Appendix B

Differential Geometry 101

In the main text, we have gone through the elegant formulation of differential geometry based on the concept of bundles. A little more down-to-earth approach to the same subject is often useful and heavily employed in General Relativity. In this appendix, we will recall basic notions there such as manifolds, charts, covariant derivatives, and curvatures with the emphasis on Riemannian geometry. Content of this Appendix is borrowed from main text of a separate volume “Gravitation for Theorists” published separately, which in some sense plays the role of a classical precursor for the current volume.

B.1 Manifolds, Charts, and Tensors

Euclidean geometry, say, $\mathbb{R}^d$ allows an intuitive notion of straight lines, a pair of which parallel to each other never intersect with each other. Such a global property cannot be expected to hold for more general curved space. On the other hand, even on curved spaces, we can focus on an infinitesimal neighborhood of any given point where locally, one can say that the deviation from such Euclidean geometry is arbitrarily small. If a space can be patched up from such small neighborhoods, each of which can be regarded as small open subset of $\mathbb{R}^d$, the result is called a manifold. We will generally denote such objects as $\mathcal{M}_d$ where d is the dimension, meaning the manifold looks like $\mathbb{R}^d$ locally.

Therefore, the actual content of the manifold starts with the collection of the local regions each of which could be embedded into $\mathbb{R}^d$ and how these patches glue together to form the entire space as $\mathcal{M}_d$. Each such a patch is called the "chart" $\mathcal{U}_a$. As a set the manifold is union of such patches covered by local charts,

$$\mathcal{M}_d = \cup_a \mathcal{U}_a \tag{B.1.1}$$

with each $\mathcal{U}_a$ equipped with coordinate systems $x^\mu_{(a)}$. A simple example is a sphere embedded into $\mathbb{R}^3$ by $x^2 + y^2 + z^2 = R^2$. Positions in the upper and the lower hemispheres are respectively labeled by (x, y), and the vertical position is automatically determined as $z = \pm\sqrt{R^2 - x^2 - y^2}$. As with this example, one cannot cover the entire manifold with a map to $\mathbb{R}^d$.

A rule must be given how to identify things when a pair of chart overlap with each other. For instance, a point p on $\mathcal{M}_d$ could be assigned values of coordinates on any local chart that contain p, and the latter does not change simply because we change the chart, so

$$p(x^\mu_{(a)}) = p(x^\mu_{(b)}) \tag{B.1.2}$$

for any pair of charts and their coordinate systems $x^\mu_{(a)}$ and $x^\mu_{(b)}$. This gives relation between the two charts, or a transition map,

$$x^\mu_{(b)} = x^\mu_{(b)}(x^\nu_{(a)}) \tag{B.1.3}$$

for any pair that share a common region.

A manifold $\mathcal{M}_d$ is said to be continuous if $x_{(b)}$ are continuous functions of $x_{(a)}$, and differentiable if $x_{(b)}$ are differentiable functions of $x_{(a)}$. One can further distinguish by how many time these are differentiable, but for physics purpose here, we may as well assume that these transition functions are smooth, meaning differentiable infinite number of times. In reality, we will also encounter many examples of manifolds below that admits exceptional places called curvature singularities, where this smoothness breaks down, so we will proceed with this possibility in mind as well.

A function, which is a map from points on the manifold to numbers, should also be glued between different charts such that

$$f(x_{(a)}(p)) = \tilde{f}(x_{(b)}(p)) \tag{B.1.4}$$

This demands, simply, that the function takes a well-defined value at a given position

$p \in \mathcal{M}_d$, and does not rely on the choice of the coordinates, or the charts, when multiple coordinate systems and/or charts can be used to define the same position. For functions, the gluing rule between overlapping charts is simple enough.

As such, we will often drop the tilde in the latter $\tilde{f}$, even though the functional form of $f(x_{(a)})$, given $x_{(a)}$, and that of $\tilde{f}(x_{(b)})$, given a different coordinate system, $x_{(b)}$, are not the same. Perhaps more appropriately, we should consider a function more abstractly as a map,

$$f: \quad \mathcal{M}_d \quad \rightarrow \quad \mathbb{R} \text{ or } \mathbb{C} , \tag{B.1.5}$$

We will go back and forth between the point p and the coordinate values x^μ that label p on $\mathcal{M}$. As such, we will mix notations for value of functions,

$$f\Big|_p , \qquad f\Big|_{p(x)} , \qquad f(x) \tag{B.1.6}$$

etc., where $p = p(x)$ is the point on $\mathcal{M}_d$ represented by the coordinate value x.

The main question is what are the derivatives and the integrations one may use on $\mathcal{M}_d$. Given a local chart, an immediate object we can define is a partial derivative of functions, namely

$$\frac{\partial f(x)}{\partial x^\mu} = \lim_{\delta \to 0} \frac{f(x^1, \cdots, x^\mu + \delta, \cdots, x^d) - f(x^1, \cdots, x^\mu, \cdots, x^d)}{\delta} , \tag{B.1.7}$$

where, as usual, the derivative is computed with all other $x^{\nu \neq \mu}$ held fixed. This means that the partial derivative requires an entire coordinate system, collectively, near that point; it cannot be defined if the other coordinates $x^{\nu \neq \mu}$ are not properly specified.

B.1.1 Vectors as Directional Derivatives

In the Euclidean space, the Cartesian coordinate and the vectors often share the common notation of d-tuple, leaving the impression that the two are interchangeable. However, this was possible only because the Euclidean space is itself is a vector space, where addition and subtraction between elements are well-defined. No such luck for general manifolds, since two positions cannot be added together to yield a third position, so coordinates do not extend the notation of vectors naturally, although we label them by $\mu = 1, \cdots, d$ (or $\mu = 0, 1, \cdots, d-1$). If something is moving along a trajectory $x^\mu(\mathbf{s})$ on $\mathcal{M}_d$, its velocity

$$\frac{dx^\mu(\mathbf{s})}{d\mathbf{s}} \tag{B.1.8}$$

can have a meaning as an arrow sitting at $x^\mu(\mathbf{s})$. But more generally, we need a working model of tangent vectors on $\mathcal{M}_d$, inherent to the manifold, rather than by the presence of moving objects confined to the manifold.

On general manifolds, vectors make sense only as an infinitesimal deviation from a given point, and the linear space built out of such objects is called a tangent space. Indeed a velocity vector of an object at a point belong to such a tangent space sitting at that point. One can visualize a tangent space at point x as Euclidean space, $\mathbb{R}^d$, sitting at x and grazing the manifold. On each such tangent space, we can add and subtract its elements as in the Euclidean geometry of $\mathbb{R}^d$. The very definition of the manifold is that, at any given point thereof, such a tangent space can be placed and locally used to label points smoothly in the small neighborhood. This is a nice visualization but does not immediately give us the computational power of the vector notations in the Euclidean space. For general manifolds, however, this picture does lead to a powerful tool called the calculus on manifold.

Given a natural map from the tangent space at x to its immediate neighborhood, a vector V at x can be phrased in terms of how one compute directional derivatives of arbitrary function on $\mathcal{M}_d$. Given coordinate charts, and arbitrary function $f(x)$, we define

$$V[\,f\,] \equiv V^\mu \frac{\partial}{\partial x^\mu} f \ , \tag{B.1.9}$$

which, given completely arbitrary but differentiable f's defined in an immediate neighborhood of x, gives an unambiguous characterization of V. Since the function f is taken to be completely arbitrary, the object $V[\cdot]$ has nothing to do with particular functions, so this leads us to write abstractly,

$$V = V^\mu \frac{\partial}{\partial x^\mu} \tag{B.1.10}$$

where the partial derivatives, $\partial/\partial x^\mu$, serves as convenient basis for vector fields. This can be easily extended to all points on $\mathcal{M}_d$, whereby a vector field, or collection of V at all x's, is determined by $V[\,f\,]$ everywhere for arbitrary function f's on $\mathcal{M}_d$.

Under coordinate transformations $x \to \tilde{x}(x)$, transformation of V follows from the chain rule as

$$V^\mu \frac{\partial}{\partial x^\mu} = V^\mu \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial}{\partial \tilde{x}^\alpha} = \tilde{V}^\alpha \frac{\partial}{\partial \tilde{x}^\alpha} \ , \tag{B.1.11}$$

so

$$\tilde{V}^\alpha = V^\mu \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \ , \qquad \tilde{V}^\alpha \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} = V^\mu \ . \tag{B.1.12}$$

What must be noted here is that even though the component transforms nontrivially when we switch between different coordinate systems, the vector V itself does not. Although we often say that vector fields transform under the coordinate transformation as above, this actually refers to the components V^μ's.

