Lecture Notes on
Classical Gauge Field Theory

Jonas T. Hartwig

Version 0.5 from Feb 20, 2025

Revision	**Date**	**Author(s)**	**Description**
0.1	Jan 21, 2025	JH	Added Lecture 1 (Geometries)
0.2	Jan 28, 2025	JH	Added Lecture 2 (Tensor Fields)
0.3	Feb 6, 2025	JH	Added Lecture 3 (Covariant Derivatives)
0.4	Feb 15, 2025	JH	Added Lecture 4 (Curvature, Metric)
0.5	Feb 20, 2025	JH	Added Lecture 5 (Invariant Integration).

Contents

1 Lecture 1: Geometries

1.1 On the Origin of Geometry in Physics

In one word, the origin of geometry in physics is *observer-independence*. Physics is the ongoing attempt to describe "nature", which is to say, the collection of phenomena that seem "external" to us; phenomena which we feel would take place even without the presence of humans. It is therefore clear that physical laws, to the extent possible, should be formulated in a way that is independent of any one person. This includes past, current, and future people, all over the world. More generally, due to the feeling that we have a choice about our future, we include hypothetical people. The umbrella term for these different perspectives is that of an *observer*. Since we are not really concerned with individual details related only to the observer itself, what matters is how the observer interacts with the "external" environment. This is sometimes described as an observer carrying a clock and a marked rod. This is referring to the observer's notion of the passage of time and local measure of distance. More generally, we can imagine an observer carrying other collections of measuring apparatuses, or gauges, through which they can probe the world. In fact, it is the only thing we care about when it comes to an observer. Thus, an **observer** can be abstracted to *a choice of coordinate system for spacetime, along with a choice of gauges*. Collectively, we refer to coordinate systems and choice of gauge as a **reference frame**. These choices extend only locally, as the observer cannot perform measurements far away.

Now we encounter the strange dilemma of how an observer would actually make such a "choice". Consider the situation of a lone observer floating in empty space, carrying a clock and nothing else. Suppose the clock is equipped with a dial, enabling the clock hands to speed up or slow down. Consider two different settings, a slow and a fast mode. What is the difference? You might say that the observer knows roughly how long a second is, and could therefore tell the two settings apart. Surely their heartbeat, if nothing else, can be used to compare the two settings? But this requires internal details about the observer which we are not allowed to refer to. There is no heartbeat, no person, only a clock. You might say, the faster setting would wear out the gears (or battery) of the clock faster and would allow us to distinguish between the two settings. But that again refers to internal structure of the observer (now just a clock). Therefore the observer must be regarded as a disembodied clock without internal structure. There is therefore *no physical difference between the settings of the clock*. The same reasoning applies to choosing measuring rods, or calibrating various gauges. The inescapable conclusion is that there is actually **no content to a particular observer themself, only the relative comparison between**

two reference frames has physical meaning. The manner in which we relate reference frames to each other therefore takes center stage. Typically this is done using the language of coordinates and coordinate transformations. It is in this way that geometry becomes a theoretical foundation for physics.

1.2 Coordinate Transformations and Geometries

Definition 1.1. Let n be a positive integer and M be a set.

i) A (global, n-dimensional) *coordinate system* on M is a mapping from M to $\mathbb{R}^n$, denoted

$$x \mapsto x^i \stackrel{\text{def}}{=} (x^1, x^2, \ldots, x^n) \tag{1.1}$$

which is injective (i.e. one-to-one).

ii) If $x \mapsto x^{i'}$ is another coordinate system, the map

$$x^i \mapsto x^{i'} \tag{1.2}$$

is called a *coordinate transformation* or a *change of coordinates*. It is a one-to-one and onto map from a subset of $\mathbb{R}^n$ (the image of $x \mapsto x^i$) to a subset of $\mathbb{R}^n$ (the image of $x \mapsto x^{i'}$).

iii) An (n-dimensional) *geometry* $\mathscr{G}$ is a collection of one-to-one and onto functions between subsets of $\mathbb{R}^n$, closed under taking inverses, and function compositions (when defined).

iv) A (global, n-dimensional) $\mathscr{G}$*-space* is a set M equipped with a set $\mathscr{C}$ of coordinate systems such that

1. The coordinate transformations between any coordinate systems from $\mathscr{C}$ belong to $\mathscr{G}$.
2. Compositions defines a map $\mathscr{G} \times \mathscr{C} \to \mathscr{C}$.

Broadly speaking, there are three geometries that play important roles in physics:

1.2.1 The Galilean/Newtonian geometry

Here $n = 3+1$ and we define $\mathscr{G}_{\text{Gal}}$ to consist of the following three kinds of transformations and their compositions:

i) Rotations/reflections:

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & A_i^j & \\ 0 & & & \end{bmatrix} \cdot \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} \tag{1.3}$$

where $(A_i^j)_{i,j=1}^3$ is an orthogonal 3×3-matrix (see Homework #1, Problem 1):

$$\sum_{j,l=1}^{3} A_i^j A_k^l \delta_{jl} = \delta_{jk}. \tag{1.4}$$

ii) *Galilean boosts* parametrized by $v^i = (v^1, v^2, v^3) \in \mathbb{R}^3$ (called *relative velocity*):

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ v^1 & 1 & 0 & 0 \\ v^2 & 0 & 1 & 0 \\ v^3 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} \tag{1.5}$$

which can also be written $\begin{cases} t' = t \\ x^{i'} = x^i + v^i \cdot t \end{cases}$

iii) Translations:

$$\begin{cases} t' = t + b \\ x^{i'} = x^i + a^i \end{cases} \tag{1.6}$$

where $b, a^1, a^2, a^3 \in \mathbb{R}$.

Remark 1.2. We give three examples of notions that are *invariant* under Galilean coordinate transformations. It means that these objects "live on M" (where $(M, \mathscr{C})$ is some fixed $\mathscr{G}$-space, $\mathscr{G}$ being the Galilean geometry defined above).

If p and q are two points on M (since M model a region if spacetime, we also call such p and q *events*), choose one of the coordinate systems on M and write $p^\mu = (t, p^1, p^2, p^3)$ and $q^\mu = (u, q^1, q^2, q^3)$. Then we can define the following two quantities (which depend on p and q):

$$\Delta t = t - u, \qquad \Delta s = \Big(\sum_{i,j=1}^{3} (p^i - q^i)(p^j - q^j)\delta_{ij} \Big)^{1/2} = \Big(\sum_{i=1}^{3} (p^i - q^i)^2 \Big)^{1/2}. \tag{1.7}$$

Then, for fixed p and q, the quantity Δt take the same value in all coordinate systems. To see this, recall that by definition of a $\mathscr{G}$-space, all coordinate systems are related to each other by some sequence of coordinate changes of the above three types. For each of those coordinate changes we have $\Delta t' = t' - u' = t - u = \Delta t$ (the only nontrivial case is translations). It means that the *duration between the events* p *and* q *is coordinate independent* and therefore is a notion all observers agree on. Consequently we can consider the duration between two events as part of reality (in Galilean physics).

Similarly, but now using that A_i^j are entries from an orthogonal matrix, the quantity Δs can be seen to be independent of coordinate changes and therefore the *(spatial) distance between two events* p *and* q is something "real" in Galilean physics.

Lastly, the Laplace differential operator

$$\Delta = \sum_{i,j=1}^{3} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \delta_{ij} \tag{1.8}$$

is also a coordinate independent quantity. This follows from the chain rule for the coordinate change of type i) above:

$$\frac{\partial}{\partial x^{i'}} = \sum_j A_{i'}^j \frac{\partial}{\partial x^j} \tag{1.9}$$

and that $A_{i'}^j$ are entries of an orthogonal matrix. (Details left as homework problem.)

1.2.2 The Lorentzian geometry (special relativity)

Again, $n = 3 + 1$. Here space and time will be considered on equal footing. To facilitate this we need a dimensionfull conversion factor between time and space. We call it c:

$$x^0 = ct \tag{1.10}$$

When we get to Maxwell's equations, c will be identified with the speed of light in vaccuum. The Lorentzian geometry $\mathscr{G}_{\text{Lor}}$ consists of compositions of three types of coordinate transformations:

i) Rotations/reflections in space:

$$\begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & A_i^j & \\ 0 & & & \end{bmatrix} \cdot \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{1.11}$$

where $(A_i^j)_{i,j=1}^3$ is an orthogonal 3×3-matrix as before.

ii) *Lorentz boosts* parametrized by $\lambda \in \mathbb{R}$ (called the *rapidity*):

$$\begin{bmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{bmatrix} = \begin{bmatrix} \cosh\lambda & \sinh\lambda & 0 & 0 \\ \sinh\lambda & \cosh\lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \tag{1.12}$$

(Here for simplicity we only consider Lorentz boost in the x^1-direction. Replacing $(\beta, 0, 0)$ by an arbitrary 3-vector β^i we get abitrary direction. We leave the details to the reader.) By hyperbolic trig identity we have $\gamma^2 - \beta^2\gamma^2 = 1$, therefore

$$\gamma = (1 - \beta^2)^{-1/2}. \tag{1.13}$$

Substituting $x^0 = ct$, we may write the Lorentz boost in terms of time and space coordinates as follows

$$\begin{cases} t' = \gamma(t + (\beta/c)x^1) \\ x^{1'} = \gamma(\beta ct + x^1) \\ x^{2'} = x^2 \\ x^{3'} = x^3 \end{cases} \tag{1.14}$$

from which we see that

$$v^1 := \frac{dx^{1'}}{dt'} = \frac{dx^{1'}/dt}{dt'/dt} = \beta c \tag{1.15}$$

Thus, substituting $\beta = v^1/c$ into (1.13),(1.14) and taking the limit as $c \to \infty$ we obtain a Galilean boost with relative velocity $v^i = (v^1, 0, 0)$. In this way we see that *Galilean physics is a limiting case of Lorentzian physics.*

iii) Translations, $x^{\mu'} = x^\mu + a^\mu$, $(\mu = 0, 1, 2, 3)$.

The following theorem gives an alternative definition, that is often much simpler to work with (but the connection to Galilean geometry is less clear).

Theorem 1.3. *The two types of Lorentz transformations i) and ii) above can be summarized in a single type of transformation of the form*

$$x^{\mu'} = \sum_{\nu=0}^{3} \Lambda_\nu^{\mu'} x^\nu \tag{1.16}$$

where $\Lambda_\nu^{\mu'}$ *are the entries of a* 4×4*-matrix satisfying*

(a)

$$\sum_{\nu',\tau'} \Lambda_\mu^{\nu'} \Lambda_\sigma^{\tau'} \eta_{\nu'\tau'} = \eta_{\mu\sigma} \tag{1.17}$$

where $\eta_{\mu\nu}$ *are the entries of the so called* flat metric, *by definition a diagonal matrix with entries* $(+1, -1, -1, -1)$ *(the "mostly minus" convention. The opposite sign convention yields the same Lorentz transformations.)*

(b) $\Lambda_0^{0'} > 0$ *(we say* (Λ_μ^ν) *is* orthochronous, *meaning time-direction-preserving, if this holds)*

(c) $\det(\Lambda_\mu^\nu) = 1$ *(we say* (Λ_μ^ν) *is* proper, *if this holds).*

Proof. Exercise for the reader. □

We introduce some notation:

$$\begin{aligned} \mathrm{O}(1,3) &= \{\Lambda \mid (1.17) \text{ holds}\} \\ \mathrm{SO}(1,3) &= \{\Lambda \in \mathrm{O}(1,3) \mid \det \Lambda = 1\} \\ \mathrm{O}^+(1,3) &= \{\Lambda \in \mathrm{O}(1,3) \mid \Lambda_0^0 > 0\} \\ \mathrm{SO}^+(1,3) &= \mathrm{SO}(1,3) \cap \mathrm{O}^+(1,3). \end{aligned}$$

Each of these is a group under matrix multiplication. $\mathrm{SO}^+(1,3)$ is called the *Lorentz group*. It is connected.

Remark 1.4. We can easily extend Lorentz geometry to more space dimensions. The same applies to Galilean geometry.

1.2.3 The smooth (or general) geometry

This n-dimensional geometry, denoted $\mathscr{G}_{\mathrm{sm}}$ consists of all one-to-one and onto functions $f : U \to V$ where U and V are open subsets of $\mathbb{R}^n$ such that all partial derivatives of f (and f^{-1}) exist to all orders.

Denoting such a function f by

$$x^i \mapsto x^{i'} = x^{i'}(x^j) = x^{i'}(x^1, x^2, \ldots, x^n) \tag{1.18}$$

we let $J^{i'}_j$ denote the entries of the Jacobian matrix:

$$J^{i'}_j = \frac{\partial x^{i'}}{\partial x^j} \tag{1.19}$$

and we let J' denote the determinant

$$J' = \det\left(J^{i'}_j\right)^n_{i',j=1} \tag{1.20}$$

Entries of the Jacobian matrix of the *inverse* coordinate change $x^{i'} \mapsto x^i(x^{j'})$ are consequently denoted

$$J^i_{j'} = \frac{\partial x^i}{\partial x^{j'}} \tag{1.21}$$

with determinant

$$J_{'} = (J')^{-1} = \det\left(J^i_{j'}\right)^n_{i,j'=1}. \tag{1.22}$$

(Note that $'$ is used a subscript in $J_{'}$ although we will not use it much.)

