CHAPTER 1

# Invitation: BRST "Symmetry"

If physics on a manifold with non-trivial structure is of interest to you, you may want to minimize the use of indices and Clifford algebras and think instead in terms of differential forms on fiber bundles. In this introductory section, we illustrate the use of the bundle formalism to explain the Becchi-Rouet-Stora-Tyutin transformation and its role in the quantization of a non-Abelian field theory.

We proceed very briskly on the assumption that the reader is already familiar with the structures we introduce, so we do not provide formal definitions of all syntax, rigorous existence proofs, precise definitions of smoothness, and so forth. However, most structures on a manifold are defined with a high degree of generality (at least for smooth orientable finite-dimensional manifolds) from the beginning. This is different from the common approach in which structures are first defined on $\mathbb{R}^n$ and then mapped onto a coordinate patch on a general manifold. The intention is to reveal the subtle without belaboring the obvious.

## 1.1. Fields, Tangents, and Lie Algebras

Among physicists, a smooth map $f : M \to \mathbb{R}$ is called a *(real) scalar field*. The space $\mathbb{R}^M \equiv \mathbb{R}C^\infty(M)$ of real scalar fields forms a ring with addition and multiplication operations given by pointwise addition (with the identity written $0^M$ or simply 0) and pointwise multiplication (with the identity $1^M$). This ring is commutative but not a rational domain, since a scalar field that is zero in some places but not everywhere is not invertible even though it is not equal to $0^M$. Likewise, if the range of a field is a vector space $V$, we have the space $V^M$ of $V$-*valued fields* over $M$. (Anticipating the notation of differential forms, we will also write $V\Lambda^0(M)$ for the space $V^M$ and $\Lambda^0(M)$ for $\mathbb{R}^M$.)

We will also use the term field when the range of a map $u$ is not exactly a group or vector space $F$ but a collection of spaces, one for each point in $M$, individually isomorphic to $F$. $F$ is then called the *fiber* of the range of the map $u$, which is called a *lift* or *section*. (A graph on paper of $f(x) \in F \subset \mathbb{R}$ versus $x \in X \subset \mathbb{R}$ is a lift from $X$ to $X \times F$.) Later, we will elaborate this idea and construct examples whose natural topology (the definition of continuity and smoothness on the union of the copies of $F$) is not globally equivalent to $M \times F$.

Given a space $X$ of fields whose range or fiber is a vector space $V$, the obvious pointwise addition operation and the global scaling operation $(\lambda, v) \in \mathbb{R} \times X \to \lambda v \in X : [\lambda v](x) = \lambda \cdot v(x)$ form a vector space structure on $X$. We adjoin the *pointwise scaling* operation $(f, v) \in \mathbb{R}^M \times X \to fv \in X : [fv](x) = f(x) \cdot v(x)$, which has the associative and distributive properties and therefore forms a module structure on $X$ over the commutative ring $\mathbb{R}^M$. A map between modules over $\mathbb{R}^M$ that respects the operation of pointwise scaling is called *pointwise linear*.

If $V$ is a (unitary, associative) algebra, the space of fields with range or fiber $V$ is also a (unitary, associative) algebra with its multiplication operation defined pointwise. Clearly any pointwise linear algebra structure on a space of fields must correspond to an algebra structure on the range or fiber. However, we can sometimes equip a space of fields with a *global* algebra structure using a product operation that is not pointwise linear, even if its range or fiber is not an algebra. This product will generally be written using a bracket notation $[A, B]$. We shall examine the canonical example, the Lie bracket of (tangent) vector fields, from several perspectives.

A curve $\gamma$ is a continuous map from (some open interval $I \ni 0$ of) $\mathbb{R}$ into some topological space such as an $n$-manifold $M$. Let us call $\gamma$ a curve *through* the point $x$ only if $\gamma(0) = x$; if we restrict ourselves to non-self-intersecting curves whose domain is the entirety of $\mathbb{R}$, we can define for any $\gamma_0$ a family of uniformly parameter-shifted curves $\gamma_t(s) = \gamma_0(s + t)$ through each point in the *locus* $\gamma_0(\mathbb{R})$. For any given point $x$, we may divide the category of smooth curves through $x$ into equivalence classes $[\gamma]_x$ such that the first derivative of the scalar field $f$ along any two curves in the same equivalence class is the same for all $f \in \mathbb{R}^M$. Each such equivalence class is a particular *tangent vector* at $x$, and we may shift the curve parameter as above to speak of the family of tangents to a curve at each point it passes through.

More formally, we may say that two curves $\gamma_a$ and $\gamma_b$ are *locally equivalent up to first order* ($\gamma_a \equiv_1 \gamma_b$) if, for every smooth real scalar field $f$ over $M$, the Taylor expansion of $f \circ \gamma_a - f \circ \gamma_b$ in a neighborhood of zero has no zeroth or first order term. Denoting the collection of all curves in $M$ by $\Gamma M$ and the subcollection of curves through $x$ by $\Gamma_x M$, we may form the *tangent space* $T_x M = (\Gamma_x M \backslash \equiv_1)$ to $M$ at the point $x$, consisting of the set of equivalence classes of $\Gamma_x M$ under the relation $\equiv_1$. The *tangent bundle* $TM = (\Gamma M \backslash \equiv_1)$ consists of all possible tangent vectors at all points of $M$, and may also be thought of as the union of the tangent spaces at each point. It has a natural topology, locally isomorphic to $\mathbb{R}^{2n}$, under which the tangents at each point of any smooth curve in $M$ form a continuous curve in $TM$. Each $T_x M$ is isomorphic to the fiber $\mathbb{R}^n$, and for any simply connected region $\mathcal{U} \subset M$ the natural topology on $T_{\mathcal{U}} M = \bigcup_{x \in \mathcal{U}} T_x M$ coincides with that on $\mathcal{U} \times \mathbb{R}^n$.

However, $TM$ is not isomorphic to $M \times \mathbb{R}^n$ for a general manifold $M$, and need not admit a family of isomorphisms $\phi_x : T_x M \to \mathbb{R}^n$ that form a smooth (or even continuous) map from $TM$ with its natural topology to $\mathbb{R}^n$. (If such a family $\{\phi_x\}$ exists, then one can obtain a smooth tangent vector field, nowhere zero, by taking $\phi_x^{-1}(v)$ for some fixed $v \in \mathbb{R}^n$. The famous "hedgehog theorem" denies the existence of such a field in the case of the 2-sphere.)

A *tangent vector field* $\xi \in \mathrm{vect}[M]$ is therefore not a map from $M$ to $\mathbb{R}^n$ but a smooth map $\xi : x \in M \to \xi(x) \in T_x M \subset TM$. Any smooth homomorphism $\phi : M \to N$ between manifolds induces a map $T_x\phi : T_x M \to T_{\phi(x)} N$, since equivalence classes of curves through $x$ map to equivalence classes of curves through $\phi(x)$. If $\phi$ is injective ($\phi(N)$ is a submanifold of $N$), $T_x\phi$ in turn induces a map $\phi_* : \mathrm{vect}[M] \to \mathrm{vect}[\phi(N)]$ called the *pushforward* map along $\phi$. Whether or not $\phi$ is injective, it defines a *pullback* map $\phi^* : \Lambda^0(N) \to \Lambda^0(M)$ such that $\phi^* f|_x = f|_{\phi(x)}$. The pullback operation can be applied to any map $f : N \to X$ to obtain a map $\phi^* f = f \circ \phi : M \to X$, and will also be used to pull a functional such as $\omega : \text{vect}[N] \to \Lambda^0(N)$ back to the induced functional $\phi^*\omega : \xi \in \text{vect}[M] \to \phi^*(\omega(\phi_*\xi)) \in \Lambda^0(M)$.

We can associate with a tangent vector field $\xi$ an operation $\mathcal{L}_\xi : \Lambda^0(M) \to \Lambda^0(M)$, called the *Lie derivative* with respect to $\xi$, such that if $\gamma(t)$ is any curve with $\gamma(0) = x$ which lies in the equivalence class $\xi(x)$, then $\mathcal{L}_\xi f\big|_x = \frac{d}{dt}(f \circ \gamma)\big|_0$. The theory of ordinary differential equations assures us that any smooth tangent vector field $\xi$ is *integrable*, meaning that for each $x \in M$ there is a unique *integral curve* $\gamma_x : \mathbb{R} \to M$ that obeys $\gamma_x(0) = x$ and lies in the equivalence class $\xi(\gamma_x(t))$ for all $t$. The integral curves form a smooth map $(t, x) \in \mathbb{R} \times M \to \gamma_x(t) \in M$ called the *exponential* of $\xi$, and the implicitly defined *flow map* $\xi_t(x) = \gamma_x(t)$ is a smooth automorphism for sufficiently small $t$. We can therefore pull back the scalar field $f$ along the flow map $\xi_t$ ($t$ small) and define the Lie derivative equivalently as $\mathcal{L}_\xi f = \frac{d}{dt}(\xi_t^* f)\big|_{t=0}$.

The definition of the Lie derivative using a flow map allows us to define a vector space structure on $\text{vect}[M]$, under which $\mathcal{L}_{\xi+\eta} f = \frac{d}{dt}(\xi_t^* \eta_t^* f)\big|_{t=0}$ and $\mathcal{L}_{\lambda\xi} f = \frac{d}{dt}(\xi_{\lambda t}^* f)\big|_{t=0}$ (the *fixed scaling* operation for a given $\lambda \in \mathbb{R}$). These equations state that the expressions $\xi_t \circ \eta_t$ and $\xi_{\lambda t}$ define, for sufficiently small $t$, smooth automorphisms of $M$, and therefore implicitly define curves through each $x \in M$, which identify which equivalence classes should be called $[\xi + \eta](x)$ and $[\lambda\xi](x)$ respectively. (The assumption that $\xi$, $\eta$, and $f$ are smooth is necessary for these constructions to be valid and to produce operations that are associative, distributive, and commutative.)

We can even replace our fixed $\lambda \in \mathbb{R}$ with a real scalar field $\lambda(x)$ that varies smoothly over $x \in M$ by defining $\mathcal{L}_{\lambda\cdot\xi} f = \frac{d}{dt}\left((\lambda \cdot \xi)_t^* f\right)\big|_{t=0}$. This definition uses the flow map $(\lambda \cdot \xi)_t : x \to \gamma_x^\lambda(t)$, where the curve $\gamma_x^\lambda$ is defined by the constraint $\gamma_x^\lambda(0) = x$ and the rule that, at each $x_t = \gamma_x^\lambda(t)$, $\gamma_x^\lambda$ lies in the same equivalence class as the integral curve through $x_t$ of the vector field $\lambda(x_t)\xi$ defined by the fixed scaling operation above. Since this definition implies that $\mathcal{L}_{\lambda\cdot\xi} f\big|_x = \mathcal{L}_\xi f\big|_x \cdot \lambda(x)$, the Lie derivative is now demonstrated to be pointwise linear in $\xi$. Hence the space $\text{vect}[M]$ is a module over $\mathbb{R}^M$, and therefore the individual tangent spaces $T_xM$ are vector spaces. (This is the reverse of the usual construction, in which the additive group structure on $T_xM$ is exhibited via a local isomorphism between a region of $\mathbb{R}^n$ and a region $\mathcal{U} \subset M$ containing $x$, followed by a pointwise definition of scaling and thence a module structure on $\text{vect}[M]$.)

Given that $\text{vect}[M]$ is a module over $\mathbb{R}^M$, the space $\Lambda^1(M)$ of pointwise linear maps from $\text{vect}[M]$ to $\Lambda^0(M)$ is also a module over $\mathbb{R}^M$. The Lie derivative $\mathcal{L}_\xi f$ is pointwise linear in $\xi$, so there must be some element $df \in \Lambda^1(M)$ such that $\mathcal{L}_\xi f = df(\xi)$. We call $df$ the *gradient* of $f$ and general elements of $\Lambda^1(M)$ *gradient vector fields* or *1-forms* over $M$. When we return to the topic of differential forms, we will speak of $d : \Lambda^0(M) \to \Lambda^1(M)$ as the *exterior derivative operator* and extend its definition to a wider domain. For the present, we point out that the Lie derivative $\mathcal{L}_\xi f$, and therefore the exterior derivative $df$, is linear but not pointwise linear in $f$.