An alternative is to consider the vector as collection of components V^μ with the accompanying transformation rules among them; Most of earlier physics text starts out this way. On the other hand, if we introduce a vector as the directional derivative as above, the object is entirely inert under coordinate transformations or any other changes of basis. The transformation rules among components between coordinate charts follow from the definition rather than being part of the definition. This way, a vector and more generally tensors may be defined more intrinsically without having to refer to coordinate systems.

B.1.2 Differential 1-Forms

On the other hand, given a specific coordinate system, the basis $\partial/\partial x^\mu$ are themselves vectors in some restrictive sense. A vector, or a tensor more generally, is an object we try to extend to the entire manifold. This would be the right definition for most purpose. However, sometimes we need to also think about objects that are defined on specific chart. Ignoring such global issues, $\partial/\partial x^\mu$ has the defining property,

$$\frac{\partial}{\partial x^\mu}[x^\lambda] = \delta_\mu{}^\lambda \ . \tag{B.1.13}$$

The coordinates x^λ are defined on a local chart of the manifold covered by the coordinate system in question, but this is no problem as long as the above operation is performed within that chart.

Given this, we may invent differential 1-forms, dx^λ, with natural bilinear pairing against vectors as

$$\langle dx^\lambda, \frac{\partial}{\partial x^\mu}\rangle \equiv \frac{\partial}{\partial x^\mu}[x^\lambda] = \delta_\mu{}^\lambda \ . \tag{B.1.14}$$

Given the coordinate system, dx^λ and $\partial/\partial x^\mu$ are said to be "dual" to each other. An immediate generalization is differential 1-form df, given a differentiable function f,

$$df \equiv \frac{\partial f}{\partial x^\mu}\, dx^\mu \ . \tag{B.1.15}$$

An 1-form of this type is said to be "exact" if f is defined globally on the manifold. The pairing we have introduced above computes the directional derivative,

$$\langle df, V\rangle = V[\,f\,] = V^\mu \partial_\mu f \tag{B.1.16}$$

for general vector V.

As with the vectors, which make use of $\partial/\partial x^\mu$ as a basis, we can invent a more general differential 1-form as

$$\Lambda = \Lambda_\mu\, dx^\mu \ , \tag{B.1.17}$$

The above natural pairing against vectors works as

$$\langle \Lambda, V\rangle = V^\mu \Lambda_\mu \ . \tag{B.1.18}$$

inducing a contraction between these two types of indices. As with vectors, we want this object to make sense independent of coordinate system choice,

$$\Lambda_\mu\, dx^\mu = \tilde{\Lambda}_\alpha\, d\tilde{x}^\alpha \ , \tag{B.1.19}$$

which demands

$$\Lambda_\mu = \tilde{\Lambda}_\alpha \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \tag{B.1.20}$$

since $d\tilde{x}^\alpha = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} dx^\mu$. Note how the pairing $\langle \Lambda, V\rangle$ is also invariant under such transformations. This is exactly the opposite chain rule relative to that for vectorial indices, so the pairing $\langle \Lambda, V\rangle$ is inert under such coordinate transformation.

Contravariant and Covariant Tensors

This leads to more general objects with multiple vector indices and 1-form indices. The former leads to

$$W^{\mu_1 \cdots \mu_m} \frac{\partial}{\partial x^{\mu_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{\mu_m}} . \tag{B.1.21}$$

known as the contravariant tensors in older literature. Also, this gives the transformation of W under $x \to \tilde{x}(x)$,

$$\frac{\partial}{\partial \tilde{x}^{\alpha}} = \left(\frac{\partial x^{\mu}}{\partial \tilde{x}^{\alpha}} \right) \frac{\partial}{\partial x^{\mu}} . \tag{B.1.22}$$

W itself does not depend on coordinate choices, just as with V, so

$$\tilde{W}^{\alpha_1 \cdots \alpha_m} = W^{\mu_1 \cdots \mu_m} \frac{\partial \tilde{x}^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial \tilde{x}^{\alpha_m}}{\partial x^{\mu_m}} \tag{B.1.23}$$

component-wise.

Just as how the basis $\partial / \partial x^{\mu}$ was used to build higher-rank contravariant tensors, one can imagine higher-rank covariant tensor,

$$S = S_{\mu_1 \cdots \mu_k} dx^{\mu_1} \otimes \cdots \otimes dx^{\mu_k} \tag{B.1.24}$$

as well, whose components transform similarly as

$$S_{\mu_1 \cdots \mu_k} = \tilde{S}_{\alpha_1 \cdots \alpha_k} \frac{\partial \tilde{x}^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial \tilde{x}^{\alpha_k}}{\partial x^{\mu_k}} , \tag{B.1.25}$$

so that S itself, with the basis 1-forms attached, remains inert.

Even more generally we may consider tensors of mixed type,

$$\Sigma_{\lambda_1 \cdots \lambda_k}{}^{\mu_1 \cdots \mu_m} \, dx^{\lambda_1} \otimes \cdots \otimes dx^{\lambda_k} \otimes \frac{\partial}{\partial x^{\mu_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{\mu_m}} . \tag{B.1.26}$$

The transformation rule should be clear by now.

B.2 Lie Derivative

Is there a way to take the directional derivative of such tensors along a vector V? So far, we have learned only how to take a directional derivative of functions. To extend the directional derivative to vectors and tensors, a useful first step is to take directional derivatives on a function twice. That is, let us consider the antisymmetric combination of two repeated directional derivatives,

$$\begin{aligned} V[W[f]] - W[V[f]] &= V^\mu \frac{\partial}{\partial x^\mu}\left(W^\nu \frac{\partial}{\partial x^\nu} f\right) - W^\mu \frac{\partial}{\partial x^\mu}\left(V^\nu \frac{\partial}{\partial x^\nu} f\right) \\ &= \underbrace{(V^\mu \partial_\mu W^\nu - W^\mu \partial_\mu V^\nu)}_{\text{Again a vector}} \partial_\nu f \ , \end{aligned} \tag{B.2.1}$$

where terms with two partial derivatives acting on f dropped out due to the universal property of the partial derivatives,

$$\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}\frac{\partial}{\partial x^\mu} \ . \tag{B.2.2}$$

Since the first vector used for the directional derivative is differentiated by the second vector, we are in effect working with vector fields rather than isolated tangent vectors.

As such, the "commutator" of the two directional derivatives produces another directional derivative, defining a new vector field,

$$[V, W]^\nu = V^\mu \partial_\mu W^\nu - W^\mu \partial_\mu V^\nu \ . \tag{B.2.3}$$

Of course, we need to make sure that this is a vector, i.e., it transforms correctly as a vector would under coordinate changes. To see this, we start with

$$\tilde{V}^\alpha \tilde{\partial}_\alpha \tilde{W}^\beta = \frac{\partial \tilde{x}^\beta}{\partial x^\nu}(V^\mu \partial_\mu W^\nu) + V^\mu W^\nu \left(\frac{\partial^2 \tilde{x}^\beta}{\partial x^\mu \partial x^\nu}\right) \ , \tag{B.2.4}$$

which by itself does not transform correctly as a vector, due to the last double-derivative term.

On the other hand, the above commutator does transform correctly as a vector,

$$\left(\tilde{V}^\alpha \tilde{\partial}_\alpha \tilde{W}^\beta - \tilde{W}^\alpha \tilde{\partial}_\alpha \tilde{V}^\beta\right) = \frac{\partial \tilde{x}^\beta}{\partial x^\nu}(V^\mu \partial_\mu W^\nu - W^\mu \partial_\mu V^\nu) \ , \tag{B.2.5}$$

since the problematic piece with a double derivative of $\tilde{x}$ by x cancels out. Using this, we can define a directional derivative of a vector with respect to another vector called the Lie derivative,

$$\mathfrak{L}_V[W] = -\mathfrak{L}_W V \equiv [V, W] \ , \tag{B.2.6}$$

sometime written as a commutator like the last expression. With this, note that the basis vectors obey

$$\left[\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\right] = 0 \ . \tag{B.2.7}$$

Can we extend this Lie derivative to 1-form Λ? For this, we start with the function $\langle \Lambda, V\rangle$, on which the Lie derivative is nothing but the directional derivative on a function,

$$\mathfrak{L}_V\langle \Lambda, W\rangle = V^\mu \partial_\mu \langle \Lambda, W\rangle = \left(V^\beta \partial_\beta W^\alpha\right)\Lambda_\alpha + \left(V^\beta \partial_\beta \Lambda_\alpha\right) W^\alpha \ . \tag{B.2.8}$$