We can get more restricted geometries by imposing conditions on J':

- $J' > 0$ is called *oriented geometry*. It is useful for integration.

Although less important, other examples are:

- $|J| = 1$, leading to *unimodular geometry*, and
- $J = 1$, giving *proper geometry* (in analogy with Lorentz transformations).

1.3 Homework #1

1. Show that a matrix $A = (A_{ij})^n_{i,j=1}$ satisfies $A^T \cdot A = I_n$ (where A^T denotes the transpose) if and only if

$$\sum_{j,l=1}^{n} A_{ij} A_{kl} \delta^{jl} = \delta_{ik}, \qquad \forall i, k \tag{1.23}$$

where $\delta_{jl} = \delta^{jl} = \begin{cases} 1, & j = l \\ 0, & j \neq l \end{cases}$ denotes the Kronecker delta.

2. Let $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a symmetric bilinear form on $\mathbb{R}^n$ and define $B_{ij} = B(e_i, e_j)$ where $(e_i)_{i=1}^n$ is the standard ordered basis for $\mathbb{R}^n$. Further, let $(A_i^j)_{i,j=1}^n$ be a real $n \times n$-matrix. Show that B is A-invariant, that is,

$$B(Av, Aw) = B(v, w), \qquad \forall v, w \in \mathbb{R}^n \text{ (regarded as column vectors)} \tag{1.24}$$

if and only if the following identity holds:

$$\sum_{j,l} A_i^j A_k^l B_{jl} = B_{ik}, \qquad \forall i, k. \tag{1.25}$$

3. Let $(M, \mathscr{C})$ be a $\mathscr{G}$-space where $\mathscr{G}$ is any geometry whose coordinate transformations are smooth (i.e. $\mathcal{C}^\infty$) maps between open subsets of $\mathbb{R}^n$. Let $\phi : M \to \mathbb{R}$ be a function. Show that if ϕ is smooth in some coordinate system, then it is smooth in all coordinate systems. [That is, let $M \ni x \mapsto x^i \in \mathbb{R}^n$ be a coordinate system from the collection $\mathscr{C}$. Abusing notation, denote the corresponding function from (a subset of) $\mathbb{R}^n$ to $\mathbb{R}$ by $\phi(x^i) = \phi(x^1, x^2, \ldots, x^n)$. Show that if $\phi(x^i)$ is smooth then $\phi(x^{i'})$ is smooth for every other coordinate system $x^{i'}$ from $\mathscr{C}$.] (In this case we call ϕ *smooth*.)

4. Show that the Laplace operator

$$\Delta = \sum_{i,j=1}^{3} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \delta_{ij} \tag{1.26}$$

is invariant under Galilean coordinate transformations. (Conclude that if $\phi, \psi : M \to \mathbb{R}$ are smooth functions, in the sense of previous problem, then the equation $\Delta\phi = \psi$ makes sense without choosing coordinates. See the Newton-Laplace problem below.)

5. Show that the d'Alembert operator

$$\Box = \sum_{\mu,\nu=0}^{3} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} \eta_{\mu\nu} \tag{1.27}$$

is invariant under Lorentz transformations. [Hint: Use Theorem 1.3.]

6. Consider a $\mathscr{G}_{\mathrm{Gal}}$-space $(M, \mathscr{C})$. The *Newton-Laplace* equation for gravity is

$$\Delta\phi = 4\pi G\rho \tag{1.28}$$

where

- $\phi : M \to \mathbb{R}$ (in coordinates, $\phi(t, x, y, z)$) is the so-called *gravitational potential*,
- $\rho : M \to \mathbb{R}$ is a mass density distribution,
- G is Newton's constant.

The physical interpretation of ϕ is that $\nabla\phi = -\vec{a}$ where $\vec{a}$ is the acceleration vector field for a test particle.

Show that for $\rho(t, x, y, z) = \delta(x, y, z)M$ (where $\delta(x, y, z)$ denotes the delta distribution on $\mathbb{R}^3$ centered at the origin) we recover Newton's Law of Universal Gravitation.

[Integrate both sides of (1.28) over a closed ball of radius r centered at the origin: $\iiint_{B(0,r)} \Delta\phi \, d^3x = \iint_{\partial B(0,r)} \nabla\phi \cdot d\vec{S}$ by the divergence theorem. The latter equals $-\big|(\nabla\phi)(r)\big| 4\pi r^2$ by symmetry. On the other hand, $\iiint_{B(0,r)} 4\pi G\rho \, d^3x = 4\pi GM$. If the test particle has mass m is at distance r from the origin we thus get $F = m|\vec{a}(r)| = G\frac{mM}{r^2}$.]

7. In Lorentz geometry, show that if a curve is given by $x^i = x^i(t)$, $i = 1, 2, 3$, in some coordinate system such that $\sum_{i=1}^{3}(dx^i/dt)^2 = c^2$ then the same is true in any other coordinate system. (This shows that the speed of light in vacuum (the constant c) is the same in all reference frames.)

8. In Lorentz geometry, if two distinct events $x, y \in M$ have the same time coordinates, i.e. $x^0 = y^0$, in some coordinate system, then there is another coordinate system in which $x^{0'} \neq y^{0'}$. (Thus, simultaneity is lost as a "real" concept, when we go from Galilean to Lorentzian physics.)

9. In Lorentz geometry, the *interval* between two events $x, y \in M$ is defined by

$$\Delta\tau = \Big(\sum_{\mu,\nu=0}^{3} (x^\mu - y^\mu)(x^\nu - y^\nu)\eta_{\mu\nu} \Big)^{1/2} \tag{1.29}$$

Show that $\Delta\tau$ only depends on the events x and y, and not on the particular choice of coordinate system we used to write down the right hand side.

2 Lecture 2: Tensor Fields

2.1 Vector Fields and Contravariant Tensor Fields

In this lecture, we assume $\mathscr{G} \subseteq \mathscr{G}_{\text{sm}}$ is a smooth geometry, i.e. consisting of smooth maps between open subsets of $\mathbb{R}^n$ (but not necessarily *all* such maps), and we fix a $\mathscr{G}$-space $M = (M, \mathscr{C})$.

2.1.1 Curves, Tangent Vectors, and Vector Fields

Definition 2.1. A *(smooth) curve* in M is a function

$$\gamma : \mathbb{R} \to M \tag{2.1}$$

such that, for some coordinate system x^μ (from $\mathscr{C}$) the functions

$$\gamma^\mu : \mathbb{R} \to \mathbb{R}^n, \qquad \gamma^\mu(t) = x^\mu\big(\gamma(t)\big), \tag{2.2}$$

are smooth functions of t.

Remark 2.2. If $x^{\mu'}$ is another coordinate system (from $\mathscr{C}$), then the corresponding functions $\gamma^{\mu'}(t)$ defined by $\gamma^{\mu'}(t) = x^{\mu'}\big(\gamma(t)\big)$ are also smooth. This follows because any two coordinate systems are related by a smooth coordinate transformation. Explicily, for any $x \in M$, we have $x^{\mu'}(x) = x^{\mu'}(x^\nu(x))$, where in the left hand side $x^{\mu'}$ is the new coordinate system and in the right hand side $x^{\mu'}$ stands for the coordinate transformation from x^μ to $x^{\mu'}$. We thus have

$$\gamma^{\mu'}(t) = x^{\mu'}\big(\gamma^\nu(t)\big). \tag{2.3}$$

Since the composition of smooth maps is smooth, $\gamma^{\mu'}$ is indeed smooth.

Definition 2.3. The *tangent vector* of a curve γ at a point $\gamma(t)$ in a coordinate system x^μ is given by

$$\dot{\gamma}^\mu(t) = \frac{d}{dt}\gamma^\mu(t). \tag{2.4}$$

Remark 2.4. If $x^{\mu'}$ is another coordinate system then, by the chain rule,

$$\dot{\gamma}^{\mu'}(t) = \sum_{\nu=1}^{n} \frac{\partial x^{\mu'}}{\partial x^\nu}\dot{\gamma}^\nu(t). \tag{2.5}$$

This can be written more succinctly:

$$\dot{\gamma}^{\mu'} = J_{\nu}^{\mu'} \dot{\gamma}^{\nu} \tag{2.6}$$

using the *Einstein summation convention* that **for repetead indices, one upper and one lower, summation over the index range is implied**.

To summarize, tangent vectors $\dot{\gamma}^{\mu}$ are given in each coordinate system, and are related by the transformation law (2.6). We say that the components $\dot{\gamma}^{\mu}$ *transform according to* (2.6) *under coordinate transformations* $x^{\mu} \mapsto x^{\mu'}$.

Since any vector in $\mathbb{R}^n$ is the tangent vector of some curve, the following definition makes sense. (From now on we drop the adjective "smooth".)

Definition 2.5. A *vector field* on M assigns to each coordinate system x^{μ} an n-tuple of functions $A^{\mu} = (A_{\mu})_{\mu=1}^{n}$ (each defined on the image of x^{μ}) such that, for any coordinate transformation $x^{\mu} \mapsto x^{\mu'} = x^{\mu'}(x^{\nu})$ we have

$$A^{\mu'} = J_{\nu}^{\mu'} A^{\nu} \tag{2.7}$$

where $A^{\mu'}$ is the n-tuple corresponding to the coordinate system $x^{\mu'}$. The functions A^{μ} are the *components* of the vector field.

For $n = 2$, if we write out (2.7), it reads

$$\begin{cases} A^{1'}(x^{1'}, x^{2'}) = \dfrac{\partial x^{1'}}{\partial x^{1}}(x^1, x^2) \cdot A^1(x^1, x^2) + \dfrac{\partial x^{1'}}{\partial x^{2}}(x^1, x^2) A^1(x^1, x^2) \\ A^{2'}(x^{1'}, x^{2'}) = \dfrac{\partial x^{2'}}{\partial x^{1}}(x^1, x^2) \cdot A^1(x^1, x^2) + \dfrac{\partial x^{2'}}{\partial x^{2}}(x^1, x^2) A^1(x^1, x^2) \end{cases} \tag{2.8}$$

2.1.2 Addition and Scaling of Vector Fields

If A^{μ} and B^{μ} are (the components in an arbitrary coordinate system x^{μ} of) two vector fields, then

$$A^{\mu} + B^{\mu}, \qquad cA^{\mu}, \qquad (c \in \mathbb{R} \text{ a constant}) \tag{2.9}$$

define vector fields. To check this, we must see that they obey the transformation law (2.7). We have

$$A^{\mu'} + B^{\mu'} = J_{\nu}^{\mu'} A^{\nu} + J_{\nu}^{\mu'} B^{\nu} = J_{\nu}^{\mu'} (A^{\nu} + B^{\nu})$$
$$cA^{\mu'} = cJ_{\nu}^{\mu'} A^{\nu} = J_{\nu}^{\mu'} (cA^{\nu})$$

It is straightforward to show that the set of all vector fields on M, with these two operations, forms itself a vector space. This vector space is denoted by $\mathrm{Vect}(M)$.

2.1.3 Contravariant Tensor Fields

Perhaps it is natural to wonder whether we can multiply vector fields as well. We immediately run into problems, however. If we try with $A^\mu B^\mu$ (without summation convention) then these objects do not transform according to (2.7) (check this!). On the other hand, $A^\mu B^\nu$ have too many indices (two upper). The best solution is to accept that we obtain not a vector field, but a new type of object. We do have the following transformation law for $A^\mu B^\nu$:

$$A^{\mu'} B^{\nu'} = J^{\mu'}_{\lambda} J^{\nu'}_{\rho} A^\lambda B^\rho \tag{2.10}$$

(summation over both λ and ρ implied, per convention). Similarly, if we multiply three vector fields we get three factors of the Jacobian matrix. We take this as basis for a new definition.

Definition 2.6. A *contravariant r-tensor (field)* assigns to each coordinate system x^μ an n^r-tuple of functions $T^{\mu_1 \mu_2 \cdots \mu_r}$ (defined on the image of x^μ) such that under $x^\mu \mapsto x^{\mu'}$,

$$T^{\mu'_1 \mu'_2 \cdots \mu'_r} = J^{\mu'_1}_{\nu_1} J^{\mu'_2}_{\nu_2} \cdots J^{\mu'_r}_{\nu_r} T^{\nu_1 \nu_2 \cdots \nu_r}. \tag{2.11}$$

Thus, for example, if A^μ and B^μ are vector fields, then $A^\mu B^\nu$ is a contravariant 2-tensor. However, not every 2-tensor has this form. Just like for vector fields, r-tensors for fixed r can be added and scaled, and form a vector space. For example,

$$A^\mu B^\nu + A^\nu B^\mu$$

is a contravariant 2-tensor.