The Lie derivative obeys the defining property of a *derivation*, a linear endomorphism $D$ on an algebra $A$ such that $D(ab) = (Da)b + a(Db)$. (This property is called the *Leibniz rule.*) The *Lie bracket* $[D_1, D_2] = D_1 D_2 - D_2 D_1$ of two derivations is also a derivation:

$$[D_1, D_2](ab) = ([D_1, D_2] a) b + a [D_1, D_2] b \overbrace{+ D_2 a D_1 b + D_1 a D_2 b - D_1 a D_2 b - D_2 a D_1 b}^{0}$$

It turns out that every derivation on $\Lambda^0(M)$ is the Lie derivative with respect to some tangent vector field and therefore that $[\mathcal{L}_\xi, \mathcal{L}_\eta] = \mathcal{L}_{[\xi,\eta]}$ for some tangent vector field $[\xi, \eta]$. We postpone the proof of this statement until we can prove a more general version in the context of a broader algebra of derivations.

Given the canonical representation $L_g : h \to gh$ of a Lie group $G$ on itself we have the induced map $T_h L_g$ from the tangent space $T_h G$ to the tangent space $T_{gh} G$. A *(left) invariant* tangent vector field $a$ on $G$ has the property that $a(gh) = T_h L_g a(h)$ for all $g \in G$. This implies that $a(g) = T_e L_g a(e)$, and in fact this formula generates a one-to-one correspondence between elements of $T_e G$ and left invariant tangent vector fields on $G$. This space is called the *Lie algebra* $\mathfrak{g}$ of $G$, since it has the structure of an additive group and also the multiplication operation given by the Lie bracket $[a, b]$ of tangent vector fields on $G$. The reader may verify that the Lie bracket of two invariant tangent vector fields is also left invariant. Any representation $\rho_g$ of a Lie group on manifold $M$ induces a representation $\widetilde{\rho}_a$ of the Lie algebra on $M$, i. e., a map from $a \in \mathfrak{g}$ to a left invariant tangent vector field on $M$; we shall exhibit this construction while discussing fiber bundles below.

For the present, we cite the maps induced by $\mathrm{aut}_g : h \to ghg^{-1}$. Since $\mathrm{aut}_g e = e$ for all $g \in G$, the map $T_e \mathrm{aut}_g : T_e G \to T_e G$ is well defined and induces the *adjoint representation* $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ of $G$ on $\mathfrak{g}$. Given any fixed $g \in G$, $\mathrm{Ad}_g$ is a Lie algebra automorphism of $\mathfrak{g}$. (If $G$ is a matrix group, the matrix product $gag^{-1}$ is well defined for $a \in \mathfrak{g}$ and coincides with $\mathrm{Ad}_g a$; we used this above as the defining property of $\mathrm{Ad}_g$ on $\mathbb{R}(n)$.) The corresponding Lie algebra representation $\mathrm{ad}_a = \widetilde{\mathrm{Ad}}_a : \mathfrak{g} \to \mathfrak{g}$ is also called adjoint representation, and has the property $\mathrm{ad}_a b = [a, b]$. (This is easily confirmed for a matrix group $G$ by Taylor expanding $\mathrm{Ad}_g a$ around $g = e$.)

## 1.2. Projections and Fiber Bundles

A linear endomorphism $P$ of a vector space $V$ which has the property $P \circ P = P$ is called a *linear projection* to distinguish it from other sorts of map which we shall call projections (generally smooth surjective maps from a manifold onto one of lower dimension). The notion of a *fibered space* $(B, M, \pi)$ is given by the latter sort of projection, the *bundle projection* $\pi$ from a *bundle manifold* $B$ to a *base manifold* $M$, with the additional requirement of a *standard fiber* $F$ diffeomorphic to each preimage $\pi^{-1}(x)$ (the *fiber over* $x$). The direct product $\mathcal{U} \times F$ of a region $\mathcal{U}$ with the standard fiber $F$ has the natural bundle projection $\pi : (x, y) \to x$. We shall speak of a map $f : A \to B$ between fibered spaces for which $\pi_A(p_1) = \pi_A(p_2)$ implies $\pi_B(f(p_1)) = \pi_B(f(p_2))$ as respecting the fiber structure. Unless otherwise stated, any map between fibered spaces said to be a homomorphism (isomorphism, automorphism, etc.) must respect the fiber structure.

We require an additional structure on $(B, M, \pi)$ for it to be called a *fiber bundle.* On any open cover $\{\mathcal{U}_r\}$ of $M$, the preimages $\pi^{-1}(\mathcal{U}_r) \subset B$ must each possess a diffeomorphism $f_r : \mathcal{U}_r \times F \to \pi^{-1}(\mathcal{U}_r)$ such that $\pi f_r(x, y) = x$ and $f_s^{-1} \circ f_r$ is a smooth automorphism on $(\mathcal{U}_r \cap \mathcal{U}_s) \times F$. This structure of "local diffeomorphisms" (true diffeomorphisms of a fragment of the fiber bundle) is called a *local trivialization* and induces a set of *transition functions* $g_{sr}(x) : y_r \in F \to y_s \in F$ defined by $f_s(x, y_s) = f_r(x, y_r)$. If we have a Lie group $G$ that is represented effectively on $F$ and each $g_{sr}(x)$ coincides with an element of $G$, then we refer to $G$ as the *structure group* and the whole structure as a *$G$-bundle*.

One way of looking at the relationship between linear projections and bundle projections is to consider the linear projection $P : V \to V$ as a special case of a bundle projection for bundle space $V$. Clearly the requirements of a fibered space are satisfied, with the standard fiber given by the subspace $\ker P = P^{-1}(0)$ annihilated by $P$ and the base space given by range$[P]$ (alternatively, by the quotient space of $V$ by the equivalence relation $Px = Py$). The local trivialization is given on an open cover of the base space consisting of the single region $\mathcal{U}_0 \cong \text{range}[P]$, with the diffeomorphism $f_0 : (x, y) \in \mathcal{U}_0 \to (x + y) \in V$ given by the vector space structure on $V$. In the absence of a vector space structure on a general bundle space, an atlas $\{\mathcal{U}_r\}$ on $M$ may require numerous regions to cover $M$ and the standard fiber may not even be a vector space (e. g., the frame bundle discussed below).

For instance, we shall speak of the *tangent bundle* $TM$ to an $n$-dimensional manifold $M$ as the union of tangent spaces $T_xM$ at each point $x \in M$, with the standard fiber $\mathbb{R}^n$ and the bundle projection $\pi : TM \to M$ defined by $\pi : v \in T_xM \to x$. The usual construction of *holonomic bases* $\left\{ (e_r)_\mu(x) = \frac{\partial}{\partial x_r^\mu}\Big|_x \right\}$ on coordinate patches gives a local trivialization with structure group $GL_n$, such that $f_r$ takes $(x \in \mathcal{U}_r, \xi_r^\mu(x) \in \mathbb{R}^n)$ to the point $\xi_r^\mu(x)\,(e_r)_\mu(x) \in \pi^{-1}(x)$ and the transition function $g_{sr}(x)$ takes $\xi_r^\mu(x)$ to $\xi_s^\mu(x) = \frac{\partial x_s^\mu}{\partial x_r^\nu} \xi_r^\nu(x)$.

The basis vectors $\frac{\partial}{\partial x_r^\mu}\Big|_x$ on which we have decomposed $T_xM$ are equivalence classes of "basis curves" $[\gamma_{r\mu}]_x(t)$ through $x$ such that $\frac{d}{dt}\left(x_r^\nu \circ [\gamma_{r\mu}]_x\right) = \delta_\mu^\nu$, i. e., 1 if $\mu = \nu$ and 0 otherwise. The Lie derivative with respect to $\xi$ is then represented by the component sum of the derivatives along a set of basis curves. We cannot really use the more elegant pullback construction here, since the basis vectors are only defined on the region $\mathcal{U}_r$. The reader may verify the formula $[\xi, \eta]^\mu = \xi^\nu \frac{\partial}{\partial x^\nu} \eta^\mu - \eta^\nu \frac{\partial}{\partial x^\nu} \xi^\mu$ in a coordinate basis on a simply connected region of $M$ and confirm that it is a proper tangent vector field (i. e., the real-valued functions $\xi_r^\mu(x)$ transform appropriately under a change of basis and under a local diffeomorphism). The tensor calculus approach to differential geometry uses this sort of patchwise definition of objects using the transition functions on a local trivialization.

We can extend the local trivialization to general bases $\left\{ e_i(x) = e_i^\mu(x) \frac{\partial}{\partial x^\mu}\big|_x \right\}$ and general linear transition functions $\xi_s^i(x) = [g_{sr}(x)]^i_j\, \xi_r^j(x)$. We can then define local trivializations limited to more closely related local diffeomorphisms (e. g., mutually related by a smaller structure group such as $O_n^+$, as discussed below). Such a "coordinate system" on a patch of the bundle does not generally contain holonomic bases.

Given a fiber bundle $(B, M, \pi)$ and a trivial region $\mathcal{U} \subset M$, we define a *local section* $u : \mathcal{U} \to \pi^{-1}(\mathcal{U})$ as a differentiable map obeying $\pi(u(x)) = x$, i. e., $\pi \circ u = \text{Id}_{\mathcal{U}}$. If the standard fiber has a vector space structure which the transition functions respect, then we may use a *smooth partition of unity* (a set of scalar fields $\{\mathcal{I}_r\}$, each with support on a single region $\mathcal{U}_r$, that sum to 1) to construct a *global section* $u : M \to B$ from a set of local sections on $\{\mathcal{U}_r\}$. We shall sometimes refer to such a fiber bundle as a *vector bundle* and its global sections as *vector fields*; we shall reserve the term *tangent vector field* for an element of vect $[M]$, i. e., a global section of the tangent bundle $TM$. Note that local sections exist on any fiber bundle, but a bundle whose standard fiber is not a vector space need not admit global sections.

One important type of bundle that is not a vector bundle is a *principal $G$-bundle*, a $G$-bundle $(B, M, \pi)$ whose standard fiber $F$ is isomorphic to its structure group $G$ and whose transition functions coincide with the canonical representation (the left action $L_g : f \to gf$) of $G$ on itself. This means that, given a point $x \in M$ and two diffeomorphisms $f_\alpha$ and $f_\beta$ from the local trivialization $\{f_\alpha\}$ over $\{\mathcal{U}_\alpha\}$, the $y$-coordinates of any point $f_\beta(x, y_\beta) = f_\alpha(x, y_\alpha) \in \pi^{-1}(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ in the $\alpha$ and $\beta$ systems are related by $y_\beta = g_{\beta\alpha}(x)(y_\alpha) = L_g(y_\alpha) = gy_\alpha$ for some $g \in G$ independent of the particular $y_\alpha$. (We have switched to labeling the principal bundle's coordinate patches with $\alpha, \beta, \ldots$ to reduce confusion in the discussion of associated bundles below.)

Since left translation commutes with right translation $R_g : f \to fg$, the right action of $G$ on $\pi^{-1}(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ induced by $f_\alpha$ coincides with that induced by $f_\beta$. Therefore a local trivialization induces a consistent, smooth, free right action $\widetilde{R}_g$ of $G$ on the principal bundle. This gives us the *fundamental representation* $\rho_g(p) = \widetilde{R}_{g^{-1}}(p) = pg^{-1}$ of $G$ on $B$, a left action that is well defined independent of the choice of coordinate system within the local trivialization.

For instance, the tangent bundle is not a principal bundle, but the collection of all possible bases $\left\{e_i(x) = e_i^\mu(x) \frac{\partial}{\partial x^\mu}\big|_x\right\}$ for the tangent spaces $T_xM$ can be given a principal bundle structure, which we will call the *frame bundle* $FM$. The transition functions $[g_{\beta\alpha}(x)]^\mu_\nu = \frac{\partial x^\mu_\beta}{\partial x^\nu_\alpha}$ act on the basis components according to

$$(e_\beta)^\mu_i(x) = [g_{\beta\alpha}(x)]^\mu_\nu (e_\alpha)^\nu_i(x)$$

which is the canonical left action of $GL_n$ on itself. The left action $L_g$ of $GL_n$ on a the basis components $\{e_i^\mu(x)\}$ is called a *passive transformation*, since it changes not the basis $\{e_i(x)\}$ but the coordinate system in which its components are expressed. The fundamental representation $\rho_g$ represents an *active transformation* of the frame bundle, in which each point $p = f_\alpha(x, (e_\alpha)^\mu_i(x))$ of the frame bundle is carried to a new point $pg^{-1} = f_\alpha\left(x, (e'_\alpha)^\mu_i(x) = (e_\alpha)^\mu_j(x)\left[g^{-1}\right]^j_i\right)$.