On the other hand, a derivative must obey the Leibniz rule, so

$$\begin{aligned} \mathfrak{L}_V\langle \Lambda, W\rangle &= \langle \Lambda, \mathfrak{L}_V W\rangle + \langle \mathfrak{L}_V \Lambda, W\rangle \\ &= \left(V^\beta \partial_\beta W^\alpha - W^\beta \partial_\beta V^\alpha\right)\Lambda_\alpha + W^\alpha (\mathfrak{L}_V \Lambda)_\alpha \ . \end{aligned} \tag{B.2.9}$$

Equating the two right hand sides, we find

$$W^\alpha \left(\mathfrak{L}_V \Lambda\right)_\alpha = W^\alpha \left(V^\beta \partial_\beta \Lambda_\alpha\right) + W^\alpha \left(\partial_\alpha V^\beta\right)\Lambda_\beta \ , \tag{B.2.10}$$

which defines the Lie derivative of the 1-form as

$$(\mathfrak{L}_V \Lambda)_\alpha = V^\beta \partial_\beta \Lambda_\alpha + \left(\partial_\alpha V^\beta\right)\Lambda_\beta \ . \tag{B.2.11}$$

Similarly with the Lie derivative of a vector, one can see that this Lie derivative of a 1-form is again a 1-form, transforming properly,

$$V^\beta \partial_\beta \Lambda_\alpha + \left(\partial_\alpha V^\beta\right)\Lambda_\beta \;=\; \left[(\tilde{V}\cdot\tilde{\partial})\tilde{\Lambda}_\mu + (\tilde{\partial}_\mu \tilde{V}^\lambda)\tilde{\Lambda}_\lambda\right]\cdot \frac{\partial \tilde{x}^\mu}{\partial x^\alpha}$$

$$+ \tilde{V}^{\mu}\tilde{\Lambda}_{\nu} \underbrace{\left[\tilde{\partial}_{\mu}\left(\frac{\partial \tilde{x}^{\nu}}{\partial x^{\alpha}}\right) + \partial_{\alpha}\left(\frac{\partial x^{\beta}}{\partial \tilde{x}^{\mu}}\right)\frac{\partial \tilde{x}^{\nu}}{\partial x^{\beta}}\right]}_{=\tilde{\partial}_{\mu}\left(\frac{\partial \tilde{x}^{\nu}}{\partial x^{\alpha}}\right)-\tilde{\partial}_{\mu}\left(\frac{\partial \tilde{x}^{\nu}}{\partial x^{\alpha}}\right)=0}, \tag{B.2.12}$$

where we used

$$\partial_{\alpha}\left(\frac{\partial \tilde{x}^{\nu}}{\partial x^{\beta}}\frac{\partial x^{\beta}}{\partial \tilde{x}^{\mu}}\right) = \partial_{\alpha}(\delta^{\nu}_{\mu}) = 0 \ , \tag{B.2.13}$$

and the chain rule.

Given how $\mathfrak{L}_V$ acts on vectors and 1-forms, generalizations to all tensors follow,

$$\begin{aligned}(\mathfrak{L}_V\Sigma)_{\lambda_1\lambda_2\cdots}{}^{\mu_1\mu_2\cdots} \quad &= \quad V^{\alpha}\partial_{\alpha}\Sigma_{\lambda_1\lambda_2\cdots}{}^{\mu_1\mu_2\cdots} \\ &\quad + \left(\partial_{\lambda_1}V^{\alpha}\Sigma_{\alpha\lambda_2\cdots}{}^{\mu_1\mu_2\cdots} + \partial_{\lambda_2}V^{\alpha}\Sigma_{\lambda_1\alpha\cdots}{}^{\mu_1\mu_2\cdots} + \cdots\right) \\ &\quad - \left(\partial_{\alpha}V^{\mu_1}\Sigma_{\lambda_1\lambda_2\cdots}{}^{\alpha\mu_2\cdots} + \partial_{\alpha}V^{\mu_2}\Sigma_{\lambda_1\lambda_2\cdots}{}^{\mu_1\alpha\cdots} + \cdots\right)\end{aligned} \tag{B.2.14}$$

The tensorial transformation of $\mathfrak{L}_V\Sigma$ follows again by repeating the above computation index by index. Partial derivatives on a tensor do not give a new tensor, generally, but we have identified how a directional derivative defined by a vector can take a derivative on a tensor and map it to another tensor.

Pull-Back and Push-Forward

A useful concept to learn is a map between manifolds, which can be used in turn to related tensors on the respective manifold. A pair of manifolds $\mathcal{M}$ and $\mathcal{N}$ may be sometimes connected by a map,

$$\sigma \ : \ \mathcal{M} \quad \longrightarrow \quad \mathcal{N} \tag{B.2.15}$$

With $p \in \mathcal{M}$ and $\sigma(p) \in \mathcal{N}$, the map pulls back any function h on $\mathcal{N}$ to a function on $\mathcal{M}$,

$$\sigma^* h\,(p) = h(\sigma(p)) \tag{B.2.16}$$

When the dimension of the two manifolds are equal this pulls back local charts $\{y^{\alpha}\}$ on $\mathcal{N}$ to ones on $\mathcal{M}$, say, $Y^{\alpha} \equiv \sigma^*(y^{\alpha})$, which allows pull-back of differential forms and other covariant tensors via

$$dY^\alpha = \sigma^*(dy^\alpha) \tag{B.2.17}$$

In particular, with $\mathcal{M} = \mathcal{N}$, such a map relates different points of the manifold.

Now we can understand the Lie derivative $\mathfrak{L}_V$ more geometrically. For this, take $\mathcal{M} = \mathcal{N}$ and consider an infinitesimal shift of position,

$$\sigma_{\epsilon V} : x^\mu \longmapsto x^\mu + \epsilon V^\mu \tag{B.2.18}$$

or more precisely,

$$\sigma_{\epsilon V} : p(x^\mu) \longmapsto p(x^\mu + \epsilon V^\mu) , \tag{B.2.19}$$

where $p(x)$ is a position on the manifold, represented by particular values of x^μ in the given coordinate system. One should be mindful that this $\sigma_{\{\epsilon V\}}$ is not a coordinate transformation but a map from the manifold onto the same manifold. We will also write this, in a slight abuse of notation, also as

$$p \longmapsto p + \epsilon V , \tag{B.2.20}$$

which makes sense in the infinitesimal limit of ϵ.

Given a function f, this shift can be used to define a new function $f_{\epsilon V}$, given f, as

$$f_{\epsilon V}\Big|_p \equiv \sigma^*_{\epsilon V} f \Big|_p = f \Big|_{p+\epsilon V} . \tag{B.2.21}$$

This is a special case of a more general operation called the pull-back. The Lie derivative of a function f along V is then

$$\mathfrak{L}_V f \Big|_p = \lim_{\epsilon\to 0} \frac{f_{\epsilon V} - f}{\epsilon}\Big|_p = V^\mu \partial_\mu f \Big|_p \tag{B.2.22}$$

as advertised.

The above infinitesimal pull-back is easily generalized to covariant tensors S,

$$S = S_{\alpha\beta\,\cdots\,\gamma}\, dx^\alpha \otimes dx^\beta \otimes \;\cdots\; \otimes dx^\gamma \tag{B.2.23}$$

as

$$S_{\epsilon V}\bigg|_p \equiv \sigma^*_{\epsilon V} S\bigg|_p = S\bigg|_{p+\epsilon V} \ . \tag{B.2.24}$$

For infinitesimal ϵ, one can see that

$$\begin{aligned} S_{\epsilon V}\bigg|_p &= S_{\alpha\beta\,\cdots\,\gamma}(x+\epsilon V)\; d(x+\epsilon V)^\alpha \otimes d(x+\epsilon V)^\beta \otimes\ \cdots\ \otimes d(x+\epsilon V)^\gamma \\ &= S\bigg|_p + \epsilon\left(\mathfrak{L}_V S\right)_{\alpha\beta\,\cdots\,\gamma}\; dx^\alpha \otimes dx^\beta \otimes\ \cdots\ \otimes dx^\gamma + O(\epsilon^2)\ , \end{aligned} \tag{B.2.25}$$

which defines

$$\mathfrak{L}_V S = \lim_{\epsilon\to 0}\frac{S_{\epsilon V}-S}{\epsilon} \tag{B.2.26}$$

similarly. Writing things out in terms of the components,

$$\left(\mathfrak{L}_V S\right)_{\alpha\beta\,\cdots\,\gamma} = V^\mu \partial_\mu S_{\alpha\beta\,\cdots\,\gamma} + S_{\mu\beta\,\cdots\,\gamma}\partial_\alpha V^\mu + \cdots + S_{\alpha\beta\,\cdots\,\mu}\partial_\gamma V^\mu \tag{B.2.27}$$

follows immediately, which confirms the earlier generalization of the Lie derivative to all contravariant tensors.