2.2 Covector Fields and Covariant Tensors Fields

There is a dual notion to the previous subsection. Instead of starting with a curve, which is a function from $\mathbb{R}$ to M, we start with a *scalar field* ϕ which is a function from M to $\mathbb{R}$. We assume ϕ is smooth, in the sense of Problem 3 of Homework #1, and denote by $\phi(x^\mu)$ the function $\mathbb{R}^n \to \mathbb{R}$ obtained by composing the inverse of the coordinate system by ϕ.

Definition 2.7. The *gradient* of a scalar field ϕ is

$$\partial_\mu \phi = \frac{\partial}{\partial x^\mu} \phi(x^\mu). \tag{2.12}$$

Just as with the tangent vector, the gradient depends on the choice of coordinate system. If $x^{\mu'}$ is another coordinate system, the gradient transforms as follows:

$$\partial_{\mu'}\phi = \frac{\partial}{\partial x^{\mu'}}\phi(x^\nu) = \frac{\partial x^\nu}{\partial x^{\mu'}}\frac{\partial}{\partial x^\nu}\phi(x^\nu) = J^\nu_{\mu'}\partial_\nu\phi \tag{2.13}$$

where $J^\nu_{\mu'}$ is now the inverse of the Jacobian matrix $J^{\mu'}_\nu$. That is:

$$J^\nu_{\mu'}J^{\mu'}_\lambda = \delta^\nu_\lambda, \qquad J^\nu_{\mu'}J^{\lambda'}_\nu = \delta^{\lambda'}_{\mu'} \tag{2.14}$$

where we used the Kronecker delta (see example below). We take the transformation law (2.13), which is opposite to the vector field (2.7), as the basis for a new definition.

Definition 2.8. A *covector field* on M assigns to each coordinate system x^μ an n-tuple of functions $B_\mu = (B_\mu)^n_{\mu=1}$ (each defined on the image of $x \mapsto x^\mu$) which transform as

$$B_{\mu'} = J^\nu_{\mu'}B_\nu \tag{2.15}$$

under a coordinate change $x^\mu \mapsto x^{\mu'}$.

Following the same procedure as for vector fields leads us to the dual notion to contravariant tensors.

Definition 2.9. A *covariant s-tensor (field)* assigns to each coordinate system x^μ an n^s-tuple $B_{\mu_1\mu_2\cdots\mu_s}$ of functions (defined on the image of $x \mapsto x^\mu$) which transform as

$$B_{\mu'_1\mu'_2\cdots\mu'_s} = J^{\nu_1}_{\mu'_1}J^{\nu_2}_{\mu'_2}\cdots J^{\nu_s}_{\mu'_s}B_{\nu_1\nu_2\cdots\nu_s} \tag{2.16}$$

under a coordinate change $x^\mu \mapsto x^{\mu'}$.

2.3 Mixed Tensors

Lastly, there is nothing preventing us from multiplying contravariant and covariant tensor fields. Since upper and lower indices obey opposite transformation laws, we are led to the following general definition. For non-negative integers r, s we write $\binom{r}{s}$ for the column vector (not a binomial coefficient).

Definition 2.10. An $\binom{r}{s}$*-tensor*, also called a *tensor of type* $\binom{r}{s}$, assigns to each coordinate system x^μ on M an n^{r+s}-tuple

$$T^{\mu\cdots}_{\nu\cdots}$$

with r upper and s lower subscripts, which transform as

$$T^{\mu'\cdots}_{\nu'\cdots} = J^{\mu'}_{\lambda} \cdots J^{\rho}_{\nu'} \cdots T^{\lambda\cdots}_{\rho\cdots} \tag{2.17}$$

under coordinate transformations $x^\mu \mapsto x^{\mu'}$. The functions $T^{\mu\cdots}_{\nu\cdots}$ are the *components* of the tensor. The *rank* of the tensor is the number $r + s$. The upper indices are called *contravariant* and the lower indices are called *covariant*.

The set of all $\binom{r}{s}$-tensors forms a vector space, that we could denote by $\mathscr{T}\binom{r}{s}$. Furthermore, the product of an $\binom{r}{s}$-tensor by a $\binom{t}{u}$-tensor is an $\binom{r+t}{s+u}$-tensor.

Example 2.11 (Kronecker Delta)**.** In all coordinate systems, define

$$\delta^{\nu}_{\mu} = \begin{cases} 1, & \mu = \nu \\ 0, & \mu \neq \nu \end{cases}. \tag{2.18}$$

Then δ^ν_μ is a $\binom{1}{1}$-tensor field on M. (This is one of very few "universal" tensors.) To check this, we start with the right hand side of (2.17):

$$J^{\mu'}_{\lambda} J^{\rho}_{\nu'} \delta^{\lambda}_{\rho} = J^{\mu'}_{\lambda} J^{\lambda}_{\nu'} = \delta^{\mu'}_{\nu'} \tag{2.19}$$

by one of the two Jacobian identities (2.14). This shows δ^μ_ν satisfies the transformation law for a $\binom{1}{1}$-tensor.

2.4 Operations on Tensors: Contraction, Relabeling and Permuting

The process of *contraction* involves setting one upper index equal to a lower index and then summing over that index. The result turns an $\binom{r+1}{s+1}$-tensor into an $\binom{r}{s}$-tensor.

Example 2.12. If A^μ_ν is a $\binom{1}{1}$-tensor, then A^μ_μ is a $\binom{0}{0}$-tensor, also known as a scalar field. To check this, consider:

$$A^{\mu'}_{\mu'} = J^{\mu'}_{\lambda} J^{\rho}_{\mu'} A^{\lambda}_{\rho} = \delta^{\rho}_{\lambda} A^{\lambda}_{\rho} = A^{\lambda}_{\lambda} \tag{2.20}$$

As another application, for a vector field ξ^μ and a scalar field ϕ, we can define the *directional derivative of* ϕ *along* ξ^μ to be

$$\xi^\mu \partial_\mu \phi. \tag{2.21}$$

Since this is a contraction of the $\binom{1}{1}$-tensor $\xi^\mu \partial_\nu \phi$, it is a scalar function describing the rate of change of ϕ relative to the vector field x^μ. This quantity is also known as the *Lie derivative of ϕ along ξ^μ* and is denoted $\pounds_\xi \phi$. (One can define the Lie derivative of tensors of rank > 0, but they are more complicated.)

A trivial operation is to simply permute the indices of a tensor. For example, if $A^{\mu\nu\lambda}$ is a tensor (of type $\binom{3}{0}$, needless to say) then, so is $A^{\nu\lambda\mu}$. Combining this with addition we can construct more interesting tensors, for example

$$A^{\mu\nu\lambda} + A^{\nu\lambda\mu} + A^{\lambda\mu\nu}.$$

Similarly, if B_i^j and C_k^l are tensors then a somewhat nontrivial combination of them is the tensor

$$B_i^j C_k^l + B_i^l C_k^j.$$

2.5 Connections to Multilinear Algebra

In this subsection we assume the reader is familiar with some elements of multilinear algebra. It can be safely skipped; no future sections or problems will rely on it.

If V is a real n-dimensional vector space and we pick a basis $\boldsymbol{e}_i$ [1] for V, and let $\boldsymbol{e}^i$ be the dual basis in V^*, then $V^{\otimes r} \otimes (V^*)^{\otimes s}$ has a basis $(\boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r}) \otimes (\boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s})$. Consequently, any element $T \in V^{\otimes r} \otimes (V^*)^{\otimes s}$ can be written

$$T = T^{i_1 \cdots i_r}_{j_1 \cdots j_s} \cdot (\boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_r}) \otimes (\boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_s}) \tag{2.22}$$

for some real numbers $T^{i_1 \cdots i_r}_{j_1 \cdots j_s}$. It is starting to look like a tensor field! (At least, evaluated at a point.) If $\boldsymbol{e}_{i'}$ is any other basis, we have

$$\boldsymbol{e}_j = J_j^{i'} \boldsymbol{e}_{i'} \tag{2.23}$$

for some invertible matrix $J_j^{i'}$. Let $\boldsymbol{e}^{i'}$ be the basis for V^* dual to $\boldsymbol{e}_{i'}$. It is straightforward to check that

$$\boldsymbol{e}^j = J_{i'}^j \boldsymbol{e}^{i'}, \tag{2.24}$$

where $J_{i'}^j$ is the inverse of the matrix $J_j^{i'}$. The same element T can now be expanded in the new basis $(\boldsymbol{e}_{i'_1} \otimes \cdots \otimes \boldsymbol{e}_{i'_r}) \otimes (\boldsymbol{e}^{j'_1} \otimes \cdots \otimes \boldsymbol{e}^{j'_s})$ for $V^{\otimes r} \otimes (V^*)^{\otimes s}$:

$$T = T^{i'_1 \cdots i'_r}_{j'_1 \cdots j'_s} (\boldsymbol{e}_{i'_1} \otimes \cdots \otimes \boldsymbol{e}_{i'_r}) \otimes (\boldsymbol{e}^{j'_1} \otimes \cdots \otimes \boldsymbol{e}^{j'_s}) \tag{2.25}$$

[1] We use tuple and summation conventions meaning that the $(\cdot)_{i=1}^n$ are dropped around $\boldsymbol{e}_i$. Similarly, A_i^j is a matrix, and $A_i^j e_j$ means $(\sum_{j=1}^n A_i^j e_j)_{i=1}^n$ etc.

for some $T^{i'_1\cdots i'_r}_{j'_1\cdots j'_s} \in \mathbb{R}$. Substituting (2.23) and (2.24) into (2.22), using bilinearity, and then equating coefficients with those in (2.25), we find that

$$T^{i'_1\cdots i'_r}_{j'_1\cdots j'_s} = \big(J^{i'_1}_{a_1}\cdots J^{i'_r}_{a_r}\big)\big(J^{b_1}_{j'_1}\cdots J^{b_s}_{j'_s}\big)T^{a_1\cdots a_r}_{b_1\cdots b_s}. \tag{2.26}$$

This is exactly the same transformation law as for a $\binom{r}{s}$-tensor field, (2.17), except that here there is no dependence on a point $x \in M$. The coefficients of T that we discussed in this section (which are assigned to each basis for V, and transform according to (2.26) under a change of basis) are known as *numerical tensors*.

2.6 Homework #2

1. Convert to "tensor index notation" using tuple and summation conventions.
   (i) $\big(\sum_{i=1}^n A^{jk}_{il}B^i - 2A^{ij}_{li}B^k\big)^n_{j,k=1}$
   (ii) $(X_{11}Y^1 + X_{12}Y^2,\, X_{21}Y^1 + X_{22}Y^2)$ assuming $n = 2$
   (iii) $\mathrm{Tr}(ABC)$ for square matrices A, B, C
2. Let us define the "trace" of a vector field T^μ to be $T^1 + T^2 + \cdots + T^n$. What is wrong with this definition?
3. Assuming all objects involved are tensors, which of the following expressions define tensors?
   (i) $F_{\mu\nu}F_{\lambda\rho}g^{\mu\lambda}g^{\nu\rho}$
   (ii) $A_{ij}A_{ji}$
   (iii) $X^{ij} + Y^{jk} + Z^{ki}$
   (iv) $W^a_\mu T^i_{aj}\psi^j$
4. Show that the *covariant Kronecker delta* $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ (remaining the same in all coordinate system) is a $\binom{0}{2}$-tensor if and only if the geometry consists of orthogonal transformations along with translations. (That is, show that all coordinate transformations must have the form $x^{i'} = A^{i'}_j x^j + b^{i'}$, where $A^{i'}_j$ are the entries of an orthogonal matrix.) Consider the analogous problem for the flat metric $\eta_{\mu\nu}$.

5. (a) A $\binom{0}{2}$-tensor $F_{\mu\nu}$ is *symmetric* if $F_{\mu\nu} = F_{\nu\mu}$. Similarly, we say $F_{\mu\nu}$ is *anti-symmetric* if $F_{\mu\nu} = -F_{\mu\nu}$. Analogous definitions apply to $\binom{2}{0}$-tensors. Show that these notions are coordinate-independent.

   (b) If $F_{\mu\nu}$ is anti-symmetric and $G^{\mu\nu}$ is symmetric, show that $F_{\mu\nu}G^{\mu\nu} = 0$.

   (c) Show that any $\binom{0}{2}$-tensor is the sum of one symmetric and one anti-symmetric $\binom{0}{2}$-tensor.

6. If A_μ is any covector field and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\nu$, show that $F_{\mu\nu}$ transforms as a $\binom{0}{2}$-tensor field under general (smooth) coordinate transformations. (Hint: You will have to use the product rule and the definition of $J^{\mu}_{\nu'}$.)

7. Show that if two $\binom{r}{s}$-tensors are equal in one coordinate system, then they are equal in all coordinate systems. In particular, conclude that whether "a tensor field vanishes on M (or some subset thereof)" is a coordinate-independent property.

