0\end{pmatrix} \right\} \subset \mathbb{R}^8 \tag{36}$$

as

$$a\begin{pmatrix}1\\0\\0\\0\\0\\0\\1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\\0\\0\\0\\0\\0\\-1\end{pmatrix} + c\begin{pmatrix}0\\0\\1\\0\\-1\\0\\0\\0\end{pmatrix} + d\begin{pmatrix}0\\0\\0\\1\\0\\1\\0\\0\end{pmatrix} = \begin{pmatrix}a\\b\\c\\d\\-c\\d\\e\\-f\end{pmatrix} \mapsto \begin{pmatrix}a+bi & c+di\\ -(c-di) & a-bi\end{pmatrix} \tag{37}$$

and all elements of $SU(2)$ take the form of the last term. In particular, if we identify this subspace with $\mathbb{R}^4$ via the above basis, $SU(2)$ coincides with S^3, in particular by the map

$$SU(2) \ni \begin{pmatrix}a+bi & c+di\\ -(c-di) & a-bi\end{pmatrix} \mapsto \begin{pmatrix}a\\b\\c\\d\end{pmatrix} \in S^3 \tag{38}$$

We proceed by identifying $SU(2) \cong S^3$ via this map.

We now exploit this identification to understand tangent spaces to $SU(2)$. Pick $g = \begin{pmatrix}a & b & c & d\end{pmatrix}^\top \in S^3$. Letting w, x, y, z be the coordinate functions, T_pS^3 is given by the equation

$$a(w-a) + b(x-b) + c(y-c) + d(z-d) = 0 \tag{39}$$

Or, equivalently,

$$aw + bx + cy + dz = 1 \tag{40}$$

Let us briefly emphasize the case we are ultimately most interested in, $g = e$, or $g = \begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix}^\top$. The linear tangent space (i.e., $T_eS^3 - e$) is given by the null space of the matrix

$$\begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix} \tag{41}$$

This null space is spanned by $\{e_2, e_3, e_4\}$, which are identified with the following matrices

$$\left\{ \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}0 & i\\ i & 0\end{pmatrix} \right\} \tag{42}$$

Hence, an arbitrary element of $T_eSU(2)$ takes the form

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \kappa\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix} + \lambda\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} + \mu\begin{pmatrix}0 & i\\ i & 0\end{pmatrix} \tag{43}$$

for $\kappa, \lambda, \mu \in \mathbb{R}$. As an aside, our basis matrices are but a factor of $-i$ away from the Pauli spin matrices.

Now we are equipped to show the desired result. Let $v = \begin{pmatrix} q & r & s & t \end{pmatrix}^\top \in T_p S^3$. Note that g, v are identified with the following matrices.

$$g \mapsto \begin{pmatrix} a+bi & c+di \\ -(c-di) & a-bi \end{pmatrix}, v \mapsto \begin{pmatrix} q+ri & s+ti \\ -(s-ti) & q-ri \end{pmatrix} \tag{44}$$

Note that g^{-1} takes the following form.

$$g^{-1} = \begin{pmatrix} a-bi & -(c-di) \\ c-di & a+bi \end{pmatrix} \tag{45}$$

For our purposes it suffices to show that $g^{-1}v \in T_e SU(2)$. Consider this product.

$$\begin{aligned} g^{-1}v &= \begin{pmatrix} a-bi & -(c-di) \\ c-di & a+bi \end{pmatrix} \begin{pmatrix} q+ri & s+ti \\ -(s-ti) & q-ri \end{pmatrix} \\ &= \begin{pmatrix} (a-bi)(q+ri)+(c+di)(s-ti) & (a-bi)(s+ti)-(c+di)(q-ri) \\ -(a+bi)(s-ti)+(c-di)(q+ri) & (a+bi)(q-ri)+(c-di)(s+ti) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (aq+br+cs+dt) + \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} (-ar+bq+ct-ds) \\ &\quad + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (as+bt-cq-dr) + \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} (at-bs+cr-dq) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} (-ar+bq+ct-ds) \\ &\quad + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (as+bt-cq-dr) + \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} (at-bs+cr-dq) \end{aligned}$$

This is evidently an element of $T_e SU(2)$ per our earlier discussion, concluding our argument.

Our transformation law for local connection one-forms further illustrates why Prop. 3.28 makes sense. In particular, the $\xi^*_{\alpha\beta}\theta$ term prevents the law from describing a legitimate tensor on M via Prop. 2.8. However, because that term is constant on a fixed overlap, the transformation law for a difference $\omega_\alpha - \omega'_\alpha$ sees the $\xi^*_{\alpha\beta}\theta$ term cancel, leaving merely that

$$(\omega_\beta - \omega'_\beta)(u) = \mathrm{ad}(\xi^{-1}_{\alpha\beta}(u)) \circ (\omega_\alpha - \omega'_\alpha)(u) \tag{46}$$

By Prop. 2.8, then, these differences of local connections describe a single one-form tensored with single section of a fibre bundle with base manifold M and transition functions given by $\mathrm{ad}(\xi^{-1}_{\alpha\beta}(u))$: this latter bundle is $\mathrm{ad}\, P$, hence the differences of local connection one-forms sew together to form a $\mathrm{ad}\, P$-valued one-form on M, or an element of $\Omega^1(M, \mathrm{ad}\, P)$. It suffices at this point to recall Prop. 3.25, which entails differences of connection one-forms are indeed elements of $\Omega^1(M, \mathrm{ad}\, P)$.

3.2.5 Exterior Covariant Derivative

We're now interested in constructing differentiation operations that are compatible with our newfound connections. We will begin this endeavor now with principal bundles and continue to pursue it in the context of (associated) vector bundles in Section 3.3. Recall the exterior derivative $d : \Omega^k(M) \mapsto \Omega^{k+1}(M)$ for arbitrary M. We can extend this to an operation $d : \Omega^k(P, V) \mapsto \Omega^{k+1}(P, V)$ by observing that $\alpha \in \Omega^k(P, V)$ takes the form $\alpha = \sum_i e_i \otimes \alpha_i$ (where $\{e_i\}$ is a basis for V and $\alpha_i \in \Omega^k(P)$) and the definition $d\alpha = \sum_i e_i \otimes d\alpha_i$. This vector-valued exterior derivative has all of the same desired properties.

Now, given a connection on P, let $\boldsymbol{\hbar} : TP \to HP$ be the canonical projection.

Definition 3.34 (Exterior Covariant Derivative)**.** The *exterior covariant derivative* induced by a connection is the map $D\alpha = d\alpha \circ \boldsymbol{\hbar}$ for $\alpha \in \Omega^k(P)$.

Here, $\boldsymbol{\hbar}$ is understood to be applied to all $k+1$ vector field arguments of $d\alpha$. This definition extends to $\Omega^k(P, V)$ in a manner identical to similar to d. While we can apply this operation to any V-valued differential k-form on P, it bears special relevance for the specific kinds of forms discussed in the previous section.

First, note that equivariant forms maintain equivariance when precomposed with $\boldsymbol{\hbar}$ in all arguments, and moreover the commutativity of the exterior derivative with pullbacks ensures that the exterior derivative preserves equivariance as well. Thus, in conjunction, the exterior covariant derivative maps equivariant forms to equivariant forms.

Now, observe that horizontal forms need not have horizontal exterior derivatives, but this is fixed by precomposing with $\boldsymbol{\hbar}$: thus, the exterior covariant derivative is designed precisely to preserve horizontality. In conjunction, this discussion exhibits the following

Proposition 3.35. The exterior covariant derivative is a well-defined as an operation $\Omega^k_\rho(P, V) \to \Omega^{k+1}_\rho(P, V)$.

Recall that we do have the following useful formula for $d\beta$ for $\beta \in \Omega^k(P)$ which is a general result from differential geometry [13, Th. 20.14].

$$\begin{aligned} d\beta(v_0, \dots, v_k) = &\sum_{i=0}^{k} (-1)^i v_i(\beta(v_0, \dots, \hat{v_i}, \dots, v_k)) \\ &+ \sum_{i=0}^{k-1} \sum_{j=i+1}^{k} (-1)^{i+j} \beta([v_i, v_j], v_0, \dots, \hat{v_i}, \dots, \hat{v_j}, \dots, v_k) \end{aligned} \tag{47}$$

When we apply this formula to V-valued forms α, we are implicitly applying it to each α_i and regrouping at the end.

Def. 3.34, on the other hand, provides us with no insight on how to efficiently compute D. We therefore endeavor now to understand precisely how D differs from this formula[9]. In particular, we focus on the case for where $\alpha \in \Omega^k_\rho(P, V)$, for which there is a satisfying answer to be had. By linearity, it suffices to consider

$$D\alpha(p)(v_0, \dots, v_k) - d\alpha(p)(v_0, \dots, v_k) \tag{48}$$

where each $v_i \in T_pP$ is either vertical or horizontal. By alternativity, without loss of generality we can assume v_i is vertical if $i < \ell$ and horizontal otherwise for some $\ell \in \{0, \dots, k\}$. From here, we actually choose to reinterpret this computation by considering

$$D\alpha(V_0, \dots, V_k) - d\alpha(V_0, \dots, V_k) \tag{49}$$

[9]The results we prove to this end are collectively inspired by [14, Th. 31.19]

where each $V_i \in \Gamma(P)$ is an extension of v_i (i.e., $V_i(p) = v_i$) and is strictly vertical or horizontal. In particular, for $i < \ell$, v_i can be extended to $V_i = A^{\#}$ by recognizing that $v_i = A_i^{\#}(p)$ for some $A_i \in \mathfrak{g}$ by Prop. 2.23, while for $i \geq \ell$, v_i can be extended to the horizontal lift $\tilde{B}$ associated to some vector field $B \in \Gamma(M)$, recalling that vectors can always be extended to global vector fields[10] (in this case, we are extending $(\pi_P)_*(v_i)$ on M to form B).