Note that the frame bundle on a general manifold $M$ usually does not admit global sections, but no global section is needed for $\rho_g$ to be smoothly defined. (Consider the example of the 2-sphere: as discussed above, its frame bundle does not admit a global section, but the operation of rotating the frame 90° clockwise is well defined throughout the frame bundle.) The tangent bundle is then said to be an *associated bundle* to the frame bundle, since its transition functions $\xi^i_s(x) = [g_{sr}(x)]^i_j\, \xi^j_r(x)$ correspond on the fiber over $x$ to the active transformations $e'_i(x) = \rho_{g_{sr}(x)}\{e_i(x)\} = e_j(x)\left[g^{-1}_{sr}(x)\right]^j_i$ of the fiber of $FM$ over $x$.

In this last expression we have dropped the label $\alpha$ and index $\mu$ associated with the coordinate system on a patch of the frame bundle, since we are focusing on the correspondence between an *active* change in the choice of local section $\{e_i(x)\}$ on a portion of the frame bundle and a *passive* change of frame (coordinate system) on the corresponding portion of the associated tangent bundle. When we label $\{(e_r)_i(x)\}$ with the Latin label $r$, we are calling attention to the choice of local section on $FM$ and hence of coordinate system on associated bundles; we will try not to confuse this with the occasional use of Greek labels when a particular expression involves an explicit coordinate system on $FM$.

Note also that two local trivializations $\{f_r\}$ and $\{f_{r'}\}$ on the same fibered space may have the same structure group $G$ but not be $G$-related (i. e., $g_{rr'}(x)$ does not coincide with the representation of any $g \in G$), in which case they define distinct $G$-bundles and, if $G \approx F$, distinct fundamental representations of $G$ on the bundle space. For instance, if $M$ is an orientable manifold of dimension $n$ and we choose a Riemannian metric $\eta_x(v, v')$ on each tangent space $T_xM$ (smoothly dependent on $x$), we may define a reduced frame bundle $O^+FM$ (the collection of all possible bases $\{e_i(x)\}$ for $T_xM$ which are orthonormal with respect to $\eta_x$). Its structure group is the special orthogonal group $O_n^+$. Each choice of Riemannian metric gives a distinct reduced frame bundle, although we can demonstrate that any two reduced frame bundles over $M$ are isomorphic.

## 1.3. Alternating Forms and Differential Forms

We next introduce the notion of an *alternating $k$-form* $\Phi \in V\Omega^k(W)$, a smooth, skew-symmetric multilinear map from $W^k$ (the direct product of $k$ copies of a vector space $W$) to a vector space $V$. The nonnegative integer $k$ is called the *rank* of $\Phi$. Clearly $V\Omega^0(W) \equiv V$ and $V\Omega^1(W) \equiv VC^\infty(W)$, the space of smooth functions from $W$ to $V$. We will write $w \llcorner \Phi$ for the $(k-1)$-form defined by $(w \llcorner \Phi)(w_1, \ldots, w_{k-1}) = \Phi(w, w_1, \ldots, w_{k-1})$, with the convention that $w \llcorner f = 0$ for any 0-form $f$. Then the properties of skew symmetry and multilinearity are simply stated: $w \llcorner z \llcorner \Phi + z \llcorner w \llcorner \Phi = 0$ and $(\lambda w) \llcorner \Phi = \lambda (w \llcorner \Phi)$. *Differential forms* are alternating forms whose domain is $\mathrm{vect}[M]$, whose range is a module over $\mathbb{R}^M$ such as $V\Lambda^0(M)$ or $\mathrm{vect}[M]$, and whose action is *pointwise* multilinear, i. e., $[(f\xi) \lrcorner \Phi]_x = f(x)[\xi \lrcorner \Phi]_x$. (Note that we use the symbol $\llcorner$ when $W$ is a general vector space, reserving the conventional notation $\xi \lrcorner$ (the *inner derivative* with respect to $\xi$) for use with differential forms.)

We have already encountered the notations $\Lambda^0(M)\ \left(\equiv \mathbb{R}^M\Omega^0(\mathrm{vect}[M])\right)$ for the space of (real) *0-forms* (smooth scalar fields) on a manifold $M$ and $\Lambda^1(M)$ for the space of (real-valued) *1-forms* (smooth gradient vector fields) on $M$. In general, a $V$-valued 1-form $\omega \in V\Lambda^1(M)\ \left(\subset V^M\Omega^1(\mathrm{vect}[M])\right)$ is a smooth pointwise linear map from the space $\mathrm{vect}[M]$ of tangent vector fields over $M$ to the space $V\Lambda^0(M)$ of 0-forms taking values in the vector space $V$. We will use the notation $\xi \lrcorner \omega$ for the image of a tangent vector field $\xi$ under the map $\omega$. Since $\omega$ is by construction pointwise linear, it has a value $\omega(x) \in V\Lambda_x^1 : T_xM \to V$ at each $x \in M$, a linear map given by $\omega(x)(\xi(x)) = \xi \lrcorner \omega|_x$ for all $\xi \in \mathrm{vect}[M]$. Given a basis $\{e_i(x)\}$ for $T_xM$, we have a corresponding *dual basis* $\left\{e^i(x)\right\}$ for $V\Lambda_x^1$, defined so that $e^i(x)(e_j(x)) = \delta_j^i \equiv [\mathrm{Id}(\mathbb{R}^n)]_j^i$. If $\omega(x) = \omega_i(x) e^i(x)$ and $\xi(x) = \xi^i(x) e_i(x)$, then $\omega(x)(\xi(x)) = \xi^i(x)\omega_i(x)$. (Here we are using the Einstein summation convention over repeated indices, which we shall do consistently unless otherwise stated.)

Since the expression $\xi \lrcorner \omega|_x = \xi^i(x)\omega_i(x)$ must hold for any basis on $T_xM$, we see that the components $\omega_i(x)$ amount to coordinates in a local trivialization of the global section $\omega$ of the ($V$-valued) *gradient (cotangent) bundle* $VT^{*1}M$, an associated bundle to the frame bundle. Its transition functions must be given by $(\omega_s)_i(x) = (\omega_r)_j(x)\left[g_{sr}^{-1}(x)\right]_i^j$ for $\xi \lrcorner \omega$ to be a well-defined $V$-valued 0-form. It is more usual to start with a dual coordinate basis on each cotangent space $V\Lambda^1_x$, define components smoothly on coordinate patches, construct transition functions, etc., but we shall generally prefer to define fields in terms of maps between other spaces of fields and then express their coordinates in a local trivialization as we have done.

Note carefully that a basis $\{e_i(x)\}$ for $T_xM$ is a point of the fiber over $x$ in $FM$. The fiber $\pi^{-1}_{FM}(x) \approx GL_n$ is not a vector space and the expression $\xi(x) = \xi^i(x)\, e_i(x)$ is meaningful in $TM$, not $FM$. The components $\xi^i(x) = e^i(x)(\xi(x))$ are simply real numbers relating $\xi(x) \in T_xM$ to a particular set of basis vectors. Given a local section $\{(e_r)_i(x \in \mathcal{U}_r)\}$, the components $\xi^i_r(x \in \mathcal{U}_r)$ are smooth functions on the domain of the local section. But since the frame bundle over a general manifold does not possess global sections, these functions cannot be extended to true scalar fields over $M$. When we evaluate a true scalar field such as $\xi \lrcorner \omega$ at a point $x$, we may write $\xi\lrcorner\omega|_x = \omega(x)(\xi(x)) = \xi^i(x)\,\omega_i(x)$, which comes out the same in any basis. But we prefer to write equations so that both sides are globally defined objects and to extend the summation convention to summation over a set $\{(e_r)_i(x)\}$ of local sections of $FM$ using a partition of unity $\{\mathcal{I}_r\}$ as necessary.

As another example, consider the space $T\Lambda^1(M)\,\big(\subset \mathrm{vect}[M]\,\Omega^1(\mathrm{vect}[M])\big)$ of *tangent-valued 1-forms*, smooth pointwise linear maps from $\mathrm{vect}[M]$ to $\mathrm{vect}[M]$. An element $K \in T\Lambda^1(M)$ has components $(K_r)^i_j(x)$ , i. e., the fiber coordinates in the system $r$ of a point $K(x) = f_r\Big(x, (K_r)^i_j(x)\Big)$ on the associated bundle $TT^{*1}(M)$. Its defining relation has the schematic coordinate expression $\xi\lrcorner K|_x = e_i(x)\, K^i_j(x)\,\xi^j(x)$, which really stands for

$$\xi\lrcorner K|_x = \sum_r \mathcal{I}_r(x)\,(e_r)_i(x)\,(K_r)^i_j(x)\,(\xi_r)^j(x)$$

The appropriate transition functions on $TT^{*1}(M)$ are those for which this expression results in a well-defined tangent vector field independent of the open cover $\{\mathcal{U}_r\}$ and partition of unity $\{\mathcal{I}_r\}$ used to calculate it. We must therefore have $(K_s)^i_j(x) = [g_{sr}(x)]^i_k\,(K_r(x))^k_l\,\big[g^{-1}_{sr}(x)\big]^l_j = \big[\mathrm{aut}_{g_{sr}(x)}K_r(x)\big]^i_j$. The identity map on $\mathrm{vect}[M]$ corresponds to the *Maurer-Cartan form* $\zeta \in T\Lambda^1(M)$, the 1-form for which $\xi\lrcorner\zeta = \xi$; its components are $(\zeta_r)^i_j(x) = \delta^i_j$ in any system.

We have previously defined $\mathcal{L}_\xi : \Lambda^0(M) \to \Lambda^0(M)$ point-by-point using the definition of a tangent vector $\xi(x)$ as an equivalence class of curves through $x$. When we took the first derivative along a curve $\gamma(t) \in \xi(x)$, we implicitly identified the tangent space to $\Lambda^0(M)$ at $f$ with $\Lambda^0(M)$ itself. Writing $\frac{\delta}{\delta(x_0)}$ for the basis vector on $T_f\big[\Lambda^0(M)\big]$ along which $f(x_0)$ increases at unit pace while $f(x \neq x_0)$ does not change, the tangent vector is written

$$\oint_M dx_0\,[\mathcal{L}_\xi f](x_0)\,\frac{\delta}{\delta(x_0)}$$

Likewise, we have previously defined the Lie bracket of two tangent vector fields in terms of the bracket of the associated derivations (Lie derivative operators) on $\Lambda^0(M)$ and asserted its closure on $\mathrm{vect}[M]$, again identifying $T_f\big[\Lambda^0(M)\big]$ with $\Lambda^0(M)$. This allowed us to develop the notion of (left) invariant vector fields associated with a group representation on a manifold, which we needed for our definition of Lie algebras.

Now that we have defined Lie groups and Lie algebras, we could instead exhibit a Lie group structure on the space $\mathrm{Diff}\,[M]$ of diffeomorphisms from $M$ to $M$ and define $\mathrm{vect}\,[M]$ as its Lie algebra, represented on $\Lambda^0(M)$ by $\rho_\xi(f) = \mathcal{L}_\xi f \in T_f\left[\Lambda^0(M)\right]$. We could then turn around and define the Lie bracket in terms of the adjoint representation of $\mathrm{vect}\,[M]$ on itself, i. e., $[\xi, \eta] = \mathrm{ad}_\xi \eta$. We would continue by verifying that $\mathcal{L}_\xi$ is a derivation (obeys $\mathcal{L}_\xi(ab) = (\mathcal{L}_\xi a)\, b + a\,(\mathcal{L}_\xi b)$) for each $\xi$, that $\mathcal{L}_\xi f$ is linear in $f$ and pointwise linear in $\xi$, and that $[\xi, \eta]$ is anti-symmetric. The equivalence of these two approaches is fairly obvious but rather laborious to prove, so we simply assert the deep results and move on.

Since $\mathcal{L}_\xi f$ is pointwise linear in $\xi$, there must be for each 0-form $f \in \Lambda^0(M)$ a 1-form $df \in \Lambda^1(M)$ such that $\mathcal{L}_\xi f = \xi \lrcorner df$. The *exterior derivative* operator $d : \Lambda^0(M) \to \Lambda^1(M)$ is linear but not pointwise linear; in fact, it has the defining property of a derivation, the Leibniz rule $d(a \cdot b) = (da) \cdot b + a \cdot (db)$. Given that we have defined $\xi \lrcorner f = 0$ for any 0-form $f$, the *inner derivative* operator $\xi \lrcorner$ also obeys the Leibniz rule $\xi \lrcorner (a \cdot b) = (\xi \lrcorner a) \cdot b + a \cdot (\xi \lrcorner b)$ when $a$ is a 0-form. If we could extend $\xi \lrcorner$, $\mathcal{L}_\xi$, and $d$ from their original domains to the entire direct sum $\Lambda^0(M) \oplus \Lambda^1(M)$ of these operators' domains and ranges, then these operators would all be derivations on $\Lambda^0(M) \oplus \Lambda^1(M)$. (The direct sum $A \oplus B$ of two vector spaces consists of linear combinations of an element of $A$ and an element of $B$, identifying their zero elements; it differs subtly from the direct product $A \times B$ in that $A \oplus \{0\}$ and $\{0\} \oplus A$ are considered not merely isomorphic to $A$ but identical to $A$.)