8. Let X^μ be a vector field. Suppose that, in some coordinate system, $\partial_\mu X^\nu = 0$ on M, that is to say, all partial derivaties of all components of the vector field vanishes everywhere. Must the same be true in all other coordinate systems? If not, find some sufficient conditions which ensures $\partial_\mu X^\nu$ vanishes in all coordinate systems.

9. A $\binom{1}{1}$-tensor A^ν_μ is *invertible* if there is a $\binom{1}{1}$-tensor B^ν_μ such that $A^\nu_\mu B^\lambda_\nu = \delta^\lambda_\mu$ and $B^\nu_\mu A^\lambda_\nu = \delta^\lambda_\mu$ (by Problem 7, this is a coordinate-independent property). Show that in this case,
$$T^\mu_\lambda A^\nu_\mu = S^\nu_\lambda \iff T^\mu_\lambda = S^\nu_\lambda B^\mu_\nu$$
for any tensor fields S^ν_μ and T^ν_μ. Similarly, show that for any coordinate change $x^\mu \mapsto x^{\mu'}$ we have $S_{\mu'} J^{\mu'}_\nu = T_\nu$ if and only if $S_{\mu'} = T_\nu J^\nu_{\mu'}$.

10. A symmetric tensor $g_{\mu\nu}$ is *non-degenerate* if there is a symmetric tensor $h^{\mu\nu}$ such that $g_{\mu\nu}h^{\nu\lambda} = \delta^\lambda_\mu$. A *metric* is a symmetric non-degenerate $\binom{0}{2}$-tensor. Show that any metric $g_{\mu\nu}$ gives a well-defined bijection between vector fields and covector fields, sending A^μ to $g_{\mu\nu}A^\nu$.

11. With $n = 2$, consider the (polar) coordinate change
$$\begin{cases} x^{1'} = x^1 \cos x^2, \\ x^{2'} = x^1 \sin x^2, \end{cases}$$

where $x^1 > 0$ and $x^2 \in (0, \pi/2)$. (x^2 is the second coordinate, not the square of x)

(a) Find the matrices $J^a_{b'}$, $J^{b'}_a$, and their respective determinants $J\prime$ and J'.

(b) If a vector field A^μ is given by $A^1(x^1, x^2) = x^1$ and $A^2(x^1, x^2) = 1$, find the components $A^{\mu'}(x^{1'}, x^{2'})$ of the vector field in the coordinates $x^{\mu'}$.

(c) In each coordinate system, sketch the vector field. Sketch the coordinate curves $t \mapsto (t, 0)$ and $t \mapsto (0, t)$ of one system in the other.

3 Lecture 3: Covariant Differentiation

Fix a smooth geometry $\mathscr{G} \subseteq \mathscr{G}_{\text{sm}}$ and a $\mathscr{G}$-space $(M, \mathscr{C})$.

3.1 The Problem of Constant Fields

If $\phi : M \to \mathbb{R}$ is a scalar field, we say ϕ is *(locally) constant* if

$$\partial_\mu \phi = 0 \tag{3.1}$$

Remark 3.1. It makes sense to say that $\phi : M \to \mathbb{R}$ is constant, as a function. This does not require a coordinate system on M. Every such constant function is locally constant in the above sense. The converse holds if the image of one (hence all) coordinate system $x \mapsto x^\mu$ is a connected subset of $\mathbb{R}^n$.

Remark 3.2. The condition (3.1) is coordinate independent, because $\partial_\mu \phi$ is a tensor field (see Problem 7 of Section 2.6).

On the other hand, if A^μ is a vector field on M, the condition

$$\partial_\mu A^\nu = 0 \tag{3.2}$$

is **not coordinate independent** (see Problem 8 of Section 2.6). In fact, let us see precisely what goes wrong:

$$\begin{aligned} \partial_{\mu'} A^{\nu'} &= J^\alpha_{\mu'} \partial_\alpha (J^{\nu'}_\beta A^\beta) \\ &= J^\alpha_{\mu'} J^{\nu'}_\beta \partial_\alpha A^\beta + J^\alpha_{\mu'} J^{\nu'}_{\alpha\beta} A^\beta \end{aligned} \tag{3.3}$$

where we introduce the notation

$$J^{\nu'}_{\alpha\beta} = \partial_\alpha J^{\nu'}_\beta = \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\beta}. \tag{3.4}$$

The first term in (3.3) would say that $\partial_\mu A^\nu$ is a tensor of type $\binom{1}{1}$, but the second term,

$$J^\alpha_{\mu'} J^{\nu'}_{\alpha\beta} A^\beta, \tag{3.5}$$

spoils this. We therefore think of this as an "error term" that needs to be corrected for.

Remark 3.3. The same problem appears for covector fields. However, the expression $F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ does in fact transform as a $\binom{0}{2}$-tensor (see Problem 6 in Section 2.6), even though neither term does. The lesson we draw here is that *sometimes the sum of two non-tensorial terms is tensorial.*

These remarks leads us to the idea that we might be able to add a *counteracting term* to $\partial_\mu A^\nu$, which is also non-tensorial but transform in such a way as to counteract the error term. Note that the error term (3.5) has the index structure $\binom{1}{2}$ and contracts against the vector field A^β in (3.3).

3.2 Affine Connections

The discussion in the previous subsection leads us to make the following Ansatz:

$$\nabla_\mu A^\nu := \partial_\mu A^\nu + \Gamma^\nu_{\mu\beta} A^\beta \tag{3.6}$$

and the question we face is: *How should $\Gamma^\nu_{\mu\beta}$ transform under coordinate changes, so as to counteract the error term in* (3.3)*, making the object $\nabla_\mu A^\nu$ into a tensor of type $\binom{1}{1}$?*

Under a change of coordinates $x^\mu \mapsto x^{\mu'}$, we have by (3.6) and (3.3),

$$\begin{aligned}\nabla_{\mu'} A^{\nu'} &= \partial_{\mu'} A^{\nu'} + \Gamma^{\nu'}_{\mu'\beta'} A^{\beta'} \\ &= J^\alpha_{\mu'} J^{\nu'}_\beta \partial_\alpha A^\beta + \left(J^\alpha_{\mu'} J^{\nu'}_{\alpha\beta} + \Gamma^{\nu'}_{\mu'\alpha'} J^{\alpha'}_\beta \right) A^\beta \end{aligned} \tag{3.7}$$

where we also used that A^μ are the components of a vector field. For $\nabla_\mu A^\nu$ to be a tensor field type $\binom{1}{1}$, we want (3.7) to equal

$$J^\alpha_{\mu'} J^{\nu'}_\beta \nabla_\alpha A^\beta = J^\alpha_{\mu'} J^{\nu'}_\beta \partial_\alpha A^\beta + J^\alpha_{\mu'} J^{\nu'}_\lambda \Gamma^\lambda_{\alpha\beta} A^\beta . \tag{3.8}$$

Equating the coefficients of A^β in (3.7) and (3.8) gives

$$J^{\alpha'}_\beta \Gamma^{\nu'}_{\mu'\alpha'} = J^\alpha_{\mu'} J^{\nu'}_\lambda \Gamma^\lambda_{\alpha\beta} - J^\alpha_{\mu'} J^{\nu'}_{\alpha\beta} \tag{3.9}$$

or, after multiplying both sides by $J^\tau_{\sigma'}$ and contracting along $\tau = \beta$ (see Homework 9 in Section 2.6):

$$\boxed{\Gamma^{\nu'}_{\mu'\rho'} = J^\alpha_{\mu'} J^\beta_{\rho'} J^{\nu'}_\gamma \Gamma^\gamma_{\alpha\beta} - J^\alpha_{\mu'} J^\beta_{\rho'} J^{\nu'}_{\alpha\beta}} \tag{3.10}$$

We have discovered a new object, of fundamental importance in geometry and physics.

Definition 3.4. An *affine connection* on M assigns to each coordinate system x^μ an n^3-tuple of functions $\Gamma^\nu_{\mu\rho}$ which transforms according to (3.10) under changes of coordinates $x^\mu \mapsto x^{\mu'}$.

Definition 3.5. An *affine connection space* is a $\mathscr{G}$-space equipped with an affine connection. We denote it by $(M, \Gamma^\nu_{\mu\rho})$ or just (M, Γ) for short.

3.3 Covariant Derivatives

Definition 3.6. Let (M, Γ) be an affine connection space. The *covariant derivative* of a vector field A^μ is defined to be the $\binom{1}{1}$-tensor field $\nabla_\mu A^\nu$ defined in (3.6).

Because $\nabla_\mu A^\nu$ is a tensor field, we can now define:

Definition 3.7. A vector field A^μ on an affine connection space (M, Γ) is said to be *(locally) constant* if $\partial_\mu A^\nu = 0$.

For a scalar field ϕ, for the sake of consistent notation, we put

$$\nabla_\mu \phi = \partial_\mu \phi \tag{3.11}$$

No counteracting term is needed here, because $\partial_\mu \phi$ is already a tensor field.

To summarize, we have:

$$\begin{array}{rcl} \text{a } \binom{0}{0}\text{-tensor } \phi & \rightsquigarrow & \text{a } \binom{0}{1}\text{-tensor } \nabla_\mu \phi \\ \text{a } \binom{1}{0}\text{-tensor } A^\mu & \rightsquigarrow & \text{a } \binom{1}{1}\text{-tensor } \nabla_\mu A^\nu \end{array}$$

What about $\nabla_\mu T^{\sigma\tau}$? We want the product rule to hold. So if $T^{\sigma\tau}$ happens to be a simple tensor, i.e., a product of two vector fields $T^{\sigma\tau} = A^\sigma B^\tau$ then we would want

$$\begin{aligned} \nabla_\mu (A^\lambda B^\rho) &= (\nabla_\mu A^\lambda) B^\rho + A^\lambda (\nabla_\mu B^\rho) \\ &= (\partial_\mu A^\lambda + \Gamma^\lambda_{\mu\alpha} A^\alpha) B^\rho + A^\lambda (\partial_\mu B^\rho + \Gamma^\rho_{\mu\alpha} B^\alpha) \\ &= \partial_\mu (A^\lambda B^\rho) + \Gamma^\lambda_{\mu\alpha} A^\alpha B^\rho + \Gamma^\rho_{\mu\alpha} A^\lambda B^\alpha. \end{aligned}$$

This leads to the guess

$$\boxed{\nabla_\mu T^{\lambda\rho} = \partial_\mu T^{\lambda\rho} + \Gamma^\lambda_{\mu\alpha} T^{\alpha\rho} + \Gamma^\rho_{\mu\alpha} T^{\lambda\alpha}} \tag{3.12}$$

This does indeed work! That is, if $T^{\lambda\rho}$ are the components of a $\binom{2}{0}$-tensor, then $\nabla_\mu T^{\lambda\rho}$ are the components of a $\binom{2}{1}$-tensor. The same logic leads to covariant derivatives of higher rank contravariant tensor fields. Before stating the general formula, let's address the same question for covariant tensor fields.

3.4 Contravariant tensor fields

Similarly to the case of $\binom{2}{0}$-tensors, there is a natural way to arrive at the definition of the covariant derivative $\nabla_\mu A_\nu$ of a covector field A_μ. If we assume B^μ is some vector field, then $A_\mu B^\mu$ is a scalar field, and therefore

$$\nabla_\mu (A_\nu B^\nu) = \partial_\mu (A_\nu B^\nu) = (\partial_\mu A_\nu) B^\nu + A_\nu (\partial_\mu B^\nu). \tag{3.13}$$

On the other hand, we would like the covariant derivative to also satisfy the product rule. Therefore would like

$$\begin{aligned}\nabla_\mu(A_\nu B^\nu) &= (\nabla_\mu A_\nu)B^\nu + A_\nu(\nabla_\mu B^\nu)\\ &= (\nabla\mu A_\nu)B^\nu + A_\nu(\partial_\mu B^\nu + \Gamma^\nu_{\mu\alpha}B^\alpha)\end{aligned} \tag{3.14}$$

Equating (3.13) and (3.14), we see that a sufficient condition is that

$$\boxed{\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\alpha_{\mu\nu}A_\alpha} \tag{3.15}$$

Again, this does indeed give a $\binom{0}{2}$-tensor (check!).

3.5 General Formula for the Covariant Derivative of a Tensor Field

Given the previous subsections, the following is not surprising.

Definition 3.8. Let (M, Γ) be an affine connection space. The *covariant derivative* of an $\binom{r}{s}$-tensor field on $T^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s}$ is

$$\nabla_\lambda T^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} = \partial_\lambda T^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s} + \sum_{i=1}^{r}\Gamma^{\mu_i}_{\lambda\alpha}T^{\mu_1\cdots\overset{i}{\alpha}\cdots\mu_r}_{\nu_1\cdots\nu_s} - \sum_{j=1}^{s}\Gamma^{\beta}_{\lambda\nu_j}T^{\mu_1\cdots\mu_r}_{\nu_1\cdots\underset{j}{\beta}\cdots\nu_s} \tag{3.16}$$

The result is a tensor field of type $\binom{r}{s+1}$ (verify this!).