Given this set up, we have the following intermediate result.

Lemma 3.36. For $\alpha \in \Omega^k_\rho(P, V)$ and $V_0, \ldots, V_k \in \Gamma(P)$ (vertical for $i < \ell$, horizontal otherwise),

$$D\alpha(V_0, \ldots, V_k) - d\alpha(V_0, \ldots, V_k) = \begin{cases} 0 & \ell \neq 1 \\ V_0(\alpha(V_1, \ldots, V_k)) & \ell = 1 \end{cases} \tag{50}$$

Proof. We begin with the simpler case, $\ell \neq 1$. If $\ell = 0$, all vector fields are horizontal, meaning $\boldsymbol{\hbar}$ has no impact upon the vector field arguments and the expression is evidently the difference of identical quantities. On the other hand, if $\ell \geq 2$, both terms vanish. In particular, that the D term vanishes follows from the observation that $\boldsymbol{\hbar}(V_1) = 0$. To see the same for the d term, we inspect Eq. 47 and note that each term in either sum involves the evaluation of the horizontal α upon at least one vertical vector field: namely, V_0, V_1 or $[V_0, V_1] = [A_0, A_1]^{\#}$ (recalling Prop. 3.18).

Now we consider the case $\ell = 1$. Because $\boldsymbol{\hbar}(V_0) = 0$, $D\alpha(V_0, \ldots, V_k)$ vanishes: it thus suffices to show that $d\alpha(V_0, \ldots, V_k) = V_0(\alpha(V_1, \ldots, V_k))$. In examining Eq. 47, we indeed see that only the term $V_0(\alpha(V_1, \ldots, V_k))$ survives because all other terms entail the evaluation of the horizontal α upon the vertical V_0. Finally, each term in the second sum vanishes because either α again takes on a vertical argument V_0 or an argument $[V_0, V_i] = 0$ which vanishes by Prop. 3.21 (in particular, per Ex. 3.22). ☐

Thus, the difference $D - d$, when applied to horizontal and vertical arguments, is non-vanishing only when a singular argument is vertical. This difference can be characterized specifically in the following way.

Theorem 3.37. Let V be a vector space upon which G acts by a representation ρ: D takes the following form on $\alpha \in \Omega^k_\rho(P, V)$.

$$D\alpha = d\alpha + \omega \wedge_{\mathrm{ad}} \alpha \tag{51}$$

Proof. Continuing to use the nomenclature we have been employing, it suffices to show that

$$(\omega \wedge_{\mathrm{ad}} \alpha)(V_0, \ldots, V_k) = \begin{cases} 0 & \ell = 0 \\ -V_0(\alpha(V_1, \ldots, V_k)) & \ell = 1 \end{cases} \tag{52}$$

where we recall that ℓ is the integer such that V_i is vertical if $i < \ell$ and horizontal otherwise.

Once more, we begin with $\ell \neq 1$. If $\ell = 0$, all vector field arguments are horizontal and, in particular, in considering Eq. 22, we find that the argument of ω is always horizontal. On the other hand, if $\ell \geq 2$, inspecting the same sum reveals that the horizontal form α always has at least one vertical argument. In either case, the expression vanishes as a whole.

[10]Through the use of a local chart and a bump function.

Now we turn to $\ell = 1$. Looking again at Eq. 22, we first find that only terms satisfying $\sigma(0) = 0$[11] survive (otherwise $V_{\sigma(0)}$ is horizontal and $\omega(V_{\sigma(0)})$ vanishes). Second, we observe that for all such σ,

$$\operatorname{sgn}(\sigma)\rho_*(\omega(V_0))(\alpha(V_{\sigma(1)}, \ldots, V_{\sigma(k)})) = \rho_*(\omega(V_0))(\alpha(V_1, \ldots, V_k)) \tag{53}$$

and there are $k!$ such terms, hence

$$(\omega \wedge_{\text{ad}} \alpha)(V_0, \ldots, V_k) = \rho_*(\omega(V_0))(\alpha(V_1, \ldots, V_k)) \tag{54}$$

It suffices now to show that $\rho_*(\omega(V_0))(\alpha(V_1, \ldots, V_k)) = -V_0(\alpha(V_1, \ldots, V_k))$. These are each functions $P \to V$: we demonstrate that they coincide at $p \in P$. To see this, we manipulate the right hand side. Let $f : P \to V$ be $p \mapsto \alpha(V_1, \ldots, V_k)(p)$: this is an equivariant map, which we can see as follows, exploiting the right-invariance of each V_i and the equivariance of α.

$$\begin{aligned} f(pg) &= \alpha(V_1, \ldots, V_k)(pg) = \alpha((R_g)_* V_1, \ldots, (R_g)_* V_k)(pg) \\ &= g^{-1}\alpha(V_1, \ldots, V_k)(p) = g^{-1} f(p) \end{aligned}$$

Now, let $V_0 = A^\#$ for $A \in \mathfrak{g}$, recall $\sigma_p : G \to P$ given by $g \mapsto pg$, and let $a : \mathbb{R} \to G$ be $t \mapsto \exp(At)$ (as in the proof of Prop. 2.29). If we and apply the chain rule, we see

$$V_0(\alpha(V_1, \ldots, V_k))(p) = (\sigma_p \circ a)'(0)(f) = (f \circ \sigma_p \circ a)'(0) \tag{55}$$

By the equivariance of f we are assured

$$(f \circ \sigma_p \circ a)(t) = f(pa(t)) = (\rho \circ a^{-1})(t) f(p) \tag{56}$$

meaning

$$\begin{aligned} V_0(\alpha(V_1, \ldots, V_k))(p) &= (f \circ \sigma_p \circ a)'(0) = (\rho \circ a^{-1})'(0) f(p) \\ &= \rho_*((a^{-1})'(0)) f(p) = -\rho_*(a'(0)) f(p) = -\rho_*(A) f(p) \\ &= -\rho_*(A)\alpha(V_1, \ldots, V_k)(p) \end{aligned}$$

where we have exploited that $\frac{d}{dt} e^{-tA}|_{t=0} = -\frac{d}{dt} e^{tA}|_{t=0}$. Hence,

$$\begin{aligned} \rho_*(\omega(V_0))(\alpha(V_1, \ldots, V_k))(p) &= \rho_*(\omega(A^\#))(\alpha(V_1, \ldots, V_k))(p) \\ &= \rho_*(A)(\alpha(V_1, \ldots, V_k))(p) \\ &= \rho_*(A)(\alpha(V_1, \ldots, V_k))(p) \\ &= -(-\rho_*(A)\alpha(V_1, \ldots, V_k)(p)) \\ &= -V_0(\alpha(V_1, \ldots, V_k))(p) \end{aligned}$$

□

[11] Here, we've discreetly changed the (co)domain of permutations σ from $\{1, \ldots, k+1\}$ to $\{0, \ldots, k\}$.

We can summarize our findings regarding $D\alpha$, $d\alpha$, $\omega \wedge_\rho \alpha$ as follows.

	$D\alpha(V_0, \ldots, V_k)$	$d\alpha(V_0, \ldots, V_k)$	$(\omega \wedge_\rho \alpha)(V_0, \ldots, V_k)$
$\ell = 0$	Eq. 47	Eq. 47	0
$\ell = 1$	0	$V_0(\alpha(V_1, \ldots, V_k))$	$-V_0(\alpha(V_1, \ldots, V_k))$
$\ell \geq 2$	0	0	0

3.2.6 Curvature

Definition 3.38 (Curvature Two-Form)**.** The *curvature* of a connection is $\Omega = D\omega$, where ω is the connection one-form.

Theorem 3.39. The following *structure equation* holds.

$$\Omega = d\omega + \frac{1}{2}\omega \wedge_{\mathrm{ad}} \omega = d\omega + \frac{1}{2}[\omega, \omega] \tag{57}$$

Proof. Let $X, Y \in \Gamma(P)$: it suffices to let each of X, Y be either fundamental or horizontal. If both are horizontal, the second right-hand side term vanishes and D, d agree, giving the desired equality. If both are vertical, $X = A^\#$ and $Y = B^\#$ for $A, B \in \mathfrak{g}$. The left-hand side certainly vanishes by the horizontality of Ω from which we can deduce

$$\begin{aligned} d\omega(X,Y) + \frac{1}{2}[\omega,\omega](X,Y) &= A^\#\omega(B^\#) - B^\#\omega(A^\#) - \omega([A^\#, B^\#]) + \frac{1}{2}[\omega,\omega](A^\#, B^\#) \\ &= -\omega([A,B]^\#) + [A,B] \\ &= -[A,B]^\# + [A,B] \qquad = 0 \end{aligned}$$

Finally, if X is horizontal while Y is vertical, we have that $X = A^\#$ for $A \in \mathfrak{g}$. Again, the left-hand side vanishes, and

$$\begin{aligned} d\omega(X,Y) + \frac{1}{2}[\omega,\omega](X,Y) &= A^\#\omega(Y) - Y\omega(A^\#) - \omega([A^\#, Y]) + \frac{1}{2}[\omega,\omega](A^\#, Y) \\ &= Y\omega(A^\#) - \omega([A^\#, Y]) \end{aligned}$$

where terms vanish because they contain $\omega(Y)$. The first term here is the derivation Y acting on a constant function $\omega(A^\#) = A$, so it vanishes; likewise, it follows from Prop. 3.19 that $[A^\#, Y]$ is horizontal, meaning the second term vanishes as well, yielding the desired result and concluding the proof. □

Theorem 3.40. The *Bianchi identity* holds.

$$D\Omega = 0 \tag{58}$$

Proof.

$$\begin{aligned} D\Omega &= D\left(d\omega + \frac{1}{2}[\omega,\omega]\right) \\ &= d^2\omega + \frac{1}{2}d[\omega,\omega] + [\omega, d\omega] + \frac{1}{2}[\omega, [\omega,\omega]] \\ &= \frac{1}{2}[d\omega,\omega] - \frac{1}{2}[\omega, d\omega] + [\omega, d\omega] \\ &= 0 \end{aligned}$$

The fourth term in line two vanishes by the Jacobi identity. □

Finally, we can prove the result concerning Lie brackets of horizontal vector fields discussed earlier.