In fact, to obtain a domain on which $\xi \lrcorner$, $\mathcal{L}_\xi$, and $d$ can be extended to endomorphisms, we will need to go further than $\Lambda^0(M) \oplus \Lambda^1(M)$. We will need the entire hierarchy of spaces $\Lambda^k(M) \equiv \mathbb{R}^M \Omega^k(\mathrm{vect}\,[M])$ of (alternating) differential $k$-forms. (Note that $\Lambda^k(M) \equiv \{0\}$ if $k > n$, since there are only $n$ independent tangent vectors in $T_x M$ for an $n$-dimensional manifold $M$.) We will define an associative algebra structure on $\Lambda^*(M) \equiv \bigoplus_k \Lambda^k(M)$ via a suitable definition of the *wedge product* $\wedge : \Lambda^k(M) \times \Lambda^l(M) \to \Lambda^{k+l}(M)$. (This definition must be suitable in two respects: it must be closed on $\Lambda^*(M)$, and a modified version of the Leibniz rule must hold for $\xi \lrcorner$ distributed over the wedge product.) We can then extend the definitions of $\mathcal{L}_\xi$ and $d$ to differential forms of higher rank in such a way that this modified Leibniz rule holds also for $\mathcal{L}_\xi$ and $d$ distributed over the wedge product.

We therefore define the *algebra of graded derivations* on $\Lambda^*(M)$ as follows. The *rank* of an operator $D : \Lambda^l(M) \to \Lambda^{k+l}(M)$ is $\mathrm{rank}\,(Da) - \mathrm{rank}\,(a) = k$. The *graded Lie bracket* of two operators of well-defined rank is $[D_1, D_2] = D_1 D_2 - (-1)^{\mathrm{rank}(D_1)\cdot\mathrm{rank}(D_2)} D_2 D_1$. It obeys the *graded Jacobi identity*:

$$[D_1, [D_2, D_3]] - [[D_1, D_2], D_3] = (-1)^{\mathrm{rank}(D_1)\cdot\mathrm{rank}(D_2)} [D_2, [D_1, D_3]]$$

The crucial property of the wedge product is that $a\wedge$ is an operator of the same rank as the (real) $k$-form $a$. We therefore define the *graded Leibniz rule* $[D, a\wedge]\, b = D(a \wedge b) - (-1)^{\mathrm{rank}(D)\cdot\mathrm{rank}(a)} a \wedge (Db) = (Da) \wedge b$ and extend it using the vector space structure of $\Lambda^*(M)$ to forms and operators of mixed rank. A *graded derivation* is any linear endomorphism of $\Lambda^*(M)$ that obeys the graded Leibniz rule.

The space of graded derivations is given a (non-associative) algebra structure using the obvious vector space operations (addition, additive inverse, and global scaling) and the graded Lie bracket as the product operation. The resulting operator algebra is a *graded Lie algebra* structure on the space of graded derivations on

$\Lambda^*(M)$. We assert that the algebra of graded derivations on $\Lambda^*(M)$ includes $\xi\lrcorner$ (rank $-1$) and $d$ (rank 1). We will see that $\mathcal{L}_\xi$ (rank 0) lies in the subalgebra of this algebra generated by $\xi\lrcorner$ and $d$.

We already have a definition of $\xi\lrcorner \equiv \xi\llcorner$ given by the general alternating form structure on $\Lambda^k(M) \equiv \mathbb{R}^M\Omega^k(\mathrm{vect}[M])$. The extension of $\mathbb{R}^M$-module structure to $\Lambda^k(M)$ is naturally given inductively by $[\xi\lrcorner, f\cdot]\,a \equiv 0$, making $f\cdot$ a graded derivation of rank 0. We shall define the wedge product inductively via $f \wedge a \equiv f \cdot a$ (given $f \in \Lambda^0(M)$ and $a \in \Lambda^k(M)$) and $[\xi\lrcorner, a\wedge]\,b \equiv (\xi\lrcorner a) \wedge b$. The explicit distributive property for the inner derivative $\xi\lrcorner$ over the wedge product is therefore $\xi\lrcorner(a \wedge b) = (\xi\lrcorner a) \wedge b + (-1)^{\mathrm{rank}(a)}\,a \wedge (\xi\lrcorner b)$. The relation $[\xi\lrcorner, \eta\lrcorner]\,a = \xi\lrcorner(\eta\lrcorner a) + \eta\lrcorner(\xi\lrcorner a) = 0$ holds on the wedge product of two forms, implying that the wedge product is closed on $\Lambda^*(M)$.

We can then extend the definitions of $\mathcal{L}_\xi$ and $d$ to forms of higher rank in such a way that a version of the Leibniz rule holds for $\xi\lrcorner$, $\mathcal{L}_\xi$, and $d$ distributed over the wedge product. Using the definition $\xi\lrcorner f = 0$ for $f \in \Lambda^0(M)$, we can restate $\mathcal{L}_\xi f = \xi\lrcorner df$ (the defining relation for $df$) in the form $[\xi\lrcorner, d]\,a = \xi\lrcorner(da) + d(\xi\lrcorner a) \equiv \mathcal{L}_\xi a$. Allowing $a \in \Lambda^k(M)$, we use this relation to extend the definition of $d$ from $\Lambda^{k-1}(M)$ to $\Lambda^k(M)$ (given a definition of $\mathcal{L}_\xi$ on $\Lambda^k(M)$), starting with $k = 0$, $\Lambda^{-1}(M) \equiv \{0\}$, and $d0 = 0$.

We extend the domain of $\mathcal{L}_\xi$ to $d\left(\Lambda^k(M)\right) \subset \Lambda^{k+1}(M)$ by defining $[\mathcal{L}_\xi, d]\,a = \mathcal{L}_\xi(da) - d(\mathcal{L}_\xi a) \equiv 0$. This immediately implies $[d,d]\,a = d(da) + d(da) \equiv 0$ (consider the graded Jacobi identity for $[\xi\lrcorner, [d,d]]$). The distributive property $[\mathcal{L}_\xi, a\wedge]\,b = (\mathcal{L}_\xi a)\wedge b$ expands to $\mathcal{L}_\xi(a \wedge b) = (\mathcal{L}_\xi a)\wedge b + a \wedge(\mathcal{L}_\xi b)$ and further extends the domain of $\mathcal{L}_\xi$ to all of $\Lambda^{k+1}(M)$. (This assumes that the vector space $\Lambda^{k+1}(M)$ is spanned pointwise by the *exact* $(k+1)$-forms $d\left(\Lambda^k(M)\right) \equiv \left\{d\Phi : \Phi \in \Lambda^k(M)\right\}$, which is true at least for a finite-dimensional manifold $M$.) Hence we can use $[\xi\lrcorner, d]\,a \equiv \mathcal{L}_\xi a$ again to extend $d$ to $\Lambda^{k+1}(M)$, and so forth.

Given the definition $[\mathcal{L}_\xi, d]\,f = 0$, we find that $[\xi,\eta]\lrcorner df = \mathcal{L}_{[\xi,\eta]}f = [\mathcal{L}_\xi, \mathcal{L}_\eta]\,f = [\mathcal{L}_\xi, \eta\lrcorner]\,df$. Clearly this identifies $[\mathcal{L}_\xi, \eta\lrcorner]$ on $d\left(\Lambda^0(M)\right)$ with the derivation $[\xi,\eta]\lrcorner$, and so its extension to $\Lambda^*(M)$ must be $[\mathcal{L}_\xi, \eta\lrcorner]\,a = \mathcal{L}_\xi(\eta\lrcorner a) - \eta\lrcorner(\mathcal{L}_\xi a) = [\xi,\eta]\lrcorner a$. We could have used this formula to extend $\mathcal{L}_\xi$ from $\Lambda^k(M)$ to $\Lambda^{k+1}(M)$ without explicitly using formulas containing $d$. This formula implies the extension of the relation $[\mathcal{L}_\xi, \mathcal{L}_\eta]\,a = \mathcal{L}_{[\xi,\eta]}a$ to all $a \in \Lambda^*(M)$ (consider the graded Jacobi identity for $[d, [\mathcal{L}_\xi, \eta\lrcorner]]$). Note also that the construction of $d$ as an element of the algebra of graded derivations on $\Lambda^*(M)$ allows us immediately to assert that $[d, a\wedge]\,b = d(a \wedge b) - (-1)^{\mathrm{rank}(a)}\,a \wedge (db) = (da) \wedge b$ (graded Leibniz rule).

We have already defined the spaces $T\Lambda^0(M) \equiv \mathrm{vect}[M]$ of tangent vector fields and $T\Lambda^1(M)$ of tangent-valued 1-forms; the extension to general tangent-valued $k$-forms is obvious. We cannot yet extend the operator algebra to include "exterior" operations like $\xi\wedge : \Lambda^k(M) \to T\Lambda^k(M)$, since we would have to extend the field algebra $\Lambda^*(M)$ and the definitions of the existing operators to include $T\Lambda^k(M)$. This is not trivial, as the only reasonable definition of $\mathcal{L}_\xi : T\Lambda^0(M) \to T\Lambda^0(M)$ is given by $(\mathcal{L}_\xi\eta)\lrcorner = [\mathcal{L}_\xi, \eta\lrcorner] = [\xi,\eta]\lrcorner$, which is not pointwise linear in $\eta$ and hence does not lead to a suitable extension of the exterior derivative $d$ to $T\Lambda^0(M)$. (The construction of a pointwise linear derivation $\nabla_\xi : \eta \in \mathrm{vect}[M] \to \nabla_\xi\eta \in \mathrm{vect}[M]$ and a "covariant" exterior derivative operator $D : T\Lambda^0(M) \to T\Lambda^1(M)$ will have to await the connexion on $M$ defined below.)

However, we can extend the parameter $\xi$ of the inner derivative operation to a tangent-valued form of higher rank. Let us continue to reserve the symbol $\lrcorner$ for the simple case of a tangent vector field, defining an alternate notation $j_{\_} : (K, a) \in T\Lambda^k (M) \times \Lambda^l (M) \to j_K a \in \Lambda^{k+l-1} (M)$ for the general inner derivative. We will define it inductively, with the base cases $j_\xi a \equiv \xi \lrcorner a$ ($\xi \in T\Lambda^0 (M)$) and $j_K f \equiv 0$ ($f \in \Lambda^0 (M)$) and the induction rule $[\xi\lrcorner, j_K]\, a \equiv j_{\xi\lrcorner K} a$. (This is of course the graded Lie bracket, and the rank of $j_K$ is $k-1$ when $K \in T\Lambda^k (M)$.) This induction rule is an extension of $[\xi\lrcorner, j_\eta]\, a = [\xi\lrcorner, \eta\lrcorner]\, a = 0$, since of course $\xi\lrcorner\eta = 0$. The natural definition of the Lie derivative with respect to a tangent-valued form is then $\mathcal{L}_K a \equiv [j_K, d]\, a$, implying $[\mathcal{L}_K, d]\, a = 0$.

Let us illustrate the use of these inductive rules using the important case of the Maurer-Cartan form $\zeta \in T\Lambda^1 (M)$. The rank of $j_\zeta$ is $1 - 1 = 0$ and hence $\xi\lrcorner\, j_\zeta a - j_\zeta (\xi\lrcorner a) = [\xi\lrcorner, j_\zeta]\, a = j_\xi a = \xi\lrcorner a$. Given that $\xi\lrcorner a = 0$ when $a$ is a 0-form, $j_\zeta a$ could be any 0-form, if it were not for the second base case above. When $a$ is a 1-form, we have $\xi\lrcorner\, j_\zeta a = \xi\lrcorner a$; this would imply that $j_\zeta a = a$ up to a 0-form, if it were not for the requirement that $j_K a$ be of uniform rank when $K$ and $a$ are. If we hypothesize that $j_\zeta a = la$ for $l$-form $a$, then for $(l+1)$-form $a$ we have $\xi\lrcorner\, j_\zeta a = j_\zeta (\xi\lrcorner a) + \xi\lrcorner a = (l+1)(\xi\lrcorner a)$; this is the inductive case we need in order to prove that $j_\zeta a = \operatorname{rank}(a)\, a$ for all $a \in T\Lambda (M)$. This leads immediately to $\mathcal{L}_\zeta a \equiv [j_\zeta, d]\, a = da$.