3.6 Locally Inertial Coordinate Systems

In this section, for simplicity we assume $\mathscr{G} = \mathscr{G}_{\text{sm}}$ so that we have all smooth coordinate transformations available. (As the reader will see, we actually only need polynomial coordinate changes of degree two or less.)

Definition 3.9. An affine connection $\Gamma^\lambda_{\mu\nu}$ is *symmetric* if $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu}$.

Symmetric connections are particularly nice, in that, for any point $x \in M$ there is a coortinate system in which the connection vanishes. This property plays a crucial role in the phenomenology of general relativity, and can also greatly simplify mathematical calculations.

Theorem 3.10. *Let $\Gamma^{\lambda}_{\mu\nu}$ be a symmetric affine connection on M, and let $y \in M$ be any point. Then there exists a coordinate system x^{μ} on M in which $\Gamma^{\lambda}_{\mu\nu}(y) = 0$.*[2]

Proof. After changing coordinates using a translation, we may without loss of generality assume that $y^{\mu} = 0$. Consider now the coordinate transformation

$$x^{\mu'} = A^{\mu'}_{\alpha} x^{\alpha} + \frac{1}{2} B^{\mu'}_{\alpha\beta} x^{\alpha} x^{\beta}$$

where $A^{\mu'}_{\alpha}$ and $B^{\mu'}_{\alpha\beta}$ are constant arrays of real numbers, to be determined. We may without loss of generality assume $B^{\mu'}_{\alpha\beta} = B^{\mu'}_{\beta\alpha}$ (because its anti-symmetric piece must vanish due to $x^{\alpha}x^{\beta} = x^{\beta}x^{\alpha}$). Computing the first and second Jacobian of this coordinate change we get

$$J^{\mu'}_{\alpha} = A^{\mu'}_{\alpha} + B^{\mu'}_{\alpha\beta} x^{\beta}, \qquad J^{\mu'}_{\alpha\beta} = B^{\mu'}_{\alpha\beta}$$

(recall the second Jacobian from (3.4)). Thus, at the point y we have $J^{\mu'}_{\alpha}(y) = A^{\mu}_{\alpha}$. Therefore, as long as $A^{\mu'}_{\alpha}$ is an invertible matrix, the coordinate change is a bijection in a neighborhood of y. The transformation law for an affine connection can be written

$$J^{\mu'}_{\alpha} J^{\nu'}_{\beta} \Gamma^{\lambda'}_{\mu'\nu'} = J^{\lambda'}_{\gamma} \Gamma^{\gamma}_{\alpha\beta} - J^{\lambda'}_{\alpha\beta}.$$

(See Problem 1) At the point y we have $y^{\mu} = 0$, so the right hand side vanishes at y if we choose

$$B^{\lambda'}_{\alpha\beta} = A^{\lambda'}_{\gamma} \Gamma^{\gamma}_{\alpha\beta}(y)$$

Note that this requires $\Gamma^{\gamma}_{\alpha\beta}$ to be symmetric. □

Definition 3.11. A coordinate system x^{μ} on an affine connection space (M, Γ) is *locally inertial at $y \in M$* if the components of the connection vanish at y:

$$\Gamma^{\lambda}_{\mu\nu}(y) = 0.$$

Such coordinate systems are also called *normal* and the point y is a *pole* of the coordinate system.

[2]Note that we here regarded $\Gamma^{\lambda}_{\mu\nu}$ as a function on M rather than on the image of x^{μ}. We can of course switch back and forth between viewpoints because x^{μ} is bijective onto its image.

3.7 Homework #3

1. Show that the transformation law for an affine connection can be written
$$J^{\mu'}_{\alpha} J^{\nu'}_{\beta} \Gamma^{\lambda'}_{\mu'\nu'} = J^{\lambda'}_{\gamma} \Gamma^{\gamma}_{\alpha\beta} - J^{\lambda'}_{\alpha\beta}.$$

2. Show that if $\Gamma_{\mu}{}^{\nu}{}_{\lambda}$ and $\gamma_{\mu}{}^{\nu}{}_{\lambda}$ are two affine connections then their difference $\Gamma_{\mu}{}^{\nu}{}_{\lambda} - \gamma_{\mu}{}^{\nu}{}_{\lambda}$ is a $\binom{1}{2}$-tensor. (This suggests that an affine connection is a kind of potential.)

3. The *torsion* of an affine connection $\Gamma_{\mu}{}^{\nu}{}_{\lambda}$ is defined by $T_{\mu}{}^{\nu}{}_{\lambda} = \Gamma_{\mu}{}^{\nu}{}_{\lambda} - \Gamma_{\lambda}{}^{\nu}{}_{\mu}$. Show that the torsion is a $(1,2)$-tensor. In particular, conclude that the property of an affine connection to be symmetric is coordinate independent.

4. Differentiate $J^{\mu}_{\alpha'} J^{\alpha'}_{\nu} = \delta^{\mu}_{\nu}$ with respect to $x^{\lambda'}$ to show the following sometimes useful identity involving the second Jacobian:
$$J^{\mu}_{\alpha'\lambda'} J^{\alpha'}_{\nu} = -J^{\mu}_{\alpha'} J^{\beta}_{\lambda'} J^{\alpha'}_{\beta\nu}.$$
Use this to rewrite the transformation law for an affine connection so that the second term is a (slightly different looking) term with a plus sign.

5. Show that, in general, the set of affine connections on a space is not closed under multiplication by a scalar.

6. Given a covector field A_{μ} on an affine connection space, compare the two $\binom{0}{2}$-tensors $G_{\mu\nu} = \nabla_{\mu} A_{\nu} - \nabla_{\mu} A_{\nu}$ and $F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu}$. When are they the same?

7. Compute the second order covariant derivative $\nabla_{\mu} \nabla_{\nu} A^{\lambda}$ of a vector field A^{λ}. (Hint: First apply ∇_{μ}, using that $\nabla_{\nu} A^{\lambda}$ is a tensor field.)

8. Verify that the formula (3.15) does define a tensor field of type $\binom{0}{2}$.

4 Lecture 4: Curvature and Torsion; Metric and Vielbein

4.1 The Curvature and Torsion of an Affine Connection

Let $\mathscr{G} \subset \mathscr{G}_{\mathrm{sm}}$ be a smooth geometry and $(M, \mathscr{C})$ a $\mathscr{G}$-space equipped with an affine connection $\Gamma^{\lambda}_{\mu\nu}$. (In other words, let M be an affine connection space in a smooth setting.)

As we saw in the previous lecture, we then have available the covariant derivative ∇_μ, which operates on the space of all tensor fields. It is natural to ask to what extent the different components of ∇_μ commute.

To check this, we let A^λ be a vector field and compute $\nabla_\mu \nabla_\nu A^\lambda$ in order to compare it with $\nabla_\nu \nabla_\mu A^\lambda$. Since $\nabla_\nu A^\lambda$ is a $\binom{1}{1}$-tensor, the general formula of covariant derivative (3.16) tells us we get two correction terms, one with a minus sign corresponding to the covariant index ν, and one with a plus sign corresponding to the contravariant index λ:

$$\nabla_\mu \nabla_\nu A^\lambda = \partial_\mu(\nabla_\nu A^\lambda) - \Gamma^\alpha_{\mu\nu} \nabla_\alpha A^\lambda + \Gamma^\lambda_{\mu\alpha} \nabla_\nu A^\alpha. \tag{4.1}$$

Since the middle term already has a clear dependence on switching μ and ν, we keep it as is. For the other two terms we have, by the formula (3.6) for the covariant derivative of a vector field,

$$\begin{aligned} \partial_\mu(\nabla_v A^\lambda) &= \partial_\mu(\partial_\nu A^\lambda + \Gamma^\lambda_{\nu\alpha} A^\alpha) \\ &= \partial_\mu \partial_\nu A^\lambda + \partial_\mu \Gamma^\lambda_{\nu\alpha} A^\alpha + \Gamma^\lambda_{\nu\alpha} \partial_\mu A^\alpha \end{aligned} \tag{4.2}$$

where we used the product rule for the ordinary partial derivative, and

$$\Gamma^\lambda_{\mu\alpha} \partial_\nu A^\alpha + \Gamma^\lambda_{\mu\alpha} \Gamma^\alpha_{\nu\beta} A^\beta. \tag{4.3}$$

Substituting these into (4.1) we obtain

$$\begin{aligned} \nabla_\mu \nabla_\nu A^\lambda = {} & \partial_\mu \partial_\nu A^\lambda + \partial_\mu \Gamma^\lambda_{\nu\alpha} A^\alpha + \Gamma^\lambda_{\nu\alpha} \partial_\mu A^\alpha \\ & - \Gamma^\alpha_{\mu\nu} \nabla_\alpha A^\lambda \\ & + \Gamma^\lambda_{\mu\alpha} \partial_\nu A^\alpha + \Gamma^\lambda_{\mu\alpha} \Gamma^\alpha_{\nu\beta} A^\beta. \end{aligned} \tag{4.4}$$

When we switch μ and ν and substract, the red terms (the first, third, and fifth) cancel, and the remaining terms get anti-symmetrized in μ and ν:

$$[\nabla_\mu, \nabla_\nu] A^\lambda = \big(\underbrace{\partial_\mu \Gamma^\lambda_{\nu\alpha} - \partial_\nu \Gamma_{\mu\alpha^\lambda} + \Gamma^\lambda_{\mu\beta} \Gamma^\beta_{\nu\alpha} - \Gamma^\lambda_{\nu\beta} \Gamma^\beta_{\mu\alpha}}_{\stackrel{\text{def}}{=} R^\lambda_{\mu\nu\alpha}} \big) A^\alpha - \big(\underbrace{\Gamma^\alpha_{\mu\nu} - \Gamma^\lambda_{\nu\mu}}_{\stackrel{\text{def}}{=} T^\alpha_{\mu\nu}} \big) \nabla_\alpha A^\lambda$$

Definition 4.1. $R^\lambda_{\mu\nu\alpha}$ is the *Riemann Curvature tensor* and $T^\alpha_{\mu\nu}$ is the *torsion tensor* of the affine connection.

In a previous homework problem, it was shown that $T^{\alpha}_{\mu\nu}$ is a tensor. Therefore $T^{\alpha}_{\mu\nu}\nabla_\alpha A^\lambda$ is a tensor. Since the left hand side of (4.5) is also a tensor, it follows that $R^{\lambda}_{\mu\nu\alpha}A^\alpha$ is a tensor for every vector field A^α. One can from this draw the conclusion that $R^{\lambda}_{\mu\nu\tau}$ is a tensor of type $\binom{1}{3}$, see Problem 3 of Section 4.5. Alternatively, one can directly verify that $R^{\lambda}_{\mu\nu\tau}$ is a tensor (see Problem 4 of Section 4.5).

Remark 4.2.

1. The torsion tensor vanishes if and only if the affine connection is symmetric in its two lower indices.

2. The curvature tensor $R^{\lambda}_{\mu\nu\tau}$ is anti-symmetric in its first two covariant indices μ and ν.

3. The formula for $[\nabla_\mu, \nabla_\nu]$ acting on other tensors is similar in form to the covariant derivative, in that we get a + curvature term for each contravariant index, and a − curvature term for each covariant index, though only a single − torsion term. For example:

$$\begin{aligned}[\nabla_\mu, \nabla_\nu]B^{\sigma\tau}_{\pi} = & R^{\sigma}_{\mu\nu\alpha}B^{\alpha\tau}_{\pi} + R^{\tau}_{\mu\nu\alpha}B^{\sigma\alpha} \\ & - R^{\alpha}_{\mu\nu\pi}B^{\sigma\tau}_{\alpha} - T^{\alpha}_{\mu\nu}\nabla_\alpha B^{\sigma\tau}_{\pi}\end{aligned} \tag{4.5}$$

   (We leave it to the reader to verify this.)

4. The *Ricci tensor* is obtained by contracting the curvature tensor:

$$R_{\mu\nu} = R^{\alpha}_{\alpha\mu\nu}. \tag{4.6}$$

4.2 Metrics and Vielbeins

Definition 4.3. A *metric* $g_{\mu\nu}$ on M is a symmetric non-degenerate $\binom{0}{2}$-tensor. Symmetric means $g_{\mu\nu} = g_{\nu\mu}$, while non-degenerate means there is a contravariant tensor $\widetilde{g}^{\mu\nu}$ (sometimes called the *inverse metric*) such that

$$g_{\mu\nu}\widetilde{g}^{\nu\lambda} = \delta^{\lambda}_{\mu}. \tag{4.7}$$

Example 4.4. The flat metric η_{ab} is a metric in the Lorentzian geometry.

If X^μ and Y^μ are two vector fields, then

$$g_{\mu\nu}X^\mu Y^\nu \tag{4.8}$$

is a scalar field. Thus a metric is a kind of "dot product". More precisely, at each point on M, a metric $g_{\mu\nu}$ defines a non-degenerate bilinear form on the tangent space (defined as the space of all vector fields evaluated at that point). By Gram-Schmidt's orthonormalization process from linear algebra, we know that such a bilinear form can be diagonalized. In our notation, it means there exists a matrix $e^a_\mu(x)$ defined at each point x on M, such that

$$g_{\mu\nu}(x) = e^a_\mu(x) e^b_\nu(x) \eta_{ab} \tag{4.9}$$

where η_{ab} is diagonal with ± 1 on the diagonal (since square roots are available in $\mathbb{R}$). The pair (t, s) where t is the number of $+1$'s and s the number of -1's, is the *signature* of $g_{\mu\nu}$.