Proof of Prop. 3.20.

$$[X, Y] \text{ is horizontal} \iff \omega([X, Y]) = 0 \iff d\omega(X, Y) = 0 \iff \Omega(X, Y) = 0 \tag{59}$$

The second step follows from the observation $d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y])$ (see Eq. 47), where the first two terms vanish because X, Y are horizontal. The third step follows because D, d agree on horizontal arguments by definition. ☐

3.2.7 Local Curvature

Analogously to the connection one-form ω, the curvature two-form Ω induces local curvature two-forms Ω_α on U_α, the local trivializations of P. In particular, $\Omega_\alpha = \psi_\alpha^* \Omega$. We are also interested in the transformation laws for these Ω_α on overlaps $U_\alpha \cap U_\beta$. To derive this, we require a classical result from Lie theory.

Proposition 3.41. The Maurer-Cartan one-form θ satisfies $d\theta = -\frac{1}{2}[\theta, \theta]$.

Proof. Letting $X, Y \in \mathfrak{g}$ and recalling Eq. 47, we have that

$$\begin{aligned} d\theta(X, Y) &= X(\theta(Y)) - Y(\theta(X)) - \theta([X, Y]) \\ &= -\theta([X, Y]) \\ &= -[\theta(X), \theta(Y)] \\ &= -\frac{1}{2}[\theta, \theta](X, Y) \end{aligned} \tag{60}$$

The second equality exploits the fact that $\theta(X), \theta(Y)$ are constant on G, while the third equality uses the definition of the Lie bracket on the Lie algebra. Finally, because left-invariant vector fields span $\Gamma(G)$, we see that this formula holds in general by utilizing the linearity of θ to decompose arbitrary elements on $\Gamma(G)$ as linear combinations of left-invariant vector fields, applying the formula, then recombining. ☐

Having established this, we can prove the desired result.

Theorem 3.42. Given local connection one-forms Ω_α, Ω_β defined on U_α and U_β, respectively, the following transformation law holds for $u \in U_\alpha \cap U_\beta$.

$$\Omega_\beta(u) = \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \Omega_\alpha(u) \tag{61}$$

Proof. First, observe the following

$$\begin{aligned} d\omega_\beta(u) &= d(\mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha(u)) + d(\xi_{\alpha\beta}^* \theta)(u) \\ (d\omega_\beta(u)) &= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ d\omega_\alpha(u) + \xi_{\alpha\beta}^* d\theta(u) \\ &= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ d\omega_\alpha(u) - \frac{1}{2}\xi_{\alpha\beta}^*[\theta, \theta](u) \end{aligned}$$

From here, we can see that, exploiting the antisymmetry of the bracket to cancel terms,

$$\begin{aligned}
\Omega_\beta(u) &= d\omega_\beta(u) + \frac{1}{2}[\omega_\beta, \omega_\beta](u) \\
&= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ d\omega_\alpha(u) - \frac{1}{2}\xi_{\alpha\beta}^*[\theta, \theta](u) \\
&\quad + \frac{1}{2}\Big[\mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha(u) + \xi_{\alpha\beta}^*\theta(u), \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha(u) + \xi_{\alpha\beta}^*\theta(u)\Big] \\
&= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ d\omega_\alpha(u) - \frac{1}{2}\xi_{\alpha\beta}^*[\theta, \theta](u) \\
&\quad + \frac{1}{2}\left(\Big[\mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha, \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \omega_\alpha\Big](u) + \Big[\xi_{\alpha\beta}^*\theta, \xi_{\alpha\beta}^*\theta\Big](u)\right) \\
&= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ d\omega_\alpha(u) + \frac{1}{2}\mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ [\omega_\alpha, \omega_\alpha](u) \\
&= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \left(d\omega_\alpha(u) + \frac{1}{2}[\omega_\alpha, \omega_\alpha](u)\right) \\
&= \mathrm{ad}(\xi_{\alpha\beta}^{-1}(u)) \circ \Omega_\alpha(u)
\end{aligned}$$

This concludes the proof[12]. □

3.3 Covariant Derivatives on Vector Bundles

3.3.1 Covariant Derivatives

Definition 3.43 (Covariant Derivative)**.** A *covariant derivative* on a vector bundle $E \to M$ is an $\mathbb{R}$-bilinear map $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ that satisfies the Leibniz rule as follows for $X \in \Gamma(E)$ and functions $f : M \to \mathbb{R}$.

$$\nabla(fX) = df \otimes X + f\nabla X \tag{62}$$

Given $Y \in \Gamma(TM)$ in addition to $X \in \Gamma(E)$, we let $\nabla_Y(X) \in \Gamma(E)$ denote the contraction of the $\Gamma(T^*M)$ part of $\nabla(X)$ with Y. Additionally, given $y \in T_mM$, we let $\nabla_y(X) \in \pi_E^{-1}(m)$ denote the contraction of the $(T_mM)^*$ part of $\nabla(X)(m) \in (T_mM)^* \otimes \pi_E^{-1}(m)$.

Let $\mathscr{D}(E)$ denote the space of covariant derivatives on E and consider the sum and the difference of $\nabla, \nabla' \in \mathscr{A}(E)$ on fX.

$$(\nabla + \nabla')(fX) = df \otimes X + f\nabla(X) + df \otimes X + f\nabla'(X) = 2\,df \otimes X + f(\nabla + \nabla')(X) \tag{63}$$
$$(\nabla - \nabla')(fX) = df \otimes X + f\nabla(X) - df \otimes X - f\nabla'(X) = f(\nabla - \nabla')(X) \tag{64}$$

Evidently, neither the sum nor the difference remains a covariant derivative; moreover, the difference is actually a C^∞-linear map $\Gamma(E) \to \Gamma(T^*M \otimes E)$, or an element of $\Gamma(E^* \otimes T^*M \otimes E) \cong \Omega^1(M, E^* \otimes E)$ (isomorphic as modules). If $C \in \Omega^1(M, E^* \otimes E)$, then $(\nabla + C)(fX) = df \otimes X + f\nabla(X) + fC(X) = df \otimes X + f(\nabla + C)(X)$ (interpreting C as $\Gamma(E) \to \Gamma(T^*M \otimes E)$), hence the sum of a covariant derivative and an element of $\Omega^1(M, E^* \otimes E)$ remains a covariant derivative, enabling the following conclusion.

Proposition 3.44. $\mathscr{D}(E)$ is an affine $C^\infty(M)$-module modelled on $\Omega^1(M, E^* \otimes E)$.

[12] This proof is inspired by [3, Sec. 2.1.3]

In the special case $E = TM$, in a chart $U \subset M$ we have that a section $X \in \Gamma(TU)$ can be expressed in terms of the basis $\frac{\partial}{\partial x^i}$, giving rise to a one-dimensional array $[X]^\mu$. We can define a local covariant derivative ∂ by defining

$$[\partial T]^\mu_\nu = \frac{\partial [T]^\mu}{\partial x^\nu} \tag{65}$$

where x^ν are the coordinate functions on U. The tensor field $\Gamma = \nabla - \partial \in \Omega^1(U, T^*U \otimes TU) \cong \mathcal{T}^1_2(M)$ (isomorphic as $C^\infty(M)$ modules) for some covariant derivative $\nabla \in \mathcal{D}(TM)$ is called the *Christoffel symbol*.

Note that these covariant derivatives can be extended to be a map

$$\Gamma(E \otimes T^*M \otimes \cdots \otimes T^*M) \to \Gamma(E \otimes T^*M \otimes \cdots \otimes T^*M \otimes T^*M) \tag{66}$$

by allowing the covariant derivative to act only on the $\Gamma(E)$ part.

3.3.2 Relationship to Connections

Let Φ denote the isomorphism $\Omega^k_\rho(P, V) \to \Omega^k(M, E)$ as exhibited in Prop. 3.25. The exterior covariant derivative D induces a map $\Omega^k(M, E) \to \Omega^{k+1}(M, E)$ by $\Phi^{-1} \circ D \circ \Phi$. We will abuse notation and refer to both the exterior covariant derivative and this induced map on $\Omega^k(M, E)$ by D.

$$\begin{array}{ccc} \Omega^k_\rho(P, V) & \xrightarrow{D} & \Omega^{k+1}_\rho(P, V) \\ {\scriptstyle \Phi}\downarrow & & \downarrow{\scriptstyle \Phi} \\ \Omega^k(M, E) & \xrightarrow[\nabla]{} & \Omega^{k+1}(M, E) \end{array}$$

In particular, this induced map restricts to a map $\nabla : \Omega^0(M, E) \to \Omega^1(M, E)$, or equivalently, $\Gamma(E) \to \Gamma(T^*M \otimes E)$. Hence, as the notation was constructed to suggest, this ∇ is a candidate for a covariant derivative on the associated vector bundle E. We show that this map indeed satisfies the desired properties.