More generally, given any smooth, pointwise linear transformation $\lambda \in T\Lambda^1 (M)$ of $\operatorname{vect}[M]$, on a 1-form $a$ we obtain $\xi\lrcorner\, j_\lambda a = \lambda(\xi)\lrcorner a$. The extension of this expression to $a \in T\Lambda^2 (M)$ is given by $\eta\lrcorner\,\xi\lrcorner\, j_\lambda a = \eta\lrcorner\, j_\lambda (\xi\lrcorner\, a) + \eta\lrcorner\,\lambda(\xi)\lrcorner\, a = \lambda(\eta)\lrcorner\,\xi\lrcorner\, a + \eta\lrcorner\,\lambda(\xi)\lrcorner\, a$, and it may be further extended to $a \in T\Lambda^l (M)$ to obtain:

$$\xi_l\lrcorner \dots \xi_1\lrcorner\, j_\lambda a = \sum_{j=1}^{l} \xi_l\lrcorner \dots \lambda(\xi_j)\lrcorner \dots \xi_1\lrcorner\, a$$

In the case $\lambda = \zeta$, the above formula reduces to $j_\zeta a = la$ as expected. Denoting the omission of an index in a series of inner derivatives by $\hat{\xi}_j\lrcorner$, we may write out the formula for Lie derivation with respect to $\lambda$:

$$\begin{aligned}
\xi_l \lrcorner \ldots \xi_0 \lrcorner \mathcal{L}_\lambda a &= \xi_l \lrcorner \ldots \xi_0 \lrcorner j_\lambda (da) - \xi_l \lrcorner \ldots \xi_0 \lrcorner d (j_\lambda a) \\
&= \sum_{j=0}^{l} \left[ \begin{array}{l} \xi_l \lrcorner \ldots \lambda (\xi_j) \lrcorner \ldots \xi_0 \lrcorner da \\ -(-1)^j \xi_l \lrcorner \ldots \xi_{j+1} \lrcorner \mathcal{L}_{\xi_j} (\xi_{j-1} \lrcorner \ldots \xi_0 \lrcorner j_\lambda a) \end{array} \right] \\
&= \sum_{j=0}^{l} (-1)^j \left[ \begin{array}{l} \xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner \lambda (\xi_j) \lrcorner da \\ -\mathcal{L}_{\xi_j} \left( \xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner j_\lambda a \right) \\ +\sum_{i>j} \xi_l \lrcorner \ldots \hat{\xi}_i \lrcorner [\xi_j, \xi_i] \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner j_\lambda a \end{array} \right] \\
&= \sum_{j=0}^{l} (-1)^j \left[ \begin{array}{l} \xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner \mathcal{L}_{\lambda(\xi_j)} a \\ -\xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner d (\lambda (\xi_j) \lrcorner a) \\ -\mathcal{L}_{\xi_j} \left( \sum_{k \neq j} \xi_l \lrcorner \ldots \hat{\xi}_k \lrcorner \lambda (\xi_k) \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner a \right) \\ + \sum_{i>j,\, k \neq i,j} \xi_l \lrcorner \ldots \hat{\xi}_k \lrcorner \lambda (\xi_k) \lrcorner \ldots \hat{\xi}_i \lrcorner [\xi_j, \xi_i] \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner a \\ +\sum_{i>j} \xi_l \lrcorner \ldots \hat{\xi}_i \lrcorner \lambda ([\xi_j, \xi_i]) \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner a \end{array} \right]
\end{aligned}$$

The three middle terms in the above expression cancel after rearranging some indices (exercise for the reader) and we obtain:

$$\xi_l \lrcorner \ldots \xi_0 \lrcorner \mathcal{L}_\lambda a = \sum_{j=0}^{l} (-1)^j \left[ \begin{array}{l} \xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \xi_0 \lrcorner \mathcal{L}_{\lambda(\xi_j)} a \\ -\sum_{i=0}^{j-1} (-1)^i \xi_l \lrcorner \ldots \hat{\xi}_j \lrcorner \ldots \hat{\xi}_i \lrcorner \ldots \xi_0 \lrcorner \lambda ([\xi_i, \xi_j]) \lrcorner a \end{array} \right]$$

This is the generalization of a similar expansion for $da = \mathcal{L}_\zeta a$, in which $\zeta$ is the identity on vect $[M]$. Note that, although the Lie derivative $\mathcal{L}_\lambda a$ is not pointwise linear in $\lambda$, it is linear under global scaling and over the sum of tangent-valued 1-forms. So if we have two projections $\lambda_0$ and $\lambda_1$ on vect $[M]$ that sum to the identity, then we can write $d = \mathcal{L}_{\lambda_0} + \mathcal{L}_{\lambda_1}$. We shall see this expression again when discussing the BRST operator.

We have built up our operator algebra from a set of recursive definitions based on a few primitive operations: scaling by a real scalar field $f$, the Lie derivative $\mathcal{L}_\xi$ broken into a $\xi$-independent operator $d$ and a pointwise linear operator $\xi \lrcorner$, and the Lie bracket $[\xi, \eta]$ of two tangent vector fields defined by $\mathcal{L}_{[\xi,\eta]} = [\mathcal{L}_\xi, \mathcal{L}_\eta]$. The relation $[\mathcal{L}_\xi, \eta \lrcorner] = [\xi, \eta] \lrcorner$ followed from a set of definitions consistent with the relation $[\xi \lrcorner, \eta \lrcorner] = 0$ implied by the alternating property of the field algebra. We did not need to define any new primitive operations on fields in order to complete (to the current extent) our algebra of graded derivations. (Even the wedge product was defined inductively starting from $f \wedge a = f \cdot a$.)

Similarly, the algebra of differential forms can be built up from a single base case (the multiplicative identity 1, which is annihilated by every element of the operator algebra) and just two operations, pointwise scaling $(f \cdot)$ and the exterior derivative $(d)$. We will now use this fact to classify all derivations on $\Lambda^* (M)$, incidentally providing the existence proof for$[\xi, \eta]$ that we postponed earlier. We will find that it makes sense to extend the domain of the Lie bracket to $T\Lambda^* (M)$ by defining $\mathcal{L}_{[K,L]} = [\mathcal{L}_K, \mathcal{L}_L]$, exhibiting a formula for a unique tangent-valued form $[K, L]$

that satisfies this relation (sometimes called the Frölicher-Nijenhuis bracket). This formula is complicated by the fact that generally $[j_K, j_L] = j_{[K,L]^\wedge}$ for a non-zero $[K, L]^\wedge \in T\Lambda^{k+l-1}(M)$, the *algebraic bracket* of $K \in T\Lambda^k(M)$ and $L \in T\Lambda^l(M)$.

(Insert proof, along the lines of Cap et al. 1994, that every derivation on $\Lambda^*(M)$ is of the form $j_K + \mathcal{L}_L$. Demonstrate that all algebraic (i. e., pointwise linear) derivations are of the form $j_K$. Point out that $[j_\zeta, D] = \operatorname{rank}(D)\, D$.)

We will work through these steps sequentially again when we come back to the topic of tangent vector field valued forms on fiber bundles, using a broader space of forms and obtaining a wider subalgebra of the algebra of graded derivations. In that context, we will prove that the space $\Lambda^{k+1}(M)$ of $(k+1)$-forms is spanned pointwise both by $\left\{\omega \wedge \Phi : \omega \in \Lambda^1(M), \Phi \in \Lambda^k(M)\right\}$ and by the *exact* $(k+1)$-forms $d\left(\Lambda^k(M)\right) \equiv \left\{d\Phi :\ \Phi \in \Lambda^k(M)\right\}$. For now, we will simply collate the results for real-valued (differential) forms in $\Lambda^*(M)$ and hint at the extension to tangent vector field valued (alternating) forms.

We find that $\xi\llcorner$, $\mathcal{L}_\xi$, and $d$ can be extended to the space $j(\mathrm{vect}[M])$ of inner derivative operations and thence to a portion of the space $j(\mathrm{vect}[M])\,\Omega^*(\mathrm{vect}[M])$, including the space of Lie derivative operations $\mathcal{L}(\mathrm{vect}[M]) \subset j(\mathrm{vect}[M])\,\Omega^1(\mathrm{vect}[M])$. But as we have indicated by using the symbols $\llcorner$ and $\Omega$, these tangent vector field valued forms are not generally pointwise linear, and their algebra product is not associative.

Here is a summary of all our inductive formulas on differential forms:

| | |
|---|---|
| $f \wedge b \equiv f \cdot b \ \left(f \in \Lambda^0(M),\, b \in \Lambda^k(M)\right)$ | (base case for an inductive definition of the wedge product) |
| $[\xi\lrcorner, a\wedge]\, b \equiv (\xi\lrcorner a) \wedge b$ | (defines the wedge product; note that $\operatorname{rank}[a\wedge] = \operatorname{rank}[a]$) |
| $\xi\lrcorner (a \wedge b) = (\xi\lrcorner a) \wedge b + (-1)^{\operatorname{rank}(a)}\, a \wedge (\xi\lrcorner b)$ | (explicit distribution of $\xi\lrcorner$ over the wedge product) |
| $[\xi\lrcorner, \eta\lrcorner]\, a = \xi\lrcorner (\eta\lrcorner a) + \eta\lrcorner (\xi\lrcorner a) = 0$ | (follows from the above, implying $\wedge$ is closed on $\Lambda^*(M)$) |
| $[\xi\lrcorner, d]\, a \equiv \xi\lrcorner (da) + d(\xi\lrcorner a) \equiv \mathcal{L}_\xi a$ | (extends $d$ from $\Lambda^{k-1}$ to $\Lambda^k$ given $\mathcal{L}_\xi$ on $\Lambda^k$) |
| $[\mathcal{L}_\xi, d]\, a \equiv \mathcal{L}_\xi (da) - d(\mathcal{L}_\xi a) \equiv 0$ | (extends $\mathcal{L}_\xi$ from $\Lambda^k$ to $d\left(\Lambda^k\right)$ given $d$ on $\Lambda^k$) |
| $[d, d]\, a = d(da) + d(da) = 0$ | (implied by graded Jacobi identity for $[\xi\lrcorner, [d, d]]$) |
| $[\mathcal{L}_\xi, a\wedge]\, b \equiv (\mathcal{L}_\xi a) \wedge b$ | (extends $\mathcal{L}_\xi$ from $d\left(\Lambda^k\right)$ to $\Lambda^{k+1}$) |
| $\mathcal{L}_\xi (a \wedge b) = (\mathcal{L}_\xi a) \wedge b + a \wedge (\mathcal{L}_\xi b)$ | (explicit distribution of $\mathcal{L}_\xi$ over the wedge product) |
| $[\mathcal{L}_\xi, \eta\lrcorner]\, a = \mathcal{L}_\xi (\eta\lrcorner a) - \eta\lrcorner (\mathcal{L}_\xi a) = [\xi, \eta]\lrcorner a$ | (equivalent extension of $\mathcal{L}_\xi$ from $\Lambda^k$ to $\Lambda^{k+1}$) |
| $[\mathcal{L}_\xi, \mathcal{L}_\eta]\, a = \mathcal{L}_\xi (\mathcal{L}_\eta a) - \mathcal{L}_\eta (\mathcal{L}_\xi a) = \mathcal{L}_{[\xi,\eta]} a$ | (axiomatic on $\Lambda^0$, provable on higher ranks) |
| $[d, a\wedge]\, b = (da) \wedge b$ | (follows from the formulas for $\mathcal{L}_\xi (a \wedge b)$, etc.) |
| $d(a \wedge b) = (da) \wedge b + (-1)^{\operatorname{rank}(a)}\, a \wedge (db)$ | (explicit distribution of $d$ over the wedge product) |
| $j_\xi a \equiv \xi\lrcorner a \ \left(\xi \in T\Lambda^0(M)\right)$ | (base case for an inductive definition of $j_K$) |
| $[\xi\lrcorner, j_K]\, a \equiv j_{\xi\lrcorner K} a$ | (extends $j_K$ from $\xi\lrcorner K \in T\Lambda^k$ to $K \in T\Lambda^{k+1}$) |
| $\mathcal{L}_K a \equiv [j_K, d]\, a$ | (defines $\mathcal{L}_K$ for $K \in T\Lambda^k$) |
| $[\mathcal{L}_K, d]\, a = 0$ | (follows from the definitions of $\mathcal{L}_K$ and $j_K$) |
| $\mathcal{L}_\xi (\eta\lrcorner a) \equiv (\mathcal{L}_\xi [\eta\lrcorner])\, a + \eta\lrcorner (\mathcal{L}_\xi a)$ | (extends $\mathcal{L}_\xi$ to $j(\operatorname{vect}[M])$ such that $\mathcal{L}_\xi [j(\eta)] \equiv j([\xi, \eta])$) |
| $\xi\lrcorner (\eta\lrcorner a) \equiv \overbrace{(\xi\llcorner[\eta\lrcorner])}^{0}\, a - \eta\lrcorner (\xi\lrcorner a)$ | (trivial extension of $\xi\llcorner$ to $j(\operatorname{vect}[M])$) |
| $\mathcal{L}_\xi (\mathcal{L}_\eta a) \equiv (\mathcal{L}_\xi [\mathcal{L}_\eta])\, a + \mathcal{L}_\eta (\mathcal{L}_\xi a)$ | (extends $\mathcal{L}_\xi$ to $\mathcal{L}(\operatorname{vect}[M])$ such that $\mathcal{L}_\xi [\mathcal{L}_\eta] \equiv \mathcal{L}_{[\xi,\eta]}$) |
| $\xi\lrcorner (\mathcal{L}_\eta a) \equiv (\xi\llcorner[\mathcal{L}_\eta])\, a + \mathcal{L}_\eta (\xi\lrcorner a)$ | (extends $\xi\llcorner$ to $\mathcal{L}(\operatorname{vect}[M])$ such that $\xi\llcorner[\mathcal{L}_\eta] \equiv j([\xi, \eta])$) |
| $d(\xi\lrcorner a) \equiv (d[\xi\llcorner])\, a - \xi\lrcorner (da)$ | (extends $d$ to $j(\operatorname{vect}[M])$ such that $d[j(\xi)] \equiv \mathcal{L}_\xi$) |
| $[\xi\llcorner, d](\eta\lrcorner) = \xi\llcorner[\mathcal{L}_\eta] = [\xi, \eta]\lrcorner = \mathcal{L}_\xi [\eta\lrcorner]$ | (not pointwise linear, i. e., $d(j(\operatorname{vect}[M])) \not\subseteq j(\operatorname{vect}[M])\,\Lambda^1$) |
| $d(\mathcal{L}_\xi a) \equiv \overbrace{(d[\mathcal{L}_\xi])}^{0}\, a + \mathcal{L}_\xi (da)$ | (trivial extension of $d$ to $\mathcal{L}(\operatorname{vect}[M]) = d(j(\operatorname{vect}[M]))$) |
| $d(f\mathcal{L}_\xi a) = df \wedge \mathcal{L}_\xi a + f\mathcal{L}_\xi (da)$ | (illustration of $d:\ j(\operatorname{vect}[M])\,\Omega^1 \to j(\operatorname{vect}[M])\,\Omega^2$) |