For contravariant tensors, with the basis $\frac{\partial}{\partial x}$, on the other hand, σ induces something different, called "push-forward" and denoted as σ_*. Consider a vector field W on $\mathcal{M}$ and a function h on $\mathcal{N}$. Using the pull-back above, we can bring h onto $\mathcal{M}$ and then take the directional derivative,

$$W[\sigma^* h] \tag{B.2.28}$$

on $\mathcal{M}$ where σ^* denotes the pull-back we have used above on the functions and covariant tensors above. This also means that, given W on $\mathcal{M}$, one can define $\sigma_* W$ on $\mathcal{N}$ such that

$$(\sigma_* W)[f]\bigg|_{\sigma(p)} = W[\sigma^* f]\bigg|_p \ . \tag{B.2.29}$$

Recall that the vectors are defined precisely by such directional derivatives. Hence this defines $\sigma_* W$ unambiguously on $\mathcal{N}$ as long as we demand this to hold for all differentiable f's on $\mathcal{N}$.

Now let us consider again $\sigma_{\epsilon V}$ which merely shift position infinitesimally on $\mathcal{M}$ along V direction with $\mathcal{N} = \mathcal{M}$. Unlike the covariant tensors, which can be pulled-back from the forward position and compared, we must do things backward, computing a derivative as

$$\lim_{\epsilon \to 0} \frac{W\big|_p - W_{-\epsilon V}\big|_p}{\epsilon} \;, \qquad W_{-\epsilon V}\Big|_p \equiv (\sigma_{\epsilon V})_* W\Big|_p = W\Big|_{p-\epsilon V} \;. \tag{B.2.30}$$

Does this equal the Lie derivative defined above purely on the basis of the tensorial property? One can see this rather easily from

$$\begin{aligned} W_{-\epsilon V}(x) &= W^\mu(x-\epsilon V)\frac{\partial}{\partial(x-\epsilon V)^\mu} \\ &= W^\mu(x)\partial_\mu - \epsilon\left[V^\mu\partial_\mu W^\alpha - W^\mu\partial_\mu V^\alpha\right]\partial_\alpha \;. \end{aligned} \tag{B.2.31}$$

Therefore,

$$\lim_{\epsilon \to 0}\left(\frac{W - W_{-\epsilon V}}{\epsilon}\right) = (V^\alpha\partial_\alpha W^\mu - W^\alpha\partial_\alpha V^\mu)\frac{\partial}{\partial x^\mu} = \mathfrak{L}_V W \tag{B.2.32}$$

as expected. The generalization to higher-rank contravariant tensors should be immediate, which gives

$$(\mathfrak{L}_V W)^{\alpha\beta\,\cdots\,\gamma} = V^\mu\partial_\mu W^{\alpha\beta\,\cdots\,\gamma} - W^{\mu\beta\,\cdots\,\gamma}\partial_\mu V^\alpha - \cdots - W^{\alpha\beta\,\cdots\,\mu}\partial_\mu V^\gamma \tag{B.2.33}$$

component-wise, again confirming the general formulae earlier for the case of contravariant tensors.

The Lie derivative $\mathfrak{L}$ is also a derivative; therefore, it must obey the Leibniz rule. This means that when acting on a tensor product of two different types of tensors,

$$\mathfrak{L}_V(S\otimes W) = \mathfrak{L}_V(S)\otimes W + S\otimes\mathfrak{L}_V(W) \;. \tag{B.2.34}$$

Since the partial derivative obeys the Leibniz rule, the rotating action of $\mathfrak{L}_V$ on various types of indices occurs linearly and independently. All of this brings us back to (B.2.14) for general tensors.

By the way, the pull-back and the push-forward operations we used above is available not only for infinitesimal shift on $\mathcal{M}$ but for any smooth map between a pair of (differentiable) manifolds. Given such a map,

$$\sigma : p \longmapsto \sigma(p) \tag{B.2.35}$$

from $\mathcal{M}$ and $\mathcal{N}$, a covariant tensor S on $\mathcal{N}$ can be used to define a pull-back $\sigma^* S$ as

$$\sigma^* S \bigg|_p = S \bigg|_{\sigma(p)} . \tag{B.2.36}$$

The push-forward σ_* can be similarly defined between contravariant tensors on $\mathcal{M}$ and $\mathcal{N}$.

Diffeomorphisms vs. Coordinate Transformations

Consider smooth bijective maps for $\mathcal{M} = \mathcal{N}$. One potential outfall of such a "diffeomorphism" is a coordinate change, although not strictly necessary. Given a chart $\{y^\mu\}$ over $\mathcal{N} = \mathcal{M}$ and a map $p \mapsto \sigma(p)$, the pull-back of the coordinates effectively induces a new set of coordinates on $\mathcal{M}$ as

$$y^\mu(p) \quad \longmapsto \quad x^\mu(p) \equiv \sigma^* y^\mu(p) = y^\mu((\sigma(p)) \tag{B.2.37}$$

However, as noted earlier, although σ^* may be thus used to induce a coordinate change, σ^* itself is not the same as the nominal coordinate transformation. Unlike coordinate transformation, σ^* relate coordinates of two different points p and $\sigma(p)$.

On the other hand, recall that, given a local chart, we may perform a coordinate transformation, say from $\{x^\mu\}$ to $\{\tilde{x}^\mu\}$, which has nothing to do with such diffeomorphisms. Functions (and tensors) do not change its value at a given point p, say,

$$f(x(p)) = \tilde{f}(\tilde{x}(p)) \tag{B.2.38}$$

where we revive the notation $\tilde{f}$ for $\tilde{f}(\tilde{x})$ to emphasize that the different functional form of $\tilde{f}$ written as a function of $\tilde{x}$. What does change with the coordinate transformation is the functional dependence f, for instance, on the coordinates simply because the coordinate values for a given point p have changed. This should be contrasted against how the pull-back/push-forward operations under the diffeomorphism do actually change values of functions and tensors at a given point p.

On a real line $\mathcal{M} = \mathbb{R}^1$, for instance, suppose we have $f(x) = x^3$. Using a new coordinate $\tilde{x} = x^3$, we find the same function is represented by $\tilde{f}(\tilde{x}) = \tilde{x}$. An even simpler case of $f(x) = x$ and $\tilde{x} = x - 1$, we fine $\tilde{f}(\tilde{x}) = \tilde{x} + 1$. These are different from what happens with the diffeomorphisms

$$\sigma(x) = x^3 \,, \qquad \sigma(x) = x - 1 \tag{B.2.39}$$

Yet, they are often said to be "equivalent". Exactly what is the nature of the claimed equivalence?

Consider an infinitesimal shift $\tilde{x}(x) = x - \epsilon\xi(x)$ with $\epsilon \ll 1$. Since the function value itself should be the same in the end, we cannot use the same functional form in the new coordinates and let us represent the necessary change as $\tilde{f} = f + \epsilon\boldsymbol{\delta}_\xi f$. Then,

$$f(x) = (f + \epsilon\boldsymbol{\delta}_\xi f)(\tilde{x}) = f(x) + \epsilon\left(-\xi^\mu\partial_\mu f + \boldsymbol{\delta}_\xi f\right)\bigg|_x + O(\epsilon^2) \tag{B.2.40}$$

so that we obtain

$$\boldsymbol{\delta}_\xi f(x) = \xi^\mu\partial_\mu f\,\bigg|_x = \mathfrak{L}_\xi f\,\bigg|_x \tag{B.2.41}$$

as the "variation" of the function, bringing us back to the Lie derivative again. The same works for more general tensors, and again we recover a Lie derivative dictating the "variation". We leave this as an exercise.

Since the function f as a map from $\mathcal{M}$ to $\mathbb{R}$ does not change at all under such coordinate changes, $\boldsymbol{\delta}_\xi f$ computes the difference of the same f at two different places in the given chart; these two different points p and p' on $\mathcal{M}$ are determined such that the values of the coordinates $x(p)$ happen to be equal to those of $\tilde{x}(p')$. Note how we again end up comparing the function in two nearby places, although for a very different reason. In the end, on a local chart U, we are employing a local diffeomorphism $p' \mapsto p$, which is of course why we again end up with a Lie derivative.

This "variation" makes sense only locally, since the map relies on the coordinate values. Even though it cannot be extended to a proper diffeomorphism, the infinitesimal transformation rule under it is no different than those under globally defined diffeomorphism, so the invariance of a theory under one implies the invariance under the other. In the literature, these two types of transformations are called the active, for the diffeomorphism, and the passive, for the coordinate transformation. The passive one represents an ambiguity in the description and is a close analog to the gauge ambiguity of Maxwell theory.