Proposition 3.45. The map $\nabla = (\Phi^{-1} \circ D \circ \Phi)|_{\Omega^0(M,E) \cong \Gamma(E)}$ (isomorphic as $C^\infty(M)$ modules) is a covariant derivative.

We refer to ∇ as the *covariant derivative induced by* D, or the *covariant derivative induced by a connection* on P. We might hope that ∇ acts on sections of E in a matter analogous to D upon elements of $\Omega^k(P, V)$ (i.e., in the nice form exhibited by Th. 3.37). It turns out that this is indeed the case, but only locally.

Proposition 3.46. Given a local trivialization (U_α, ϕ_α) on the bundle E associated to P by a representation ρ, the covariant derivative ∇ induced by D takes the following form on a section $V \in \Gamma(E)$ which reduces to $V_\alpha : U \to V$ on U.

$$\nabla(V_\alpha) = dV_\alpha + \omega_\alpha \wedge_\rho V_\alpha \tag{67}$$

Proof of Prop. 3.45 and Prop. 3.46. See [5, Prop 5.9.4], although this is written to some extent from a more physical standpoint. □

3.4 Connections and Bundle Automorphisms

The set $\mathcal{G}(P)$ of automorphisms of P forms a group. We begin by investigating the transformation behavior of characterizations of $\Psi \in \mathcal{G}(P)$ in local trivializations[13]. This is not unlike the procedure we've undergone for ω and Ω.

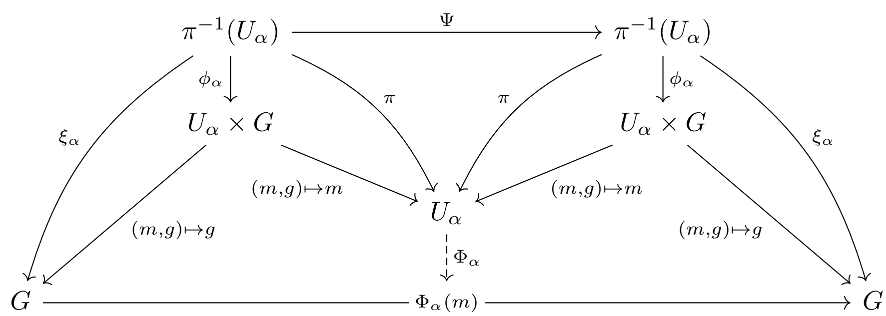
Recall the local trivialization $\phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times G$ acting by $p \mapsto (\pi(m), \xi_\alpha(p))$. Then Ψ restricts to an automorphism on the bundle $\pi^{-1}(U_\alpha)$, and $\phi_\alpha \circ \Psi$ is $p \mapsto (\pi(p), (\xi_\alpha \circ \Psi)(p))$. We are free to define $\Psi_\alpha : \pi^{-1}(U) \to G$ by $p \mapsto (\xi_\alpha \circ \Psi)(p)\xi_\alpha(p)^{-1}$, in which case

$$(\phi_\alpha \circ \Psi)(p) = (\pi(p), \Psi_\alpha(p)\xi_\alpha(p)). \tag{68}$$

Now, recalling the G-equivariance of Ψ, ϕ_α, and ξ_α, note that

$$\begin{aligned}
(\pi(pg), \Psi_\alpha(p)\xi_\alpha(p)g) &= (\pi(pg), \Psi_\alpha(p)\xi_\alpha(p))g \\
&= (\phi_\alpha \circ \Psi)(pg) \\
&= (\pi(pg), \Psi_\alpha(pg)\xi_\alpha(pg)) \\
&= (\pi(pg), \Psi_\alpha(pg)\xi_\alpha(p)g)
\end{aligned}$$

from which it follows that, in particular, $\Psi_\alpha(p) = \Psi_\alpha(pg)$, hence Ψ_α is really a function $U_\alpha \to G$ and, more generally, Ψ acts on a local trivialization by $(m, g) \mapsto (m, \Psi_\alpha(m)g)$, where $\Psi_\alpha(m) \in G$.



It's worth noting the similarity between the diagrams describing the application of a bundle automorphism and the one describing changes of local trivializations on fibre bundles more generally: indeed, we will exploit this similarity in structure imminently.

In the meantime, though, we return to our objective of understanding the local behavior of Φ as follows. In particular, we pose the question "how are Ψ_α, Ψ_β related on an overlap $U_\alpha \cap U_\beta$?"

Proposition 3.47. Given an automorphism $\Psi \in \mathcal{G}(P)$ with associated maps $\Psi_\alpha : U_\alpha \to G$ and $\Psi_\beta : U_\beta \to G$, the following transformation law holds on an overlap $U_\alpha \cap U_\beta \ni u$.

$$\Psi_\alpha(u) = A_{\xi_{\alpha\beta}(u)}\Psi_\beta(u) \tag{69}$$

[13]The results of this section are inspired by [3, Sec. 1.5 and 1.6]

Proof. Let $\pi(p) = (\pi \circ \Psi)(p) = m$.

$$\begin{aligned}\Psi_\alpha(m) &= (\xi_\alpha \circ \Psi)(p)\xi_\alpha(p)^{-1}) \\ &= (\xi_\alpha \circ \Psi)(p)\Big[(\xi_\beta \circ \Psi)(p)^{-1}(\xi_\beta \circ \Psi)(p)\Big]\Big[\xi_\beta(p)^{-1}\xi_\beta(p)\Big]\xi_\alpha(p)^{-1} \\ &= \xi_{\alpha\beta}(m)\Psi_\beta(m)\xi_{\beta\alpha}(m) \\ &= A_{\xi_{\alpha\beta}(m)}\Psi_\beta(m)\end{aligned}$$

□

This transformation behavior motivates the following characterization of $\mathcal{G}(P)$.

Definition 3.48. Given a principal bundle P, we define $\operatorname{Ad} P$ as the bundle associated to P by the adjoint action of G upon itself: namely, the representation $g \mapsto (h \mapsto ghg^{-1})$.

Corollary 3.49. $\mathcal{G}(P) = \Gamma(\operatorname{Ad} P)$.

Proof. Follows from Prop. 3.47 and Prop. 2.8. □

Thus, principal bundle automorphisms are locally merely a (smooth) choice of $g \in G$ at each point in M.

We now note that the group $\mathcal{G}$ acts naturally on $\mathcal{A}(P)$ by pullback, but it perhaps isn't immediately obvious that $\mathcal{G}$ preserves $\mathcal{A}(P)$. We begin investigating this matter by asking how local connection forms ω_α defined on U_α transform under pullback by Φ. It is at this point we will apply the observation that bundle automorphisms behave similarly to changes in local trivializations.

Proposition 3.50. We have the following transformation law for local connection-one forms ω_α and any $u \in U_\alpha$.

$$\Phi^*\omega_\alpha(u) = \operatorname{ad}(\Psi_\alpha(u)) \circ \omega_\alpha(u) + (\Psi_\alpha^{-1})^*\theta(u) \tag{70}$$

Proof. We can understand $\Phi^*\omega_\alpha$ as ω_β, the local connection one-form for a distinct local trivialization (U_β, ϕ_β). In particular, $U_\beta = U_\alpha$ and $\psi_\beta = \psi_\alpha \circ \Phi$. Then $U_\alpha \cap U_\beta = U_\alpha$ (so Eq. 70 is well-defined on all U_α). Eq. 68 implies that $\xi_{\beta\alpha}(m) = \Psi_\alpha(p)$, so we can directly apply Th. 3.32 to arrive at Eq. 70. □

Here, we have exploited the coincidence between the action of $\mathcal{G}$ on $\mathcal{A}(P)$ and the transformation law of local connection-one forms to show that this $\mathcal{G}$-action is well-defined. In particular, a local bundle automorphism, or an automorphism of $\pi^{-1}(U)$ for some local trivialization (U, ϕ), is identifiable with a change in local trivialization.

4 Gauge Theory Dictionary

This section is a qualitative comparison between the terminology employed in the less formal use of gauge theoretic objects in particle physics and that employed in their more formal use in mathematics. The latter of these has been our focus thus far: hence, this section will necessarily involve a brief, casual discussion of some of the ideas of the former.

In physics, (classical) field theories describe physical systems through *fields*, which assign a particular object to each point in spacetime (or some subset). Formally, these are sections of fibre bundles whose base manifold is spacetime M (some four-dimensional Lorentzian manifold). In particular, in particle physics, the fields associated with fermions (or matter particles), denoted *matter fields*, are characterized by sections of vector bundles $V \hookrightarrow E \to M$. However, rather than beginning with the "global" definition (Def. 2.7) and deriving a "local" characterization (Prop. 2.8), physics approaches fields inversely, choosing to assume that sections are, by definition, maps from $M \to V$ which transform in a particular way (we will clarify this shortly). Physicists can get away with this because M is nearly always contractible, hence by Th. 2.2 local trivializations (U, ϕ) are free to satisfy $U = M$ and local sections thus can have domain M.