## 1.4. $\widetilde{R}_g$ Invariance, the Invariant Algebra, and the Connexion

We return to the general principal $G$-bundle $(B, M, \pi, G)$ to discuss operations on the bundle which commute with the fundamental representation $\rho_g(p) = \widetilde{R}_{g^{-1}}(p) = pg^{-1}$ of $G$ on $B$. We have, for a given $p \in B$, the map $\operatorname{bit}_p : g \in G \to \rho_{g^{-1}}(p) = pg \in B$ and its range $\operatorname{orbit}(p) = \operatorname{bit}_p(G)$; clearly any two $\rho_g$-related points share the same orbit, which coincides with the fiber $\pi^{-1}(\pi(p))$. We may associate with the map $\operatorname{bit}_p$ from manifold $G$ to manifold $B$ a map $T_e\operatorname{bit}_p$ from the tangent space $T_eG$ at $e \in G$ to the tangent space $T_pB$. (This is similar to the construction that is used to define the Lie algebra of Lie group $G$ by extending each tangent vector at the identity to a tangent vector field.) The corresponding representation of the Lie algebra $\mathfrak{g}$ of $G$ on $B$ is given by $\widetilde{\rho}_a(p) = T_e \mathrm{bit}_p a|_e$.

The tangent space at $p$ to the fiber $\pi^{-1}(\pi(p))$ is called the *vertical subspace* $V_p \subset T_p B$ of the tangent space at $p$ to the bundle space $B$. The representation $\widetilde{\rho}_a$ induces the map $\lambda_p : a \in \mathfrak{g} \to \lambda_p a|_p = \widetilde{\rho}_a(p) \in V_p$, which is a vector space isomorphism between $\mathfrak{g}$ and $V_p$; we shall denote its inverse $\zeta_p$. We have written $\lambda_p a|_p$ because we wish to extend $\lambda_p a$ to a $\rho_{g^{-1}}$-invariant (i. e., $\widetilde{R}_g$-invariant) tangent vector field $\lambda_p a|_{pg} \in V_{pg}$ on $\mathrm{orbit}(p)$. Let $h_t = \exp[t\lambda_p a] : \mathrm{orbit}(p) \to \mathrm{orbit}(p)$ be the flow of $\lambda_p a$, i. e., the family of curves with $h_0(p) = p$ whose tangent at each point $h_t(p)$ is given by $\lambda_p a|_{h_t(p)}$. Since $\mathrm{bit}_p$ is surjective on $\mathrm{orbit}(p)$, we have $h_t(p) = \rho_{\Gamma_t(p)}(p) = p\Gamma_t^{-1}(p)$ for some $\Gamma_t(p) \in G$; similarly, $\lambda_p a|_{pg} = \widetilde{\rho}_{\gamma(pg)}(pg)$ for some $\gamma(pg) = \zeta_{pg}\left(\lambda_p a|_{pg}\right) \in \mathfrak{g}$.

The condition of $\widetilde{R}_g$-invariance on the flow $h_t$ may then be written $pg\Gamma_t^{-1}(pg) = p\Gamma_t^{-1}(p)\,g$, which implies $\Gamma_t(pg) = g^{-1}\Gamma_t(p)\,g = \mathrm{aut}_{g^{-1}}\Gamma_t(p)$. Therefore the condition of $\widetilde{R}_g$-invariance on the tangent vector field $\lambda_p a|_{pg}$ is $\gamma(pg) = \mathrm{Ad}_{g^{-1}}\gamma(p)$. Let us restate this using the family of maps $\zeta_p : V_p \to \mathfrak{g}$ and the pullback notation $\widetilde{R}_g^*\zeta\Big|_p = \zeta_{pg} : V_{pg} \to \mathfrak{g}$. Let $\mathrm{vert}[B] \subset \mathrm{vect}[B]$ be the space of tangent vector fields on $B$ which are everywhere vertical and let $V\mathfrak{E} \subset \mathrm{vert}[B]$ be the space of $\widetilde{R}_g$-invariant tangent vector fields $v$, which must coincide with some $\lambda_p a$ on $\mathrm{orbit}(p)$. We therefore have:

$$\widetilde{R}_g^*\zeta\Big|_p(v) = \mathrm{Ad}_{g^{-1}}\zeta_p(v) \Rightarrow \widetilde{R}_g^*\zeta = \mathrm{Ad}_{g^{-1}}\zeta$$

We interpret the last version of this formula as follows. We combine the maps $\zeta_p$ defined by $\zeta_p(\widetilde{\rho}_a(p)) = a$ for all $p \in B$, $a \in \mathfrak{g}$ to obtain a pointwise linear map $\zeta : \mathrm{vert}[B] \to \mathfrak{g}\Lambda^0[B]$ such that $\zeta(v)|_p = \zeta_p(v(p))$. Since $\widetilde{R}_g$ is a smooth automorphism of $B$, we can define the pullback $\left[\widetilde{R}_g^*\zeta\right](v)$, a $\mathfrak{g}$-valued field on $B$ whose value at $p$ depends only on the value of $v$ at $pg$. If we restrict the domain of $\zeta$ to $V\mathfrak{E}$, the value of $v$ at $pg$ is determined by its value at $p$, so both the restricted map $\zeta : V\mathfrak{E} \to \mathfrak{g}\Lambda^0[B]$ and its pullback $\widetilde{R}_g^*\zeta$ are pointwise linear. We have demonstrated that a $\widetilde{R}_g$-invariant vertical tangent vector field $v \in V\mathfrak{E}$ which coincides with $\widetilde{\rho}_a$ at $p$ must coincide with $\widetilde{\rho}\left(\mathrm{Ad}_{g^{-1}}a\right)$ at $pg$. So we must have $\widetilde{R}_g^*\zeta = \mathrm{Ad}_{g^{-1}}\zeta$.

Since the derivation above holds fiber-by-fiber, we may replace the parameter $g \in G$ with an element $\Gamma$ of the *gauge group* $G^M$, whose right action on $B$ is given by $\widetilde{R}_\Gamma : p \to p\Gamma(\pi(p))$. The map $\zeta$ is an important example of a *gauge covariant* object, one whose values on the range of a local section $u_r : \mathcal{U}_r \subset M \to \pi^{-1}(\mathcal{U}_r)$ can be related to those on a second local section $\widetilde{R}_\Gamma \circ u_r$ by a formula involving $\Gamma$. Similarly, any $\widetilde{R}_g$-invariant vertical tangent vector field $v \in V\mathfrak{E}$ is also $\widetilde{R}_\Gamma$-invariant and may be characterized on the range of a local section $u_r : x \in \mathcal{U}_r \to u_r(x) \in \pi^{-1}(x)$ by the $\mathfrak{g}$-valued function $[v \llcorner \zeta_r](x) \equiv \zeta[v](u_r(x))$.

Let us work with a particular local trivialization of $B$ over the open cover $\{\mathcal{U}_r\}$ of $M$, given by $f_r : (x, g) \in \mathcal{U}_r \times G \to \widetilde{R}_g(u_r(x)) \in \pi^{-1}(x)$. Then we may view $\{\zeta_r\}$ as a set of maps taking $v \in V\mathfrak{E}$ to a set of $\mathfrak{g}$-valued functions on $\{\mathcal{U}_r\}$, related on each region of overlap $\mathcal{U}_r \cap \mathcal{U}_s$ (on which $u_s(x) = u_r(x)\,\Gamma_{rs}(x)$ for some $G$-valued $\Gamma_{rs}(x)$) by $\zeta_s(x) = \mathrm{Ad}_{\Gamma_{rs}^{-1}}\zeta_r(x)$. Hence we can speak of a set of $\mathfrak{g}$-valued functions

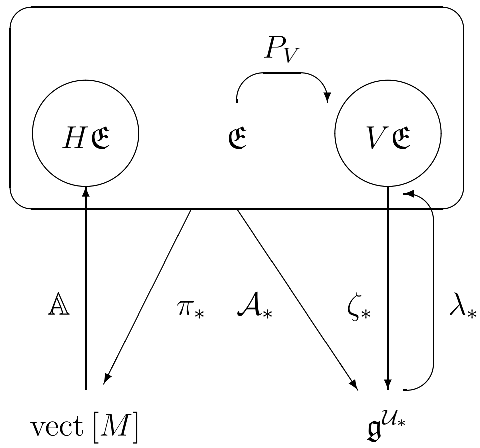


FIGURE 1.4.1. Maps to and from the Lie algebra $\mathfrak{E}$.

$\{a_r\}$ which satisfy $a_s(x) = \mathrm{Ad}_{\Gamma_{rs}^{-1}} a_r(x)$ as a "member $a_*$ of the space $\mathfrak{g}^{\mathcal{U}_*}$" and write $\zeta_* : v \in V\mathfrak{E} \to v \llcorner \zeta_* = \{v \llcorner \zeta_r\} \in \mathfrak{g}^{\mathcal{U}_*}$. Likewise, $a_* \in \mathfrak{g}^{\mathcal{U}_*}$ defines a unique $\lambda_* a_* \in V\mathfrak{E}$ whose value at any $p \in \pi^{-1}(x)$ is given by $\lambda_{u_r(x)} a_r(x)\big|_p \in V_p$.