Be mindful that when we perform physics computations the explicit form of quantities are expressed in concrete functional forms. For instance, $V^\mu \partial_\mu$ is a perfectly sensible invariant object, yet we often deal with the components $V^\mu(x)$'s whose functional form is also equally important. The choices we make for the sake of such explicit form are redundant in view of the final fully invariant quantities, but often we can perform more detailed manipulation by dealing with $V^\mu(x)$'s. Of course, along the way, we need to make much effort to be sure of whatever the final results being independent of coordinate and basis choices. This is a sort of tautology but an important one; the final sensible answers we extract should not depend on the redundant choices, yet the redundant description is often more convenient than otherwise. Such redundances associated with coordinate and basis choices here are part of much bigger concept called the gauge principle which we will encounter many times over as we develop General Relativity and other related theories.

B.3 Exterior Calculus

B.3.1 Differential Forms and the Exterior Derivative

Another interesting derivative is called the exterior derivative. For this, we need to introduce differential forms, generalizing differential 1-forms, which are covariant tensors whose components are totally antisymmetric. Namely, any exchange of a pair of indices results in the sign flip of the components in question,

$$\Omega_{\cdots\,\alpha\,\cdots\,\beta\,\cdots} = -\Omega_{\cdots\,\beta\,\cdots\,\alpha\,\cdots} \ . \tag{B.3.1}$$

One can impose such constraints more simply by inventing the wedge product between differential 1-forms, such that

$$df \wedge dg = -dg \wedge df \ . \tag{B.3.2}$$

which may be translated to ordinary tensor product as

$$df \wedge dg \equiv df \otimes dg - dg \otimes df \ . \tag{B.3.3}$$

With this, a differential k-form $\Omega^{(k)}$ has the form,

$$\Omega^{(k)} = \frac{1}{k!}\Omega_{\alpha_1\,\cdots\,\alpha_k} dx^{\alpha_1} \wedge\, \cdots\, \wedge dx^{\alpha_k} \ . \tag{B.3.4}$$

Differential forms obey simple multiplicative rules,

$$\Omega^{(k)} \wedge \Omega'^{(p)} = (-1)^{kp}\, \Omega'^{(p)} \wedge \Omega^{(k)} \tag{B.3.5}$$

derived from the basic relation (B.3.2).

On this special class of covariant tensors, one can define another type of derivatives that maps a k-form to $(k+1)$-form,

$$d\Omega \ = \ \frac{1}{k!}\partial_\mu \Omega_{\alpha_1\,\cdots\,\alpha_k} dx^\mu \wedge dx^{\alpha_1} \wedge\, \cdots\, \wedge dx^{\alpha_k} \ , \tag{B.3.6}$$

or equivalently,

$$(d\Omega)_{\alpha_1\,\cdots\,\alpha_{k+1}} = \partial_{\alpha_1}\Omega_{\alpha_2\,\cdots\,\alpha_{k+1}} - \partial_{\alpha_2}\Omega_{\alpha_1\alpha_3\,\cdots\,\alpha_{k+1}} + \cdots + (-1)^k \partial_{\alpha_{k+1}}\Omega_{\alpha_1\,\cdots\,\alpha_k} \tag{B.3.7}$$

The simplest example is $k = 0$, where we obtain the differential 1-form df out of a function.

Does $d\Omega^{(k)}$ make sense as a tensor for $k \geq 1$? Under the coordinate transformation $x \to \tilde{x}$, a potential problem comes from pieces where ∂_μ acts on the transformation matrix rather than on components of Ω such that an extra term with factors like

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial \tilde{x}^\beta}{\partial x^\alpha}\right) = \frac{\partial^2 \tilde{x}^\beta}{\partial x^\mu \partial x^\alpha} \tag{B.3.8}$$

show up and may potentially interfere with the tensorial property of $d\Omega$. However, in $d\Omega$, all such pieces are contracted with

$$dx^\mu \wedge dx^\alpha \ , \tag{B.3.9}$$

and the latter's antisymmetric property removes any such terms, as partial derivatives of any given coordinate system always commute among themselves.

One of the most important properties of this so-called exterior derivative d is that it is nilpotent. That is

$$dd = 0 \tag{B.3.10}$$

identically. Again this follows from the commuting property of the partial derivative since

$$dd\Omega \;=\; \frac{1}{k!}\partial_{[\mu}\partial_{\nu]}\Omega_{\alpha_1\ \cdots\ \alpha_k}\, dx^\mu \wedge dx^\nu \wedge dx^{\alpha_1} \wedge\ \cdots\ \wedge dx^{\alpha_k} \;=\; 0\ . \tag{B.3.11}$$

Another useful fact about the exterior derivative is

$$d\left(\Omega^{(k)} \wedge \Omega'^{(p)}\right) = d\Omega^{(k)} \wedge \Omega'^{(p)} + (-1)^k \Omega^{(k)} \wedge d\Omega'^{(p)}\ , \tag{B.3.12}$$

whose sign follows from the antisymmetric property of the wedge product.

One example of this we encountered earlier is $F_{\mu\nu}$ of Maxwell in flat spacetime,

$$F \;\equiv\; \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu \;=\; F_{0i}dx^0 \wedge dx^i + \frac{1}{2}F_{ij}dx^i \wedge dx^j\ . \tag{B.3.13}$$

Recall that one can introduce the gauge field A such that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\ . \tag{B.3.14}$$

In turn, if the gauge field is regarded as 1-form, although this is not quite correct since A is really a connection,

$$A \;\equiv\; A_\mu dx^\mu\ , \tag{B.3.15}$$

upon which we can write

$$F = dA\ . \tag{B.3.16}$$

The tensorial property of F is guaranteed by the above, but we may also spell out the transformation rule and see this more explicitly,

$$\partial_\beta A_\alpha - \partial_\alpha A_\beta \;=\; \frac{\partial \tilde{x}^\mu}{\partial x^\beta}\frac{\partial \tilde{x}^\nu}{\partial x^\alpha}\left(\tilde{\partial}_\mu \tilde{A}_\nu - \tilde{\partial}_\nu \tilde{A}_\mu\right)\ . \tag{B.3.17}$$

One simple consequence of the above new formalism is

$$dF = ddA = 0 \tag{B.3.18}$$

identically, which gives half of the Maxwell equations,

$$\epsilon^{\alpha\beta\gamma\lambda}\partial_\beta F_{\gamma\lambda} = 0 \ , \tag{B.3.19}$$

which do not come with source terms on the right-hand side. One can see now that this half is a trivial consequence of $dd = 0$ and $F = dA$. This type of identities is called the Bianchi identity. In fact, this example offers the simplest possible prototype for the notion of curvature. We will come back to the gauge field in a later section as a motivation of the Riemann curvature.

Finally, we say that Ω is closed if it obeys

$$d\Omega = 0 \ . \tag{B.3.20}$$

If Ω can be written as

$$\Omega = dw \tag{B.3.21}$$

for some differential form w, we say it is exact. As such, 2-form field strength F of the Maxwell theory is closed since $dF = 0$.

However, $F = dA$ does not mean that F is exact because A is not a differential form, per se. A is not a tensor but a connection, or a gauge field well-defined only modulo $A \to A + d\theta$ for arbitrary gauge function θ. When we glue A across overlapping local charts, this additional shift is allowed in addition to the usual tensorial coordinate transformation via chain rules.

B.3.2 Lie Derivative on Differential Forms

Given a differential form Ω, we define a contraction with a vector field V

$$V \lrcorner \Omega \ , \tag{B.3.22}$$

say, for example,

$$\Omega = \Omega_{123}\, dx^1 \wedge dx^2 \wedge dx^3 \tag{B.3.23}$$

$$\Rightarrow \quad V \lrcorner \Omega = V^1 \Omega_{123}\, dx^2 \wedge dx^3 - V^2 \Omega_{123}\, dx^1 \wedge dx^3 + V^3 \Omega_{123}\, dx^1 \wedge dx^2$$

The same can be more generally written in terms of the antisymmetrized summation form (B.3.4) as

$$(V \lrcorner \Omega)^{(k-1)}_{\mu_1\mu_2\cdots\mu_{k-1}} \equiv V^\alpha \Omega^{(k)}_{\alpha\mu_1\mu_2\cdots\mu_{k-1}} \ . \tag{B.3.24}$$

The combinatorial factors $1/k!$ and $1/(k-1)!$ are taken into account in the definition of the components in the latter general form, while for the former samples these factors and the summations are used up for the formula for Ω and the contraction rule for $V \lrcorner \Omega$.

With this, the Lie derivative along V has a succinct form

$$\mathfrak{L}_V \Omega = d(V \lrcorner \Omega) + V \lrcorner d\Omega \ . \tag{B.3.25}$$

We can read off an important identity from (B.3.25)

$$d\, \mathfrak{L} = \mathfrak{L}\, d \ , \tag{B.3.26}$$

as

$$[d, \mathfrak{L}_V]\Omega = d[d(V \lrcorner \Omega) + \cancel{V \lrcorner d\Omega}] - [d(\cancel{V \lrcorner d\Omega}) + V \lrcorner dd\Omega] = 0 \ , \tag{B.3.27}$$

with $dd = 0$, for any vector field V and any differential form Ω.