So what do we mean when we say matter fields transform in a certain way? *Symmetries* are of the utmost importance across physics: Noether's theorem is perhaps the primary manifestation of this. In particle physics symmetries abound, but two of the more important instances are the Lorentz symmetry (the equivalence of inertial reference frames, or invariance under transformations mapping between inertial reference frames) and gauge symmetries (so-called "internal symmetries" which give rise to the fundamental forces: strong, weak, and electromagnetic). Formally, symmetries in general are groups: these particular symmetries are Lie groups.

In particular, for the field theories to be invariant under certain symmetries, there must be a sensible way of applying a symmetry transformation to the fields, the dynamical objects. Moreover, physicists demand that the Lorentz/gauge symmetries are "local," meaning that the symmetry holds not merely when a single $g \in G$ is applied uniformly to a field (which a physicists would call a global symmetry), but also when that g is allowed to vary smoothly across the manifold (i.e., it becomes something like $g(m) : M \to G$). This kind of transformation—wherein a symmetry's group acts upon a field in a smoothly varying way–is referred to as a *gauge transformation*.

Formally, this is handled by allowing the group G associated with a given symmetry (termed the *gauge group* by physicists) to be the structure group of the bundle E. It's worth commenting now that, because matter fields frequently have multiple symmetries, matter fields are most often sections of tensor products of vector bundles, each with a distinct structure group such that all of the relevant symmetries are accounted for. In particular, locally gauge transformations are changes of local trivialization, whereas globally they are bundle automorphisms. In the event of local trivializations (U, ϕ) satisfying $U = M$, we recall that these notions coincide. Physicists refer to $\mathcal{G}(E)$ as the *group of gauge transformations.* Local trivializations themselves are denoted *gauges* and a choice of a particular local trivialization is referred to as *fixing a gauge* (which physicists do often: indeed, recall that they see sections as most fundamentally being expressed in "fixed gauges"). Although the word "gauge" is the one repeatedly employed here, on curved spacetime Lorentz symmetry behaves in precisely the same way.

Physicists summarize their theories through *Lagrangians*, which are functions mapping the set of relevant matter fields to a real-valued function on the manifold (i.e., $f : M \to \mathbb{R}$).

From here, physicists have a well-known method for extracting partial differential equations which describe the dynamics of the fields (the Euler-Lagrange equations, defined through the calculus of variations). Physicists say a theory is *gauge invariant* (*Lorentz invariant*) if its associated Lagrangian is invariant under the group of gauge transformations on the G-vector bundle E (where G is the Lorentz group in the Lorentz invariance case).

In particular, physicists hope to include derivatives of their matter fields in their Lagrangians (and in their theories more broadly), and they expect the derivatives of matter fields to be matter fields themselves. But the naive approach—going into a local trivialization and differentiating components of local sections with respect to the spacetime base manifold coordinates—fails: specifically, matter fields differentiated in this way fail to transform properly under local gauge transformations if they are acted upon non-trivially by the gauge group. Some algebra reveals that this naive derivative (often denoted ∂_μ) requires a counter-term A_μ. This coincides with what we have found mathematically: namely, that a well-defined covariant derivative along a vector field (or, alternatively, a Dirac operator) is locally the sum of d and the local connection one-form ω_α.

Products of distinct fields in a Lagrangian are physically interpreted as *interactions* (a notion backed up by the dynamics of the ensuing Euler-Lagrange equations). The necessity of covariant differentiation (versus ordinary differentiation) mandates terms in the Lagrangian that are products of matter fields and the connection one-form. Physicists perceive this as the introduction of a brand new dynamical field, which they term the *gauge field* or *gauge potential*. In this way, all matter fields which transform non-trivially with respect to a given structure group G interact with that structure group's associated gauge field, making it natural for physicists to understand that gauge group as being responsible for a force whose interactions are between matter fields acted upon non-trivially by the gauge group, for which the gauge field is a mediator (as it is always involved in the interaction). Moreover, terms involving just the gauge field also can be naturally included in the Lagrangian: in particular, it turns out that the curvature of the connection can easily be made into a Lagrangian term. Physicists refer to the curvature as the *field strength*, as indeed the curvature of a connection ultimately corresponds to "stronger" fields with more non-trivial dynamics.

At this point, principal G-bundles P become a very natural thing to introduce. Suddenly gauge fields and field strengths can be understood as legitimate sections on P (although this wasn't a problem for the field strength), and all of the matter field-housing vector bundles with structure group G are merely bundles associated to P by a representation ρ.

We conclude by summarizing our discussion with a dictionary enabling translation between the gauge theoretic vocabulary of physics and mathematics.

Physics		**Mathematics**
(Matter) field	$\longleftrightarrow$	Section of a (vector) fibre bundle
Symmetry	$\longleftrightarrow$	Group
Gauge group	$\longleftrightarrow$	Structure group
Local gauge transformation	$\longleftrightarrow$	Change of local trivialization
Global gauge transformation	$\longleftrightarrow$	Bundle automorphism
Gauge fixing	$\longleftrightarrow$	Choice of local trivialization
Gauge	$\longleftrightarrow$	Local trivialization
Gauge field/potential	$\longleftrightarrow$	Connection one-form
Lagrangian	$\longleftrightarrow$	Map from sections of bundles to real-valued function on base
Interaction	$\longleftrightarrow$	Product of distinct sections in the Lagrangian
Field strength	$\longleftrightarrow$	Curvature of connection

5 Review of Lie Representation Theory

In this section, we review important Lie theoretical facts regarding representations whose theory extends beyond the scope of this text but which remain relevant for our culminating section.

5.1 Introductory Definitions and Theorems

We recall some basic definitions of Lie group representation theory. A representation ρ of a group G on a (complex) vector space V is a homomorphism $\rho : G \to \mathrm{GL}(V)$. For Lie groups, we require ρ be smooth: for the remainder of this text, all representations are Lie group representations. The representation ρ is unitary if its image lies within $\mathrm{U}(V)$, the group of unitary linear transformations on V. Additionally, ρ is irreducible if it does not restrict to a well-defined representation on a non-trivial linear subspace of V.

More broadly, Lie representation theory is often vastly simplified by descending from the non-linear structure of a Lie group G to the much more tractable, linear structure of its Lie algebra $\mathfrak{g}$.

Theorem 5.1 (Lie's Third Theorem)**.** There is an equivalence between the category of simply-connected Lie groups and the category of Lie algebras.

Proof. See [16, Th. 3.28], which in turn cites [6, Ch. 6] where Ado's theorem is proven. □

Corollary 5.2. Given a simply-connected Lie group G with Lie algebra $\mathfrak{g}$, there is a bijection between representations of G and representations of $\mathfrak{g}$

Finally, we would like to reduce the study of Lie group (or Lie algebra) representations to the study of their building blocks: irreducible representations (irreps). In general, arbitrary representations are not the direct sum of irreps. However, this property does hold for a superset of the groups we are ultimately interested in. We recall that a Lie algebra is simple if it is non-Abelian and contains no non-trivial proper ideals (analogous to the definition of a simple group).

Definition 5.3. A Lie algebra is *semisimple* if it is the direct sum of simple Lie algebras.

We say that a Lie group is semisimple if its Lie algebra is semisimple.

Theorem 5.4. Representations of semisimple Lie groups and compact Lie groups decompose as the direct sum of irreps.

Proof. For the semisimple case, see [6, Ch. 3.7]. For the compact case, see [4, Prop 4.28]; while this result is for compact matrix Lie groups, this in fact exhausts compact Lie groups per [7, Cor 4.22]. □

We will exclusively consider groups that are either semisimple or compact[14], thus these results assure us that we have access to decomposability, motivating the study of irreps. We conclude this introduction by noting the simplicity of the Abelian case.

[14] It turns out that enforcing that gauge groups be both semisimple and compact is a useful way to ensure physically viable gauge theories: in particular, these conditions assure you that the Yang-Mills Lagrangian terms defined in the first line of Eq. 84 are positive definite, which it should be, as this is supposed to be a kinetic term.

Proposition 5.5. Irreps of Abelian groups are one-dimensional.

Proof. See [4, Cor. 4.31]. □

Corollary 5.6. The irreps of $U(1)$ are exactly the maps $\theta \mapsto e^{in\theta}$ for $n \in \mathbb{Z}$.

We refer to the irrep $\theta \mapsto e^{in\theta}$ by n.

5.2 Important Representations

We state the classification results for the finite dimensional irreps of the groups we care about (which haven't already been discussed). Namely, these are $SU(2)$, $SU(3)$ and $\mathrm{Spin}(3,1)$.

Theorem 5.7. The irreps of $SU(2)$ are indexed by a non-negative integer ℓ such that the irrep ℓ has dimension $\ell + 1$.

Proof. See [12, Sec. 6.4]. □

Theorem 5.8. The irreps of $SU(3)$ are indexed by a pair of non-negative half-integers (ℓ, k) such that that the irrep (ℓ, k) has dimension $\ell + 1$.

Proof. See [4, Th. 6.7]. □

Notations abound for irreps of the special unitary groups. For example, physicists frequently prefer to denote $SU(2)$ representations by $\frac{\ell}{2}$ (e.g., spin-$\frac{1}{2}$ particles). However, in accordance with popular particle physics notation, we will avoid use of the integer(s) used above to characterize $SU(N)$ irreps[15] (or their fractions) and choose to instead refer to irreps by their dimension, labelled in boldface (and possible with an overline: we will elaborate on this in a moment).

Intuitively, for every irrep of G on a complex vector space V we should have both a dual irrep on V^* and a complex conjugate irrep on $\overline{V}$ (the complex conjugate vector space). In general, each of these representations may or may not be equivalent. We state a few results that clarify this in some of the special cases we're interested in.