That e^a_μ are smooth functions on M follows from the fact that the operations in the orthonormalization process are algebraic. They transform as covectors in μ.

The matrix $e^a_\mu(x)$ is uniquely determined (by $g_{\mu\nu}$ and the relation (4.9)) up to a (not necessarily proper) Lorentz tranformation acting on the top index: if Λ^b_a is any matrix such that $\Lambda^a_i \Lambda^b_j \eta_{ab} = \eta_{ij}$ then $e^a_\mu(x)$ may be replaced by $\Lambda^a_b e^b_\mu(x)$. Conversely, any two solutions to (4.9) are related this way.

Thus, the Vielbein is subject to two distinct transformation laws:

$$e^a_{\mu'} = J^\nu_{\mu'} e^a_\nu, \qquad e^{a'}_\mu = \Lambda^{a'}_b e^b_\mu. \tag{4.10}$$

The fact that the matrix is invertible says there is a field e^μ_a transforming as a general vector and as a Lorentz covector such that

$$e^a_\mu e^\mu_b = \delta^a_b, \qquad e^a_\mu e^\nu_a = \delta^\nu_\mu. \tag{4.11}$$

In e^a_μ we call a a *flat index* and μ is a *curved index*. Conversely, given any such field e^a_μ, we may use (4.9) to define a metric.

The field e^a_μ is called the *Vielbein* (German for many-legged) of the metric. Other common names for e^a_μ is *soldering form*, *orthogonal frame section*, *tetrad* or *Vierbein* ($n = 4$), *Dreibein* ($n = 3$), *Tweibein* ($n = 2$).

The Vielbein plays an important role in physics and geometry. For example, it can serve as an (equivalent) starting point for gravity, wherein the Vielbein takes over the role of the graviton field. It also provides a way to define spinor fields in curved spacetime.

4.3 Raising and Lowering of Indices using a Metric

Let $g_{\mu\nu}$ be a metric. If Y^μ is a vector field, then we define the covector field

$$Y_\mu := g_{\mu\alpha} Y^\alpha. \tag{4.12}$$

We say that we have *lowered* the index μ (with respect to the metric $g_{\mu\nu}$). Similarly, we may turn a covector field A_μ into a vector field using the contravariant (what some call the "inverse") metric:

$$A^\mu = \widetilde{g}^{\mu\alpha} A_\alpha, \tag{4.13}$$

and we say we have *raised* the index μ in A_μ. These processes are mutually inverse:

$$Y^\mu \rightsquigarrow g_{\mu\alpha} Y^\alpha \rightsquigarrow \widetilde{g}^{\mu\beta} g_{\beta\alpha} Y^\alpha = \delta^\mu_\alpha Y^\alpha = Y^\mu.$$

For a tensor $A^{\mu\nu}$ we wish to be able to distinguish between the result of lowering μ or ν. It is natural to define

$$A_\mu{}^\nu = g_{\mu\alpha} A^{\alpha\nu}$$
$$A^\mu{}_\nu = g_{\nu\alpha} A^{\mu\alpha}$$

Unless $A^{\mu\nu}$ is symmetric, these are two different $\binom{1}{1}$-tensors. This makes **horizontal placement of indices** important, when we have a metric present and plan to raise and lower indices. Similarly, when a metric is present, we should avoid writing B^ν_μ because it is completely unclear what $B_{\mu\nu}$ means: $g_{\mu\alpha} B^\alpha_\nu$ or $g_{\nu\alpha} B^\alpha_\mu$? The rule is to simply **avoid writing one index directly above another**, and either denote a $\binom{1}{1}$-tensor by $B_\mu{}^\nu$ or $B^\nu{}_\mu$. This is purely a notational distinction.

There are two conventions that have some logic to them, but many texts mix and match:

In the **prefix** notation, the "default" form of writing a tensor (or connection, etc) is $T_{\nu\dots}{}^{\mu\cdots}$. This form is preserved by application of a differential operator from the left:

$$\nabla_\mu A_\nu{}^\lambda = B_{\mu\nu}{}^\lambda$$
$$2\Gamma_{\mu\nu}{}^\lambda = (\partial_\mu g_{\nu\alpha} + \cdots) g^{\alpha\lambda}$$
$$R_{\mu\nu\lambda}{}^\rho = \partial_\mu \Gamma_{\nu\lambda}{}^\rho + \cdots$$

We will mostly follow this convention.

In the **postfix** notation, one defaults to $T^{\mu\cdots}{}_{\nu\dots}$. In this convention, it is logical to write the result of derivatives with an index on the right (to preserve the default index form). For partial derivatives, one defines

$$A^\mu{}_{\nu,\lambda} \stackrel{\text{def}}{=} \partial_\lambda A^\mu{}_\nu$$

and use ; or | for covariant derivatives:

$$A^\mu{}_{\nu;\lambda} = A^\mu{}_{\nu|\lambda} \stackrel{\text{def}}{=} \nabla_\lambda A^\mu{}_\nu$$

The second covariant derivative of a vector field X^μ is written $X^\mu{}_{|\nu\lambda}$ or $X^\mu{}_{|\nu|\lambda}$. Connections and curvature tensors are then typically written

$$2\Gamma^\lambda{}_{\mu\nu} = g^{\lambda\alpha}(g_{\alpha\mu,\nu} + \cdots)$$

$$Y^\lambda{}_{|\mu|\nu} - Y^\lambda{}_{|\nu|\mu} = Y^\alpha R^\lambda{}_{\alpha\mu\nu} - Y^\lambda{}_{|\alpha} T^\alpha{}_{\mu\nu}$$

$$R^\lambda{}_{\mu\nu\rho} = \Gamma^\lambda{}_{\mu\nu,\rho} + \cdots$$

Here there one needs to use slight caution: $\nabla_\mu \nabla_\nu Y^\lambda = Y^\lambda{}_{|\nu|\mu}$. Thus, one can encounter a different sign convention for the curvature tensor, justified by the notation. (In addition, there is a different sign convention in the definition of the Ricci tensor). Ultimately these signs only matter once we compute the energy component of the energy-momentum tensor of our lagrangian density: we want the energy to be non-negative.

Remark 4.5. A curious fact is revealed when we raise both indices in the metric itself:

$$g^{\mu\nu} = \widetilde{g}^{\mu\alpha}\widetilde{g}^{\nu\beta} g_{\alpha\beta} = \widetilde{g}^{\mu\alpha}\delta^\nu_\alpha = \widetilde{g}^{\mu\nu}.$$

In other words, raising the metric gives the contravariant metric. For this reason, the contravariant metric is denoted by the simpler form $g^{\mu\nu}$. Similarly, the raising μ (using $g_{\mu\nu}$) and lowering a (using the flat metric) in the Vielbein e^a_μ gives the inverse Vielbein e^μ_a.

4.4 The Levi-Civita Connection Associated to a Metric

Definition 4.6. (i) A $\mathscr{G}$-space M equipped with a metric $g_{\mu\nu}$ is a *pseudo-Riemannian* $\mathscr{G}$-space. If the metric is positive definite (i.e. its diagonal form η_{ab} has all +1's on the diagonal) then M is a Riemannian $\mathscr{G}$-space.

(ii) A connection $\Gamma_{\mu\nu}{}^\lambda$ on a pseudo-Riemannina $\mathscr{G}$-space $(M, g_{\mu\nu})$ is called *metric* (or *metric-compatible*) if

$$\nabla_\lambda g_{\mu\nu} = 0 \tag{4.14}$$

where the covariant derivative is computed using $\Gamma_{\mu\nu}{}^\lambda$.

Theorem 4.7 (The Fundamental Theorem of (pseudo-)Riemannian Geometry)**.** *Let $(M, g_{\mu\nu})$ be a pseudo-Riemannian $\mathscr{G}$-space (where $\mathscr{G} \subset \mathscr{G}_{\mathrm{sm}}$ is any smooth geometry). Then there exists a unique affine connection $\Gamma_{\mu\nu}{}^\lambda$ on M satisfying the following two conditions:*

1) *The torsion tensor* $T_{\mu\nu}{}^{\lambda} = \Gamma_{\mu\nu}{}^{\lambda} - \Gamma_{\nu\mu}{}^{\lambda}$ *vanishes identically.*

2) $\Gamma_{\mu\nu}{}^{\lambda}$ *is metric-compatible.*

This affine connection is explicitly given by

$$\Gamma_{\mu\nu}{}^{\lambda} = \frac{1}{2}\big(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu}\big) g^{\lambda\alpha}. \tag{4.15}$$

Proof. Outlined in the Homework #4 below. □

Definition 4.8. The affine connection is called the *Levi-Civita connection* on $(M, g_{\mu\nu})$. It is sometimes denoted by

$$\overset{g}{\Gamma}_{\mu\nu}{}^{\lambda} \quad \text{or} \quad \gamma_{\mu\nu}{}^{\lambda} \quad \text{or} \quad \begin{Bmatrix} \lambda \\ \mu \;\; \nu \end{Bmatrix}$$

and the components called *Christoffel symbols*.

4.5 Homework #4

1. Let S^μ and T^μ be two vector fields on a pseudo-Riemannian $\mathscr{G}$-space $(M, g_{\mu\nu})$. Using notation of lowering and raising, show that $S^\mu T_\mu = S_\mu T^\mu$.

2. Show that the Jacobian $J^{\nu'}_{\mu}$ (for fixed μ) transforms as a vector in ν', and (for fixed ν') as a covector in μ:

$$J^{\nu''}_{\mu} = J^{\nu''}_{\alpha'} J^{\alpha'}_{\mu}, \qquad J^{\nu'}_{\mu''} = J^{\alpha'}_{\mu''} J^{\nu'}_{\alpha'} \tag{4.16}$$

[Thus, on the one hand, for a fixed coordinate system $x \mapsto x^\mu$, the Jacobian defines n vector fields, denoted $\boldsymbol{\partial}_\mu$. The components of $\boldsymbol{\partial}_\mu$ in an arbitrary coordinate system $x \mapsto x^{\nu'}$ are $(\boldsymbol{\partial}_\mu)^{\nu'} = J^{\nu'}_{\mu}$. These are the *basis vector fields* relative to the coordinate system x^μ. On the other hand, for a fixed coordinate system $x \mapsto x^\mu$, the Jacobian defines n covector fields, denoted dx^μ with components in an arbitrary coordinate system $x \mapsto x^{\mu'}$ given by $(dx^\mu)_{\nu'} = J^{\mu}_{\nu'}$. These are the *basis covector fields* (or *basis* 1*-forms*) relative to the coordinate system x^μ. Thus,

$$(\boldsymbol{\partial}_{\nu'})^{\mu} = (dx^\mu)_{\nu'} = J^{\mu}_{\nu'}. \tag{4.17}$$

3. Suppose that a tuple of functions $R^{\lambda}_{\mu\nu\tau}$ are given in each coordinate system, but no transformation law is assumed between different coordinate systems. But, suppose that we do know that the quantities $R^{\lambda}_{\mu\nu\alpha}A^{\alpha}$ define a $\binom{1}{3}$-tensor for each vector field A^{α}. Prove that in that case the quantities $R^{\lambda}_{\mu\nu\tau}$ also transform as a $\binom{1}{3}$-tensor. (The same is true for any other type, but this case is of particular relevance to the curvature tensor.) (Hint: Fix a coordinate system $x \mapsto x^{\mu'}$ and take the vector field A^{μ} to be one of the corresponding coordinate vector fields: $A^{\mu} = (\boldsymbol{\partial}_{\nu'})^{\mu} = J^{\mu}_{\nu'}$.)

4. Show directly, using the transformation laws for an affine connection, that the Riemann curvature tensor $R^{\lambda}_{\mu\nu\tau}$ is in fact a $\binom{1}{3}$-tensor. (Hint: Calculations can be cut in half by writing "$R^{\lambda}_{\mu\nu\tau} = \partial_{\mu}\Gamma^{\lambda}_{\nu\tau} + \Gamma^{\lambda}_{\mu\alpha}\Gamma^{\alpha}_{\nu\tau} - \{\mu \leftrightarrow \nu\}$".)

5. Prove the Fundamental Theorem of (pseudo-)Riemannian Geometry as follows:

   (a) Suppose that $g_{\mu\nu}$ is a metric, and that ${\Gamma_{\mu\nu}}^{\lambda}$ is any symmetric affine connection such that $\nabla_{\lambda}g_{\mu\nu} = 0$. Using the definition of the covariant derivative (see e.g. (3.16)), find an equation involving the metric and the affine connection. It should have the symbolic form $\partial g - \Gamma g - \Gamma g = 0$.