We will be working extensively with the *invariant algebra* $\mathfrak{E} \subset \mathrm{vect}\,[B]$ of all smooth $\widetilde{R}_g$-invariant tangent vector fields, of which $V\mathfrak{E}$ is an ideal (i. e., $[v, \varepsilon] \in V\mathfrak{E}$ for all $v \in V\mathfrak{E}$ and $\varepsilon \in \mathfrak{E}$). A non-vertical field $\varepsilon \in \mathfrak{E}$ is not $\widetilde{R}_\Gamma$-invariant, since the flow $\exp[\varepsilon]$ of $\varepsilon$ carries some points $p$ to points $\exp[\varepsilon](p)$ outside $\mathrm{orbit}\,(p)$, where it is possible that $\Gamma(\pi(\exp[\varepsilon](p))) \neq \Gamma(\pi(p))$. We may use the map $T_p\pi : T_pB \to T_{\pi(p)}M$ to define a pushforward map $\pi_* : \mathfrak{E} \to \mathrm{vect}\,[M]$, since the values of $\varepsilon \in \mathfrak{E}$ on each $\mathrm{orbit}\,(p)$ have the same image $[\pi_*\varepsilon](\pi(p)) = T_p\pi(\varepsilon(p)) = T_{pg}\pi(\varepsilon(pg))$. Clearly $\ker \pi_* = V\mathfrak{E}$, i. e., the subspace of $\mathfrak{E}$ that $\pi_*$ maps to $0 \in \mathrm{vect}\,[M]$ consists of the smooth vertical $\widetilde{R}_g$-invariant tangent vector fields.

To complete the description of $\varepsilon \in \mathfrak{E}$ in terms of the given local trivialization over $\{\mathcal{U}_r\}$, we need to choose a pointwise linear map $\mathbb{A} : \mathrm{vect}\,[M] \to \mathfrak{E}$ from among the right inverses of $\pi_*$. We designate by $H\mathfrak{E}$ the image of $\mathrm{vect}\,[M]$ under a particular choice of $\mathbb{A}$. Then every $\varepsilon \in \mathfrak{E}$ decomposes uniquely into the sum of some $\xi \lrcorner \mathbb{A} \in H\mathfrak{E}$ and some $\lambda_* a_* \in V\mathfrak{E}$. Alternately, we can choose a smoothly varying $\widetilde{R}_g$-invariant *horizontal subspace* $H_p$ complementary to $V_p$ in each $T_pB$, defined by a linear projection $P_V : T_pB \to V_p$ that annihilates $H_p$. (This is equivalent, since for each $\xi \in \mathrm{vect}\,[M]$, the induced projection $P_V : \mathfrak{E} \to V\mathfrak{E}$ annihilates exactly one element $\xi \lrcorner \mathbb{A}$ of $\pi_*^{-1}(\xi) \subset \mathfrak{E}$.) The linear map $\mathcal{A}_p \equiv \zeta_p \circ P_V$ is then an extension of $\zeta_p$ from $V_p$ to $T_pB$ that coincides with $\zeta_p$ on $V_p$ and annihilates the complementary subspace $H_p$; given a local trivialization as above, we have the corresponding definition $\mathcal{A}_* \equiv \zeta_* \circ P_V : \mathfrak{E} \to \mathfrak{g}^{\mathcal{U}_*}$.

We have the following relationships among maps to and from $\mathfrak{E}$:

$$\begin{aligned}
&\pi_* \circ \mathbb{A} = \mathrm{Id}\,[\mathrm{vect}\,[M]] && \mathbb{A} \circ \pi_* = \mathrm{Id}\,[\mathfrak{E}] - P_V \\
&\zeta_* \circ \lambda_* = \mathrm{Id}\left[\mathfrak{g}^{\mathcal{U}_*}\right] && \lambda_* \circ \zeta_* = \mathrm{Id}\,[V\mathfrak{E}] \\
&\pi_* \circ \lambda_* = 0 && P_V \circ P_V = P_V \\
&P_V \circ \mathbb{A} = 0 && \mathcal{A}_* \equiv \zeta_* \circ P_V \\
&\mathcal{A}_* \circ \mathbb{A} = 0 && \lambda_* \circ \mathcal{A}_* = P_V
\end{aligned}$$

This set of constructs is called a *connexion* on $B$ and is uniquely defined by any of the objects $\mathbb{A}$ (the *Ehresmann connexion* or *horizontal lift*), $\{H_p\}$ (the *horizontal subspaces* of $T_pB$), $P_V$ (the *vertical projection* on $T_pB$ and $\mathfrak{E}$), or $\mathcal{A}_*$(the *1-form of the connexion* or *Faddeev-Popov ghost*). We will discuss the identification of $\mathcal{A}_*$ with the physicist's "ghost field" later; for now, observe that $\mathfrak{E}$ is an abstract algebra on which we can construct spaces $\Lambda^k(\mathfrak{E})$ of alternating $k$-forms, and $\mathcal{A}_*$ is a 1-form on $\mathfrak{E}$ which takes values in the space $\mathfrak{g}^{\mathcal{U}_*}$ of $\mathrm{Ad}_{\Gamma_{rs}^{-1}}$-related $\mathfrak{g}$-valued fields on $\{\mathcal{U}_r\}$.

## 1.5. Identifying the BRST Operator

(Insert context from Göckeler and Schücker chapter 12, which is largely consistent with that of Schücker 1987.)

In physics it is common to postulate fields of unknown origin which "carry a representation" of a Lie algebra $\mathfrak{g}$, or more properly of $\mathcal{E} = \mathfrak{g}^{\mathcal{U}_*}$. The (left) action of $E \in \mathcal{E}$ on a polynomial in these fields and their derivatives is characterized by a linear approximation $W(E)$ called the Ward operator. We rewrite (12.56), expressing the action of the BRST coboundary operator $s$ on the $l$-form $Q$ in terms of the Ward operator $W(E)$, in our notation for alternating forms:

$$E_l\lrcorner \ldots E_0\lrcorner\, sQ = \sum_{j=0}^{l}(-1)^j \left[\begin{array}{l} E_l\lrcorner \ldots \hat{E}_j\lrcorner \ldots E_0\lrcorner\, W(E_j)\, Q \\ +\sum_{i=0}^{j-1}(-1)^i E_l\lrcorner \ldots \hat{E}_j\lrcorner \ldots \hat{E}_i\lrcorner \ldots E_0\lrcorner\, [E_i,\, E_j]\lrcorner\, Q \end{array}\right]$$

The Lie derivative with respect to a vector field has the axiomatic relationship $[\mathcal{L}_\xi,\, \eta\lrcorner]\, a = \mathcal{L}_\xi(\eta\lrcorner\, a) - \eta\lrcorner(\mathcal{L}_\xi a) = [\xi, \eta]\lrcorner\, a$ to the Lie bracket and inner derivative. (This is axiomatic in the sense that it may be taken as the recursive definition of the Lie derivative on forms of rank 1 and higher.) Using this and the property $[\xi\lrcorner,\, \eta\lrcorner]\, a = \xi\lrcorner(\eta\lrcorner\, a) + \eta\lrcorner(\xi\lrcorner\, a) = 0$ of the inner derivative, we may write:

$$\begin{aligned}
E_l\lrcorner \ldots E_0\lrcorner\, sQ &= \sum_{j=0}^{l}(-1)^j \left[\begin{array}{l} E_l\lrcorner \ldots \hat{E}_j\lrcorner \ldots E_0\lrcorner\, W(E_j)\, Q \\ -\sum_{i=0}^{j-1} E_l\lrcorner \ldots \hat{E}_j\lrcorner \ldots \hat{E}_i\lrcorner \left[\mathcal{L}_{E_j},\, E_i\lrcorner\right] \ldots E_0\lrcorner\, Q \end{array}\right] \\
&= \sum_{j=0}^{l}(-1)^j E_l\lrcorner \ldots E_{j+1}\lrcorner \left[\begin{array}{l} E_{j-1}\lrcorner \ldots E_0\lrcorner\, W(E_j)\, Q \\ -\mathcal{L}_{E_j}(E_{j-1}\lrcorner \ldots E_0\lrcorner\, Q) \\ +E_{j-1}\lrcorner \ldots E_0\lrcorner\, \mathcal{L}_{E_j} Q \end{array}\right]
\end{aligned}$$

The exterior derivative has the axiomatic relationship $[\xi\lrcorner,\, d]\, a = \xi\lrcorner(da) + d(\xi\lrcorner\, a) = \mathcal{L}_\xi a$, again axiomatic in the sense that it may be taken as the recursive definition of the exterior derivative on forms of rank 1 and higher. This gives us a new formula for one of the terms in the above expression:

$$\sum_{j=0}^{l} (-1)^j E_l \lrcorner \ldots E_{j+1} \lrcorner \mathcal{L}_{E_j} (E_{j-1} \lrcorner \ldots E_0 \lrcorner Q) = (-1)^l d (E_l \lrcorner \ldots E_0 \lrcorner Q) + E_l \lrcorner \ldots E_0 \lrcorner dQ$$

Since $Q$ is an $l$-form, the result of applying $l+1$ inner derivatives to $Q$ is zero. Hence our formula for the BRST operator may be rewritten:

$$E_l \lrcorner \ldots E_0 \lrcorner [s+d] Q = \sum_{j=0}^{l} (-1)^j E_l \lrcorner \ldots \hat{E}_j \lrcorner \ldots E_0 \lrcorner \left[W(E_j) + \mathcal{L}_{E_j}\right] Q$$

This equation is trivially satisfied if $\mathcal{L}_{E_j} = -W(E_j)$ and $d = -s$, i. e., if we identify (up to a sign) the Lie derivative on $\Lambda(\mathcal{E}, Pl_n)$ with the Ward operator and the exterior derivative with the BRST operator. In this picture, the Faddeev-Popov ghost field is identified with the Maurer-Cartan form, the unique element $z \in \Lambda^1(\mathcal{E}, Pl_0)$such that $E \lrcorner z = E$. When we take $\mathcal{E}$ to be the Lie algebra of infinitesimal gauge transformations on a fiber bundle $P$ over manifold $M$, we recover the BRST transformation familiar from non-abelian QFT. Our goal is to extend this construction to the full Lie algebra $\mathfrak{E}$ of right-invariant vector fields on a principal (gauge) bundle.

We follow Schücker in postulating a fixed auxiliary connexion $\dot{\mathcal{A}}$ represented on a local section over the region $\mathcal{U} \subset M$ by the $\mathfrak{g}$-valued 1-form $\dot{A}$ with field strength $\dot{F} = \left(\partial_\mu \dot{A}_\nu + \frac{1}{2}\left[\dot{A}_\mu, \dot{A}_\nu\right]\right) dx^\mu \wedge dx^\nu$. We also follow him to the extent of defining the "covariant Lie derivative" $\dot{\mathcal{L}}_{(\Omega,v)}$ on $\Lambda^0(\mathfrak{E}, Pl_{\mathcal{U}})$ to be $\dot{\mathcal{L}}_{(\Omega,v)} Q = -W(\Omega) Q + v \lrcorner \dot{D} Q$ and postulating this form for the Lie bracket:

$$[(\Omega', v'), (\Omega, v)] = \left([\Omega', \Omega] + v' \lrcorner \dot{D}\Omega - v \lrcorner \dot{D}\Omega' - v' \lrcorner v \lrcorner \dot{F}, [v', v]\right)$$

Although it is perhaps not obvious from these expressions, they do not depend essentially on the auxiliary connexion, which is only used to lift $v$ to the local section over $\mathcal{U}$ and hence to associate a right-invariant vector field on $\pi^{-1}(\mathcal{U})$ with a (vertical, horizontal) pair $(\Omega, v)$ on $\mathcal{U}$. We can therefore take these equations to define the true Lie derivative $\mathcal{L}_\varepsilon$ on 0-forms and the Lie bracket $[\varepsilon_i, \varepsilon_j]$. We use the same axiomatic technique as above to build a graded Lie algebra of derivations on $\Lambda(\mathfrak{E}, Pl)$: $\mathcal{L}_\varepsilon$ of degree 0, $\varepsilon \lrcorner$ of degree -1, and $d$ of degree 1, followed by $j_K$ and $\mathcal{L}_K$.