Proposition 5.9. For compact Lie group representations, the dual representation and the complex conjugate representation are equivalent.

Proof. In the proof of [4, Th. 4.28], we see that representations of compact Lie groups can always be made into unitary representations. The inner product in this case maps vectors in V to elements of $\overline{V}^*$, the space of all antilinear maps $V \to \mathbb{C}$. We therefore have a dual map $\overline{V} \to V^*$, which is what we need for the desired equivalence of representations. □

This means, for $SU(3)$ and $SU(2)$, we can use overlines to denote either the dual or the complex conjugate representation.

In this thesis, we are ultimately interested in arbitrary $U(1)$ representations, the $SU(2)$ irreps $\mathbf{2} \cong \overline{\mathbf{2}}, \mathbf{1}$ (which correspond to $1, 0$ in the notation of Th. 5.8), the $SU(3)$ irreps $\mathbf{3}, \overline{\mathbf{3}}, \mathbf{1}$ (equivalent representations; as it turns out, these correspond to $(1,0)$, $(0,1)$ and $(0,0)$ in the notation of Th. 5.8). We will also be interested in certain representations of spin groups, but this will be discussed in the next section.

[15] These are the so-called *highest weights* of the irrep, arising from classification theory for semisimple Lie algebra representations. Entire chapters are dedicated to this theory in, e.g., [4, 12].

6 Spin

This exposition is inspired by [2, Sec. 4.6] and [5, Sec. 6]

6.1 Clifford Algebras and Spin Groups

Let $\mathcal{T}V$ denote the tensor algebra of an $\mathbb{F}$-vector space V ($\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$) with a symmetric non-degenerate bilinear form[16] $B : V \times V \to \mathbb{F}$. That is, $\mathcal{T}(V) = \mathcal{F} \oplus V \oplus V \otimes V \oplus \dots$. We let $\mathcal{T}^0(V), \mathcal{T}^1(V)$ denote the even and odd tensor powers: i.e., $\mathcal{T}(V) = \mathcal{T}^0(V) + \mathcal{T}^1(V)$. Let the signature of B be (k, ℓ) such that $k + \ell = n = \dim(V)$

Definition 6.1. The *Clifford algebra* of V is $Cl(V, B) = \mathcal{T}V$ modulo the equivalence relation $u \otimes v + v \otimes u \sim -2B(u, v)$[17] for $u, v \in V$. We let $Cl^0(V, B)$ and $Cl^1(V, B)$ denote quotients of $\mathcal{T}^0(V)$ and $\mathcal{T}^0(V)$, respectively (i.e., $Cl(V, B) = Cl^0(V, B) + Cl^1(V, B)$.

In $Cl(V, B)$ we typically suppress the $\otimes$ symbol. We now introduce some notation: namely, $\mathbb{C}l(n) = Cl(\mathbb{C}^n, B)$ (B the complex linear Euclidean dot product), $Cl(n) = Cl(\mathbb{R}^n, B)$ (B the Euclidean dot product), and $Cl(k, \ell) = Cl(\mathbb{R}^{k,\ell}, B)$ (B a signature (k, ℓ) bilinear form). We have the following relationships. We adopt a similar convention for Cl^0 and Cl^1.

Proposition 6.2. We have the following algebra isomorphisms.

$$\mathbb{C}l(n) \cong \begin{cases} \mathrm{End}(\mathbb{C}^{2^{n/2}}) & n \text{ is even} \\ \mathrm{End}(\mathbb{C}^{2^{(n-1)/2}}) \oplus \mathrm{End}(\mathbb{C}^{2^{(n-1)/2}}) & n \text{ is odd} \end{cases} \tag{71}$$

$$\mathbb{C}l^0(n) \cong \begin{cases} \mathrm{End}(\mathbb{C}^{2^{n/2}/2}) \oplus \mathrm{End}(\mathbb{C}^{2^{n/2}/2}) & n \text{ is even} \\ \mathrm{End}(\mathbb{C}^{2^{(n-1)/2}}) & n \text{ is odd} \end{cases} \tag{72}$$

Proof. See [5, Th. 6.3.21] □

Proposition 6.3. We have an algebra isomorphism $\mathbb{C}l(k + \ell) \cong Cl(k, \ell) \otimes \mathbb{C}$. In particular, complex representations of $Cl(k, \ell)$ coincide with the complex representations of $\mathbb{C}l(k + \ell)$.

We now turn our attention and observe that, given an orthonormal basis $\{e_1, \dots, e_n\}$ for V (using B) we have that $e_i e_j = -e_j e_i$: thus intrinsic to Clifford algebras is a certain antisymmetric behavior. We also have $e_i^2 = -B(e_i, e_i) \in \{\pm 1\}$: hence, there will be k instances of 1 and ℓ instances of -1 across the basis e_i[18]. More generally, $v^{-1} = -v/B(v, v)$ for unit vectors $v \in V$: in particular, unit vectors are invertible, enabling the following definition.

Definition 6.4 (Pin Group)**.** The *pin group* $\mathrm{Pin}(V, B) \subset Cl(V, B)$ is the (Lie) group generated by unit vectors in $V \subset Cl(V, B)$.

There is an action of $\mathrm{Pin}(V, B)$ on V.

[16] Note that in the case $\mathbb{F} = \mathbb{C}$, this is complex-linear, not Hermitian

[17] Sometimes the sign on the right hand side here is omitted, but more frequently and canonically it is preserved.

[18] Notice how the minus sign on the right-hand side of the fundamental equivalence relation entails a swapping between k and ℓ from their usual sign affiliations

Proposition 6.5. The action of $\text{Pin}(V,B)$ on V given by conjugation, or

$$v(w) = -vwv^{-1} \tag{73}$$

for v a generator of $\text{Pin}(V,B)$ (i.e., $v \in V$ is a unit vector) and $w \in V$, is well-defined.

Proof. Observe the following.

$$\begin{aligned}
-vwv^{-1} &= \frac{vwv}{B(v,v)} \\
&= -\frac{v(vw + 2B(v,w))}{B(v,v)} \\
&= -\frac{-B(v,v)w + 2B(v,w)v}{B(v,v)} \\
&= w - 2\frac{B(v,w)}{B(v,v)}w
\end{aligned} \tag{74}$$

This final expression is evidently a vector in V. □

Intuitively, Eq. 74 describes reflection of w across v (as this is equivalent to subtracting twice the projection of w onto v). Thus, this isn't a generic action: $\text{Pin}(V,B)$ acts by compositions of reflections. In particular, these are B-preserving maps: we formalize this property of the action as follows.

Definition 6.6 (Orthogonal Group)**.** The *orthogonal group* $O(V,B)$ is the (Lie) subgroup of $GL(V)$ preserving B: namely, given $A \in O(V,B)$ and $w_1, w_2 \in V$, $B(A(w_1), A(w_2)) = B(w_1, w_2)$.

Proposition 6.7. The action specified in Prop. 6.5 is by elements of $O_{k,\ell}(V)$.

Proof. Let v be a generator of $\text{Pin}(V,B)$ and $w_1, w_2 \in V$: it suffices to show that

$$B(v(w_1), v(w_2)) = B(w_1, w_2) \tag{75}$$

as if this holds for generators, it will hold for arbitrary elements of $\text{Pin}(V)_{k,\ell}$, because $O_{k,\ell}$ is a group. Recalling Eq. 74, we have

$$\begin{aligned}
B(v(w_1), v(w_2)) &= B(w_1, w_1) - 2B\left(w, \frac{2B(v,w)}{B(v,v)}v\right) + B\left(\frac{2B(v,w)}{B(v,v)}v, \frac{2B(v,w)}{B(v,v)}v\right) \\
&= B(w_1, w_1) - \frac{4B(v,w)}{B(v,v)}B(w,v) + \left(\frac{2B(v,w)}{B(v,v)}\right)^2 B(v,v) \\
&= B(w_1, w_1) - \frac{4B(v,w)^2}{B(v,v)} + \frac{4B(v,w)^2}{B(v,v)} \\
&= B(w_1, w_1)
\end{aligned}$$

□

Moreover, appealing to classical result of linear algebra, we have the following fact.

Proposition 6.8. The homomorphism $p : \text{Pin}(V,B) \to O(V,B)$ given by the action of Prop. 6.5 is a surjection.

Proof. This follows from the earlier observation that the generators of $\text{Pin}(V, B)$ act by reflection and the linear algebraic result that all orthogonal transformations are some finite composition of reflections. ☐

From the Pin group we can progress to Spin as follows.

Definition 6.9 (Special orthogonal group)**.** The *special orthogonal group* $SO(V, B)$ is the (Lie) subgroup of $O(V, B)$ given by transformations with determinant 1. In particular, we let $SO^+(V, B)$ denote the identity component of $SO(V, B)$.

Definition 6.10 (Spin Group)**.** The *spin group* $\text{Spin}(V, B)$ is the (Lie) subgroup of $\text{Pin}(V, B)$ given by the preimage of $SO(V, B)$ by p as defined in Prop. 6.8. Moreover, we let $\text{Spin}(V, B)^+$ denote the preimage of $SO^+(V, B)$.

Theorem 6.11. $\text{Spin}^+(V, B)$ is the universal cover of the identity component of $SO^+(V, B)$.

Proof. See [5, Cor. 6.5.16] ☐

We conclude by expanding some of our earlier notation in a natural way. Namely, we let $\text{Pin}(k, \ell) = \text{Pin}(\mathbb{R}^n, B)$ where B is a bilinear form with signature (k, ℓ), and we adopt a similar convention for Spin and Spin^+.