   (b) Let $\Gamma_{\mu\nu\lambda} = {\Gamma_{\mu\nu}}^{\alpha}g_{\alpha\lambda}$ be the so-called *covariant connection*.[3] Use the part (a) to show that
   $$\partial_{\mu}g_{\nu\lambda} + \partial_{\nu}g_{\mu\lambda} - \partial_{\lambda}g_{\mu\nu} = 2\Gamma_{\mu\nu\lambda}.$$

   (c) Conclude that the connection is therefore determined by the metric:
   $${\Gamma_{\mu\nu}}^{\lambda} = \frac{1}{2}g^{\lambda\alpha}\big(\partial_{\mu}g_{\nu\alpha} + \partial_{\nu}g_{\mu\alpha} - \partial_{\alpha}g_{\mu\nu}\big) \tag{4.18}$$

   (d) Conversely, prove that for any metric $g_{\mu\nu}$, the formula (4.18) actually defines an affine connection. This is the Levi-Civita connection on a pseudo-Riemannian space $(M, g_{\mu\nu})$.

6. Let e^{a}_{μ} be a Vielbein field. Show that $\Gamma^{\lambda}_{\mu\nu} = e^{\lambda}_{a}\partial_{\mu}e^{a}_{\nu}$ defines an affine connection. [This is actually a much easier formula than the traditional one (4.18) in terms of the metric! Note that the torsion does not necessarily vanish here.]

7. Let ${\Gamma_{\mu\nu}}^{\lambda}$ be the Levi-Civita connection of a metric $g_{\mu\nu}$. Fix a coordinate system x^{μ} and a point $x \in M$. Show that the following two statements are equivalent:

[3] It is only "covariant" in the sense that it has all indices downstairs. It is not a $\binom{0}{3}$-tensor but satisfies some other transformation law.

(a) ${\Gamma_{\mu\nu}}^{\lambda}(x) = 0$ (that is, this is a locally inertial coordinate system at x, see Definition 3.11)

(b) $\partial_\lambda g_{\mu\nu}(x) = 0$ (that is, in this coordinate system, the metric is constant to first order at x).

8. Let $g_{\mu\nu}$ be a metric and ${\Gamma_{\mu\nu}}^{\lambda}$ the corresponding Levi-Civita connection. Let $R_{\mu\nu\lambda\rho} = g_{\rho\alpha}{R_{\mu\nu\lambda}}^{\alpha}$ be the *covariant curvature tensor*. Prove that in a locally inertial coordinate system at x,

$$R_{\mu\nu\lambda\rho} = \frac{1}{2}\big(\partial_\mu\partial_\lambda g_{\nu\rho} + \partial_\nu\partial_\rho g_{\mu\lambda} - \partial_\nu\partial_\lambda g_{\mu\rho} - \partial_\mu\partial_\rho g_{\nu\lambda}\big) \tag{4.19}$$

9. Show that the covariant curvature tensor of the Levi-Civita connection satisfies the C_2-symmetries[4]

$$R_{\mu\nu\lambda\rho} = -R_{\nu\mu\lambda\rho}, \qquad R_{\mu\nu\lambda\rho} = -R_{\mu\nu\rho\lambda}, \qquad R_{\mu\nu\lambda\rho} = R_{\lambda\rho\mu\nu}, \tag{4.20}$$

and the C_3-symmetry

$$R_{\mu\nu\lambda\rho} + R_{\mu\lambda\rho\nu} + R_{\mu\rho\nu\lambda} = 0. \tag{4.21}$$

[Hint: The first one in (4.20) holds by definition of the curvature tensor. For the second one, by the previous problem $R_{\mu\nu\lambda\rho}(x) = R_{\mu\lambda\rho\nu}(x)$ at the pole x of an inertial coordinate system. Since that identity is tensorial, it must hold (at x) in any coordinate system. But x was arbitary so it holds everywhere. The other identity is proved similarly.]

10. Use (4.20) to show that the Ricci tensor of the Levi-Civita connection is symmetric.

11. (a) Show that if x, y, z are three linear operators on a vector space, then

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

where $[a, b] = a \circ b - b \circ a$ is the commutator. (This is the *Jacobi identity*.)

(b) Assuming $\Gamma^{\lambda}_{\mu\nu}$ is a symmetric affine connection, apply part (a) to the covariant derivatives $x = \nabla_\mu, y = \nabla_\nu, z = \nabla_\lambda$, to prove the *Bianchi Identity*

$$\nabla_\mu R^{\sigma}_{\nu\lambda\tau} + \nabla_\nu R^{\sigma}_{\lambda\mu\tau} + \nabla_\lambda R^{\sigma}_{\mu\nu\tau} = 0. \tag{4.22}$$

12. Give a second proof of the Bianchi Identity using Hint (2) in Problem 8.

[4] By C_k we mean the cyclic group of order k.

13. Show that under a spatial spherical change of coordinates $(x^0, x^1, x^2, x^3) \to (t, r, \theta, \phi)$, the Lorentzian flat metric is transformed to

$$ds^2 = dt^2 - dr^2 - r^2 d\Omega^2$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. [Here we use the *line element notation* $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ to specify the metric.]

14. For constant r and t, we have $dt = dr = 0$ in the previous problem, and we get (after negating for simplicity) the metric $ds^2 = r^2 d\Omega^2$ on the static 2-sphere. In other words, $n = 2$, and

$$g_{\theta\theta} = r^2, \qquad g_{\phi\phi} = r^2 \sin^2\theta, \qquad g_{\phi\theta} = g_{\theta\phi} = 0.$$

Compute the Levi-Civita connection, curvature tensor, Ricci tensor. Show that the curvature scalar R here is constant, equal to $\frac{2}{r^2}$.

5 Lecture 5: Invariant Integration

5.1 Invariant Integration, Scalar Densities and Relative Tensors

Let $K \subset M$ and consider the idea of an integral

$$I = \int_K \phi \; d^n x \tag{5.1}$$

where ϕ is some object to be determined. To make sense of this, we choose a coordinate system $x \mapsto x^\mu$ (from the collection $\mathcal{C}$ that M comes equipped with), and we define

$$\int_K \phi \; d^n x \stackrel{\text{def}}{=} \int_{K_1} \phi(x^\mu) \; dx^1 dx^2 \cdots dx^n \tag{5.2}$$

where $\phi(x^\mu)$ is a function expressing ϕ in the coordinate system $x \mapsto x^\mu$, and $K_1 = \{x^\mu \mid x \in K\}$. But, we want this definition to be **independent of the choice of coordinate system**. This imposes a relationship between the different coordinate expressions of ϕ. More precisely, suppose $x \mapsto x^{\mu'}$ is another coordinate system (from $\mathcal{C}$) on M. Then we need

$$\int_K \phi \; d^n x = \int_{K_2} \phi(x^{\mu'}) \; dx^{1'} dx^{2'} \cdots dx^{n'}.$$

The two coordinate systems are related by a transformation belonging to the geometry $\mathscr{G}$ under consideration. We will make the assumption on $\mathscr{G}$ that all coordinate transformations have a positive Jacobian determinant $J_{\prime} = \det(\partial x^{\mu}/\partial x^{\nu'})$. Making the substitution $x^{\mu} = x^{\mu}(x^{\mu'})$ in (5.2), we obtain

$$\int_K \phi \, d^x = \int_{K_2} \phi(x^{\mu}(x^{\nu'})) \cdot J_{\prime} \, dx^{1'} dx^{2'} \cdots dx^{n'} \tag{5.3}$$

Comparing (5.2) with (5.3), we see that we want

$$\phi(x^{\mu'}) = \phi(x^{\mu}(x^{\nu'})) \cdot J_{\prime} \tag{5.4}$$

We arrive at the following definition.

Definition 5.1. A *scalar density* ϕ assigns to each coordinate system x^{μ} a (real-valued, smooth) function $\phi(x^{\mu})$, called the *expression of* ϕ *in the coordinates* x^{μ}. Furthermore, if x^{μ} and $x^{\mu'}$ are any two coordinate systems on M, their respective expressions are related by the transformation law (5.4).

Example 5.2. Suppose that the expression of $\phi : M \to \mathbb{R}$ in a coordinate system (x, y) is $\phi(x, y) = x^3 + xy^2$. Suppose $K \subset M$ is a region whose image under the coordinate map $p \mapsto (x(p), y(p))$ is $K_1 = \{(x, y) \mid x^2 + y^2 \leq 1\}$ (the unit disk). Then, by definition,

$$\int_K \phi \, d^2x = \iint_{K_1} (x^3 + xy^2) \, dxdy$$

Suppose further that (r, θ) is another coordinate system, related to the first by polar change of coordinates $x = r\cos\theta$, $y = r\sin\theta$. The Jacobian determinant is $\det\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{smallmatrix}\right) = r$ and therefore, with $K_2 = \{(r, \theta) \mid r \leq 1\}$ and the "naive substitution" is $\phi(x(r, \theta), y(r, \theta)) = r^3 \cos\theta$. As we know from calculus, therefore,

$$\int_K \phi \, d^n x = \iint_{K_2} r^3 \cos\theta \cdot r \, drd\theta.$$

This illustrates that the *expression* $\phi(r, \theta)$ *of* ϕ *in* (r, θ) *is not merely the "naive substitution", but we also have to multiply by a Jacobian determinant.* In other words: $\phi(r, \theta) = r^3 \cos\theta \cdot r$.

Note that if ϕ and ψ are two scalar densities, then ϕ/ψ is a scalar field. Therefore we also call them relative scalars. However, the product $\phi\psi$ transforms as $\phi'\psi' = \phi\psi(J_{\prime})^2$. We say that $\phi\psi$ is a relative scalar of weight 2. Multiplying a relative scalar and a tensor gives a "relative tensor". The definition that cover everything we need is as follows.

Definition 5.3. A *relative tensor (field) of type* $\binom{r}{s}$ *and weight* w is an n^{r+s}-tuple of functions $S_{\nu\cdots}{}^{\mu\cdots}$ given in each coordinate system x^μ, which are related by the transformation law

$$S_{\nu'\cdots}{}^{\mu'\cdots} = (J_{\prime})^w \cdot J^{\alpha}_{\nu'} \cdots J^{\mu'}_{\beta} \cdots S_{\alpha\cdots}{}^{\beta\cdots}. \tag{5.5}$$

Relative tensors/vectors/covectors of weight 1 are called *tensor/vector/covector densities*. (We will only need the case when w is an integer. In geometries where all Jacobian determinants are positive, w could in principle be any real number.)

Theorem 5.4. *Relative tensors have the following properties.*

(i) *The set of all relative tensors of type* $\binom{r}{s}$ *and weight* w *forms a vector space (i.e. it is closed under addition and multiplication by constant scalars).*

(ii) *Relative tensors can be multiplied; the resulting weight is the sum of the respective weights (just like the types).*

(iii) *The* covariant derivative *of a relative tensor field of type* $\binom{r}{s}$ *and weight* w *is defined by*

$$\begin{aligned}\nabla_\lambda T_{\mu\cdots}{}^{\nu\cdots} &= \partial_\lambda T_{\mu\cdots}{}^{\nu\cdots} \\ &+ \mathit{\Gamma}_{\lambda\alpha}{}^{\nu} T_{\mu\cdots}{}^{\alpha\cdots} + \cdots \quad && \textit{(one term for each contravariant index)} \\ &- \mathit{\Gamma}_{\lambda\mu}{}^{\alpha} T_{\alpha\cdots}{}^{\nu\cdots} - \cdots \quad && \textit{(one term for each covariant index)} \\ &- w\mathit{\Gamma}_{\alpha\lambda}{}^{\alpha} T_{\mu\cdots}{}^{\nu\cdots} \quad && \textit{(a single extra term)}\end{aligned} \tag{5.6}$$

The result is a tensor of type $\binom{r}{s+1}$ *and weight* w.

(iv) *The product rule works as usual for taking the covariant derivative of any products of relative tensors. Dropping all indices for brevity, the rule reads*

$$\nabla(TS) = (\nabla T)S + S(\nabla T). \tag{5.7}$$

Example 5.5. If j^μ is a vector density, then its "covariant divergence" is

$$\nabla_\mu j^\mu = \partial_\mu j^\mu + \mathit{\Gamma}_{\mu\alpha}{}^{\mu} j^\alpha - \mathit{\Gamma}_{\beta\mu}{}^{\beta} j^\mu = \partial_\mu j^\mu \tag{5.8}$$

In particular, the usual partial derivative $\partial_\mu j^\mu$ is a scalar density! [5]

[5] Furthermore, this fact — although very special to the vector density case — can be a useful way of remembering the formula for the covariant derivative of a relative tensor.

As a consequence of this remarkable fact is that the integral of the covariant divergence of a vector density has coordinate independent meaning. An important result related to this is the divergence theorem. We only need the case when j^μ vanishes on the boundary of K.