Two special limits of this is when $k = 0$ and $k = d$. For the former, Ω is a function, say f,

$$\mathfrak{L}_V f = V \lrcorner df = V[f] \ , \tag{B.3.28}$$

giving the directional derivative along V as expected. For a differential form of the highest possible rank,

$$\Omega^{(d)} = \mu(x)\ dx^1 \wedge \cdots \wedge dx^d \ , \tag{B.3.29}$$

we find

$$\mathfrak{L}_V \Omega^{(d)} = d(V \lrcorner\, \Omega^{(d)}) = \partial_\alpha(\mu(x)\, V^\alpha)\, dx^1 \wedge \cdots \wedge dx^{\,d} \ , \tag{B.3.30}$$

since $d\Omega^{(d)} = 0$ identically.

B.3.3 Integration and Stokes' Theorem

Differential forms are naturally integrable. For instance, given a 1-form

$$\int \Omega^{(1)}_\mu(x)\, dx^\mu \ , \tag{B.3.31}$$

let us choose a path $\mathrm{C}_1 : \mathbf{s} \to x^\mu(\mathbf{s})$, with $\mathbf{s} \in \mathbf{I} \equiv [0,1]$, whereby one can define the integration of the 1-form along D_1 which is the image of C_1 as

$$\int_{\mathrm{D}_1} \Omega^{(1)}_\mu(x)\, dx^\mu = \int_0^1 \underbrace{\left(\Omega_\mu(x(\mathbf{s}))\frac{dx^\mu}{d\mathbf{s}}\right) d\mathbf{s}}_{\text{pull-back of } \Omega_\mu dx^\mu \text{ to } \mathbf{I}} \ , \tag{B.3.32}$$

which brings us to the ordinary integral by pulling-back $\Omega^{(1)}$ back to $\mathbf{I} \subset \mathbb{R}$.

If $\Omega^{(2)}$ is a 2-form,

$$\Omega^{(2)} = \frac{1}{2}\Omega_{\mu\nu}(x)\, dx^\mu \wedge dx^\nu \ . \tag{B.3.33}$$

We define the integral over $\mathrm{C}_2 : \mathbf{s}^i \mapsto x^\mu(\mathbf{s}^1, \mathbf{s}^2)$ similarly via the pull-back of Ω to $\mathbf{I} \times \mathbf{I}$ by C_2 as

$$\int_{\mathrm{D}_2} \Omega^{(2)} = \int_{\mathrm{D}_2} \frac{1}{2}\Omega_{\mu\nu}(x)\, dx^\mu \wedge dx^\nu = \int_{\mathbf{I}\times\mathbf{I}} \frac{1}{2}\Omega^{(2)}_{\mu\nu}(x(\mathbf{s}))\, \frac{\partial x^\mu}{\partial \mathbf{s}^i}\frac{\partial x^\nu}{\partial \mathbf{s}^j}\ d\mathbf{s}^i \wedge d\mathbf{s}^j. \tag{B.3.34}$$

The generalization to higher-rank forms $\Omega^{(k)}$ and higher-dimensional surfaces D_k should be clear,

$$\int_{\mathrm{D}_k} \Omega^{(k)} = \int_{\mathbf{I}^k} \mathrm{C}^*_k\left(\Omega^{(k)}\right) \ , \tag{B.3.35}$$

where C_k is a map from $\mathbf{I}^k$ onto D_k and C^*_k is the pull-back.

Thus, the local form of the integration on curved manifold is inherited from that on $\mathbb{R}^k$. We have not yet declared what we mean by the integration of a differential k-form on $\mathbb{R}^k$. Most of it follows from the usual multi-integration rule,

$$\int_{\mathbf{I}^k} f(\mathbf{s})\, d\mathbf{s}^1 \wedge \cdots \wedge d\mathbf{s}^k = \int_0^1 d\mathbf{s}^1 \cdots \int_0^1 d\mathbf{s}^k\; f(\mathbf{s}) \;. \tag{B.3.36}$$

The only new parts are (1) the normalization chosen such that the integration of

$$w \equiv \frac{1}{k!}\, w_{i_1\cdots i_k}\, d\mathbf{s}^{i_1} \wedge \cdots \wedge d\mathbf{s}^{i_k} \tag{B.3.37}$$

over $\mathbf{I}^k$ is

$$\int_{\mathbf{I}^k} w = \int_0^1 d\mathbf{s}^1 \cdots \int_0^1 d\mathbf{s}^k\; w_{123\cdots k} \tag{B.3.38}$$

and (2) the choice of the sign can which is fixed by the ordering of $d\mathbf{s}$'s as above. The latter choice is called the orientation. We usually assume that such a self-consistent orientation can be chosen; if a manifold does not allow one, we call the manifold unorientable.

What shall we do if the integration region cannot be covered by D_k inside a single chart? For this, we invoke a concept of the partition of unity. The latter is merely a set of smooth functions, u_a supported on local charts $\mathcal{U}_a$ such that at any given point p, we have

$$1 = \sum_a u_a \bigg|_p \;. \tag{B.3.39}$$

With such a device, we can define integral over a k-dimensional surface Σ_k, covered by multiple coordinate patches so that $\Sigma_k = \cup_a \mathrm{D}^{(a)}$, and the maps $\mathrm{C}_{(a)} : \mathbf{I}^k \to \mathrm{D}^{(a)}$, as

$$\int_{\Sigma_k} \Omega^{(k)} = \int_{\Sigma_k} \left(\sum_a u_a\right) \times \Omega^{(k)} = \sum_a \int_{\mathbf{I}^k} \mathrm{C}^*_{(a)} \left(u_a \times \Omega^{(k)}\right) \;. \tag{B.3.40}$$

One immediate consequence of this definition is that all the usual integration rules on Euclidean planes will carry over to general manifold.

Among one such is the Stokes theorem. In the language of differential forms, this theorem relates an integration of $d\Omega^{(k)}$ over Σ_{k+1} to that of $\Omega^{(k)}$ over the boundary of Σ_{k+1}, say, $(\partial\Sigma)_k$,

$$\int_{\Sigma_{k+1}} d\Omega^{(k)} = \int_{(\partial\Sigma)_k} \Omega^{(k)} \ . \tag{B.3.41}$$

There is a potential sign confusion with this in terms of how one defines the orientation of $(\partial\Sigma)_k$, given Σ_{k+1}; however, this is easily resolved by recalling how this theorem worked in Euclidean spaces over the rectangle $\mathbf{I}^k$, by the above definitions.

One immediate consequence is that by taking the special case of $\Sigma_d = \mathcal{M}_d$,

$$\int_{\mathcal{M}_d} d\Omega^{(d-1)} = \int_{(\partial\mathcal{M})_{d-1}} \Omega^{(d-1)} \quad \rightarrow \quad 0 \ \text{ if } \mathcal{M}_d \text{ has no boundary} \ , \tag{B.3.42}$$

so that, for instance,

$$\int_{\mathcal{M}_d} \mathfrak{L}_V \Omega^{(d)} = \int_{(\partial\mathcal{M})_{d-1}} V \lrcorner\, \Omega^{(d)} \quad \rightarrow \quad 0 \ \text{ if } \mathcal{M}_d \text{ has no boundary} \ , \tag{B.3.43}$$

for any V and $\Omega^{(d)}$.

B.4 Riemannian Geometry

Our discussion so far involves the differential structure on manifolds but fall shy of the Riemannian geometry proper. Various objects we have accumulated, such as partial derivatives and vector fields, differential forms and tensors, the Lie derivative, the exterior calculus, and Stokes' theorem, belong to a study commonly termed as Calculus on Manifolds. The Riemannian geometry builds upon these structures by adding the metric and the accompanying Levi-Civita connection.

Strictly speaking, between the differential structure and the Riemannian geometry sit some more flexible and universal concepts of bundles and connections that we went through quickly in the main text. The main difference of the Riemannian geometry from the general geometry of (co-)tangent bundles and affine connection is of course the existence of the metric whose covariantly constancy, in the language of the holonomy, forces the structure group to shrink from $GL(d)$ to $O(d)$. The main aim of this last Section is to recall more practical and elemantary formulae of the Riemannian geometry. In next Chapter, we will continue on but shift gears somewhat to the so-called Cartan-Maurer formulation and the accompanying theory of spinors.