6.2 Spin Bundles

We now endeavor to define a spin representation. This notion takes on different meanings across the mathematical and physics literatures: we chose a particular approach here to achieve some level of clarity, but we recognize that our terminology will inevitably differ from other pedagogical and scholarly conventions.

Definition 6.12. The Dirac spinor representation of $\mathbb{C}l(n)$ is the complex representation $\rho : \mathbb{C}l(n) \to \mathbb{C}^N$ per the algebra isomorphism described by Prop. 6.2. In particular,

$$N = \begin{cases} 2^{n/2} & n \text{ is even} \\ 2^{(n-1)/2} & n \text{ is odd} \end{cases} \tag{76}$$

and in the case where n is odd, we merely adopt the restriction of the isomorphism to the first term in the direct sum.

Corollary 6.13. The Dirac spinor representation for $\mathbb{C}l(n)$ restricts to an irrep

$$\rho|_{\mathbb{C}l^0(n)} : \mathbb{C}l^0(n) \to \text{End}(\mathbb{C}^{2^{(n-1)/2}}) \tag{77}$$

if n is odd, otherwise it restricts to the reducible representation with invariant subspaces as follows

$$\rho|_{\mathbb{C}l^0(n)} : \mathbb{C}l^0(n) \to \text{End}(\mathbb{C}^{2^{n/2}/2}) \oplus \text{End}(\mathbb{C}^{2^{n/2}/2}) \tag{78}$$

Proof. Follows from Prop. 6.2. ☐

In the case where n is even, we refer to the two irreducible representations arising from restricting to $\mathbb{C}l^0(n)$ as the *left and right Weyl spin representations.*

Corollary 6.14. The Dirac spin representation on $\mathbb{C}l^0(n)$ induces a complex representation $\rho : \text{Spin}^+(k, \ell) \to \mathbb{C}^N$ for all $k + \ell = n$.

Proof. Follows from Prop. 6.3. □

In an abuse of notation, we also refer use the Dirac/Weyl nomenclature for these $\mathrm{Spin}(k,\ell)^+$ representations. Note that for fixed k, ℓ there is a unique Dirac spinor representation.

Definition 6.15 (Spinor)**.** A *spinor* is an element of a vector space acted upon by a spinor representation.

We now turn our attention toward constructing associated bundles acted on by these representations. It turns out that this need not be possible for all base manifolds M.

Definition 6.16 (Spin Structure)**.** Given a manifold M of dimension n with a (k, ℓ) signature metric g, a *spin structure of type* (k, ℓ) on M is a map $\Pi : \tilde{P} \to P$ where $P \to M$ is the principal bundle of oriented orthonormal frames (i.e., a principal $SO(k, \ell)$-bundle) and $\tilde{P} \to M$ is a principal $\mathrm{Spin}(k, \ell)$-bundle such that Π is a double covering and equivariant with respect to the principal bundle right actions.

Theorem 6.17. A manifold M admits a spin structure if and only if the second Steifel-Whiteney class of TM vanishes.

Proof. See [8, Th. II.1.7]. □

We say a manifold is *spin* if it admits a spin structure.

Definition 6.18 (Dirac Spin Bundle)**.** Given a manifold M with spin structure $\Pi : \tilde{P} \to P$, a *Dirac spin bundle*, or just a spin bundle, is a vector bundle associated to the principal bundle $\tilde{P}$ by a spinor representation of Spin.

Sections of spin bundles are referred to as spin fields, or even (by an abuse of notation) just spinors. Immediately we have two properties of Dirac spin bundles.

Proposition 6.19. Given a spin bundle $S \to M$, there is a bilinear Clifford multiplication $\Gamma(TM) \times \Gamma(S) \to \Gamma(S)$.

Proof. See [8, Sec. II.3] or [5, Prop. 6.9.13] □

This multiplication can be understood as a generalization of multiplication of elements of a Clifford algebra by elements of the underlying vector space.

Proposition 6.20. Given a spin bundle $S \to M$ for even-dimensional M with spin structure $\Pi : \tilde{P} \to P$, there is a splitting of $S = S_+ \oplus S_-$ where $S_\pm$ are the associated bundles to $\tilde{P}$ by the two Weyl representations. Clifford multiplication on the spinor bundle is then a map $\Gamma(S_\pm) \to \Gamma(S_\mp)$.

Proof. See [5, Prop. 6.9.13]. □

We refer to these two complementary subbundles of spin bundles for even dimensional base manifold as the left $(+)$ and right $(-)$ Weyl spin bundles.

It is upon spin bundles that we will be able to define a special derivative operator, which we turn our attention toward in the next section. Before moving beyond this section, we note the following regarding covariant differentiation on spin bundles.

Proposition 6.21 (Spin Connection)**.** Given a spinor bundle $S \to M$ for a pseudo-Riemannian spin manifold M, there is a unique covariant derivative ∇ on S that is both compatible with the metric and with Clifford multiplication in the following sense.

$$\frac{d}{dt}\langle X, Y\rangle = \langle \nabla_v(X), Y\rangle + \langle X, \nabla_v(Y)\rangle \tag{79}$$

$$\nabla_v(\alpha \cdot X) = \nabla_v^{LC}(\alpha) \cdot X + \alpha \cdot \nabla_v(X) \tag{80}$$

Here $X, Y \in \Gamma(S)$, $\frac{d}{dt}$ is differentiation along a curve with tangent vector field $v \in \Gamma(M)$, $\alpha \in \Omega^1(M)$, and ∇^{LC} is the Levi-Civita covariant derivative on M. We refer to ∇ as the *spin connection.*

6.3 Dirac Operator

Definition 6.22 (Dirac Operator)**.** Let M be a manifold with spin structure $\Pi : \tilde{P} \to P$, a connection on $\tilde{P}$, and a vector bundle $V \hookrightarrow E \to M$ associated to $\tilde{P}$ with induced covariant derivative ∇[19]. Let $X \in \Gamma(E)$: in a local trivialization (U_α, ϕ_α), this is a map $X_\alpha U_\alpha \to V$. Choose an orthonormal basis $\{e_1, e_2, e_3, e_4\}$ for $\Gamma(U_\alpha)$: then the *Dirac operator* on E is $\not\nabla : \Gamma(E) \to \Gamma(E)$ given by

$$(\not\nabla(X))_\alpha(m) = \sum_{i=1}^{4} e_i(m)\Big(\nabla_{e_i(m)}(X_\alpha)\Big) \tag{81}$$

Proposition 6.23. The Dirac operator is well-defined.

Proof. See pages 31 and 32 of [1]. □

The Dirac operator was originally motivated by a desire to construct a first-order differential operator which squared to the Laplacian. However, in a general differential geometric environment, there are many Laplacians to choose from: for instance, one has the connection Laplacian $\Delta : \Gamma(E) \to \Gamma(E)$ given by

$$\Delta(X) = \mathrm{Tr}(\nabla^2(X)) \tag{82}$$

where $X \in \Gamma(E)$, hence $\nabla^2(X) \in \Gamma(E \otimes T^*M \otimes T^*M) \cong \Gamma(E \otimes \mathrm{End}(TM))$ (isomorphic as C^∞-modules using the metric on M) and Tr refers to taking the (point-wise) trace on the $\Gamma(\mathrm{End}(TM))$ part. There is additionally the closely related Bochner Laplacian: both are closely related to the square of a Dirac operator, yet the strict equality arises when one considers the Hodge Laplacian (or the Laplace–de Rham operator) defined as $\Delta = d\delta + \delta d = (d+\delta)^2$, where δ is the codifferential, or the adjoint to the exterior derivative with respect to the norm induced on differential k-forms via the Hodge star operator arising from the metric. For a review of these objects, the reader may refer to, for instance, [8, Pg. 123][20].

Proposition 6.24. Given a manifold M with spin structure $\Pi : \tilde{P} \to P$, $\not\nabla^2 = \Delta$ on an associated vector bundle E to $\tilde{P}$ for the Hodge Laplacian Δ.

Proof. See in [8, Th. 5.12]. □

---

[19] In particular, given our restriction to the case $(k, \ell) = (1, 3)$, M is necessarily a four-dimensional Lorentzian manifold (i.e., M has a metric with signature $(1, 3)$); also, Spin acts on V.

[20] Note that the connection Laplacian and the Bochner Laplacian depend on a choice of connection, whereas the Hodge Laplacian is intrinsic.