In the previous section, we identified $\mathcal{A}_* \equiv \zeta_* \circ P_V : \mathfrak{E} \to \mathfrak{g}^{\mathcal{U}_*}$ with the physicist's "ghost field"; now we can justify this claim. The map $\zeta_*$ from the vertical ideal of $\mathfrak{E}$ to the physicist's gauge algebra $\mathcal{E} = \mathfrak{g}^{\mathcal{U}_*}$is the Maurer-Cartan form on $V\mathfrak{E}$. Since the Lie bracket on the full Lie algebra $\mathfrak{E}$ is not pointwise linear, it does not make sense to extend the Maurer-Cartan form to an "$\mathfrak{E}$-valued" 1-form on $\mathfrak{E}$. However, we can use a connexion on $B$ (in the form of the vertical projection $P_V$) to define $\mathcal{A}_*$ such that it coincides with $\zeta_*$ on $V\mathfrak{E}$ and annihilates vector fields that are horizontal with respect to the connexion, i. e., $\varepsilon \lrcorner \mathcal{A}_* = \zeta_*(P_V(\varepsilon))$. That is exactly how a physicist expects a Lie algebra valued "Lorentz scalar" to behave.

With the aid of the Frölicher-Nijenhuis construction, we can go farther. Having identified the BRST operator $s$ with the "restriction" of $-d$ to the vertical ideal of $\mathfrak{E}$, we may write $sQ = -\mathcal{L}_{P_V} Q$, where $P_V$ is taken to be a $V\mathfrak{E}$-valued 1-form on $\mathfrak{E}$. Having identified $W(E_j)$ with $-\mathcal{L}_{E_j}$ on the physicist's gauge algebra, we may write $W(\varepsilon)\,Q = -\mathcal{L}_{P_V(\varepsilon)}Q$. Expanding this formula, we have:

$$\begin{aligned} W(\varepsilon)\,Q &= -\left[[\varepsilon\lrcorner,\, j_{P_V}],\, d\right]Q \\ &= -\left[\varepsilon\lrcorner,\, [j_{P_V},\, d]\right]Q + \left[j_{P_V},\, [\varepsilon\lrcorner,\, d]\right]Q \\ &= \left[\varepsilon\lrcorner,\, -\mathcal{L}_{P_V}\right]Q + \left[j_{P_V},\, \mathcal{L}_\varepsilon\right]Q \end{aligned}$$

Now this is an interesting result. An informal extrapolation from Schücker's original formula for the action of the BRST operator might produce the first term alone, viz. $W(\varepsilon)\,Q = [\varepsilon\lrcorner,\, s]\,Q$. This simpler formula is correct if $Q$ is a horizontal form (on which the action of $j_{P_V}$ coincides with that of the rank operator $j_\zeta$), because $\mathcal{L}_\varepsilon$ is of rank 0. But when $Q$ is not horizontal and $\varepsilon$ is not vertical, we have a novel term in the Ward operator to contend with.

(Once the Frölicher-Nijenhuis proof is inserted above, we will be able to use the lemma on $[j_L,\, \mathcal{L}_K]$ to demonstrate that $[j_{P_V},\, \mathcal{L}_\varepsilon] = -j\left([P_V,\, \varepsilon]\right)$, which is zero when $\varepsilon$ is vertical. I think it will turn out that $[P_V,\, \varepsilon] = -P_H\varepsilon$.)

Let's look at this from another angle, applying our earlier example of a Frölicher-Nijenhuis calculation to the example of $P_V$:

$$\varepsilon_l\lrcorner \ldots \varepsilon_0\lrcorner\,\mathcal{L}_{P_V}Q = \sum_{j=0}^{l}(-1)^j \left[\begin{array}{l} \varepsilon_l\lrcorner \ldots \hat{\varepsilon_j}\lrcorner \ldots \varepsilon_0\lrcorner\,\mathcal{L}_{P_V(\varepsilon_j)}Q \\ -\displaystyle\sum_{i=0}^{j-1}(-1)^i\,\varepsilon_l\lrcorner \ldots \hat{\varepsilon_j}\lrcorner \ldots \hat{\varepsilon_i}\lrcorner \ldots \varepsilon_0\lrcorner\,P_V\left([\varepsilon_i,\, \varepsilon_j]\right)\lrcorner\,Q \end{array}\right]$$

What's interesting about this is that, even if $Q$ is "vertical" relative to some connexion (annilates the connexion-dependent horizontal lift of vect$[M]$) and so is its Lie derivative $\mathcal{L}_{P_V(\varepsilon_j)}Q$ along any vertical vector field, $\mathcal{L}_{P_V}Q$ is not. The vertical subspace of vect$[B]$ forms an ideal under the Lie bracket inside $P_V\left([\varepsilon_i,\, \varepsilon_j]\right)$ but the horizontal subspace does not (unless the bundle is trivial and therefore admits a flat connexion). This can be seen in the formula for the Lie bracket above, which shows that $\lambda_*\left[\mathcal{A}_*(\varepsilon_i),\, \mathcal{A}_*(\varepsilon_j)\right] = [\varepsilon_i,\, \varepsilon_j]$ on vertical vector fields but $\mathbb{A}\left[\pi_*(\varepsilon_i),\, \pi_*(\varepsilon_j)\right] \neq [\varepsilon_i,\, \varepsilon_j]$ on general horizontal vector fields.

## 1.6. The BRST Transformation

One of the most important physical applications of the BRST operator is its role in defining the "BRST transformation". This is a variational prescription $\delta\phi_i = \varepsilon s\phi_i$ for all fields $\phi_i$ in a quantum theory, including those introduced to break gauge invariance, under which the Lagrangian is invariant (up to a global divergence). (Any gauge fixing procedure which fails to exhibit such an invariance spells trouble for a perturbative expansion, essentially because the "horizontal" slice through Hilbert space that satisfies the gauge fixing condition is not of uniform thickness along the "vertical" axis of gauge transformations.) Physicists speak of such a Lagrangian-preserving infinitesimal transformation as a "continuous symmetry" and of $s$ as a "conserved charge".

Conserved charges play an important role in the Hamiltonian picture of field theory, and hence in S-matrix calculations, because the fact that they commute with the time evolution operator $\exp(-iHt)$ implies that initial states within one eigenspace of a conserved charge cannot evolve into final states within a different eigenspace. (This assumes, of course, that the charge operator also commutes with $i$; more about this later.) Perturbative calculations and other approximation procedures in which $H \to \widetilde{H}$ must be designed to preserve this property of the time evolution operator, and more generally to preserve the commutation property $\left[\widetilde{H}, s\right] = 0$. Otherwise they would violate the continuous symmetry associated with $s$, resulting in calculations that produce different answers in physically indistinguishable situations. This is usually not an acceptable sort of approximation, especially when the experimental scale is so far above the Planck scale that calculation “anomalies” of this kind come out infinite (or at least much larger than the desired result).

The BRST charge $s$ is an unusual sort of conserved charge because it is nilpotent (of degree 2), i. e., $s^2 = 0$. The usual notion of eigenspaces preserved by the Hamiltonian breaks down when the conserved charge is singular. Instead, we find that any particular space on which $s$ is an endomorphism can be broken down in terms of the subspaces $\ker s = \{Q : sQ = 0\}$ and $\text{img } s = \{Q : Q = sQ'\} \subset \ker s$. An initial state inside $\ker s$ (i. e., $s\,|\Psi\rangle = 0$) cannot evolve to a final state outside $\ker s$ (i. e., $s \exp(-iHt)\,|\Psi\rangle \neq 0$) and vice versa. Likewise, an initial state inside img $s$ (i. e., $|\Psi\rangle = s\,|\Psi'\rangle$ for some $|\Psi'\rangle$) cannot evolve to a final state outside img $s$ (i. e., $\exp(-iHt)\,|\Psi\rangle \neq s\,|\Psi'\rangle$for any $|\Psi'\rangle$) and vice versa.

Assuming that the state space and its inner product are defined such that the action of $s$ on states is Hermitian or anti-Hermitian (i. e., $(s\,|\Psi'\rangle)^\dagger = \pm \langle\Psi'|\, s)$, the inner product of any state in img $s$ with any state in $\ker s$ is zero. If we have reason to think that the initial states of physical interest all lie within $\ker s$, then no S-matrix calculation can distinguish between two initial (or final) states that differ only by an element of img $s$, and we can say that the asymptotic physical states of our theory lie in the quotient space $\ker s \setminus \text{img } s$. If we choose to represent each initial physical state by a particular representative of its equivalence class, we may say that we are picking out a subspace “transverse” to img $s$. The time evolution operator generally will not preserve this subspace, and therefore we will have to deal with intermediate states in all of $\ker s$; but we may limit S-matrix calculations to final states in our “transverse” subspace.

The alert reader will have perceived that the replacement of the quotient space in which physical states actually live by a “transverse” subspace is the same sort of delta-function constraint that haunted our gauge fixing procedure in the first place. The difference is that this constraint lives at the asymptotic limits of time, at which (by the conventions of S-matrix calculations) the gauge coupling is “turned off” ($g \to 0$). This affects the action of $s$ on the space of asymptotic states in such a way that the entire sector of state space containing “physical” fermions (not ghosts) and transverse polarizations of gauge bosons is “annihilated in both directions” by $s$ (i. e., within its kernel but not part of its image).

As long as our Lagrangian obeys $sL = 0$ for all values of $g$, we can adiabatically “turn on” the gauge coupling to obtain the S-matrix of the interaction Lagrangian. It remains to demonstrate that the remaining degrees of freedom in the asymptotic state space (in a generalized Lorentz gauge, “forward” and “backward” polarizations of the gauge bosons, together with the ghost and “antighost” introduced in order to construct a BRST-invariant Lagrangian) can be suppressed in S-matrix calculations. This means that they are all either inside img$s$ or outside $\ker s$ and that their “volume” (according to the functional measure used in functional quantization) does not depend on the connexion.

In fact, the connexion independence of the “no-ghost” slice taken through the asymptotic state space is guaranteed when $sL = 0$. In the framework of functional quantization, this is because the functional determinant of the Jacobian of the embedding of the reduced state space into the full state space, which we need in order to regularize the constraint embodied in the gauge fixing term, is supplied by the ghost term added to the Lagrangian at the same time. In our fiber bundle framework, we arrive at a more geometric interpretation. Look at the effect of the BRST "gauge fixing" procedure on a typical Yang-Mills Lagrangian (written in physicist style):

$$L = -\frac{1}{4}\left(F^a_{\mu\nu}\right)^2 + \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi + \frac{\xi}{2}\left(B^a\right)^2 + B^a\partial^\mu A^a_\mu + \bar{c}^a\left(-\partial^\mu D^{ac}_\mu\right)c^c$$

The key is to realize that this Lagrangian is actually still gauge invariant up to a total divergence $s\left(\bar{c}^a\partial^\mu A^a_\mu\right)$, *if* $s$ is taken to anti-commute with $\bar{c}$ (and therefore to commute with $B \equiv s\bar{c}$). There is nothing unreasonable about this prescription; the action of $s$ on $c$ and $A_\mu$ (and, as we shall see, on $\psi$) is fixed by their geometrical significance on the fiber bundle, but $\bar{c}$ exists only as a formal dual to $c$ and may be given the algebraic properties of our choice. This total divergence is not gauge invariant on individual fibers, but its integral over any almost-complete chart reduces to a surface term on the boundary of the chart. The integral of this surface term is not necessarily zero, because the edges of the chart do not necessarily meet on the fiber bundle! But it is a topological invariant, inaccessible via perturbation theory relative to a non-interacting Lagrangian.

We have combined the BRST perspective on the Faddeev-Popov Lagrangian with the fiber bundle view of the BRST operator to obtain an essential insight: the Faddeev-Popov prescription breaks the local gauge invariance of the Lagrangian in order to obtain a usable framework for perturbative calculations, but preserves (up to a topological term) the global gauge invariance of the action functional. The Lagrangian is BRST closed both before and after gauge fixing.

Recognizing the relationship between the BRST operator and the exterior derivative on the fiber bundle, we may ask: Can we write a Lagrangian that is not only BRST closed but entirely closed on the fiber bundle? If so, we have a theory that is not only gauge invariant but diffeomorphism invariant, depending only on the topology of the fiber bundle and its connexion. A fiber bundle of particular interest is the $GL^+(4, \mathbb{R}) \times O^+(4)$ principal bundle over a 4-manifold, and a theory of particular interest has only a few objects before gauge fixing: an orthonormal frame and the two connexions it relates. Restricting the connexions to those compatible with the orthonormal frame (and torsion-free) recovers the intrinsic Riemannian geometry of the 4-manifold, and imposing invariance under infinitesimal changes of metric leaves one with a Lagrangian containing only topological invariants of the base 4-manifold. Since every compact orientable 4-manifold may be given a Riemannian metric, one obtains a perturbative framework for the calculation of topological invariants on arbitrary (compact, orientable) 4-manifolds.