Theorem 5.6 (Special Case of the Divergence Theorem from Calculus)**.** *For any vector density j^μ vanishing on the boundary ∂K of a region $K \subset M$, we have*

$$\int_K \partial_\mu j^\mu \, d^n x = 0. \tag{5.9}$$

Remark 5.7. In practice, we will only need to consider regions $K \subset M$ which in some coordinate system is a closed ball.

5.2 The Levi-Civita Symbols, and Determinants

We define, in all coordinate systems,

$$\varepsilon_{\mu_1 \cdots \mu_n} = \varepsilon^{\mu_1 \mu_2 \cdots \mu_n} = \begin{cases} +1, & \text{if } \mu_1 \cdots \mu_n \text{ is an even permutation of } 12 \cdots n, \\ -1, & \text{if } \mu_1 \cdots \mu_n \text{ is an odd permutation of } 12 \cdots n, \\ 0, & \text{if } \mu_1 \cdots \mu_n \text{ is not a permutation of } 12 \cdots n. \end{cases} \tag{5.10}$$

These are called the *Levi-Civita symbols*. Note that the number of indices is equal to the dimension n of the space M.

Example 5.8. When $n = 2$, we have $\varepsilon_{12} = 1, \varepsilon_{21} = -1, \varepsilon_{11} = \varepsilon_{22} = 0$.

When $n = 3$, $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{132} = \varepsilon_{321} = \varepsilon_{213} = -1$, and all other components (having at least two indices equal) are zero.

This notation is convenient for expressing determinants. For example,

$$\det({A_\mu}^\nu) = \varepsilon^{\nu_1 \cdots \nu_n} {A_{\nu_1}}^1 \cdots {A_{\nu_n}}^n. \tag{5.11}$$

(Check this for $n = 2, 3!$) It can also be written in a more tensorial way:

$$\det({A_\mu}^\nu) \varepsilon^{\mu_1 \cdots \mu_n} = \varepsilon^{\nu_1 \cdots \nu_n} {A_{\nu_1}}^{\mu_1} \cdots {A_{\nu_n}}^{\mu_n} \tag{5.12}$$

Something interesting is learned when this formula is applied to the Jacobian matrix $J_\nu^{\mu'}$ associated to a coordinate change:

$$J_{\nu_1}^{\mu_1'} \cdots J_{\nu_n}^{\mu_n'} \varepsilon^{\nu_1 \cdots \nu_n} = J' \cdot \varepsilon^{\mu_1' \cdots \mu_n'}$$

Multiplying both sides by $J_{\prime} = (J')^{-1}$ we get

$$J_{\prime} \cdot J^{\mu_1'}_{\nu_1} \cdots J^{\mu_n'}_{\nu_n} \varepsilon^{\nu_1 \cdots \nu_n} = \varepsilon^{\mu_1' \cdots \mu_n'}. \tag{5.13}$$

This shows the first part of the following proposition.

Proposition 5.9.

(a) *The contravariant Levi-Civita symbol*

$$\varepsilon^{\mu_1 \cdots \mu_n}$$

is a relative tensor of type $\binom{n}{0}$ *and weight* 1 *(i.e. an* $\binom{n}{0}$*-tensor density).*

(b) *The covariant Levi-Civita symbol*

$$\varepsilon_{\mu_1 \cdots \mu_n}$$

is a relative tensor of type $\binom{n}{0}$ *and weight* -1*.*

An importan application of this is that by anti-symmetrizing (i.e. contracting against a Levi-Civita symbol) we can create scalar densities, which can be integrated in a coordinate-independent manner:

Corollary 5.10. *If* $C_{\mu_1 \mu_2 \cdots \mu_n}$ *is any* $\binom{0}{n}$*-tensor, then*

$$\varepsilon^{\mu_1 \cdots \mu_n} C_{\mu_1 \cdots \mu_n} \tag{5.14}$$

is a scalar density.

5.3 The Scalar Density $\sqrt{-g}$

Another important way to build scalar densities comes about when we have a metric available. The determinant of the metric $g_{\mu\nu}$ is denoted by g. It can be written

$$g = \det(g_{\mu\nu}) = \frac{1}{n!} \varepsilon^{\mu_1 \cdots \mu_n} \varepsilon^{\nu_1 \cdots \nu_n} g_{\mu_1 \nu_1} \cdots g_{\mu_1 \nu_1}. \tag{5.15}$$

Since this is a fully contracted product of two relative tensors of weight 1 and a one of weight 0, it follows that g is a relative scalar of weight 2. That is,

$$g' = (J_{\prime})^2 g, \tag{5.16}$$

where $g' = \det(g_{\mu'\nu'})$. Assuming that $J_{\prime} > 0$ for all coordinate changes, taking absolute values on both sides and then the square roots, we conclude that

$$\sqrt{|g'|} = (J_{\prime}) \sqrt{|g|}. \tag{5.17}$$

This proves the following result about the metric.

Proposition 5.11. *Let $g_{\mu\nu}$ be a metric on a $\mathscr{G}$-space M, where $\mathscr{G}$ is an oriented geometry (i.e. all coordinate transformations in $\mathscr{G}$ have positive determinant). Then the field*

$$\sqrt{|g|} \tag{5.18}$$

is a scalar density. In particular, when $n = 4$ and $g_{\mu\nu}$ has Lorentzian signature (either $+---$ or $-+++$), then $g < 0$ and consequently

$$\sqrt{-g} \tag{5.19}$$

is a scalar density.

5.4 Useful Formulas Involving the Metric

In this section we provide a number of useful formulas involving the metric.

Let $g_{\mu\nu}$ be any metric and g its determinant.

Lemma 5.12.

$$\boxed{\partial_\mu g = g g^{\alpha\beta} \partial_\mu g_{\alpha\beta}} \tag{5.20}$$

Proof. By cofactor expansion of the determinant,

$$g\delta^\rho_\lambda = g_{\lambda\alpha} C^{\alpha\rho} \tag{5.21}$$

where

$$C^{\lambda\rho} = g \cdot g^{\lambda\rho} \tag{5.22}$$

Regarding the g as a polynomial in n^2 variables $g_{\mu\nu}$, we have by the chain rule [6]

$$\partial_\mu g = \frac{\partial g}{\partial g_{\alpha\beta}} \partial_\mu g_{\alpha\beta} \tag{5.23}$$

By (5.21), and that each variable $g_{\mu\nu}$ only once in a single monomial of g, we have

$$\partial_\mu g = C^{\alpha\beta} \partial_\mu g_{\alpha\beta}$$

Now, using (5.22) we get

$$\partial_\mu g = g g^{\alpha\beta} \partial_\mu g_{\alpha\beta}.$$

□

[6] Note that $\partial g / \partial g_{\alpha\beta}$ should be regarded as a contravariant object, i.e. it has two *upper* indices. This is for three reasons: In this way the contraction against $\partial_\mu g_{\alpha\beta}$ is valid. Secondly, the formula (5.23) actually shows that $\partial g / \partial g_{\alpha\beta}$ is a $\binom{2}{0}$-tensor in proper Lorentz geometry (or any geometry with unimodular linear transformations). Third, it is analogous to how $\partial_\mu \phi = \partial \phi / \partial x^\mu$ is a $\binom{0}{1}$-tensor for scalars ϕ.

Let $|g|$ be the absolute value of the determinant g of the metric.

Lemma 5.13.

$$\boxed{\partial_\mu \sqrt{|g|} = \frac{1}{2}\sqrt{|g|} \cdot g^{\alpha\beta}\partial_\mu g_{\alpha\beta}} \tag{5.24}$$

Proof. We write the proof when $|g| = -g$.

$$\begin{aligned}
\partial_\mu \sqrt{-g} &= \frac{1}{2}(-g)^{-1/2}(-\partial_\mu g) && \text{by the chain rule} \\
&= \frac{1}{2}(-g)^{-1/2} \cdot (-g g^{\alpha\beta}\partial_\mu g_{\alpha\beta}) && \text{by Lemma 5.12} \\
&= \frac{1}{2}\sqrt{-g} g^{\alpha\beta}\partial_\mu g_{\alpha\beta}.
\end{aligned}$$

□

Lemma 5.14. *Let ${\Gamma_{\mu\nu}}^{\lambda}$ be the Levi-Civita connection corresponding $g_{\mu\nu}$. Then*

$$\boxed{{\Gamma_{\alpha\nu}}^{\alpha} = \frac{\partial_\nu \sqrt{|g|}}{\sqrt{|g|}} = \partial_\nu \ln \sqrt{|g|}} \tag{5.25}$$

Proof. The covariant Levi-Civita connection is given by

$$2\Gamma_{\mu\nu\lambda} = \partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}.$$

Since the first plus the third term is $\partial_\mu g_{\nu\lambda} - \partial_\lambda g_{\nu\mu}$, those terms vanish upon contraction against the symmetric tensor $g^{\mu\lambda}$:

$${\Gamma_{\alpha\nu}}^{\alpha} = g^{\mu\lambda}\Gamma_{\mu\nu\lambda} = \frac{1}{2}g^{\mu\lambda}\partial_\nu g_{\mu\lambda} = \frac{\partial_\nu\sqrt{|g|}}{\sqrt{|g|}}, \tag{5.26}$$

where we used Lemma 5.13 in the last equality. □

Lemma 5.15. *Let ∇_μ be the covariant derivative with respect to the Levi-Civita connection corresponding to $g_{\mu\nu}$. Then*

$$\boxed{\nabla_\mu \sqrt{|g|} = 0.} \tag{5.27}$$

Proof. Since $\sqrt{|g|}$ is a scalar density,

$$\nabla_\mu \sqrt{|g|} = \partial_\mu \sqrt{|g|} - {\Gamma_{\alpha\mu}}^{\alpha}\sqrt{|g|}$$

which is zero by Lemma 5.14. □

Lemma 5.16. *Let ∇_μ be the covariant derivative with respect to the Levi-Civita connection corresponding to $g_{\mu\nu}$. Then, for any scalar field ϕ, we have*

$$\boxed{\nabla_\mu(\phi\sqrt{|g|}) = (\partial_\mu\phi)\sqrt{|g|}.} \tag{5.28}$$

Proof. By the product rule for the covariant derivative,

$$\nabla_\mu(\phi\sqrt{|g|}) = (\nabla_\mu\phi)\sqrt{|g|} + \phi \cdot \nabla_\mu\sqrt{|g|}$$

Since ϕ is a scalar field, $\nabla_\mu\phi = \partial_\mu\phi$, and the other term is zero by Lemma 5.15. □

5.5 Homework #5

1. Show that when $n = 3$, the curl of a covector field A_j can be written $\varepsilon^{ijk}\nabla_i A_j$. What is the type and weight of this relative tensor? If you had to, how would you define the "curl" of a covector field A_μ in n dimensions?

2. Suppose we are given an affine connection $\Gamma_{\mu\nu}{}^\lambda$. Let A_μ be a relative covector of weight w. Show that the formula (5.6) for the covariant derivative of A_μ does indeed give a relative $\binom{0}{2}$-tensor of weight w.

3. Let $\Gamma_{\mu\nu}{}^\lambda$ be an affine connection and ∇_μ the corresponding covariant derivative. Show that if ϕ is a relative scalar of weight w and A_μ is a relative covector field of weight v, then the following product rule holds:

$$\nabla_\mu(\phi A_\nu) = (\nabla_\mu\phi)A_\nu + \phi\nabla_\mu A_\nu.$$

4. Prove the claim that $\varepsilon_{\mu_1\cdots\mu_n}$ is a relative tensor type $\binom{0}{n}$ and weight -1.

5. (Raised and lowered Levi-Civitas) Suppose a metric $g_{\mu\nu}$ has been chosen. Define the *raised Levi-Civita symbol* $\varepsilon_g^{\mu_1\cdots\mu_n}$ as the relative tensor obtained by raising all the indices in $\varepsilon_{\mu_1\cdots\mu_n}$ using the metric. Show that this gives a relative $\binom{n}{0}$-tensor of weight 1 and that

$$\varepsilon_g^{\mu_1\cdots\mu_n} = g^{-1}\varepsilon^{\mu_1\cdots\mu_n}$$

   Similarly, the *lowered Levi-Civita symbol* $\varepsilon^g_{\mu_1\cdots\mu_n}$ is obtained by lowering all indices in the contravariant Levi-Civita symbol and $\varepsilon^g_{\mu_1\cdots\mu_n} = g\varepsilon_{\mu_1\cdots\mu_n}$. [Warning: Some authors use $\varepsilon^{\mu_1\cdots\mu_n}$ to denote the raised Levi-Civita symbol.]

6. Let η_{ab} be the flat metric (or, if you want, any constant metric, regarded as a Lorentz $\binom{0}{2}$-tensor), $g_{\mu\nu}$ a general metric, and e_μ^a be a Vielbein field connecting the two as in (4.9). Let $\eta = \det(\eta_{ab})$, $g = \det(g_{\mu\nu})$, and $e = \det(e_\mu^a)$ be the respective determinants. Show that

$$g = \eta e^2 \tag{5.29}$$

In particular, for $n = 4$ and Lorentzian signature, conclude that

$$e = \sqrt{-g}. \tag{5.30}$$