B.4.1 Covariant Exterior Differential and the Curvature

The Yang-Mills curvature we have encountered and used repeatedly in this volume shares the common structures with the Riemann curvature tensor in that both arise from a commutator of covariant derivatives and how each measures rotation around an infinitesimal loop of in the spacetime. Although the Yang-Mills theories became relevant for fundamental physics much later that the Riemannian geometry, motivating the Riemann curvature as a special case of Yang-Mills curvature gives us a much simpler and unifying view on the matter.

One simple way to incorporate this universal structure of the connection and the curvature is the language of differential forms. For this, recall the exterior derivative d, which maps k-form Ω to $(k+1)$-form $d\Omega$ and obey $dd = 0$. When Ω carries additional indices transforming homogeneously when gluing between adjacent charts, i.e. a section of some vector bundle in abstract geometrical language,

$$\Omega^I \tag{B.4.1}$$

one can extend d to a covariant version, say,

$$d_{\mathcal{A}} \equiv d + \mathcal{A} \ . \tag{B.4.2}$$

with the connection $\mathcal{A}$ is matrix-valued in a Lie Algebra $\mathfrak{g}$ of some Lie group. The exterior derivative is thus elevated to a covariant exterior derivative

$$(d_{\mathcal{A}}\Omega)^I = d\Omega^I + \mathcal{A}^I{}_J \wedge \Omega^J \tag{B.4.3}$$

on the differential form Ω^I belonging to some representation $\mathbf{R}$ of $\mathfrak{g}$. When Ω^I is real, $\mathcal{A}^I{}_J$ is antisymmetric while for complex Ω^I, $\mathcal{A}^I{}_J$ may be taken as anti-Hermitian.

What happens if we take $d_{\mathcal{A}}$ twice? Does it vanish like dd? One can see clearly that this is not going to happen in general, since

$$\begin{aligned} d_{\mathcal{A}}d_{\mathcal{A}}\Omega &= d(\mathcal{A}\wedge\Omega) + \mathcal{A}\wedge d\Omega + \mathcal{A}\wedge\mathcal{A}\wedge\Omega \\ &= (d\mathcal{A} + \mathcal{A}\wedge\mathcal{A})\wedge\Omega \ , \end{aligned} \tag{B.4.4}$$

where the surviving d acts only on $\mathcal{A}$ and not on Ω. Therefore $(d_{\mathcal{A}})^2$ produces a matrix-valued differential 2-form,

$$\mathcal{F}^I{}_J \;=\; (d\mathcal{A} + \mathcal{A}\wedge\mathcal{A})^I{}_J \;. \tag{B.4.5}$$

The same may be written component-wise as,

$$\mathcal{F}^I{}_J \;=\; \frac{1}{2}\mathcal{F}^I{}_{J\mu\nu}dx^\mu \wedge dx^\nu \;, \tag{B.4.6}$$

with

$$\mathcal{F}^I{}_{J\mu\nu} = (\partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu])^I{}_J \;, \tag{B.4.7}$$

which is precisely the field strength of the general gauge connection we have encountered when motivating the Riemann curvature tensor.

An important fact is that the curvature transform homogeneously under the "gauge" transformation

$$d + \mathcal{A} \;\to\; U^{-1}(d+\mathcal{A})U \qquad \Rightarrow \qquad \mathcal{F} \;\to\; U^{-1}\mathcal{F}U \tag{B.4.8}$$

even though the former means $\mathcal{A} \to U^{-1}\mathcal{A}U + U^{-1}dU$. This occurs because $(d+\mathcal{A})^2$ contains no left-over derivative operation that can act on U on the right. An infinitesimal version of the same is also well known, which, with $U = e^\Phi$, gives

$$d + \mathcal{A} \;\to\; e^{-\Phi}(d+\mathcal{A})e^\Phi \qquad \to \qquad \boldsymbol{\delta}_\Phi\mathcal{A} = d\Phi + \mathcal{A}\Phi - \Phi\mathcal{A} \tag{B.4.9}$$

Note that we did not specify to which group U belongs, nor in which algebra $\mathcal{A}$ takes value. The above definitions and relations hold regardless these details.

Another such universal fact is that the field strength obeys

$$\begin{aligned} d_\mathcal{A}\mathcal{F} \;&\equiv\; d\mathcal{F} + \mathcal{A}\wedge\mathcal{F} - \mathcal{F}\wedge\mathcal{A} \\ &=\; d(\mathcal{A}\wedge\mathcal{A}) + \mathcal{A}\wedge(d\mathcal{A} + \mathcal{A}\wedge\mathcal{A}) - (d\mathcal{A} + \mathcal{A}\wedge\mathcal{A})\wedge\mathcal{A} \\ &=\; \cancel{d\mathcal{A}\wedge\mathcal{A}} - \cancel{\mathcal{A}\wedge d\mathcal{A}} + \cancel{\mathcal{A}\wedge d\mathcal{A}} - \cancel{d\mathcal{A}\wedge\mathcal{A}} \;=\; 0 \;, \end{aligned} \tag{B.4.10}$$

called the Bianchi identity. In the Maxwell theory, the Bianchi identity simplifies to $dF = 0$.

The reason why $\mathcal{A} \wedge \mathcal{F} - \mathcal{F} \wedge \mathcal{A}$ enter the definition of $d_{\mathcal{A}}\mathcal{F}$ is because $\mathcal{F}$ itself carries two indices I and J, belonging to the "adjoint" representation. Or more straightforwardly, the above form of $d_{\mathcal{A}}\mathcal{F}$ may be motivated from a Leibniz rule,

$$\begin{aligned}(d_{\mathcal{A}}\mathcal{F}) \wedge \Omega &= d_{\mathcal{A}}(\mathcal{F} \wedge \Omega) - \mathcal{F} \wedge d_{\mathcal{A}}\Omega \\ &= (d\mathcal{F}) \wedge \Omega + (\mathcal{A} \wedge \mathcal{F} - \mathcal{F} \wedge \mathcal{A}) \wedge \Omega \ . \end{aligned} \tag{B.4.11}$$

B.4.2 Levi-Civita Connection and Riemann Tensor

It is worthwhile to note how one can adapt this general structure of the connection and the curvature 2-form in a manner consistent with the existence of the metric. Imagine a manifold equipped with a metric tensor, a symmetric covariant tensor of rank two,

$$g = g_{\alpha\beta}\, dx^\alpha \otimes dx^\beta \ . \tag{B.4.12}$$

The Riemannian geometry starts from the statement that the metric should be covariantly constant,

$$\nabla_\mu g_{\alpha\beta} = 0 \tag{B.4.13}$$

in any coordinate system $\{x^\mu\}$.

Starting with the general form of the covariant derivative such that

$$\begin{aligned}\nabla_\mu V^\alpha &= \partial_\mu V^\alpha + \Gamma^\alpha{}_{\mu\lambda} V^\lambda \ , \\ \nabla_\mu W_\beta &= \partial_\mu W_\beta - \Gamma^\lambda{}_{\mu\beta} W_\lambda \ , \end{aligned} \tag{B.4.14}$$

there always exist a universal solution of the above vanishing condition as

$$\begin{aligned}&\Gamma^\mu{}_{\beta\gamma} = g^{\mu\alpha}\,\Gamma_{\alpha\beta\gamma} \ , \\ &2\Gamma_{\alpha\beta\gamma} \equiv \partial_\beta g_{\alpha\gamma} + \partial_\gamma g_{\alpha\beta} - \partial_\alpha g_{\beta\gamma}\end{aligned} \tag{B.4.15}$$

called the Christoffel symbols. This is not the most general solution to the requirement $\nabla_\mu g_{\alpha\beta} = 0$ but assumes an additional requirement known as the torsion-free condition. This torsion-free and metric-preserving covariant derivative is called the Levi-Civita connection.

The Riemann curvature tensor R is constructed from a commutator of two covariant derivatives ∇_μ and ∇_ν such that

$$[\nabla_\mu, \nabla_\nu]\, V^\alpha = R^\alpha{}_{\beta\mu\nu} V^\beta \ , \tag{B.4.16}$$

whose detailed form is

$$R^\alpha{}_{\beta\mu\nu} = \partial_\mu \Gamma^\alpha{}_{\beta\nu} - \partial_\nu \Gamma^\alpha{}_{\beta\mu} + \Gamma^\alpha{}_{\mu\lambda}\Gamma^\lambda{}_{\beta\nu} - \Gamma^\alpha{}_{\nu\lambda}\Gamma^\lambda{}_{\beta\mu} \ , \tag{B.4.17}$$

as one can see from some of the above.