7 The Standard Model

7.1 The Domain of the Mathematician's Standard Model Lagrangian

Let M be a four dimensional Lorentzian manifold with spin structure $\Pi : \tilde{P} \to P$. Thus, we have the spin bundle $S = \tilde{P} \times_\rho \mathbb{C}^4$ where $\rho : \mathrm{Spin}(3,1) \to \mathbb{C}^4$ is the spin representation of the universal cover of the identity component of the Lorentz group, $SO^+(3,1)$. Because M is even dimensional, we have that $S = S_+ \oplus S_-$. From here, let

$$\begin{aligned} SU(3) &\hookrightarrow P^3 \to M \\ SU(2) &\hookrightarrow P^2 \to M \\ U(1) &\hookrightarrow P^1 \to M \end{aligned}$$

be principal bundles, from which we define the associated bundles

$$\begin{aligned}
E^3_{\mathbf{3}} &= P^3 \times_{\mathbf{3}} \mathbb{C}^3 \\
E^3_{\mathbf{1}} &= P^3 \times_{\overline{\mathbf{3}}} \mathbb{C}^3 \\
E^3_{\mathbf{1}} &= P^3 \times_{\mathbf{1}} \mathbb{C} = M \times \mathbb{C}^3 \\
E^2_{\mathbf{2}} &= P^2 \times_{\mathbf{2}} \mathbb{C}^2 \\
E^2_{\mathbf{1}} &= P^2 \times_{\mathbf{1}} \mathbb{C} = M \times \mathbb{C}^2 \\
E^1_{\frac{1}{6}} &= P^1 \times_{\frac{1}{6}} \mathbb{C} \\
E^1_{\frac{2}{3}} &= P^1 \times_{\frac{2}{3}} \mathbb{C} \\
E^1_{\text{-}\frac{1}{3}} &= P^1 \times_{\text{-}\frac{1}{3}} \mathbb{C} \\
E^1_{1} &= P^1 \times_{1} \mathbb{C} \\
E^1_{\text{-}\frac{1}{6}} &= P^1 \times_{\frac{1}{6}} \mathbb{C} \\
E^1_{\text{-}\frac{2}{3}} &= P^1 \times_{\frac{2}{3}} \mathbb{C} \\
E^1_{\frac{1}{3}} &= P^1 \times_{\text{-}\frac{1}{3}} \mathbb{C} \\
E^1_{1} &= P^1 \times_{1} \mathbb{C} \\
E^1_{0} &= P^1 \times_{0} \mathbb{C} = M \times \mathbb{C}
\end{aligned}$$

From here, we construct the actual matter bundles we follows[21].

$$\begin{aligned}
Q_L &= S_+ \otimes E^3_{\mathbf{3}} \otimes E^2_{\mathbf{2}} \otimes E^1_{\frac{1}{6}} && \text{(left-handed quarks)} \\
U^C_L &= S_+ \otimes E^3_{\mathbf{3}} \otimes E^2_{\mathbf{1}} \otimes E^1_{\frac{1}{3}} && \text{(left-handed antiquarks of charge } \tfrac{1}{3}\text{)} \\
D^C_L &= S_+ \otimes E^3_{\mathbf{3}} \otimes E^2_{\mathbf{1}} \otimes E^1_{\text{-}\frac{2}{3}} && \text{(left-handed antiquarks of charge } -\tfrac{2}{3}\text{)} \\
E_L &= S_+ \otimes E^3_{\mathbf{1}} \otimes E^2_{\mathbf{2}} \otimes E^1_{\text{-}\frac{1}{2}} && \text{(left-handed leptons)} \\
P^C_L &= S_+ \otimes E^3_{\mathbf{1}} \otimes E^2_{\mathbf{1}} \otimes E^1_{1} && \text{(left-handed antileptons)} \\
H &= S_0 \otimes E^3_{\mathbf{1}} \otimes E^2_{\mathbf{2}} \otimes E^1_{1} && \text{(Higgs boson)}
\end{aligned}$$

[21] We have proceeded by defining principal bundles, their associated bundles, and then tensor products of these associated bundles. Equivalently, we could have defined associated bundles to the tensor product of the principal bundles.

Here, S_0 is the trivial spin bundle $M \times_\rho \{0\}$. For X_L (or X_L^C) ($X \in \{Q, U, D, E, P\}$) we define X_R^C (or X_R) by swapping S_+ with S_- and taking the dual of each $SU(3), SU(2), U(1)$ representation.

Note that a choice of connection on $\tilde{P}, P^1, P^2, P^3$ induces a covariant derivative and a Dirac operator on all of these bundles by taking the tensor product of the induced covariant derivatives and Dirac operators induced in each constituent associated bundle.

The Standard Model Lagrangian is a map sending the sections of (three copies of) the bundles

$$Q_L, Q_R^C, U_L^C, U_R, D_L^C, D_R, E_L, E_R^C, P_L^C, H \tag{83}$$

along with the connections on $\tilde{P}, P^1, P^2, P^3$ (which feature both explicitly as curvatures and implicitly inside Dirac operators) to a real-valued function on M.

7.2 Interpreting the Physicist's Standard Model Lagrangian

As a conclusion to this thesis, we briefly present the way a physicist typically writes the Standard Model Lagrangian and make some qualitative statements regarding what each of these objects are mathematically (classically).

$$\begin{aligned} \mathscr{L} = &-\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} B^a_{\mu\nu} B^{a\mu\nu} - -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} \\ &+ D_\mu \phi^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \\ &+ \sum_\psi i \overline{\psi} \gamma^\mu D_\mu \overline{\psi} \\ &+ \sum_{\psi_i, \psi_j} h_{ij} \overline{\psi}_i \phi \psi_j \end{aligned} \tag{84}$$

In this first row, we have the local curvatures G, B, F on P^3, P^2, P^1, respectively (these are the so-called *Yang-Mills* terms). The Einstein summation notation is expressing that they are being contracted on all indices, enabled by a metric on the principal bundle (induced by the Lorentzian spacetime metric on the base manifold) and a metric on the Lie algebra. This Lie algebra metric is the Killing form, which exploits the identification of the Lie algebra with endomorphisms on the Lie algebra through the adjoint representation $A \mapsto \mathrm{ad}_*(A) = B \mapsto [A, B]$.

In the second row, we have terms involving sections ϕ of the associated bundle H, understood as the Higgs field. On this bundle we have induced covariant derivatives D_μ (which are not Dirac operators, as the Higgs transforms trivially under Lorentz transformations). The variables μ, λ are merely experimentally determined real values.

In the third row, the terms $\overline{\psi} \gamma^\mu D_\mu \overline{\psi}$ can be understood as the application of a Hermitian inner product on the sections of associated (matter) bundles (minus the Higgs bundle)—induced by G-invariant metrics on the groups $SU(3), SU(2), U(1)$—of the form

$$\langle \psi, \gamma^\mu D_\mu \psi \rangle. \tag{85}$$

That is, the overline is notation to denote the conjugate-linear argument to the Hermitian inner product. The $\gamma^\mu D_\mu$ is the Dirac operator on the relevant associated bundle: in particular, D_μ is the induced covariant derivative and γ^μ are representations of Clifford algebra elements in a way that Eq. 81 is reproduced.

Finally, the fourth row denotes more Hermitian inner products on associated (matter) bundle sections, where h_{ij} is a (Clifford-valued) matrix and ϕ is once again the Higgs field.

References

[1] Christian Bär. *Spin Geometry*. Jan. 2011. URL: https://www.math.uni-potsdam.de/fileadmin/user_upload/Prof-Geometrie/Dokumente/Lehre/Veranstaltungen/SS11/spingeo.pdf.

[2] Edson de Faria and Welington de Melo. *Mathematical Aspects of Quantum Field Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511760532.

[3] Josè Figueroa-O'Farrill. *Gauge Theory*. Feb. 2006. URL: https://empg.maths.ed.ac.uk/Activities/GT/.

[4] B. Hall and B.C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer, 2003. ISBN: 9780387401225. URL: https://books.google.com/books?id=m1VQi8HmEwcC.

[5] M.J.D. Hamilton. *Mathematical Gauge Theory: With Applications to the Standard Model of Particle Physics*. Universitext. Springer International Publishing, 2018. ISBN: 978-3-319-68438-3. URL: https://books.google.com/books?id=DkaYswEACAAJ.

[6] N. Jacobson. *Lie Algebras*. Dover books on advanced mathematics. Dover Publications, 2013. ISBN: 9780486136790. URL: https://books.google.com/books?id=jTSoAAAAQBAJ.

[7] A.W. Knapp. *Lie Groups Beyond an Introduction*. Progress in Mathematics. Birkhäuser Boston, 2013. ISBN: 9781475724530. URL: https://books.google.com/books?id=J8EGCAAAQBAJ.

[8] H.B. Lawson and M.L. Michelsohn. *Spin Geometry (PMS-38)*. Princeton Mathematical Series v. 38. Princeton University Press, 1990. ISBN: 9780691085425. URL: https://books.google.com/books?id=pZx%5C_wAEACAAJ.

[9] K. Mackenzie. *Lie Groupoids and Lie Algebroids in Differential Geometry*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1987. DOI: 10.1017/CBO9780511661839.

[10] T. Okubo. *Differential Geometry*. Monographs and textbooks in pure and applied mathematics. M. Dekker, 1987. ISBN: 9780824777005. URL: https://books.google.com/books?id=jVbvAAAAMAAJ.

[11] N. Steenrod. *The Topology of Fibre Bundles. (PMS-14), Volume 14*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 2016. ISBN: 9781400883875. URL: https://books.google.com/books?id=JNODDAAAQBAJ.

[12] R. Ticciati et al. *Quantum Field Theory for Mathematicians*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999. ISBN: 9780521632652. URL: https://books.google.com/books?id=l-ruCWBpK8AC.

[13] L.W. Tu. *An Introduction to Manifolds*. Universitext. Springer New York, 2007. ISBN: 9780387481012. URL: https://books.google.com/books?id=v71RwPHlxx8C.

[14] L.W. Tu. *Differential Geometry: Connections, Curvature, and Characteristic Classes*. Graduate Texts in Mathematics. Springer International Publishing, 2017. ISBN: 9783319550848. URL: https://books.google.com/books?id=bmsmDwAAQBAJ.

[15] Robert M. Wald. *General Relativity*. Chicago, USA: Chicago Univ. Pr., 1984. DOI: 10.7208/chicago/9780226870373.001.0001.

[16] F.W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Graduate Texts in Mathematics. Springer, 1983. ISBN: 9780387908946. URL: https://books.google.com/books?id=iaeUqc2yQVQC.