A more succinct formulation of this Levi-Civita connection and the curvature can be modeled after the gauge connection and the field strength thereof as follows. We start by defining the Christoffel connection 1-form as

$$\mathbb{\Gamma}_\beta{}^\alpha \equiv -\Gamma^\alpha{}_{\mu\beta}\, dx^\mu \ . \tag{B.4.18}$$

The matrix-valued curvature 2-form, similar to the gauge analog we have encountered earlier,

$$\mathbb{R}_\beta{}^\alpha = (d\mathbb{\Gamma} + \mathbb{\Gamma} \wedge \mathbb{\Gamma})_\beta{}^\alpha \ , \tag{B.4.19}$$

translates to the usual Riemann curvature as

$$\mathbb{R}_\beta{}^\alpha{}_{\mu\nu} = -R^\alpha{}_{\beta\mu\nu} \tag{B.4.20}$$

component-wise.

Given the widely known combinatorial properties of the Riemann tensor and thus of $\mathbb{R}$ under exchanges of indices,

$$R_{\alpha\beta\mu\nu} = R_{\mu\nu\alpha\beta} \ , \qquad R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} \ , \tag{B.4.21}$$

the two are really one and the same

$$R^\alpha{}_{\beta\mu\nu} = \mathbb{R}^\alpha{}_{\beta\mu\nu} \ , \tag{B.4.22}$$

when the indices are raised and lowered via the metric.

One important fact about this connection 1-form is its transformation property under an infinitesimal coordinate shift, $x \to x - \epsilon\xi$, or equivalently under an infinitesimal diffeomorphism by $+\epsilon\xi$. One can compute the latter by demanding a Leibniz-type rule, for example,

$$\mathfrak{L}_\xi(\nabla V) = \boldsymbol{\delta}_\xi(\nabla)V + \nabla(\mathfrak{L}_\xi V) \ . \tag{B.4.23}$$

where we used the notation $\boldsymbol{\delta}_\xi$ on connections in place of the usual $\mathfrak{L}_\xi$ to emphasize that the connection is not a tensor by itself.

Reading off how the diffeomorphism acts on the Levi-Civita connection Γ. In terms of $\mathbb{\Gamma}$, we find the following suggestive form,

$$\mathfrak{L}'_\xi\mathbb{\Gamma} + d(-\partial\xi) + [\mathbb{\Gamma}, -\partial\xi] \;=\; \mathfrak{L}'_\xi\mathbb{\Gamma} + d_{\mathbb{\Gamma}}(-\partial\xi) \ , \tag{B.4.24}$$

where both $\mathfrak{L}'_\xi$ and d treat $\mathbb{\Gamma}$ as if the latter is a (matrix-valued) 1-form tensor. Note that this is a combination of translation by $\mathfrak{L}'_\xi$ and an infinitesimal gauge transformation of $\mathcal{A} = \mathbb{\Gamma}$, as in (B.4.9), under a $GL(d)$ gauge function $\Phi_\beta{}^\alpha = -\partial_\beta\xi^\alpha$. Despite the appearance of a $GL(d)$ gauge transformation here, the condition $\nabla g = 0$ implies that the action of any metric-preserving connection is $O(d)$ in secret (or $O(1, d-1)$ if Lorentzian).

B.4.3 Volume Form, Signatures, and Orientations

The Levi-Civita connection is designed to keep g constant under the covariant derivative thereof. Once this is done, it turns out one gets another covariantly constant object called the volume-form. For this, let us recall the totally antisymmetric symbol $\epsilon_{\alpha_1 \cdots \alpha_d}$ and try to build a d-form of the type

$$\frac{1}{d!}\epsilon_{\alpha_1 \cdots \alpha_d}\, dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_d} \ . \tag{B.4.25}$$

This expression if one try to use it across all coordinate systems does not define a tensor, but one can compensate the unwanted Jacobian by multiplying,

$$\sqrt{g} \equiv \sqrt{|\mathrm{det}g|} \tag{B.4.26}$$

which has the transformation property

$$\sqrt{g} \quad \xrightarrow{x\to\tilde{x}} \quad \left|\det\frac{\partial\tilde{x}}{\partial x}\right| \sqrt{\tilde{g}}\ , \tag{B.4.27}$$

regardless of the sign of det$g$.

For most physics applications, the signature of the metric is either the Lorentzian, sometimes denoted as $(-+\cdots+)$, or the Euclidean, $(++\cdots+)$. We thus have a tensorial object,

$$\mathcal{V} \equiv \frac{1}{d!}\sqrt{g}\,\epsilon_{\alpha_1\,\cdots\,\alpha_d}\,dx^{\alpha_1}\wedge\cdots\wedge dx^{\alpha_d}\ , \tag{B.4.28}$$

which remains inert under coordinate transformations, provided that the Jacobian,

$$J \equiv \det\left(\frac{\partial\tilde{x}}{\partial x}\right) \ > \ 0\ . \tag{B.4.29}$$

is everywhere positive. This $\mathcal{V}$ measures the volume element of the manifold, and thus is called the volume-form.

One can also show this volume-form $\mathcal{V}$ is covariantly constant under the Levi-Civita connection, $\nabla\mathcal{V}=0$. For this, it suffices to compute

$$\begin{aligned}\nabla_\mu(\sqrt{g}\epsilon_{123\,\cdots\,d}) \ &= \ (\partial_\mu\sqrt{g})\epsilon_{123\,\cdots\,d} - \Gamma^\alpha{}_{\mu 1}\sqrt{g}\epsilon_{\alpha 23\,\cdots\,d} - \Gamma^\alpha{}_{\mu 2}\sqrt{g}\epsilon_{1\alpha 3\,\cdots\,d} - \ \cdots \\ &= \ (\partial_\mu\sqrt{g})\epsilon_{123\,\cdots\,d} - \sqrt{g}(\Gamma^\alpha{}_{\mu\alpha})\epsilon_{123\,\cdots\,d}\ .\end{aligned} \tag{B.4.30}$$

On the other hand, we find

$$\Gamma^\alpha{}_{\mu\alpha} = \frac{1}{2}g^{\alpha\lambda}(\partial_\mu g_{\lambda\alpha} + \cancel{\partial_\alpha g_{\lambda\mu}} - \cancel{\partial_\lambda g_{\mu\alpha}}) = \frac{1}{2}g^{\alpha\lambda}\partial_\mu g_{\lambda\alpha} = \frac{1}{\sqrt{g}}\partial_\mu\sqrt{g} \tag{B.4.31}$$

so that

$$\nabla_\mu(\mathcal{V}_{123\,\cdots\,d}) \ = \ (\partial_\mu\sqrt{g})\epsilon_{123\,\cdots\,d} - (\partial_\mu\sqrt{g})\epsilon_{123\,\cdots\,d} \ = \ 0\ . \tag{B.4.32}$$

Despite all these, there is a fundamental question of whether this so-called volume-form exists globally. Note that $\mathcal{V}$'s sign gets flipped under a simple innocuous operation of $x^1\leftrightarrow x^2$. Performing the coordinate transformations sequentially but coming back to the same chart in the end, we would come back to the same volume form modulo a sign,

$$\mathcal{V} \quad \to \quad \mathcal{V}' \quad \to \quad \mathcal{V}'' \quad \to \quad \cdots \quad \to \quad \pm\mathcal{V} \,. \tag{B.4.33}$$

If one can find a collection of charts that cover the entire manifold and choose orderings in each local chart, such that one ends up + sign for all possible such sequences, the manifold is called "orientable". If not, the sign of $\mathcal{V}$ cannot be defined unambiguously, so $\mathcal{V}$ does not exist.

For the current volume, we will be content with orientable manifolds for classical theory of gravity in this note. This overall sign choice for $\mathcal{V}$ thus made is called the Orientation. With $\mathcal{V}$ chosen, one can now perform integrations over the entire manifold. Since the differential d-form $\mathcal{V}$ is a quantity that can be naturally integrated over the entire manifold, the following integration is also naturally defined

$$\int \mathcal{V} f \tag{B.4.34}$$

where f is any piece-wise continuous function on the manifold. The same integration is written at times as

$$\int \mathcal{V} f \quad \to \quad \int d^d x \, \sqrt{g} \, f \tag{B.4.35}$$

which we also use in this note. In this form, the combination $\sqrt{g}\, f$ is sometimes called a "tensor density", reflecting how it transforms nontrivially under general coordinate change by a factor of Jacobian. The name is misleading since "$\sqrt{g}\, f$" is not a sensible stand-alone object that can be separated from "$d^d x$", It should suffice to remember that the only sensible integral over the manifold is via the integration of a d-form; since the volume form $\mathcal{V}$ exists universally on any orientable manifold with a metric, this also means that a function f may be integrated over an oriented manifold via the associated d-form, $\mathcal{V} f$.

In particular, a Lie derivative on the volume form is such that,

$$\int \mathfrak{L}_\xi(\mathcal{V} f) = \int d(\xi \lrcorner \mathcal{V} f) = 0 \tag{B.4.36}$$

on any closed manifold, i.e., on any manifold without boundary.